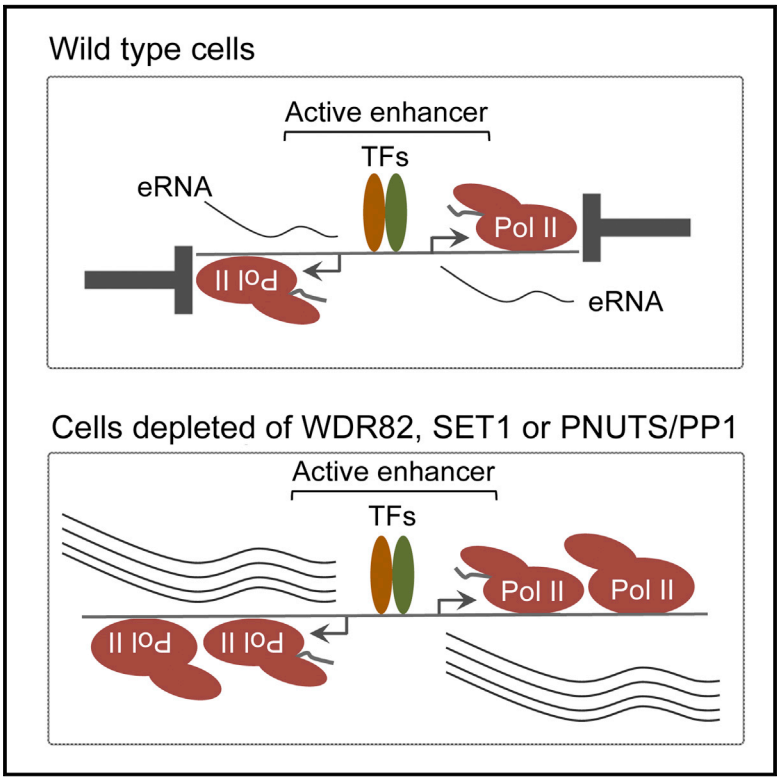
Article

Molecular Cell

Transcription of Mammalian *cis*-Regulatory Elements Is Restrained by Actively Enforced Early Termination

Graphical Abstract



Authors

Liv M.I. Austenaa, Iros Barozzi, Marta Simonatto, ..., Mattia Pelizzola, Wouter de Laat, Gioacchino Natoli

Correspondence

liv.austenaa@ieo.eu (L.M.I.A.), gioacchino.natoli@ieo.eu (G.N.)

In Brief

Upon recruitment to enhancers and promoters, RNA polymerase II synthesizes short and poorly abundant non-coding transcripts. Austenaa et al. show that the inefficient synthesis of such ncRNAs results from actively enforced early transcription termination, whose reversal leads to the abundant production of extended ncRNAs from active *cis*-regulatory elements.

Highlights

- Enhancers and promoters produce Pol II-dependent short ncRNAs
- WDR82 targets to Pol II both the SET1 H3K4 methyltransferase and the PP1 phosphatase
- Depletion of WDR82, SET1, or PP1 impaired Pol II transcription termination
- Enhancers and promoters generated long and abundant ncRNAs upon WDR82 depletion

Accession Numbers

GSE66955

Austenaa et al., 2015, Molecular Cell *60*, 460–474
November 5, 2015 ©2015 Elsevier Inc.
http://dx.doi.org/10.1016/j.molcel.2015.09.018

CellPress

Molecular Cell
Article

Transcription of Mammalian *cis*-Regulatory Elements Is Restrained by Actively Enforced Early Termination

Liv M.I. Austenaa,[1,4,*] Iros Barozzi,[1,4,5] Marta Simonatto,[1,4] Silvia Masella,[1] Giulia Della Chiara,[1] Serena Ghisletti,[1] Alessia Curina,[1] Elzo de Wit,[2,6] Britta A.M. Bouwman,[2] Stefano de Pretis,[3] Viviana Piccolo,[1] Alberto Termanini,[1] Elena Prosperini,[1] Mattia Pelizzola,[3] Wouter de Laat,[2] and Gioacchino Natoli[1,*]

[1]Department of Experimental Oncology, European Institute of Oncology (IEO), Via Adamello 16, 20139 Milan, Italy
[2]Hubrecht Institute-KNAW & University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands
[3]Center for Genomic Sciences of IIT@SEMM, Fondazione Istituto Italiano di Tecnologia (IIT), Via Adamello 16, 20139 Milan, Italy
[4]Co-first author
[5]Present address: Genomics Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[6]Present address: Division of Gene Regulation, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands
*Correspondence: liv.austenaa@ieo.eu (L.M.I.A.), gioacchino.natoli@ieo.eu (G.N.)
http://dx.doi.org/10.1016/j.molcel.2015.09.018

SUMMARY

Upon recruitment to active enhancers and promoters, RNA polymerase II (Pol II) generates short non-coding transcripts of unclear function. The mechanisms that control the length and the amount of ncRNAs generated by *cis*-regulatory elements are largely unknown. Here, we show that the adaptor protein WDR82 and its associated complexes actively limit such non-coding transcription. WDR82 targets the SET1 H3K4 methyltransferases and the nuclear protein phosphatase 1 (PP1) complexes to the initiating Pol II. WDR82 and PP1 also interact with components of the transcriptional termination and RNA processing machineries. Depletion of WDR82, SET1, or the PP1 subunit required for its nuclear import caused distinct but overlapping transcription termination defects at highly expressed genes and active enhancers and promoters, thus enabling the increased synthesis of unusually long ncRNAs. These data indicate that transcription initiated from *cis*-regulatory elements is tightly coordinated with termination mechanisms that impose the synthesis of short RNAs.

INTRODUCTION

Mammalian genomes produce a large number of non-coding RNAs (ncRNAs), in addition to mRNAs (Carninci et al., 2005). ncRNAs are generated by bona fide genes but also as a consequence of RNA polymerase II (Pol II) transcription of *cis*-regulatory regions, including promoters (Kapranov et al., 2007; Preker et al., 2008) and enhancers (De Santa et al., 2010; Kim et al., 2010; Koch et al., 2011). Generation of enhancer RNAs (eRNAs), which include short species generated by bidirectional transcription from enhancer centers and more infrequently long unidirectional transcripts (Natoli and Andrau, 2012), correlates with enhancer acetylation and activity (Wang et al., 2011). Transcription initiated at *cis*-regulatory elements was shown to cause functional effects mediated either by the ncRNA (Wang et al., 2008; Lam et al., 2013; Li et al., 2013; Melo et al., 2013; Ørom et al., 2010) or by the act of transcription (Schmitt et al., 2005). However, other studies showed this transcription to be unsuppressed noise (Hah et al., 2013).

Promoters and enhancers share similar mechanisms of transcription initiation and common architectural principles, including the presence of the same core promoter sequence elements (Core et al., 2014) and their nucleosome depletion (Barozzi et al., 2014; Schones et al., 2008). The outcome of Pol II recruitment to active *cis*-regulatory elements is determined by the interplay between the mechanisms that control transcription initiation, elongation, and termination and those that control the stability of the nascent transcripts (Jensen et al., 2013). At promoters, poly(A) signals (PAS) in the nascent divergent transcripts (promoter proximal transcripts, or PROMPTs) are recognized by the cleavage and polyadenylation specificity factor (CPSF), resulting in transcript cleavage, transcriptional termination (Almada et al., 2013; Ntini et al., 2013), and transcript degradation by the nuclear exosome (Flynn et al., 2011). Both early termination and exosome-mediated degradation are stimulated by the recognition of the 5′ cap of the nascent PROMPT by a cap-binding complex (CBC) and its associated subunits (Andersen et al., 2013; Hallais et al., 2013), pointing to a model whereby the proximity of the PAS to the 5′ cap determines the short length and the instability of PROMPTs. Instead, in the mRNA coding direction, PASs are depleted and their usage is suppressed (Almada et al., 2013; Ntini et al., 2013). When Pol II reaches the gene 3′ end, PAS recognition by CPSF triggers transcription termination. In this case, however, the distance of the PAS from the 5′ cap and the associated CBC would explain why, differently from PROMPTs, the cleaved pre-mRNAs are not degraded by the exosome.

Whereas eRNAs, like PROMPTs, are capped, short and unstable molecules (De Santa et al., 2010; Andersson et al., 2014), it is not known whether their instability is similarly linked to premature

transcription termination and why enhancer-initiated transcription is biased toward the production of short RNAs.

Several studies (mainly in yeast) suggest a role of chromatin in limiting pervasive non-coding transcription. First, chromatin configurations that are non-permissive for transcription are generated by chromatin remodelers (Hainer et al., 2015; Whitehouse et al., 2007), chromatin assembly factors (Marquardt et al., 2014), and histone deacetylases (Carrozza et al., 2005; Pinskaya et al., 2009). Second, in yeast, H3K4 methylation by the SET1/COMPASS complex was shown to control transcription termination (Terzi et al., 2011). SET1/COMPASS, the only yeast H3K4 methyltransferase (Shilatifard, 2012), was linked to a termination pathway responsible for termination of cryptic unstable transcripts (CUTs), the yeast equivalent of PROMPTs (Porrua and Libri, 2015).

The mammalian SET family includes MLL1-4 and SET1A/B. The SET1A/B complexes specifically contain CFP1, which is required for H3K4me3 deposition at CpG islands (CpGi) (Thomson et al., 2010; Clouaire et al., 2012) and WDR82 (Lee and Skalnik, 2008; Wu et al., 2008), which recognizes Ser5 Pol II and mediates cotranscriptional SET1A/B recruitment and deposition of H3K4me3 (Ng et al., 2003). WDR82 is also found in a protein phosphatase 1 (PP1) complex that is directed to the nucleus by the PP1 nuclear targeting subunit (PNUTS) (Lee et al., 2010; van Nuland et al., 2013). In *Drosophila*, PNUTS is required for PP1 recruitment to active transcription sites, where it also dephosphorylates the Pol II carboxy-terminal repeat domain (CTD) (Ciurciu et al., 2013). Overall, while WDR82 specifically recognizes the initiating Pol II and mediates its association with the SET1A/B and PNUTS/PP1 complexes, the functional impact of tethering these two enzymatic activities to Pol II is unknown.

Additional evidence points to the involvement of WDR82 and PNUTS/PP1 in transcription termination. The *S. cerevisiae* WDR82 and PP1 orthologs (as well as the PNUTS ortholog in *S. pombe*) (Vanoosthuyse et al., 2014) are associated with the cleavage and polyadenylation factor (CPF) (Dichtl et al., 2002; He et al., 2003) and their absence causes polyadenylation and termination defects (Cheng et al., 2004; Dichtl et al., 2004; Nedea et al., 2008). Moreover, PP1 and PNUTS are components of the molecular machinery processing the pre-mRNA 3′ end in human cells (Shi et al., 2009). A summary of the complexes mentioned is shown in Figure S1A.

