www.sciencemag.org/cgi/content/full/science.1190831/DC1

Supporting Online Material for

Sex-Specific Parent-of-Origin Allelic Expression in the Mouse Brain

Christopher Gregg, Jiangwen Zhang, James E. Butler, David Haig, Catherine Dulac*

*To whom correspondence should be addressed. E-mail: dulac@fas.harvard.edu

Published 8 July 2010 on *Science* Express
DOI: 10.1126/science.1190831

This PDF file includes

Other Supporting Online Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/science.1190831/DC1)

Supplementary Materials and Methods

Experimental Procedures

Illumina Sequencing Sample Preparation

Imprinting was assessed by Illumina sequencing as described in our companion study (S1). Male and female F1 hybrid mice were generated by reciprocal crossings of CAST/EiJ and C57BL/6J mouse strains (Jackson Laboratories) to generate F1i (initial) and F1r (reciprocal) cross offspring. These samples were used to assess imprinting. In addition, samples were prepared from the parents of the F1 hybrid offspring for the adult mPFC and POA. The CAST and C57 parent samples were used to identify SNPs in genes expressed in the different brain regions as detailed in our companion paper (S1).

The POA and mPFC were microdissected from fresh 500 µm vibratome sections of adult brain from 5 week old adult male and female mice. POA dissections were landmarked based on the anterior commissure and included the AVPV. mPFC dissections were landmarked based on the relative morphology and position of the corpus callosum as guided by the Allen Brain Atlas reference atlas. RNA was isolated from adult POA and mPFC samples using Trizol (Sigma) according to manufacturers instructions and pooled samples were further purified using an RNeasy Kit (Qiagen). RNA was pooled from 8 different F1 hybrid adults. RNA quality was assessed using an Agilant Bioanalyzer and samples with an RNA integrity score >9 were used for sequencing. For each sample, mRNA was purified from 10 µg of total RNA using two rounds of hybridization to oligo(dT) beads (Dynal). RNA was fragmented and cDNA was generated and prepared for sequencing according to the Illumina RNA-Seq protocol for the Genome Analyzer sequencing platform (Illumina Inc.). cDNA fragment size selection for sequencing was 200±25bp.

Sequence Depth

Two technical replicates were run for each sample and the data for each replicate was pooled. For each adult sample, we generated 1.3-2.8 Gb of sequence. Sequence metrics are as described in our companion study (S1).

Data Analysis

SNP sites were uncovered using sequence from the adult CAST and C57 mPFC and POA (S1). The 32-mer reads for each F1 hybrid sample were aligned to the transcriptome using Novoalign. Reads were sorted by NovoAlign into candidate SNP containing reads (1 or more mismatches), perfectly matched reads, and

those that did not align. Reads with unique alignments and reads with repeated alignments in the transcriptome (generally due to multiple isoforms) were maintained separate. Reads that did not align to the transcriptome, but did align to the genome indicated detection of unannotated transcription (often associated with UTRs). We cross-referenced the F1 hybrid reads with the SNP sites annotated in our CAST/C57 SNP Reference Table (data deposited in gene expression omnibus (GSE22131). SNP site base calls in the reference tables were used to assign the reads as CAST or C57 specific for each SNP site in a given expressed transcript. Reads in the F1i and F1r cross were analyzed to assess imprinting. Imprinting was assessed by a $\chi 2$ test in the initial and reciprocal cross ($P<0.05$) and by requiring a paternal or maternal bias greater than 50% in both crosses. A minimum read depth of 10 in each cross was required for analysis of male versus female imprinting. Imprinted genes had one or more significantly imprinted SNP sites. To analyze imprinting of X-linked genes in female samples, no minimum read depth was required, but weak SNP sites were filtered by removing any SNPs with 1 or more paternal reads in either the F1i or F1r male samples. For the X chromosome, the expected values for the $\chi 2$ were adjusted for strain and maternal bias in X chromosome expression according to the overall biases calculated from the chrX datasets. Note that the $\chi 2$ test is an approximation at read depths below 10.

For the analysis of sex specific imprinting on the autosomes, this approach assessed sex specific imprinting for over 9000 genes expressed in the POA (male: 9235, female: 11241) or mPFC (male: 9584, female: 10488) with an average of 100 reads or more per SNP site.

