Protein Abundance Control by Non-coding Antisense Transcription

**Highlights**
- The impact of antisense transcription on protein levels is measured for 188 genes
- Antisense has mostly weak suppressive effects on ~25% of the genes
- Regulation by antisense correlates with promoter overlap and H3K4 methylation
- Antisense-regulated genes have lower than expected noise levels

**Accession Numbers**
E-MTAB-4731

**Authors**
Florian Huber, Daria Bunina, Ishaan Gupta, ..., Patrick Theer, Lars M. Steinmetz, Michael Knop

**Correspondence**
m.knop@zmbh.uni-heidelberg.de

**In Brief**
Huber et al. conduct a systematic study of the function of antisense RNAs in yeast by selectively suppressing the antisense transcripts of 188 GFP-tagged genes and quantifying the resulting changes in protein expression levels. The authors use this information to identify features that distinguish functional from non-functional antisense RNAs.
Protein Abundance Control by Non-coding Antisense Transcription

Florian Huber,1,6 Daria Bunina,1,6 Ishaan Gupta,2 Anton Khmelinskii,1 Matthias Meurer,1 Patrick Theer,1,3 Lars M. Steinmetz,2,4,5 and Michael Knop1,3,*
1Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), University of Heidelberg, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany
2Genome Biology Unit, European Molecular Biology Laboratory (EMBL), Meyerhofstraße 1, 69117 Heidelberg, Germany
3Deutsches Krebsforschungszentrum (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
4Stanford Genome Technology Center, Stanford University, Palo Alto, CA 94304, USA
5Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA
6Co-first author
*Correspondence: m.knop@zmbh.uni-heidelberg.de

http://dx.doi.org/10.1016/j.celrep.2016.05.043

SUMMARY

Stable unannotated transcripts (SUTs), some of which overlap protein-coding genes in antisense direction, are a class of non-coding RNAs. While case studies have reported important regulatory roles for several of such RNAs, their general impact on protein abundance regulation of the overlapping gene is not known. To test this, we employed seamless gene manipulation to repress antisense SUTs of 162 yeast genes by using a unidirectional transcriptional terminator and a GFP tag. We found that the mere presence of antisense SUTs was not sufficient to influence protein abundance, that observed effects of antisense SUTs correlated with sense transcript start site overlap, and that the effects were generally weak and led to reduced protein levels. Antisense regulated genes showed increased H3K4 di- and trimethylation and had slightly lower than expected noise levels. Our results suggest that the functionality of antisense RNAs has gene and condition-specific components.

INTRODUCTION

High-throughput technologies such as tiling arrays and deep sequencing enable genome-wide and strand-specific detection of RNAs and have revealed the pervasive nature of transcription in eukaryotic genomes (Bertone et al., 2004; David et al., 2006; Nagaralakshmi et al., 2008), resulting in the identification of many classes of non-coding RNAs (ncRNAs). In Saccharomyces cerevisiae, ncRNAs typically originate from nucleosome-depleted regions (NDRs), which are frequently associated with bidirectional promoters of protein-coding genes (Neil et al., 2009; Xu et al., 2009). However, such pervasive transcription from NDRs is limited by a combination of transcriptome surveillance mechanisms such as transcription attenuation mediated by the Nrd1-Nab3-Sen1 (NNS) termination complex (Arigo et al., 2006; Schulz et al., 2013), suppression of divergent transcription via histone marks (Churchman and Weissman, 2011; Marquardt et al., 2014; Upreti et al., 2016), or rapid degradation of the resulting transcripts by the exosome (Davis and Ares, 2006; van Dijk et al., 2011; Wyers et al., 2005).

In contrast to exosome-sensitive cryptic unstable transcripts (CUTs), stable unannotated transcripts (SUTs) are readily detectable in wild-type cells, and more than 600 of such transcripts have been annotated in yeast (Xu et al., 2009). When ncRNAs are transcribed in antisense direction to an open reading frame (ORF), they are also referred to as antisense RNAs (asRNAs).

In a number of detailed studies, asRNAs were found to exert important biological functions by regulating the expression of the overlapping gene. For example, the asRNA RME2 blocks entry into meiosis in haploid yeast cells by repressing transcription elongation of the IME4 gene (Gelfand et al., 2011; Hongay et al., 2006). Lenstra and colleagues demonstrated that an RNA antisense to GAL10 prevents transcriptional leakage of both GAL10 and GAL1, thus modulating the responsiveness of the underlying metabolic switch (Lenstra et al., 2015). In addition, strong regulatory functions of asRNAs in yeast have been shown for several genes, including CDC28 (Nadal-Ribelles et al., 2014), PHO84 (Camblong et al., 2009, 2007; Castelnuovo et al., 2013), PHO5 (Uhler et al., 2007), and IME1 (van Werven et al., 2012), each with individual mechanistic characteristics different from RNAi, as S. cerevisiae lacks a functional RNAi machinery (Dringenberg et al., 2011, 2009).

While these cases are well established, our understanding of which asRNAs serve a biological function and whether those share certain characteristics remains incomplete (Pelechano and Steinmetz, 2013). Several high-throughput transcriptome studies generated correlative data on transcript levels of sense-antisense pairs on a genome-wide scale. From these studies, a picture emerges where sense levels are anticorrelated with antisense. However, not all sense transcript levels are affected by changes in antisense transcript levels, and the anticorrelation is often weak (Alcid and Tsukiyama, 2014; Castelnuovo et al., 2013; Schulz et al., 2013; Xu et al., 2011, 2009). Importantly, others have pointed out that there is a lack of anticorrelation at the level of nascent transcription (Murray et al., 2013).
In the latter study, Murray and colleagues detected specific chromatin signatures associated with antisense transcription. From this, they hypothesized that the main biological function of asRNAs lies in affecting traits such as expression noise or gene silencing rather than affecting bulk protein abundances. Indeed, studies showing the on/off switching of sense-antisense pairs in different conditions (Lenstra et al., 2015; Nguyen et al., 2014) and suggesting effects of antisense transcription on noise levels (Xu et al., 2011) exist. However, how common such roles are among all antisense transcripts remains to be determined.

