Cell

Transcriptome Surveillance by Selective Termination of Noncoding RNA Synthesis

Daniel Schulz,[1,2] Bjoern Schwalb,[1,2] Anja Kiesel,[1] Carlo Baejen,[1] Phillipp Torkler,[1] Julien Gagneur,[1] Johannes Soeding,[1,*] and Patrick Cramer[1,*]
[1]Gene Center Munich and Department of Biochemistry, Center for Integrated Protein Science CIPSM, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany
[2]These authors contributed equally to this work
*Correspondence: soeding@lmb.uni-muenchen.de (J.S.), cramer@lmb.uni-muenchen.de (P.C.)
http://dx.doi.org/10.1016/j.cell.2013.10.024

SUMMARY

Pervasive transcription of eukaryotic genomes stems to a large extent from bidirectional promoters that synthesize mRNA and divergent noncoding RNA (ncRNA). Here, we show that ncRNA transcription in the yeast *S. cerevisiae* is globally restricted by early termination that relies on the essential RNA-binding factor Nrd1. Depletion of Nrd1 from the nucleus results in 1,526 Nrd1-unterminated transcripts (NUTs) that originate from nucleosome-depleted regions (NDRs) and can deregulate mRNA synthesis by antisense repression and transcription interference. Transcriptome-wide Nrd1-binding maps reveal divergent NUTs at most promoters and antisense NUTs in most 3′ regions of genes. Nrd1 and its partner Nab3 preferentially bind RNA motifs that are depleted in mRNAs and enriched in ncRNAs and some mRNAs whose synthesis is controlled by transcription attenuation. These results define a global mechanism for transcriptome surveillance that selectively terminates ncRNA synthesis to provide promoter directionality and to suppress antisense transcription.

INTRODUCTION

Sequencing and microarray technologies revealed that the genomes of eukaryotic cells are pervasively transcribed. About 74% of the human genome gives rise to RNA transcripts, although only about 2% correspond to protein-coding mRNAs (Djebali et al., 2012). In the yeast *Saccharomyces cerevisiae* (*Sc*), 85% of the genome is transcribed (David et al., 2006), and hundreds of noncoding RNAs (ncRNAs) exist in addition to the classical 4 rRNAs, 42 tRNAs, 6 snRNAs, and 77 snoRNAs (*Saccharomyces* genome database) (Hani and Feldmann, 1998). Pervasive transcription stems to a large extend from bidirectional RNA polymerase (Pol) II transcription (Core et al., 2008; Neil et al., 2009; Seila et al., 2008; Xu et al., 2009) that in yeast was shown to originate from two adjacent preinitiation complexes (PICs) within a nucleosome-depleted region (NDR) (Murray et al., 2012; Rhee and Pugh, 2012).

Three mechanisms have been identified that restrict the extent of pervasive transcription. First, transcription initiation can be biased to the mRNA direction by gene looping that limits initiation of divergent ncRNA transcription (Rhee and Pugh, 2012; Tan-Wong et al., 2012). Second, chromatin remodeling and deacetylation can restrict ncRNA transcription initiation (Churchman and Weissman, 2011; Whitehouse et al., 2007). Third, ncRNAs are rapidly removed by RNA degradation. In yeast, 925 ncRNAs called cryptic unstable transcripts (CUTs) are degraded from their 3′ end by the exosome, and deletion of the nuclear exosome subunit Rrp6 stabilizes these ncRNAs (Wyers et al., 2005; Xu et al., 2009). Other studies even detected 1,496 CUTs that we refer to as CUT*s (Neil et al., 2009) and full inactivation of the exosome resulted in 1,600 CUTs (Gudipati et al., 2012). Degradation of ncRNAs also occurs from the 5′ end, since deletion of the 5′ exonuclease Xrn1 stabilizes 1,658 Xrn1-dependent unstable transcripts (XUTs) (van Dijk et al., 2011). Thus pervasive transcription is limited at the level of transcription initiation and RNA degradation.

Global ncRNA synthesis may generally be restricted by selective termination of ncRNA transcription, which was recently shown in mammalian cells to provide promoter directionality (Almada et al., 2013; Ntini et al., 2013). In yeast, the termination of Pol II transcription occurs via two distinct pathways (Hsin and Manley, 2012; Kim et al., 2006; Mischo and Proudfoot, 2013). Termination of mRNA genes requires the cleavage and polyadenylation factor that binds a polyadenylation signal (pA) in the nascent RNA (Mandel et al., 2008). Termination of snRNAs and snoRNAs, however, depends on Nrd1, an essential protein that interacts with Pol II via the serine-5 phosphorylated form of its C-terminal domain (CTD) and contains an RNA recognition motif (RRM) (Steinmetz and Brow, 1996; Vasiljeva et al., 2008). Nrd1 binds a tetramer motif in the RNA (Carroll et al., 2004; Creamer et al., 2011; Porrua et al., 2012; Wlotzka et al., 2011) and interacts with Nab3 and Sen1 to promote termination (Steinmetz et al., 2001). Transcription termination of several CUTs (Arigo et al., 2006b; Thiebaut et al., 2006) and a few stable unannotated transcripts (SUTs) depends on Nrd1 (Marquardt et al., 2011). Nrd1 is also required for the removal of aberrant *Sc* transcripts resulting from expression of a prokaryotic factor (Honorine et al., 2011).

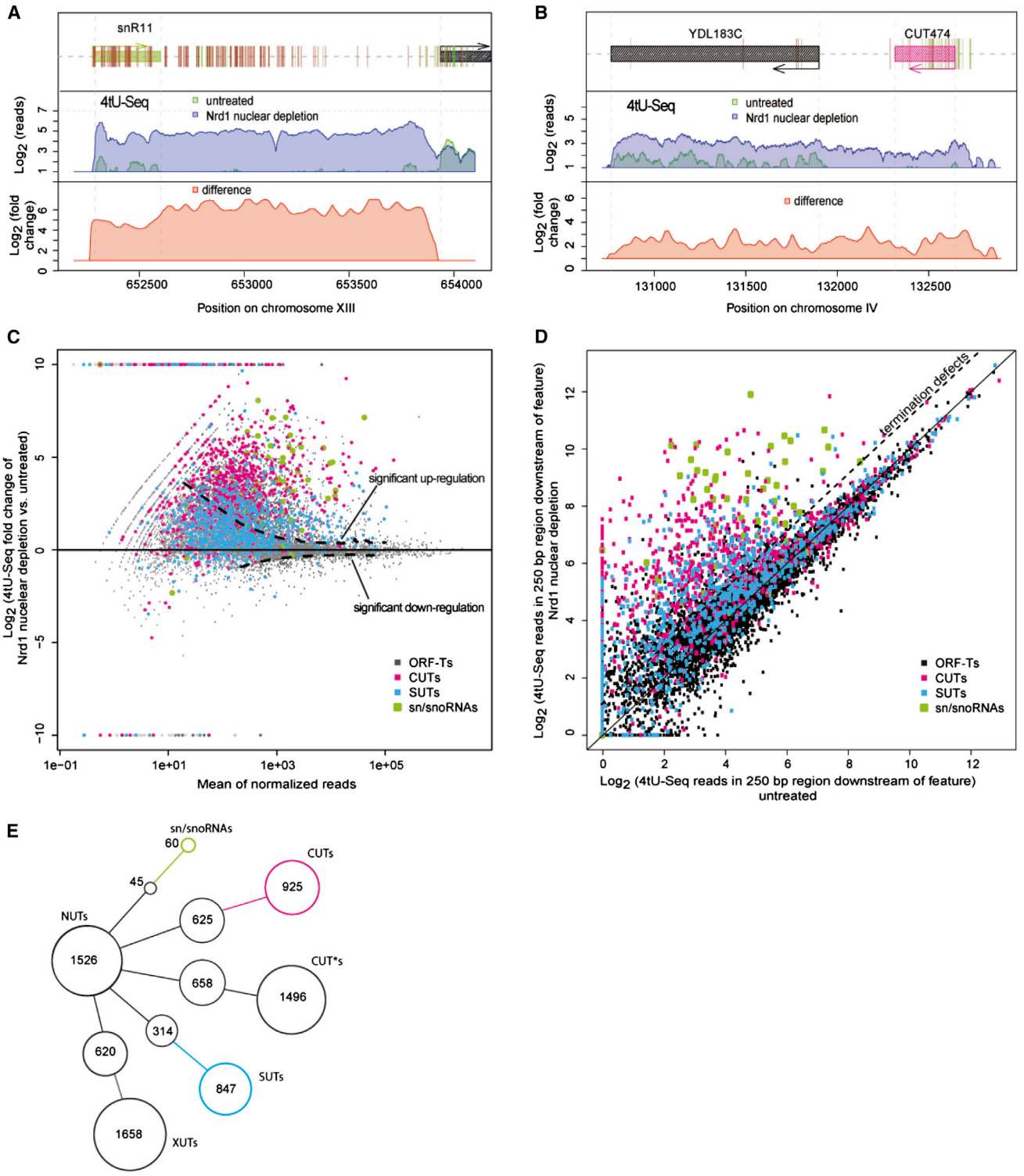

Figure 1. Nrd1 Globally Terminates ncRNAs

(A) Nuclear depletion of Nrd1 leads to defective termination of snRNA11 transcription. Genome browser view of log2 counts of reads measured by 4tU-seq before (green) and after (blue) nuclear depletion of Nrd1, and the fold-change between these signals (red) for every genomic position. Vertical green and brown lines depict RNA-binding sites of Nrd1 and Nab3 as determined by PAR-CLIP.

(B) Genome browser view as in A but for CUT474.

