Cell Systems, Volume *3*

Supplemental Information

Single-Cell Transcriptomics Reveals that Differentiation and Spatial Signatures Shape Epidermal and Hair Follicle Heterogeneity

Simon Joost, Amit Zeisel, Tina Jacob, Xiaoyan Sun, Gioele La Manno, Peter Lönnerberg, Sten Linnarsson, and Maria Kasper

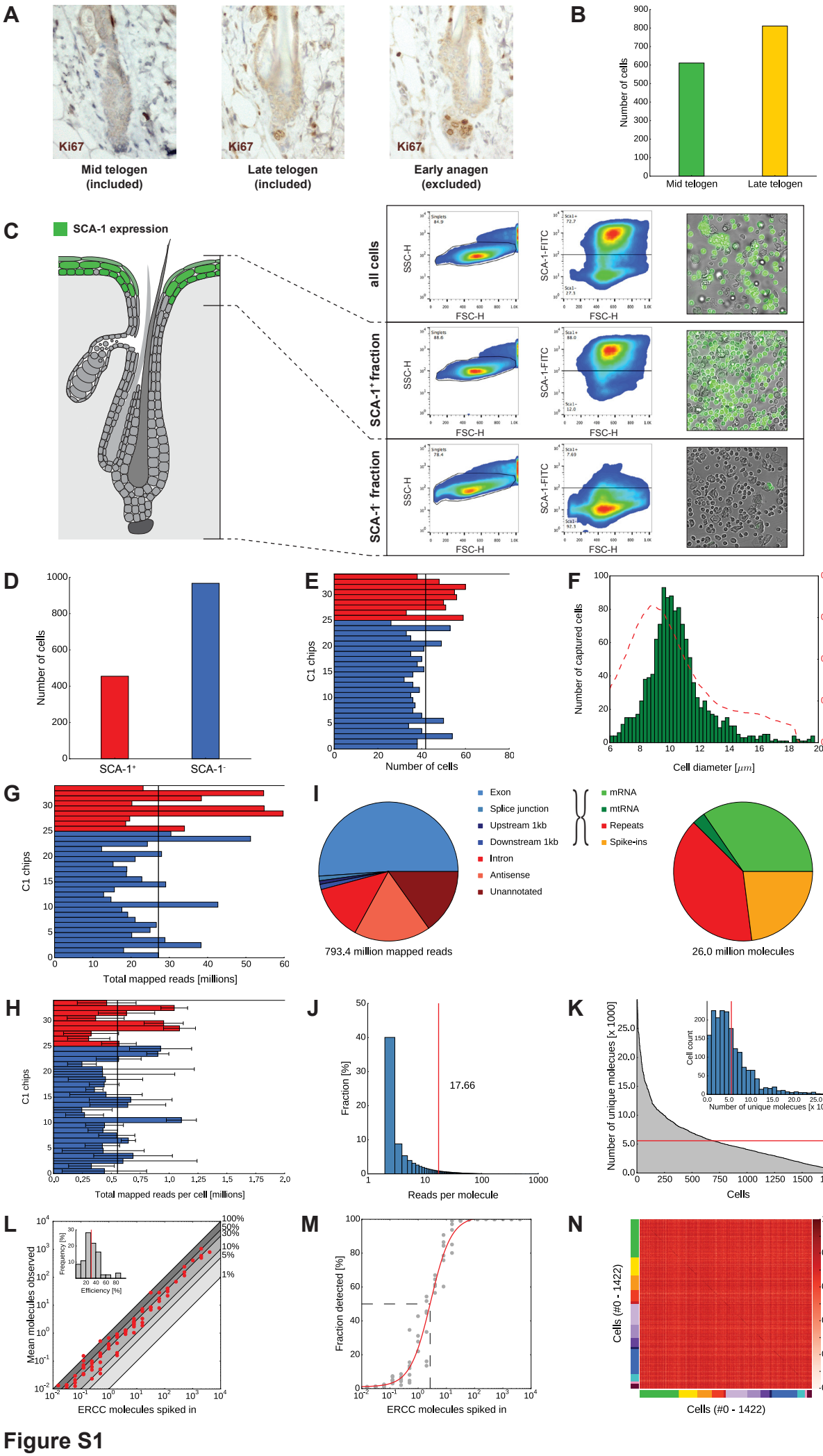
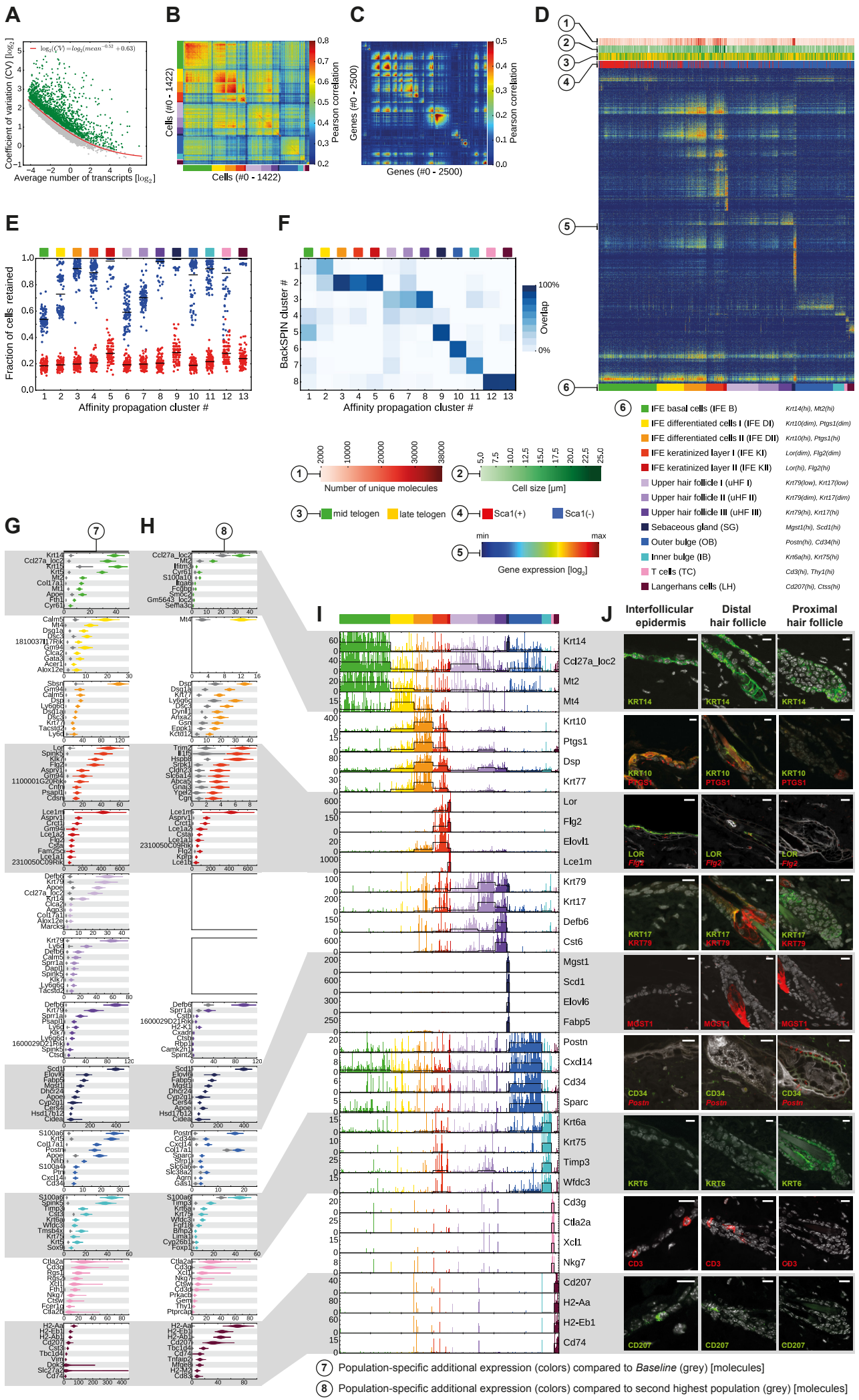

A
Ki67
Mid telogen (included)
Late telogen (included)
Early anagen (excluded)
B
Number of cells
Mid telogen
Late telogen
C
SCA-1 expression
all cells
SCA-1+ fraction
SCA-1- fraction
SSC-H
SCA-1-FITC
FSC-H
D
Number of cells
SCA-1+
SCA-1-
E
C1 chips
Number of cells
F
Number of captured cells
Size distribution of cell input
Cell diameter [μm]
G
C1 chips
Total mapped reads [millions]
I
Exon
Splice junction
Upstream 1kb
Downstream 1kb
Intron
Antisense
Unannotated
mRNA
mtRNA
Repeats
Spike-ins
793.4 million mapped reads
26.0 million molecules
H
C1 chips
Total mapped reads per cell [millions]
J
Fraction [%]
17.66
Reads per molecule
K
Number of unique molecules [x 1000]
Cell count
Number of unique molecules [x 1000]
Cells
L
Mean molecules observed
Frequency [%]
Efficiency [%]
ERCC molecules spiked in
M
Fraction detected [%]
ERCC molecules spiked in
N
Cells (#0 - 1422)
Pearson correlation
Cells (#0 - 1422)

Figure S1

Figure S1. Cell isolation and technical performance. Related to STAR Methods.

(**A**) Representative pictures of Ki67 immunostainings used to stage the dorsal epidermis of all experimental mice. Mice in mid or late telogen (1st and 2nd panel) were included in the study, while mice in anagen (3rd panel) were excluded.
(**B**) Number of captured cells from experimental mice in mid (green bar) and late telogen (yellow bar) that were included in the dataset.
(**C**) Separation of SCA-1+ and SCA-1− cells using microbeads. Left side: illustration of murine telogen epidermis including HF and SG showing the expected expression pattern of SCA-1. Right side: flow cytometry and immunomicroscopy images of epidermal cell suspensions prior to ('all cells') and after ('SCA-1+ fraction', 'SCA-1− fraction') cell separation with SCA-1 microbeads.
(**D**) Number of SCA-1+ (red bar) and SCA-1− cells (blue bar) included in the dataset.
(**E**) Number of cells passing the quality control per C1 chip. 34 C1 chips were sequenced for this study, and in total 1,422 single-cell transcriptomes passed the quality control. Red and blue bars signify chips loaded with either SCA-1+ or SCA-1− cells. Black line: mean over all C1 chips.
(**F**) Size distribution of captured cells included in the dataset (green bars) compared to input cell suspension (red line). Size of captured cells was determined based on the cell area in the microphotographs of cells in the C1 chip. Size distribution of input cell suspensions was measured using a Millipore Scepter Cell Counter and averaged over all experiments. While the single-cell capturing exhibits a minor bias towards larger cells, single cells in the dataset represent the whole size range of the cell input.
(**G–H**) Total mapped reads (G) and total mapped reads per cell (H) for cells passing quality criteria in each sequenced C1 chip. Red and blue bars signify chips loaded with either SCA-1+ or SCA-1− cells. Black lines denote the mean over all C1 chips; error bars in (H) show the standard deviation between individual cells.
(**I**) Census of all reads included in the data based on their alignment to the genome (left side) and census of all unique mRNA molecules based on their class (right side).
(**J**) Number of sequenced reads per molecule (unique molecular identifiers [UMI]). Red line: mean value.
(**K**) Number of unique mRNA molecules (RNA spike-ins and repeats excluded) sequenced from every cell of the initial dataset. Cells were ordered according to number of unique mRNA molecules from highest (38,000 unique molecules) to lowest and cells with less than 2,000 unique molecules were excluded (leaving 1,422 cells in the final dataset). Inset: histogram of cells according to mRNA yield. Red lines represent the average number of unique mRNA molecules over all cells.
(**L**) Efficiency of RNA spike-in detection. For each ERCC spike-in species, the number of molecules added to the reaction is plotted against the average number of molecules detected over all 1,422 cells included in the dataset. The diagonal lines demarcate the efficiency boundaries. Inset: histogram aggregating the detection efficiency of each ERCC spike-in species. Red line: median value.
(**M**) RNA spike-in detection limit. For each ERCC spike-in species, the number of molecules added to the reaction is plotted against the fraction of cells in the dataset in which the molecule was detected at least once. The red line represents a logistic curve fitted to the data. The dotted line marks the inferred minimal number of spike-in molecules necessary for detection in 50% of cases.
(**N**) Uniformity of single-cell cDNA synthesis reaction environment. All cells included in the dataset (ordered based on 1st level clustering) were correlated according to their ERCC spike-in data.



(7) Population-specific additional expression (colors) compared to *Baseline* (grey) [molecules]

(8) Population-specific additional expression (colors) compared to second highest population (grey) [molecules]

Figure S2

Figure S2. First-level clustering of epidermal cells. Related to Figure 1.

