RSC-mediated nucleosome positioning and GRFs form barriers in promoters to limit divergent noncoding transcription

Andrew Wu[1], Claudia Vivori[1], Harshil Patel[2], Theodora Sideri[1] and Folkert J. van Werven[1]

[1]Cell Fate and Gene Regulation Laboratory, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK

[2]Bioinformatics and Biostatistics, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK

Corresponding author: folkert.vanwerven@crick.ac.uk

Abstract (200 words)

The directionality of gene promoters - the ratio of protein-coding over divergent noncoding transcription - is highly variable and regulated. How promoter directionality is controlled remains poorly understood. We show that the chromatin remodelling complex RSC and general regulatory factors (GRFs) dictate promoter directionality by attenuating divergent transcription. Depletion of RSC increased divergent noncoding transcription and decreased protein-coding transcription at promoters with strong directionality. Consistent with RSC's role in regulating chromatin, RSC depletion impacts nucleosome occupancy upstream of the nucleosome depleted region where divergent transcription initiates, suggesting that nucleosome positioning at the 5' border of gene promoters physically blocks the recruitment of the transcription machinery and inhibits initiation of divergent transcription. Highly directional promoters were also enriched for the binding of GRFs such as Reb1 and Abf1. Furthermore, ectopic targeting of divergent transcription initiation sites with GRFs or the dCas9 protein can suppress divergent transcription. Our data suggest that RSC-mediated nucleosome positioning and GRFs play a pervasive role in repressing divergent transcription. We propose that any DNA binding factor, when stably associated with cryptic transcription start sites, can form a barrier for repressing divergent transcription. Our study provides an explanation as to why certain promoters are more directional than others.

Introduction

Transcription is highly pervasive and results in intergenic and intragenic noncoding transcription events. Noncoding transcription and noncoding RNAs play diverse roles in gene and genome regulation (Ard et al., 2017; Gil and Ulitsky, 2020).

Transcriptionally active coding gene promoters are a major source of noncoding transcription. At these genomic locations, noncoding transcription initiates in the divergent direction, a process known as divergent or bidirectional transcription, which generates upstream transcripts in the opposing direction to the coding gene from a distinct core promoter (Neil et al., 2009; Seila et al., 2008; Sigova et al., 2013; Xu et al., 2009). Divergent noncoding transcription is an intrinsic property of coding gene transcription, present across eukaryotes, and widespread across actively transcribed regions in the genome (Seila et al., 2009). It is worth noting that divergent transcription differs from antisense transcription, which typically initiates at the 3' end of gene bodies and represses sense coding transcription (Pelechano and Steinmetz, 2013).

Our understanding of the function of divergent transcription is still incomplete. There is evidence that the noncoding transcripts emanating from divergent transcription events can help to promote gene transcription in the coding direction and to facilitate cell fate control (Frank et al., 2019; Luo et al., 2016; Yang et al., 2021). Others have suggested that divergent transcription may reflect transcriptional noise of active promoters (Seila et al., 2009; Struhl, 2007). Accordingly, promoters with high levels of divergent transcription (*i.e.* those with high bidirectionality and low directionality) likely represent a transcriptional ground state (Jin et al., 2017; Wu and Sharp, 2013).

Conversely, highly directional promoters (*i.e.* promoters with little divergent transcription but high levels of sense coding gene transcription) are likely more evolved.

Aberrant noncoding transcription, including divergent transcription, can negatively impact cell fitness. For example, induced noncoding transcription can cause R-loop formation and DNA damage (Nojima et al., 2018). In addition, aberrant divergent transcription events can affect coding gene transcription leading to mis-regulation of gene expression (Chiu et al., 2018; du Mee et al., 2018). This especially impacts species with gene-dense genomes such as *Saccharomyces cerevisiae*. Hence, there are various molecular mechanisms that either limit noncoding transcription or the accumulation of noncoding RNAs. Indeed, to reduce the accumulation of divergent RNAs, these noncoding transcripts are often rapidly degraded via the RNA degradation machinery (Flynn et al., 2011; Malabat et al., 2015; Neil et al., 2009; van Dijk et al., 2011; Xu et al., 2009). Divergent transcription units are also typically short, due to enrichment in transcription termination signals (Schulz et al., 2013). Additionally, divergent transcription can be repressed by chromatin remodellers and histone modifying enzymes (Gowthaman et al., 2021; Marquardt et al., 2014). Lastly, general regulatory factors (GRFs) can repress initiation of aberrant noncoding transcription events (Challal et al., 2018; Wu et al., 2018).

In previous work, we showed that the GRF and pioneer transcription factor Rap1 represses divergent transcription in the promoters of highly expressed ribosomal protein and metabolic genes (Challal et al., 2018; Wu et al., 2018). These promoters are highly directional, and Rap1 is key for both promoting transcription in the coding

direction and for limiting transcription in the divergent direction (Challal et al., 2018; Wu et al., 2018). We proposed that Rap1 limits divergent transcription by interfering with recruitment of the transcription machinery. In the same study, we also identified a role for RSC, a major chromatin remodelling complex important for chromatin organization (Cairns et al., 1996). We showed for two gene promoters that RSC promotes divergent transcription in Rap1 depleted cells. RSC activity increases nucleosome positioning, which, in turn, is important for maintaining nucleosome depleted regions (NDRs) in promoters (Badis et al., 2008; Hartley and Madhani, 2009; Lorch et al., 1998; Parnell et al., 2008). RSC also acts with GRFs to stimulate transcription in the protein-coding direction (Brahma and Henikoff, 2019; Kubik et al., 2017; Kubik et al., 2018). Specifically, RSC positions the +1 nucleosome in gene promoters to allow pre-initiation complex (PIC) assembly and transcription start site (TSS) scanning (Klein-Brill et al., 2019). However, the role of RSC in controlling promoter directionality through divergent transcription remains to be elucidated.

Here, we examined the role of RSC and GRFs in controlling promoter directionality. Our analysis revealed that RSC depletion leads to increased divergent transcription at many promoters. These promoters tend to be highly directional and are enriched for GRFs such as Abf1 and Reb1. Consistent with the role of RSC in chromatin organization, its depletion affects nucleosome positioning in upstream regions of directional promoters. Finally, we demonstrate that ectopic targeting of GRFs and dCas9 to divergent core promoters can also repress divergent noncoding transcription. We propose that nucleosomes positioned by RSC activity and GRFs constitute physical barriers in promoters which limit aberrant divergent transcription, thereby increasing promoter directionality.

Results

RSC represses divergent transcription independently of Rap1

In *Saccharomyces cerevisiae*, the chromatin remodelling complex RSC acts at most gene promoters, including promoters by regulated by GRFs such as Reb1, Abf1 and Rap1 (Brahma and Henikoff, 2019; Klein-Brill et al., 2019; Kubik et al., 2015; Kubik et al., 2018). Previously, we demonstrated that RSC and Rap1 control divergent transcription in opposing ways (Wu et al., 2018). While Rap1 depletion induced divergent transcription at the *RPL43B* and *RPL40B* promoters, co-depletion of RSC and Rap1 partially reduced divergent RNA expression when compared to Rap1 depleted cells.

To investigate how RSC and GRFs promote directionality more closely, we performed RNA-seq in *S. cerevisiae* cells after depletion of both RSC and Rap1. Since RSC affects gene transcription genome-wide (Klein-Brill et al., 2019; Parnell et al., 2008), we added a spike-in control of *S. pombe* cells to all *S. cerevisiae* samples in a defined ratio, and used the *S. pombe* RNA-seq signal to normalize our RNA-seq data (Figure 1A, bottom panel). We also performed the “standard” internal normalization procedure according to DESeq size factors. Finally, to deplete RSC and Rap1 we used auxin-inducible degron alleles (*RAP1-AID* and *STH1-AID*) and treated cells with indole-3-acetic acid (IAA) (Nishimura et al., 2009). Rap1 and Sth1 were efficiently depleted in cells harbouring either one or both degron alleles after treatment (Figure 1A, top panel).

First, we examined how regulated protein-coding transcripts were affected by Rap1 and RSC depletion. We focused on a set of 141 Rap1-regulated gene promoters described previously (Wu et al., 2018). As expected, almost all Rap1-regulated genes decreased in expression after Rap1 depletion (*RAP1-AID* + IAA vs *RAP1-AID* + DMSO) (Figure 1B and S1A, central panel). Expression of Rap1-regulated genes also decreased in Sth1-depleted cells, when the signals were normalized using *S. pombe* spike-in controls (*sp*, compared to normalization over *S. cerevisiae*, *sc*), confirming that RSC has genome-wide impact on transcription (Figure 1B and S1A, left and right side). Second, we examined two previously well-characterized promoters (*RPL43B* and *RPL40B*) at which Rap1 is known to repress divergent transcription (Figures 1C and S1B). As expected, expression of divergent noncoding transcripts (*IRT2* and *iMLP1*) was upregulated upon Rap1 depletion, while coding transcription was decreased. Upon co-depletion of Rap1 and Sth1, we observed that *IRT2* and *iMLP1* levels were reduced when signals were normalized over *sp*. While Sth1 depletion by itself did not affect *IRT2* levels, levels of *iMLP1* were already upregulated in Sth1-depleted cells (Figure S1B). This suggests that RSC may repress divergent transcription at this locus. Indeed, other loci also displayed increased divergent transcription upon depletion of Sth1 (Figure 1D). For example, the *RPS10* and *RPS3* genes showed increased divergent transcription in both Rap1- and Sth1-depleted cells. Interestingly, the divergent transcription in Sth1-depleted cells initiated further upstream to the coding gene than in Rap1-depleted cells, suggesting that Rap1 and RSC limit divergent transcription through different mechanisms (Figure 1D).

Next, we examined Rap1-regulated gene promoters more closely for changes in divergent and sense expression upon depletion of Sth1 and Rap1 (Figure 1E, S2A and S2B). As expected, Rap1-regulated gene promoters displayed increased levels of divergent transcripts upon Rap1 depletion (Figures 1E, 1F and S2B). In Sth1-depleted cells we also observed an increase in divergent RNAs using either *S. cerevisiae* and *S. pombe* normalization (Figures 1E, 1F, S2B and S2C). The increase in divergent RNAs was higher when Sth1 depleted cells (*STH1-AID* + *IAA)* were compared to wild-type (WT) cells instead of the mock treated control (*STH1-AID* + DMSO), suggesting the AID tag on Sth1 has a small negative impact on RSC activity in the cells (Figure S2D). The changes in divergent RNA expression after Sth1 and Rap1 co-depletion were comparable to Rap1 depletion, suggesting that RSC does not promote or repress divergent transcription at Rap1-regulated genes after Rap1 depletion. We also observed increased RNA levels downstream of the Rap1 site after co-depletion of RSC and Rap1, which likely reflects increased antisense transcription which initiates further downstream within gene bodies, particularly from 3' ends of genes (Alcid and Tsukiyama, 2014). We also observed increased RNA expression in the protein-coding direction in Rap1-depleted and Rap1/Sth1 co-depleted cells (Figures 2E, 2F, S2A, S2B and S2C). Indeed, it is known that Rap1 depletion can lead to inappropriate transcription initiation in both the protein-coding and divergent directions (Challal et al., 2018).

RSC plays a prominent role in positioning nucleosomes in promoters, which in turn affects PIC recruitment and TSS scanning (Hartley and Madhani, 2009; Klein-Brill et al., 2019). We assessed the nucleosome occupancy pattern at Rap1-regulated promoters using a published MNase-seq dataset (Figure S3) (Kubik et al., 2015;

Kubik et al., 2018). In Sth1-depleted cells, there was a small shift in upstream nucleosome positioning towards to the Rap1 binding site, while the size and depth of the NDRs remained relatively unchanged (Figure S3, left). The subtle shift observed in the absence of RSC may reflect less stable upstream nucleosomes, allowing divergent PIC formation and thus initiation of divergent transcription upstream of Rap1 binding sites (Figure 1D). This model for RSC role in controlling divergent transcription is further explored in Figure 4. In contrast, Rap1 depletion caused a large shift of upstream nucleosomes towards the Rap1 binding site for clusters where divergent transcription increased (Figure S3C, middle) (Wu et al., 2018). A similar shift was observed in Rap1/Sth1 co-depleted cells. However, the NDR was more reduced and consensus nucleosome positioning was less clear in these cells (Figure S3C, right). The different effects of Rap1 and RSC depletion on nucleosome positioning may explain the differences in divergent TSS usage as observed at the *RPS3* and *RPS10* loci (Figure 1D). We conclude that RSC represses the expression of divergent RNAs likely independently of Rap1.

RSC controls promoter directionality by repressing divergent transcription

Bulk RNA measurements are not an accurate representation of levels of transcription. Hence, as a proxy for transcriptional activity, we determined the levels of nascent transcribed RNA by measuring RNA Pol II-associated transcripts using an adapted native elongating transcript sequencing protocol (Churchman and Weissman, 2011). In short, we affinity purified RNA Pol II using Rpb3-FLAG and quantified the Pol II-associated RNAs in wild-type cells, and after depletion of RSC and/or Rap1 (Figure S4A). First, we assessed the changes in transcription level around Rap1 binding sites (Figure 2A). In line with bulk RNA measurements, we

found that divergent transcription was increased in Sth1-depleted cells. Likewise, Rap1 depletion and Rap1/Sth1 depletion led to an increase in divergent transcription (Figure 2A).

We next assessed how promoter directionality is controlled by Rap1 and RSC. Promoter directionality can be calculated by taking the logarithm of the ratio between transcription in the sense direction over the divergent direction (Figure 2B) (Jin et al., 2017). We computed the directionality score for all protein-coding gene promoters, based on the experimentally-determined TSS wherever possible (Park et al., 2014). We focused our analysis on a set of tandem, non-overlapping gene promoters where the upstream gene is positioned in the tandem orientation (n=2609 promoters), which excludes all bi-directional gene promoters (S4B and S4C). As expected, Rap1-regulated genes are highly directional when compared to all tandem genes (Figure 2C). Upon Rap1 depletion, directionality was strongly reduced due to increased divergent transcription and decreased sense transcription (Figures 2D and S4D). Upon Sth1 depletion, the promoter directionality scores were reduced. Indeed, Sth1-depleted cells displayed a relative increase in divergent transcription and decrease in sense transcription (Figures 2E and S4D). The combination of Rap1 and Sth1 depletion rescued the effect of Rap1 depletion, in part (Figure 2F, 2G, and S4D). This suggests that RSC can have opposing effects on promoter directionality at certain genes. At highly directional Rap1-regulated genes, RSC ensures promoter directionality when Rap1 is present by repressing divergent transcription. However, when promoter directionality was instead strongly reduced due to Rap1 depletion, RSC activity maintains transcription in the divergent direction. This is consistent with

the findings for the divergent noncoding RNA *IRT2* expressed from the *RPL43B* promoter (Figure 1C).

Our analyses of Rap1-regulated gene promoters suggest that RSC normally contributes to repressing divergent transcription. This prompted to us to investigate the role of RSC in controlling promoter directionality genome wide. We identified gene promoters (*STE24* and *HIR3*) with increased divergent transcription upon RSC depletion (Figure 3A). We also assessed the chromatin structure at these promoters by examining published MNase-seq data (Kubik et al., 2015; Kubik et al., 2018). The MNase-seq signal for the nucleosome positioned upstream of the NDR in these promoters, also known as the -1 nucleosome (+1 nucleosome in the divergent direction), was decreased at both promoters, suggesting that RSC-mediated positioning of promoter nucleosomes is important for repressing divergent transcription (Figure 3A, further discussed in Figure 4).

