Supporting Information

Ho et al. 10.1073/pnas.0812888106

SI Methods

Sequences of shRNA hairpins:

Brg shRNA #1: ccggcggctcaagaaggaagttgaactcgagttcaacttccttcttgacgnttttg (TRCN0000071383; Open Biosystems).

Brg shRNA #2: ccggcgcccgacacattattgagaactcgagttctcaataatgtgtcgggcgtttttg (TRCN0000071384; Open Biosystems).

BAF155 shRNA: ccggcctgaaatatacttggcatatctcgagatatgccaagtatatttcaggttttttg (TRCN0000071388; Open Biosystems).

shGFP Control: (Cat. no. RHS4459; Open Biosystems).

pLKO empty vector control: (Cat. no. RHS4080; Open Biosystems).

Culture of ES Cells. E14 ES cells were cultured in Knockout Dulbecco's Modified Eagle's Medium (Cat. no. 10829; GIBCO) supplemented with 15% ES-qualified FBS (Cat. no. 16141-079; Invitrogen), 2 mM L-glutamine (Cat. no. 25030-081; GIBCO), 10 mM Hepes (Cat. no. 15630-080; GIBCO), 100 units/mL penicillin/streptomycin (Cat. no. 15070-063; GIBCO), 0.1 mM nonessential amino acids (Cat. no. 11140-050GIBCO), 0.1 mM 2-mercaptoethanol (Cat. no. 21985-023; GIBCO) and 1,000 units/mL LIF (Esgro, Cat. no. ESG1107; Chemicon). ES cells were maintained at 37 °C, 7% carbon dioxide, fed with fresh media daily, and passaged onto new plates after trypsin dissociation.

Genome-Wide ChIP-Seq and Microarray Data Analysis. Sequenced 25-bp reads were mapped to the mouse genome (mm8 assembly) using the Solexa Analysis Pipeline. Only those ≈12.2 million reads that mapped to unique genomic locations were retained for further analysis. The average length F of the sequenced DNA fragments was estimated to be 229 bp [from the mapped reads (tags)] using the algorithm presented in Jothi et al. (1). The tags were then assigned to the centers of the sequenced DNA fragments, to which they correspond, by shifting the tag positions by $+F/2$ and $-F/2$ bp for those that map to the positive and negative strands, respectively. Using the TIROE (Tool to Identify Regions of Enrichment) algorithm developed by Jothi and colleagues (unpublished), we defined Brg binding regions as genomic regions enriched with Brg tags. TIROE uses a clustering framework to cluster tags that are within $F + G$ bp of each other, where G (set to F) is the number of base pairs that accounts for the region with no tags that may otherwise be mapped with tags if the sample was sequenced to saturation. Each cluster is assigned a P value using a random background model based on Poisson probabilities. A stringent P value cutoff (10^{-10}) was used to identify 10,559 Brg-binding regions. Genome-wide distribution of Brg binding regions was determined in reference to RefSeq genes downloaded from the UCSC genome browser (2).

The 17,030 UCSC known genes with expression data in ES cells, obtained from a published microarray data of wild-type E14 ES cells (3), were grouped into 10 bins based on expression. To generate the gene body profile of tag density distribution, all of the genes in each bin were aligned relative to their TSS and TES and the tag density within the gene body was calculated by using windows, whose length is 5% of the gene length. The 5-kb region upstream of TSS was profiled using 5 1-kb windows. The tag density near TSS (±5 kb) was profiled by using 100-bp windows after aligning the genes in each bin relative to their TSS. The relationship between expression and Brg occupancy of a gene was assessed by determining the fraction of genes in each bin that was bound by Brg. A gene was defined to be bound by Brg if 1 or more Brg binding regions are present in the 5-kb region upstream of TSS or in the gene body. Summary windows, displaying the number of tags in 400-bp windows, were used for viewing the ChIP-Seq data on the UCSC genome browser, and to generate screenshots.

