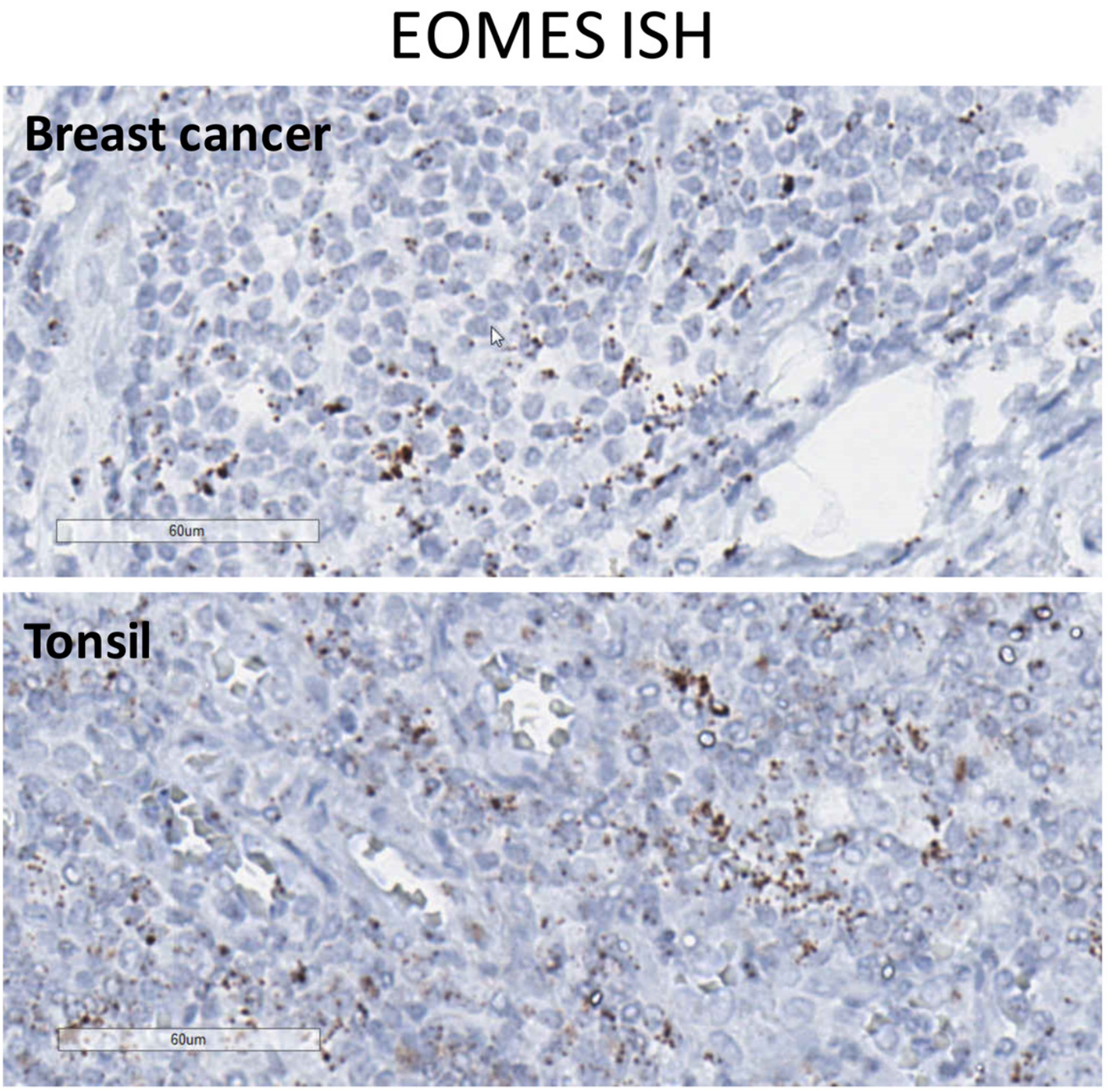
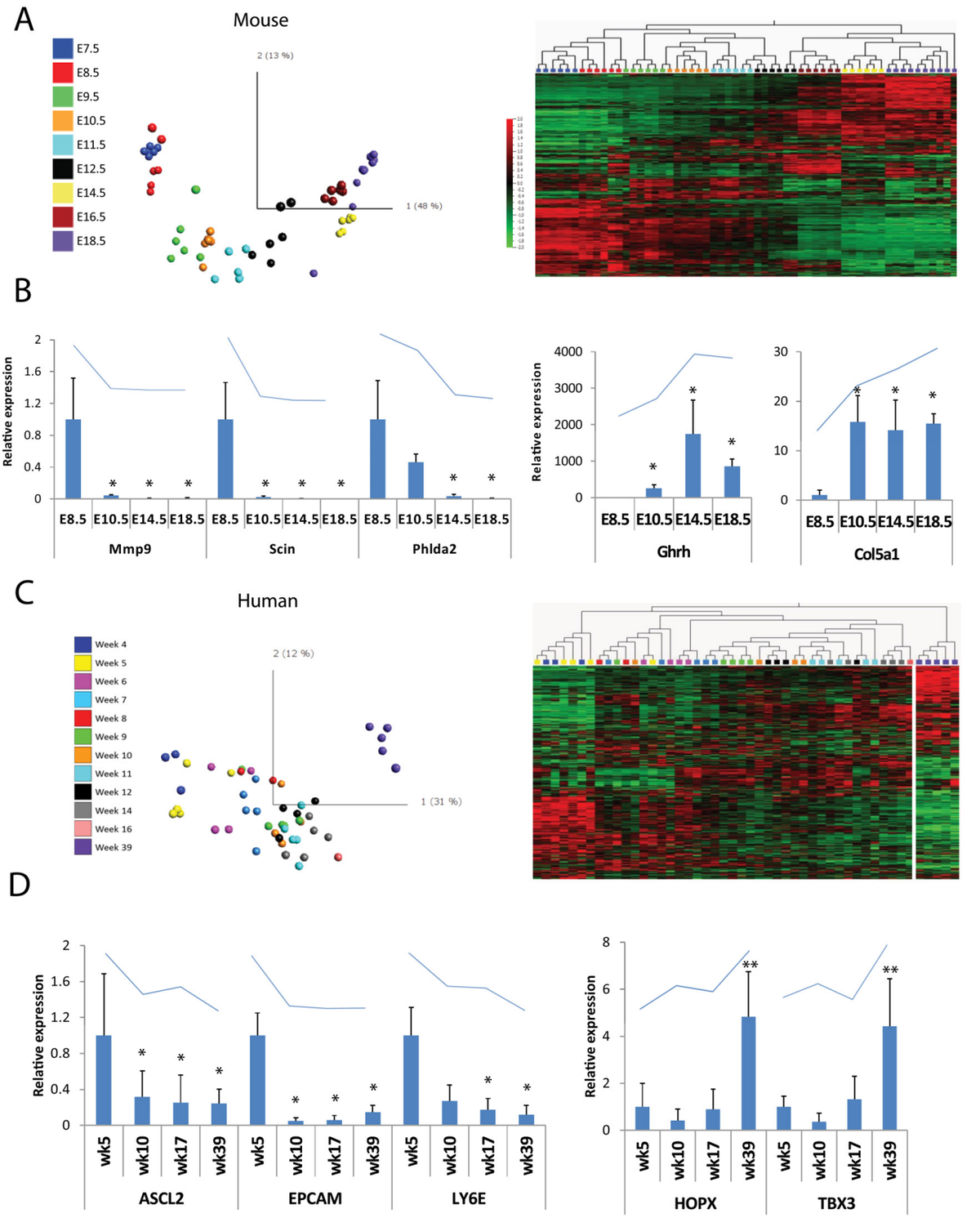
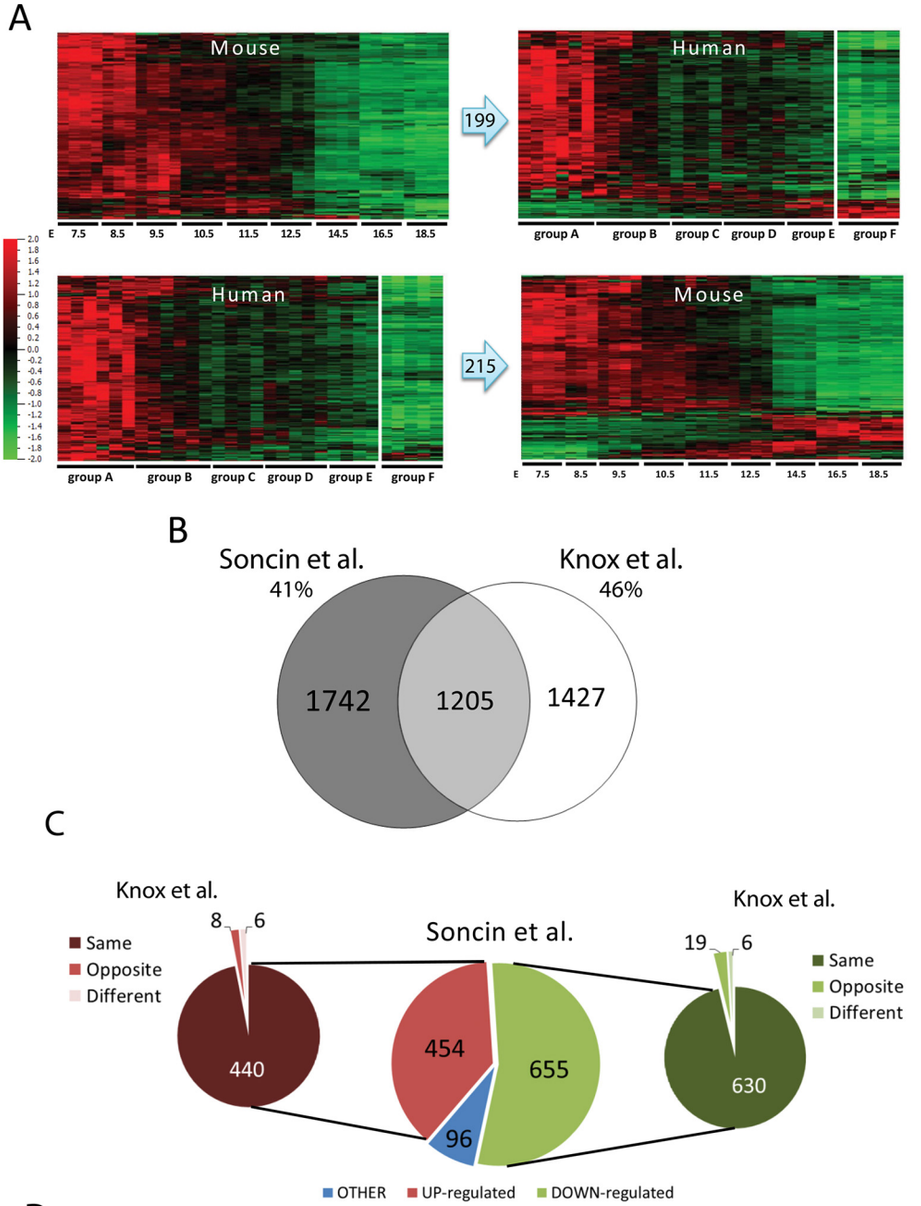


