Published in final edited form as:

Cell Rep. 2017 June 13; 19(11): 2371–2382. doi:10.1016/j.celrep.2017.05.057.

H3K36 Methylation Regulates Nutrient Stress Response in *Saccharomyces cerevisiae* by Enforcing Transcriptional Fidelity

Stephen L. McDaniel[1,2], **Austin Hepperla**[1], **Jie Huang**[1,3], **Raghuvar Dronamraju**[2], **Alexander T. Adams**[2], **Vidyadhar G. Kulkarni**[3], **Ian J. Davis**[1,4,5,*], and **Brian D. Strahl**[1,2,4,*]

[1]Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, NC 27599

[2]Department of Biochemistry, University of North Carolina, Chapel Hill, NC 27599

[3]Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC 27599

[4]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599

[5]Departments of Pediatrics and Genetics, University of North Carolina, Chapel Hill, NC 27599

SUMMARY

Set2-mediated histone methylation at H3K36 regulates diverse activities including DNA repair, mRNA splicing and suppression of inappropriate (cryptic) transcription. Although failure of Set2 to suppress cryptic transcription has been linked to decreased life span, the extent to which cryptic transcription influences other cellular functions is poorly understood. Here, we uncover a role for H3K36 methylation in the regulation of the nutrient stress response pathway. We found the transcriptional response to nutrient stress was dysregulated in *SET2*-deleted (*set2*Δ) cells and was correlated with genome-wide bi-directional cryptic transcription that originated from within gene bodies. Antisense transcripts arising from these cryptic events extended into the promoters of the genes from which they arose and were associated with decreased sense transcription under nutrient stress conditions. These results suggest that Set2-enforced transcriptional fidelity is critical to the proper regulation of inducible and highly regulated transcription programs.

IN BRIEF

McDaniel et al. find that Set2-mediated H3K36 methylation is necessary for proper TOR signaling and the nutrient stress response. Cells lacking Set2 display a disrupted transcriptional response to

*Corresponding Authors: brian_strahl@med.unc.edu (lead contact) and ian_davis@med.unc.edu.

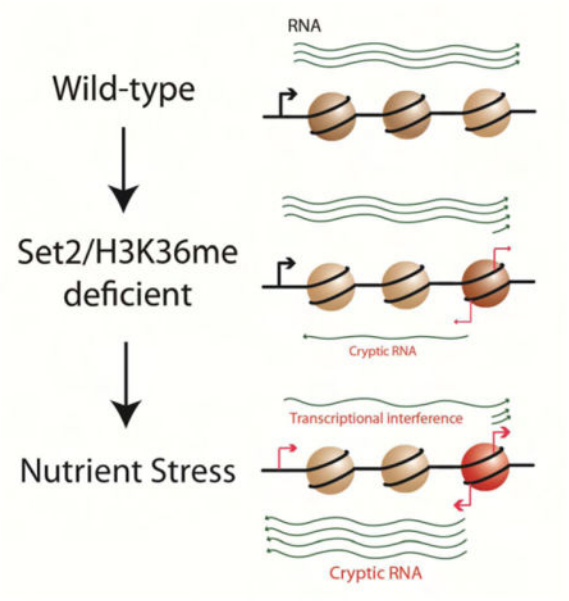
ACCESSION NUMBERS

The GEO accession number for the RNA-seq datasets reported in this paper is GSE89265.

AUTHOR CONTRIBUTIONS

S.L.M, A.H., J.H., V.G.K., R.D. A. A., I.J.D. and B.D.S. designed research and analyzed the data. S.L.M, A.H., R. D., A. A. and J.H. conducted the experiments. S.L.M, A.H., I.J.D. and B.D.S. wrote the manuscript with comments from all authors.

nutrient stress that is correlated with increased intragenic bi-directional transcription within genes that interferes with proper transcription.



Keywords

Histone; H3K36 methylation; Set2; Transcriptional Regulation; Cryptic Transcription; Nutrient Stress

INTRODUCTION

A vast array of post-translational modifications occur on histones and provide distinct binding sites for a wide variety of effector proteins to interact with the genome and direct essential DNA-based functions such as gene expression (Rothbart and Strahl, 2014; Strahl and Allis, 2000). Set2 is a highly conserved histone methyltransferase that methylates histone H3 at lysine 36 (H3K36) (McDaniel and Strahl, 2017; Strahl et al., 2002; Venkatesh and Workman, 2015). In contrast to higher eukaryotes, Set2 is the sole H3K36 methyltransferase in *Saccharomyces cerevisiae,* and is responsible for modifying H3K36 with up to three methyl groups, creating mono-, di-, and tri-methylated H3K36. Significantly, evidence from yeast and human cells suggest that each H3K36 methyl state recruits distinct effector proteins (Carrozza et al., 2005; Gilbert et al., 2014; Keogh et al., 2005; McDaniel and Strahl, 2017; Smolle et al., 2012; Vermeulen et al., 2010), thereby increasing the signaling capacity of this highly conserved histone residue.

Major functions of H3K36me in yeast are to repress bi-directional transcription at promoters and to suppress the aberrant localization of RNA polymerase II (RNAPII) in gene bodies, as the failure of either function causes production of “cryptic” transcripts (Carrozza et al., 2005; Churchman and Weissman, 2011; Keogh et al., 2005; Xu et al., 2009). In large part, H3K36me accomplishes these activities by recruiting the ISW1b chromatin-remodeling

complex, via binding of the PWWP domain of Ioc4 to H3K36me3 (Maltby et al., 2012; Smolle et al., 2012), and recruiting the Rpd3S histone deacetylase complex, via binding of the chromodomain of Eaf3 to H3K36me2/3 (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). Together, these complexes remodel chromatin structure in the wake of elongating RNAPII, thereby preventing RNAPII from aberrantly binding to intragenic regions that can function as sites of transcription initiation.

Recent evidence shows that antisense transcripts that arise from wild-type (WT) cells can create transcriptional interference and decrease normal transcript levels from the canonical 5′ gene promoter, particularly if the antisense transcripts overlap with the native promoter for that gene (Huber et al., 2016). Despite the potential for such deleterious effects, it is surprising that a loss of Set2, which leads to the increased production of cryptic transcripts across the genome, does not result in more severe growth defects or changes to the transcriptome (Lenstra et al., 2011); note, however, that the narrow alteration to the transcriptome in *SET2* delete (*set2*Δ) cells was identified only in asynchronously growing cells and not by RNA-seq. Regardless, it is clear that cryptic transcripts impact cellular physiology, as the loss of Set2 or H3K36me leads to loss of transcriptional fidelity, defective dynamics of mRNA/ncRNA transcription and decreased lifespan in yeast and worms (Kim et al., 2016; Sen et al., 2015).

In this report, we examine the importance of preventing cryptic transcription by Set2. Given the mild phenotypes found in *set2*Δ cells, we surmised that the importance of maintaining transcriptional fidelity by Set2 would be more apparent following activation of an inducible and highly-tuned transcription program, such as the nutrient response system. Indeed, we found that *set2*Δ cells were sensitive to inhibitors of Tor1, Tor2 and MAP kinases (Heitman et al., 1991; Reinke et al., 2006). Consistent with a role in nutrient response, we demonstrated that *set2*Δ cells show synthetic genetic interactions with the *TOR1, TOR2,* flocculation, and protein kinase C (PKC) MAP kinase genetic pathways. Critically, TORC1 and PKC signaling were also disrupted in *set2*Δ cells. Finally, in the absence of Set2 and H3K36me, we determined that the transcriptional response to nutrient stress was significantly altered, and we used a genome-wide stranded RNA-seq analysis to characterize the cryptic transcription profile of *set2*Δ cells. Surprisingly, we found a widespread and time-dependent increase in bi-directional cryptic transcription within the bodies of genes. The antisense transcripts arising from these new promoters traversed the promoters from which sense transcripts arise, and antisense transcripts correlated with decreased transcription from those promoters, implying transcriptional interference. Collectively, our work reveals that Set2 maintains transcriptional fidelity during the execution of precise and rapidly changing transcriptional programs.

RESULTS

Cells Lacking Set2 or H3K36 Methylation are Sensitive to Caffeine and Rapamycin

H3K36 methylation plays a critical role in suppressing cryptic transcription. Yet in asynchronously growing cells, the deletion of *SET2* results in surprisingly few growth defects and changes to the overall transcriptome (Lenstra et al., 2011). Thus, we reasoned that Set2-enforced transcriptional fidelity might be more important to the regulation of rapid

and highly tuned transcription programs such as that of the nutrient stress response pathway. To test this hypothesis, we exposed *set2*Δ cells to caffeine (which inhibits Tor1, Tor2 and MAP kinases) and rapamycin (a potent and specific inhibitor of the Tor1 complex (TORC1)). As shown in Figure 1A, *set2*Δ cells from three distinct genetic backgrounds showed a slow growth phenotype in the presence of caffeine and rapamycin, indicating that the stress response pathway was disrupted in the absence of Set2.

