SUPPLEMENTARY MATERIAL

Fig. S1. Whole mount ISH, TUNEL staining and cell proliferation analysis of *Hesx1*$^{-/-}$ mutants. (A) ISH with *Atx*. The rostral limit of *Atx* expression domain appears anteriorised in the *Hesx1*$^{-/-}$ mutant (right) when compared with the wild-type embryo (left) (arrowheads) (B) ISH with *Pax6*. Expression of *Pax6* in the dorsal region of the diencephalon and pretectum (bracket) appear expanded in the *Hesx1*$^{-/-}$ mutant (right) when compared with the wild-type littermate (left). (C-J) TUNEL staining of wild-type embryo (C,E,G,I) and *Hesx1*$^{-/-}$ mutant (D,F,H,J). Cell death around the forebrain-midbrain boundary (arrowheads) is enhanced in the *Hesx1*$^{-/-}$ mutant (D) when compared with the wild-type littermate (C). (E,F) Frontal view of the embryos depicted in C,D, showing enhanced apoptosis in the forebrain-midbrain boundary and decreased cell death in the roof plate of the telencephalon (arrowhead). Transverse section of wild-type embryo (G) and *Hesx1*$^{-/-}$ mutant (H) showing that cell death occurs within the neuroepithelium. White line in C indicates the plane of section. Cell death within the anterior neural plate is undistinguishable between genotypes at 8.5 dpc (I,J). (K,L) No significant differences are observed in cell proliferation between wild-type (K) and *Hesx1*$^{-/-}$ (L) embryos at early somite stages.

Fig. S2. Genetic fate mapping of *Hesx1* expressing cells in *Hesx1*$^{Cre/+}$;*R26*$^{Cond\text{-}lacZ/Cond\text{-}lacZ}$ embryos. All embryos were X-gal stained. (A) Frontal view of a 7.5 dpc embryo showing staining in the visceral endoderm of the extraembryonic region (arrowhead). Staining is beginning to appear in the anterior foregut (arrows). (B) Dorsal view of a 1-2 somite embryo showing strong *lacZ* expression in the anterior neural plate. (C) Fronto-ventral

view of the embryo depicted in B, showing *lacZ* expression in the foregut pocket (arrowhead) and anterior neural plate (arrow). (D) Lateral view of a 6-7 somite embryo. *LacZ* expression is observed in the anterior neural plate (arrowhead), anterior foregut (black arrow) and in scattered cells of the yolk sac (white arrow). (E) 8-10 somite embryo, lateral view. *LacZ* positive cells are observed in the AFB region. Arrowhead indicates the boundary between anterior and posterior forebrain. (F-H) Lateral (F), frontal (G), and ventral (H) views of the head of a 10.5 dpc embryo. The vast majority of *lacZ* positive cells are located within the AFB: telencephalon (black arrow) and eyes (black arrowhead). Only scattered *lacZ* positive cells are observed in the posterior forebrain: dorsal diencephalon (white arrow) and pretectum (white arrowhead). Note the absence of *lacZ* positive cells in the midbrain (red arrow). Additionally, X-gal staining is observed in the olfactory placodes (white arrowhead in G) and the developing Rathke's pouch (arrowhead in H). (I) *LacZ* expression is also detectable in the liver bud (arrowhead). (J-M) frontal sections of the10.5 dpc embryo depicted in F-I. Code: dd, dorsal diencephalon; dt, dorsal telencephalon; ey, eye; ht, heart; hy, hypothalamus; lb, liver bud; op, olfactory placode; rp, Rathke's pouch; vd, ventral diencephalon; vt, ventral telencephalon.

Fig. S3. Analysis of *Cre* expression in *Hesx1*$^{Cre/Cre}$, *Hesx1*$^{Cre/-}$ and *Hesx1*$^{Cre/+}$ embryos. (A-C) ISH with *Hesx1* (A) and *Cre* (B,C) on *Hesx1*$^{+/+}$ and *Hesx1*$^{Cre/Cre}$ embryos at the 10-12 (A,B) and 6-7 somite stage (C). The *Cre* expression domain appears slightly smaller than the *Hesx1* expression domain, but it is restricted to the ventral forebrain. (D,E) ISH with *Cre* on *Hesx1*$^{Cre/+}$ (D) and *Hesx1*$^{Cre/-}$ (E) embryos. *Cre* expression is restricted to the

ventral forebrain in both genotypes. Staining was detectable only after five days in NBT/BCIP staining solution. (F) Quantitative Real Time PCR of *Hesx1*$^{Cre/+}$ and *Hesx1*$^{Cre/-}$ embryos at the 2, 5 and 8 somite stage. At the 2 somite stage, *Cre* expression levels are similar in both genotypes, but lower in *Hesx1* deficient embryos at the 5 and 8 somite. Error bars represent standard error of the mean. Significant differences are observed only at the 5 somite stage. *Gapdh* was used as internal control. *Cre* expression in the *Hesx1*$^{Cre/+}$ embryo was assigned a value of 1.0 for each stage. At least two embryos were analysed for each stage. The graph shows the summarised results of five repeats.

