

Supplementary information

Benchmarking brain organoid recapitulation of fetal corticogenesis

Cristina Cheroni^{1,2,3,#}, Sebastiano Trattaro^{1,2,3,#}, Nicolò Caporale^{1,2,3}, Alejandro López-Tobón^{1,2,3,4}, Erika Tenderini^{1,5}, Sara Sebastiani^{1,6}, Flavia Troglio^{1,2,3}, Michele Gabriele^{1,2,7}, Raul Bardini Bressan⁸, Steven M Pollard⁸, William T Gibson^{9,10}, Giuseppe Testa^{1,2,3,*}

¹ Department of Experimental Oncology, European Institute of Oncology IRCCS, Milan, Italy.

² Department of Oncology and Hemato-Oncology, University of Milan, Via Santa Sofia 9, 20122, Milan, Italy.

³ Human Technopole, Viale Rita Levi-Montalcini 1, 20157, Milan, Italy.

⁴ Current address: Herophilus, Inc., San Francisco, CA 94107, USA.

⁵ Current address: Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Naples, Italy, Via Campi Flegrei, 34 - 80078.

⁶ Current address: Department of Molecular Neurobiology, Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen, The Netherlands.

⁷ Current address: Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; The Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA; Koch Institute for Integrative Cancer Research, Cambridge, MA 02139, USA.

⁸ Centre for Regenerative Medicine, Institute for Regeneration and Repair, and Edinburgh Cancer Research UK Centre, University of Edinburgh, Edinburgh, UK. EH16 4UU.

⁹ BC Children's Hospital Research Institute, 950 West 28th Avenue, Vancouver, BC, Canada V5Z 4H4.

¹⁰ Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada V6T 1Z4.

These authors contributed equally.

*Correspondence: giuseppe.testa@fht.org; giuseppe.testa@unimi.it; giuseppe.testa@ieo.it.

Content:

- Supplementary figure and table legends
- Supplementary figures

Supplementary Figure 1

(A-B) Correlation analysis of BrainSpan prenatal and postnatal cortices. The heatmaps show the correlation coefficient calculated by Spearman correlation across stages, for prenatal (A) and postnatal (B) timepoints, respectively. (C-D) Top-35 genes according to positive and negative loading values for the first (C) and second (D) components from PCA analysis. Results of gene ontology enrichment analysis performed for the Biological Process domain on the top genes selected on the basis of their loading value are visualized in the bar plots.

Supplementary Figure 2

(A) Gene dendrogram generated on the Topological Overlap Dissimilarity matrix for BS cortex samples; each branch of the dendrogram corresponds to a module of highly interconnected genes, for a total of 17 modules. The row below the dendrogram depicts the module assignment, each identified by a different colour. (B) Sample dendrogram showing the clustering of Brainspan samples included in WGCNA analyses. The rows below the dendrogram depict post-conceptional week associated to each sample, either as continuous or categorical variable (first and other rows respectively). (C) Brainspan WGCNA diagnostic plots for the selection of beta parameter (set to 18) according to mean connectivity and scale independence values. (D) Ribbon plot, GO enrichment analysis and network reconstruction for BS_Midnightblue module.

Supplementary Figure 3

(A) Widefield images showing the expression of apical progenitor (PAX6) and neural stem cell (SOX2) markers in day 25 CBO of a representative control line. (B) Widefield images showing the expression of cell cycle (KI67), neural stem cell (SOX2), apical progenitor (PAX6), intermediate progenitor (EOMES) and low layer neuron (BCL11B) markers in day 50 CBO of representative control lines.

Supplementary Figure 4

(A) Widefield images showing the expression of apical progenitor (PAX6), neural stem cell (SOX2), low layer neuron (BCL11B), up layer neuron (SATB2) and astrocyte (GFAP) markers in day 100 CBO of representative control lines.

Supplementary Figure 5

(A) PCA on the CBO cohort, with differentiation stage shown as dot color. (B-C) Top-35 genes according to positive and negative loading values for the first (B) and second (C) components from PCA analysis and barplots showing functional enrichment results for the Biological Process domain on the top genes. (D-E) Functional analysis of up-regulated (D) and down-regulated (E) genes for Day150 vs Day100 and Day200 vs Day150 stage-wise differential expression analysis on CBO. (F) Scatterplots representing the relationship of the fold-change across successive time windows. Genes resulting differentially expressed ($FDR < 5\%$ and absolute $\log_2FC > 1$) in both the two examined comparisons are reported in yellow, while the ones specific for one of the two comparisons are in blue or purple. Gene symbols are reported for the top 10 protein-coding genes for each quadrant according to fold-change.

Supplementary Figure 6

(A) Gene dendrogram generated on the Topological Overlap Dissimilarity matrix for CBO samples; each branch of the dendrogram corresponds to a module of highly interconnected genes, for a total of 14 modules. The row below the dendrogram depicts the module assignment, each identified by a different colour. (B) Sample dendrogram showing the clustering of CBO samples. The rows below the dendrogram depict the differentiation day associated to each sample, either as continuous or categorical variable (first and other rows respectively). (C) CBO WGCNA diagnostic plots for the selection of beta parameter (set to 15) according to mean connectivity and scale independence values. (D-E) Ribbon chart, bubble plot and network reconstruction for the green (D) and red (E) modules.

Supplementary Figure 7

(A) Schematic representation of the differentiation steps of the four BO protocols analyzed. (B-D) PCA of the MGO (B), FO (C) and TA (D) datasets.

Supplementary Figure 8

Overlap between genes retrieved as differentially expressed by stage-wise differential expression analysis in cortical brain organoids (CBO) compared to external brain organoid dataset (Minimally Guided Organoids in panel **A**; Forebrain Organoids in panel **B**; Telencephalic Aggregates in panel **C**). P-value and Odds Ratio are reported for the overlap with P-value < 0.01 and Odds Ratio > 3. Cell color is assigned according to Odds Ratio values. **(D)** Heatmap depicting the results of the whole-transcriptome correlation analysis between CBO data and external BO protocols.

Supplementary Figure 9

Lollipops showing the expression levels of gene signatures related to brain areas, off-target tissues and cell stress in BS **(A)**, CBO **(B)**, MGO **(C)**, FO **(D)** and TA **(E)**. Expression levels (Log2Fpkms or Log2Rpm) along timepoints are reported in each dataset as the mean value across replicates. Each bar colour corresponds to a specific signature, as reported in the plot legend. For brain organoids (B-E), asterisks identify genes that are retrieved as significantly changed in stage-wise differential expression analyses (comparison of each differentiation stage with the previous) reported in Figure 2 for CBO and Figure 4 for the other protocols. **: FDR < 0.05; *: Conventional pvalue < 0.05.

Supplementary Figure 10

Lollipops showing the expression levels of gene signatures related to neurotransmission systems in BS **(A)**, CBO **(B)**, MGO **(C)**, FO **(D)** and TA **(E)**. Expression levels (Log2Fpkms or Log2Rpm) along time-points are reported in each dataset as the mean value across replicates. Each bar colour corresponds to a specific signature, as reported in the plot legend. For brain organoids (B-E), asterisks identify genes that are retrieved as significantly changed in stage-wise differential expression analyses (comparison of each differentiation stage with the previous) reported in Figure 2 for CBO and Figure 4 for the other protocols. **: FDR < 0.05; *: Conventional P-value < 0.05.

Supplementary Figure 11

Overlap between BrainSpan and Cortical Brain Organoid gene modules identified by WGCNA. Dot plot displaying the overlap between BS (y-axis) and CBO (x-axis) gene modules identified by WGCNA. Numbers represent shared genes and are shown for overlaps with odds ratio (OR) > 1, while dots are reported for those having also P-value < 0.05. Dot colour is assigned according to OR values and dot size according to P-value.

Supplementary Figure 12

Overlap between BrainSpan (A, C) or Cortical Brain Organoid (B, D) WGCNA relevant gene modules and gene-disorder knowledge bases. Panels A and B show the overlap with SFARI genes (SFARI) and Development Disorder Genotype - Phenotype Database (DD), either for all organs or selectively for brain (DD_B). Panels C-D represent the overlap with risk genes retrieved from GWAS catalogue for the following disorders: Attention Deficit Hyperactive Disorder (ADHD), Autism Spectrum Disorder (ASD), Diabetes Mellitus (DM), Inflammatory Bowel Disease (IBD), Schizophrenia (SCZ), Unipolar Depression (UD). In each dot plot WGCNA gene sets are reported on the x-axis and disorder gene sets on the y-axis. Numbers represent shared genes and are shown for overlaps with odds ratio (OR) > 1, while dots are reported for those having also P-value < 0.01. Dot colour is assigned according to OR values and dot size according to P-value as shown by each legend.

Supplementary Figure 13

Behaviour of the module eigengene for Cortical Brain Organoids gene co-expression modules in fetal cortex and brain organoids datasets. (A) Visualization of CBO_Turquoise and CBO_Black eigengene module in prenatal fetal cortex as well as in the brain organoids' datasets. PCA was performed to calculate the module eigengene on the CBO dataset and then applied to BS prenatal cortex, as well as each of the other brain organoid datasets. Each dot represents a data point, while the line connects the median value for each postconceptional week (BS) or differentiation day (BO). The same analysis and visualization are applied to CBO_Brown and CBO_Blue (B, decreasing during differentiation) and CBO_Green and CBO_Red (C, nonmonotonic behaviour).

Supplementary Table 1

Differentially expressed genes with $FDR < 0.05$ and $abs(logFC) > 1$ from differential expression analyses.

Supplementary Table 2

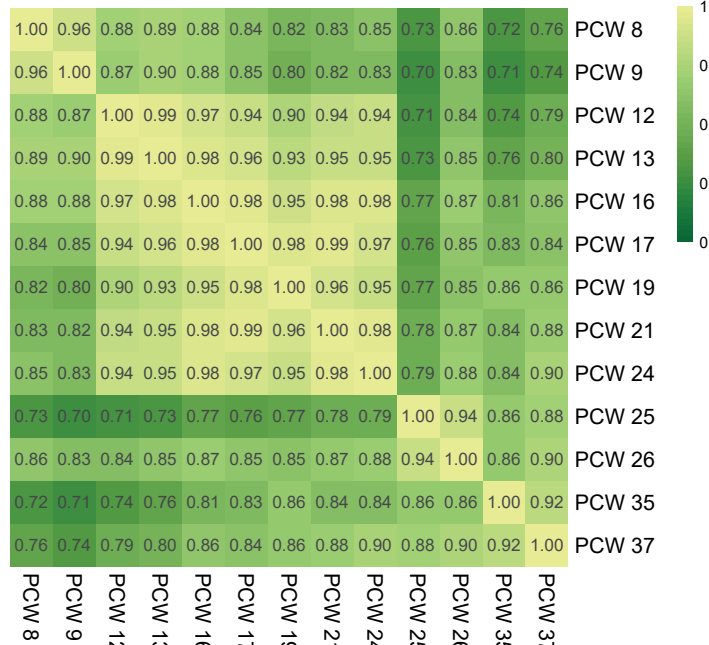
Neurotransmission signature genes with the specification of their role in the relative neurotransmission system (Enzyme, receptor or transporter).