Although non-histone substrates for yeast Set2 have not been reported, it was formally possible that the growth defects we observed in the *set2*Δ cells on caffeine and rapamycin were due to lack of modification of non-histone targets of Set2. To examine this possibility, we employed histone mutant strains in which H3K36 was mutated to alanine or arginine (H3K36A and H3K36R), thereby creating non-methylatable forms of H3. As shown in Figure 1B, the H3K36A and H3K36R mutant strains showed a slow growth phenotype when plated on caffeine and rapamycin, similar to the phenotype observed for *set2*Δ. Thus, Set2-mediated H3K36 methylation, not methylation of a nonhistone substrate, is necessary for yeast cells to respond properly to nutrient stress.

H3K36me3, but not H3K36me2, is Dispensable for the Nutrient Stress Response

We next assessed which methyl states of H3K36 were necessary for cells to respond properly to nutrient stress. We employed a variety of Set2 point mutants and truncations/deletions that either affect the association of Set2 with RNAPII (*set2-1-618*), the ability of Set2 to interact with the nucleosome (*set2-*Δ*31-39*) or that impact its catalytic function (*set2-R195C* and *set2-H199L*) – all of which result in limiting degrees of H3K36 methylation *in vivo* (Figure 1D). We confirmed that the phenotypes of *set2*Δ cells were specific to a loss of *SET2*, as a WT *SET2* expression construct was able to rescue the *set2*Δ growth defect on caffeine and rapamycin (Figure 1C). Interestingly, mutants that lacked H3K36me3, but not H3K36me1 or H3K36me2, rescued the phenotype (*set2-*Δ*31-39* and *set2-R195C*). In contrast, the catalytically-dead *SET2* mutant (*set2-H199L*) (Sorenson et al., 2016) and a construct with a truncation of the RNAPII interaction domain (*set2-1-618*) (Kizer et al., 2005) failed to rescue the phenotype of *set2*Δ cells. These results highlight an important role for H3K36me1/me2, but not H3K36me3, in mediating proper nutrient response.

Multiple H3K36me Effector Complexes are Necessary to Properly Respond to Nutrient Stress

Given that H3K36 methylation is critical for responding to caffeine and rapamycin stress, we next asked which of the known H3K36me-specifc effector proteins are required for the nutrient stress response. We first examined the association of caffeine sensitivity with two key members of the Rpd3S histone deacetylase complex, *EAF3* or *RCO1*, – singly or in combination with *SET2*. Unexpectedly, when either gene was deleted alone, neither *eaf3*Δ nor *rco1*Δ cells showed significant sensitivity to caffeine, though *rco1*Δ cells do exhibit a subtle slow growth phenotype (Supplemental Figure 1A). However, both *eaf3*Δ and *rco1*Δ double deletions with *set2*Δ nearly phenocopy the deletion of *SET2*, displaying slightly more robust growth than a deletion of *SET2* alone (Supplemental Figure 1A). These data indicated that the Rpd3S complex alone does not mediate the cellular response to nutrient stress.

We next expanded our search beyond Rpd3S to determine what other effector proteins are required for the nutrient stress response. Ioc4 of the Isw1b chromatin-remodeling complex (Smolle et al., 2012), and Pdp3, a member of the NuA3b histone acetyltransferase complex (Gilbert et al., 2014) both bind to H3K36me3 via a PWWP domain. Chd1 is a chromatin remodeler that does not directly associate with H3K36me, but functions in a parallel genetic pathway with Set2/H3K36me (Biswas et al., 2007; Park et al., 2014). Similar to the deletions of *EAF3* and *RCO1*, single deletions of *PDP3, IOC4,* and *CHD1* did not result in sensitivity to caffeine (Supplemental Figure 1B). The *rco1*Δ *ioc4*Δ double mutant strain showed a subtle sensitivity to caffeine, but it did not fully recapitulate the phenotype observed for *set2*Δ cells. However, the *rco1*Δ *ioc4*Δ *pdp3*Δ triple mutant strain, encompassing all three major effector protein complexes, did fully phenocopy a deletion of *SET2* (Supplemental Figure 1B). Interestingly, the *rco1*Δ *ioc4*Δ *chd1*Δ mutant showed a more severe phenotype than the *rco1*Δ *ioc4*Δ *pdp3*Δ strain (Supplemental Figure 1B), supporting the notion that *CHD1* functions in a parallel genetic pathway from *SET2*. Together, these data suggest that the combined actions of Rpd3S, Isw1b and NuA3b are required for nutrient stress response.

Nutrient and PKC Signaling Pathways Genetically Interact with *SET2*

Having established that Set2/H3K36me play a role during the caffeine and rapamycin stress response, we employed a genetic suppressor screen to define the genes and pathways that function with Set2 and H3K36 methylation in the nutrient response. We transformed *set2*Δ cells with a 2μ plasmid overexpression library of genomic fragments (Carlson and Botstein, 1982) and exposed the cells to a concentration of caffeine (20 μM) at which *set2*Δ cells are normally inviable. Genes identified by this screen fell into one of three distinct pathways: the nutrient signaling pathway (Pho85 and Bmh1), the flocculation MAPK pathway (Sfl1), and the PKC signaling pathway (Lre1, Kkk1 and Slt2) (Figure 2A). Retransformation with the isolated plasmids successfully rescued viability (Figure 2B). To further validate the results of the suppression screen, we deleted each of these genes alone or together with *SET2* and analyzed their genetic interactions on caffeine-containing plates. Genes recovered from all three major pathways in the screen displayed negative genetic interactions with *SET2*, thus further confirming that these pathways genetically interact with *SET2* (Figure 2C).

SET2 Genetically Interacts with *TOR1* and *TOR2*

Although components of the *TOR1* and *TOR2* pathways demonstrated genetic interactions with *SET2* in the suppressor screen above, neither kinase was identified in the screen (likely because the screen was not saturating). To specifically test whether *TOR1* and *TOR2* genetically interact with *SET2*, we deleted *SET2* together with either a deletion of *TOR1* or a temperature sensitive mutant of *TOR2, tor2-1,* (*TOR2* is essential in budding yeast) and plated these strains on caffeine and rapamycin. Consistent with our synthetic genetic interactions in Figure 2C, deletion of both *tor1*Δ and *set2*Δ produced a synthetic growth defect in the presence of either caffeine or rapamycin (Figure 3A). This growth defect was also observed with the *tor2-1 set2*Δ double mutant strain (Figure 3B). Notably, lower concentrations of caffeine and rapamycin were used for the analyses of *tor1* and *tor2* alleles, because these strains are extremely sensitive to rapamycin and caffeine; however, at the

concentrations used, *set2*Δ cells were below the threshold for observing a growth defect using either drug.

We next asked if we could rescue the sensitivity of *set2*Δ cells to caffeine and rapamycin by expressing a hyperactive allele of *TOR1, TOR1*^{L2134M} (Takahara and Maeda, 2012). When expressed alone, this allele shows a subtle resistance to both caffeine and rapamycin (Figure 3C and see (Takahara and Maeda, 2012)). Consistent with a role for Set2/H3K36me in the nutrient response pathway, we found that the *TOR1*^{L2134M} allele was able to partially rescue the sensitivity of *set2*Δ cells to rapamycin, but not caffeine. This is consistent with rapamycin and caffeine having different modes of inhibition on the TOR1C (Reinke et al., 2006).

Given the differential sensitivities of the *SET2* deletion compared to the *TOR1* and *TOR2* mutants, we next further verified that the sensitivity of *set2*Δ cells to rapamycin was occurring through TORC1 and not through an off-target effect of the drug. We combined a *SET2* deletion with a deletion of *FPR1*, the key mediator of rapamycin inhibition that binds the drug and inhibits TORC1 kinase activity (Heitman et al., 1991). When plated on rapamycin, *set2*Δ *fpr1*Δ cells were not rapamycin sensitive, indicating that the decreased growth rate of *set2*Δ cells was due to inhibition of the TORC1 kinase (Supplemental Figure 2).

Tor1 Activity is Aberrant in *set2*Δ Cells

Given that *SET2* genetically interacts with both *TOR1* and *TOR2*, we hypothesized that *set2*Δ cells would show significant disruptions in TORC1 and nutrient response signaling. We shifted WT and *set2*Δ cells from nutrient rich (YPD) medium to medium lacking amino acids (SD) and collected samples every 30–60 minutes over a 2-hour time period. We examined the protein and phosphorylation levels of Rps6, the *Saccharomyces cerevisiae* S6K homolog (Gonzalez et al., 2015), a key downstream target of TORC1. As shown in Figure 3D, the level of Rps6 was much lower in *set2*Δ cells compared with their isogenic WT counterparts, and, further, Rps6 had a drastically increased level of phosphorylation at the start of the time course, continuing through 30 minutes after nutrient stress. The increased phosphorylation of Rps6 in *set2*Δ cells was consistent with aberrant TORC1 kinase signaling. These results show that Set2 is necessary for proper nutrient response signaling, which is consistent with our genetic analyses described above.

