

Online Only

Supplementary Table S3: Functional categories for basal and luminal expressed genes

Functional Categories	P-Value	Example genes
<b>BASAL LAYER</b>		
Positive regulation of physiological process	0.0381	<i>IL6, VEGFC, CBL, JARID1A, CIDEA, CIDEA,</i>
Receptor binding	0.0479	<i>TAPBP, IL6, VEGFC, DOCK2, CXCL1, INHBC, NCOR2,</i>
Positive regulation of biological process	0.0480	<i>IL6, VEGFC, CBL, JARID1A, CIDEA, CIDEA, HTR2B,</i>
Protein binding	0.0008	<i>DOCK2, ITGAV, CRHBP, DICER1, CXCL1, TEAD1, IL6, CBL, RFP</i>
RNA processing	0.0127	<i>MTO1, C1ORF19, DICER1, SNRPB2, NOLC1, TXNL4A, SFRS1,</i>
Enzyme regulator activity	0.0456	<i>TRIO, ITIH4, DOCK2, ARL1, A2M, CXCL1, RGS4, RGS11,</i>
RNA metabolism	0.0343	<i>MTO1, C1ORF19, DICER1, SNRPB2, NOLC1, TXNL4A, SFRS1,</i>
Protein complex assembly	0.0170	<i>TAPBP, A2M, CREBBP, TXNL4A, SFRS1, KIFAP3,</i>
Positive regulation of cellular process	0.0231	<i>IL6, VEGFC, CBL, JARID1A, CIDEA, CIDEA, HTR2B,</i>
Positive regulation of cellular physiological process	0.0322	<i>IL6, VEGFC, CBL, JARID1A, CIDEA, CIDEA,</i>
Helicase activity	0.0388	<i>C9ORF102, DICER1, SMARCA3, CHD2,</i>
Growth factor activity	0.0131	<i>IL6, VEGFC, CXCL1, INHBC,</i>
Ligase activity, forming carbon-nitrogen bonds	0.0367	<i>RAPSN, UBE2I, CBL, RFP, SMARCA3, HAL,</i>
Cellular process	0.0377	<i>GSTZ1, ITGAV, CXCL1, TEAD1, NOLC1, IFNAR2, IL6, C1ORF19, FKBP6</i>
<b>LUMINAL LAYER</b>		
Calmodulin binding	0.0030	<i>EWSR1, SPTBN1, CAMKK2, MYH10, IQGAP2, PPP3CA, MYLK,</i>
Nucleotide binding	0.0346	<i>SYK, PDK1, HNRPU, RDBP, SNRK, PABPC5, ATAD1, PRKCI, FGFR2, SFRS11</i>
Regulation of biological process	0.0401	<i>CDKN1B, NBL1, CDKN1A, SNAP25, SPRY1, ITGB3BP, ENO1, CUL3, SOX6, PHB</i>
Protein transporter activity	0.0162	<i>ARFGAP3, A2M, KPNA6, EIF4ENIF1, IPO11,</i>
Hydrolase activity	0.0230	<i>PLA2G12B, ATAD1, SHFM1, ASAH1, SND1, HDHD2, CDC14A, ASTE1, ENO1</i>
Protein transport	0.0488	<i>A2M, ARFGAP3, PRKCI, VPS35, SEC23B, CADPS, SCAMP2, KPNA6, RANBP6, RANBP3</i>
Ion binding	0.0496	<i>PLA2G12B, CDKN1A, RNF144, BNC2, AMFR, ZNF228, RNF167, ENO1, ENO2</i>
Cell cycle	0.0348	<i>CCNK, CDKN1B, CUL3, CDC14A, NBL1, PLCB1, PRM2, HRASLS3, CDKN1A, PHB</i>
Cellular physiological process	0.0015	<i>NBL1, CDKN1A, ITGB3BP, SYK, CCNK, CUL3, FGFR2, SOX6, SNAP25, ENO1, PHB</i>
Alcohol metabolism	0.0110	<i>PDK1, TPI1, UGP2, N-PAC, ALDH9A1, PFKM, ENO1, FBP1, ENO2, SCAP,</i>
Cellular localization	0.0463	<i>TUBA1, VPS4B, A2M, ARFGAP3, PRKCI, MAL, VPS35, SEC23B, TUBE1, SCAMP2</i>
Protein localization	0.0368	<i>ARFGAP3, PRKCI, MAL, VPS35, SEC23B, CADPS, SCAMP2, KPNA6, RANBP6, RANBP3</i>
Enzyme regulator activity	0.0096	<i>CDKN1B, ARL2BP, SERPINE2, PLN, ARFGAP3, IQGAP2, PPP1R12B</i>
Protein binding	0.0000	<i>CDKN1B, PLN, NBL1, CDKN1A, SNAP25, ITGB3BP, RPL24, SYK, CCNK, IQGAP2</i>
Transcriptional activator activity	0.0046	<i>PHB, SMARCA4, TRIP4, THRAP1, BRD8, FHL2, TCERG1, CRSP7, PRRX1,</i>
Carbohydrate metabolism	0.0317	<i>PDK1, UGP2, TPI1, N-PAC, PFKM, SDHC, OXCT1, TDG, SLC3A2, MGAM</i>
Transcription factor binding	0.0242	<i>SMARCA4, TRIP4, THRAP1, BRD8, FHL2, TCERG1, CRSP7, SND1, PRRX1,</i>
Actin binding	0.0231	<i>SPTBN1, MYH10, ABLIM1, IQGAP2, CAPZA1, MSN, KLHL5, PPP1R9A,</i>
Cell proliferation	0.0384	<i>C2ORF29, SYK, CDKN1B, CAPNS1, CUL3, CDC14A, CDKN1A, PHB, KLF11, RHOG</i>
Muscle contraction	0.0377	<i>KCNH2, PLN, VIPR1, PPP1R12B, MRCL3, MRCL2, C4A,</i>
Cell organization and biogenesis	0.0024	<i>A2M, H2AFV, PRKCI, FGFR2, SOX6, PHB, SPTBN1, BRD8, ABLIM1, SATB1</i>
Cell motility	0.0101	<i>SYK, CSPG3, AMFR, SERPINE2, CAPZA1, RTN4, MSN, PSG7, TSPAN2,</i>
Organic acid metabolism	0.0488	<i>ABAT, BBOX1, TPI1, LTA4H, GCH1, PTGES, SHMT1, ENO1, GNPAT, MTHFR</i>

In order to identify key biological processes, basal and luminal gene lists were individually entered into the web based ontology program DAVID (<http://david.abcc.ncifcrf.gov/>). This program categorises genes based on their biological or molecular function and assigns a p-value to each category to indicate its overall significance. Filtering on a p-value less than or equal to 0.05 returned 14 basal cell categories and 23 luminal categories. Table S3 shows the final list of categories for both populations, including up to 10 example genes from each category and the category p-value.