Supplemental Material to:

Kirthi Pulakanti, Luca Pinello, Cary Stelloh, Steven Blinka, Jeremy Allred, Samuel Milanovich, Sid Kiblawi, Jonathan Peterson, Alexander Wang, Guo-Cheng Yuan, and Sridhar Rao

Enhancer transcribed RNAs arise from hypomethylated, Tet-occupied genomic regions

Epigenetics 2013; 8(12)
http://dx.doi.org/10.4161/epi.26597

http://www.landesbioscience.com/journals/epigenetics/article/26597

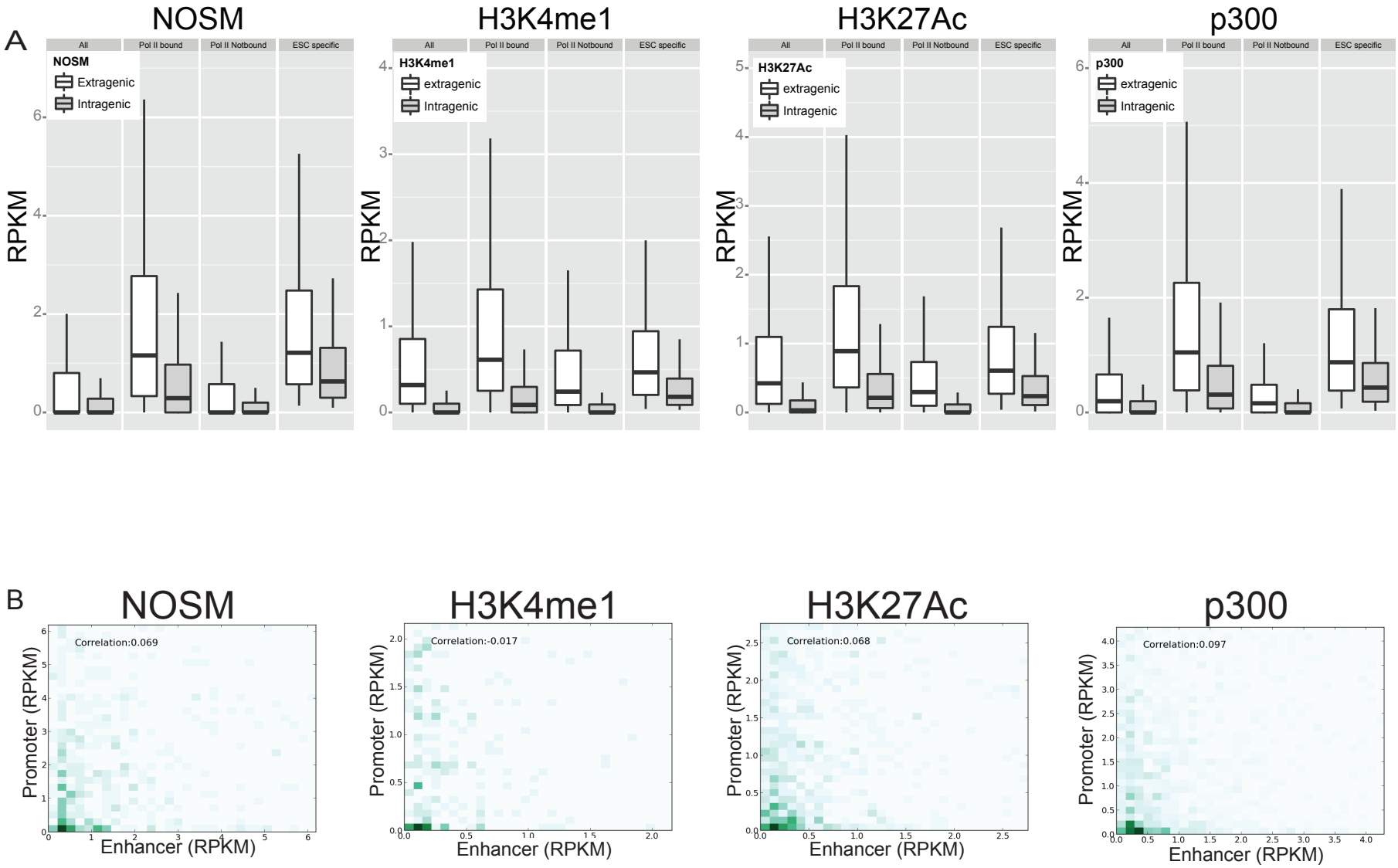
http://www.landesbioscience.com/journals/epigenetics/article/26597/2013EPI0305R_TableS1.xlsx

http://www.landesbioscience.com/journals/epigenetics/article/26597/2013EPI0305R_TableS2.xlsx

http://www.landesbioscience.com/journals/epigenetics/article/26597/2013EPI0305R_TableS3.xlsx