Therefore, our understanding of the general principles that govern antisense-dependent gene regulation remains incomplete (Pelechano and Steinmetz, 2013). Some of the main reasons concern methodological difficulties. First, studies that measure protein rather than RNA levels are lacking. Second, datasets typically rely on correlative data and use mutants that interfere with ncRNA stability, since direct ncRNA abrogation is difficult given the overlapping arrangement of sense-antisense pairs. Hence, the causality of sense-antisense regulation is often unclear. In addition, determining the proportion of antisense transcripts that have a biological function is difficult. Finally, it remains an open question which features are predictive of asRNA functionality and whether such features are shared among many asRNAs.

Here, we interfered with antisense transcription of 162 genes and assessed the impact of this disturbance on the level of the expressed protein using single-cell microscopy. The strategy used is based on seamless genomic manipulations (Khmelinskii et al., 2011) and a specific DNA element from the PHO5 termination region that blocks transcription in a unidirectional manner (Irniger et al., 1991). We applied this strategy to 188 genes in S. cerevisiae, 162 of which had been annotated with antisense SUTs. We then assessed the resulting changes of protein levels of GFP-tagged variants of the overlapping genes by high-throughput fluorescence microscopy in four growth conditions and by using flow cytometry. This allowed us to investigate the general impact of asRNAs on protein abundance and gene expression regulation.

RESULTS

Antisense Library Construction

A previous study on yeast polyadenylation sites reported that a short sequence of ~100 bp from the 3′ intergenic region of the PHO5 gene acts as a unidirectional terminator (Irniger et al., 1991). To explore whether such an element could be used for the specific abrogation of antisense transcription, we tailored a strategy based on seamless gene tagging (Khmelinskii et al., 2011). We inserted this fragment, termed PHO5\textsubscript{Tscr}, and a scrambled control, PHO5\textsubscript{Tscr}, directly downstream of the stop codon of genes so that no auxiliary sequences such as marker genes were left behind after insertion and seamless marker excision. In order to measure protein levels, we simultaneously fused superfolder GFP (sfGFP; Pédelacq et al., 2006) to the C terminus of the proteins (Figures 1A and 1B). We tested this strategy using IME4, a gene that is suppressed by the asRNA RME2 in haploid yeast (Gelfand et al., 2011; Hongay et al., 2006). We observed Ime4-sfGFP mRNA and protein upregulation and asRNA down-regulation upon insertion of the PHO5\textsubscript{Tscr} element, but not with the PHO5\textsubscript{Tscr} element or when sfGFP was inserted alone (Figure 1C). In diploid yeast, transcription of RME2 is suppressed (Hongay et al., 2006). Accordingly, using strand-specific qRT-PCR on IME4 in a diploid context, we could not detect asRNA. Notably, there were no significant differences in sense RNA levels between the three constructs, indicating that PHO5\textsubscript{T} and PHO5\textsubscript{Tscr} do not, in general, lead to changes in RNA stability or transcription rates when compared to the control (Figure S1A).

We also applied our strategy to RSC58, a gene with no reported asRNA. Sense RNA and protein levels did not change (Figures S1B and S1C), indicating that the insertion of PHO5\textsubscript{T} and PHO5\textsubscript{Tscr} did not affect transcription of this gene. Interestingly, for this gene, qRT-PCR detected low levels of an antisense transcript, which was repressed by the PHO5\textsubscript{T} element. This indicates that current SUT/CUT annotations are probably a conservative estimate dictated by the sensitivity of the specific assays used for their detection and that antisense transcription did not influence sense levels of this particular gene.

We conclude that our strategy represents a suitable approach to study the impact of antisense transcription on protein levels by specifically disrupting antisense transcription while leaving sense transcription unaffected.

The above results encouraged us to apply the PHO5\textsubscript{T} strategy to study the function of a larger number of asRNAs. To this end, we tagged approximately one-quarter of the 613 yeast ORFs that are annotated with antisense SUTs starting downstream of the STOP codon (Xu et al., 2009), using sfGFP-PHO5\textsubscript{T}, sfGFP-PHO5\textsubscript{Tscr} or sfGFP alone. A total of 188 ORFs were chosen randomly from the subset of genes predicted to be not affected by the C-terminal sfGFP tag (based on previous genome-wide datasets; Gavin et al., 2006; Ghaemmaghami et al., 2003; Huh et al., 2003; Khmelinskii et al., 2014). The selected genes differ in the extent to which the asRNA overlaps the sense. For 81 genes, the annotated asRNA overlaps the transcript start site (TSS). For another 81 genes, the asRNA terminates earlier. In addition, we selected 26 genes that contain no annotated asRNA (Figure 1D; Table S1). To control for PCR errors associated with PCR targeting during library construction, we used two biological replicates of each strain for subsequent experiments (Experimental Procedures).

