Current Biology, Volume *34*

Supplemental Information

Molecular organization of neuronal cell types and neuromodulatory systems in the zebrafish telencephalon

Lukas Anneser, Chie Satou, Hans-Rudolf Hotz, and Rainer W. Friedrich

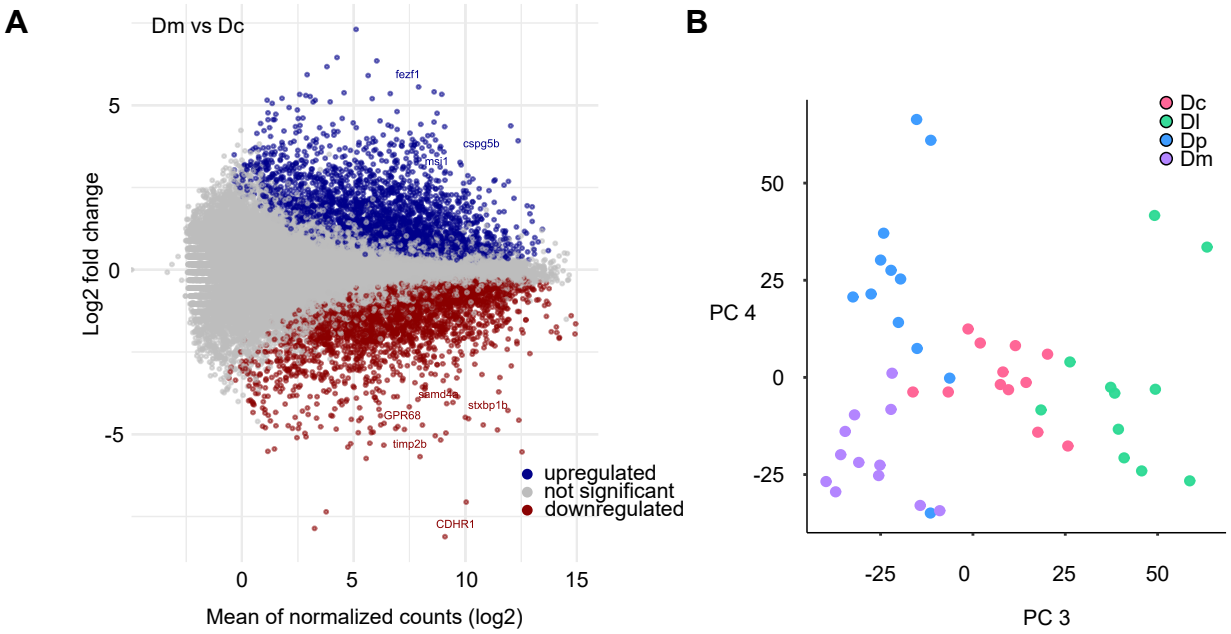
A
Dm vs Dc
fezf1
cspg5b
msi1
semd4a
stxbp1b
GPR68
timp2b
CDHR1
upregulated
not significant
downregulated
Log2 fold change
Mean of normalized counts (log2)
5
0
-5
0
5
10
15
B
Dc
Dl
Dp
Dm
PC 4
PC 3
50
25
0
-25
-25
0
25
50

Figure S1. Bulk sequencing of pallial microdissected areas, related to Figure 1.
(**A**) Example of differential gene expression in two different telencephalic brain areas (Dm and Dc). (**B**) Projection of gene expression patterns (2.000 most variable genes; same as in Figure 1B) onto PCs 3 and 4. Each data point is one replicate from a pallial region (color-coded). Note separation of Dp and Dm. See Figure 1B for projection of the same data onto PCs 1 and 2.