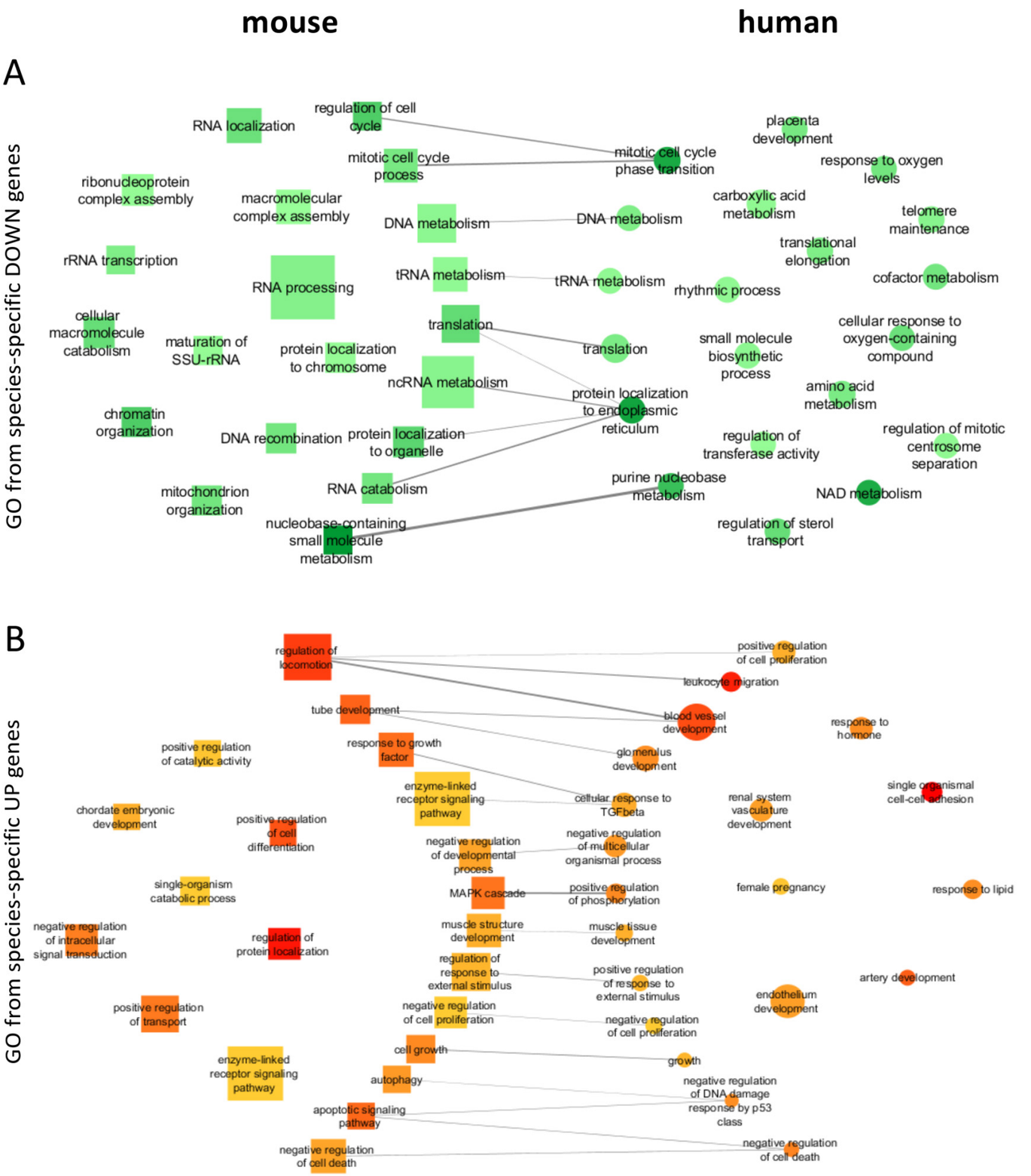
Supplementary Figure S1. In-situ hybridization using EOMES-specific probes on positive control human tissues.