Analysis of the Antisense Library Using Quantitative Microscopy

To determine the effect of antisense transcription on protein abundance, we compared sfGFP intensities of the PHO5\textsubscript{T}, PHO5\textsubscript{Tscr} and control strains in our library. For this purpose, we established and validated a high-throughput quantitative fluorescence microscopy pipeline and quantified total cellular sfGFP fluorescence in >100 cells for each of the strains using automated image acquisition and analysis (Figure 2A; Experimental Procedures). We used mid-log phase cells grown at 30°C using the four growth conditions in which the SUTs had been annotated (Xu et al., 2009): rich medium with either glucose, galactose, or ethanol (YPAD, YPGal, and YPE) and synthetic complete medium (SC). Approximately 5% of the data points were removed based on different quality control criteria (Experimental Procedures; Figures S2A–S2C; Tables S2 and S3).
Because the original RNA expression dataset used to annotate the SUTs was obtained from a diploid strain (Xu et al., 2009), we repeated the strand-specific tiling array assay using the haploid library background strain and using the same four growth conditions. Microscopy-based protein measurements and RNA levels measured by tiling arrays exhibited the correlation expected from the literature ($R^2$ between 0.49 and 0.62; Figure S2D; Csárdi et al., 2015). Depending on the growth condition, between 121 and 139 genes showed sfGFP fluorescence above background (Figure 2B), even though our library was enriched for low-expressed genes (Figure S2E). This demonstrates the validity and sensitivity of our approach.

Next, we identified genes where the suppression of antisense transcription led to significant changes in protein abundance (Figures 2B and 2C; the genes are listed in Table S4). We calculated...
Figure 2. High-Throughput Quantitative Microscopy Identifies Antisense-Regulated Genes

(A) Cells were imaged, and whole-cell fluorescence intensities were measured using co-cultured non-fluorescent cells (identified by differential labeling with an mCherry reporter) for background subtraction and fluorescence normalization on a per-well basis (Experimental Procedures). Example cells for IME4 in YPAD are shown. Gray, non-fluorescent cells; green, sfGFP-expressing sample cells.

(B) Venn diagram showing the number of antisense-regulated genes under the different growth conditions. Numbers in parentheses indicate the total number of genes with sfGFP levels above background (for PHO5T and/or PHO5T:scr). In total, 41 genes were found to be antisense regulated under at least one condition.

(C) A volcano plot of p values versus log2-fold changes: ratio of sfGFP levels between PHO5T:scr (with antisense) and PHO5T (no antisense) strains. Each dot represents one gene in one growth condition. Red indicates genes that met our criteria for being regulated by antisense (see text).

(legend continued on next page)
the average fold change of the sfGFP levels between the PHO5\textsubscript{T\_scr} and the PHO5\textsubscript{T} constructs. These values were well reproducible, as judged by a repetition of the experiment using a subset of the strains ($R^2 = 0.81$; Figure S2F). $p$ values were obtained using a linear modeling approach (Experimental Procedures), and the significance threshold was set to $p < 0.01$. In addition, we stipulated that protein expression values of PHO5\textsubscript{T\_scr} and sfGFP (wt) constructs be within a 50% expression range of each other, which was true for all but six cases, and that protein abundances between the PHO5\textsubscript{T} and PHO5\textsubscript{T\_scr} constructs differed more than biological replicates (Experimental Procedures). Depending on the growth condition, we found between 14 out of 121 (12%) and 31 out of 125 (25%) genes with significant differences between the PHO5\textsubscript{T} and PHO5\textsubscript{T\_scr} constructs. This corresponds to 41 genes where significant regulation by antisense was observed under at least one condition (Figures 2B and 2C).

On average, the presence of antisense led to an $\sim$2-fold reduction in protein abundance, with effect sizes ranging from 1.35-fold (CHS7) to 6.3-fold (YKL068W-A; Figure 2D). Only one gene, AMS1, was positively regulated by antisense (Figures 2C and 2D). Depending on the growth condition, between 1 and 9 of the antisense-regulated genes had in the PHO5\textsubscript{T\_scr} background sfGFP levels that were below the threshold we determined for calling a gene “expressed,” suggesting a potential antisense-dependent on/off switch for these genes (Figure 2D; Table S4).

Next, we looked for genes that were regulated by antisense under only a subset of the conditions where expression was observed (Experimental Procedures). We found five such genes (Figure 2E). Interestingly, in all of those cases regulation was absent in YPE. Mostly, however, regulation between conditions did not change significantly.

To validate our findings, we used northern blotting to detect the mRNAs of a selected number of antisense-regulated genes (Figure S3A). In agreement with microscopy results, the mRNA levels were higher in the strains with the PHO5\textsubscript{T} than in the strains with the control or PHO5\textsubscript{T\_scr} constructs. In some instances, we noted that the sense transcripts were slightly shorter for PHO5\textsubscript{T\_scr} background sfGFP levels that were below the threshold we determined for calling a gene “expressed,” suggesting a potential antisense-dependent on/off switch for these genes (Figure 2D; Table S4).

These data shed light on which antisense transcripts influence sense protein levels in a native context and provide evidence that SUTs can change the protein abundance of their parent genes in at least 12%–25% of the cases. The majority of the genes was found to be negatively regulated by antisense.

The Regulatory Effect of Antisense Correlates with Antisense Levels and Anticorrelates with Sense Levels

In order to better understand the relationship of antisense–dependent effects and gene expression, we first explored whether there is a correlation between the strength of the effect of antisense regulation (microscopy data) and sense/antisense RNA levels (tiling array data). We found that the repressive effect of antisense is stronger with increasing antisense levels (Spearman’s $p = -0.20$; Figures 3A and S4A). However, the antisense to sense ratio for a particular gene was a stronger predictor for the repressive effect ($p = -0.43$; Figures S4B and S4C). Finally, we found that the repressive effect of antisense declines with increasing protein expression levels ($p = 0.44$; Figures 3B and S4D) and that highly expressed genes (sfGFP\textsubscript{norm} of PHO5\textsubscript{T\_scr} > 1.63) did not exhibit regulation by antisense. These trends suggest that sense and antisense inhibit one another. However, there is also a considerable variability between genes in their susceptibility for antisense dependent regulation.