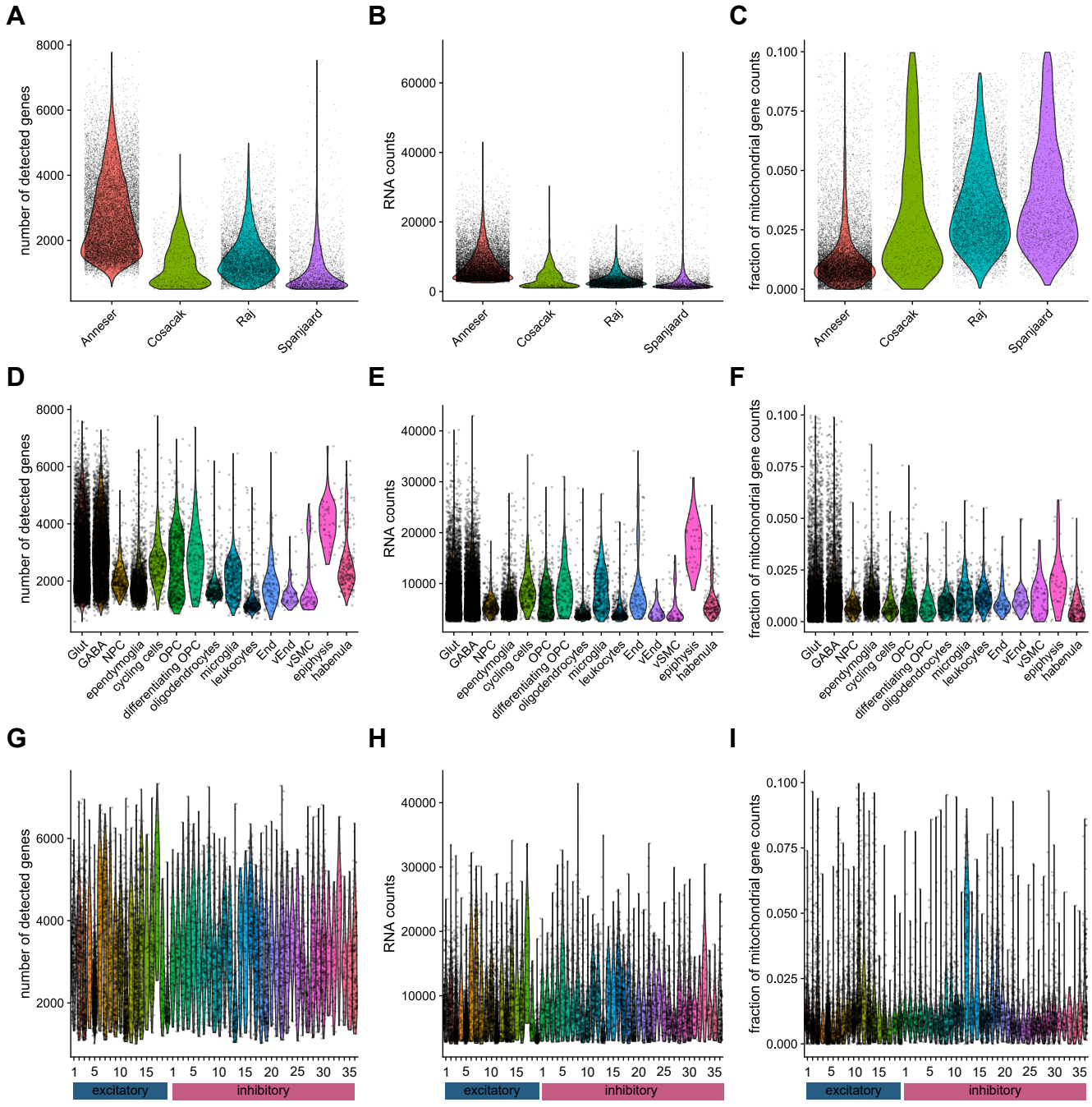
A
number of detected genes
8000
6000
4000
2000
Anneser
Cosacak
Raj
Spanjaard
B
RNA counts
60000
40000
20000
0
C
fraction of mitochondrial gene counts
0.100
0.075
0.050
0.025
0.000
D
E
F
Glut
GABA
NPC
ependymoglia
cycling cells
OPC
differentiating OPC
oligodendrocytes
microglia
leukocytes
End
vEnd
vSMC
epiphysis
habenula
G
H
I
1 5 10 15 1 5 10 15 20 25 30 35
excitatory
inhibitory

Figure S2. Quality assessment of scRNA-seq dataset, related to Figure 2.
(**A-C**) Comparison of quality-related parameters of our single-cell RNA sequencing (scRNA-seq) dataset ("Anneser") to previously published scRNA-seq datasets from the zebrafish telencephalon ("Cosacak"[S1], "Raj"[S2], "Spanjaard"[S3]). Each dot represents one sequenced cell; violin plots show distributions. In our dataset, the number of detected genes and RNA counts are comparably high (**A**, **B**) while the fraction of mitochondrial gene reads, which usually indicates cell death, is relatively low (**C**). (**D-F**) Same quantifications stratified by superclusters in our dataset (Figure 2A). (**E-I**) Same quantification stratified by neuronal subtypes.