Xegfp Transgenic Analysis

Xegfp transgenic mice have been previously described (S2). Xegfp mice were crossed with 129 wildtype mice to generate 5 week old heterozygous females for analysis. EGFP was visualized by immunohistochemistry using rabbit anti-GFP 1:100 (Inivtrogen) and double-labeled with mouse anti-NeuN 1:100 (Sigma) or mouse anti-vGLUT2 1:200 (Millipore). Secondary antibodies were 555 or 488 Alexa dye conjugated donkey anti-mouse or rabbit (Invitrogen). The number of EGFP+ cells was quantified by counting two 150 μm^2 fields in each brain hemisphere for four 20 µm cryosections per animal. Numbers represent mean±SEM.

Sequenom

A cohort of F1 hybrid animals that was distinct from the Illumina studies was generated for all iPLEX Sequenom studies. For each sample, 3 biological replicates and 3 technical replicates were run. Results indicate the average allele frequency of all replicates. Each biological replicate consisted of RNA

pooled from two animals from which a cDNA library was prepared. Primers were designed to target a 200 bp region of sequence with the target SNP site in the middle according to published methods (S3). The Sequenom MALDI-TOF iPLEX sample preparation and analysis was carried by the Partners HealthCare Center for Personalized Genetic Medicine (Cambridge, MA, USA)(http://www.hpcgg.org/).

Il18 Mice

Previously described Il18 homozygous -/- mice (S4) were obtained from Jackson Laboratories (Bar Harbor, Maine, USA). Mutant mice were crossed to C57BL/6J controls to generate Il18 -/+ and Il18 +/- offspring. Male and female animals were sacrificed at 30 weeks of age and RNA was harvested from the mPFC and POA using Trizol (Sigma). A cDNA library was generated using the Invitrogen Superscript II RT-PCR Kit according to manufacturer's instructions (Invitrogen). Quantitative PCR was performed to assess expression levels of Il18 using the Quantitect SYBR Green PCR kit (Qiagen) and an MJ Opticon 2 thermal cycler according to the manufacturers instructions. Three technical replicates were run for each sample. The results were normalized to GAPDH levels (similar results were obtained with normalization to beta-actin levels). The primer sequences used are: Il18 F- 5' CACTTCTCCCCTGTGGTGTGCTG 3', Il18 R- 5' ATCAGTCATATCCTCGAACACAGGCTGTC 3'; Gapdh F – 5' GCTGAACGGGAAGCT CAC T 3', R – 5' GGGAGTTGCTGTTGAAGTCG 3'.

Supplementary References

S1. C. Gregg et al., companion study (2010).

S2. A. K. Hadjantonakis, M. Gertsenstein, M. Ikawa, M. Okabe, A. Nagy, Nat Genet 19, 220 (1998).

S3. Jurinke et al., Mutat Res 573, 83 (2005).

S4. Zorrilla et al., Proc Natl Acad Sci USA 104, 11097 (2007).

Supplementary Figures and Tables

Fig. S1. Schematic depicting possible modes of sex specific parental effects. (A-C) Schematic descriptions of sex-specific genomic imprinting, including biased X-inactivation (A), imprinting of X-linked genes leading to differential gene expression from the active Xp versus the Xm (B), or sex-specific imprinting of autosomal genes (C), an example of a male-specific imprinted gene is shown.

Fig. S2. Mosaic expression in postnatal and adult non-neural tissues of heterozygous X^{egfp} transgenic mice confirms that the X-linked *egfp* reporter undergoes random X-inactivation. (**A**) EGFP expression is uniformly detected in the skin of postnatal day 5 (P5) male $Xm^{egfp}Y$ pups. (**B-D**) Analysis of EGFP expression in the skin (B) of individual P5 female $Xm^{egfp}Xp$ and $XmXp^{egfp}$ pups (1, 2 and 3 indicate examples of different individual pups; red is $Xm^{egfp}Xp$, blue is $XmXp^{egfp}$) clearly demonstrates the stochastic mosaic expression pattern expected from random X-inactivation. Similarly, EGFP expression was observed in a random mosaic pattern in the cardiac (C) and skeletal muscle (D) of adult female $Xm^{egfp}Xp$ and $XmXp^{egfp}$ mice (1 and 2 indicate examples of different individual adult females).