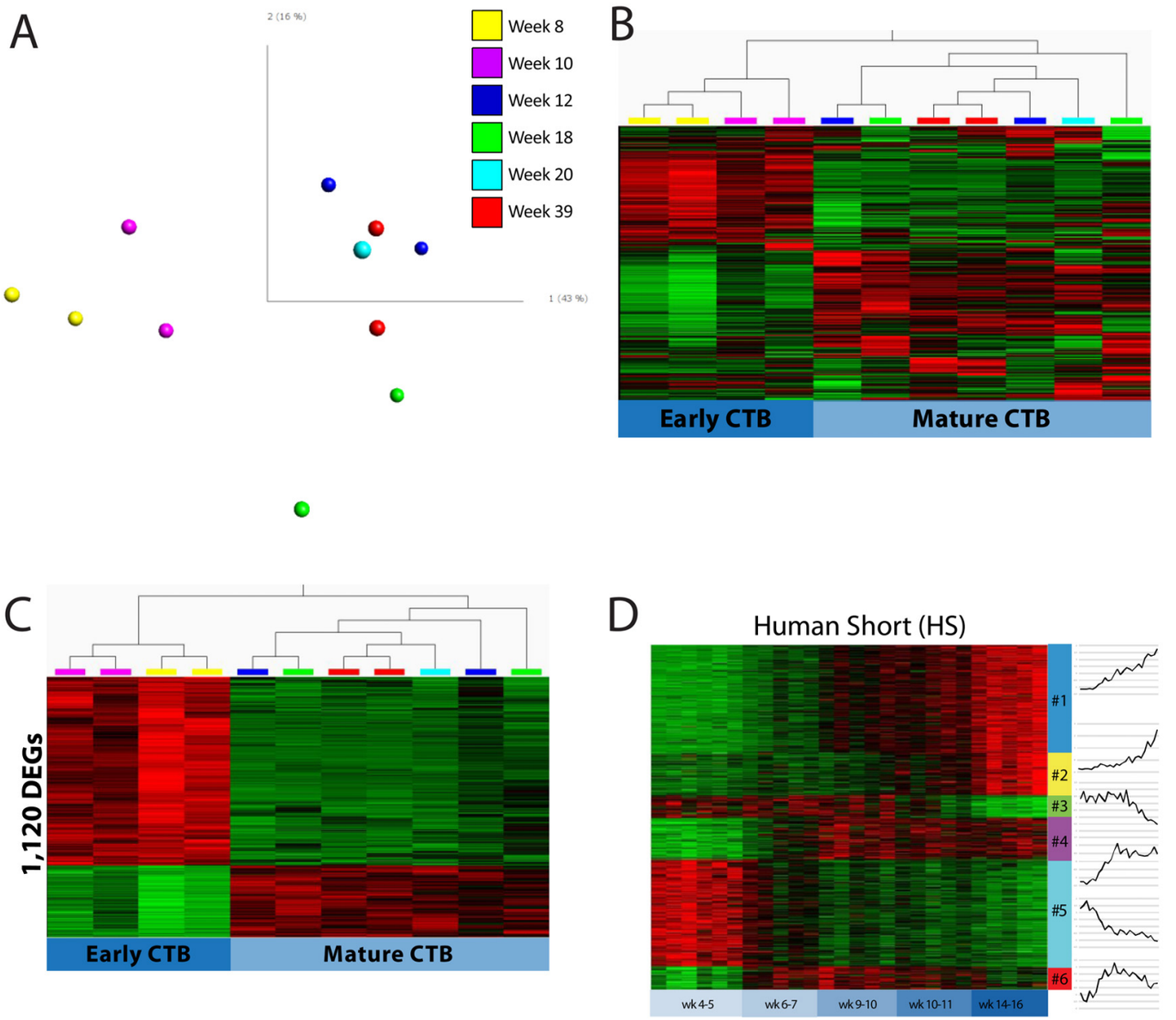
Supplementary Figure S2. Microarray data analysis of mouse and human placenta samples. A. Principle component analysis (PCA) and heatmap of mouse dataset using Qlucore (variance 0.02). Samples were colour-coded according to gestational age. **B.** qRT-PCR analysis (bar graph) for Mmp9, Scin, Phlda2, Ghrh, and Col5a1 in mouse placentas at E8.5, E10.5, E14.5 and E18.5 confirmed expression pattern observed in the microarray (line above bar graph). **C.** PCA and heatmap of human dataset using Qlucore (variance 0.02). Samples were colour-coded according to gestational age. **D.** qRT-PCR analysis (bar graph) for ASCL2, EPCAM, LY6E, HOPX, and TBX3 in human placenta from gestational week 5, 10, 17, and 39 confirmed expression pattern observed in the microarray (line above bar graph). n=3 * $p<0.05$ compared to E8.5 for mouse and Week 5 for human; ** $p<0.05$ compared to week 10.



