Figure S1. Validation of zebrafish QRFP, Hcrt, and DBH transgenic reporter lines. All cells labeled by fluorescent ISH for *qrfp* and *dbh* are also labeled by anti-GFP immunostaining in *Tg(qrfp:EGFP, hcrt:mRFP)* embryos at 24 hpf (A) and *Tg(dbh:EGFP)* larvae at 120 hpf (C). All mRFP expressing cells in 24 hpf *Tg(qrfp:EGFP, hcrt:mRFP)* embryos are labeled by anti-Hcrt immunostaining (B). Anterior (A,B) and dorsal (C) views are shown. Scale = 10 μm.

Figure S2. Strategy to identify genes enriched in embryonic zebrafish Hcrt neurons. Transgenic zebrafish embryos were dissociated into single cells at 26 hpf and cells of interest were purified by FACS. RNA from purified neurons was amplified and hybridized to microarrays. Signals from Hcrt neurons were compared to other neuron types and genes most highly enriched in Hcrt neurons were identified. Highly enriched transcription factors and secreted peptides were selected for verification by ISH and overexpression analysis in zebrafish.

Figure S3. Expression patterns of genes identified as enriched in Hcrt neurons. ISH was performed using wild-type zebrafish embryos at 24 hpf. Anterior is left. Most candidate genes are expressed in a domain that is similar to, or overlaps with, the *hcrt* expression domain. Images in the bottom row were obtained from the ZFIN ISH database (Liu and Patient, 2008; Thisse et al., 2004).

Figure S4. Time course of endogenous *hcrt* and *lhx9* expression. Confocal projections of *Tg(hcrt:EGFP)* embryos at 24, 48, 72, 96, and 120 hpf show that all Hcrt cells express

lhx9 throughout development, as determined by fluorescent ISH for *lhx9* followed by anti-GFP immunostaining. Ventral images are shown. Scale = 10 μm.

Figure S5. Overexpression of candidate genes does not affect the number of Hcrt neurons in the hypothalamus. Overexpression of each candidate gene was induced by heat shock at 24 hpf and larvae were fixed at 120 hpf for ISH with a *hcrt* probe. Mean ± s.e.m. number of Hcrt cells in the entire hypothalamus is shown. Control larvae were injected with a HS-EGFP plasmid. n indicates number of larval brains analyzed. No significant difference was detected between any overexpressed gene and control ($p>0.05$ by one-way ANOVA). Note that ectopic Hcrt neurons (i.e. not in the hypothalamus) were only observed following overexpression of *lhx9* (Fig. 3).

Figure S6. *lhx9* overexpression induces ectopic QRFP neurons. (A-D) Ectopic *qrfp:EGFP* expressing neurons that persist until 96 hours post HS (120 hpf) also express ectopic *lhx9* (arrowheads in C, D). (E-H) Double fluorescent ISH for *lhx9* and *qrfp* in WT 120 hpf larvae shows an absence of *lhx9* expression in the hindbrain region that contains *lhx9*-induced ectopic *qrfp* neurons in (A-D). Boxed regions in (A, B, E, F) are shown at higher magnification in (C, D, G, H). (A, B, E, F) show 95 μm thick confocal maximum intensity projections containing both endogenous and ectopic QRFP neurons. (C, D, G, H) show 43 μm thick confocal maximum intensity projections including only the region containing ectopic QRFP neurons. *lhx9*-expressing neurons that appear close to ectopic QRFP neurons in (B) are located 30 μm ventral to the ectopic QRFP neurons, and are thus not observed in (D). Scale = 10 μm.

Figure S7. Expression of endogenous *qrfp* and *lhx9*. Confocal projections of WT embryos at 24 and 120 hpf show that all *qrfp*-expressing cells express *lhx9*, as determined by double fluorescent ISH. Ventral images are shown. Scale = 10 μm.

Figure S8. Hcrt neurons do not express *vglut2a* at 120 hpf. Confocal projections of a 120 hpf *Tg(vglut2a:mRFP, hcrt:EGFP)* larva shows that *vglut2a* cells labeled with mRFP do not colocalize with endogenous (A) or ectopic (C) Hcrt neurons labeled with EGFP. The regions shown in (A) and (C) are indicated with dashed boxes in (B) and (D), respectively. Scale = 10 μm.

Figure S9. Few or no Hcrt neurons express *vglut2b* at 120 hpf. Confocal projections of a 120 hpf *Tg(hcrt:EGFP)* larva shows that most Hcrt neurons immunostained with a GFP-specific antibody do not co-localize with *vglut2b*-expressing cells labeled by fluorescent ISH. Weak co-labeling was occasionally observed in endogenous and ectopic Hcrt neurons (white arrowheads). The approximate regions shown in (A) and (C) are indicated with dashed boxes in (B) and (D), respectively. Scale = 10 μm.

Figure S10. Molecular analysis of *lhx9* morpholino knockdown. (A) Mature *lhx9* mRNA lacks part or all of the second exon after injection with a splice blocking morpholino, reducing the PCR product size by 96 bp or 203 bp, respectively. A 5 bp mismatch control morpholino and the apoptosis suppressing p53 morpholino have no effect on *lhx9* splicing. Gene knockdown persists until at least 72 hpf. Note that a small

amount of correctly spliced *lhx9* is present at all time points, indicating incomplete knockdown. (B) Diagram of the five exons of wild-type *lhx9* and the two variants caused by the *lhx9* morpholino. A cryptic splice site in exon 2 produces an in-frame 32 amino acid deletion that removes most of the LIM1 domain (Band 2). A second variant (Band 3) lacks exon 2, contains an early stop codon and lacks both LIM domains and the DNA-binding homeodomain (HD).

Figure S11. Endogenous *hcrt*- and *pdyn*-expressing neurons are reduced in *lhx9* morphants. *Tg(hcrt:EGFP)* embryos injected with either *lhx9* morpholino (C, D) or *lhx9* mismatch control morpholino (A, B) were fixed at 72 hpf and probed for *vglut1* or *pdyn* expression by fluorescent ISH. No obvious defects in *vglut1* expression were observed in *lhx9* morphants, but the total number of cells with intense, punctate *pdyn* expression was reduced by approximately 40% (E). This result suggests that the morpholino-induced reduction in Hcrt neurons is caused by cell loss, rather than silencing of *hcrt* expression. n indicates number of morphant brains analyzed. ***, $p<0.001$ compared to embryos injected with the control morpholino by Student's t-test. Scale = 10 μm.

Figure S12. The *lhx9* morpholino does not induce apoptosis and can be rescued by *lhx9* overexpression. *Tg(hcrt:mRFP)* embryos were stained with acridine orange at 24 hpf to quantify apoptotic cells (A-D). There was no increase in apoptosis in embryos injected with *lhx9* morpholino (C) compared to embryos injected with the *lhx9* mismatch control morpholino (B) or wild-type embryos (A), suggesting that the reduced number of Hcrt neurons in *lhx9* morphants is not due to apoptosis. Acridine orange labeled fewer

cells in embryos injected with the p53 morpholino (D), presumably because apoptosis that normally occurs during development was suppressed. (E) Mean ± s.e.m. number of apoptotic cells in the white boxed region. At least 4 embryo brains were analyzed for each condition. **, $p<0.01$ compared to WT by one-way ANOVA followed by Bonferroni's correction for multiple comparisons. (F) To rescue the morpholino phenotype, *Tg(hcrt:EGFP)* embryos were co-injected with the *lhx9* morpholino and the *hs:lhx9* plasmid. Following heat shock at 24 hpf, *lhx9*-overexpressing cells located in the endogenous *hcrt* expression domain also expressed *hcrt* (white arrowheads), indicating that *lhx9* overexpression can rescue the *lhx9* morpholino phenotype. (G) Mean ± s.e.m. number of endogenous and rescued *hcrt* cells per brain. Following rescue, the total number of *hcrt*-expressing cells is similar to the number observed in wild-type embryo brains (see Fig. 2G). *hcrt* expression in rescued cells was weaker than in endogenous Hcrt neurons, presumably because the rescued cells only received a pulse of *lhx9* while endogenous Hcrt neurons continuously express *lhx9*. Three embryo brains with ectopic *lhx9* expression in the endogenous Hcrt region were quantified. Anterior views of 26 hpf embryos are shown. Scale = 10 μm.

Figure S13. CRISPR/Cas9 targeting of *lhx9* affects Hcrt and QRFP neuron specification. (A) Embryos injected with Cas9+*lhx9* sgRNA subset 1 (sgRNAs 1, 3, 5, 7, 9) or Cas9+*lhx9* sgRNA subset 2 (sgRNAs 2, 4, 6, 8, 10) (supplementary material Table S2) show a 5.1-fold and 3.7-fold reduction, respectively, in the number of Hcrt neurons per brain hemisphere compared to embryos injected with Cas9 alone. (C) Embryos injected with Cas9 and all 10 *lhx9* sgRNAs show a 2.7-fold reduction in the number of

QRFP neurons per brain hemisphere. This reduction was significant, but less dramatic than the reduction observed for Hcrt neurons (Fig. 6F). (E) Embryos injected with Cas9+*lhx9* sgRNA subset 1 or Cas9+*lhx9* sgRNA subset 2 have significantly fewer QRFP neurons in each brain hemisphere. Injection of the two non-overlapping *lhx9* sgRNA subsets each causes the same phenotype as injection of all 10 sgRNAs, suggesting that loss of Hcrt and QRFP cells is due to *lhx9* knockout and not due to off-target effects of particular sgRNAs. Histograms in (B, D, F) show the percentage of brain hemispheres containing the indicated number of Hcrt neurons measured in (A, C, E). ***, $p<0.001$ compared to control embryos injected with Cas9 alone by one-way ANOVA followed by Tukey's correction for multiple comparisons. n indicates number of brain hemispheres analyzed.

Figure S14. Time course of *lhx9* and *hcrt* expression after heat shock.

Double fluorescent ISH for *lhx9* and *hcrt* are shown at 1 hour, 8 hours, and 24 hours after heat shock (HS)-induced *lhx9* overexpression. Fluorescent ISH for *lhx9* and immunostained *hcrt:EGFP* is shown at 96 hours post HS. (B) At 1 hour post HS, widespread *lhx9* mRNA is detected throughout the embryo. Nearly all *lhx9*-expressing cells also express *hcrt*. (D) At 8 hours post HS, the number of cells with *lhx9* overexpression and ectopic *hcrt* expression is reduced by approximately four-fold. Most *lhx9*-expressing cells still express *hcrt*. (F) At 24 hours post HS, little ectopic *lhx9* or *hcrt* expression is observed, except for ectopic *lhx9*- and *hcrt*-co-expressing neurons in the medial hindbrain. Ectopic *hcrt*-expressing neurons that persist until 96 hours post HS (120 hpf) also express ectopic *lhx9* (H, J). In contrast, control embryos injected with

empty heat shock vector only exhibit *hcrt* expression in the hypothalamus (A, C, E, G, I). Arrowheads indicate examples of ectopic *lhx9* and *hcrt* co-expression. (G, H) show 91 μm thick confocal maximum intensity projections containing both endogenous and ectopic Hcrt neurons. (I, J) show 45 μm thick confocal maximum intensity projections including only the region containing ectopic Hcrt neurons. *lhx9*-expressing neurons that appear close to ectopic Hcrt neurons in (H) are located 30 μm ventral to the ectopic Hcrt neurons, and are thus not observed in (J). Embryos fixed at 1 hour and 8 hours post HS are shown in side view. Embryos fixed at 24 hours and 96 hours post HS are shown dorsally and ventrally, respectively. Boxed regions in (G, H) are shown at higher magnification in (I, J). Scale indicates 50 μm (A-F) and 10 μm (G-J).

