Published online 28 July 2021

Nucleic Acids Research, 2021, Vol. 49, No. 14 ***8097–8109***
https://doi.org/10.1093/nar/gkab639

Core promoter activity contributes to chromatin-based regulation of internal cryptic promoters

Bo Bae Lee[1], Hyeonju Woo[1], Min Kyung Lee[1], SeoJung Youn[1], Sumin Lee[1], Jae-Seok Roe[2], Soo Young Lee[1] and TaeSoo Kim[1,*]

[1]Department of Life Science and the Research Center for Cellular Homeostasis, Ewha Womans University, Seoul 03760, Korea and [2]Department of Biochemistry, College of Life Science and Biotechnology, Yonsei University, Seoul 03722, Korea

Received March 20, 2021; Revised July 01, 2021; Editorial Decision July 13, 2021; Accepted July 16, 2021

ABSTRACT

During RNA polymerase II (RNA Pol II) transcription, the chromatin structure undergoes dynamic changes, including opening and closing of the nucleosome to enhance transcription elongation and fidelity. These changes are mediated by transcription elongation factors, including Spt6, the FACT complex, and the Set2-Rpd3S HDAC pathway. These factors not only contribute to RNA Pol II elongation, reset the repressive chromatin structures after RNA Pol II has passed, thereby inhibiting aberrant transcription initiation from the internal cryptic promoters within gene bodies. Notably, the internal cryptic promoters of infrequently transcribed genes are sensitive to such chromatin-based regulation but those of hyperactive genes are not. To determine why, the weak core promoters of genes that generate cryptic transcripts in cells lacking transcription elongation factors (e.g. *STE11*) were replaced with those from more active genes. Interestingly, as core promoter activity increased, activation of internal cryptic promoter dropped. This associated with loss of active histone modifications at the internal cryptic promoter. Moreover, environmental changes and transcription elongation factor mutations that downregulated the core promoters of highly active genes concomitantly increased their cryptic transcription. We therefore propose that the chromatin-based regulation of internal cryptic promoters is mediated by core promoter strength as well as transcription elongation factors.

INTRODUCTION

The elongation rate, processivity, and transcriptional fidelity of RNA Pol II are regulated by multiple factors, including those that evict or deposit histones within the gene body, thereby allowing RNA Pol II to pass; after PNA Pol II has passed, these factors then reset a repressive chromatin structure that blocks aberrant transcription. Other factors include those that mediate post-translational modifications of histones at gene bodies. These modifications including acetylation and methylation, which positively or negatively regulate elongation and spurious transcription from internal cryptic promoters (1,2). Histone acetylation can also directly stimulate transcription by reducing the interaction between histones and DNA and by recruiting factors that control chromatin structure and transcription. Histone acetylation is a highly dynamic modification regulated by the opposing functions of histone acetyltransferases (HATs) and histone deacetylases (HDACs). Histone methylation at specific sites also regulates histone acetylation, either by recruiting HATs or HDACs or by shaping their activities (3).

During transcription elongation, the carboxy-terminal domain (CTD) of Rpb1, the largest subunit of RNA Pol II, is phosphorylated at different timepoints and locations by CTD kinases. In yeast, these kinases include Kin28 (which complexes with the general transcription factor TFIIH) and Ctk1. Thus, during the early elongation step, Kin28 phosphorylates serine 5 of the CTD. Subsequently, during elongation, Ctk1 phosphorylates serine 2. The latter phosphorylation is important because it is recognized and bound by the methyltransferase Set2, which then induces co-transcriptional histone methylation at histone H3K36 during elongation (4–7). Set2-mediated H3K36 methylation peaks at the 3′ end of genes and is recognized by both HDACs and HATs; as a result, it plays important roles in regulating the histone acetylation/deacetylation at the gene body (3,8–11). Specifically, H3K36-methylated nucleosomes are deacetylated by Rpd3 small (Rpd3S) HDAC; this process is mediated by domains on two subunits of Rpd3S, the chromodomain of Eaf3 and the PHD finger domain of Rco1, which enhance Rpd3S binding to the chromatin and its deacetylation of histones in the gene body (12–15). This deacetylation activity by the Set2-Rpd3S HDAC pathway negatively regulates both RNA Pol II elongation and tran-

*To whom correspondence should be addressed. Tel: +82 2 3277 6807; Fax: +82 2 3277 3760; Email: tskim@ewha.ac.kr

scription from internal cryptic promoters (12,14,16,17). The H3K36 methylation is also a binding site for the NuA3 and NuA4 HATs: it is bound by the PWWP domain of the Pdp3 protein in NuA3 HAT (3,18,19) and by the Eaf3 subunit in NuA4 HAT (as well as in Rpd3S HDAC) (3,20). Thus, optimal transcription elongation and internal cryptic promoter regulation are likely to depend on a H3K36 methylation-regulated balance between acetylation and deacetylation within coding regions (19). It should be noted that the transcription from internal cryptic promoters is also regulated by other elongation factors and histone chaperones, including Spt6, FACT complex, Asf1, Paf1 complex, and Bur1-Bur2 kinase: all are required for normal Set2-mediated H3K36 methylation (21–25). Consequently, deletions in *SET2* or other elongation factors spur massive transcription from internal cryptic promoters in the gene body; this is particularly evident when the cells are confronted with environmental shifts (21,26–29).

While it is clear that the Set2–Rpd3S pathway represses cryptic transcription, many questions remain. First, what are the other factors that regulate internal cryptic promoters? Second, the Set2–Rpd3S pathway is known to primarily affect internal cryptic promoters at infrequently transcribed genes (17). Why highly active genes do not produce cryptic transcripts is unclear. Addressing this question may help to understand the commonalities of internal cryptic promoters, which remain poorly defined. It is possible that hyperactive genes may simply lack internal cryptic promoters in their gene bodies. Alternatively, the internal cryptic promoters in these genes may be repressed by as yet unknown mechanisms. We addressed these questions in the present study. We showed that high core promoter activity directly repressed internal cryptic promoters, even in mutants for the transcription elongation factors that displayed ample cryptic transcription. Highly active core promoters also blocked the histone modifications that were observed at active internal cryptic promoters. When the weak core promoter of genes with active internal cryptic promoters was replaced with a hyperactive one, histone H3 and H4 acetylation and trimethyl H3K4 (H3K4me3) were not seen. Interestingly, internal cryptic transcription was not only observed when transcription elongation factors were mutated, it also emerged transiently when the cells were exposed to environmental shifts that downregulated core promoters. We therefore propose that the core promoter plays an important role in the chromatin-based regulation of internal cryptic promoters.

MATERIALS AND METHODS

Yeast strains

Yeast strains used in this study are listed in Supplementary Table S1. To replace the core promoter of *STE11*, *PCA1* and *FLO8* with the core promoter of *YEF3*, *SSE1*, *CYC1*, or *HXT5*, the *delitto perfetto* strategy was used (30). The *kanMX4* and *KlURA3* CORE cassette (30) was amplified by PCR and then integrated at base pair position -500, -518, and -625 of *STE11*, *PCA1* and *FLO8* (relative to the corresponding +1 ATG start codon), respectively. The resulting strain was transformed with PCR products harboring the *YEF3*, *SSE1*, *CYC1* or *HXT5* core promoter region (*i.e.* the respective base pairs –695 to –1, –587 to –1, –1082 to –1 and –937 to –1 relative to the + 1 ATG start codon). To generate the *STE11-HIS3* reporter construct used in Figure 2A, the *kanMX4* and *KlURA3* CORE cassette was integrated at the 3′ end of *STE11* (+1871 to + 2154 of *STE11* relative to the + 1 ATG start codon) as described previously (25). The resulting strain was transformed with the entire *HIS3* gene. For the experiment shown in Supplementary Figure S2E, the *kanMX4* and *KlURA3* CORE cassette was integrated at the +1 ATG start codon (Met) of the *STE11*-TAP strain. The resulting strain was transformed with the 100 bp oligonucleotide that changed +1 ATG to a CGT codon.

Yeast culture conditions

For Figure 5D and Supplementary Figure S5B, time-course experiments with carbon-source shifts were conducted with the indicated strains as previously described (28,29). The *spt6-1004* temperature-sensitive mutants from Cheung *et al.* (21) were grown to an OD_{600} of ∼0.6 in YPD medium at 30°C, after which half were shifted to 37°C for an additional 80 min. For Figure 5F, the indicated strains were grown in YPD and then transferred to minimal media (SC) for 60 min, and then back to YPD for 120 min.

Northern blot analysis

Total RNA was isolated from yeast cells with hot acidic phenol/chloroform (1:1) (Sigma). 10 μg of total RNA or 1 μg of mRNA purified with Poly(A)Purist™ MAG Kit (Thermo, AM1922) was separated on agarose gel. The RNA was then transferred onto membranes (Bio-Rad, 1620159) and hybridized to radioactive probes as described previously (31). The strand-specific radioactive probes were generated by unidirectional PCR with [α-32P] dATP (PerkinElmer) with only one primer listed in Supplementary Table S2. Hybridization was conducted in a buffer containing 1% BSA, 7% SDS, 1 mM EDTA (pH 8.0) and 300 mM sodium phosphate buffer (pH 7.2).

Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation (ChIP) was carried out as previously described (32) with the oligonucleotides listed in Supplementary Table S2. Thus, anti-H3 (1.0 μl; Abcam, 1791), anti-acetyl H4 (1.0 μl; Millipore, 06–598), anti-acetyl H3 (1.0 μl; Millipore, 06-599), or anti-trimethyl H3K4 (0.5 μl; Millipore, 07-473) was bound to Protein A (GE-Healthcare, 17078001) agarose beads. Binding for anti-acetyl H4, anti-acetyl H3 and anti-H3K4me3 or for anti-H3 was done overnight in FA lysis buffer containing 1 M NaCl or in FA lysis buffer with 275 mM NaCl, respectively. Precipitates were washed with the same buffer, once with FA lysis buffer containing 1.5 M NaCl for anti-acetyl H4, anti-acetyl H3 and anti-H3K4me3 or with FA lysis buffer containing 500 mM NaCl for anti-H3, once with 10 mM Tris–HCl (pH 8.0), 0.25 M LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% sodium deoxycholate, and once with TE (10 mM Tris–HCl [pH 8.0], 1 mM EDTA). Precipitated DNAs were analyzed by real-time quantitative PCR using SYBR qPCR

Master mix (TOYOBO, QPS-201T) and CFX96 cycler (Bio-Rad).