Fig. S3. The X-linked EGFP transgene is expressed by glutamatergic neurons in the CNS of X^{egfp} transgenic mice. (**A**) EGFP+NeuN double-labeling in the CNS of male $Xm^{egfp}Y$ mice revealed that the *egfp* transgene expression is confined to a subpopulation of neurons, comprising 30-40% of NeuN+ neurons in the mPFC and POA. Expression was also observed in a subpopulation of ependymal cells. (**B**) EGFP+vGLUT2 double-labeling revealed EGFP expression by a subpopulation of glutamatergic neurons, as shown by images of staining in the mPFC and POA.

Fig. S4. Neurons preferentially express the maternally inherited X chromosome in the mPFC and POA. (**A** and **B**) Pooled Xm (red) and Xp (blue) reads for the neuron specific X-linked genes *Syp, Dcx, Nlgn3, Bex1, Flna, Dlg3, and Syn1* were normalized for strain effects in X-inactivation and revealed that the general neuronal populations of the mPFC (A) and POA (B) preferentially express the Xm. Summed Xm and Xp reads suggest a 21% and 15% Xm bias in the mPFC and POA, respectively. Statistical analysis for Xm versus Xp expression biases for normalized, pooled reads was assessed by a two-sided Fisher's Exact Test for an association between strain and cross ($P<0.0001$**)**. Significant differences in expression between summed Xm and Xp reads were further assessed using $\chi 2$ analysis (***$P<0.0001$).

Fig. S5. UCSC browser tracks for parental expression biases at Mrpl48 and Il18 SNP sites. (A) Six paternally biased SNP sites identified imprinting of Mrpl48 in the female, but not male POA as indicated by UCSC browser tracks of the parental expression bias aligned to the UCSC gene model (paternal: blue, maternal: red, raw expression from genomic alignment: black). (B) An imprinted SNP site in exon 5 of Il18 indicates preferential expression of the maternal allele in the female, but not male mPFC, as shown by UCSC genome browser tracks. Annotated genebank mRNAs suggest multiple Il18 isoforms.

Fig. S6. A prospective cluster of sex specific imprinted genes that includes *Bcdo2, Il18* and *SDHD.* (**A**) Gene models indicate *Bcdo2, Il18, Tex12*, and *SDHD* are adjacent in the mouse (and human) genome. (**B**) Three SNP sites indicate preferential expression of the *Bcdo2* paternal allele in female, but not male mPFC. One SNP is located in exon 1 and two SNPs are in exon 10. One SNP indicates a statistically significant paternal effect in females (**D,** and highlighted in **B**). (**C**) Three SNP sites in the distal region of the 3'UTR of *SDHD* indicate a maternal bias in the male, but not female mPFC. One SNP indicates a statistically significant maternal effect in males (**D,** and highlighted in **C**). Genebank mRNAs suggest alternative 3'UTRs for *SDHD*. Statistical analysis was performed by Fisher's Exact for an association between strain and cross (***P<0.001, **P<0.01). *Il18* results are described in the text.

Table S1. Genes with imprinted SNP sites identified on the X chromosome in females. Imprinted SNP sites are presented and in separate worksheets, all ChrX SNP data is presented for the POA and mPFC.

Table S2. All sex specific imprinted gene SNP sites detected in the POA and mPFC. Includes worksheets for female maternal SNPs, female paternal SNPs, male maternal SNPs, and male paternal SNPs.

Table S3. Full datasets for all genes associated with sex specific imprinted features. For each exon with a statistically imprinted SNP site, an agreement ratio (AR) is calculated for all SNP sites in the exon to determine the percentage that agree with the parental expression bias of the statistically imprinted SNP. Gene lists are separated to present genes with exons that have an AR greater or equal to 75%, genes with exons with an AR below 75%, and genes with exons with a single SNP. See the tabs in the worksheet to find genes with a particular AR for the POA or mPFC for males versus females.

