Inventory of Supplemental Information

We present 2 Supplementary Figures and 2 Supplementary Tables as further support of the main Figures 1-7.

Figure S1, related to Figure 2 and Figure 4. The neuroepithelium remains intact following AF collection at E8.5.

Figure S2, related to Figure 6. Availability of receptors in differentiating cerebrospinal fluid.

Table S1, related to Figure 4. Spectral counts, GProX clustering, enrichment and pathway analysis of the 961 proteins detected in E8.5 AF, E10.5 CSF and E14.5 CSF

Table S2, related to Figure 7. Spectral counts, GProX clustering and enrichment analysis of the 844 proteins detected in E8.5 AF, E10.5 AF and E14.5 AF

Figure S1

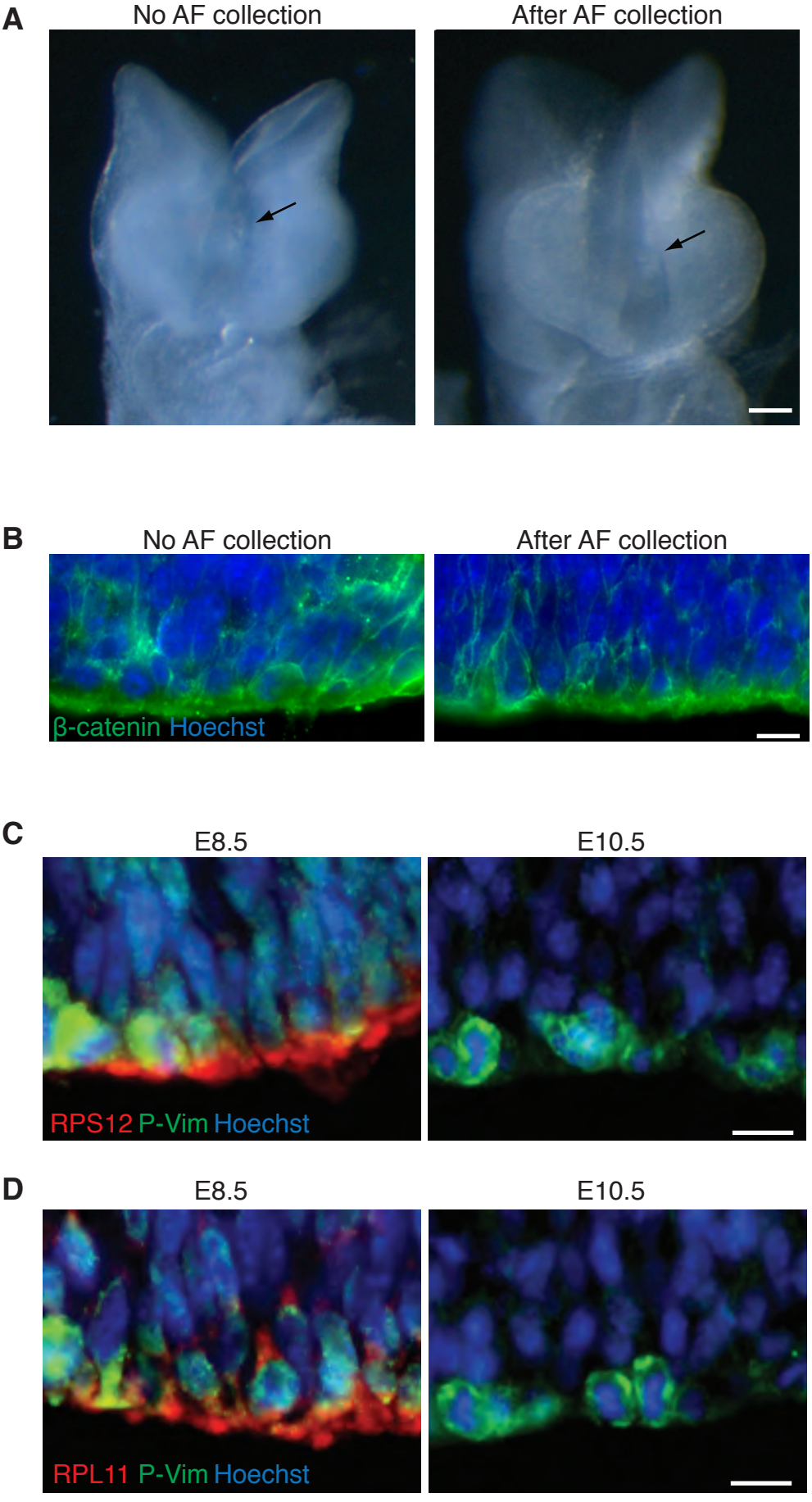
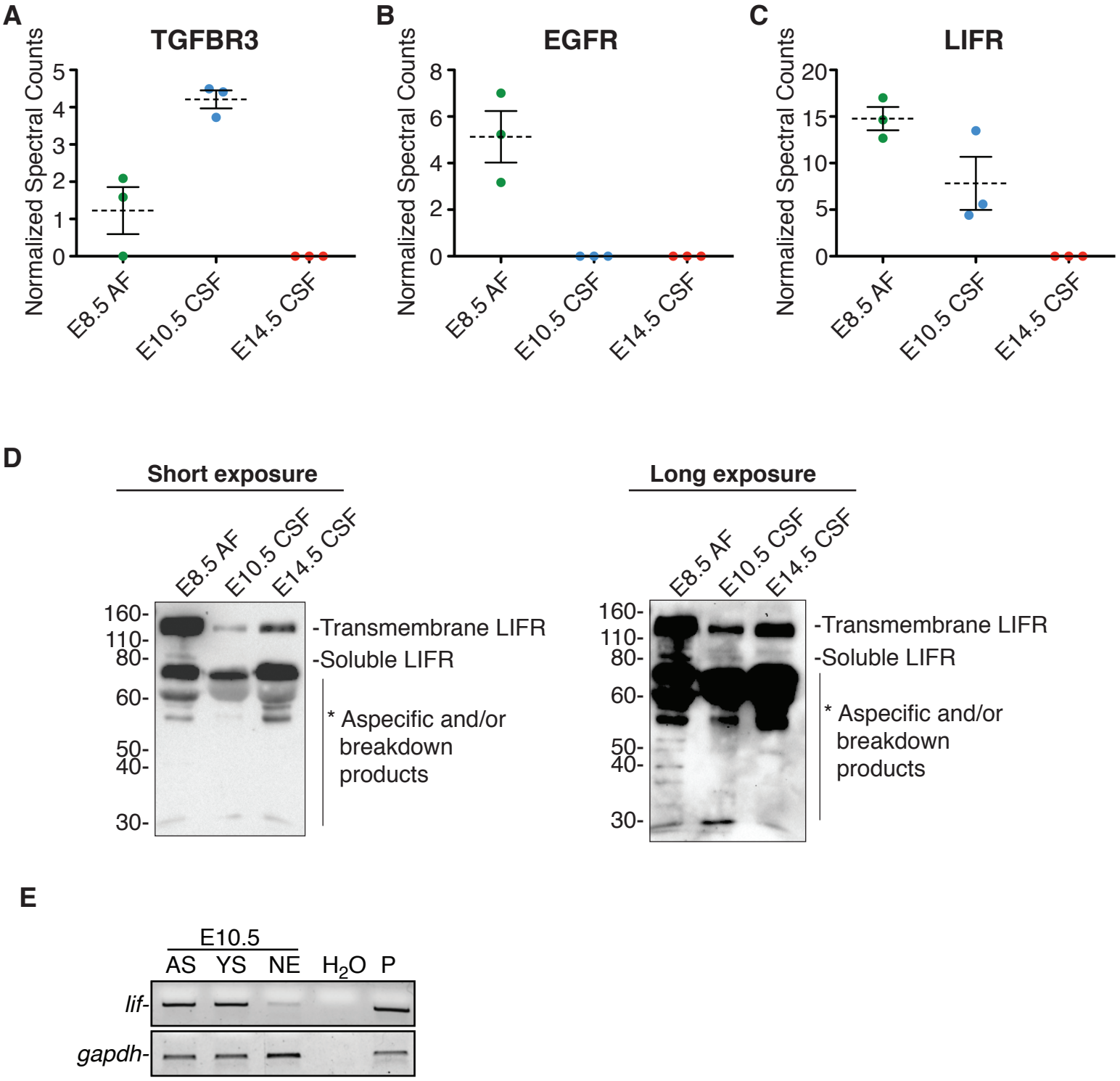