Fig. S4. X-gal staining on cryosections of *Hesx1*$^{Cre/-}$*;R26*$^{Cond\text{-}lacZ/+}$ (A-D) and *Hesx1*$^{Cre/+}$*;R26*$^{Cond\text{-}lacZ/+}$ (E) embryos at 16.5 dpc. X-gal positive cells colonise the nasal septum (A,B) and the follicle primordium of the vibrissae (C,D) (arrowheads). A,B were photographed at 50x magnification and C,D at 400x. (E) *lacZ* expression in the pituitary is strong in the anterior and intermediate lobes (arrow), but hardly detectable in the posterior lobe (arrowhead). (F) X-gal staining of *Hesx1*$^{Cre/-}$*;R26*$^{Cond\text{-}lacZ/+}$ mutant at the 3 somite stage. No significant differences are observed when compared with *Hesx1*$^{Cre/+}$*;R26*$^{Cond\text{-}lacZ/+}$ embryo (Fig. S2B).

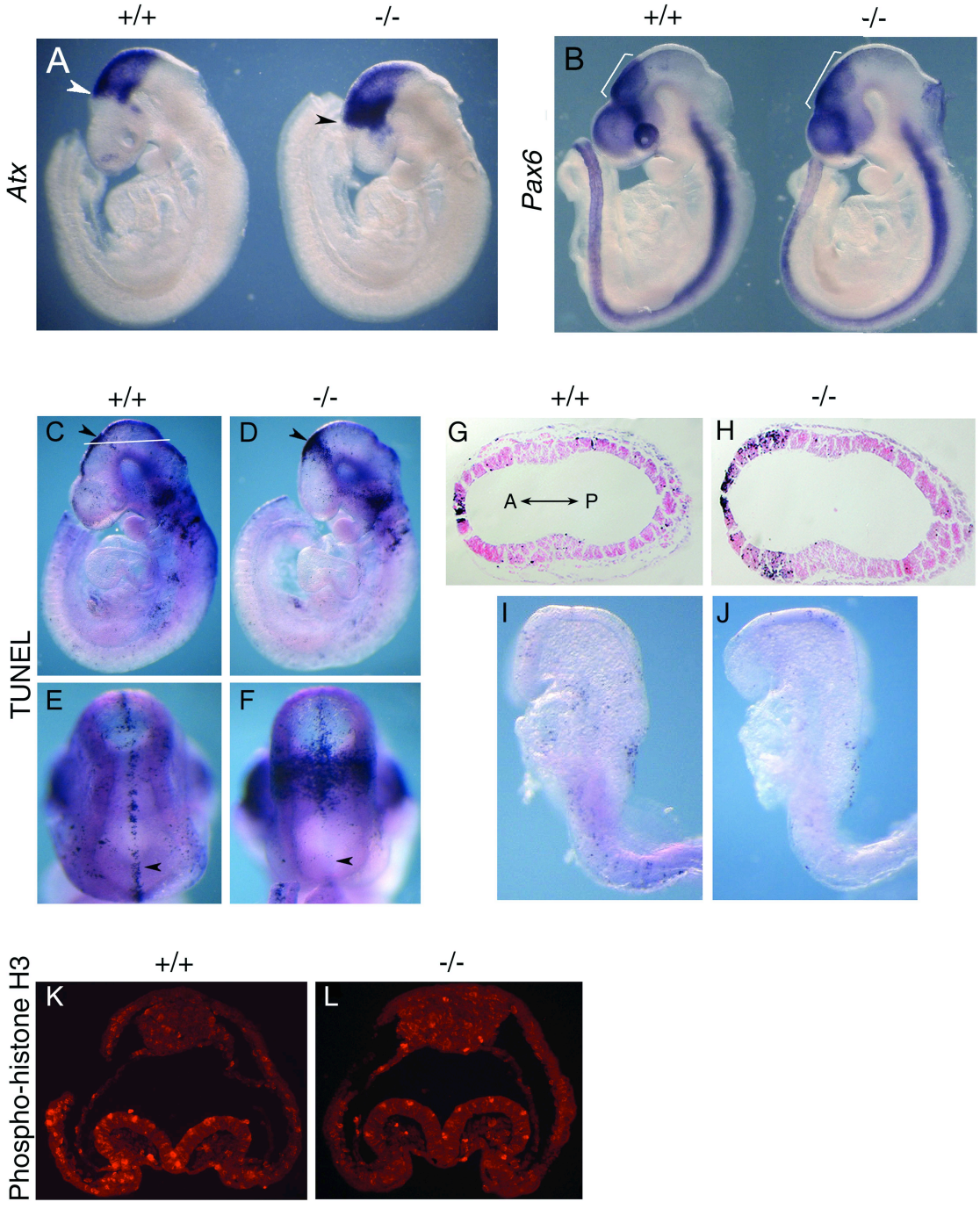
Fig. S5. X-gal positive cells do not colonise the midbrain in *Hesx1*$^{Cre/+;}$*R26*$^{Cond\text{-}lacZ/+}$ or *Hesx1*$^{Cre/-}$*;R26*$^{Cond\text{-}lacZ/+}$ embryos. (A-D) 12.5 dpc brains from *Hesx1*$^{Cre/+}$*;R26*$^{Cond\text{-}lacZ/+}$ (A,C) and *Hesx1*$^{Cre/-}$*;R26*$^{Cond\text{-}lacZ/+}$ (B,D) embryos. Embryos were X-gal stained first, followed by double ISH with *Shh* (purple) and *Pax6* (red) (A.B), or single ISH with *Pax6* (C,D) . Note that many more *lacZ* positive cells (arrow) (light blue) colonise the brain region posterior