Noncoding transcription is an inherent feature of gene regulation across species (Jacquier, 2009; Pelechano and Steinmetz, 2013; Seila et al., 2009). Accordingly, many classes of noncoding transcripts have been identified. These include: cryptic unstable transcripts (CUTs) which are processed by exome complex, Xrn1-sensitive antisense non-coding RNA (XUTs), Nrd1-unterminated transcripts (NUTs) and stable unannotated transcripts (SUTs) (Neil et al., 2009; Schulz et al., 2013; van Dijk et al., 2011; Xu et al., 2009). These noncoding transcripts (CUTs/SUTs/XUTs/NUTs) are often found at gene promoters and expressed in the divergent direction, but are also found at antisense transcription units which overlap protein-coding genes. We

examined how expression levels of CUTs/SUTs/XUTs/NUTs were affected upon RSC depletion. While protein-coding gene transcription was relatively decreased, noncoding transcription (CUTs/SUTs/XUTs/NUTs) was up-regulated (Figures 3B). Indeed, RSC depletion led to increased divergent transcription upstream of a large fraction of genes (Figure 3C). We found that more than 900 gene promoters displayed a 1.5 fold or more increase in divergent transcription, out of 2609 tandem and non-overlapping genes (Figure S5A). We also observed increased antisense transcription over gene bodies for a large group of genes (ASc2), which is line with previous findings (Figure 3C) (Alcid and Tsukiyama, 2014). We also computed the directionality score for all tandem genes of Sth1-depleted cells. There was a clear reduction in promoter directionality upon Sth1 depletion (S*TH1-AID* + IAA) compared to wild-type or control cells (*STH1-AID* +DMSO) (Figure 3D). Notably, control cells showed a decreased directionality score compared to the wild type, indicating that the AID tag on Sth1 had a small negative impact RSC function (Figure 3D). The changes in directionality score were due to increased transcription in the divergent direction and decreased transcription in the coding or ‘sense’ direction (Figures 3E, S5A and S5B). Taken together, these data suggest that RSC is important for repressing divergent noncoding transcription and thereby controls promoter directionality of a large set of gene promoters in yeast.

RSC mediated nucleosome organization limits TBP recruitment upstream of directional promoters

Next, we examined whether there are features that could explain the effect of RSC on promoter directionality. We reasoned that RSC could specifically drive

directionality by only promoting transcription in the coding direction. Indeed, it is known that RSC acts on highly transcribed gene promoters and coding sequences (Biernat et al., 2021; Kubik et al., 2015; Rawal et al., 2018). Alternatively, it is possible that RSC limits divergent transcription at gene promoters to control directionality. Our data indicate that both divergent and sense transcription levels change upon Sth1 depletion, indicating that control of promoter directionality by RSC is, at least in part, mediated through repression of divergent transcription (Figure 3E). We divided gene promoters into quintiles (Q1-Q5) based on transcription levels in sense direction or on directionality score, wherein Q5 represents the group of promoters with the highest sense transcription or directionality score and Q1 the lowest. Subsequently, we computed the changes in divergent and sense direction transcription, as well as changes in directionality score for each quintile (Figure 4A and S6). Since promoter directionality and sense transcription levels are linked, there was about fifty percent overlap between the promoters in Q5 for sense transcription and Q5 for directionality score. Genes important for rRNA processing and transcription were highly enriched in both of these overlapping sets of most highly transcribed (Q5) and most directional promoters (Q5) (Table S1), which perhaps is not surprising because ribosomal gene promoters are highly transcribed and display little divergent transcription (Wu et al., 2018). Gene promoters with the highest directionality score in wild-type cells (Q5) showed the largest relative increase in divergent transcription after Sth1 depletion (Figure 4A, left panel). The increase in divergent transcription of promoters in Q5 for directionality score was larger when compared to gene promoters with the highest sense transcription levels (Q5, sense transcription), suggesting that RSC-mediated repression of divergent transcription is more prominent at highly directional promoters (Figures 4A, 4B, and S6A). Control

cells (*STH1-AID* +DMSO) displayed a marginal increase in divergent transcription in Q5 (highest sense transcription levels and most directional promoters) compared to wild-type cells, which consistent with the finding that the AID tag partially affects RSC activity and had a small negative impact on promoter directionality (Figures 3D, S6D and S6E). Overall, these analyses suggest that RSC represses divergent transcription at promoters that are highly directional.

To investigate the mechanisms underlying these effects of RSC on promoter directionality, we assessed the chromatin structure at the different groups of promoters in Sth1 depleted cells (Figure 4C). Specifically, we compared profiles of MNase protection in cells treated with high and low amounts of MNase using a published dataset (Kubik et al., 2015). The extraction of nucleosome particles using different MNase concentrations identified the presence of so-called fragile nucleosomes (Kubik et al., 2015)}(Brahma and Henikoff, 2019; Kubik et al., 2018). These nucleosome particles appear when extracts are treated with low concentrations of MNase and likely represent partially unwrapped nucleosomes, but also have been suggested represent non-histone proteins (Brahma and Henikoff, 2019; Chereji et al., 2017; Kubik et al., 2015; Kubik et al., 2018). First, we find that the group of gene promoters with the highest directionality score (Q5) displayed more defined nucleosome peaks which are also narrower in width at the +1, -1 and -2 position compared to gene promoters with the lowest directionality score which are more broad (Q1) (Figure 4C, left panel). Second, upon depletion of Sth1 from the nucleus, nucleosome positioning becomes broader and less defined for the -1 and -2 positions. Third, the +1, -1 and -2 nucleosomes more are sensitive to MNase concentration for Q5 promoters compared to Q1 (Figure 4C, right panel). Therefore,

the nucleosome at the -1 position can be considered, at least to some extent, a fragile nucleosome particle. The sensitivity to MNase of the nucleosomes at the -2 position suggests that these nucleosome particles are also partly unwrapped and thus represent fragile nucleosomes. Similar trends, albeit less strong, were observed for gene promoters sorted by transcription levels in the sense direction (Q5 versus Q1) (Figure S7A). We also compared the effect of promoter directionality on a subset of promoters with comparable sense coding gene transcription levels (Figure 4D), to reduce the potential influence of coding gene transcription on nucleosome positioning. This analysis revealed that the upstream -1 nucleosome is more defined and sensitive to MNase concentrations for the promoters with the highest directionality score (highest 20%) compared to group of promoters in lowest 20% directionality score group within this subset of promoters with comparable coding gene transcription, further supporting the idea that positioning of upstream nucleosomes mediates promoter directionality (Figure 4E). We also observed the RSC-dependent differences in chromatin structure at individual genes. For example, the *PDE2* and *YBR238C* promoters display increased divergent transcription and reduced MNase-seq signal for the -1 nucleosomes in Sth1 depleted cells (Figure 4F). Notably, the *YBR238C* promoter also harbours a fragile nucleosome (Figure 4F, labelled “FN”).

The increase in divergent transcription, as observed in Sth1 depleted cells, is likely a consequence of increased PIC formation at divergent promoters. We assessed whether association of TATA-binding protein (TBP) was increased upon Sth1 depletion using a published ChIP-seq dataset (Kubik et al., 2018), as a proxy for PIC assembly. Despite that the ChIP-seq data does not have high spatial resolution, we

found that upon Sth1 depletion, TBP was more enriched in the region approximately 400 bp upstream of the coding gene TSS in the group of gene promoters with highest directionality (Q5) (Figure 4G, right panel), compared to the group of gene promoters with the lowest directionality score (Q1) (Figure 4G, left panel).

Finally, we examined whether DNA sequence motifs associated with RSC activity are differentially enriched and or distributed between promoters with low and high directionality. Both A stretches and the CG(C/G)G motif are associated with RSC binding (Badis et al., 2008; Krietenstein et al., 2016; Kubik et al., 2018; Lorch et al., 2014). We examined the distribution of these sequences in both the sense and divergent direction (*i.e.* on opposite DNA strands) (Figure S7B). We found that A stretches and the CG(C/G)G motif were enriched in the divergent direction in the group of promoters with the highest directionality (Q5) compared to the group of promoters with lowest directionality (Q1) (Table S2). Interestingly, the A track on the divergent strand showed two peaks in promoter with highest directionality, indicating that localized RSC activity likely facilitates the positioning of FNs and -1 nucleosomes at highly directional promoters (Figure 4H and S7B). These data further support a model where RSC is recruited to multiple, distinct positions to promote positioning of upstream nucleosomes in promoters, thereby limiting divergent transcription initiation. The highly positioned nucleosomes, in turn, can form physical barriers which inhibit recruitment of the PIC and RNA Pol II. Taken together, these data suggest that RSC mediated positioning of nucleosomes upstream in promoters plays direct role in limiting initiation of divergent transcription.

RSC and GRFs are enriched at highly directional promoters

RSC and GRFs are key players for controlling promoter architecture and regulating gene transcription (Brahma and Henikoff, 2019; Kubik et al., 2017; Kubik et al., 2018). We assessed how RSC and the GRFs Abf1 and Reb1 associate with groups of genes with the lowest or highest promoter directionality (Q1 vs Q5, Figure 4A) using a published dataset (Brahma and Henikoff, 2019). In line with role of RSC in controlling divergent transcription, we found that RSC was more enriched in highly directional gene promoters (compare Q5 to Q1, Figure 5A and S8). Moreover, we found that Abf1 and Reb1 were also more enriched for the group of promoters with the highest directionality score (Figure 5A and S8). This suggests that, apart from Rap1, other GRFs may repress divergent transcription. To further investigate this hypothesis, we examined a published dataset that measured RNA crosslinked to RNA Pol II (RNA Pol II CRAC) upon Reb1 depletion (Candelli et al., 2018b; Zentner et al., 2015). We found that depletion of Reb1 increased levels of divergent RNA associated with Pol II at the *BAP2* and *CSG2* promoters (Figure 5B).

Targeting GRFs or dCas9 to divergent core promoters is sufficient to repress divergent noncoding transcription.

Our analysis suggests that Abf1 and Reb1 can repress divergent transcription, possibly using a similar mechanism as Rap1 (Figure 5A). As we previously demonstrated, Rap1 limits initiation of divergent noncoding transcription by occupying target motifs nearby or adjacent to cryptic divergent core promoters (Wu et al., 2018). To investigate whether Reb1, Abf1, and transcription factors (TFs) including GRFs can repress divergent transcription, we introduced ectopic

transcription factor binding sites adjacent to a divergent core promoter (Marquardt et al., 2014; Wu et al., 2018). We used an established fluorescent reporter assay based on the *PPT1* promoter which normally has a noncoding transcript *SUT129* in the divergent direction, with both the coding and the divergent noncoding core promoters driving the transcription of fluorescent reporter genes encoding mCherry and YFP, respectively. We cloned the binding site sequences of several transcription factors 20 base pairs (bp) upstream of the *SUT129*/YFP TSS in the *PPT1* reporter (Figure 6A). We selected GRFs (Cbf1, Abf1, and Reb1) and TFs (Gal4, Gcn4, Cat8, Gcr1) that have well-defined DNA binding sequence motifs. To establish that the changes in the reporter signal was dependent on the transcription factor's presence in the cells and not due to the alteration in the underlying promoter DNA sequence, we also measured the reporter signal after deleting or depleting the same transcription factors. For example, for the reporter construct with Reb1 binding sites we measured YFP/mCherry levels in Reb1 control and depleted cells using the auxin induced degron (*REB1-AID* +IAA), whereas for the reporter construct with Cbf1 binding sites we measured the YFP/mCherry levels in wild-type and *cbf1*Δ cells.

Our data indicate that several GRFs were able to repress transcription in the divergent direction and promote promoter directionality, suggesting that they behave similarly to Rap1 (Figure 6B) (Wu et al., 2018). Specifically, Cbf1, Abf1 and Reb1 showed a reduced signal in the noncoding transcription in response to protein depletion (Figure 6B). Introduction of binding site motifs for one transcription factor, Gcr1, did not affect noncoding (YFP) direction relative to the wild-type control (wt), but the YFP signal increased in *gcr1*Δ cells suggesting that Gcr1 also can repress divergent transcription (Figure 6A and 6B). The TFs that did not negatively affect the

noncoding transcript signal (Gal4, Gcn4, and Cat8) were possibly less active under the growth conditions where we performed the reporter assay (Figure 6B). Notably, Gal4 and Gcn4 showed increased signal in the coding direction, suggesting that they are still somewhat active in the conditions tested (Figure S9A). We conclude that, as we reported previously for Rap1 (Wu et al., 2018), other transcription factors can repress divergent transcription when targeted to regions near a divergent gene promoter.

Our data indicate that RSC-mediated nucleosome positioning and GRFs can repress initiation of divergent transcription, and thereby promote directionality of gene promoters (Challal et al., 2018; Wu et al., 2018). One explanation is that these proteins are barriers for PIC formation and Pol II recruitment. If true, then any protein stably associated with DNA near divergent TSSs should be able to interfere with divergent transcription. To test this, we targeted an exogenous protein to a divergent TSS. We used the catalytically inactivated version of *S. pyogenes* Cas9 (dCas9), which is a bulky protein (over 150 kDa) which has been widely used as tool to modulate transcription and chromatin (Gilbert et al., 2013). We induced divergent transcription by depleting Rap1 and used gRNAs to target dCas9 to the core promoter of *IRT2*, which is divergently transcribed from *RPL43B* and strongly induced upon Rap1 depletion, Figures 6C-E). As expected, Rap1 depletion induced expression of *IRT2* (Figure 6D and 6E). Importantly, we did not detect *IRT2* expression when dCas9 was targeted to the *IRT2* TSS (Figure 6F and 6G). Expression of *IRT2* was still detectable in cells where dCas9 was targeted to a control locus (Figure 6G). We obtained a similar result with targeting dCas9 fused mammalian transcriptional repressor domain Mxi1 (dCas9-Mxi1) with gRNAs

targeting *IRT2*, but not with two control gRNAs (Figure S9C and S9D)(Gilbert et al., 2013). We conclude that steric hindrance by DNA-binding proteins (such as dCas9) is sufficient for repression of divergent transcription.

Discussion

Transcription from promoters is bidirectional, leading to sense coding direction and divergent noncoding transcription. However, the directionality (sense over divergent transcription) of promoters greatly varies indicating that there are mechanisms in place to control directionality of promoters (Jin et al., 2017). Why some promoters are more directional than others is not well understood. Here, we showed that RSC mediated nucleosome positioning represses divergent noncoding transcription. Furthermore, we provided evidence that a wide range of DNA-binding proteins, including nucleosomes, GRFs, and dCas9, are capable of repressing divergent transcription. We propose that such RSC-positioned nucleosomes and other DNA-binding factors form physical barriers at promoters which prevent transcription in the divergent direction, thereby promoting promoter directionality.