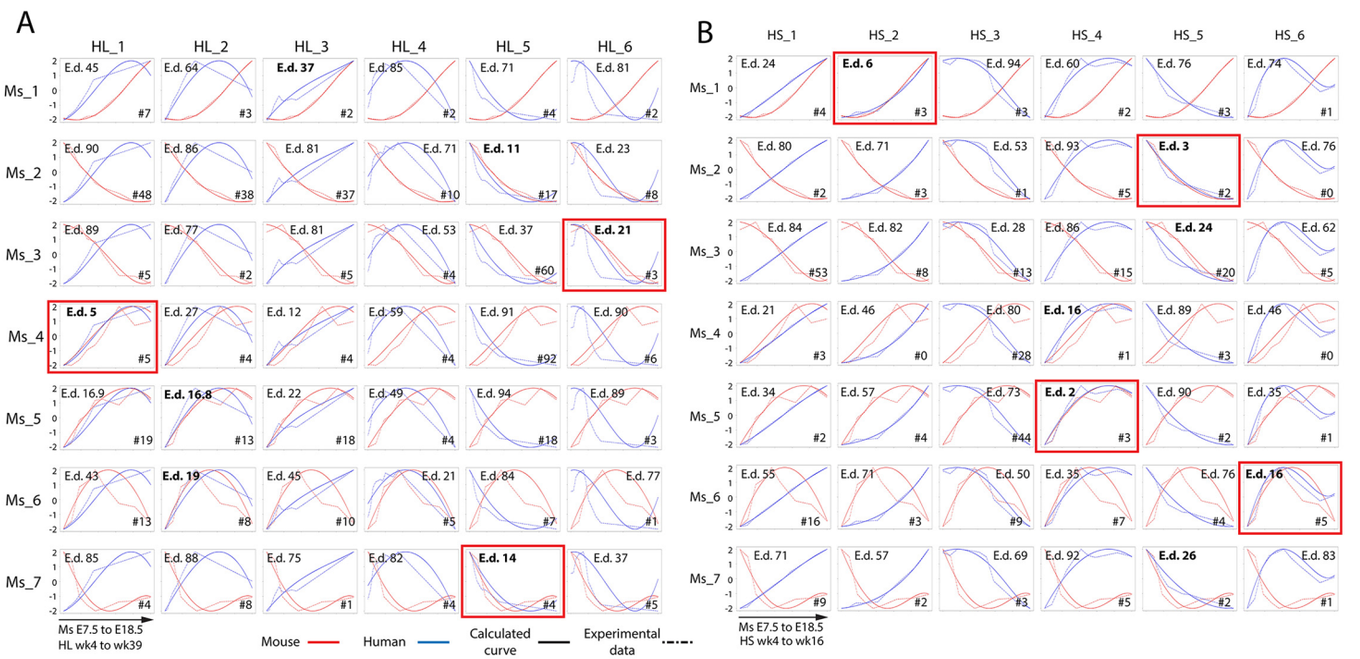
Supplementary Figure S3. Comparison between datasets. A. Top: Heatmap of mouse differentially expressed genes (DEG) downregulated with gestational age (199 genes, left) and their expression pattern in human (heatmap on the right). Bottom: Heatmap of human DEGs (215 genes, left) downregulated with gestational age and their expression pattern in mouse (heatmap on the right). These data were used to create the pie chart in Fig.3D. **B.** Venn diagram representing the overlapping of DEGs between our mouse data and the Knox et al. dataset. We analysed the mouse dataset published by Knox and Baker (2008) using similar parameters applied to our inter-species analysis (v 0.02, $q<0.05$, FC>2.0). We identified 2,632 DEGs in the Knox et al. dataset, of which 1,205 were in common with our analysis, representing over 40%. **C.** Pie charts showing the comparison in expression pattern direction of the common DEGs (1,205) between our data and that of Knox et al. Central pie chart shows the expression pattern in the current study. Genes with expression pattern other than UP and DOWN-regulation across gestation were combined and labelled as "Other." The side charts show the expression pattern comparison between the sub-classes UP- (red) and DOWN- (green) regulated genes in our analysis and their expression in the Knox et al. dataset. Over 96% of the common DEGs showed the same pattern of expression (either UP or DOWN) in both datasets. **D.** Expression levels of CDX2 and ELF5 in human placentas across gestation from our microarray dataset. Note the high variability of CDX2 expression between samples.