Set2 Loss Disrupts the Transcriptional Response to Nutrient Stress

Having established the importance of Set2 for a proper nutrient stress response, we analyzed transcriptional changes that accompany nutrient depletion in the presence and absence of Set2. Stranded RNA-seq was performed at four time points following nutrient depletion (0, 30, 60, and 120 minutes). We identified 1,449 differentially expressed genes in WT cells over this time course (702 upregulated and 747 downregulated; Figure 4A, left panel; Supplemental Table 3). Suggestive of the rapidity of nutrient sensing and its transcriptional response, the majority of gene expression changes (891 genes, 61.5%) occurred between 0 and 30 minutes. We then identified four clusters of differentially regulated genes based on variation in expression: genes that appeared to decrease in expression throughout the time

course, genes that achieved an expression nadir early and then recovered slightly, genes that decreased expression early and remained underexpressed, and genes that increased after 0 minutes (top to bottom, respectively). We used gene ontology analysis to characterize the genes that compose each group (Figure 4B) and found that downregulated genes were associated with cell cycle regulation and rRNA/RNA processing. Indicative of a metabolic response to nutrient deprivation, upregulated genes were strongly associated with oxoacid and amine metabolism.

We then examined how Set2 loss affects the transcriptional response to nutrient depletion (Figure 4A, right panel). Examining the 1,499 nutrient stress-responsive genes in WT cells, we noted that the nutrient stress response appeared to have increased variance and was less defined. To explore this observation further, we identified the fraction of genes within each of the four regulatory clusters that no longer exhibited the expression characteristic of each group in *set2*Δ cells (black lines; Figure 4A, right panel). The group associated with the cell cycle had the highest fraction of genes no longer identified as differentially expressed in *set2*Δ cells (70.2%). The rRNA/RNA processing group showed the lowest fraction of non-responsive genes in *set2*Δ cells (38.3%).

We then identified genes that demonstrated a transcriptional change associated with nutrient deprivation in the *set2*Δ cells, independent of our previous analyses in WT cells. Fewer genes were differentially regulated in the absence of Set2 relative to WT cells (943 total differentially expressed genes, with 523 upregulated and 420 downregulated genes) (Figure 4C, Supplemental Table 3). Overall, genes in the up- and downregulated classes identified in the *set2*Δ cells were associated with similar gene ontologies when compared to those identified in WT cells. However, in *set2Δ* cells, an enrichment for cell cycle-associated genes was no longer observed. Of the 332 genes in the cell cycle enriched group in WT cells, only 29.8% demonstrated regulation in the absence of Set2 (Figure 4A, right panel). Although there were differences in gene expression between WT and *set2Δ* cells after nutrient depletion, random permutation among all expressed genes showed that a significant fraction of both up- and downregulated nutrient responsive genes were shared by WT and *set2*Δ cells ($p < 0.001$; Figure 4D). Taken together, these analyses suggests that Set2 mediates a subset of the transcriptional responses to nutrient stress, for example cell cycle genes, whereas, regulation of rRNA/RNA processing genes is largely not dependent on Set2.

The above analyses reported on the overall gene responses after nutrient starvation; however, the analyses did not examine the precise dynamics of these responses. To specifically explore relative levels and temporal shifts in gene expression in the absence of Set2, we scaled the median expression of each time point between 0 and 1 for each differentially regulated gene in WT and *set2*Δ cells following nutrient starvation. For each gene, we then subtracted the scaled expression value of *set2*Δ from WT, meaning a gene with the same relative expression at a given time point in WT and *set2*Δ cells would have a score of zero (Figure 4E). Strikingly, despite being identified as similarly differentially expressed in both WT and *set2*Δ cells, the patterns and relative levels of expression of individual genes varied greatly, with the largest individual class demonstrating increased expression in the absence of Set2. Taken together, these data indicate that Set2 loss disrupts the transcriptional

response after nutrient depletion, affecting a wide range of genes by suppressing induction and altering the timing and magnitude of expression.

Nutrient Depletion Promotes Widespread Cryptic Transcription in the Absence of Set2

We next assessed the relationship between Set2 deficiency and nutrient depletion on aberrant transcription. The stranded RNA-seq data from 120 minutes post nutrient depletion revealed a subset of genes exhibiting an obvious, focal, intragenic signal increase exclusively in *set2*Δ, resulting in relatively higher 3′ signals (Figure 5A). Signal increases at the 3′ relative to the 5′ end were validated by qPCR for two genes and revealed a significant increase associated with time following nutrient depletion (Figure 5B). This pattern seemed suggestive of sense strand cryptic transcription, such that the location of signal change reflected cryptic initiation sites (CISs).

We then developed an approach to systematically identify all genes that could harbor a CIS. Because of the local signal variance observed in RNA-seq data, we modeled the signal as a queue-length process in an M(t)/G/∞ queue (Eick et al., 1993) based on a 50 bp read length and on the assumption that the starting positions for each read constitutes a non-homogeneous Poisson process (see Supplemental Methods and Supplemental Figure 3). Briefly, we scored every base along genes for the probability of increased signal in the RNA from *set2*Δ cells that continued to the end of the gene, when compared with WT (mean square root loss difference, MSRLD).

Based on our hypothesis that cryptic initiation would necessitate alteration of a gene body nucleosome, we selectively analyzed genes that were greater than 700 bp (thus avoiding 2- and 3-nucleosome genes). Also, to avoid ambiguity in read mapping, we filtered genes with introns, overlapping genes, and genes on chromosome M, the mitochondrial chromosome (Supplementary Table 4). Finally, we required that genes were expressed (average normalized RNA-seq signal > 5) and demonstrated signal throughout the gene. We applied our model to identify genes containing this focal signal increase in *set2*Δ cells (compared with WT) at 30, 60, and 120 minutes following nutrient starvation. The mean square root loss difference (MSRLD) was calculated by our model for each gene, inferring an estimated magnitude of a CIS (high 4+, intermediate 2–4, low 0–2).

From approximately 3,400 genes that passed our filtering criteria, we identified 121 high scoring and 318 intermediate scoring genes following nutrient depletion (Figure 6A, Supplementary Table 5). High and intermediate CIS scores are likely to represent true cryptic transcription. The much larger group with a low CIS score may contain limited cryptic transcription, and as such, was used as a control for further analyses. Overall, the number of genes demonstrating a CIS increased during the course of nutrient deprivation. CISs occurred most often in genes greater than 1700 bp, consistent with the model that long genes would be more susceptible to CISs (Supplementary Figure 4A) (Li et al., 2007). CISs occurred primarily in the latter half of the genes (66.5% of high and intermediate genes with CIS), typically at positions within 40%–80% of the gene lengths (Supplementary Figure 4B). Also, genes with a CIS showed a similar expression distribution compared with all expressed genes, indicating that expression levels were not artificially influencing the model (Supplementary Figure 4C).

Cryptic Transcription is Associated with Antisense Transcription and Decreased RNA Abundance

The identification of CISs enabled us to evaluate transcription around these sites by comparing RNA-seq signal in high, intermediate, and low scoring genes throughout the time course in WT and *set2*Δ cells (Figure 6B). At sites of CIS detected based on sense transcription, we observed a striking increase in antisense transcription preceding the CISs at both high and intermediate scoring genes. Whereas increases in sense strand signals occurred mostly from 0 to 30 minutes, antisense signals increased throughout the time course.

To further characterize the antisense transcription, we determined how often relative increases in 5′ antisense signal occurred. Among high CIS scoring genes, 92 of the 121 genes (76%) exhibited antisense transcription. We observed that antisense transcription largely persisted to the beginning of the gene, regardless of gene length (Figure 6C). Recent studies indicate that antisense transcription extending into the sense promoter of a gene can reduce sense transcription of the full-length gene (Huber et al., 2016). To test this association in nutrient stress-associated antisense transcription, we compared sense and all high and intermediate scoring genes between WT and *set2*Δ. We focused on the RNA-seq signal prior to the CIS, as this signal should reflect full-length message by avoiding reads originating from sense strand cryptic transcription. Significantly, 91% of high and intermediate magnitude genes displayed both decreased sense signal and increased antisense signal in *set2*Δ relative to WT (Figure 6D, full dataset set in Supplementary Figure 5A). This pattern was not observed in random genes (Supplementary Figure 5B). Together, these data suggest that long antisense transcripts originating from sites of cryptic transcriptional initiation decrease the sense transcription of their host genes, which likely contributes to, at least in part, the differential expression of genes in *set2*Δ cells following nutrient stress.

To explore our findings in relation to other studies that have assessed cryptic transcription, we analyzed a recently published stranded RNA-seq dataset that was used to reveal Set2-repressed antisense transcripts (SRAT) (Venkatesh et al., 2016). Using our approach, we identified 22 sense cryptic transcripts (MSRLD ≥ 2). As these data were not obtained under conditions of nutrient depletion, our analysis further supports that Set2 loss leads to low-level cryptic transcription (Figure 6E, Supplementary Fig 6A). We were reassured to find that genes with evidence of cryptic initiation using our approach matched those genes they showed had SRATs (Supplementary Figure 6B). We then compared MSRLD scores derived from our approach using their data with results from analysis of our CISs after nutrient depletion. We found that genes with a higher MSRLD in the Venkatesh et al. data demonstrated a high fractional overlap with genes that we identified to be associated with cryptic transcripts after nutrient stress. Higher MSRLD in the *set2Δ* data trended with the MSRLD in our data following nutrient deprivation (Figure 6E). In other words, analysis of this independent data set revealed an association with both the number and the degree of cryptic transcription detected in our data. Notably, antisense transcription at genes with cryptic transcription in the Venkatesh et al. data set was similar to those detected in our data in the absence of nutrient depletion (t = 0 min) (Supplementary Figure 6C and 6D). However, relatively few genes showed antisense signal originating at the CIS that continued

to the start of the gene in the Venkatesh et al. data set (Figure 6F). Taken together, these observations indicate that nutrient stress both establishes new cryptic transcripts and enhances cryptic transcription resulting from Set2 loss alone.