(legend continued on next page)

It was proposed that Nrd1-dependent termination of divergent transcription can generate promoter directionality in *Sc* (Buratowski, 2009; Jacquier, 2009; Seila et al., 2009; Wei et al., 2011). This was supported by global in vivo RNA crosslinking of Nrd1 and Nab3 to CUTs (Wlotzka et al., 2011) and to RNA produced antisense of weakly expressed genes (Creamer et al., 2011). Nrd1 however also crosslinks to many mRNAs (Creamer et al., 2011; Wlotzka et al., 2011) and is recruited to mRNA genes according to chromatin immunoprecipitation (ChIP) (Mayer et al., 2012). This raised the question whether Nrd1-dependent termination functions widely in the attenuation of mRNA transcription, as observed for mRNA genes NRD1, HRP1, and IMD2 (Arigo et al., 2006a; Steinmetz et al., 2006), URA2, URA8, and ADE12 (Kuehner and Brow, 2008; Thiebaut et al., 2008), and FKS2 (Kim and Levin, 2011).

To investigate the global functions of Nrd1-dependent termination, we depleted Nrd1 from the yeast nucleus and monitored changes in RNA synthesis and Pol II occupancy. We found that Nrd1 rarely attenuates mRNA transcription but that it is required for the selective termination of ncRNA synthesis, including transcription antisense to known genes and divergent transcription from bidirectional promoters. Comprehensive mapping of Nrd1 and Nab3 onto the transcriptome revealed divergent ncRNAs at most promoters and a depletion of the preferred RNA-binding motifs in mRNAs. Our results show that selective Nrd1-dependent termination of ncRNA synthesis acts as a mechanism for transcriptome surveillance that provides promoter directionality and prevents transcriptome deregulation.

RESULTS

Nrd1 Nuclear Depletion Is Lethal

To investigate the roles of pA-independent transcription termination in genome expression, we conditionally depleted Nrd1 from the *Sc* nucleus using the anchor-away method (Haruki et al., 2008). Nrd1 was tagged with the FKBP12 rapamycin-binding domain (FRB) and depleted from the nucleus by rapamycin treatment. Rapamycin forms a ternary complex with Nrd1-FRB and FKBP12-RPL13A fusion proteins that is exported from the nucleus. Strains expressing Nrd1-FRB from the endogenous *NRD1* promoter grew normally, but failed to grow in the presence of 1 μg/ml rapamycin (Figure S1 available online). Fluorescence microscopy showed that rapamycin treatment led to nuclear depletion of the Nrd1-FRB fusion protein after 60 min (Figure S1), indicating that nuclear depletion of Nrd1 is lethal.

Nrd1 Generally Terminates ncRNA Transcription

To monitor RNA synthesis in yeast, we metabolically labeled newly synthesized RNA for 6 min with 4-thiouracil (4tU), purified labeled RNA (Sun et al., 2012), and subjected purified labeled RNA to sequencing (Experimental Procedures). We refer to this sensitive method for monitoring global transcription activity as 4tU-seq, in accordance with 4sU-seq that uses 4-thiouridine (4sU) labeling in human cells (Rabani et al., 2011; Windhager et al., 2012). High correlations between biological replicates demonstrated the reproducibility of 4tU-seq (Figure S2).

To follow changes in RNA synthesis upon nuclear depletion of Nrd1, we carried out 4tU-seq before and after addition of 1 μg/ml rapamycin for 60 min. After rapamycin addition at OD_{600} = 0.6, cells grew only to OD_{600}~3. 4tU-seq revealed that sn/snoRNAs were generally not terminated and ncRNA signals were increased (Figures 1A and 1B). The normalized read counts for annotated genomic features (Anders and Huber, 2010) showed increased RNA synthesis for 80% of sn/snoRNAs and many CUTs by >1.5-fold (adjusted p value 0.1) but only of 4% of transcribed protein-coding regions (ORF-Ts) (Figure 1C).

To examine the termination defects globally, we determined the amount of read-through transcription upon nuclear depletion of Nrd1 by calculating the difference in the number of reads in a 250 bp window downstream of each feature (Figure 1D). Whereas mRNAs were generally not affected, termination defects were observed for 80% of sn/snoRNAs, 68% of CUTs, and 58% of SUTs (Xu et al., 2009). This indicates that the Nrd1 pathway generally terminates ncRNA transcription.

NUTs Are Extended ncRNA Transcripts

To describe changes in the transcriptome upon nuclear depletion of Nrd1, we annotated a total of 1,526 new transcripts and called them Nrd1-dependent unterminated transcripts, or NUTs (Figure S2E) (Extended Experimental Procedures). Many annotated ncRNAs overlapped by at least 50% with NUTs, namely 625 CUTs, 314 SUTs, 620 XUTs, 45 sn/snoRNAs, and 658 CUT*s (Figure 1E). NUTs showed 4tU-seq signals similar to the overlapping ncRNAs but were on average 3.8-fold longer (Figure S2), consistent with a termination defect. Only 120 NUTs (8%) overlapped with mRNAs and 213 NUTs (14%) did not overlap with known genomic features. Therefore NUTs are distinct from, although often overlapping with, previously annotated ncRNAs and arise from defective Nrd1-dependent termination of ncRNA transcription.

NUTs Originate from Distinct PICs in NDRs

The majority of NUTs (896, 59%) originated from 5′ and 3′ NDRs flanking known genes (Mavrich et al., 2008), whereas 339 NUTs originated from intergenic regions, and 291 from ORF-Ts (Figure 2A). NUTs originate from NDRs with similar levels of nucleosome depletion (Figure 2B) that were almost as high as for NDRs containing the transcription start sites (TSSs) of ORF-Ts

(C) Most sn/snoRNAs and CUTs, but few ORF-Ts, show increased RNA synthesis upon nuclear depletion of Nrd1. Points mark each transcript's log2 fold-change upon Nrd1 depletion versus the normalized mean read count across replicates and conditions (Anders and Huber, 2010). Transcripts above or below the dashed line are significantly up- or downregulated as calculated by DE-seq. SUTs, CUTs, sn/snoRNAs, and mRNAs from ORF-Ts are in magenta, blue, green, and gray, respectively.

(D) Most sn/snoRNAs and CUTs but few T-ORFs show termination defects upon nuclear depletion of Nrd1. Log2 of normalized read counts in a 250 bp region downstream of annotated genomic feature upon nuclear depletion of Nrd1 versus the same measure in untreated cells.

(E) Overlap of NUTs with CUTs and SUTs from Xu et al. (2009), XUTs from van Dijk et al. (2011), snRNAs, and CUT*s from Neil et al. (2009). NUTs were counted to be overlapping when they covered at least 50% of a previously annotated transcript. See also Figure S1 and S2.

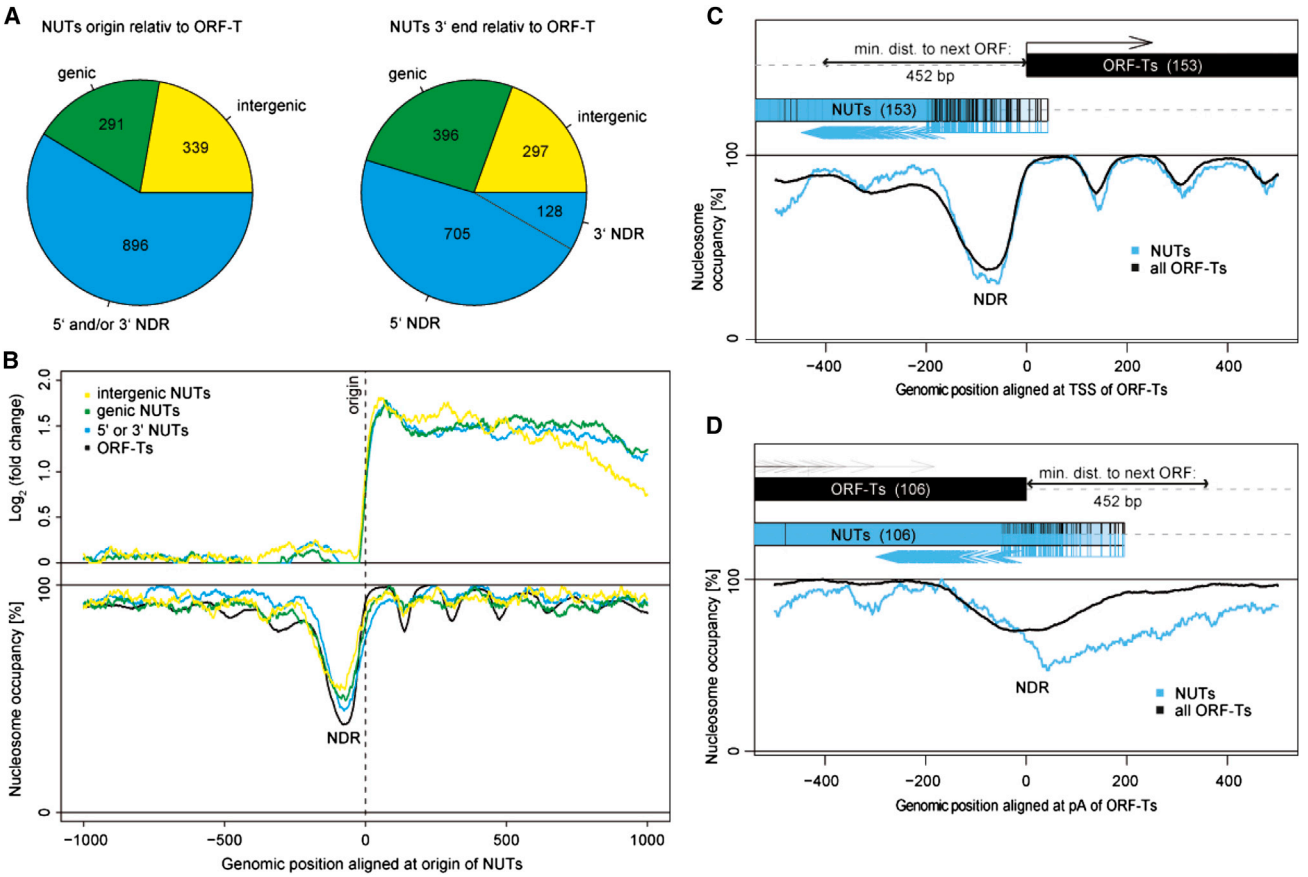

Figure 2. NUTs Originate from NDRs

(A) Fraction of NUT 5′ origins and 3′ ends in NDRs (blue), genic regions in ORF-Ts (green, from TSS + 50 bp to pA site – 50 bp) or intergenic regions (yellow). For 3′ end positions in NDRs those in the 5′ NDR of ORF-Ts in antisense or sense direction are shown.