Figure S15. Quantification of Hcrt neuron specification following *in utero* overexpression of *lhx9*. E10.5 mouse embryos were electroporated with CAG-EYFP (A) or CAG-Lhx9 + CAG-EYFP (B) *in utero* into the right brain hemisphere lateral hypothalamus and analyzed at P6 for *hcrt* expression by ISH in 6 adjacent coronal sections. Sections were co-stained with DAPI to visualize individual cells to facilitate accurate quantification of *hcrt* ISH (DAPI co-stain is shown for all three brains in (A) and for two brains in (B)). White dots in (A) are included to show how quantification was performed. White numbers indicate the number of *hcrt*-expressing neurons in a brain hemisphere in each section. (C) Mean ± s.e.m. number of *hcrt*-expressing neurons for three experimental and three control brain hemispheres in each section. *, $p<0.05$ compared to control by Student's t-test. These data were combined to generate the graph in Fig. 8F.

Figure S16. Hypothalamic *lhx9* overexpression coincides with increased *hcrt*-expressing neurons. Six serial coronal sections of exemplar brains electroporated with CAG-EYFP or CAG-Lhx9 + CAG-EYFP in the right brain hemisphere lateral hypothalamus at E10.5 and processed for *hcrt* (A) or *lhx9* (B) ISH at P6 are shown. *hcrt* ISH sections were co-stained with DAPI to facilitate quantification of *hcrt*-expressing neurons. The majority of ectopic *lhx9*-expressing neurons are observed in sections 4 and 5 (B, compare boxed region in right column to left column), which coincides with sections that show the greatest increase in *hcrt*-expressing neurons compared to CAG-EYFP controls (A, compare boxed regions in right column to left column). See Fig. S12 for quantification of *hcrt*-expressing neurons in each section. The CAG-EYFP and CAG-Lhx9 + CAG-EYFP images shown are the same as brain 2 and brain 1 in Fig. S12, respectively.

Figure S17. *lhx9* overexpression in the zona incerta does not induce Hcrt neuron specification. Mice electroporated with CAG-Lhx9 + CAG-EYFP in the zona incerta (ZI) (A, red arrow), show no ectopic Hcrt neurons in the ZI (B). *hcrt* ISH was allowed to develop longer for these samples than those shown in Figs 8, S12 and S13 to ensure that any faint *hcrt* expression in the ZI could be detected. Dashed box indicates the lateral hypothalamus.

Table S1. Genes identified by microarray as enriched in Hcrt neurons. This table contains three sheets containing processed microarray data used to identify Hcrt enriched

genes. Sheet 1: Alphabetized list of the genes enriched in Hcrt neurons in multiple pair-wise comparisons. Sheet 2: Summary of data used to generate the gene list in Sheet 1, including probe name, the number of pair-wise comparisons for which a probe reached significance, the significance rank for a probe in each comparison, and detailed information on the gene represented by a probe. Sheet 3: R analysis used to obtain Sheet 2. Each probe was independently ranked by fold change and Bayesian significance and the combined P value was computed and adjusted. Raw data used to generate the values in Sheet 3 are available upon request. Note that the microarray did not contain probes for QRFP.

Table S2. sgRNAs used to target *lhx9* using the CRISPR/Cas9 system. This table lists the sequences of the 10 sgRNAs used to target *lhx9* using CRISPR/Cas9.

Supplementary References

Liu, F. and Patient, R. (2008). Genome-wide analysis of the zebrafish ETS family identifies three genes required for hemangioblast differentiation or angiogenesis. *Circ. Res.* **103**, 1147–1154.

Thisse, B., Heyer, V., Lux, A., Alunni, V., Degrave, A., Seiliez, I., Kirchner, J., Parkhill, J.-P. and Thisse, C. (2004). Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. *Methods Cell Biol.* **77**, 505–519.

Figure S1

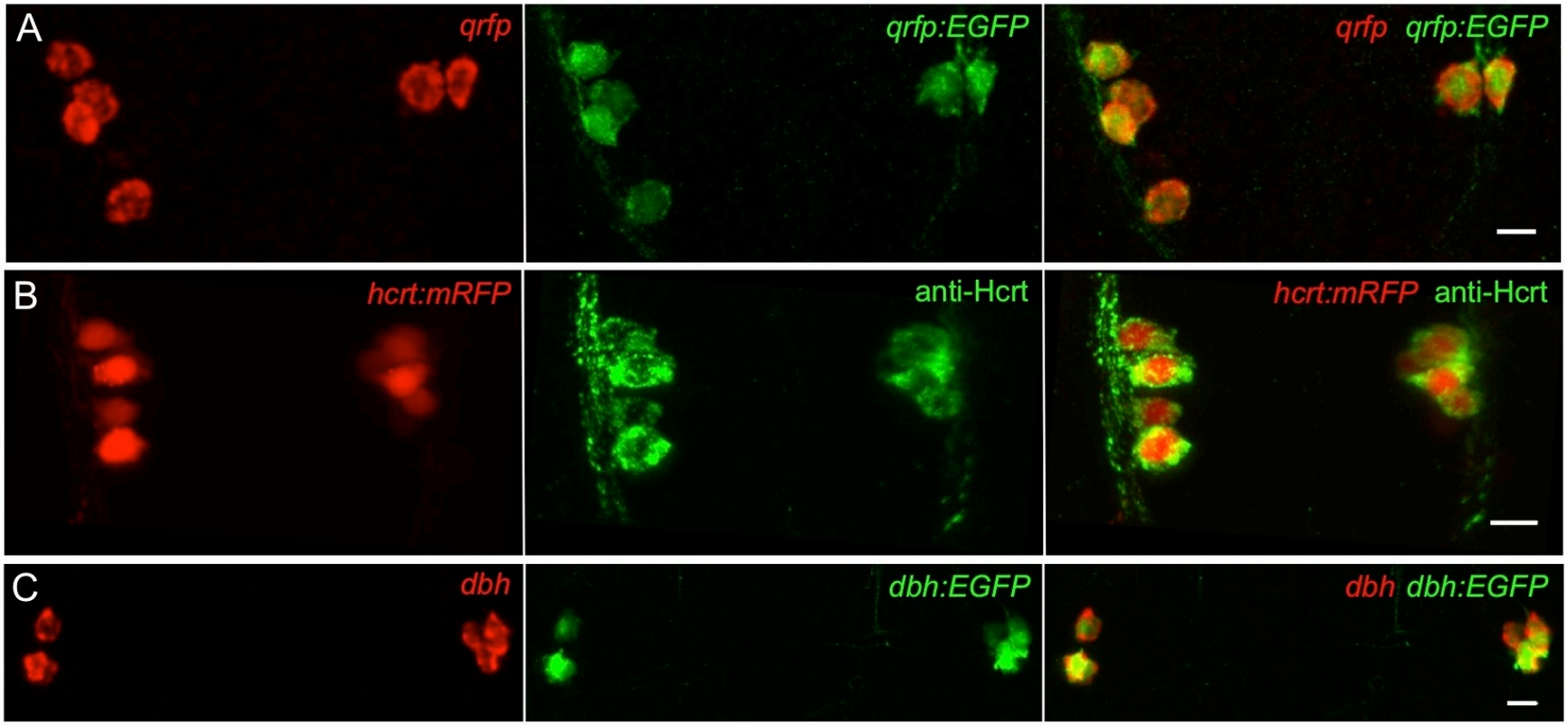

A
qrfp
qrfp:EGFP
qrfp qrfp:EGFP
B
hcrt:mRFP
anti-Hcrt
hcrt:mRFP anti-Hcrt
C
dbh
dbh:EGFP
dbh dbh:EGFP

Figure S2

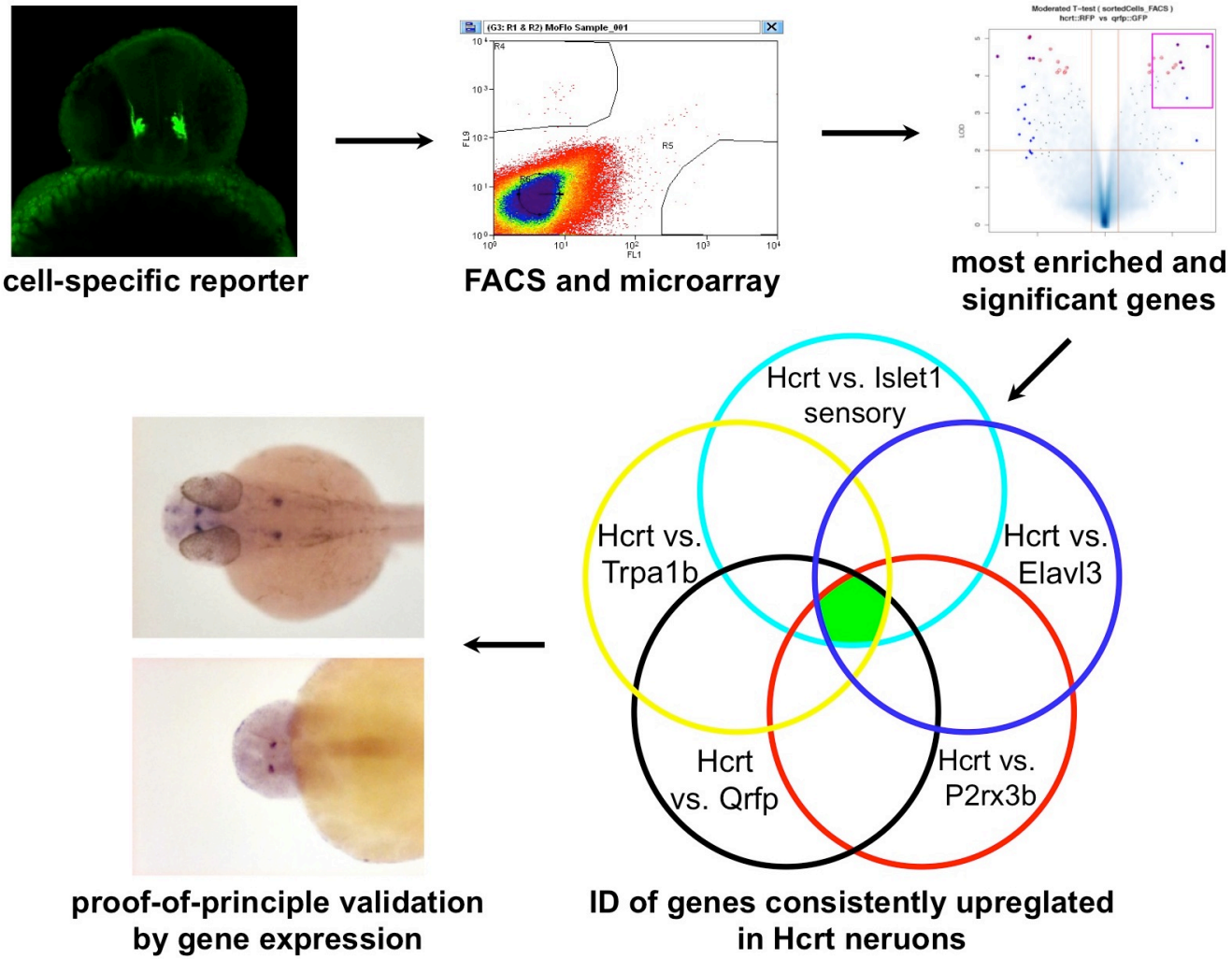

cell-specific reporter
FACS and microarray
most enriched and
significant genes
Hcrt vs. Islet1
sensory
Hcrt vs.
Trpa1b
Hcrt vs.
Elavl3
Hcrt
vs. Qrfp
Hcrt vs.
P2rx3b
proof-of-principle validation
by gene expression
ID of genes consistently upreglated
in Hcrt neruons

