Supplemental Data

Supplemental Figures and Legends:

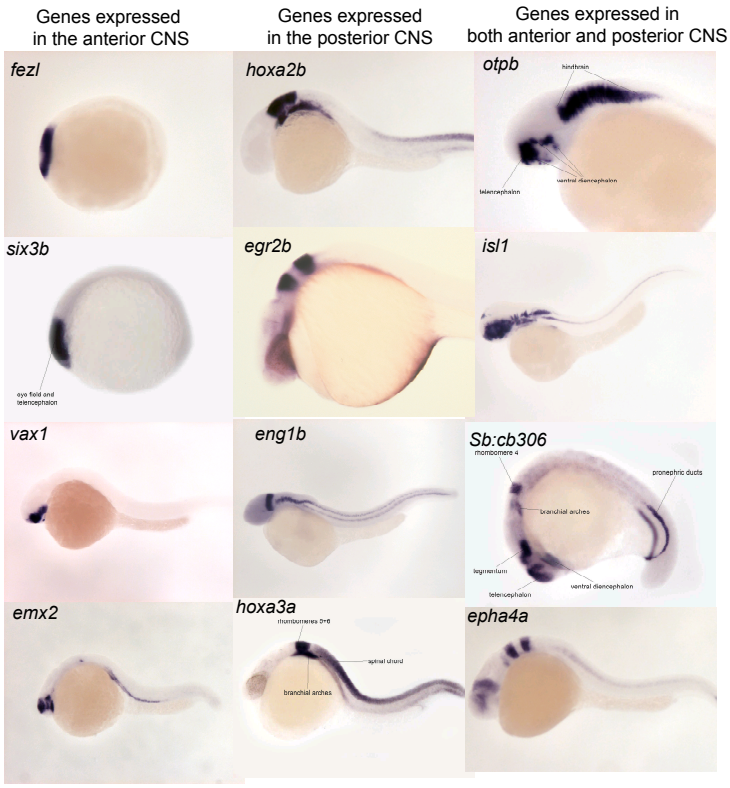