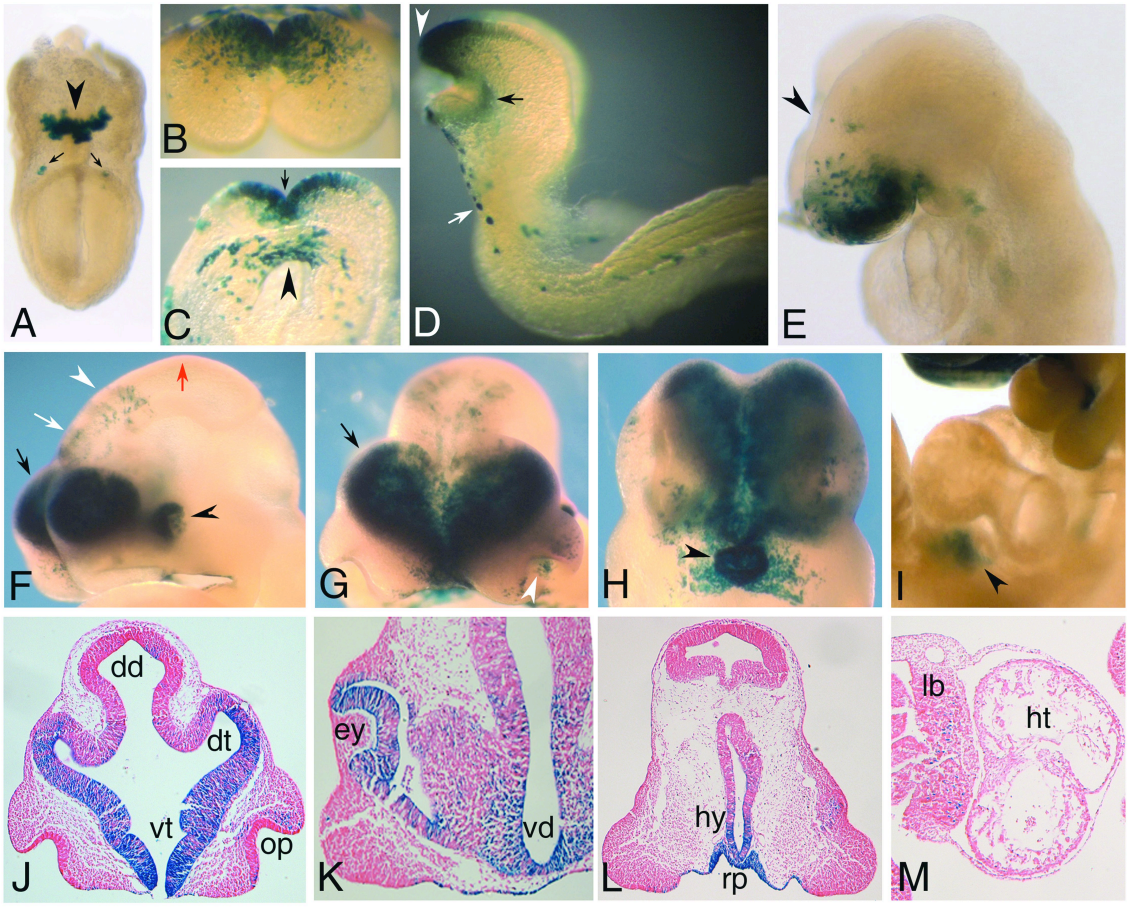
to the ZLI (arrowheads) (A,B). The region of the embryos depicted in C,D is boxed in A,B. Note that *lacZ* positive cells are located within the *Pax6* expression domain in the posterior forebrain and pretectum and do not cross the posterior boundary of the *Pax6* expression domain. *Shh* was stained with NBT/BCIP and *Pax6* was stained either with INT/BCIP (A,B) or NBT/BCIP (C,D).