DISCUSSION

In this report, we provide compelling evidence that a critical function of Set2 is to enforce the transcriptional fidelity and regulation of genes that are rapidly activated or repressed during nutrient stress. In the absence of Set2/H3K36 methylation, we observed two significant defects in the transcriptome: 1) gene expression changes between WT and *set2*Δ cells, encompassing the suppression of gene induction and altering the timing and magnitude of nutrient response gene expression and 2) widespread bi-directional transcription within gene bodies that associated with down regulation of the genes. These findings agree with our genetic and biochemical analyses that show *set2*Δ cells are sensitive to drugs that impact the Tor1/2 and MAP kinase pathways and that *set2*Δ cells have defects in nutrient signaling. In addition, these results emphasize the physiological importance of suppressing cryptic transcripts by Set2, and, further, the results suggest that Set2-enforced transcriptional fidelity, both in terms of regulating absolute gene expression levels and by repressing intragenic cryptic transcription, may be a central theme to the proper regulation of rapidly induced transcription programs.

Unlike the distinctive transcriptional and phenotypic defects observed during nutrient stress, *set2*Δ cells do not show large transcriptional defects in rich growth medium (Lenstra et al., 2011). One possible reason for the lack of defects under favorable conditions, and why dramatic changes are observed largely under stress conditions, may be due to the functional redundancy existing between many chromatin factors. It is likely that the loss of Set2 and downstream transcriptional fidelity is compensated for by the actions of other chromatin regulatory pathways, which, under non-stressed conditions, can preserve, to a great extent, the proper regulation of genes that maintain normal growth. However, cellular stresses that elicit a rapid transcriptional response may expose a chromatin “Achilles heel” that is not readily apparent in normal growth conditions. Thus, it may be that as chromatin undergoes significant structural changes to deal with a new transcriptional program, the regulatory machinery becomes displaced or reorganized leaving hyperacetylated and/or promoter-like elements in *set2*Δ cells exposed and susceptible to aberrant transcription initiation events.

Our studies employed strand-specific sequencing of ribosomally-depleted RNA to identify both sense and antisense cryptic transcription. By using this approach, however, we had to address several challenges to enable robust and reliable analyses. Notably, RNA-seq signal is dependent on many factors including sequencing and fragmentation bias, and RNA stability, all of which lead to local signal variances. Also, other analyses may identify transcripts *de novo* and compare them to annotated transcripts to identify novel and cryptic transcripts. However, sense cryptic transcripts initiate within previously annotated regions, thus making that method of cryptic transcript discovery unusable in our analyses. These challenges led to the generation of a new model for stranded analyses, which accounted for the above variables.

One of the striking observations made from our RNA analyses is that sense cryptic transcription within gene bodies in *set2*Δ cells is almost always associated with bi-directional transcription. Notably, we found that the majority of these antisense transcripts arise in genes that are not being actively induced, which is largely is achieved 30 minutes after stress. Interestingly, while sense cryptic transcripts appear robustly after 30 minutes of nutrient stress and remain relatively constant throughout the time course, antisense transcripts are not robustly observed until 60 minutes after nutrient stress, and increase further at 120 minutes after stress. Taken together, these results suggest that a broad number of genes (many being unrelated to nutrient response) are predisposed with the potential for cryptic transcription that Set2/H3K36me is suppressing. Without Set2/H3K36me, and under stress, these potential cryptic sites become activated and show a directionality that ultimately becomes lost over time. How these transcription events initiate and how bi-directional transcription is achieved will be interested to determine in future studies.

Mechanistically, cryptic transcripts that overlap the promoters of genes are likely to negatively impact the transcript levels of those genes (Huber et al., 2016). Here, we demonstrate that not only do *set2*Δ cells produce intragenic antisense transcripts, but also the abundance of the antisense transcripts increases over time during nutrient stress. Further, as most of these transcripts, once initiated, extend back to the canonical 5′ promoters of their host genes, conceivably such overlap would impact sense transcription levels where this overlap occurs. Critically, this phenomenon only seems to occur during nutrient stress. Analysis of *set2Δ* data sets from unstressed cells uncovered limited levels of antisense transcription, with very few of these transcripts extending to the promoter (Figure 6F). Further, our data suggest that during cellular stress, once CISs initiate, they impact the levels of sense transcription. One explanation for this transcriptional interference might be the inability of two RNAPII complexes to occupy the same space along the gene at the same time, and/or that there is a dominating polymerase that prevents the productive transcription elongation of the other overlapping polymerase. Further, it is possible, that, as the new CIS promoter becomes increasingly utilized over time, incorporation of H2A.Z and H3K4me will reflect the establishment of more canonical promoter environment within the gene body as opposed to the natural 5′ promoter. Future studies into the mechanism of Set2- dependent transcriptional interference are currently underway.

Finally, the function of Set2 and H3K36me in the nutrient response pathway and in regulating transcriptional fidelity is likely to be highly conserved across eukaryotes. For example, recent studies on SETD2, the human homolog of yeast Set2, reveal that its loss leads to widespread intragenic transcription defects (Grosso et al., 2015; Simon et al., 2014). In addition, TOR inhibitors combined with *SETD2* knockdown is a lethal combination to leukemic cells (Zhu et al., 2014), suggesting that SETD2 and its role in enforcing transcriptional fidelity is likely conserved and important to the growth of cancer cells. As SETD2 is found mutated in a variety of cancers (McDaniel and Strahl, 2017), it will be of significant interest to determine the degree to which transcriptional fidelity regulated by SETD2/H3K36me3 contributes to its role in cancer development.

EXPERIMENTAL PROCEDURES

Yeast Strains and Plasmids

Yeast strains were created using standard methods (Janke et al., 2004) employing PCR-amplified cassettes with ~50 base pairs of homology to the genes of interest (Janke et al., 2004). Yeast strains used in this study are listed in Supplemental Table 1. Plasmids used in this study are listed in Supplemental Table 2.

Spotting Assays

Strains were grown in YPD (1% yeast extract, 2% peptone, and 2% glucose) and diluted to an OD_{600} of 0.5 prior to spotting 5-fold serial dilutions on the indicated plates at 30° C for 2–3 days.

H3K36 Methylation Analyses

Protein was isolated from 5×10^7 cells as previously described (Gilbert et al., 2014). Extracts were loaded onto 15% SDS-PAGE gels and transferred to PVDF. Membranes were incubated overnight with H3 C-term (EpiCypher), H3K36me1 (Abcam 9048), H3K36me2 (Active Motif 39255), H3K36me3 (Abcam 9050), Set2 (in house), or G6PDH (Sigma A9521) antibodies. Membranes were then washed in TBS-Tween (50 mM Tris, 150 mM NaCl, and 0.5% Tween 20), incubated in secondary antibody (Jackson Labs) and then probed with ECL reagent (GE Healthcare).

Genome-wide Suppressor Screen

Three isolates of a high-copy number 2μ library of Sau3AI digested genomic fragments (Carlson and Botstein, 1982) were transformed into 1×10^8 *set2*Δ cells. Transformations were re-suspended in 1 ml of SC-Ura medium and plated onto 20 SC-Ura + 20 mM caffeine plates. Plasmids were recovered from the indicated yeast strains using Qiagen mini-prep columns and sequenced by Sanger sequencing using one of two primers flanking the site of insertion: pBR-1: CACTATCGACTACGCGATCA or pBR-2: CGATGCGTCCGGCGTAGA.

Phospho-Protein Analysis

WT or *set2*Δ cells were grown overnight in YPD and diluted to an OD_{600} of 0.2 and grown to an OD_{600} of ~1.0. Upon reaching log phase, cells were washed twice with water and resuspended in SD media. Ten ODs of cells were collected at each time point and protein was isolated via TCA extraction as previously described (Fillingham et al., 2008). Extracts were then loaded onto 10% SDS-PAGE gels and transferred to PVDF membrane. Membranes were incubated overnight with the following antibodies: Rps6 (Abcam ab40820), ph-S6K (Cell Signaling 2211S), Set2 (in house), and G6PDH (Sigma A9521). Membranes were then washed in TBS-Tween (50 mM Tris, 150 mM NaCl, and 0.5% Tween 20), incubated in secondary antibody (Jackson Labs), and probed with SuperSignal West Fempto Maximum Sensitivity Substrate (ThermoFisher). Two biological replicates were quantified using ImageJ and the changes in ph-Rps6 signal were averaged.

RNA-seq Library Preparation and Sequencing

WT and *set2*Δ cells were grown overnight in YPD and diluted to an OD_{600} of 0.2 and grown to an OD_{600} of ~1.0. Upon reaching log phase, cells were isolated and washed twice with water and resuspended in SD medium (Cheung et al., 2008). 1×10^8 cells were isolated at each time point and RNA was isolated by acid phenol extraction. 10 μg of RNA was treated with DNase (Promega) and purified (RNeasy column, Qiagen). 2.5–5 μg of RNA was then processed using yeast-specific rRNA depletion beads (Illumina). Strand-specific barcoded sequencing libraries were then prepared (TruSeq Stranded Total RNA library preparation kit, Illumina). Libraries were pooled and sequenced across two lanes (Hi-Seq 2500, Illumina).

