# c-Myc regulates transcriptional pause release

Peter B. Rahl [1], Charles Y. Lin [1,2], Amy C. Seila [3], Ryan A. Flynn [3], Scott McCuine [1], Christopher B. Burge [2], Philip A. Sharp [2,3], Richard A. Young [1,2]

## Supplemental Information

### Supplemental tables



### Supplementary figures

**Supplemental experimental procedures**

**Supplemental Tables**

| Dataset | total number of unique reads used for analysis |
|---|---|
| Pol II Ser5P | 3543456 |
| Pol II Ser2P | 10231400 |
| Spt5 | 4639835 |
| NelfA | 3406514 |
| Ctr9 | 4672291 |
| Pol II (all) sh Control | 5156732 |
| Pol II (all) sh NelfA | 3785780 |
| Pol II (all) sh Spt5 | 3369501 |
| Pol II (all) DMSO control | 1991231 |
| Pol II (all) flavopiridol-treated | 3001750 |
| Pol II (all) DMSO control (for 10058-F4 experiment) | 6320713 |
| Pol II (all) 10058-F4-treated | 4429350 |
| Pol II (all) sh Control (for sh c-Myc) | 6027028 |
| Pol II (all) sh c-Myc | 7031951 |
| Pol II (all) 0 hr doxycyclin (ZHBTc4 cells) | 9820076 |
| Pol II (all) 12 hr doxycyclin (ZHBTc4 cells) | 9211878 |
| Pol II (all) 24 hr doxycyclin (ZHBTc4 cells) | 7392695 |

**Table S1. Total number of reads for each ChIP-seq dataset. Related to Figures 1, 2, 3, 4, 6 and 7.** A list of the total number of ChIP-seq mapped reads used from each dataset for analysis in this study.

**Table S2. Enriched Genes and Regions, at 1e-9, for each ChIP-seq dataset. Related to Figures 1, 2, 3, 4, 6 and 7.** The first tab is a list of the genes tested for enrichment for each ChIP-seq dataset. 1 denotes enriched over background and 0 denotes not enriched at 1e-9 over background. Binding is determined at

p=1e-9, a stringent cutoff to minimize false positives. Enrichment was determined 1000bp upstream or downstream of each transcriptional start site. The genomic coordinates refer to mouse genome build mm8. The minimum enrichment for calling a bound gene is 5 fold enriched over background reads with a statistical significance of p=1e-9. The additional tables list the regions enriched above background for each factor where the enriched region start and enriched region end is listed. The genomic coordinates correspond to mouse genome build mm8. The minimum enrichment for calling an enriched region is 5 fold enriched over background at a statistical significance of p=1e-9, a stringent cutoff to minimize false positives. Additionally, the nearest transcript is listed, if it is within 1000 bp of the enriched region. The enriched region for each factor is represented as separate tabs in the excel file.

| | Number of Genes | Number of Genes with TR>2 | Number of Genes with TR<2 |
|---|---|---|---|
| Pol II bound | 11456 | | |
| Pol II bound with H3K79me2 (active) | 6842 | 6252 | 590 |
| Pol II bound without H3K79m2 (non-productive) | 4614 | 4184 | 430 |
| Inactive | 10407 | NA | NA |

**Table S3. Number of genes classified as active, non-productive and inactive. Related to Figures 1, 2, 3, and 4.** This table lists the number of genes in each gene class (active, non-productive and inactive). 21865 total RefSeq genes were used for this analysis. Active genes are bound by Pol II and H3K79me2 (marker of elongation), non-productive genes are bound by Pol II but not H3K79me2 and inactive genes are not bound by either. Binding is determined at p=1e-9, a stringent cutoff to minimize false positives. In Figure 1c, we determine that approximately 91% of genes have a TR greater than 2 and 9% of genes had a TR less than or equal to 2. We list the number of genes in each of these classes as they fall into the active and non-productive gene classes.

## Supplemental Figures

**Figure S1. Example genes with varying traveling ratios. Related to Figure 1.**

(A) A plot of the percent of Pol II bound genes with the indicated TR value or less. Each example gene depicted below is shown in red.

(B) Gene plots showing the Pol II occupancy at a given gene and the corresponding TR value to provide the reader with a general idea of how Pol II occupancy at a gene is related to TR. There is a very small correlation (R^2 = 0.17) between gene length and TR, but gene length normalized versions of the TR produce no meaningful changes to our results.

(C) Traveling ratios plots for Pol II occupancy in two mES control datasets (mES cells +shControl (from shNelfA and shSpt5 experiment) and mES cells +DMSO).

(D) Traveling ratio as a function of gene expression, finding no statistically significant correlation between TR and gene expression.

(E) Amount of Pol II Ser2P in the gene end region (+/-1kb from the 3' end of gene) as a function of gene expression, finding a weak correlation between the two.

**Figure S2. Flavopiridol treatment in mES cells results in loss of phosphorylation at P-TEFb targets. Related to Figure 2.**

(A) mES cells were treated with the indicated flavopiridol concentration or DMSO alone (-), for 60 minutes. Nuclear extracts were analyzed by Western blot using specific antibodies against Ser2P Pol II, Ser5P Pol II, Spt5, Cdk9 and Brg1

(loading control). ** - higher molecular weight Spt5 species, most likely the phosphorylated C-terminal repeat form as reported in (Yamada et al., 2006), that is flavopiridol sensitive. * - lower molecular weight Spt5 species.

(B) Pol II density at the gene end (region defined by +/-1kb from the 3' end of the gene) is plotted from mES cells treated with DMSO (X axis) versus flavopiridol (Y axis) for all active, non-overlapping genes. AUC is area under the curve, representing the ChIP-seq density for each gene. The vast majority of genes have reduced Pol II occupancy in this region following P-TEFb inhibition.

**Figure S3. DSIF and NELF co-occupy Pol II in the promoter proximal region following P-TEFb inhibition. Spt5, NelfA and Pol II ChIP-seq occupancy is highly correlative. Related to Figure 3.**

(A) Pairwise correlation analysis of Spt5, NelfA and Pol II (all) ChIP-seq occupancy in mES cells. This demonstrates that Spt5 and NelfA ChIP-seq occupancy positively correlates with Pol II occupancy.

(B) *left* - Western blot analysis of mES cells treated with 1μM flavopiridol or control for 60 minutes prior. Nuclear extracts were analyzed with specific antibodies against Ser2P Pol II, Spt5, Cdk9 and Brg1 (loading control). ** - higher molecular weight Spt5 species, most likely the phosphorylated C-terminal repeat form, that is flavopiridol sensitive. * - lower molecular weight Spt5 species. *right* - Average gene binding for NelfA and Spt5 following 1μM flavopiridol (red) or control (black) for 60 minutes in mES cells. The average ChIP-chip enrichment

was determined in each bin (250bp) in each condition and plotted from –4kb to +2kb.

**Figure S4. Spt5 associates with chromatin following NelfA knockdown and NelfA associates with chromatin following Spt5 knockdown. Spt5 and NelfA knockdown cause a decrease in promoter-proximal Pol II occupancy at a subset of genes. Related to Figure 4.**

(A) Individual gene examples of NelfA binding in shControl (blue) and shSpt5 (red) mES cells using ChIP-chip.

(B) Average gene binding for NelfA in shControl (blue) and shSpt5 (red) mES cells. The average NelfA enrichment was determined in each bin (250bp) in each cell type and plotted from –4kb to +2kb.

(C) Individual gene examples of Spt5 binding in shControl (blue) and shNelfA (red) mES cells using ChIP-chip. Fold enrichment is plotted over the indicated chromosomal region.

(D) Average gene binding for Spt5 in shControl (blue) and shNelfA (red) mES cells. The average Spt5 enrichment was determined in each bin (250bp) in each cell type and plotted from –4kb to +2kb.

(E) Pol II TR analysis at active genes in mES cells following shControl (black), shSpt5 (orange) and shNelfA (blue) knockdown. Lower TR values indicate a lower degree of pausing and a shift in TR curve to the left indicates a general trend in this gene class to become less paused.

(F) Pol II TR analysis at non-productive genes in mES cells following shControl (black), shSpt5 (orange) and shNelfA (blue) knockdown. Lower TR values indicate a lower degree of pausing and a shift in TR curve to the left indicates a general trend in this gene class to become less paused. Spt5 knockdown effects Pol II occupancy both active and non-productive genes and NelfA knockdown mainly effects Pol II occupancy at non-productive genes.

(G) RNA Pol II (all) ChIP-seq occupancy at three non-productive genes in mES cells in shControl (black), shSpt5 (orange) or shNelfA (blue).

**Figure S5. c-Myc occupies regions near the transcriptional start site. Related to Figure 5.**

(A) c-Myc (red) ChIP-seq occupancy is close to the TSS and the average Pol II promoter proximal peak (black).

(B) Distribution of the canonical E-Box core sequence motif (CACGTG), which is recognized by c-Myc, near the TSS.