To study Brg's role in the expression of genes in ES cells and differentiated cells, we first computed the fold change for 17,030 known genes in Brg shRNA compared with a GFPshRNA control. To avoid including genes with artificial fold changes, we removed ≈15% of the genes with low expression levels in both the test and the control samples. The resulting genes were then grouped into 10 bins based on the fold change. Using the same strategy, we also grouped genes into 10 bins based on the fold change in embryoid body (EB) cells compared with ES cells (4), Oct4 RNAi compared with a parental vector pSuper-puro control (3), and Sox2-KO (72 h) compared with a Sox2-KO (0 h) (5). The relationship between 2 datasets (10-bin sets) was assessed by performing pairwise comparisons of bins from the 2 datasets, and counting the number of genes in common between every pair of bins. The resultant matrix containing the number of common genes was plotted as a heat map to infer the interplay between the functional roles of Brg, Oct4, and Sox2.

Previously published ChIP-Seq datasets for Oct4, Sox2, Nanog, STAT3, and Smad1 (6) and Suz12 (7) were used to determine Brg's cooccupancy with these factors. A gene unit, defined as gene body plus 5 kb upstream of the TSS of an annotated gene, is said to be cooccupied by Brg and a factor if they both bind anywhere on it. Alternatively, we defined cooccupancy as strict physical colocalization, i.e., the binding regions of Brg and factors must overlap within the gene unit. Statistical significance of overlap was calculated by using a hypergeometric distribution calculator available at http://elegans.uky.edu/MA/progs/overlap_stats.html

1. Jothi R, Cuddapah S, Barski A, Cui K, Zhao K (2008) Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Aacids Res* 36(16):5221–5231.
2. Karolchik D, et al. (2008) The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Res* 36:D773–D779.
3. Loh Y, et al. (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 38(4):431–440.
4. Sene K, et al. (2007) Gene function in early mouse embryonic stem cell differentiation. *BMC Genomics* 8(1):85.
5. Masui S, et al. (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* 9(6):625–635.
6. Chen X, et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133(6):1106–1117.
7. Marson A, et al. (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134(3):521–533.
8. Novak A, Guo C, Yang W, Nagy A, Lobe CG (2000) Z/EG, a double reporter mouse line that expresses enhanced green fluorescent protein upon Cre-mediated excision. *Genesis* 28(3–4):147–155.