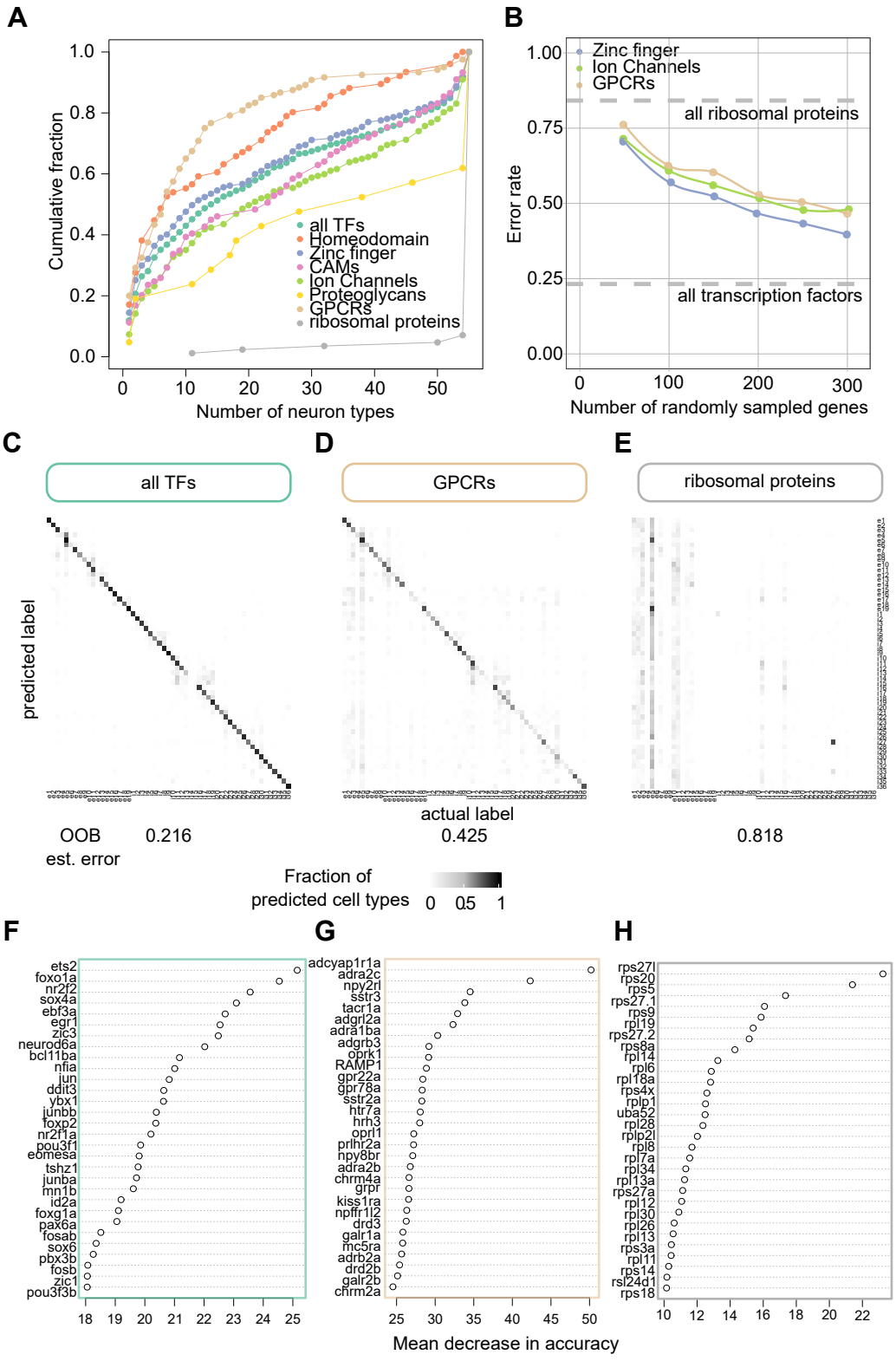
A
Cumulative fraction
Number of neuron types
all TFs
Homeodomain
Zinc finger
CAMs
Ion Channels
Proteoglycans
GPCRs
ribosomal proteins
B
Error rate
Number of randomly sampled genes
Zinc finger
Ion Channels
GPCRs
all ribosomal proteins
all transcription factors
C
D
E
all TFs
GPCRs
ribosomal proteins
predicted label
actual label
OOB est. error
0.216
0.425
0.818
Fraction of predicted cell types
0
0.5
1
F
G
H
ets2
foxo1a
nr2f2
sox4a
ebf3a
egr1
zic3
neurod6a
bcl11ba
nfia
jun
ddit3
ybx1
junbb
foxp2
nr2f1a
pou3f1
eomesa
tshz1
junba
mn1b
id2a
foxg1a
pax6a
fosab
sox6
pbx3b
fosb
zic1
pou3f3b
adcyap1r1a
adra2c
npy2rl
sstr3
tacr1a
adgrl2a
adra1ba
adgrb3
oprk1
RAMP1
gpr22a
gpr78a
sstr2a
htr7a
hrh3
oprl1
prlhr2a
npy8br
adra2b
chrm4a
grpr
kiss1ra
npffr1l2
drd3
galr1a
mc5ra
adrb2a
drd2b
galr2b
chrm2a
rps27l
rps20
rps5
rps27.1
rps9
rpl19
rps27.2
rps8a
rpl14
rpl6
rpl18a
rps4x
rplp1
uba52
rpl28
rplp2l
rpl8
rpl7a
rpl34
rpl13a
rps27a
rpl12
rpl30
rpl26
rpl13
rps3a
rpl11
rps14
rsl24d1
rps18
Mean decrease in accuracy

Figure S3. Cell-type specificity of gene expression, related to Figure 2.
(**A**) Cumulative fraction of information about neuron types contained in expression patterns of different gene families (TFs, transcription factors; Homeodomain TFs; Zinc finger TFs, CAMs, cell adhesion molecules; ion channels; proteoglycans; ribosomal proteins). Information content of individual gene families was assessed by classifying the 55 neuronal cell types which were Louvain-clustered based on the entire transcriptome. This analysis is similar to Figure 1D but follows the computation in Hain et al. 2022[S4]. (**B**) Plot of randomly subsampled sets of different gene families and the resulting accuracy of the corresponding random forest classifier, bounded by the accuracy of the ribosomal gene classifier and the classifier trained on the entire set of transcription factors. Different accuracy in random forest classification might thus not only be attributable to degree of heterogeneity in the expression pattern, but simply to the number of genes available per gene family. (**C**-**E**) Cell type classification performance of random forest classifiers trained on expression patterns of transcription factors, GPCRs, or ribosomal genes. The number below each plot indicates the out-of-bag error estimate. Note high performance of classifiers trained on TFs and GPCRs but not ribosomal genes. (**F-H**) Most informative genes for each of the classifier, ranked by mean decrease in accuracy when left out of the classifier.