Supplementary Table 3

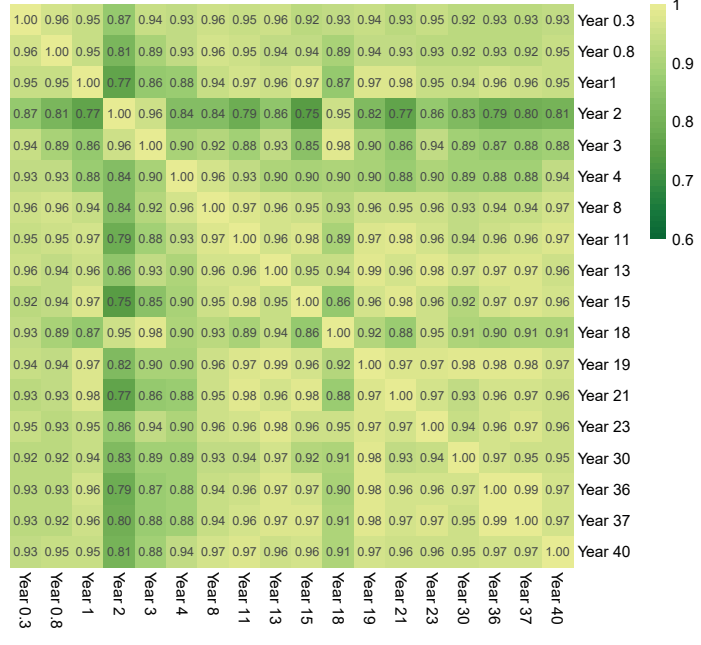
Results of Gene Ontology analyses.

Figure S1

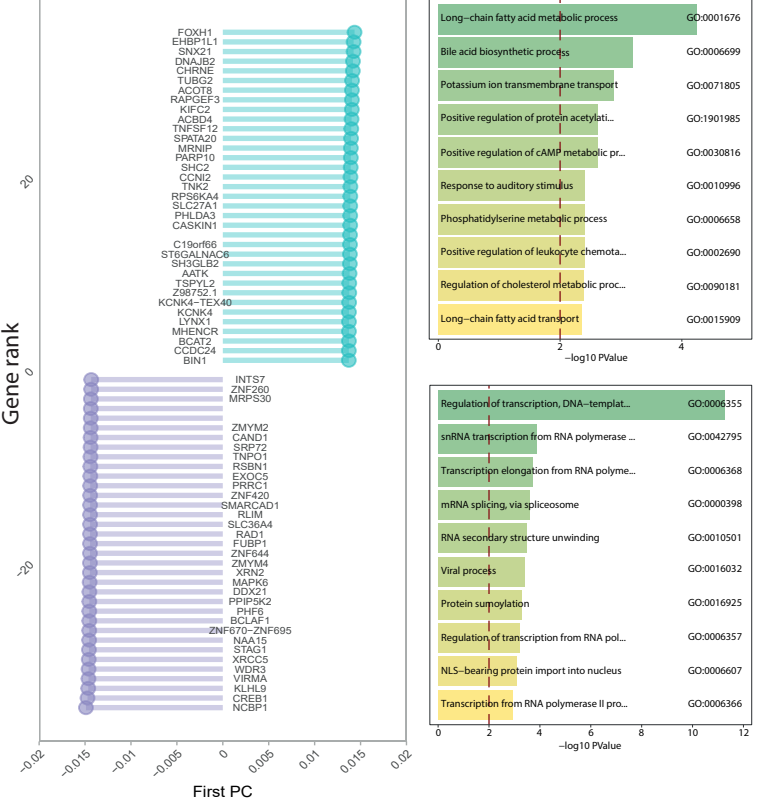
A



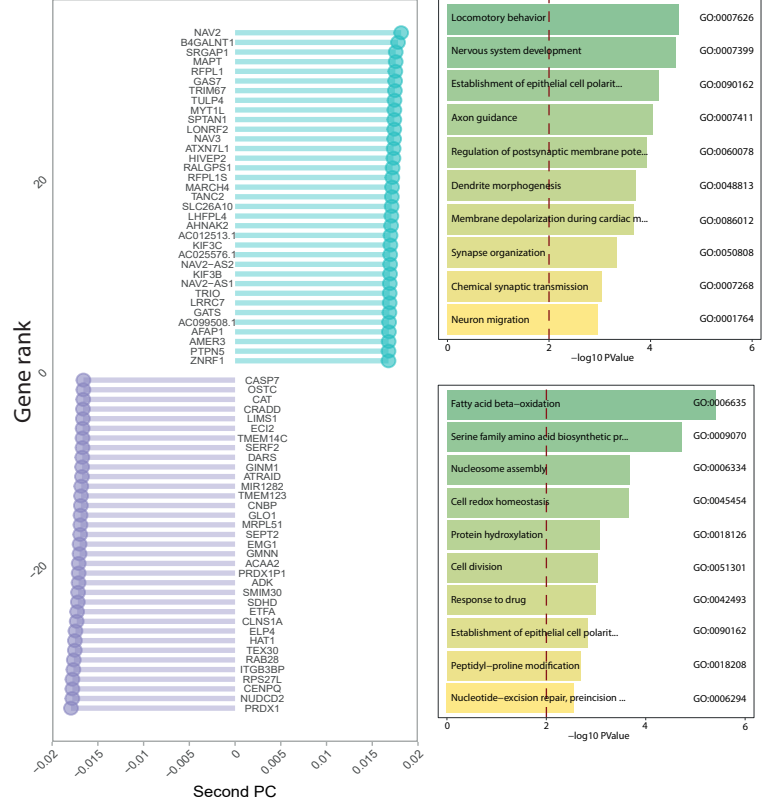
B



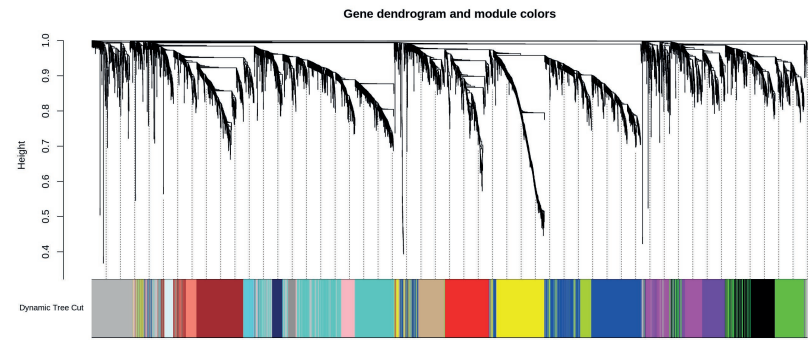
C



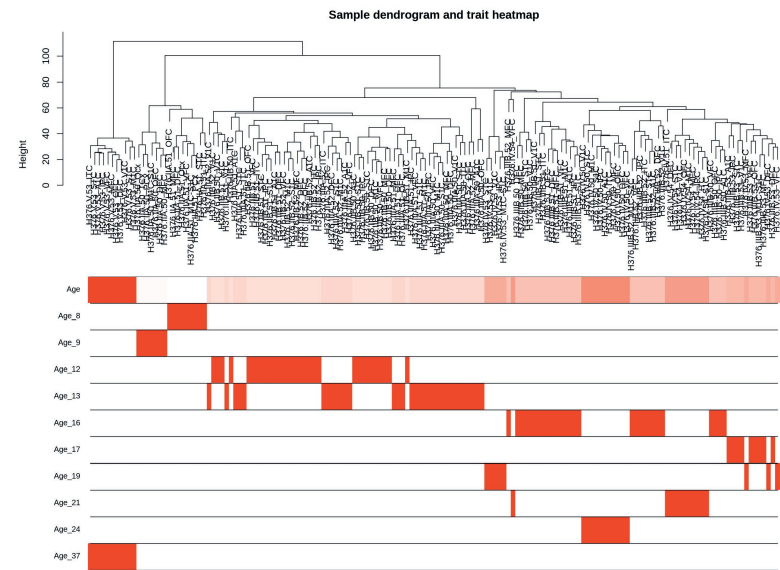
D



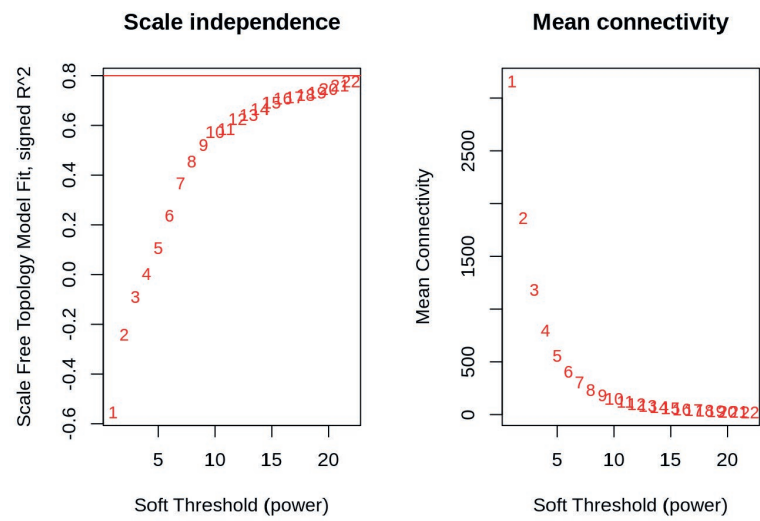
A



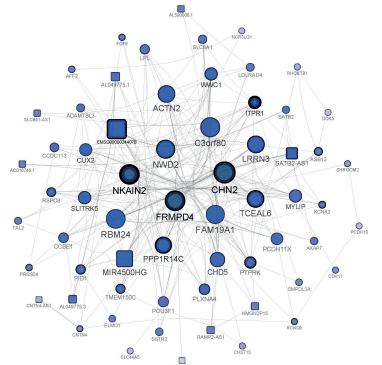
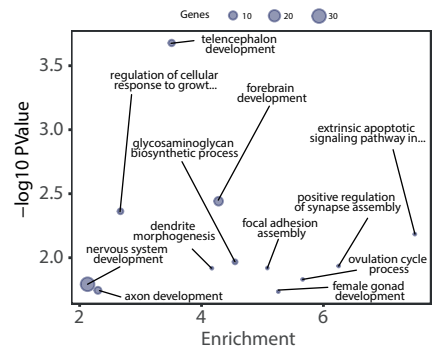
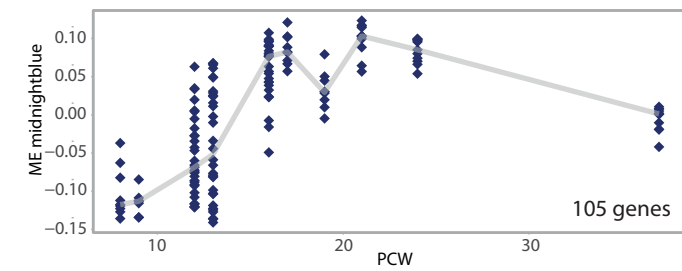
B