Here, we show that the depletion of WDR82, PNUTS, or SET1A/B resulted in distinct but overlapping Pol II termination defects affecting highly expressed genes and active enhancers and promoters. Our data indicate that actively enforced early transcription termination is an essential mechanism limiting transcription of *cis*-regulatory sequences in mammals.

RESULTS

Wdr82 Depletion Reduced H3K4me3 at Non-CpGi Promoters

We first depleted WDR82 in mouse primary macrophages using retrovirus-mediated short hairpin RNA (shRNA) delivery. Consistent with previous observations (Lee and Skalnik, 2008; Wu et al., 2008), WDR82 depletion reduced SET1A protein levels without affecting its mRNA (Figure S1B). Levels of RBBP5, a SET1 and MLL core complex component, were not affected, while PNUTS was upregulated (Figure S1B). The impact of WDR82 depletion on H3K4me3 was analyzed by chromatin immunoprecipitation sequencing (ChIP-seq). In untreated macrophages, 550 out of the 15,662 (3.5%) constitutive H3K4me3 peaks were reduced upon WDR82 depletion (MACS [model-based analysis of chIP-seq] score $\geq$ 100), while effects on histone acetylation were minimal (Figure 1A and Table S1). Compared to the reduction of H3K4me3 bulk levels reported in WDR82-depleted HeLa cells (Wu et al., 2008), such limited effects are likely to occur because macrophages divide poorly in culture, implying an overall lower turnover of histones and their modifications compared to cycling cells. In response to lipopolysaccharide (LPS) stimulation, however, 60.6% of the induced H3K4me3 peaks were reduced in cells depleted of WDR82 (Figure 1A), indicating that most de novo deposited H3K4me3 depends on SET1A/B. Instead, WDR82 depletion had minimal consequences on LPS-inducible histone acetylation (Figure 1A). Reduction of H3K4me3 correlated with transcriptional activity, as indicated by 4-thiouridine (4sU) labeling of nascent RNAs (Figure 1B, p = 1.6e-11, Wilcoxon rank-sum test), indicating that genes with higher transcription-related nucleosome turnover rely on SET1A/B to maintain H3K4me3 levels.

Loss of H3K4me3 anti-correlated with the promoter G+C content (Figure 1C), and less than 20% of the genes with H3K4me3 loss contained a CpGi, as opposed to 65% of the genes that lost H3K4me3 in *Mll2*$^{-/-}$ macrophages (Figure 1D) (Austenaa et al., 2012). Genes (or gene clusters) in which H3K4me3 was lost upon WDR82 depletion were frequently bracketed by CpGi-associated genes with unaffected levels of H3K4me3 (Figure 1E). Consistent results were obtained with a distinct shRNA (shRNA #2, Figures S1B and S1C). Antibodies for WDR82 did not generate robust ChIP data, thus preventing the analysis of its genomic distribution.

CpG islands enable constitutive and homogeneous gene expression across tissues. Reciprocally, the absence of a CpGi enables both tight gene regulation and broad dynamic range of expression across tissues or upon stimulation (Bhatt et al., 2012). Consistently, genes at which H3K4me3 required WDR82 displayed broad expression variation across tissues and were highly expressed in cells of the hematopoietic system, while the expression of genes with MLL2-dependent H3K4me3 (mainly CpGi-containing genes, Figure 1) was constant across tissues (Figure S1D).

Transcriptional Consequences of Wdr82 Depletion

WDR82 depletion altered the expression of a few hundred poly(A)+ transcripts, although these effects were of limited magnitude when compared to LPS-induced gene expression changes (Figure S2A and Table S2). Because steady-state transcripts do not reflect ongoing transcription, we used 4sU incorporation (Rabani et al., 2011) to analyze nascent transcripts. We identified 426 differentially expressed genes (DEGs) in WDR82-depleted cells, with an overwhelming majority of upregulated genes: 90.6% of the DEGs were more than 2-fold upregulated at least at one time point, while 39.2% of them were upregulated at all time points (Figure S2B).

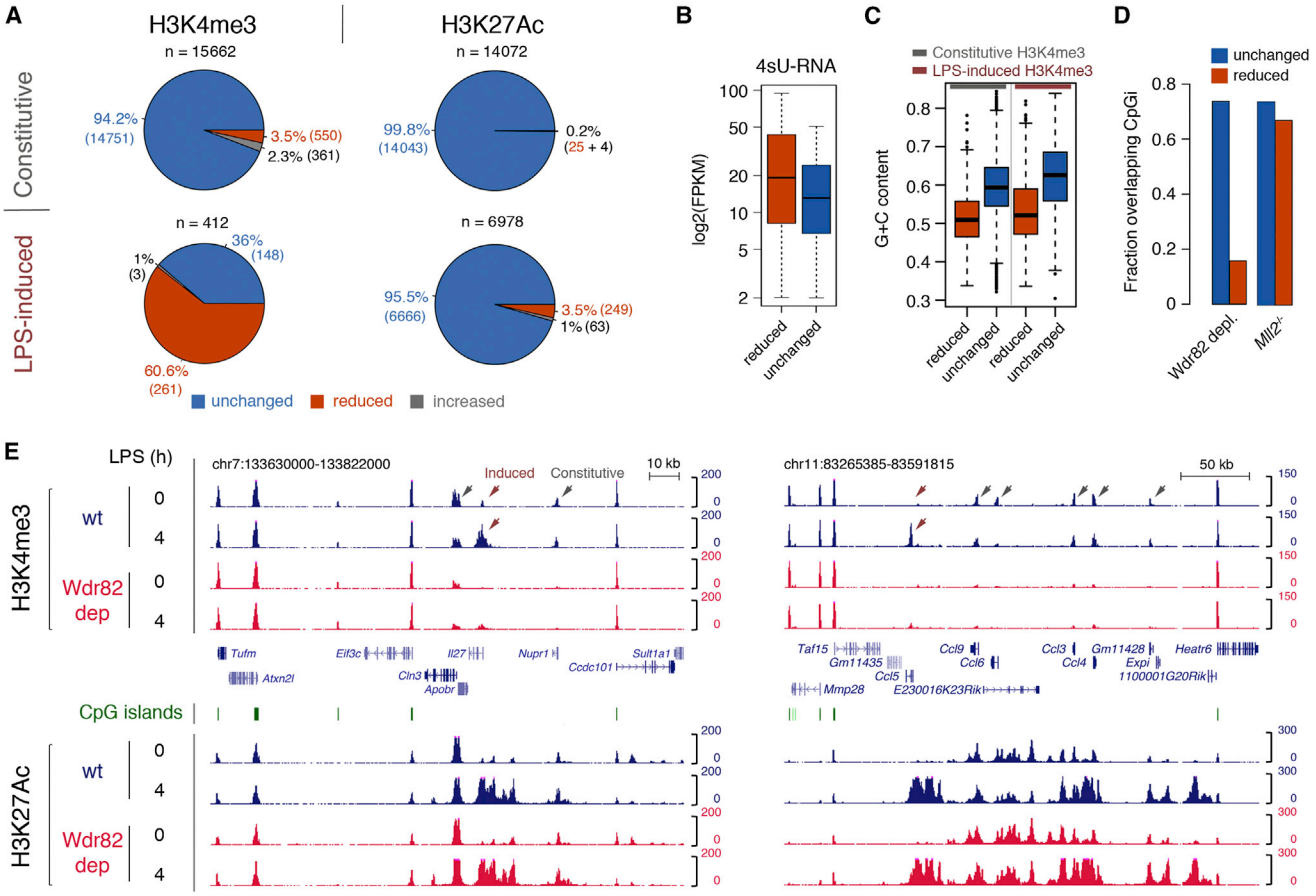


Figure 1. Effects of WDR82 Depletion on Histone H3K4me3

(A) Constitutive and LPS-inducible H3K4me3 and H3K27Ac in WDR82-depleted macrophages.

(B) Correlation between gene transcription (4sU) and H3K4me3 changes in WDR82-depleted untreated macrophages.

(C) Correlation between G+C content at promoters and H3K4me3 changes induced by WDR82 depletion in untreated and LPS-treated macrophages.

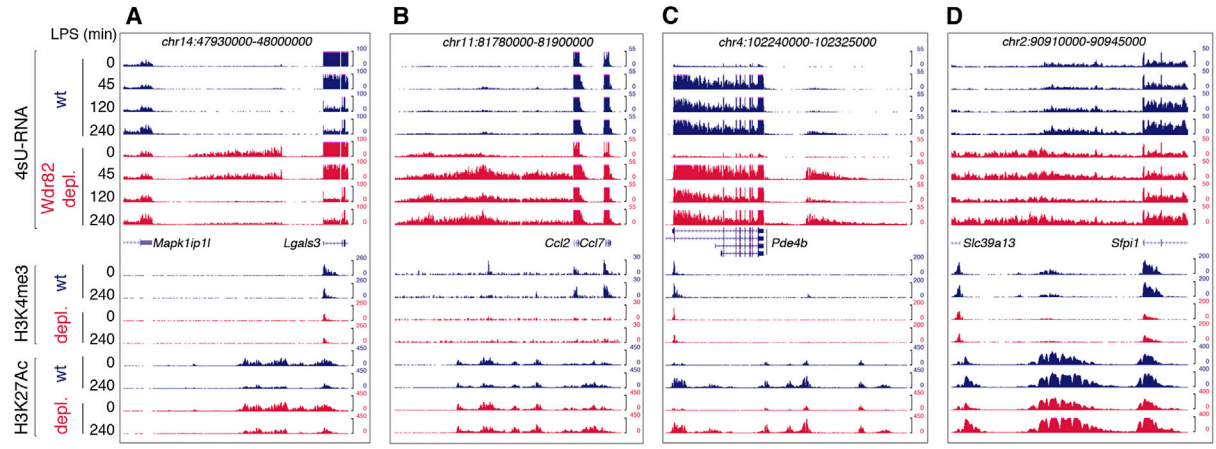
(D) Histogram showing the fraction of genes with reduced or unaffected H3K4me3 in WDR82-depleted or *Mll2*$^{-/-}$ macrophages that contain a CpGi.

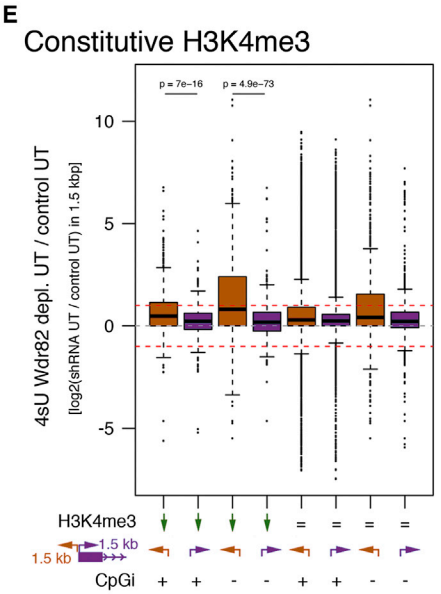
(E) Two representative ChIP-seq snapshots showing the effects of WDR82 depletion (dep) on constitutive or LPS-inducible H3K4me3.

