doi: 10.1038/nature08924

1. Supplementary Figures

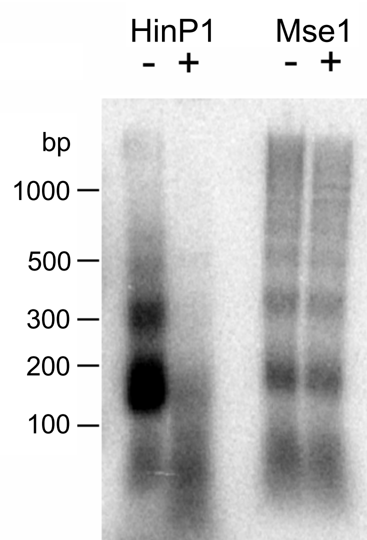


Figure S1.Digestion of mouse brain nuclei with the methyl-CpG sensitive restriction enzyme HinP1 releases predominantly CpG island chromatin. Isolated nuclei were digested with HinP1 or control Mse1 (see Methods) and released chromatin was deproteinised. The resulting DNA was end-labelled with α-^{32}P dCTP using the Klenow fragment of DNA polymerase and digested (+) or mock digested (-) with HpaII. The “nucleosomal ladder” released by HinP1 is collapsed by HpaII indicating the presence of many non-methylated CCGG sites. Chromatin released by MseI, however, is resistant to HpaII digestion as CpGs are deficient and methylated in bulk genomic DNA.

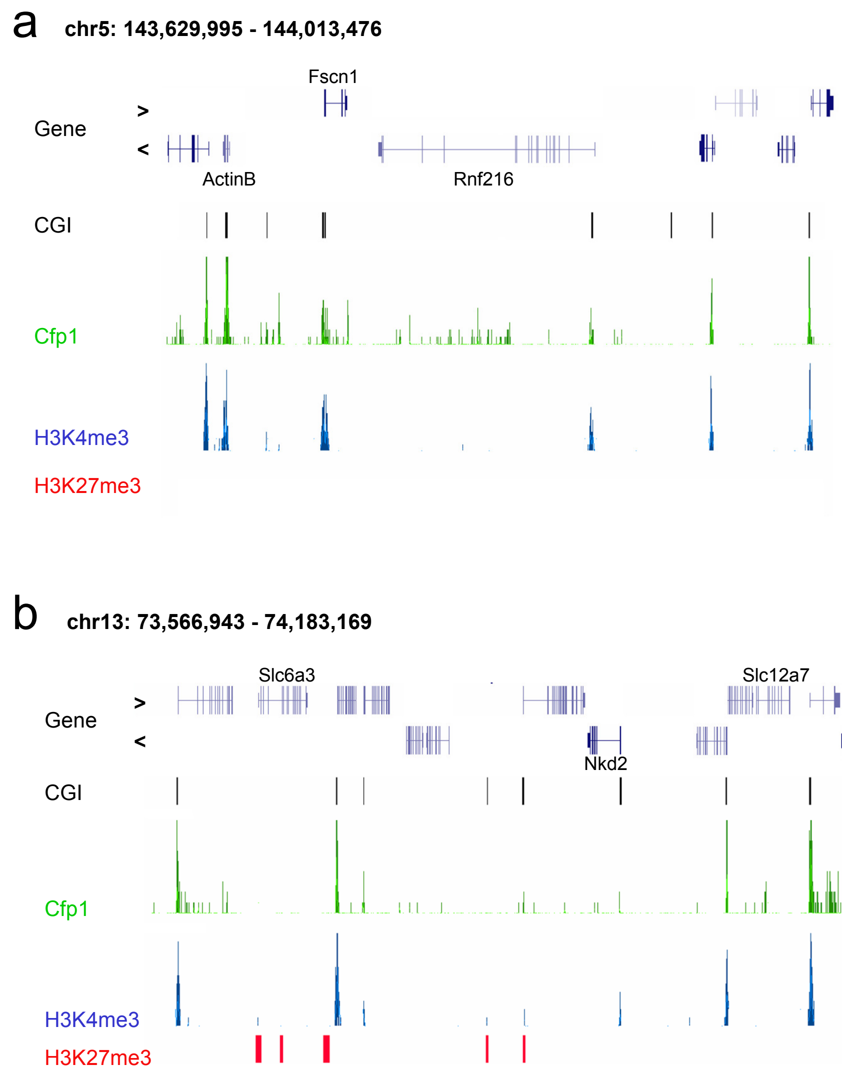
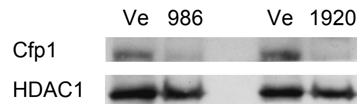


Figure S2. Examples of Cfp1 ChIP-sequence maps of mouse brain aligned with non-methylated CGIs, H3K4me3 and H3K27me3. a, A region of chromosome 5 showing typical coincidence of Cfp1 with CGIs and the H3K4me3 mark. **b,** A region of chromosome 3 showing non-methylated CGIs without detectable Cfp1 binding or H3K4me3, but which are marked H3K27me3 (red signal data of Meissner and co-workers[1] for mouse brain).