**Figure S6. Max occupies c-Myc binding sites and Pol II occupancy is altered in the transcribed region following c-Myc shRNA knockdown. Related to Figure 6.**

(A) Max ChIP enrichment in mES cells at seven c-Myc binding sites (*Prdx1, Josd3, Actb, Eef1g1, Nol5, Mat2a* and *Ybx1*) and one non c-Myc binding site (*Gata1*). Max ChIP DNA was analyzed using qPCR at the selected binding sites and enrichment was calculated for replicate PCR reactions over input DNA at

each region. This demonstrates that Max binds to c-Myc binding sites in mES cells, consistent with a model where c-Myc and Max heterodimerize and bind the same sites. Error bars represent s.d from duplicate PCR reactions.

(B) Levels of c-Myc mRNA levels in mES cells infected with sh Control or sh c-Myc shRNA constructs. ES cells were infected for 24 hours, then selected for 72 hours prior to harvesting. c-Myc levels were determined using RT-PCR analysis and normalized against a Gapdh control.

(C) Levels of c-Myc protein levels in mES cells infected with sh Control or sh c-Myc shRNA constructs. ES cells were infected for 24 hours, then selected for 72 hours prior to harvesting. c-Myc levels were determined using Western blot analysis using an antibody against c-Myc. Brg1 is used as a loading control.

(D) Pol II ChIP-seq binding plots in mES cells with sh Control or sh c-Myc at three c-Myc target genes (*Npm1*, *Ncl,* and *Nol5*) and two non c-Myc taget genes (*Txnip* and *Nanog*). This panel is demonstrating that at these genes c-Myc knockdown has a similar phenotype to 10058-F4 treatment where Pol II density is reduced in the gene body and the density in the promoter proximal region is unaffected.

(E) Protein levels of P-TEFb components Cyclin T1 and Cdk9 with and without 10058-F4 treatment for 6 hours. Protein extracts were analyzed using Western blots and probed with antibodies against CycT1 and Cdk9. TBP was used as a loading control. This analysis demonstrates that the phenotype observed following 10058-F4 treatment (similar to P-TEFb inhibition with flavopiridol) is not a result of decreased levels of P-TEFb.

(F) TR analysis on c-Myc target genes in mES cells + sh Control (black) and sh c-Myc (blue), displaying the percent of genes with a given TR. This analysis shows that genes generally become more paused following c-Myc shRNA knockdown, as indicated by the shift in TR to the right.

(G) c-Myc and Oct4 binding at active genes ordered by amount of binding at the promoter (c-Myc +/-1kb, Oct4 +/-5kb). Target (black) and non-target (grey) gene sets are demarcated, which were used for subsequent analysis on Pol II ChIP-seq occupancy.

**Supplemental experimental procedures**

**mES cell culture**

V6.5 (C57BL/6-129) murine ES cells were grown under typical mES conditions on irradiated mouse embryonic fibroblasts (MEFs). In summary, cells were grown on gelatinized tissue culture plates in Dulbecco's modified Eagle medium supplemented with 15% fetal bovine serum (characterized from Hyclone), 1000 U/ml leukemia inhibitory factor (LIF, Chemicon; ESGRO ESG1106), non-essential amino acids, L-glutamine, Penicillin/Streptomycin and ß-mercaptoethanol. For location analysis, cells were grown for two passages off of MEFs, on gelatinized tissue-culture plates. Mouse embryonic fibroblasts were prepared and cultured from DR-4 strain mice.

For location analysis following Spt5 and NelfA knockdown, shRNA plasmids targeting the mouse Spt5 and NelfA mRNAs and an empty plasmid (control) (Open Biosystems, Huntsville, AL RMM4534-NM_013676, RMM4534-NM_011914, and RMM4534_NM_010849, RHS4080) were used. We purchased and tested each set of shRNA hairpins for ability of each hairpin to knockdown the mRNA of the factor of interest. We then selected the hairpin that performed the best for use in ChIP-seq analysis. For Spt5, hairpin TRCN0000092761 was used. For NelfA, hairpin TRCN0000124874 was used. For c-Myc, hairpin TRCN0000042516 was used. 293T cells were plated in 6-well dishes at 6x10^5 cells/well. The shRNA plasmids and lentiviral components were co-transfected into 293T cells. V6.5 mES cells were plated in T-75 flasks at 2x10^6 cells/flask. Viral media was collected 48 hours after co-transfection and the V6.5 mES cells were directly infected with the viral media 24 hours after initial plating of the mES cells. The infection media was 1:2 viral media:mES cell media with 2mM polybrene. The efficiently infected cells were selected for 24 hours post infection with mES cell media containing 2μM puromycin. V6.5 cells were cross-linked 72 hours post selection and frozen for ChIP-Seq experiments and western blotting. To assess mRNA knockdown a small fraction of cells were collected and mRNA was prepared (using a Qiagen RNeasy Mini Kit, Qiagen, Valencia, CA). RT-qPCR was used to determine relative gene expression between mock knockdown and shRNA knockdown samples. Taqman gene expression assays from Applied Biosystems, Foster City, CA were ordered for the Spt5, NelfA, and GAPDH (control) genes (Mm01217228_m1, Mm01170629_m1,

Mm99999915_g1, Mm00487804_m1). Western blot analysis (described below) was done to assess knockdown at the protein level prior to ChIP-seq analysis. c-Myc knockdown was also assessed at the mRNA level (described below) using RT-PCR.

For location analysis on mES cells following treatment with small molecule inhibitors, cells were grown two passages off feeders and prior to formaldehyde crosslinking, the cells were treated by addition of the indicated final concentration of flavopiridol (1μM for 1 hour for ChIP-chip and ChIP-seq experiments, or the indicated concentration and time for Western blot analysis), or c-Myc/Max inhibitor 10058-F4 (50μM for 6 hours for ChIP-seq experiments or the indicated time for Western blot analysis), both dissolved in DMSO, to the growth medium. 50μM is within the concentration range commonly used for 10058-F4 *in vivo* to investigate c-Myc function (Arabi et al., 2005; Fang et al., 2008; Faumont et al., 2009; Follis et al., 2008; Hammoudeh et al., 2009; Khanna et al., 2009; Lee et al., 2009; Sampson et al., 2007; Wang et al., 2007). As a control, vehicle alone (DMSO) was added to the growth medium at the same final volume as with drug. Small molecule inhibitors used were: Flavopiridol (Sigma cat #F3055), and c-Myc inhibitor 10058-F4 (Sigma cat #F3680).

For location analysis following Oct4 shutdown, ZHBTc4 mES cells (Niwa et al., 2000) were grown under standard mES cell culture conditions and expanded for two passages off MEF feeders. ES cell culture media with 2 μg/ml doxycycline

was added to the cells for 0 hours, 12 hours and 24 hours prior to formaldehyde crosslinking. Loss of Oct4 was assessed using Western Blot analysis (see below). Oct4 protein was essentially depleted at the 24 hour time point and we noticed that the cells appeared morphologically different from mES cells and the 0 hr and 12 hr time points. It is well established that loss of Oct4 causes ES cell differentiation. Therefore, in order to minimize measuring secondary effects, we performed our TR analysis on Oct4 targets and non-targets on the earliest time points: 0 hr and 12 hr following doxycycline treatment.

**Chromatin immunoprecipitation (ChIP)**

ChIP was done following the Agilent Mammalian ChIP-on-chip protocol (version 9.1, Nov 2006). In summary, mES cells were grown as described above and cross-linked for 15 minutes at room temperature by the addition of one-tenth of the volume of 11% formaldehyde solution (11% formaldehyde, 50mM Hepes pH7.3, 100mM NaCl, 1mM EDTA pH8.0, 0.5mM EGTA pH8.0) to the growth media followed by two washes with PBS. Cells were scraped and frozen in liquid nitrogen. 100ul of Dynal magnetic beads (Sigma) were blocked with 0.5% BSA (w/v) in PBS. Magnetic beads were bound with 10ug of the indicated antibody. Antibodies used are as follows: Pol II (all; Rpb1 N-terminus): Santa Cruz sc-899; Ser5P Pol II: Abcam ab5131; Ser2P Pol II: Abcam (H5 clone) ab24758 with Upstate IgG-IgM linker antibody 12-488; Spt5: gift from Yuki Yamaguchi and Hiroshi Handa (Wada et al., 1998); NelfA: Santa Cruz (A-20) sc-23599;; Ctr9: Bethyl labs A301-395; and Max: Santa Cruz sc-197. For all of the experiments

analyzing Pol II occupancy following shRNA-mediated knockdown, flavopiridol, or 10058-F4 treatment, the Pol II (all; Santa Cruz sc-899, Pol II N-20) antibody was used. Crosslinked cells were lysed with lysis buffer 1 (50mM Hepes pH7.3, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40, and 0.25% Triton X-100) and washed with lysis buffer 2 (10mM Tris-HCl pH8.0, 200mM NaCl, 1mM EDTA pH8.0 and 0.5mM EGTA pH8.0).