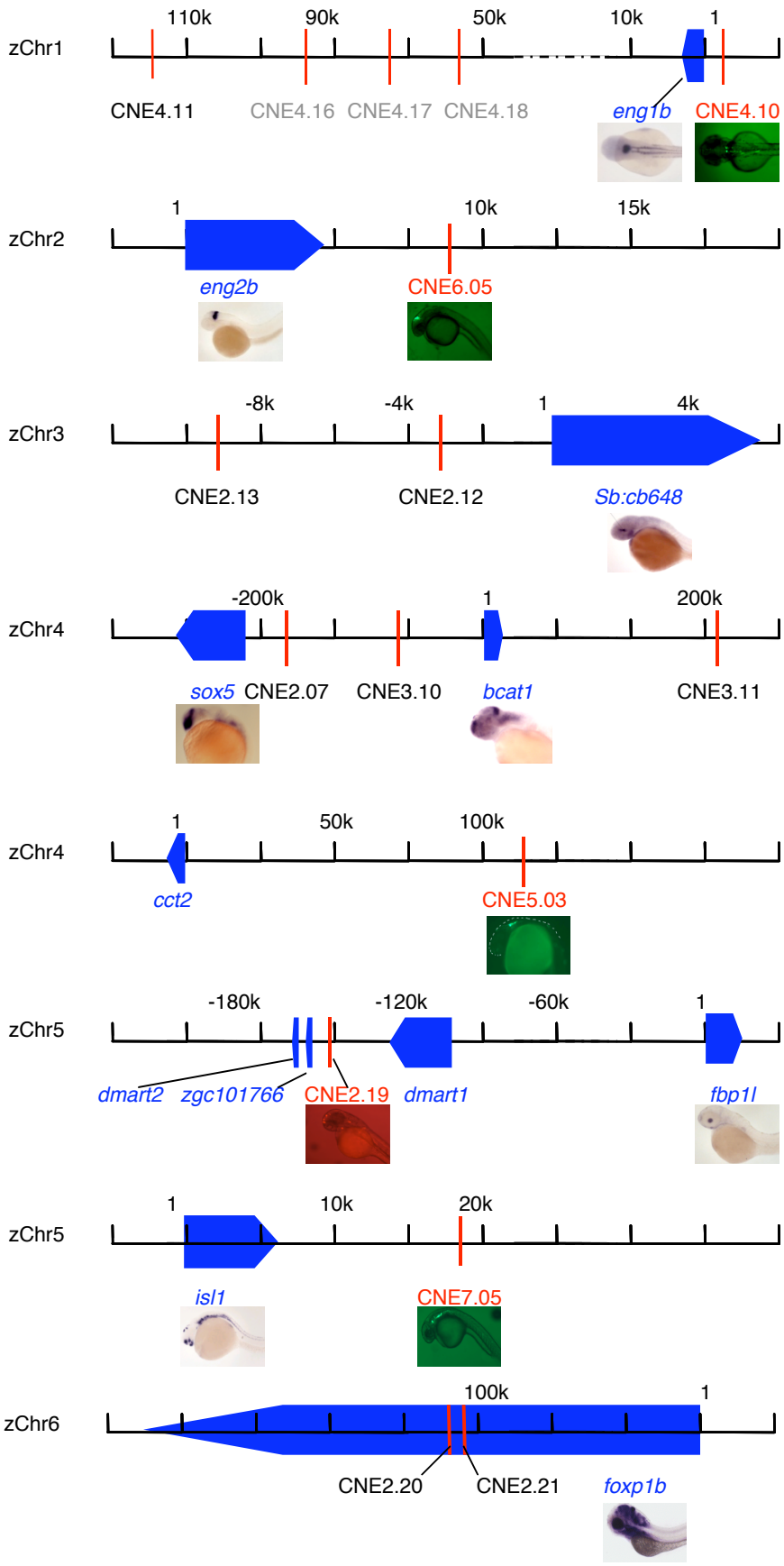
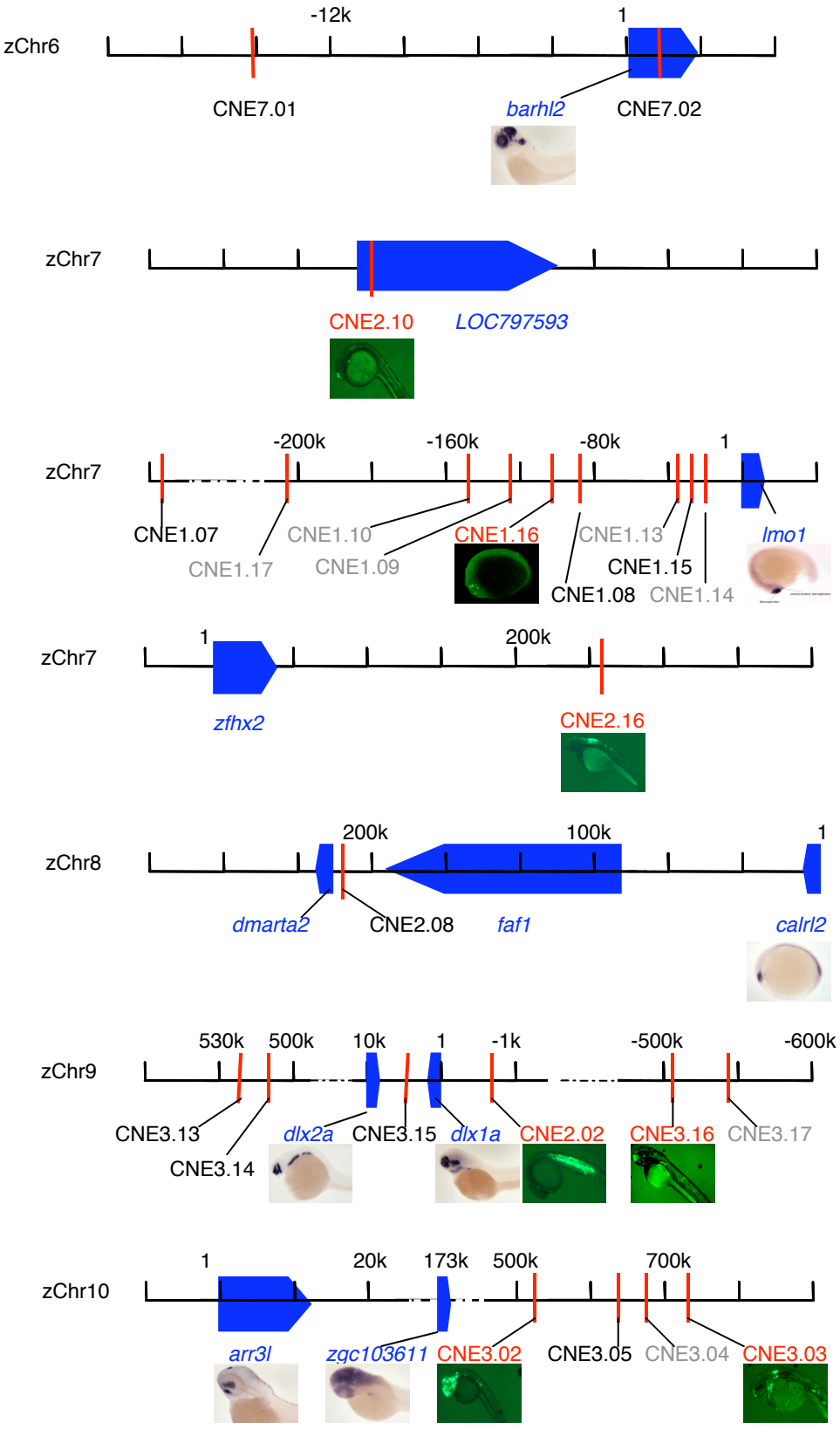