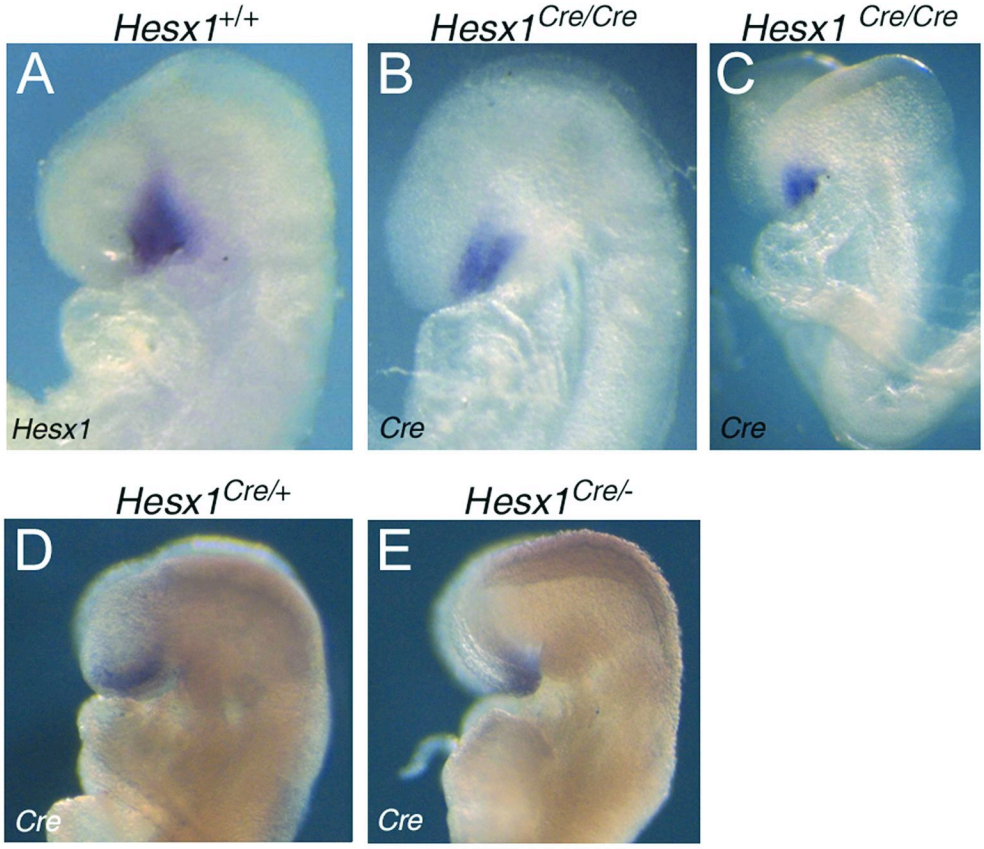
Fig. S6. Generation of the *R26-Cond-Hesx1* targeted allele. (A) Top to bottom: *pBigT-Hesx1*, a plasmid containing a *loxP*-flanked cassette with *PGK-Neo* selectable marker and a *tpA* transcriptional stop sequence, into which the *Hesx1* coding region was cloned; *pR26PA*, containing genomic *R26* sequences for homologous recombination; wild-type *R26* locus, with the location of the probe indicated (red line); the structure of the targeted locus prior to and after Cre-mediated excision of the *loxP*-flanked (*PGK-Neo*, *tpA*) cassette. (B) Southern blot hybridisation of DNA isolated from wild-type (+/+) and two heterozygous ($R26^{Cond\text{-}Hesx1/+}$) ES clones after digestion with *Eco*RV. (C) Representative examples of PCR genotyping of wild-type, $R26^{Cond\text{-}Hesx1/+}$ and $R26^{Cond\text{-}Hesx1/Cond\text{-}Hesx1}$ embryos. (D-E) RNA ISH with *Hesx1* detects transcripts in the ventral diencephalon and developing Rathke's pouch in the $R26^{Cond\text{-}Hesx1/+}$ embryo (left), but they are ubiquitously expressed in the $R26^{Cond\text{-}Hesx1/+}$*;β-Actin-Cre* compound embryo (right). Note the failure in the closure of the cranial neural folds in the $R26^{Cond\text{-}Hesx1/+}$*;β–Actin-Cre* compound embryo (arrowheads). (F) Coomassie staining (top) and western blot using a specific HESX1 antibody (bottom) of protein extracts isolated from: (1) heads or (2) tails from $R26^{Cond\text{-}Hesx1/+}$ embryos (10 micrograms); (3) heads or (4) tails from $R26^{Cond\text{-}Hesx1/+}$*;β-Actin-Cre* compound embryos (2 micrograms). HESX1 protein is approximately 22 kDa.