Western blot analysis

Cells expressing TAP-tagged proteins were grown in YPD at 30°C to mid-log phase. Cells were lysed using lysis buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.1% NP-40) with protease inhibitors (2 μM Pepstatin A, 0.6 μM Aprotinine, 2 μM Leupeptin, 1 mM PMSF) and glass beads (Sigma). Protein concentration was quantitated by Bradford assay. For SDS-PAGE and western blot analyses, 5–50 μg of protein was used. Proteins were separated by SDS-PAGE and transferred onto a PVDF membrane (Millipore). Membranes were incubated with primary antibody for 3 h and then washed with PBST. After incubation with HRP-conjugated secondary antibody for 1 h, membranes were washed in three times with PBST. The blots were visualized on film (AGFA) with Super Signal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific).

Spot assay

Cells were resuspended and diluted to $OD_{600} = 1.0$ in synthetic complete (SC) media lacking any carbon sources and then subjected to 3-fold serial dilutions. Two microliters of each dilution were spotted on the indicated plates.

Tiling array data analysis

Regarding the experiments shown in Figure 1B, the *spt6-1004* and *spt16-197* genes that produced sense cryptic transcripts were obtained from Cheung *et al*. (21), while the genes producing sense or antisense cryptic transcripts in *SET2*-deleted cells were from Kim *et al*. (28). For Figure 1B, C and Supplementary Figure S1B, all genes were divided into five groups on the basis of the Rpb3 occupancy profiles described by Mayer *et al*. (33).

RESULTS

Transcription frequency correlates negatively with cryptic initiation

The Set2-Rpd3S HDAC pathway, FACT complex, and Spt6 not only regulate RNA Pol II elongation, but they also inhibit transcription initiation from internal cryptic promoters within gene bodies by reinstituting repressive chromatin (12–14,21,23). Examples of the genes having the internal cryptic promoters are the *STE11*, *PCA1* and *FLO8* genes: these genes respectively bear two, three, and one cryptic promoter(s) that are activated in transcription elongation mutants (21). Indeed, as shown in Figure 1A, loss of Set2 resulted in the activation of the internal cryptic promoters in all three genes, which produced 1–3 short transcripts. Full-length transcript levels were unaffected.

It has been suggested that the Set2-Rpd3S HDAC pathway mainly affects the internal cryptic promoters of infrequently transcribed long genes (17). However, it remains unclear whether RNA Pol II transcription frequency negatively correlates with transcription initiation from internal cryptic promoters. To address this, we analysed the data of previous studies that identified the many yeast genes that produce cryptic transcripts when Set2 is deleted (674 genes) or Spt6 or the FACT complex subunit Spt16 are mutated (960 and 1130 genes, respectively) (21,28). Venn diagram analysis showed that the *spt6-1004* and *spt16-197* mutants shared many cryptic transcript genes (74% and 63%, respectively) whereas the Set2 mutant shared relatively fewer genes with these mutants (Supplementary Figure S1A). We then investigated the relationship between the core promoter activity and cryptic initiation by dividing all of the yeast genes into five groups (G1 to G5) on the basis of their RNA Pol II occupancy (Supplementary Figure S1B) and then analysing how many of these genes produced cryptic transcripts in the *set2Δ*, *spt6-1004* or *spt16-197* mutants. Indeed, in *set2Δ*, the genes producing cryptic transcripts were more likely to be the less active genes: 78% were in the G3–G5 groups (Figure 1B). Interestingly, this pattern was also seen in the *spt6-1004* and *spt16-197* mutants (Figure 1B): 80% and 84% of the cryptic transcript genes in *spt6-1004* and *spt16-197* had intermediate–low RNA Pol II occupancy (G3–G5), respectively. Notably, the three cryptic transcript genes *STE11*, *PCA1*, and *FLO8* also fell into the less active gene groups (G4–G5). We next examined the relationship between the gene length and cryptic initiation as the internal cryptic promoters within long genes were known to be more sensitive to the Set2-Rpd3S pathway (17). As expected, a majority of genes producing cryptic transcripts in *set2Δ* were longer than 1000 bp. The same pattern was observed in mutants for Spt6 and Spt16 (Supplementary Figure S1C). Importantly, the genes from 1000 bp to 2000 bp in length generating cryptic transcripts in mutants for transcription elongation factors were more likely to be the less active genes (Supplementary Figure S1D).

To further confirm the negative correlation between RNA Pol II occupancy and cryptic initiation, the hyperactive genes in the G1 groups of *set2Δ*, *spt6-1004*, and *spt16-197* were subdivided into five subgroups according to their core promoter activity and the genes with cryptic transcription were counted. As shown in Figure 1C, in all three mutants, the genes with cryptic transcripts were most likely to be the least active of the G1 genes (G1-3 to G1-5). These data suggest that indeed, transcription initiation from internal cryptic promoters associates with low RNA Pol II transcription frequency.

Enhanced core promoter activity suppresses cryptic initiation

An important question is, why do highly transcribed genes tend to lack cryptic transcripts in transcription elongation factor mutants? There are two possible answers: either hyperactive genes simply lack internal cryptic promoters within the gene body, or enhanced core promoter activity inhibits transcription initiation from internal cryptic promoters, possibly *via* transcriptional interference. To test this idea, the core promoter of the *STE11* gene, which belongs to the relatively inactive G4 group, was replaced with the promoters from highly active genes (*i.e.* the G1 genes *YEF3* and *SSE1*) or a slightly less active gene (i.e. the G2 gene *CYC1*) (Figure 1D). The RNA Pol II occupancy on the *STE11* variants was determined by the ChIP assay. The *YEF3* promoter (pYE-*STE11*) was the strongest one, fol-

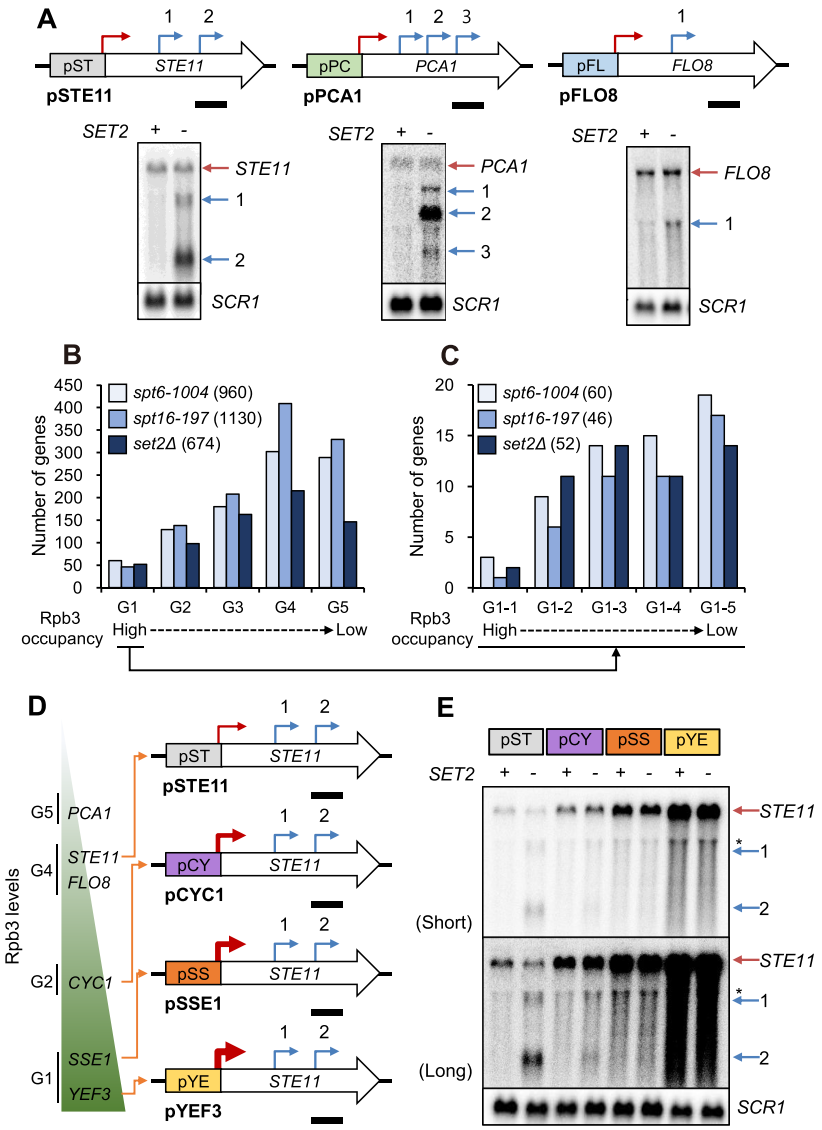

Figure 1. RNA Pol II transcription frequency correlates negatively with cryptic initiation. (**A**) Set2 suppresses the internal cryptic transcription of *STE11*, *PCA1*, and *FLO8*. The top images show schematic depictions of *STE11*, *PCA1*, and *FLO8*. The red arrow shows the core promoter while the blue arrows indicate the internal cryptic promoters that produce short cryptic transcripts when *SET2* is deleted (bottom panels). The wild-type and *set2Δ* mutant cells were grown to an OD_{600} of ∼0.6 in YPD medium at 30°C and then subjected to northern blot analysis. Two independent experiments showed the same results. (B, C) The genes that produce cryptic transcripts in the *set2Δ*, *spt6-1004* and *spt16-197* mutants tend to have lower Rpb3 occupancy. (**B**) All yeast genes were divided into five groups based on their Rpb3 occupancy, as reported by Mayer *et al.* (33). The genes in each group that produced cryptic transcripts in *set2Δ*, *spt6-1004* or *spt16-197* were then counted. Total numbers of genes in each group are 1343. **(C)** The G1 group from (B) was subdivided into five groups on the basis of their Rpb3 occupancy. The genes that produced cryptic transcripts in each mutant were counted. While G1-1 and G1-2 include 268 genes, total numbers of genes for three other groups are 269. (D, E) Strong core promoter activity reduces cryptic initiation in *SET2*-deleted cells. (**D**) Schematic representation of the *STE11* gene in which the original promoter (wild-type gene, pST-*STE11*, grey) was replaced by the promoter from two strongly transcribed genes (G1 group; *YEF3* and *SSE1*) or a more weakly transcribed gene (G2 group; *CYC1*). These variants are respectively designated pYE-*STE11* (yellow), pSS-*STE11* (orange), and pCY-*STE11* (violet). The red and blue arrows respectively indicate the core promoter and the *STE11* internal cryptic promoters. The thickness of the red promoters indicates the transcriptional activity of the indicated core promoters. A bar underneath indicates the position of the probe used for northern blot analysis. (**E**) Wild-type and *set2Δ* cells with the indicated *STE11* variants were grown as described in (A) and then subjected to northern blot analysis of the *STE11* and cryptic transcripts. The red and blue arrows indicate the core and the internal cryptic promoters of *STE11*, respectively. The effect of long and short exposure time is shown. The asterisks indicate nonspecific bands. *SCR1* was used as a loading control. Two independent experiments showed the same results.