Figure S3

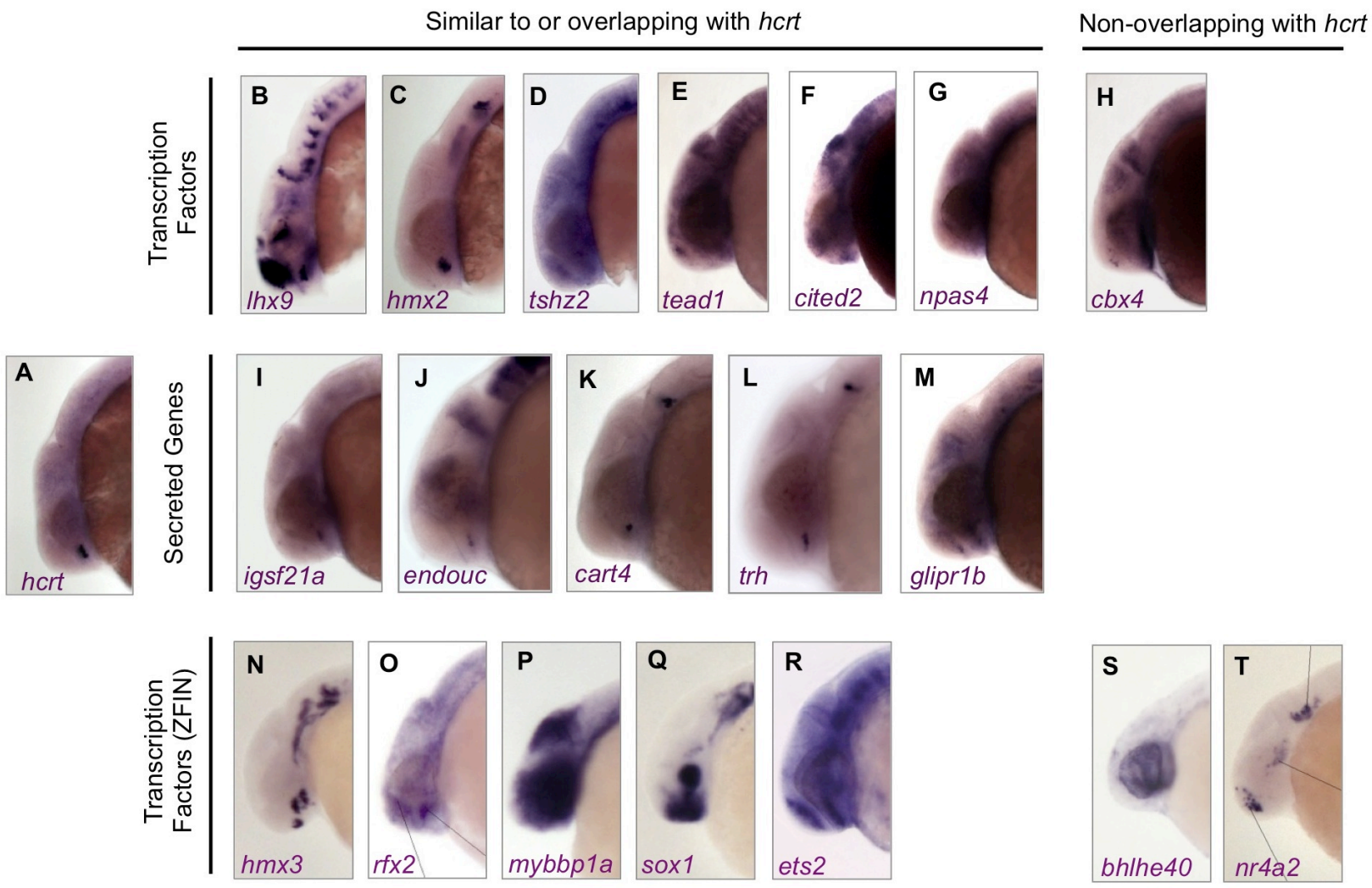

Similar to or overlapping with *hcrt*
Non-overlapping with *hcrt*
Transcription Factors
Secreted Genes
Transcription Factors (ZFIN)
A
hcrt
B
lhx9
C
hmx2
D
tshz2
E
tead1
F
cited2
G
npas4
H
cbx4
I
igsf21a
J
endouc
K
cart4
L
trh
M
glipr1b
N
hmx3
O
rfx2
P
mybbp1a
Q
sox1
R
ets2
S
bhlhe40
T
nr4a2