(**A**) Selection of genes with high variance for unsupervised clustering using a negative binomial noise model. Scatter represents average number of transcripts (log_2) versus coefficient of variation (CV; log_2) for all genes with average number of transcripts > 0.05 and at least five highly correlated neighbors (7,123 genes). The red line denotes the expected CV as a function of transcript mean according to our fitted noise model $log_2(CV) = log_2(mean^{alpha} + k)$ with alpha = -0.52 and k = 0.63. The 2,500 genes with the largest difference between expected and observed CV are colored in green and were used for 1st level clustering.
(**B–C**) Unsupervised clustering of cells (B) and genes (C). Pearson correlation of cells or genes was used as input for clustering with affinity propagation (AP) and Ward's linkage was subsequently used to define in-cluster order (see STAR Methods). Ordered Pearson correlation matrices of 1,422 cells (B) and 2,500 genes (C) with different color scale cut-offs (right panels) are shown. Group membership for cells in (B**)** is marked in left- and bottom-panels.
(**D**) Heatmap showing the expression of 2,500 genes (rows) in 1,422 single cells (columns). Cells and genes are clustered as in (B) and (C), respectively. Group membership of cells is color-coded in the bottom panel and explained in Figure 1C. The four top panels show cell-specific metadata: number of unique mRNA molecules and size of every cell are visualized in the first and second panel while telogen stage and SCA-1 expression are categorized in the third and fourth panel.
(**E**) Robustness of 1st level clustering was evaluated by resampling (100 iterations) of the dataset and randomly excluding 25% of all cells per iteration. Each subset was reclustered, and the percentage of cells from each cell population that were assigned to the same group was determined (blue dots). The red dots represent the percentage of cells from each group that end up together by pure chance after permutation of cell labels. The black lines show the group means.
(**F**) Comparison of affinity propagation (AP) clustering and unsupervised clustering with backSPIN. Shown is the relative distribution of cells within each AP cluster over all clusters defined by backSPIN.
(**G**) Identification of genes that are most highly expressed over *Baseline* in each population based on negative binomial regression of 1st level clustering data. For each population, the ten genes whose population-specific expression coefficient exceeds the *Baseline* coefficient with 99.9% posterior probability and who show the largest gap to the *Baseline* (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the *Baseline*) are reported. The gray and colored violin plots show the posterior probability distribution of the *Baseline* and population-specific coefficients respectively (scale in molecules).
(**H**) Identification of genes that are most highly uniquely expressed in each population based on negative binomial regression of 1st level clustering data. For each population, the ten genes whose population-specific expression coefficient exceeds the *Baseline* and all other populations-specific coefficients with 99.9% posterior probability and who show the largest gap to the second highest coefficient (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the second highest coefficient) are reported. The gray and colored violin plots show the posterior probability distribution of the second highest and population-specific coefficients respectively (scale in molecules).
(**I**) Barplots showing the absolute expression of selected marker genes in each cell. Cells are ordered into groups according to the clustering in (D). Group membership is color-coded in the upper panel and explained in Figure 1C. Black lines show the average expression over each group.
(**J**) Remapping of cell populations according to marker gene expression. For every group, either one or two marker genes were selected and antibodies and single molecule FISH probes (gene names in italics) were used to determine their spatial expression pattern in telogen skin. Counterstaining displayed in gray: DAPI (nuclei) or WGA (cell membranes). Scale bars, 10μm.

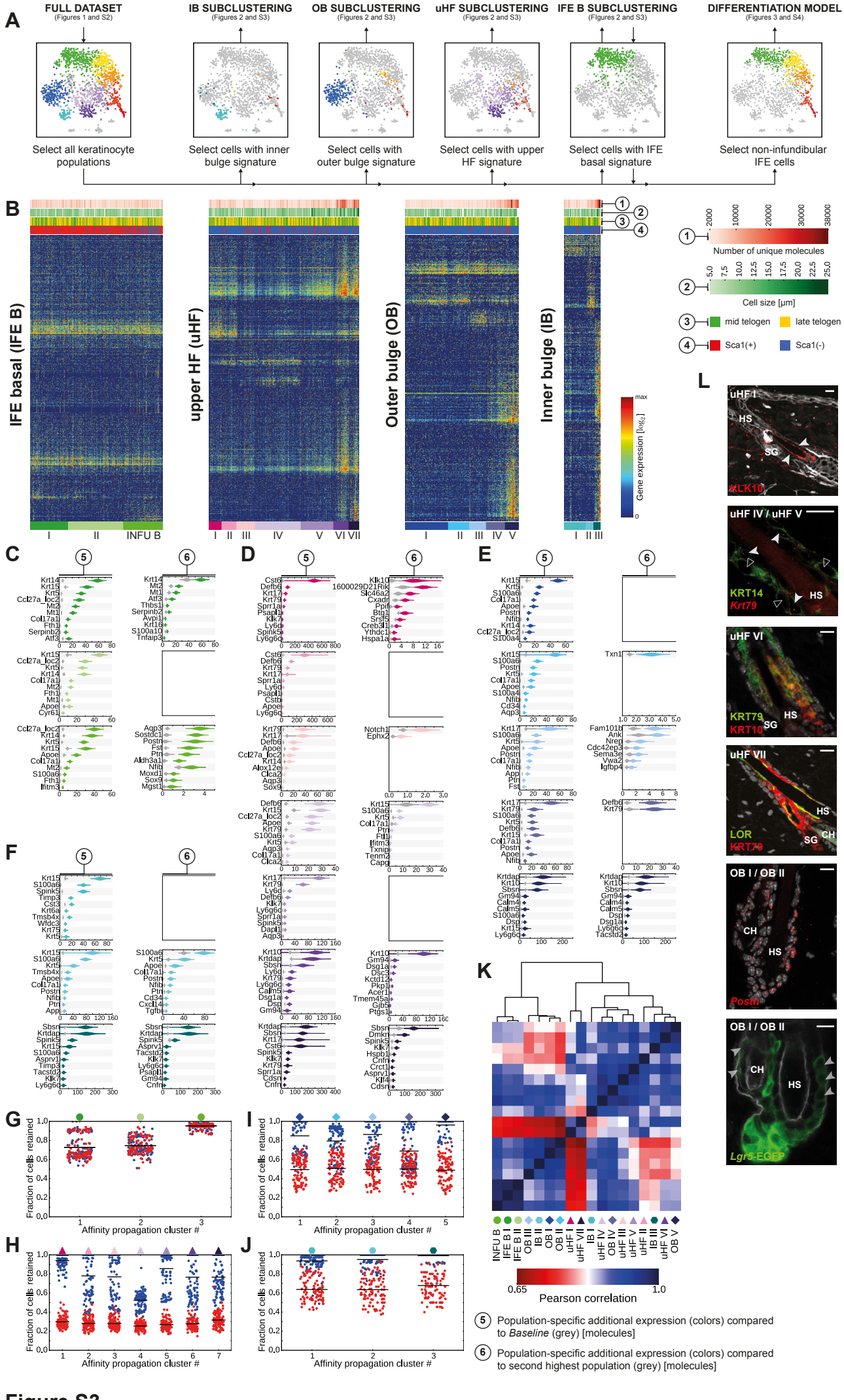
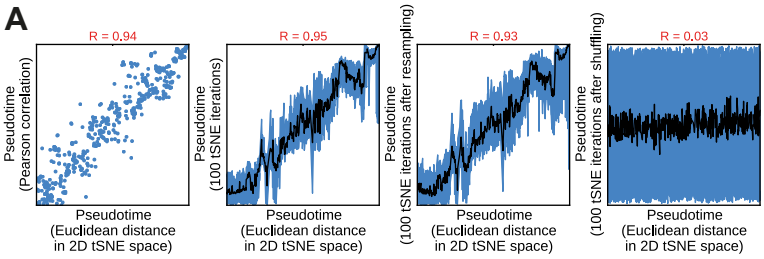
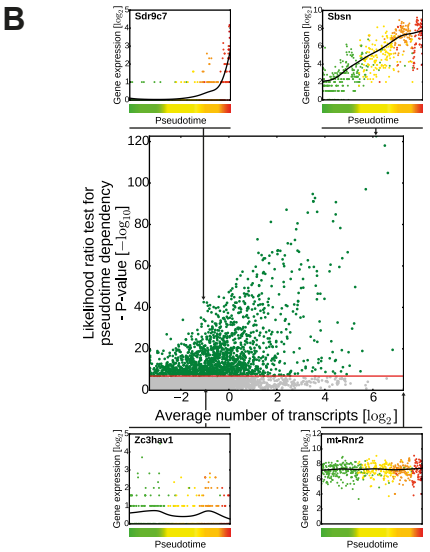

A
FULL DATASET
(Figures 1 and S2)
IB SUBCLUSTERING
(Figures 2 and S3)
OB SUBCLUSTERING
(Figures 2 and S3)
uHF SUBCLUSTERING
(Figures 2 and S3)
IFE B SUBCLUSTERING
(Figures 2 and S3)
DIFFERENTIATION MODEL
(Figures 3 and S4)
Select all keratinocyte populations
Select cells with inner bulge signature
Select cells with outer bulge signature
Select cells with upper HF signature
Select cells with IFE basal signature
Select non-infundibular IFE cells
B
IFE basal (IFE B)
upper HF (uHF)
Outer bulge (OB)
Inner bulge (IB)
I II INFU B
I II III IV V VI VII
I II III IV V
I II III
Gene expression [log2]
max
0
Number of unique molecules
Cell size [µm]
mid telogen
late telogen
Sca1(+)
Sca1(-)
C
D
E
F
G
H
I
J
Fraction of cells retained
Affinity propagation cluster #
K
INFU B
IFE B I
IFE B II
OB III
IB II
OB I
OB II
uHF I
uHF VII
IB I
uHF IV
OB IV
uHF III
uHF V
IB III
uHF VI
OB V
0.65
1.0
Pearson correlation
L
uHF I
HS
SG
XLK10
uHF IV / uHF V
KRT14
Krt79
HS
uHF VI
KRT79
KRT10
HS
SG
uHF VII
LOR
KRT79
HS
SG
CH
OB I / OB II
CH
HS
Postn
OB I / OB II
CH
HS
Lgr5-EGFP
5 Population-specific additional expression (colors) compared to Baseline (grey) [molecules]
6 Population-specific additional expression (colors) compared to second highest population (grey) [molecules]

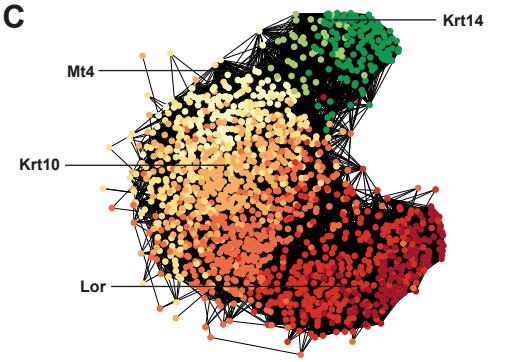
Figure S3

Figure S3. Second-level clustering of epidermal cells. Related to Figure 2.

(**A**) Summary of the cell reselection approach for 2nd level clustering. All cells from the dataset, excluding the sebaceous gland and immune cells, were selected and cells with inner bulge, outer bulge, upper HF and IFE basal signatures were chosen in order of primacy. For pseudotemporal ordering, IFE basal cells, excluding infundibular cells, were combined with the remaining IFE cells (intermediate, mature and terminally differentiated).
(**B**) Heatmaps showing the subclustering of IFE basal, upper HF, outer, and inner bulge cells. Group membership of cells is color-coded in the bottom panel and explained in Figures 2F–2G. The four top panels show cell-specific metadata: number of unique mRNA molecules and size of every cell are visualized in the first and second panel while telogen stage and SCA-1 expression are categorized in the third and fourth panel.
(**C–F**) Identification of genes that are most highly expressed over *Baseline* and most highly uniquely expressed in each IFE basal (C), upper HF (D), outer bulge (E) and inner bulge (F) subpopulation based on negative binomial regression of 2nd level clustering data. Left panels: for each subpopulation, the ten genes whose population-specific expression coefficient exceeds the *Baseline* coefficient with 95% posterior probability and who show the largest gap to the *Baseline* (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the *Baseline*) are reported. The gray and colored violin plots show the posterior probability distribution of the *Baseline* and population-specific coefficients respectively (scale in molecules). Right panels: for each subpopulation, the ten genes whose subpopulation-specific expression coefficient exceeds the *Baseline* and all other subpopulations-specific coefficients (limited to either the subpopulations of the IFE basal, uHF, OB or IB) with 95% posterior probability and who show the largest gap to the second highest coefficient (difference between the 25th percentile of the subpopulation-specific coefficient and the 75th percentile of the second highest coefficient) are reported. The gray and colored violin plots show the posterior probability distribution of the second highest and subpopulation-specific coefficients respectively (scale in molecules).
(**G–J**) Robustness of IFE basal (G), upper HF (H), outer bulge (I) and inner bulge (J) clustering was evaluated by resampling (100 iterations) of the dataset and randomly excluding 25% of all cells per iteration. Each subset was reclustered, and the percentage of cells from each cell population that were assigned to the same group was determined (blue dots). The red dots represent the percentage of cells from each group that end up together by pure chance after permutation of cell labels. The black lines show the group means.
(**K**) Transcriptomic similarity of 2nd level subclusters visualized by Ward's linkage hierarchical clustering of single-cell gene expression data averaged over each group.
(**L**) Remapping of subpopulations to their spatial location in the epidermis by immunostaining or single molecule FISH (gene symbols in italics). A summary of the populations' spatial localization can be found in Figure 2G. Arrowheads highlight the positions of the populations. uHF IV (empty arrowhead) / uHF V (filled arrowhead). OB II (arrowhead marks Lgr5(dim) cells in *Lgr5-EGFP-Ires-CreERT2* mice using anti-EGFP staining). HS, hair shaft. SG, sebaceous gland. CH, club hair. Scale bars, 10μm.