(B) Average 4tU-seq log2 fold-changes upon nuclear depletion of Nrd1 (upper chart) and averaged nucleosome occupancies (lower chart) (Mavrich et al., 2008) around the NUT 5′ origin for the three categories defined in (A) and for all ORF-Ts (black line, lower chart).

(C) Nucleosome occupancies (as in B) for 153 selected NUTs (blue box with arrows indicating origins) that originate strictly from 5′ NDRs and run antisense to ORF-Ts (black box). Plot was aligned at the TSS of those 153 ORF-Ts.

(D) Nucleosome occupancies (as in B) for 106 selected NUTs that originate strictly from 3′ NDRs of ORF-Ts.

(Figure 2B). When we analyzed ORF-T pairs with a distance of at least 452 bp, which enables separation into 5′ and 3′ NDRs (Xu et al., 2009), 153 NUTs originated from the 5′ region of ORF-Ts, whereas 106 NUTs originated from 3′ regions (Figures 2C and 2D). Thus, NUTs generally originate from NDRs, and about half of them terminate in promoter-associated NDRs (Figure 2A).

We could assign the origins of 690 NUTs (45%) to experimentally mapped PICs (Rhee and Pugh, 2012). Of these, 318 were previously assigned to CUTs and 147 to SUTs, but 257 were unassigned, corresponding to one-third of all PIC orphans. NUTs with mapped PICs showed a 1.6-fold higher median RNA synthesis than NUTs lacking mapped PICs (Figure S3). The 3′ ends of 60% of all NUTs were found in a 5′ NDR of an ORF-T, maybe due to the presence of a PIC for ORF-T transcription (Figure 2A).

Many NUTs Are Divergent and Antisense Transcripts

A total of 845 NUTs (55%) were divergent transcripts arising from bidirectional promoters. There was no correlation between levels of divergent NUT and ORF-T transcription arising from the same bidirectional promoter (data not shown). This is consistent with previous findings (Murray et al., 2012; Yassour et al., 2010) and with the suggestion that transcription activity is set by independent PICs for divergent transcripts and not by the level of nucleosome depletion (Rhee and Pugh, 2012). Many NUTs originated upstream and antisense of ORF-Ts from the 5′ NDR or an overlapping 3′ NDR of an upstream ORF-T. The NUT origin in 5′ NDRs is on average 180 bp upstream of the TSS of ORF-Ts. Thus NUTs often run antisense to known genes and often originate from bidirectional promoters as divergent transcripts.

Nrd1 and Nab3 Preferentially Bind Divergent and Antisense ncRNAs

To examine why Nrd1 preferentially terminates divergent transcription, but not sense transcription, we comprehensively mapped RNA interactions of Nrd1 and its partner Nab3 in growing yeast with the use of photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) (Hafner et al.,

2010), similar to Creamer et al. (2011). We optimized PAR-CLIP for the yeast system (Extended Experimental Procedures) and developed a computational pipeline for data analysis (Extended Experimental Procedures). We defined an RNA-binding event as the occurrence of at least two overlapping reads with T-C nucleotide conversion (Hafner et al., 2010). This identified approximately 267,000 Nrd1- and 184,000 Nab3-binding sites in the yeast transcriptome.

To estimate relative binding affinities of Nrd1 and Nab3 over the transcriptome, we normalized the PAR-CLIP data with transcript occurrence (Kishore et al., 2011; König et al., 2011). Normalization is necessary because the cellular concentration of regular transcripts is much higher than that of rapidly degraded ncRNAs. Normalization was carried out with 4tU-seq data upon nuclear depletion of Nrd1 because ncRNAs transcripts are barely detected under normal conditions (Extended Experimental Procedures). Analysis of the normalized PAR-CLIP data revealed that the relative binding affinity of Nrd1 and Nab3 to 5,123 ORF-T transcripts was low, whereas it was higher to divergent antisense ncRNA transcripts, in particular within the first few hundred nucleotides (Figures 3 and S4). We also observed strong binding of Nrd1 to antisense transcripts originating from the 3′ region of ORF-Ts (Figure 3D). When we normalized our PAR-CLIP data with NET-seq data from Churchman and Weissman (2011) that measures the amount of nascent RNA emerging from polymerase, we obtained very similar results (Figures S4G and S4H and S3), confirming that Nrd1 and Nab3 preferentially bind divergent and antisense ncRNAs.

Nrd1-Binding RNA Motifs Are Depleted in mRNA

We speculated that the different Nrd1/Nab3-binding densities observed between ORF-T transcripts, antisense ncRNAs, and intergenic ncRNAs may be a result of different motif compositions of these transcript classes. Analysis of the PAR-CLIP sites revealed several tetrameric RNA-binding motifs, including known motifs UGUA and GUAG for Nrd1, and UCUU and CUUG for Nab3 (Carroll et al., 2004; Creamer et al., 2011; Porrua et al., 2012; Wlotzka et al., 2011), which were strongly overrepresented (Figure S5). The best binding motif for Nab3 (UCUU) frequently occurred in a window of 21 bp around Nrd1-binding sites (Figure S5), consistent with a Nrd1-Nab3 complex. Thus, as suggested by Porrua et al. (2012), Nrd1 and Nab3 have RNA-binding preferences rather than strict specificity for a single motif.

To investigate whether mRNAs contain fewer Nrd1-binding motifs than ncRNAs, we calculated an apparent Nrd1-binding affinity for each of the 256 tetrameric motifs from their relative frequency at PAR-CLIP sites, and then mapped apparent Nrd1-binding affinities along the yeast genome. Strikingly, mRNAs were markedly depleted in additive apparent Nrd1-binding affinity, a factor of 1.5 lower than antisense ncRNAs (Figure 3E). Intergenic ncRNAs were also enriched in overall Nrd1-binding affinity with respect to mRNAs, by a factor of 1.3 (Figure 3F). The real binding preference in vivo is likely much higher than the observed differences in apparent binding affinity because multiple copies of Nrd1 likely bind cooperatively, and because Nrd1 and Nab3 cooperate to bind neighboring sites.

Consistent with this, Nab3 also showed an increased apparent binding affinity for ncRNA, with values similar to that for Nrd1 (Figure S4). A positive control is further provided by a higher number of detected Nrd1 sites in NUTs and sn/snoRNAs compared to ORF-Ts, SUTs, and other ncRNAs after normalization with either 4tU-seq or NET-seq data (Figure S3). The observed site density matched the occurrence of Nrd1 motifs in these RNA classes (Figure S3). Similar results were obtained when nonnormalized PAR-CLIP data were compared for different RNAs with similar RNA synthesis rates (Figure S3).

These results indicate that Nrd1 preferentially binds to ncRNAs, because the preferred Nrd1-binding motifs are depleted from mRNAs. Nrd1 may have evolved to bind RNA motifs that do not occur in coding mRNA, or yeast genes may have evolved to preferentially use codons that do not give rise to Nrd1 motifs, or both. The higher motif occurrence explains why ncRNAs are preferred substrates for Nrd1-dependent termination. Higher motif occurrence and PAR-CLIP site density was also detected downstream of ORF-Ts, which can account for a known fail-safe mechanism for mRNA termination (Rondón et al., 2009).

Yeast Promoters Are Generally Bidirectional

Of all 5,123 ORF-Ts in the annotation file we used (Xu et al., 2009), 1,712 are divergent ORF-T pairs with a maximum distance of 452 bp between them. Among the remaining 3,411 ORF-Ts we detected at least two PAR-CLIP sites upstream and antisense of 1,898 ORF-Ts, which had no other ORF-T annotated upstream and divergent within 452 bp. At least one PAR-CLIP site was observed upstream of 2,272 ORF-Ts. The PAR-CLIP sites show that divergent ncRNAs must have existed at this position. Thus 3,984 *Sc* promoters (78%) are bidirectional and show at least one PAR-CLIP site on the divergent ncRNA. Consistent with this, a divergent NUT was observed for 845 of the 3,411 ORF-Ts. Thus the sensitivity for detecting a short-lived divergent ncRNA of low abundance is higher for PAR-CLIP than for 4tU-seq.

Nrd1 Is Required for Promoter Directionality

The above results provide strong evidence that yeast promoters generate both mRNA and divergent ncRNA, and that the divergent ncRNA preferentially binds Nrd1. This is consistent with the idea that selective Nrd1-dependent termination of divergent ncRNA transcription is important for setting promoter directionality. To investigate whether Nrd1 depletion leads to a partial loss of promoter directionality, we plotted sense and antisense 4tU-seq signals around all TSSs before and after Nrd1 depletion (Figures 3G and 3H). This revealed that Nrd1 depletion leads to a 2-fold average increase in divergent transcription, demonstrating a partial loss of promoter directionality. Nrd1 depletion also increases antisense transcription in ORF-Ts and sense transcription upstream of ORF-Ts, consistent with a global transcriptome surveillance mechanism that restricts ncRNA synthesis by Nrd1-dependent termination.