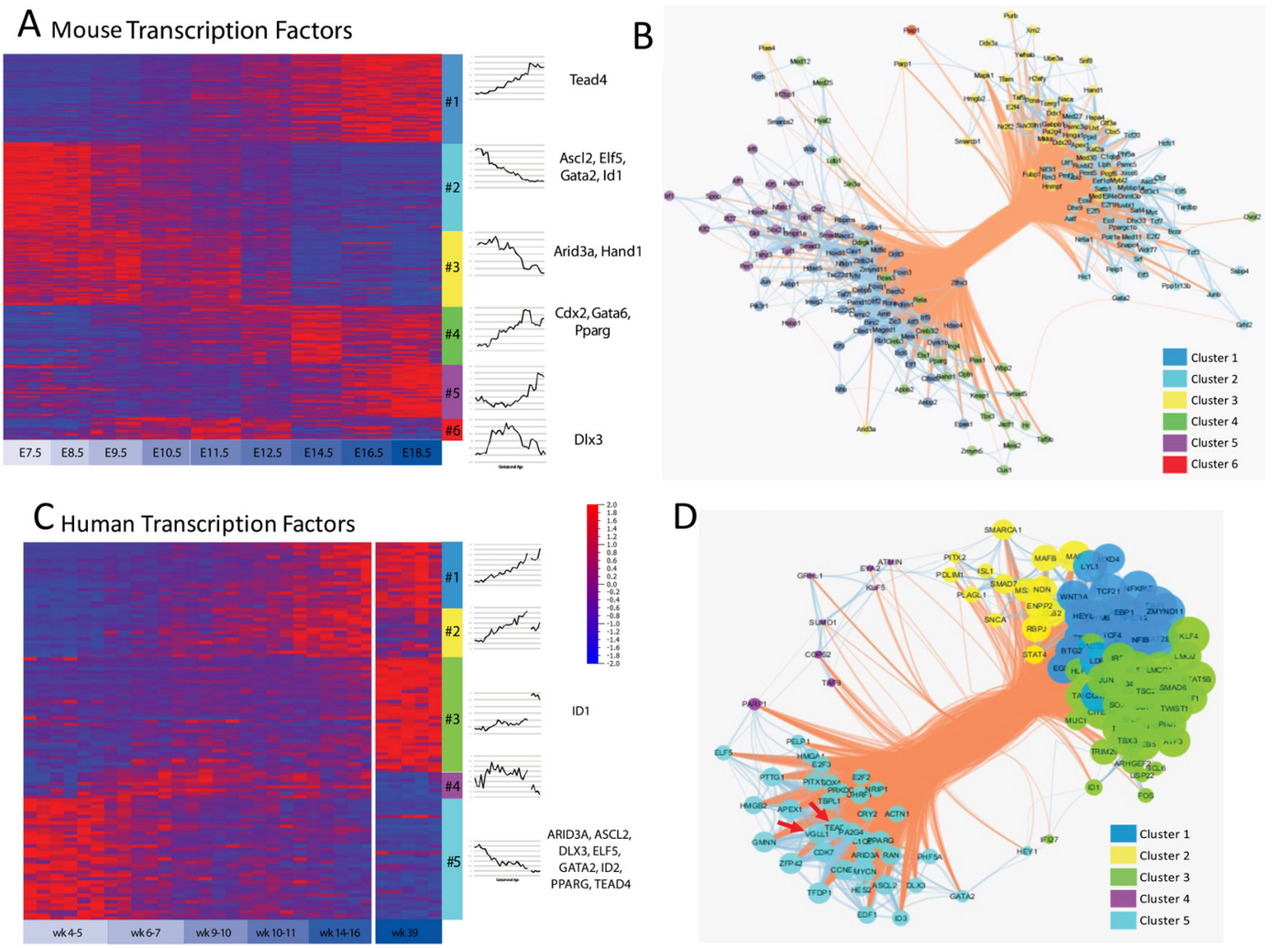
Supplementary Figure S4. Enrichment analysis of Gene Ontology Biological Functions of species-specific up and down-regulated genes in mouse and human placentas. A. Gene ontology of genes downregulated across gestation. **B.** Gene ontology of genes upregulated across gestation. Node shape represents the species (mouse – square; human – circle). The size of the nodes is correlated to the log(p) of the enriched term. Lines (edges) connect mouse and human terms that include overlapping functions. The thickness of the edges is proportional to the number of terms in common between the species.