For Spt5 ChIPs, cells were resuspended and sonicated in lysis buffer 3 (10mM Tris-HCl pH8.0, 100mM NaCl, 1mM EDTA pH8.0, 0.5mM EGTA pH8.0, 0.1% Na-Deoxycholate and 0.5% N-lauroylsarcosine) for 8 cycles at 30 seconds each on ice (18 watts) with 60 seconds on ice between cycles. Triton X-100 was added to a final concentration of 1% to the sonicated lysates. Sonicated lysates were cleared and incubated overnight at 4$^{o}$C with magnetic beads bound with antibody to enrich for DNA fragments bound by the indicated factor. Beads were washed four times with RIPA (50mM Hepes pH7.3, 500mM LiCl, 1mM EDTA, 1% NP-40 and 0.7% Na-Deoxycholate) and once with TE + 50mM NaCl. Bound complexes were eluted in elution buffer (50mM Tris-HCl pH8.0, 10mM EDTA pH8.0, 1% SDS) at 65$^{o}$C for 15 minutes with occasional vortexing. Cross-links were reversed overnight at 65$^{o}$C. RNA and protein were digested using RNAse A and Proteinase K, respectively and DNA was purified with phenol chloroform extraction and ethanol precipitation.

For Pol II Ser5P, Pol II (all), NelfA, Ctr9 and Max ChIPs, cells were resuspended and sonicated in sonication buffer (50mM Tris-HCl pH7.5, 140mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS) for 8 cycles at 30 seconds each on ice (18 watts) with 60 seconds on ice between cycles. Sonicated lysates were cleared and incubated overnight at 4$^{o}$C with magnetic beads bound with antibody to enrich for DNA fragments bound by the indicated factor. Beads were washed three times with sonication buffer, one time with sonication buffer with 500mM NaCl, one time with LiCl wash buffer (20mM Tris pH8.0, 1mM EDTA, 250mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate) and one time with TE. DNA was eluted in elution buffer. Cross-links were reversed overnight. RNA and protein were digested using RNAse A and Proteinase K, respectively and DNA was purified with phenol chloroform extraction and ethanol precipitation.

For Pol II Ser2P ChIP, cells were resuspended and sonicated in sonication buffer II (50mM Tris-HCl pH7.5, 140mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS) for 8 cycles at 30 seconds each on ice, at 18 watts with 60 seconds on ice between cycles. Sonicated lysates were cleared and incubated overnight at 4$^{o}$C with magnetic beads bound with antibody to enrich for DNA fragments bound by Pol II Ser2P. Beads were washed two times with sonication buffer II, one time with LiCl wash buffer (2mM Tris pH8.0, 0.02mM EDTA, 50mM LiCl, 0.1% NP-40, 0.1% Na-deoxycholate) and one wash with TE. DNA was eluted in elution buffer. This protocol is similar to that used in

(Stock et al., 2007). Cross-links were reversed overnight. RNA and protein were digested using RNAse A and Proteinase K, respectively and DNA was purified with phenol chloroform extraction and ethanol precipitation.

**ChIP-PCR analysis**

Max ChIP DNA was analyzed using SYBR Green real-time PCR analysis (Applied Biosystems). Fold enrichment was determined from replicate PCR reactions at seven c-Myc binding sites (Prdx1, Josd3, Actb, Eef1g1, Nol5, Mat2a, Ybx1) and one non-c-Myc binding site (Gata1), as determined by previously published c-Myc ChIP-seq (Chen et al., 2008), over input DNA. The oligos used for this analysis are:

Prdx1 fwd: ttagttcccggacctgttg

Prdx1 rev: acaaactcgtcccaccaag

Josd3 fwd: cctggagggcgtttttagt

Josd3 rev: acccttttcggaacgtaacc

Actb fwd: gatcactcagaacggacacc

Actb rev: acacgctaggcgtaaagttg

Eef1g1 fwd: CTGGGTCTCCATTGTCTGG

Eef1g1 rev: AGTTCCACCAACCTGCTCA

Nol5 fwd: GGCTCCGAAAAGATGTGAA

Nol5 rev: AGCAGAGGTCGCCCTAAAT

Mat2a fwd: GTCTCCGAAGGTCCCATCT

Mat2a rev: TGAAGGCTAAAGGGCATGT

Ybx1 fwd: AGATCCTGGACCGACTTCC

Ybx1 rev: GTTCCCAAAACCTTCGTTG

Gata1 fwd: agagcctaaaaggtcctcca

Gata1 rev: caccttctccctcctctttc

**Hybridization to DNA microarray**

Purified immunoprecipitated DNA was amplified using two rounds of ligation mediated PCR (LM PCR), as described in (Lee et al., 2006b) and the Agilent Mammalian ChIP-on-chip protocol (version 9.1, Nov 2006). LM-PCR immunoprecipitated DNA was labeled with Cy5, LM-PCR input DNA was labeled with Cy3 using Invitrogen Bioprime random primer labeling kit. For microarray hybridization, the Agilent Mammalian ChIP-on-chip protocol (version 9.1, Nov 2006) was followed. In brief, mouse Cot1 DNA, Agilent blocking buffer (1X final conc.), Agilent hybridization buffer (1X final conc.), was added to Cy5- and Cy3-labeled DNA. The mixture was incubated at 95$^{o}$C for 3 minutes, followed by 37$^{o}$C for 30 minutes. Sample was centrifuged for 1 min and sample was hybridized to Agilent DNA microarray. For experiments testing effects of flavopiridol on NelfA and Spt5 occupancy, and NelfA or Spt5 occupancy following Spt5 or NelfA knockdown, respectively, ChIP samples were hybridized to Agilent arrays MTvB (44,000 features covering the promoter region, from approximately -6kb to +2kb, of approximately 10% of mouse genes – MTvB). DNA was hybridized to microarray for 40 hours at 65$^{o}$C. Microarray was washed and scanned following the Agilent Mammalian ChIP-on-chip protocol. The Agilent

DNA microarray scanner BA was used. PMT settings were set manually to normalize bulk signal in the Cy3 and Cy5 channel. Data was processed as described in (Lee et al., 2006a) to calculate enrichment ratios and determine bound regions. Figures presented in this manuscript using ChIP-chip data display the chromosomal coordinates from mouse genome build mm6.

**Solexa/Illumina sequencing**

All protocols used for Solexa/Illumina ChIP-seq analysis (sample preparation, polony generation on Solexa flow-cells, sequencing, and Solexa data analysis) are described in (Guenther et al., 2008; Marson et al., 2008). A summary of the protocol used is described below.

*Sample preparation*

Purified immunoprecipitated DNA was prepared for sequencing according to a modified version of the Illumina/Solexa Genomic DNA protocol. Fragmented DNA was prepared by repairing the ends and adding a single adenine nucleotide overhang to allow for directional ligation. A 1:100 dilution (in water) of the Adaptor Oligo Mix (Illumina) was used in the ligation step. A subsequent PCR step with limited (18) amplification cycles added additional linker sequence to the fragments to prepare them for annealing to the Genome Analyzer flow-cell. Following amplification, the library was size selected to a narrow range of fragment sizes by separation on a 2% agarose gel and a band between 150-300 bp (representing shear fragments between 50 and 200nt in length and ~100bp of

primer sequence) was excised. The DNA was purified from the agarose and this DNA library was subsequently used for polony generation and sequencing.

*Polony Generation and Sequencing*

The DNA library (2-4 pM) was applied to the flow-cell (8 samples per flow-cell) using the Cluster Station device from Illumina. The concentration of library applied to the flow-cell was calibrated such that polonies generated in the bridge amplification step originate from single strands of DNA. Multiple rounds of amplification reagents were flowed across the cell in the bridge amplification step to generate polonies of approximately 1,000 strands in 1μm diameter spots. Double stranded polonies were visually checked for density and morphology by staining with a 1:5000 dilution of SYBR Green I (Invitrogen) and visualizing with a microscope under fluorescent illumination. Validated flow-cells were stored at $4^{o}C$ until sequencing.

Flow-cells were removed from storage and subjected to linearization and annealing of sequencing primer on the Cluster Station. Primed flow-cells were loaded into the Illumina Genome Analyzer 1G. After the first base was incorporated in the Sequencing-by-Synthesis reaction the process was paused for a key quality control checkpoint. A small section of each lane was imaged and the average intensity value for all four bases was compared to minimum thresholds. Flow-cells with low first base intensities were re-primed and if signal was not recovered the flow-cell was aborted. Flow-cells with signal intensities

meeting the minimum thresholds were resumed and sequenced for 26 or 32 cycles.

*Solexa Data Analysis*

Images acquired from the Illumina/Solexa sequencer were processed through the bundled Solexa image extraction pipeline, which identified polony positions, performed base-calling and generated QC statistics. Sequences were aligned using ELAND software to NCBI Build 36 (UCSC mm8) of the mouse genome. Only sequences that mapped uniquely to the genome with zero or one mismatch were used for further analysis. When multiple reads mapped to the same genomic position, a maximum of two reads mapping to the same position were used. Refer to Table S1 for a list of the total number of mapped reads used for analysis of each ChIP-seq dataset.

Analysis methods were derived from previously published methods (Johnson et al., 2007; Mikkelsen et al., 2007; Marson et al., 2008; Guenther et al., 2008). Sequence reads from multiple flow cell runs were combined for Pol II Ser2P ChIP-seq dataset. Each read was extended 100bp, towards the interior of the sequenced fragment, based on the strand of the alignment. The number of ChIP-Seq reads across the genome, in 25bp bins within a 1kb window surrounding each bin (+/- 500bp) was tabulated. The 25bp genomic bins that contained statistically significant ChIP-Seq enrichment was identified by comparison to a Poissonian background model. Assuming background reads

are spread randomly throughout the genome, the probability of observing a given number of reads in a 1kb window can be modeled as a Poisson process in which the expectation can be estimated as the number of mapped reads multiplied by the number of bins (40) into which each read maps, divided by the total number of bins available (we estimated 70%). Enriched bins within 1kb of one another were combined into regions. The complete set of RefSeq genes was downloaded from the UCSC table browser (http://genome.ucsc.edu/cgi-bin/hgTables?command=start) on December 20, 2008. Genes with enriched regions within 1kb to their transcription start site to annotated stop site were called bound.