Fig. S7. HESX1 cannot repress transcription from *Wnt1* reporter plasmids. (A) Schematic diagram of the *Wnt1* locus, boxes represent exons, black boxes represent the coding sequence of the gene. Thickened lines underneath correspond to regions upstream (165, 166 and 167) and downstream (2C, 1A, 147 and 39del) of the *Wnt1* coding region which contain putative binding sites for HESX1. The grey line for constructs 2C, 1A and 147 corresponds to the sequence shown in D (+8138 to +8249) containing three putative binding sites for HESX1. (B,C) Co-transfection of reporter constructs with a plasmid expressing *Hesx1 w*as unable to significantly repress activation of the reporter in any of the constructs tested. Co-transfection of a positive control reporter construct containing six repeats of a paired homeodomain consensus binding element (P3) showed that HESX1 was able to significantly repress transcription of this reporter by approximately 40% ($p = 0.002$). pRLSV40 *Renilla* luciferase reporter vector (Promega) was used to control for transfection efficiency. Experiments were performed in triplicate and data presented as mean ± SD. (D) Binding of HESX1 to a putative enhancer element downstream of *Wnt1* in vitro. (Left) Sequence of region downstream of *Wnt1* (+8138 to +8249), indicated by grey region in A, containing three partial putative paired homeodomain binding sites shown underlined (W1, W2 and W3). It has been shown that mutation of W3 causes ectopic activation of the *Wnt1* locus in the forebrain of transgenic embryos (Iler et al., 1995). (Right) EMSA shows that HESX1 is capable of binding to two of these sites, W2 and W3 (lower band, monomer and upper band, dimer). No binding was observed to the W1 sequence. In the EMSA assay, ^{32}P-labelled probes were

incubated with either: in vitro translation extract (lane 1); increasing amounts of in vitro translated (IVT) HESX1 (lanes 2 and 3); IVT HESX1 and 100 times excess of cold unlabeled probe (lane 4). W2 mut (ATTTA***GG***A) and W3 mut (T***GG***TTGAGGTAA***GG***A) ablated binding of HESX1 to the mutant probes, confirming the specificity of binding to these sites in vitro.