Figure S4

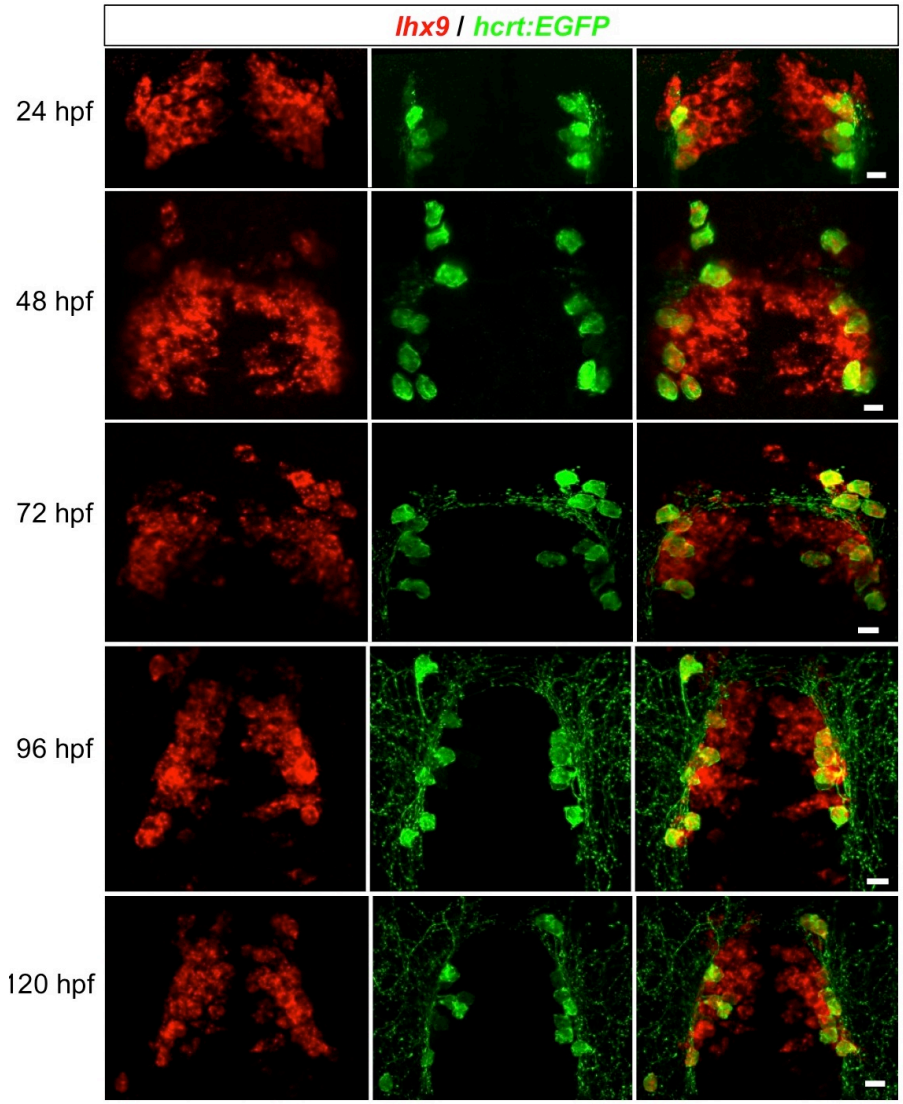


Figure S5

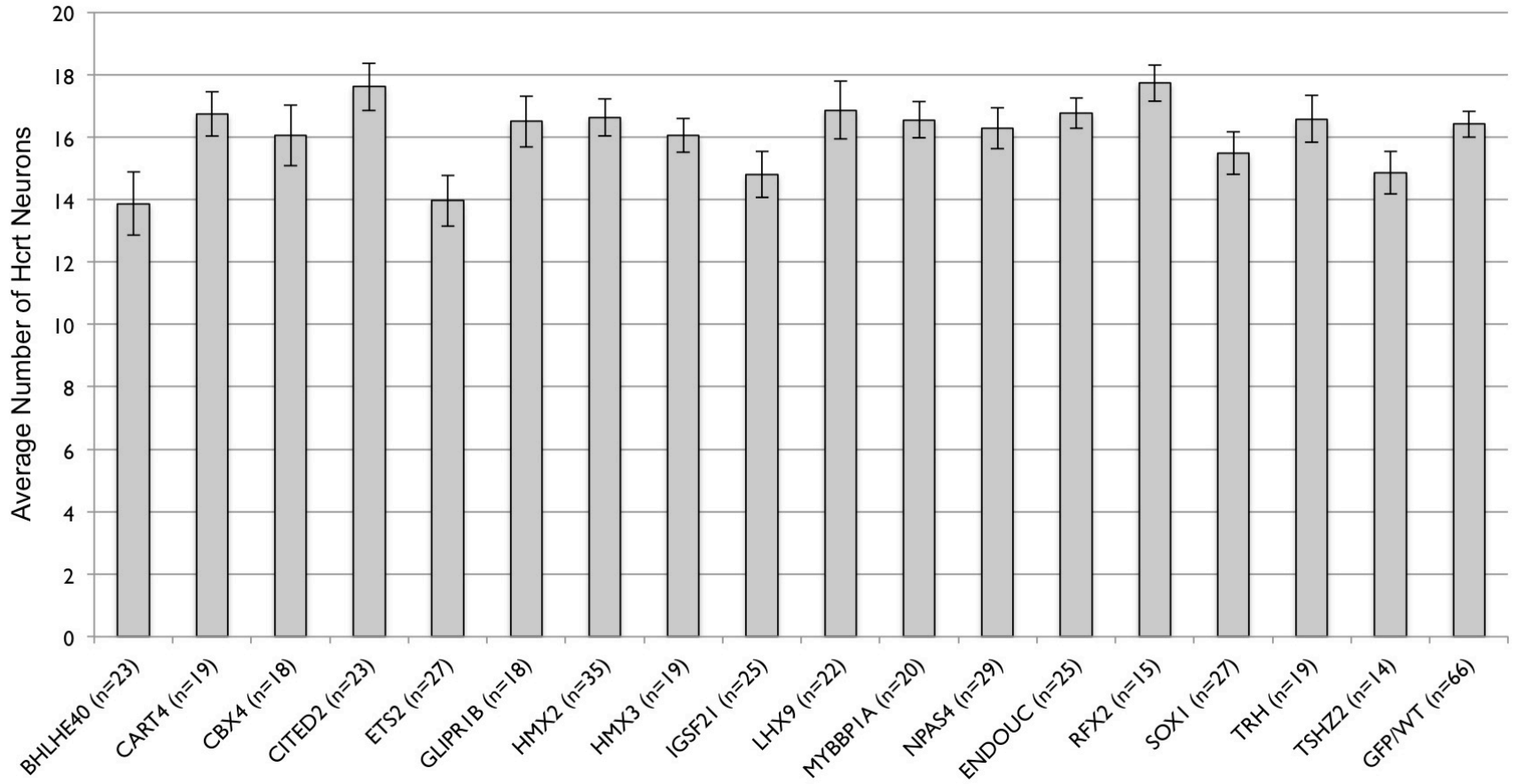

Average Number of Hcrt Neurons
0
2
4
6
8
10
12
14
16
18
20
BHLHE40 (n=23)
CART4 (n=19)
CBX4 (n=18)
CITED2 (n=23)
ETS2 (n=27)
GLIPR1B (n=18)
HMX2 (n=35)
HMX3 (n=19)
IGSF21 (n=25)
LHX9 (n=22)
MYBBP1A (n=20)
NPAS4 (n=29)
ENDOUC (n=25)
RFX2 (n=15)
SOX1 (n=27)
TRH (n=19)
TSHZ2 (n=14)
GFP/WT (n=66)

Figure S6

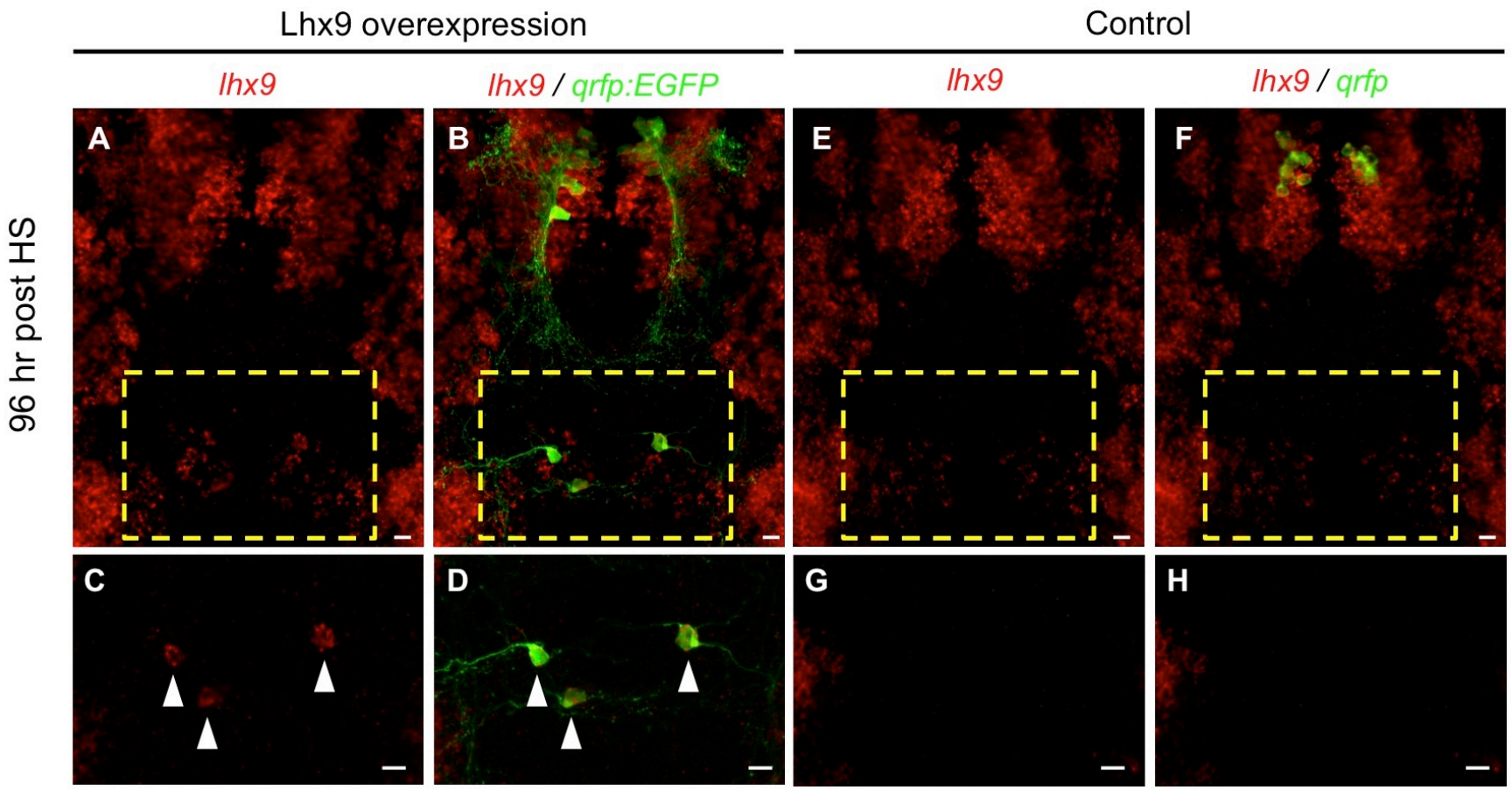

Lhx9 overexpression
Control
lhx9
lhx9 / qrfp:EGFP
lhx9
lhx9 / qrfp
96 hr post HS
A
B
C
D
E
F
G
H