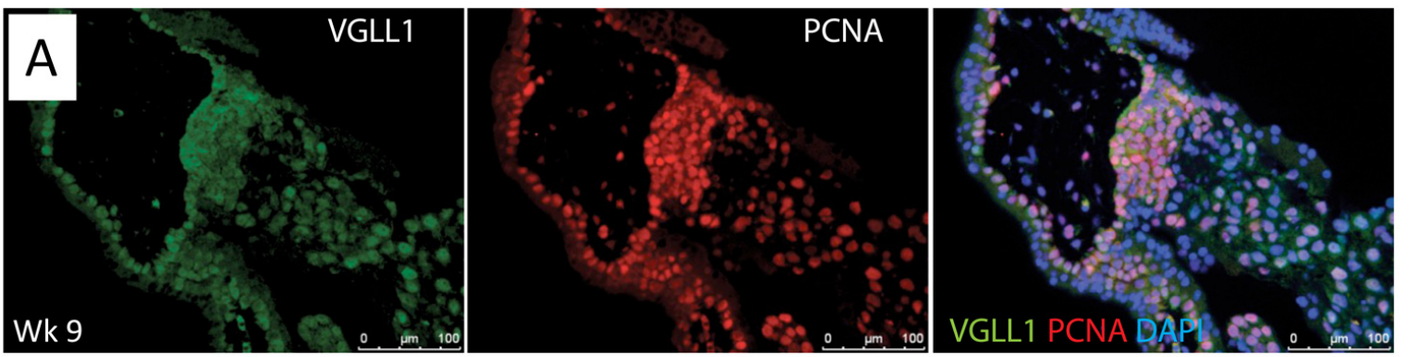
Supplementary Figure S5. Details of microarray data analysis of human primary CTB and placental tissues. A. Principal component analysis (PCA) of primary CTB expression profiles. Note separation into two main groups along PCA1: weeks 8 and 10 on one side and weeks 12, 18, 20, and 39 on the other. **B.** Heatmap of human primary CTB genes filtered for variance (0.05). Note hierarchical clustering of samples into two groups, labelled "early" CTB and "mature" CTB, which were used for subsequent differential expression analysis. **C.** Heatmap of 1,120 differentially expressed genes (DEGs) between the two identified groups of CTB (two-group analysis $q<0.01$). Of these DEGs, 197 genes were also differentially expressed in human placentas across gestation and included in the AP analysis (see Figure 6B). **D.** Affinity Propagation analysis of the 665 DEGs in the "short" human dataset, which did not include term samples. AP distinguished 6 clusters of co-regulated genes. Heatmap represents expression values of genes; line graphs (right) represent the average trace for each cluster.

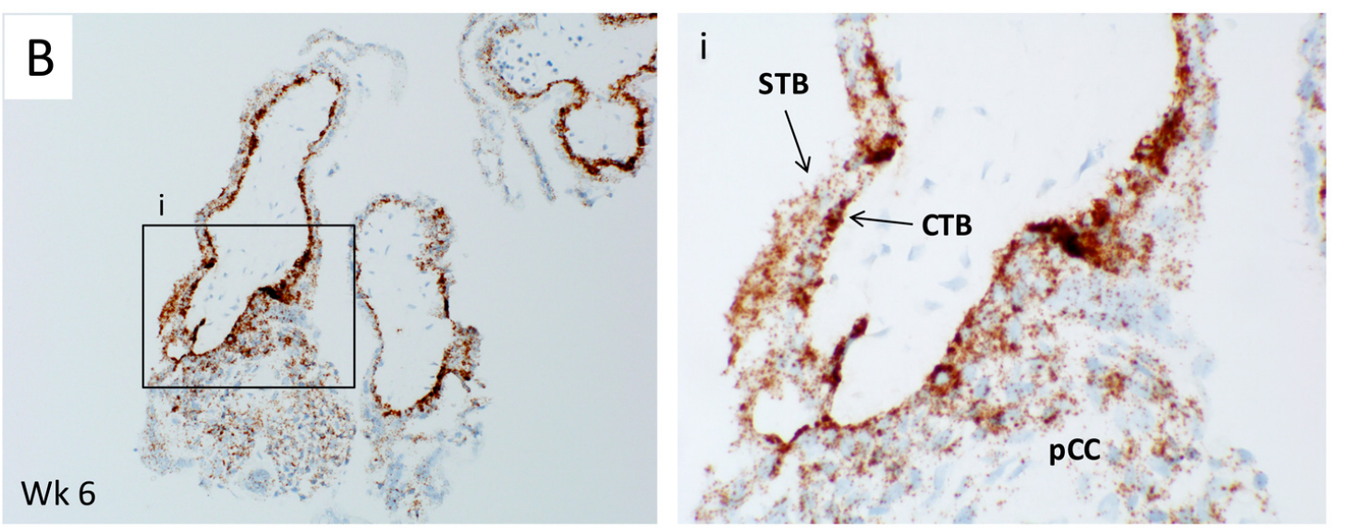


Supplementary Figure S6. Comparison of expression profiles of all mouse and human co-expression clusters. A. Matrix of graphs showing the relationship between each mouse and human AP co-expression cluster (from the HL dataset, which included data from term human placentae). **B.** Matrix of graphs showing the relationship between each mouse and human AP co-expression cluster (from the HS dataset, which did not include data from term human placentae). For both matrices, both the experimental data (dashed lines) and corresponding best-fit curves (solid lines) for each AP cluster are shown. Data from mouse clusters are shown in red and data from human clusters are shown in blue. The Euclidean distance (E.d) values for each pair of curves is shown. # = number of genes in common between the mouse and human clusters. For each mouse cluster, the HL or HS cluster with the most similar pattern of expression of gestation, as measured by the Euclidean distance, is boxed in red.