lowed by the *SSE1* promoter (pSS-*STE11*), the *CYC1* promoter (pCY-*STE11*), and, finally, the original *STE11* promoter (pST-*STE11*) (Supplementary Figure S1E).

We next measured the cryptic transcripts from *STE11* in the presence and absence of Set2. In both the wild-type and *set2Δ* cells, the full-length *STE11* transcript levels were highest when the core promoter of *STE11* was replaced with the *YEF3* promoter. The next highest levels were observed with the *SSE1* promoter and then the *CYC1* promoter. The original *STE11* promoter very weakly produced full-length transcripts (Figure 1E). These findings are consistent with the Rpb3 levels in Supplementary Figure S1E. As observed above (Figure 1A), the deletion of *SET2* produced two differently sized cryptic transcripts at *STE11* transcribed from its own promoter. Interestingly, however, as the core promoter activity increased, these cryptic transcript levels dropped. Specifically, while the modestly strong promoter in pCY-*STE11* produced detectable levels of cryptic transcripts in *set2Δ*, this was not observed for pSS-*STE11* or pYE-*STE11*, which have strong promoters (Figure 1E, Supplementary Figure S1F and 1G). Although a weak band was detected in both wild type and *set2Δ* in pYE-*STE11* background, this signal was significantly reduced in northern blot analyses using purified mRNAs suggesting that this might be from non-specific hybridization to a different RNA (Figure 1E, Supplementary Figure S1H and 1I). We speculated that the failure to detect cryptic transcripts in the latter variants could perhaps reflect their rapid degradation: some cryptic transcripts are known to be removed by the nuclear exosome (28,34). However, this possibility is unlikely: when we examined *STE11* transcripts from the *YEF3* promoter in *set2Δrrp6Δ*, which lacks nuclear exosome activity, cryptic transcripts were also not detected (Supplementary Figure S2A).

To confirm these results, we constructed a reporter system that allowed us to detect the activation of the second cryptic promoter in *STE11* (Figure 2A). This reporter consists of the open reading frame of *HIS3* inserted into the 3′ part of *STE11*: thus, when the second *STE11* cryptic promoter is activated, *HIS3* is expressed. This reporter was employed with *set2Δ* or wild-type cells during a growth assay on plates lacking histidine. The *set2Δ* strain (but not the wild-type strain) containing the normal *STE11* promoter was able to grow on the plates lacking histidine because the second cryptic promoter was activated. By contrast, neither *set2Δ* nor the wild-type strain grew in the absence of histidine if they contained pYE-*STE11-HIS3*. Thus, when the core promoter is strong, cryptic initiation in *set2Δ* is blocked (Figure 2B). Northern blot analysis confirmed that *HIS3* transcripts were only observed in *set2Δ* cells when pST-*STE11-HIS3* was present; these transcripts were lacking when the *STE11-HIS3* construct had a strong promoter (Supplementary Figure S2B). Similar effects were observed with a mutant for Rco1, a Rpd3S HDAC-specific subunit. Thus, like *set2Δ*, the *RCO1* deleting cells generated two cryptic transcripts from *STE11* when the normal *STE11* promoter, but not the *YEF3* promoter, was present (Figure 2C). Similarly, the histidine growth assay showed that histidine was only produced by the *STE11-HIS3* reporter in the *RCO1* mutant when it was transcribed from the original *STE11* promoter, although the effect was not as pronounced as for the *set2Δ* mutant. This effect was also observed when Eaf3, another Rpd3S HDAC subunit, was mutated (Supplementary Figure S2C).

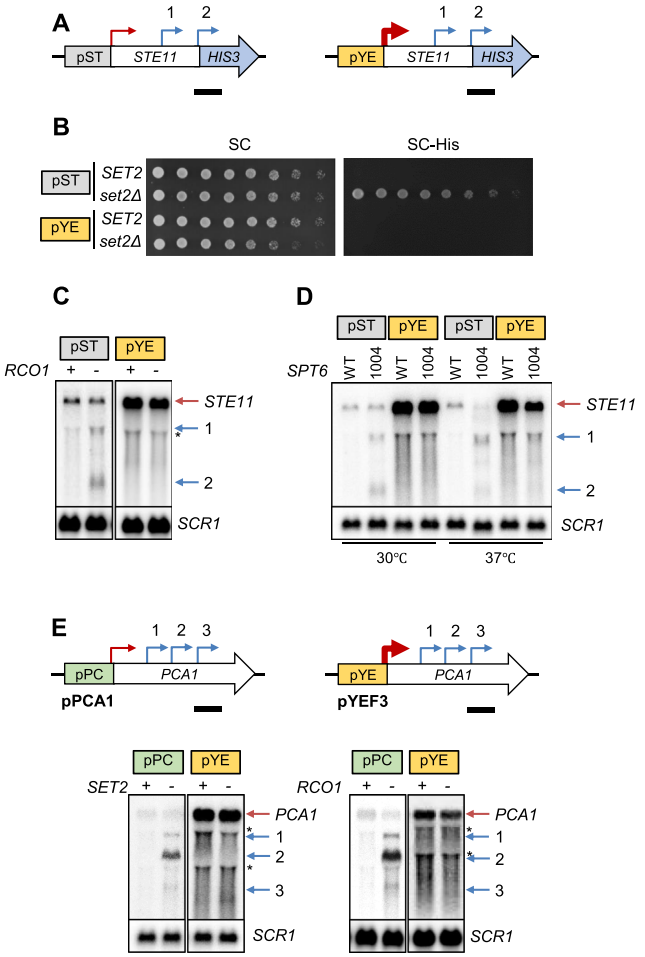

Figure 2. Enhanced core promoter activity inhibits transcription from internal cryptic promoters in Set2-Rpd3S pathway or Spt6 mutants. (A, B) A strong (*YEF3*) but not weak (*STE11*) core promoter causes the growth in medium without histidine *of SET2* deleting cells bearing a reporter construct composed of *STE11* with *HIS3* inserted just downstream of the second cryptic promoter. (**A**) Schematic representation of the *STE11-HIS3* reporter constructs. A bar underneath indicates the position of the probe used for northern blot analysis. (**B**) The wild-type and *set2Δ* cells bearing the indicated reporter constructs were spotted in 3-fold serial dilutions onto plates with synthetic complete (SC) medium (2 days growth shown) or SC medium lacking histidine (3 days growth shown). (C, D) A strong promoter also blocks cryptic transcription from *STE11* in *rco1Δ* (**C**) and *spt6-1004* (**D**) cells. *RCO1* and *rco1Δ* cells were grown as described in Figure 1A. *SPT6* and *spt6-1004* cells were grown to an OD_{600} of ∼0.6 in YPD at 30°C, after which half were shifted to 37°C for an additional 80 min. Northern blot analyses of *STE11* were conducted as described in Figure 1A. Two independent experiments showed the same results. (**E**) The strong *YEF3* promoter also inhibits cryptic transcription of *PCA1*, which has an inactive (G5) promoter. The top images show the *PCA1* gene with its original weak promoter and with the *YEF3* promoter. The bottom panels show the cryptic transcripts of *PCA1* on northern blot analysis, which was conducted as described in Figure 1A. The red arrows indicate core promoters while the blue arrows depict the internal cryptic promoters. Two independent experiments showed the same results.

Similar results were observed when Spt6, which also represses cryptic transcription (12–14,21,23), was mutated.

Thus, the *spt6-1004* strain, like mutants for the Set2-Rpd3S HDAC pathway, produced both *STE11* cryptic transcripts at the permissive and non-permissive temperatures if the normal *STE11* promoter was present. Interestingly, however, if *STE11* was transcribed from the *YEF3* promoter, the *STE11* internal cryptic promoters were not activated, even at the non-permissive temperature (Figure 2D and Supplementary Figure S2D). Another interesting point relating to this mutant is that it produces a *STE11* cryptic transcript that is translated into a short protein lacking the N-terminal region (21). We found that this short Ste11 protein was also present in *set2Δ* containing the normal *STE11* promoter. It was even observed in *set2Δ* even in the absence of the +1 ATG start codon of *STE11* (Supplementary Figure S2E). However, this short protein was not observed when the *STE11* core promoter was replaced with the *YEF3* core promoter (Supplementary Figure S2F). These findings together show clearly that if the *STE11* promoter is replaced with that of *YEF3*, cryptic transcription at *STE11* is completely blocked in the Set2-Rpd3S HDAC pathway and Spt6 mutants.

To test that this effect is not specific for the *STE11* internal cryptic promoters, the core promoters of the inactive G4–G5 genes, *PCA1* and *FLO8* were replaced with that of *YEF3*. The cryptic transcripts of *PCA1* and *FLO8* detected in *set2Δ* and *rco1Δ* were not seen when the core promoters of these two genes were hyperactive (Figure 2E and Supplementary Figure S2G–H). Taken together, these findings suggest that the activity of the core promoter plays an important role in the chromatin-based regulation of the internal cryptic promoters in gene bodies.