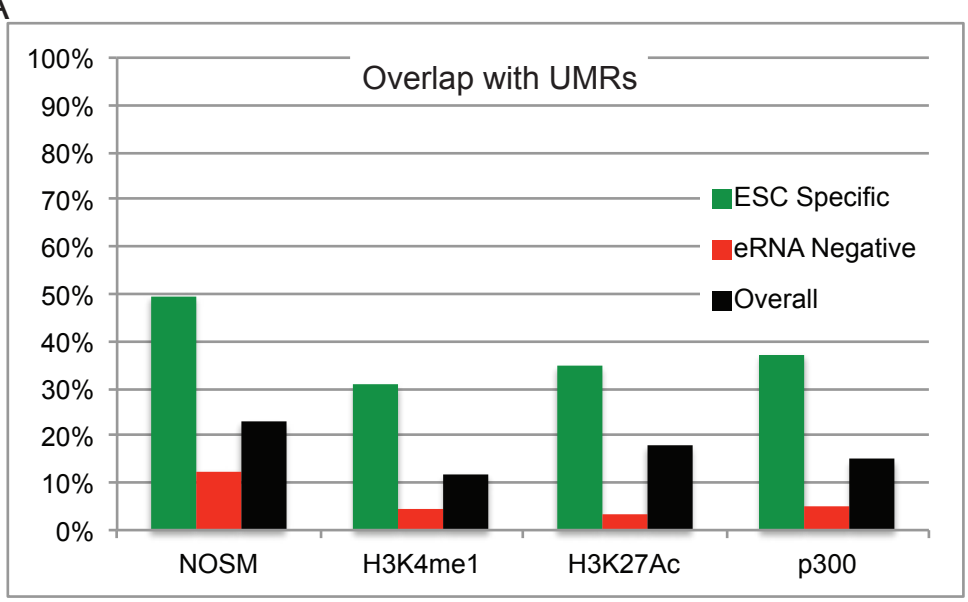
http://www.landesbioscience.com/journals/epigenetics/article/26597 2013EPI0305R_TableS4.xlsx

