Supplemental Inventory

Figure S1. This figure supports Figure 1 and shows additional information on our cell sorting method, marker immunofluorescence and reporter expression, and qRT-PCR on sorted populations.

Figure S2. This figure supports Figure 3 and shows qRT-PCR validation of known markers identified in the RNA-seq, and the relationships and overlaps of our new signatures with previously published molecular signatures.

Figure S3. This figure supports Figure 4 and shows qRT-PCR validation for new markers, comparison of our DP niche cell population signatures with previously published signatures, and gene ontology analysis.

Figure S4. This figure supports Figure 5 and shows qRT-PCR validation of known and novel signature genes for SC precursors.

Figure S5. This figure supports Figure 7 and shows heat maps of selected signaling pathways for which factors are expressed in heterogeneous cell types.

Supplemental tables legends.

Supplemental experimental procedures. These procedures include information on immunofluorescence, cell isolation, library production and RNA-sequencing analysis, Real-Time Quantitative Reverse Transcription PCR, and details on the intercellular network analysis. Primers are listed for qRT-PCR.

Supplemental references.

Figure S1

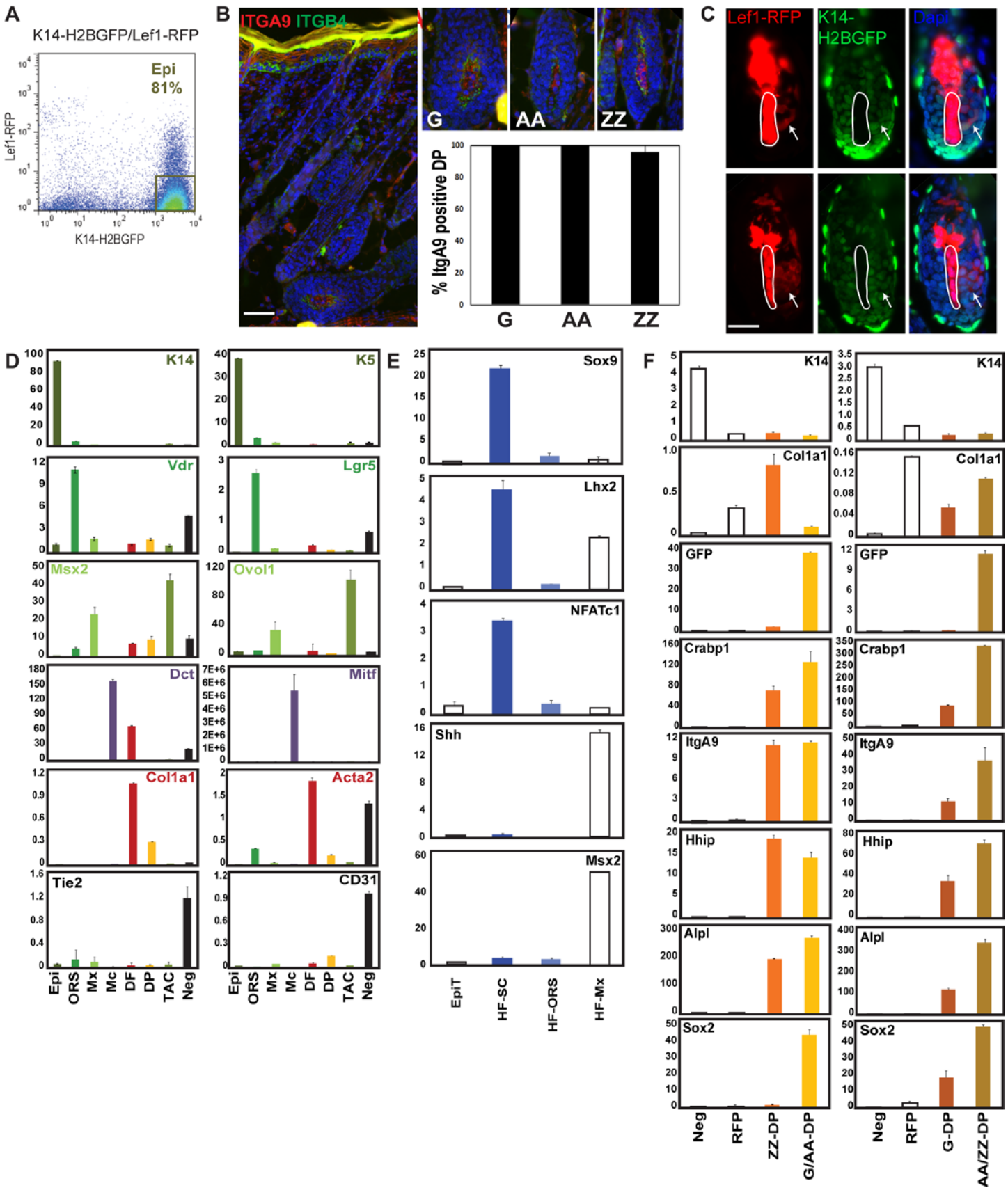

Figure S1. Supports Figure 1 data. Cell sorting method and qRT-PCR validation.
(A) FACS plot and gate for isolation of Epi from the epidermal preparation of P5 *K14-H2BGFP;Lef1-RFP* back skins.
(B) Immunofluorescence staining on P5 back skin for ITGA9 showed expression in 100% of all DPs of all hair types. G, guard hair; AA, awl and auchene hair; ZZ, zigzag hair. Scale bar=100 μm.