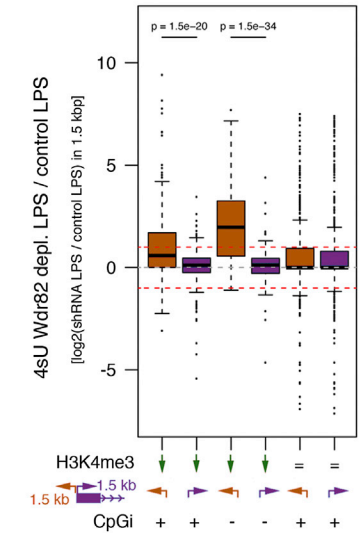
See also Figure S1 and Table S1.

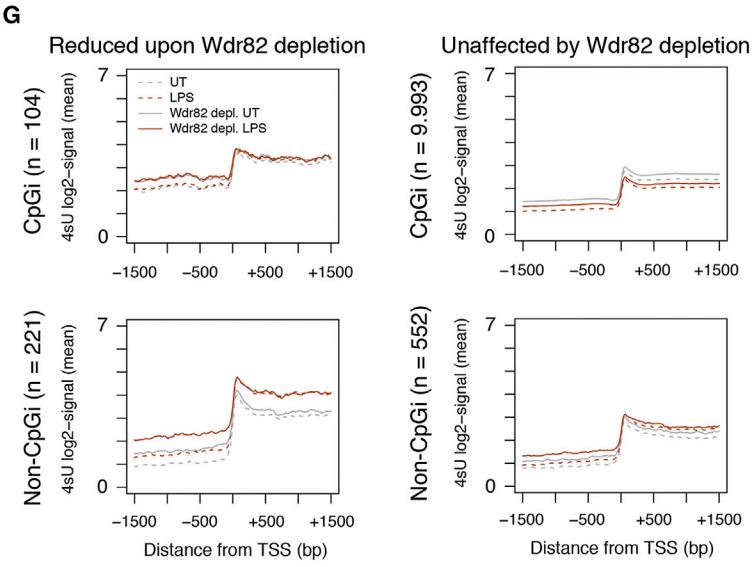
Visual browsing of the 4sU data revealed that the most common effect of WDR82 depletion was the occurrence of basal (Figure 2A) or LPS-inducible (Figure 2B) extragenic transcription at genomic regions that in wild-type (wt) macrophages were not transcribed. In other instances, extragenic regions with low but detectable transcription in normal macrophages showed either a higher (Figure 2C) or a more extended (Figure 2D) 4sU signal. RT-PCR validations (Figure S2C) confirmed the increased extragenic transcription outside of the boundaries of genes whose expression was unaffected. The rescue of WDR82 expression with an shRNA-resistant FLAG-tagged WDR82 (Figure S2D) reversed PNUTS upregulation and the increased expression of all extragenic RNAs tested. An shRNA that inefficiently depleted WDR82 (shRNA #3) caused no effects on extragenic transcription (Figure S2E and Table S3). Finally, an shRNA depleting MENIN, a subunit of MLL1/2 complexes, was devoid of effects at the genomic regions where WDR82 depletion increased extragenic transcription (Figure S2E and Table S3).

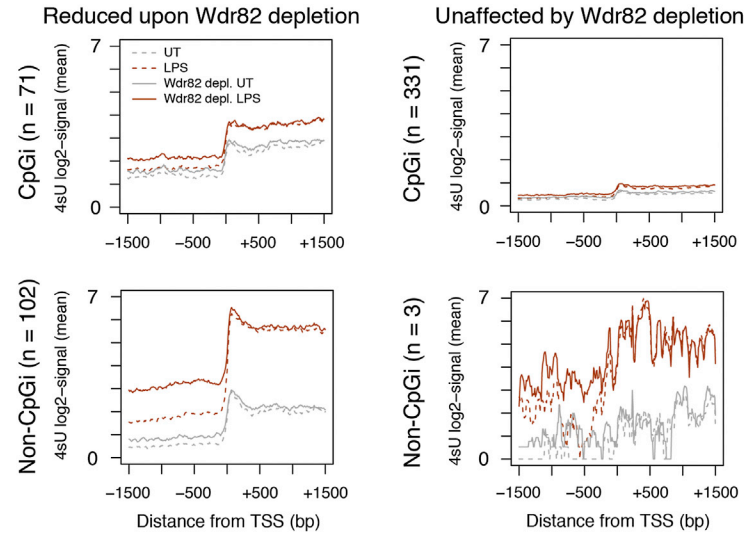
Increased extragenic transcription often occurred in the vicinity of genes whose H3K4me3 was reduced by WDR82 depletion (Figures 2A–2D). To determine the correlation between H3K4me3 reduction and increased extragenic transcription, we divided Ensembl genes based on (1) the presence of constitutive or LPS-inducible H3K4me3, (2) the impact of WDR82 depletion on H3K4me3, and (3) the presence of a CpGi. Then we analyzed the 4sU reads in a ±1.5 kb window centered on gene transcription start sites (TSSs). WDR82 depletion had no major consequences on transcription downstream of the TSS, irrespective of its impact on H3K4me3 (Figures 2E and 2F). Conversely, it significantly increased non-coding transcription upstream of the TSSs of those genes whose H3K4me3 levels were reduced (Figures 2E and 2F). The increase in upstream non-coding transcription was stronger at genes with LPS-inducible H3K4me3 and lacking a CpGi (Figure 2F). Figures 2G and 2H describe the 4sU signal trends around TSSs in the different classes of genes. The one showing the highest dysregulation of











(legend on next page)

upstream non-coding transcription (non-CpGi genes with LPS-induced and WDR82-dependent H3K4me3) (Figure 2H, bottom left panel) also showed the highest level of genic transcription. LPS-inducible, CpGi-negative genes are specifically characterized by a high dynamic range with low basal activity and high induction by LPS (Bhatt et al., 2012), suggesting substantial de novo recruitment of Pol II. Therefore, our data indicate that following massive LPS-induced recruitment to CpGi-negative promoters, Pol II in WT cells efficiently elongated inside the coding region, while the transcriptional activity in the extragenic regions upstream of the promoter was minimal. In the absence of WDR82, however, transcription of such extragenic regions was greatly enhanced. The increased extragenic transcription was usually not associated with changes in the downstream coding region, suggesting that the absence of WDR82 specifically impaired mechanisms that negatively regulate non-coding transcription initiated at promoters.

WDR82 Depletion Increases the Synthesis of Extragenic Transcripts

Attempts to characterize the nascent extragenic transcripts produced in WDR82-depleted cells using Cufflinks were not conclusive (Figure S3), possibly because most of these transcripts are long and unspliced RNAs that differ from the standard transcripts for which Cufflinks was designed. Therefore, we resorted to Spatial Cluster Identification of ChIP-Enriched Regions (SICER) (Zang et al., 2009), an algorithm devised to detect extended regions of signal enrichment and thus suitable to identify transcribed extragenic regions (blocks). The number of extragenic blocks retrieved by SICER in WDR82-depleted macrophages largely exceeded the number of those identified in their matched control (Table S4), indicating that the most common consequence of WDR82 depletion was the upregulation of transcription outside of gene borders. In the analyzed window of time, the number of regions with increased extragenic RNAs peaked at 45 min post-LPS stimulation (2,632 regions); at any time point, only a negligible number of extragenic regions (<50) showed reduced 4sU signals in WDR82-depleted cells (Figure 3A).

The genomic regions with increased 4sU in WDR82-depleted cells had a median size between 6 and 7.5 kb, depending on the time point, with some of them spanning more than 100 kb (Figure 3B).

We next classified the extragenic regions where the 4sU signal increased upon WDR82 depletion using a set of genomic annotations (Figure 3C). We first determined the overlap with long intervening non-coding RNAs (lincRNAs) (Guttman et al., 2009). At 45 min, 10% of the regions in our dataset overlapped annotated mouse lincRNAs. Almost none of these lincRNAs is annotated as expressed in macrophages. Among the remaining regions, 9% of them overlapped super-enhancers (Whyte et al., 2013), 9% overlapped active (acetylated) enhancers, and 17% overlapped poised (H3K4me1-positive but not acetylated) enhancers (Creyghton et al., 2010; Rada-Iglesias et al., 2011). Because many of these aberrant RNAs extended for several kilobases, they overlapped different classes of enhancers. The remaining regions (55%) were associated either with the 10 kb region upstream of gene TSSs or with the 10 kb downstream of gene transcription termination sites (TTSs) (Figure 3C).

The high frequency of increased transcription downstream of TTS pointed to an inefficient transcription termination as the unifying mechanism underlying widespread dysregulated extragenic transcription in WDR82-depleted cells. To explore this possibility, we calculated the rates of synthesis and degradation of genic and extragenic RNAs in WT and WDR82-depleted macrophages (de Pretis et al., 2015). Comparison of 4sU-labeled nascent transcripts and steady-state total RNA (Figure 3D), showed that extragenic RNAs were characterized by a higher degradation rate compared to the genic ones. WDR82 depletion did not affect degradation rates; instead, it greatly increased the synthesis of extragenic RNAs while causing low-magnitude effects at genic regions (Figure 3D).

Taken together, these data suggest that reduced WDR82 levels trigger transcription termination defects leading to the increased synthesis of read-through genic RNAs, as well as of unusually long extragenic transcripts at enhancers and promoters.

Correlation between Dysregulated Extragenic Transcription and Gene Expression Changes

We next correlated the increased extragenic transcription with the observed changes in gene expression (Figure S2). We generated strand-specific 4sU sequencing (4sU-seq) datasets to unambiguously characterize the different types of genic or extragenic RNAs detected. We first focused on the genes showing dysregulated transcription within 10 kb from their TSS or TTS. In most cases, dysregulated non-coding transcription at a TSS or TTS was not associated with detectable changes in the abundance of the corresponding coding RNA; in those cases in which gene expression changes were detected, they were nearly always increases (Figure 4A).

Next, we analyzed all extragenic regions associated with increased transcription in WDR82-depleted cells and measured the distance of each of them from the nearest gene whose transcription was upregulated, unchanged, or downregulated. Increased extragenic transcription occurred closer to genes with increased transcription compared to unaffected ones (Figure 4B).

Figure 2. Correlation between Reduced H3K4me3 and Increased Gene-Proximal Non-coding Transcription in WDR82-Depleted Macrophages

(A–D) Four representative snapshots showing 4sU-seq, H3K4me3, and H3K27Ac sequencing data in control and WDR82-depleted (depl.) macrophages.