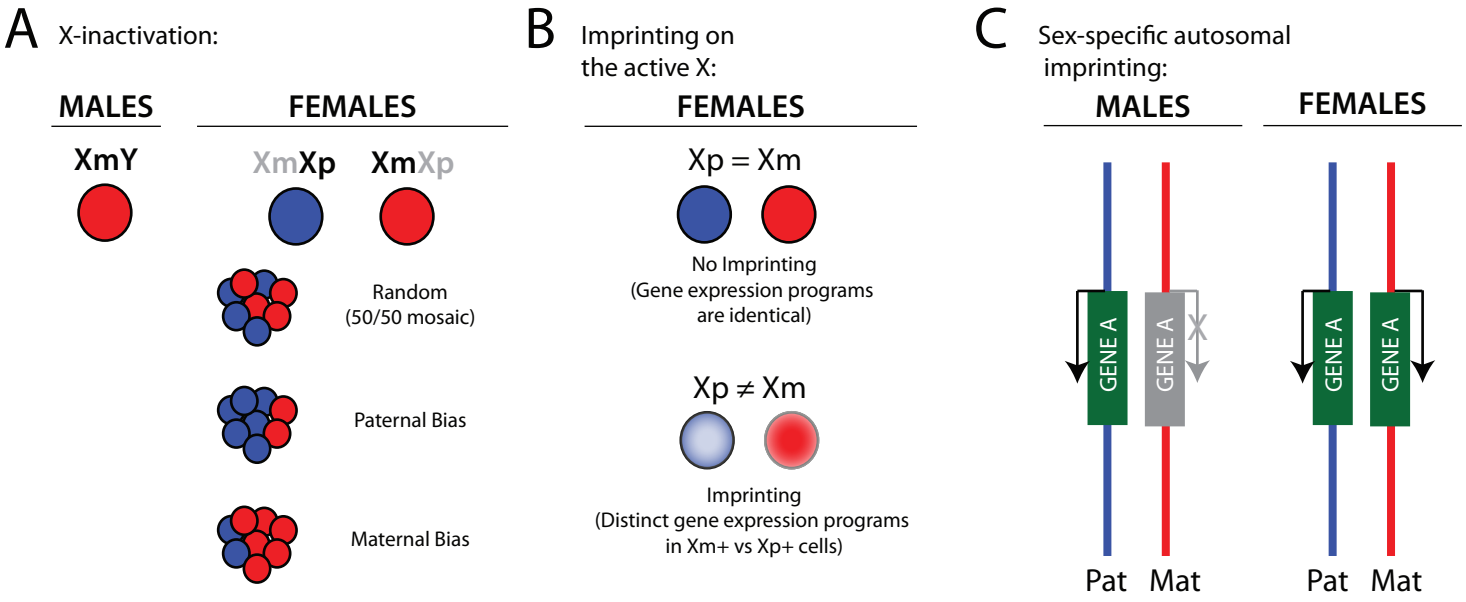
A
X-inactivation:
MALES
FEMALES
XmY
XmXp
XmXp
Random
(50/50 mosaic)
Paternal Bias
Maternal Bias
B
Imprinting on
the active X:
FEMALES
Xp = Xm
No Imprinting
(Gene expression programs
are identical)
Xp ≠ Xm
Imprinting
(Distinct gene expression programs
in Xm+ vs Xp+ cells)
C
Sex-specific autosomal
imprinting:
MALES
FEMALES
GENE A
GENE A
X
GENE A
GENE A
Pat
Mat
Pat
Mat