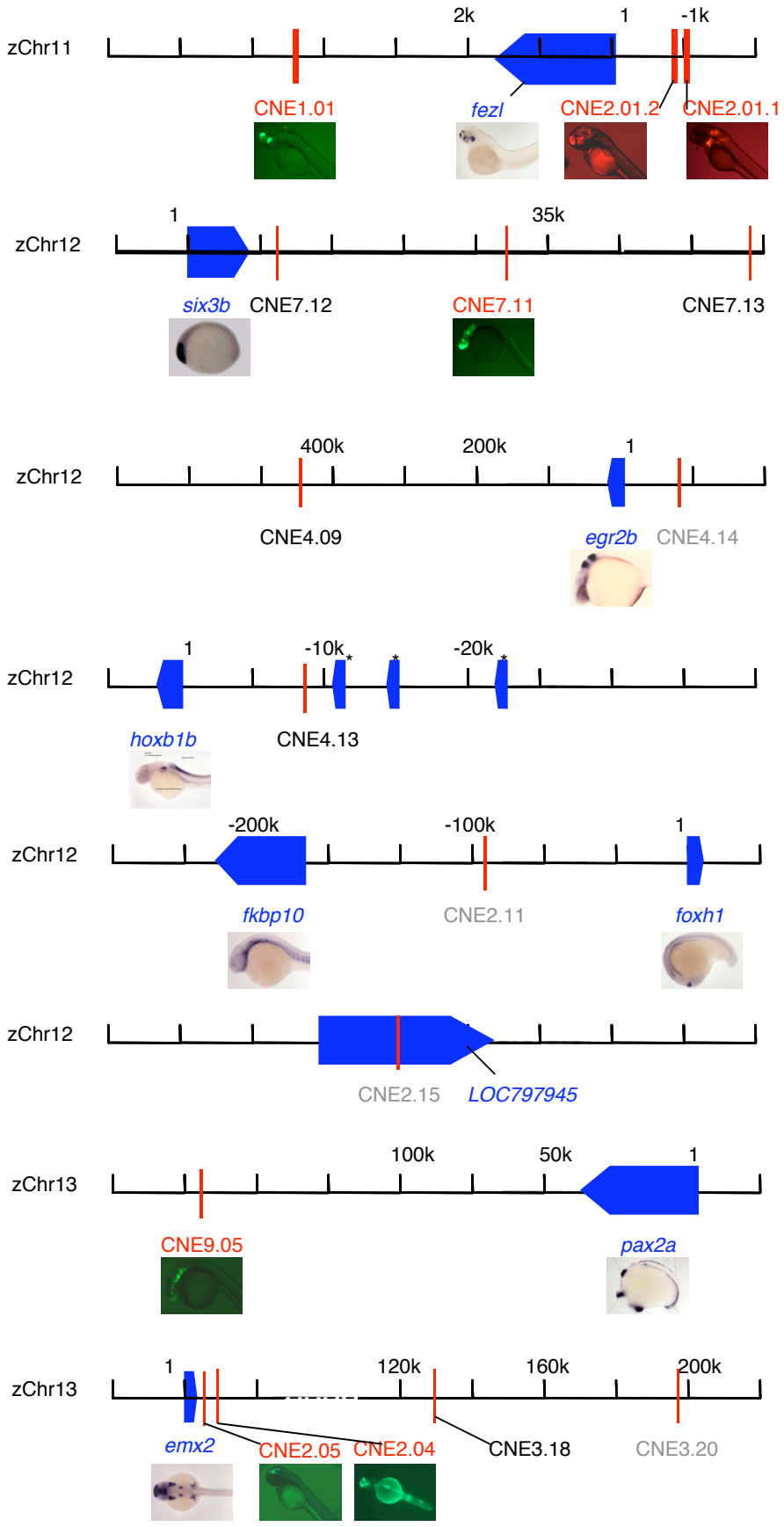
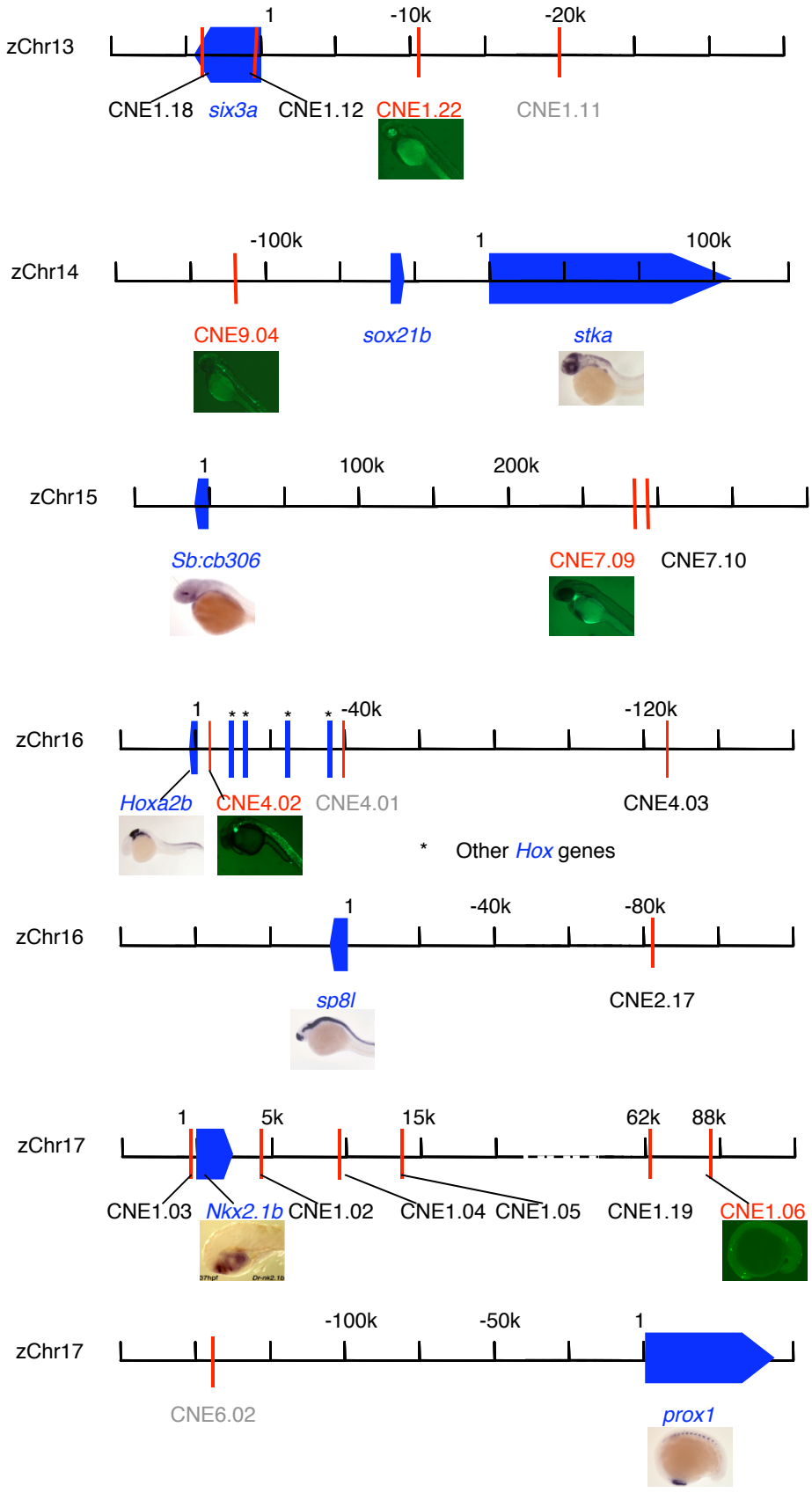