Antisense ncRNA Synthesis Can Downregulate Transcription

To investigate whether defects in Nrd1-dependent ncRNA termination deregulate genome transcription, we tested whether

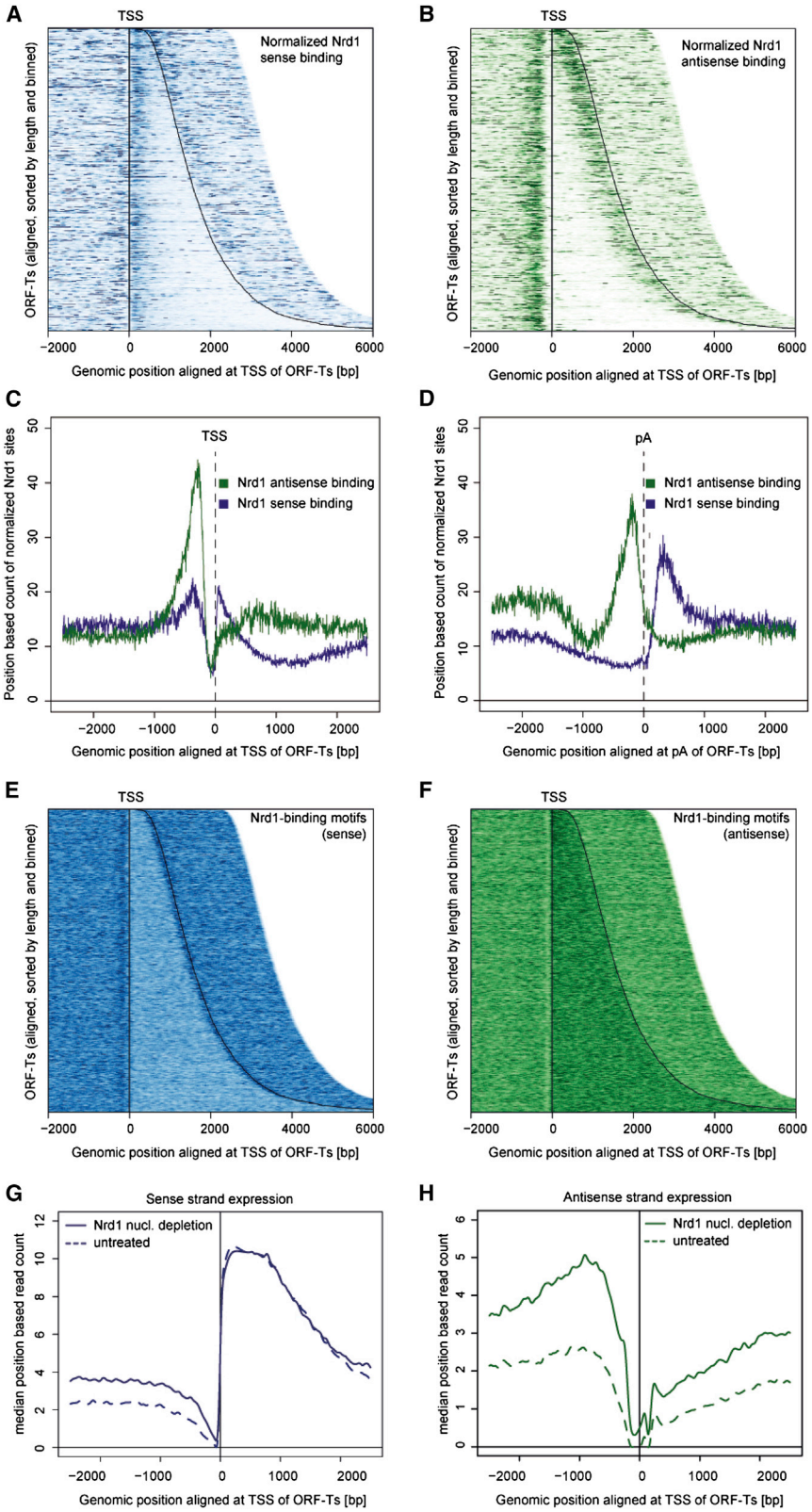

Figure 3. Nrd1 Preferentially Binds Divergent and Antisense ncRNAs

(A) Expression-normalized heat map of Nrd1 RNA-binding sites derived from PAR-CLIP in sense direction for all ORF-Ts. ORF-Ts were sorted by length and aligned at their TSS and binned (Xu et al., 2009). The curved line on the right represents the pA sites. Signals were plotted until 2,000 bp after the pA site. Strength of binding is coded from white (no binding) to dark blue (strong binding).

(B) Expression-normalized heat map of Nrd1 RNA-binding sites as in A but for the antisense direction. Strength of binding is coded from white (no binding) to dark green (strong binding).

(C) Expression-normalized meta-plot of the data from (A). Nrd1 RNA-binding site distribution around the TSS of all ORF-Ts for the sense (blue) and antisense direction (green) with respect to ORF-Ts. The y-values are proportional to the occupancy of Nrd1 on the transcripts.

(D) As in C but around the pA site of all ORF-Ts.

(E and F) Heat maps of tetrameric motif binding preference of Nrd1 in sense (E) and antisense (F) direction for all ORF-Ts. The occurrence of tetramers was weighted by the likelihood of Nrd1 binding. ORF-T alignment and coloring like in (A) and (B).

(G) Sense strand 4tU-seq signals (median position based read count) of ORF-Ts aligned at their TSS before (dashed line) and after (solid line) nuclear depletion of Nrd1.

(H) As in G but with antisense strand expression. See also Figures S3, S4, and S5.

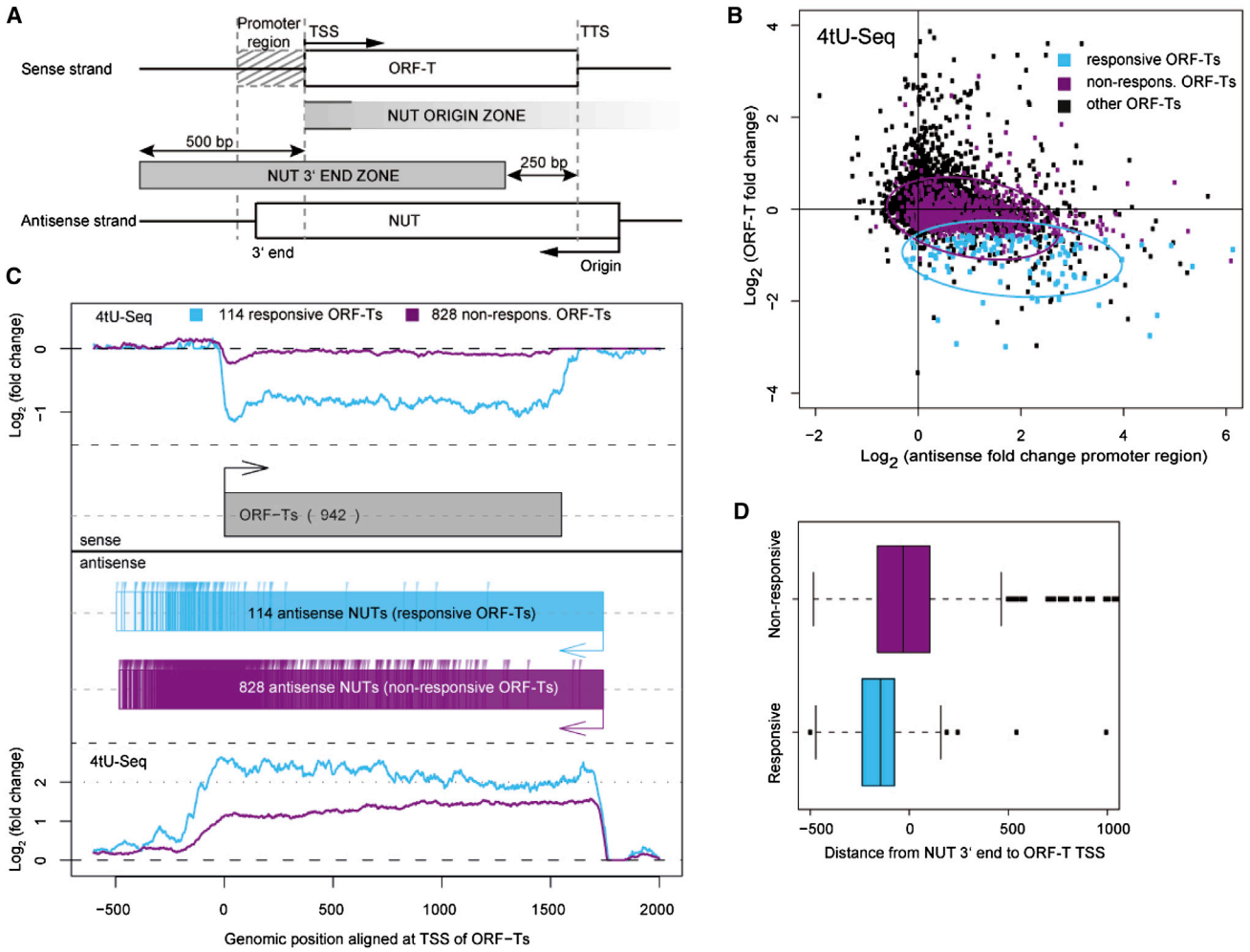

Figure 4. Antisense NUTs Interfere with ORF-T Transcription

(A) Definition of antisense NUT class. A NUT belongs to the antisense class when its origin is downstream of the TSS of the overlapping ORF-T (NUT ORIGIN ZONE). The 3′ end of a NUT has to be 250 bp upstream of the TTS of the ORF-T and maximal 500 bp upstream of the TSS of the ORF-T (NUT 3′ END ZONE). Dashed vertical lines delineate the ORF-T promoter region (−200 bp to TSS).