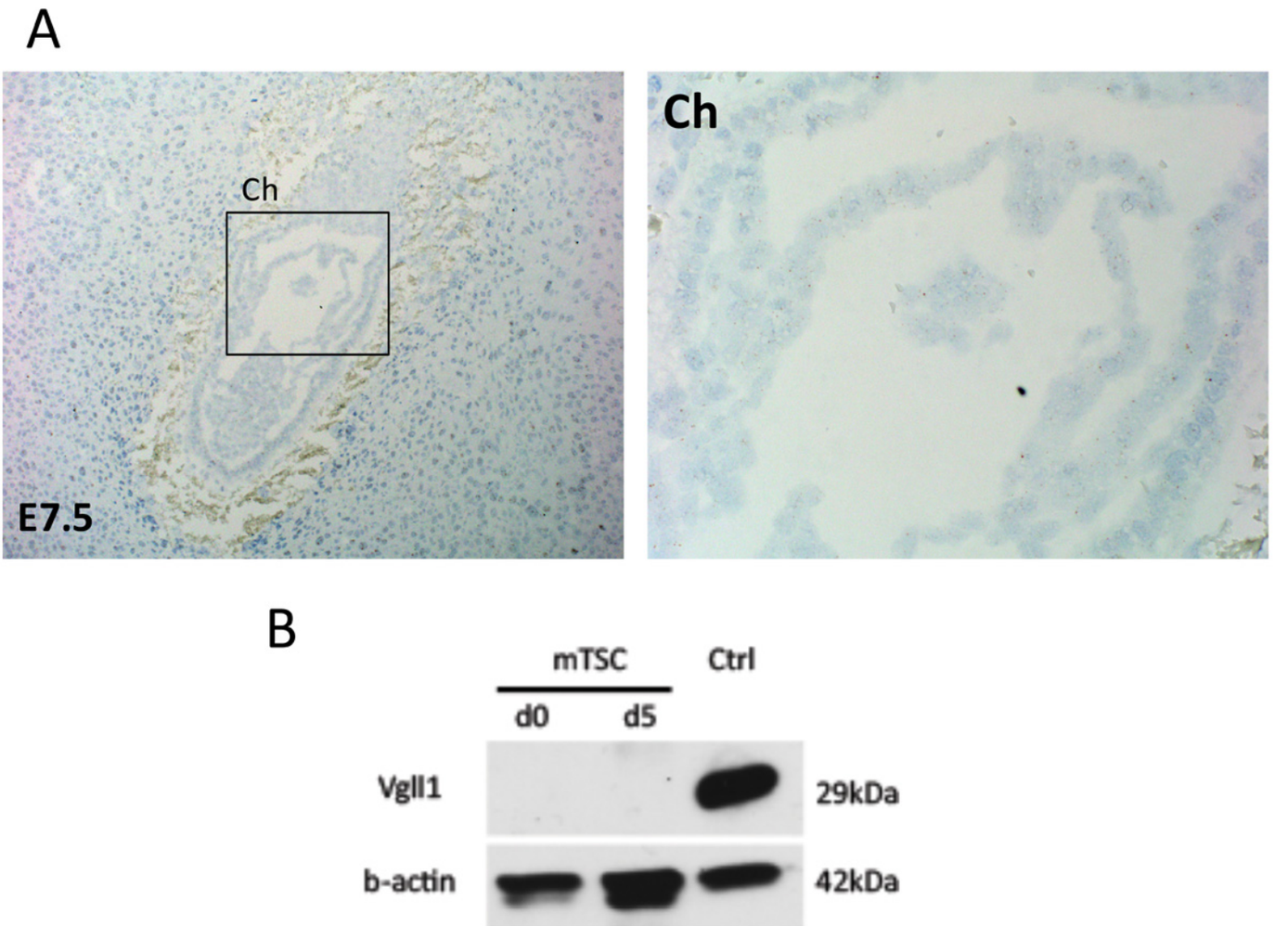


Supplementary Figure S7. A. Affinity Propagation (AP) analysis of transcription factors (TFs) in the mouse dataset distinguished 6 clusters of co-regulated genes. Heatmap represents expression values of TFs; line graphs (right) represent the average trace for each cluster. Representative gene(s) in each cluster are shown. **B.** Mouse TF cluster network. For clarity, node size was set at a fixed value. Nodes are colour-coded according to the cluster they belong to; red edges indicate negative correlation, while blue edges show positive correlation between nodes. **C.** Affinity Propagation (AP) analysis of transcription factors in the human dataset distinguished 5 clusters of co-regulated genes. Heatmap representing expression values of TFs; line graphs (right) represent the average trace for each cluster. Representative gene(s) in each cluster are shown. **D.** Human TF cluster network. Nodes are colour-coded according to the AP cluster number, while their size represents the positive sum scores. Red edges indicate negative correlation, while blue edges show positive correlation between nodes. Edge thickness identifies Pearson correlation absolute value. Red arrows show the position of TEAD4 and VGLL1.