1. Meissner, A. *et al*. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454,** 766–770 (2008).

a



b

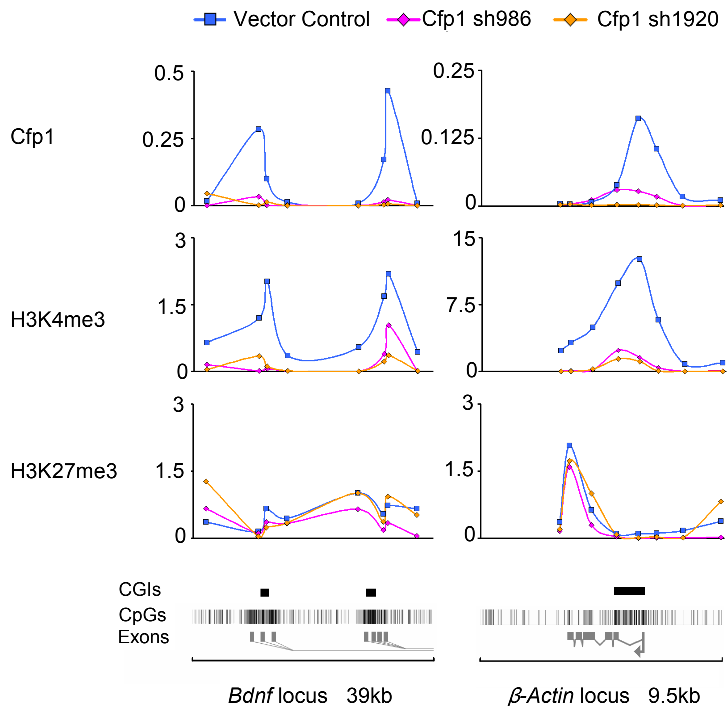


Figure S3. Cells expressing single shRNAs (either sh986 or sh1920) are depleted for Cfp1 and show reduced Cfp1 and H3K4me3 across CGIs. a. Western blot of Cfp1 in NIH3T3 cells stably expressing sh986, sh1920 or empty vector (Ve). Blots from the same gel were probed by an HDAC1 antibody as a loading control. **b.** ChIP using antibodies against Cfp1, H3K4me3 and H3K27me3 assayed by quantitative PCR with primers distributed through the *ß-Actin* and *Bdnf* loci. CGIs, CpGs and exon structure at each locus is shown below the respective plots.