(B) Some ORF-Ts are repressed upon antisense NUT transcription. Fold-change of ORF-T 4tU-seq signal versus fold-change in antisense signal over the ORF-T promoter region (compare A). ORF-Ts with a decrease in their 4tU-seq signal (114 responsive ORF-Ts) are in cyan, 828 nonresponsive ORF-Ts are in purple, and all others are in black.

(C) Top: Distribution of log2 fold-changes in 4tU-seq signal upon nuclear depletion of Nrd1 for responsive (cyan) and nonresponsive ORF-Ts (purple) defined in (B). Bottom: 4tU-seq signals for antisense NUTs corresponding to responsive (cyan) and nonresponsive (purple) ORF-Ts. All 942 ORF-Ts are scaled to a median length and aligned at their TSS. NUTs are aligned at their median origin (blue and purple boxes with arrows). Vertical lines indicate NUT 3′ ends.

(D) Extension of antisense NUTs into the promoter region of their corresponding sense ORF-T correlates with transcription repression. Color code as in (B). The distance between the antisense NUT 3′ end and the TSS of its corresponding sense TSS is plotted on the horizontal axis.

NUT transcription antisense to ORF-Ts influences sense transcription (Figure 4A). Antisense transcription was shown to regulate several yeast loci (Camblong et al., 2007; Castelnuovo et al., 2013; Hongay et al., 2006; Houseley et al., 2008; Xu et al., 2011). A total of 942 NUTs were antisense to annotated ORF-Ts (antisense NUT class). We plotted changes in 4tU-seq signals in ORF-Ts over changes in antisense signals in the promoter region of ORF-Ts upon nuclear depletion of Nrd1 (Figure 4B).

We found that increasing levels of antisense transcription in the promoter region of ORF-Ts correlates with downregulation of ORF-T transcription (correlation coefficient −0.21, 95% confidence interval [-0.18,-0.23], p value < 2.2×10^{-16}). Of the 202 significantly downregulated genes, 114 (56%, "responsive") showed an antisense NUT that explained downregulation. These 114 responsive ORF-Ts showed stronger antisense transcription (Figure 4C), and antisense NUTs extended further into the promoter region than for the 828 nonresponsive ORF-Ts (Figure 4D). These results are consistent with previous findings (Xu et al., 2011) and show that antisense NUT transcription can downregulate sense ORF-T transcription when it reaches a certain level, maybe by interfering with sense transcription due to converging

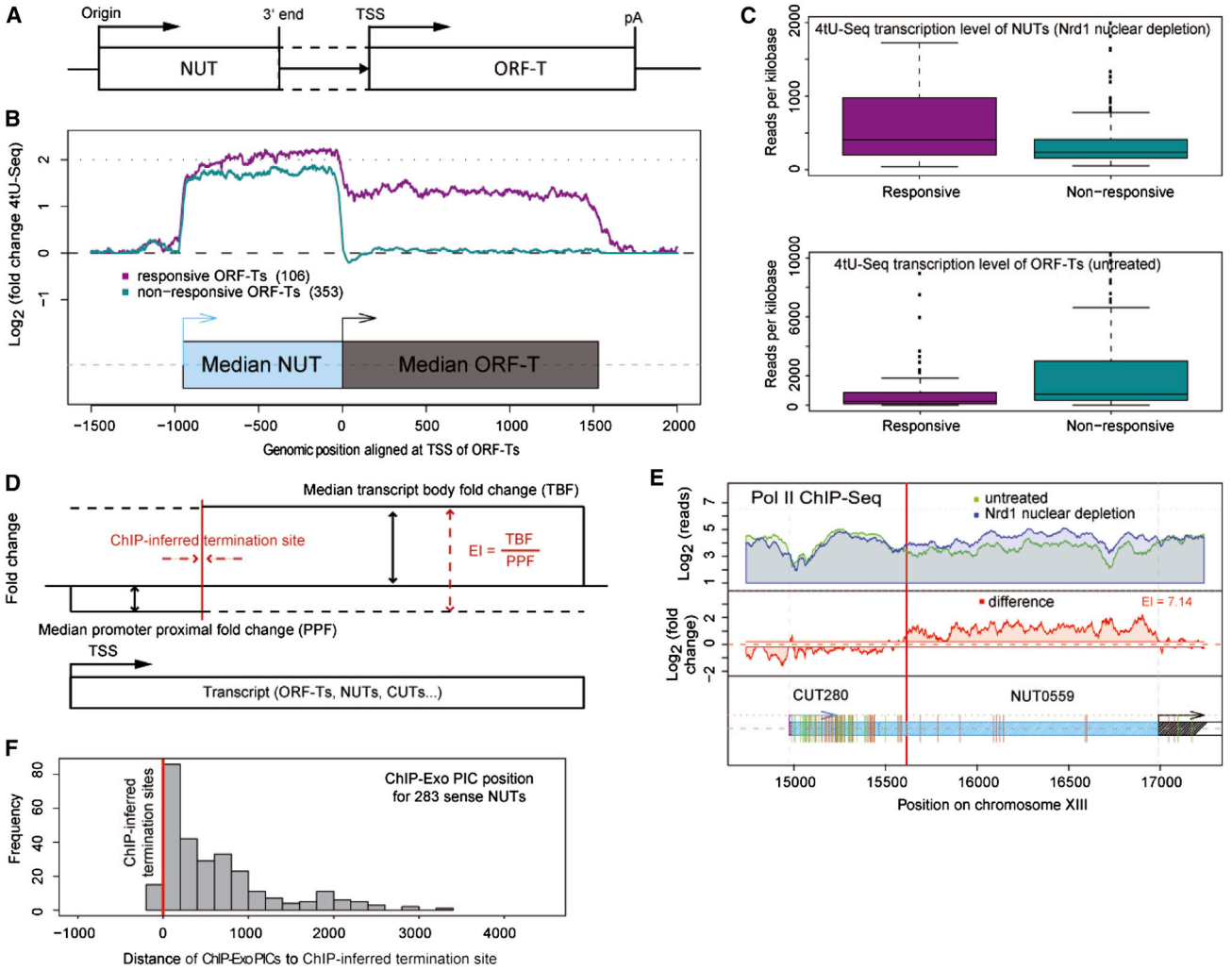

Figure 5. Sense NUTs Can Deregulate Transcription of Downstream ORF-Ts

(A) Definition of the sense NUT class. A NUT belongs to the sense class when its origin is upstream of the TSS of the downstream ORF-T and the distance between the NUT and the ORF-T was not more than 100 bps.

(B) A fraction of sense class NUTs upregulates downstream ORF-T transcription. Median log2 fold-change in 4tU-seq signal upon nuclear depletion of Nrd1 was plotted for responsive (purple) and nonresponsive (green) ORF-Ts of the sense NUT class.

(C) ORF-Ts that are responsive to upstream sense NUT transcription are weakly transcribed, whereas the corresponding NUTs are highly transcribed. Color code as in (B)

(D) Scheme illustrating the determination of termination site and Escape Index (EI) from ChIP-seq data. EIs were calculated as the median fold-change in the transcribed gene body divided by the median fold-change in the proximal TSS region upon nuclear depletion of Nrd1.

(E) Pol II occupancies around the CUT280 locus measured by ChIP-seq in wild-type conditions (green) and upon nuclear depletion of Nrd1 (blue). The position of sign change in the occupancy fold-change difference profile (red) defines the termination site of the CUT (red vertical line). RNA-binding sites of Nrd1 and Nab3 as determined by PAR-CLIP are shown as green and brown vertical lines over the blue bar at the bottom.

(F) Sense NUTs are generally terminated before the promoter of downstream ORF-Ts. The plot shows the frequency of distances from the ncRNA termination site (vertical red line) to the PIC location defined by ChIP-Exo of TFIIB (Rhee and Pugh, 2012). See also Figure S6.

Pol II enzymes (Hobson et al., 2012), although this is considered a rare event in vivo.

Upstream ncRNA Synthesis Can Upregulate Transcription

Another possible mechanism for transcription deregulation involves upstream synthesis of ncRNAs that can interfere with ORF-T transcription (Colin et al., 2011). One-third of all NUTs is found upstream of ORF-Ts with a median NUT origin approximately 1,000 bp upstream of ORF-Ts. We selected all NUTs upstream of ORF-Ts with a maximum distance of 100 bp between the ORF-T and the NUT (sense NUT class, 459 NUTs) (Figure 5A). Downstream of sense NUTs, 106 ORF-Ts showed a 3- to 4-fold increase in 4tU-seq signals, whereas the remaining 353 ORF-Ts were unchanged (Figure 5B). The upregulated ORF-Ts showed lower median RNA synthesis than unchanged

ORF-Ts (Figure 5C), and associated upstream NUTs showed higher levels (Figure 5C). Upstream NUT synthesis was responsible for upregulation of 37% of a total of 287 significantly upregulated ORF-Ts. Only in 28 cases, when the ORF-T was transcribed at high levels, NUT synthesis repressed ORF-T transcription slightly (Figure S6).