(C) Representative images of *K14-H2BGFP*;*Lef1-RFP* hair follicles at P5. TAC progenitors co-express H2BGFP and RFP (arrows). DP is RFP^{+} (demarcated by white line). Mc capping the DP strongly express RFP as previously described (Rendl et al., 2005). Few cells in the upper differentiating compartment express H2BGFP/RFP as well.
(D) qRT-PCR validation of sorted populations from *K14-H2BGFP*;*Lef1-RFP* P5 back skins. Note that detection of epithelial markers (K14, K5, Vdr, Lgr5, Msx2, Ovol1) within the Neg population comes from H2BGFP negative epithelial cells due to sporadic mosaic expression of the *K14-H2BGFP* transgene. Data are mean ± SD from 2 measurements.
(E) qRT-PCR validation of sorted populations from *K14-RFP*;*Sox9-GFP* P5 back skins. Data are mean ± SD from 2 measurements.
(F) qRT-PCR validation of sorted populations from $Sox2^{GFP}$;*Lef1-RFP* and *Crabp1-GFP*;*Lef1-RFP* P5 back skins. Data are mean ± SD from 2 measurements.

Figure S2

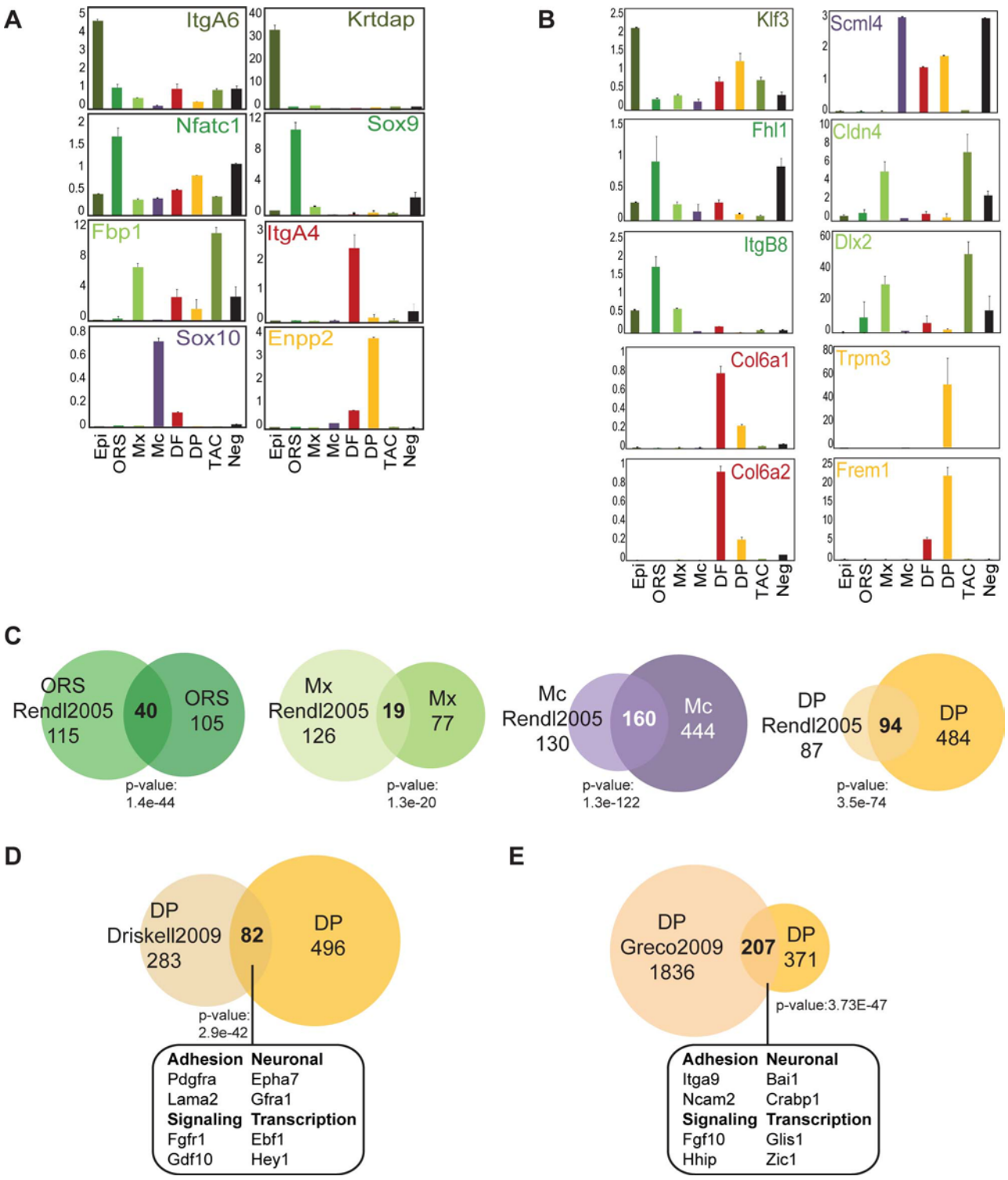

Figure S2. Supports Figure 3 data. Molecular characterization of the main HF populations and qRT-PCR validation of new signature genes.
(A) qRT-PCR validation of known cell-type specific markers. Data are mean ± SD from 2 measurements.
(B) qRT-PCR validation of new signature genes. Data are mean ± SD from 2 measurements.
(C) Comparison of the new molecular signatures with previously published HF signatures at P5 (Rendl et al, 2005).
(D) Comparison of the new DP signature with previously published DP signature at P2 (Driskell et al, 2009).
(E) Comparison of novel P5 DP signature with previously published adult DP signature (Greco et al, 2009).