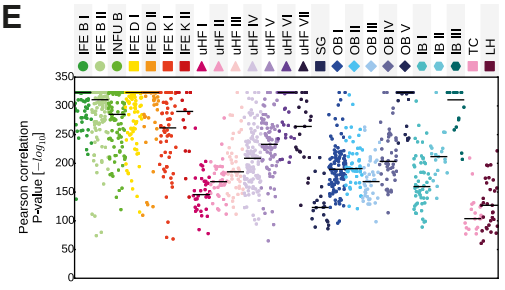


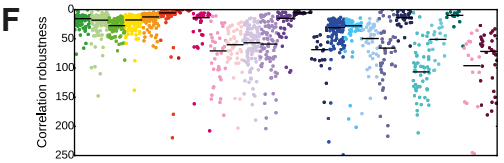




D

	Pseudotime	Network	TFs	Mol. signatures
I	Pseudotime		Bhlhe40, Snai2, Tsc22d1, Klf9	• Extracellular matrix • Focal adhesion • ECM-receptor interaction • Tissue development
II	Pseudotime		Sox6, Trp63, Zfp36l2	• Cellular component assembly • Protein transport • Chloride channel activity
III	Pseudotime		Hes1, Gata3, Pou3f1, Tsc22d3, Ikzf2	• Lipid metabolic process • Ribosome • Cell development • Apoptosis
IV	Pseudotime		Phb, Aff4, Tet3, Zmat3, Irf6, Srebf2, Nfatc2, Creb3, Ppp1r13l	• Steroid / cholesterol biosynthesis • Immune system • Adherens junction • Tight junction • Nectin pathway
V	Pseudotime		Cbfb, Pou2f3, Tcf25, Kdm5b, Zfp750, Mta3, Grhl3, Cebpa, Zbtb7a, Nfe2l1, Tsc22d4, Zfp91, Casz1, Ovol1, Zc3h15'	• Protein transport • Secretory pathway • Wnt signaling • Biopolymer metabolic process • Immune system
VI	Pseudotime		Klf4, Mafk, Nfat5, Bach1, Rora, Jarid2, Mllt4, Klf3, Xbp1, Grhl1, Hivep1, Lrrfip2, Rel, Egr2, Hopx, Sp6, Maff, Csrnp2, Cebpb	• Protein metabolic process / Proteolysis • Signal transduction • Lipid metabolism • Tissue development • Protein transport
VII	Pseudotime		Mxi1, Tead1, Rorc, Zfp706, Max, Hbp1, Tsc22d2, Mxd1, Prdm1	• Proteolysis • Endocytosis • MAPK signaling pathway • Epidermis development • Met pathway
VIII	Pseudotime		Id4	• Sphingolipid metabolism • MAPK signaling pathway • Cytokine / cytokine receptor interaction





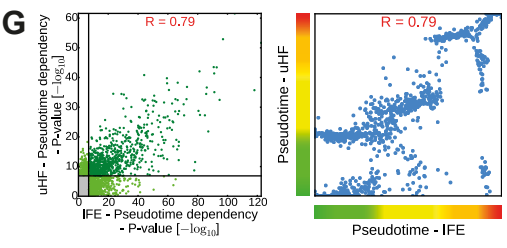


Figure S4

Figure S4. Modeling of the epidermal differentiation process. Related to Figure 3.

(**A**) Robustness of pseudotemporal ordering. Far left panel: comparison of the pseudotemporal ordering selected for Figures 3 and S4 (x-axis) to a pseudotemporal ordering acquired without dimensional reduction through t-SNE (y-axis). Center left panel: comparison of the selected pseudotemporal ordering (x-axis) to one hundred randomly acquired alternative orderings based on different initial t-SNE plots (y-axis). Center right panel: comparison of the selected pseudotemporal ordering (x-axis) to one hundred alternative orderings after randomly removing 25% of cells (y-axis). Far right panel: comparison of the selected pseudotemporal ordering (x-axis) to one hundred alternative orderings after shuffling cell labels (y-axis). The black line in the last three plots shows the median position of each cell over all one hundred iterations, while the blue areas cover the range between the 5^{th} and 95^{th} percentile.
(**B**) Center: average expression of 7,345 genes expressed during IFE differentiation plotted against pseudotime-dependency. Pseudotime-dependency was tested against a pseudotime-independent restricted model using an approximate likelihood ratio test (see STAR Methods) and the p-values are reported. The 1,627 genes (green) with p-values below a Bonferroni-corrected significance threshold of 0.001 (red line) were used for further analysis. Upper/lower panels: examples of low and high expressed genes with or without pseudotime-dependency, respectively.
(**C**) Shared nearest neighbor network of 1,627 pseudotime-dependent genes. Genes are colored according to group membership as established in Figure 3C.
(**D**) Characteristics of different subgroups of differentiation-related genes as defined in Figure 3C. Pseudotime: averaged expression of subgroup-specific genes over pseudotime. The red line shows the median while the gray lines demarcate the 25^{th} and 75^{th} percentile. Network: position of subgroup-specific genes in the shared nearest neighbor network established in (C). TFs: transcription factors included in each subgroup of genes. Mol. signatures: molecular and functional signatures linked to each subgroup of genes.
(**E**) P-values corresponding to the best correlation (highest correlation coefficient) of each cell to the pseudotime model as shown in Figure 3F. Cells are grouped according to (sub) population membership as defined by 2^{nd} level clustering. Black lines denote the median over each group.
(**F**) Robustness of each cell's correlation to the pseudotime model. To measure robustness of correlation, cells were re-correlated to the pseudotime model for one hundred times after randomly removing 75% of pseudotime-dependent genes. Shown is the average distance between a cell's pseudotime position in the full model and its position in the re-correlations. It is assumed that a small average distance is indicative of a more robust link to a particular stage in the pseudotime model.
(**G**) Comparison of pseudotemporal ordering of cells in the IFE and the uHF. Left panel: comparison of differentiation-dependency of genes involved in IFE and uHF differentiation. Right panel: comparison of the pseudotime-positions of epidermal cells derived from an IFE- and uHF-based model of differentiation.

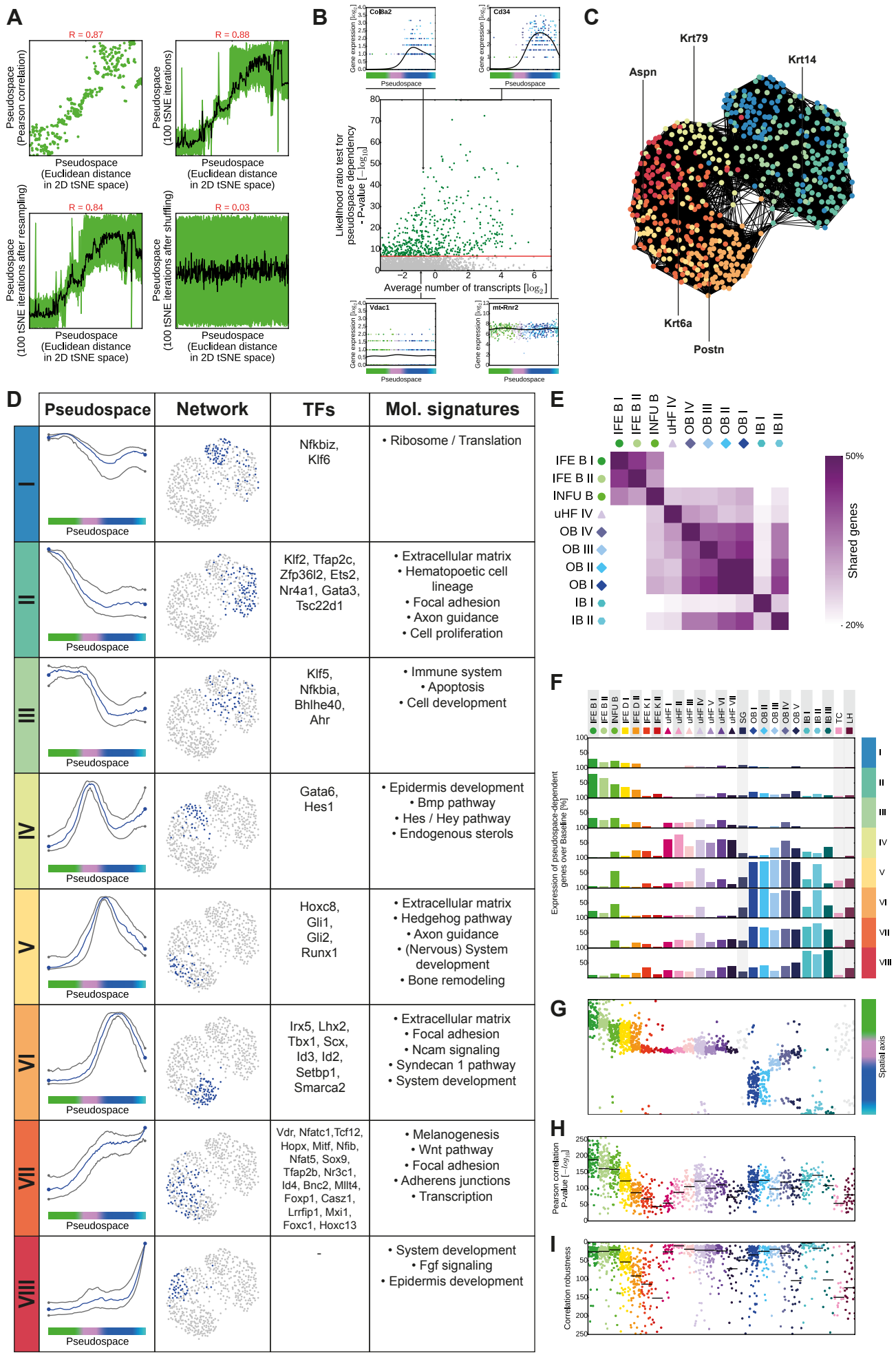

A
R = 0.87
R = 0.88
R = 0.84
R = 0.03
Pseudospace (Pearson correlation)
Pseudospace (100 tSNE iterations)
Pseudospace (100 tSNE iterations after resampling)
Pseudospace (100 tSNE iterations after shuffling)
Pseudospace (Euclidean distance in 2D tSNE space)
B
Col8a2
Cd34
Vdac1
mt-Rnr2
Gene expression [log2]
Pseudospace
Likelihood ratio test for pseudospace dependency - P-value [−log10]
Average number of transcripts [log2]
C
Krt79
Krt14
Aspn
Krt6a
Postn
D
Pseudospace
Network
TFs
Mol. signatures
I
Nfkbiz, Klf6
• Ribosome / Translation
II
Klf2, Tfap2c, Zfp36l2, Ets2, Nr4a1, Gata3, Tsc22d1
• Extracellular matrix
• Hematopoetic cell lineage
• Focal adhesion
• Axon guidance
• Cell proliferation
III
Klf5, Nfkbia, Bhlhe40, Ahr
• Immune system
• Apoptosis
• Cell development
IV
Gata6, Hes1
• Epidermis development
• Bmp pathway
• Hes / Hey pathway
• Endogenous sterols
V
Hoxc8, Gli1, Gli2, Runx1
• Extracellular matrix
• Hedgehog pathway
• Axon guidance
• (Nervous) System development
• Bone remodeling
VI
Irx5, Lhx2, Tbx1, Scx, Id3, Id2, Setbp1, Smarca2
• Extracellular matrix
• Focal adhesion
• Ncam signaling
• Syndecan 1 pathway
• System development
VII
Vdr, Nfatc1,Tcf12, Hopx, Mitf, Nfib, Nfat5, Sox9, Tfap2b, Nr3c1, Id4, Bnc2, Mllt4, Foxp1, Casz1, Lrrfip1, Mxi1, Foxc1, Hoxc13
• Melanogenesis
• Wnt pathway
• Focal adhesion
• Adherens junctions
• Transcription
VIII
-
• System development
• Fgf signaling
• Epidermis development
E
IFE B I
IFE B II
INFU B
uHF IV
OB IV
OB III
OB II
OB I
IB I
IB II
Shared genes
50%
20%
F
Expression of pseudospace-dependent genes over Baseline [%]
G
Spatial axis
H
Pearson correlation P-value [−log10]
I
Correlation robustness

Figure S5

Figure S5. Modeling of spatial gene expression signatures. Related to Figure 4.

(**A**) Robustness of pseudospatial ordering. Upper left panel: comparison of the pseudospatial ordering selected for Figures 4 and S5 (x-axis) to a pseudospatial ordering acquired without dimensional reduction through t-SNE (y-axis). Upper right panel: comparison of the selected pseudospatial ordering (x-axis) to one hundred randomly acquired alternative orderings based on different initial t-SNE plots (y-axis). Lower left panel: comparison of the selected pseudospatial ordering (x-axis) to one hundred alternative orderings after randomly removing 25% of cells (y-axis). Lower right panel: comparison of the selected pseudospatial ordering (x-axis) to one hundred alternative orderings after shuffling cell labels (y-axis). The black line in the last three plots shows the median position of each cell over all one hundred iterations, while the green areas cover the range between the 5^{th} and 95^{th} percentile.
(**B**) Center: average expression of 6,788 genes expressed in basal cells plotted against pseudospace-dependency. Pseudospace-dependency was tested against a restricted model using an approximate likelihood ratio test (see STAR Methods) and the p-values are reported. The 547 genes (green) with p-values below a Bonferroni-corrected significance threshold of 0.001 (red line) were used for further analysis. Upper/lower panels: examples of low and high expressed genes with or without pseudospace-dependency, respectively.
(**C**) Shared nearest neighbor network of 547 pseudospace-dependent genes. Genes are colored according to group membership as established in Figure 4C.
(**D**) Characteristics of different subgroups of spatial genes as defined in Figure 4C. Pseudospace: averaged expression of subgroup-specific genes over pseudospace. The blue line shows the median while the gray lines demarcate the 25^{th} and 75^{th} percentile. Network: position of subgroup-specific genes in the shared nearest-neighbor network established in (C). TFs: transcription factors included in each subgroup of genes. Mol. signatures: molecular and functional signatures linked to each subgroup of genes.
(**E**) Overlap of genes expressed over *Baseline* in basal-cell populations (and IB I). Genes were called from the negative binomial regression model of 2^{nd} level clustering if the population-specific regression coefficient exceeded *Baseline* with 95% posterior probability.
(**F**) Expression of spatial genes in all epidermal (sub) populations defined by either 1^{st} or 2^{nd} level clustering. A gene was considered expressed in a population if its population-specific coefficient in the negative binomial regression model exceeded *Baseline* with 95% posterior probability. Genes are ordered according to group membership introduced in Figure 4C. The shaded populations exhibit gene expression inconsistent with any distinct spatial signature.
(**G**) Position of epidermal cells from each population on the spatial axis as determined by highest Pearson correlation. The shaded cells belong to populations that show gene expression inconsistent with any spatial signature.
(**H**) P-values corresponding to the best correlation of each cell to the pseudospace model as shown in (G). Black lines denote the median over each group.
(**I**) Robustness of each cell's correlation to the pseudospace model. To measure robustness of correlation, cells were re-correlated to the spatial axis for one hundred times after randomly removing 75% of pseudospace-dependent genes. Shown is the average distance between a cell's pseudospace position in the full model and its position in the re-correlations. It is assumed that a small average distance is indicative of a more robust link to a particular stage in the pseudospace model.