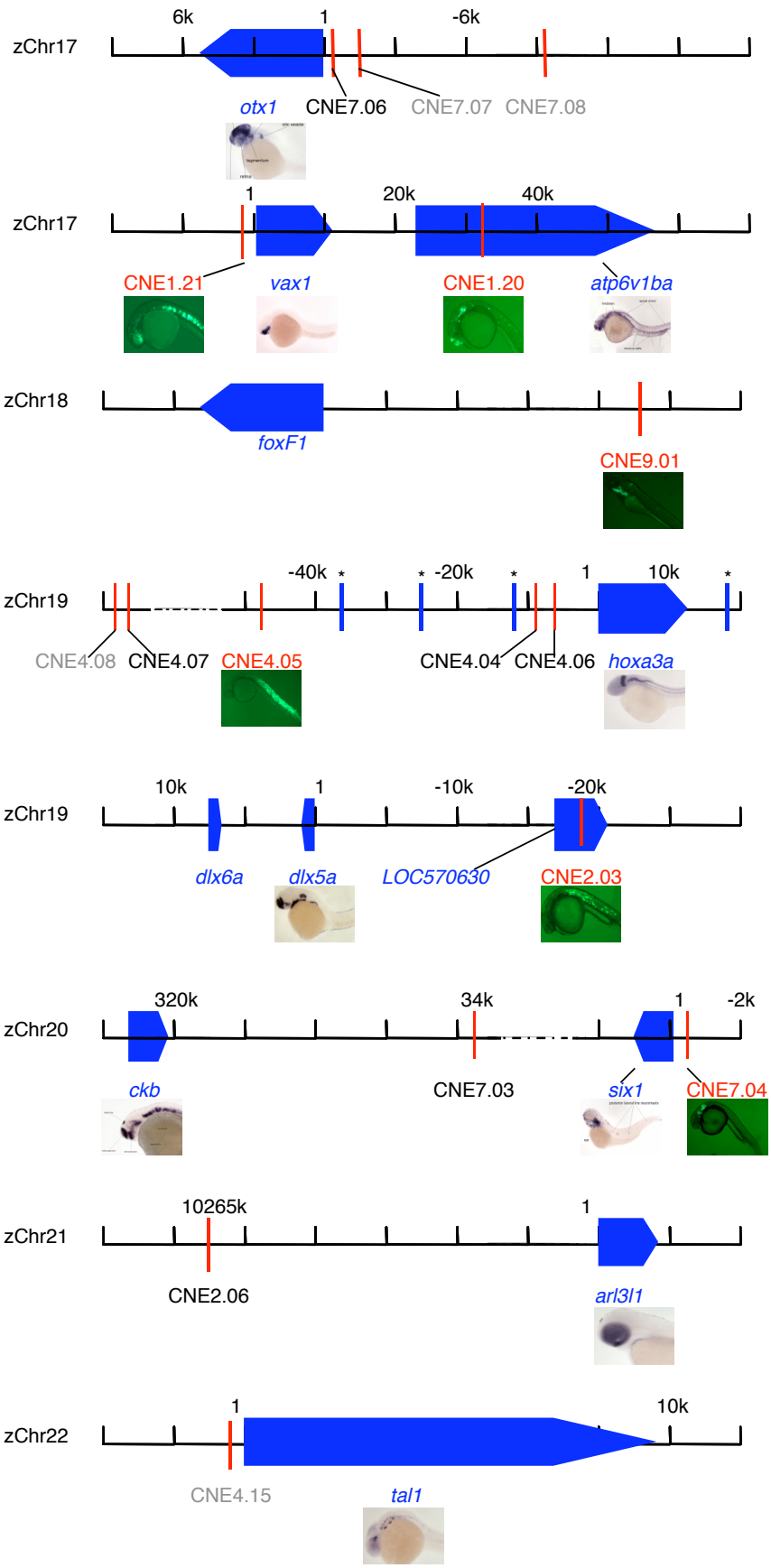
Figure S1. Selected images showing brain region-specific gene expression. Images were taken with permission from zfin gene expression database (Thisse and Thisse, 2005).