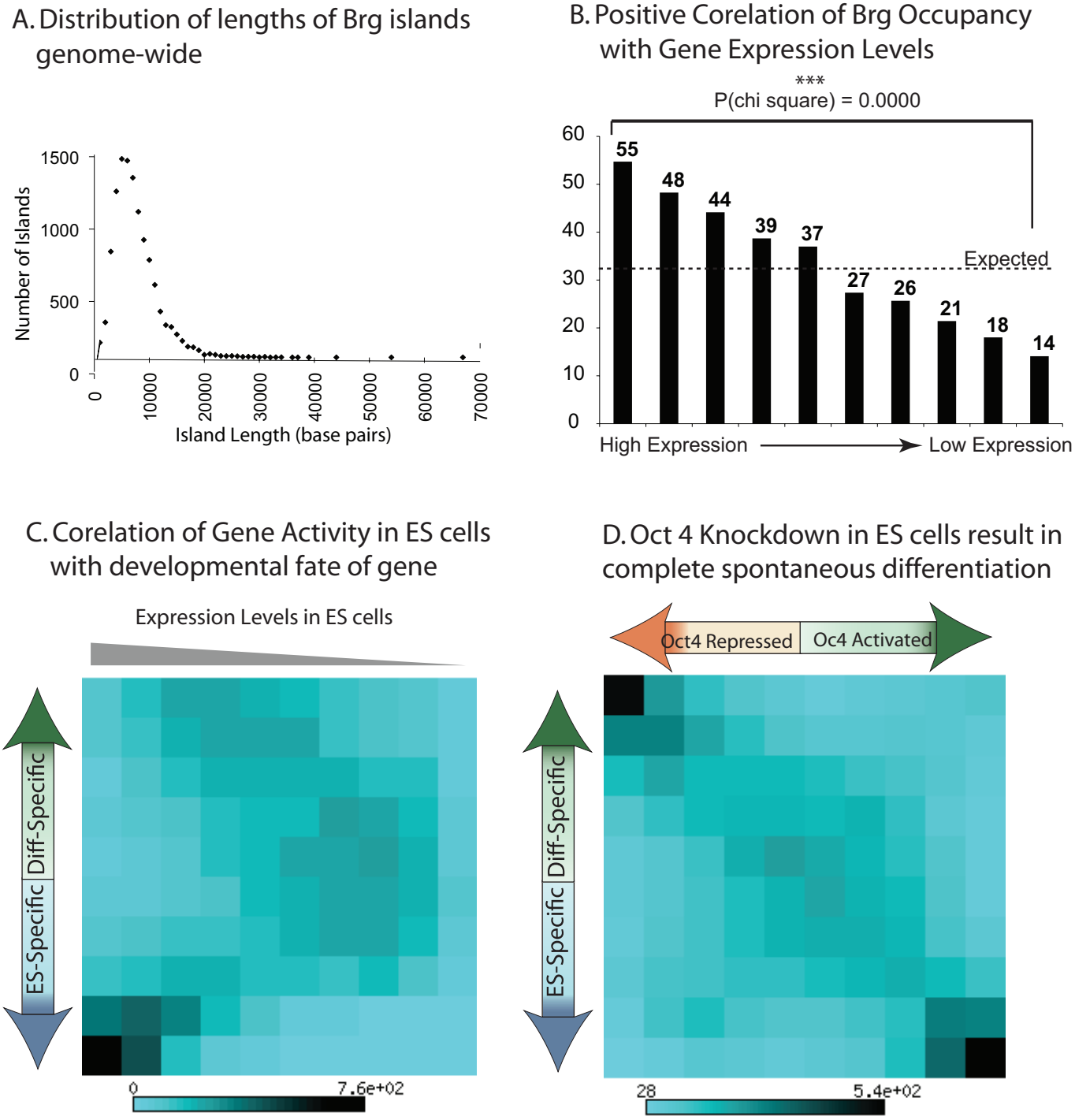


Fig. S1. ChIP-Seq Analysis of Brg-bound Regions in ES cells. (*A*) Frequency of observed lengths of Brg-enriched regions. (*B*) Genes (n = 17,030 for which expression data were available) were grouped into 10 bins from highest expression to lowest expression. The percentage of genes in each group that had Brg-enriched regions was plotted. Chi-square test confirms that group with the highest expression was significantly enriched over the group with the lowest expression. (*C*) Matrix of all genes in ES cells grouped according to their expression in ES cells (*x* axis) against their developmental fate (day 14 embryoid body/ES fold change on the *y* axis representing ES or differentiation-specific). (*D*) Corelation of gene expression after Oct4 knockdown in ES cells against their normal developmental fate. For both *C* and *D*, intensity in matrix reflects the number of genes falling in each square, as measured by color bars below the matrices.