C

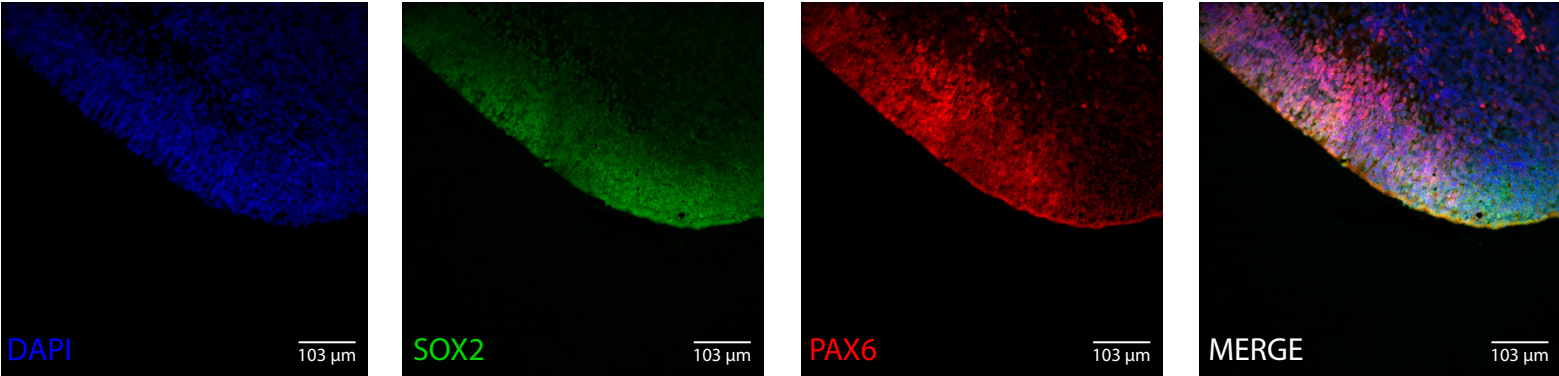


D



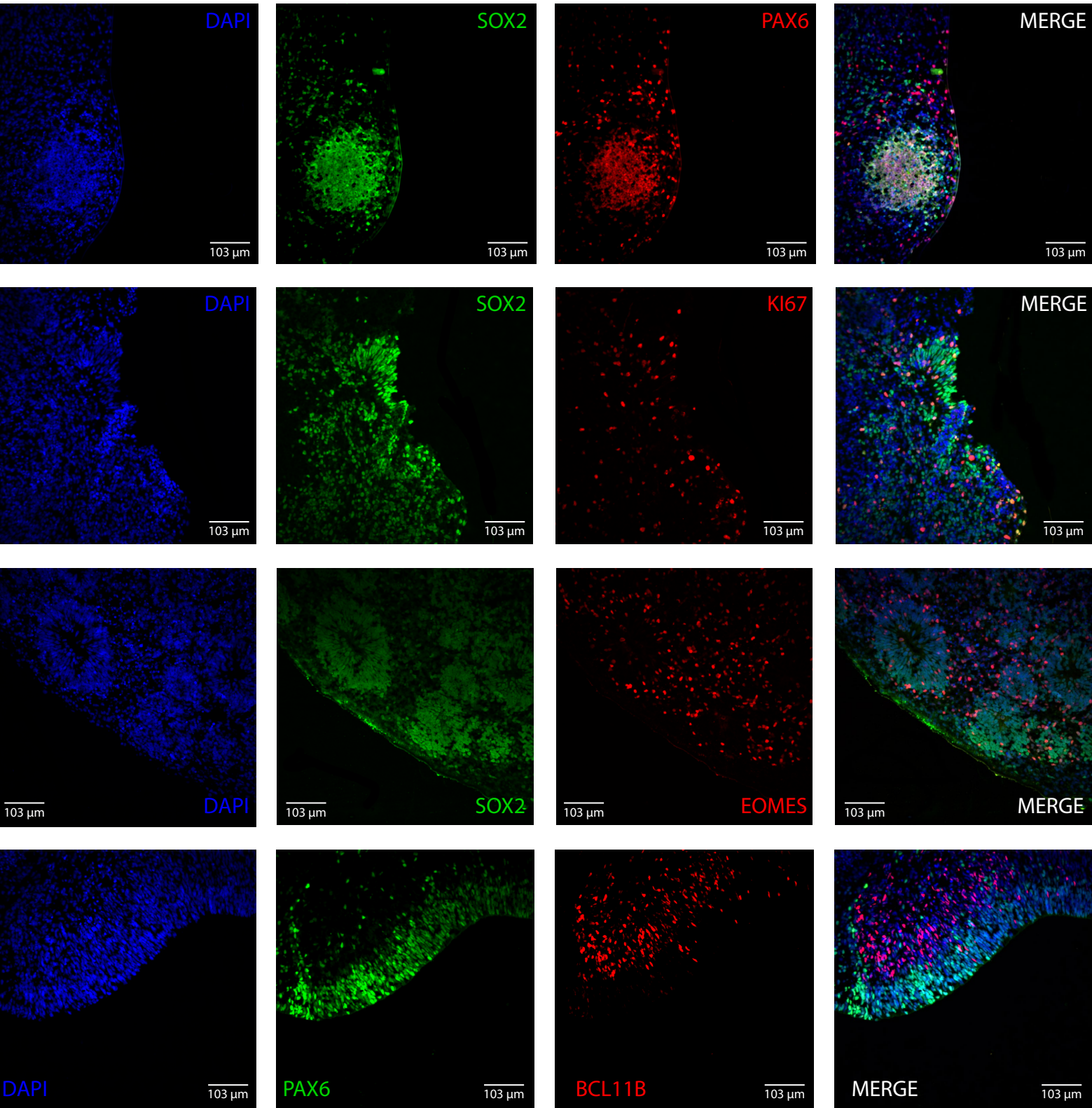
A

DAY 25



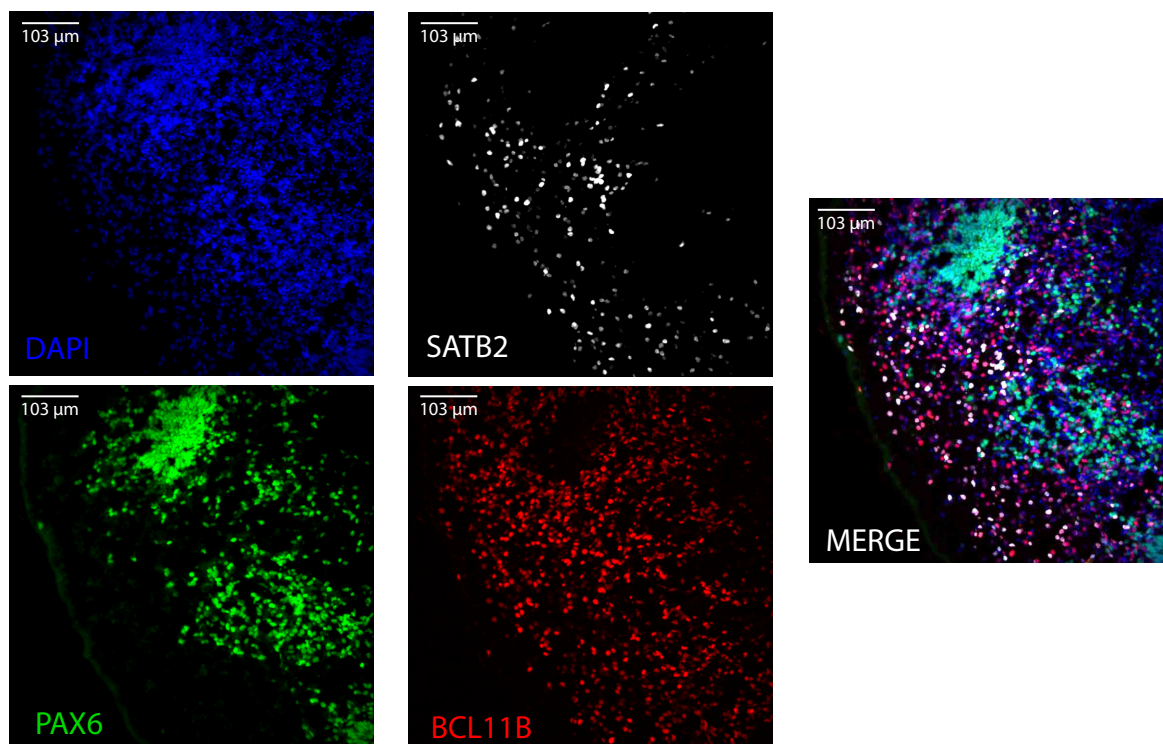
B

DAY 50