Regulation by Antisense Correlates with Transcript Start Site Overlap

Previous studies showed that antisense transcription across the TSS of the overlapping gene may cause repression of the latter mediated by a variety of chromatin-dependent mechanisms (Castelnuovo et al., 2014, 2013; van Werven et al., 2012). Consistent with this, we found that genes whose antisense transcript overlaps the TSS had a higher chance of being antisense regulated than genes whose antisense transcript terminated earlier (Figure 4A). While the size of the antisense effect was not higher in the case of TSS overlap (Figure 4B), more genes showed an “all or nothing” behavior (11/26 in set 1 versus 2/12 in set 2; Table S4).

TSS overlaps were classified based on tiling array data. Native elongating transcript sequencing (NET-seq; Churchman and Weissman, 2011) is a complementary method that reports on the positions of actively transcribing RNA polymerase complexes. It is thus capable of detecting transcriptional events not visible by bulk analysis methods such as tiling arrays. We made use of available NET-seq data (Churchman and Weissman, 2011) and plotted their read numbers in antisense direction relative to the position of the TSS or the stop codon of the genes in our dataset (Figure 4C). This revealed for antisense-regulated genes the presence of active antisense transcription at the sense TSS even if the respective SUTs were annotated to not overlap the TSS. Similarly, regulated genes without annotated SUTs also showed reads near the TSS. In contrast, non-regulated genes in those two groups did not show any reads near the TSS. For those genes where TSS overlap was annotated, both regulated and non-regulated genes showed reads near the TSS. This suggests that TSS overlap is not a sufficient determinant of regulation. As an alternative explanation, we noted that the peak of antisense transcription for the non-regulated genes was clearly shifted into the coding sequence at the stop codon. This could mean that some of

(D) Log2-fold changes between PHO5\textsubscript{T\_scr} and PHO5\textsubscript{T} for each growth condition. Selected examples are labeled. Short black bars indicate the median. Colored circles indicate genes whose sfGFP intensities could not be reliably distinguished from background for the PHO5\textsubscript{T\_scr} construct. To be able to visualize these genes in the diagram, a small offset for the PHO5\textsubscript{T\_scr} expression value was introduced if necessary (see Experimental Procedures for details).

(E) Box plots of sfGFP intensities (sfGFP\textsubscript{norm}) normalized to the fluorescence of the control cells without sfGFP are shown for genes found to exhibit condition-specific regulation by antisense. Only the conditions differing in their regulatory strength are shown. Numbers indicate the log2-fold change of PHO5\textsubscript{T\_scr} over PHO5\textsubscript{T}. Red numbers indicate that the gene was found to be significantly regulated in that condition and fill colors represent the construct (PHO5\textsubscript{T} or PHO5\textsubscript{T\_scr}). Populations of both biological replicates were pooled in the boxplots for simplicity.
the antisense transcripts in this group initiated in the coding region, which would result in a failure of termination by the PHO5T construct.

In summary, the NET-seq data analysis further strengthens the notion of a functional connection between antisense regulation and antisense transcription across the sense TSS.

Antisense-Regulated Genes Show Increased H3K4 Di- and Trimethylation

We proceeded to identify more features correlating with regulation by antisense. To ensure that potential differences are not due to differences in transcript levels, we did not consider genes without annotated antisense for this and subsequent comparisons. Of the remaining genes, antisense regulated and non-regulated genes showed similar sense and antisense levels, as measured by tiling arrays (Figures S5A and S5B). Since histone modifications have been implicated in antisense-dependent gene regulation (Camblong et al., 2007; Castelnuovo et al., 2014; Uhler et al., 2007; van Werven et al., 2012), we first tested whether specific chromatin modifications are associated with functional antisense transcripts. We made use of available datasets (Experimental Procedures) and compared histone traces of antisense-regulated and non-regulated genes relative to their TSSs and their stop codons. This revealed increased H3K4 di- and trimethylation (H3K4me2/3) at the 3' end of the regulated genes (Figures 5A and 5B). H3R2me2 is known to counteract H3K4me3 (Kirmizis et al., 2007). Consistent with this, H3R2me2 was decreased in this region (Figure S5C). We could not detect significant changes for other histone modifications (Table S5). This implies a functional relevance of H3K4 methylation in gene regulation by antisense transcripts.

Lack of Obvious Sequence Features Associated with Regulation by Antisense

Next, we searched for sequence motifs associated with genes in either the antisense regulated or the non-regulated group. Extensive searches using databases of annotated DNA-protein binding sites, a catalog of RNA-protein interaction sites, and de novo motif identification tools did not reveal any significant results (Bailey et al., 2009; Hogan et al., 2008; Shannon, 2015; Vectors et al., 2011; see Experimental Procedures for further details). Of special interest in this respect are motifs specific for transcription attenuation by the NNS complex known to limit non-coding transcription of CUTs in yeast (Arigo et al., 2006; Schulz et al., 2013). The involvement of the NNS complex in early termination and the correlation of regulation by antisense and TSS overlap suggested that non-regulated genes might be enriched in Nrd1-Nab3 binding sites. To test this, we used previously published data on Nrd1 binding sites (Hogan et al., 2008). The fraction of genes that displayed at least one Nrd1 binding site in antisense direction was increased in non-regulated genes versus regulated ones (25 out of 92 versus 5 out of 39), but this increase was not significant (Figure 5C) and may in part be due to the increased length of non-regulated genes versus regulated ones (Figure 5D). A similar comparison using strand-specific photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) data (Creamer et al., 2011) did not reveal any significant differences either (data not shown).