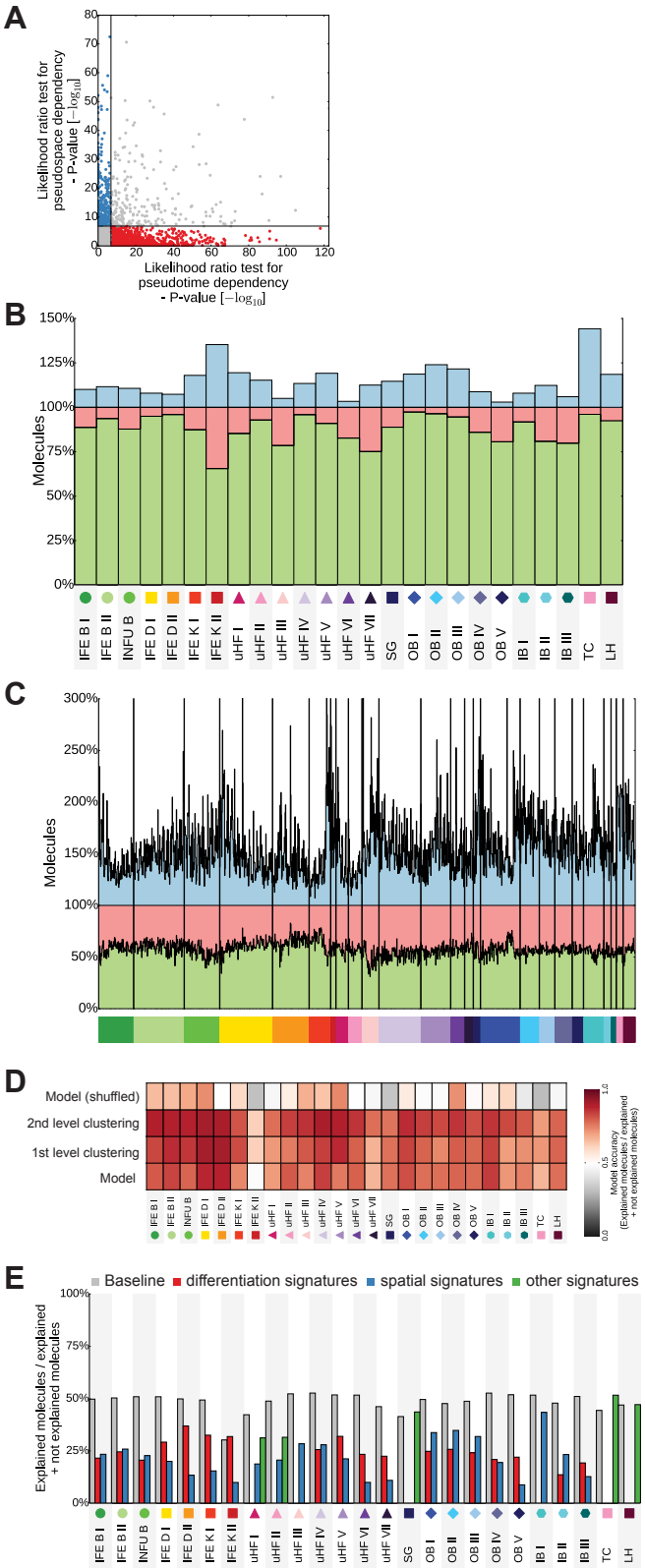

A
Likelihood ratio test for pseudospace dependency - P-value [−log10]
Likelihood ratio test for pseudotime dependency - P-value [−log10]
B
Molecules
C
Molecules
D
Model (shuffled)
2nd level clustering
1st level clustering
Model
Model accuracy (Explained molecules / explained + not explained molecules)
E
Baseline
differentiation signatures
spatial signatures
other signatures
Explained molecules / explained + not explained molecules
IFE B I
IFE B II
INFU B
IFE D I
IFE D II
IFE K I
IFE K II
uHF I
uHF II
uHF III
uHF IV
uHF V
uHF VI
uHF VII
SG
OB I
OB II
OB III
OB IV
OB V
IB I
IB II
IB III
TC
LH

Figure S6

Figure S6. Explaining cellular heterogeneity using differentiation and spatial signatures. Related to Figure 5.

(**A**) Pseudotime- and pseudospace-dependency of genes. Black lines mark the Bonferroni-corrected significance threshold of 0.001. Of 7,893 genes, 1,409 were uniquely pseudotime-, 329 uniquely pseudospace- and 218 both pseudotime- and pseudospace-dependent. Only the uniquely pseudotime- or pseudospace-dependent genes were considered in the pseudospacetime model.
(**B**) Percentage of molecules explained (green), underexplained (red) or overexplained (blue) by the pseudospacetime model. Molecules were pooled across the cells per population.
(**C**) Percentage of molecules explained (green), underexplained (red) or overexplained (blue) by the pseudospacetime model in each single cell per population.
(**D**) Accuracy of pseudospacetime, 1^{st} level clustering, 2^{nd} level clustering and shuffled pseudospacetime model stratified according to cell populations defined in either 1^{st} or 2^{nd} level clustering.
(**E**) Fraction of explained molecules contributed by *Baseline*, differentiation axis, spatial axis and other (sebaceous gland, immune) signatures for each cell population.

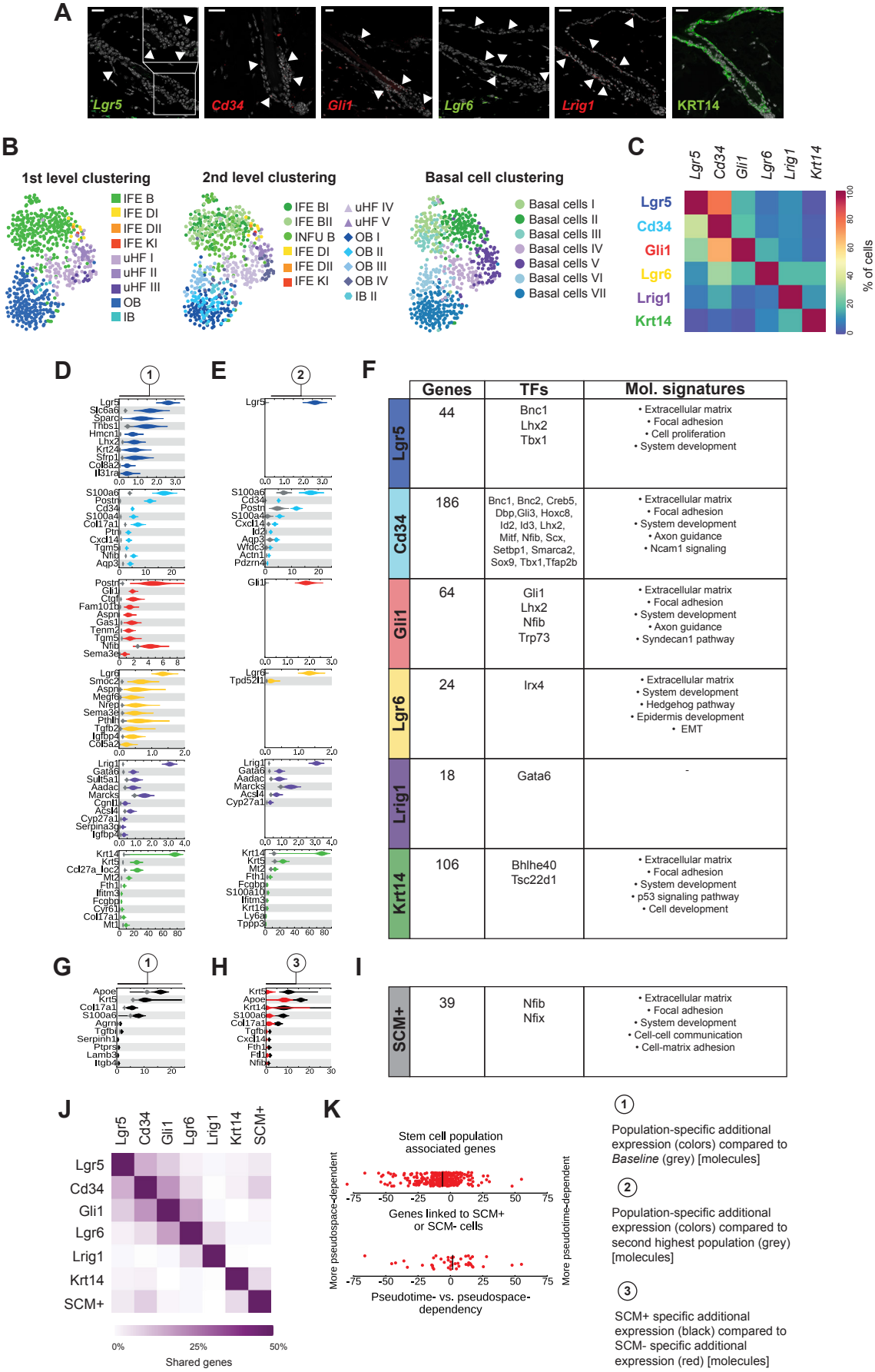

A
Lgr5
Cd34
Gli1
Lgr6
Lrig1
KRT14
B
1st level clustering
IFE B
IFE DI
IFE DII
IFE KI
uHF I
uHF II
uHF III
OB
IB
2nd level clustering
IFE BI
IFE BII
INFU B
IFE DI
IFE DII
IFE KI
uHF IV
uHF V
OB I
OB II
OB III
OB IV
IB II
Basal cell clustering
Basal cells I
Basal cells II
Basal cells III
Basal cells IV
Basal cells V
Basal cells VI
Basal cells VII
C
Lgr5
Cd34
Gli1
Lgr6
Lrig1
Krt14
% of cells
D
E
F
Genes
TFs
Mol. signatures
Lgr5
44
Bnc1
Lhx2
Tbx1
• Extracellular matrix
• Focal adhesion
• Cell proliferation
• System development
Cd34
186
Bnc1, Bnc2, Creb5, Dbp,Gli3, Hoxc8, Id2, Id3, Lhx2, Mitf, Nfib, Scx, Setbp1, Smarca2, Sox9, Tbx1,Tfap2b
• Extracellular matrix
• Focal adhesion
• System development
• Axon guidance
• Ncam1 signaling
Gli1
64
Gli1
Lhx2
Nfib
Trp73
• Extracellular matrix
• Focal adhesion
• System development
• Axon guidance
• Syndecan1 pathway
Lgr6
24
Irx4
• Extracellular matrix
• System development
• Hedgehog pathway
• Epidermis development
• EMT
Lrig1
18
Gata6
-
Krt14
106
Bhlhe40
Tsc22d1
• Extracellular matrix
• Focal adhesion
• System development
• p53 signaling pathway
• Cell development
G
H
I
SCM+
39
Nfib
Nfix
• Extracellular matrix
• Focal adhesion
• System development
• Cell-cell communication
• Cell-matrix adhesion
J
Shared genes
K
Stem cell population associated genes
Genes linked to SCM+ or SCM- cells
Pseudotime- vs. pseudospace-dependency
More pseudospace-dependent
More pseudotime-dependent
1
Population-specific additional expression (colors) compared to Baseline (grey) [molecules]
2
Population-specific additional expression (colors) compared to second highest population (grey) [molecules]
3
SCM+ specific additional expression (black) compared to SCM- specific additional expression (red) [molecules]

Figure S7

Figure S7. Cellular heterogeneity of stem cell populations. Related to Figure 6.

(**A**) Immunostaining and single molecule FISH (gene symbols in italics) of SCMs in epidermis. Note that most markers are expressed in several epidermal compartments. Scale bars, 20μm.
(**B**) Cells projected onto the t-SNE map of basal cells (see Figure 6B), colored according to 1st level, 2nd level and selective clustering of basal cells.
(**C**) Matrix showing the overlap in SCM expression. Percentage of cells expressing each SCM *Lgr5*, *Cd34*, *Gli1*, *Lgr6*, *Lrig1,* or *Krt14* (rows) co-expressing additional SCM genes (columns).
(**D–F**) Analyses of gene expression signatures for SCM-expressing populations. (D) Identification of the top ten genes that were most highly expressed over *Baseline* in each stem cell population based on negative binomial regression. For each population, the ten genes whose population-specific expression coefficient exceeds the *Baseline* coefficient with 95% posterior probability and which show the largest gap to the *Baseline* (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the *Baseline*) are reported. The gray and colored violin plots show the posterior probability distribution of the *Baseline* and population-specific coefficients, respectively (scale in molecules). (E) Identification of genes that are most highly and uniquely expressed in each stem cell population based on negative binomial regression. For each population, the ten genes whose population-specific expression coefficient exceeds the *Baseline* and all other populations-specific coefficients with 95% posterior probability and that show the largest gap to the second highest coefficient (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the second highest coefficient) are reported. The gray and colored violin plots show the posterior probability distribution of the second highest and population-specific coefficients, respectively (scale in molecules). (F) Characteristics of different stem cell populations. Genes: number of genes expressed over *Baseline* with 95% posterior probability. TFs: transcription factors included in each set of genes. Mol. signatures: molecular and functional signatures linked to each subgroup of genes.
(**G–I**) Analyses of shared gene expression signatures in SCM+ cells. (G) Identification of the top ten genes that were most highly expressed over *Baseline* among all SCM+ cells. In contrast to (D), a 90% posterior probability cut-off was chosen. (H) Identification of the top ten genes expressed in all SCM+ cells that were most highly expressed compared to SCM− cells. A 90% posterior probability cut-off was chosen and only genes whose SCM+ cell-specific expression exceeded 0.25 molecules were chosen. The black and red violin plots show the posterior probability distribution of the SCM+ and SCM− specific coefficients respectively. (I) Characteristics of the SCM+ population. Genes: number of genes with higher expression than in the SCM− population based on a 90% posterior probability cut-off. TFs: transcription factors. Mol. signatures: molecular and functional signatures.
(**J**) Percentage of shared genes between the specific signatures of *Lgr5*+, *Cd34*+, *Gli1*+, *Lgr6*+, *Lrig1*+, *Krt14*^{hi}, and SCM+ cells. The specific signatures were defined as specified in (D), (F), (G) and (I).
(**K**) Pseudotime- vs. pseudospace-dependency of stem cell-specific genes. Plotted is the difference between the $-\log_{10}$ transformed p-value of pseudotime- and pseudospace-dependency. "Stem cell population associated genes" include all genes, which are expressed over *Baseline* in at least one stem/progenitor population. "Genes linked to SCM+ or SCM– cells" are the 44 genes that are expressed differently between SCM+ and SCM– cells as specified in Figures 6G and S7G – S7I. The black lines denote the median.