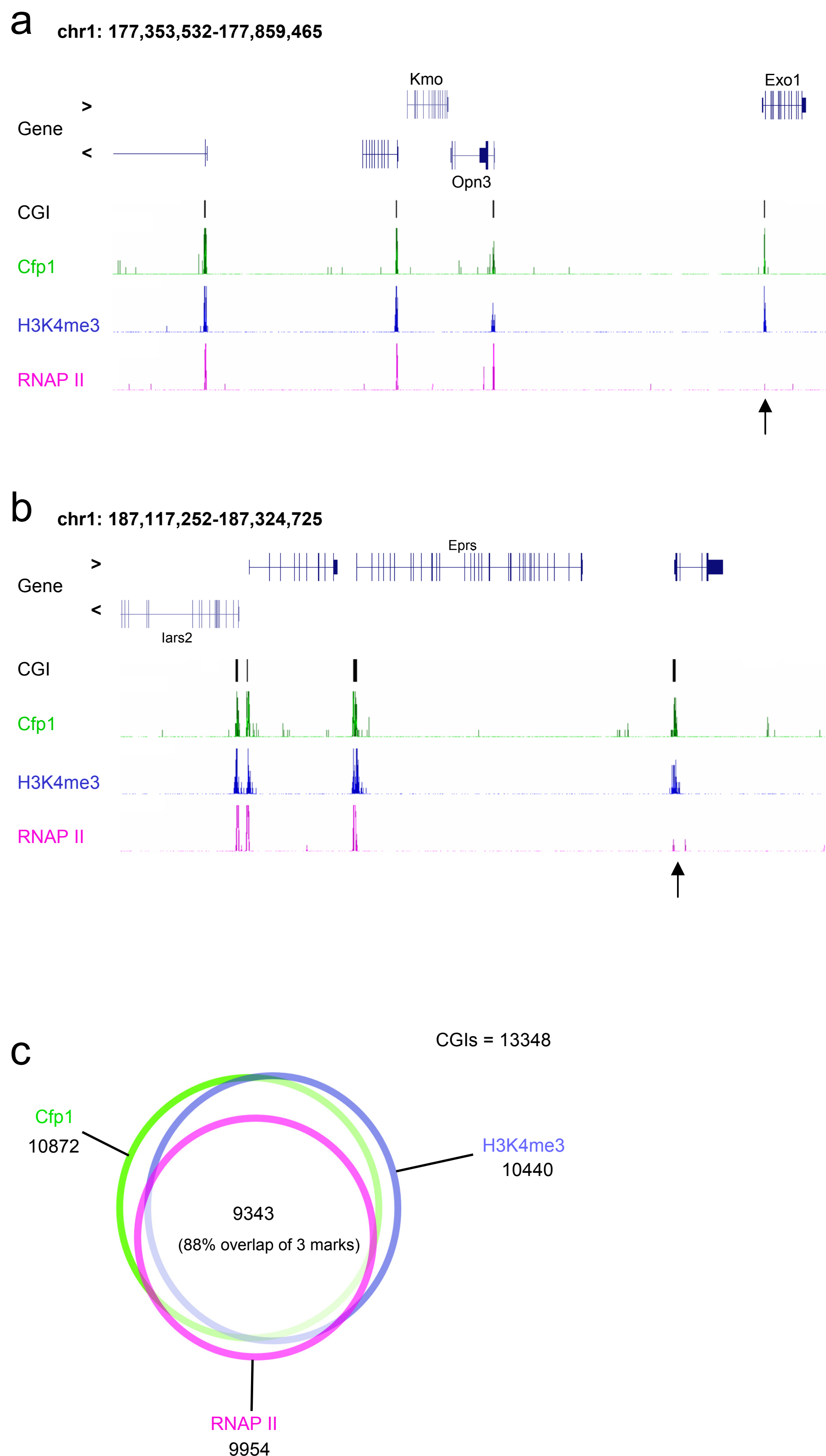


Figure S4. The majority of non-methylated CGIs in mouse brain display coincident Cfp1, H3K4me3 and RNA polymerase II peaks, although a minority (~7%) have Cfp1 and H3K4 without significant RNA polymerase II binding. a &b, In each panel the arrowed CGI has bound Cfp1 and H3K4me3, but lacks detectable RNA polymerase II (RNAPII). **c,** Pronounced overlap of H3K4me3, Cfp1 and RNA polymerase II at non-methylated CGIs.

2. Supplementary Tables

Table S1. Association of Cfp1, histone H3K4me3, H3K27me3 and RNA polymerase II with non-methylated CGIs. Regions of significant ChIP-seq signal were identified using custom bioinformatic tools (see Methods). Co-localisation of these regions was then scored as overlapping (1) or non-overlapping (0) with non-methylated CGIs. All chromosomal coordinates are based on the mouse genome mm9 build. The published H3K27me3 data (extracted from Meissner et al.,2008) was converted from mm8 to mm9 to allow comparison. Data attached as an .xls file

Table S2. Antibodies used in this study

Histone H3	Rabbit polyclonal to Histone H3 (ab1791 – Abcam)
Histone H3K4me3	Rabbit polyclonal to Histone H3K4me3 (ab8580 – Abcam)
Histone H3K4me3	Rabbit polyclonal to Histone H3K4me3 (07-473 – Upstate)
Histone H3K9me3	Rabbit polyclonal to Histone H3K9me3 (ab8898 – Abcam)
Histone H3K27me3	Rabbit polyclonal to Histone H3K27me3 (07-449 – Millipore)
Histone H3K36me3	Rabbit polyclonal to Histone H3 K36me3 (ab9050 – Abcam)
Histone H3 pan acetyl	Rabbit polyclonal to acetyl Histone H3 (06-599 – Upstate)
RNAP II	Mouse monoclonal to unphosphorylated RNA polymerase II (ab817 - Abcam)
CFP1 [westerns / ChIP]	Rabbit polyclonal to CFP1 (kind gift of D. Skalnik)
CFP1 [ChIP]	Rabbit polyclonal to CGBP: H-120 (Sc-25391 – Santa Cruz)

Table S3. Parameters used for Solexa sequence data processing

	Background Subtraction			Peak finding		
Sequence Dataset	Depth of sequence	length[1]	gap[2]	Depth of sequence	length[3]	gap[4]
Cfp1	2	90	20	4	45	600
RNApolII	2	90	20	4	90	250
H3K4me3	2	90	20	3	90	700
H3K27me3	Published in Meissner et al., 2008.					
Non-methylated CGIs	Illingworth et al. (Manuscript in preparation)					

[1]Minimum run of sequence below the stated depth required to maintain read data during background subtraction.
[2]Allowed gap in the run of sequence outlined for length1.
[3]Minimum run of sequence at the stated depth required for identification as an enriched genomic locus.
[4]Individual peaks that are closer together than this value are knitted together as a single locus.