Sequencing Alignment and Analysis

FASTQ files were filtered for adapters using TagDust (v1.12) (Lassmann et al., 2009), then aligned to the sacCer3 genome using Bowtie (v1.1.2) (Langmead et al., 2009). For the stranded RNA-seq, Bowtie options included -m 1, --seed=123.456, --nomaqround, and --best. Samtools (v1.3.1) (Li et al., 2009) and bedtools (v2.25.0) (Quinlan and Hall, 2010) were used to sort files and filter for properly mapped/paired reads, as well as generate signal-based wiggle files. To remove potential PCR artifacts, duplicate reads were removed from the RNA-seq data. However, only two identical reads were allowed for the RNA-seq. R (v3.1.1) (Team, 2014) was used for downstream analyses and generating line plots. Specifically, the R package ‘Vennerable’ (v3.0) (Haibe-Kains et al., 2013) was used to make weighted Venn diagrams.

Reads per kilobase per million mapped reads (RPKMs) were calculated for each gene in each sample, and R package ‘DESeq2’ (v1.6.3) (Love et al., 2014) was used to identify differentially expressed genes. Genes with a p-value < 0.05, a minimum of $\log_2$(Fold Change) > 1 (as defined by DESeq2), and an average RPKM > 1 were considered differentially expressed. In WT cells, the 0 minute time point had four biological replicates, whereas 30, 60, and 120 minute time points had three biological replicates. In *set2*Δ cells, the 0 minute time point had two replicates, whereas the 30, 60, and 120 minute time points had 3 replicates. Differentially expressed nutrient stress response gene groups were identified as the union set of all differentially expressed genes from all possible pairwise comparisons between time point within WT or *set2*Δ treatments. Heatmaps were generated using Morpheus (Broad Institute). DAVID (v6.7) (Huang da et al., 2009b) (Huang da et al., 2009a) was used to generate gene ontologies.

RNA extraction and Real time PCR

RNA was extracted using acid phenol method from yeast cells that were collected after different time points of nutrient shift as described above. 10 μg of RNA was treated with DNase (Promega) and purified (RNeasy column, Qiagen). 1 μg of total RNA was used to generate cDNA using superscript reverse transcriptase (First strand cDNA synthesis system, Thermo Scientific). cDNA was diluted 1:25 and subjected to quantitative real-time PCR (qRT-PCR) using SYBR green reagent (Biorad) using manufacturer’s instructions. The relative quantities of the transcript were calculated using $\Delta\Delta C_t$ method (Livak et al., 2013). Data presented are the mean and the standard deviations of three independent experiments.

Detecting Sense Cryptic Transcription

Due to local signal variance observed in RNA-seq data, we modeled the signal as a queue-length process in an M(t)/G/∞ queue (Eick et al., 1993) based on a 50 bp read length and on the assumption that the starting positions for each read constitute a non-homogeneous Poisson process. We also assumed that at each gene position, the signal of *set2*Δ genes are a mixture of full length mRNAs and cryptic initiated mRNAs starting at a single position. Based on these assumptions, we determined the CIS position by minimizing mean square root loss (MSRL) between the predicted *set2*Δ signal and observed *set2*Δ signal. The mean square root loss difference (MSRLD) was calculated for each gene as to infer the estimated magnitude of a CIS (high 4+, intermediate 2–4, low 0–2). The same method was applied for all genes across all time course (for complete model description, see Supplemental Methods). The model algorithm can be found at: http://github.com/jiehuang2000/cryp_init_caller.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Michael Hall for the *tor2* temperature sensitive allele, Tatsuya Maeda for the *tor1* mutant allele, Patrick Brennwald for the isolates for the genome-wide overexpression libraries, and Howard Fried for editorial suggestions. S.L.M was supported in part by a grant from the National Institute of General Medical Sciences under award 5T32 GM007092. B.D.S. acknowledges support from NIH grant GM68088 and I.J.D. from NIH grant CA166447, CA198482 and the Corn-Hammond Fund for Pediatric Oncology.

References

Biswas D, Dutta-Biswas R, Stillman DJ. Chd1 and yFACT act in opposition in regulating transcription. Mol Cell Biol. 2007; 27:6279–6287. [PubMed: 17620414]

Carlson M, Botstein D. Two differentially regulated mRNAs with different 5′ ends encode secreted with intracellular forms of yeast invertase. Cell. 1982; 28:145–154. [PubMed: 7039847]

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP, et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell. 2005; 123:581–592. [PubMed: 16286007]

Cheung V, Chua G, Batada NN, Landry CR, Michnick SW, Hughes TR, Winston F. Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. PLoS Biol. 2008; 6:e277. [PubMed: 18998772]

Churchman LS, Weissman JS. Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature. 2011; 469:368–373. [PubMed: 21248844]

Eick SG, Massey WA, Whitt W. The Physics of the M(T)/G/Iota Queue. Operations Research. 1993; 41:731–742.

Fillingham J, Recht J, Silva AC, Suter B, Emili A, Stagljar I, Krogan NJ, Allis CD, Keogh MC, Greenblatt JF. Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. Mol Cell Biol. 2008; 28:4342–4353. [PubMed: 18458063]

Gilbert TM, McDaniel SL, Byrum SD, Cades JA, Dancy BC, Wade H, Tackett AJ, Strahl BD, Taverna SD. A PWWP domain-containing protein targets the NuA3 acetyltransferase complex via histone H3 lysine 36 trimethylation to coordinate transcriptional elongation at coding regions. Mol Cell Proteomics. 2014; 13:2883–2895. [PubMed: 25104842]

Gonzalez A, Shimobayashi M, Eisenberg T, Merle DA, Pendl T, Hall MN, Moustafa T. TORC1 promotes phosphorylation of ribosomal protein S6 via the AGC kinase Ypk3 in Saccharomyces cerevisiae. PLoS One. 2015; 10:e0120250. [PubMed: 25767889]

Grosso AR, Leite AP, Carvalho S, Matos MR, Martins FB, Vitor AC, Desterro JM, Carmo-Fonseca M, de Almeida SF. Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. Elife. 2015:4.

Haibe-Kains B, El-Hachem N, Birkbak NJ, Jin AC, Beck AH, Aerts HJ, Quackenbush J. Inconsistency in large pharmacogenomic studies. Nature. 2013; 504:389–393. [PubMed: 24284626]

Heitman J, Movva NR, Hall MN. Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. Science. 1991; 253:905–909. [PubMed: 1715094]

Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009a; 37:1–13. [PubMed: 19033363]

Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009b; 4:44–57. [PubMed: 19131956]

Huber F, Bunina D, Gupta I, Khmelinskii A, Meurer M, Theer P, Steinmetz LM, Knop M. Protein Abundance Control by Non-coding Antisense Transcription. Cell Rep. 2016; 15:2625–2636. [PubMed: 27292640]

Janke C, Magiera MM, Rathfelder N, Taxis C, Reber S, Maekawa H, Moreno-Borchart A, Doenges G, Schwob E, Schiebel E, et al. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast. 2004; 21:947–962. [PubMed: 15334558]

Joshi AA, Struhl K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell. 2005; 20:971–978. [PubMed: 16364921]

Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, et al. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell. 2005; 123:593–605. [PubMed: 16286008]

Kim JH, Lee BB, Oh YM, Zhu C, Steinmetz LM, Lee Y, Kim WK, Lee SB, Buratowski S, Kim T. Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. Nat Commun. 2016; 7:13534. [PubMed: 27892458]

Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol. 2005; 25:3305–3316. [PubMed: 15798214]

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10:R25. [PubMed: 19261174]

Lassmann T, Hayashizaki Y, Daub CO. TagDust--a program to eliminate artifacts from next generation sequencing data. Bioinformatics. 2009; 25:2839–2840. [PubMed: 19737799]

Lenstra TL, Benschop JJ, Kim T, Schulze JM, Brabers NA, Margaritis T, van de Pasch LA, van Heesch SA, Brok MO, Groot Koerkamp MJ, et al. The specificity and topology of chromatin interaction pathways in yeast. Mol Cell. 2011; 42:536–549. [PubMed: 21596317]

Li B, Gogol M, Carey M, Pattenden SG, Seidel C, Workman JL. Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. Genes Dev. 2007; 21:1422–1430. [PubMed: 17545470]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]

Livak KJ, Wills QF, Tipping AJ, Datta K, Mittal R, Goldson AJ, Sexton DW, Holmes CC. Methods for qPCR gene expression profiling applied to 1440 lymphoblastoid single cells. Methods. 2013; 59:71–79. [PubMed: 23079396]

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. [PubMed: 25516281]

Maltby VE, Martin BJ, Schulze JM, Johnson I, Hentrich T, Sharma A, Kobor MS, Howe L. Histone H3 lysine 36 methylation targets the Isw1b remodeling complex to chromatin. Mol Cell Biol. 2012; 32:3479–3485. [PubMed: 22751925]

McDaniel SL, Strahl BD. Shaping the cellular landscape with Set2/SETD2 methylation. Cell Mol Life Sci. 2017

Park D, Shivram H, Iyer VR. Chd1 co-localizes with early transcription elongation factors independently of H3K36 methylation and releases stalled RNA polymerase II at introns. Epigenetics Chromatin. 2014; 7:32. [PubMed: 25395991]