A
No AF collection
After AF collection
B
No AF collection
After AF collection
β-catenin Hoechst
C
E8.5
E10.5
RPS12 P-Vim Hoechst
D
E8.5
E10.5
RPL11 P-Vim Hoechst

Figure S2


A
TGFBR3
Normalized Spectral Counts
0
1
2
3
4
5
E8.5 AF
E10.5 CSF
E14.5 CSF
B
EGFR
Normalized Spectral Counts
0
2
4
6
8
E8.5 AF
E10.5 CSF
E14.5 CSF
C
LIFR
Normalized Spectral Counts
0
5
10
15
20
E8.5 AF
E10.5 CSF
E14.5 CSF
D
Short exposure
E8.5 AF
E10.5 CSF
E14.5 CSF
160-
110-
80-
60-
50-
40-
30-
-Transmembrane LIFR
-Soluble LIFR
* Aspecific and/or breakdown products
Long exposure
E8.5 AF
E10.5 CSF
E14.5 CSF
160-
110-
80-
60-
50-
40-
30-
-Transmembrane LIFR
-Soluble LIFR
* Aspecific and/or breakdown products
E
E10.5
AS
YS
NE
H2O
P
lif-
gapdh-

SUPPLEMENTARY FIGURE LEGENDS

Figure S1, related to Figure 2 and Figure 4. The neuroepithelium remains intact following AF collection at E8.5. (A) Whole mount images show presumptive forebrain region in E8.5 embryos that are un-manipulated (left) and following AF collection (right). Note that the forebrain neuroectoderm (black arrows) remains intact after AF collection. Scale bar, 100 μm. **(B)** Apical membrane staining of β-catenin (green) shows that the ventricular surface remains intact following AF collection at E8.5. Nuclei counterstained with Hoechst. Scale bar, 10 μm. **(C, D)** E8.5 forebrain neuroectoderm shows stronger expression of two ribosomal protein subunits RPS12 and RPL11 (red), in P-Vimentin-positive (green) progenitor cells. Nuclei counterstained with Hoechst. Scale bar, 10 μm.

Figure S2, related to Figure 6. Availability of receptors in differentiating cerebrospinal fluid. (A) Availability of TGFBR3 peaked at E10.5 CSF, and was not detected in E14.5 CSF (E8.5 AF = 1.225 ± 0.630; E10.5 CSF = 4.211 ± 0.243; E14.5 CSF = not detected). Circles represent individual samples, n=3 for all samples). **(B)** EGFR was detected in E8.5 AF but not in E10.5 CSF or E14.5 CSF (E8.5 AF = 5.132 ± 1.107; E10.5 CSF = not detected; E14.5 CSF = not detected). **(C)** The availability of LIFR decreased from E8.5 AF to E10.5 CSF, and was not detected in E14.5 CSF by spectral counting (E8.5 AF = 14.771 ± 1.249; E10.5 CSF = 7.829 ± 2.847; E14.5 CSF = not detected). **(D)** LIFR was most abundant in E8.5 AF by immunoblotting. The transmembrane isoform (122.57kDa) was detected in all three fluid compartments, whereas the soluble isoform (81.29kDa) was only detected in E8.5 AF under low exposure setting (left). Under long exposure (right), soluble LIFR isoform was also detected in E10.5 CSF but not in

E14.5 CSF. **(E)** RT-PCR showed higher *lif* expression in amniotic sac (AS) and yolk sac (YS) compared to neuroepithelium (NE). H_2O as negative control; placenta (P) as positive control.

Table S1, related to Figure 4. Spectral counts and GProX clustering of all 961 proteins detected in E8.5 AF, E10.5 CSF and E14.5 CSF (worksheet 1); Functional annotation clustering of proteins in GProX clusters 1, 2 and 3 using DAVID (worksheet 2, 3 and 4); KEGG pathway analysis of proteins in GProX cluster 2 using DAVID (worksheet 5).

Table S2, related to Figure 7. Spectral counts and GProX clustering of all 844 proteins detected in E8.5 AF, E10.5 AF and E14.5 AF (worksheet 1); Functional annotation clustering of proteins in GProX clusters 1, 2 and 3 using DAVID (worksheet 2, 3 and 4).

