Supplemental Figures and Tables

Figure S1 (related to Figure 1)

(A-I) Expression patterns of *Pbx1* (A-F) and *Pbx2* (G-I). (A) Expression of *Pbx1* RNA at E11.7 in sagittal view (Allen Brain Atlas). (B) Expression of Pbx1 protein by immunohistochemistry at E12.5 in sagittal view. (C) Expression of *Pbx1* RNA at E14.5 in sagittal view (GenePaint). (D-F) expression of *Pbx1* RNA in coronal view at E12.5 (D), E13.5 (E) and E15.5 (F). (G) Expression of *Pbx2* RNA at E11.5 in sagittal view (Allen Brain Atlas). (H) Expression of *Pbx2* RNA at E13.5 in sagittal view (Allen Brain Atlas). (I) Expression of *Pbx2* RNA at E15.5 in sagittal view (Allen Brain Atlas). (J-P) PBX1A protein expression in the developing cortex. (J) High magnification of PBX1 (red) co-staining with PAX6 (green). (K) Quantification of co-localization of PBX1A and PAX6 signal in the ventricular zone of the E12.5 cortex was performed as previously described (Bolte and Cordelieres, 2006). Pearson's coefficient of 0.937 supports significant co-localization. (L) High magnification of PBX1A (red) co-staining with TBR2 (green). (M) Quantification of co-localization of PBX1A and TBR2 signal in the subventricular zone of the E13.5 cortex. Pearson's coefficient of 0.1 implies very little to no co-localization. (N) Quantification of co-localization of PBX1A and TBR1 signal in the cortical plate of the E15.5 cortex. (O) Quantification of co-localization of PBX1A and TBR1 signal in the region of the cortical plate delineated by white rectangle in (N). (P) Quantification of co-localization of PBX1A and TBR1 signal in the region of the cortical plate delineated by blue rectangle in (N). Pearson's coefficients of 0.758 (O) and 0.659 (P) supports co-localization. (Q-R) *Pbx1* RNA expression in control and conditional mutants at E12.5. (S): Ki67 (green):Cre (red) double immunofluoresence in the E15.5 cortex of E15.5 *Nex-Cre* shows lack of Cre expression in dividing ($Ki67^{+}$) cells. (T-Z) *Pbx;Emx1-cre P8* histological phenotypes in Nissl-stained coronal sections. (T,T') Loss of the anterior commissure in mutants is indicated by red arrows. Red arrow in control indicates position of the normal commissure. (U,U') Thinning corpus callosum is indicated by red arrows in mutants. Red arrow in control section indicates position of the normal corpus callosum. Green arrow in mutants points to indusium griseum. (V-W') Thinning of superficial cortical layers. (X-Z') Abnormal lamination of the hippocampal CA fields. (Z) High magnification of areas inside black rectangles in (Y).

Figure S2 (related to Figure 2)

(A-B) Quantification of cortical region changes in *Pbx1* mutants as analyzed whole mount *in situ* hybridization. Red line in (A) measures the distance from the frontal pole to the beginning of the caudal *Lmo4*-posistive domain. Green line in (A) measures the length of the whole cortex. (A') Quantification of the measurements in (A) presented as % of cortical length. *Emx1-cre* mutants show a reduction in the cortical measurement delineated by red line in (A), whereas *Nex-cre* mutants don't. (B) Red line measures the width of NT3-positive medial domain. Green line measures the width of the cortex. (B') Quantification of the measurements in (B) presented as % of cortical width. *Emx1-cre* mutants show a substantial increase in the width of the *Nt3*-positive medial domain, whereas in *Nex-cre* mutants this increase is not as pronounced. (C-G) Analysis of the cortical patterning by whole mount in situ hybridization in control *Pbx1F/+ (C), Pbx1F/-; Emx1-cre (D), Pbx1F/-; Pbx2-/+; Emx1-cre (E), Pbx2+/+ (F)* and *Pbx2-/+ (*G) mice using *Lmo4* in situ probe. *Pbx1F/-; Emx1-cre* (D), *Pbx1F/-; Pbx2-/+; Emx1-cre* (E) show a loss of *Lmo4*-positive rostral domain as compared to control *Pbx1F/+* mice. *Lmo4*-positive rostral domain in *Pbx2+/+* (F) and *Pbx2-/+* (G) mice is similar to control *Pbx1F/+* (C). (H-O) *In situ* hybridization analysis and quantification of *Dbx1* (H and I), *Pdzrn3* (J and K), *Cux2* (L and M) and *Lmo3* (N and O) signal between *Pbx1F/+* and *Pbx1F/-; Emx1-cre* genotypes and *Pbx1F/+;Pbx-/+* and *Pbx1F/-; Pbx2-/+; Emx1-cre* genotypes. (I'', K'', M'' and O'') Quantification of the *in situ* signal in (H-O') between different genotypes showing exacerbation of the molecular phenotype in *Pbx2-/+* mutants (mean+/-SD). (P-GG) *In situ* hybridization analysis of *Nt3* (P-R''), *Lmo4* (S-U''), *Cux2* (V-X''), *Lmo3* (Y-AA'), *Nurr1* (BB-DD'') and *Er81* (EE-GG'') on coronal sections of P8 control, *Pbx;Emx1-cre* and *Pbx;Nex-cre.* Red arrowheads show expression boundaries that shift in their expression in the *Pbx;Emx1-cre* mutant; red stars indicate loss of expression in the *Pbx;Emx1-cre* mutant. Green arrows in (S-U'') and (BB-DD'') indicate additional shifts in medial gene expression boundaries in mutant genotypes.

Figure S3 (related to Figure 3)

(A-B) Reelin overexpression in the *Pbx;Emx1-cre* mutant at E15.5 by immunohistochemistry; white arrowhead shows increased Reelin in the rostrodorsal cortex. (C) Quantification of Reelin (green);calretinin (red) co-expression in the region delineated by the white rectangle in (B''). Pearson's coefficient of 0.074 supports little to no co-expression, suggesting that ectopic Reelin^{+} cells are not Cajal Retzius cells.

Figure S4 (related to Figure 4)

Inversion of cortical layers in the *Pbx;Emx1-cre* mutant at E18.5 indicated using BrdU pulse-chase labeling. (A-A') E11.5 pulse; (B,B') E15.5 pulse. Coronal hemisections were stained using anti-BrdU histochemistry. Red arrowhead in (A') shows that early-born neurons are in a superficial position in the mutant (position of subplate in mutant is indicated by red dotted line). Red arrowhead in (B') shows that late-born neurons are in a deep position in the mutant. Red arrowhead in control section indicates normal position of labeled neurons. (A'') ad (B'') quantification of BrdU-positive cells in different layers of the cortex in control vs mutants. At least 5 measurements were taken for each cortical layer. Data was analyzed using ImageJ. * Indicates $p<0.05$ (mean+/-SD).