A

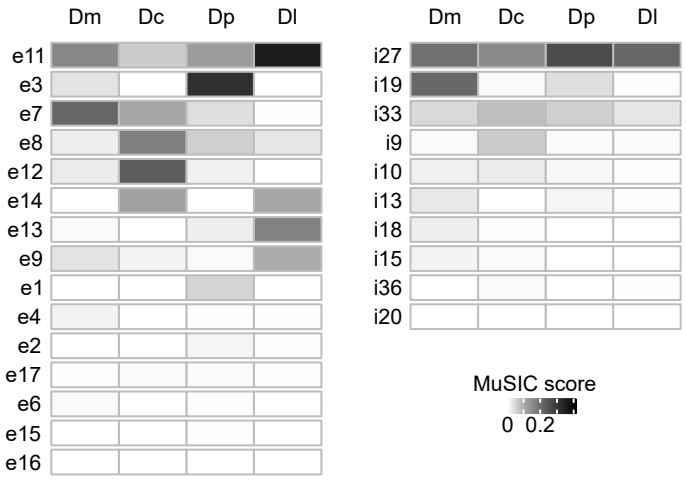
Dm
Dc
Dp
Dl
e11
e3
e7
e8
e12
e14
e13
e9
e1
e4
e2
e17
e6
e15
e16
Dm
Dc
Dp
Dl
i27
i19
i33
i9
i10
i13
i18
i15
i36
i20
MuSIC score
0
0.2

Figure S4. Contribution of individual cell types to pallial areas, related to Figure 4.
(**A**) Deconvolved contribution of individual cell types to pallial areas, computed by the MuSIC algorithm[S5]. Heatmaps are separated into glutamatergic and GABAergic cells and sorted by maximum contribution score.