Together, this suggests that antisense transcripts do not contain sequence information sufficient to predict their regulatory potential.
Antisense-Regulated Genes Show Reduced Protein Expression Noise

Previous studies discussed a correlation between the presence of an overlapping antisense transcript and the degree of cell-to-cell variability (“noise”) in protein abundance (Pelechano and Steinmetz, 2013; Xu et al., 2011). We followed the flow cytometry based noise measurement strategy of Newman et al. (Newman et al., 2006) to investigate differences in noise levels between \( \text{PHO5}^+ \) and \( \text{PHO5}^\text{scr} \) strains. Briefly, a special gating procedure is applied to select for a homogeneous population of unbudded G1 cells (Figure S6A). Consequently, sources of extrinsic noise are minimized and the resulting coefficient of variation (CV % = (SD/mean) \times 100) values are composed of intrinsic noise and extrinsic noise specific to the pathway that regulates that gene’s expression (Newman et al., 2006; Raser and O’Shea, 2005). The limited sensitivity of flow cytometry meant that we could only measure noise levels for 66 genes, including 14 antisense-regulated ones (see Figure S6B for a comparison with microscopy). As previously reported, we observed an inverse proportional relationship between CV\(^2\) and protein abundance (Figure 6A; Newman et al., 2006; Paulsson, 2004). For the antisense-regulated genes, protein abundances of the \( \text{PHO5}^\text{scr} \) strains are on average smaller than for the \( \text{PHO5}^+ \) strains (Figure S6B). This also applies for the genes considered to be not regulated by antisense, albeit in a non-significant manner. Because of this slight deviation from zero, a direct comparison of noise between the two categories will be strongly affected by changes in protein abundances (Figure S6C). To obtain noise levels independent of this confounding influence, we adapted the analysis procedure of Newman et al., 2006 (Newman et al., 2006). We calculated the residuals of each CV value (CV\(_\text{res}\)) to a robust regression model of the CV values (Figure 6A). By calculating the difference in the CV\(_\text{res}\) values of \( \text{PHO5}^\text{scr} / \text{PHO5}^+ \) one can compare the noise levels of the two constructs for a given gene and for different categories in general. Interestingly, the distributions of the gene-wise differences in CV\(_\text{res}\) values were significantly smaller for antisense regulated genes than for non-regulated ones (Wilcoxon’s rank sum test, \( p < 0.05 \); Figure 6B) and were mainly negative. We conclude that, provided that a gene is antisense regulated, noise levels are reduced in the presence of antisense transcription.

DISCUSSION

Reports that provide experimental evidence for the implications of antisense regulation on the expressed protein amounts of the overlapping sense gene are scarce, with most studies focusing on RNA levels or on individual genes. Here, we investigated the impact of SUT antisense transcripts on expressed protein amounts for a larger group of genes. We employed yeast
high-throughput strain construction and a strategy to specifically terminate antisense transcription and used this to investigate the function of >150 SUTs in regulating protein levels under four growth conditions. We found that SUTs led to a significant reduction of protein abundance for approximately one quarter of the genes. Therefore, the mere presence of an antisense SUT for a given ORF is no indication for a functional role of this ncRNA. Our results provide experimental confirmation for the previous observations from transcriptomics studies that antisense transcription seems to generally repress sense expression. For example, a certain number of antisense transcripts may be initiated within the ORF of the gene, in which case the PHO5T strategy does not work. This could be due to premature termination or selection of an alternative poly(A) site as a function of the PHO5T element. Equally well, however, this could be caused as a function of the abrogated antisense RNA. Nevertheless, we are confident that the PHO5T element functions as expected in a majority of the cases. First, there was a neutral effect of both the PHO5T and the PHO5T:scr construct on IME4-sfGFP RNA levels in a diploid background, where the RME2 antisense transcript is efficiently repressed. Second, we would expect that general premature termination of transcripts results in decreased stability. The results for IME4, the absence of mRNA degradation products in our northern blots and the fact that protein levels for the PHO5T strains were usually increased indicate that this is not a general concern.

Alternative approaches to study antisense transcription employ mutants in genes that globally affect the stability and transcription of non-coding RNAs. However, such mutations are likely to influence a cell’s RNA homeostasis globally, and hence the predictive value for the assignment of regulatory
functions of antisense transcripts is limited. Interestingly, the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 system has been recently used in yeast to suppress a ncRNA in a strand-specific and position-dependent manner (Lenstra et al., 2015). This offers the advantage of not manipulating the endogenous locus. However, it requires gene-specific optimization in order to identify guide RNAs that are effective in quenching the antisense transcript (Lenstra et al., 2015), which makes it difficult to use this strategy in systematic studies.

Antisense regulation might affect gene noise with or without affecting the average expression level of a gene. Interestingly, we found that regulation of protein amounts by antisense results in a concomitant reduction in protein expression noise. Our fluorescence-activated cell sorting (FACS) measurement could only assess genes that are well expressed. For this subset of antisense-regulated genes, our result is in contrast to the view that antisense transcription increases the noise of genes in the "on" state (Pelechano and Steinmetz, 2013; Xu et al., 2011). It will be interesting to study whether the effect we observed from antisense-regulated genes being intrinsically less noisy or because the pathways involved in antisense-dependent repression act as extrinsic low-level noise sources.