Fig. S1

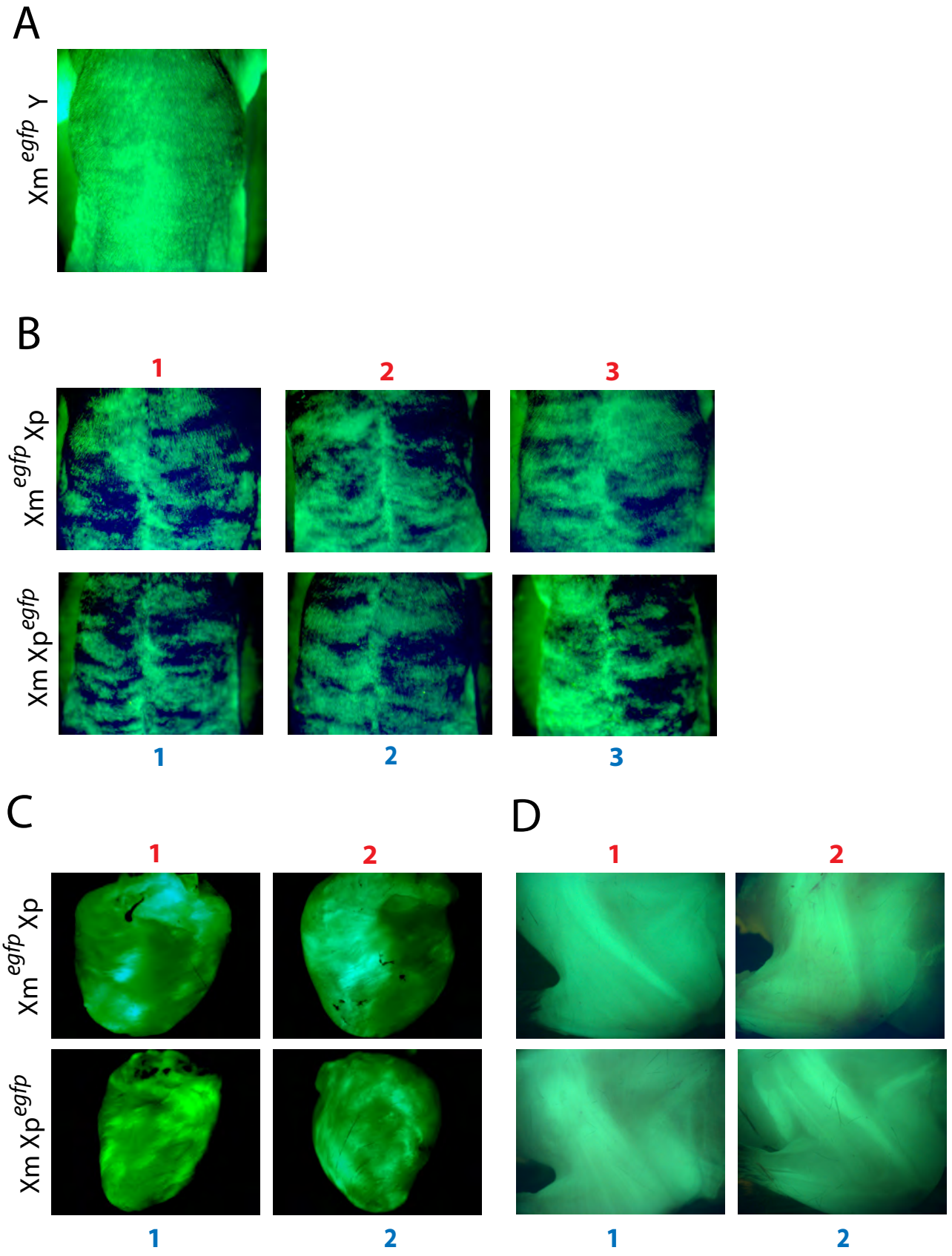
A
Xm egfp Y
B
1
2
3
Xm egfp Xp
Xm Xp egfp
1
2
3
C
D
1
2
Xm egfp Xp
Xm Xp egfp
1
2
1
2
1
2