Figure S5 (related to Figure 5)

(A-I') Genes whose expression appears normal in the *Pbx;Emx1-cre* mutant at E12.5 and E13.5 by section in situ hybridization; coronal hemisections are shown. (J,J'): *Pbx1* expression in the cortical VZ appears unchanged in the *Fgf8*$^{neo/neo}$ hypomorphic mutant at E11.5 (note: reduced *Pbx1* expression in neuronal layers of the ventral cortex and LGE). (K,K') *Pbx1* expression in the cortical VZ appears unchanged in the *Pax6*$^{sey/sey}$ mutant at E12.5. (L-O') Progenitor defects in the *Pbx;Emx1-cre* mutant at E12.5 and E15.5. (L-L'') Normal phosphohistone-3 (PH3) in the VZ and reduced in the SVZ (250%); (M-M'') Reduced *Svet1* in the SVZ (~50%); (N-N'') reduced *Cxcl12* in the SVZ (~50%); (O-O'') Increased *Cav1* in the VZ (~50%). (P-T') Progenitor properties in the *Pbx;Nex-cre* mutant at E13.5 and E15.5.(P-Q') Normal phosphohistone-3 (PH3) in the VZ and SVZ; (R-T') Normal *Svet* and *Lhx2* expression. (U-V') *Pbx;Nex-cre* mutant has normal expression of Smoc1 and pSMAD1/5 at E15.5. * Indicates $p<0.05$ (mean+/-SD).

Figure S6 (related to Figure 6)

Distribution of PBX ChIP-Seq peaks with respect to the transcription start site (TSS) at E12.5 (A) and E15.5 (B). (C): Examples of E11.5 transient transgene mice that express LacZ in cortical progenitors (Vista Enhancer Browser) under the control of enhancer elements that are bound by PBX at E12.5 (ChIP-Seq peaks).

Figure S7 (related to Figure 7)

(A-C) Short and Long PBX motifs variants found in PBX ChIP-Seq peaks. (D-E) Hierarchical clustering dendrograms showing co-occurrence of motfis within PBX ChIP-seq peaks, with red boxes indicating significant co-occurrence clustering. AU (Approximately Unbiased) and BP (Bootstrap Probability) values are p-value representing how strong the cluster is supported by the co-occurrence data.

Table S1 (related to Figure 6)

Enhancers active at E11.5 (Visel et al., 2013) that have PBX ChIP-Seq peaks.

Table S2 (related to Figure 6)

Enhancers active in the forebrain

Supplemental experimental procedures

Mice and genotyping

Mice were housed and used in accordance with National Institutes of Health and UCSF guidelines. *Pbx1*^{flox} and *Pbx2*$^{+/-}$ mice were obtained from Dr. Licia Selleri, and were genotyped as in (Capellini et al., 2006; Koss et al., 2012).

Image analysis

ImageJ was used as previously described (McCloy et al., 2014) to analyze intensity and calculate Integrated Density of *in situ* and immunostaining signals. Briefly, images were converted to 8-bit files and inverted. For a given selected area within a brain section the mean gray value measurement was collected along with integrated density measurement. Similar measurements were collected for the adjacent background areas. Corrected integrated density (reported in figures as "Integrated Density") was calculated as: integrated density – (selected area * mean gray value of the background reading). To analyze co-localization of fluorescent signal we used the JACoP ImageJ plugin (http://rsb.info.nih.gov/ij/ plugins/track/jacop.html) (Bolte and Cordelieres, 2006)

Statistical methods

Statistical analysis of data collected during image analysis was performed using GraphPad prism software. Parametric unpaired t-test was perform to determine statistical significance. Data in all graphs represent mean values +/- SD.

Informatics Methods

ChIP-seq reads were adaptor-trimmed and quality filtered, and mapped to mm9 using BWA (Li and Durbin, 2009). Peaks were called using MACS1.4 with default settings. For PBX datasets, we used the union of peaks called against input control and negative blocking peptide control. PBX peaks were filtered to remove regions where the majority of supporting reads originated in repeat sequence and annotated to genomic features.

All comparisons between peak datasets were performed via overlapping called regions. Data from these experiments has been deposited in the GEO repository (accession TBD). Testing of functional enrichment of PBX regions was performed using GREAT (McLean et al., 2010).

Sequences for the core 1000bp within peaks from PBX E12.5 and E15.5 ChIP-Seq datasets were downloaded from the UCSC mm9 genome build. Motif analysis was performed using peak-motifs (Thomas-Chollier et al., 2012), with default parameters selected and all four motif finding algorithms applied (dyads, local words, positions, and oligos) identifying the top ten motifs of lengths six, seven, and eight base pairs. After *de novo* identification, enriched sequence motifs were clustered according to motif correlation using RSAT and co-occurrence using pvclust. Peaks were scored for presence of each consensus motif. Correlations between motif occurrences were tested using the pvclust package in R, producing a hierarchical clustering dendrogram with significance scores across the tree generated via bootstrap analysis. Motifs were matched to known TF binding sites from JASPAR (core vertebrates, mouse, and homeo mouse), footprintDB, and CisBP (mouse). Motifs with normalized correlation between binding motifs greater than 0.5 were clustered into sets with the most highly correlated motif across the set used as the consensus. Motifs were manually correlated to cluster motifs that were highly related but clustered independently. Pbx E12.5 and E15.5 sequences were analyzed against each other using peak-motifs to identify motifs that showed temporal specificity. PBX binding sites were scored for presence/absence of the clustered motifs, and testing for difference in distribution of motif binding site relative to gene promoters was performed.

Immunohistochemistry

20μm-thick frozen sections on slides were allowed to thaw and dry, washed three times with PBS (5 min each) and blocked for 1 hour at room temperature using Blocking buffer containing 10% goat serum and 0.1% TritonX-100 in PBS. Primary antibody staining was performed overnight at 4 degrees. After 3 washes (10 min each) with PBS containing 0.1% Triton, sections were incubated with secondary antibody labeled with Alexa fluorescent dyes (Life technologies). Nuclei were counterstained with DAPI. Both primary and secondary antibodies were diluted in Blocking buffer. After secondary

antibody staining sectioned were washed three times with PBS containing 0.1% Triton. Slides were then rinsed with PBS and mounted using hardening mounting media (Dako). PBX1a and BrdU staining required antigen retrieval. Antigen retrieval step was done by incubating slides at ~100C in a steaming bath in Citric acid buffer (10mM Sodium Citrate, 0.05% Tween 20, pH 6.0) for 5min. Slides were allowed to cool to room temperature and were washed with PBS and processed as described above. PBX1a antibody was a gift from Dr. Michael Cleary (Stanford). Antibodies for immunostaining were Pax6 (rabbit, Millipore), Tbr1 (rabbit, Abcam), Tbr2 (rabbit, Abcam), Ctip2 (rat, Abcam), Reelin (mouse, Millipore), Calretinin (rabbit, Swant), and BrdU (mouse, BD Biosciences).

Nissl staining

20µm frozen coronal and sagittal sections were allowed to air dry, rehydrated through an ethanol series and incubated in 0.1% Cresyl Violet solution for several minutes until the desired intensity of staining was achieved. Sections were then washed, differentiated through an ethanol series, cleared in Xylenes, and mounted with Permount.

Microarray analysis

Cortices (includes medial, dorsal, and lateral pallium) were dissected from E12.5 and E15.5 embryos and preserved in RNA Later (Qiagen) awaiting genotyping. RNA was extracted from the cortex using Qiagen RNAeasy kit according to the vendor's instruction. DNAse digestion was included as part of the RNA extraction. Microarray and data analysis was performed at the Sandler Center Functional Genomics core at UCSF. Four control and four mutant cortices were used for the analysis. Data was deposited to the GEO database: GSE73286; GSE73286.