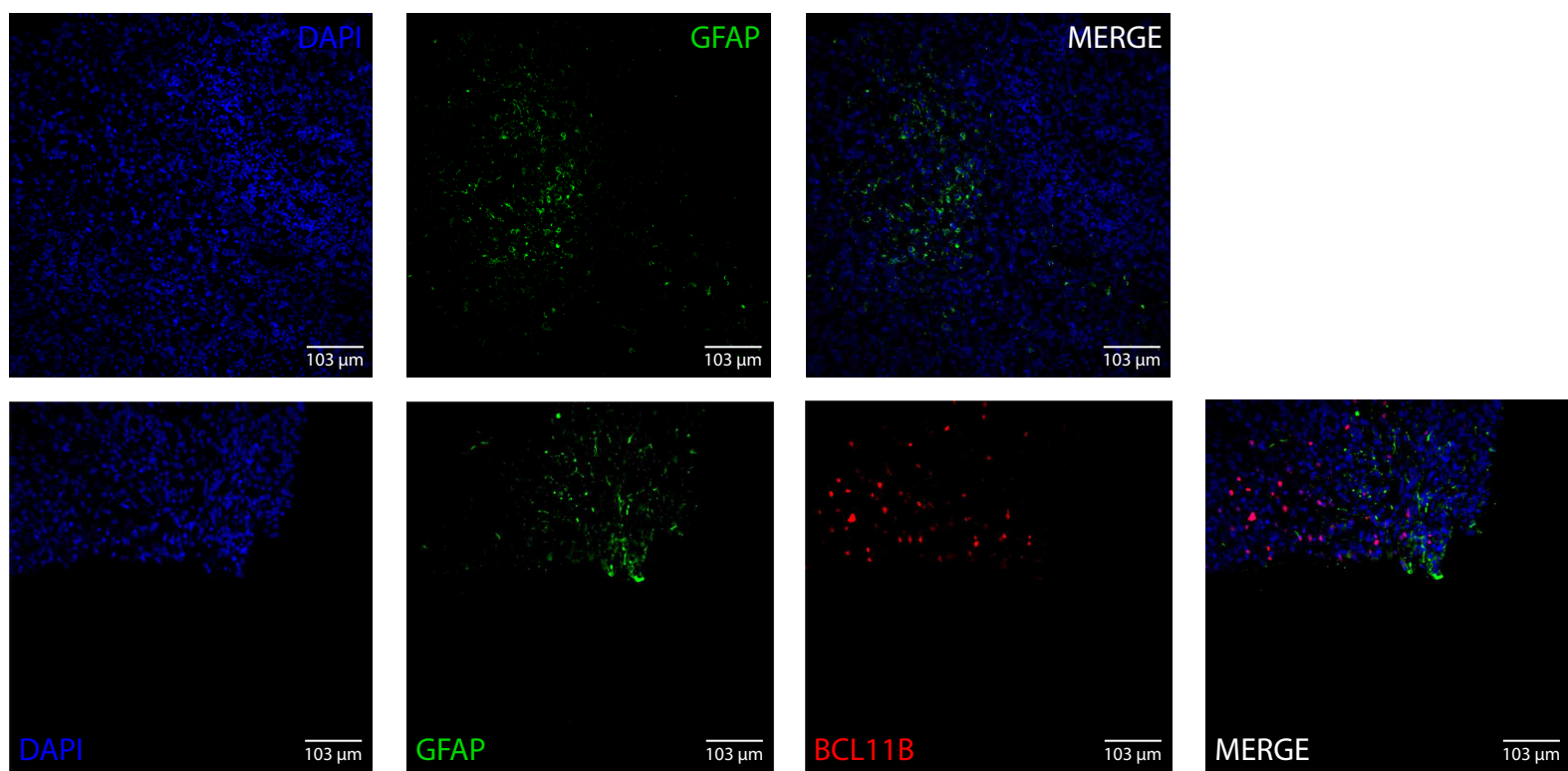
A

DAY 100



B

DAY 100



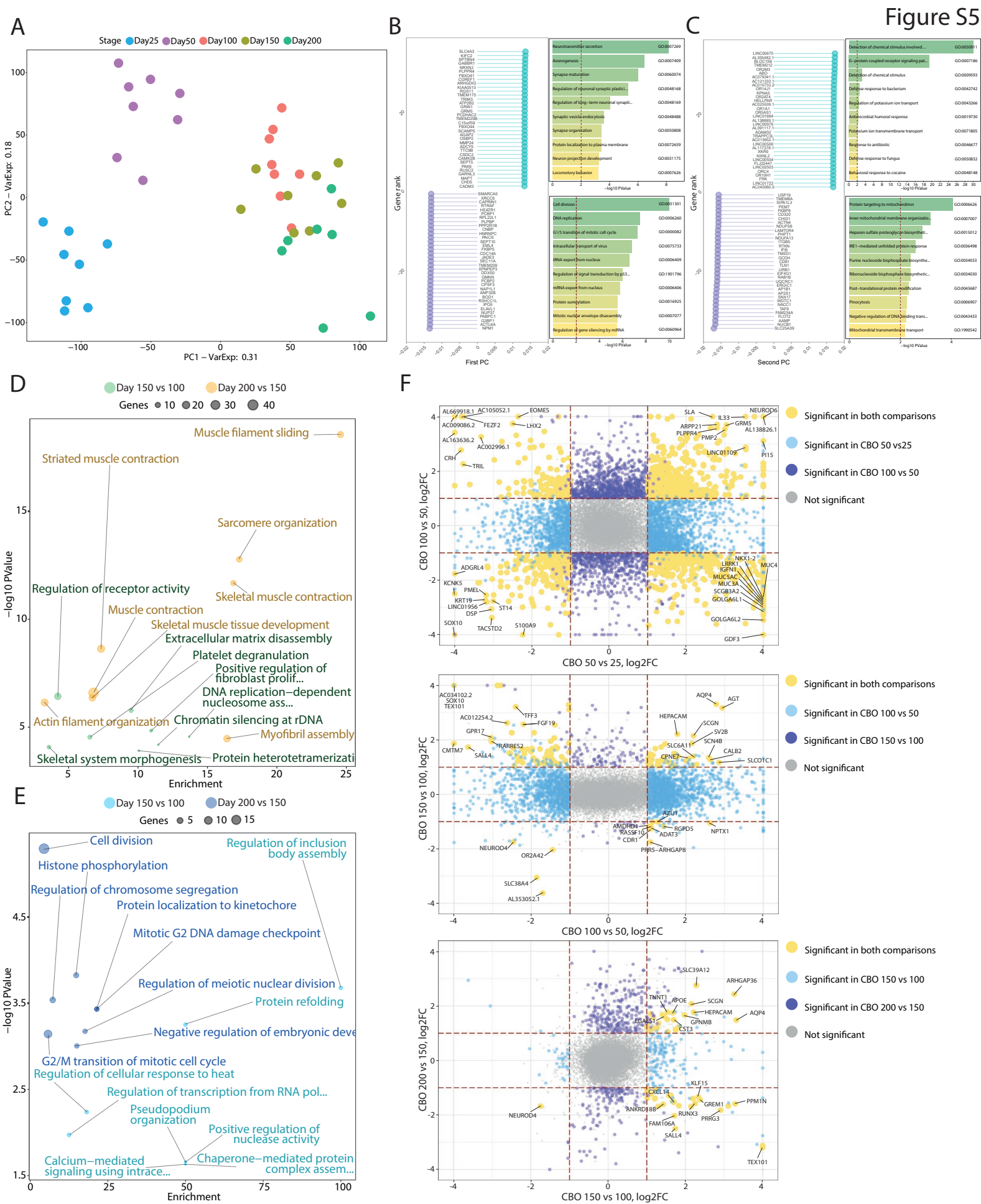
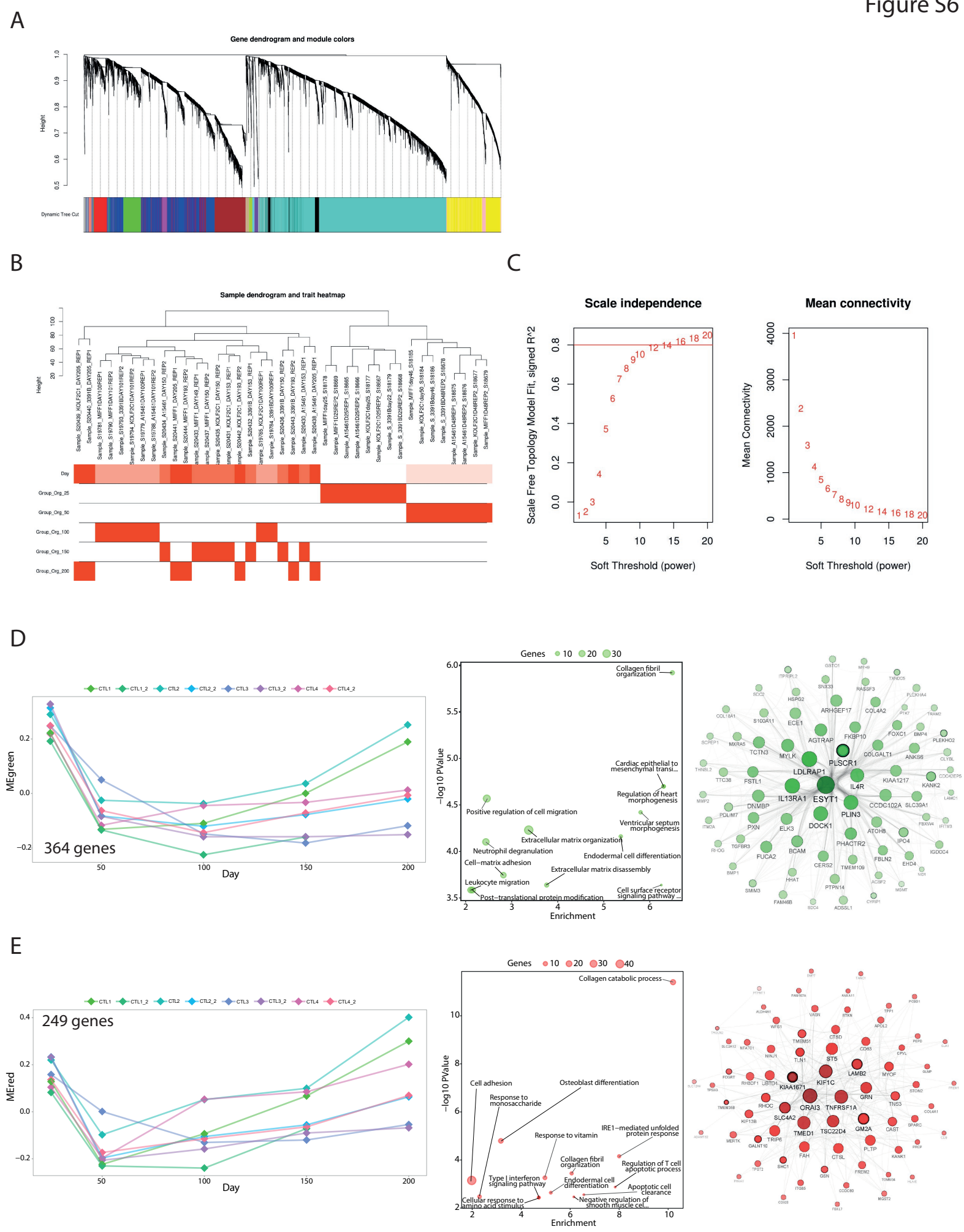
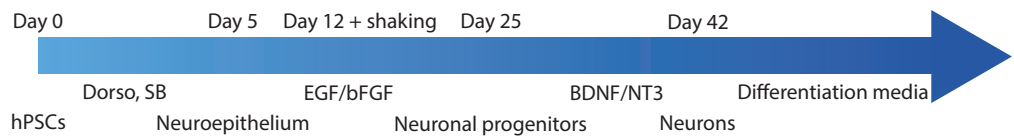