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–842. [PubMed: 20110278]

Reinke A, Chen JC, Aronova S, Powers T. Caffeine targets TOR complex I and provides evidence for a regulatory link between the FRB and kinase domains of Tor1p. J Biol Chem. 2006; 281:31616–31626. [PubMed: 16923813]

Rothbart SB, Strahl BD. Interpreting the language of histone and DNA modifications. Biochim Biophys Acta. 2014; 1839:627–643. [PubMed: 24631868]

Sen P, Dang W, Donahue G, Dai J, Dorsey J, Cao X, Liu W, Cao K, Perry R, Lee JY, et al. H3K36 methylation promotes longevity by enhancing transcriptional fidelity. Genes Dev. 2015; 29:1362–1376. [PubMed: 26159996]

Simon JM, Hacker KE, Singh D, Brannon AR, Parker JS, Weiser M, Ho TH, Kuan PF, Jonasch E, Furey TS, et al. Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. Genome Res. 2014; 24:241–250. [PubMed: 24158655]

Smolle M, Venkatesh S, Gogol MM, Li H, Zhang Y, Florens L, Washburn MP, Workman JL. Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. Nat Struct Mol Biol. 2012; 19:884–892. [PubMed: 22922743]

Sorenson MR, Jha DK, Ucles SA, Flood DM, Strahl BD, Stevens SW, Kress TL. Histone H3K36 methylation regulates pre-mRNA splicing in Saccharomyces cerevisiae. RNA Biol. 2016; 13:412–426. [PubMed: 26821844]

Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000; 403:41–45. [PubMed: 10638745]

Strahl BD, Grant PA, Briggs SD, Sun ZW, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF, et al. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol Cell Biol. 2002; 22:1298–1306. [PubMed: 11839797]

Takahara T, Maeda T. Transient sequestration of TORC1 into stress granules during heat stress. Mol Cell. 2012; 47:242–252. [PubMed: 22727621]

Team, R.C. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2014.

Venkatesh S, Li H, Gogol MM, Workman JL. Selective suppression of antisense transcription by Set2-mediated H3K36 methylation. Nat Commun. 2016; 7:13610. [PubMed: 27892455]

Venkatesh S, Workman JL. Histone exchange, chromatin structure and the regulation of transcription. Nat Rev Mol Cell Biol. 2015; 16:178–189. [PubMed: 25650798]

Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, et al. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell. 2010; 142:967–980. [PubMed: 20850016]

Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Munster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM. Bidirectional promoters generate pervasive transcription in yeast. Nature. 2009; 457:1033–1037. [PubMed: 19169243]

Zhu X, He F, Zeng H, Ling S, Chen A, Wang Y, Yan X, Wei W, Pang Y, Cheng H, et al. Identification of functional cooperative mutations of SETD2 in human acute leukemia. Nat Genet. 2014; 46:287–293. [PubMed: 24509477]

HIGHLIGHTS

- Set2 is required for the proper function of and genetically interacts with TORC1/2
- Set2/H3K36me loss impairs nutrient stress response signaling and transcription
- Without Set2, nutrient stress results in bi-directional intragenic transcription
- Antisense transcripts arising after nutrient stress cause transcriptional interference

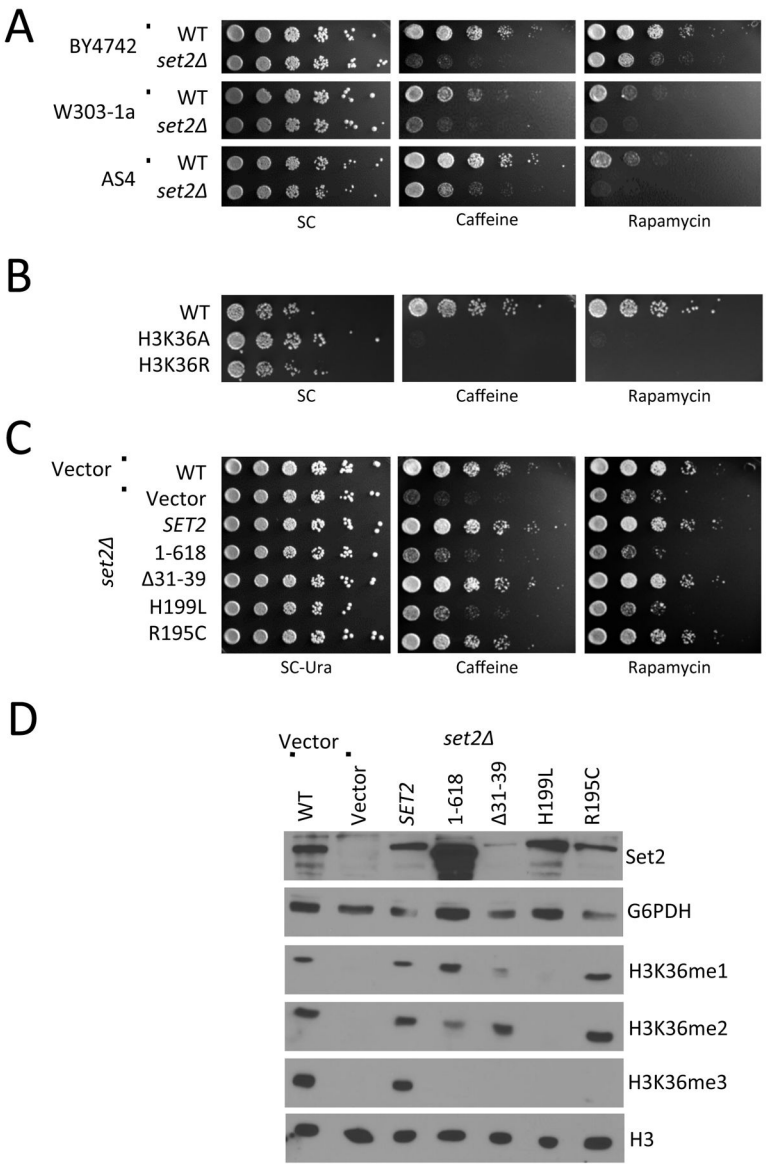


Figure 1. Cells lacking H3K36 methylation are sensitive to caffeine and rapamycin

A. – C. Five-fold serial dilutions of the indicated strains were plated on SC plates or plates containing caffeine (7–10 mM) or rapamycin (8–12.5 nM). D. Immunoblots of the indicated strains were probed with different H3K36me antibodies. H3 and G6PDH served as loading controls. The genetic background for the histone mutant strain is W303 and the background for the *SET2* allele analysis is BY4742.

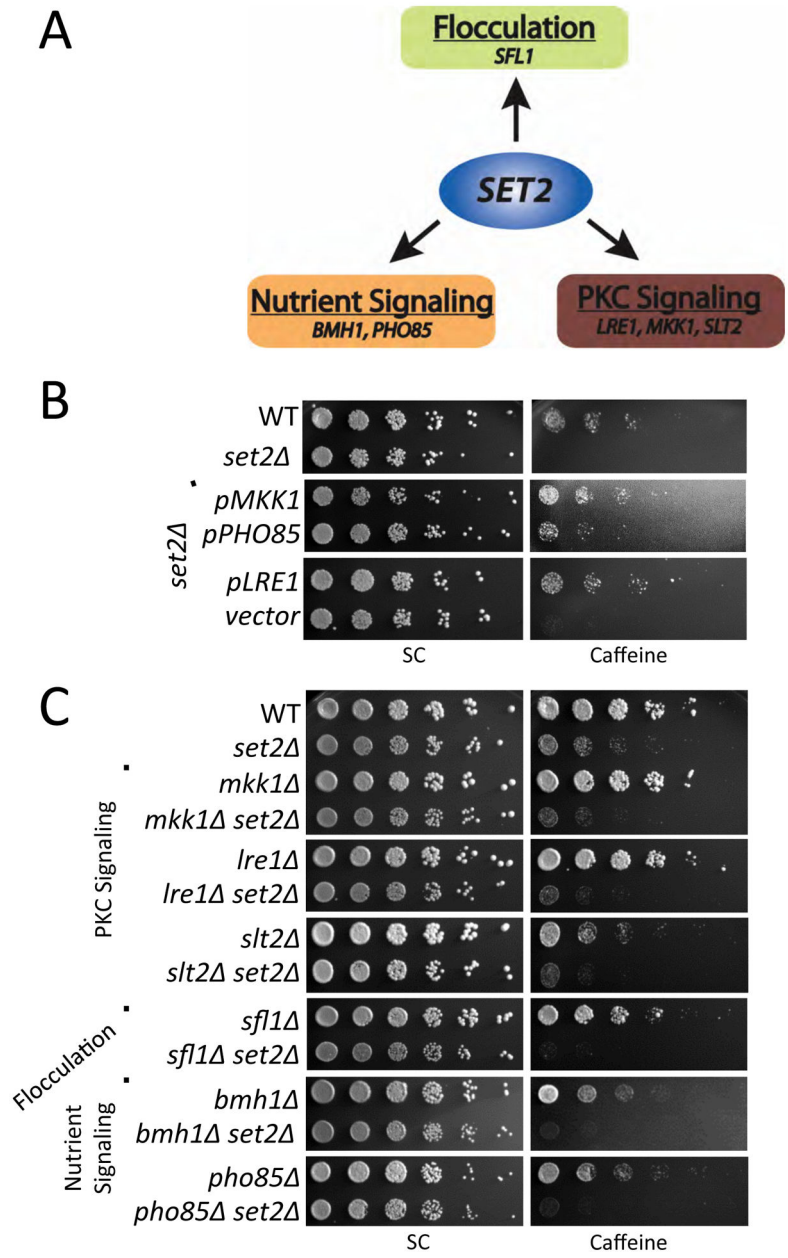


Figure 2. Nutrient response and PKC signaling pathways genetically interact with *SET2*

A. A genome-wide high copy suppressor screen revealed three pathways that suppressed the lethality of *set2*Δ cells on a high concentration of caffeine (20 mM). B. Five-fold serial dilutions of the indicated strains were plated on control or 20 mM caffeine plates. C. Five-fold serial dilutions of the indicated strains were plated on control or 5–10 mM caffeine plates.