Figure S3

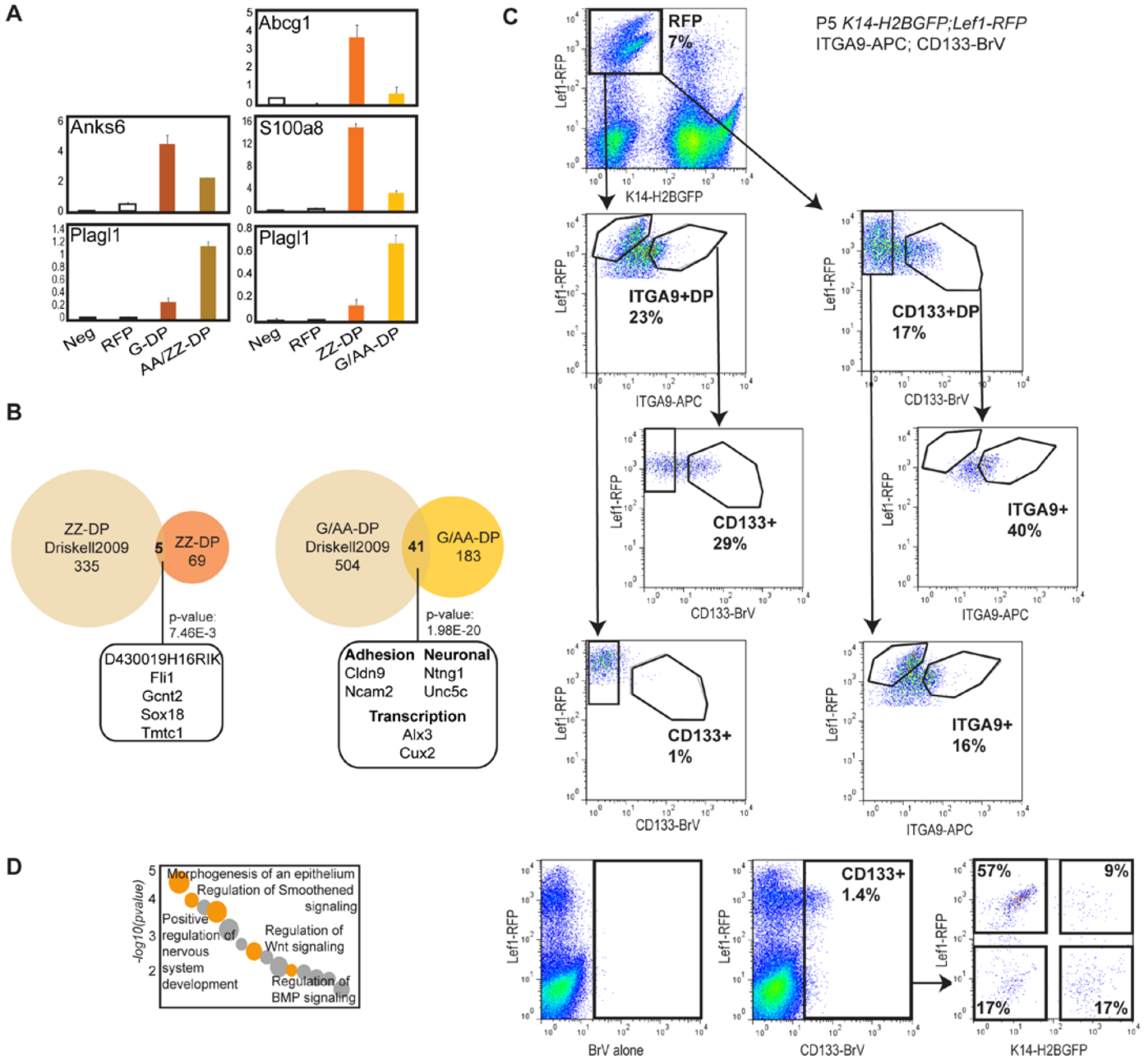

Figure S3. Supports Figure 4 data. Identification of hair type-specific DP signature genes.
(A) qRT-PCR validation of new hair type specific signature genes. G, guard hair; AA, awl and auchene hair; ZZ, zigzag hair. Data are mean ± SD from 2 measurements.
(B) Comparison of novel ZZ-DP and G/AA-DP signatures with previously published DP signatures at P2 (Driskell et al, 2009).
(C) Comparison of ITGA9 and CD133 stainings in a dermal preparation from P5 *K14-H2BGFP;Lef1-RFP* back skins. Top: DP cells are highly enriched when selected as $RFP^{+}ITGA9^{+}$ population where >90% of cells are AP^{+}, alkaline phosphatase, an exclusive DP marker at P5 (see also Figure 1D,E). Only 30% of these DP cells are $CD133^{+}$ suggesting that CD133 is expressed in a subpopulation of DP cells. Likewise $ITGA9^{+}$ DP