BrdU birthdating

BrdU solution was prepared in PBS. Pregnant dams received a single 50mg/kg dose intraperitoneally at different gestational time points. Embryos were collected at E18.5 and analyzed for BrdU label distribution in the cortex.

References

Bolte, S., and Cordelieres, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. J Microsc *224*, 213-232.

Capellini, T.D., Di Giacomo, G., Salsi, V., Brendolan, A., Ferretti, E., Srivastava, D., Zappavigna, V., and Selleri, L. (2006). Pbx1/Pbx2 requirement for distal limb patterning is mediated by the hierarchical control of Hox gene spatial distribution and Shh expression. Development *133*, 2263-2273.

Koss, M., Bolze, A., Brendolan, A., Saggese, M., Capellini, T.D., Bojilova, E., Boisson, B., Prall, O.W., Elliott, D.A., Solloway, M.*, et al.* (2012). Congenital asplenia in mice and humans with mutations in a Pbx/Nkx2-5/p15 module. Dev Cell *22*, 913-926.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.

McCloy, R.A., Rogers, S., Caldon, C.E., Lorca, T., Castro, A., and Burgess, A. (2014). Partial inhibition of Cdk1 in G 2 phase overrides the SAC and decouples mitotic events. Cell Cycle *13*, 1400-1412.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28, 495-501.

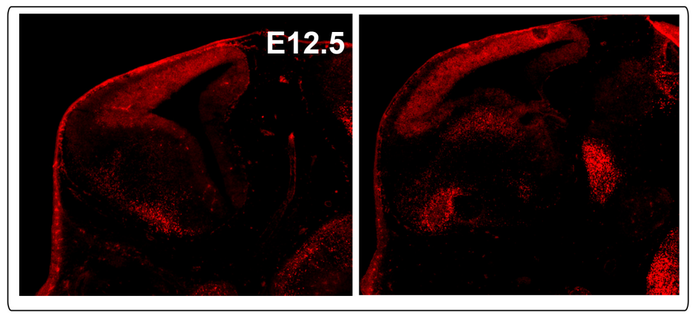
Thomas-Chollier, M., Herrmann, C., Defrance, M., Sand, O., Thieffry, D,, van Helden J. (2012). RSAT peak-motifs: motif analysis in full-size ChIP-seq datasets. Nucleic Acids Res. *40*, e31.

Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R.V., McKinsey, G.L., Pattabiraman, K., Silberberg, S.N., Blow, M.J.*, et al.* (2013). A high-resolution enhancer atlas of the developing telencephalon. Cell *152*, 895-908.