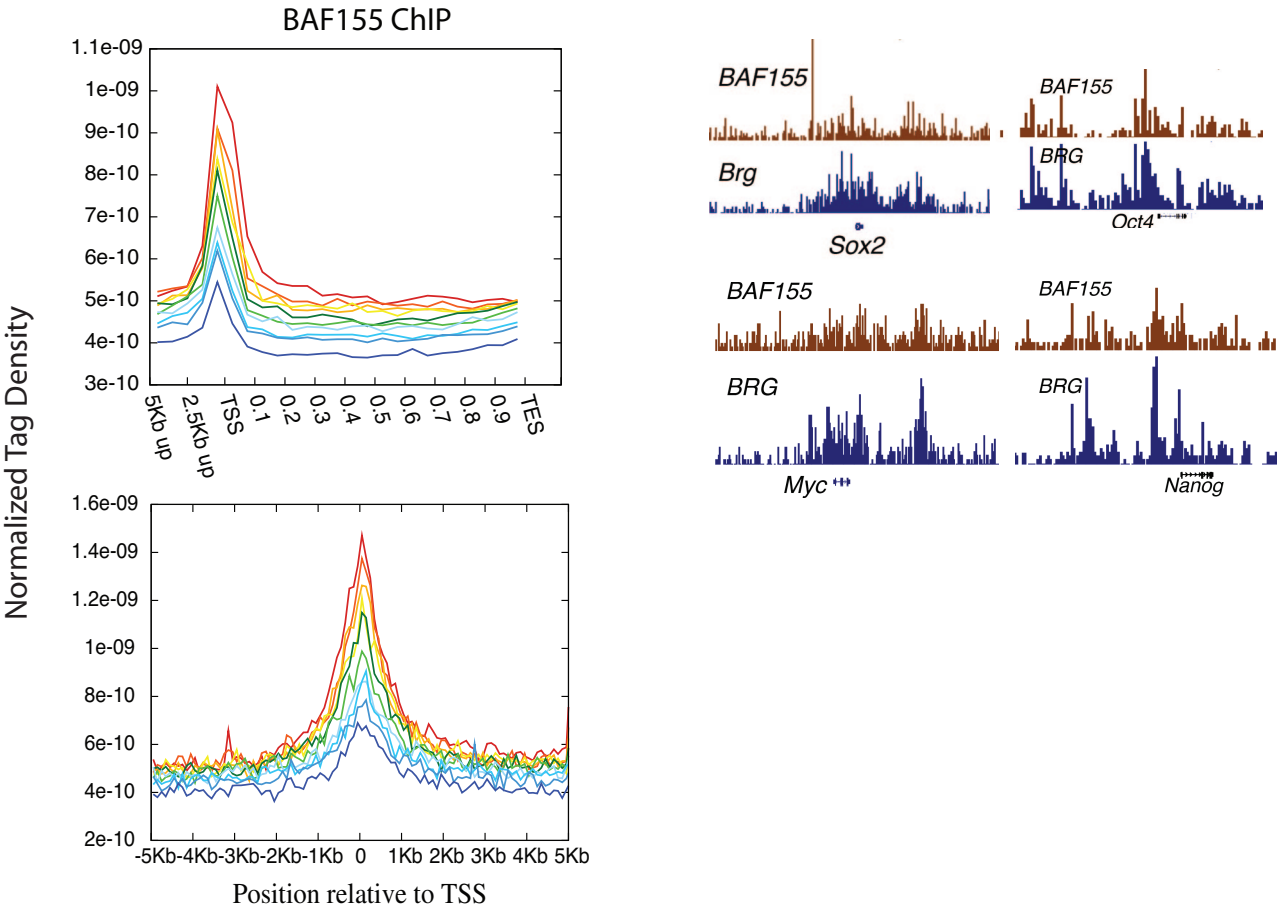
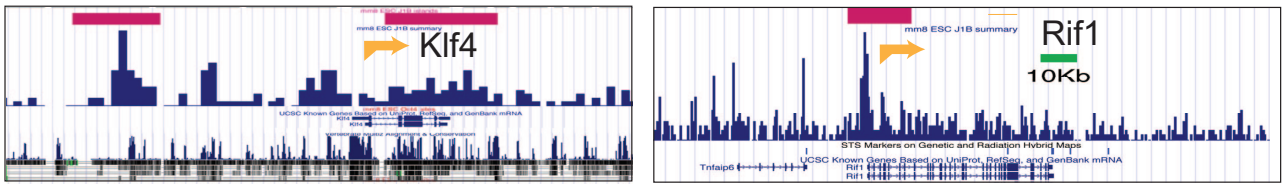


Fig. S2. ChIP-Seq Analysis of BAF155 in ES cells. (*A*) Average distribution of Brg BAF155 tag density across a gene unit (*Upper*). Higher-resolution analysis of average tag density surrounding the TSS (*Lower*). In each plot, genes were classified into 10 groups based on expression levels in ES cells (highest to lowest represented by colored lines), and average tag density across the gene unit was plotted for each group. (*B*) Brg and BAF155 display similar binding profiles on target genes. The Sox2, Oct4, Nanog, and Myc loci are shown here as examples.