cells (16%) are found in the $RFP^{+}CD133^{-}$ fraction. The $RFP^{+}ITGA9^{-}$ gate does not contain $CD133^{+}$ cells suggesting that within the RFP^{+} cell population CD133 is indeed specific for DP cells. Bottom: Note some $CD133^{+}$ cells are found within the $H2BGFP^{+}$ and RFP^{-} gates, suggesting that CD133 is also expressed by other cell types and is not specific to DP.
(D) Gene ontology analysis of core DP signature. Notable terms are highlighted; all terms are listed in Table S5.

Figure S4

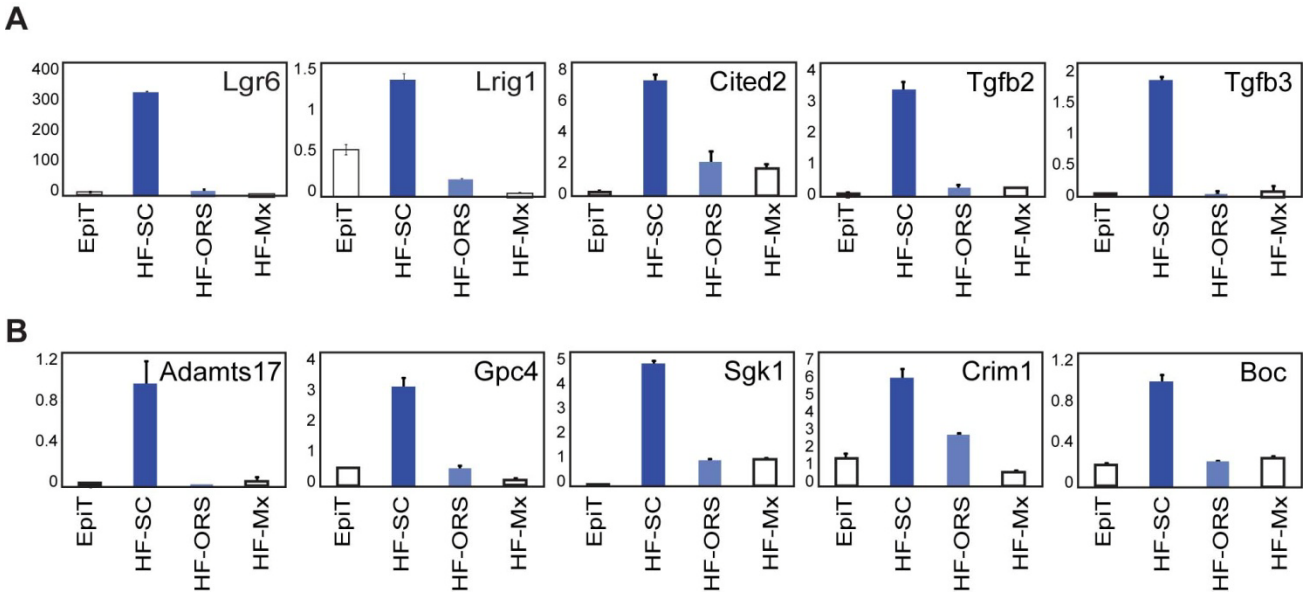

Figure S4. Supports Figure 5 data. Validation of signature genes in HF-SC precursors at P5.
(A) qRT-PCR validation of known HF-SC precursor-specific signature genes. Data are mean ± SD from 2 measurements.
(B) qRT-PCR validation of new signature genes in HF-SC precursors. Data are mean ± SD from 2 measurements.