SUPPLEMENTAL TABLES

Table S1. Marker genes – 1st level clustering. Related to Figure 1.

Lists of genes that are most highly expressed over *Baseline* (vs. *Baseline,* left column) or the second highest population (vs. other groups, right column). Identification of genes is based on negative binomial regression of 1st level clustering data. For each population, genes whose population-specific expression coefficient exceeded the *Baseline* coefficient (left column) or all other populations-specific coefficients (right column) with 99.9% posterior probability and which show the largest gap to the *Baseline* / the second highest population (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the *Baseline* / the second highest population) are listed.

(Supplied as Excel file: Joost Table S1.xlsx)

Table S2. Marker genes – 2nd level clustering. Related to Figure 2.

Lists of genes that are most highly expressed over *Baseline* (vs. *Baseline,* left column) or the second highest population (vs. other groups, right column). Identification of genes is based on negative binomial regression of 2nd level clustering data. For each population, genes whose population-specific expression coefficient exceeded the *Baseline* coefficient (left column) or all other populations-specific coefficients (right column) with 95% posterior probability and which show the largest gap to the *Baseline* / the second highest population (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the *Baseline* / the second highest population) are listed.

(Supplied as Excel file: Joost Table S2.xlsx)

Table S3. Description of cell populations and comparison to literature. Related to Figures 1, 2 and 6.

(**A**) Description of cell populations defined during 1st and 2nd level clustering. Described are the molecular and spatial characteristics of each population defined in this study and their relation to previous work.
(**B**) Previously defined murine epidermal (stem) cell populations and their relation to populations defined in this study.

A. Description of cell populations from 1st and 2nd level clustering

Populations	Molecular and spatial description in Joost et al.	Previous descriptions
Interfollicular basal I (IFE B I)	IFE basal cell population marked by *Thbs1* expression and higher than average expression of IFE basal genes such as *Krt14*, *Mt1* and *Mt2*. IFE B I cells are interspersed with IFE B II cells in the IFE basal layer.	A THBS1-positive subpopulation of the IFE has not been previously described. This population is not congruent with the *Ivl*+ population or *Lgr6*+ population described by Mascré et al., 2012 and Füllgrabe et al., 2015.
Interfollicular basal II (IFE B II)	IFE basal cell population marked by absence of *Thbs1* and lower than average expression of IFE basal genes such as *Krt14*, *Mt1*, *Mt2*. IFE B II cells are interspersed with IFE B I cells in the IFE basal layer.	A distinct THBS1-negative subpopulation of the IFE has not been previously described. This population is not congruent with the *Ivl*+ population or *Lgr6*+ population described by Mascré et al., 2012 and Füllgrabe et al., 2015.
Infundibular basal (INFU B)	A population of cells dominated by typical IFE basal markers such as *Krt14*, *Mt1* and *Mt2*, which additionally expresses pan and upper HF markers such as *Sostdc1*, *Aqp3*, *Postn* and *Krt79*. Located in the infundibular region close to the HF opening. The specific spatial position of this population is clearly identifiable by *Postn* expression.	While it has been shown that pan HF markers such as *Sostdc1* reach into the infundibular region (Collette et al., 2013), the particularly IFE basal character of this population in combination with low-level gradual expression of HF genes has not been clarified before.
IFE differentiated cells I (IFE D I)	Transient population of cells marked by low level expression of both (IFE) basal (*Krt14*, *Mt2*) and (IFE) suprabasal (*Krt10*, *Sbsn*) markers. Expresses *Mt4* as one of a few specific markers.	Although the existence of this basal-to-suprabasal transient population is expected according to the accepted model of epidermal differentiation, it has never been resolved at a transcriptional level.
IFE differentiated cells II (IFE D II)	Mature, suprabasal population expressing high levels of well-established spinous layer markers such as *Krt10* and *Sbsn*.	The cells of the spinous layer are well described and characterized. See for instance Fuchs, 1990.
IFE keratinized layer I (IFE K I)	A transient population of cells which is marked by decreasing levels of spinous layer markers such as *Krt10* and *Sbsn* (when compared to IFE D II) and increasing levels of granular layer markers including *Lor* and *Flg2*.	Although the existence of this suprabasal/spinous-to-terminal/granular transient population is expected according to the accepted model of epidermal differentiation, it has never been resolved at a transcriptional level.
IFE keratinized layer II (IFE K II)	A population of flat, keratinized cells expressing high levels of well-established granular layer markers such as *Lor* and *Flg2*.	The cells of the granular layer are well described and characterized. See for instance Fuchs, 1990.
Upper HF I (uHF I)	A population of cells marked by a typical uHF signature (*Krt79*, *Krt17*, *Cd44*) in combination with expression of a gene module distinguished by high *Rbp1*, *Defb6* and *Cst6* expression. Additionally epidermis- and compartment-unique markers such as *Klk10* and *Cryab* are expressed. Could be mapped to two rings of suprabasal cells above and below the SG opening based on highest CST6 expression. Accordingly, KLK10 is mostly found secreted into the hair canal at the corresponding positions.	Has not been previously described in molecular detail. Although Zeeuwen et al., 2002 and Veniaminova et al. 2013 described CST6 expression in the upper HF, they were unable to differentiate a set of uHF populations with strong CST6 expression and a set of uHF populations with weak or absent CST6 expression or subdivide those populations further. It is not clear whether this population contains cells of BLIMP1 / *Prdm1* population described by Horsley et al., 2006.
Upper HF II (uHF II)	A population of cells marked by a typical uHF signature (*Krt79*, *Krt17*, *Cd44*) in combination with high expression of a gene module distinguished by high *Rbp1*, *Defb6* and *Cst6* expression. This population is additionally distinguished by a *Krt14* / *Lrig1* signature while the *Krt5* / *Ptn* module is absent. Can be mapped to the SG opening based on the location of highest *Defb6* expression in the epidermis. Additionally, the absence of *Krt5* / *Ptn* expression in the population does not allow its location in the basal layer of the upper HF / junctional zone while the presence of *Krt14* argues against suprabasal localization (e.g. adjacent to uHF I).	Has not been previously described. Although it has been shown previously that *Lrig1*+ cells are located in different compartments of the skin including the uHF and the SG (Page et al., 2013), and KRT79 and CST6 are expressed in the uHF (Veniaminova et al., 2013; Zeeuwen et al., 2002), these populations were never dissected on a molecular / transcriptional level.

Upper HF III (uHF III)	Marked by an uHF signature (*Krt79*, *Krt17*, *Cd44*) and a *Krt14* / *Lrig1* module. However, in contrast to basal cells in the upper HF / junctional zone (uHF IV), *Krt5* / *Ptn* are absent. Can only be mapped based on exclusion. While the presence of *Krt14* / *Lrig1* indicated basal location, the absence of *Krt5* / *Ptn* did not allow mapping to the *Krt5*(hi) uHF / junctional zone basal layer. Absence of high *Defb6* / *Cst6* levels excluded localization in the SG opening while non-expression of typical SG markers such as *Scd1* / *Mgst1* excluded mapping to the proximal half of the SG. The population is in consequence most likely located in the distal half of SG.	See uHF II.
Upper HF IV (uHF IV)	This population expressed low levels of uHF markers (*Krt79*, *Krt17*, *Cd44*) in combination with both a *Krt14* / *Lrig1* and a *Krt5* / *Ptn* basal signature. Its low level expression of *Krt79* and *Krt17*, which could both be found expressed in small numbers in the basal layer of the uHF and SG, and its positivity for both *Krt14* and *Krt5* (which excludes the basal cells of the SG) linked this population to the basal layer of the uHF including isthmus.	See uHF II. This population most likely corresponds to the main population of Lrig1+ cells described by Jensen et al., 2009 and Page et al., 2013.
Upper HF V (uHF V)	A population with a strong uHF signature (*Krt79*, *Krt17*, *Cd44*), a fading basal signature (*Krt14* / *Lrig1* / *Krt5* / *Ptn*) and low-level expression of suprabasal markers such as *Krt10* / *Sbsn*. Most likely contains cells during their transition from uHF basal (uHF IV) to uHF suprabasal (uHF VI).	Differentiation in the uHF and the presence of cells expressing both uHF markers (e.g. *Krt79*, *Krt17*) and spinous cell (*Krt10*, *Sbsn*) or granular cell markers (*Lor*, *Flg2*) has been previously described (e.g. by Veniaminova et al., 2013)
Upper HF VI (uHF VI)	Cells distinguished by an uHF signature (*Krt79*, *Krt17*, *Cd44*) in combination with a high expression of suprabasal markers (*Krt10* / *Sbsn*). Located in the 2nd cell layer of the uHF excluding the suprabasal cells which express a *Cst6* / *Defb6* / *Rbp1* signature.	See uHF IV.
Upper HF VII (uHF VII)	Cells distinguished by an uHF signature (*Krt79*, *Krt17*, *Cd44*) in combination with a high expression of terminal markers (*Lor* / *Flg2*). Could be mapped to a 3rd layer of flat, keratinized cells which line the hair canal in the upper HF.	See uHF IV.
Sebaceous gland (SG)	A population of cells, which is very distinct from all other keratinocyte populations, and distinguished by a large signature of genes mainly involved in lipid metabolism. Marked by well-established SG markers such as *Scd1*. MGST1 staining and a relative large heterogeneity in cell size suggested that this population includes both cells from the proximal basal layer of the SG and the inner parts of the gland.	Well described.
Outer bulge I (OB I)	A cell population, which is dominated by an outer bulge signature (*Postn*, *Cd34*). In contrast to OB II, it expressed slightly higher levels of *Lgr5*, *Krt24*, lower levels of *Cd34* and contained more *Gli1* expressing cells. This suggests that the cells of this population tend to be located more proximally than the cells of OB II. Furthermore, OB I cells are limited to the bulge and the cell layer between club hair and bulge.	*Lgr5*^{hi} cells located in the proximal bulge area including the hair germ are well established (Jaks et al., 2008). Most intriguingly, our single-cell transcriptional data indicate that the differences between the proximal and distal bulge populations are small and that the transition between both populations is "fluid".
Outer bulge II (OB II)	The counterpart to OB I. The cells are likely to be located more distally in the outer bulge than the cells of OB I. Cells are spread over the bulge, the cell layer between club hair and bulge as well as the club hair.	See OB I.
Outer bulge III (OB III)	A population, which, in addition to a pan outer bulge signature (*Cd34*, *Postn*), was negative for more proximal outer bulge markers (*Lgr5*, *Krt24*) and instead expressed a unique signature including *Aspn* and *Nrep*. The expression of *Gli1*, *Lgr6* and *Krt17* and the low levels of *Krt15* link these cells to a particular position at the upper edge of the outer bulge region.	Most likely corresponding to the *Gli1*+ population in the isthmus described by Brownell et al. 2011.
Outer bulge IV (OB IV)	Transient population of cells, which expressed both outer bulge (*Cd34*, *Postn*) and upper HF (*Krt79* / *Krt17*) markers. Located at the interface of bulge and upper HF compartment distal of OB III and proximal of uHF IV. Interestingly, a subset of cells co-expressed *Lgr6*.	Although spatially congruent with the isthmus *Lgr6*+ population (Snippert et al., 2010), not previously described as a transient population exhibiting both outer bulge and uHF features.
Outer bulge V (OB V)	Outer bulge population (*Postn*, *Cd34*) additionally expressing differentiation markers including *Krt10* and *Sbsn*, that could be linked to a group of suprabasal cells at the level of OB III and OB IV.	Most likely the CD34+ / ITGA6- population identified by Blainpain et al., 2004.
Inner bulge I (IB I)	Population of cells that are solely distinguished by an inner bulge signature (*Krt6a*, *Krt75*, *Timp3*, *Fgf18*). Represent most of the cells located in the inner bulge.	Well-described epidermal population (Hsu et al., 2011).
Inner bulge II (IB II)	Cells, which expressed an outer bulge signature (*Postn*, Cd34) in combination with an inner bulge signature (*Krt6a*, *Krt75*, *Timp3*, *Fgf18*). This population could be mapped to a group of basal, KRT6^{hi} and *Postn*+ cells in the outer bulge region.	The separation of the club hair bulge into two populations has not been previously described.
Inner bulge III (IB III)	Cells expressing both an inner bulge signature (*Krt6a*, *Krt75*, *Timp3*, *Fgf18*) and differentiation markers. Could be located in the upper edge of the inner bulge.	Not previously described.
T cells (TC)	Population of immune cells. Their specific gene expression (*Cd3g*, *Cd3d*) distinguished them as γδ T cells.	Well described.
Langerhans cells (LH)	Immune cells that expressed typical Langerhans cell markers (*Cd207*, *Cd74*).	Well described.