Figure S7

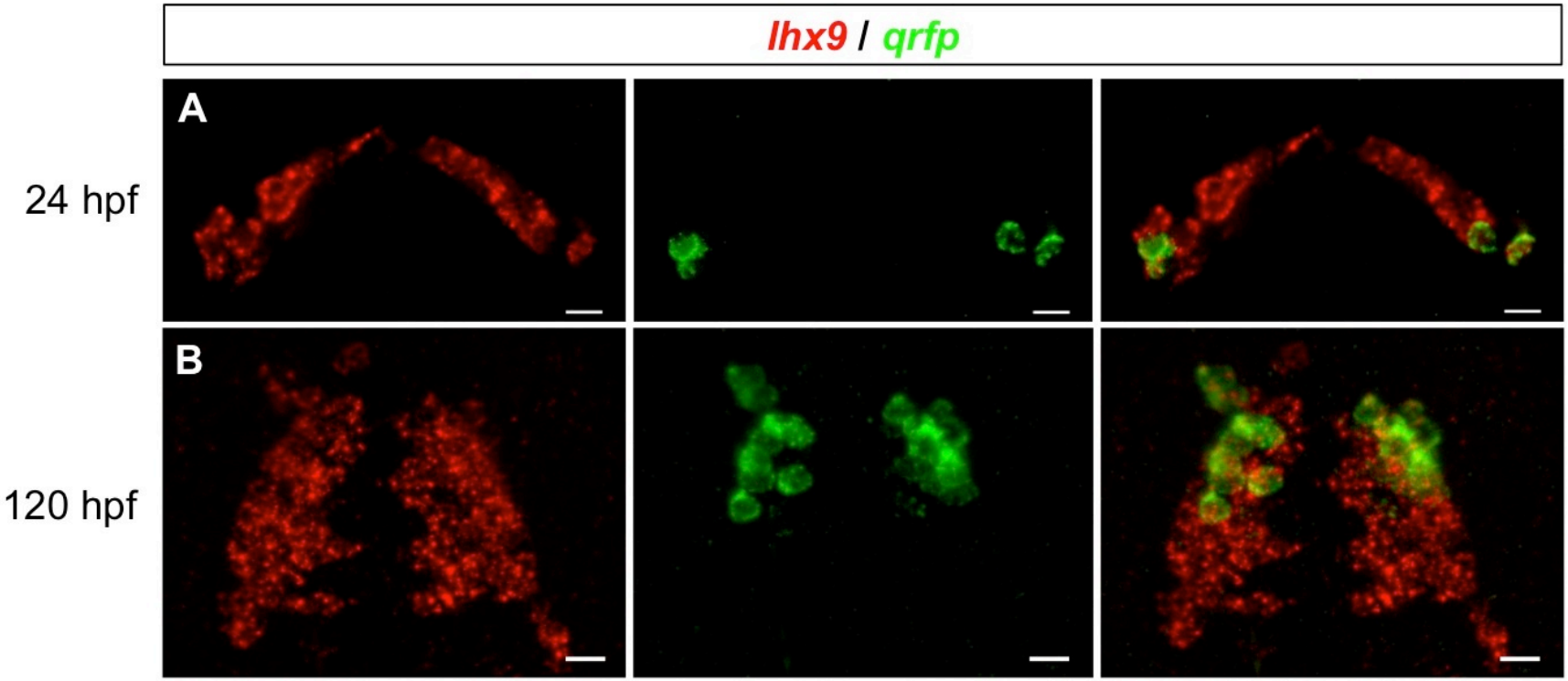


Figure S8

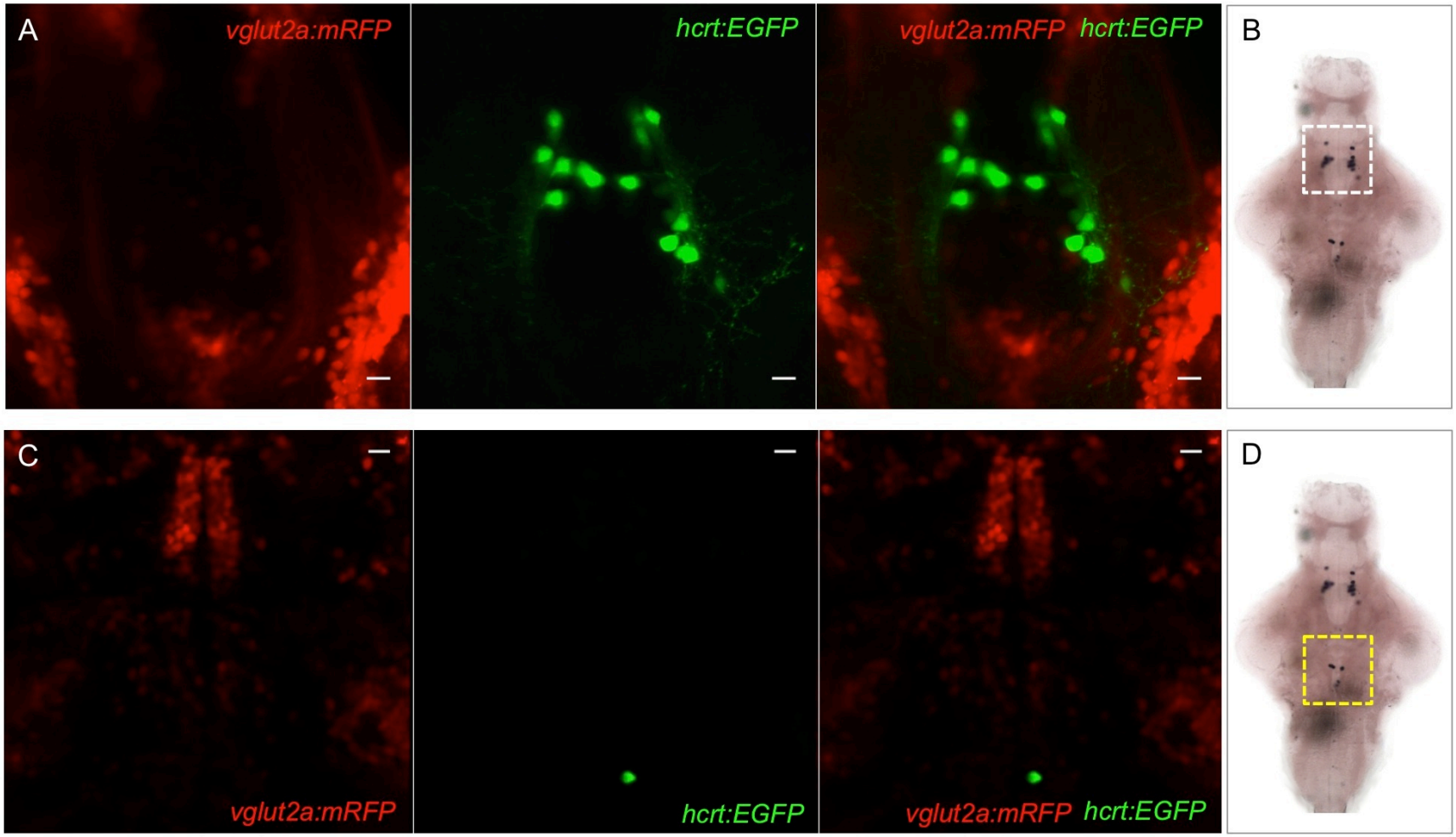

A
vglut2a:mRFP
hcrt:EGFP
vglut2a:mRFP hcrt:EGFP
B
C
vglut2a:mRFP
hcrt:EGFP
vglut2a:mRFP hcrt:EGFP
D

Figure S9

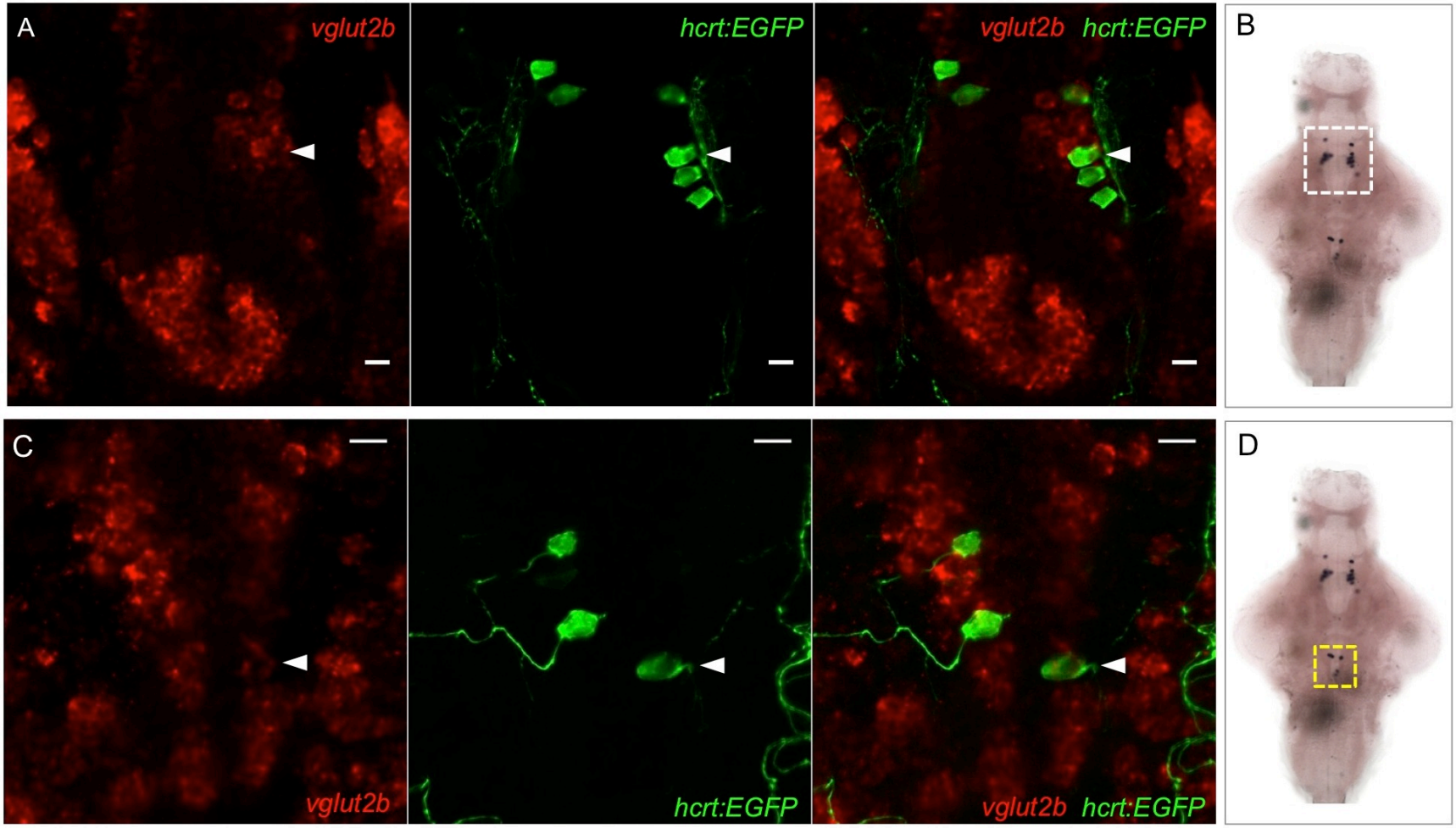

A
vglut2b
hcrt:EGFP
vglut2b hcrt:EGFP
B
C
vglut2b
hcrt:EGFP
vglut2b hcrt:EGFP
D

Figure S10

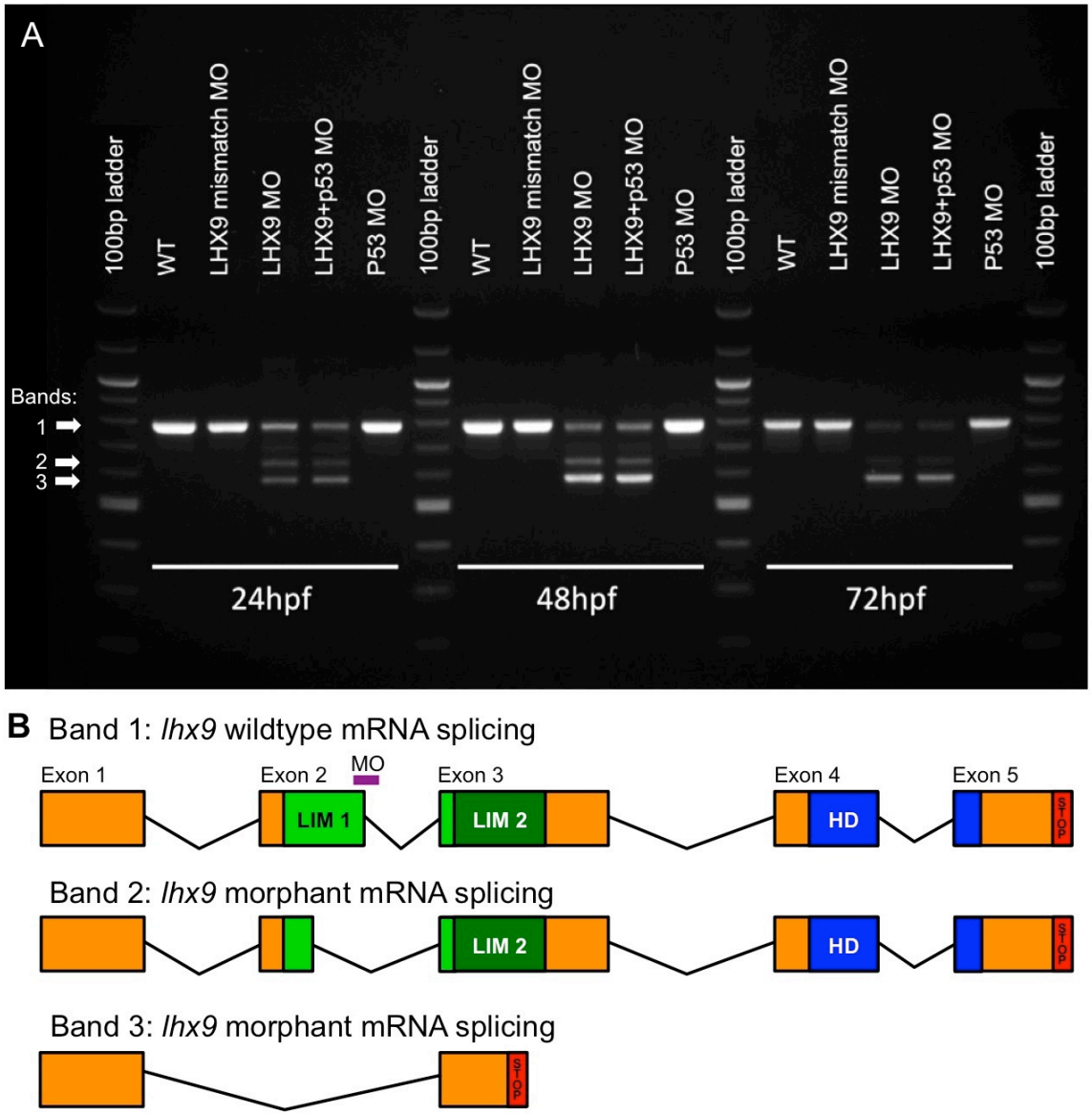

A
100bp ladder
WT
LHX9 mismatch MO
LHX9 MO
LHX9+p53 MO
P53 MO
Bands:
1
2
3
24hpf
48hpf
72hpf
B
Band 1: lhx9 wildtype mRNA splicing
Exon 1
Exon 2
MO
Exon 3
Exon 4
Exon 5
LIM 1
LIM 2
HD
Band 2: lhx9 morphant mRNA splicing
Band 3: lhx9 morphant mRNA splicing