The Poissonian background model assumes a random distribution of background reads, however we have observed significant deviations from this expectation. Some of these non-random events can be detected as sites of apparent enrichment in negative control DNA samples and can create many false positives in ChIP-Seq experiments. To remove these regions, we compared genomic bins and regions that meet the statistical threshold for enrichment to a set of reads obtained from Solexa sequencing of DNA from whole cell extract (WCE) in matched cell samples. We required that enriched bins and enriched regions have five-fold greater ChIP-Seq density in the specific IP sample, compared with the control sample, normalized to the total number of reads in each dataset. This served to filter out genomic regions that are biased to having a greater than

expected background density of ChIP-Seq reads. A summary of the bound regions and genes for each antibody is provided in Table S3 and Table S4.

For comparison of Pol II occupancy following either shRNA-mediated knockdown or small molecule inhibition with a control dataset, rank normalization was used to normalize the datasets to be compared. This normalization method is described in (Bilodeau et al., 2009). Briefly, a quantile normalization method was used for analysis. For each dataset compared, the genomic bin with the greatest ChIP-Seq density was identified. The average of these values was calculated and the highest signal bin in each dataset was assigned this average value. This was repeated for all genomic bins from the greatest signal to the least, assigning each the average ChIP-Seq signal for all bins of that rank across all datasets.

**ChIP-seq binding tracks available on GEO database**

ChIP-seq datasets generated in this study can be downloaded from the GEO datasets database (http://www.ncbi.nlm.nih.gov/gds) under the accession number GSE20485. Included under this accession number is ChIP-seq binding tracks that can uploaded onto the UCSC genome browser. Descriptions for the datasets contained in each of the WIG files are listed below.

*RNAPolII_Phosphorylation_Tracks.WIG.gz* - contains three ChIP-seq datasets mapped to mouse genome mm8 using a +200bp extension model that can be uploaded to the UCSC genome browser to view binding events

(http://genome.ucsc.edu/). The datasets include Pol II Ser5P in wildtype V6.5 mES cells, Pol II Ser2P in wildtype V6.5 mES cells, and Pol II (all) in wildtype V6.5 mES cells + shControl. Enrichment is shown as counts per million reads in 25bp genomic bins and the track floor is at 1 count per million reads. Additional tracks identify genomic regions that have been determined enriched above background at the given p value.

*Factor_Tracks.WIG.gz* - contains three ChIP-seq datasets mapped to mouse genome mm8 using a +200bp extension model that can be uploaded to the UCSC genome browser to view binding events. The datasets include Spt5, NelfA, and Ctr9 in wildtype V6.5 mES cells. Enrichment is shown as counts per million reads in 25bp genomic bins and the track floor is at 1 count per million reads. Additional tracks identify genomic regions that have been determined enriched above background at the given p value.

*RNAPolII_PauseFactorKnockdown_Tracks.WIG.gz* – contains three ChIP-seq datasets mapped to mouse genome mm8 using a +200bp extension model that can be uploaded to the UCSC genome browser to view binding events. The ChIP experiments were done using the N20 Pol II antibody (sc-899). The datasets include Pol II (all) mES V6.5 + shSpt5, Pol II (all) in mES V6.5 + shNelfA, and Pol II (all) mES V6.5 + shControl. Binding enrichment tracks for all datasets are shown as rank normalized counts in 25bp genomic bins and the track floor is at 2 count per million reads. Rank normalized data is indicated with

the title of the track (title of the track is, for example, mES_Pol2_shControl_norm). Additional tracks identify genomic regions that have been determined enriched above background at the given p value.

*RNAPolII_FlavopiridolTreated_Tracks.WIG.gz* – contains two ChIP-seq datasets mapped to mouse genome mm8 using a +200bp extension model that can be uploaded to the UCSC genome browser to view binding events. The ChIP experiments were done using the N20 Pol II antibody (sc-899) to determine total Pol II occupancy. The datasets include Pol II (all) in mES V6.5 + DMSO control, and Pol II (all) in mES V6.5 + flavopiridol (P-TEFb inhibitor). In these experiments, cells were treated with either flavopiridol (1μM) or DMSO alone for 60 minutes prior to crosslinking and ChIP with Pol II antibody. Binding enrichment tracks for all datasets are shown as rank normalized counts in 25bp genomic bins and the track floor is at 2 count per million reads. Rank normalized data is indicated with the title of the track (title of the track is, for example, mES_Pol2_Flavo_norm). Additional tracks identify genomic regions that have been determined enriched above background at the given p value.

*RNAPolII_10058F4Treated_Tracks.WIG.gz* - contains two ChIP-seq datasets mapped to mouse genome mm8 using a +200bp extension model that can be uploaded to the UCSC genome browser to view binding events. The ChIP experiments were done using the N20 Pol II antibody (sc-899) to determined total Pol II occupancy. The datasets include Pol II (all) in mES V6.5 + DMSO control,

and Pol II (all) in mES V6.5 + 10058-F4 (c-Myc/Max inhibitor). In these experiments, cells were treated with either 10058-F4 (50μM) or DMSO alone for 6 hours prior to crosslinking and ChIP with Pol II antibody. Binding enrichment tracks for all datasets are shown as rank normalized counts in 25bp genomic bins and the track floor is at 2 count per million reads. Rank normalized data is indicated with the title of the track (title of the track is, for example, mES_DMSO_norm). Additional tracks identify genomic regions that have been determined enriched above background at the given p value.

*RNAPolII_cMycKnockdown_Tracks.WIG.gz* - contains two ChIP-seq datasets mapped to mouse genome mm8 using a +200bp extension model that can be uploaded to the UCSC genome browser to view binding events. The ChIP experiments were done using the N20 Pol II antibody (sc-899) to determine total Pol II occupancy. The datasets include Pol II (all) mES V6.5 + shc-Myc and Pol II (all) mES V6.5 + shControl. Binding enrichment tracks for all datasets are shown as rank normalized counts in 25bp genomic bins and the track floor is at 2 count per million reads. Rank normalized data is indicated with the title of the track (title of the track is, for example, mES_Pol2_shControl_norm). Additional tracks identify genomic regions that have been determined enriched above background at the given p value.

*RNAPolII_Oct4Shutdown_Tracks.WIG.gz* - contains three ChIP-seq datasets mapped to mouse genome mm8 using a +200bp extension model that can be

uploaded to the UCSC genome browser to view binding events. The ChIP experiments were done using the N20 Pol II antibody (sc-899) to determine total Pol II occupancy. The datasets include Pol II (all) mES ZHBTc4 + doxycycline 0 hours, Pol II (all) mES ZHBTc4 + doxycycline 12 hours, and Pol II (all) mES ZHBTc4 + doxycycline 24 hours. Binding enrichment tracks for all datasets are shown as rank normalized counts in 25bp genomic bins and the track floor is at 2 count per million reads. Rank normalized data is indicated with the title of the track (title of the track is, for example, mES_Pol2_Dox0hr_norm). Additional tracks identify genomic regions that have been determined enriched above background at the given p value.

**Active and non-productive gene classes in mES cells**

The active and non-productive genes were classified in mES cells using H3K4me3 (initiation-associated chromatin modification) and H3K79me2 (elongation-associated chromatin modification), as determined by ChIP-seq (Marson et al. 2008), as markers of transcriptional state. Active genes with both H3K4me3 and H3K79me2 chromatin modifications, non-productive genes had only H3K4me3 and inactive genes did not have H3K4me3 or H3K79me2 chromatin modifications. When showing *Rpl3* as the example of the active gene and *Surb7* as the example of the non-productive gene in Figure 1A and Figure 3A, the Pol II (all) dataset used for generating the gene plots was Pol II (all) shControl (from shSpt5 and shNelfA experiment).

**Traveling ratio calculation**

Pol II levels peak in the 5' region of many genes. To quantify this effect, we have developed a measure called Traveling Ratio (TR) that compares the ratio between Pol II density in the promoter and in the gene region.

We first defined the promoter region from -30 to +300 relative to the TSS and the gene body as the remaining length of the gene. We next calculated the average density/nt from rank normalized ChIP-seq density files (described in Bilodeau et al. 2009) for each region and computed the TR as the ratio between the two. Following perturbation from either shRNA knockdown or small molecule inhibition, TR can shift either through changes in the density of promoter proximal Pol II or changes in the gene body Pol II density. For example, Figure S4 shows examples of genes where knockdown of Spt5 or NelfA cause a decrease in promoter proximal Pol II density while Figure 4B shows examples where knockdown of Spt5 causes increases in Pol II density in gene bodies, sometimes accompanied by lower amounts of promoter proximal density.