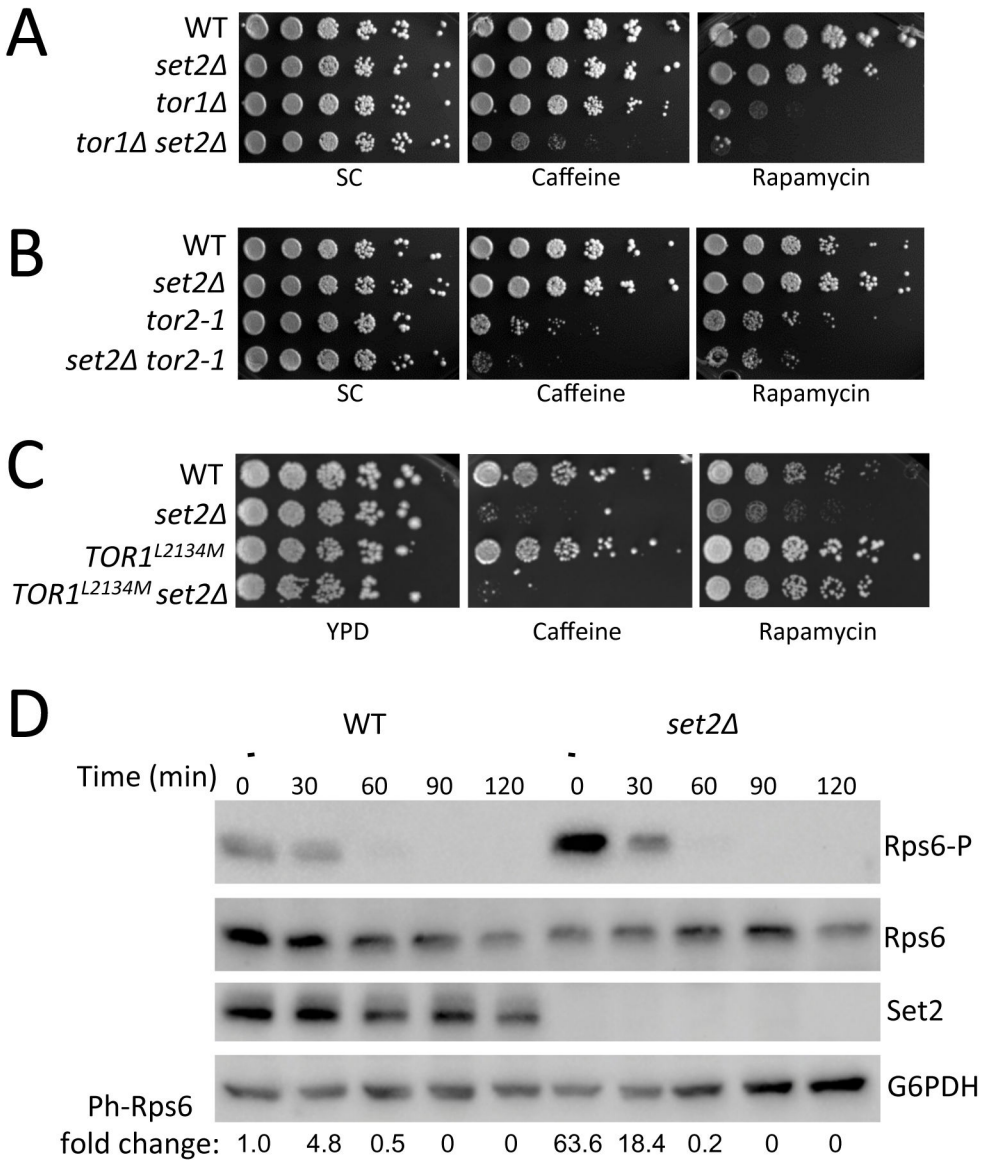


Figure 3. Set2 is required for proper nutrient stress response signaling

A.–B. Five-fold serial dilutions of the indicated strains were plated on control, caffeine (5 mM) and rapamycin (8 nM) plates. Given the extreme sensitivity of the *tor1* and *tor2* alleles to caffeine and rapamycin, lower concentrations of these drugs were used to examine synthetic interactions; however, these concentrations were just under that needed to observe growth defects in *set2*Δ cells. C. Five-fold serial dilutions of the indicated strains were plated on control, caffeine (5 mM), and rapamycin (30 ng/ml) plates. D. Log phase cells were transferred from a nutrient rich medium (YPD) to a medium lacking amino acids (SD). 1×10^8 cells were isolated at the indicated times and submitted to immunoblot with the indicated antibodies. G6PDH served as a loading control. An average of the change in ph-Rps6 levels in two biological replicates is quantified below.

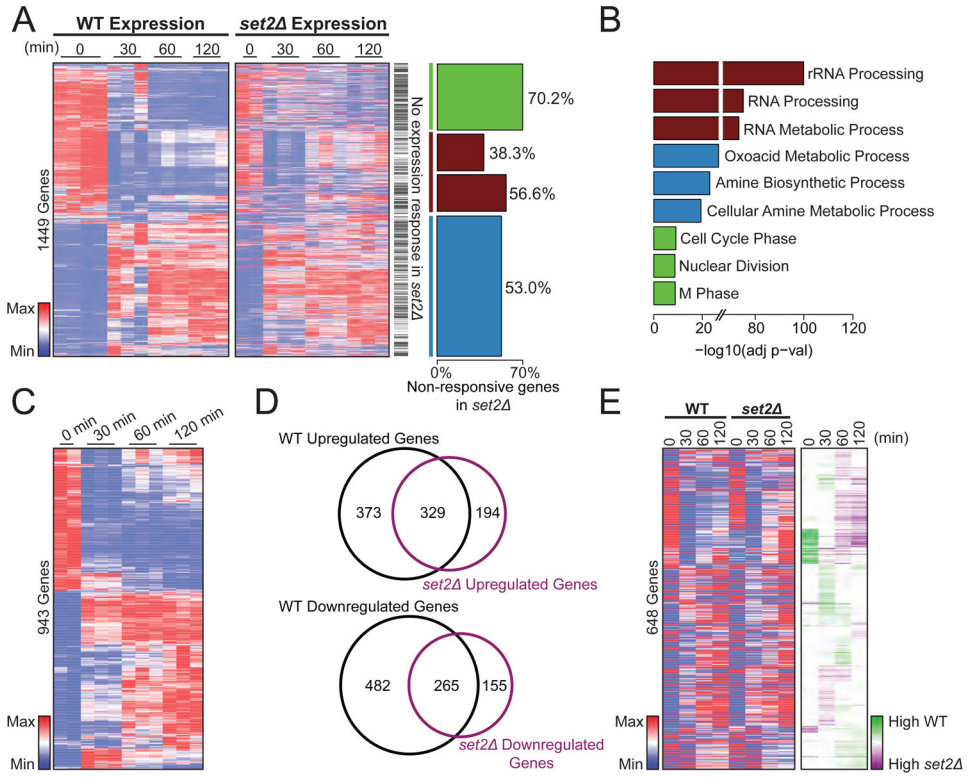


Figure 4. Set2 loss disrupts the transcriptional response to nutrient stress

A. (Left) Heat map showing relative WT RPKM values across biological replicates for 1,499 differentially expressed genes after nutrient depletion in WT cells. A. (Right) Heat map displaying relative *set2*Δ RPKM values for the same 1,499 differentially expressed genes identified in WT cells after nutrient depletion ($p < 0.5$ and $\log_2$ fold change > 0.5). Black lines to the right mark genes that no longer met statistical threshold for differential expression in *set2*Δ cells after nutrient depletion. The horizontal bar-plot identifies groups of differentially regulated of genes, displaying the percentage of genes in each group that were no longer differentially expressed in *set2*Δ cells. B. The three gene ontology classes with the greatest significance for groups identified in (A). The ontology bar colors match the gene groups defined in panel (A). C. Heatmap showing relative *set2*Δ RPKM values for the identified 943 differentially expressed genes in *set2*Δ cells. D. Venn diagrams comparing the identified up- and downregulated genes in WT and *set2*Δ cells. E. (Left) Heatmaps displaying scaled median RPKM values across the time course within each treatment, for genes identified as differentially expressed in both WT and *set2*Δ cells E. (Right) Heatmap showing subtracted scaled values (WT – *set2*Δ) for each gene identified as differentially expressed in both WT and *set2*Δ cells.