B. Previously defined murine epidermal (stem) cell populations

References	Representation in Joost et al. dataset
Krt14+/Ivl- and Krt14+/Ivl+ IFE basal cells (Mascré et al., 2012)	While the infundibular (INFU B) and upper HF (uHF IV) basal cells showed a higher *Ivl* baseline expression, *Ivl* expression was rarely found in the IFE basal populations (IFE B I – II). Instead, *Ivl* expression in the IFE was predominantly detected in the differentiating and terminally differentiated populations. It is possible that Mascré et al., 2012 targeted cells, which are in the process of transition from basal to suprabasal and show both basal and suprabasal characteristics.
Spinous layer cells (Miscellaneous)	Cells that show a spinous layer signature could be found in the IFE (IFE DII), the upper HF (uHF VI), the outer bulge (OB V) and the inner bulge compartment (IB III).
Granular layer cells (Miscellaneous)	Cells distinguished by a granular layer signature are present in the IFE (IFE K I and II) and the upper HF compartment (uHF VII).
Lrig1+ cells located in the upper HF (Jensen et al., 2009, Page et al., 2013)	*Lrig1* was expressed in a variety of populations primarily located in the upper HF and SG including uHF I, uHF II, uHF III, and uHF IV. Although most highly expressed in the basal layer, *Lrig1* molecules were also sporadically detected in differentiated uHF cells. Furthermore, low-level sporadic *Lrig1* expression was found in the IFE basal layer and cells of the outer bulge.
Mts24+ cells located in the isthmus (Nijhof, 2006)	Expression of MTS24 (*1600029D21Rik*) could be detected in all populations of the upper HF and in terminally differentiated cells of the IFE. In the upper HF, *1600029D21Rik* expression peaked in the terminally differentiated populations (uHF VI and uHF VII) and in the *Defb6*^{hi}/*Cst6*^{hi} populations uHF I / uHF II.
Lgr6+ cells located in the isthmus and IFE (Snippert et al., 2010, Füllgrabe et al., 2015)	*Lgr6* was expressed only sporadically over the whole dataset. The highest *Lgr6* expression was found in the Gli1+ cells of the outer bulge (OB III). Lgr6 expressing cells were also found in the outer bulge / upper HF transitional population OB IV and the upper HF basal population uHF IV. A distinct *Lgr6*^{hi} isthmus (sub) population could not be resolved. Neither was it possible to clearly demarcate *Lgr6*+ and *Lgr6*- populations in the IFE basal layer.
Gli1+ cells in the upper bulge (Brownell et al., 2011)	Could be identified (OB III) as cells with an outer bulge signature and a unique set of co-expressed genes (*Gli1*, *Aspn*, *Nrep*).
Krt15+ / Cd34+ mid bulge cells (Cotsarelis et al., 1990, Morris et al., 2004)	CD34 is the most prominent marker of the bulge and in our dataset we confirmed *Cd34* as a pan outer bulge marker found in all outer bulge populations. Cells of the mid bulge (CD34+ / *Lgr5*-) were most likely included in OB I, and OB II. However, it was not possible to clearly delineate the CD34+ / *Lgr5*+ and CD34+ / *Lgr5*- populations by unsupervised clustering of our data.
Krt15+ / Cd34+ / Lgr5+ lower bulge cells (Jaks et al., 2008)	See above.
P-cadherin+ cells of the hair germ (Greco et al., 2009)	Although some P-cadherin (*Cdh3)*-expressing cells could be found in OB I, OB II and OB III, it was not possible to resolve those cells as a distinct population.
Suprabasal Cd34+ / Itga6- cells (Blanpain et al., 2004)	Most likely represented by the population of cells with both an outer bulge and a spinous layer signature (OB V).
Egfl6+ bulge population which provides attachment to the arrector pili muscle (Fujiwara et al., 2011)	*Egfl6* is expressed consistently over all outer bulge populations and sporadically in the IFE. No distinct *Egfl6*^{hi} population could be resolved.
Krt6+ inner bulge cells (Miscellaneous)	Inner bulge cells formed a highly distinct 1st level cluster (Krt6+) which could be further divided into three subpopulations (IB I – III)

Table S4. Differentiation-related genes. Related to Figure 3.

List of significantly pseudotime-dependent genes. Genes are grouped according to clustering shown in Figure 3C. Within each cluster, genes are ordered from lowest to highest p-value.