RSC mediated nucleosome positioning represses divergent transcription

The RSC complex largely contributes to generating NDRs by ‘pushing’ the nucleosomes positioned at +1 and -1 positions ‘outwards’, widening the NDR (Ganguli et al., 2014; Hartley and Madhani, 2009; Parnell et al., 2008). RSC also binds directly to DNA via A-track sequence motifs (Krietenstein et al., 2016; Kubik et al., 2018; Lorch et al., 2014). How does RSC mediate the position of the -1 nucleosome to repress divergent transcription? Interestingly, we found that A-tracks

were abundant in promoters on both strands indicating that directional positioning of nucleosomes upstream in promoters may contribute to preventing divergent transcription. RSC also promotes the partial unwinding of nucleosomes, which has been also referred to as “fragile” nucleosomes as these nucleosome particles are sensitive to MNase concentration (Brahma and Henikoff, 2019; Kubik et al., 2018). These sub-nucleosome complexes are also bound by GRFs. We found that at directional promoters the -1 nucleosome and FN are more defined and sensitive to MNase concentration compared to promoters with less directionality, suggesting that the formation of sub-nucleosome complexes controls promoter directionality (Figure 6H). We further found that depletion of RSC resulted in less defined consensus nucleosome patterns and reduced sensitivity to MNase. We propose that nucleosomes strongly positioned by RSC inhibit PIC formation and Pol II recruitment for the divergent direction (Figure 6C). Upon RSC depletion, the positioning of the -1 nucleosome and FN is less strong and PIC formation can occur. Possibly, the fragile nature of the -1 and FN nucleosome increases the propensity for PIC assembly and Pol II recruitment relative to a non-fragile nucleosome, because it may be already partially unwrapped and therefore it is easier for TBP and other initial PIC components to bind and form the PIC.

RSC controlled opening of chromatin facilitates the recruitment of sequence-specific TFs to promoters (Floer et al., 2010). It may be possible that GRFs or other non-histone proteins mediate the repression of divergent transcription via RSC, as has been proposed by others (Brahma and Henikoff, 2019; Chereji et al., 2017). Our analysis identified examples of divergent transcription events in the same promoter which were repressed by Rap1 and RSC, but utilised distinct TSSs, suggesting that

RSC can repress divergent transcription independently of GRFs (Figures 1D and 6H). How other GRFs act with RSC in controlling divergent transcription remains to be dissected further. Based on our analysis, it appears more likely that RSC and GRFs act together to repress divergent transcription and promote transcription in the protein-coding direction.

Model for promoter directionality

Nucleosomes form barriers for transcription. In the context of repressing divergent transcription, chromatin assembly factor I (CAF-I) plays a widespread role in yeast, demonstrating that chromatin assembly upstream in promoters is important in repressing divergent transcription (Marquardt et al., 2014). On the other hand, opening of chromatin in promoters increases divergent transcription. For example, increased histone lysine 56 acetylation (H3K56ac) leads to more divergent transcription in yeast (Marquardt et al., 2014). Conversely, deacetylation of histone H3 by Hda1 deacetylase complex (Hda1C) facilitates repression of divergent transcription (Gowthaman et al., 2021). The role of RSC at promoters is likely different from CAF-I, H3K56ac, and Hda1C. We propose that RSC-mediated positioning of NDR-flanking nucleosomes in promoters generates barriers for recruiting RNA Pol II and initiation of divergent transcription. In line with this, differential TSS usage has been observed in RSC-depleted cells (Kubik et al., 2019).

In previous work we showed that Rap1 can repress divergent transcription at gene promoters likely by occupying cryptic divergent promoters (Wu et al., 2018). Our new analysis suggests that other GRFs have potentially the same ability in a model

system. Moreover, we found that Abf1 and Reb1 were more enriched at highly directional promoters, suggesting that these GRFs also perform this role *in vivo*. dCas9 has the same capacity to repress aberrant noncoding transcription when targeted to divergent core promoters. We propose that proteins, physically interfere with divergent transcription when bound to DNA. With this view, RSC activity is important to position crucial nucleosomes and ensure that cryptic divergent promoters are ‘protected’ and transcriptionally repressed. Thus, GRFs and positioned nucleosomes constitute essential components for promoter organization, which promote sense coding transcription and limit divergent noncoding transcription (Figure 6H).

RSC is part of the SWI/SNF family of chromatin remodellers and related to the mammalian PBAF and BAF complexes (Mohrmann and Verrijzer, 2005). The BAF complex in mouse embryonic stem cells (esBAF), has related functions in repressing noncoding transcription to some extent (Hainer et al., 2015). Like RSC, esBAF positions the nucleosomes flanking the NDR. Depletion of esBAF leads to less stable nucleosome positioning, and consequently widespread increases in noncoding transcription. Some of this resultant increased noncoding transcription occurs at promoters, resulting in divergent noncoding transcription events. We anticipate that the role of RSC and BAF complexes in controlling divergent transcription and promoter directionality is likely conserved.

Further work will be required to dissect the interplay between GRFs and chromatin states in control of aberrant noncoding transcription and promoter directionality.

These regulatory complexes are present and conserved across eukaryotes. Thus, our study may provide insight how promoter directionality controlled in multi-cellular organisms.

Materials and Methods

Strains and plasmids and growth conditions

Strains isogenic to the Saccharomyces cerevisiae BY4741 strain background (derived from S288C) were used for this investigation. A one-step tagging procedure was used to generate endogenous carboxy (C)-terminally tagged auxin-inducible degron (AID) alleles of Rap1 and Sth1 (Nishimura et al., 2009). The *RAP1-AID* strain was previously described, which contains three copies of the V5 epitope and the IAA7 degron (Wu et al., 2018). For tagging *RPB3* with the FLAG epitope, we used a one-step tagging procedure using a plasmid (3xFLAG-pTEF-NATMX-tADH1, gift from Jesper Svejstrup). The single copy integration plasmids for expression of *Oryza sativa TIR1* (*osTIR1*) ubiquitin E3 ligase were described previously (Wu et al., 2018).

For auxin-inducible degron (*AID*) depletion experiments, induction of AID-tagged protein depletion was performed with 3-indole-acetic acid (IAA) (Sigma-Aldrich) during the exponential growth phase (approx. OD_{600} 0.8).

The *pPPT1-pSUT129* (*pPS*, mCherry-YFP) fluorescent reporter plasmid was described previously (gift from Sebastian Marquardt, University of Copenhagen) (Marquardt et al., 2014). The target motifs for each transcription factor (obtained from the YeTFaSCo database (de Boer and Hughes, 2012)) were introduced by blunt-end cloning into unique restriction site (SspI) proximal to the *SUT129* TSS and verified by Sanger sequencing. After linearisation by digestion with EcoRI, the

reporter construct was transformed and integrated to replace the endogenous *PPT1-SUT129* locus.

For CRISPR interference (CRISPRi) experiments, yeast expression constructs for nuclease-inactivated Cas9 (D10A/H840A mutations) from *Streptococcus pyogenes* – dCas9 (Addgene #46920) and dCas9-Mxi1 (Addgene #46921) (Gilbert et al., 2013; Qi et al., 2013) – were sub-cloned into single copy integration plasmids by restriction cloning. The 3xFLAG epitope was introduced by Gibson-style cloning (NEBuilder HiFi, NEB) at the C-terminus of each construct, in-frame with the ORF. dCas9-3xFLAG and dCas9-Mxi1-3xFLAG expression constructs on single copy integration plasmids were transformed into yeast after linearisation with PmeI. Single guide RNA (sgRNA) plasmids (gift from Elçin Ünal, UC Berkeley) (Chen et al., 2017) were generated by site-directed mutagenesis (Q5 Site-Directed Mutagenesis Kit, NEB), using specific primers containing the sgRNA 20-mer target sequence (split in half across the two primers). sgRNA target sequence selection was guided by availability of “NGG” protospacer adjacent motif (PAM) sites and transcription start-site sequencing (TSS-seq) data (Wu et al., 2018). sgRNA expression plasmids were integrated at the *SNR52* locus after linearisation by XbaI digestion. Strains, plasmids, and oligonucleotide sequences are listed in Table S3, S4, and S5, respectively.

RNA extraction

Yeast cells were harvested from cultures for RNA extraction by centrifugation, then washed once with sterile water prior to snap-freezing in liquid nitrogen. RNA was extracted from frozen yeast cell pellets using Acid Phenol:Chloroform:Isoamyl alcohol (125:24:1, Ambion) and Tris-EDTA-SDS (TES) buffer (0.01 M Tris-HCl pH

7.5, 0.01 M EDTA, 0.5% w/v SDS), by rapid agitation (1400 RPM, 65 °C for 45 minutes). After centrifugation, the aqueous phase was obtained and RNA was precipitated at -20 °C overnight in ethanol with 0.3 M sodium acetate. After centrifugation and washing with 80% (v/v) ethanol solution, dried RNA pellets were resuspended in DEPC-treated sterile water and subsequently stored at -80 °C.

RT-qPCR

Reverse transcription and quantitative PCR for *IRT2* performed as described previously (Tam and van Werven, 2020). In short, total RNA was purified using the NucleoSpin RNA kit (Macherey-Nagel) according to the manufacturer's instructions. For reverse transcription, ProtoScript II First Strand cDNA Synthesis Kit (New England BioLabs) was used and 500 ng of total RNA was provided as template in each reaction. qPCR reactions were prepared using EXPRESS SYBR GreenER SuperMix (Thermo Fisher Scientific) and *IRT2* levels were quantified on Applied Biosystems 7500 Fast Real-Time PCR System (Thermo Fisher Scientific). The result of n=3 biological replicates are shown. Signals were normalised over *ACT1*. Primer sequences are listed in Table S5.

Western blot

Western blots were performed as described previously (Chia et al., 2017; Chiu et al., 2018). Protein extracts were prepared from whole cells after fixation with trichloroacetic acid (TCA). Samples were pelleted by centrifugation and incubated with 5% w/v TCA solution at 4 °C for at least 10 minutes. Samples were washed with acetone, pelleted, and dried. Samples were then resuspended in protein breakage buffer (50 mM Tris (pH 7.5), 1 mM EDTA, 2.75 mM dithiothreitol (DTT)) and

subjected to disruption using a Mini Beadbeater (Biospec) and 0.5 mm glass beads. Protein extract samples were mixed with SDS-PAGE sample buffer (187.5 mM Tris (pH 6.8), 6.0% v/v β-mercaptoethanol, 30% v/v glycerol, 9.0% v/v SDS, 0.05% w/v Bromophenol blue) in a 2:1 ratio by volume, and protein samples were denatured at 95 °C for 5 minutes.

SDS-PAGE (polyacrylamide gel electrophoresis) was performed using 4-20% gradient gels (Bio-Rad TGX) and samples were then transferred onto PVDF membranes by electrophoresis (wet transfer in cold transfer buffer: 3.35% w/v Tris, 14.9% w/v glycine, 20% v/v methanol). Membranes were incubated in blocking buffer (1% w/v BSA, 1% w/v non-fat powdered milk in phosphate buffered saline with 0.01% v/v Tween-20 (PBST) buffer) before primary antibodies were added to blocking buffer for overnight incubation at 4 °C. For probing with secondary antibodies, membranes were washed in PBST and anti-mouse or anti-rabbit IgG horseradish peroxidase (HRP)-linked antibodies were used for incubation in blocking buffer (1 hour, room temperature). Signals corresponding to protein levels were detected using Amersham ECL Prime detection reagent and an Amersham Imager 600 instrument (GE Healthcare).

Antibodies

The following antibodies were used for western blotting.

Antibody	**Dilution**	**Source**	**Reference**
Anti-hexokinase rabbit IgG	1:8000	US Biological	H2035
Anti-FLAG mouse monoclonal IgG_1, clone M2	1:2000	Sigma-Aldrich (Merck)	F3165
Amersham ECL anti-mouse IgG,	1:10000	GE Life Sciences	NA931V5

HRP-linked whole antibody (from sheep)			
Amersham ECL anti-rabbit IgG, HRP-linked whole antibody (from donkey)	1:10000	GE Life Sciences	NA934V

Northern blot

Northern blots were performed as described previously (Chia et al., 2017; Wu et al., 2018). Briefly, RNA samples (10 µg per lane) were incubated in sample denaturation buffer (1 M deionised glyoxal, 50% v/v DMSO, 10 mM sodium phosphate (NaP_i) buffer (pH 6.8)) at 70 °C for 10 minutes, loading buffer (10% v/v glycerol, 2 mM NaP_i buffer, 0.4% w/v bromophenol blue) was added, and RNA samples were subjected to electrophoresis (2 hours at 80 V) on an agarose gel (1.1% v/v agarose, 0.01 M NaP_i buffer). Capillary transfer was used to transfer total RNA onto positively charged nylon membranes (GE Amersham Hybond N+). Bands corresponding to mature rRNA were visualised by staining with methylene blue solution (0.02% w/v methylene blue, 0.3 M sodium acetate).

The nylon membranes were incubated for at least 3 hours at 42 °C in hybridisation buffer (1% w/v SDS, 40% v/v deionised formamide, 25% w/v dextran sulfate, 58 g/L NaCl, 200 mg/L sonicated salmon sperm DNA (Agilent), 2 g/L BSA, 2 g/L polyvinyl-pyrolidone, 2 g/L Ficoll 400, 1.7 g/L pyrophosphate, 50 mM Tris pH 7.5) or ULTRAhyb Ultrasensitive Hybridization Buffer (Thermo Fisher Scientific) to minimise non-specific probe hybridisation. Probes were synthesised using a Prime-it II Random Primer Labeling Kit (Agilent), 25 ng of target-specific DNA template, and

radioactively labelled with dATP [α-32P] (Perkin-Elmer or Hartmann Analytic). The oligonucleotides used to amplify target-specific DNA templates for *IRT2* and *SNR190* northern blot probes by PCR can be found in Table S5.

Blots were hybridised overnight with radioactively labelled probes at 42 °C, and then washed at 65 °C for 30 minutes each with the following buffers: 2X saline-sodium citrate (SSC) buffer, 2X SSC with 1% w/v SDS, 1X SSC with 1% SDS, and 0.5X SSC with 1% SDS. For image acquisition, membranes were exposed to storage phosphor screens before scanning on the Typhoon FLA 9500 system (GE Healthcare Life Sciences). To strip membranes prior to re-probing for different transcripts, membranes were washed with stripping buffer (1 mM Tris, 0.1 mM EDTA, 0.5% w/v SDS) at 85 °C until negligible residual signal remained on the blots.

Fluorescence microscopy

Yeast cells were grown in YPD media (small batch cultures) to the exponential growth phase and fixed with formaldehyde (3.7% w/v), incubating at room temperature for 15 minutes. Fixed cells were washed with phosphate-sorbitol buffer (0.1 M KP_i (pH 7), 0.05 M $MgCl_2$, 1.2 M sorbitol), and resuspended in the same buffer prior to imaging. Images were acquired using a Nikon Eclipse Ti-E inverted microscope imaging system (Nikon) equipped with a 100x oil objective (NA 1.4), SOLA SE light engine (Lumencor), ORCA-FLASH 4.0 camera (Hamamatsu) and NIS-Elements AR software (Nikon). 500 ms exposure time was specified and GFP and mCherry filters were used to detect YFP and mCherry signals, respectively.

To quantify whole cell fluorescence signals, measurements were performed using ImageJ (version 1.52i, NIH) (Schneider et al., 2012) for the YFP and mCherry channels. Regions of interest (ROIs) were manually drawn around the border of each cell. Mean signal represents the mean intensity in each channel per cell multiplied by the cell area, and the signal for YFP and mCherry was corrected for cell-free background fluorescence in a similar manner. Wild-type cells without integrated fluorescent reporter plasmids were also measured to determine auto-fluorescence signal. 50 cells were quantified for each sample.