Figure S5

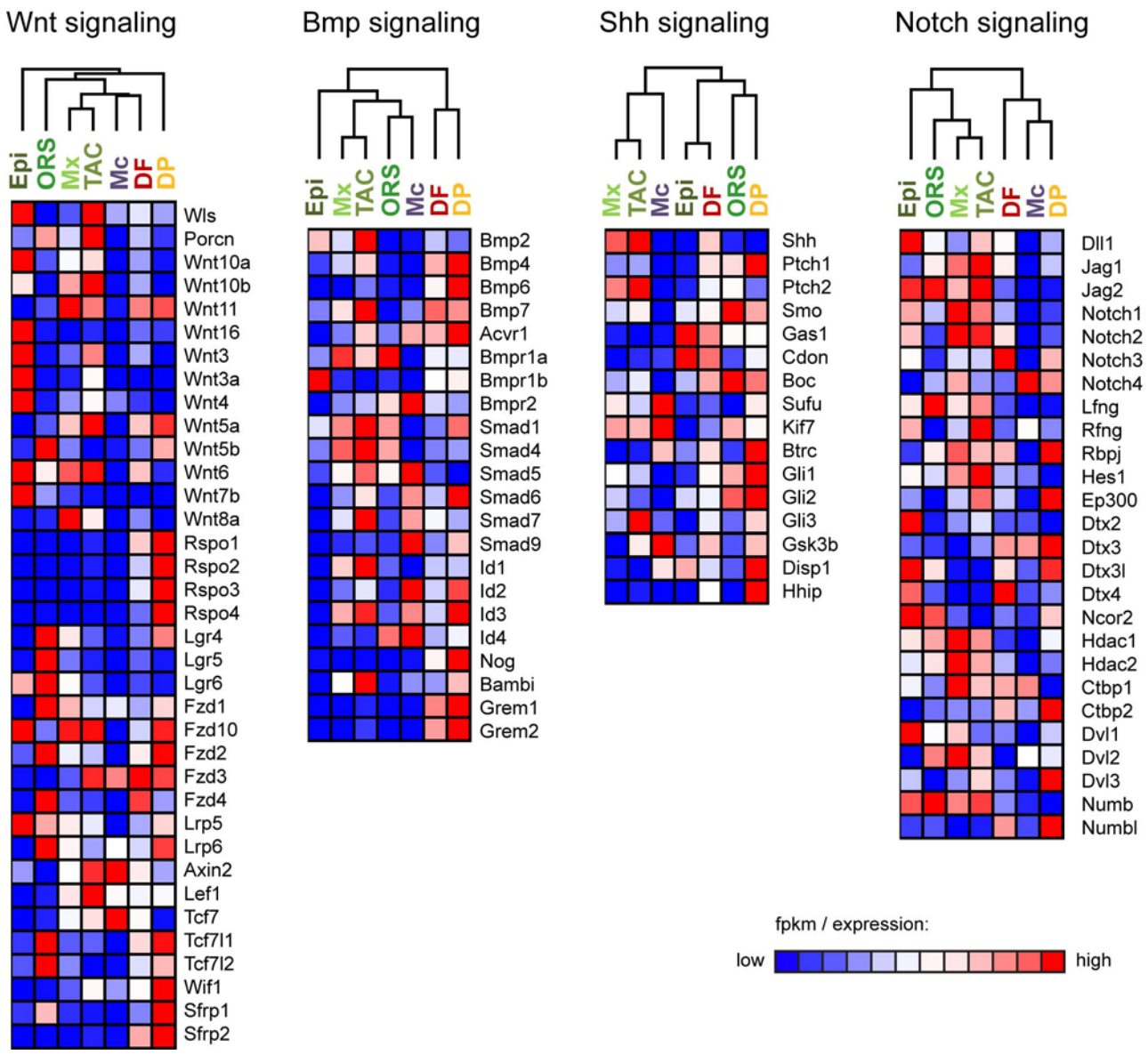

Figure S5. Supports Figure 7. Multiple factors from major signaling pathways are specifically expressed by distinct skin and hair follicle cell types in back skin at P5.
Genes involved in Wnt, Bmp, Shh and Notch signaling were mined from the KEGG database and represented in a heat map if expressed (FPKM >1) in P5 back skin.

SUPPLEMENTAL TABLE LEGENDS

Supplemental tables available as xls files:

Table S1. Supports Figure 1. RNA-Seq QC.
Table S2. Supports Figure 2. Differentially expressed genes.
Table S3. Supports Figures 3 to 6. Signature gene lists.
Table S4. Supports Figure 3. Enriched genes in epithelial and dermal cells.
Table S5. Supports Figures 3 to 6. Results of Enrichr GO analysis.
Table S6. Supports Figures 3 to 6. Overlap between P5 signatures and previously identified signatures.
Table S7. Supports Figure 7. Intercellular network analysis.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Immunofluorescence staining

Antibodies used included GFP (Rabbit, 1:500 Millipore), GPX3 (Goat, 1:50 R&D Systems), ITGA9 (Goat 1:100, R&D Systems), ITGB8 (Kind gift from Dr J.McCarty), MC3R (Rat, 1:100 R&D Systems), WNT5B (Rat, 1:20 R&D Systems). For GPX3 immunofluorescence staining, slides were microwave boiled for 10 minutes in antigen unmasking solution (10nM Na-citrate, 0.05% Tween-20, pH 6.0), then treated with Triton X-100 (0.5% in PBS) for 15 minutes before the blocking step and GFP antibodies were used to detect H2BGFP. After incubation with primary antibodies, samples were washed, then incubated with Rhodamine Red-X- or AF488-conjugated donkey anti-goat or rat secondary antibodies (Jackson Immunoresearch). Nuclei were counterstained with DAPI. Slides were analyzed using a Leica SP5 DM confocal microscope driven by the Leica LASAF software.