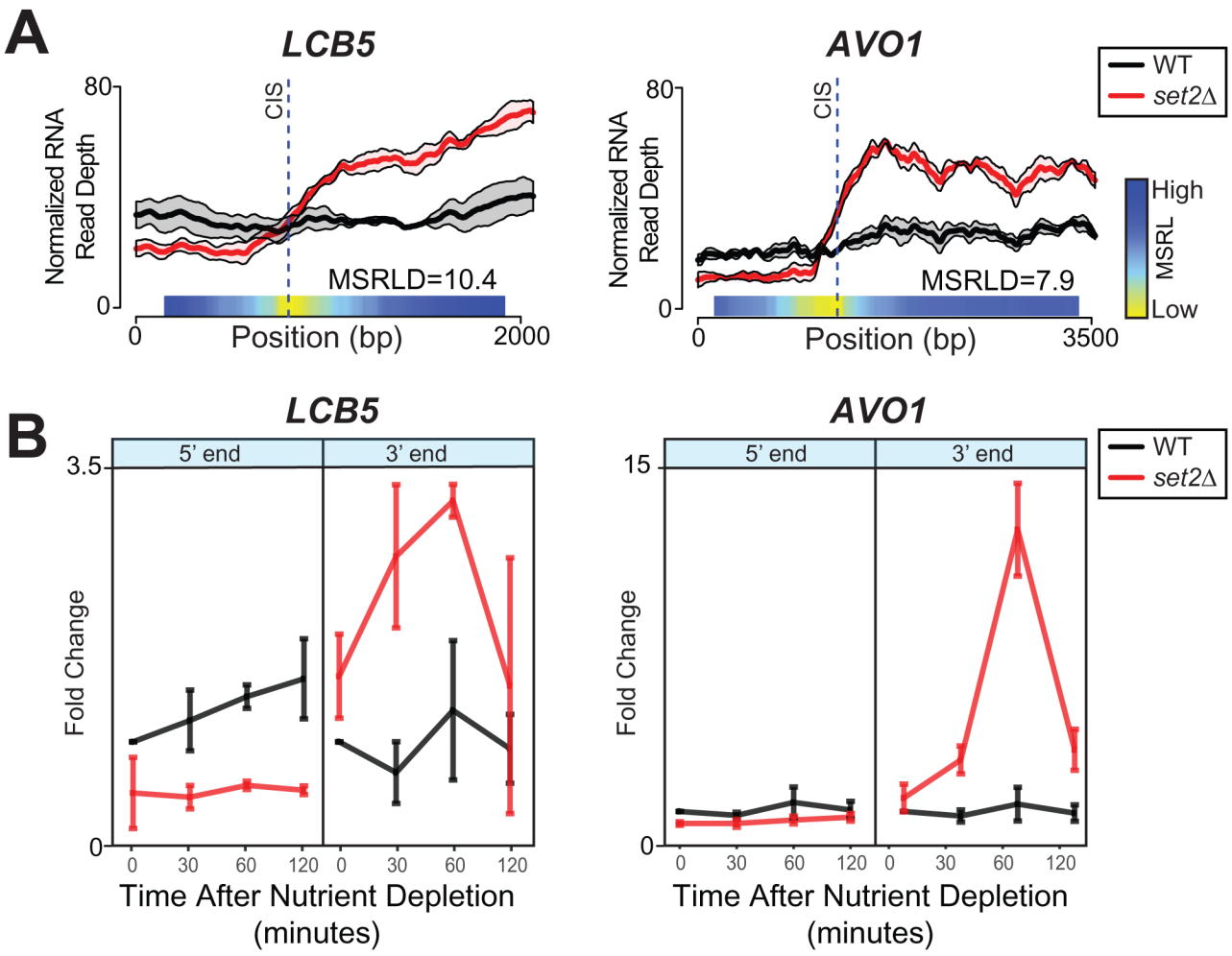


Figure 5. Nutrient depletion leads to sense cryptic initiation in the absence of Set2

A. Stranded RNA-seq signal for two representative genes (*LCB5* and *AVO1*) that demonstrate sense cryptic transcription in *set2*Δ (black: WT read depth normalized RNA-seq signal, red: *set2*Δ read depth normalized RNA-seq signal). The light shading represents standard deviation across replicates. The blue/yellow gradient represents the calculated mean square root loss (MSRL) for each gene. The CIS position is identified by the global MSRL minima. Mean square root loss difference (MSRLD) is the difference between the maximum and minimum MRSL for each gene. MSRLD score related to the magnitude and likelihood of CIS: MSRLD ≥ 4 (high), 2 ≤ MSRLD < 4 (intermediate), and MSRLD < 2 (low). B. Quantitative PCR signal for two representative genes (*LCB5* and *AVO1*) across the nutrient deprivation time course. Amplicons were located at the 5′ and 3′ ends of the genes. Fold change indicates qPCR signal relative to *SCR1* (WT black, *set2*Δ red).

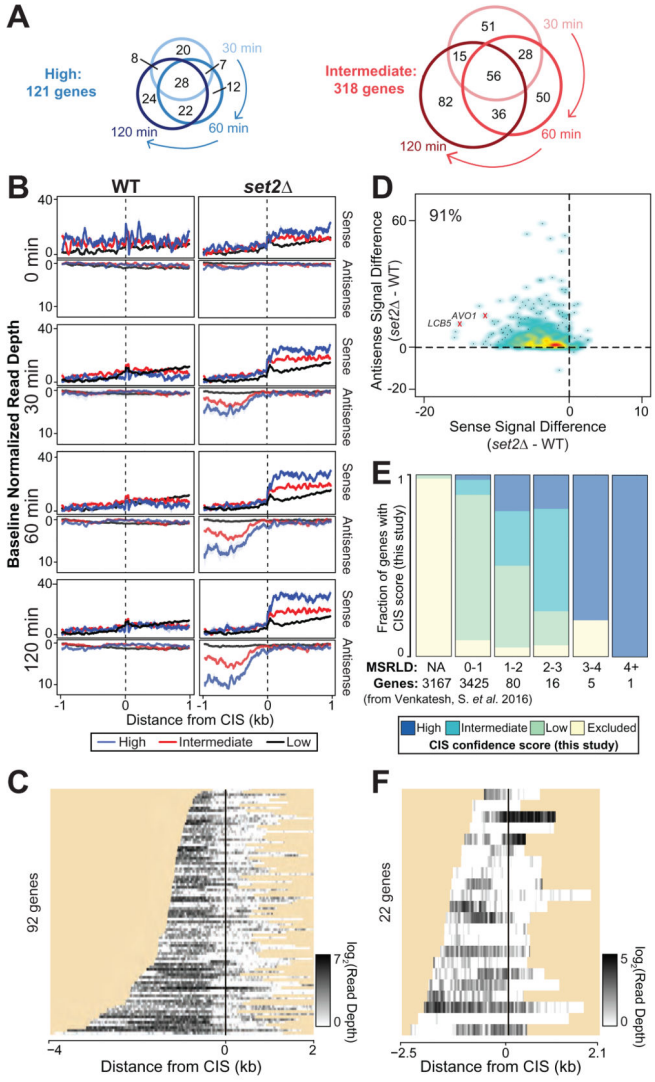


Figure 6. Nutrient depletion in the absence of Set2 leads to sense and antisense transcription affecting the expression of genes

A. Venn diagram associating the high (121 genes, blue) and intermediate (318 genes, red) scoring CISs during nutrient deprivation. B. Read depth normalized sense and antisense RNA-seq signal ±1 kb from the identified CISs with high (blue), intermediate (red) or low (black) scores in WT and *set2*Δ cells. To remove the effects of overall gene expression differences for each set, the minimum value for each line was adjusted to zero (y-axis). This correction enables a comparison of the magnitude of cryptic transcription between the three groups. The lighter colors represent standard deviations for each line. C. Heatmap of antisense RNA-seq signal (black). Of the 121 genes containing high scoring CISs, 92 (76%) had antisense transcription (average antisense signal 5′ of CIS > average antisense signal 3′ of CIS). Regions outside of the gene are masked (yellow) and ordered based on distance from CIS to the start of the gene. D. Density scatterplot of the median centered antisense and sense signal differences (*set2*Δ - WT) 5′ of the 439 identified CISs. For the sake of visualization, outlying points were not plotted (full dataset in Supplementary Figure 5A). E. Comparison of MSRLD based on data from Venkatesh, S. et al. 2016 and CIS scores for

genes from this study. Each column represents the number of genes identified from the Venkatesh data that are within the listed MSRLD range. The colors of in each column indicate the fractional overlap of genes from Venkatesh et al. 2016 and the CIS scores of genes identified in this study after nutrient depletion (high CIS (dark blue), intermediate CIS (light blue), low CIS (light green) and genes that are excluded from this study (light yellow)). The leftmost bar (marked as “NA”) represents genes that are excluded from the algorithm for reasons listed in the Supplemental Methods. F. Antisense RNA-seq signal (black) for the 22 genes that have MSRLD ≥ 2 based on data from Venkatesh, S. et al. 2016. Regions outside of the gene are masked (yellow) and ordered based on distance from CIS to the start of the gene.