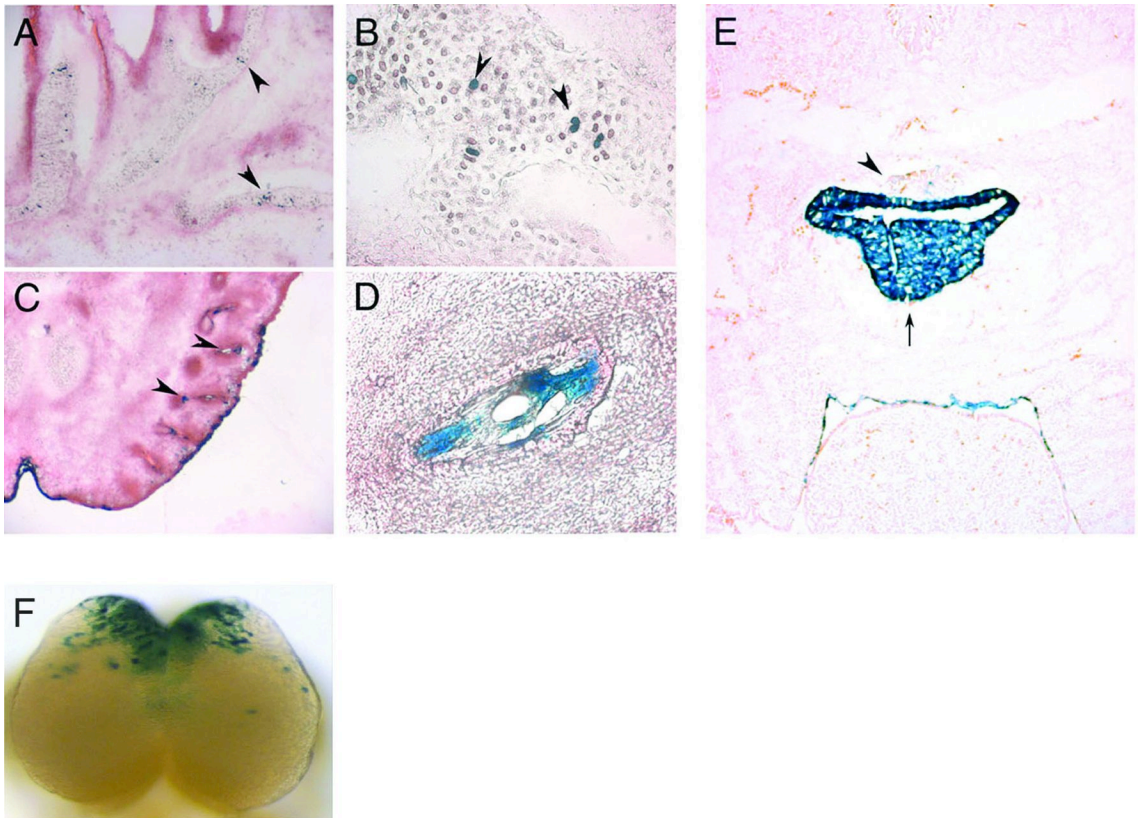

+/+
-/-
A
Atx
+/+
-/-
B
Pax6
+/+
-/-
C
D
TUNEL
E
F
+/+
-/-
G
A
P
H
I
J
+/+
-/-
Phospho-histone H3
K
L