zChr1
110k
90k
50k
10k
1
CNE4.11
CNE4.16
CNE4.17
CNE4.18
eng1b
CNE4.10
zChr2
1
10k
15k
eng2b
CNE6.05
zChr3
-8k
-4k
1
4k
CNE2.13
CNE2.12
Sb:cb648
zChr4
-200k
1
200k
sox5
CNE2.07
CNE3.10
bcat1
CNE3.11
zChr4
1
50k
100k
cct2
CNE5.03
zChr5
-180k
-120k
-60k
1
dmart2
zgc101766
CNE2.19
dmart1
fbp1l
zChr5
1
10k
20k
isl1
CNE7.05
zChr6
100k
1
CNE2.20
CNE2.21
foxp1b


zChr6
-12k
1
CNE7.01
barhl2
CNE7.02
zChr7
CNE2.10
LOC797593
zChr7
-200k
-160k
-80k
1
CNE1.07
CNE1.10
CNE1.16
CNE1.13
lmo1
CNE1.17
CNE1.09
CNE1.15
CNE1.08
CNE1.14
zChr7
1
200k
zfhx2
CNE2.16
zChr8
200k
100k
1
dmarta2
CNE2.08
faf1
calrl2
zChr9
530k
500k
10k
1
-1k
-500k
-600k
CNE3.13
CNE3.14
dlx2a
CNE3.15
dlx1a
CNE2.02
CNE3.16
CNE3.17
zChr10
1
20k
173k
500k
700k
arr3l
zgc103611
CNE3.02
CNE3.05
CNE3.04
CNE3.03


