Figure S1, related to Figure 1. Zic2 is largely dispensable for the expression of key stem cell genes. (A) Genome browser track examples of the colocalization of Zic2 with the enhancer marks H3K4me1, H3K27ac and/or p300 at Pou5f1, Nanog, Sox2 and Klf4. The co-bound regions are highlighted with gray bars. (B) Fold expression analyses of Zic2, known Mbd3-NuRD targets Tbx3 and Gpr83, and the pluripotency genes Pou5f1, Nanog, Sox2 and Klf4 in Zic2 knockdown ES cells versus control cells. Error bars indicate s.d. from the analysis of RNA-Seq replicates. (C) Alkaline Phosphatase (AP) staining of the control and Zic2 knockdown ES cells.
Figure S2, related to Figure 2. Interaction of Mbd3-NuRD with Zic2 in ES cells. (A)

The peptide count, uniquely detected peptide count, distributed spectral count and sequence coverage of Zic2 and the subunits of Mbd3-NuRD complex in Zic2 purifications were analyzed by MudPIT. (B) Flag-Zic2 purification followed by Western blotting confirms that the subunits of the Mbd3-NuRD complex can be co-purified with Zic2. (C-D) Western blot analysis of the protein levels of Chd4, Mbd3, Mta2, Zic2 and Sall4 after depletion of Mbd3, Mta2 (C) or Zic2 (D). Knockdown of either Mbd3 or Mta2 leads to reduced levels of Mbd3, Mta2 and Zic2. No effect on protein levels of NuRD subunits was observed upon Zic2 RNAi.
**Figure S3, related to Figure 3. Zic2 and Mbd3-NuRD function together in ES cells.**  
(A-B) Pie charts showing the genome-wide distribution of Chd4 and Mbd3 peaks, respectively, with both proteins predominantly occupying non-TSS regions. (C-D) Validation of our Chd4 (C), Mbd3 (D) ChIP-seq analysis of published data using different Mbd3 antibodies (Hnisz et al., 2013; Whyte et al., 2012). (E) Venn diagram showing the overlap between Zic2, Chd4 and Mbd3 binding sites in ES cells. (F) Fold expression analyses of selected genes after Zic2, Mbd3 or Mta2 RNAi. Error bars indicate s.d. from the analysis of RNA-Seq replicates. (G) RNA-seq track examples of selected genes with expression change in Mbd3, Mta2 and Zic2 RNAi-treated cells. (H) MA plot showing differentially expressed genes after Zic2 knockdown in ES cells. Significantly changed genes are shown in blue. Among these significantly changed genes after Zic2 RNAi, genes up-regulated in both Mbd3 knockdown and knockout datasets are highlighted by a red circle, and genes down-regulated in both Mbd3 knockdown and knockout datasets are highlighted by a green circle. (I) Scatter plot showing differentially expressed genes, as assessed by RNA-seq in Zic2 knockdown versus Mbd3 knockout cells. The genes significantly misregulated are shown in brown. Correlation coefficients are shown. RNA-seq after Zic2 knockdown was performed and analyzed in four replicates. RNA-seq in the Mbd3 knockout cells were analyzed in duplicates (Reynolds et al., 2012a). (J) Expression heatmap for the genes nearest to Zic2, Chd4 and Mbd3 co-bound peaks after Zic2, Mbd3 or Mta2 knockdown. Genes are sorted by decreasing fold change expression in the Zic2 knockdown vs. the control.
Figure S4, related to Figure 4. Zic2 is dispensable for the bulk level of histone marks in ES cells. Western blot analysis of histone marks after depletion of Zic2. No significant changes were observed for any of the tested histone modifications following Zic2 RNAi.
Figure S5, related to Figure 5. Zic2 is essential for the differentiation of ES cells toward neuronal lineage. (A) qRT-PCR analyses of the expression of *Pou5f1*, *Nestin*, *Tubb3* in control cells, which were induced to differentiate towards a neuronal lineage by culturing in N2B27 media for 1-5 days, as indicated. The expression levels were normalized to *Actb*. Error bars represent the s.d. of independent measurements of a representative experiment. (B) Zic2 depletion leads to cell death during neural differentiation of ES cells.
Table S1, related to Figure 3. Gene expression changes in Zic2 and Mbd3 and statistics for NGS datasets.

This table includes sheets for:

1) Co-regulated by Zic2 and Mbd3. This sheet includes those genes differentially expressed (and changed in the same direction) in the Zic2 knockdown, the Mbd3 knockdown, and Mbd3 knockout ES cells.

2) Zic2 shRNA affected genes. This sheet includes all differentially expressed genes after Zic2 knockdown.

3) Mbd3 shRNA affected genes. This sheet includes all differentially expressed genes after Mbd3 knockdown.

In the three sheets described above the following descriptors and their definitions are: Log2FC indicates the log2 fold change in expression in the indicated RNAi (shRNA) or knockout (KO) condition; “PValue”, the p-value as determined by EdgeR for changed expression in the respective condition; “Nearest”, True or False for whether a ChIP-seq peak for the indicated factor is nearer to the gene than any other gene; “at TSS” True or False for whether the indicated ChIP-seq factor was found at an annotated TSS for the gene; “at Enhancer” True or False for whether a non-TSS peak for the indicated factor is nearest to the gene, but this is irrespective of whether this factor is also found at the TSS of this gene.