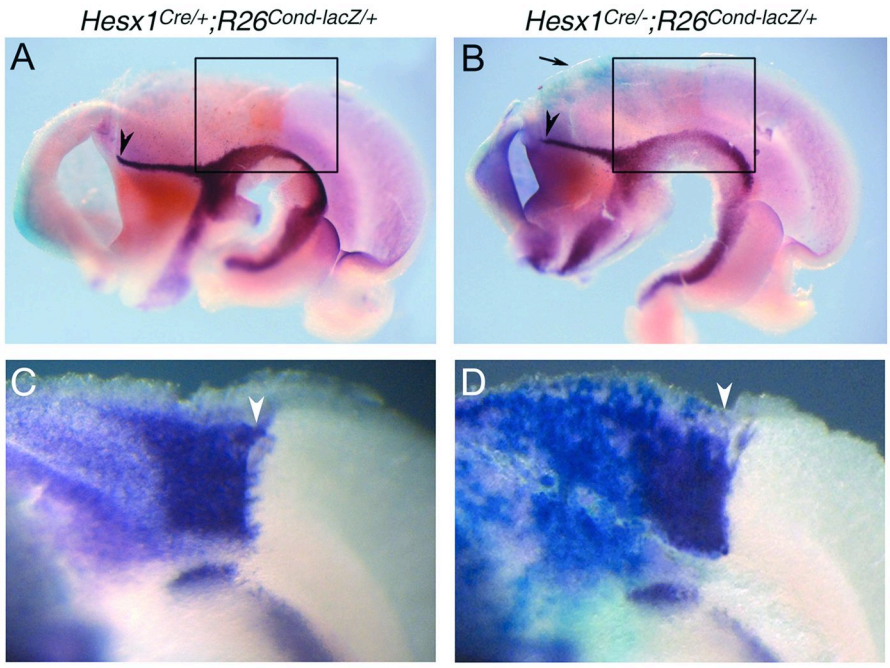

A
B
C
D
E
F
G
H
I
dd
dt
vt
op
J
ey
vd
K
hy
rp
L
lb
ht
M

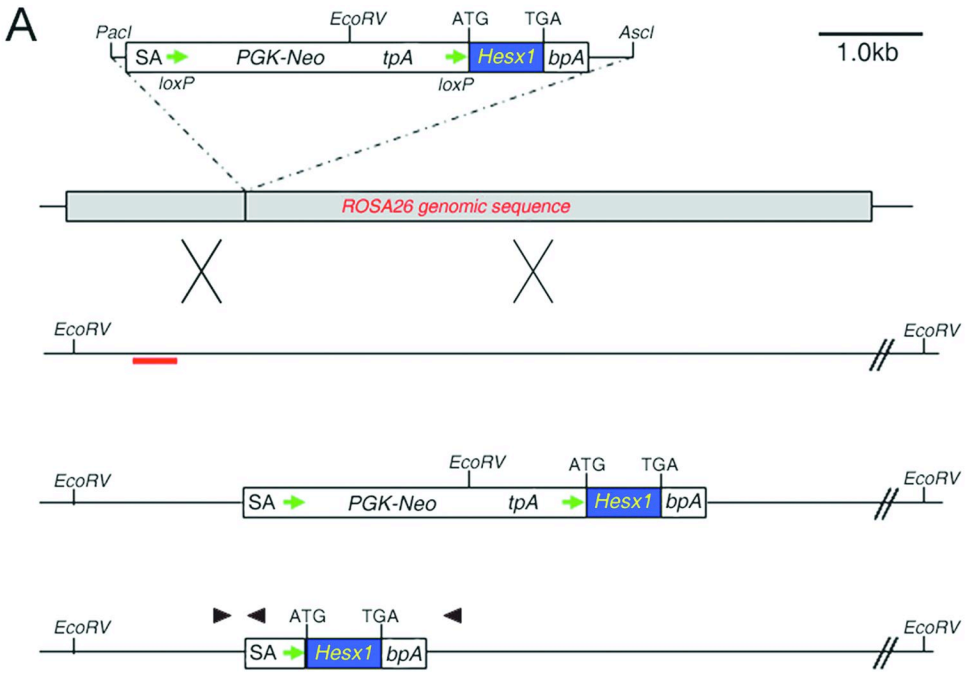

Hesx1+/+
A
Hesx1
Hesx1Cre/Cre
B
Cre
Hesx1 Cre/Cre
C
Cre
Hesx1Cre/+
D
Cre
Hesx1Cre/-
E
Cre

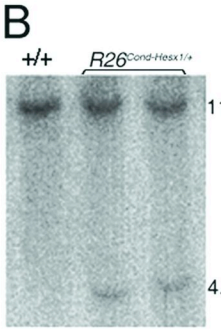
F

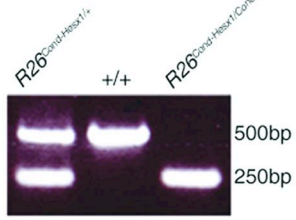

Graph comparing realtive Cre expression between Hesx1Cre/+ and Hesx1Cre/- embryos
Cre/+
Cre/-
Relative Quantification
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
0.985
1
0.612
1
0.893
2 somites
5 somites
8 somites

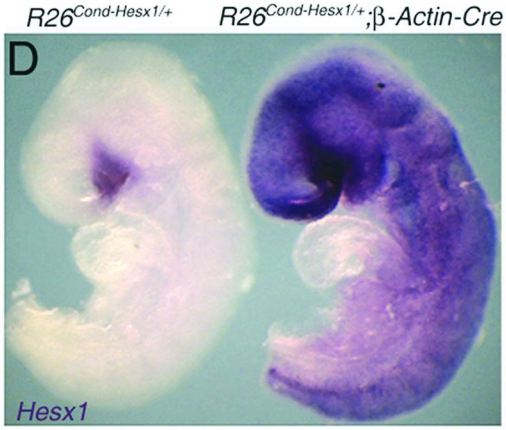

A
B
E
C
D
F

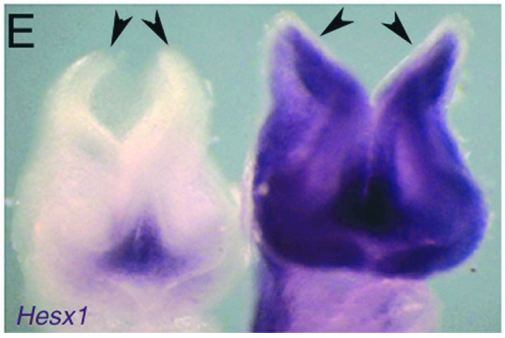

Hesx1Cre/+;R26Cond-lacZ/+
Hesx1Cre/-;R26Cond-lacZ/+
A
B
C
D

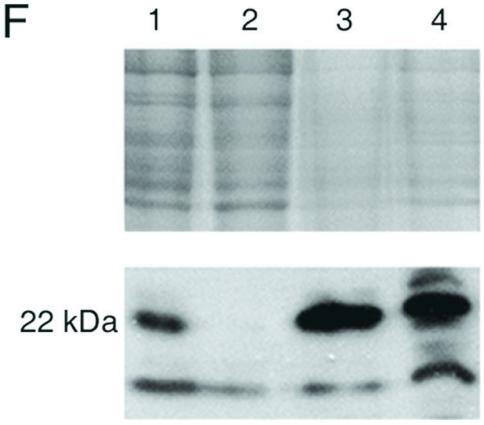

A
PacI
EcoRV
ATG
TGA
AscI
SA
PGK-Neo
tpA
Hesx1
bpA
loxP
loxP
1.0kb
ROSA26 genomic sequence
EcoRV
EcoRV
EcoRV
EcoRV
ATG
TGA
SA
PGK-Neo
tpA
Hesx1
bpA
EcoRV
EcoRV
ATG
TGA
SA
Hesx1
bpA