Enhanced core promoter activity blocks deposition of active histone marks at the internal cryptic promoters

Set2 participates in the repression of internal cryptic promoters by methylating H3K36, which is recognized by Rpd3S HDAC and promotes its histone deacetylation activity (12,14,17). By contrast, transcription initiation at internal cryptic promoters associates with trimethylation at H3K4 (H3K4me3) and histone acetylation. We therefore determined the levels of acetylation of histone H4 (with an antibody that recognizes tetra-acetyl H4), acetylation of histone H3 (with an antibody recognizing di-acetyl H3 on K9 and K14), and H3K4me3 (with an antibody that recognizes trimethylated K4 of H3) in *set2Δ* and the wild-type strain containing pST-*STE11* or pYE-*STE11*. The levels of these modifications were normalized to total histone content, as determined by histone H3 levels. When the normal *STE11* promoter was present, *SET2* deletion increased H4 and H3 acetylation and particularly H3K4me3 levels at the *STE11* internal cryptic promoter. However, when the *YEF3* promoter was present, there was no H4 or H3 acetylation and the H3K4me3 levels were significantly decreased, albeit still present at higher levels in *set2Δ* relative to the wild-type cells (Figure 3A and B). The latter observation may be due to the spreading of H3K4me3 into coding regions in the absence of Set2 (35). Increased histone acetylation but not H3K4me3 was also seen in a mutant for Rco1 (Figure 3C). We then determined occupancy of Sua7, a component of TFIIB. While loss of Set2 showed 2.5 fold increase

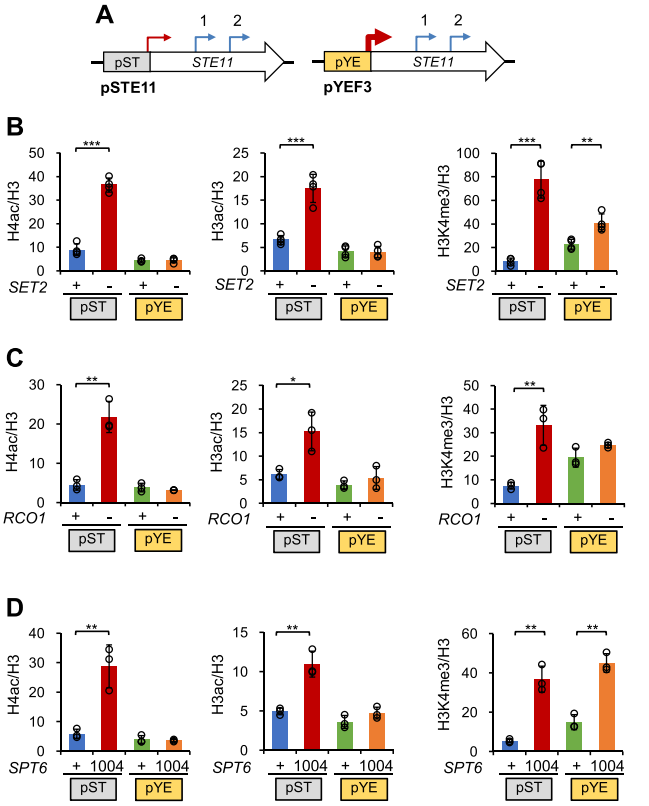

Figure 3. Core promoter activity shapes the histone modifications at internal cryptic promoters. (**A**) Schematic representation of pST-*STE11* and pYE-*STE11* strain. The *STE11* and *YEF3* promoters are shown in grey and yellow, respectively. The core and internal cryptic promoters are indicated by the red and blue arrows, respectively. (**B–D**) Effect of Set2-Rps3S HDAC pathway and Spt6 mutations on histone modifications at the *STE11* internal cryptic promoter. Cross-linked chromatin from the indicated cells grown in YPD at 30°C was precipitated with anti-acetyl H4, anti-acetyl H3, anti-H3K4me3, or anti-H3 antibodies. The precipitated DNA was subjected to PCR analysis of the second internal cryptic promoter region of *STE11*. A non-transcribed region near the telomere of chromosome VI was used as an internal control. The signals for acetyl-H4, acetyl-H3, or H3K4me3 were quantitated and normalized to the H3 signal, and the ratios were graphed. Error bars show the standard deviation (S.D.) calculated from four or three biological replicates, each with three technical replicates. $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$ (two-tailed unpaired Student's *t* tests).

of Sua7 binding at the *STE11* internal cryptic promoter in pST-*STE11* strain, this was not seen in pYE-*STE11* (Supplementary Figures S3A and S3B).

Consistent with the slight production of cryptic transcripts in *set2Δ* with pCY-*STE11* (Figure 1E and Supplementary Figure S1F), *set2Δ* cells with pCY-*STE11* exhibited slightly more histone acetylation and H3K4me3 levels at the internal cryptic promoters compared to when pYE-*STE11* was present (Supplementary Figures S3C and S3D). The same patterns were observed when *PCA1* and *FLO8* bearing their own promoter or the *YEF3* promoter were used instead of *STE11* in *set2Δ* or *rco1Δ*: the original promoter associated with histone acetylation and H3K4me3 at the internal cryptic promoters in the mutants but this was largely abolished by the strong *YEF3* promoter (Supplementary Figure S3E–I). These findings indicate that hy-

peractive core promoters both block transcription initiation and the deposition of active histone marks at internal cryptic promoters in mutants for the Set2-Rpd3S HDAC pathway.

This was also largely observed in *spt6-1004* mutant. Consistent with the cryptic initiation phenotypes of this mutant with pST-*STE11* or pYE-*STE11* (Figure 2D), *spt6-1004* exhibited increased histone acetylation at the internal cryptic promoter of *STE11* when the normal *STE11* promoter was present but not when the core promoter of *STE11* was replaced with that of *YEF3*. Interestingly, however, *spt6-1004* showed higher H3K4me3 levels regardless of whether the normal *STE11* or *YEF3* promoter was present (Figure 3D). This suggests that Spt6 may have an additional function in regulation of H3K4me3 pattern. These findings suggest that the activity of the core promoter helps regulate the histone modifications at internal cryptic promoters.

Cryptic initiation in hyperactive genes may require downregulation of core promoter activity

The data presented above suggest that elevated transcription from the core promoter inhibits the activation of internal cryptic promoters within gene bodies. An obvious question is, how are internal cryptic promoters at highly active genes regulated? To address this question, we initially sought to identify the factors that specifically control the activity of the internal cryptic promoters of highly transcribed genes. Thus, we examined the effect of mutations in factors that are known to regulate cryptic transcription. Asf1, Bur2, and Paf1 help repress cryptic promoters (22,24,36). However, deleting these factors in *set2Δ* had no effect on cryptic transcription when the *STE11* gene was transcribed from the *YEF3* promoter (Figure 4A). Other factors examined were Rtf1, Chd1, Spt2 and Rtt106, which have been found to regulate cryptic transcription in the *FLO8* gene. Specifically, Cheung *et al*. created a *FLO8:HIS3* reporter construct in which *HIS3* is located 3′ of *FLO8* and the normal promoter of *FLO8* is replaced by the strong *GAL1* promoter, which is activated by culture in medium containing galactose. Cells that bore this reporter and lacked Spt2, Chd1, or Rtf1 were able to grow on plates lacking histidine. This suggested that these factors may regulate the internal cryptic promoters in highly active genes (21). However, mutations in Spt2, Chd1, and Rtf1 did not induce cryptic transcription at *STE11* when the *YEF3* promoter was present (Supplementary Figure S4A). We also asked whether treating *set2Δ* with 6-AU (6-Azauracil) would induce *STE11* cryptic transcription in the presence of the *YEF3* promoter: since 6-AU inhibits transcription elongation by reducing NTP pools, it could therefore activate internal cryptic promoters at highly active genes (37). However, 6-AU treatment did not induce cryptic transcription from *STE11* in *set2Δ* when the *YEF3* promoter was present, as shown by the histidine growth assay (Supplementary Figure S4B and 4C). Thus, the specific factors and mechanisms that regulate internal cryptic promoters in hyperactive genes remain to be elucidated. As shown in Figure 1C, *set2Δ*, *spt6-1004* and *spt16-197* respectively had 52, 60 and 46 genes that produced cryptic transcripts despite having strong promoters (G1 group). To determine why these highly active genes still had cryptic transcription, we analysed the 52 genes with Set2-repressed internal cryptic promoters. We showed previously that while many cryptic promoters are constitutively active in *set2Δ* cells, ∼50% of all Set2-regulated cryptic promoters are only induced in medium containing galactose (28). Interestingly, the vast majority (47/52, 96%) of the Set2-regulated genes in the G1 group had inducible internal cryptic promoters that only generated cryptic transcripts during galactose incubation (Figure 4B). Two of these genes were *KRS1* and *VTS1*: both had an internal cryptic promoter that produced antisense transcript in *set2Δ* cells when they were grown in media containing galactose. This transcription increased with the duration of galactose exposure but was absent when the cells were grown in raffinose. Importantly, both *KRS1* and *VTS1* were downregulated during galactose incubation (Figure 4C and D). This raises the interesting possibility that some internal cryptic promoters are only expressed when core promoter activity is downregulated by environmental shifts. This notion is supported by our analysis of the 47 Set2-regulated genes in the G1 group that had galactose-inducible internal cryptic promoters: the mRNA expression of 42 (89%) was significantly downregulated during galactose incubation (Figure 4E). Similarly, the 60 and 46 hyperactive G1 group genes that produced cryptic transcripts in the Spt6 and Spt16 mutants, respectively (Figure 1C) were also significantly downregulated (probe 1 in Figure 4F–H) (21); by contrast, the cryptic transcription of these genes tended to be increased in mutants for Spt6 and Spt16 (probe 6 in Figure 4F–H). These findings suggest that the core promoter activity must be downregulated before internal cryptic promoters can be activated.