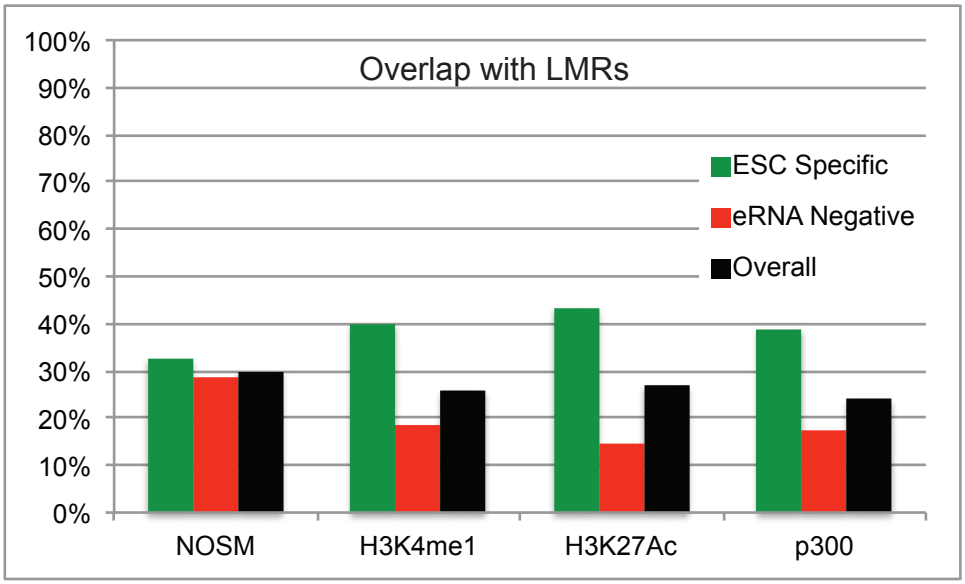
Supplemental Figure 1



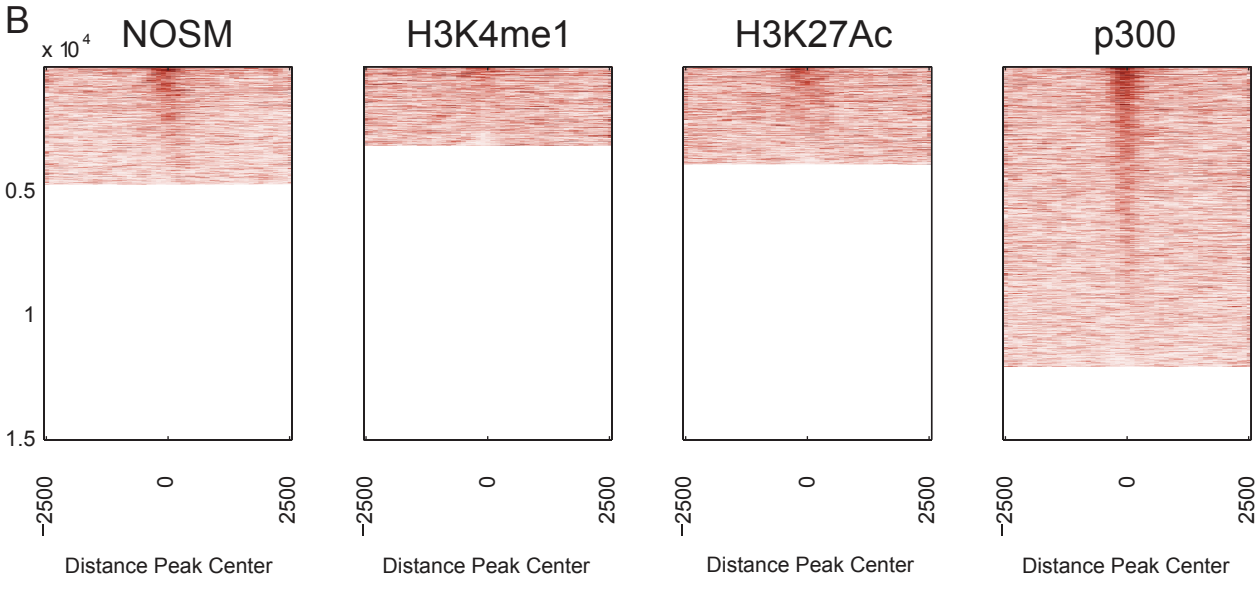
Supplemental Figure 2

A



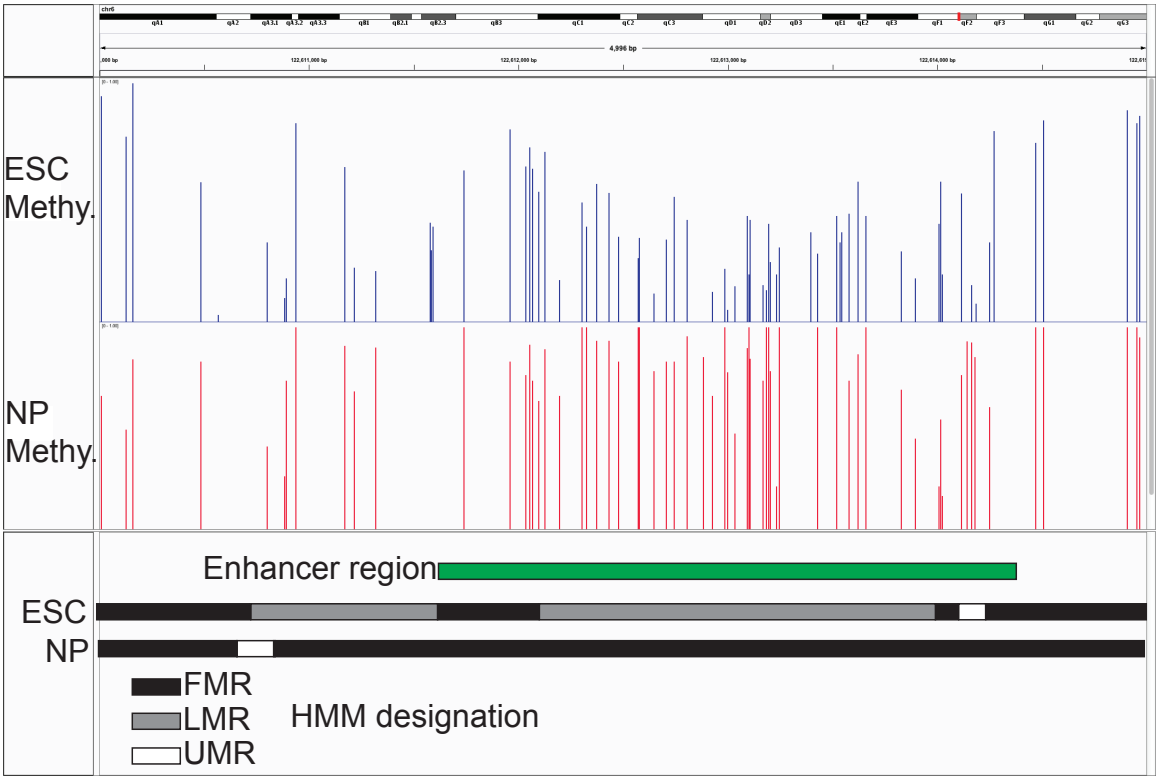


B

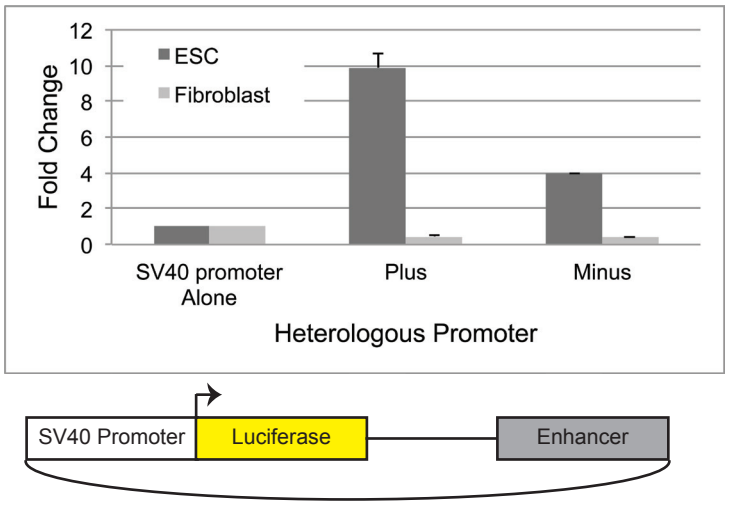


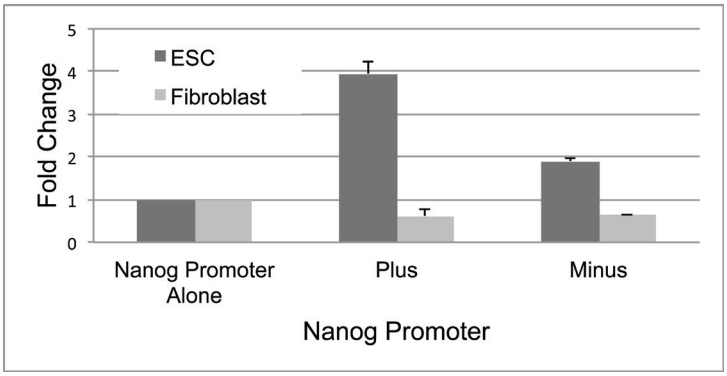