Figure S11

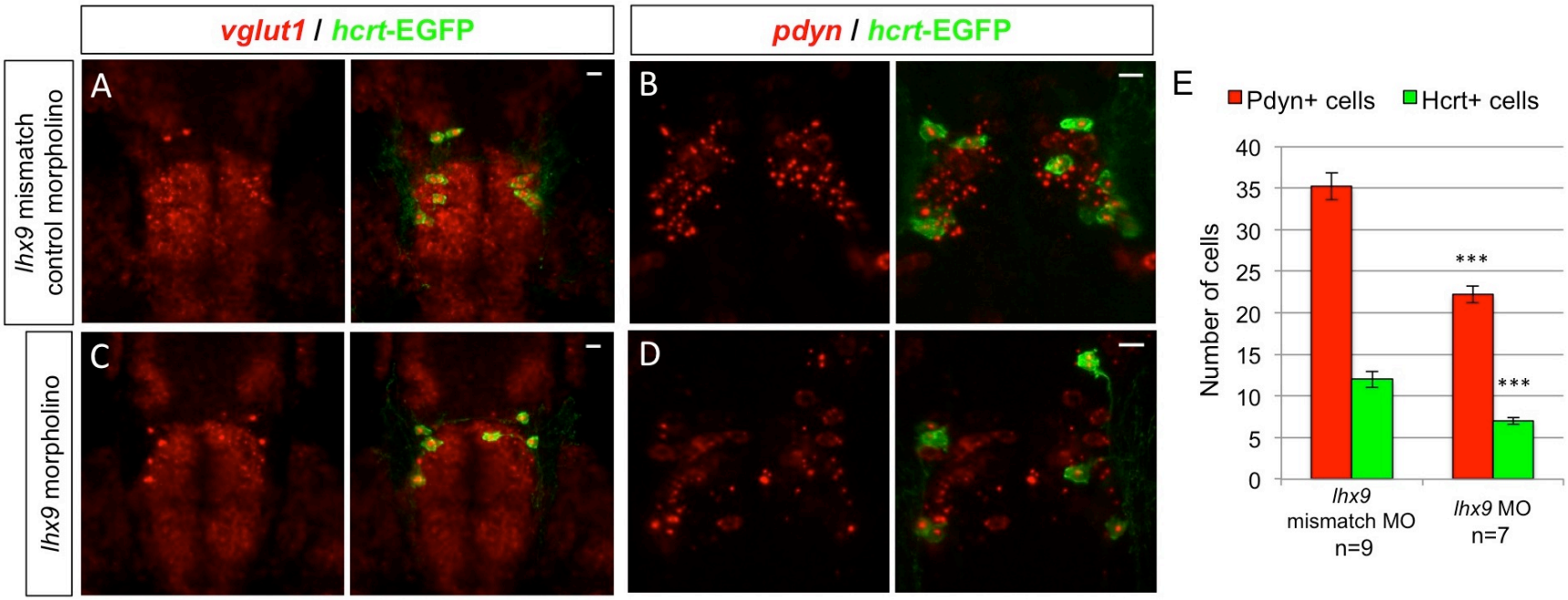

vglut1 / hcrt-EGFP
pdyn / hcrt-EGFP
lhx9 mismatch control morpholino
lhx9 morpholino
A
B
C
D
E
Pdyn+ cells
Hcrt+ cells
Number of cells
40
35
30
25
20
15
10
5
0
***
***
lhx9 mismatch MO n=9
lhx9 MO n=7

Figure S12

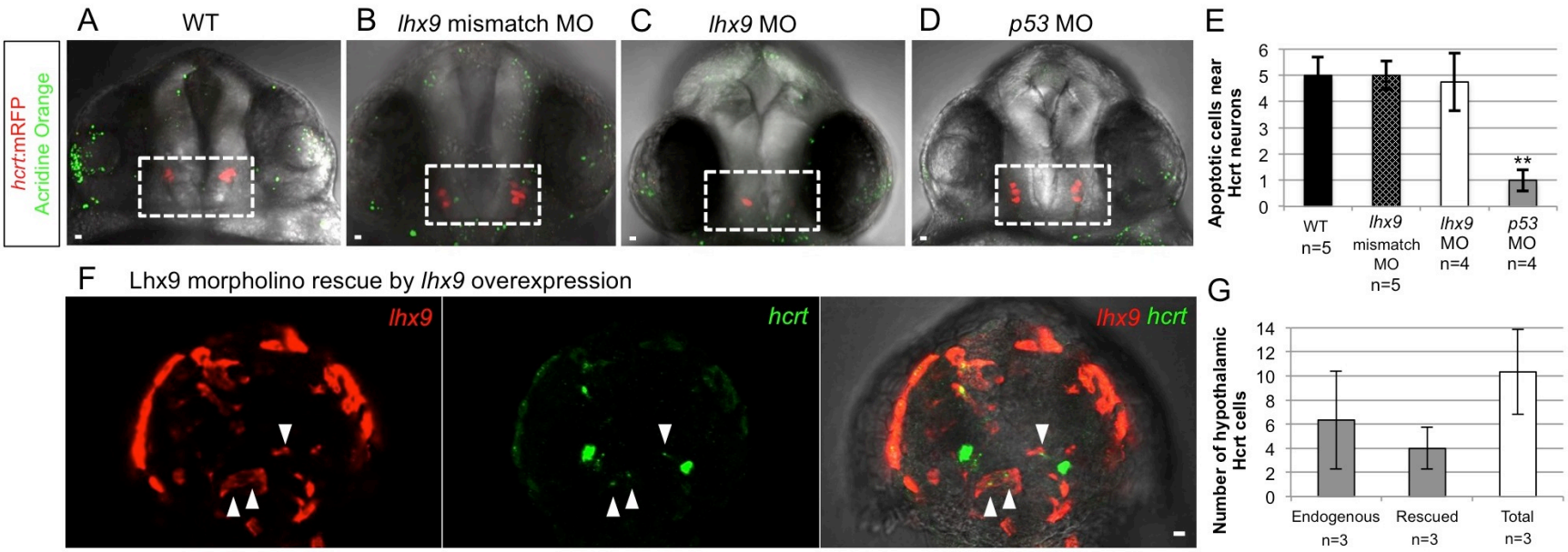

A WT
B lhx9 mismatch MO
C lhx9 MO
D p53 MO
hcrt:mRFP
Acridine Orange
E
Apoptotic cells near Hcrt neurons
WT n=5
lhx9 mismatch MO n=5
lhx9 MO n=4
p53 MO n=4
**
F Lhx9 morpholino rescue by lhx9 overexpression
lhx9
hcrt
lhx9 hcrt
G
Number of hypothalamic Hcrt cells
Endogenous n=3
Rescued n=3
Total n=3

Figure S13

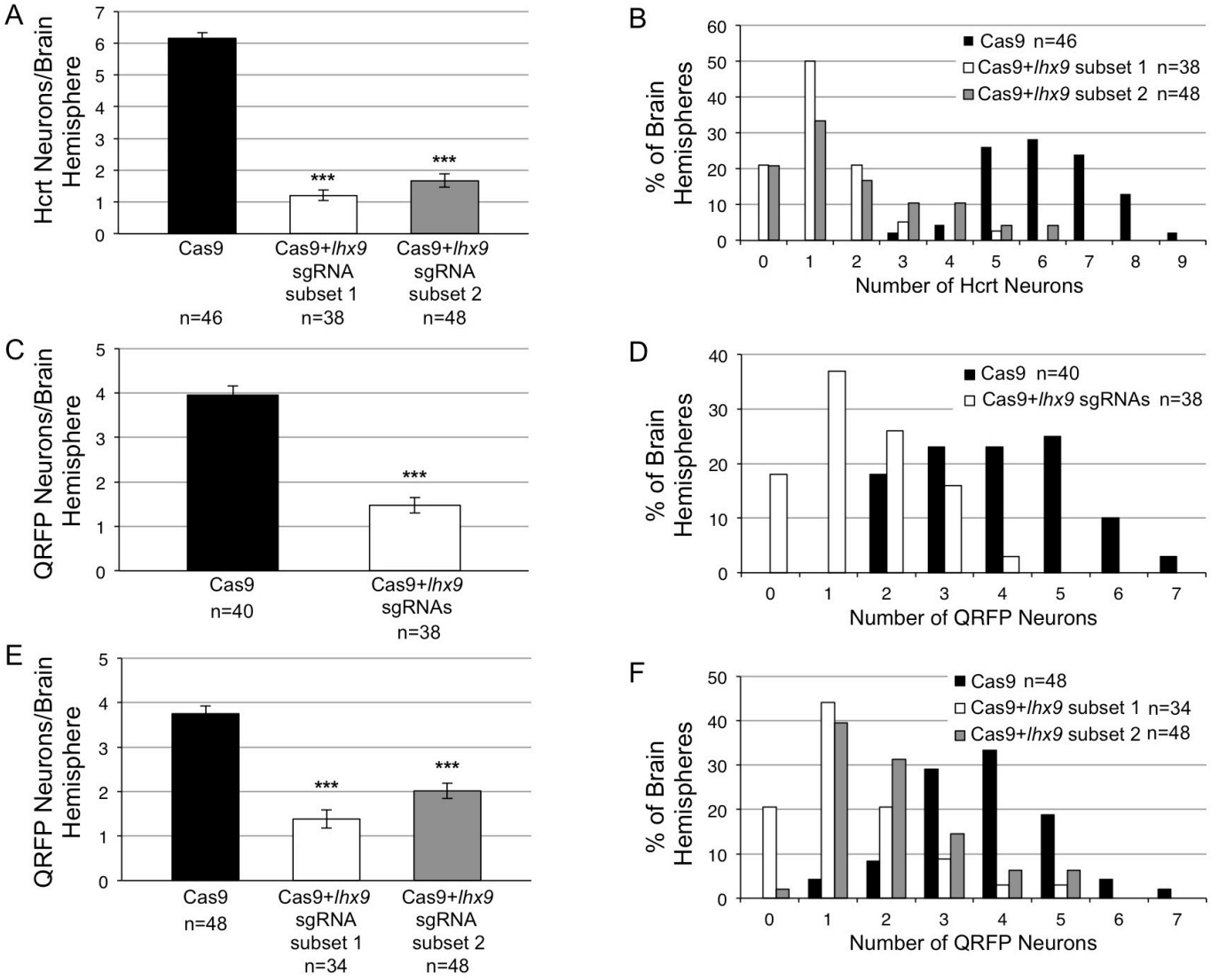

A
Hcrt Neurons/Brain Hemisphere
***
***
Cas9
n=46
Cas9+lhx9 sgRNA subset 1
n=38
Cas9+lhx9 sgRNA subset 2
n=48
B
% of Brain Hemispheres
Cas9 n=46
Cas9+lhx9 subset 1 n=38
Cas9+lhx9 subset 2 n=48
Number of Hcrt Neurons
C
QRFP Neurons/Brain Hemisphere
***
Cas9
n=40
Cas9+lhx9 sgRNAs
n=38
D
% of Brain Hemispheres
Cas9 n=40
Cas9+lhx9 sgRNAs n=38
Number of QRFP Neurons
E
QRFP Neurons/Brain Hemisphere
***
***
Cas9
n=48
Cas9+lhx9 sgRNA subset 1
n=34
Cas9+lhx9 sgRNA subset 2
n=48
F
% of Brain Hemispheres
Cas9 n=48
Cas9+lhx9 subset 1 n=34
Cas9+lhx9 subset 2 n=48
Number of QRFP Neurons