TR values were calculated for all Pol II bound genes. Pol II occupancy profiles and their corresponding TR value for several example genes are shown in Figure S1. For this figure, the Pol II (all) shControl (from shNelfA and shSpt5 experiment) was used for the gene plots. TR values also show strong agreement between the two control datasets (Pol II shControl and Pol II DMSO), as

distributions of TR values are not statistically different (p-Value <0.5) as determined by a Welch's T-Test (Figure S1).

TR values were plotted as a function of gene expression in Figure S1D. Expression data used was from (Hailesellasse Sene et al., 2007). No statistically significant correlation was found between TR and gene expression. In contrast, when Pol II Ser2P occupancy at the gene end (the region +/- 1kb from the 3' end of the gene) is plotted as a function of gene expression (Figure S1E), we see a weak correlation between the two variables.

**Mapping DSIF and NELF peaks relative to Pol II peaks and correlation calculations**

We find that NelfA and Spt5 enrichment spatially overlaps distributions of Pol II near TSS. To determine the spatial relationships between Pol II, NELF, and DSIF, at individual genes, we mapped the location of significant peaks of Spt5 and NelfA relative to locations of downstream Pol II peaks, as determined by a high resolution peak finding algorithm. This algorithm is designed to operate on top of traditional bound region enrichment models—such as the one used in this study—to precisely map binding sites from ChIP-Seq data at higher resolution. The method operates on the assumption that broader "bound regions" of statistical enrichment have already been defined, and that within these "bound regions", exists one or more true binding events. Our method is similar to methods used to map ChIP-Seq peaks (Zhang et al., 2008, Marson et al. 2008)

in that it primarily utilizes the requirement of a paired forward read and reverse read peak. We searched 500nt upstream and downstream of the Pol II peak for the most significant peak in NelfA and Spt5 ChIP-Seq enrichment that was called enriched above 1e-9 threshold by our gene calling algorithm. If a gene was not enriched for NelfA or Spt5 at this threshold, then it was automatically determined as not co-localized with the Pol II peak.

88% of downstream Pol II peaks were co-localized by a corresponding peak of NelfA. Similarly, 61% of Pol II peaks were co-localized by a corresponding Spt5 peak. By visual inspection, we found a majority of Pol II peaks where the peak-finding algorithm failed to identify co-localized Spt5 and NelfA peaks to be co-occupied by NelfA or Spt5 peaks, just below cutoff thresholds for calling a bound gene. This suggests that the high threshold of the peak finding algorithm under reported the number of true instances of Pol II and NELF/DSIF co-localization, and indeed NelfA and Spt5 occupancy positively correlate with Pol II occupancy (Figure S3A) further support for general co-localization of NELF and DSIF to Pol II. Almost all identified promoter proximal peaks of Spt5 and NelfA occurred within +/- 50nt of the Pol II peak. In contrast, when we applied the same analysis to nearest peaks of H3K4me3, the majority of peaks were found more than +/- 50nt from the Pol II peak.

**Heatmap analysis on ChIP-seq occupancy**

ChIP-seq enrichment for the indicated factor or modification was determined in 50bp bins (enrichment in the bin as counts per million), centered on each transcriptional start site. Generally, the gene list for each representation was rank ordered based on the amount of Pol II (all) in mES cells, from most to least to correlate the enrichment of the given factor with the amount of Pol II at each gene. Cluster 3.0 (http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm) and Java Treeview (www.jtreeview.sourceforge.ne) were used to visualize the data and generate figures shown in this manuscript.

To determine the transcriptional state of Oct4, Nanog and c-Myc target genes in mES cells, enriched genes were determined for each factor at p value 1e-8. Many studies have found that the distance between an Oct4 and Nanog binding site and the gene transcriptional start site can vary widely (Boyer et al., 2005; Chen et al., 2008; Kim et al., 2008; Marson et al., 2008). However, c-Myc binding is generally centered within 1kb of the transcriptional start site (Chen et al., 2008; Kim et al., 2008). To take this into account Oct4 and Nanog target genes were determined +/-5kb around each transcriptional start site and c-Myc target genes were determined +/-1kb around each transcriptional start site. The transcriptional state was then determined for each bound gene set through determining the occupancy of Pol II Ser5P, H3K4me3 (initiation-associated chromatin modification), H3K79me2 (elongation-associated chromatin modification) and H3K27me3 (repression-associated chromatin modification).

The data was displayed at each gene centered on the TSS from -2.5kb to +3kb, rank ordered based amount of Pol II Ser5P, using Cluster3.0 and Treeview (described above).

**Determining target and non-target gene sets**

To determine c-Myc and Oct4 target and non-target genes for analysis on changes in Pol II ChIP-seq occupancy, we first ordered all active genes by c-Myc or Oct4 occupancy near the promoter (+/-1kb for c-Myc and +/-5kb for Oct4). For c-Myc genes, the top 1000 and lowest 1000 genes by c-Myc occupancy were demarcated as target and non-target genes respectively. For Oct4, we selected as targets the top 100 Oct4 bound genes that also showed at least a -0.2 fold change in expression following Oct4 shutdown (Matoba et al., 2006). The Oct4 non-targets were determined by taking the lowest 100 Oct4 bound genes for which expression data existed.

**Global analysis of Pol II occupancy changes following cMyc/Max inhibition**

To determine the effect and significance of changes in Pol II occupancy at c-Myc target genes following treatment with 10058-F4, we first derived sets of actively elongating c-Myc targets and non-targets. Actively elongating genes were defined as those bound by Pol II within 1kb of the TSS at a p-Value of 1e-9 and also bound by H3K79me3 in the first 5kb of the gene, again at a p-Value of 1e-9.

Actively elongating genes were ranked by c-Myc levels within +/- 1kb of the TSS, as cMyc generally associates with its target genes close to the TSS (Chen et al.,

2008; Kim et al., 2008). We analyzed the effects on transcription following c-Myc/Max inhibition by generating average Pol II occupancy in the promoter or the gene body region for the high confidence c-Myc targets using rank normalized datasets. For the average Pol II occupancy analysis shown in Figure 6C, we used the high confidence set of c-Myc targets and non-c-Myc targets containing the top 100 c-Myc target genes and the bottom 100 genes from the rank ordering. This set of target genes and non-target genes were also used for the Pol II TR analysis following 10058-F4 treatment using rank normalized ChIP-seq density data (Figure 6D).

In order to determine the significance of these changes, we calculated the Pol II density for each gene in the promoter, and the gene body plus gene end region, expanding the target genes to the top 1000 and bottom 1000 genes ranked by c-Myc levels were respectively used as a target and non-target set for further statistical analysis. Regions are delineated in Figure 6. Averages of the Pol II occupancy at c-Myc targets show a loss of Pol II in the gene body and gene end following 10058-F4 treatment. No such loss is seen in non-target genes. Additionally, levels of Pol II at the promoter appear unchanged following 10058-F4 treatment in both target and non-target genes. We used a Welch's T-Test to determine whether distributions of gene densities in control versus 10058-F4 treated cells could be generated by the same distribution in each region for target and non-target genes. We found significant changes in Pol II occupancy only in the gene body of cMyc target genes ($p=7.341e-06$) but not the promoter region

(p=0.4536) of c-Myc genes. Non c-Myc target genes did not have a statistically significant change in either gene region ($p<0.001$). These data suggest that 10058-F4 causes a loss in Pol II occupancy specifically at c-Myc target genes.

**Co-immunoprecipitation analysis**

Co-immunoprecipitation studies were done in mES cells using IgG: Upstate/Millipore, anti-c-Myc: Santa Cruz sc-764, anti-Max: Santa Cruz sc-197, anti-Cdk9 Santa Cruz sc-484, anti-CycT1: gift from David Price (Peng et al., 1998). Max was immunoprecipitated from mES cell lysates made in Lysis Buffer (20mM Tris pH8.0, 150mM NaCl, 10% glycerol, 1% NP-40, 2mM EDTA and protease inhibitors) using anti-IgG control, anti-c-Myc, anti-Max, anti-Cdk9 and anti-CycT1 and incubated overnight at $4^{o}$C. CycT1 and Cdk9 were immunoprecipitated with mES cell lysates made in Lysis Buffer using anti-IgG control, anti-c-Myc, anti-Max and anti-Cdk9 and incubated for 3 hours at $4^{o}$C. Immunoprecipitates were washed three times with lysis buffer and proteins were analyzed by SDS-PAGE gel electrophoresis followed by Western blot analysis.

**Western blots**

Western blots were done following standard protocols. Antibodies used for Western blots were as follows: Spt5: gift of Yuki Yamaguchi and Hiroshi Handa (Wada et al., 1998); NelfA: gift of Yuki Yamaguchi and Hiroshi Handa; Cdk9: Santa Cruz sc-484 and sc-8338, Ser2P Pol II: Abcam (H5 clone) ab24758; Ser5P Pol II: Abcam ab5131; Brg1: Santa Cruz sc-10768; Max: Santa Cruz sc-

197; c-Myc: Santa Cruz-764; Cyclin T1: gift of David Price; Oct4: Santa Cruz sc-5279; and TBP: Abcam ab818. Secondary antibodies used were anti-rabbit IgG (HRP-conjugated; GE Healthcare - NA934V), anti-mouse IgG (HRP-conjugated; GE Healthcare – NA931V), anti-Protein A (HRP-conjugated; GE Healthcare – NA9120V), anti-sheep (HRP-conjugated; Santa Cruz – sc2770) and anti-mouse IgG+IgM (HRP-conjugated; Jackson ImmunoResearch – 115-035-044).