Cell isolation by FACS

P5 back skins from *K14-H2BGFP*;*Lef1-RFP*, *Sox2*^{GFP};*Lef1-RFP* and *Crabp1-GFP*;*Lef1-RFP* transgenic mice were treated with dispase (Invitrogen) 4ºC overnight to separate epidermis from dermis, as previously described (Rendl et al., 2005). Dermis was digested with 0.2% collagenase (Sigma-Aldrich) and 20U/ul of DNAse (Roche) at 37ºC for 50 min. Intact follicles and dermal cells were sedimented at 300g and the preparation was then enriched in follicles by centrifuging twice at 20g. Following trypsinization for 5 min at 37ºC, cell suspensions were filtered through 40µm cell strainers. Resuspended cell pellets were stained with primary antibodies against CD117-biot (Rat, 1:50, BDPharmingen) and/or ITGA9 (Goat, 1:50, R&D Systems), followed by staining with Donkey anti-goat APC secondary antibodies (1:200, Jackson Immunoresearch) and Streptavidin-Brilliant Violet 421 (1:200, Biolegend). DAPI was added for dead cell identification and exclusion. Cell isolations were performed on a BD Influx cell sorter at the ISMMS flow core facility. For isolations from P4 *K14–RFP*;*Sox9–GFP* double transgenic mice, the full thickness skin was treated with dispase for 2h at 37 °C to remove the epidermal sheets. Single cells were resuspended in PBS with 3% chelexed FBS and incubated with Alexa-Fluor-647-conjugated ITGA6 antibodies (1:75, AbD Serotec) for 45min on ice. Cell isolations were performed on Moflo XDP cell sorter (Beckman Coulter).

Library production and RNA-sequencing analysis

Total RNA obtained from FACS-sorted cells was purified with the Absolutely RNA Nanoprep kit (Agilent), quantified with the NanoDrop spectrophotometer (Thermo Scientific) and measured for quality control by the Agilent Bioanalyzer. Samples with RIN (RNA integrity number) scores 8 and higher were further processed. 6 ng starting material was reverse transcribed and amplified to 5-7µg cDNA with the RNA Ovation RNA-seq System V2 (NuGEN). From 100ng amplified cDNA sequencing libraries were generated with 36 unique barcoded adapters using the Ovation Ultralow DR Library System (NuGEN). Library concentration and quality was quantified by Qbit (Invitrogen) and Bioanalyzer (Agilent) and subsequently sequenced on the IlluminaHiSeq 2000 platform using a 100nt single-read setting.

All raw RNA sequencing reads were mapped to the mouse genome (mm10) with TopHat v2.0.3 (Trapnell et al., 2009) coupled with the Bowtie2 (Langmead and Salzberg, 2012) aligner with default parameters. Transcriptomes were assembled and fragments per kilo-base per million reads (FPKM) for each gene were computed with Cufflinks v2.1.1 (Trapnell et al., 2010) with default parameters. Differentially expressed genes (DEGs) were identified using Cuffdiff (with default parameters except for the library normalization method was upper quartile normalization, where FPKMs were scaled via the ratio of the 75 quartile fragment counts to the average 75 quartile value across all libraries) and ANOVA, and the Fisher exact test was used for enrichment analysis with the Benjamini-Hochberg correction for multiple hypotheses testing with FDR significance cut off $q<0.05$. Hierarchical clustering analyses for samples were performed with an FPKM matrix of either all detected genes or DEGs. The FPKM matrix was $\log_{10}$ transformed and standardized across each gene using z-scores so that the relative gene expression values across samples were 0 centered. Hierarchical clustering was performed for both genes and samples with Euclidean distance and average linkage functions. Principle component analyses (PCA) were performed for samples with the scikit-learn Python package and visualized in 3D plot using the Matplotlib Python package. Gene ontology enrichment analysis was carried out using Enrichr (Chen et al., 2013). Signature genes were defined as genes from the Cuffdiff or ANOVA analyses with an FPKM $\geq$ 1 and an FC of $\geq$ 2 compared to other populations, unless described otherwise in the main text. Heatmaps were generated by integrating data from the KEGG pathways database (Kanehisa and Goto, 2000; Kanehisa et al., 2014) and with Genepattern (Reich et al., 2006). KEGG pathway enrichment analysis was performed with Enrichr (Chen et al., 2013).

Real-Time Quantitative Reverse Transcription PCR

Total RNA obtained from FACS-sorted cells was purified with the Absolutely RNA Nanoprep kit (Agilent), quantified with the NanoDrop spectrophotometer (Thermo Scientific) and reverse transcribed using oligo(dT) primers (Superscript III First-Strand Synthesis System, Invitrogen). Real-time quantitative reverse transcription PCR (qRT-PCR) was performed with a LightCycler 480 (Roche) instrument with LightCycler DNA Master SYBR Green I reagents.

Intercellular network analysis

To uncover signaling interaction between populations, we built a ligand-receptor database for the mouse system consisting of 342 ligand genes mapped to 470 refSeq IDs and 353 receptor genes mapped to 486 refSeq IDs using BioMart (Kasprzyk, 2011). Unique interaction pairs were identified using iRefWeb (Turner et al., 2010).This current database is derived from the human system database previously used in Qiao et al. (Qiao et al., 2014). KEGG pathway was accessed in August 2015.