The fact that 13 of 41 genes showed full repression of protein expression under at least one growth condition (within the limits of sfGFP detection; Figure 2D; Table S4) supports the idea that antisense is frequently involved in the switching between "on" and "off" states and serves to suppress the leakiness of gene expression (Castelnovo et al., 2013; Hongay et al., 2006; Lenstra et al., 2015; Xu et al., 2011). However, whether these potential on/off switches are associated with a physiological function remains to be determined for each individual gene.

What distinguishes antisense-regulated genes from non-regulated ones? First, our data indicate that increasing antisense levels led to a stronger regulation by antisense while higher sense levels reduced such an effect. This favors the idea that sense and antisense competitively inhibit each other as studied in detail in a previous report using an artificial gene construct in yeast (Buetti-Dinh et al., 2009). However, in contrast to that study, we could not observe a similar decrease in the repressive effect for weakly expressed genes.

Next, we found that antisense transcription overlapping the TSS is a strong predictor for the functionality of antisense transcription. NET-seq data revealed reads around the TSS even for those genes where tiling array experiments did not detect transcripts in that region. This suggests that the act of antisense transcription rather than the amount of the antisense transcript is of functional importance. This would also imply that the distinction between different types of ncRNAs such as CUTs/XUTs/NUTs/SUTs, while interesting in terms of RNA metabolism, may be less important from a functional point of view. For example, Castelnovo and colleagues presented evidence that led them to hypothesize that the frequency of CUTs escaping early termination and extending into the sense TSS is inversely correlated with sense levels (Castelnovo et al., 2014, 2013). We propose that similar mechanisms are at work for SUTs. In this respect, it is unclear which role the NNS complex plays in regulating protein abundance by antisense SUTs. While we found a slight increase in the number of Nrd1 binding sites in the non-regulated gene set, this increase was not significant. At the same time, early termination events are a mechanism that could help to explain how antisense SUTs are prevented from reaching the sense TSS. The fact that non-regulated genes are longer than antisense regulated ones could further decrease the chance of TSS overlap. We speculate that the impact of antisense may be modulated by a combination of strand-specific termination signals and a relative shift of TSS positions, for example by changing the lengths or initiation sites of transcription units. We also did not find any other motifs or protein binding sites associated with antisense-dependent gene regulation. Considering the importance of where in the gene antisense transcription takes place, it is possible that narrowing motif searches
down to certain positions in the gene will reveal sequence motifs missed in our analysis. When comparing histone modification profiles across the genes of interest we found a significant increase in H3K4me2/3 distribution at the end of antisense-regulated genes. H3K4me3 is a modification associated with actively transcribed genes (Pokholok et al., 2005). Thus, one explanation might be that transcription of the antisense RNA in the regulated genes is more active in certain regions and thus redistributes this mark into the gene body, as was proposed in a recent study (Murray et al., 2015). H3K4me3 has also been linked to SET1-dependent antisense transcription (Castelnuovo et al., 2014; Margaritis et al., 2012). Our data provide evidence for a widespread functional role of this modification in gene regulation by SUTs. No other histone marks were found to be associated with antisense dependent gene regulation, suggesting that they are not involved in the regulation of protein amounts.

Even genes whose antisense SUTs overlapped the TSS were not always regulated by antisense transcription. This indicates that TSS overlap is either not sufficient to impose antisense-dependent regulation or that certain factors blocked such a repression under the conditions tested. Other possibilities include that the regulated genes were more responsive to certain chromatin modifications under the conditions tested, as exemplified by the genes subject to condition-specific regulation.

In summary, we conclude that the majority of antisense transcripts are unlikely to be effector molecules whose synthesis and presence is involved in regulating the abundance of the sense genes. A smaller fraction of antisense transcripts exhibit weak suppressive and denoising functions that may in some cases lead to a complete shutdown of the sense gene. Given that cases of strong and functional antisense-dependent gene regulation have been observed in yeast, this argues that single antisense transcripts are unlikely to be effector molecules whose synthesis is amplified by the genes subject to condition-specific regulation. This indicates that the regulated genes were more responsive to certain chromatin modifications under the conditions tested, as exemplified by the genes subject to condition-specific regulation.

EXPERIMENTAL PROCEDURES

Yeast Strains and Culturing Conditions

Yeast cells were grown according to standard methods (Sherman, 2002). Cultures were grown to logarithmic phase (optical density 600 [OD600] between 0.5 and 1.0) unless otherwise stated. See Supplemental Experimental Procedures for a list of strains and growth media.

Antisense Library Construction

Details about plasmids and strains are listed in Supplemental Experimental Procedures, Briefly, gene-specific oligos containing S2/S3 annealing sites (Knop et al., 1999) for the template cassettes (pMaM201 with sfGFP-PHOS, pMaM203 with sfGFP-PHOS2, and pMaM175 with sfGFP) were used to generate PCR products using a high-fidelity polymerase. Strain YM4M30 was used for transformation, and for each transformation, six clones were picked, single out, and validated using colony PCR.

RNA Extractions and Northern Blots

For all methods, total RNA was extracted using a hot phenol protocol (Collart and Olivier, 2001). Remaining DNA was removed using the TURBO DNA-free Kit (Life Technologies). Northern blotting is explained in Supplemental Experimental Procedures.

Tiling Arrays

The background strain of the antisense library (YMaM330) was grown to mid-log phase in one of four growth media (YPAD, YPGal, YPE, or SC), and total RNA was extracted and hybridized to tiling arrays as described previously (Kuret et al., 2009). The dataset for the entire genome is available in a searchable web database (http://steinmetzlab.embl.de/cgi-bin/viewKnopLabArray.pl?showSamples=KnopHaploidTest2&type=heatmap).