**Expression analysis**

RNA was extracted from mES cells treated with 10058-F4 or DMSO alone with biological replicates using Qiagen Qiashredder and RNeasy kits. Residual DNA was degraded using DNA-free kit for DNAse treatment (Ambion). cDNA was generated from the DNA-free RNA using Superscript III First Strand reverse transcriptase PCR kits. Expression was determined using quantitative PCR analysis using Taqman assays in triplicate against Gapdh (Mm99999915_g1), Rnf2 (Mm00803321_m1), Brg1 (Mm01151944_m1), Atic (Mm00546566_m1), Bax (Mm00432050_m1), Nol5 (Mm00479705_m1), Brca2 (Mm00464784_m1), and Zfp451 (Mm00659728_m1). Expression was normalized against Gapdh internal control and displayed as a percentage of expression in the DMSO alone control. c-Myc mRNA levels were quantitated following sh c-Myc knockdown analysis using duplicate Taqman assays for c-Myc (Mm00487804_m1) and normalized against Gapdh.

**Supplemental References**

Adelman, K., Wei, W., Ardehali, M.B., Werner, J., Zhu, B., Reinberg, D., and Lis, J.T. (2006). Drosophila Paf1 modulates chromatin structure at actively transcribed genes. Mol Cell Biol *26*, 250-260.

Arabi, A., Wu, S., Ridderstrale, K., Bierhoff, H., Shiue, C., Fatyol, K., Fahlen, S., Hydbring, P., Soderberg, O., Grummt, I.*, et al.* (2005). c-Myc associates with ribosomal DNA and activates RNA polymerase I transcription. Nat Cell Biol *7*, 303-310.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev *23*, 2484-2489.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G.*, et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J.*, et al.* (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106-1117.

Fang, Z.H., Dong, C.L., Chen, Z., Zhou, B., Liu, N., Lan, H.F., Liang, L., Liao, W.B., Zhang, L., and Han, Z.C. (2008). Transcriptional regulation of survivin by c-Myc in BCR/ABL-transformed cells: implications in antileukemic strategy. J Cell Mol Med.

Faumont, N., Durand-Panteix, S., Schlee, M., Gromminger, S., Schuhmacher, M., Holzel, M., Laux, G., Mailhammer, R., Rosenwald, A., Staudt, L.M.*, et al.* (2009). c-Myc and Rel/NF-kappaB are the two master transcriptional systems activated in the latency III program of Epstein-Barr virus-immortalized B cells. J Virol *83*, 5014-5027.

Follis, A.V., Hammoudeh, D.I., Wang, H., Prochownik, E.V., and Metallo, S.J. (2008). Structural rationale for the coupled binding and unfolding of the c-Myc oncoprotein by small molecules. Chem Biol *15*, 1149-1155.

Guenther, M.G., Lawton, L.N., Rozovskaia, T., Frampton, G.M., Levine, S.S., Volkert, T.L., Croce, C.M., Nakamura, T., Canaani, E., and Young, R.A. (2008). Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. Genes Dev *22*, 3403-3408.

Hailesellasse Sene, K., Porter, C.J., Palidwor, G., Perez-Iratxeta, C., Muro, E.M., Campbell, P.A., Rudnicki, M.A., and Andrade-Navarro, M.A. (2007). Gene function in early mouse embryonic stem cell differentiation. BMC Genomics *8*, 85.

Hammoudeh, D.I., Follis, A.V., Prochownik, E.V., and Metallo, S.J. (2009). Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. J Am Chem Soc *131*, 7390-7401.

Khanna, A., Bockelman, C., Hemmes, A., Junttila, M.R., Wiksten, J.P., Lundin, M., Junnila, S., Murphy, D.J., Evan, G.I., Haglund, C.*, et al.* (2009). MYC-dependent regulation and prognostic role of CIP2A in gastric cancer. J Natl Cancer Inst *101*, 793-805.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049-1061.

Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M.*, et al.* (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell *11*, 721-729.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006a). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006b). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat Protoc *1*, 729-748.

Lee, W.H., Liu, F.H., Lin, J.Y., Huang, S.Y., Lin, H., Liao, W.J., and Huang, H.M. (2009). JAK pathway induction of c-Myc critical to IL-5 stimulation of cell proliferation and inhibition of apoptosis. J Cell Biochem *106*, 929-936.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J.*, et al.* (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.

Matoba, R., Niwa, H., Masui, S., Ohtsuka, S., Carter, M.G., Sharov, A.A., and Ko, M.S. (2006). Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. PLoS One *1*, e26.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.

Peng, J., Zhu, Y., Milton, J.T., and Price, D.H. (1998). Identification of multiple cyclin subunits of human P-TEFb. Genes Dev *12*, 755-762.

Pokholok, D.K., Hannett, N.M., and Young, R.A. (2002). Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. Mol Cell *9*, 799-809.

Sampson, V.B., Rong, N.H., Han, J., Yang, Q., Aris, V., Soteropoulos, P., Petrelli, N.J., Dunn, S.P., and Krueger, L.J. (2007). MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. Cancer Res *67*, 9762-9770.

Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. Nat Rev Mol Cell Biol *7*, 557-567.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol *9*, 1428-1435.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F.*, et al.* (1998). DSIF, a novel transcription

elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev *12*, 343-356.

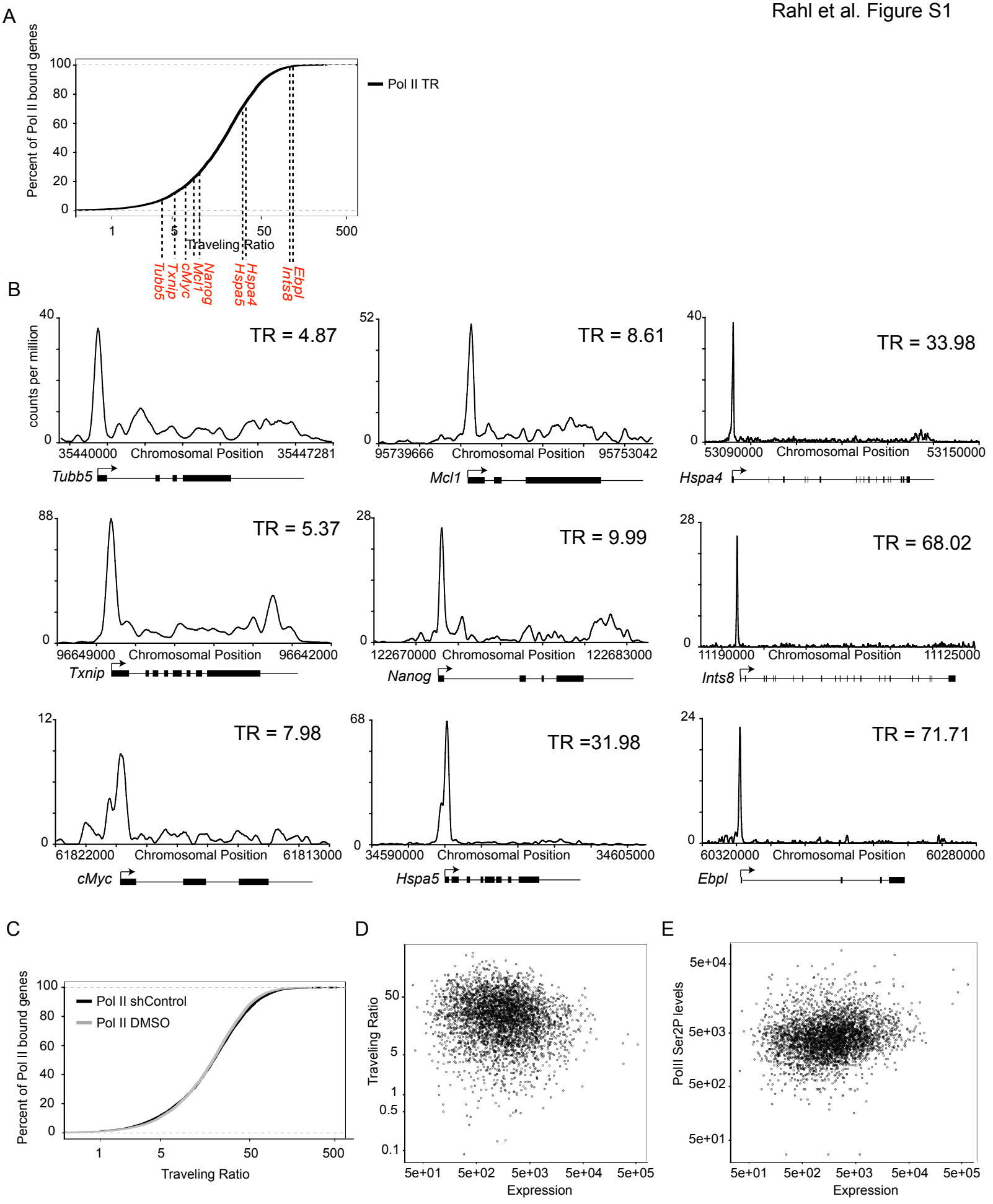
Wang, H., Hammoudeh, D.I., Follis, A.V., Reese, B.E., Lazo, J.S., Metallo, S.J., and Prochownik, E.V. (2007). Improved low molecular weight Myc-Max inhibitors. Mol Cancer Ther *6*, 2399-2408.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Mol Cell *21*, 227-237.