A.



B.

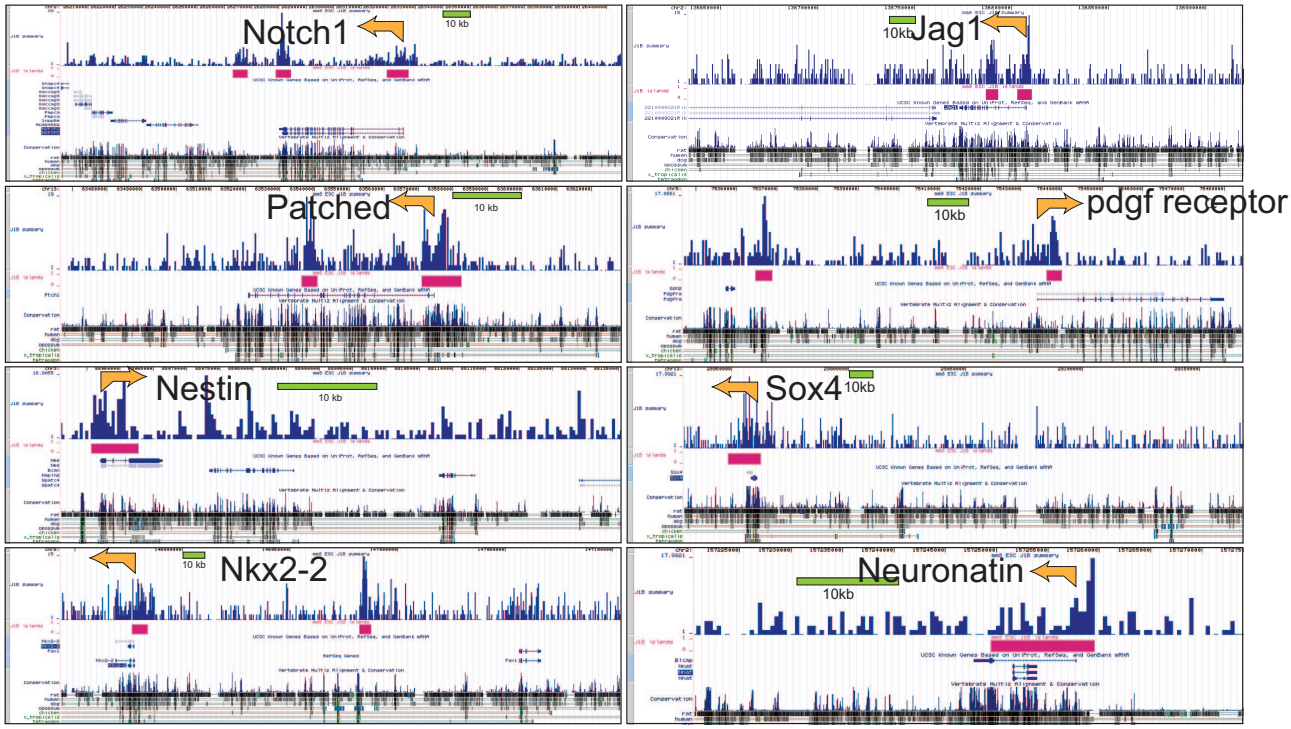


Fig. S3. Examples of Brg occupancy at ES cell genes and neuronal development-related genes. Brg occupancy at Klf4 and Rif1 (*A*) and neural progenitor genes (*B*). Snapshots from UCSC genome browser of Brg-enriched regions ($P < 10E-10$; pink squares).