zChr11
2k
1
-1k
CNE1.01
fezl
CNE2.01.2
CNE2.01.1
zChr12
1
35k
six3b
CNE7.12
CNE7.11
CNE7.13
zChr12
400k
200k
1
CNE4.09
egr2b
CNE4.14
zChr12
1
-10k
-20k
hoxb1b
CNE4.13
zChr12
-200k
-100k
1
fkbp10
CNE2.11
foxh1
zChr12
CNE2.15
LOC797945
zChr13
100k
50k
1
CNE9.05
pax2a
zChr13
1
120k
160k
200k
emx2
CNE2.05
CNE2.04
CNE3.18
CNE3.20


zChr13
1
-10k
-20k
CNE1.18
six3a
CNE1.12
CNE1.22
CNE1.11
zChr14
-100k
1
100k
CNE9.04
sox21b
stka
zChr15
1
100k
200k
Sb:cb306
CNE7.09
CNE7.10
zChr16
1
-40k
-120k
Hoxa2b
CNE4.02
CNE4.01
CNE4.03
* Other Hox genes
zChr16
1
-40k
-80k
sp8l
CNE2.17
zChr17
1
5k
15k
62k
88k
CNE1.03
Nkx2.1b
CNE1.02
CNE1.04
CNE1.05
CNE1.19
CNE1.06
zChr17
-100k
-50k
1
CNE6.02
prox1


zChr17
6k
1
-6k
otx1
CNE7.06
CNE7.07
CNE7.08
zChr17
1
20k
40k
CNE1.21
vax1
CNE1.20
atp6v1ba
zChr18
foxF1
CNE9.01
zChr19
-40k
-20k
1
10k
CNE4.08
CNE4.07
CNE4.05
CNE4.04
CNE4.06
hoxa3a
zChr19
10k
1
-10k
-20k
dlx6a
dlx5a
LOC570630
CNE2.03
zChr20
320k
34k
1
-2k
ckb
CNE7.03
six1
CNE7.04
zChr21
10265k
1
CNE2.06
arl3l1
zChr22
1
10k
CNE4.15
tal1