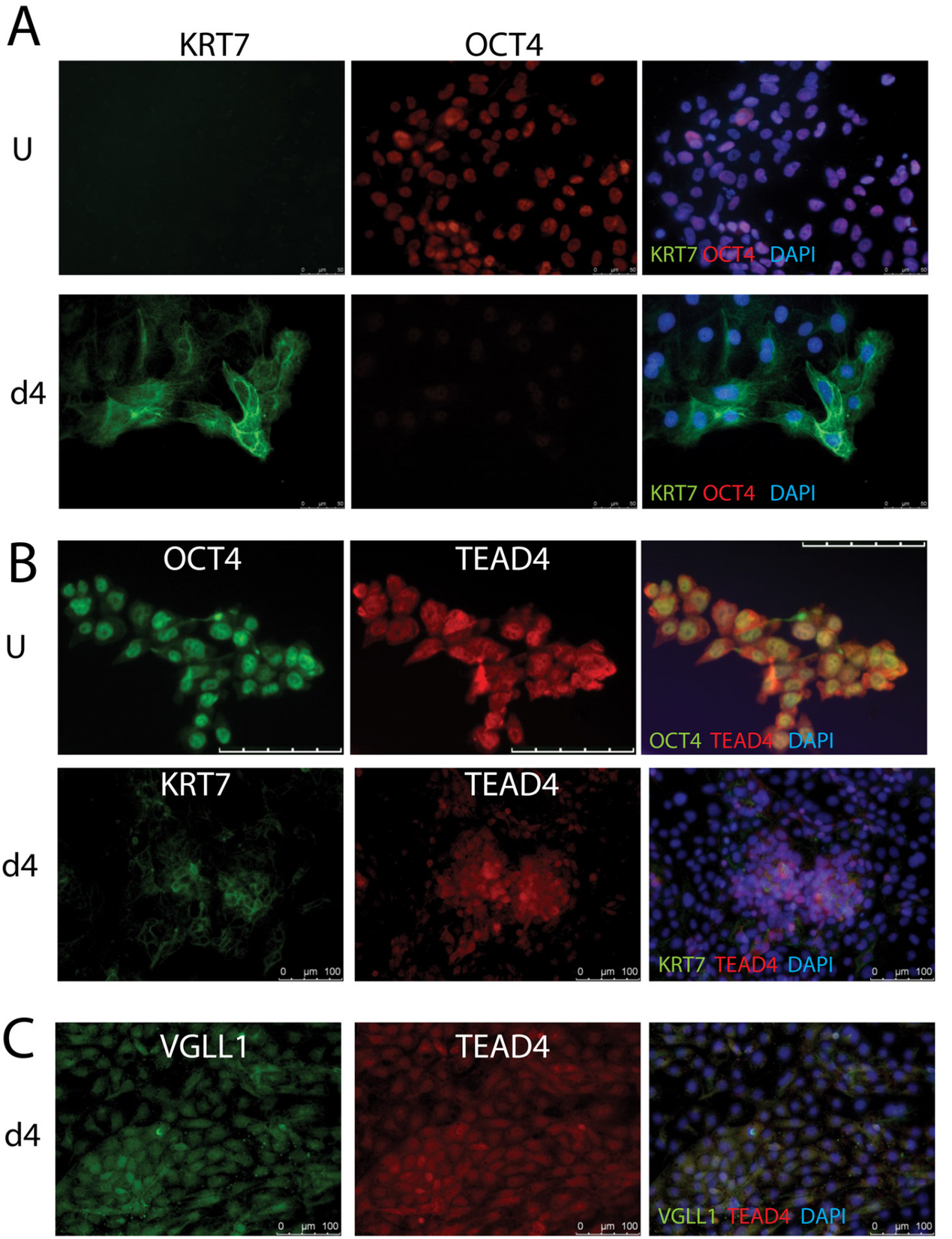




Supplementary Figure S8. IHC and ISH of first trimester human placentas. A. Co-staining of week 9 placenta with VGLL1 (green), the proliferation marker PCNA (red), and Dapi (blue). **B.** In situ hybridization with VGLL1-specific probes in week 6 placenta. Expression of VGLL1 mRNA is noted in cytotrophoblast (CTB), syncytiotrophoblast (STB), as well as in trophoblast of the proximal cell column (pCC), although it is most enriched in CTB. Magnification: Left – 100x; Right: 300x.



Supplementary Figure S9. Vgll1 is not expressed in mouse placenta and trophoblast. A. No Vgll1 RNA was noted in E7.5 placentas by in-situ hybridization. Ch= chorion. Magnification: 50x (left); 200x (right). **B.** Western blot analysis of undifferentiated (d0) and differentiated (day 5/d5) mouse trophoblast stem cells (mTSC) for Vgll1 and β-actin (loading control). Commercially available mouse whole stomach lysate (Abcam) was used as positive control.



Supplementary Figure S10. Expression of OCT4, KRT7 and TEAD4 in the hESC-based *in vitro* model of human trophoblast differentiation. A. Immunostaining for OCT4 (red) and KRT7 (green) in pluripotent (undifferentiated) H9 hESC (U) and in hESC-derived CTB-like cells (d4). OCT4 and KRT7 in this model are mutually exclusive. **B.** Immunostaining for TEAD4, along with either OCT4 in undifferentiated hESC (U) or with KRT7 in hESC-derived CTB-like cells (d4). **C.** Double immunostaining of hESC-derived CTB-like cells (d4) with VGLL1 and TEAD4, showing co-localization in nuclei of these cells.

Table S1. Genelists

Click here to Download Table S1

Table S2. GO enrichment

Click here to Download Table S2

Table S3. q-RT-PCR primers

Gene_ID	Sequence
h18S F	CGCCGCTAGAGGTGAAATTCT
h18S R	CGAACCTCCGACTTTCGTTCT
hASCL2 F	CACTGCTGGCAAACGGAGAC
hASCL2 R	AAAACTCCAGATAGTGGGGGC
hCDX2 F	TTCACTACAGTCGCTACATCACC
hCDX2 R	TTGATTTTCCTCTCCTTTGCTC
hEPCAM F	AATCGTCAATGCCAGTGTACTT
hEPCAM R	TCTCATCGCAGTCAGGATCATAA
hHOPX F	GACAAGCACCCGGATTCCA
hHOPX R	GTCTGTGACGGATCTGCACTC
hLY6E F	GGGAATCTCGTGACATTTGGC
hLY6E R	ACACCAACATTGACGCCTTCT
hOCT4 F	TGGGCTCGAGAAGGATGTG
hOCT4 R	GCATAGTCGCTGCTTGATCG
hTBX3 F	CCCGGTTCCACATTGTAAGAG
hTBX3 R	GTATGCAGTCACAGCGATGAAT
hTEAD4 F	CAGTATGAGAGCCCCGAGAA
hTEAD4 R	TGCTTGAGCTTGTGGATGAA
hTP63 F	CTGGAAAACAATGCCCAGA
hTP63 R	AGAGAGCATCGAAGGTGGAG
hVGLL1 F	CTCCCGGCTCAGTTCACTATAA
hVGLL1 R	CCCAGTGGTTTGGTGGTGTA
m18S F	CGCGGTTCTATTTTGTTGGT
m18S R	AACCTCCGACTTTCGTTCTTG
mCol5a1 F	AGGGAGCCAGAATCACTTCTT
mCol5a1 R	GCATCCACATAGGAGAGCAGT
mGhrh F	CACAACATCACAGAGTCCCACC
mGhrh R	CTGGTGAGGATGAGGATCACAA
mMmp9 F	GCCGACTTTTGTGGTCTTCC
mMmp9 R	TACAAGTATGCCTCTGCCAGC
mPhlda2 F	TGCAAGACTTTCCCCGCTAC
mPhlda2 R	GTGCGTTTCACGGACCCA
mScin F	CCAGAATTGTGGAGGTTGACG
mScin R	GCCATTGTTTCGTGGCAGTT

Table S4. Model errors

Cluster #	Error_value	Cluster #	Error_value	Cluster #	Error_value
Ms_1	1.7058	**HL_1**	0.2514	**HS_1**	0.2867
Ms_2	0.2037	**HL_2**	0.4030	**HS_2**	0.5153
Ms_3	0.6422	**HL_3**	1.5269	**HS_3**	1.1920
Ms_4	1.2510	**HL_4**	1.8015	**HS_4**	1.2262
Ms_5	1.6053	**HL_5**	1.7795	**HS_5**	1.3352
Ms_6	1.4172	**HL_6**	2.2254	**HS_6**	1.3348
Ms_7	2.6252				