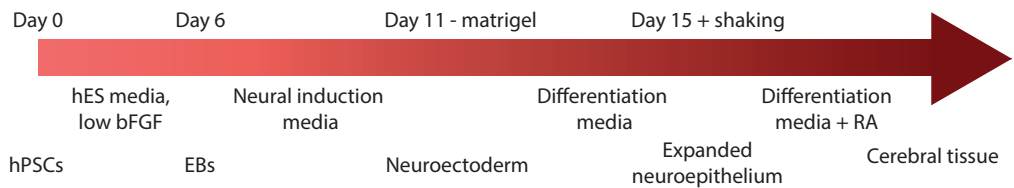
Figure S6



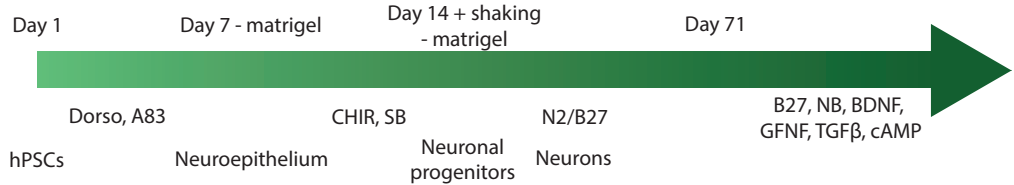
A Cortical brain organoids



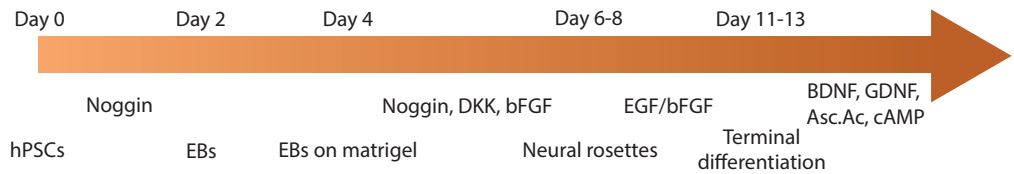
Minimally-guided neural organoids



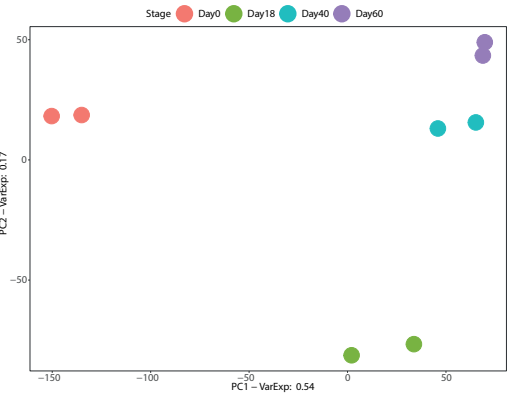
Forebrain organoids



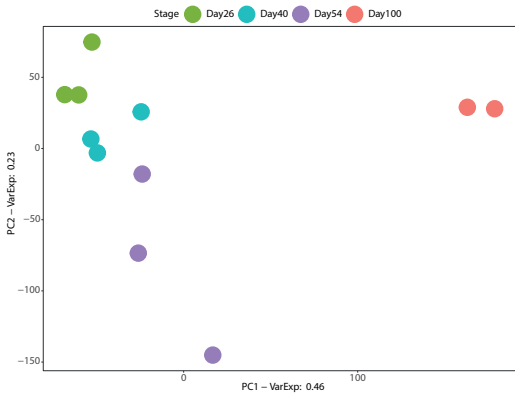
Telencephalic aggregates



B Minimally-guided neural organoids



C Forebrain organoids



D Telencephalic aggregates

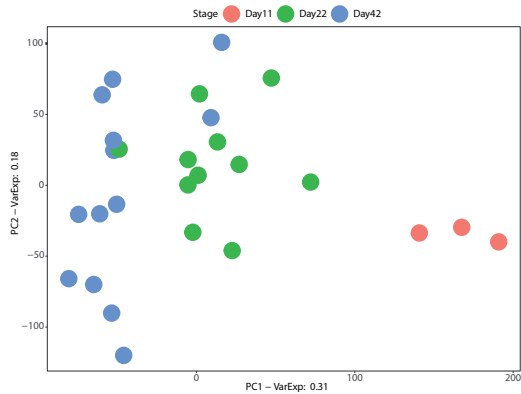
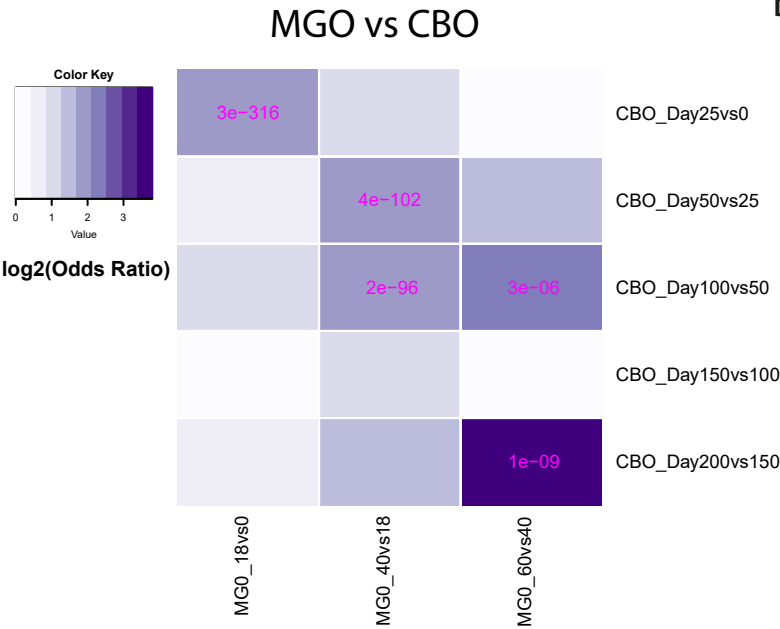
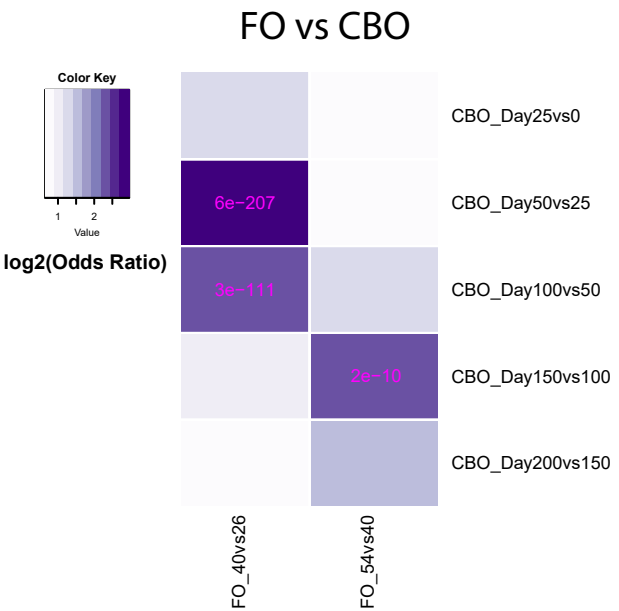