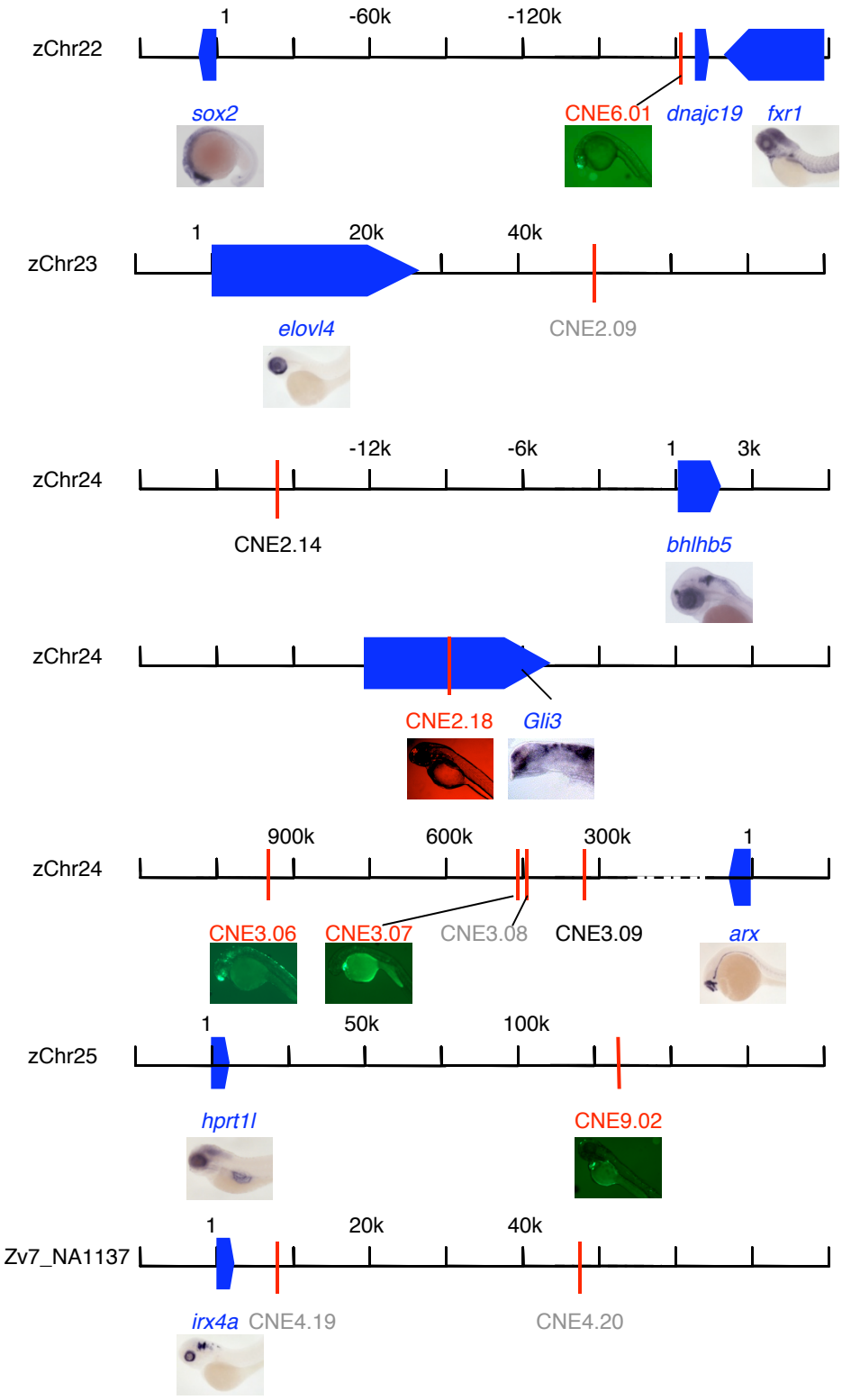


Figure S2. Schematic diagrams depicting the chromosomal location and activity of analyzed CNEs, as well as their nearby brain-expressed genes.
An GFP image of living zebrafish embryos was shown for the CNEs with region-specific enhancer activity. CNEs with no enhancer activity are labeled in gray. Images were not shown for CNEs with broad enhancer activity.

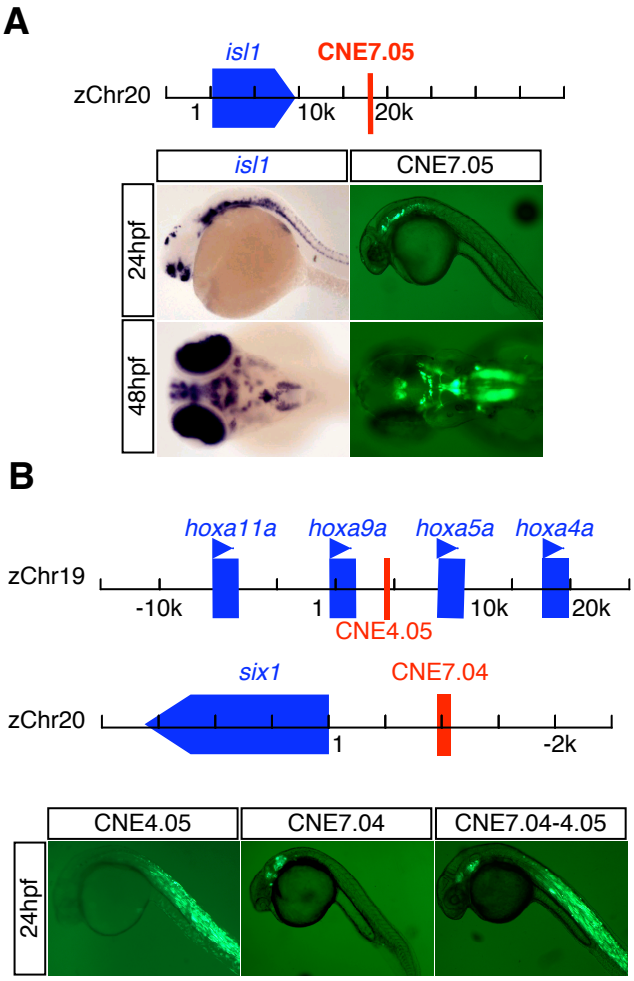


Figure S3. CNEs as versatile tools to driving gene expression in desired tissues.
(A) CNE7.05 displays activity in subsets of cells expressing the *islet1* gene.
(B) Combined activity is observed when CNE7.04 and 4.05 were used together to drive reporter expression.

Supplementary online text:

TRANSFAC predicted factors that bind to motifs (Table 4)