RNA-seq with *S. pombe* spike in

Total RNA was extracted from *S. cerevisi*ae pellets spiked in with *S. pombe* cells in a 10:1 ratio of *S. cerevisiae*:*S. pombe* as described above, treated with rDNase in solution (Machery-Nagel), and purified by spin column (Machery-Nagel) prior to preparation of sequencing libraries. 500 ng of RNA input material was used for polyadenylated (polyA) RNA sequencing. Libraries were prepared using TruSeq Stranded mRNA kit (Illumina) according to the manufacturer's instructions (10 or 13 PCR cycles). The libraries were multiplexed and sequenced on either the HiSeq 2500 or 4000 platform (Illumina), and generated ~45 million 101 bp strand-specific paired-end reads per sample on average.

Nascent RNA sequencing (Nascent RNA-seq)

For nascent RNA sequencing (nascent RNA-seq), RNA fragments associated with RNA polymerase II subunit Rpb3 endogenously tagged with 3xFLAG epitope at the C-terminus were isolated by affinity purification as described previously (Churchman and Weissman, 2011, 2012). Small batch cultures of yeast cells grown in YPD media

were collected by centrifugation, the supernatant was aspirated, and cell pellets were immediately snap-frozen by submerging in liquid nitrogen to minimise changes in nascent transcription activity (e.g. in response to cell resuspension in cold lysis buffer with high concentrations of salts and detergents). Frozen cell pellets were dislodged from centrifuge tubes and stored at -80 °C. Cells were subjected to cryogenic lysis by freezer mill grinding under liquid nitrogen (SPEX 6875D Freezer/Mill, standard program: 15 cps for 6 cycles of 2 minutes grinding and 2 minutes cooling each). Yeast “grindate” powder was stored at -80 °C. 2 g of yeast grindate was resuspended in 10 mL of 1X cold lysis buffer (20 mM HEPES pH 7.4, 110 mM potassium acetate, 0.5% v/v Triton-X-100, 1% v/v Tween-20) supplemented with 10 mM $MnCl_2$, 1X Roche cOmplete EDTA-free protease inhibitor, and 50 U/mL SUPERase.In RNase Inhibitor. Chromatin-bound proteins were solubilised by incubation with 1320 U of DNase I (RQ1 RNase-free DNase I, Promega) on ice for 20 minutes. Lysates containing solubilised chromatin proteins were clarified by centrifugation at 20,000 x *g* for 10 minutes (at 4 °C), and the supernatant was taken as input for immunoprecipitation using 500 µL of anti-FLAG M2 affinity gel suspension (A2220, Sigma-Aldrich) per sample (2.5 hours at 4 °C). After immunoprecipitation, the supernatant was removed and beads were washed 4 times with 10 mL of cold wash buffer each time (1X lysis buffer with 50 U/mL Superase.In RNase inhibitor and 1 mM EDTA). After the last wash, agarose beads were transferred to small chromatography spin columns (Pierce Spin Columns, Thermo Fisher), and competitive elution of protein complexes containing Rpb3-3xFLAG protein from the resin was performed by incubating beads with 300 µL of elution buffer (1X cold lysis buffer with 2 mg/mL 3xFLAG peptide) for 30 minutes at 4 °C (3xFLAG peptide provided by the Peptide Chemistry Science Technology Platform,

The Francis Crick Institute). Elution was performed twice and 600 µL of eluate was subjected to acid phenol:chloform RNA extraction and ethanol precipitation. A significant amount of 3xFLAG peptide co-precipitates with the RNA as a contaminant and is later removed by spin column purification.

Purified RNA was fragmented to a mode length of ~200 nucleotides using zinc ion-mediated fragmentation (Ambion AM870, 70 °C for 4 minutes). Fragmented RNA was purified using miRNeasy spin columns (miRNeasy mini kit, QIAGEN), which retain RNAs approximately 18 nucleotides or more in length. Purified RNA was quantified by Qubit (Thermo Fisher) and approximately 150 ng of RNA was subjected to rRNA depletion using the Ribo-Zero Gold rRNA Removal Kit (Yeast) (Illumina MRZY1324, now discontinued). Libraries were prepared using the TruSeq Stranded Total RNA kit (Illumina) according to the manufacturer's instructions (14 PCR cycles). The libraries were multiplexed and sequenced on the HiSeq 4000 platform (Illumina), and generated ~28 million 101 bp strand-specific paired-end reads per sample on average.

RNA-seq data analysis

Adapter trimming was performed with cutadapt (version 1.9.1) (Martin, 2011) with parameters "--minimum-length=25 --quality-cutoff=20 -a AGATCGGAAGAGC -A AGATCGGAAGAGC". BWA (version 0.5.9-r16) (Liao et al., 2014) using default parameters was used to perform the read mapping independently to both the *S. cerevisiae* (assembly R64-1-1, release 90) and *S. pombe* (assembly ASM294v2, release 44) genomes. Genomic alignments were filtered to only include those that were primary, properly paired, uniquely mapped, not soft-clipped, maximum insert

size of 2kb and fewer than 3 mismatches using BamTools (version 2.4.0; (Barnett et al., 2011)). Read counts relative to protein-coding genes were obtained using the featureCounts tool from the Subread package (version 1.5.1) (Liao et al., 2014). The parameters used were “-O --minOverlap 1 --nonSplitOnly --primary -s 2 -p -B -P -d 0 -D 1000 -C --donotsort”.

Differential expression analysis was performed with the DESeq2 package (version 1.12.3) within the R programming environment (version 3.3.1) (Love et al., 2014). The spiked-in *S. pombe* transcripts were used to assess the size factors, potentially mitigating any impact a global shift in the *S. cerevisiae* read counts would have on DESeq2’s usual normalisation procedure. An adjusted p-value of <= 0.01 was used as the significance threshold for the identification of differentially expressed genes.

Nascent RNA-seq data analysis

Adapter trimming was performed with cutadapt (version 1.9.1) (Martin, 2011)) with parameters "--minimum-length=25 --quality-cutoff=20 -a AGATCGGAAGAGC -A AGATCGGAAGAGC -u -50 -U -50". The RSEM package (version 1.3.0) (Li and Dewey, 2011) in conjunction with the STAR alignment algorithm (version 2.5.2a) (Dobin et al., 2013) was used for the mapping and subsequent gene-level counting of the sequenced reads with respect to all *S. cerevisiae* genes downloaded from the Ensembl genome browser (assembly R64-1-1, release 90; (Kersey et al., 2016)). The parameters used were "--star-output-genome-bam --forward-prob 0", and all other parameters were kept as default. Differential expression analysis was performed with the DESeq2 package (version 1.12.3) (Love et al., 2014) within the R

programming environment (version 3.3.1). An adjusted p-value of <= 0.01 was used as the significance threshold for the identification of differentially expressed genes.

STAR genomic alignments were filtered to include only unspliced, primary, uniquely mapped, and properly paired alignments with a maximum insert size of 500 bp. The featureCounts tool from the Subread package (version 1.5.1) (Liao et al., 2014) was used to obtain insert fragment counts within defined intervals (e.g. +100 bp) around the 141 Rap1 binding sites (separate intervals for Watson and Crick strand alignments, 242 intervals total) using the parameters "-O --minOverlap 1 --nonSplitOnly --primary -s 2 -p -B -P -d 0 -D 600 -C". Strand-specific reads were only counted if they overlapped with the interval on the corresponding strand. DESeq2 was used to perform differential expression analysis for intervals around Rap1 binding sites as described above, and the DESeq2 size factors calculated with respect to the transcriptome or S. pombe transcriptome were used for normalisation of the per-sample counts. Genomic annotations of CUTs, SUTs, XUTs, and NUTs (various noncoding RNA species) was obtained from (Wery et al., 2016) (http://vm-gb.curi.e.fr/mw2/), and filtered for transcripts greater than 200 nt in length to match the insert size distribution of RNA-seq libraries.

MNase-seq/TBP ChIP-seq/CUT&RUN analysis

Publicly available datasets were obtained from NCBI Gene Expression Omnibus (GEO) MNase-seq (GSE73337, GSE98260) (Challal et al., 2018; Kubik et al., 2015). Adapters were trimmed from MNase-seq reads using cutadapt as described above. Adapter-trimmed reads were mapped to the *S. cerevisiae* genome (Ensembl assembly R64-1-1, release 90) (Zerbino et al., 2018) with BWA (version 0.5.9-r16) (Li and Durbin, 2009) using default parameters. Only uniquely mapped and properly

paired alignments that had no more than two mismatches in either read of the pair and an insert size 120 – 200 bp were kept for paired-end MNase-seq alignnments. To generate genome-wide coverage tracks for nucleosome occupancy from MNase-seq data, the DANPOS2 dpos command (version 2.2.2) (Chen et al., 2013) was used with parameters "--span 1 --smooth_width 20 --width 40 --count 1000000".

TBP ChIP-seq data were obtained from GSE98260 (Kubik et al., 2018). CUT&RUN data were obtained from GSE116853 (Brahma and Henikoff, 2019). The binding of TBP from the above-mentioned ChIP-seq data and the binding of Sth1/Abf1/Reb1 from the above-mentioned CUT&RUN-ChIP data were plotted for genes belonging to the specified categories with deepTools (version 3.3.0) (Ramirez et al., 2016), using the computeMatrix parameters "reference-point --referencePoint TSS --upstream 500 --downstream 500". The bedgraph files of CUT&RUN-ChIP data were previously clipped to the size of sacCer3 chromosomes, values were normalised over the sum of each sample and transformed to log2.

Promoter directionality score analysis

To calculate directionality scores for coding gene promoters, a curated list of coding gene TSSs was obtained from published TSS sequencing data as described (Park et al., 2014). Any missing TSS coordinates for coding genes were supplemented with the TSS annotation from Ensembl (assembly R64-1-1, release 90) (Zerbino et al., 2018), generating a list of 6,646 *S. cerevisiae* coding gene TSSs. To avoid quantification of divergent direction transcription that constituted coding transcription for an upstream divergent gene, overlapping and divergent gene pairs were removed from the analysis resulting in 2,609 non-overlapping and tandem genes. To simplify

the counting analysis, the coverage for each paired-end read from nascent RNA-seq was reduced to the single 3' terminal nucleotide of the strand-specific read using genomeCoverageBed function within BEDTools (version 2.26.0) (Quinlan and Hall, 2010) with parameters "-ibam stdin -bg -5 -scale %s -strand %s". "Sense" direction windows encompassed nucleotide positions +1 to +500 in the coding direction relative to the TSS, and "antisense" direction windows encompassed nucleotide positions -1 to -500 in the divergent direction relative to the TSS. The total number of reads with 3' end positions falling within the sense and antisense direction windows was quantified for each gene using the computeMatrix tool within deepTools (version 2.5.3) (Ramirez et al., 2016) using parameters "reference-point --referencePoint center --upstream 500 --downstream 500 --binSize 1 --scale 1". A value of 1 was added to all counts to avoid dividing by zero. For each window, the mean read count was calculated from three biological replicate experiments, and used for subsequent calculation and plotting.

Stratification of promoters and sequence analysis

Gene promoters were stratified in quintiles (Q1 to Q5) using Matt (version 1.3.0) (Gohr and Irimia, 2019), according to sense transcription levels or directionality score (log10) in the wild-type strain. Analyses were repeated on only promoters with intermediate gene transcription levels (belonging to Q3 according to sense transcription level). Analysis for enrichment and distribution of A-tracks (AAAAAAA) and GC-rich motifs (CG(C/G)G) as defined in (Kubik et al., 2015) was performed with Matt (version 1.3.0), using the *test_regexp_enrich* and *get_regexp_prof* functions (Gohr and Irimia, 2019).

Data plotting and visualisation

Bar plots, scatter plots, and volcano plots were generated using GraphPad Prism (version 7 or 8). Screenshots of sequencing data were captured using the Integrative Genomics Viewer (IGV, Broad Institute, version 2.4.15) (Robinson et al., 2011). The RStudio integrated development environment (version 1.0.143) was used within the R statistical computing environment (version 3.4.0) for data analysis and visualisation. Software packages within the tidyverse (version 1.2.1) collection were used for data analysis and plotting. The following functions within ggplot (version 3.0.0) were used for plotting: violin plots, geom_violin with parameters “scale = count”;, geom_boxplot with parameters “outlier.colour = NA”; smoothed density plots, geom_density with default parameters; scatter plots, geom_point with default parameters; marginal density histogram plots, ggMarginal with parameters “type = “histogram”, bins = 40, size = 8). The Cairo graphics library (version 1.17.2) was used to generate heat map plots for the RNA-seq data after fold-change values were calculated for bins of 5 nt within defined intervals, comparing between two samples.

Statistical analysis

Information regarding any statistical tests used, number of samples, or number of biological replicate experiments is stated in the corresponding figure legends. For Students’ *t*-tests, calculated *p* values less than 0.05 were considered significant. Plotted error bars in individual figures are stated in the figure legend, as either standard error of the mean (SEM) or 95% confidence intervals (CI).

Publicly available datasets used in this study

MNase-seq and TBP-ChIP seq data were obtained from GSE73337 (Kubik et al., 2015) and GSE98260 (Kubik et al., 2018), respectively. TSS annotation was

described in GSE49026 (Park et al., 2014). Rap1 regulated gene promoters were described in (Wu et al., 2018). The annotations of CUTs, SUTs, NUTs, XUTs were described in http://vm-gb.curi.e.fr/mw2/ (Neil et al., 2009; Schulz et al., 2013; van Dijk et al., 2011; Xu et al., 2009). CUT&RUN data were obtained GSE116853 (Brahma and Henikoff, 2019). Pol II CRAC data were described in GSE97913 and GSE97915 (Candelli et al., 2018b). Reb1-ChEC-seq data were described in GSE67453 (Zentner et al., 2015).

Competing Interests

The authors declare no competing interests.

Author Contributions

F.J.v.W. and A.C.K.W. conceived the project. A.C.K.W. performed most experiments. C.V., H.P. and A.C.K.W. performed the bioinformatic analyses. D.S. and A.C.K.W performed the *s. pombe* normalized RNA-seq experiment. A.C.K.W. and F.J.v.W. designed the experiments. F.J.v.W. wrote the manuscript with input from the other authors. F.J.v.W. supervised the project.

Acknowledgements

We thank the Crick Advanced Sequencing and Genomics Equipment Park Facilities for experimental support, and Sebastian Marquardt and Jesper Svejstrup for sharing reagents. This research was funded in whole, or in part, by the Wellcome Trust (FC001203). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. This work was supported by the Francis Crick Institute (FC001203),

which receives its core funding from Cancer Research UK (FC001203), the UK Medical Research Council (FC001203), and the Wellcome Trust (FC001203).

Data Availability

The accession number for the RNA sequencing and TSS sequencing data reported in this paper is GSE179256.