Figure S8

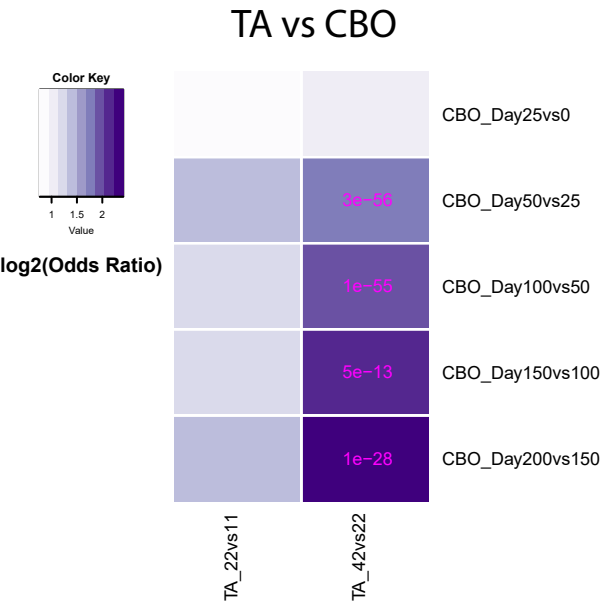
A



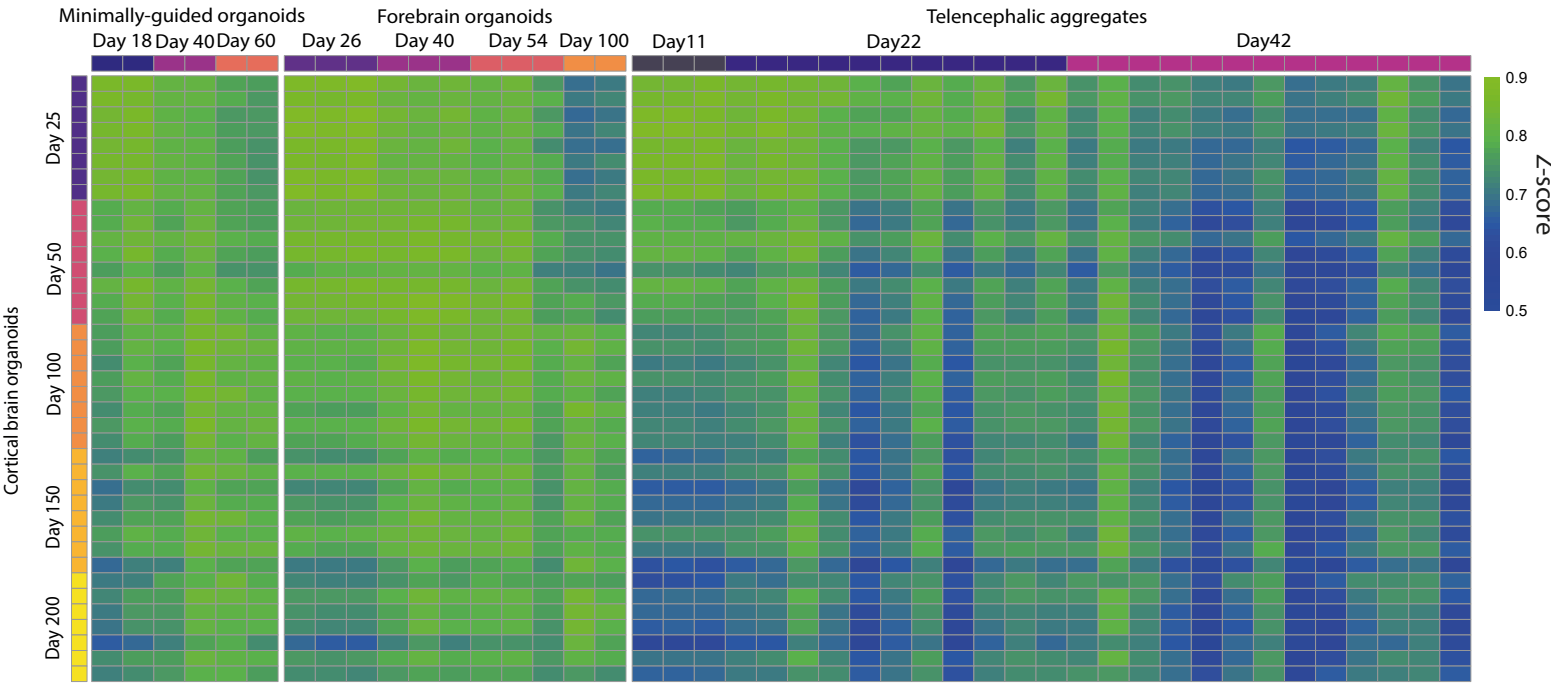
B



C



D



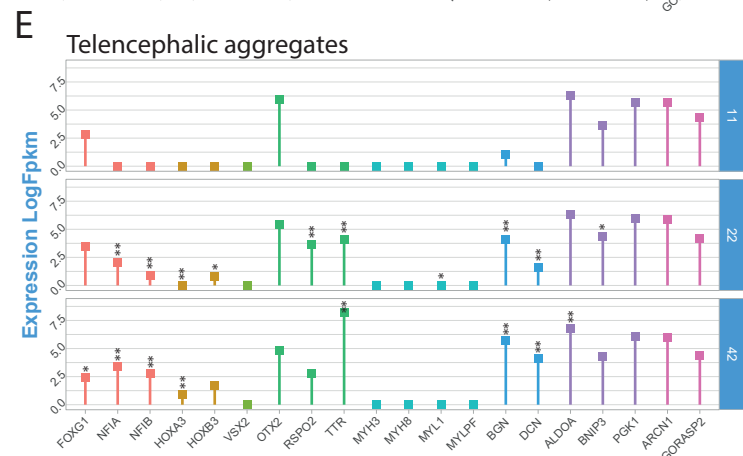
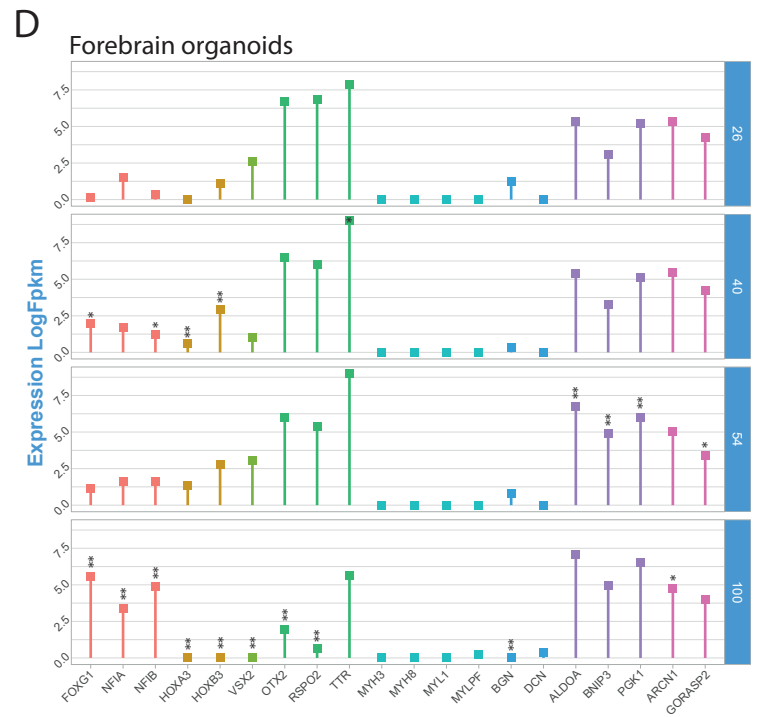
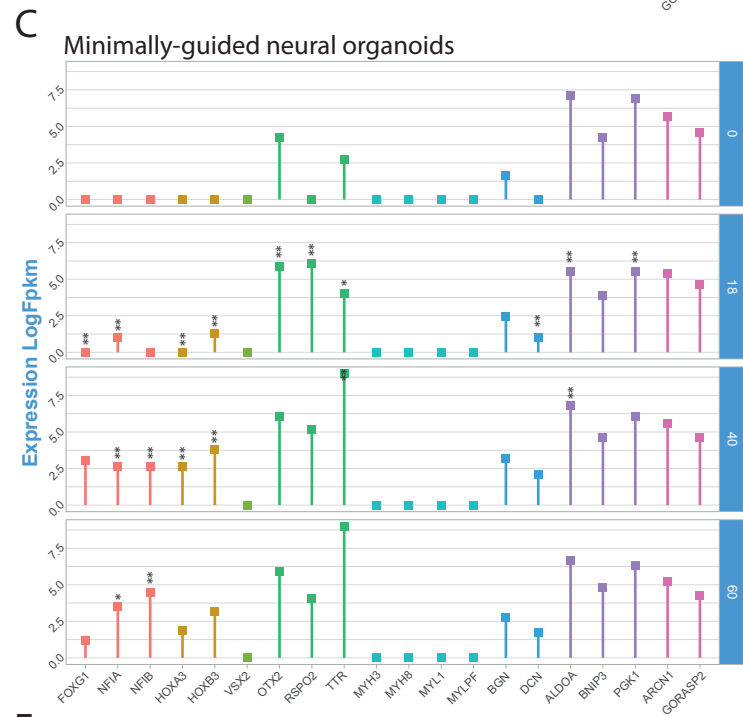
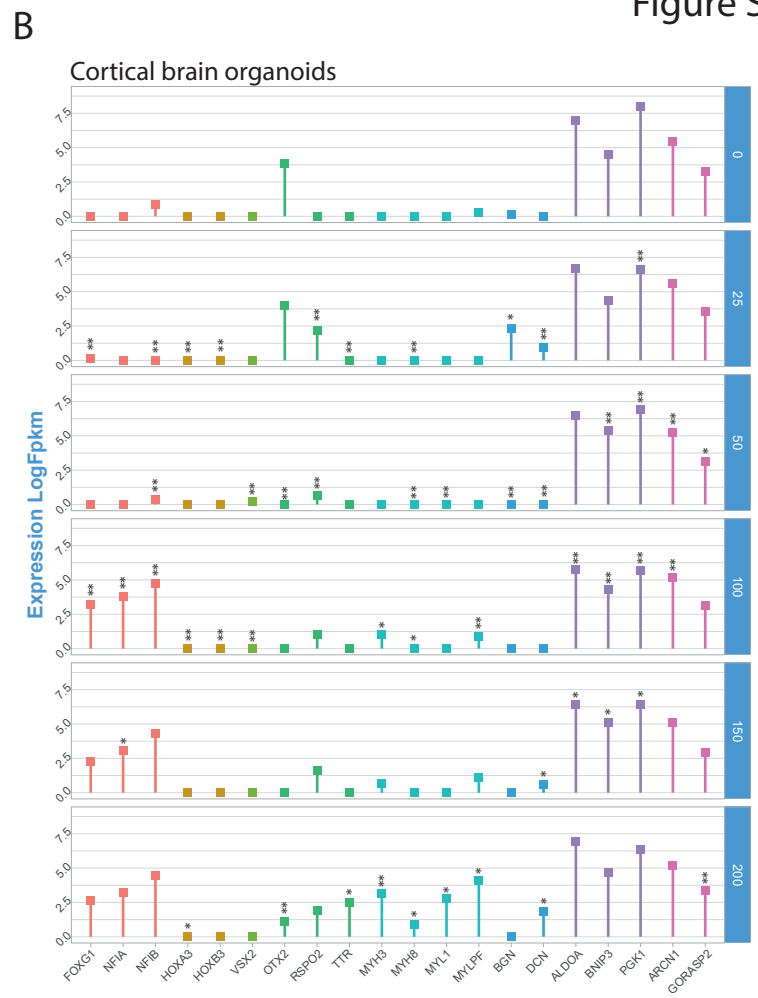
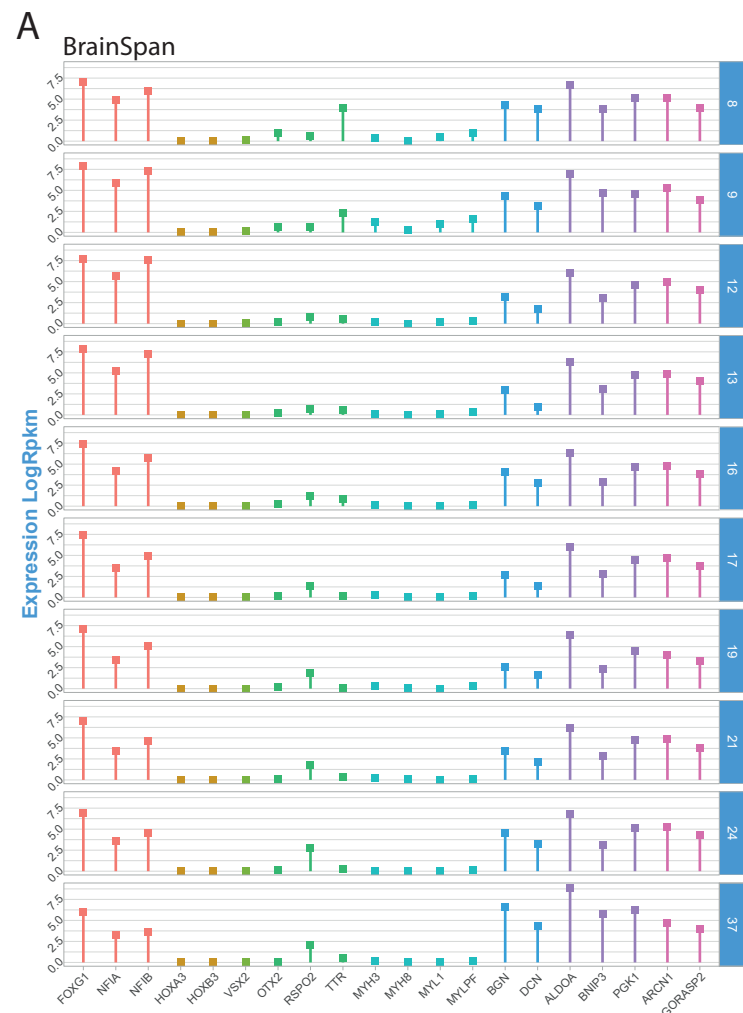
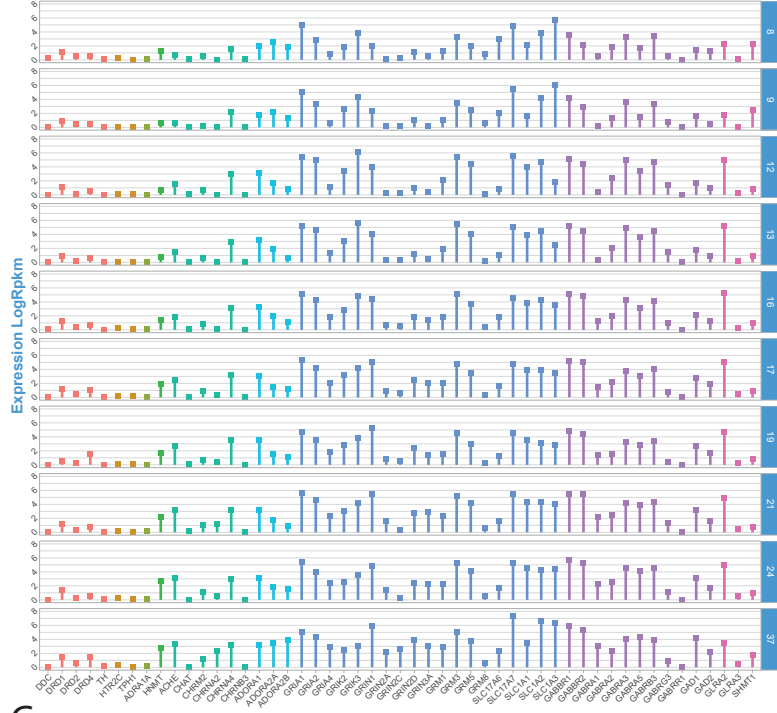