SUPPLEMENTARY EXPERIMENTAL PROCEDURES

Cell culture assays. *Pair cell assay.* The E10.5 dorsal telencephalic neuroepithelium was isolated as for explants (see above). Tissue was dissociated using papain (Worthington kit), and cells were plated at clonal density on poly-D-lysine coated Terasaki plates (Shen et al., 2002; Tucker et al., 2010). Cell suspensions were prepared in Neurobasal medium with antibiotics and glutamine, and supplemented with LIF (200ng/ml reported in E11 mouse CSF; Hatta et al., 2006), its vehicle, 20% E10.5 CSF or artificial CSF. For LIF signaling inhibition experiments, E10.5 CSF was pre-incubated with antibodies for 4-6 hours. Single cell adherence to plates was visually confirmed after 2 hour of incubation, and after 18 hours, cells were fixed and immunostained with Sox2 and Tuj1, and counterstained with Hoechst. Expression of cell identity markers in pairs of cells was scored in a blinded manner. Two-way ANOVA statistical analyses were performed. *Indicator cell experiments:* Measurement of BMP and RA activity using luciferase-based assay in reporter cell lines were performed as previously described (Lehtinen et al., 2011).

Mass Spectrometry. Samples were prepared as described (Lun et al., 2015a). 30µg of protein from fluid reduced in 0.1M DTT was filtered twice using a 30kDa spin filter (Millipore). The nominal molecular weight cut off does not apply to denatured proteins (Wiśniewski et al., 2009). Samples were alkylated, washed, and digested overnight with 2µg trypsin (Promega) at 37°C, and the digested proteins were eluted, acidified, desalted, and dried. Samples were reconstituted and analyzed by a nanoLC system (Eksigent) equipped with LC-chip system (cHiPLCnanoflex, Eksigent) coupled online to a Q-Exactive mass spectrometer (Thermo Scientific). Peptides were separated by a linear gradient using a gradient length of 120 minutes. Data were acquired in the

"data dependent acquisition" (DDA) mode, selecting the 10 most abundant species for fragmentation (Top 10). Thermo ".raw" files were converted into the Mascot generic format (MGF-files) using the ProteoWizard software tool, and then submitted to database searches using Mascot (v2.3.01 Matrix Science, London, UK) using a concatenated mouse protein sequence database (UNIPROT.MOUSE, downloaded 06/2014). The protein sequence database was searched with tryptic cleavage specificity, a fragment ion mass tolerance of 20 mmu and a parent ion tolerance of 10 ppm; carbamidomethylated cysteine as fixed modification, and N-terminal pyroglutamic acid formation for E and Q, deamidation for N and Q, oxidation for M as variable modification. The instrument selection was set to Orbi MS, Orbi MS/MS. Scaffold (v4.2.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications, resulting in an overall peptide and protein false discovery rates (FDR) of 1%. Subsequent output provided protein IDs with spectral counts corresponding to unique peptides identified for the protein per sample. Proteins were accepted as present if identified in any of the biological replicates.

Bioinformatic Analysis. Before any analysis, spectral counts of all samples were first normalized to the sample with the highest number of total spectral counts. Heatmap generation and hierarchical clustering were performed using the "gplots" package in R: normalized spectral counts were scaled across each protein so that the sample with the highest spectral count for that protein is set as 1; distance was measured using the Euclidean method and clustering was done by complete linkage. Venn diagrams were generated using BioVenn (Hulsen et al., 2008). Clustering by temporal expression pattern was performed using GProX (Rigbolt et al., 2011): for each protein, its average normalized spectral count at each fluid compartment was calculated,

scaling was then performed so that the fluid with the highest average spectral count for that protein is set as 1; proteins were not included in the clustering if their scaled average spectral counts showed < 2 fold difference between the ages; only proteins with membership value > 0.35 were displayed on the plots. Functional annotation clustering was performed using DAVID v6.7 (Huang et al., 2008; Huang et al., 2009): enrichment (p-value < 0.05) is determined by comparing proteins in each GProX cluster to the reference mouse genome; enrichment score is the geometric mean of the individual enrichment p-values of all annotation terms within a group. KEGG pathway analysis was performed in DAVID v6.7 using a cutoff of p-value < 0.05. Network analysis was performed using GeneMania (Warde-Farley et al., 2010). Principal component analysis was performed using the "FactoMineR" package in R.