Fig. S2

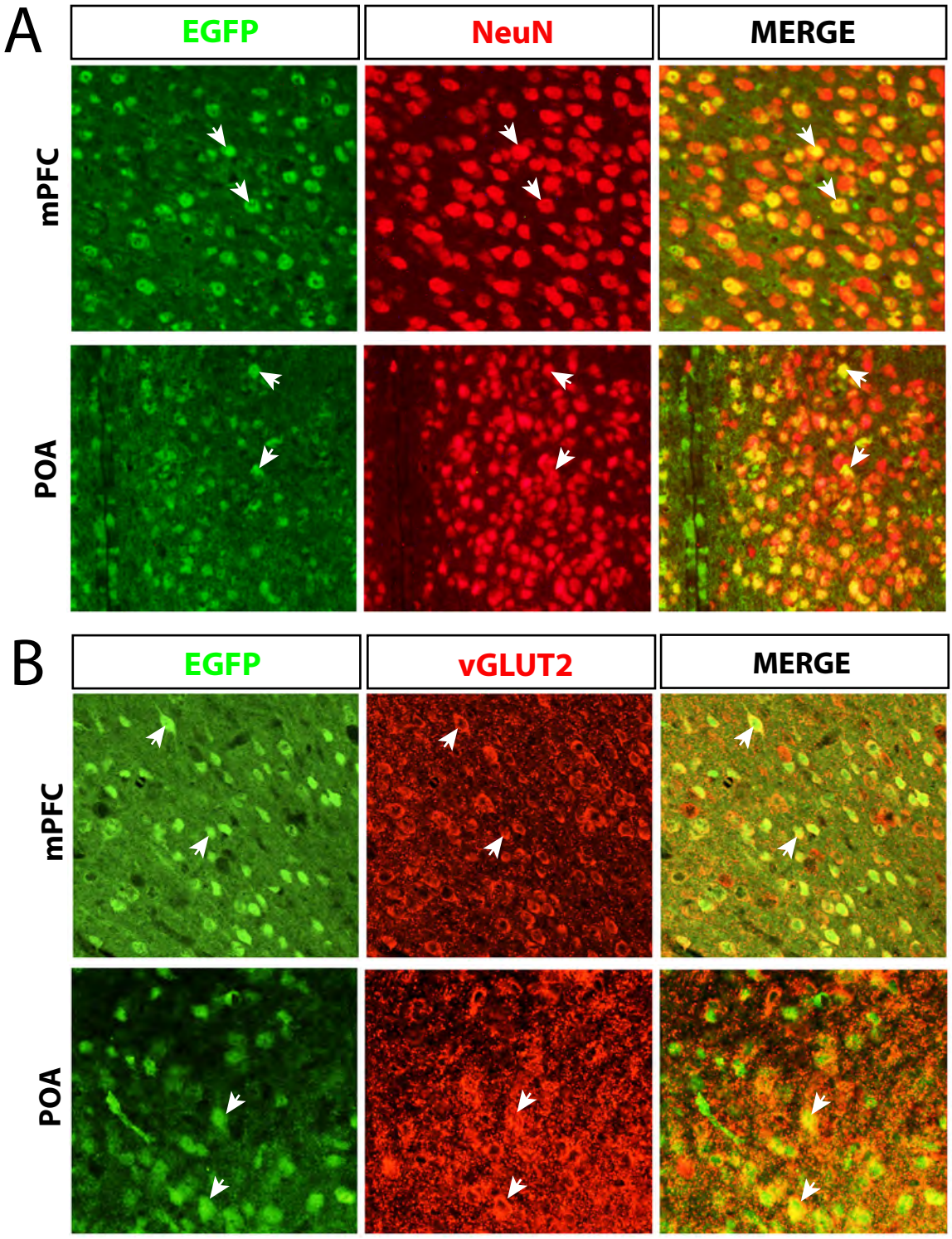

Fig. S3

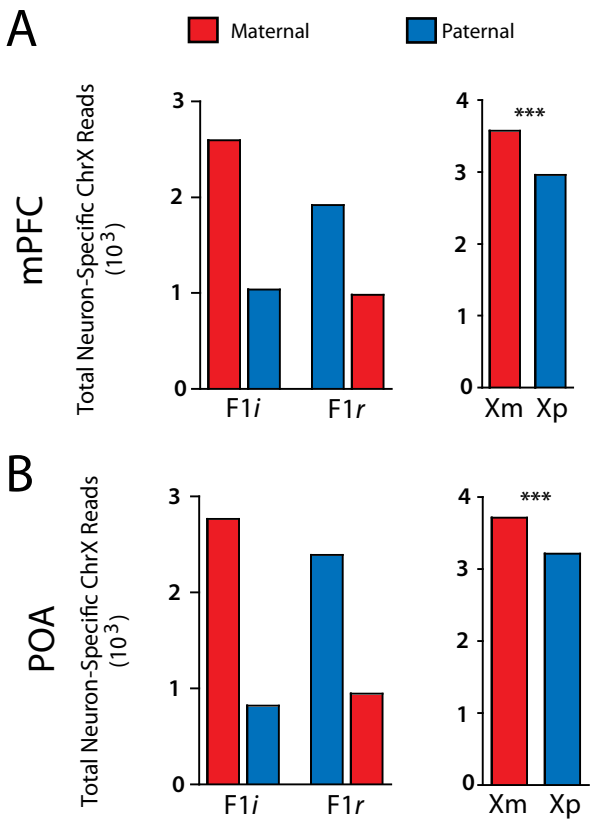
A
Maternal
Paternal
mPFC
Total Neuron-Specific ChrX Reads
(10^3)
3
2
1
0
F1i
F1r
4
3
2
1
0
***
Xm
Xp
B
POA
Total Neuron-Specific ChrX Reads
(10^3)
3
2
1
0
F1i
F1r
4
3
2
1
0
***
Xm
Xp