Primer list for qRT-PCR

Gene	Forward primer	Reverse primer
Abcg1	GAGTCCCTGCTTTCTGGCTT	CCGCCTTCTCTTTGAGTGGT
Acta2	CCCTCCAGAACGCAAGTACTC	GCGCTGATCCACAAAACGTTC
Adamts17	CGGGCTTGATCTAGCTTTGTG	GGCTCCTATGTCTCCCTGTG
Alpl	TTGGGCAGGCAAGACACAGA	TTGGCAACCCTGGGTAGACAG
Anks6	CTCTAAGTCCGGCAGACACG	GAATCACGCCCTCGTGTAGA
Boc	GGTGGAAGGACCCTAAGCAC	TCTGGGGCTTTTCTCGTGAC
CD31	TTGTCAAGCGAAGGATAGATAAGA	TGTTGCTGGGTCATTGGAGGTCA
CD34	CCAGCTTATGGAAAAGCACC	CTCCAGCCTTTCTCCTGTAG
Cited2	TTCCTCACTTTCTCCCGTGC	ACTGACGACATTCCACACCC
Cldn4	CTAACTGCTTGCTGGGGACA	GGCTGAGTGAGGGGGATAGA
Col1a1	GTGGGAGGGAACCAGATTG	GCAACAGTCGCTTCACCTAC
Col6a1	GGACACACATGGCATGGAGA	ACGGATAGGTTAGGGGCAGT
Col6a2	CTCTGAACCCAAAGCCCCTT	GTGGATGCATGAAGGGGGAA
Crabp1	ATGCGAGTTCCCCTGAGGAGT	GGTACAATCATGCAAATGCCAAACC
Crim1	AAGCCAGGGAGATGGAAAGC	AGTGTAGCCCCTGGAAATGC
Dct	AGCAGCCAACGACCCTGTGT	CCTTTGCGAAGCCTTCTGTATTG
Dlx2	GGCCACTTTTAGGCCATCCT	TCACGGGGGTAGGTGATAGG
Egr3	AGCACCTCACTGTTTTCCCC	GGGTAGACACAGCCCAGTTC
Enpp2	CGCCCTGATGTCCGTGTATCT	ACGGCTAGTCTTCCGGTAGAAATC
Fbp1	TTGCCATGGATTGTGGTGTCA	CTTGTTGGCGGGGTATAAAAAGAT
Fhl1	AGCTAGGAAGTCCCCAGTGT	TGGAGCCTTTACCAAACCCG
Foxi3	AACTCCATCCGCCACAACCTAT	CTGCCCTTCTGATTTTGATGTCC
Foxn1	CATCTACCCTGGGACCAAACTGA	TGGGGCTTGAACTATATGAACCTCT
Frem1	GCGCGGCACAGCTAGTTTATTAC	TCCTGAGTGGGCATTCAAGTTAGA
Frzb	ATGGCCTGGGGTTAGATCCT	CCTGGTGTTGGGCTACTCTG
Fzd7	GGGGCGAGAGATGGTTTTGA	AGGCTACAGACAGAGCGGTA
Gapdh	CGTAGACAAAATGGTGAAGGTCGG	AAGCAGTTGGTGGTGCAGGATG
GFP	CCGCGCCGAGGTGAAGT	CGCCGATGGGGGTGTTC
Gjb1	TCCCTCTCCCTGAAACGAGG	GATGTGTCCTCTTCTCCCCC
Gpc4	CAGCAGTGCCCTTCGGAGTT	CCCCGCTGTCCCCTCAA
Gpx3	CCCAGTCATTCTACCCCCAC	GGACCATCCCTGGGTTTCAA
Hhip	TCAGTAACGGCCCTTTGGTTG	TGGGCAGGTTGAACTGTGACTC
Hoxc13	GGTGAGGTTGGGAACTAGGC	ATGTCACCCCTCCTCAGTCA
Igf1	CACGCTGTACATCCCCTGATTACT	AGGCCACCTATGTGTTTTTCTGG
ItgA4	CTTGGGTGCAGCCTATAAGGA	TTAGGTCCTTGCTGGCTTGTG
ItgA6	TTTGGAGCCCCAGGGACTTAC	CCCCACTGTGATTGGCTCTTG
ItgA9	GCAGCAGGCACTAAACCCTCTT	AGGCGTGGCATCACAAAAGTC
ItgB8	AGAGTGTGCCATGAAGACTGG	TGAAGCTCAGGGCTGACAGG
K14	GGCCCAGATCCAGGAGATGAT	CAGGGGCTCTTCCAGCAGTATC
K5	GATATCCTTTTGGAGCCCCC	CCAGGAACCATCATGTCTCG
Kitl	TGTTCTTGCTACCCGTGACC	CCCACCATCCAGGCTGAAAT
Klf3	GAAGCTCATGTGGAGGGCTT	AGATGTTGATGGTCCAGCCG
Krtdap	CAACAGAGGGCCTTAACAATGAGTT	TCCATGCTTGCCTCCTCTTCTAC