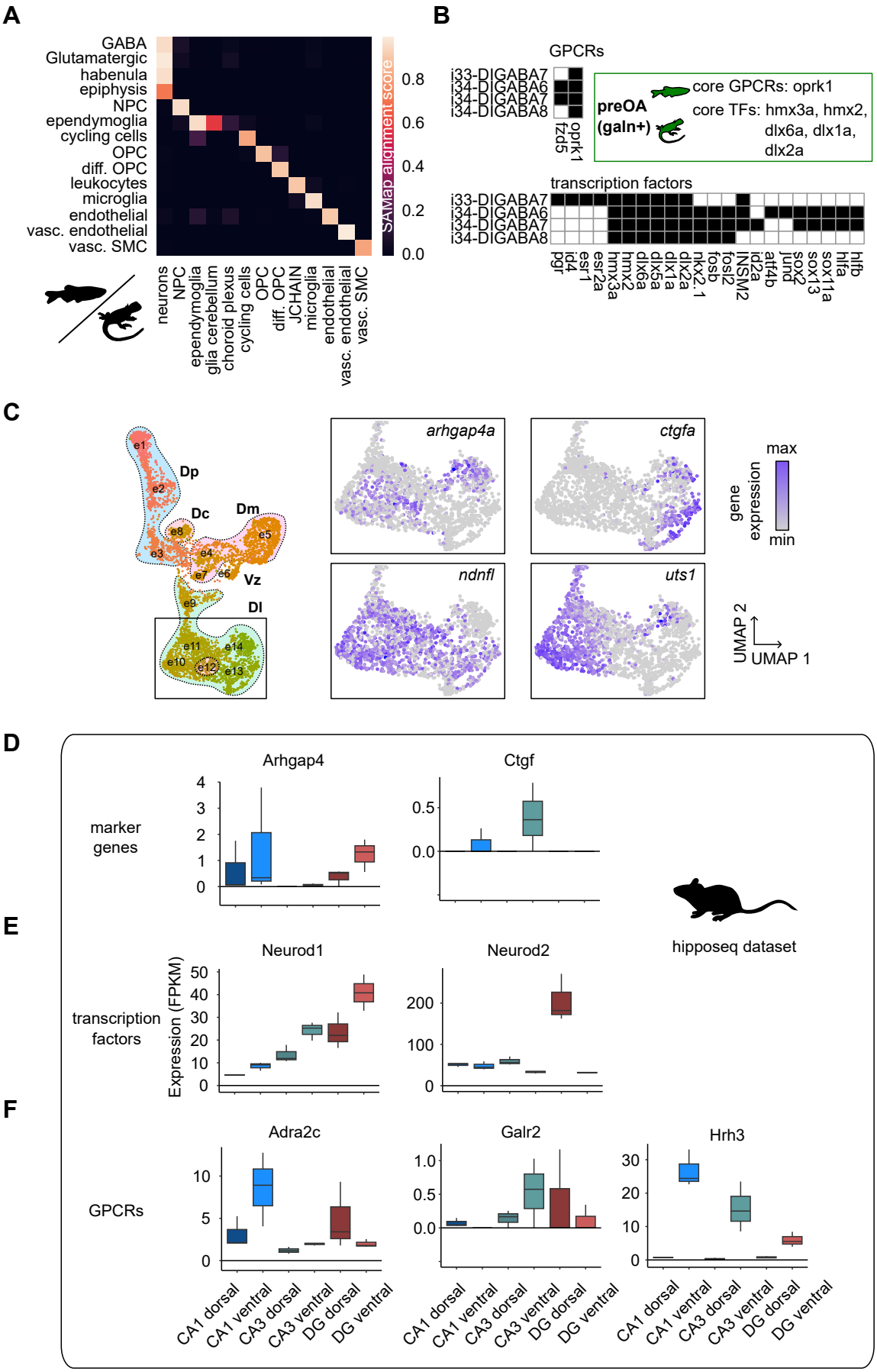
A
GABA
Glutamatergic
habenula
epiphysis
NPC
ependymoglia
cycling cells
OPC
diff. OPC
leukocytes
microglia
endothelial
vasc. endothelial
vasc. SMC
SAMap alignment score
0.8
0.6
0.4
0.2
0.0
neurons
NPC
ependymoglia
glia cerebellum
choroid plexus
cycling cells
OPC
diff. OPC
JCHAIN
microglia
endothelial
vasc. endothelial
vasc. SMC
B
GPCRs
i33-DIGABA7
i34-DIGABA6
i34-DIGABA7
i34-DIGABA8
oprk1
fzd5
preOA
(galn+)
core GPCRs: oprk1
core TFs: hmx3a, hmx2, dlx6a, dlx1a, dlx2a
transcription factors
i33-DIGABA7
i34-DIGABA6
i34-DIGABA7
i34-DIGABA8
pgr
id4
esr1
esr2a
hmx3a
hmx2
dlx6a
dlx5a
dlx1a
dlx2a
nkx2.1
fosb
fosl2
INSM2
id2a
atf4b
jund
sox2
sox13
sox11a
hifa
hlfb
C
e1
Dp
e2
Dc
Dm
e8
e5
e3
e4
e7
e6
Vz
e9
Dl
e11
e14
e10
e12
e13
arhgap4a
ctgfa
ndnfl
uts1
max
gene expression
min
UMAP 2
UMAP 1
D
marker genes
Arhgap4
Ctgf
E
transcription factors
Neurod1
Neurod2
Expression (FPKM)
hipposeq dataset
F
GPCRs
Adra2c
Galr2
Hrh3
CA1 dorsal
CA1 ventral
CA3 dorsal
CA3 ventral
DG dorsal
DG ventral

Figure S5. Evolutionary conservation of cell superclusters between zebrafish and bearded dragon, related to Figure 6.

(**A**) Alignment score between corresponding superclusters in zebrafish and the bearded dragon. The sparsity of the matrix implies specificity of alignment. vasc. SMC, vascular smooth muscle cells; OPC, oligodendrocyte precursor cells; diff. OPC, differentiating OPC; NPC; neural progenitor cells, JCHAIN, plasma cells. (**B**) Binarized expression heatmap of GPCRs and TFs significantly contributing to the alignment of galanin-expressing cells in the preoptic area of the zebrafish and hypothalamic galanin+ cells in the lizard. Both appear to be susceptible to modulation via the k-opioid receptor in both species. (**C**) Enlargement of the area in UMAP space where Dl-associated cell types are localized (e10, e11, e13, e14). Expression levels of the four marker genes *arhgap4a*, *ctgfa*, *ndnfl*, and *uts1* within this area are shown. (**D** – **E**) Expression levels of the mouse genes orthologous to zebrafish marker genes, TFs and GPCRs enriched in Dl taken from the hipposeq dataset[S6].