(E and F) Impact of WDR82 depletion on nascent promoter-proximal transcripts. 4sU-seq reads in ±1.5 kb relative to annotated TSSs were counted and normalized across datasets. Genes were first divided based on the constitutive (E) or LPS-inducible (F) character of H3K4me3 and then classified based on the presence of a CpGi and on H3K4me3 changes (reduction or no effect) in WDR82-depleted macrophages. (F) Genes without a CpGi and no changes in LPS-inducible H3K4me3 are not shown because of their numerical exiguity (n = 3). p values were computed using a Wilcoxon signed-rank test.

(G and H) Cumulative distributions of 4sU-seq signals in control and WDR82-depleted macrophages in ±1.5 kb relative to annotated TSSs of genes with constitutive (G) or LPS-inducible (H) H3K4me3.

See also Figure S2 and Tables S2 and S3.

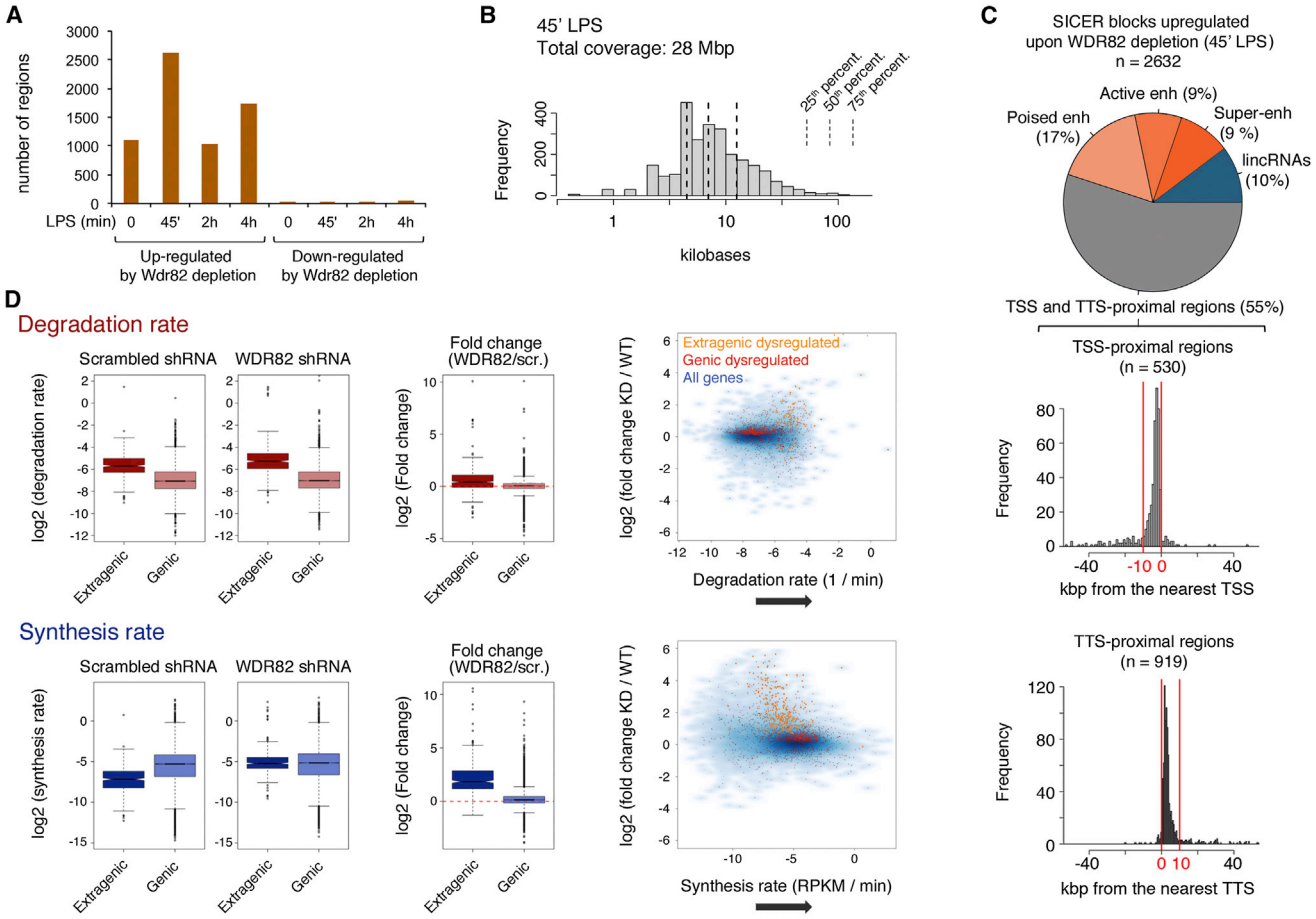


Figure 3. Increased Synthesis of Long Extragenic Transcripts and Genic Read-Through Transcription in WDR82-Depleted Macrophages

(A) Bar plot showing the number of extragenic regions detected by SICER with an increased or decreased 4sU signal in WDR82-depleted macrophages relative to control cells.

(B) Size distribution of extragenic regions covered by upregulated transcription in WDR82-depleted macrophages at 45 min of LPS stimulation. Median, 25th, and 75th percentiles (percent.) are indicated by dashed lines.

(C) Classification of the non-coding transcripts induced in response to WDR82 depletion based on their overlap with annotated genomic features (lincRNAs, super-enhancers, active enhancers, or poised enhancers). Regions not included in these categories (55% of total) were proximal to gene TSS or TTS (histograms at the bottom).

(D) Genic and extragenic synthesis and degradation rates measured by comparing steady-state total RNAs and 4sU-labeled transcripts. The box plots indicate synthesis and degradation rates in control and WDR82-depleted macrophages. The scatter plots on the right report changes in degradation and synthesis rates for individual genes or extragenic regions in untreated macrophages.

See also Figure S3 and Table S4.

Finally, genes were classified based on the aberrant transcription patterns detected at their boundaries in cells depleted of WDR82. We observed four common patterns (Figure 4C). In addition to promoter-divergent transcription and transcription extending beyond the gene 3′ end, we found abundant promoter-convergent and antisense genic transcription. In all four cases, changes in transcription of the associated genes were uncommon and, if detected, they were nearly exclusively upregulations (Figure 4C).

Aberrant Transcription from Highly Acetylated Enhancers

We next focused on changes in extragenic transcription occurring at gene-distal regions, mainly represented by putative enhancers (Figure 3C). Extragenic regions showing detectable histone H3 lysine 27 acetylation (H3K27ac) were clustered together based on their proximity on the linear genome and then ranked according to their total level of H3K27ac (Whyte et al., 2013) (Figure 5A). Acetylation clusters were then used to determine correlations among acetylation levels and other genomic signals: nascent transcription (4sU), total and elongating Pol II (phosphorylated at Ser2 in the CTD) (Komarnitsky et al., 2000), and H3K36me3 (which is associated with transcription elongation and restricted to the body and 3′ end of genes) (Bannister et al., 2005). A correlation was detected between the levels of enhancer acetylation and the increase in the 4sU signal in WDR82-depleted macrophages (Figure 5A), the levels of acetylation being an accurate predictor of the increase in nascent

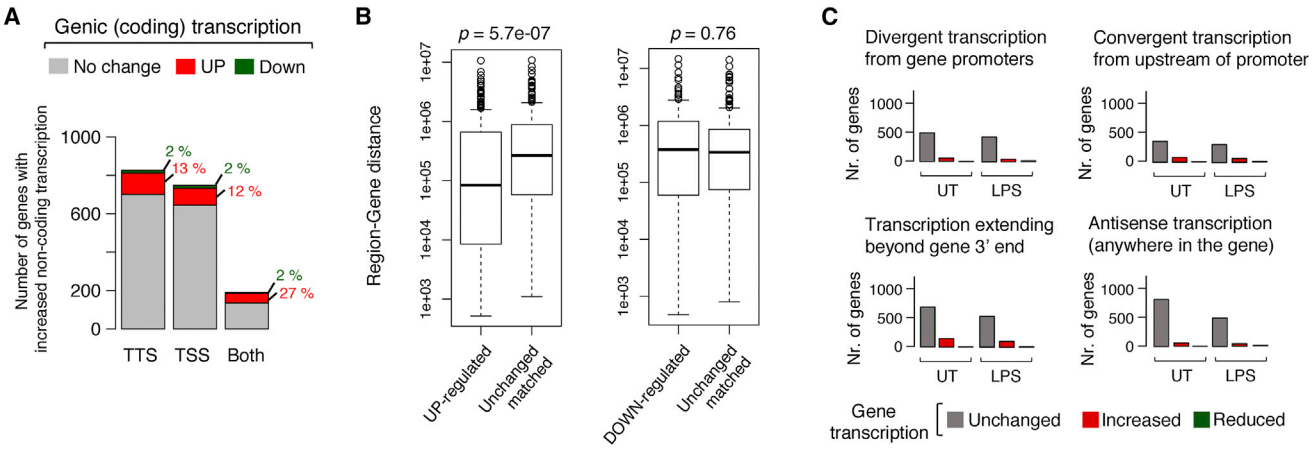


Figure 4. Correlation between Increased Extragenic Transcription and Gene Expression Changes in WDR82-Depleted Macrophages

(A) Number of genes associated with increased TSS- or TTS-proximal extragenic transcription or both. Data were obtained using strand-specific 4sU-seq in macrophages stimulated with LPS (1 h). Colors indicate whether coding transcription was also affected.

(B) Distance (bp) between extragenic regions of increased non-coding transcription and the nearest genes whose transcription is upregulated, downregulated, or unchanged. p values were computed using Wilcoxon rank-sum test.

(C) Four most common types of aberrant non-coding transcription (based on strand-specific 4sU data) occurring within 10 kb from borders of expressed genes in cells depleted of WDR82. Colors indicate whether genic transcription was increased (red), reduced (green), or unaffected (gray). UT, untreated.