I	Tgfbi, Krt14, Krt5, Mt2, Fth1, Dst, Smoc2, Itga6, Cbr2, Ifitm3, Tspo, Ccnd2, Fcgbp, S100a10, Cav1, Ftl1, Phgdh, Antxr1, Il33, Capg, Itga3, Serpinh1, Serpinb10, Scin, Itgb4, Lamb3, Slc6a6, Ly6c1, Bhlhe40, Slco2a1, Oat, Cdh13, Ly6a, Paics, Sema3c, Fbxo32, Agrn, Cxcl14, Ltbp4, Serpinb6a, Csrp1, Cyr61, Tnfrsf18, S100a6, Gja1, Jag2, Ptrf, Itgb1, Col4a6, Slc27a3, Col23a1, Ackr3, Slc2a12, Igfbp3, Crlf1, Pdlim1, Ier3, 9530053A07Rik, Col7a1, Sdcbp, Dsg2, S100a13, Ctnnal1, Sult5a1, Fgfr2, Col16a1, Ccnd1, Rassf1, H2-Q9, Rbbp8, Icam1, Dll1, Ephx1, Thbs1, Bmp4, Atf3, Gga2, Tagln2, Dsg3, Mgll, Limch1, Lama5, Slc38a2, Adamts14, Tubb5, Lamc2, Slc12a4, Phldb2, Marveld1, Arhgdib, H3f3b, Htra1, 4930523C07Rik, Gapdh, Pmp22, Frem2, Slc2a1, Efemp1, 1810011O10Rik, Fat2, Ccdc3, Snai2, Klf9, Slc4a8, Tsc22d1, Guca2a, Card10, Hmgn2, Fpgs, Sat1, Cyp4b1, Il20ra, Gm13305_loc1, Socs3, Lmna, Wls, mt-Tc, Sh3pxd2b, mt-Tf, Rab30, Ptprs, Ackr4, Man1c1, Mapkbp1, Cisd3, Anxa5
II	Ccl27a_loc2, Col17a1, Mt1, Krt15, Apoe, Clca2, Cav2, Gpx1, Gas5, Snhg1, 2410006H16Rik, Zfp36l2, Tns4, Clca4, Tmsb4x, Gpt, Cd59a, Clca1, Slc7a6, Trp63, Plp2, Myof, Gm5643_loc2, Sox6, Rnase4, Gm11974, Snrpg, Capns2, Rps12, Cyb5r3, Hmgb1, Ggta1, 1110038B12Rik, 1500012F01Rik, Slc22a4, Npm1, Acp5, Atpif1, Prdx6, Oxct1
III	Acer1, Fabp5, Gsn, Dsc3, Lgals3, Pkp1, Alox12e, Sdr16c5, Micu1, Sdc1, Tmprss4, Mt4, Dbi, Pycard, Tesc, Slco3a1, Rab25, Anxa8, 1810037I17Rik, Camsap3, Pvrl1, Tecr, 4833423E24Rik, Galnt1, Slc12a2, Crip2, Ephb6, Rab3d, Gsta2, Lgals7, Ly6d, Serpinb12, Iqgap1, Degs1, Rdh12, Clu, Eef2k, Cd109, Krt1, Tmem254b_loc2, Aldh3b2, Alox12, Smim5, Hes1, Paqr5, Man2a1, Gcat, Tmem254b_loc3, Tmem254b_loc1, Rhov, Vdac1, Pkp3, Cd9, Tacc2, Hpgds, Sfn, Mir703, Kctd15, Gata3, Atp6v1a, Anxa1, Cyb5, Mif, Npl, Serbp1, Sepp1, Pla2g4f, Atp2b4, Ap1b1, Cat, F2rl1, Bag1, Il17rc, Dock6, Fxyd3, Ell3, Lrp4, Rps2, Epha4, Tm9sf3, Rplp0, S100a11, Epn2, Ctnnd1, Pgrmc1, Bmpr2, Acad9, Atp1b3, Anxa7, Syngr2, Ikzf2, Pdcd4, Urah, Clca5, Ghr, Tsc22d3, Ache, Cst3, Acaa2, Ufsp1, Srd5a1, Zfp185, Ppia, Cyp2b10, Rassf3, Lamp2, Pou3f1, Prdx1, Pkib, Apoc1, Notch3, Serpinb3b, Tmem208, Rell1, Nod1, Ywhaq, Eef1a1, Agpat3, Mbnl2, Rpl10, Sh3bgrl3, Cyb561, 5430435G22Rik, Ralbp1, Ndufv3, Rpl14, Rps26, Nat8l, Dync1li2, Memo1, Hs3st6
IV	Itpr3, Hba-a2_loc2, Sqle, Hba-a2_loc1, Ankrd35, Sptlc3, Tmem154, Ptprf, Gjb5, Gdpd2, Fdps, Sema3d, Crabp2, Dsg1c, Itgb5, Chl1, Xpnpep1, Ttk, Cdh1, Tpm4, Mvk, Acer3, Myh14, Wdr47, Atp2b1, Cyp51, BC064078, Tmem123, Irf6, Add3, Emp1, Cdc42, Aacs, Hmgcs1, Arhgap24, Actr2, Idi1, Sc5d, Dhcr24, Lad1, Snrnp25, Pcyt1a, Tnik, Canx, Lamp1, Cyp2j6, Daglb, Ywhaz, Niacr1, Flrt1, Hspa1b, Flnb, Ldlr, Ahcyl2, Sult2b1, Actg1, Pdlim5, Mvd, Cnn2, Ermp1, Scd2, Adtrp, Magt1, Hn1, Abhd6, Slc6a19, Kif5b, Rab1, Dbnl, Gng5, Eif2s2, Eif1ad, Atp2c2, Jak1, Sar1b, Rnase1, Smagp, Hsd17b7, Pcdhgb4, Fam20b, Gm7334, Hsdl2, Gldc, Hsp90ab1, B4galnt2, Lss, Fut8, Ept1, Anxa2, Serpinb2, P4hb, H2-D1, Creb3, Prom2, 2310022B05Rik, Erbb2, Taf13, Cldn1, Mapkapk3, Map7d1, Osbpl5, Sltm, Slc16a6, Klc3, Tpm3, 2810025M15Rik, Wwp1, Palm, Actn4, Rangap1, BC031181, Cltc, Akirin1, Selt, Atp6ap1, Fam210b, Pgs1, N4bp3, Sh3d19, Rab6a, Slc9a9, Efna5, Fermt3, Ppp1r13l, Fasn, Dusp7, Ctnna1, Prpf18, Zdhhc5, Clic1, Rmnd5b, Llgl2, 2210404O07Rik, F11r, Spry1, Sypl, Plekha6, Rpn2, Tm7sf2, Adam17, Arf4, Timm23, Arpc2, Rai14, Tubb2b, Mtch1, Hsd17b12, Neurl3, Nans, Nsun2, Neu3, Atf4, Fam111a, Gnb1, Slc5a1, Gnb2, Atp11b, Pvrl2, Tjp1, Naa15, Gm15987, Eif3c, Manf, Fnbp1, Ktn1, Ralgapa1, Rdh11, Cyfip1, Sbno1, Eif3a, Shisa5, Tmod3, Ubxn4, Wwc2, Flrt2, Rhbdf1, Rb1, Sec24a, Chit1, Ptma, Phactr2, Prob1, Herc4, Aff4, Prodh, 1110008P14Rik, Fam63b, Cept1, Zmat3, Fkbp1a, Il15, Hsp90b1, Nkd2, Slc39a6, Eif6, BC016579, Pon3, Kdm1b, Vimp, Sumf2, 2510039O18Rik, Gapvd1, Ywhae, Dlg2, Coro1c, Tram1, Mapre1, Slc24a3, Ddx5, Vps41, Srebf2, Itgav, Pla2g4a, Snx3, Bcap31, Psma2, Prrc2b, Tmem35, Scrib, Cand1, Kpna4, Cnbp, Psmb3, Tmem41a, Tet3, Aktip, Emc10, Psmd3, Axl, Nedd8, Slain2, Gab1, Sacm1l, Cdc14b, Ier3ip1, Gnas, C77080, Fgfbp1, Atl2, Nfatc2, Gipc2, Gfod1, Phb, Akr1a1, Vps37a, Ccdc25, Sec31a, Tmem33
V	Sbsn, Krt10, Krtdap, Dmkn, Dsg1a, Calm4, Dsp, Ly6g6c, Gm94, Elovl4, Pvrl4, Susd2, Kctd12, Krt77, Fam25c, Pdzk1ip1, Skint5, Emp2, Ptgs1, Lypd3, Prdx5, Calm5, Gltp, Dapl1, Tmem45a, Ly6g6e, Acsl1, Skint9, Krt78, Them5, Scd1, Perp, Serpinb5, Eppk1, Clip4, Dynll1, Dstn, Kif21a, Skint6, Kif1c, Gng12, Ahnak, Evpl, Pla2g16, Ptgr1, Scel, Cdkn1a, Prnp, Jup, Hebp2, Cebpa, Ablim1, Mfsd6, Cmtm6, Nupr1, Lsr, Sgpp1, Sptssa, Gpr115, Itm2b, Cyp4f39, Skint4, Dsg1b, Me1, Ggh, App, Arpc1a, Pak3, Rab11a, Casz1, Tanc1, Trim29, Ptbp3, Gjb3, Mpzl2, Scamp1, Acly, Ace2, Chmp5, Rcn1, Sort1, Zfp750, Pon2, Rab10, Fuca1, Kank1, Pgrmc2, Pea15a, Nrarp, Metrnl, Nfe2l1, Rab18, Mocs2, Dynlt3, Fut2, Ezr, Tuba4a, Micu2, Ank3, Lrrc58, Tmbim6, Tmed2, Pim3, Ormdl2, Sppl3, Rhbg, Al314180, Plekhn1, mt-Co3, Tollip, Ppp1cb, Sptan1, Arf1, Eif4g2, Notch2, Tom1, Cs, Elovl7, Pglyrp4, Slc6a8, Kif1b, Creg1, Abhd5, Polb, Erlin1, Kremen1, Nhp2, Ptplb, Psen1, Ss18, Golgb1, Ppfia3, Grhl3, Esrp1, Acap2, Scamp2, Hdac5, Rnh1, Cltb, Fam213b, Tuba1a, Nlrc5, Map4k4, Rfwd2, Sptbn2, Atp6v0b, mt-Nd4, Psmd7, Mpp7, H2afy, Tmem79, Wbp1l, Ptpra, Ppp2ca, Nagk, Atp6v0e, Spint2, Hr, Mfsd5, Zfp266, Mob4, Eif5a, Sun2, Fut1, Gpr125, Capn5, Tprgl, Cbx4, Fam57a, Klk11, Baz2a, Ppp1r2, Csnk1a1, Wdfy3, Itfg3, Rnasek, Rtn3, Oas1f, Chmp2b, Rab14, Il20rb, Sphk2, Slc31a2, Ywhab, Stard5, Laptm4a, Acbd3, Dusp22, Atp6v1g1, Jhdm1d, mt-Rnr1, Wnk1, Pdlim2, Ppp2r1a, Ubqln1, Acsbg1, Psmd2, Atp5d, Arf5, Cyth1, mt-Atp6, Ppp2r5e, Rhbdf2, Atp6v1e1, Srgap2, 0610009D07Rik, Abhd12, Tsc22d4, Cope, Adh1, Gp1bb, Ocln, Ovol1, Acaca, 6430548M08Rik, Rab2a, Ddx3x, Tmem134, Aftph, Arap2, Atp7a, Sptlc1, Synrg, Tmbim1, Ttc39c, Mycbp, Kdm5b, S1pr5, Ostf1, Nsdhl, Usp20, Spop, Mta3, Rer1, Usp32, Tspan9, Zmiz1, Lars2, Erp29, Git1, Ztbt7a, Cldn25, Cox7c, Tmem66, Npepps, Clint1, mt-Co2, Pcyox1, Ugt1a2, Smpdl3a, Tmem159, Car13, Ist1, Arpc5, Larp1, Tec, BC048507, Pou2f3, Dctn1, Psmc2, Tmem109, Gbf1, Fuca2, Fbxo6, Usp4, Fbxl3, Aadacl2, Dync1i2, Zmiz2, mt-Ti, Rap2c, Kdsr, Napg, Chmp1a, mt-Nd3, Tmem141, Tiam1, Slc9a3r1, Znrf2, Tmem254c_loc1, 1110057K04Rik, S100a16, Bnipl, Rapgefl1, Pex3, Yif1a, Psmc1, Nsmaf, Rnpepl1, Hdlbp, Zfp91, Tmem254c_loc3, Mtmr1, Ube2v1, Uchl3, Tpd52, Zc3h15, Pls3, Ppp2cb, Dynlrb1, Mcur1, Agfg1, Chmp1b, Tcf25, Mcu, Arpc5l, Wwp2, Cbfb, Ghitm, Emc3, Ppp3r1, Ndufc1, Gnptg, Herc2, Ctage5, Crb3, Lta4h, Celsr2, Irak2, Ppp2r2a
VI	Spink5, Psapl1, Ssfa2, Orm1, Clic3, Hal, Dgat2, Calm1, Mboat2, Skint10, Homer2, Pla2g2f, Ank, Ivl, Agpat4, Grhl1, Ephx3, Ppl, Skint3, Dkkl1, Klk8, Ano9, Tspan8, Plxdc2, Hiatl1, Hspb1, Elovl6, Fam129b, Gjb4, Ybx3, Lipm, Dnase1l3, Tubb2a, Trim2, Klk5, Gdpd3, Blmh, Ccdc64b, Gsdma, Adipor2, Csnk2a2, Rgs2, Myo18a, Calm2, Mllt4, Defb1, Ift20, Cd82, Sdr16c6, Ptpn21, Macf1, Emb, Dsc1, Tmem43, Csnk2a1, Faah, 'Sept8', Rarres1, Abtb2, Klf4, Ccdc50, Sc4mol, Adrbk1, Lgalsl, Pmvk, Pepd, Rhou, Cldn4, Sec11c, Ly6e, Adipor1, Fam107b, Skint11, Ano10, Setd8, Tuft1, Mal2, Hmgcr, Sgms1, Tprg, Sprr1a, Tacstd2, Ndel1, AA986860, Nudt17, Tmed5, Map2k4, Ifitm6, Desi1, Usp7, Cpeb2, Gm11992, Gdi2, Areg, Prdx2, Cox17, Rab27b, Mall, Fdft1, Rdh9, Xbp1, Sorbs1, Stx19, Cast, Spint1, Stk40, 4833439L19Rik, Ceacam19, Klf3, Chmp4b, Inpp5b, Liph, Paip2, Adh6a, Actb, Reep3, Ifrd1, Ppfibp1, Dcun1d1, Pdzd11, Cebpb, Otop3, Spag1, Pld3, Gsk3b, Ptdss1, Arf6, Arhgef10l, Gba, Sp6, Eif1, Rdh1, Atp6v0d1, Sft2d2, Atp6v1d, Stard4, Nt5e, Slc22a23, Bicd2, Gal3st1, Ttc7, C130079G13Rik, Arsa, Tmppe, Gstm5, Scnn1a, Csde1, Golga4, Dpp3, Ttc39b, Defb6, Tgm5, Iffo2, Cab39, H2-T9, Prrc1, Gltpd1, Cep170b, Fam188a, Pdcd6ip, Rhod, Otud1, Kifc3, Cdr2, Chmp2a, Plxnb2, Ndrg2, Il1rn, Ube2b, Jarid2, Atp6v1b2, D15Ertd621e, Grina, Gng13, H2-K1, Eea1, Slc11a2, Tjp3, Clcn3, Atp2a2, Eif4g1, Riok3, Erp44, Sh3bgrl2, Ndfip2, Psmc5, Tmem65, Tmed3, Atp6v0c-ps2, Mknk2, Cstb, Ankrd27, Sphk1, Rnf39, Ptges3, Vcp, Cmpk1, Atp9a, Lyst, Ppfibp2, Baiap2, Usp25, Icmt, Clip1, Slc6a9, Atp6v0a4, Pnpla8, Tyro3, Hspa5, Slk, Ndrg1, Gpr111, Cmtm8, Ssh3, Fam135a, Abcd3, Arl4d, Clasp1, Cyb5r4, Tspan5, Sqstm1, Lrrfip2, Plekha5, St14, Krcc1, Suds3, Usp54, Pum2, H2-Ab1, Pfdn1, Mpzl3, Suco, Syngr1, Stk17b, Tsg101, Rnf13, Ppp1r12a, Golga2, Ykt6, Nipbl, Csrnp2, Cdkn2b, Itpkb, Akap13, Mapre2, Dohh, Tmem50a, Camk2d, Cxcl16, Abi1, Pcdh1, Ilvbl, Mrgprb3, Arl5b, Tmem170b, Bsg, Nfat5, Egr2, Cryl1, Lrrc8e, Maff, Tob2, Slc39a2, Tmem9b, Vamp8, BC100530, Gch1, Pgd, Mad2l2, Gm1821_loc1, Hivep1, H2-Eb1, Mir22hg, Ccdc6, Rab32, Tmed10, Nras, H2-Aa, Rasal2, Leprel4, Selk, Rbx1, Napa, Fbrs, Lig3, Psmc3, Derl1, Rgs1, Hopx, Gns, Golt1b, Srgap1, Il18r1, Guk1, Frmd4b, Gal3st4, Capza1, Mafk, Rec8, Tmed9, Arhgap17, Irf2bp2, Senp2, Pitpna, Glrx2, Man1a2, Emc4, Ell2, Uba1, Vamp2, Mcf2l, Slc30a7, Macc1, Rel, Tmem167, Mtmr6, Mgea5, Atxn7l3b, Mroh6, Hyou1, Errfi1, Lrrc8a, Leprotl1, Kbtbd4, Bach1, Wsb2, Cyth3, Map2k2, Arhgef12, Uso1, Ogdh, Rora
VII	Asprv1, Flg2, Lor, Elovl1, Lce1m, Cst6, Klk7, Pof1b, Crct1, Krt23, 2310002J15Rik, Ide, 2310050C09Rik, Gabarapl2, Nrd1, Ndufa4, Serpinb7, C2cd2, Bpifc, Il1f5, Cnn3, Cnfn, Cdsn, Hrnr, Tgm1, Agpat5, Abca12, Ggct, Eps8l1, Nlrp10, Ugcg, Map2, Ctsd, 2310042E22Rik, Alox12b, Kprp, Tmem54, Arhgef5, Ereg, Sdcbp2, Sec62, Nipal4, Laptm4b, Ctnnbip1, Pcsk6, Osbpl1a, Gnai3, Trex2, Wfdc5, Pvrl3, Serpinb8, Rab5a, Cers3, Vgll3, D8Ertd738e, Rab11fip1, Ubl3, Amotl2, Slc46a2, Ctsh, Atp6v1c2, Tex264, Txnrd1, Mprip, Calml3, Aloxe3, Tmprss13, Cyp2b19, Srpk1, Daam1, Zdhhc9, Cpm, Tead1, Wipf2, 'March7', Rnf11, Nccrp1, Cgn, Hspb8, Abca5, Cldn23, Dhcr7, Il22ra1, Zfand6, Kctd11, Myo6, Mapk3, Slc6a14, Pcmtd1, Dhrs1, 'March8', Serinc3, Mxi1, Map1lc3a, Stk4, Gan, Tmem40, Ypel2, Fcho1, F3, Yod1, Sh3kbp1, Cobl, Fam89a, Rab8b, Por, Ctsa, Epcam, Wdr26, Nipal1, Ifi30, Smpd1, Cln8, Cul3, Tob1, Hook3, Pdzd2, Psmd8, Appl2, Peli1, Tuba8, B2m, Arhgap23, Pnpla1, Fndc9, Dlgap4, Hectd1, 4933426M11Rik, Hbp1, Slc30a4, AI661453, Mxd1, Taok1, Rps6ka6, Hip1r, Zswim4, Ptrh1, Pim1, Hdgf, Cds1, Blnk, Gramd3, Pard3, Eif5, Card14, Zfp706, Slc15a1, Abca3, Chp1, Arg1, Ypel3, Hars, Kctd5, Shroom3, Ube2d3, Atp6v0a2, Asah2, Sidt2, Capn9, Sertad1, Rmnd5a, Endou, Ccdc120, Ier5, Fis1, Ppp1cc, Fhdc1, Tma7, Map3k9, Sucla2, Casp14, Ppp1r13b, Slc27a4, Mid1ip1, Faf1, Tsc22d2, Pnpla3, Dusp16, Prdm1, Itsn2, BC021767, Stx12, Eps8l2, Tpm1, Lrp10, Chmp3, Lmo7, Cited4, Slc19a2, Ube2k, Snx2, Ttpal, Pdxdc1, Huwe1, Vasn, Myo5b, Krt79, Xkrx, Cited2, Aim1, Midn, Tmcc3, Fam84a, Ranbp9, Pink1, Gdpd1, Cyfip2, Sppl2a, Gpr87, Tab2, Aim1l, Acpp, Mreg, Abhd12b, Usp6nl, Slmo2, Tnfaip8, Spag9, Endod1, Zfand3, Fam73b, Golga7, Dnaja4, 1600029D21Rik, Usp16, Epn3, Abhd17c, Ankrd44, Prkch, Paf1, Rtfdc1, Eps15, Tjp2, Jund, Ypel5, Tnfaip1, Shfm1, Litaf, Ammecr1, Pacsin2, Mtss1, Dnaja1, Arhgap29, Mlf2, Ehd1, Dab2ip, Sdr42e1, Dusp14, Sdf2, Amotl1, Arrb1, Bin3, Sash1, Trp53inp2, Wipi2, Eif4h, Rnf208, Mbnl1, Neat1, Max, Prkx, Lypd6b, S100a14, Pip5k1a, Eif4e, Tpra1, Sec14l1, Hmgcl, Ube2n, Sh3gl1, Med4, Wasf2, Timp2, Rorc
VIII	Csta, Ctsl, 1100001G20Rik, A030009H04Rik, Anxa9, Sdr9c7, Lypd5, Ube2r2, Serpina12, Elmod1, Nt5c3, Aif1l, 2310014L17Rik, Krt17, Hsbp1l1, Il18, Mapk13, 2010109I03Rik, Lce1c, Spns2, Btc, Lce1a2, Slurp1, Lipk, Lce1b, Smpd3, Ecm1, Dusp18, Lce6a, Lipn, Bnip3l, Hmox1, 2200002D01Rik, Lce1a1, Il1f9, Cpe, Tmprss11f, Gas7, Lrrc28, Minpp1, Msrb1, Id4, Lphn2, Glrx, Hexb, Map2k3, Slc35e4, Il1r2, Malat1, Fchsd2, Usp2, Rnf222, Gm14137, Sod3, Tgfa, Plk3, Cpa4, Spsb3, Stk10, Myzap, Atg9b, Scgb1a1, Smap2, Gng4, Prss8, Tmem229a, Nck2, Fmo2, Ttc9, Lce1l, Ano8, Psors1c2, Scpep1, Actn1, Slc38a3, Crlf3, Rassf5, Robo1, Tnfrsf11b, Pqlc1, Map4, Tmem177, Klk9, Esyt3, Acot7, Kctd4, Rab3b, Nkpd1, Heca, Wdr45, Lce1f, Dnase1l2, Calcb, Ankrd23, Ralgds, Oser1, Piezo1, Il1rap, Tmem56, Ubtd1, Rapgef2, Lce1d, Ckb, Krtap3-2, Cyp26b1, Gm2a, Rnf103, Thop1, Gramd1c, Dennd1b, Ccl20, Lce1e, Cyp4v3, Glcci1, Lce1k, Lce1i, Ormdl1, Tmem120a, Lce1j, Pcdh7, Lce1g, Clec2g, Ppif, Abcg1, Strip1, Chic2, Gpld1, 3300005D01Rik, Gpsm1, Fam214a, Gba2, Myo1b, Myo10, Lrch1, Btbd3, Plcxd2, Fam110a, Map1lc3b, Mdfic, Blzf1, Asap3, Lnx1, Plcxd1, Map3k8, Ak1, Bnip3, Lpin3, Krt80, Fmn1, Tnfrsf21, Cnst, Arhgap32, Rassf7, Ubash3b, Trim16

Table S5. Spatial axis related genes. Related to Figure 4.