Figure S14

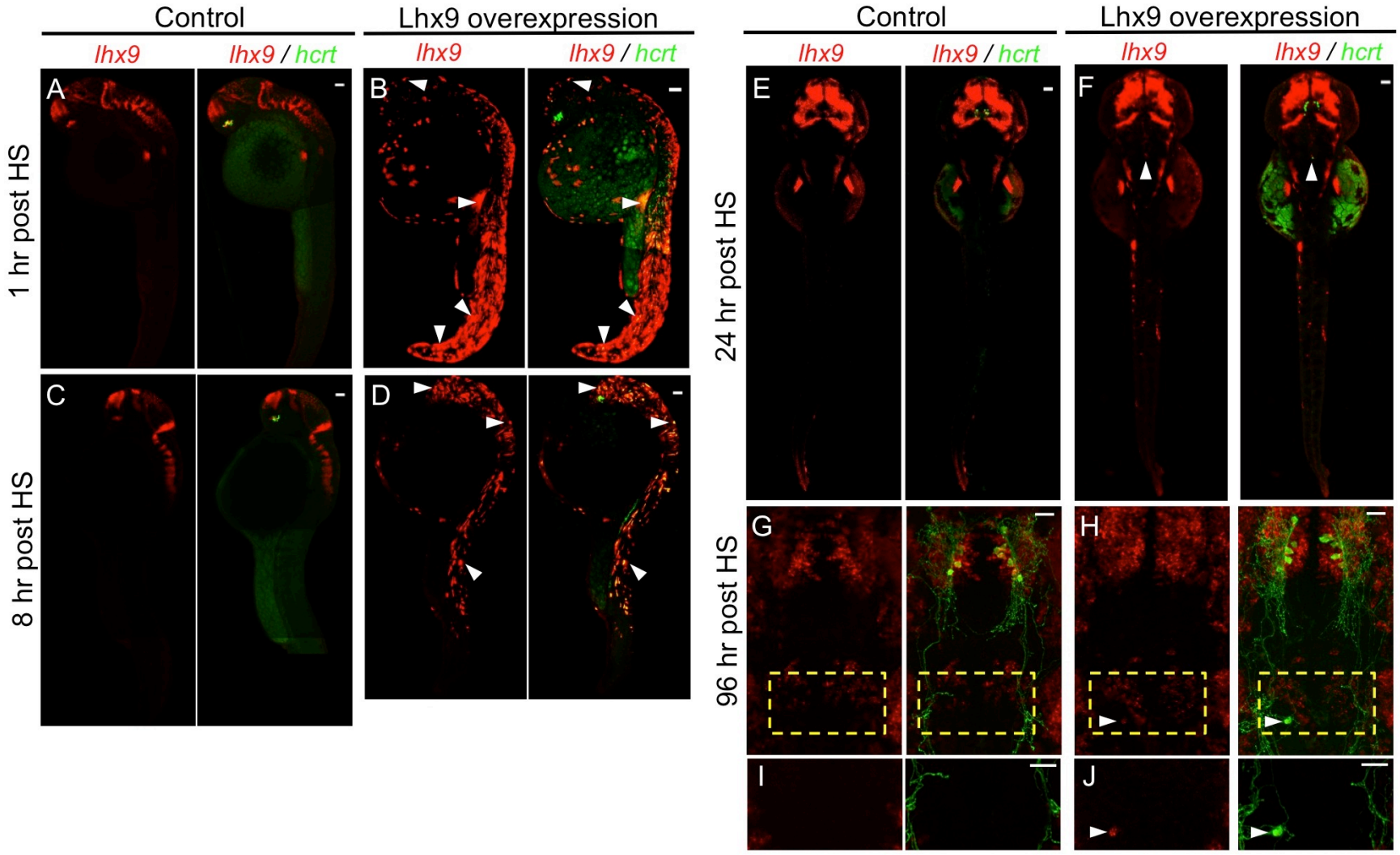

Control
Lhx9 overexpression
lhx9
lhx9 / hcrt
1 hr post HS
8 hr post HS
24 hr post HS
96 hr post HS
A
B
C
D
E
F
G
H
I
J

Figure S15

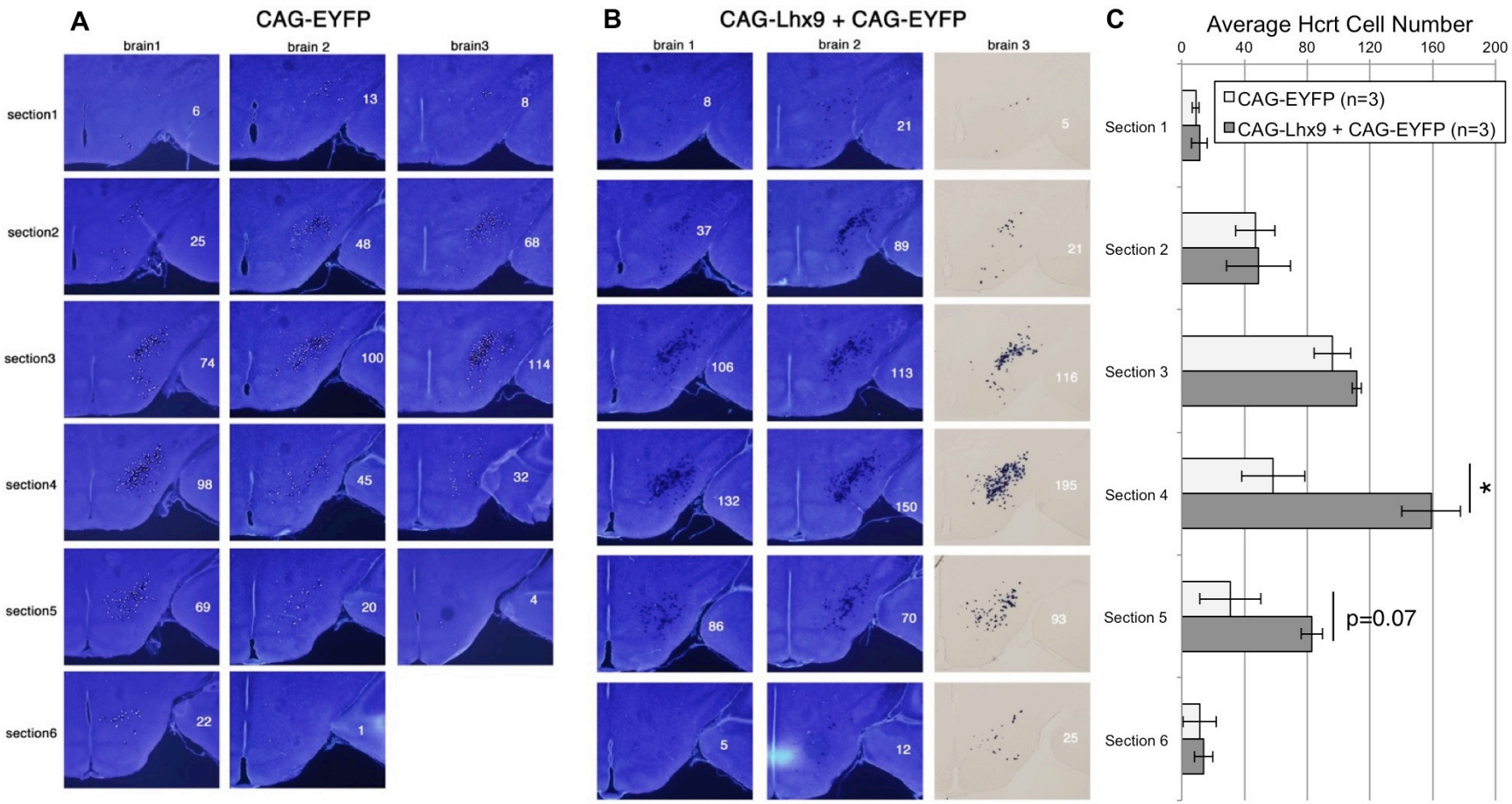

A
CAG-EYFP
brain1
brain 2
brain3
section1
section2
section3
section4
section5
section6
6
13
8
25
48
68
74
100
114
98
45
32
69
20
4
22
1
B
CAG-Lhx9 + CAG-EYFP
brain 1
brain 2
brain 3
8
21
5
37
89
21
106
113
116
132
150
195
86
70
93
5
12
25
C
Average Hcrt Cell Number
0
40
80
120
160
200
CAG-EYFP (n=3)
CAG-Lhx9 + CAG-EYFP (n=3)
Section 1
Section 2
Section 3
Section 4
Section 5
Section 6
*
p=0.07