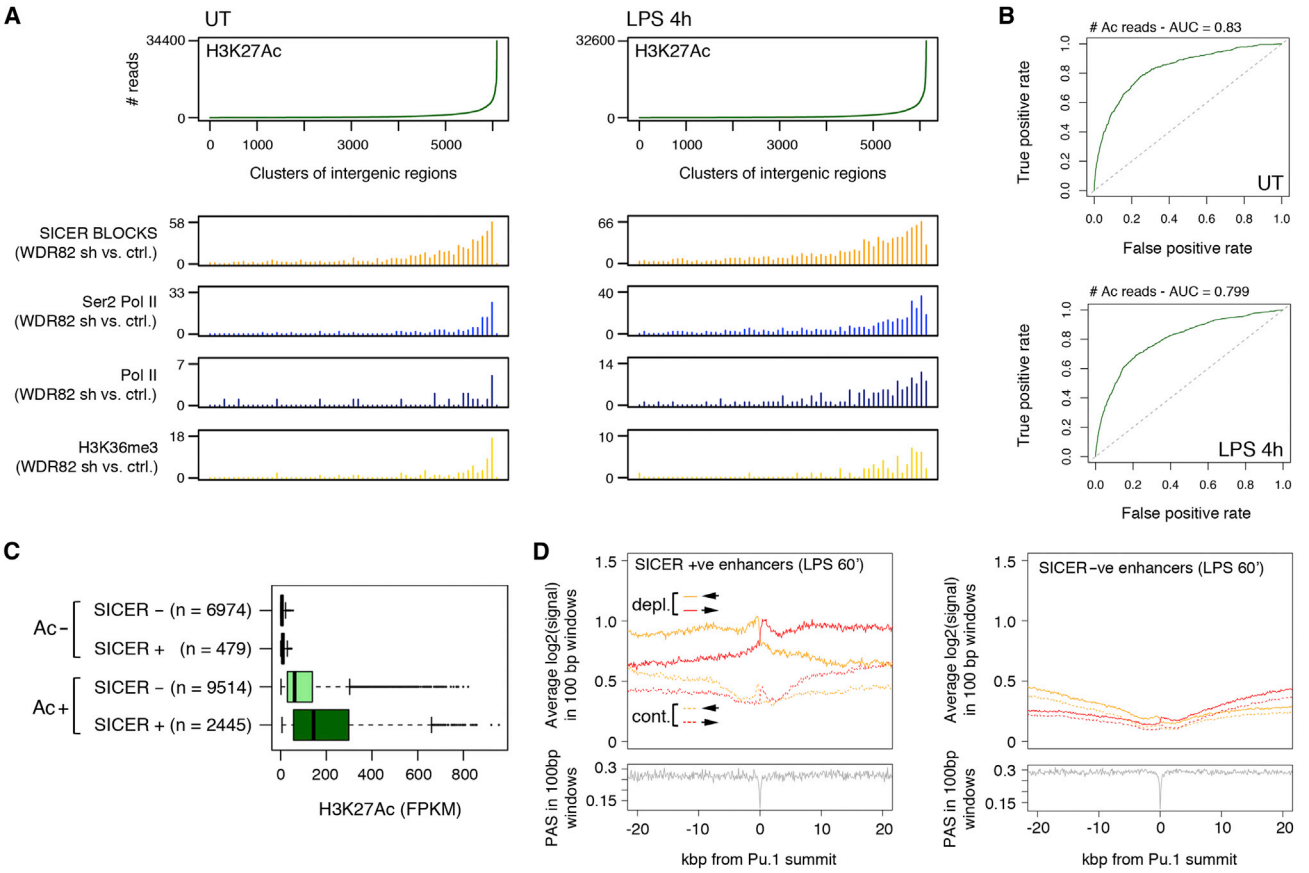
transcription upon WDR82 depletion (Figure 5B). While total and Ser2-phosphorylated Pol II showed qualitatively similar trends of induction in WDR82-depleted cells compared to control cells (Figure 5A), the elongating Pol II isoform was several-fold more enriched than the total Pol II at every level of acetylation. Therefore, the increased 4sU signal due to WDR82 depletion was accounted for mainly by enhanced elongation rather than new initiation events. This result is consistent with the notion that depletion of WDR82 rescues early terminated extragenic transcription. The overall trend of H3K36me3 was similar to that of Ser2 Pol II, although the overlap between increased Pol II and H3K36me3 was only partial. Consistent with these data, 73.3% of the extragenic RNAs induced in WDR82-depleted cells were abrogated by treatment with 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), an inhibitor of the CTD Ser2 protein kinase CDK9 (Figure S4 and Table S4).

To further clarify the link between enhancer acetylation and increased extragenic transcription, we divided the pool of extragenic enhancers, identified by their association with the lineage-determining transcription factor PU.1 (Ghisletti et al., 2010; Heinz et al., 2010) and with H3K4me1, into acetylated and non-acetylated enhancers and determined their overlap with SICER blocks upregulated upon WDR82 depletion (Figure 5C). We found that (1) the acetylated regions associated with aberrant transcription had higher levels of acetylation than those showing no defect and (2) only 6.4% of the non-acetylated enhancers were associated with abnormally increased transcription compared to 20.4% of the acetylated ones.

We next used strand-specific 4sU-seq data to determine how transcription extends from the enhancer core (summit of the PU.1 peak) in control and WDR82-depleted macrophages. In keeping with previous data (Kim et al., 2010), enhancers in control cells gave rise mainly to short and poorly abundant transcripts, while in WDR82-depleted macrophages, aberrantly long and abundant extragenic transcripts extended for many kilobases from the enhancer core (Figure 5D). Remarkably, enhancer cores were strongly depleted of PAS relative to the surrounding genomic background (Figure 5D), pointing to their possible role in eRNA biogenesis: depletion of PASs at enhancer cores may enable the initial Pol II elongation, but when Pol II transcribes the PAS-rich enhancer-flanking regions, the nascent eRNAs may be cleaved, resulting in transcription termination and transcript degradation. In the absence of WDR82, such termination signals are likely ignored, thus enabling Pol II elongation and the synthesis of extended transcripts. Representative snapshots are shown in Figure 5E.

Increased Extragenic Transcription in SET1A/B- and PNUTS-Depleted Macrophages

The effects of WDR82 depletion may be mediated by SET1, the nuclear PP1 complex, or both. We therefore analyzed nascent transcripts in macrophages depleted of SET1A/B (using an shRNA targeting both SET1A and SET1B) or PNUTS (Lee et al., 2010) (Figure S5A). SET1A/B depletion increased the transcription of 534 extragenic regions in untreated macrophages and 836 extragenic regions after LPS (45 min) (Table S5). Of these upregulated transcripts, 70.6% overlapped those upregulated by WDR82 depletion. However, the genomic distribution of the upregulated transcripts was different (Figure 6A), because depletion of SET1 increased transcription at TTS (48.3%), TSS (34.2%), lincRNA genes (7%), and super-enhancers (10%) but not at enhancers, which may relate to the very low levels of H3K4me3 at canonical enhancers. The effects of PNUTS depletion were quantitatively stronger, with 2,392 regions (49.4% TTS, 16.3% TSS, 10% lincRNAs, 17% enhancers, and 6% super-enhancers) showing upregulated transcription 45 min post-LPS treatment and an overlap with the transcripts upregulated by WDR82 depletion of 64.5% (Table S6). In both cases, only a



(legend on next page)

handful of regions showed downregulated transcription in depleted cells. Changes in 4sU signals at TSS- and TTS-proximal extragenic regions are shown in Figure 6B.

We next analyzed whether the enhancers showing increased transcription upon WDR82 depletion were also more transcribed in SET1A/B- and PNUTS-deficient cells. While the effects were quantitatively weaker, they were qualitatively similar: depletion of either SET1A/B or PNUTS resulted in increased transcription at the same PU.1-associated enhancers at which WDR82 depletion caused the appearance of aberrant transcripts (Figure 6C).

The effects of the depletion of WDR82, SET1A/B, and PNUTS were similar, with a high degree of overlap, but also substantial differences. Such differences may relate not only to experimental issues (i.e., the different effects of the residual protein levels) but also to the distinct functional role of each depleted protein. PNUTS depletion did not cause detectable H3K4me3 changes at promoters affected by WDR82 depletion (Figure S5B), ruling out a role of PNUTS in controlling SET1 activity. To better clarify this issue, we analyzed the effects of the depletion of each protein on the four most common patterns of gene-proximal non-coding transcription. WDR82 and SET1 had the strongest impact on anti-sense transcripts, while the effects of PNUTS depletion were dominated by the increase in read-through transcription at gene 3′ ends (Figure 6D). The overall higher number of 4sU SICER blocks in PNUTS-depleted cells was mainly due to such read-through transcription. As with WDR82 depletion (Figure 4A), dysregulated non-coding transcription at TSS or TTS was not associated with changes in the corresponding coding transcript (Figure 6E). Finally, a common trend in all depletion experiments was that the genes with the highest transcriptional activity were those with the highest occurrence of TSS-proximal extragenic transcription and read-through transcription (Figure 6F). Therefore, the transcriptional termination mechanisms enforced by WDR82, PNUTS/PP1, and SET1A/B are critical in conditions of high transcriptional activity. Representative snapshots are shown in Figure S5.

Abnormal Extragenic Transcription and Topological Domains

An important issue is the identity of the mechanisms that determine transcription termination in the absence of WDR82. We hypothesized that the elongating Pol II may be arrested by the physical constraints imposed by the 3D folding of chromatin in the nucleus. If this is correct, then the boundaries of the abnormally transcribed regions should coincide with the borders of the physical domains in which Pol II is transcribing. We determined the correlations between H3K27Ac and H3K4me3 reduction and increased non-coding transcription within topologically associated domains (TADs) (Dixon et al., 2012; Nora et al., 2012). Because TADs are cell invariant, we initially used the annotations previously reported for 2,200 TADs in mouse embryonic stem cells (ESCs) (Dixon et al., 2012). First, we analyzed the correlation between H3K27Ac and reduced H3K4me3 within TADs. TADs containing at least one peak of H3K4me3 in control cells were divided into five groups based on H3K4me3 changes following WDR82 depletion (TADs with no changes represent one group; the other four groups consisted of TADs with reduced H3K4me3 divided into quartiles of progressively higher hypomethylation). Consistent with the data shown previously, the number of H3K4me3 hypomethylated regions within the TADs correlated with their overall level of acetylation (Figure 7A). Next, we determined the correlation between H3K4me3 reduction and increased extragenic transcription within TADs. The number of SICER blocks induced by WDR82 depletion correlated with hypomethylation within the TADs, indicating that H3K4me3 reduction and increased non-coding transcription tend to co-occur within the same TAD (Figures 7B and 7C).

We also generated a HiC dataset in untreated macrophages. The data obtained were similar to those from ESCs (Figures 7D–7F), thus confirming the correlation between H3K4me3 loss and increased extragenic transcription.

Next, we compared the distribution of 4sU signals in WDR82-depleted cells with the physical contacts measured at two selected loci using circularized chromosome conformation capture sequencing (4C-seq), which has a higher spatial resolution. LPS stimulation did not determine major rearrangements in the physical domain at most genomic locations we tested, including the two shown here (S.M. and G.N., unpublished data). Using as viewpoint the TSS of the *Ccl2* gene, most physical interactions occurred in an intergenic region located 5′ of the gene, which coincides with the region of maximal induction of extragenic transcription upon WDR82 depletion (Figure 7G). Remarkably, the 5′ border of the *Ccl2* physical domain coincided with the border of the region of enhanced extragenic transcription. At the 3′ of the gene, aberrant transcription was of more limited intensity. Nevertheless, also in this case, it precisely ended at the border of the physical domain. In the case of the *Jun* gene (Figure 7H), both physical interactions and enhanced extragenic transcription were skewed to the gene desert 5′ of the gene. Also in this case, the limit of extragenic transcription coincided with the border of the physical domain. Overall, these data indicate that the augmented non-coding transcription in WDR82-depleted

Figure 5. Highly Acetylated Enhancers Generate Long and Abundant Non-coding Transcripts in WDR82-Depleted Macrophages

(A) Intergenic clusters of H3K27Ac regions in untreated (UT) and LPS-treated macrophages ranked based on their total acetylation level (upper panels). The overlap with upregulated non-coding transcription (SICER blocks), total, or Ser2-P Pol II and H3K36me3 in the ranked regions was quantified for consecutive bins of 100 regions each (y axis: percentage of overlap for each bin).