Figure S10

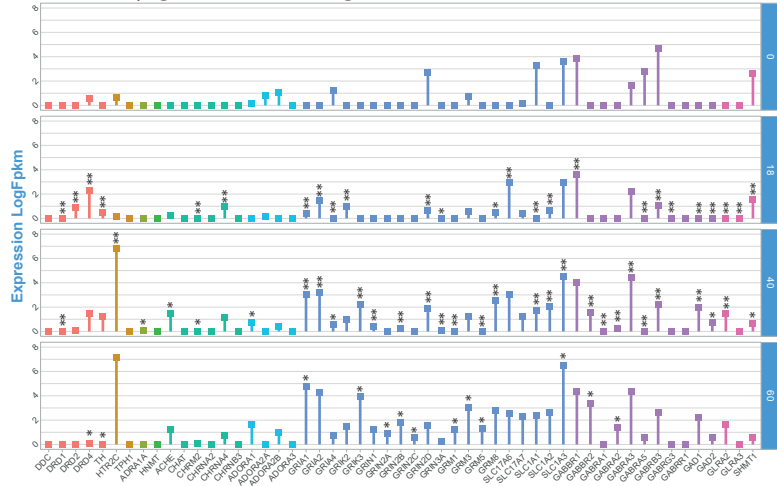
A

BrainSpan



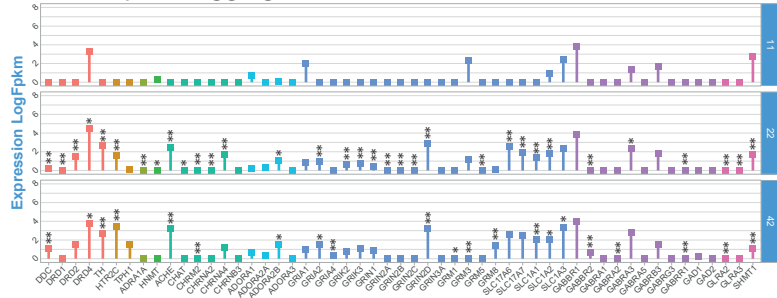
C

Minimally-guided neural organoids



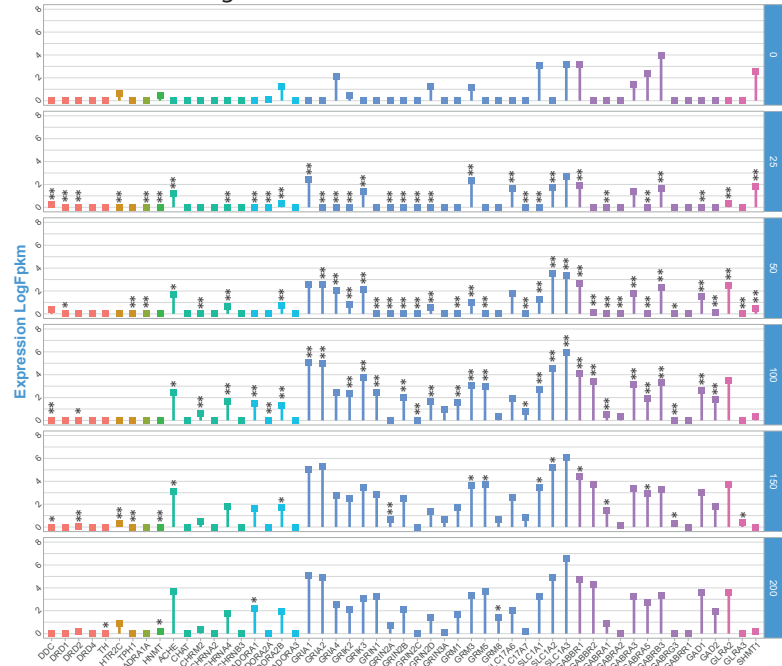
E

Telencephalic aggregates



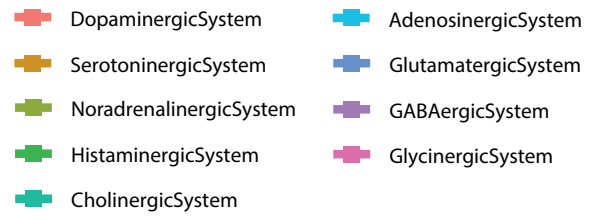
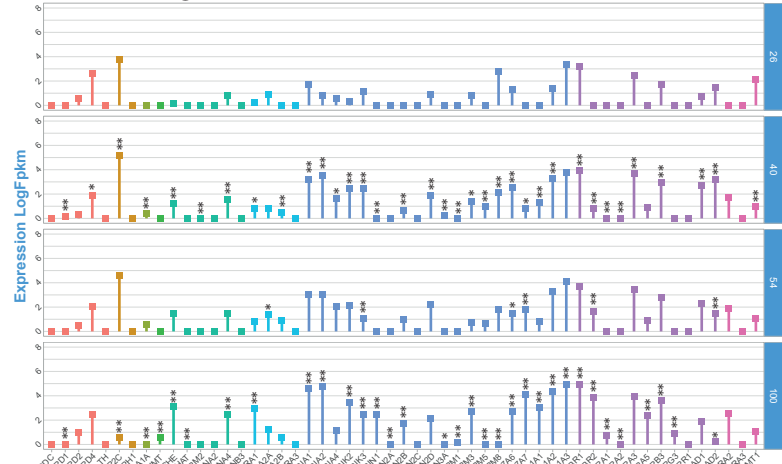
B

Cortical brain organoids



D

Forebrain organoids



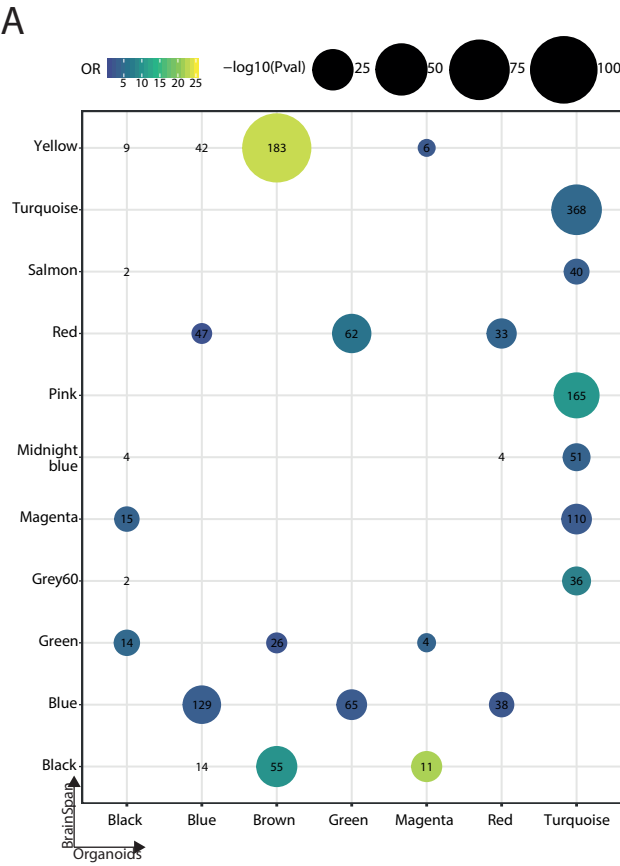
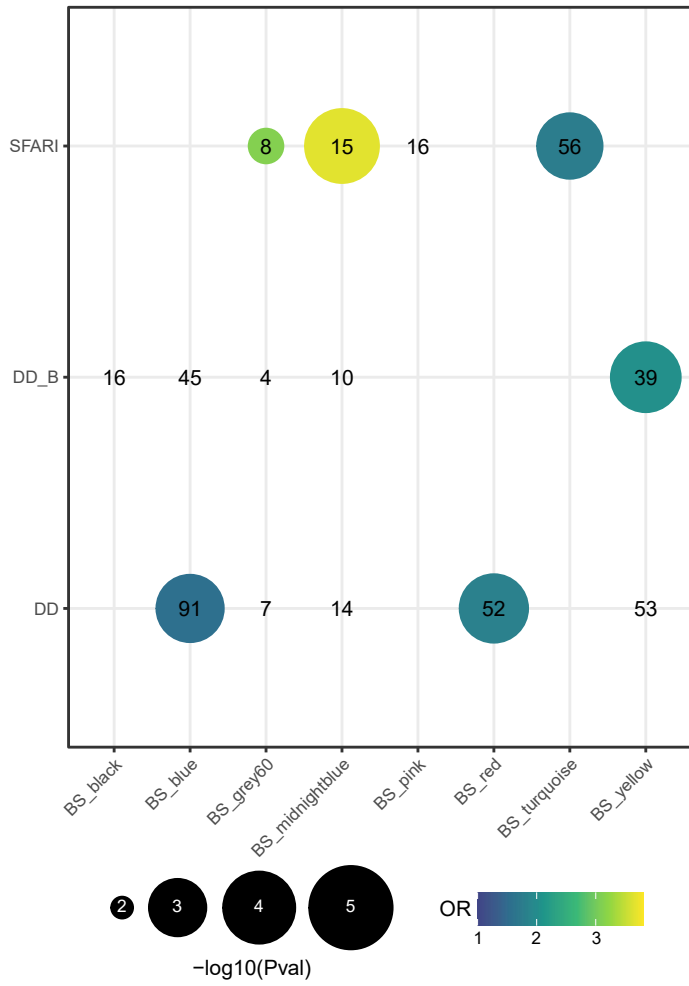
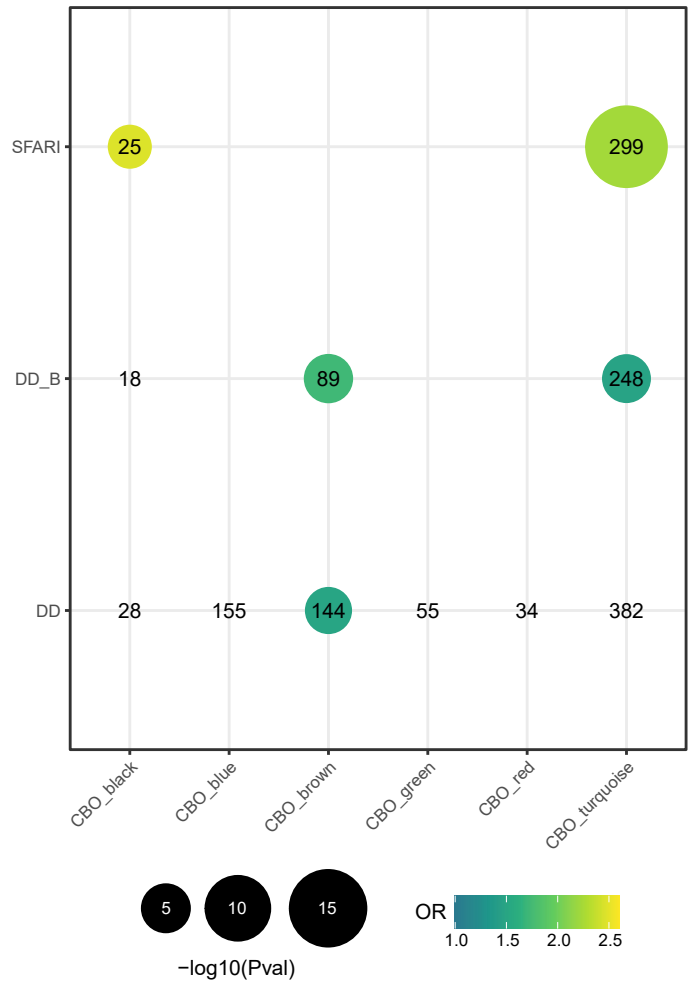