Figure S16

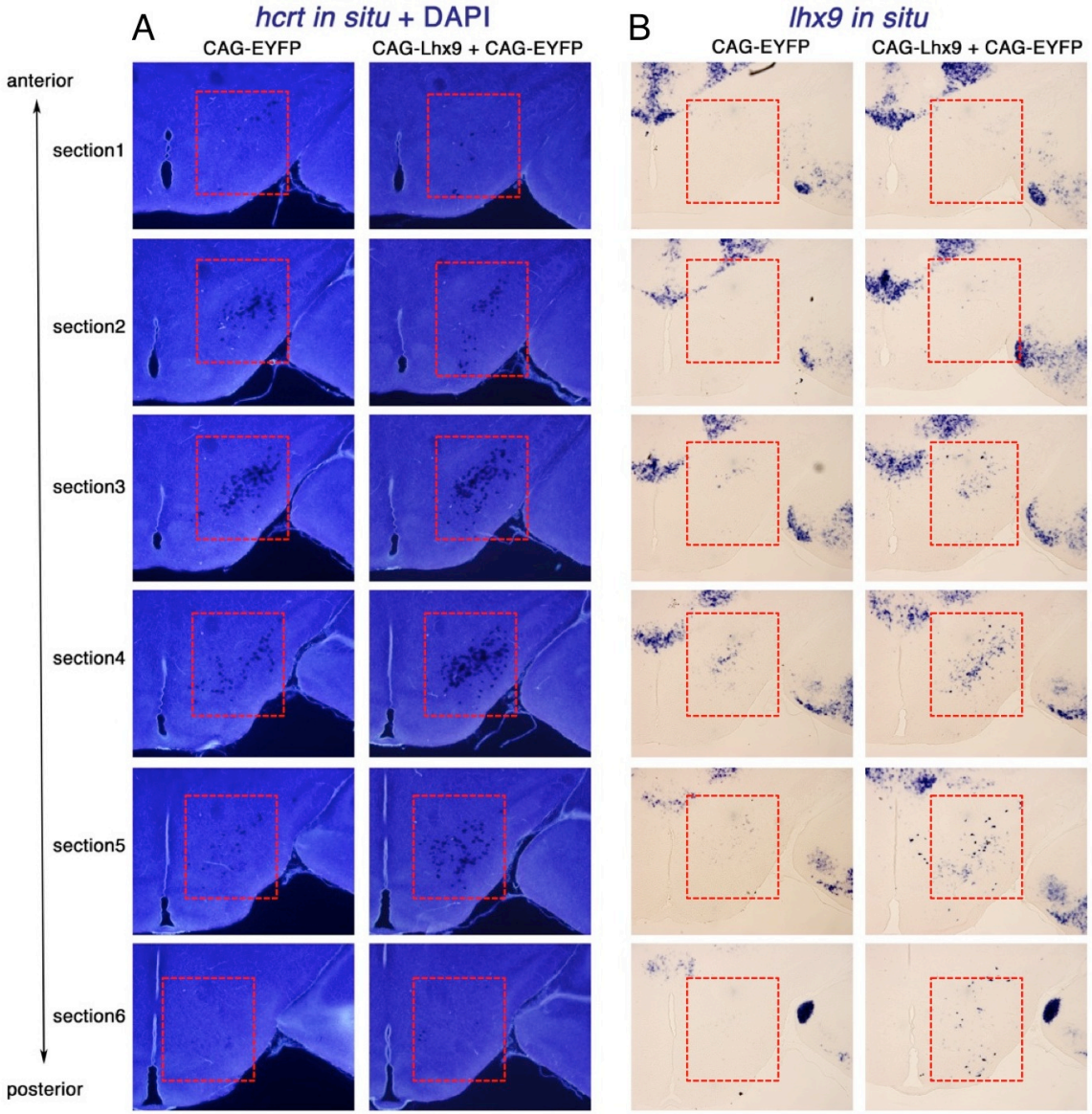
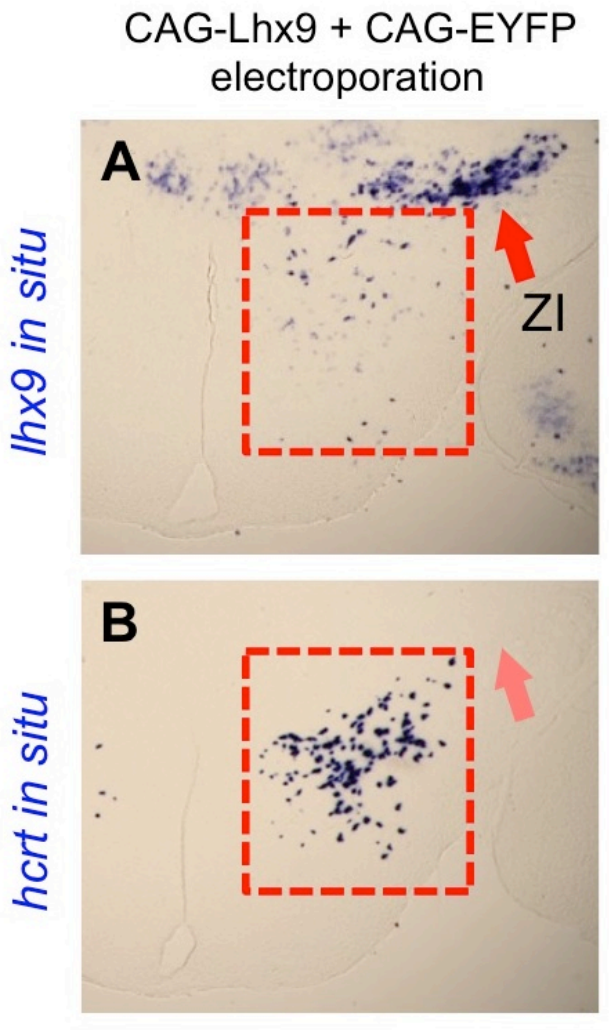

A
hcrt in situ + DAPI
CAG-EYFP
CAG-Lhx9 + CAG-EYFP
B
lhx9 in situ
CAG-EYFP
CAG-Lhx9 + CAG-EYFP
anterior
section1
section2
section3
section4
section5
section6
posterior

Figure S17


CAG-Lhx9 + CAG-EYFP
electroporation
A
Zl
lhx9 in situ
B
hcrt in situ

Movie 1. Colocalization of *hcrt-* and *lhx9*-expressing cells in the embryonic zebrafish hypothalamus. At 24 hpf, *lhx9* is expressed in the hypothalamus in bilateral clusters of 20-30 cells that overlap with *hcrt*-EGFP expressing cells. Note that all *hcrt*-EGFP cells also express *lhx9*. This movie shows the complete confocal image stack from which a single optical section is shown in Fig. 2F. Scale = 10 μm.

Movie 2. All hypothalamic Hcrt cells express *vglut1*. A 120 hpf *hcrt*-EGFP larva probed by fluorescent ISH for *vglut1* shows widespread expression of *vglut1* in the hypothalamus. After immunostaining with anti-GFP antibody, the sample was counterstained with DAPI to label cell nuclei. All Hcrt cells express *vglut1*, with particularly intense nuclear mRNA levels compared to adjacent neurons. This movie shows a confocal image stack advancing from the dorsal to the ventral hypothalamus. Scale = 10 μm.

Movie 3. All hypothalamic Hcrt cells express *pdyn*. A 120 hpf *hcrt*-EGFP larva probed by fluorescent ISH for *pdyn* shows that all Hcrt cells express *pdyn*, although *pdyn* is also expressed elsewhere in the hypothalamus. After immunostaining with anti-GFP antibody, the sample was counterstained with DAPI to label cell nuclei. Note that *pdyn* expression in Hcrt cells is observed as intense nuclear puncta that occupy a smaller area than the nucleus, potentially indicating a localized region of *pdyn* mRNA. This movie shows a confocal image stack advancing from the dorsal to the ventral hypothalamus. Scale = 10 μm.

Movie 4. Ectopic *lhx9*-expressing cells also express *hcrt* at one hour after heat shock. Embryos were injected with a HS-Lhx9 plasmid, fixed one hour after heat shock at 24 hpf, and analyzed using double fluorescent ISH with *hcrt*- and *lhx9*-specific probes. The intense green fluorescent cells are endogenous *hcrt*-expressing neurons. Endogenous *lhx9*-expressing neurons are not visible because this would require a higher laser power that would overexpose the fluorescence of ectopic *lhx9*-expressing cells. Anterior is to the left. The bright field (BF) overlay shows the position of the eye. This movie shows the complete confocal image stack from which a single optical section is shown in Fig. 7A.

Table S1

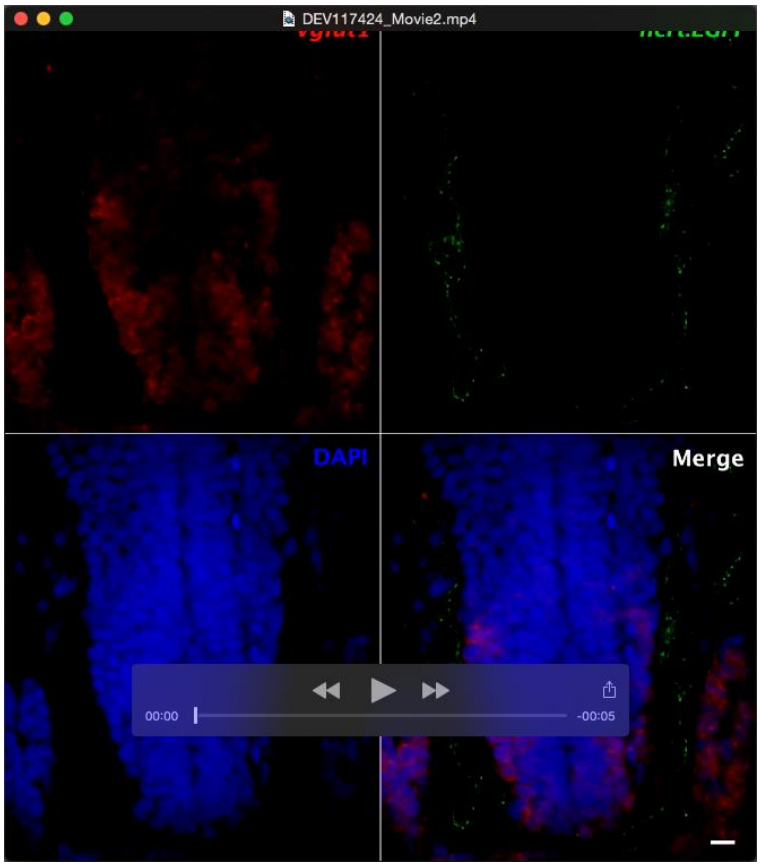
Click here to Download Table S1

Table S2

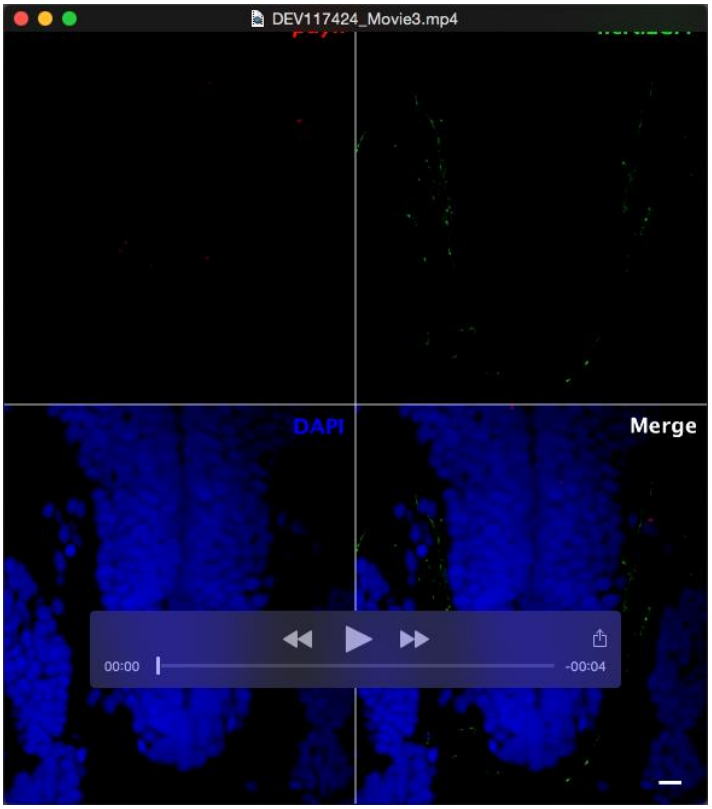
Click here to Download Table S2

Movie 1

Movie 2



Movie 3



Movie 4