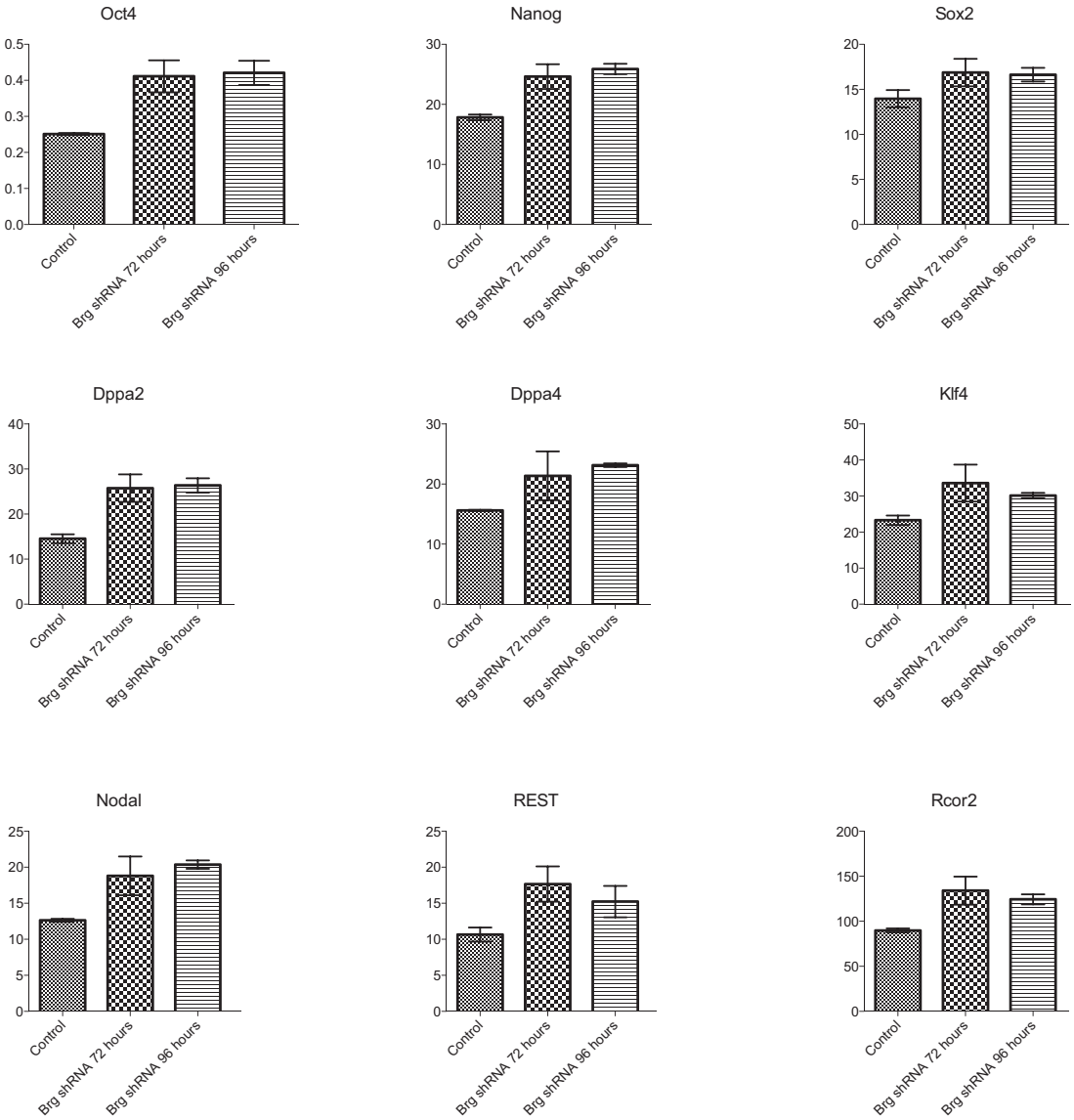


Fig. S4. QPCR verification of ES-specific genes after Brg shRNA at 2 time points compared with control pLKO vector.