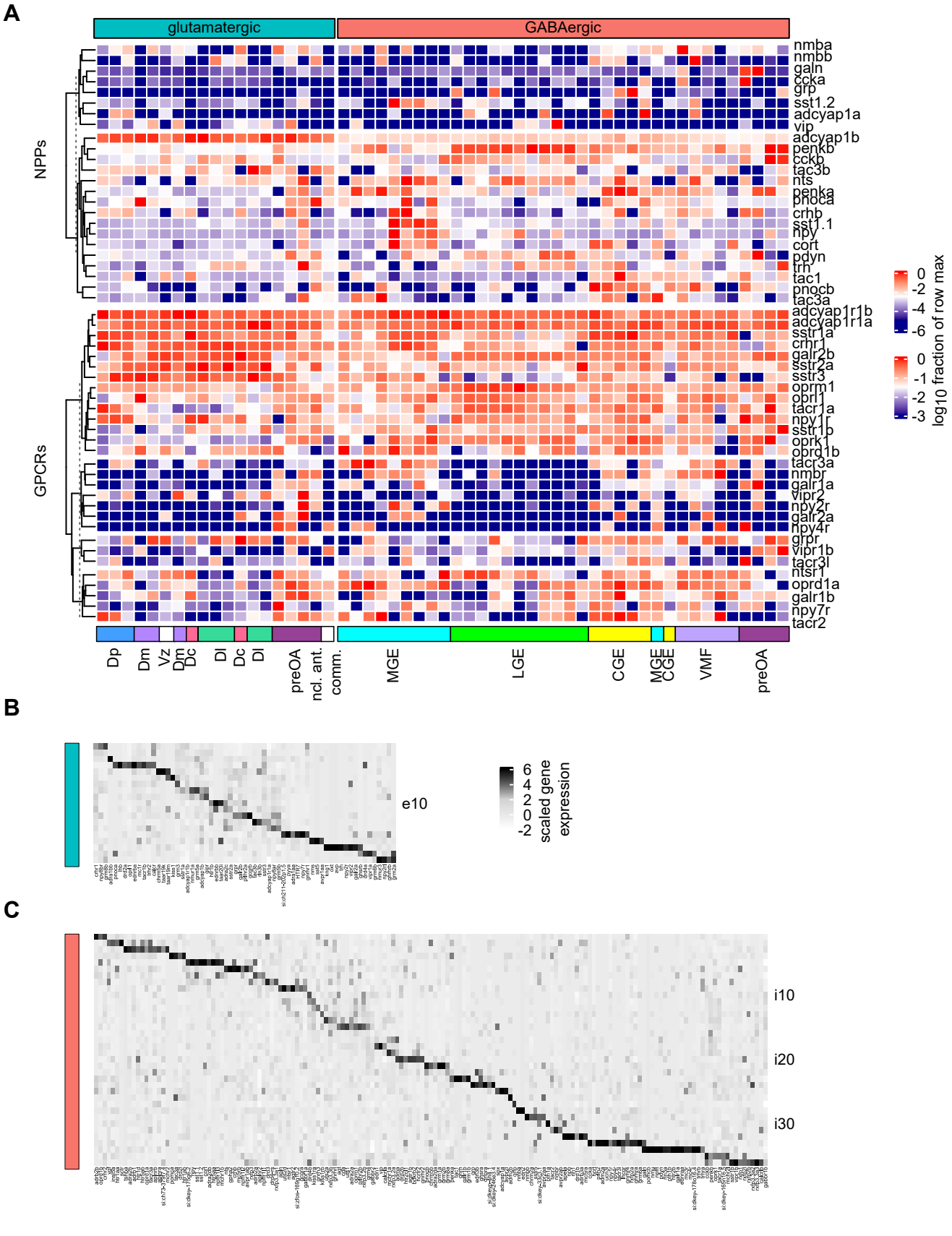
A
glutamatergic
GABAergic
NPPs
GPCRs
nmba
nmbb
galn
ccka
grp
sst1.2
adcyap1a
vip
adcyap1b
penkb
cckb
tac3b
nts
penka
pnoca
crhb
sst1.1
npy
cort
pdyn
trh
tac1
pnocb
tac3a
adcyap1r1b
adcyap1r1a
sstr1a
crhr1
galr2b
sstr2a
sstr3
oprm1
oprl1
tacr1a
npy1r
sstr1b
oprk1
oprd1b
tacr3a
nmbr
galr1a
vipr2
npy2r
galr2a
npy4r
grpr
vipr1b
tacr3l
ntsr1
oprd1a
galr1b
npy7r
tacr2
0
-2
-4
-6
0
-1
-2
-3
log10 fraction of row max
Dp
Dm
Vz
Dm
Dc
Dl
Dc
Dl
preOA
ncl. ant.
comm.
MGE
LGE
CGE
MGE
CGE
VMF
preOA
B
e10
6
4
2
0
-2
scaled gene expression
C
i10
i20
i30

Figure S6. Heterogeneity and specificity of telencephalic neuromodulatory networks, related to Figure 7.

(**A**) Normalized expression of selected NPPs and GPCRs in GABAergic and glutamatergic cell types identified in our dataset. Expression is shown using a logarithmic color scale after normalization to the maximum in each row. Rows were ordered by hierarchical clustering (left). (**B, C**) Enlargements of Figure 7H showing normalized (z-scored) expression levels of all genes associated with neuromodulation separately in glutamatergic and GABAergic cell types. Matrices were sorted by maximum expression per cell type.