References

Alcid, E.A., and Tsukiyama, T. (2014). ATP-dependent chromatin remodeling shapes the long noncoding RNA landscape. Genes Dev *28*, 2348-2360.
Ard, R., Allshire, R.C., and Marquardt, S. (2017). Emerging Properties and Functional Consequences of Noncoding Transcription. Genetics *207*, 357-367.
Badis, G., Chan, E.T., van Bakel, H., Pena-Castillo, L., Tillo, D., Tsui, K., Carlson, C.D., Gossett, A.J., Hasinoff, M.J., Warren, C.L*., et al.* (2008). A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. Mol Cell *32*, 878-887.
Barnett, D.W., Garrison, E.K., Quinlan, A.R., Stromberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691-1692.
Biernat, E., Kinney, J., Dunlap, K., Rizza, C., and Govind, C.K. (2021). The RSC complex remodels nucleosomes in transcribed coding sequences and promotes transcription in Saccharomyces cerevisiae. Genetics *217*.
Brahma, S., and Henikoff, S. (2019). RSC-Associated Subnucleosomes Define MNase-Sensitive Promoters in Yeast. Mol Cell *73*, 238-249 e233.
Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., and Kornberg, R.D. (1996). RSC, an essential, abundant chromatin-remodeling complex. Cell *87*, 1249-1260.
Candelli, T., Challal, D., Briand, J.B., Boulay, J., Porrua, O., Colin, J., and Libri, D. (2018a). High-resolution transcription maps reveal the widespread impact of roadblock termination in yeast. EMBO J.
Candelli, T., Challal, D., Briand, J.B., Boulay, J., Porrua, O., Colin, J., and Libri, D. (2018b). High-resolution transcription maps reveal the widespread impact of roadblock termination in yeast. EMBO J *37*.
Challal, D., Barucco, M., Kubik, S., Feuerbach, F., Candelli, T., Geoffroy, H., Benaksas, C., Shore, D., and Libri, D. (2018). General Regulatory Factors Control the Fidelity of Transcription by Restricting Non-coding and Ectopic Initiation. Mol Cell *72*, 955-969 e957.
Chen, J., Tresenrider, A., Chia, M., McSwiggen, D.T., Spedale, G., Jorgensen, V., Liao, H., van Werven, F.J., and Unal, E. (2017). Kinetochore inactivation by expression of a repressive mRNA. Elife *6*.
Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X., and Li, W. (2013). DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res *23*, 341-351.

Chereji, R.V., Ocampo, J., and Clark, D.J. (2017). MNase-Sensitive Complexes in Yeast: Nucleosomes and Non-histone Barriers. Mol Cell *65*, 565-577 e563.
Chia, M., Tresenrider, A., Chen, J., Spedale, G., Jorgensen, V., Unal, E., and van Werven, F.J. (2017). Transcription of a 5' extended mRNA isoform directs dynamic chromatin changes and interference of a downstream promoter. Elife *6*.
Chiu, A.C., Suzuki, H.I., Wu, X., Mahat, D.B., Kriz, A.J., and Sharp, P.A. (2018). Transcriptional Pause Sites Delineate Stable Nucleosome-Associated Premature Polyadenylation Suppressed by U1 snRNP. Mol Cell *69*, 648-663 e647.
Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368-373.
Churchman, L.S., and Weissman, J.S. (2012). Native elongating transcript sequencing (NET-seq). Curr Protoc Mol Biol *Chapter 4*, Unit 4 14 11-17.
de Boer, C.G., and Hughes, T.R. (2012). YeTFaSCo: a database of evaluated yeast transcription factor sequence specificities. Nucleic Acids Res *40*, D169-179.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.
du Mee, D.J.M., Ivanov, M., Parker, J.P., Buratowski, S., and Marquardt, S. (2018). Efficient termination of nuclear lncRNA transcription promotes mitochondrial genome maintenance. Elife *7*.
Floer, M., Wang, X., Prabhu, V., Berrozpe, G., Narayan, S., Spagna, D., Alvarez, D., Kendall, J., Krasnitz, A., Stepansky, A.*, et al.* (2010). A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. Cell *141*, 407-418.
Flynn, R.A., Almada, A.E., Zamudio, J.R., and Sharp, P.A. (2011). Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. Proc Natl Acad Sci U S A *108*, 10460-10465.
Frank, S., Ahuja, G., Bartsch, D., Russ, N., Yao, W., Kuo, J.C., Derks, J.P., Akhade, V.S., Kargapolova, Y., Georgomanolis, T.*, et al.* (2019). yylncT Defines a Class of Divergently Transcribed lncRNAs and Safeguards the T-mediated Mesodermal Commitment of Human PSCs. Cell Stem Cell *24*, 318-327 e318.
Ganguli, D., Chereji, R.V., Iben, J.R., Cole, H.A., and Clark, D.J. (2014). RSC-dependent constructive and destructive interference between opposing arrays of phased nucleosomes in yeast. Genome Res *24*, 1637-1649.
Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. Nat Rev Genet *21*, 102-117.
Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A.*, et al.* (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442-451.
Gohr, A., and Irimia, M. (2019). Matt: Unix tools for alternative splicing analysis. Bioinformatics *35*, 130-132.
Gowthaman, U., Ivanov, M., Schwarz, I., Patel, H.P., Müller, N.A., García-Pichardo, D., Lenstra, T., and Marquardt, S. (2021). Hda1C restricts the transcription initiation frequency to limit divergent non-coding RNA transcription. BioRxiv.
Hainer, S.J., Gu, W., Carone, B.R., Landry, B.D., Rando, O.J., Mello, C.C., and Fazzio, T.G. (2015). Suppression of pervasive noncoding transcription in embryonic stem cells by esBAF. Genes Dev *29*, 362-378.
Hartley, P.D., and Madhani, H.D. (2009). Mechanisms that specify promoter nucleosome location and identity. Cell *137*, 445-458.
Jacquier, A. (2009). The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. Nat Rev Genet *10*, 833-844.
Jin, Y., Eser, U., Struhl, K., and Churchman, L.S. (2017). The Ground State and Evolution of Promoter Region Directionality. Cell *170*, 889-898 e810.

Kersey, P.J., Allen, J.E., Armean, I., Boddu, S., Bolt, B.J., Carvalho-Silva, D., Christensen, M., Davis, P., Falin, L.J., Grabmueller, C.*, et al.* (2016). Ensembl Genomes 2016: more genomes, more complexity. Nucleic Acids Res *44*, D574-580.
Klein-Brill, A., Joseph-Strauss, D., Appleboim, A., and Friedman, N. (2019). Dynamics of Chromatin and Transcription during Transient Depletion of the RSC Chromatin Remodeling Complex. Cell Rep *26*, 279-292 e275.
Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C.L., Pugh, B.F., and Korber, P. (2016). Genomic Nucleosome Organization Reconstituted with Pure Proteins. Cell *167*, 709-721 e712.
Kubik, S., Bruzzone, M.J., Challal, D., Dreos, R., Mattarocci, S., Bucher, P., Libri, D., and Shore, D. (2019). Opposing chromatin remodelers control transcription initiation frequency and start site selection. Nat Struct Mol Biol *26*, 744-754.
Kubik, S., Bruzzone, M.J., Jacquet, P., Falcone, J.L., Rougemont, J., and Shore, D. (2015). Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast. Mol Cell *60*, 422-434.
Kubik, S., Bruzzone, M.J., and Shore, D. (2017). Establishing nucleosome architecture and stability at promoters: Roles of pioneer transcription factors and the RSC chromatin remodeler. Bioessays *39*.
Kubik, S., O'Duibhir, E., de Jonge, W.J., Mattarocci, S., Albert, B., Falcone, J.L., Bruzzone, M.J., Holstege, F.C.P., and Shore, D. (2018). Sequence-Directed Action of RSC Remodeler and General Regulatory Factors Modulates +1 Nucleosome Position to Facilitate Transcription. Mol Cell *71*, 89-102 e105.
Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.
Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.
Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.
Lorch, Y., Cairns, B.R., Zhang, M., and Kornberg, R.D. (1998). Activated RSC-nucleosome complex and persistently altered form of the nucleosome. Cell *94*, 29-34.
Lorch, Y., Maier-Davis, B., and Kornberg, R.D. (2014). Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. Genes Dev *28*, 2492-2497.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
Luo, S., Lu, J.Y., Liu, L., Yin, Y., Chen, C., Han, X., Wu, B., Xu, R., Liu, W., Yan, P.*, et al.* (2016). Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. Cell Stem Cell *18*, 637-652.
Malabat, C., Feuerbach, F., Ma, L., Saveanu, C., and Jacquier, A. (2015). Quality control of transcription start site selection by nonsense-mediated-mRNA decay. Elife *4*.
Marquardt, S., Escalante-Chong, R., Pho, N., Wang, J., Churchman, L.S., Springer, M., and Buratowski, S. (2014). A Chromatin-Based Mechanism for Limiting Divergent Noncoding Transcription. Cell *158*, 462.
Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal *17*, 10.
Mohrmann, L., and Verrijzer, C.P. (2005). Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. Biochim Biophys Acta *1681*, 59-73.
Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L.M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature *457*, 1038-1042.
Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat Methods *6*, 917-922.
Nojima, T., Tellier, M., Foxwell, J., Ribeiro de Almeida, C., Tan-Wong, S.M., Dhir, S., Dujardin, G., Dhir, A., Murphy, S., and Proudfoot, N.J. (2018). Deregulated Expression of Mammalian lncRNA through

Loss of SPT6 Induces R-Loop Formation, Replication Stress, and Cellular Senescence. Mol Cell *72*, 970-984 e977.
Park, D., Morris, A.R., Battenhouse, A., and Iyer, V.R. (2014). Simultaneous mapping of transcript ends at single-nucleotide resolution and identification of widespread promoter-associated non-coding RNA governed by TATA elements. Nucleic Acids Res *42*, 3736-3749.
Parnell, T.J., Huff, J.T., and Cairns, B.R. (2008). RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. EMBO J *27*, 100-110.
Pelechano, V., and Steinmetz, L.M. (2013). Gene regulation by antisense transcription. Nat Rev Genet *14*, 880-893.
Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173-1183.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.
Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165.
Rawal, Y., Chereji, R.V., Qiu, H., Ananthakrishnan, S., Govind, C.K., Clark, D.J., and Hinnebusch, A.G. (2018). SWI/SNF and RSC cooperate to reposition and evict promoter nucleosomes at highly expressed genes in yeast. Genes Dev *32*, 695-710.
Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.
Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature Methods *9*, 671-675.
Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J., and Cramer, P. (2013). Transcriptome surveillance by selective termination of noncoding RNA synthesis. Cell *155*, 1075-1087.
Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849-1851.
Seila, A.C., Core, L.J., Lis, J.T., and Sharp, P.A. (2009). Divergent transcription: a new feature of active promoters. Cell Cycle *8*, 2557-2564.
Sigova, A.A., Mullen, A.C., Molinie, B., Gupta, S., Orlando, D.A., Guenther, M.G., Almada, A.E., Lin, C., Sharp, P.A., Giallourakis, C.C*., et al.* (2013). Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. Proc Natl Acad Sci U S A *110*, 2876-2881.
Struhl, K. (2007). Transcriptional noise and the fidelity of initiation by RNA polymerase II. Nat Struct Mol Biol *14*, 103-105.
Tam, J., and van Werven, F.J. (2020). Regulated repression governs the cell fate promoter controlling yeast meiosis. Nature communications *11*, 2271.
van Dijk, E.L., Chen, C.L., d'Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoix-Ne, P., Loeillet, S*., et al.* (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature *475*, 114-117.
Wery, M., Descrimes, M., Vogt, N., Dallongeville, A.S., Gautheret, D., and Morillon, A. (2016). Nonsense-Mediated Decay Restricts LncRNA Levels in Yeast Unless Blocked by Double-Stranded RNA Structure. Mol Cell *61*, 379-392.
Wu, A.C.K., Patel, H., Chia, M., Moretto, F., Frith, D., Snijders, A.P., and van Werven, F.J. (2018). Repression of Divergent Noncoding Transcription by a Sequence-Specific Transcription Factor. Mol Cell *72*, 942-954 e947.
Wu, X., and Sharp, P.A. (2013). Divergent transcription: a driving force for new gene origination? Cell *155*, 990-996.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Munster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033-1037.
Yang, F., Tanasa, B., Micheletti, R., Ohgi, K.A., Aggarwal, A.K., and Rosenfeld, M.G. (2021). Shape of promoter antisense RNAs regulates ligand-induced transcription activation. Nature.
Zentner, G.E., Kasinathan, S., Xin, B., Rohs, R., and Henikoff, S. (2015). ChEC-seq kinetics discriminates transcription factor binding sites by DNA sequence and shape in vivo. Nature communications *6*, 8733.
Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Giron, C.G.*, et al.* (2018). Ensembl 2018. Nucleic Acids Res *46*, D754-D761.

Figure legends

Figure 1. Depletion of RSC increases divergent transcript accumulation at Rap1-regulated genes

(A) Top: Rap1 and Sth1 (co) depletion using AID tagged strains (*RAP1-AID*, *STH1-AID*, *RAP1-AID*/*STH1-AID*) (FW7238, FW7220, and FW7232). Cells were grown until the exponential growth phase and treated with IAA for 2 hours. AID tagged proteins were detected by western blot using anti-V5 antibodies. Hxk1 was used as a loading control, detected using anti-Hxk1 antibodies. Bottom: Scheme for RNA-seq experiment of Rap1 and Sth1 depletion. To control for global changes in RNA expression, we spiked in *S. pombe* cells (see Materials and Methods), and performed comparative analysis using *S. cerevisiae* (sc) transcriptome normalization or *S. pombe* (sp) normalization. **(B)** Fold change (FC) of 141 Rap1-regulated genes using *S. cerevisiae* (sc) or *S. pombe* (sp) normalization in IAA treated cells compared to mock treated cells (DMSO). Each dot represents one coding gene transcript. **(C)** RNA-seq signals at the *RPL43B* locus (coding sense direction, blue) and *IRT2* (divergent coding direction, red). Both sc and sp normalized signals are displayed. **(D)** Similar as C except that *RPS3* and *RPS10* loci and the corresponding divergent transcripts for the sp normalization are displayed. Indicated are differences

in TSS usage. **(E)** Heatmap representing changes in RNA expression levels in the divergent noncoding direction. Indicated are the three clusters (c1, c2, and c3) based on previous analysis for Rap1-regulated gene promoters (Wu et al., 2018). Both *sc*- and *sp*-normalized values are displayed. Differences with respect the DMSO control for each depletion strain are displayed. **(F)** Related to E. The log2 fold change (FC) is displayed for the antisense (AS) and sense (S) signal in 100 nucleotide windows centred on the Rap1 binding site.

Figure 2. Depletion of RSC increases divergent transcription and alters promoter directionality

(A) Heatmap representing changes in transcription in the antisense direction as detected by nascent RNA-seq after Sth1 and Rap1 (co)-depletion for 141 Rap1-regulated gene promoters (FW7238, FW7220, and FW7232). Signals of depleted cells were compared to the degron strain treated with DMSO and centred on the Rap1 binding motif. **(B)** Approach for calculating directionality score for each promoter. In short, the nascent RNA-seq signals 400 bp upstream (antisense strand) and 400 bp downstream (sense strand) of the coding gene TSS were taken. Subsequently, the ratio of sense over antisense signals (log10) was computed, which was defined as the directionality score. A comparable approach was also described in (Jin et al., 2017). **(C)** Density plot of directionality scores for tandem non-overlapping genes (n=2609) and Rap1-regulated genes (n=141). **(D)** Density plots representing promoter directionality of control (*RAP1-AID* +DMSO) and Rap1 depleted cells (*RAP1-AID* + IAA). **(E)** Similar as D, except comparing control (*STH1-AID* + DMSO) and Sth1 depleted (*STH1-AID* + IAA) cells). **(F)** Similar as E, except

comparing control (*RAP1-AID/STH1-AID* + DMSO) and (*RAP1-AID*/*STH1-AID* + IAA). **(G)** Similar as D and F, but the directionality scores from Rap1 depleted (*RAP1-AID* + IAA) with Rap1/Sth1 depleted (*RAP1-AID*/*STH1-AID* + IAA) cells were directly compared.