(B) ROC curves showing the performance of the H3K27Ac level in predicting increased transcription in WDR82-depleted cells.

(C) Extragenic enhancers divided based on the presence of H3K27Ac and on the overlap with regions of increased transcription in WDR82-depleted cells. SICER-positive acetylated enhancers showed on average a higher level of H3K27Ac compared to SICER-negative acetylated regions.

(D) Cumulative distributions of strand-specific 4sU signals in control and WDR82-depleted LPS-stimulated macrophages. Distributions were centered on the summit of PU.1 peaks in H3K4me1-positive regions. Enhancers with (SICER +ve) or without (SICER −ve) increased transcription are shown. The bottom panels show the distributions of PAS, computed in 100 bp windows, in the same regions.

(E) Representative snapshots showing Ser2 Pol II and H3K36me3 at genomic regions where WDR82 depletion increased extragenic transcription.

See also Figure S4.

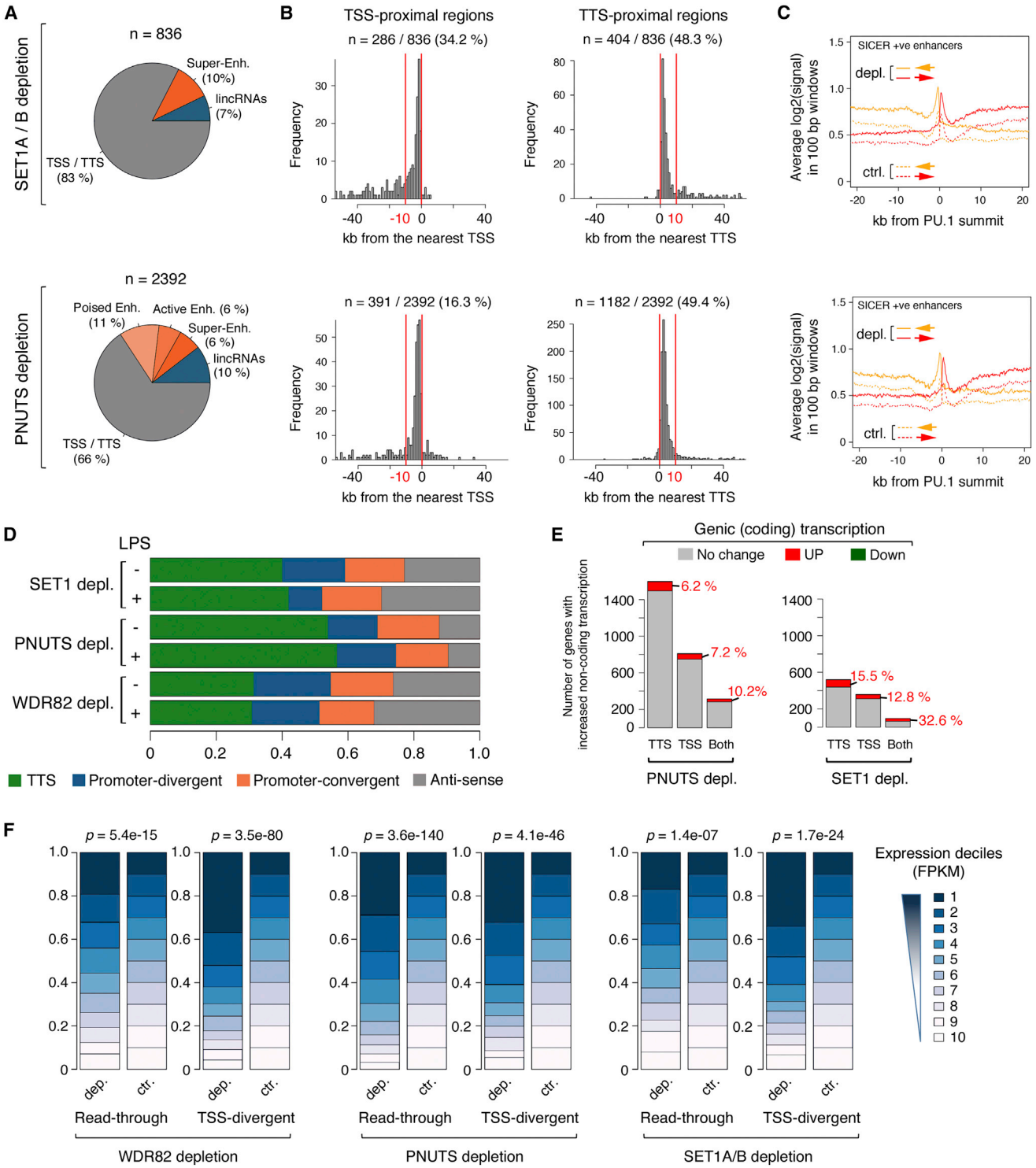


Figure 6. Aberrant Non-coding Transcription in Macrophages Depleted of SET1A/B or PNUTS

(A) Pie charts showing the overlap of non-coding transcripts induced in response to SET1A/B depletion (top) or PNUTS depletion (bottom) with annotated genomic features.

(B) Histograms showing the enrichment for blocks of increased extragenic transcription in the vicinity of TSSs or TTSs in SET1A/B- or PNUTS-depleted macrophages stimulated with LPS.

(C) Cumulative distribution of strand-specific 4sU-seq reads in SET1A/B- or PNUTS-depleted (depl.) macrophages at extragenic regions showing increased transcription in WDR82-depleted macrophages. Data were centered on the summit of PU.1 ChIP-seq peaks in H3K4me1- and H3K27Ac-positive regions.

(legend continued on next page)

cells tends to remain confined within the same chromatin spatial domain as that of the *cis*-regulatory region it originated from.

Finally, because eRNAs were shown to control enhancer-promoter looping (Li et al., 2013), we analyzed the impact of WDR82 depletion on the 3D organization of two genomic loci that are characterized by increased non-coding transcription. For both loci, we used one 5′ probe and one 3′ probe to maximize our chances of detecting changes in the 3D organization (Figure 7I). At both loci, the increase in extragenic transcription induced by WDR82 depletion did not cause measurable alterations in contacts.

DISCUSSION

This study indicates the existence of a network of regulatory complexes that in mammalian cells enforce transcription termination at *cis*-regulatory elements. Three main pathways control transcription termination in metazoans: the CPSF pathway, which is responsible for termination at mRNA coding genes; Integrator, which controls snRNA 3′ end processing and termination (Baillat et al., 2005); and the 5′ CBC-ARS2 pathway, which is involved in snRNA termination and PAS-directed termination of PROMPTs (Andersen et al., 2013; Hallais et al., 2013). The CBC interacts with the 5′ cap on the nascent transcript. In association with the nuclear exosome and the ARS2 subunit, which recruits CPSF (Hallais et al., 2013) via cleavage factor II, the CBC stimulates PAS-directed termination and degradation of PROMPTs (Andersen et al., 2013; Hallais et al., 2013). Inside the coding regions, the relative depletion of PASs and their suppression by U1 small nuclear ribonucleoprotein particle recognition sites prevent cleavage and thus enable transcript extension (Almada et al., 2013; Ntini et al., 2013).

These three pathways are intertwined, as indicated, for instance, by the role of both Integrator and CBC-ARS2 in snRNA termination and the one of CBC-ARS2 and CPSF in PROMPT termination and degradation. Although the mechanistic details remain to be addressed, our data indicate the existence of a network of complexes that affect termination mechanisms and whose depletion causes compound termination defects.

The role of WDR82, PNUTS/PP1, and SET1 in the control of transcription termination is consistent with data in yeast. The yeast WDR82 ortholog (Swd2) and PP1 ortholog (Glc7) are both CPF subunits (Dichtl et al., 2002; He et al., 2003), and their deletion caused termination defects at a subset of Pol II-dependent genes (Cheng et al., 2004; Nedea et al., 2008). In *S. pombe*, a PNUTS ortholog (Ppn1) interacts with both Swd2 and Glc7 within the CPF (Vanoosthuyse et al., 2014).

The corresponding mammalian complexes show analogies but also differences in both composition and function. The purification of the 3′ pre-mRNA processing complex resulted in the retrieval of PNUTS and PP1 but not WDR82 (Shi et al., 2009); moreover, a mass spectrometry analysis did not identify CPSF subunits in WDR82 complexes (van Nuland et al., 2013). Nevertheless, by mining data from a proteomics screening (Malovannaya et al., 2011), we retrieved an interaction between WDR82 and a CPSF ancillary complex, cleavage factor I. Therefore, the interaction of WDR82 and PNUTS/PP1 with complexes involved in 3′ end processing and transcription termination is conserved from yeast to humans.

As regards SET1/COMPASS, in *S. cerevisiae* it favors early Pol II transcription termination, leading to the synthesis of short transcripts (small nucleolar RNAs and CUTs) (Terzi et al., 2011). However, the termination complex involved, Nrd1-Nab3-Sen1, has no obvious counterpart in mammals. It is likely that the function of the Nrd1-Nab3-Sen1 complex is carried out in metazoa by CBC-ARS2 (Porrua and Libri, 2015). Clarifying the relationship between CBC-ARS2 and the complexes investigated here warrants further investigation. SET1/COMPASS has also been implicated in promoting the synthesis of anti-sense transcripts and in causing antisense RNA-mediated coding gene repression (Castelnuovo et al., 2014; Pinskaya et al., 2009). However, in our system, the synthesis of antisense transcripts was increased, not reduced, upon SET1A/B depletion, suggesting fundamental differences between yeast and mammals.

Another interesting mechanistic issue is the relationship between WDR82 and PAS-directed cleavage of nascent non-coding transcripts, which controls PROMPT termination (Almada et al., 2013; Ntini et al., 2013). Enhancer cores were depleted of PASs, suggesting that the high density of PASs encountered by the elongating Pol II outside of enhancer cores may promote transcript cleavage and early termination. Importantly, while eRNAs were initially reported to be non-polyadenylated (Kim et al., 2010), they can be detected by oligo-dT-primed RT-PCR (Hah et al., 2013) and they are retrieved in the poly(A)+ fraction in global nuclear run-on sequencing experiments (Core et al., 2014): these discrepancies are likely explained by the inefficient polyadenylation of eRNAs after cleavage, which may relate to their fast degradation.