List of significantly pseudospace-dependent genes. Genes are grouped according to clustering shown in Figure 4C. Within each cluster, genes are ordered from lowest to highest p-value.

I	Krt14, Mt2, Chit1, Mt1, Ccl27a_loc2, Rplp1, Rps12, Tnfrsf19, Rps16, Gnb2l1, Rps20, Rps15a-ps4, S100a11, Rps28, Gm6654, Rpl39, Rpl12, Rps3a1, Arhgdib, Rpsa, Txndc17, Rps24, Rpl29, Gpx1, Rpl4, Rpl23, Rplp2, Fau, Rpl18a, Rps3, Gm6402, Tnfrsf18, Rps29, Rpl5, Nme2, Nfkbiz, Rps6, Gm5643_loc2, Rpl8, Rps5, Gm13139_loc1, Hmgcs1, Rpl13a, Eef1b2, Rpl32, Slco2a1, Gm13826, Klf6, Rpl35, Eef2, Gm13139_loc2, Rpl31-ps12, Rpl24, 2410006H16Rik, Rps25, Htra1, Rps13, Cd55, Rps18, Fgfbp1, Dapl1, Calm5, Cox4i1, Gpt, Jun, Rpl36al, Limch1, Naca, Npm1, 1810037I17Rik, Dstn, Hspa5, Atox1, Pkp4, Emp1, Psme2, Gm5643_loc1, Ptger4, 1500012F01Rik, Atp5h, Litaf, Mif, LOC100861978_loc4, Btf3, Gm15421
II	Serpinb2, H3f3b, Il1r2, Avpi1, Fth1, Il33, Anxa2, Ly6a, Tnfaip3, Sat1, Wnt3, Thbs1, Gata3, Ly6c1, Itga6, Gja1, Ifngr1, Slc22a4, Adamts14, Actg1, Gdpd2, Pak6, Slc2a12, Krt16, Serpina3h, Tsc22d1, Fbxo32, Atf3, Itm2b, Adrb2, Ets2, Antxr1, Cyr61, Gm4832, Sfn, Sema3c, Mgll, Krt5, Pde4b, Sema3d, Chl1, Ipmk, Col23a1, Tubb4b, Wdr65, S100a10, Il22ra2, Wee1, Cbr2, Ackr3, 1810011O10Rik, Il20ra, Zfp36l2, Dst, Rnase4, Ifi27, Tppp3, Cpxm2, Phgdh, Igfbp3, Bzw1, Hmgn1, Cks2, Man1c1, Fam162a, Nr4a1, Bmp4, Ifi202b, Higd1a, Slc27a3, Ppp2r2c, Abcb1b, Ubc, Itga3, Crlf1, Cd59a, Efemp1, Prdx6, Gnai1, Tfap2c, Serpinb10, Fam25c, Eif1a, Atp5l, Fam213b, Il6ra, Rhoa, Atf4, Psat1, Dusp1, Klk11, Ptges, Klf2, Snhg1, Dnajb1, Wnk2, S100a13, Epgn
III	B2m, H2-K1, Ifitm3, Tacstd2, H2-Q9, Ly6d, Clca2, Fcgbp, Smoc2, H2-L, H2-D1, Scin, Clca1, Clca4, Aldh3a1, Bhlhe40, Ly6e, Atpif1, Ccnd2, Anxa1, Klf5, Calm2, Cstb, Ahr, Ptma, Oat, Ahnak, Rps9, Ivns1abp, Oas1f, Ptpn14, Nfkbia, Xist, Hmgb2, Lrig1, H2-Q6, Fam134b, Plbd1, 9530053A07Rik, Pkp1, Acsbg1, Tap1, Ifrd1, Psmb9, Slc2a1, Gm15987, Slco2b1, Tnfaip8, Car12, Cnn2, Neurl1b, Nlrc5, H2afz, Cdh1, Card10, Sik1, Pkib, Itpr2, H3f3a
IV	Krt17, Cst6, Defb6, Krt79, Fst, Aadac, Gsn, Ly6g6c, Gstm5, Sostdc1, Gata6, Epcam, Psapl1, Marcks, Efnb2, Pdzk1ip1, Tm4sf1, Pthlh, Skint4, Emp2, Skint3, Bmp7, Alox12e, Cyp1b1, Krtdap, Hes1, Tmem45a, Cxadr, Lphn2, Cyp27a1, Sprr1a, Apoe, 1600029D21Rik, Acsl4, Lmo7, Serpina3g, Klk7, Rbp1, Cwh43, Defb1, Susd2, Aldh3a2, Lgals7, Camk2n1, Tle1, Klhl8, Ttc39c, Pof1b
V	Ank, Fam101b, Fgfr1, Alcam, Gas1, Sema3e, Vwa2, Grem1, Aspn, Cspg4, Cd200, Nrep, Ltbp1, Moxd1, Robo2, Igfbp4, Wif1, Steap4, Egfl6, Runx1, Col5a2, Crim1, Nrbp2, Nudt4, Tgfb2, Tnfrsf11b, Cdc42ep3, Cald1, Gli1, Megf6, Cgnl1, Bdnf, Hoxc8, Boc, Hk2, Rbms3, Dab2, Mbnl1, Ndufa11, Crispld1, Gli2, Gpr125, Fam83d
VI	Cd34, Postn, Tgm5, S100a4, Lhx2, Sfrp1, Dkk3, Sparc, Col8a2, Fzd2, Shisa2, Col6a1, Lgr5, Slc6a6, Ctgf, Id2, Pdzrn4, Il31ra, Cxcl14, Col17a1, Krt24, Crip1, Adamtsl4, Gm973, Duoxa1, Ecm1, Igfbp5, Hmcn1, Trpv4, Ltbp2, Cadm1, Trabd2b, Ism1, Igfbp7, Fbln1, Kcnk2, Tns1, Col18a1, Ltbp3, Tbx1, Fgf1, Itm2a, Chat, Kcnma1, Scx, Setbp1, Slc38a2, Ppap2a, Id3, Fstl1, Smarca2, Col6a2, Cited2, Tgfbi, Duox1, Sardh, Gfra1, Mfge8, Cck, Nog, Col12a1, Golim4, Agrn, Sh3rf1, Dpy19l1, Prss23, Pfn2, Angptl2, Gfra2, Flrt2, Tnc, Il11ra1, Gcat, Serpinh1, Myoc, Sncg, Lrrn3, Emb, Angptl7, Cables1, Irx5
VII	Sox9, Aqp3, Wfdc3, Timp2, S100a6, Calml3, App, Sbsn, Nfib, Flnb, Actn1, Ptn, Nt5e, Serpinb11, Foxp1, Thsd1, Gpc6, Plxna2, Fzd1, Dmkn, Tfap2b, Nfatc1, Mllt4, Foxc1, Prlr, Aplp2, Hr, Wwp2, Fam132a, Pdzrn3, Tmtc1, Fzd7, Macf1, Tacc2, Tenm2, Zmiz1, Vdr, Casz1, Txn1, Lrrfip1, Ptprk, Efna5, Pi15, Nr3c1, Ptprf, Capn2, Atp13a2, Fzd3, Ppap2b, Sdc1, Lgr4, Hoxc13, Myo9a, Sdhb, Dapk2, Ctnnb1, Pdlim3, Mgst1, Rnf152, Ndrg2, Cbx6, Pcsk6, Vwa1, Dsp, Bnc2, Camsap3, Pik3r1, Tcf12, Klhl29, Myh14, Filip1l, Igdcc4, Hopx, Ftl1, Cd47, Mitf, Acot1, Grip1, Itgb5, Trim2, mt-Tp, Nfat5, Ssfa2, Arhgap44, Mxi1, Ssbp3, Pkd1, Id4, Itpr3
VIII	Spink5, Timp3, Krt75, Cst3, Cyp26b1, Fam167a, Fgf18, Krt6a, Cryab, Bmp2, Adcy1, Arg1, S100a1, Cdsn, Fam84a, Tspan2, Endod1, Cd24a, Perp, Krt15, Dap, Lima1, Ptgds, Fxyd6, Pvrl4, Sh3kbp1, Lypd3, Hspa2, Slc45a3, Fam26e, Dlgap4, Atp6v0e2, Dclk1, St3gal4, Sgk1, Fmn1, Clic3, Them5, Homer2, F3, Gm2a, Sulf2, Tmsb4x, Fgfr3

Table S6. Marker genes – Stem cell analysis. Related to Figure 6.

Lists of genes that are most highly expressed over *Baseline* (vs. *Baseline,* left column) or the second highest population (vs. other groups, right column). Identification of genes is based on negative binomial regression of stem/progenitor marker expressing populations. In the case of *Lgr5*, *Cd34*, *Gli1*, *Lgr6*, *Lrig1* and *Krt14* expressing cells, each population was compared to either a shared *Baseline* or all other populations. For each population, genes whose population-specific expression coefficient exceeded the *Baseline* coefficient (left column) or all other populations-specific coefficients (right column) with 95% posterior probability and which show the largest gap to the *Baseline* / the second highest population (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the *Baseline* / the second highest population) are listed.
The SCM+ basal cells were compared to SCM− basal cells or a *Baseline* shared by both populations. Genes which exceeded the *Baseline* coefficient (left column) or the SCM− basal cell coefficient (right column) with 90% posterior probability and which show the largest gap to the *Baseline* / the SCM− basal cells (difference between the 25th percentile of the population-specific coefficient and the 75th percentile of the *Baseline* / the SCM− basal cells) are listed.

(Supplied as Excel file: Joost Table S6.xlsx)

Table S7. Immunohistochemistry and single molecule FISH stainings. Related to STAR Methods.

Listed are the markers used for the validation and localization of the defined cell populations, the respective number of mice and analyzed images, and the corresponding figures in the manuscript.

Population	Clustering level	Staining to identify populations based on our sequencing data	Number of mice	Number of images taken (HF / HF+IFE / IFE)	Number of images showing the respective population	Corresponding figure in manuscript
IFE BI	2	KRT14(hi)/*Thbs1*(hi)	3	1 / 15 / 11	25 out of 27	2E
IFE BII	2	KRT14(dim)/*Thbs1*(lo)	3	1 / 15 / 11	25 out of 27	2E
INFU B	2	*Postn*(dim)	4	25 / 78 / 0	61 out of 103	2E
IFE DI	1	KRT10(dim)/PTGS1(dim)	2	0 / 7 / 0	7 out of 7	S2J
IFE DII	1	KRT10(hi)/PTGS1(hi)	2	0 / 7 / 0	7 out of 7	S2J
IFE KI	1	LOR(dim)/*Flg2*(dim)	3	0 / 11 / 5	16 out of 16	S2J
IFE KII	1	LOR(hi)/*Flg2*(hi)	3	0 / 11 / 5	16 out of 16	S2J
uHF I	1	KRT17(lo)/KRT79(lo)	3	0 / 17 / 0	16 out of 17	S2J
uHF II	1	KRT17(dim)/KRT79(dim)	3	0 / 17 / 0	16 out of 17	S2J
uHF III	1	KRT17(hi)/KRT79(hi)	3	0 / 17 / 0	16 out of 17	S2J
uHF I	2	KRT14(lo)/*Cst6*(hi) and KLK10 to localize	3 for Krt14/ Cst6; 3 for Klk10	0 / 16 / 0 for Krt14/Cst6 0 / 15 / 0 for Klk10	16 out of 16 for Krt14/Cst6; 14 out of 15 for Klk10	2E, S3L
uHF II	2	KRT14(hi)/*Cst6*(hi/dim)	3 for Krt14/ Cst6	0 / 16 / 0 for Krt14/Cst6	16 out of 16 for Krt14/Cst6	2E
uHF III	2	can't be stained*		not applicable	not applicable	
uHF IV	2	KRT14(dim)/*Krt79*(lo)	3	3 /12/ 0	15 out of 15	S3L
uHF V	2	KRT14(dim)/*Krt79*(hi)	3	3 /12/ 0	15 out of 15	S3L
uHF VI	2	KRT10(hi)/KRT79(hi)	3	0 / 14 / 0	11 out of 14	S3L
uHF VII	2	LOR(hi)/KRT79(hi)	3	0 / 20 / 0	17 out of 20	S3L
SG	1	MGST1(pos)	3	0 / 18 / 0	18 out of 18	1F, S2J
OB	1	CD34(hi)/*Postn*(hi)	2	7 / 7 / 0	14 out of 14	S2J
OB I	2	*Lgr5*-EGFP(hi)/*Postn*(hi)	4 for Postn; 2 for Lgr5**	25 / 78 / 0 for Postn; 43 / 0 / 0 for Lgr5	95 out of 103 for Postn; 39 out of 43 for Lgr5	S3L
OB II	2	*Lgr5*-EGFP(dim)/*Postn*(hi)	4 for Postn; 2 for Lgr5**	25 / 78 / 0 for Postn; 43 / 0 / 0 for Lgr5	95 out of 103 for Postn; 39 out of 43 for Lgr5	S3L
OB III	2	KRT15(lo)/*Postn*(hi)	3	0 / 16 / 0	9 out of 16	2E
OB IV	2	*Postn*(hi)/*Krt79*(dim)	4	5 / 20 / 0	15 out of 25	2E
OB V	2	KRT10(hi)/*Postn*(hi)	3	1 / 21 / 0	18 out of 22	2E
IB I	2	KRT6(hi)	4	6 / 24 / 0	30 out of 30	1F, S2J
IB II	2	KRT6(hi)/*Postn*(hi)	3	5 / 20 / 0	19 out of 25	2E
IB III	2	KRT6(hi)/*Krt10*(hi)	2	3 / 17 / 1	14 out of 20	2E
TC	1	CD3(pos)	3	0 / 25 / 0	24 out of 25	1F, S2J
LH	1	CD207(pos)	2	0 / 10 / 0	10 out of 10	1F, S2J

Legend

hi = high; lo = low; pos = positive

IHC / RNAscope

Populations stained together

* was located via exclusion of positive stainings

** costaining was technically not possible