Lef1	GCTGAGTGCACGCTAAAGGAGAGT	CGACAGAAGGAAGAGGTGGCAGTG
Lgr5	CCCCCAAATGCACAACACTG	CTGTCGTCCTTCCCTGTAC
Lgr6	TCAAAGGCACCACTAGCCTG	GAGACAGCTCCAGGATTCGG
Lhx2	CCTACTACAACGGCGTGGGCACTGT	GTCACGATCCAGGTGTTCAGCATCG
Lrig1	ACGTGAGGCCTTCAATCAGC	AAGGGAACTAACTTGGCGAG
Man1c1	GTGGGTGCAGCTTCTCTGTC	CACTCCTCCATCCGAGTGTG
Mc3r	CAGACGGGACGTAAAAGGGT	CACAGACAGCCCATCCTACG
Mitf	CACTGGGGAGAAGTTGATGTTGATA	AGTGCTGCGGACCATACAGAAA
Mreg	AGGCAGTGGTGTATGCTGAC	TTTCAACAGAACGAGCCCCG
Msx2	TATGCTGCCCTCAGGCTTCAG	TCAAGTGGCCCTGTCGCTTAG
Net1	GTGCTGTCCTCACCTGGTTC	TGTTCCATGCGGATCGTAGG
NFATc1	ACCGATAGCACTCTGGACCT	GTCGGGGAAAGAGACTTGGG
Ovol1	ATGTGAACAGCCGCGTGTATGT	AACCAGTGGGGGTGGAGAAAA
Pax3	ATCGGGGCCCTCCAACC	TGCCCATACTGGTAGCCTGTGA
Plagl1	AGGGTGCTTCTCTTGTGGGA	TACAGAATGCACACACGCTCA
Prr5l	CCTCCCCACTCTGGTGTCTA	CATTGGCTTCTGCCCCCTAA
S100a8	CTTTCGTGACAATGCCGTCTG	GTGGCTGTCTTTGTGAGATGC
Scml4	ATCAGGCAGTTCCCACATACAGAA	TTATTAGAGGCCTCCCCTGACTTG
Sgk1	CAGTGGCGTGAGTGTGCTAT	CACCAACGGCTTTGACTGAC
Shh	ACGAGGATGGAGCCTGTAGTTTGT	GGGTGTGTGTGGCACGCTTTATTT
Sostdc1	CCCCCATCCCAGTCATTTCTT	CAGGGGGATAATTTCACACTGAGA
Sox10	CAAGGGGCCCGTGTGCTA	GCCCGTGCCATGCTAACTCT
Sox18	CTTCATGGTGTGGGCGAAGGAC	CTCCAGCCTCCGGACCTTGC
Sox2	TACTGGCAAGACCGTTTTCGTG	TATTGGAATCAGGCTGCCGAG
Sox9	GGAACAGACTCACATCTCTC	GATCAACTTTGCCAGCTTGC
Spint1	GTCCCTCCACTCTGCAACAG	GCTTGTGGTTCGGTGTAGAGA
Steap4	GGGTAAGAGTTGAGCAGCGA	AGAAAGCACACCACCACTCC
Tgfb2	CAGGCCAGCACTCGTCATTTTA	AGAGCCTCGGCTGAAGAGCAT
Tgfb3	GTGGGTTGGTTAGGGGAAGG	TTCCAGGGGACTTTGGCTTG
Tgm6	GTGGTCCCCCAACTCATGTT	GAGTGGCTGAGTGATCTGGG
Tie2	AAGACATACGTGAACACCAC	AGTCAGAACACACTGCAGAT
Trpm3	CCGCAGACAACTGGAAAAGC	CTCCCATCACAGACCACGAC
Vdr	CAGGCCTGGAGAAAATTTGAACTT	CCCTTGGGCTGAGAAACTGGTA
Wif1	CTGCCGAAATGGAGGTAAATGC	AATTCAGGCCGGCGTTCTAAAG
Wnt5b	CCTCTCCGCCTCACAAAAGT	AGAGGGCCAGCATCTCCTAA

SUPPLEMENTAL REFERENCES

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

Kasprzyk, A. (2011). BioMart: driving a paradigm change in biological data management. Database (Oxford) *2011*, bar049.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Qiao, W., Wang, W., Laurenti, E., Turinsky, A.L., Wodak, S.J., Bader, G.D., Dick, J.E., and Zandstra, P.W. (2014). Intercellular network structure and regulatory motifs in the human hematopoietic system. Mol Syst Biol *10*, 741.

Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J.P. (2006). GenePattern 2.0. Nat Genet *38*, 500-501.

Rendl, M., Lewis, L., and Fuchs, E. (2005). Molecular dissection of mesenchymal-epithelial interactions in the hair follicle. Plos Biology *3*, 1910-1924.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

Turner, B., Razick, S., Turinsky, A.L., Vlasblom, J., Crowdy, E.K., Cho, E., Morrison, K., Donaldson, I.M., and Wodak, S.J. (2010). iRefWeb: interactive analysis of consolidated protein interaction data and their supporting evidence. Database (Oxford) *2010*, baq023.