The proximity between regions of augmented extragenic transcription and TSSs showing H3K4me3 reduction upon WDR82 depletion suggests a possible direct role of this histone mark. However, PNUTS depletion did not cause H3K4me3 changes, indirectly suggesting that H3K4me3 may not be an essential component of the mechanisms limiting non-coding transcription from *cis*-regulatory elements.

The aberrant extragenic transcription caused by WDR82 depletion was dissociated from strong effects on the transcription of adjacent genes. When changes in coding gene

(D) Major patterns of aberrant non-coding transcription caused by WDR82, SET1A/B, or PNUTS depletion at gene loci.
(E) Bar plots showing the number of genes associated with increased TSS- or TTS-proximal extragenic transcription or both. The colors indicate whether genic (coding) transcription was also affected. Downregulated genes were in all cases less than 1%, and numbers are not shown.
(F) Plots summarizing the observed and the expected distributions of the expression levels (deciles) of genes showing either increased read-through transcription at their 3′ ends or increased promoter-divergent transcription. p values were computed using chi-square test. FPKM, fragments per kilobase of exons per million mapped reads.
See also Figure S5 and Tables S5 and S6.

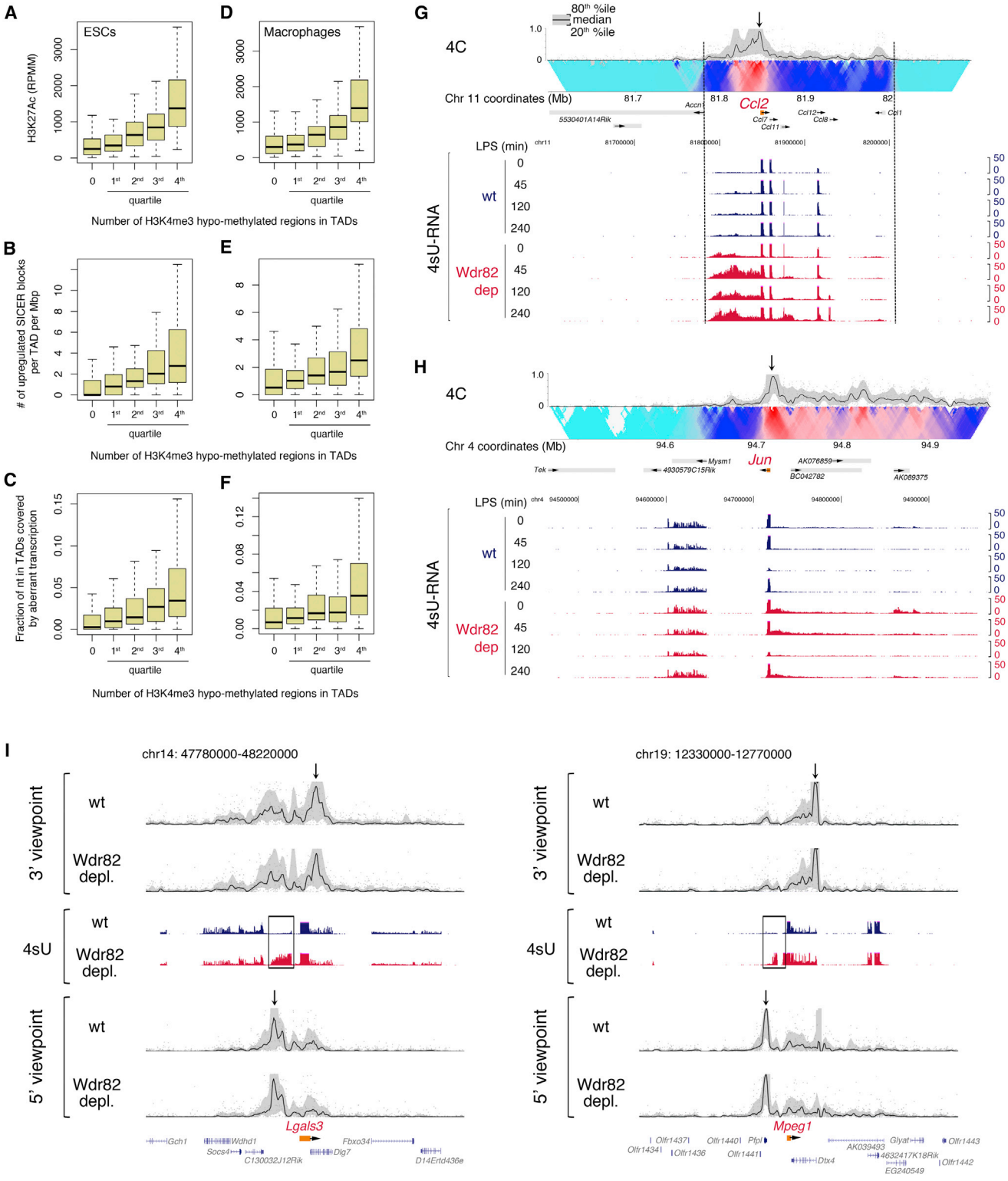


Figure 7. Relationship between Physical Domains and Increased Non-coding Transcription in WDR82-Depleted Cells

(A) Correlation between H3K27Ac levels and H3K4me3 hypomethylation within TADs. TADs in ESCs were derived from a previous publication (Dixon et al., 2012). Only TADs showing at least one H3K4me3 peak in control macrophages were considered. Regions of H3K4me3 hypomethylation in WDR82-depleted macrophages were divided into five groups (x axis): those showing no hypomethylated H3K4me3 peaks (0) and four quartiles of increasing level of H3K4me3 reduction (1st–4th). RPMM, reads per million sequenced reads per megabase.

(legend continued on next page)

expression were observed, they almost invariably consisted of increased transcription relative to control macrophages. Although changes in gene expression were limited, there are several additional reasons termination mechanisms limiting non-coding transcription may have to be in place. First, the recombinogenic effects of transcription may cause genomic instability (Kim and Jinks-Robertson, 2012). Second, abnormal long ncRNAs may promote ectopic recruitment of chromatin modifying enzymes (Rinn and Chang, 2012), thus altering the genomic distribution of histone marks. Finally, the increase in H3K36me3 associated with productive elongation may lead to de novo DNA methylation (Baubec et al., 2015), which may affect the usage of the underlying *cis*-regulatory information.

In conclusion, our data indicate that WDR82 is at the heart of regulatory mechanisms enforcing transcription termination at active promoters and enhancers, as well as at highly expressed genes. Clarifying the mechanistic bases of the phenomena described here will shed light on regulation of transcription termination in metazoa.

EXPERIMENTAL PROCEDURES

Mice, Cell Cultures, and Retroviral Infections

Macrophage cultures, retroviral infections, and LPS treatment were carried out as described (Austenaa et al., 2012) using the murine stem cell virus (MSCV)-based plasmid MSCV-LTR-miR30-Puro-resistance vector with either scrambled or gene-specific shRNAs. The sequences of the shRNAs and PCR primers used are in Table S7. The project has been approved by the Italian Ministry of Health and performed under the supervision of the Institutional Committee for Animal Welfare.

ChIP-Seq

Either 10×10^6 to 20×10^6 (histone modifications) or 100×10^6 to 150×10^6 (Pol II) fixed macrophages were lysed with RIPA (radioimmunoprecipitation assay) buffer and processed as described (Austenaa et al., 2012). The antibodies used are shown in Supplemental Experimental Procedures.

RNA-Seq and 4sU-Seq

For poly(A) RNA-seq, RNA was isolated with RNeasy (QIAGEN) with DNase I treatment and libraries prepared after oligo-dT selection using the TruSeq RNA sample preparation kit. For 4sU experiments, macrophages were labeled with 4sU (Sigma, 150 μM) for 45 min and the labeled RNA was isolated as described (Rabani et al., 2011). A total of 40–100 ng of 4sU-labeled RNA was used for cDNA-library synthesis.

4C-Seq and HiC

4C was carried out as described (van de Werken et al., 2012a) starting from 10^8 macrophages. Data were analyzed using the 4cseqpipe pipeline (van de Werken et al., 2012b). HiC was carried out according to published protocols (Belton et al., 2012).

Computational Methods

Short reads obtained from Illumina HiSeq 2000 runs were quality filtered according to the Illumina pipeline. Detailed computational methods are described in the Supplemental Information file.

ACCESSION NUMBERS

The accession number for the raw datasets reported in this paper is GEO: GSE66955.

SUPPLEMENTAL INFORMATION

Supplemental information includes Supplemental Experimental Procedures, five figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2015.09.018.

AUTHOR CONTRIBUTIONS

Conceptualization, G.N., L.M.I.A., I.B., and M.S. L.M.I.A. designed and carried out most experiments with help from S.G. and A.C. M.S. generated a subset of the ChIP-seq and RNA-seq data. S.M. carried out the 4C-seq experiments with the help of E.d.W. and B.A.M.B. and the supervision of W.d.L. G.D.C. generated the HiC data. E.P. generated plasmids and antibodies. I.B. analyzed all NGS data with contributions from V.P. and A.T. S.d.P. and M.P. designed and carried out the computational analysis of transcripts synthesis and degradation rates. G.N. wrote the manuscript with input from the authors. Supervision, G.N.; funding acquisition, G.N.

ACKNOWLEDGMENTS

The project was supported by the European Research Council (ERC grant NORM to G.N.). We thank S. Monticelli for comments on the manuscript; A. Martins and A. Siepel for the PAS PWM; L. Rotta, T. Capra, and S. Bianchi (IEO and IIT) for the preparation and processing of the sequencing libraries; and T. Kress for suggestions on the 4sU protocol.

Received: April 22, 2015
Revised: August 13, 2015
Accepted: September 17, 2015
Published: October 22, 2015

REFERENCES

Almada, A.E., Wu, X., Kriz, A.J., Burge, C.B., and Sharp, P.A. (2013). Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature *499*, 360–363.

Andersen, P.R., Domanski, M., Kristiansen, M.S., Storvall, H., Ntini, E., Verheggen, C., Schein, A., Bunkenborg, J., Poser, I., Hallais, M., et al. (2013). The human cap-binding complex is functionally connected to the nuclear RNA exosome. Nat. Struct. Mol. Biol. *20*, 1367–1376.

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al.; FANTOM Consortium (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.

Austenaa, L., Barozzi, I., Chronowska, A., Termanini, A., Ostuni, R., Prosperini, E., Stewart, A.F., Testa, G., and Natoli, G. (2012). The histone methyltransferase Wbp7 controls macrophage function through GPI glycolipid anchor synthesis. Immunity *36*, 572–585.

Baillat, D., Hakimi, M.A., Näär, A.M., Shilatifard, A., Cooch, N., and Shiekhattar, R. (2005). Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell *123*, 265–276.