Supplemental Figure 3

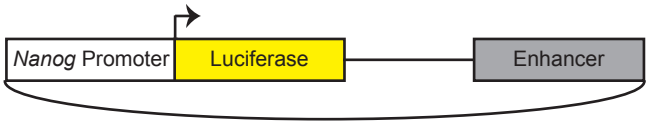
A



B

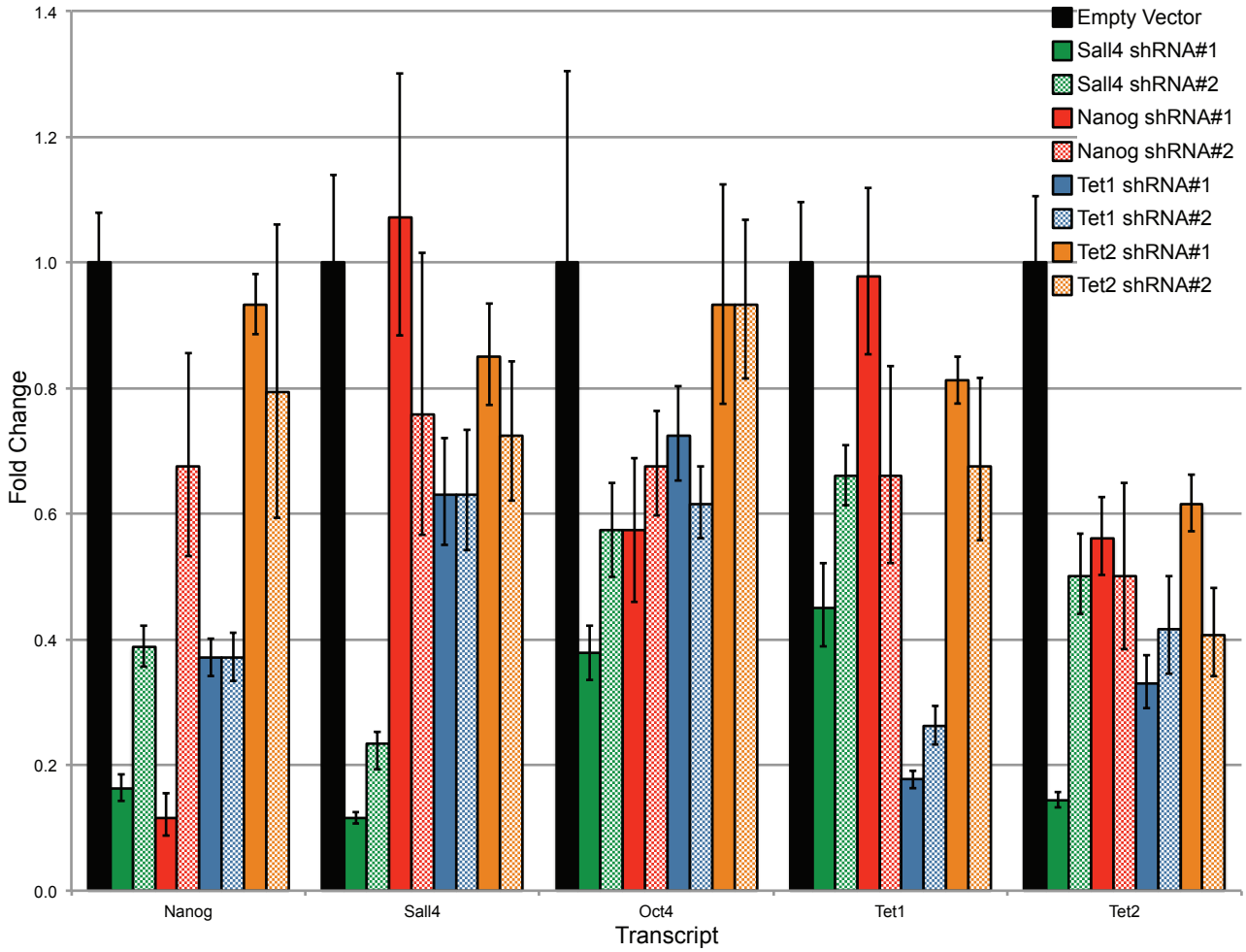




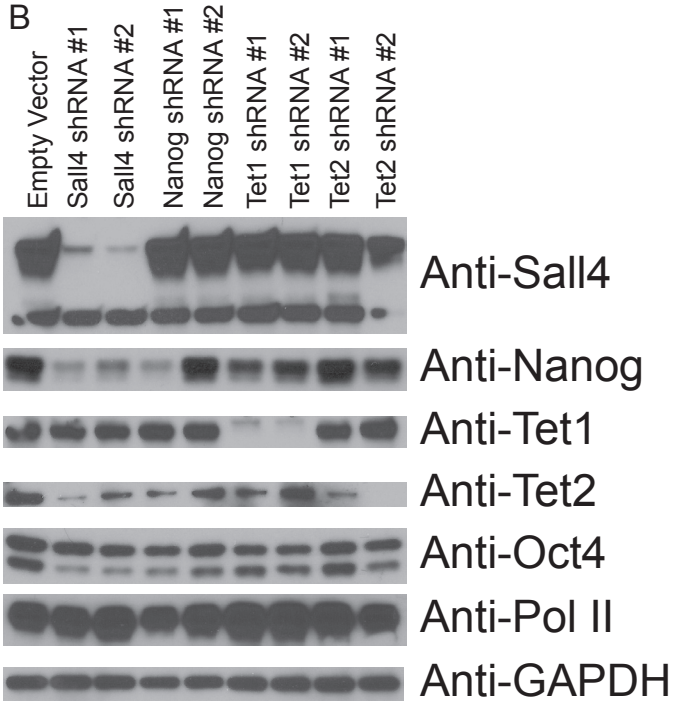


Supplemental Figure 4

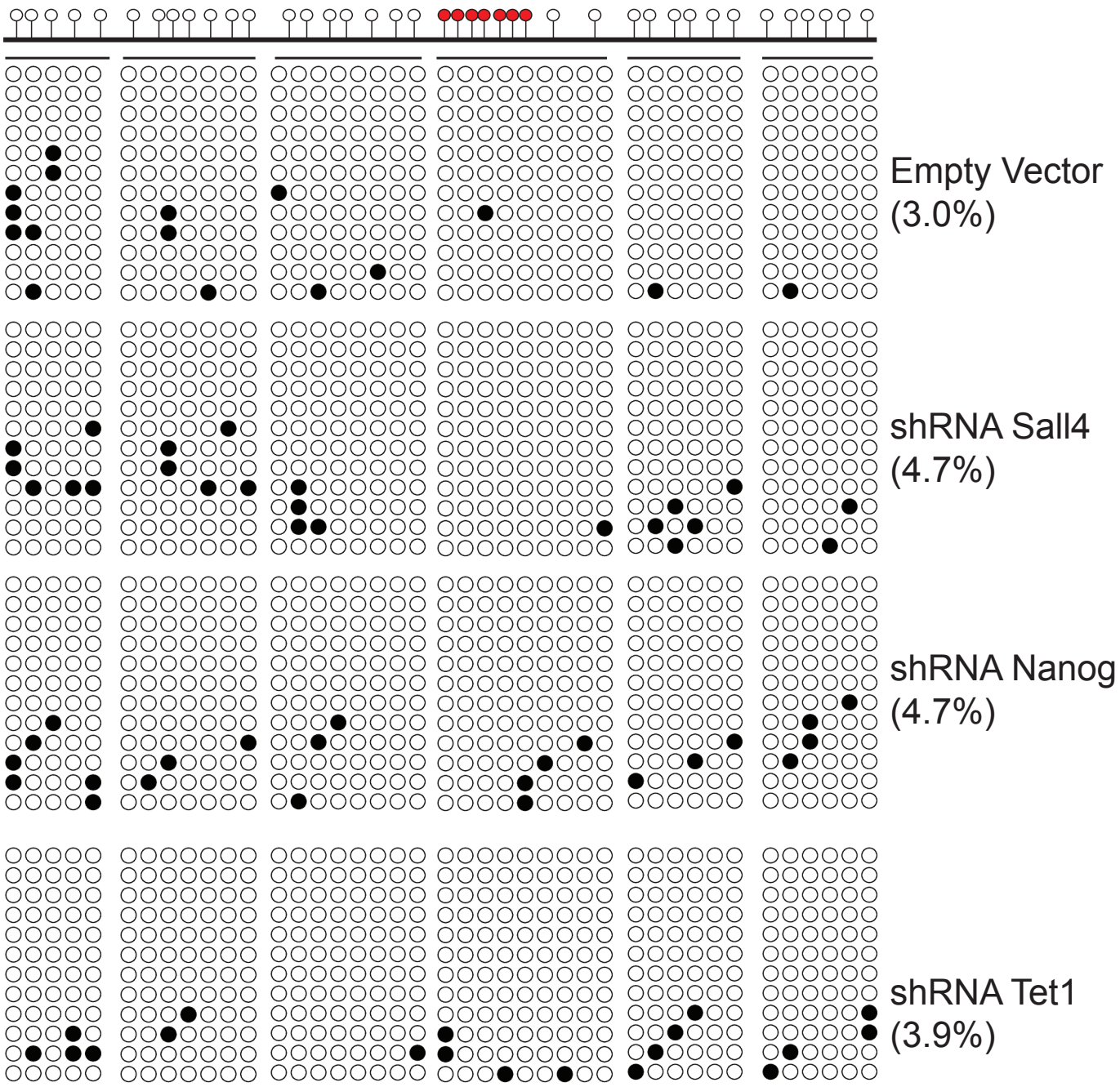
A



B



Supplemental Figure 5



Supplemental Figure/Table Legends:

Supplemental Figure S1: (A) Box Plots indicating the relative transcription rates (reads per kilobase million reads, RPKM) on the y-axis, and different classes of enhancers and/or RNAP II binding indicated on the x-axis. Outliers are omitted from the graph. (B) Scatter plots indicating to assess the correlation between enhancer and promoter transcription rates. Green Color scale indicates the # of data points present in a given bin.