Figure 3. RSC depletion affects promoter directionality genome-wide

(A) Example loci (*STE24* and *HIR3*) displaying increased divergent transcription at promoters upon depletion of Sth1. Shown are the nascent RNA-seq data of the wild-type, control (*STH1-AID* + DMSO), and Sth1 depleted (*STH1-AID* + IAA) cells. MNase-seq signals of control and Sth1 depleted cells are also displayed. The MNase-seq data were obtained from (Kubik et al., 2015; Kubik et al., 2018). **(B)** Volcano plots of nascent-RNA seq data comparing Sth1 depleted to control cells (IAA/DMSO) for different classes of transcription units. Displayed are plots for protein-coding genes, and noncoding regions: CUTs, SUTs, XUTs, and NUTs. Number of loci with increased (UP) or decreased (DOWN) transcription are displayed (P<0.05 and FC=> 2). **(C)** Heatmap of nascent-RNA-seq data comparing Sth1 depleted to control (IAA/DMSO) cells for all protein-coding genes (n=6646). The data is centred on the TSS and clustered on signals in the antisense direction (ASc1 (n=2025), ASc2 (n=1816), and ASc3 (n=2805). **(D)** Promoter directionality score of tandem and non-overlapping protein-coding genes in wild-type, control (*STH1-AID* + DMSO) and Sth1 depleted (*STH1-AID* +IAA) cells. **(E)** Violin plots displaying fold change (log2FC) in nascent-RNA-seq signal, in the sense and divergent direction for control (*STH1-AID* + DMSO) versus wild-type cells, and Sth1 depletion (*STH1-AID* + IAA) versus wild-type cells.

Figure 4. RSC acts on highly directional promoters with fragile nucleosomes

(A) Violin plots display log2 of fold change (FC) in divergent transcription, sense transcription, and directionality score in Sth1 depleted versus control cells (IAA/DMSO). Tandem genes were sorted on directionality score, and gene groups were divided in five groups (Q1 to Q5). Q1 represents the group of gene promoters with lowest directionality score, and Q5 represents the group with the highest directionality score. The range for log10 directionality scores for each group are the following: Q1(-3 - 0.9), Q2(0.9 - 1.5), Q3(1.5 - 2), Q4(2 - 2.5), Q5(2.5 - 4.2). **(B)** Violin plots displaying log2 fold change (FC) in directionality score of Sth1 depleted versus control (IAA/DMSO) for the group gene promoters (Q5) with highest sense transcription levels or highest directionality score (Q5) in WT cells. **(C)** Metagene gene analysis of MNase-seq data for Q1 and Q5 described in A. Data were centred on TSS, and displayed are MNase-seq signals for chromatin extracts treated with low MNase or high MNase concentrations. The MNase-seq data were obtained from (Kubik et al., 2015; Kubik et al., 2018). Marked are the regions representing the +1, -1, and -2 nucleosome position. **(D)** Violin plots of sense direction nascent RNA signals for a group of gene promoters with low and high directionality scores (lowest 20% vs highest 20%), but that have comparable intermediate transcription levels in the sense direction (Q3). Displayed are sense nascent RNA-seq signals for the genes with 20% lowest directionality score and top 20% directionality score within this select group of gene promoters (n=105). **(E)** Metagene analysis of MNase-seq data of the two groups of gene promoters described D in control and Sth1 depleted cells. **(F)** Example loci that display increased divergent transcription and changes in chromatin structure upon Sth1 depletion. Shown are the data from nascent RNA-seq

and MNase-seq. **(G)** Metagene analysis of TBP ChIP-seq for the groups of genes described in C. **(H)** Analysis of RSC complex sequence motifs (A-tracks and CG(C/G)G) in Q1 and Q5 groups of gene promoters sorted on directionality score. Displayed is distribution in the motif signal on the divergent or antisense strand only.

Figure 5. RSC and GRFs are enriched upstream in promoters with a high directional score

(A) Metagene analysis and heatmaps of CUT&RUN data of Sth1 (RSC), and the GRFs Abf1 and Reb1. In addition, signals of Sth1 CUT&RUN followed by histone H2B immunoprecipitation are displayed. The data was obtained from (Brahma and Henikoff, 2019). A comparison between Q1 (blue), the group of gene promoters with lowest directionality score, and Q5 (green), the group with the highest directionality score, was made. Corresponding heatmaps for Q1 and Q5 are displayed below the graph. **(B)** Data from RNA Pol II CRAC and Reb1-ChEC-seq datasets described previously (Candelli et al., 2018a; Zentner et al., 2015), representing RNA Pol II transcription in presence or absence of Reb1 and binding of Reb1, respectively. Displayed are the *BAP2* and *SCO1* loci, which show increased divergent transcription upon Reb1 depletion.

Figure 6. Targeting GRFs or dCas9 to divergent core promoters is sufficient to repress divergent noncoding transcription.

(A) Schematic overview of reporter construct. The transcription factor binding sites were cloned 20 nucleotides upstream of the YFP (*SUT129*) TSS. **(B)** YFP, noncoding, signal for constructs harbouring binding sites for Gal4, Gcn4, Cat8, Gcr1,

Cbf1, Abf1 and Reb1 cloned into the wild-type (WT) *PPT1* promoter (FW6407). The YFP activity was determined in wild-type control cells and in gene deletion stains of matching transcription factor binding site reporter constructs (FW6404, FW6306, FW6401, FW6300, FW6402, FW6302, FW6403, FW6424, FW6405, and FW6315). For Abf1 and Reb1 reporter constructs, Abf1 and Reb1 were depleted using the auxin inducible degron (*ABF1*-AID and *REB1*-AID) (FW6415 and FW6411). Cells were treated for 2 hours with IAA or DMSO. Displayed is the mean signal of at least 50 cells. The error bars represent 95% confidence intervals. **(C)** Scheme depicting targeting dCas9 to repress the divergent transcript *IRT2*. **(D)** Design of gRNAs targeted the *IRT2* TSS. Displayed are TSS-seq data from after Rap1 depletion (Wu et al., 2018). **(E)** Western blot of Rap1 and dCas9 detected with anti-V5 and anti-FLAG antibodies, respectively. Cells harbouring *RAP1*-AID were treated with IAA to deplete Rap1 and induce *IRT2* expression in either cells with empty vector (EV) (FW8477) or cells harbouring a dCas9 construct and a construct expressing the gRNA targeted to the Rap1 binding site. (**F**) *IRT2* expression detected by northern blot for strains and treatments described in E. As a loading control, the membrane was re-probed for *SNR190.* **(G)** *IRT2* expression as detected in by RT-qPCR of as described in D. A control gRNA was included in the analysis targeted to the TSS of *TEF1* (FW8527). Displayed are the mean signals of n=3 and SEM. **(H)** Model for promoter directionality.

Figure S1. Divergent transcription at *RPL40B* locus upon Rap1 and Sth1 (co) depletion

(A) Rap1 and Sth1 (co) depletion using AID tagged strains (*RAP1-AID*, *STH1-AID*, *RAP1-AID*/*STH1-AID*) (FW7238, FW7220, and FW7232). Cells were grown till exponential growth and treated with IAA for 2 hours. To control for global changes in RNA expression, we spiked in *S. pombe* cells, and performed comparative analysis using *S. cerevisiae* (sc) transcriptome normalization or *S. pombe* (sp) normalization. Fold change (FC) of 141 Rap1-regulated genes using *S. cerevisiae* (sc) or *S. pombe* (sp) normalization in depleted (+IAA) cells compared to wild-type cells (WT, FW7228). Each dot represents one coding gene transcript. **(B)** RNA-seq signals for *RPL40B* (coding sense direction, blue) and *iMLP1* (divergent coding direction, red). Both sc and sp normalized signals are displayed.

Figure S2. Rap1 and RSC control RNA accumulation in the divergent direction at Rap1-regulated genes

(A) Matching the analysis of Figure 1E, on the sense strand data is displayed. Indicated are the three clusters (c1, c2, and c3) based on previous analysis (Wu et al., 2018) for 141 Rap1-regulated gene promoters. Both sc and sp normalized transcriptome are displayed. Differential of depleted cells with respect to control cells (IAA/DMSO) are displayed. **(B)** Similar analysis as in Figure 1E, except that the data versus the wild-type (*STH1-AID* IAA vs WT, *RAP1-AID* IAA vs WT, *RAP1-AID*/*STH1-AID* IAA vs WT) are displayed. **(C)** Related to B. The log2 fold change (FC) is displayed for the antisense (AS) and sense (S) signal around 100 nucleotides up- or downstream of the Rap1 binding site. **(D)** Related to C and Figure 1F, highlighting the difference between *STH1-AID* (IAA/DMSO) and *STH1-AID*+IAA/WT.

Figure S3. Metagene analysis of MNase-seq data

Metagene analysis of MNase-seq dataset described previously gene sets described in A and B. Data are displayed for the different clusters (c1 to c3) for Sth1 depletion, Rap1 depletion, and Sth1 + Rap1 depletion. The MNase-seq signals are centred on the Rap1 binding site position. The data were obtained from (Kubik et al., 2015).

Figure S4. Nascent RNA-seq of Sth1, Rap1, and Sth1 + Rap1 depletion

(A) Rpb3-FLAG immunoprecipitation as detected by western blot. Rpb3-FLAG was immunoprecipitated from wild-type (FW7228) cells or cells harbouring AID tagged alleles for Sth1, Rap1, and Sth1 + Rap1 (*STH1-AID*, *RAP1-AID*, *RAP1-AID* + *STH1-AID*) (FW7238, FW7220, and FW7232). AID tagged cells were either treated with IAA or DMSO. Input (I), flow-through (FT), FLAG elution (E), and boiling in sample buffer from FLAG beads (B) are shown. **(B)** Scatter plots showing the sense (log10) versus divergent (log10) transcription signals for all genes (left) and tandem genes (n-2609) (left) for wild-type cells. The nascent RNA-seq signals 400 bp downstream (sense strand, labelled as sense) and 400 bp upstream (antisense strand, labelled as divergent) of the TSS were taken. **(C)** Density plots representing promoter directionality scores in wild-type cells for all genes and tandem genes (n=2609). **(D)** Relative changes in transcription levels in cells depleted for Rap1 (*RAP1-AID* + IAA), Sth1 (*STH1-AID* + IAA) or both (*RAP1-AID*/*STH1-AID* + IAA) in comparison to the control treated cells (IAA/DMSO) for 141 Rap1-regulated gene promoters. The nascent RNA-seq signals 400 bp upstream (antisense strand, labelled as div) and 400 bp downstream (sense strand, labelled as sense) the TSS were taken.

Figure S5. RSC deletion affects promoter directionality genome-wide

(A) Violin plots depicting sense and divergent transcription level changes for control (*STH1-AID* + DMSO) over wild-type (WT), Sth1 depletion (*STH1-AID* +IAA) over WT, and Sth1 (*STH1-AID* +IAA)/ (*STH1-AID* + DMSO) samples. **(B)** Same comparisons as B, but the changes in directionality score are displayed in violin plots.

Figure S6. RSC acts at highly directional promoters

Violin plots displaying fold change (log2(FC)) in divergent transcription, sense transcription, and directionality score. Tandem genes were sorted on sense nascent RNA levels (A, B and D) or directionality score (C and E), and gene groups were divided in five groups (Q1 to Q5). Displayed are the following comparisons: **(A)** Sth1 depletion (IAA/DMSO) sorted on sense nascent RNA levels. The range of sense nascent RNA levels (read counts) for each group are the following: Q1(1 - 240), Q2(240 - 747), Q3(747 - 1441), Q4(1441 - 2850), Q5(2850 - 105685). **(B)** Sth1 depletion (IAA/WT) sorted on sense nascent RNA levels. **(C)** Sth1 depletion (IAA/WT) sorted on directionality score. The range for log10 directionality scores for each group are the following: Q1(-3 - 0.9), Q2(0.9 - 1.5), Q3(1.5 - 2), Q4(2 - 2.5), Q5(2.5 - 4.2). **(D)** Control (DMSO/WT) sorted on sense nascent RNA levels. **(E)** Control (DMSO/WT) sorted on directionality score.

Figure S7. Chromatin, sequence features, sense transcription and RSC

(A) Metagene gene analysis of MNase-seq data for Q1 and Q5 described in Figure S6A. Data were centred on TSS, and displayed are MNase-seq signals for extracts

treated with low MNase or high MNase concentrations. The MNase-seq data were obtained from (Kubik et al., 2015). Marked are the regions representing the +1, -1 and -2 nucleosome positions. Metagene gene analysis of MNase-seq data of group of genes with highest sense nascent RNA levels and highest directionality score. **(B)** Distribution of RSC associated DNA sequences in promoters sorted on directionality score. Displayed are the analysis of A-tracks and CG(C/G)G motifs. Distribution of the signals with respect to the sense and divergent direction are shown.