Dynamic regulation of core and internal cryptic promoters by environmental shifts

To further investigate the effect of the core promoter activity on internal cryptic promoters, the core promoter of *STE11* was replaced with that of *HXT5*, which is differentially activated by different carbon sources. Specifically, the *HXT5* core promoter is inactive in cells cultured in synthetic complete medium containing raffinose; becomes activated when exposed to galactose for 30 or 120 min; is then rapidly downregulated when transferred to medium containing glucose for 30 or 120 min; and becomes hyperactivated when shifted from glucose back to galactose for 30 or 120 min (28,29,38) (Figure 5A and B). The latter effect is known as transcriptional memory (Figure 5B). We monitored *STE11* full-length mRNA and cryptic transcripts during carbon source shifts in *set2Δ* and wild-type cells containing pHX-*STE11* (Figure 5C). In the wild-type cells, the first and second galactose exposures (Gal 120 and 2nd Gal 120, respectively) slightly and potently induced *STE11* mRNA transcription from the *HXT5* core promoter, respectively. The *set2Δ* cells exhibited similar *STE11* mRNA transcript levels. However, the *STE11* cryptic transcripts in the *set2Δ* cells displayed the opposite pattern: they were detected in raffinose and glucose media but disappeared completely during the 2nd galactose exposure, which is when the *HXT5* promoter was hyperactivated by transcriptional memory (Figure 5D). Consistent with this, histone H4 acetylation

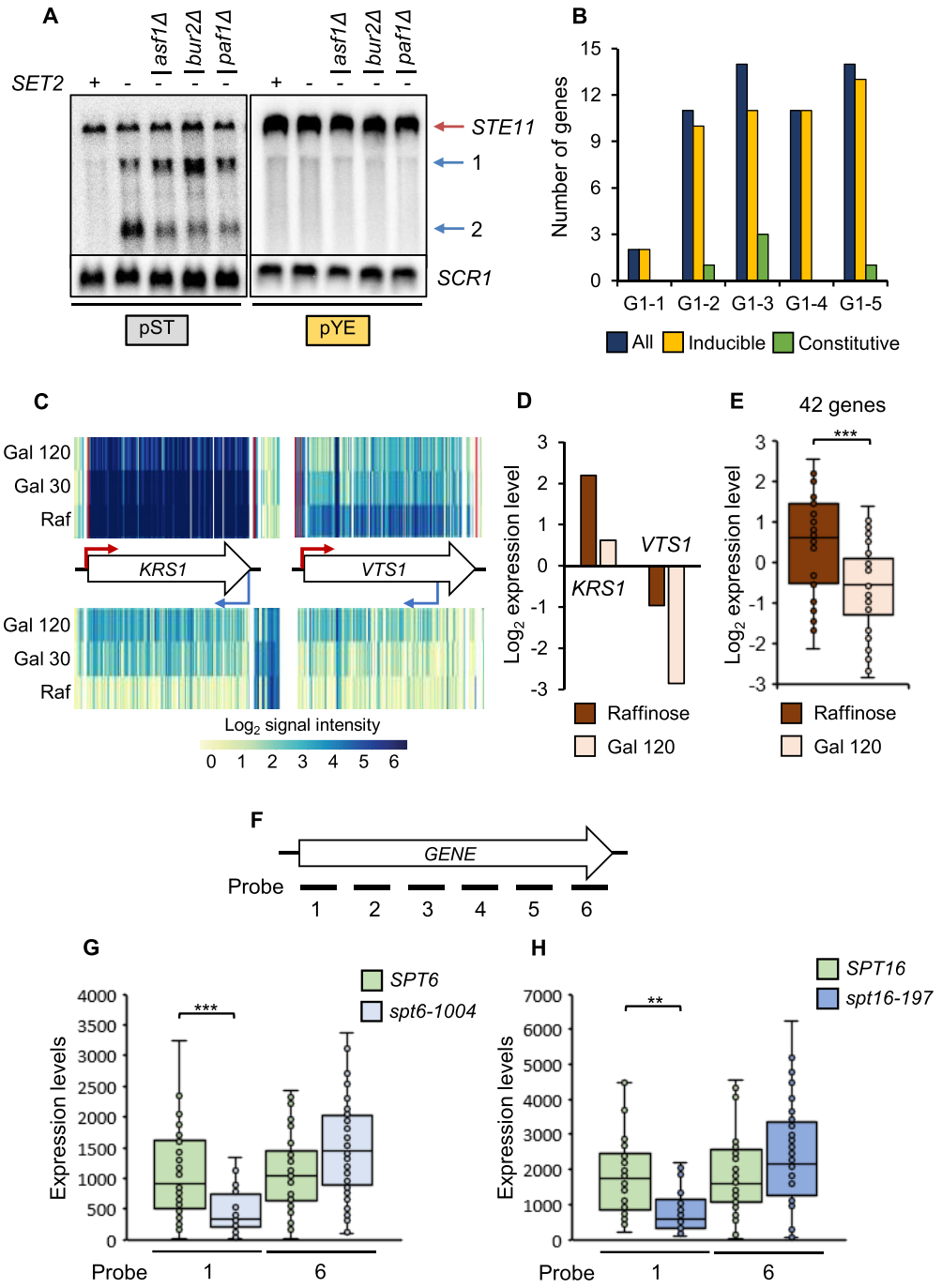

Figure 4. Downregulation of the core promoter activity enhances cryptic transcription. (**A**) Deletion of various transcription elongation factors did not affect *STE11* cryptic transcription in *set2Δ* cells that contained pYE-*STE11*. Northern blot analysis was conducted with the indicated cells as described in Figure 1A. Two independent experiments showed the same results. (**B**) Of the 52 G1 group genes with highly active core promoters that bear Set2-repressed internal cryptic promoters (Figure 1C), 47 had inducible cryptic promoters that were activated by galactose. The number of genes with inducible or constitutively active cryptic promoters are shown. (**C**) *KRS1* and *VTS1* are examples of genes with strong core promoters and inducible Set2-repressed internal cryptic promoters. The images show the microarray hybridization signals for the multiple probes arrayed along each gene (28) when *set2Δ* cells were cultured with raffinose or shifted to galactose for 30 or 120 min. Transcription from the core promoter (red arrow) is shown on top while antisense transcription from the cryptic promoter (blue arrow) is shown at the bottom. Increasing blue bars indicate more transcripts hybridizing to the array. The red lines show the annotated start and stop of the mRNA. The white box arrows show the position of the ORF. When *set2Δ* cells were exposed to galactose, *KRS1*/*VTS1* transcription from the core promoter and the cryptic promoter was decreased and increased, respectively. (**D**) Expression level of these two genes at raffinose and galactose 120 min time point. (**E**) Boxplot showing that of the 47 genes with Set2-repressed galactose-inducible internal cryptic promoters (see B), 42 were downregulated by 120 min of galactose exposure. $^{***}P < 0.001$ (two-tailed unpaired Student's *t* tests). (F–H) Downregulation of highly transcribed genes in G1 groups in mutants for Spt6 or Spt16. (**F**) Schematic representation of the microarray probe positions used to identify the cryptic promoters. Probes 1 and 6 are markers for mRNA transcription and cryptic transcription, respectively from Cheung *et al*. (21). (G, H) The microarray signals for cryptic transcripts (probe 6) were increased but the signals for mRNAs (probe 1) were decreased in *spt6-1004* (**G**) and *spt16-197* (**H**) cells. $^{**}P < 0.01$, and $^{***}P < 0.001$ (two-tailed unpaired Student's *t* tests).