Fluorescence Microscopy

Details about imaging and image processing can be found in Supplemental Experimental Procedures. Briefly, sample cells were mixed 1:1 with sfGFP-negative control cells for background subtraction and normalization on a well-by-well basis. Exponentially growing cells were fixed and seeded on 384-well microscopy plates. Imaging was performed on a Nikon Ti-E screening wide-field epifluorescence microscope using different exposure times for sfGFP and including controls to correct for shading artifacts. Image post-processing, quantification, and quality control were performed using custom scripts in ImageJ, MATLAB (MathWorks), and R (R Core Team, 2015). The whole single-cell dataset can be downloaded from the University of Heidelberg heIDATADataverse Network (http://dx.doi.org/10.11588/data/10073). Scripts and raw imaging data are available upon request.

Bioinformatic and Statistical Analyses

Data analyses were mainly performed using the open source software R (R Core Team, 2015), making extensive use of the Bioconductor framework (Huber et al., 2015) and numerous publicly available datasets. Please see Supplemental Experimental Procedures for further details. Scripts are available upon request.

Flow Cytometry

Cells were grown to mid-log phase in SC medium in 96-well plates and analyzed using a flow cytometer equipped with a high throughput stage (BD FACSScanto RUO HTS). 100,000 events were recorded per well. Gating and calculations to obtain noise estimates were done in R with the help of the flowCore package (Ellis et al., 2015), essentially using the method of Newman et al. (2006). See Supplemental Experimental Procedures for further details. Scripts and data files can be downloaded from the University of Heidelberg heIDATADataverse Network (http://dx.doi.org/10.11588/data/10073).

Western Blots

Proteins were extracted from mid log phase cell cultures grown in the respective medium using the trichloroacetic acid (TCA) method (Knop et al., 1999). Proteins were then separated by SDS-PAGE as described (Laemmli, 1970) and transferred to nitrocellulose membranes using a semi-wet blotter (XCell II Blot Module; Invitrogen). Membranes were incubated overnight with primary anti-GFP antibodies (Abcam) or anti-PGK1 antibodies (Molecular Probes). Secondary antibodies were labeled with Alexa680 (Invitrogen) or IRDye800 (Rockland Immunochemicals). Detection and quantification was performed with an Odyssey Infrared Imaging System (Li-Cor Biosciences).

qRT-PCRs

DNase-treated RNA from exponentially growing yeast cells was used as an input for reverse transcription using 2 pmol gene-specific primer and 1 μg RNA using SuperScript III reverse transcriptase (Invitrogen) according to the manufacturer’s instructions but supplemented with 20 μg/ml actinomycin D to ensure strand specificity of the reverse transcription (Perocchi et al., 2007). For qPCR, cDNA samples and -RT controls were diluted to 1 ml, and 2.5 μl were amplified using the LightCycler 480 SYBR Green I Master Mix (Roche). Actin mRNA was used as the reference gene.

ACCESSION NUMBERS

The accession number for the tiling array data reported in this paper is ArrayExpress: E-MTAB-4731.
SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.05.043.

AUTHOR CONTRIBUTIONS

M.K. conceived the project with input from A.K. and L.M.S.; A.K., M.M., and D.B. constructed the library; F.H. and D.B. designed and analyzed the experiments with input from I.G.; D.B. and F.H. performed the microscopy with technical assistance from P.T.; I.G. performed and analyzed the tiling array experiments; F.H. performed bioinformatics and statistical analyses with input from I.G.; all authors commented on the manuscript; and F.H., D.B., and M.K. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Julien Gagneur, Zhenyu Xu, and Wu Wei for their help selecting genes for antisense library construction; Gaël Yvert, Florent Chuffart, Bernd Fischer, and the anonymous reviewers for helpful advice or suggestions; and Allan Jones and Martin Steffl for comments on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (DFG-KN498/8-1 to M.K. and DFG-STE1422/3-1 to L.M.S.), the Sonderforschungsbereich 1036 (SFB1036, to F.H. and A.K.), a PhD fellowship from the Boehringer Ingelheim Fonds (to M.K.), and a PhD fellowship from the HBIGS graduate school (to F.H.).

Received: February 11, 2016
Revised: April 8, 2016
Accepted: May 10, 2016
Published: June 9, 2016