Zhu, B., Mandal, S.S., Pham, A.D., Zheng, Y., Erdjument-Bromage, H., Batra, S.K., Tempst, P., and Reinberg, D. (2005). The human PAF complex coordinates transcription with events downstream of RNA synthesis. Genes Dev *19*, 1668-1673.

**Supplemental Figure**

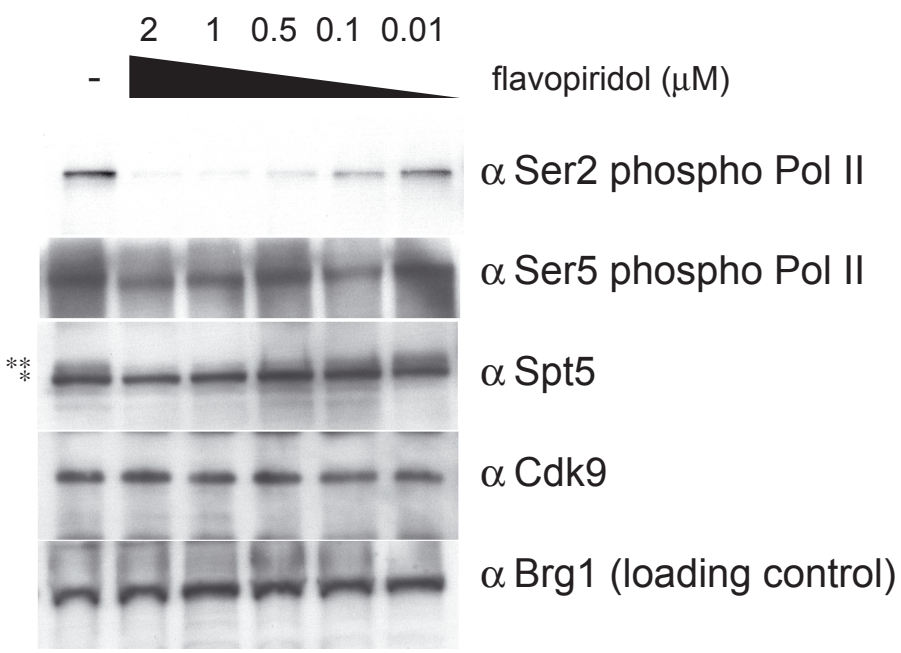
Rahl et al. Figure S1



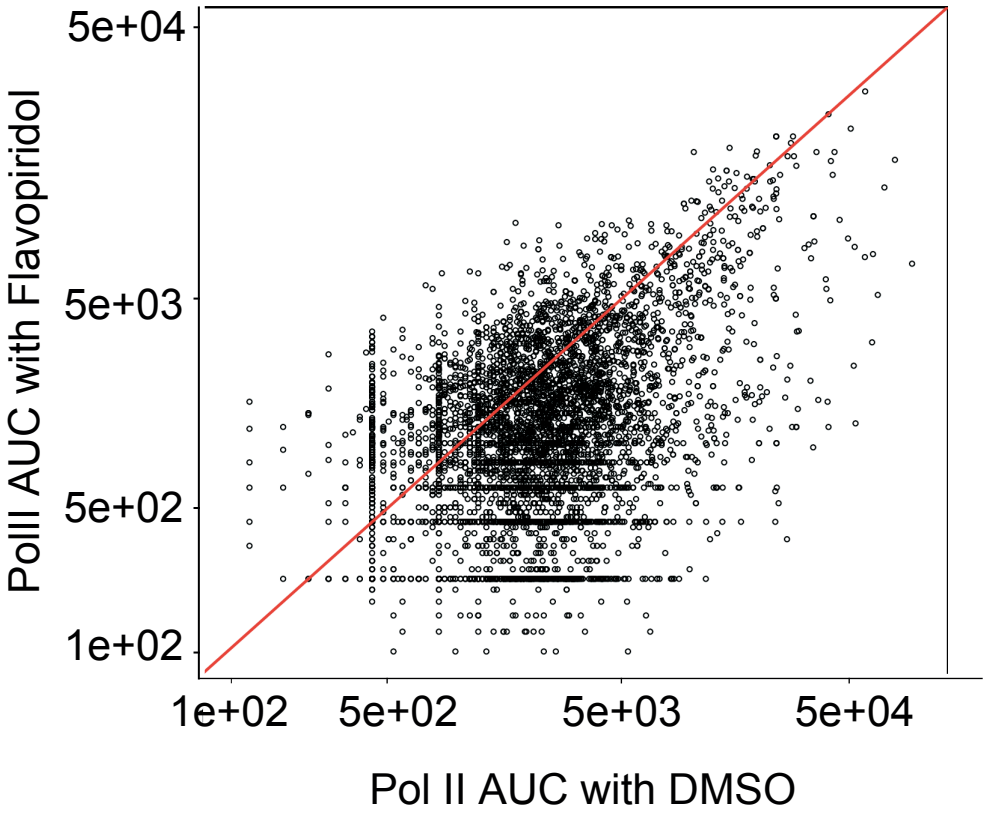
**Supplemental Figure**

Rahl et al. Figure S2

A