4) NGS parameters. This sheet includes statistics for all of the NGS data sets that were uploaded to NCBI’s gene expression omnibus (GEO) under accession number GSE61188.

Columns in the NGS parameters sheet include Geo name: the name of the data set as it appears at GEO; Total reads; uniquely mapped reads; uniquely mapped percentage, the percentage of total reads that mapped to a unique region of the genome; non-redundant reads, the number of reads for which there was not another read that was identical in sequence; Redundant rate, the percentage of reads that were identical with one or more other reads.
Supplemental Experimental Procedures

ChIP-Seq Analysis

Sequencing data was acquired through the default Illumina pipeline using Casava v1.8. Reads were aligned to the mouse genome (UCSC mm9) using the Bowtie aligner v0.12.9 allowing uniquely mapping reads only and allowing up to two mismatches (Langmead et al., 2009). Reads were extended to 150 bases toward the interior of the sequenced fragment and normalized to total reads aligned (reads per million; RPM). External sequencing data was acquired from GEO and ArrayExpress as raw reads and aligned in the same way as internally sequenced samples.

Peak detection was performed with MACS v1.4.2 (Zhang et al., 2008). Associated control samples were used to determine statistical enrichment at a p < 1e-8 and FDR < 0.05. The enrichments in external data were determined at p < 1e-5 and FDR < 0.05. The enrichment for H3K27me3 was called using the broad domain peak detector SICER v1.1 (Zang et al., 2009) at the FDR < 1e-8, window size of 200, and gap size of 600.

The overlapping high-confidence peak regions of two biological replicates were used for further analysis. The overlapping peaks were merged into one region as the union of all. Co-bound peaks were called if the peak regions of two samples are overlapping each other. Occupancies were calculated as the mean coverage under each peak region. R package edgeR 3.0.8 was used to perform differential occupancy analysis at p-value < 1e-5.
For ChIP-Seq enrichment profiles, regions of interest are shown for each factor as a binary value of enriched/not enriched and rows were sorted by the shortest distance of peak center to an annotated TSS. Regions shown are oriented 5’ to 3’ corresponding to the orientation of the nearest gene. Regions spanning 50 kb on either side of the indicated feature were binned into 200 bp windows. Each line represents a peak (peak centered).

For ChIP-Seq heatmaps, regions of interest were centered at the center of each peak and sorted by the shortest distance of peak center to an annotated TSS. Regions shown are oriented 5’ to 3’ corresponding to the orientation of the nearest gene. Regions spanning 5 kb on either side of the indicated feature were binned into 25 bp windows.

Gene annotations and transcript start site information were from Ensembl 67 utilizing only Refseq mRNA entries. Non-TSS peaks are the peaks not overlapping with any TSS of the gene annotations. GO analysis was performed with the gene list using DAVID (Huang da et al., 2009a, b).

**Total RNA-Seq Analysis**

Sequencing data was acquired through the default Illumina pipeline using Casava v1.8. Reads were aligned to the mouse genome UCSC mm9 and to gene annotations from Ensembl 67 using TopHat v2.0.9 (Trapnell et al., 2009), using option –g 1 and allowing up to two mismatches. R package edgeR 3.0.8 was used to perform differential expression analysis at p-value < 1e-5 (Robinson et al., 2010) and p-value < 0.05 for external RNA-seq data.
Data generated for this study includes ChIP-seq data for Zic2, Zic3, Chd4, Mbd3, H3K27me3, H3K4me1, H3K4me3, H3K27ac and Pol II, and the RNA-seq data. Other data sets come from previously published studies: p300, H3K4me1 and H3K27ac ChIP-seq data are from GEO accession number GSE24164 (Creighton et al., 2010); H3K4me3 and H3K27me3 ChIP-seq data are from GEO accession number GSE12241 (Mikkelsen et al., 2007); H3K27me3 ChIP-seq in Mbd3 WT and KO ES cells are from the ArrayExpress repository under accession E-MTAB-888 (Reynolds et al., 2012b); RNA-seq in Mbd3 WT and KO ES cells are from the ArrayExpress repository under accession E-MTAB-997 (Reynolds et al., 2012a); Chd4 ChIP-seq is from GEO accession number GSE27841(Whyte et al., 2012); and Mbd3 ChIP-seq is from GEO accession number GSM1246867 (Hnisz et al., 2013).

The parameters for next generation sequencing are reported in Table S1.
Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America 107, 21931-21936.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell 155, 934-947.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37, 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44-57.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10, R25.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448, 553-560.

Reynolds, N., Latos, P., Hynes-Allen, A., Loos, R., Leaford, D., O'Shaughnessy, A., Mosaku, O., Signolet, J., Brennecke, P., Kalkan, T., et al. (2012a). NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. Cell Stem Cell 10, 583-594.

Reynolds, N., Salmon-Divon, M., Dvinge, H., Hynes-Allen, A., Balasooriya, G., Leaford, D., Behrens, A., Bertone, P., and Hendrich, B. (2012b). NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. EMBO J 31, 593-605.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105-1111.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature 482, 221-225.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics 25, 1952-1958.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137.