Supplemental Figure S2: (A) The % overlap between different enhancer groups with either UMRs (upper panel) or LMRs (lower panel) originally identified.[39] (B) Heatmaps showing the normalized Tet1 signal in different classes of eRNA-negative enhancers. The x-axis corresponds to a 5kb window centered at each enhancer. Each row in the heatmap corresponds to one enhancer region.

Supplemental Figure S3: (A) 5mC methylation across the 5kb genomic region displayed in Figure 8 is shown for ESCs (blue, top track) and NP cells (red, lower track). The corresponding HMM designation (LMR, UMR, FMR) from the same work is displayed below.[39] (B) Luciferase based-reporter plasmids with either a heterologous (SV40) or ESC specific (*Nanog)* promoter were transfected into either ESCs or a fibroblast cell-line (3T3). A 2.8kb fragment of the *Nanog*-linked enhancer was cloned downstream of the Luciferase open reading frame in either the Plus or Minus orientation. Thirty-six hours after transfection, luminescence

was assessed. All samples had a Fold Change (y-axis) calculated relative to the reporter without the enhancer element transfected into the same cell line. Error bars are SEM of three experiments. A schematic of each reporter plasmid is shown below the graphs.

Supplemental Figure S4: (A) shRNAs to Sall4, Nanog, Tet1 & Tet2 or the empty lentiviral vector were used to infect ESCs and 48 hours after infection and treatment with puromycin to select cells with a high multiplicity of infection, RNA was collected and the relative expression of various transcripts determined by quantitative PCR. All samples were normalized to the expression levels of the empty lentiviral vector. Error bars represent SEM of three experiments. (B) Western blots were performed using the indicated antibodies to detect steady state protein levels in whole cell extracts after shRNA-mediated depletion at the same 48 hours time point.

Supplemental Figure S5: An approximately 1.8kb region of the *Nanog*-linked enhancer, which overlaps with the RNAP II and Tet1, Tet2 binding sites. Approximate location of each CpG and the single CpG Island (shown in red). This region was broken down into six amplicons, each approximately 250-400 bp in size, and subjected to bisulfite sequencing. 12 individual clones for each region were sequenced following depletion of Sall4, Nanog, or Tet1 by a single shRNA. An empty vector is included as a negative control. The overall percent of methylation is shown in ().

Supplemental Table S1: A list of all the datasets used for this analysis, including their specific GSE or GSM record locator from the GEO omnibus, reference, and the input, control, and *p*-value used for the analysis.

Supplemental Table S2: A modified Bed file listing the different genomic regions classified as enhancers or insulators for this study. For each genomic element, the transcription rates (RPKM) in ESCs and MEFs are indicated. RNAP II occupancy and whether the region is extragenic or intragenic are also indicated.

Supplemental Table S3: For each enhancer in a given class or insulator, the gene it was mapped to, if present, and the histone methylation status of that gene's promoter.

Supplemental Table S4: A list of all primers used for this study, separated by the technique (RT-qPCR, ChIP-qPCR, or bisulfite sequencing) they were used for. The original reference, if applicable, is indicated.

Supplemental Table S5: A listing of the number of peaks initially present after peak calling for each enhancer definition. The number of peaks retained after removal of specific genomic regions (overlap with known promoters, histone marks, etc.) is indicated.

Table S5

	NOS	H3K4me1	H3K27Ac	p300
MACS generated peaks	14,350	47,648	59,333	44,263
After removing peaks less than 1kb in size	NA	10,709	30,803	NA
After removing peaks within 2kb of RefSeq TSS	13,240	8,618	18,612	31,285
After removing peaks within 2kb of miRNA TSS	13,238	8,614	18,596	31,281
After removing H3K4me3 rich peaks	12,734	8,129	13,302	27,704
After removing H3K36me3 rich peaks	12,724	8,078	13,222	27,659
Union with Med1 or Med12 sites	8,830	NA	NA	NA
Total enhancers	**8,830**	**8,078**	**13,222**	**27,659**