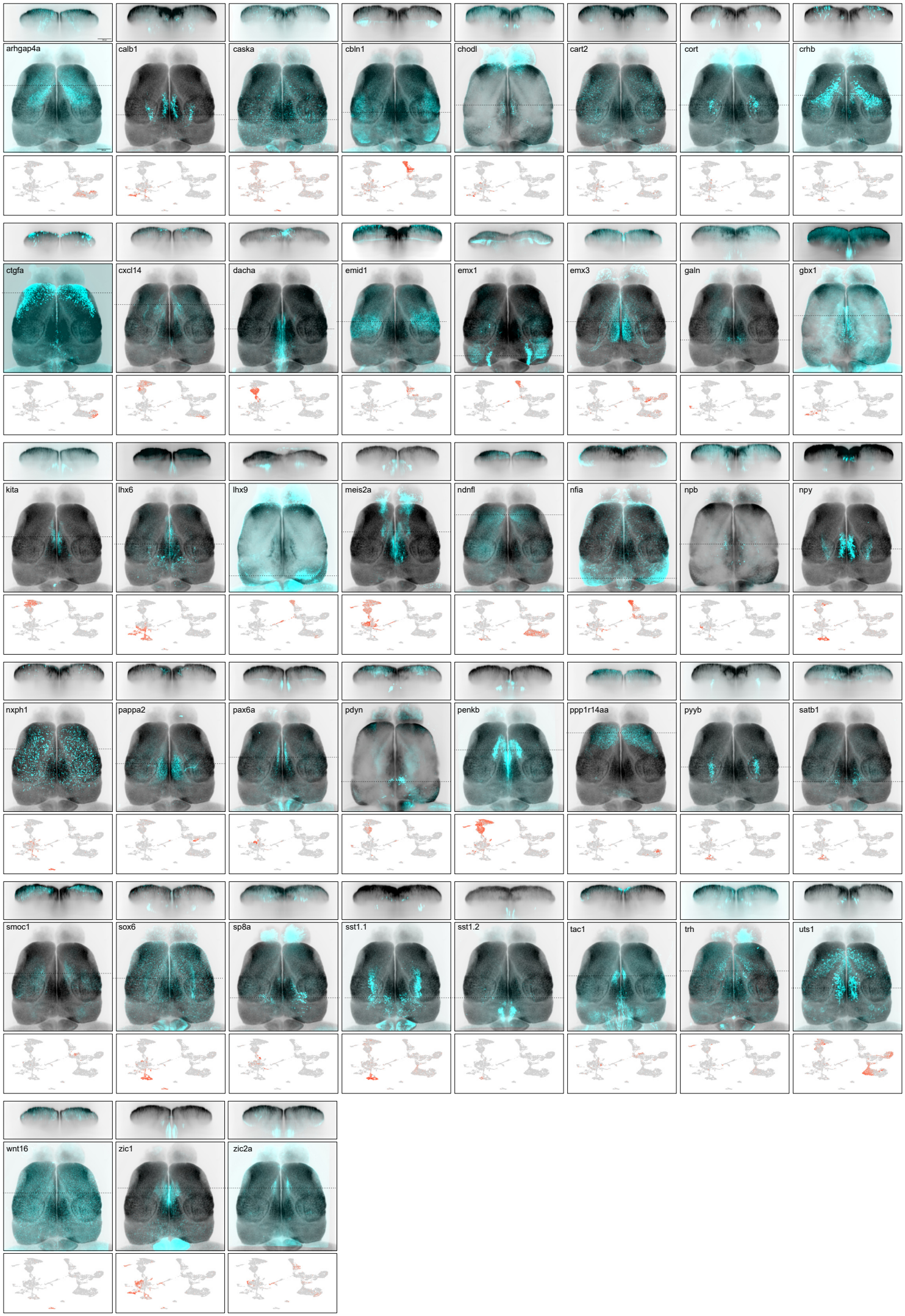
arhgap4a
calb1
caska
cbln1
chodl
cart2
cort
crhb
ctgfa
cxcl14
dacha
emid1
emx1
emx3
galn
gbx1
kita
lhx6
lhx9
meis2a
ndnfl
nfia
npb
npy
nxph1
pappa2
pax6a
pdyn
penkb
ppp1r14aa
pyyb
satb1
smoc1
sox6
sp8a
sst1.1
sst1.2
tac1
trh
uts1
wnt16
zic1
zic2a

Figure S7. HCR-based visualization of marker genes for individual cell types, related to Figures 4 – 6.

Expression patterns of 43 marker genes in the adult telencephalon. Images show maximum-intensity z-projections (dorsal views) and a coronal section at the anterior-posterior position indicated by the dashed line. Expression patterns were determined by HCR (cyan) and mapped onto a common reference brain using *snap25a* expression for registration (gray background). Scale bar: approximately 200 µm (not precise due to registration of image datasets onto the reference brain). Plots below show gene expression levels in individual cells mapped onto the first two UMAP components (Figure 2C). Genes are ordered alphabetically.