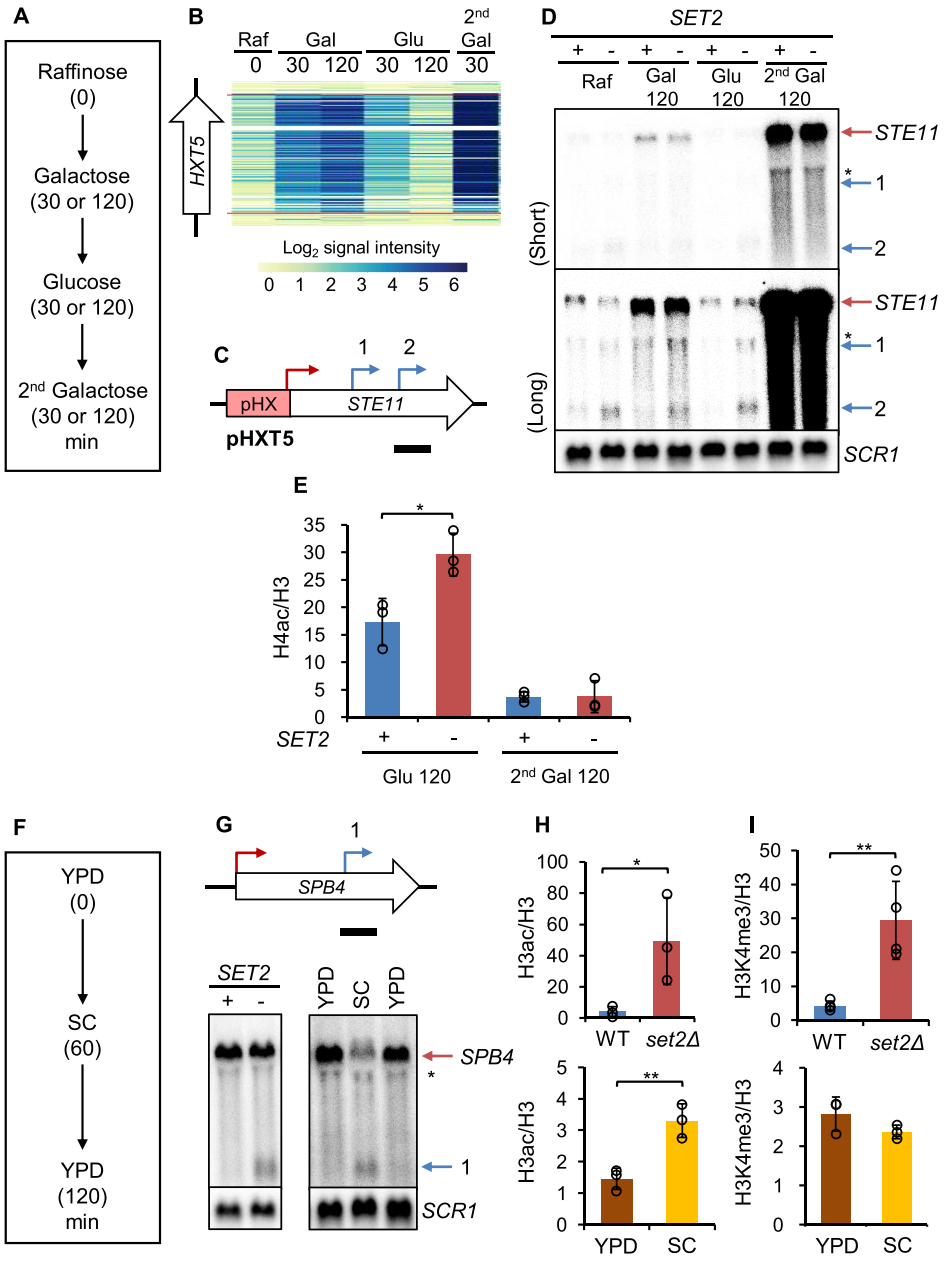

Figure 5. Dynamic regulation of core and internal cryptic promoters upon environmental shifts. (**A**) Schematic representation of the time course experiments to monitor changes in transcript levels during carbon source shifts. (**B**) Expression of *HXT5* in wild-type cells during carbon source shifts. Figure shows the changes in the microarray hybridization signals of the probes arrayed along the *HXT5* gene (28) as the carbon sources changed. Increased blue color indicates more transcription. Note that *HXT5* is hyperactivated upon the second galactose exposure. (**C**) Schematic representation of pHX-*STE11*. The core promoter of *STE11* was replaced with that of *HXT5*. A black bar underneath indicates the position of the probe used for northern blot analysis. (**D**) Northern blot analysis showing that in *set2Δ*, the first and especially the second galactose exposures activated mRNA expression of *STE11* but blocked its cryptic transcription. Two independent experiments showed the same results. (**E**) H4 acetylation patterns at the *STE11* cryptic promoter during carbon source shifts. Cross-linked chromatin from the indicated cells grown as in (A) was precipitated with anti-acetyl H4 or anti-H3 antibodies. PCR analysis of the precipitated DNA was done as in Figure 3B. Error bars show the standard deviation (S.D.) calculated from three biological replicates, each with three technical replicates. *$P < 0.05$ (two-tailed unpaired Student's *t* tests). (F–I) Effect of nutritional shifts on the mRNA and cryptic transcription of *SPB4* and histone modification at the cryptic promoter of *SPB4*. (**F**) Schematic representation of the time course experiments to monitor changes in transcript levels during nutritional shifts. The cells were grown to an OD_{600} of ~0.6 in YPD medium, shifted to SC for 60 min, and returned to YPD for 120 min. (**G**) Both loss of Set2 and nutritional shifting activated the internal cryptic promoter of *SPB4*. The upper panel shows a schematic representation of the core (red arrow) and internal cryptic (blue arow) promoters of *SPB4*. The bottom panels show the northern blot analysis of the *set2Δ* cells grown in YPD and the wild type cells before and after nutritional shifting. The full-length and cryptic transcripts of *SPB4* are indicated by red and blue arrows, respectively. Two independent experiments showed the same results. (**H**, **I**) Histone modification patterns at the internal cryptic promoter of *SPB4* in *set2Δ* cells grown in YPD and in wild type cells during the nutritional shift from YPD to SC. Cross-linked chromatin from the indicated cells was precipitated with anti-acetyl H3, anti-H3, and anti-H3K4me3 antibodies. PCR analysis of the precipitated DNA was performed as in Figure 3B. Error bars show the standard deviation (S.D.) calculated from three or four biological replicates, each with three technical replicates. *$P < 0.05$, and **$P < 0.01$ (two-tailed unpaired Student's *t* tests).

in *set2Δ* was increased during the glucose exposure but not during the second galactose exposure (Figure 5E). A converse case is the gene called *YLR419W* (28). It has at least three cryptic promoters (Supplementary Figure S5A) and, compared to *HXT5*, shows the opposite response to the carbon source shifts. In both wild-type and *set2Δ* cells, *YLR419W* mRNA was significantly upregulated by raffinose and glucose but downregulated by galactose. By contrast, the cryptic transcripts accumulated in *set2Δ* during the first and especially the second galactose exposures (Supplementary Figure S5B). This suggests that the dynamics of transcription from the core promoter directly affect internal cryptic transcription.

Some cryptic promoters can be induced by environmental changes in even wild-type cells (21,29). An example is *SPB4*, which produces a cryptic transcript in *set2Δ* cells when they are cultured in rich media (YPD). This cryptic transcript was also detected in wild-type cells when they were shifted from YPD to minimal media (SC) for one hour. Conversely, this shift significantly downregulated *SPB4* full-length mRNA levels. When the cells were returned to YPD for two hours, the original pattern was restored, upregulation of mRNA and downregulation of cryptic transcripts (Figure 5F and G). Analysis of the histone modifications in the latter experiment with *SPB4* showed that the loss of *SET2* and nutritional shifting of wild-type cells both associated with slightly increased histone H3 acetylation at the *SPB4* internal cryptic promoter (Figure 5H). Interestingly, however, while the H3K4me3 levels at the *SPB4* internal cryptic promoter were also highly increased in *set2Δ*, this was not observed in the wild-type cells after the nutritional shift (Figure 5I). No increase of H3K4me3 was also observed at internal cryptic promoters of two genes, *CHS6* and *FLO8* producing cryptic transcripts upon environmental shifts (Supplementary Figure S5C and D). Thus, the internal cryptic promoters that are transiently activated in wild-type cells by environmental shifts and those of constitutively active in mutants for transcription elongation factors may have distinct histone modification patterns.

DISCUSSION

Transcription elongation factors not only stimulate RNA Pol II elongation, but they also play important roles in the fidelity of transcription by repressing initiation from the internal cryptic promoters within gene bodies (Figure 6A). Mutants for the Set2-Rpd3S HDAC pathway, Spt6, and Spt16 have shown to activate large numbers of internal cryptic promoters (Figure 6B) (17,21,28,39). However, the commonalities of the genes that produce these cryptic transcripts have not been clearly defined. Here, we provided several lines of evidence showing that the degree of core promoter activity regulates the internal cryptic promoters in the gene body. First, most of the genes producing cryptic transcripts in mutants for the Set2–Rpd3S–HDAC pathway, Spt6, and Spt16 were infrequently transcribed (Figure 1B and C). Second, replacing the weak core promoters of *STE11*, *PCA1* and *FLO8* with the strong *YEF3* promoter directly blocked the transcription initiation from the internal cryptic promoters in mutants for the Set2-Rpd3S pathway and Spt6 (Figures 1 and 2). Third, if *STE11* was transcribed from the *YEF3* promoter rather than its original promoter, the Set2-Rpd3S pathway and Spt6 mutants failed to accumulate active histone marks, including H3 and H4 acetylation and H3K4me3, at its internal cryptic promoters (Figure 3). Lastly, when the core promoter activity was downregulated by environmental changes or by mutating transcription elongation factors, initiation from the internal cryptic promoters increased (Figures 4 and 5). These findings suggest that core promoter strength plays a critical role in the regulation of internal cryptic promoters and that the two promoters may communicate with each other to control their activities (Figure 6C and D).

An important question is, why does the internal cryptic promoter become inactivated if the core promoter is hyperactive? A simple explanation is transcriptional interference that the enhanced transcription from the core promoter by RNA Pol II blocks the formation of the preinitiation complex at the internal cryptic promoter. Although Sua7 occupancy was increased at the *STE11* cryptic promoter in mutant for Set2, but this was not seen when the *YEF3* promoter was present (Supplementary Figures S3A and S3B). This notion is further supported by previous studies showing somewhat the reverse situation that noncoding RNA or antisense RNA transcription can interfere with the binding of the transcription machinery to core promoters, thereby repressing mRNA transcription. For example, when yeast cells are grown in rich medium, the core promoter of the *SER3* gene is transcribed, thus generating the non-protein-coding RNA called *SRG1*. This transcription event directly represses downstream *SER3* expression from the promoter by blocking the binding of activators (40). However, when the *SRG1* TATA-box is mutated or positive regulators such as SWI/SNF are deleted, *SRG1* transcription stops and the repression on *SER3* transcription is lifted (40,41). Similarly, natural antisense transcription over coding gene promoters has been shown to inhibit the transcription of many genes, including quiescence-related genes, by hindering the association of the pre-initiation complex with the core promoters (42,43). Thus, hyperactive transcription from the core promoter may prevent the transcription machinery from binding to the internal cryptic promoter, thereby inhibiting cryptic transcription.