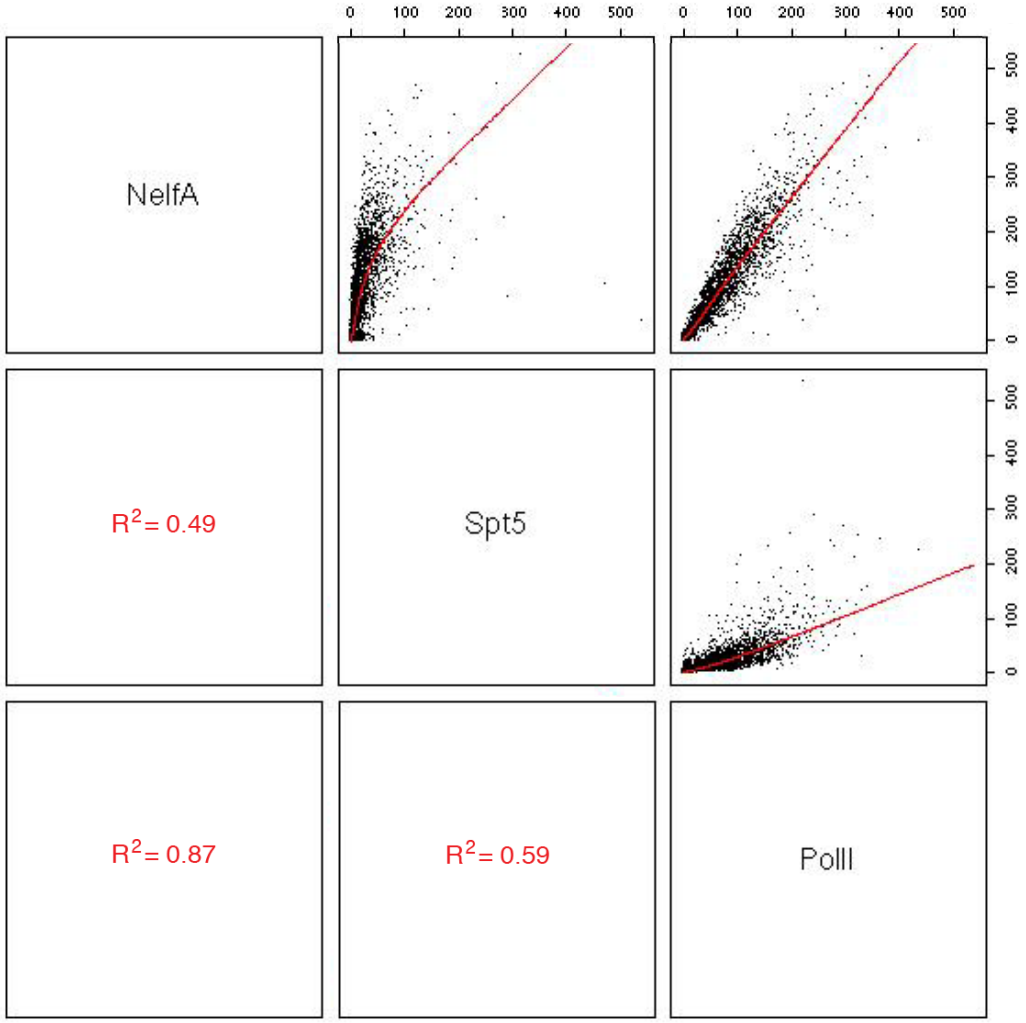

B



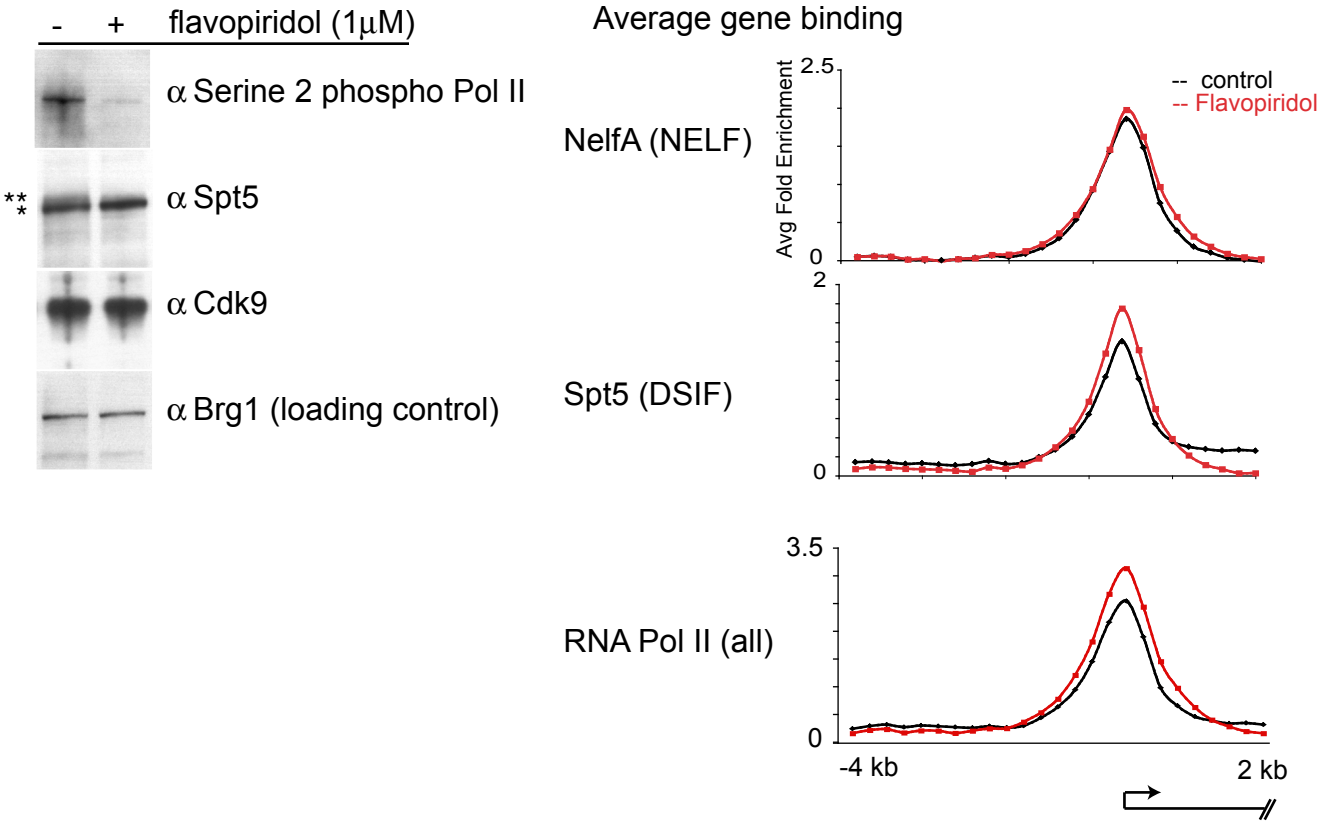
**Supplemental Figure**

Rahl et al. Figure S3

A

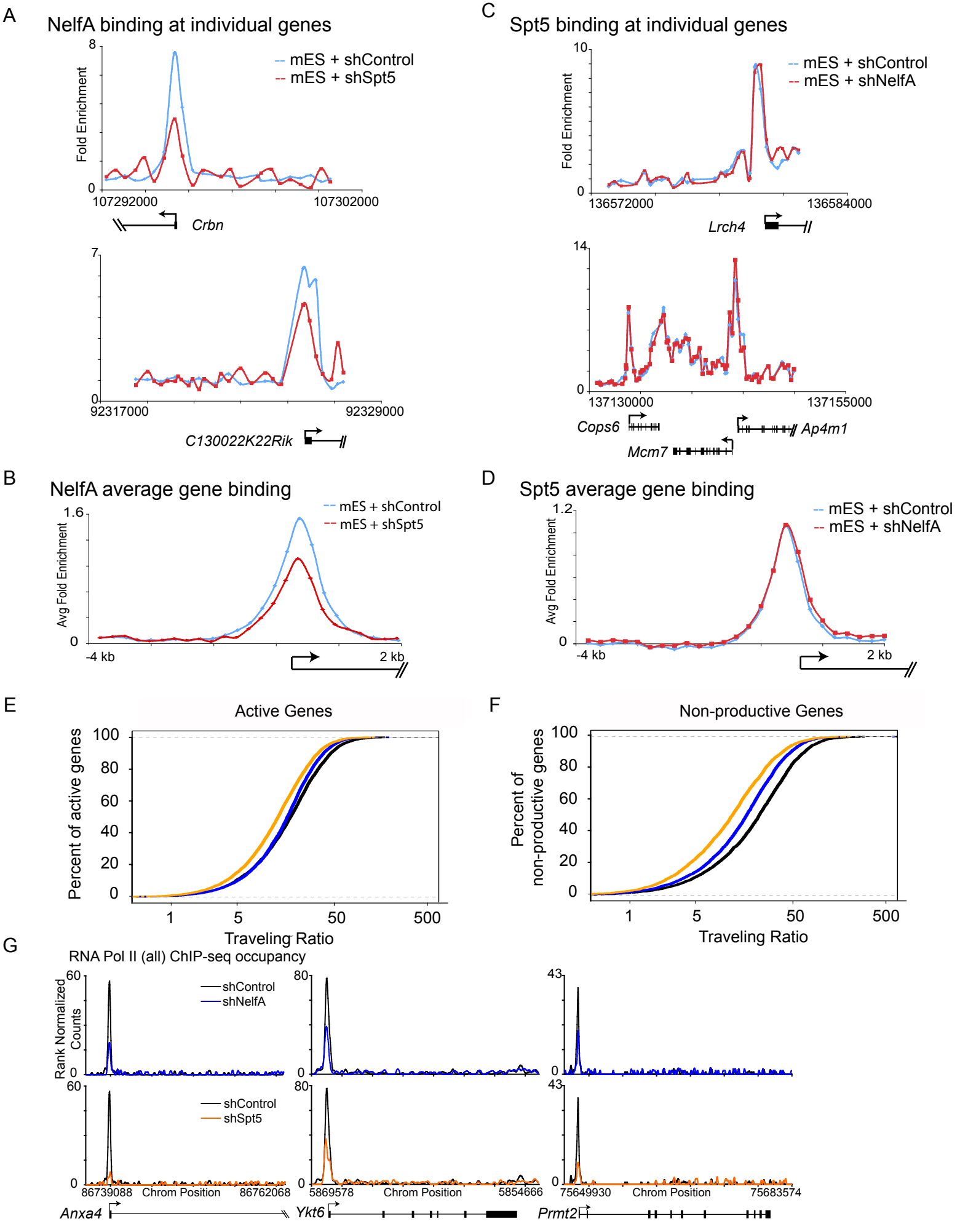



B



**Supplemental Figure**

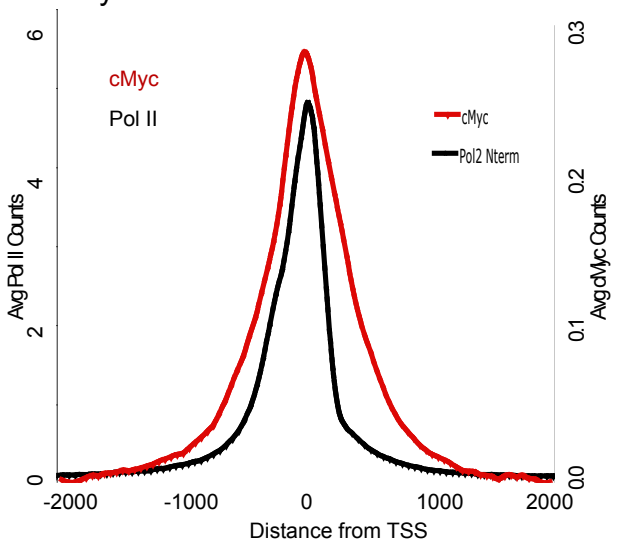
Rahl et al. Figure S4



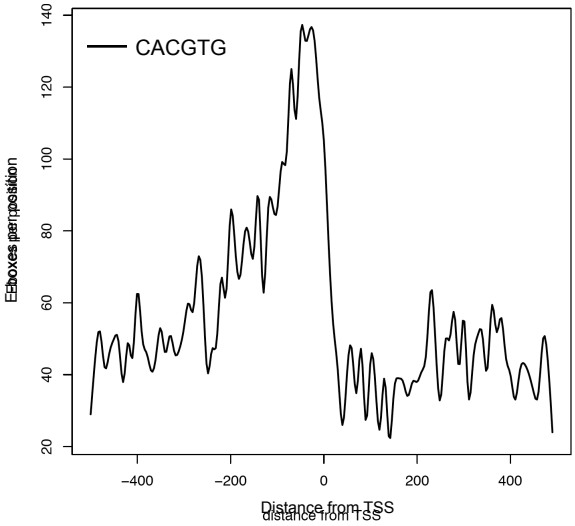
**Supplemental Figure**

Rahl et al. Figure S5

A




B



**Supplemental Figure**

Rahl et al. Figure S6