Gene	Dc	Dl	Dp	DM
adora1b	-0,95	-0,65	0,80	0,74
adora2aa	0,67	-0,58	-0,24	0,10
adora2b	1,19	-0,75	-0,45	-0,05
adra2b	-0,58	-0,63	0,62	0,54
adra2da	0,17	-0,16	0,82	-0,84
adra2db	0,84	-0,42	0,00	-0,46
adrb1	-0,49	-0,49	-0,36	1,31
agtr2	1,03	-0,17	-0,16	-0,72
aplnra	0,67	-0,08	-0,13	-0,46
aplnrb	1,11	-0,56	-0,23	-0,37
avpr2l	0,44	-0,31	0,29	-0,45
calcrla	1,14	-0,59	-0,71	0,12
chrm2a	1,23	-0,69	-0,04	-0,55
drd1b	0,89	-0,09	-0,34	-0,47
drd3	-0,76	-0,16	0,65	0,25
drd4a	-0,51	-0,64	-0,13	1,23
ednraa	0,72	-0,15	0,17	-0,74
ednrba	-0,47	-0,52	0,13	0,82
fshr	-0,33	-0,65	-0,15	1,07
galr1b	-0,64	-0,87	0,35	1,09
galr2b	-0,23	-0,52	-0,52	1,22
gnrhr1	-0,55	-0,49	0,83	0,17
gnrhr2	1,00	-0,37	-0,33	-0,33
hrh2a	-0,42	-0,43	-0,35	1,17
hrh3	0,84	0,63	-0,81	-0,61
htr1b	-0,70	-0,56	-0,10	1,31
htr1d	-0,11	-0,68	-0,37	1,10
htr2cl1	0,30	-0,72	0,83	-0,47
htr5aa	-0,67	0,28	-0,27	0,68
kiss1ra	1,20	-0,79	-0,66	0,19
lpar1	0,25	0,47	0,24	-0,92
lpar2b	0,78	0,08	-0,16	-0,69
lpar6a	0,80	-0,12	-0,10	-0,59
mc3r	-0,34	-0,33	-0,02	0,66
mtnr1aa	-0,75	-0,73	0,45	0,97
mtnr1al	-0,49	-0,50	-0,38	1,33
mtnr1bb	-0,30	-0,35	-0,14	0,76
npy1r	0,17	-0,22	0,66	-0,63
npy4r	0,65	-0,35	0,07	-0,40
npy7r	-0,45	-0,44	-0,18	1,04
npy8ar	0,98	0,23	-0,74	-0,44
npy8br	1,10	-0,96	0,29	-0,51

oprd1b	-0,66	0,93	-0,71	0,51
oprl1	0,64	0,00	-0,52	-0,12
p2ry11	0,42	-0,07	0,00	-0,36
prlhr2a	-0,53	-0,25	0,00	0,76
ptger1c	0,87	-0,76	0,57	-0,75
ptger2b	-0,55	0,03	-0,42	0,94
ptger4b	-0,35	-0,11	-0,39	0,84
rxfp2a	-0,36	-0,74	-0,32	1,36
s1pr1	-1,21	0,43	0,03	0,79
tacr1a	-0,58	-0,58	1,01	0,10
tacr2	-0,77	-0,36	0,83	0,27
tacr3a	-0,71	0,17	0,41	0,14
tbxa2r	-0,58	0,07	-0,51	1,04
trhr2	1,23	-0,17	-0,51	-0,56
trhrb	-0,68	0,06	-0,28	0,90
vipr1b	0,61	-0,79	0,56	-0,44
vipr2	1,09	-0,84	-0,06	-0,27

Table S1: list of GPCRs and their normalized bulk expression, related to Figure 1.

SUPPLEMENTAL REFERENCES

S1. Cosacak, M.I., Bhattarai, P., Reinhardt, S., Petzold, A., Dahl, A., Zhang, Y., and Kizil, C. (2019). Single-Cell Transcriptomics Analyses of Neural Stem Cell Heterogeneity and Contextual Plasticity in a Zebrafish Brain Model of Amyloid Toxicity. Cell Reports *27*, 1307-1318.e3. 10.1016/j.celrep.2019.03.090.

S2. Raj, B., Wagner, D.E., McKenna, A., Pandey, S., Klein, A.M., Shendure, J., Gagnon, J.A., and Schier, A.F. (2018). Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. Nature Biotechnology *36*, 442–450. 10.1038/nbt.4103.

S3. Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N., and Junker, J. (2018). Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. Nature biotechnology *36*, 469–473. 10.1038/nbt.4124.

S4. Hain, D., Gallego-Flores, T., Klinkmann, M., Macias, A., Ciirdaeva, E., Arends, A., Thum, C., Tushev, G., Kretschmer, F., Tosches, M.A., et al. (2022). Molecular diversity and evolution of neuron types in the amniote brain. Science *377*. 10.1126/science.abp8202.

S5. Wang, X., Park, J., Susztak, K., Zhang, N.R., and Li, M. (2019). Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. Nat. Commun. *10*, 380. 10.1038/s41467-018-08023-x.

S6. Cembrowski, M.S., Wang, L., Sugino, K., Shields, B.C., and Spruston, N. (2016). Hipposeq: a comprehensive RNA-seq database of gene expression in hippocampal principal neurons. eLife *5*, e14997. 10.7554/elife.14997.