Termination of ncRNA Synthesis Prevents Transcription Interference

The above results suggested that early Nrd1-dependent termination of aberrant ncRNAs prevents genome deregulation by NUT synthesis. To further investigate this, we determined termination sites of ncRNAs by mapping Pol II over the genome before and after nuclear depletion of Nrd1. We used chromatin immunoprecipitation (ChIP) as described (Mayer et al., 2010) coupled to deep sequencing (ChIP-seq, Experimental Procedures). The reproducible ChIP-seq replicates had a high correlation with RNA synthesis monitored by 4tU-seq, showing that Pol II ChIP occupancy is a good proxy for transcription activity (Figure S2).

We analyzed changes in Pol II occupancy upon nuclear depletion of Nrd1 by determining changes in an Escape Index (Brannan et al., 2012). For every NUT transcription unit, we calculated an Escape Index (EI) as the ratio of Pol II occupancy fold-change in the promoter-distal versus the promoter-proximal region. An increased EI after Nrd1 depletion indicates defective termination because more Pol II moves to the promoter-distal region. This was indeed observed in the average Pol II occupancy profile of NUTs (Figure S7). We defined the promoter-proximal region as the region between the NUT origin and the transcript termination site (TTS) of the ncRNA generated by normal Nrd1-dependent termination (Figure 5D).

To derive an estimated TTS for a ncRNA, we determined the point downstream of which the density of Pol II increases upon Nrd1 nuclear depletion. Specifically, we determined the point downstream of the NUT origin (maximum distance 1,000 bp) at which the profile of log2 fold change in Pol II occupancy was best approximated by a two-segment piecewise constant function (Figures 5D and 5E) (Experimental Procedures). We derived TTSs for 283 ncRNAs that upon nuclear depletion of Nrd1 gave rise to sense NUTs upstream of ORF-Ts and contain a mapped PIC (Rhee and Pugh, 2012). We then calculated the distance of each TTS to the PIC of the downstream ORF-T (Figure 5F). This revealed that ncRNA synthesis is generally terminated before the transcribing polymerase would clash with the PIC at a downstream ORF-T, apparently to prevent transcription interference.

Transcription Attenuation Is Rare

Visual inspection of our Pol II ChIP-seq data at protein-coding genes that are controlled by Nrd1-dependent attenuation (Arigo et al., 2006a; Steinmetz et al., 2006) revealed an apparent release of Pol II into promoter-distal regions after nuclear depletion of Nrd1 (Figure 6A). To search for genes controlled by attenuation, we extended the EI analysis of our ChIP-seq data to all ORF-Ts. This revealed that transcription attenuation does not generally occur under our experimental conditions (Figure 6B). Only 32 ORF-Ts were classified as attenuated mRNA genes that fulfilled the following three criteria. First, weighted EIs (Extended Experimental Procedures) had to be greater than 2.5 upon nuclear depletion of Nrd1. Second, Pol II occupancy changes in the gene body had to be greater than 1.4-fold. Third, ORF-T transcription had to be upregulated at least 1.25-fold in 4tU-seq data (adjusted p value 0.1). The attenuated genes were generally involved in biosynthetic amino acid and metabolic processes (GO pyrimidine nucleotide biosynthetic pathway, p value 1.8×10^{-3}; GO glutamine metabolic process, p value 6.4×10^{-3}).

Alignment of the 32 selected ORF-Ts at their TSS showed that the average Pol II occupancy was slightly decreased in the promoter-proximal region after nuclear depletion of Nrd1, likely reflecting a loss of early Pol II termination intermediates (Figure 6C). Average Pol II occupancy was however increased from around 400 bp downstream of the TSS (Figure 6C), reflecting an increased density of Pol II in promoter-distal regions after attenuation release. Further consistent with attenuation, PAR-CLIP detected a 3.3-fold higher density of Nrd1- and Nab3-binding sites in the promoter-proximal region of the 32 corresponding mRNAs, compared to random mRNAs (Figure 6C). We conclude that under optimum growth conditions only few genes are controlled by Nrd1-dependent attenuation, and that the main function of the Nrd1 pathway is to suppress ncRNA transcription. More genes may be under attenuation control during nonoptimum growth conditions, such as cell wall stress (Kim and Levin, 2011).

DISCUSSION

The discovery of pervasive genome transcription suggested a mechanism exists for transcriptome surveillance that is based on selective termination and degradation of ubiquitous ncRNA synthesis. This raised four questions. First, what is the origin of ncRNA transcription? Second, what is the global mechanism for ncRNA transcription termination? Third, does a failure to terminate ncRNA synthesis lead to transcriptome deregulation? Fourth, how does the termination mechanism distinguish ncRNA synthesis from mRNA transcription? Answers to these questions are required to establish the concept of transcriptome surveillance.

Here we elucidate these questions in the model eukaryote *S. cerevisiae* using 4tU-seq, PAR-CLIP, ChIP-seq, and motif analysis. We show that ncRNAs generally originate from NDRs in the yeast genome. All yeast promoters are apparently bidirectional, generating divergent ncRNAs that originate 150–200 bps upstream of the TSS of the mRNA gene. We also show that ncRNA synthesis is generally restricted by Nrd1-dependent termination. A defect in ncRNA transcription termination can lead to genome deregulation by antisense repression and transcription interference. We provide evidence that termination of ncRNA transcription is the main function of Nrd1, and attenuation control at mRNA genes is rare. Nrd1 preferentially binds to ncRNAs that frequently contain Nrd1-binding motifs, whereas mRNAs are depleted for these motifs and generally escape Nrd1 action. We conclude that Nrd1-dependent termination serves as a mechanism for transcriptome surveillance that is based on recognition and

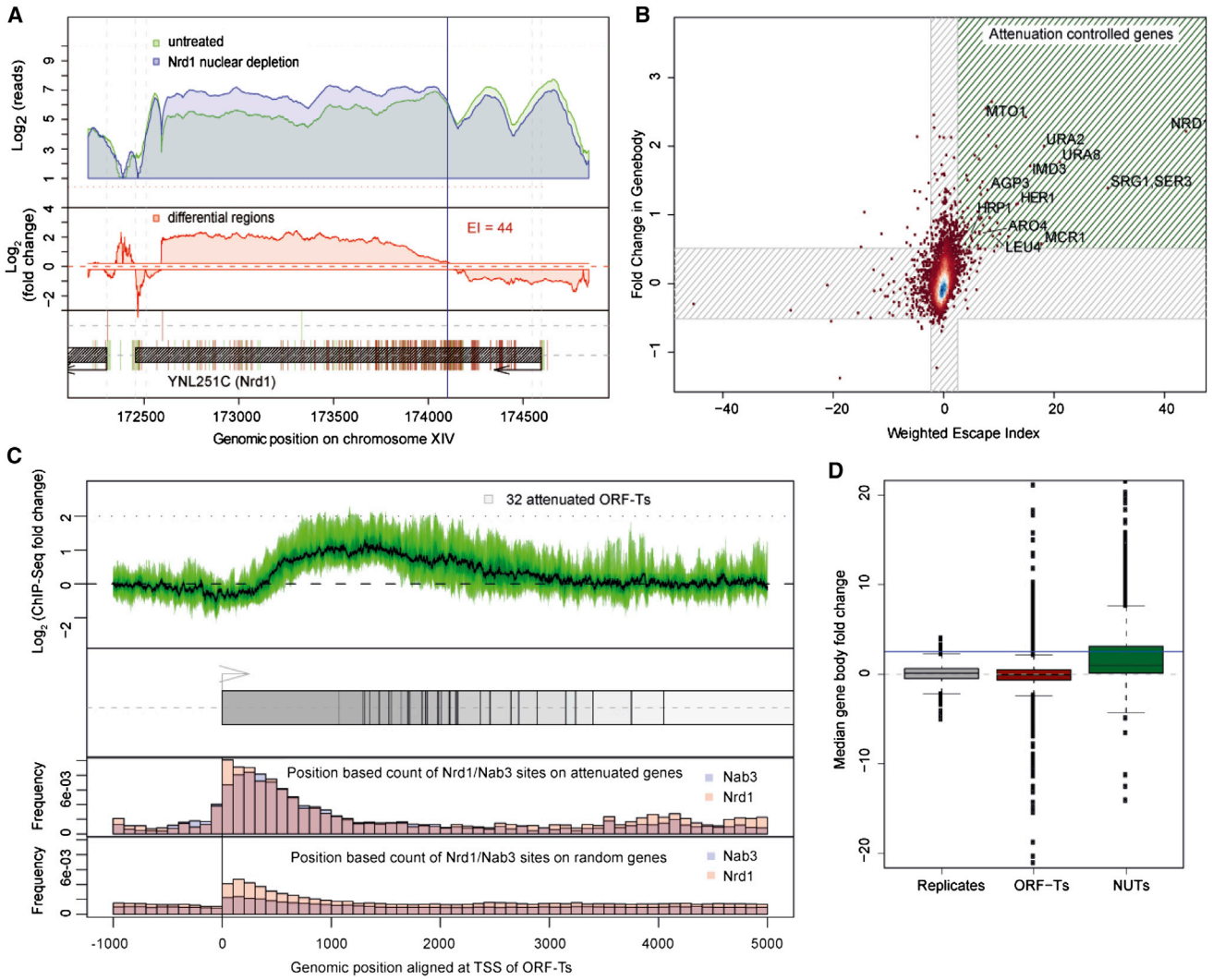

Figure 6. Nrd1-Dependent Transcription Attenuation Is Rare

(A) Log2 Pol II reads from ChIP-seq around the Nrd1 gene locus before (green) and after nuclear depletion of Nrd1 (blue) and calculated log2 differences in ChIP signal (red). The vertical black line indicates the derived early termination/attenuation site. RNA-binding sites of Nrd1 and Nab3 as determined by PAR-CLIP are shown as green and brown vertical lines at the bottom.