REFERENCES

Alcid, E.A., and Tsukiyama, T. (2014). ATP-dependent chromatin remodeling shapes the long noncoding RNA landscape. Genes Dev. 28, 2348–2360.
Arigo, J.T., Eyler, D.E., Carroll, K.L., and Corden, J.L. (2006). Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol. Cell 23, 841–851.
Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37, W202–W208.
Bertone, P., Stocl, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., et al. (2004). Global identification of human transcribed sequences with genome tiling arrays. Science 306, 2242–2246.
Buetti-Dinh, A., Ungricht, R., Kelemen, J.Z., Shetty, C., Ratna, P., and Becskei, A. (2009). Control and signal processing by transcriptional interference. Mol. Syst. Biol. 5, 300.
Camblong, J., Iglesias, N., Fickentscher, C., Diepoeis, G., and Stutz, F. (2007). Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. Cell 131, 706–717.
Camblong, J., Beyrouthy, N., Guffanti, E., Schlaepfer, G., Steinmetz, L.M., and Stutz, F. (2009). Trans-acting antisense RNAs mediate transcriptional gene cosuppression in S. cerevisiae. Genes Dev. 23, 1534–1545.
Castelnovo, 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.
Castelnovo, M., Zaug, J.B., Guffanti, E., Maffioletti, A., Camblong, J., Xu, Z., Clader-Munster, 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.
Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 469, 368–373.
Collart, M.A., and Olivier, S. (2001). Preparation of yeast RNA. In Current Protocols in Molecular Biology (John Wiley & Sons).
Creamer, T.J., Darby, M.M., Jamonnak, N., Schaugheny, P., Hao, H., Wheelan, S.J., and Corden, J.L. (2011). Transcriptionome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. PLoS Genet. 7, e1002329.
Csárdi, G., Franks, A., Choi, D.S., Airoldi, E.M., and Drummond, D.A. (2015). Accounting for experimental noise reveals that mRNA levels, amplified by post-transcriptional processes, largely determine steady-state protein levels in yeast. PLoS Genet. 11, e1005206.
David, L., Huber, W., Granovskiaia, 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.
Davis, C.A., and Ares, M. Jr. (2006). Translation of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol. Cell 23, 841–851.
Diniz, G., Brossard, A., Gunnarsson, A., Barry, J.D., Meurer, M., et al. (2014). Seamless gene tagging by endonuclease-driven homologous recombination. PLoS ONE 6, e23794.
Diniz, G., Santos-Rosa, H., Penkett, C.J., Singer, M.A., Vermeulen, M., Mann, M., Bähler, J., Green, R.D., and Kouzarides, T. (2007). Arginine methylation at histone H3K4 controls deposition of H3K4 trimethylation. Nature 449, 928–932.
Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K., and Schiebel, E. (1999). Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. Yeast 15 (109), 963–972.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.

Lenstra, T.L., Coulon, A., Chow, C.C., and Larson, D.R. (2015). Single-molecule imaging reveals a switch between spurious and functional ncRNA transcription. Mol. Cell 60, 597–610.

Margaritis, T., Oreal, V., Brabers, N., Maestroni, L., Vitaliano-Prunier, A., Benschop, J.J., van Hoof, S., van Leenen, D., Dargemont, C., Géli, V., and Holstege, F.C.P. (2012). Two distinct repressive mechanisms for histone 3 lysine 4 methylation through promoting 3′-end antisense transcription. PLoS Genet. 8, e1002952.

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.

Murray, S.C., Haenni, S., Howe, F.S., Fischl, H., Chocian, K., Nair, A., and Melior, J. (2015). Sense and antisense transcription are associated with distinct chromatin architectures across genes. Nucleic Acids Res. 43, 7823–7837.

Nadal-Ribelles, M., Solé, C., Xu, Z., Steinmetz, L.M., de Nadal, E., and Posas, F. (2014). Control of Cdc28 CDK1 by a stress-induced lncRNA. Mol. Cell 53, 549–561.

Nagalakshmi, U., Wang, Z., Warren, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320, 134–139.

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.

Newman, J.R.S., Ghemannaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature 441, 840–846.

Nguyen, T., Fischl, H., Howe, F.S., Woloszczuk, R., Serra Barros, A., Xu, Z., Brown, D., Murray, S.C., Haenni, S., Halstead, J.M., et al. (2014). Transcription mediated insulation and interference direct gene cluster expression switches. eLife 3, e03635.

Paulsson, J. (2004). Summing up the noise in gene networks. Nature 427, 415–418.

Pédelacq, J.-D., Cabantous, S., Tran, T., Tervilliger, T.C., and Waldo, G.S. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. 24, 79–88.

Pelechano, V., and Steinmetz, L.M. (2013). Gene regulation by antisense transcription. Nat. Rev. Genet. 14, 880–893.

Perocchi, F., Xu, Z., Claudio-Münster, S., and Steinmetz, L.M. (2007). Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. Nucleic Acids Res. 35, e128.

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolzheimer, E., et al. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122, 517–527.

R Core Team (2015). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

Raser, J.M., and O’Shea, E.K. (2005). Noise in gene expression: origins, consequences, and control. Science 309, 2010–2013.

Schulz, D., Schwalb, B., Kiese, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J., and Cramer, P. (2013). Transcripion surveillance by selective termination of noncoding RNA synthesis. Cell 155, 1075–1087.

Shannon, P., 2015. MotifDb: an annotated collection of protein-DNA binding sequence motifs. R package version 1.14.0.

Sherman, F. (2002). Getting started with yeast. Methods Enzymol. 350, 3–41.

Uhler, J.P., Hertel, C., and Svejstrup, J.Q. (2007). A role for noncoding transcription in activation of the yeast PHO5 gene. Proc. Natl. Acad. Sci. USA 104, 8011–8016.

Uprety, B., Kaja, A., Ferronius, J., Sen, R., and Bhaumik, S.R. (2016). Regulation of antisense transcription by NuA4 histone acetyltransferase and other chromatin regulatory factors. Mol. Cell. Biol. 36, 992–1006.

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

van Werven, F.J., Neuert, G., Hendrick, N., Lardenois, A., Buratowski, S., van Oudenaarden, A., Primig, M., and Amon, A. (2012). Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. Cell 150, 1170–1181.

Venters, B.J., Wachi, S., Mavrich, T.N., Andersen, B.E., Jena, P., Sinnenon, A.J., Jain, P., Roller, N.S., Jiang, C., Hemeryck-Walsh, C., and Pugh, B.F. (2011). A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces. Mol. Cell 41, 480–492.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.-C., Dufour, M.-E., Boulay, J., Pégault, 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., Claudio-Münster, S., Cambiong, 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., Claudio-Münster, S., Smolik, M., Huber, W., and Steinmetz, L.M. (2011). Antisense expression increases gene expression variability and locus interdependency. Mol. Syst. Biol. 7, 488.