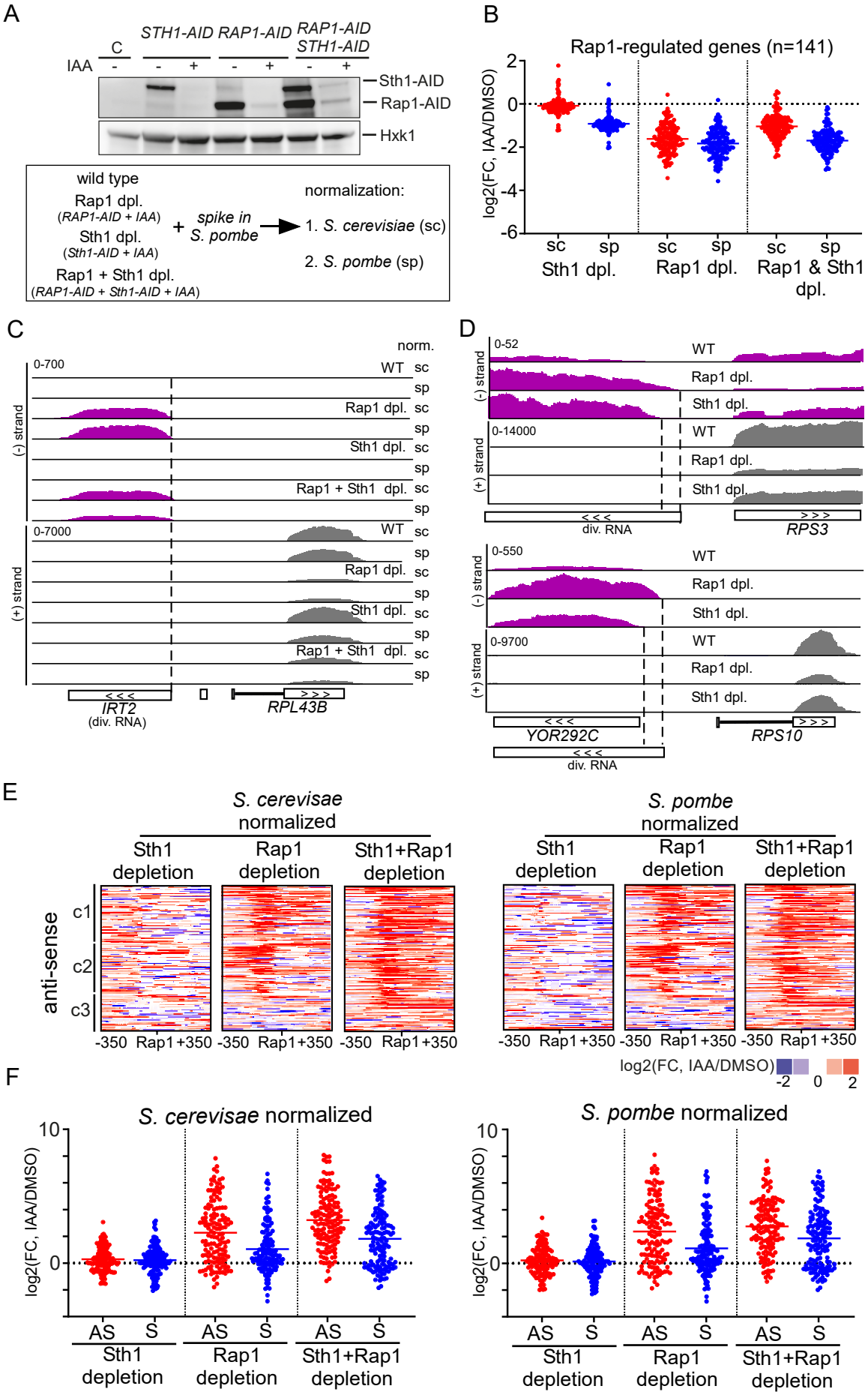
Figure S8. RSC and GRFs are enriched at upstream in promoters with a high directional score

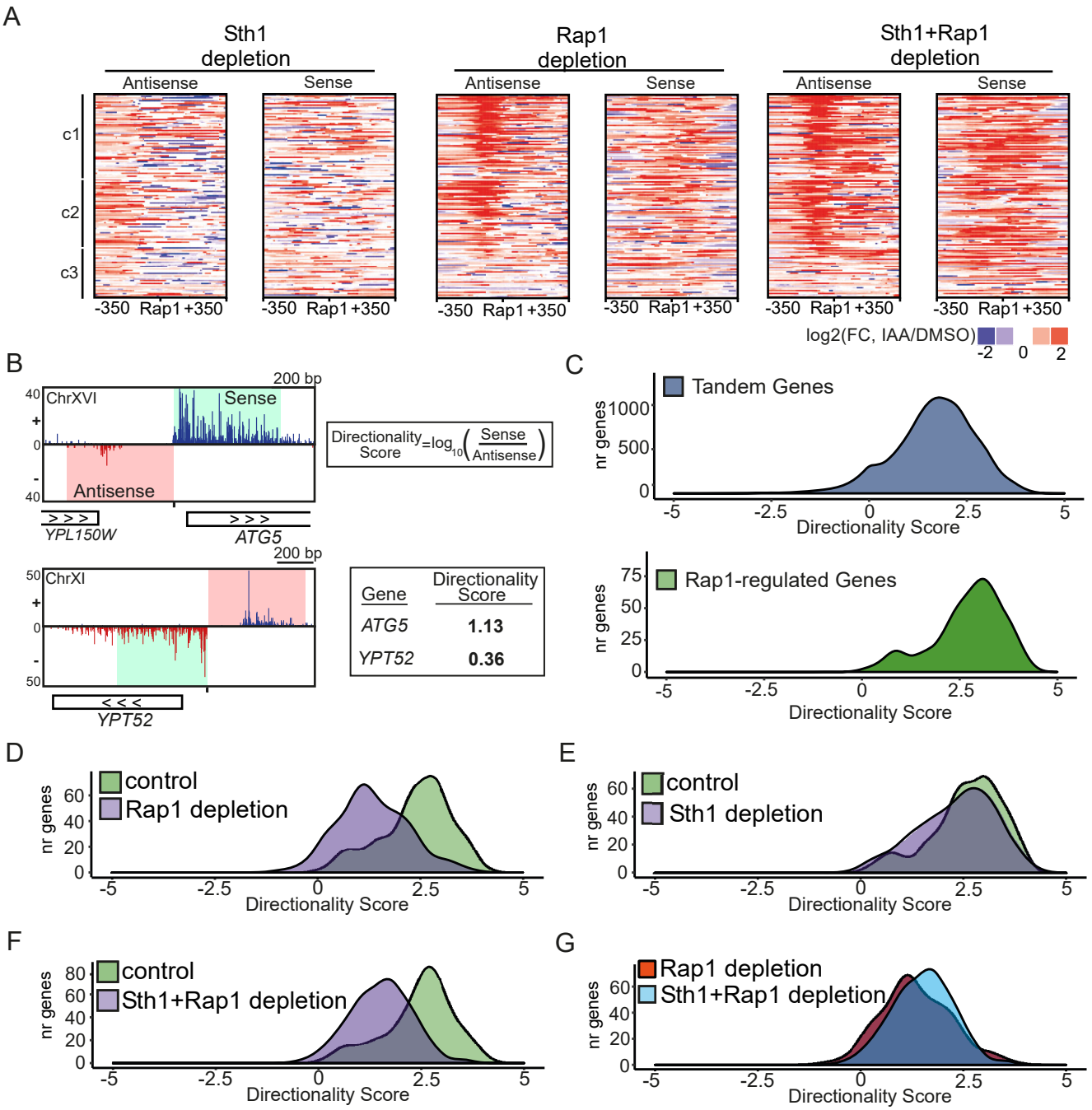
Same analysis in Figure 5A, except that the two biological repeats are displayed. Metagene analysis and heatmaps of CUT&RUN data of Sth1 (RSC), and the GRFs Abf1 and Reb1. In addition, signals of Sth1 CUT&RUN followed by histone H2B immunoprecipitation are displayed. The data was obtained from (Brahma and Henikoff, 2019). Compared Q1 (blue) being the group of gene promoter with lowest directionality score, and Q5 (green) being the group with the highest directionality score. Corresponding heatmaps for Q1 and Q5 are displayed (bottom).

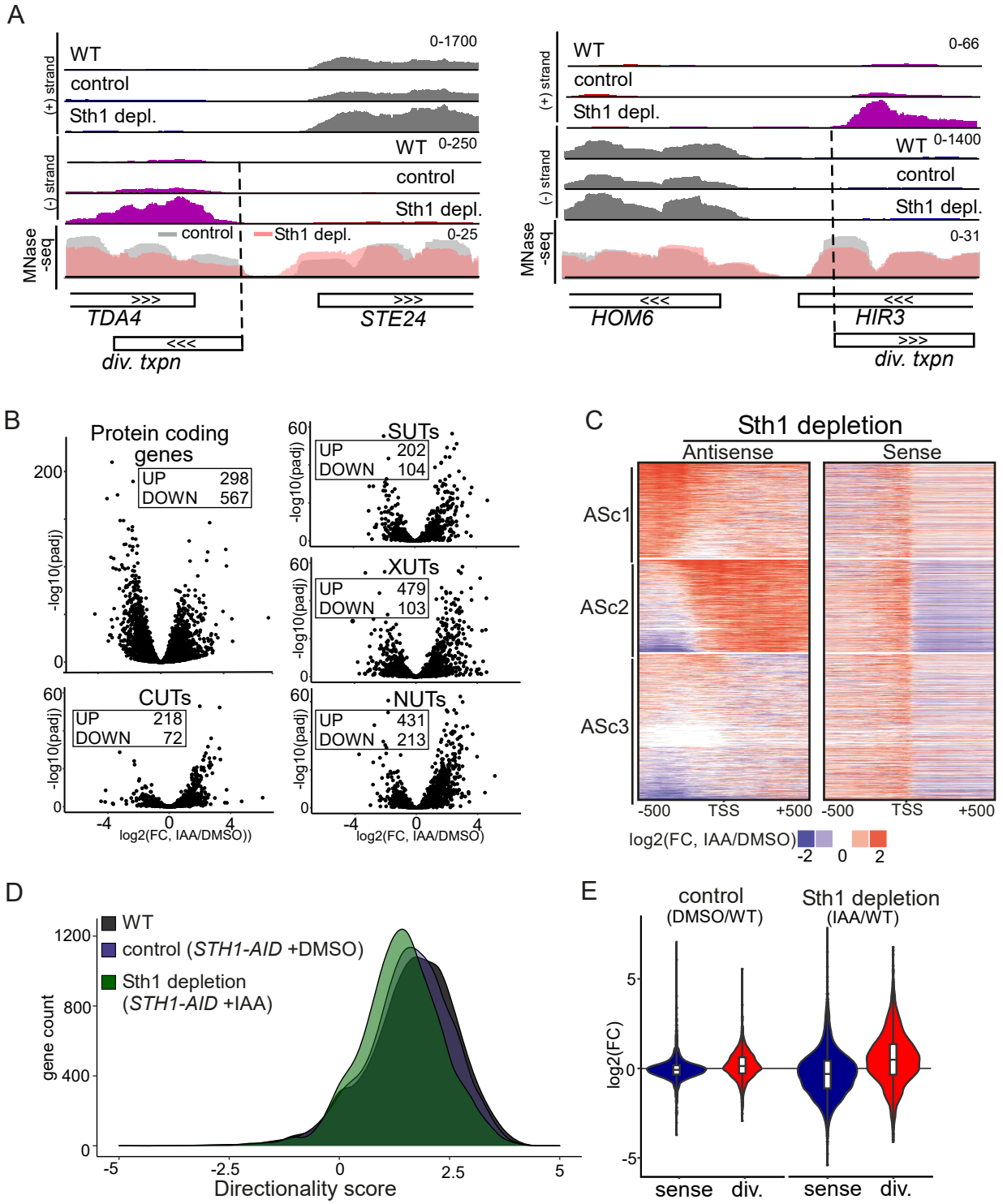
Figure S9. Targeting GRFs or dCas9 to divergent core promoters is sufficient to repress divergent noncoding transcription.

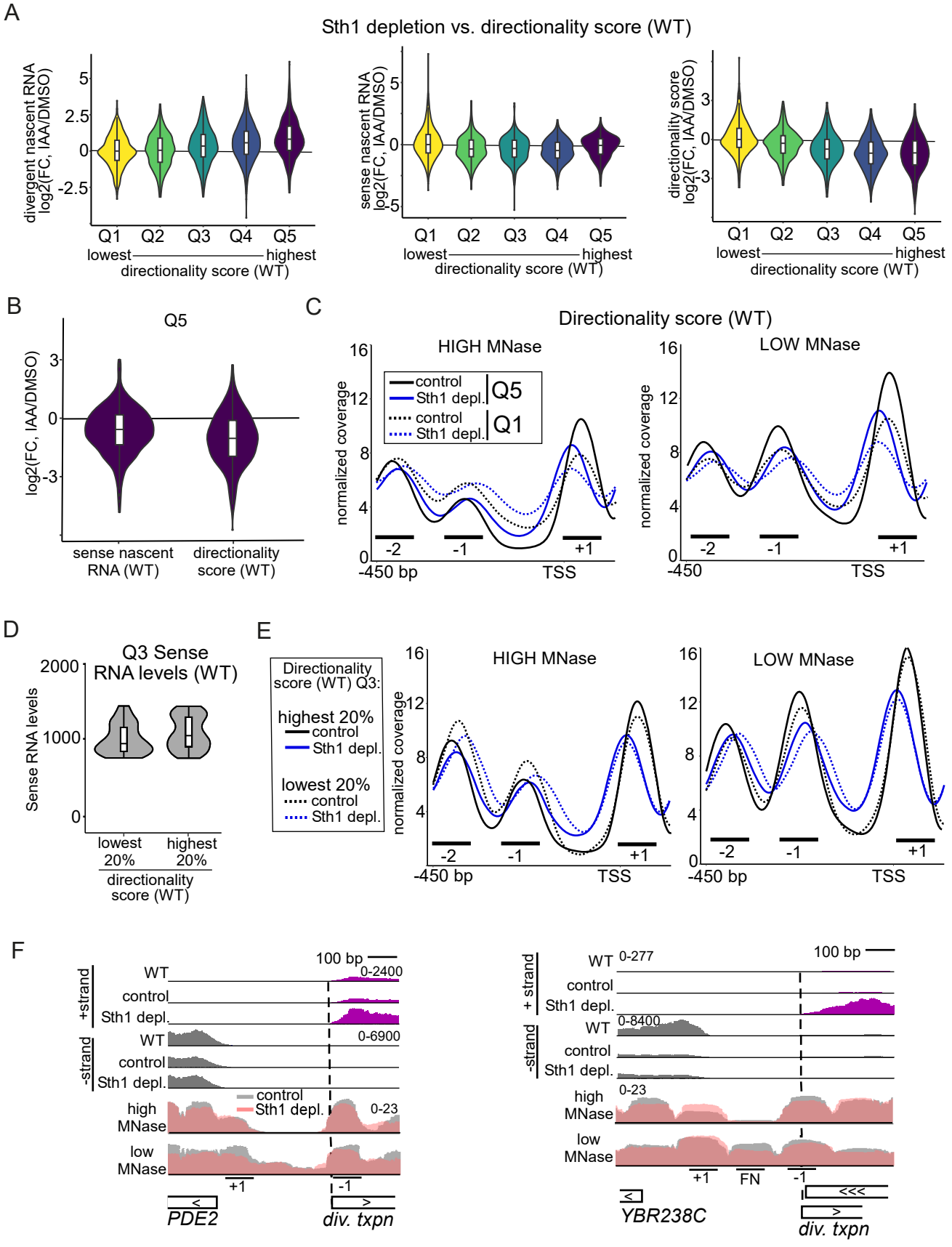
(A) mCherry (coding direction) signal for constructs harbouring binding sites for Gal4, Gcn4, Cat8, Gcr1, Cbf1, Abf1 and Reb1 cloned to the wild-type (WT) *PPT1*

promoter (FW6407). The mCherry activity was determined in wild-type control cells and in gene deletion stains of matching transcription factor binding site reporter constructs (FW6404, FVW6306, FW6401, FW6300, FW6402, FW6302, FW6403, FW6424, FW6405, and FW6315). For Abf1 and Reb1 reporter constructs, Abf1 and Reb1 were depleted using the auxin inducible degron system (*ABF1*-AID and *REB1*-AID) (FW6415 and FW6411). Cells were treated for 2 hours with IAA or DMSO. Displayed are the mean signal of at least 50 cells. The error bars represent 95% confidence intervals. **(B)** Western blot of Rap1 and dCas9-Mxi1. Cells harbouring *RAP1-AID* were treated IAA to deplete Rap1 and induce *IRT2*. Rap1 was detected with anti-V5 antibodies, and dCas9 with anti-FLAG antibodies. gRNAs targeting *IRT2*, and control gRNAs targeting *MLP1* and *TEF1* were used (FW8531, FW8529, and FW8535). **(C)** Northern blot of *IRT2* expression in the Rap1 depletion of samples strains described in C. As a loading control the membrane was re-probed with a probe directed against *SNR190*.