(B) Correlation between H3K4me3 reduction and increased extragenic transcription (SICER blocks) in WDR82-depleted cells examined within TADs.
(C) Same as (B), but with aberrant extragenic transcription (y axis) measured as the fraction of nucleotides in TADs that is covered by SICER blocks.
(D–F) Same analyses as in (A)–(C), but repeated using HiC data from macrophages.
(G and H) Representative snapshots of two genomic loci centered on the TSS (±250 Mb) of the *Ccl2* (G) or *Jun* (H) gene. 4C-seq in untreated control macrophages and 4sU-seq data in control and WDR82-depleted (dep) macrophages are shown.
(I) 4C-seq was carried out at the *Lgals3* (left) and *Mpeg1* loci in control and WDR82-depleted (depl.) untreated macrophages. Two viewpoints per locus were used. 4sU-seq data (from a separate experiment) are shown.

Bannister, A.J., Schneider, R., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2005). Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. J. Biol. Chem. *280*, 17732–17736.

Barozzi, I., Simonatto, M., Bonifacio, S., Yang, L., Rohs, R., Ghisletti, S., and Natoli, G. (2014). Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. Mol. Cell *54*, 844–857.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243–247.

Belton, J.M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., and Dekker, J. (2012). Hi-C: a comprehensive technique to capture the conformation of genomes. Methods *58*, 268–276.

Bhatt, D.M., Pandya-Jones, A., Tong, A.J., Barozzi, I., Lissner, M.M., Natoli, G., Black, D.L., and Smale, S.T. (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. Cell *150*, 279–290.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al.; FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group) (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559–1563.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., and Workman, J.L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581–592.

Castelnuovo, M., Zaugg, J.B., Guffanti, E., Maffioletti, A., Camblong, J., Xu, Z., Clauder-Münster, S., Steinmetz, L.M., Luscombe, N.M., and Stutz, F. (2014). Role of histone modifications and early termination in pervasive transcription and antisense-mediated gene silencing in yeast. Nucleic Acids Res. *42*, 4348–4362.

Cheng, H., He, X., and Moore, C. (2004). The essential WD repeat protein Swd2 has dual functions in RNA polymerase II transcription termination and lysine 4 methylation of histone H3. Mol. Cell. Biol. *24*, 2932–2943.

Ciurciu, A., Duncalf, L., Jonchere, V., Lansdale, N., Vasieva, O., Glenday, P., Rudenko, A., Vissi, E., Cobbe, N., Alphey, L., and Bennett, D. (2013). PNUTS/PP1 regulates RNAPII-mediated gene expression and is necessary for developmental growth. PLoS Genet. *9*, e1003885.

Clouaire, T., Webb, S., Skene, P., Illingworth, R., Kerr, A., Andrews, R., Lee, J.H., Skalnik, D., and Bird, A. (2012). Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes Dev. *26*, 1714–1728.

Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., and Lis, J.T. (2014). Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. Nat. Genet. *46*, 1311–1320.

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. Proc. Natl. Acad. Sci. USA *107*, 21931–21936.

de Pretis, S., Kress, T., Morelli, M.J., Melloni, G.E., Riva, L., Amati, B., and Pelizzola, M. (2015). INSPEcT: a computational tool to infer mRNA synthesis, processing and degradation dynamics from RNA- and 4sU-seq time course experiments. Bioinformatics *31*, 2829–2835.

De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. *8*, e1000384.

Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H., and Keller, W. (2002). A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. Mol. Cell *10*, 1139–1150.

Dichtl, B., Aasland, R., and Keller, W. (2004). Functions for *S. cerevisiae* Swd2p in 3′ end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. RNA *10*, 965–977.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Flynn, R.A., Almada, A.E., Zamudio, J.R., and Sharp, P.A. (2011). Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. Proc. Natl. Acad. Sci. USA *108*, 10460–10465.

Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.L., et al. (2010). Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity *32*, 317–328.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223–227.

Hah, N., Murakami, S., Nagari, A., Danko, C.G., and Kraus, W.L. (2013). Enhancer transcripts mark active estrogen receptor binding sites. Genome Res. *23*, 1210–1223.

Hainer, S.J., Gu, W., Carone, B.R., Landry, B.D., Rando, O.J., Mello, C.C., and Fazzio, T.G. (2015). Suppression of pervasive noncoding transcription in embryonic stem cells by esBAF. Genes Dev. *29*, 362–378.

Hallais, M., Pontvianne, F., Andersen, P.R., Clerici, M., Lener, D., Benbahouche, Nel.H., Gostan, T., Vandermoere, F., Robert, M.C., Cusack, S., et al. (2013). CBC-ARS2 stimulates 3′-end maturation of multiple RNA families and favors cap-proximal processing. Nat. Struct. Mol. Biol. *20*, 1358–1366.

He, X., Khan, A.U., Cheng, H., Pappas, D.L., Jr., Hampsey, M., and Moore, C.L. (2003). Functional interactions between the transcription and mRNA 3′ end processing machineries mediated by Ssu72 and Sub1. Genes Dev. *17*, 1030–1042.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Jensen, T.H., Jacquier, A., and Libri, D. (2013). Dealing with pervasive transcription. Mol. Cell *52*, 473–484.

Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermüller, J., Hofacker, I.L., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science *316*, 1484–1488.

Kim, N., and Jinks-Robertson, S. (2012). Transcription as a source of genome instability. Nat. Rev. Genet. *13*, 204–214.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182–187.

Koch, F., Fenouil, R., Gut, M., Cauchy, P., Albert, T.K., Zacarias-Cabeza, J., Spicuglia, S., de la Chapelle, A.L., Heidemann, M., Hintermair, C., et al. (2011). Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. Nat. Struct. Mol. Biol. *18*, 956–963.

Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. *14*, 2452–2460.

Lam, M.T., Cho, H., Lesch, H.P., Gosselin, D., Heinz, S., Tanaka-Oishi, Y., Benner, C., Kaikkonen, M.U., Kim, A.S., Kosaka, M., et al. (2013). Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature *498*, 511–515.

Lee, J.H., and Skalnik, D.G. (2008). Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. Mol. Cell. Biol. *28*, 609–618.

Lee, J.H., You, J., Dobrota, E., and Skalnik, D.G. (2010). Identification and characterization of a novel human PP1 phosphatase complex. J. Biol. Chem. *285*, 24466–24476.

Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516–520.

Malovannaya, A., Lanz, R.B., Jung, S.Y., Bulynko, Y., Le, N.T., Chan, D.W., Ding, C., Shi, Y., Yucer, N., Krenciute, G., et al. (2011). Analysis of the human endogenous coregulator complexome. Cell *145*, 787–799.

Marquardt, S., Escalante-Chong, R., Pho, N., Wang, J., Churchman, L.S., Springer, M., and Buratowski, S. (2014). A chromatin-based mechanism for limiting divergent noncoding transcription. Cell *157*, 1712–1723.

Melo, C.A., Drost, J., Wijchers, P.J., van de Werken, H., de Wit, E., Oude Vrielink, J.A., Elkon, R., Melo, S.A., Léveillé, N., Kalluri, R., et al. (2013). eRNAs are required for p53-dependent enhancer activity and gene transcription. Mol. Cell *49*, 524–535.

Natoli, G., and Andrau, J.C. (2012). Noncoding transcription at enhancers: general principles and functional models. Annu. Rev. Genet. *46*, 1–19.

Nedea, E., Nalbant, D., Xia, D., Theoharis, N.T., Suter, B., Richardson, C.J., Tatchell, K., Kislinger, T., Greenblatt, J.F., and Nagy, P.L. (2008). The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes. Mol. Cell *29*, 577–587.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol. Cell *11*, 709–719.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.

Ntini, E., Järvelin, A.I., Bornholdt, J., Chen, Y., Boyd, M., Jørgensen, M., Andersson, R., Hoof, I., Schein, A., Andersen, P.R., et al. (2013). Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. Nat. Struct. Mol. Biol. *20*, 923–928.

Ørom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., et al. (2010). Long noncoding RNAs with enhancer-like function in human cells. Cell *143*, 46–58.

Pinskaya, M., Gourvennec, S., and Morillon, A. (2009). H3 lysine 4 di- and trimethylation deposited by cryptic transcription attenuates promoter activation. EMBO J. *28*, 1697–1707.

Porrua, O., and Libri, D. (2015). Transcription termination and the control of the transcriptome: why, where and how to stop. Nat. Rev. Mol. Cell Biol. *16*, 190–202.

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. Science *322*, 1851–1854.

Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., Gnirke, A., Nusbaum, C., Hacohen, N., Friedman, N., et al. (2011). Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat. Biotechnol. *29*, 436–442.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. *81*, 145–166.

Schmitt, S., Prestel, M., and Paro, R. (2005). Intergenic transcription through a polycomb group response element counteracts silencing. Genes Dev. *19*, 697–708.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887–898.

Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., 3rd, Frank, J., and Manley, J.L. (2009). Molecular architecture of the human pre-mRNA 3′ processing complex. Mol. Cell *33*, 365–376.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annu. Rev. Biochem. *81*, 65–95.

Terzi, N., Churchman, L.S., Vasiljeva, L., Weissman, J., and Buratowski, S. (2011). H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. Mol. Cell. Biol. *31*, 3569–3583.

Thomson, J.P., Skene, P.J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A.R., Deaton, A., Andrews, R., James, K.D., et al. (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature *464*, 1082–1086.

van de Werken, H.J., de Vree, P.J., Splinter, E., Holwerda, S.J., Klous, P., de Wit, E., and de Laat, W. (2012a). 4C technology: protocols and data analysis. Methods Enzymol. *513*, 89–112.

van de Werken, H.J., Landan, G., Holwerda, S.J., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Oz, Y., Bouwman, B.A., et al. (2012b). Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat. Methods *9*, 969–972.

van Nuland, R., Smits, A.H., Pallaki, P., Jansen, P.W., Vermeulen, M., and Timmers, H.T. (2013). Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. Mol. Cell. Biol. *33*, 2067–2077.

Vanoosthuyse, V., Legros, P., van der Sar, S.J., Yvert, G., Toda, K., Le Bihan, T., Watanabe, Y., Hardwick, K., and Bernard, P. (2014). CPF-associated phosphatase activity opposes condensin-mediated chromosome condensation. PLoS Genet. *10*, e1004415.

Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K.A., et al. (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature *474*, 390–394.

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in *cis* to inhibit transcription. Nature *454*, 126–130.

Whitehouse, I., Rando, O.J., Delrow, J., and Tsukiyama, T. (2007). Chromatin remodelling at promoters suppresses antisense transcription. Nature *450*, 1031–1035.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Wu, M., Wang, P.F., Lee, J.S., Martin-Brown, S., Florens, L., Washburn, M., and Shilatifard, A. (2008). Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. Mol. Cell. Biol. *28*, 7337–7344.

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.