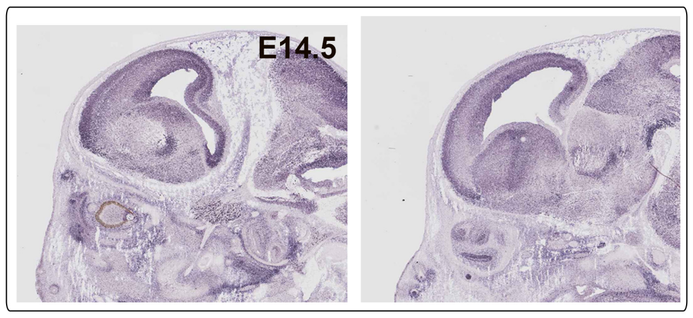
Pbx1

A

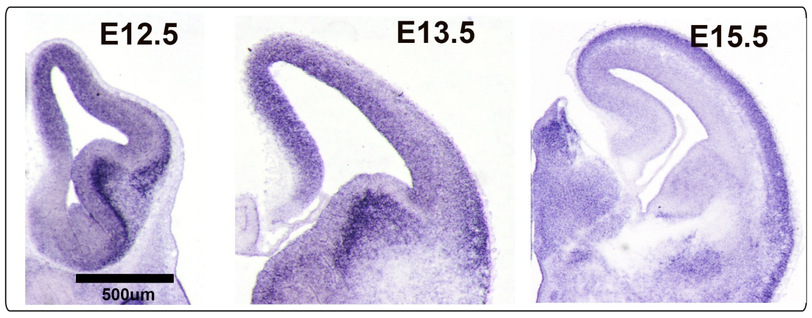
B



C

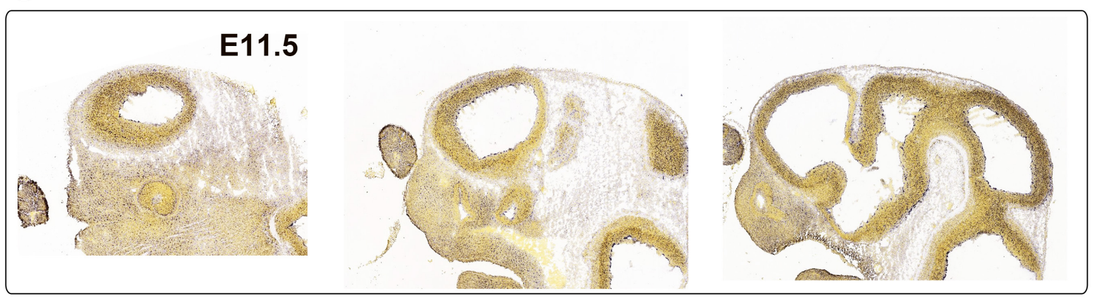


D E F

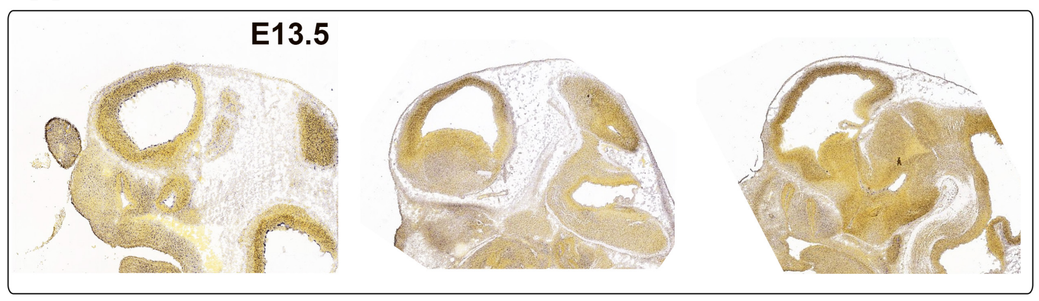


Pbx2

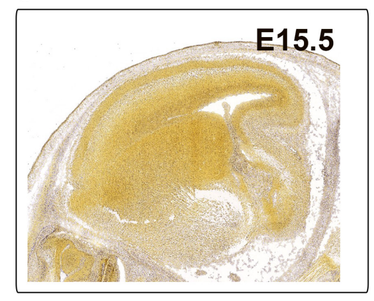
G



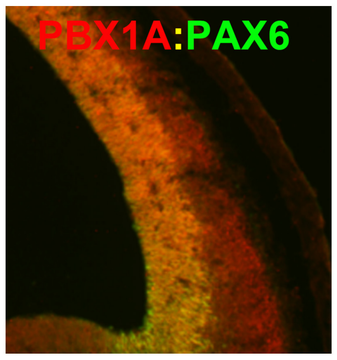
H



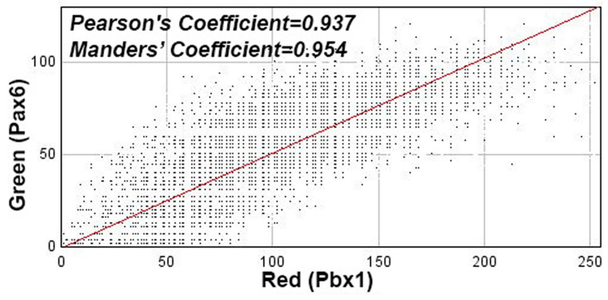
I



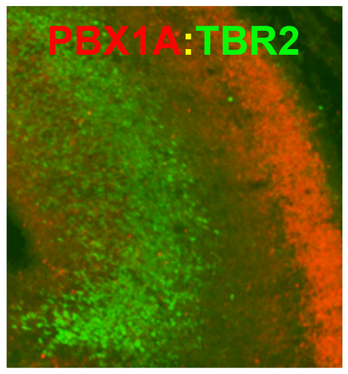
J



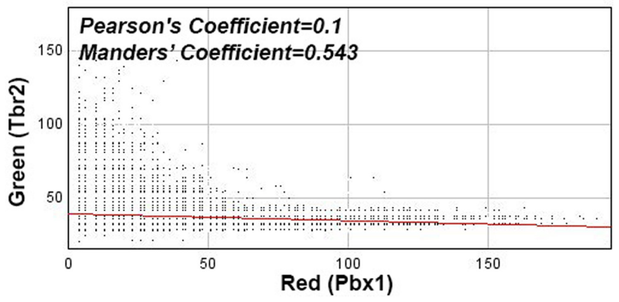
K



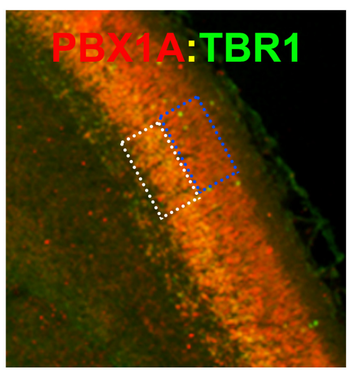
L



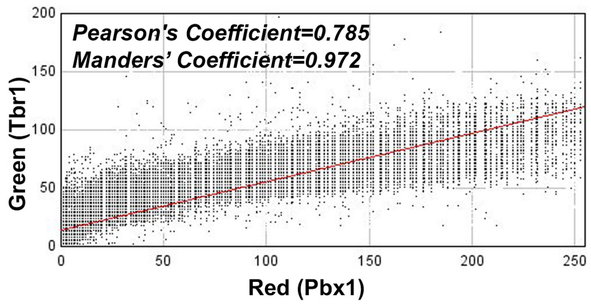
M



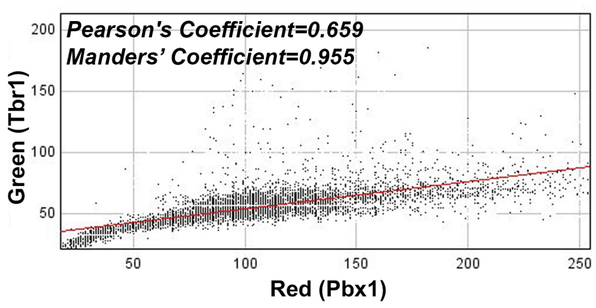
N



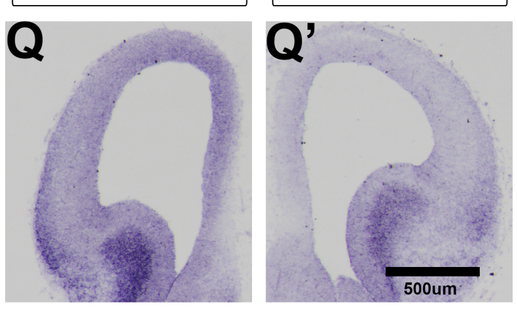
O



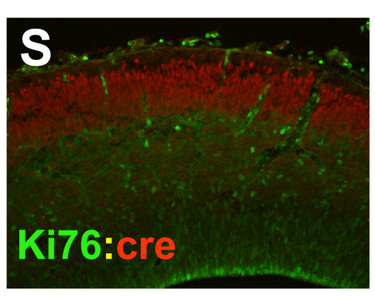
P



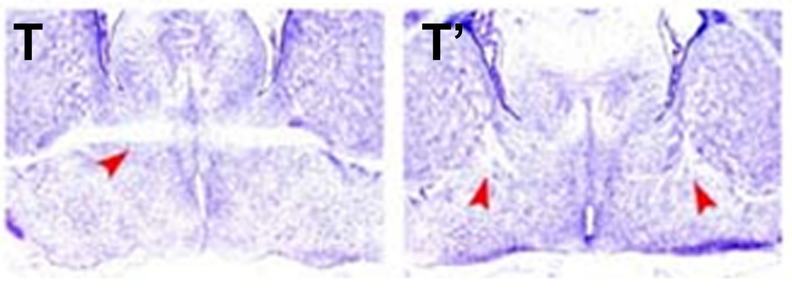
	control	Emx1-cre mutant	control	Nex-cre mutant
Pbx1	Q	Q'	R	R'