Figure S12

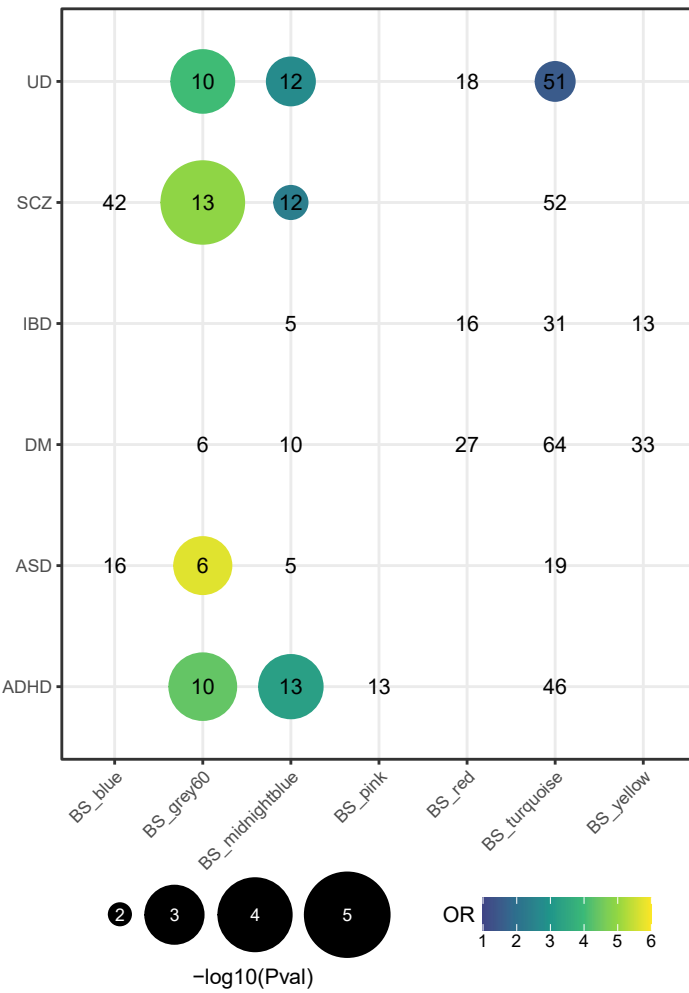
A



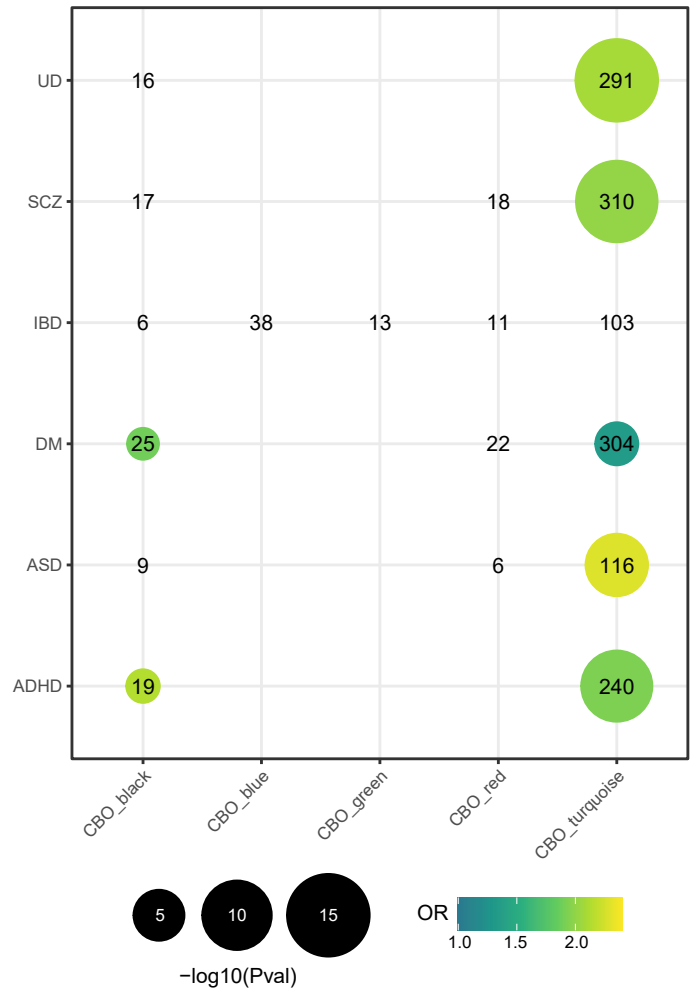
B



C

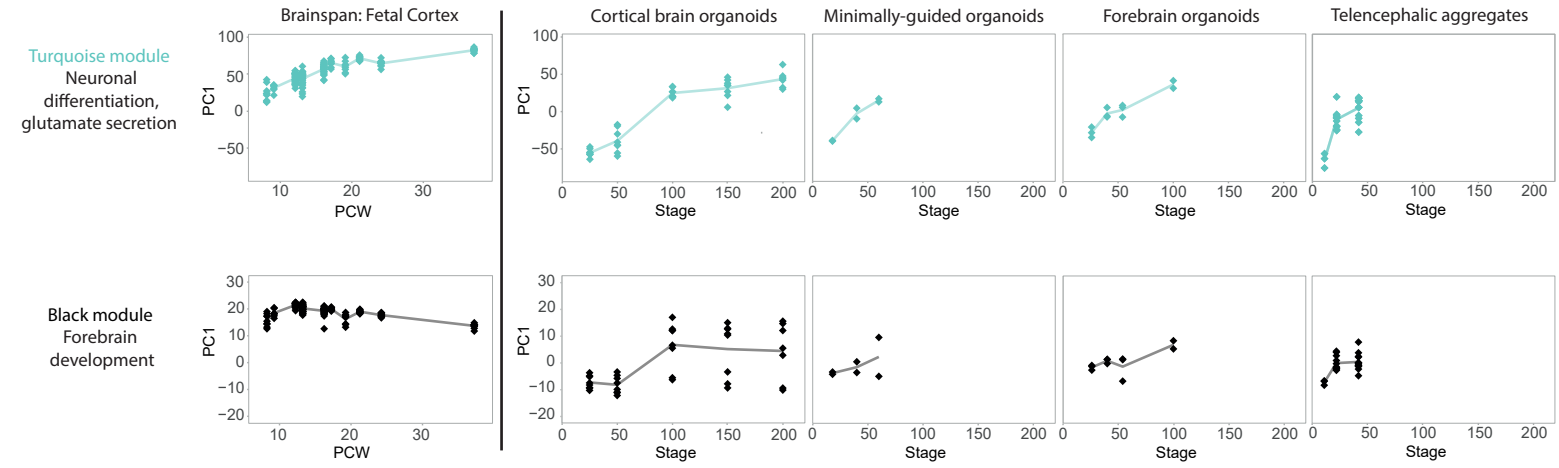


D



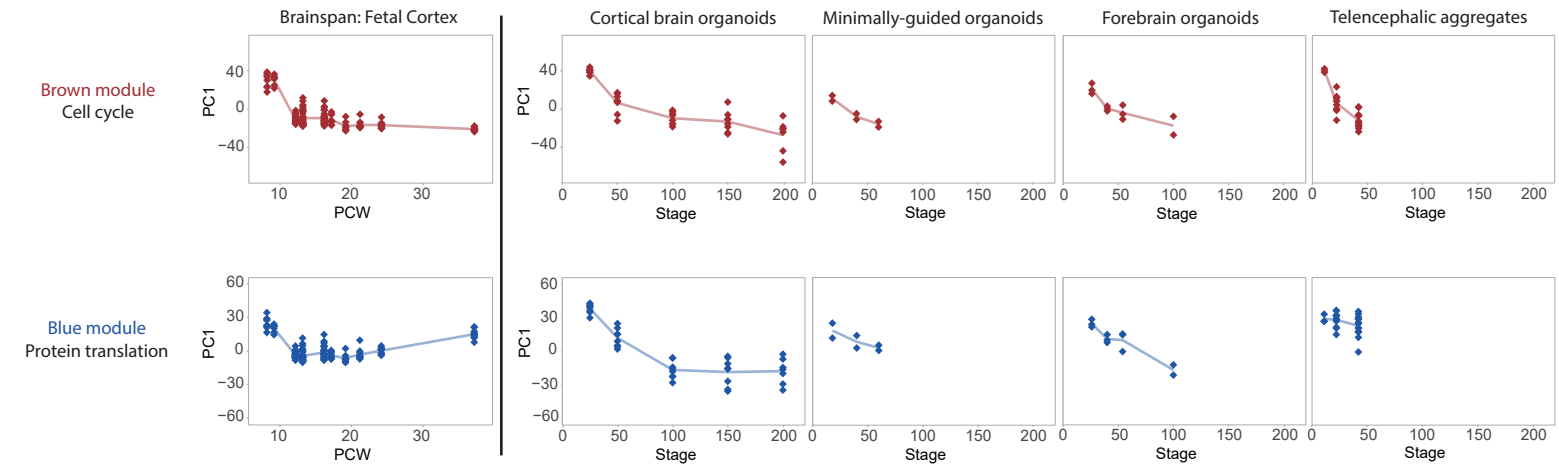
A

CBO: monotonic behaviour with increased expression over developmental time



B

CBO: monotonic behaviour with decreased expression over developmental time



C

CBO: non-monotonic behaviour