Although highly active genes tended in general to not produce cryptic transcripts in transcription elongation factor mutants, nonetheless approximately 5% did (6.3% in *spt6-1004*, 4.1% in *spt16-197* and 7.7% in *set2Δ*) (Figure 1C). Interestingly, the majority of the highly transcribed Set2 target genes had inducible cryptic promoters that were activated by galactose exposure. It was notable that the same conditions caused downregulation of the corresponding mRNA promoters (Figure 4B–D) (28). Similarly, many of the highly active genes that had Spt6- or Spt16-repressed cryptic promoters showed concomitant mRNA downregulation (Figure 4F–H). These observations are supported by a recent study showing that loss of Spt6 reduced transcription from most of the core promoters but elevated transcription at many intragenic promoters (39). Significantly, we showed when the core promoter of a gene that produces cryptic transcripts in the *spt6-1006* mutant was replaced with a strong core promoter, cryptic transcription was abolished (Figure 2D). These findings together suggest that in-

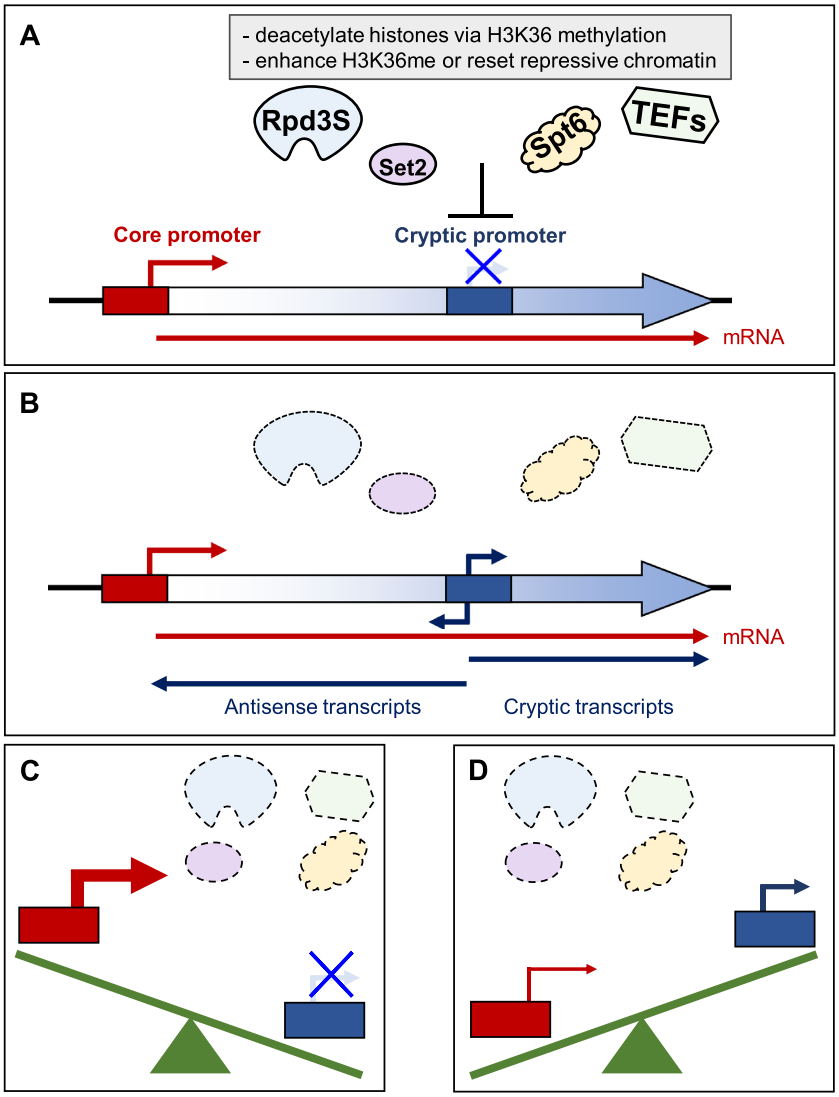

Figure 6. Models for regulation of internal cryptic promoters by core promoter activity. (**A**, **B**) In wild-type cells, transcription elongation factors such as the Set2-Rpd3S pathway and Spt6 negatively regulate the transcription from internal cryptic promoters. Set2 methylates K36 of histone H3 and the resulting methylation enhances Rpd3S HDAC-mediated histone deacetylation. Transcription elongation factors including Spt6 stimulate H3K36 methylation by Set2. When these factors are lost, the gene bodies acquire increased acetylation or abnormal chromatin structure, which activates the internal cryptic promoters. These promoters then produce short sense or antisense transcripts. (**C**) However, internal cryptic promoters can be repressed even in the absence of these transcription elongation factors if the core promoter is hyperactive. (**D**) The loss of transcription elongation factors, Spt6 and Spt16 downregulates the core promoter activity resulting in activation of the internal cryptic promoter. Environmental shifts also can reduce the core promoter activity resulting in transient activation of transcription from the internal cryptic promoter. The red and blue rectangles indicate the core promoter and the internal cryptic promoter, respectively.

ternal cryptic promoters and core promoters may crosstalk to repress each other's activity (Figure 6C and D).

Another interesting question is, why are the cryptic promoters in hyperactive genes generally insensitive to chromatin-mediated repression, such as that imposed by the Set2-Rpd3S pathway and Spt6? Nucleosome occupancy generally correlates negatively with transcription frequency: hyperactive genes have the lowest nucleosome occupancy in their gene bodies due to the rapid, continuous elongation by RNA Pol II. This suggests that the loss of transcription elongation factors will change the nucleosome occupancy of highly transcribed genes less profoundly than that of the inactive genes: this change is then probably too small to activate the internal cryptic promoters in the highly transcribed genes. This in turn suggests that the internal cryptic initiation of hyperactive genes that have very low nucleosome density is largely blocked by transcription interference. Notably, we unexpectedly observed that when the inactive core promoter of a gene is replaced with a hyperactive core promoter in Set2-Rpd3S pathway and Spt6 mutants, histone acetylation but not H3K4me3 at the internal cryptic promoter was slightly decreased (Figures 3 and 5E). It is unclear why these two active marks show these opposite, transcription frequency-dependent patterns. Nonetheless, it is possible that the reduced acetylation could inactivate the internal cryptic promoter. Indeed, we recently showed that Hda1C specifically deacetylates histone H4 at hyperactive coding regions (44,45). It should be noted that the high H4

acetylation levels in *hda1Δ* cells did not produce cryptic transcripts from *STE11* if its core promoter was replaced with the *YEF3* core promoter (data not shown). However, it remains possible that the increased H4 acetylation in *hda1Δ* cells is simply not sufficient to activate internal cryptic promoters because the H3 acetylation levels remain intact. This is supported by our previous studies showing that although loss of Hda1C increased H4 acetylation in the coding regions of the *PCA1* gene, it did not alter H3 acetylation and therefore no cryptic transcripts were produced (44,45).

These findings thus suggest that HdaC1 alone is not sufficient to regulate the internal cryptic promoters of hyperactive genes. Similarly, we were unable in the present study to identify other factors that could help regulate these internal cryptic promoters: none of the factors we tested, including Rtf1 and Spt2, affected the internal cryptic initiation of *STE11* when the core promoter was replaced with the *YEF3* promoter (Supplementary Figure S4A). It will be interesting to understand how the cryptic promoters within highly transcribed genes are regulated.

We observed that activation of internal cryptic promoters associated with potent accumulation of H3K4me3 at the cryptic promoter (Figure 3). However, the function of H3K4me3 in regulation of internal cryptic promoters remains unclear. H3K4me3 acts as a binding site for various factors that have distinct functions in transcription, including multiple HATs and HDACs. However, this specific role may not necessarily participate in H3K4me3-mediated regulation of internal cryptic promoter: we showed recently that while NuA3 HAT is required for the full activation of inducible cryptic promoters that are repressed by the Set2-Rpd3S HDAC pathway, this role does not require an interaction between its Yng1 PHD finger and H3K4me3 (19). Moreover, although H3K4me3 is believed to be an active histone mark, several studies suggest that it can also have a repressive role. For example, the binding of the Rpd3L HDAC to H3K4me3 *via* its Pho23 PHD finger plays an important role in transcriptional repression memory, which mediates the stronger and more rapid repression of unnecessary genes upon environmental changes (38,46,47). These suggest that H3K4me3 enriched at internal cryptic promoters may also function to repress cryptic transcription. Another interesting finding of this study was that while H3K4me3 accumulated at the constitutively active cryptic promoters in transcription elongation factor mutants, this accumulation was not seen at the internal cryptic promoters that were transiently induced by environmental changes (Figure 5I, Supplementary Figure S5C and 5D). It will be interesting to determine whether the latter reflects failure to deposit H3K4me3 or its rapid removal by histone demethylase.

While cryptic transcripts were initially believed to be noncoding RNAs, recent studies show that they can be translated into short proteins (21,26,48,49). Although their exact functions remain to be determined, some may function differently from their full-length counterparts. For example, when yeast cells are subjected to replication stress, the *ASE1* gene undergoes intragenic transcription that produces two short Ase1 proteins. These proteins stabilize the spindle by antagonizing the full-length protein (26). Another example is the N-terminal truncated Notch1 protein, which is produced by transcription from the alternative promoter of Notch1 and is involved in T cell development and the pathogenesis of T cell acute lymphoblastic leukemia (49). Further unravelling of the roles of internal cryptic promoters and their cryptic transcripts in protein diversity will be of great interest to the field.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Stephen Buratowski (Harvard Medical School) and Fred Winston (Harvard Medical School) for strains, and all members of the Kim lab for helpful advice and discussions.

FUNDING

National Research Foundation [NRF-2017M3A9B5060887, NRF-2017M3A9G7073033, NRF-2017M3C9A5029980, NRF-2019R1A5A6099645, NRF-2019R1A6C1010020 to T.K., NRF-2019R1A6A3A01094049 to B.B.L.]. Funding for open access charge: National Research Foundation.
Conflict of interest statement. None declared.

REFERENCES

1. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
2. Li,B., Carey,M. and Workman,J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707–719.
3. Woo,H., Dam Ha,S., Lee,S.B., Buratowski,S. and Kim,T. (2017) Modulation of gene expression dynamics by co-transcriptional histone methylations. *Exp. Mol. Med.*, **49**, e326.
4. Kizer,K.O., Phatnani,H.P., Shibata,Y., Hall,H., Greenleaf,A.L. and Strahl,B.D. (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol. Cell. Biol.*, **25**, 3305–3316.
5. Krogan,N.J., Kim,M., Tong,A., Golshani,A., Cagney,G., Canadien,V., Richards,D.P., Beattie,B.K., Emili,A., Boone,C. *et al.* (2003) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.*, **23**, 4207–4218.
6. Li,B., Howe,L., Anderson,S., Yates,J.R. 3rd and Workman,J.L. (2003) The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.*, **278**, 8897–8903.
7. Schaft,D., Roguev,A., Kotovic,K.M., Shevchenko,A., Sarov,M., Shevchenko,A., Neugebauer,K.M. and Stewart,A.F. (2003) The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. *Nucleic Acids Res.*, **31**, 2475–2482.
8. Barski,A., Cuddapah,S., Cui,K., Roh,T.Y., Schones,D.E., Wang,Z., Wei,G., Chepelev,I. and Zhao,K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823–837.
9. Bernstein,B.E., Kamal,M., Lindblad-Toh,K., Bekiranov,S., Bailey,D.K., Huebert,D.J., McMahon,S., Karlsson,E.K., Kulbokas,E.J. 3rd, Gingeras,T.R. *et al.* (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*, **120**, 169–181.
10. Liu,C.L., Kaplan,T., Kim,M., Buratowski,S., Schreiber,S.L., Friedman,N. and Rando,O.J. (2005) Single-nucleosome mapping of histone modifications in S. cerevisiae. *PLoS Biol.*, **3**, e328.
11. Pokholok,D.K., Harbison,C.T., Levine,S., Cole,M., Hannett,N.M., Lee,T.I., Bell,G.W., Walker,K., Rolfe,P.A., Herbolsheimer,E. *et al.* (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, **122**, 517–527.