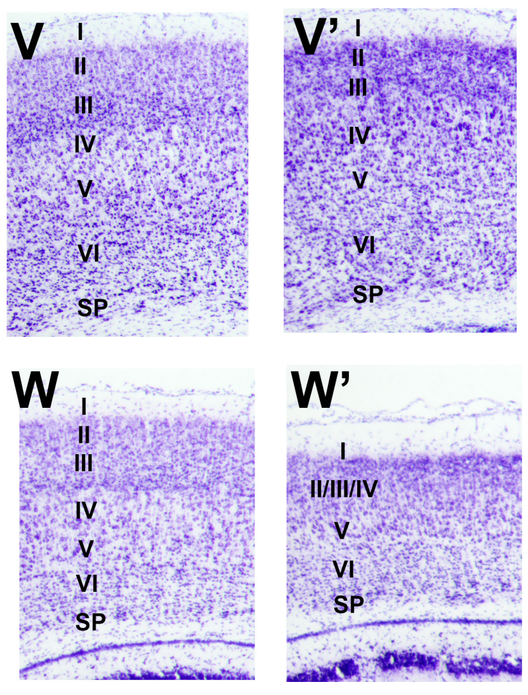
S



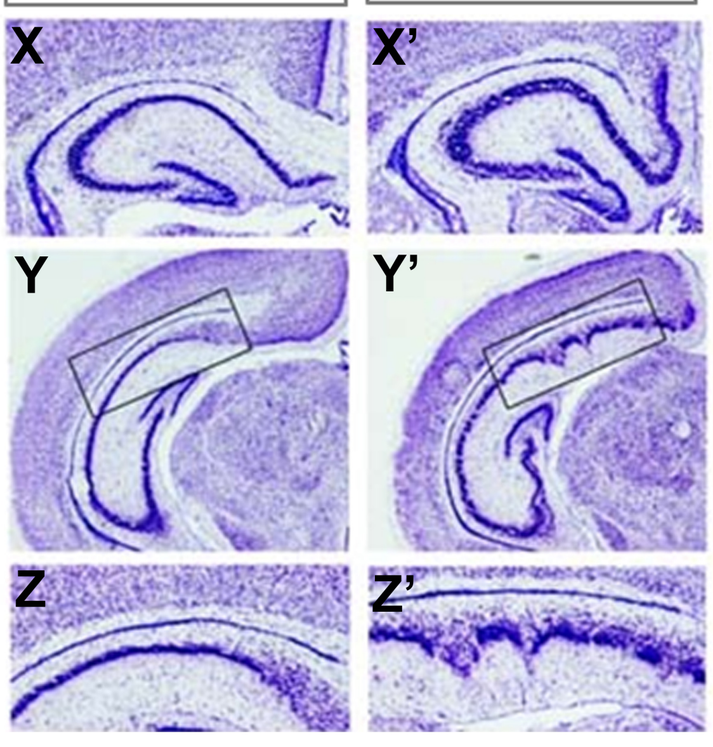
Control	Emx1-cre Mutant
anterior commissure	
T	T'
corpus callosum	
U	U'

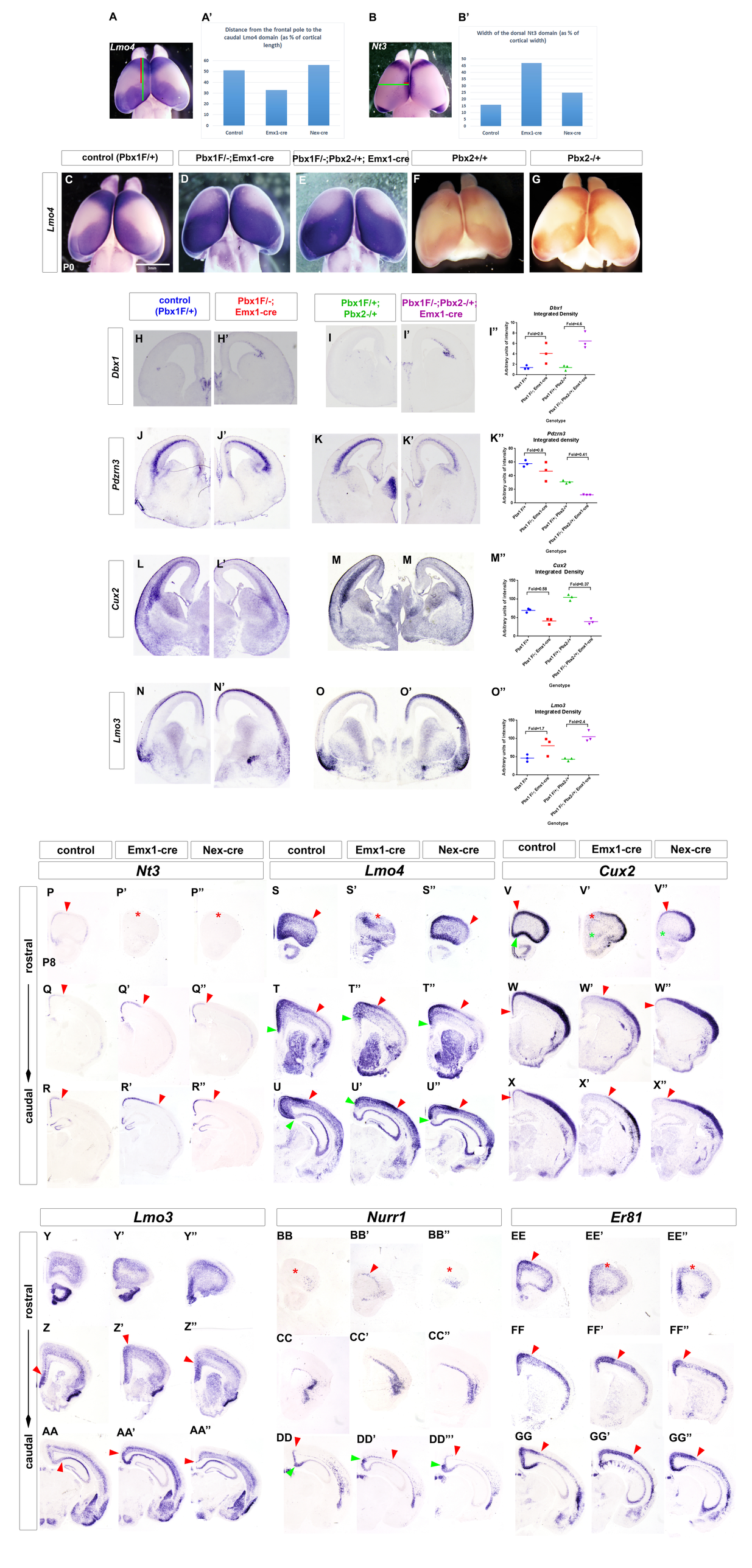


Control	Emx1-cre Mutant
V	V'
W	W'



Control	Emx1-cre Mutant
X	X'
Y	Y'
Z	Z'




A
A'
Lmo4
Distance from the frontal pole to the caudal Lmo4 domain (as % of cortical length)
Control
Emx1-cre
Nex-cre
B
B'
Nt3
Width of the dorsal Nt3 domain (as % of cortical width)
control (Pbx1F/+)
Pbx1F/-;Emx1-cre
Pbx1F/-;Pbx2-/+; Emx1-cre
Pbx2+/+
Pbx2-/+
C
D
E
F
G
P0
3mm
control (Pbx1F/+)
Pbx1F/-; Emx1-cre
Pbx1F/+; Pbx2-/+
Pbx1F/-;Pbx2-/+; Emx1-cre
Dbx1
H
H'
I
I'
I''
Dbx1
Integrated Density
Fold=2.9
Fold=4.6
Arbitrary units of intensity
Genotype
Pdzrn3
J
J'
K
K'
K''
Pdzrn3
Integrated density
Fold=0.8
Fold=0.41
Cux2
L
L'
M
M
M''
Cux2
Integrated Density
Fold=0.58
Fold=0.37
Lmo3
N
N'
O
O'
O''
Lmo3
Integrated Density
Fold=1.7
Fold=2.4
control
Emx1-cre
Nex-cre
Nt3
Lmo4
Cux2
rostral
caudal
P8
P
P'
P''
Q
Q'
Q''
R
R'
R''
S
S'
S''
T
T''
T''
U
U'
U''
V
V'
V''
W
W'
W''
X
X'
X''
Lmo3
Nurr1
Er81
Y
Y'
Y''
Z
Z'
Z''
AA
AA'
AA''
BB
BB'
BB''
CC
CC'
CC''
DD
DD'
DD'''
EE
EE'
EE''
FF
FF'
FF''
GG
GG'
GG''

Reelin:Calretinin

control

Emx1-cre mutant

A

A’

A’’