Fig. S4

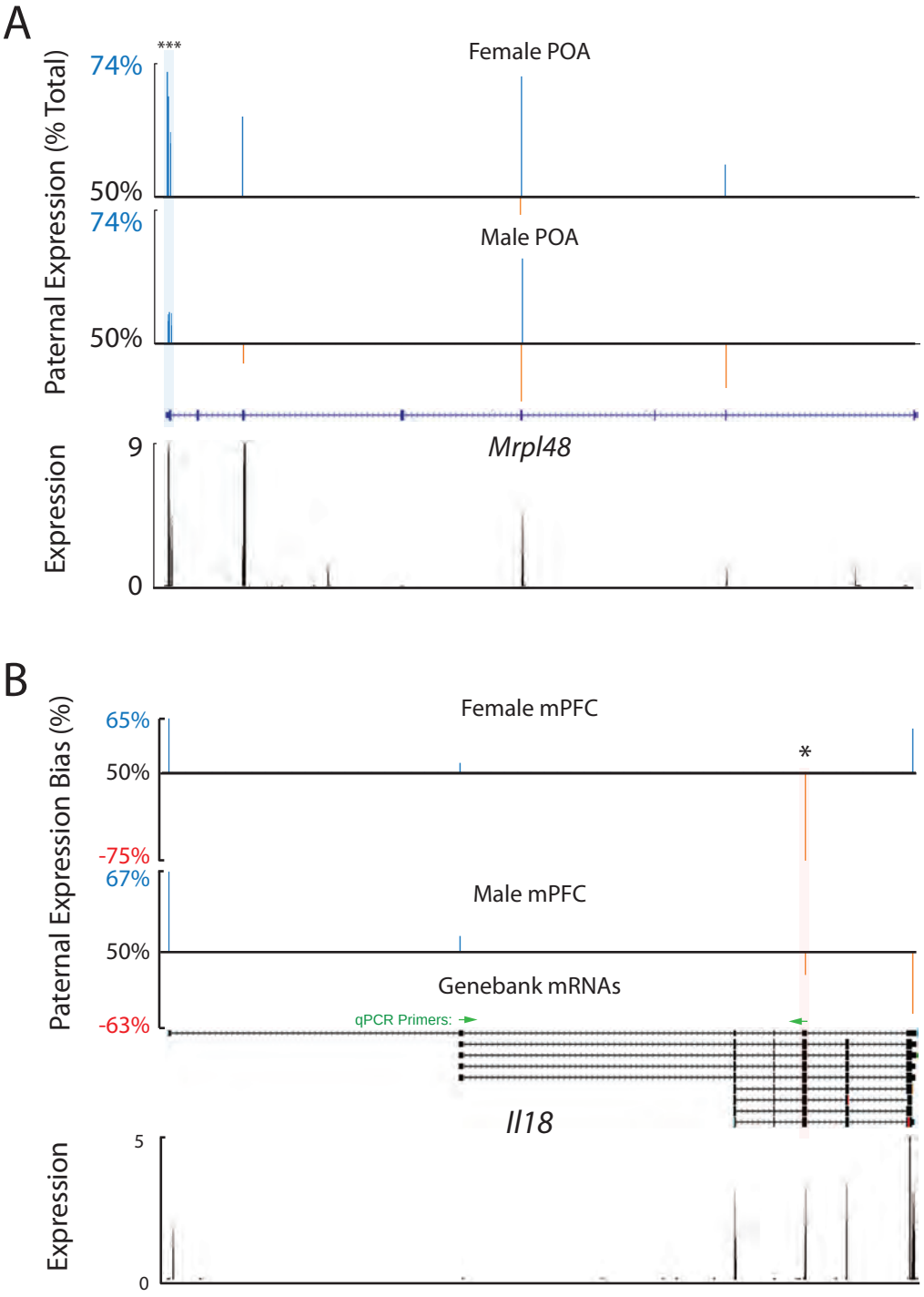
A
Paternal Expression (% Total)
***
74%
50%
Female POA
74%
50%
Male POA
Expression
9
0
Mrpl48
B
Paternal Expression Bias (%)
65%
50%
-75%
Female mPFC
*
67%
50%
Male mPFC
Genebank mRNAs
-63%
qPCR Primers:
Il18
Expression
5
0