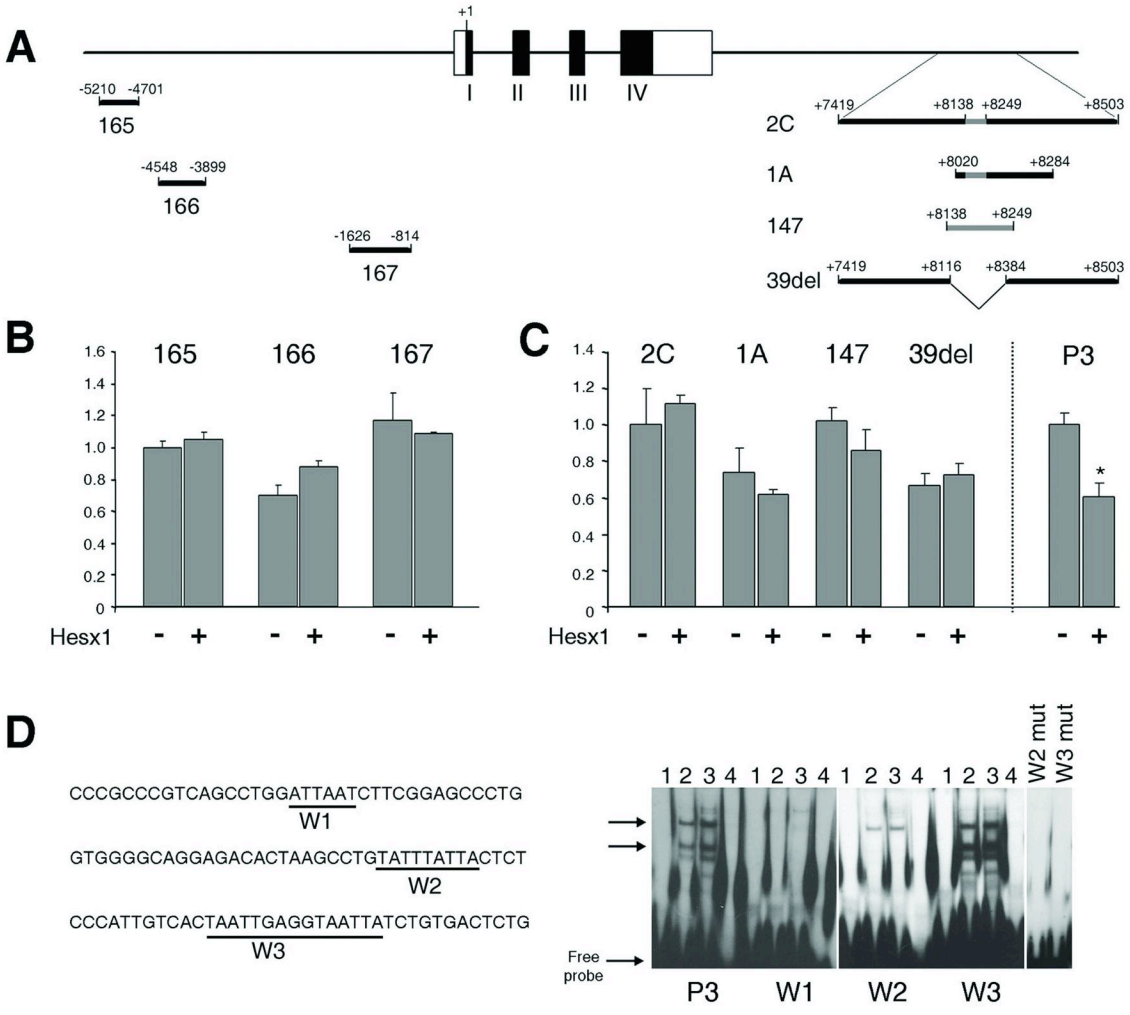

B
+/+
R26Cond-Hesx1/+
11kb
4.8kb


C
R26Cond-Hesx1/+
+/+
R26Cond-Hesx1/Cond-Hesx1
500bp
250bp


R26Cond-Hesx1/+
R26Cond-Hesx1/+;β-Actin-Cre
D
Hesx1


E
Hesx1


F
1
2
3
4
22 kDa


A
+1
I
II
III
IV
-5210 -4701
165
-4548 -3899
166
-1626 -814
167
+7419 +8138 +8249 +8503
2C
+8020 +8284
1A
+8138 +8249
147
+7419 +8116 +8384 +8503
39del
B
165
166
167
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
Hesx1 - + - + - +
C
2C
1A
147
39del
P3
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
*
Hesx1 - + - + - + - + - +
D
CCCGCCCGTCAGCCTGGATTAATCTTCGGAGCCCTG
W1
GTGGGGCAGGAGACACTAAGCCTGTATTTATTACTCT
W2
CCCATTGTCACTAATTGAGGTAATTATCTGTGACTCTG
W3
1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
W2 mut
W3 mut
Free probe
P3
W1
W2
W3