E15.5

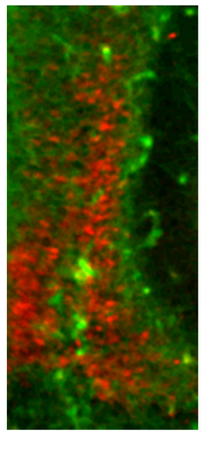
200um

B

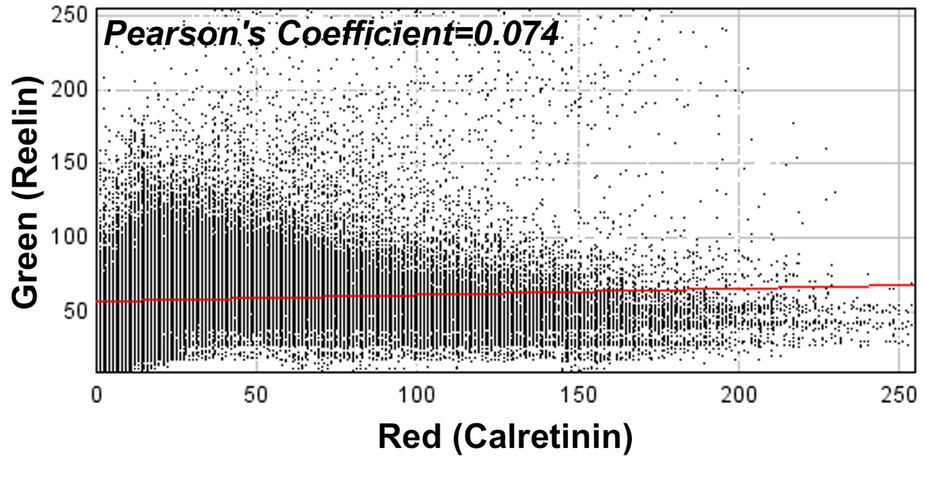
B’

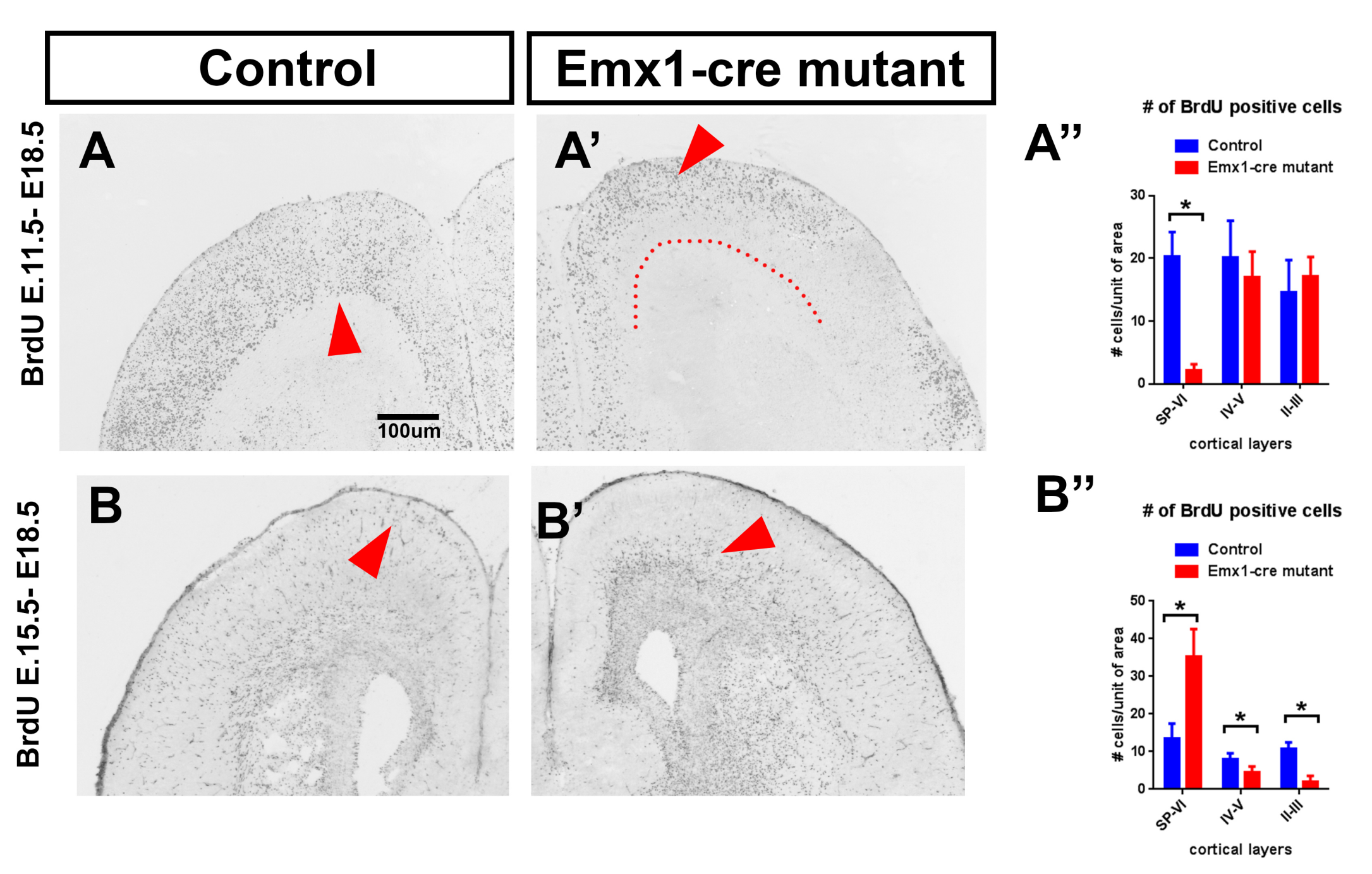
B’’

C



C’