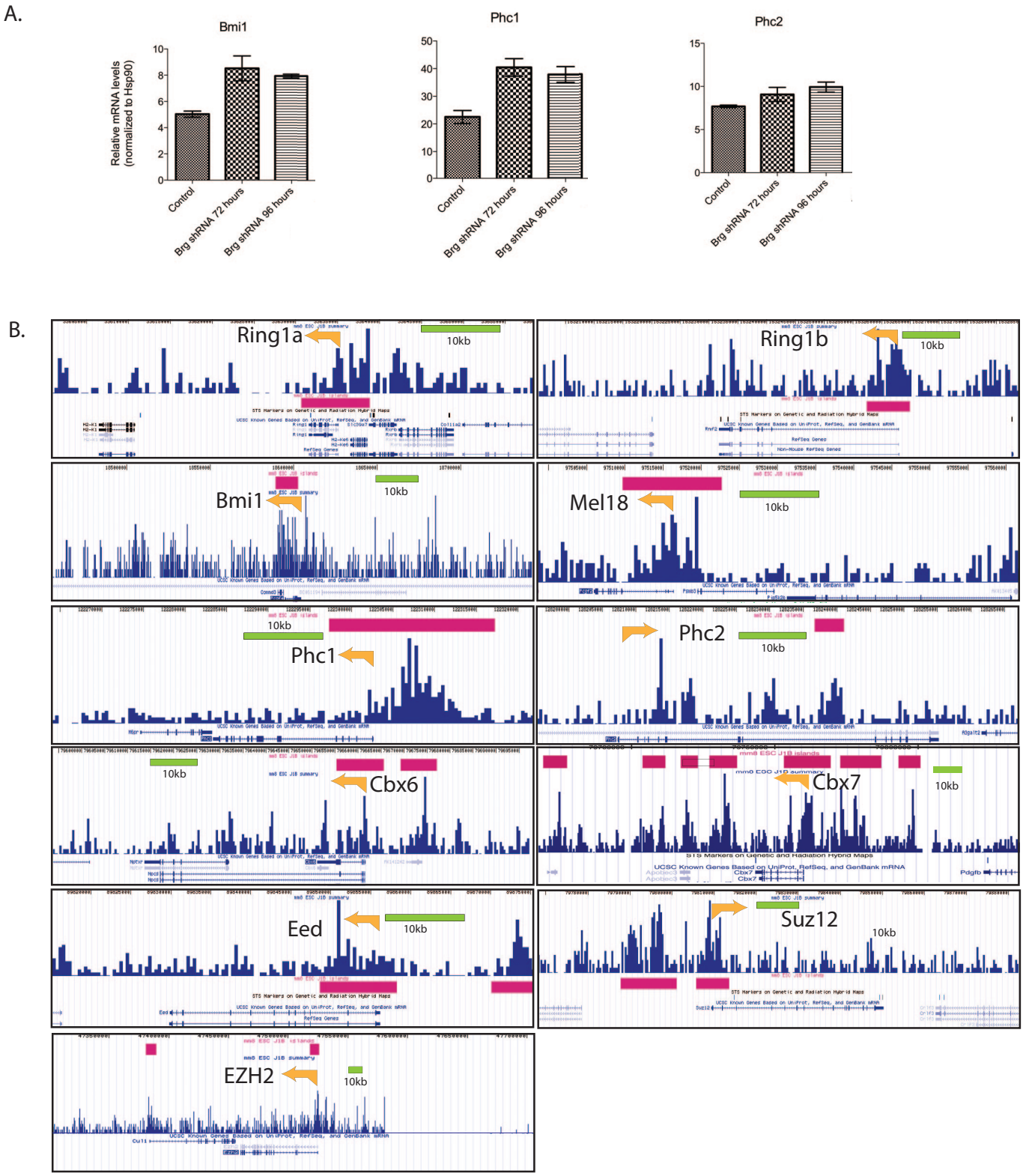


Fig. S5. (*A*) QPCR analysis of PRC1 subunit transcript levels upon Brg knockdown. (*B*) Snapshots from UCSC genome browser of Brg-enriched regions (pink squares). Brg binds to genes encoding components of Polycomb Repressive Complex 2 (*Eed*, *Suz12*, and *Ezh2*) and Polycomb Repressive Complex 1 (*Ring1a*,*b*, *Bmi1*, *Mel18*, *Phc1*,*2*, *Cbx6*,*7*).