Fluorescein measurements. E10.5 embryos were harvested into warm Neurobasal medium. The embryonic yolk sac was exposed and 0.5 μl of 2.5% fluorescein-conjugated dextran was targeted into the intra-amniotic space. Embryos were cultured *ex vivo* for 30 minutes at 37°C (5% CO_2), after which extra-embryonic tissues were removed, embryos were rinsed twice in Neurobasal medium, and CSF was collected. Pooled samples from two embryos represent one individual sample measured for fluorescence intensity (NanoDrop 3300 Fluorospectrometer). Relative fluorescence was calculated base on a standard curve generated by serial dilutions of the 2.5% 10 kDa fluorescein conjugate.

Immunostaining. Paraffin sections were dehydrated, and antigen retrieval was performed using Antigen Unmasking Solution (Vector). Sections were blocked and permeabilized (0.04% Tween-20 in PBS, 5% serum), followed by antibody incubation. Cryosections were blocked,

permeabilized, and incubated with antibodies. All tissues were counterstained with Hoechst 33258.

Antibodies. The following antibodies were purchased: SOX2 (Abcam, ab97959), P-Vimentin (Enzo Bioscience, ADI-KAM-CC249-E), β-Catenin (BD, 610154), RPS12 (Proteintech, 16490-1-AP), RPL11 (Santa Cruz, sc-50363), LIF (R&D Systems, MAB449, for Western blotting), LIF (R&D systems, AF449, for neutralization), LIFR (R&D Systems, AF249-NA, for LIFR blocking), LIFR (Santa Cruz, sc-659, for immunostaining), LIFR (Abcam, ab101228, for Western blotting), GATA6 (Santa Cruz, sc-7244x), P-STAT3 (Cell Signaling, 9145), Tuj1 (Covance, MMS-435P).

ELISA. Shh concentrations in AF and CSF were measured using the Mouse Shh N-terminus Quantikine ELISA kit (R&D Systems) according to manufacturer's instructions. 12.5 to 25 µl AF or CSF were used in each measurement. Bmp4 concentrations were similarly measured by ELISA (Biotang, Inc.).

RNA quantification. Expression of Sox2 relative to Gapdh was measured in neuroectodermal explants and olfactory placode explants by Sybr-green qPCR. The following primers were used:

Sox2 forward: AGCCCCCCCGTGGTTACCTCTTC;

Sox2 reverse: CCCCTTCTCCAGTTCGCAGTCCAG;

Gapdh forward: CTGACGTGCCGCCTGGAGAAA;

Gapdh reverse: GTTGGGGGCCGAGTTGGGATAGG

SUPPLEMENTARY REFERENCES

Hatta, T., Matsumoto, A., Ono, A., Udagawa, J., Nimura, M., Hashimoto, R., and Otani, H. (2006). Quantitative analyses of leukemia inhibitory factor in the cerebrospinal fluid in mouse embryos. Neuroreport *17*, 1863-1866.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1-13.

Hulsen, T., de Vlieg, J., and Alkema, W. (2008). BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics *9*, 488.

Lehtinen, M.K., Zappaterra, M.W., Chen, X., Yang, Y.J., Hill, A.D., Lun, M., Maynard, T., Gonzalez, D., Kim, S., Ye, P.*, et al.* (2011). The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. Neuron *69*, 893-905.

Lun, M.P., Johnson, M.B., Broadbelt, K.G., Watanabe, M., Kang, Y.-J., Chau, K.F., Springel, M.W., Malesz, A., Sousa, A.M.M., Pletikos, M.*, et al.* (2015a). Spatially Heterogeneous Choroid Plexus Transcriptomes Encode Positional Identity and Contribute to Regional CSF Production. J Neurosci *35*, 4903-4916.

Rigbolt, K.T., Vanselow, J.T., and Blagoev, B. (2011). GProX, a user-friendly platform for bioinformatics analysis and visualization of quantitative proteomics data. Mol Cell Proteomics *10*, O110.007450.

Shen, Q., Zhong, W., Jan, Y.N., and Temple, S. (2002). Asymmetric Numb distribution is critical for asymmetric cell division of mouse cerebral cortical stem cells and neuroblasts. Development *129*, 4843-4853.

Tucker, E.S., Lehtinen, M.K., Maynard, T., Zirlinger, M., Dulac, C., Rawson, N., Pevny, L., and Lamantia, A.-S.S. (2010). Proliferative and transcriptional identity of distinct classes of neural precursors in the mammalian olfactory epithelium. Development *137*, 2471-2481.

Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T.*, et al.* (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res *38*, W214-220.

Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat. Methods *6*, 359–62.