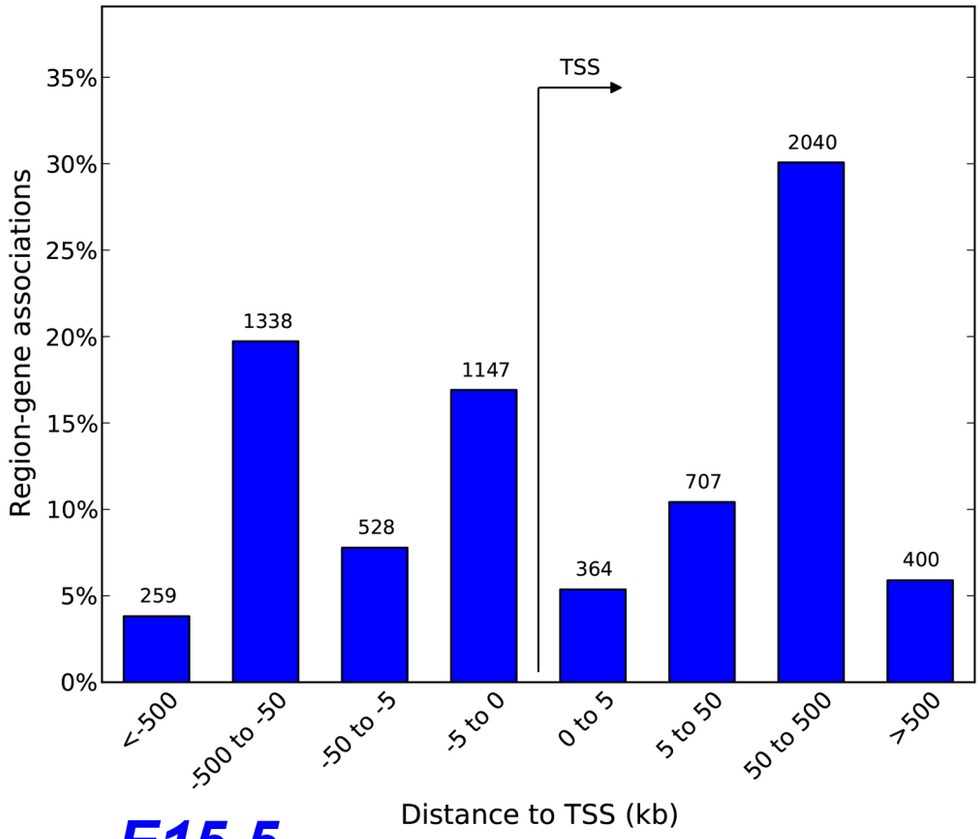

Control
Emx1-cre mutant
BrdU E.11.5- E18.5
BrdU E.15.5- E18.5
A
A'
A''
B
B'
B''
100um
# of BrdU positive cells
Control
Emx1-cre mutant
cells/unit of area
cortical layers
SP-VI
IV-V
II-III
*

control
Emx1-cre mutant
control
Emx1-cre mutant
Sp8
A
A'
E12.5
500um
Axin2
F
F'
E13.5
Pea3
B
B'
E12.5
Wnt3a
G
G'
E13.5
Erm
C
C'
E12.5
Lmx1a
H
H'
E13.5
Coup-TF1
D
D'
E12.5
Bmp4
I
I'
E13.5
Pax6
E
E'
E13.5
control
Emx1-cre
PH3
L
L'
E12.5
L''
Control
Emx1-cre
Cell count/unit area
15
10
5
0
*
VZ
SVZ
Svet1
M
M'
E15.5
500um
M''
Integrated Density
2.0
1.5
1.0
0.5
0.0
*
Cxcl12
N
N'
E15.5
N''
Integrated Density
1.0
0.8
0.6
0.4
0.2
0.0
*
Cav1
O
O'
E15.5
O''
Integrated Density
1.0
0.8
0.6
0.4
0.2
0.0
*
control
Nex1-cre mutant
PH3
P
P'
E13.5
Q
Q'
Svet1
R
R'
E15.5
control
Nex1-cre mutant
Lhx2
S
S'
E13.5
T
T'
E13.5
Smoc1
U
U'
E15.5
500um
pSMAD1/5
V
V'
E15.5
wt
Fgf8 neo/neo
wt
Pax6 sey/sey
Pbx1
J
J'
E11.5
K
K'
E12.5

A

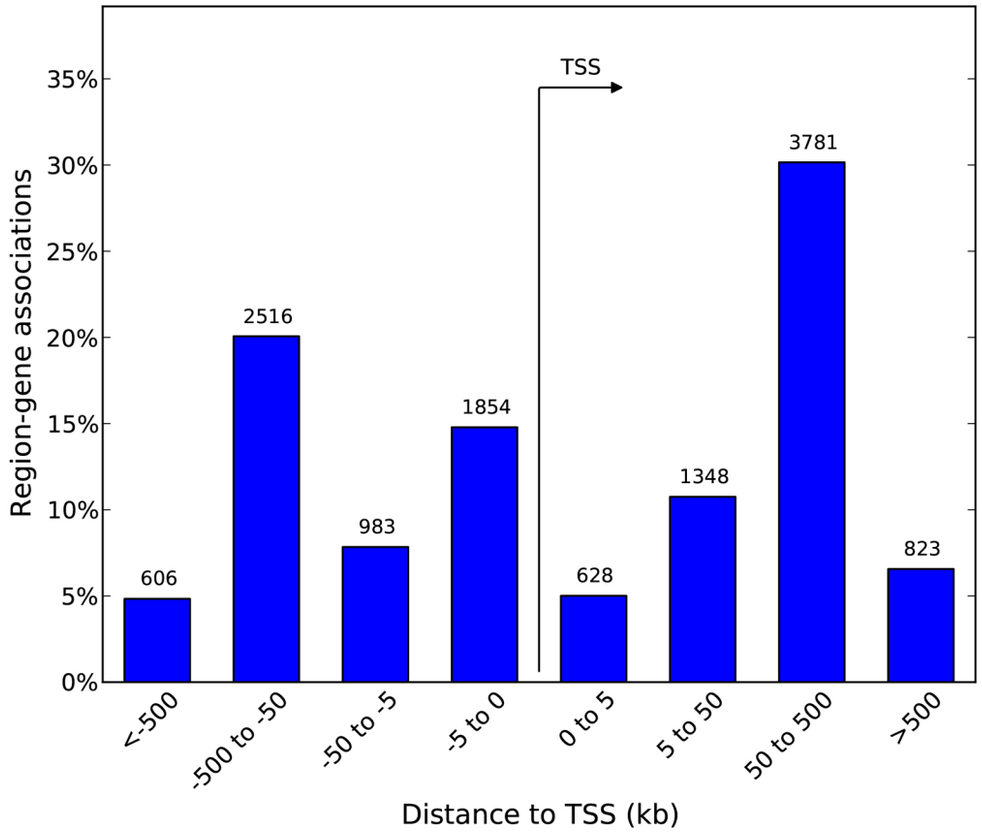
Distribution of PBX ChIP-seq peaks with respest to the transcriptional start site (TSS)