Fig S5

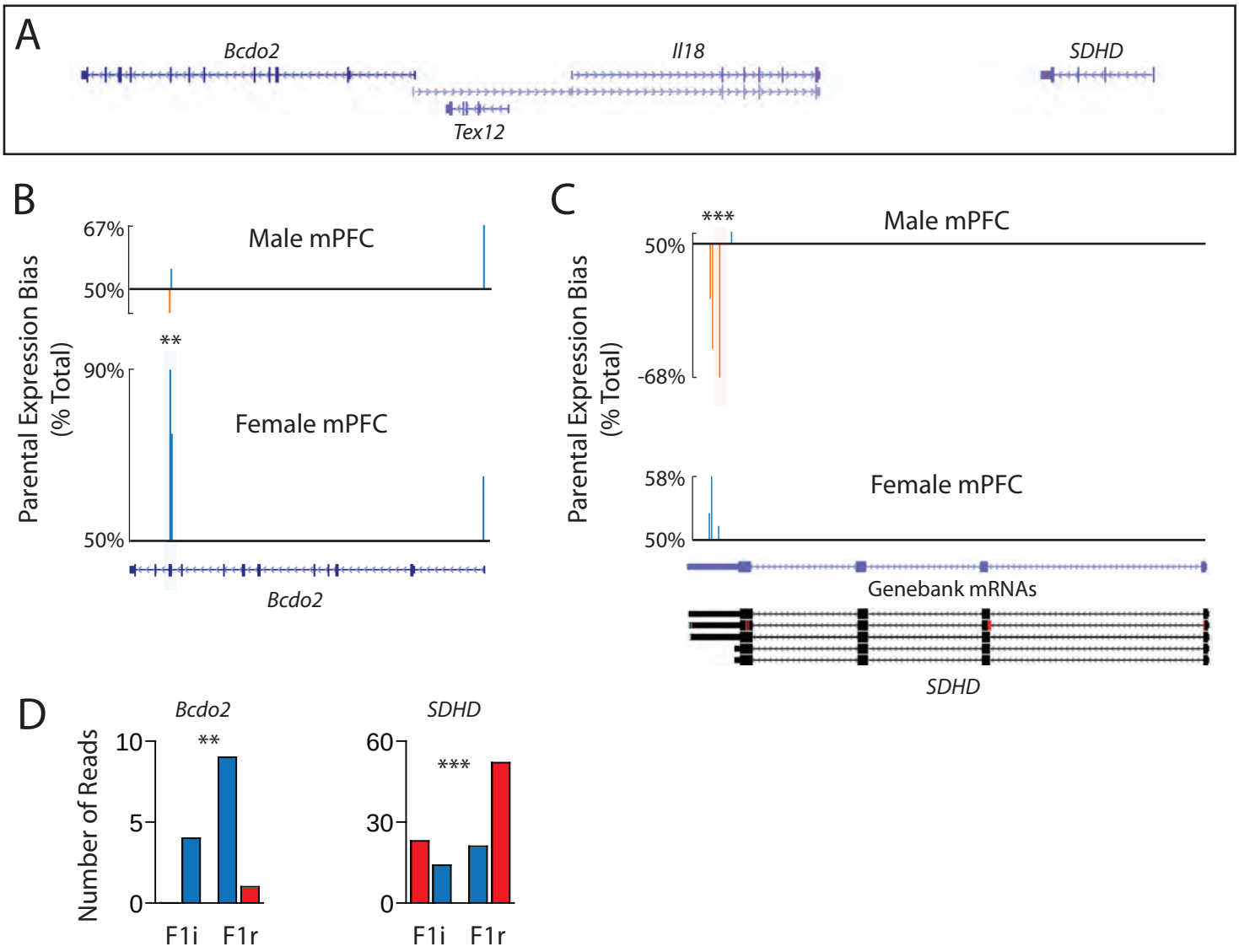
A
Bcdo2
Il18
SDHD
Tex12
B
Parental Expression Bias (% Total)
67%
50%
Male mPFC
**
90%
Female mPFC
50%
Bcdo2
C
Parental Expression Bias (% Total)
***
Male mPFC
50%
-68%
58%
Female mPFC
50%
Genebank mRNAs
SDHD
D
Bcdo2
SDHD
Number of Reads
10
5
0
**
60
30
0
***
F1i
F1r
F1i
F1r

Fig. S6