(B) Attenuation of mRNA genes upon nuclear depletion of Nrd1 is rare under optimum growth conditions. Only 32 genes show de-attenuation upon nuclear depletion of Nrd1, as indicated by a weighted EI > 2.5 and a > 1.4-fold change in ChIP-seq genebody signal (green hatched region).

(C) De-attenuation leads to Pol II accumulation. Median log2 Pol II occupancy fold-change upon nuclear depletion of Nrd1 for the 32 predicted attenuated genes. Transcripts were aligned at their TSS (gray box). Each early termination/attenuation site is represented by a vertical black line. Nrd1 and Nab3 PAR-CLIP sites are most densely distributed within the first 50 bp after the TSS. In contrast, random genes that are not under attenuation control show few PAR-CLIP sites (bottom, 100 × 32 genes were randomly chosen via resampling, the numbers of sites are normalized to the number of underlying genes and their expression).

(D) Distributions of median changes in Pol II occupancy in ORF-Ts, NUTs, and for a null distribution obtained by using two replicate measurements of Pol II ChIP-seq. The threshold to define attenuated genes is shown as a blue horizontal line. See also Figure S7.

removal of polymerases that produce aberrant nascent ncRNA (Figure 7).

Previous studies that detected divergent RNA transcripts observed bidirectional transcription at about one-third of yeast promoters (Neil et al., 2009; Xu et al., 2009). Here, we could observe many more bidirectional yeast promoters apparently because the short-lived divergent transcript was trapped by crosslinking it to Nrd1. We suggest that promoter directionality is achieved by selective termination of divergent ncRNA transcription. Selective termination may be explained by the difference in the occurrence of Nrd1- and Nab3-binding motifs in ncRNA versus mRNA, in particular because Nrd1 and Nab3 can bind cooperatively to RNA (Carroll et al., 2007). Nrd1 recruitment to early ncRNA transcription complexes may be facilitated by phosphorylation patterns in the C-terminal repeat domain of Pol II (Kubicek et al., 2012; Singh et al., 2009) and by chromatin modifications that are directional (Rando and Chang, 2009; Seila et al., 2008). The formation of mRNA gene loops

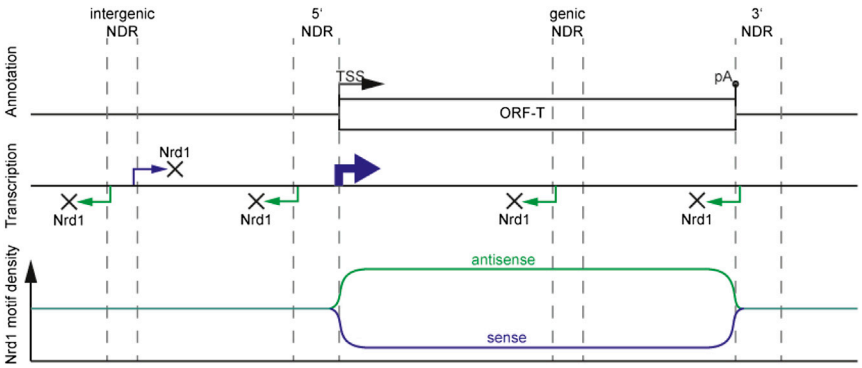

Figure 7. Model for Nrd1-Dependent Transcriptome Surveillance

Transcription of Pol II generally initiates bidirectionally in NDRs, leading to ncRNA synthesis. Transcripts that are antisense to ORF-Ts possess an elevated density of tetramer motifs with high Nrd1- and Nab3-binding affinities, leading to termination and RNA degradation. Intergenic transcripts have intermediate levels of Nrd1/Nab-binding affinity and are also removed. mRNAs originating from ORF-Ts show the lowest density of high-affinity sites, are only weakly bound by Nrd1 and Nab3, and are generally not attenuated by the Nrd1 pathway.

further contributes to transcription directionality (Tan-Wong et al., 2012).

Previous studies of ncRNA transcription in yeast used strains with a deletion of genes *rrp6* and *xrn1* that encode for a nuclear exosome subunit (Neil et al., 2009; Xu et al., 2009) and a RNA exonuclease (van Dijk et al., 2011), respectively. In these strains, RNA degradation is defective, leading to a stabilization of ncRNAs. Our approach of depleting Nrd1 provides more complete insights because Nrd1 acts upstream of Rrp6 and Xrn1 and its binding to nascent RNA is likely the first step in ncRNA transcription termination. Nuclear depletion of Nrd1 gives rise to extended ncRNAs (NUTs) that are on average four times longer than previously described ncRNAs. Whereas the deletion of degradation factors does apparently not change ncRNA transcription activity, nuclear depletion of Nrd1 deregulates the transcriptome by ncRNA transcription interference. Consistent with this, there is no significant overlap between genes that are differentially expressed after nuclear depletion of Nrd1 or *rrp6* deletion (data not shown), and the deletion strains Δ*rrp6* and Δ*xrn1* show mild phenotypes, whereas *nrd1* deletion is lethal.

Our data indicate that in yeast it is unavoidable that transcription initiates where the genome is accessible and that the resulting ncRNA synthesis must be suppressed. Other species apparently have similar transcriptome surveillance systems. In *Escherichia coli*, a termination factor-dependent mechanism for suppression of antisense transcription has been described, and proposed to be related to the Nrd1 pathway (Peters et al., 2012). A mechanism of selective transcription termination was recently shown to restrict ncRNA transcription from mammalian bidirectional promoters (Almada et al., 2013; Core et al., 2008; Ntini et al., 2013; Seila et al., 2008). These studies revealed the same principle for achieving promoter directionality by selective termination but showed that termination is due to the pA-dependent pathway. Almada et al. (2013) reported an assymetric pA site distribution around mRNA TSSs, and that mRNA transcription is protected from termination through increased densities of U1 snRNP-binding sites. The pA motif occurred more frequently upstream of TSSs of mRNAs and could direct termination when inserted into a different sequence context as observed by Ntini et al. (2013).

Although transcriptome surveillance suppresses most ncRNA production, some ncRNAs may escape rapid removal and exhibit a function. Overlap of NUTs with XUTs is less than for other ncRNAs, and 66% of XUTs are antisense to mRNAs and may be involved in gene regulation (van Dijk et al., 2011). In human cells, the fraction of ncRNAs that serve a cellular function is apparently much higher (Mercer et al., 2009). It is also likely that the process of ncRNA transcription itself serves a cellular function such as the maintenance of a chromatin state or the enhanced recruitment of polymerase-associated factors for mRNA transcription.

EXPERIMENTAL PROCEDURES

Anchor-away strains of *S. cerevisiae* containing FRB-tagged Nrd1 were used to deplete Nrd1 from the nucleus by rapamycin treatment (Haruki et al., 2008). Cultures were split at OD_{600} = 0.6 and one-half was treated with rapamycin for 60 min. Samples were taken from treated and untreated cultures for metabolic labeling of newly synthesized RNAs (Sun et al., 2012) followed by deep sequencing (4tU-seq) and for Pol II occupancy profiling by ChIP-seq. PAR-CLIP of Nrd1 was done similar to Creamer et al. (2011) with slight modifications and binding sites were called using a computational pipeline. 4tU-seq, ChIP-seq, and PAR-CLIP sequencing data were processed with Galaxy (Blankenberg et al., 2010; Giardine et al., 2005; Goecks et al., 2010) to obtain pile-ups (reads per nucleotide). Further data processing was done with R/Bioconductor. Detailed descriptions of experiments and computational analyses are found in the Extended Experimental Procedures.

ACCESSION NUMBERS

ArrayExpress Database accession number for all sequencing data is E-MTAB-1766.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures and seven figures and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2013.10.024.

AUTHOR CONTRIBUTIONS

D.S. and P.C. conceived and designed the study. D.S. performed ChIP-seq and 4tU-seq. C.B. performed PAR-CLIP. B.S., D.S., J.S., and J.G. designed data analysis. B.S., A.K., P.T., D.S., and C.B. carried out data analysis. D.S. and P.C. wrote the manuscript with input from all authors. P.C. supervised the project.

ACKNOWLEDGMENTS

We would like to thank Stefan Krebs and Alexander Graf for sequencing and Galaxy maintenance, Christophe Jung for image analysis, and Domenico Libri

for critically reading the manuscript. J.G. was supported by the Bavarian Network for Molecular Biosystems (BaySysNet). J.S. was supported by the Deutsche Forschungsgemeinschaft (SFB646, GRK1721, QBM), the Bundesministerium für Bildung und Forschung (BMBF, ebio), and the Bavarian Network for Systems Biology (BaySysNet). P.C. was supported by the DFG (SFB646, TR5, SFB960, GRK1721, CIPSM, NIM, QBM), an Advanced Investigator Grant of the European Research Council, the Deutsches Konsortium für Translationale Krebsforschung DKTK, the Jung-Stiftung, and the Vallee Foundation.

Received: July 9, 2013
Revised: September 12, 2013
Accepted: October 3, 2013
Published: November 7, 2013

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.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Arigo, J.T., Carroll, K.L., Ames, J.M., and Corden, J.L. (2006a). Regulation of yeast NRD1 expression by premature transcription termination. Mol. Cell *21*, 641–651.

Arigo, J.T., Eyler, D.E., Carroll, K.L., and Corden, J.L. (2006b). Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol. Cell *23*, 841–851.

Blankenberg, D., Von Kuster, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. Curr. Protoc. Mol. Biol. *Chapter 19*, Unit 19 10 11-21.

Brannan, K., Kim, H., Erickson, B., Glover-Cutter, K., Kim, S., Fong, N., Kiemele, L., Hansen, K., Davis, R., Lykke-Andersen, J., and Bentley, D.L. (2012). mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. Mol. Cell *46*, 311–324.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Mol. Cell *36*, 541–546.

Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G., and Stutz, F. (2007). Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. Cell *131*, 706–717.

Carroll, K.L., Pradhan, D.A., Granek, J.A., Clarke, N.D., and Corden, J.L. (2004). Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol. Cell. Biol. *24*, 6241–6252.

Carroll, K.L., Ghirlando, R., Ames, J.M., and Corden, J.L. (2007). Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. RNA *13*, 361–373.

Castelnuovo, M., Rahman, S., Guffanti, E., Infantino, V., Stutz, F., and Zenklusen, D. (2013). Bimodal expression of PHO84 is modulated by early termination of antisense transcription. Nat. Struct. Mol. Biol. *20*, 851–858.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368–373.

Colin, J., Libri, D., and Porrua, O. (2011). Cryptic transcription and early termination in the control of gene expression. Genet. Res. Int. *2011*, 653494.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Creamer, T.J., Darby, M.M., Jamonnak, N., Schaughency, P., Hao, H., Wheelan, S.J., and Corden, J.L. (2011). Transcriptome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. PLoS Genet. *7*, e1002329.

David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C.J., Bofkin, L., Jones, T., Davis, R.W., and Steinmetz, L.M. (2006). A high-resolution map of transcription in the yeast genome. Proc. Natl. Acad. Sci. USA *103*, 5320–5325.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature *489*, 101–108.

Giardine, B., Riemer, C., Hardison, R.C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., et al. (2005). Galaxy: a platform for interactive large-scale genome analysis. Genome Res. *15*, 1451–1455.

Goecks, J., Nekrutenko, A., Taylor, J., and Galaxy, T.; Galaxy Team. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. *11*, R86.

Gudipati, R.K., Xu, Z., Lebreton, A., Séraphin, B., Steinmetz, L.M., Jacquier, A., and Libri, D. (2012). Extensive degradation of RNA precursors by the exosome in wild-type cells. Mol. Cell *48*, 409–421.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., et al. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell *141*, 129–141.

Hani, J., and Feldmann, H. (1998). tRNA genes and retroelements in the yeast genome. Nucleic Acids Res. *26*, 689–696.

Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Mol. Cell *31*, 925–932.

Hobson, D.J., Wei, W., Steinmetz, L.M., and Svejstrup, J.Q. (2012). RNA polymerase II collision interrupts convergent transcription. Mol. Cell *48*, 365–374.

Hongay, C.F., Grisafi, P.L., Galitski, T., and Fink, G.R. (2006). Antisense transcription controls cell fate in Saccharomyces cerevisiae. Cell *127*, 735–745.

Honorine, R., Mosrin-Huaman, C., Hervouet-Coste, N., Libri, D., and Rahmouni, A.R. (2011). Nuclear mRNA quality control in yeast is mediated by Nrd1 co-transcriptional recruitment, as revealed by the targeting of Rho-induced aberrant transcripts. Nucleic Acids Res. *39*, 2809–2820.

Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008). A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. Mol. Cell *32*, 685–695.

Hsin, J.P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. *26*, 2119–2137.

Jacquier, A. (2009). The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. Nat. Rev. Genet. *10*, 833–844.

Kim, K.Y., and Levin, D.E. (2011). Mpk1 MAPK association with the Paf1 complex blocks Sen1-mediated premature transcription termination. Cell *144*, 745–756.

Kim, M., Vasiljeva, L., Rando, O.J., Zhelkovsky, A., Moore, C., and Buratowski, S. (2006). Distinct pathways for snoRNA and mRNA termination. Mol. Cell *24*, 723–734.

Kishore, S., Jaskiewicz, L., Burger, L., Hausser, J., Khorshid, M., and Zavolan, M. (2011). A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat. Methods *8*, 559–564.

König, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2011). Protein-RNA interactions: new genomic technologies and perspectives. Nat. Rev. Genet. *13*, 77–83.

Kubicek, K., Cerna, H., Holub, P., Pasulka, J., Hrossova, D., Loehr, F., Hofr, C., Vanacova, S., and Stefl, R. (2012). Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1. Genes Dev. *26*, 1891–1896.

Kuehner, J.N., and Brow, D.A. (2008). Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. Mol. Cell *31*, 201–211.

Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3′-end processing. Cell. Mol. Life Sci. *65*, 1099–1122.

Marquardt, S., Hazelbaker, D.Z., and Buratowski, S. (2011). Distinct RNA degradation pathways and 3′ extensions of yeast non-coding RNA species. Transcription *2*, 145–154.

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. Genome Res. *18*, 1073–1083.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J., and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nat. Struct. Mol. Biol. *17*, 1272–1278.

Mayer, A., Heidemann, M., Lidschreiber, M., Schreieck, A., Sun, M., Hintermair, C., Kremmer, E., Eick, D., and Cramer, P. (2012). CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. Science *336*, 1723–1725.

Mercer, T.R., Dinger, M.E., and Mattick, J.S. (2009). Long non-coding RNAs: insights into functions. Nat. Rev. Genet. *10*, 155–159.

Mischo, H.E., and Proudfoot, N.J. (2013). Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. Biochim. Biophys. Acta *1829*, 174–185.

Murray, S.C., Serra Barros, A., Brown, D.A., Dudek, P., Ayling, J., and Mellor, J. (2012). A pre-initiation complex at the 3′-end of genes drives antisense transcription independent of divergent sense transcription. Nucleic Acids Res. *40*, 2432–2444.

Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L.M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature *457*, 1038–1042.

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.

Peters, J.M., Mooney, R.A., Grass, J.A., Jessen, E.D., Tran, F., and Landick, R. (2012). Rho and NusG suppress pervasive antisense transcription in Escherichia coli. Genes Dev. *26*, 2621–2633.

Porrua, O., Hobor, F., Boulay, J., Kubicek, K., D'Aubenton-Carafa, Y., Gudipati, R.K., Stefl, R., and Libri, D. (2012). In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. EMBO J. *31*, 3935–3948.

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.

Rando, O.J., and Chang, H.Y. (2009). Genome-wide views of chromatin structure. Annu. Rev. Biochem. *78*, 245–271.

Rhee, H.S., and Pugh, B.F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature *483*, 295–301.

Rondón, A.G., Mischo, H.E., Kawauchi, J., and Proudfoot, N.J. (2009). Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. Mol. Cell *36*, 88–98.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849–1851.

Seila, A.C., Core, L.J., Lis, J.T., and Sharp, P.A. (2009). Divergent transcription: a new feature of active promoters. Cell Cycle *8*, 2557–2564.

Singh, N., Ma, Z., Gemmill, T., Wu, X., Defiglio, H., Rossettini, A., Rabeler, C., Beane, O., Morse, R.H., Palumbo, M.J., and Hanes, S.D. (2009). The Ess1 prolyl isomerase is required for transcription termination of small noncoding RNAs via the Nrd1 pathway. Mol. Cell *36*, 255–266.

Steinmetz, E.J., and Brow, D.A. (1996). Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. Mol. Cell. Biol. *16*, 6993–7003.

Steinmetz, E.J., Conrad, N.K., Brow, D.A., and Corden, J.L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. Nature *413*, 327–331.

Steinmetz, E.J., Warren, C.L., Kuehner, J.N., Panbehi, B., Ansari, A.Z., and Brow, D.A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol. Cell *24*, 735–746.

Sun, M., Schwalb, B., Schulz, D., Pirkl, N., Etzold, S., Larivière, L., Maier, K.C., Seizl, M., Tresch, A., and Cramer, P. (2012). Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. Genome Res. *22*, 1350–1359.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E., Ansari, A.Z., Luscombe, N.M., Steinmetz, L.M., and Proudfoot, N.J. (2012). Gene loops enhance transcriptional directionality. Science *338*, 671–675.

Thiebaut, M., Colin, J., Neil, H., Jacquier, A., Séraphin, B., Lacroute, F., and Libri, D. (2008). Futile cycle of transcription initiation and termination modulates the response to nucleotide shortage in S. cerevisiae. Mol. Cell *31*, 671–682.

Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J., and Libri, D. (2006). Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. Mol. Cell *23*, 853–864.

van Dijk, E.L., Chen, C.L., d'Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoix-Né, P., Loeillet, S., et al. (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature *475*, 114–117.

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., and Meinhart, A. (2008). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat. Struct. Mol. Biol. *15*, 795–804.

Wei, W., Pelechano, V., Järvelin, A.I., and Steinmetz, L.M. (2011). Functional consequences of bidirectional promoters. Trends Genet. *27*, 267–276.

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

Windhager, L., Bonfert, T., Burger, K., Ruzsics, Z., Krebs, S., Kaufmann, S., Malterer, G., L'Hernault, A., Schilhabel, M., Schreiber, S., et al. (2012). Ultrashort and progressive 4sU-tagging reveals key characteristics of RNA processing at nucleotide resolution. Genome Res. *22*, 2031–2042.

Wlotzka, W., Kudla, G., Granneman, S., and Tollervey, D. (2011). The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. EMBO J. *30*, 1790–1803.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.C., Dufour, M.E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., et al. (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell *121*, 725–737.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033–1037.

Xu, Z., Wei, W., Gagneur, J., Clauder-Münster, S., Smolik, M., Huber, W., and Steinmetz, L.M. (2011). Antisense expression increases gene expression variability and locus interdependency. Mol. Syst. Biol. *7*, 468.

Yassour, M., Pfiffner, J., Levin, J.Z., Adiconis, X., Gnirke, A., Nusbaum, C., Thompson, D.A., Friedman, N., and Regev, A. (2010). Strand-specific RNA sequencing reveals extensive regulated long antisense transcripts that are conserved across yeast species. Genome Biol. *11*, R87.