E12.5

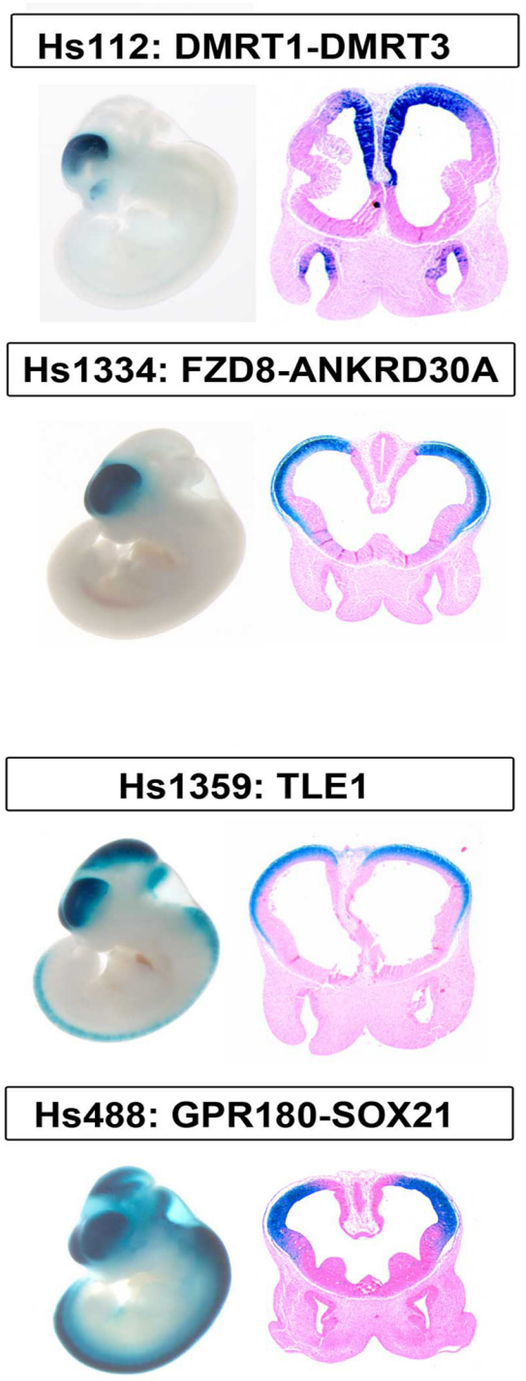


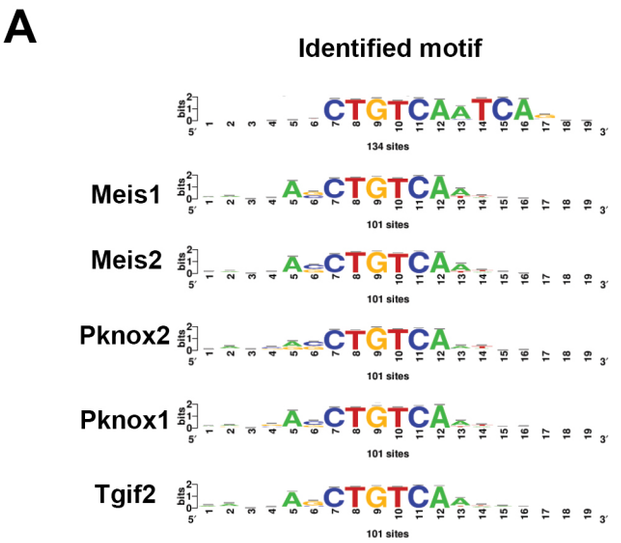
B

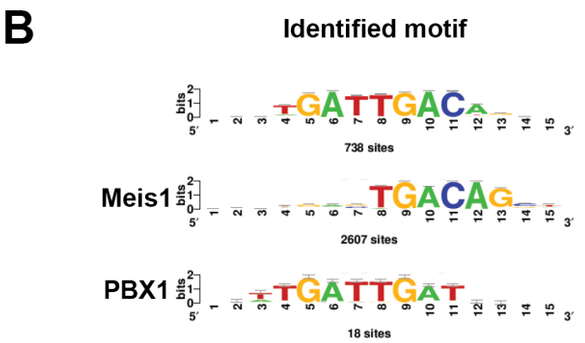
E15.5

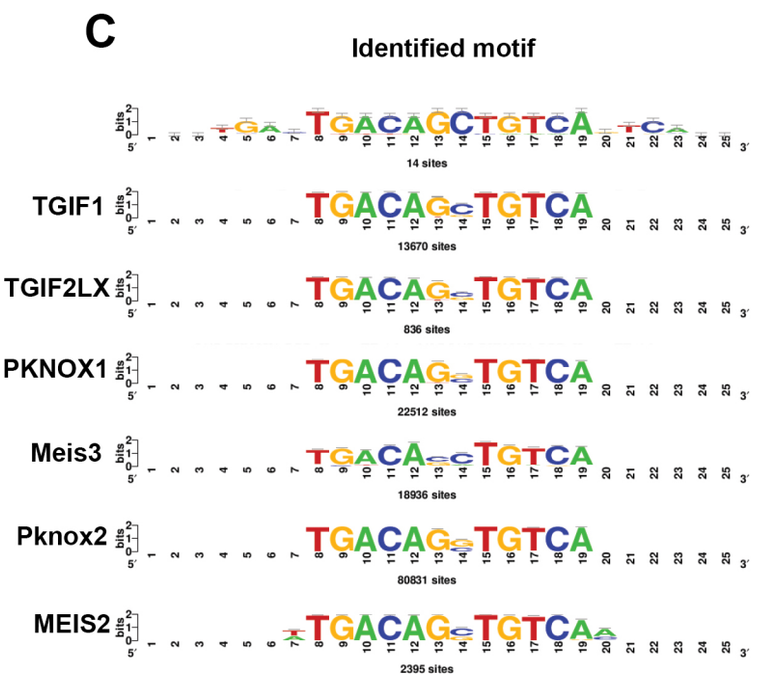


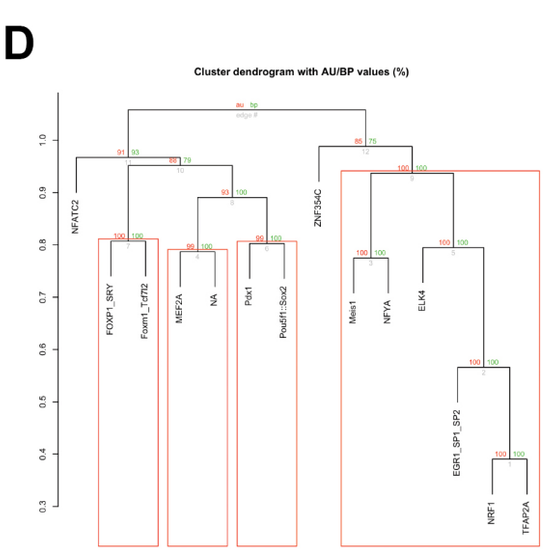
C

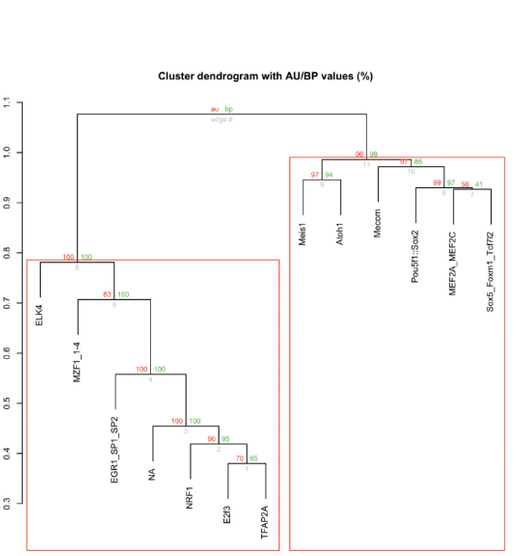



A
Identified motif
134 sites
Meis1
101 sites
Meis2
101 sites
Pknox2
101 sites
Pknox1
101 sites
Tgif2
101 sites


B
Identified motif
738 sites
Meis1
2607 sites
PBX1
18 sites


C
Identified motif
14 sites
TGIF1
13670 sites
TGIF2LX
836 sites
PKNOX1
22512 sites
Meis3
18936 sites
Pknox2
80831 sites
MEIS2
2395 sites


D
Cluster dendrogram with AU/BP values (%)
NFATC2
FOXP1_SRY
Foxm1_Tcf7l2
MEF2A
NA
Pdx1
Pou5f1::Sox2
ZNF354C
Meis1
NFYA
ELK4
EGR1_SP1_SP2
NRF1
TFAP2A


E
Cluster dendrogram with AU/BP values (%)
ELK4
MZF1_1-4
EGR1_SP1_SP2
NA
NRF1
E2f3
TFAP2A
Meis1
Ash1
Mecom
Pou5f1::Sox2
MEF2A_MEF2C
Sox5_Foxm1_Tcf7l2