A. Our functional study showed that the motif TTCATT (AATGAA in reverse complement) is essential for the telencephalic and diencpehalic enhancer activity of CNE1.01 (Figure 6B, Panel 2). In TRANSFAC, TTCATT is predicted as a binding site for ICSBP/IRF. IRFs (interferon regulatory factor) are a nine-member family of transcription factors that share a highly conserved helix-turn-helix DNA-binding domain and a less conserved protein-binding domain. The IRFs exert diverse functions through homo- or hetero-dimer formation or interaction with other transcription factors. These functions include the regulation of host defense pathways such as the innate and adaptive immune responses, and oncogenesis (Nguyen et al., 1997; Taniguchi et al., 2001). In zebrafish, the expression of IRF2 (broad expression), 2a (broad expression), 4b (brain-enriched), 5 (broad), 6 (non-neural tissue), 7 (multiple regions including telencpehalon and midbrain), 9 (broad), 10 (broad) have been characterized (ZFIN gene expression database)(Ben et al., 2005; Thisse and Thisse, 2005), but their function in brain development is not known. Based on the expression patterns, 4b and 7 represent candidates for future studies.
B. Our functional study showed that the motif TCCATT (reverse complement: AATGGA) is essential for the telencephalic and diencpehalic enhancer activity of CNE1.01 (Figure 6B, Panel 5). In TRANSFAC, TCCATT is predicted as a binding site for YY1. The transcription factor Yin Yang 1 (YY1) is a multifunctional protein that can activate or repress gene expression depending on the cellular context. YY1 is ubiquitously expressed and highly conserved between species. However, its role varies in different cell types and includes proliferation, differentiation, and apoptosis (He and Casaccia-Bonnefil, 2008). In mice, Acetylated YY1 is shown to regulate Otx2 expression in anterior neuroectoderm (Takasaki et al., 2007). In zebrafish, both YY1a and YY1b show broad but brain-enriched expression (Thisse and Thisse, 2005), although their functions are unknown. Based on this information, YY1 represents a candidate factor for future study.
C. Our functional study showed that the motif CTGAAA (reverse complement: TTTCAG) is present in two copies in CNE2.01.2. Together they are essential for the CNE enhancer activity in the forebrain (Figure 5B). In TRANSFAC, CTGAAA is predicted as a binding site for FOXP3, a forkhead transcription factor critical for the development of regulatory T cells (Curiel, 2007). In zebrafish, *foxp3* expression shows a non-specific pattern during early development. There is no report on the role of FoxP3 in zebrafish or in brain development.
D. Our functional study showed that the motif CCCCTC (reverse complement: GAGGGG) is essential to repress ectopic enhancer activity in the eyes and trunk. In TRANSFAC, CCCCTC is predicted as a binding site for MZF1 (Myeloid zinc finger 1). MZF1 is a bi-functional transcriptional regulator, repressing transcription in non-hematopoietic cells, and activating transcription in cells of

hematopoietic origins (Hromas et al., 1996). However, its expression and function is unknown in zebrafish.

E. Our functional study showed that the motif CCGCTC (reverse complement: GAGCGG) is essential for the telencephalic and diencpehalic enhancer activity of CNE1.01 (Figure 6B, Panel 4). In TRANSFAC, CCGCTC is predicted as a binding site for Pax1. Pax1 is involved in vertebral skeletal development (Bannykh et al., 2003), but no information is available for its expression or function in zebrafish.

References:

Bannykh, S. I., Emery, S. C., Gerber, J. K., Jones, K. L., Benirschke, K., Masliah, E., 2003. Aberrant Pax1 and Pax9 expression in Jarcho-Levin syndrome: report of two Caucasian siblings and literature review. . Am. J. Med. Genet. A. 120A(2)**,** 241-6.

Ben, J., Jabs, E. W., Chong, S. S., 2005. Genomic, cDNA and embryonic expression analysis of zebrafish IRF6, the gene mutated in the human oral clefting disorders Van der Woude and popliteal pterygium syndromes. . Gene Expr. Patterns. 5**,** 629-38.

Curiel, T. J., 2007. Regulatory T-cell development: is Foxp3 the decider? Nat. Med. 13**,** 250.

He, Y., Casaccia-Bonnefil, P., 2008. The Yin and Yang of YY1 in the nervous system. . J. Neurochem. 106**,** 1493-502.

Hromas, R., et al., 1996. Hematopoietic transcriptional regulation by the myeloid zinc finger gene, MZF-1. . Curr. Top. Microbiol. Immunol. 211**,** 159-64.

Nguyen, H., Hiscott, J., Pitha, P. M., 1997. The growing family of interferon regulatory factors. Cytokine Growth Factor Rev. 8**,** 293-312.

Takasaki, N., Kurokawa, D., Nakayama, R., Nakayama, J., Aizawa, S., 2007. Acetylated YY1 regulates Otx2 expression in anterior neuroectoderm at two cis-sites 90 kb apart. EMBO J. 26**,** 1649-59.

Taniguchi, T., Ogasawara, K., Takaoka, A., Tanaka, N., 2001. IRF family of transcription factors as regulators of host defense. . Annu. Rev. Immunol. 19**,** 623-655.

Thisse, C., Thisse, B., High Throughput Expression Analysis of ZF-Models Consortium Clones. ZFIN Direct Data Submission (http://zfin.org). 2005.