12. Carrozza,M.J., Li,B., Florens,L., Suganuma,T., Swanson,S.K., Lee,K.K., Shia,W.J., Anderson,S., Yates,J., Washburn,M.P. *et al.* (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*, **123**, 581–592.
13. Joshi,A.A. and Struhl,K. (2005) Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol. Cell*, **20**, 971–978.
14. Keogh,M.C., Kurdistani,S.K., Morris,S.A., Ahn,S.H., Podolny,V., Collins,S.R., Schuldiner,M., Chin,K., Punna,T., Thompson,N.J. *et al.* (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell*, **123**, 593–605.
15. Li,B., Gogol,M., Carey,M., Lee,D., Seidel,C. and Workman,J.L. (2007) Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science*, **316**, 1050–1054.
16. Govind,C.K., Qiu,H., Ginsburg,D.S., Ruan,C., Hofmeyer,K., Hu,C., Swaminathan,V., Workman,J.L., Li,B. and Hinnebusch,A.G. (2010) Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes. *Mol. Cell*, **39**, 234–246.
17. Li,B., Gogol,M., Carey,M., Pattenden,S.G., Seidel,C. and Workman,J.L. (2007) Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev.*, **21**, 1422–1430.
18. Gilbert,T.M., McDaniel,S.L., Byrum,S.D., Cades,J.A., Dancy,B.C., Wade,H., Tackett,A.J., Strahl,B.D. and Taverna,S.D. (2014) 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*, **13**, 2883–2895.
19. Kim,J.H., Yoon,C.Y., Jun,Y., Lee,B.B., Lee,J.E., Ha,S.D., Woo,H., Choi,A., Lee,S., Jeong,W. *et al.* (2020) NuA3 HAT antagonizes the Rpd3S and Rpd3L HDACs to optimize mRNA and lncRNA expression dynamics. *Nucleic Acids Res.*, **48**, 10753–10767.
20. Shi,X., Kachirskaia,I., Walter,K.L., Kuo,J.H., Lake,A., Davrazou,F., Chan,S.M., Martin,D.G., Fingerman,I.M., Briggs,S.D. *et al.* (2007) Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J. Biol. Chem.*, **282**, 2450–2455.
21. Cheung,V., Chua,G., Batada,N.N., Landry,C.R., Michnick,S.W., Hughes,T.R. and Winston,F. (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. *PLoS Biol.*, **6**, e277.
22. Chu,Y., Simic,R., Warner,M.H., Arndt,K.M. and Prelich,G. (2007) Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. *EMBO J.*, **26**, 4646–4656.
23. Kaplan,C.D., Laprade,L. and Winston,F. (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science*, **301**, 1096–1099.
24. Lin,L.J., Minard,L.V., Johnston,G.C., Singer,R.A. and Schultz,M.C. (2010) Asf1 can promote trimethylation of H3 K36 by Set2. *Mol. Cell. Biol.*, **30**, 1116–1129.
25. Gopalakrishnan,R., Marr,S.K., Kingston,R.E. and Winston,F. (2019) A conserved genetic interaction between Spt6 and Set2 regulates H3K36 methylation. *Nucleic Acids Res.*, **47**, 3888–3903.
26. McKnight,K., Liu,H. and Wang,Y. (2014) Replicative stress induces intragenic transcription of the ASE1 gene that negatively regulates Ase1 activity. *Curr. Biol.*, **24**, 1101–1106.
27. McDaniel,S.L., Hepperla,A.J., Huang,J., Dronamraju,R., Adams,A.T., Kulkarni,V.G., Davis,I.J. and Strahl,B.D. (2017) H3K36 methylation regulates nutrient stress response in *Saccharomyces cerevisiae* by enforcing transcriptional fidelity. *Cell Rep.*, **19**, 2371–2382.
28. Kim,J.H., Lee,B.B., Oh,Y.M., Zhu,C., Steinmetz,L.M., Lee,Y., Kim,W.K., Lee,S.B., Buratowski,S. and Kim,T. (2016) Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. *Nat. Commun.*, **7**, 13534.
29. Kim,T., Xu,Z., Clauder-Munster,S., Steinmetz,L.M. and Buratowski,S. (2012) Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. *Cell*, **150**, 1158–1169.
30. Storici,F. and Resnick,M.A. (2006) The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. *Methods Enzymol.*, **409**, 329–345.
31. Marquardt,S., Hazelbaker,D.Z. and Buratowski,S. (2011) Distinct RNA degradation pathways and 3′ extensions of yeast non-coding RNA species. *Transcription*, **2**, 145–154.
32. Kim,T. and Buratowski,S. (2009) Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5′ transcribed regions. *Cell*, **137**, 259–272.
33. Mayer,A., Lidschreiber,M., Siebert,M., Leike,K., Soding,J. and Cramer,P. (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nat. Struct. Mol. Biol.*, **17**, 1272–1278.
34. Tan-Wong,S.M., Zaugg,J.B., Camblong,J., Xu,Z., Zhang,D.W., Mischo,H.E., Ansari,A.Z., Luscombe,N.M., Steinmetz,L.M. and Proudfoot,N.J. (2012) Gene loops enhance transcriptional directionality. *Science*, **338**, 671–675.
35. Xiao,T., Shibata,Y., Rao,B., Laribee,R.N., O'Rourke,R., Buck,M.J., Greenblatt,J.F., Krogan,N.J., Lieb,J.D. and Strahl,B.D. (2007) The RNA polymerase II kinase Ctk1 regulates positioning of a 5′ histone methylation boundary along genes. *Mol. Cell. Biol.*, **27**, 721–731.
36. Silva,A.C., Xu,X., Kim,H.S., Fillingham,J., Kislinger,T., Mennella,T.A. and Keogh,M.C. (2012) The replication-independent histone H3-H4 chaperones HIR, ASF1, and RTT106 co-operate to maintain promoter fidelity. *J. Biol. Chem.*, **287**, 1709–1718.
37. Mason,P.B. and Struhl,K. (2005) Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. *Mol. Cell*, **17**, 831–840.
38. Lee,B.B., Choi,A., Kim,J.H., Jun,Y., Woo,H., Ha,S.D., Yoon,C.Y., Hwang,J.T., Steinmetz,L., Buratowski,S. *et al.* (2018) Rpd3L HDAC links H3K4me3 to transcriptional repression memory. *Nucleic Acids Res.*, **46**, 8261–8274.
39. Doris,S.M., Chuang,J., Viktorovskaya,O., Murawska,M., Spatt,D., Churchman,L.S. and Winston,F. (2018) Spt6 is required for the fidelity of promoter selection. *Mol. Cell*, **72**, 687–699.
40. Martens,J.A., Laprade,L. and Winston,F. (2004) Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature*, **429**, 571–574.
41. Martens,J.A. and Winston,F. (2002) Evidence that Swi/Snf directly represses transcription in *S. cerevisiae*. *Genes Dev.*, **16**, 2231–2236.
42. Nevers,A., Doyen,A., Malabat,C., Neron,B., Kergrohen,T., Jacquier,A. and Badis,G. (2018) Antisense transcriptional interference mediates condition-specific gene repression in budding yeast. *Nucleic Acids Res.*, **46**, 6009–6025.
43. Gill,J.K., Maffioletti,A., Garcia-Molinero,V., Stutz,F. and Soudet,J. (2020) Fine chromatin-driven mechanism of transcription interference by antisense noncoding transcription. *Cell Rep.*, **31**, 107612.
44. Ha,S.D., Ham,S., Kim,M.Y., Kim,J.H., Jang,I., Lee,B.B., Lee,M.K., Hwang,J.T., Roh,T.Y. and Kim,T. (2019) Transcription-dependent targeting of Hda1C to hyperactive genes mediates H4-specific deacetylation in yeast. *Nat. Commun.*, **10**, 4270.
45. Lee,M.K. and Kim,T. (2020) Histone H4-specific deacetylation at active coding regions by Hda1C. *Mol. Cells*, **43**, 841–847.
46. Kim,M.Y., Lee,J.E., Kim,L.K. and Kim,T. (2019) Epigenetic memory in gene regulation and immune response. *BMB Rep*, **52**, 127–132.
47. Shi,X., Hong,T., Walter,K.L., Ewalt,M., Michishita,E., Hung,T., Carney,D., Pena,P., Lan,F., Kaadige,M.R. *et al.* (2006) ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature*, **442**, 96–99.
48. Wei,W., Hennig,B.P., Wang,J., Zhang,Y., Piazza,I., Pareja Sanchez,Y., Chabbert,C.D., Adjalley,S.H., Steinmetz,L.M. and Pelechano,V. (2019) Chromatin-sensitive cryptic promoters putatively drive expression of alternative protein isoforms in yeast. *Genome Res.*, **29**, 1974–1984.
49. Gomez-del Arco,P., Kashiwagi,M., Jackson,A.F., Naito,T., Zhang,J., Liu,F., Kee,B., Vooijs,M., Radtke,F., Redondo,J.M. *et al.* (2010) Alternative promoter usage at the Notch1 locus supports ligand-independent signaling in T cell development and leukemogenesis. *Immunity*, **33**, 685–698.