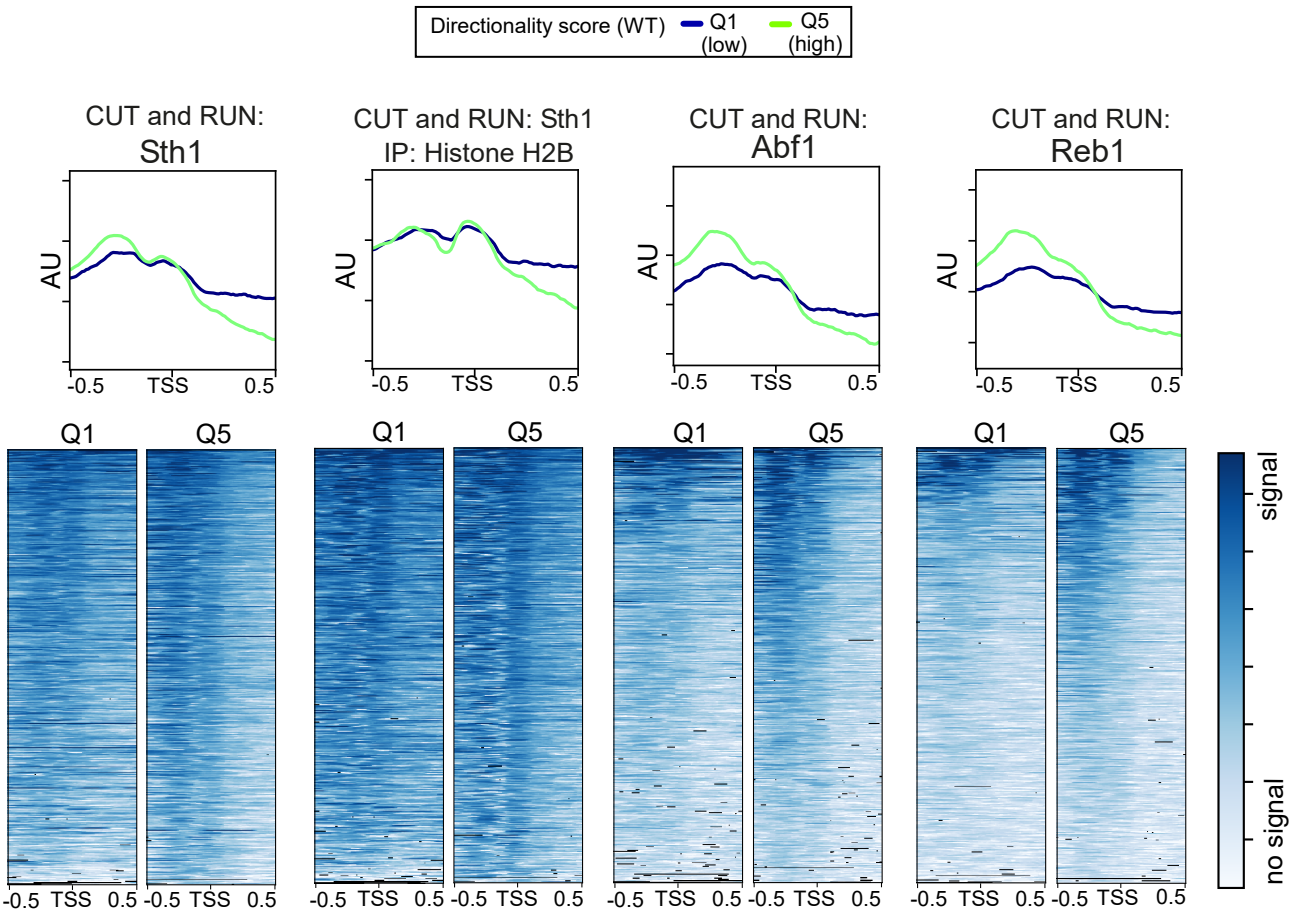

A
C
STH1-AID
RAP1-AID
RAP1-AID STH1-AID
IAA
Sth1-AID
Rap1-AID
Hxk1
wild type
Rap1 dpl.
(RAP1-AID + IAA)
Sth1 dpl.
(Sth1-AID + IAA)
Rap1 + Sth1 dpl.
(RAP1-AID + Sth1-AID + IAA)
+ spike in S. pombe
normalization:
1. S. cerevisiae (sc)
2. S. pombe (sp)
B
Rap1-regulated genes (n=141)
log2(FC, IAA/DMSO)
sc sp sc sp sc sp
Sth1 dpl.
Rap1 dpl.
Rap1 & Sth1 dpl.
C
norm.
0-700
WT
Rap1 dpl.
Sth1 dpl.
Rap1 + Sth1 dpl.
(-) strand
0-7000
(+) strand
IRT2
(div. RNA)
RPL43B
D
0-52
0-14000
0-550
0-9700
div. RNA
RPS3
YOR292C
RPS10
E
S. cerevisae normalized
S. pombe normalized
Sth1 depletion
Rap1 depletion
Sth1+Rap1 depletion
anti-sense
c1
c2
c3
-350 Rap1 +350
log2(FC, IAA/DMSO)
-2 0 2
F
S. cerevisae normalized
S. pombe normalized
AS S
Sth1 depletion
Rap1 depletion
Sth1+Rap1 depletion


A
Sth1 depletion
Rap1 depletion
Sth1+Rap1 depletion
Antisense
Sense
c1
c2
c3
-350 Rap1 +350
log2(FC, IAA/DMSO)
-2 0 2
B
200 bp
ChrXVI
Sense
Antisense
YPL150W
ATG5
Directionality Score = log10 (Sense / Antisense)
ChrXI
YPT52
Gene
Directionality Score
ATG5 1.13
YPT52 0.36
C
Tandem Genes
Rap1-regulated Genes
nr genes
Directionality Score
D
control
Rap1 depletion
E
Sth1 depletion
F
Sth1+Rap1 depletion
G
Rap1 depletion
Sth1+Rap1 depletion

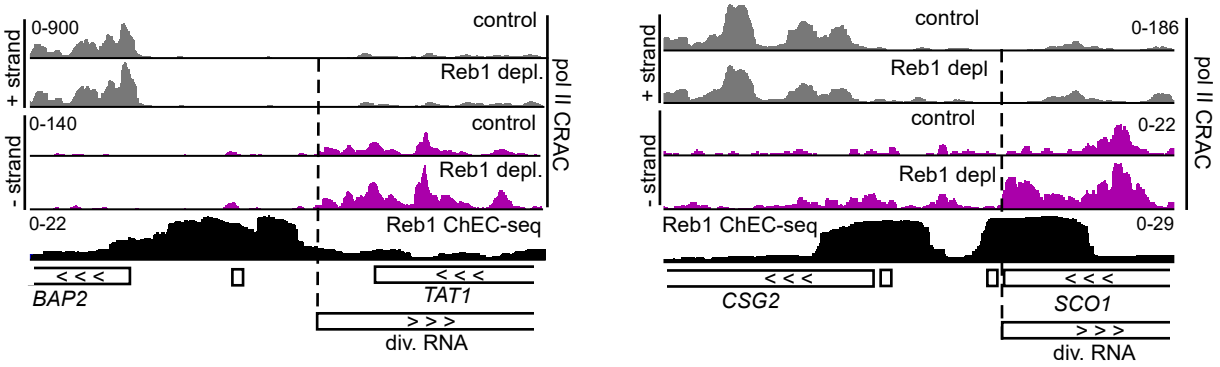

A
WT
control
Sth1 depl.
0-1700
(+) strand
(-) strand
0-250
MNase -seq
control
Sth1 depl.
0-25
TDA4
STE24
div. txpn
0-66
0-1400
0-31
HOM6
HIR3
B
Protein coding genes
UP 298
DOWN 567
SUTs
UP 202
DOWN 104
XUTs
UP 479
DOWN 103
CUTs
UP 218
DOWN 72
NUTs
UP 431
DOWN 213
-log10(padj)
log2(FC, IAA/DMSO))
log2(FC, IAA/DMSO)
C
Sth1 depletion
Antisense
Sense
ASc1
ASc2
ASc3
-500
TSS
+500
log2(FC, IAA/DMSO)
D
WT
control (STH1-AID +DMSO)
Sth1 depletion (STH1-AID +IAA)
gene count
Directionality score
E
control (DMSO/WT)
Sth1 depletion (IAA/WT)
log2(FC)
sense
div.

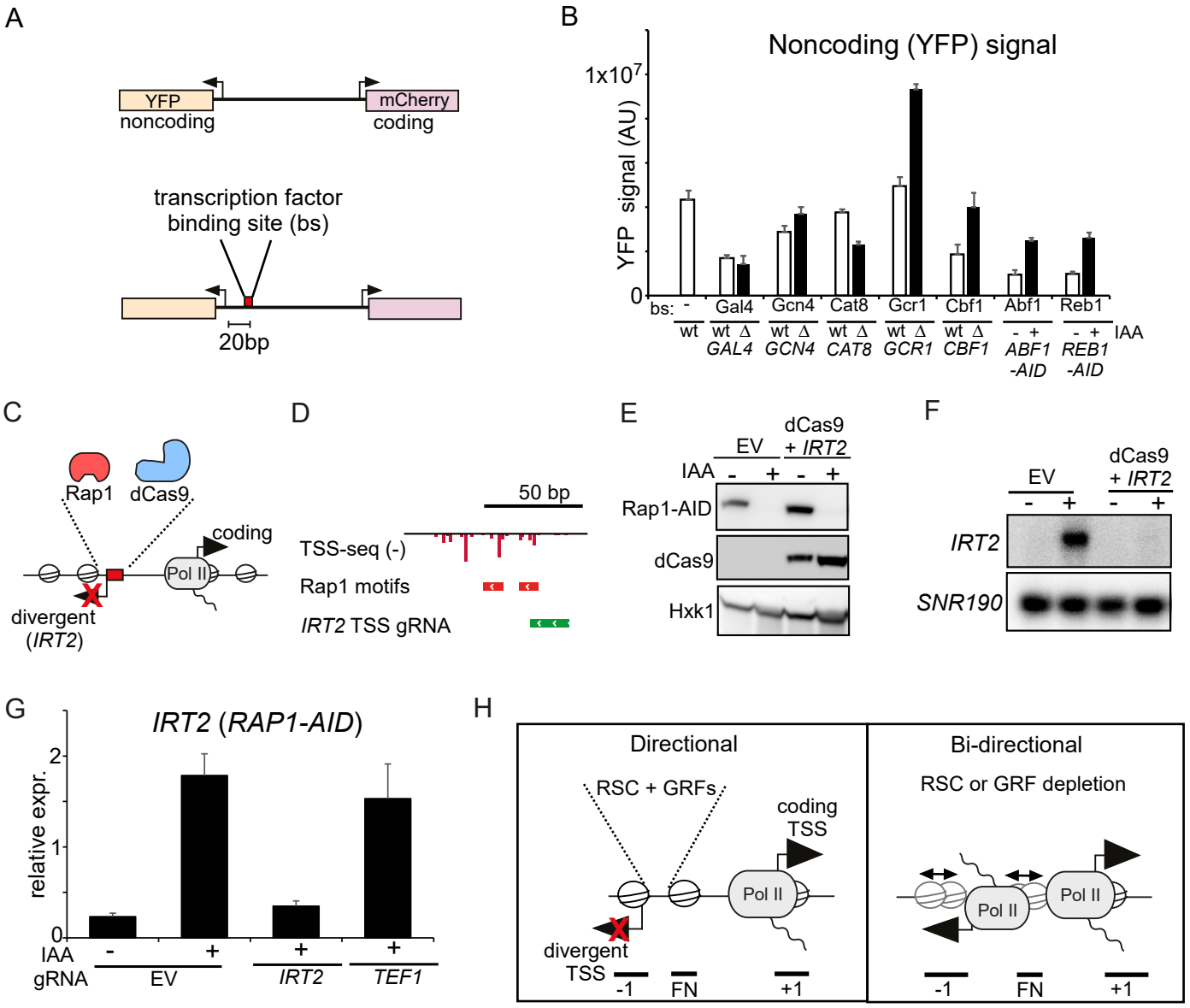

A
Sth1 depletion vs. directionality score (WT)
divergent nascent RNA log2(FC, IAA/DMSO)
sense nascent RNA log2(FC, IAA/DMSO)
directionality score log2(FC, IAA/DMSO)
Q1 Q2 Q3 Q4 Q5
lowest
highest
directionality score (WT)
B
Q5
log2(FC, IAA/DMSO)
sense nascent RNA (WT)
directionality score (WT)
C
Directionality score (WT)
HIGH MNase
LOW MNase
control
Sth1 depl.
Q5
Q1
normalized coverage
-2
-1
+1
-450 bp
TSS
D
Q3 Sense RNA levels (WT)
Sense RNA levels
lowest 20%
highest 20%
directionality score (WT)
E
Directionality score (WT) Q3:
highest 20%
lowest 20%
F
100 bp
WT
control
Sth1 depl.
+strand
-strand
high MNase
low MNase
PDE2
div. txpn
YBR238C
FN
G
Directionality score (WT)
Q1
Q5
TBP binding:
norm. cov.
-0.5 kb
0.5 kb
H
divergent direction
A-track
CG(C/G)G
AU

A

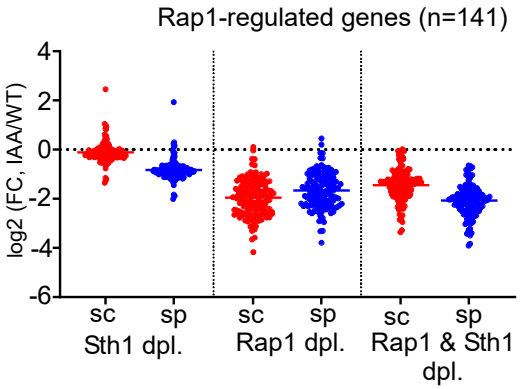

Directionality score (WT)
Q1 (low)
Q5 (high)
CUT and RUN: Sth1
CUT and RUN: Sth1 IP: Histone H2B
CUT and RUN: Abf1
CUT and RUN: Reb1
AU
-0.5
TSS
0.5
Q1
Q5
signal
no signal

B

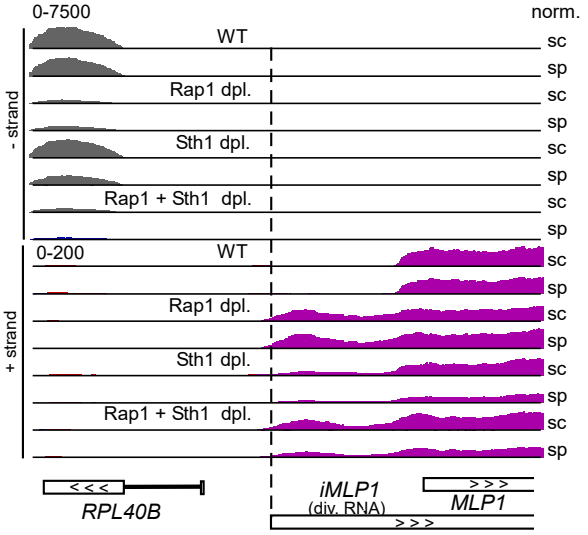

+ strand
- strand
0-900
control
Reb1 depl.
0-140
control
Reb1 depl.
pol II CRAC
0-22
Reb1 ChEC-seq
BAP2
TAT1
div. RNA
0-186
Reb1 depl
0-22
0-29
CSG2
SCO1

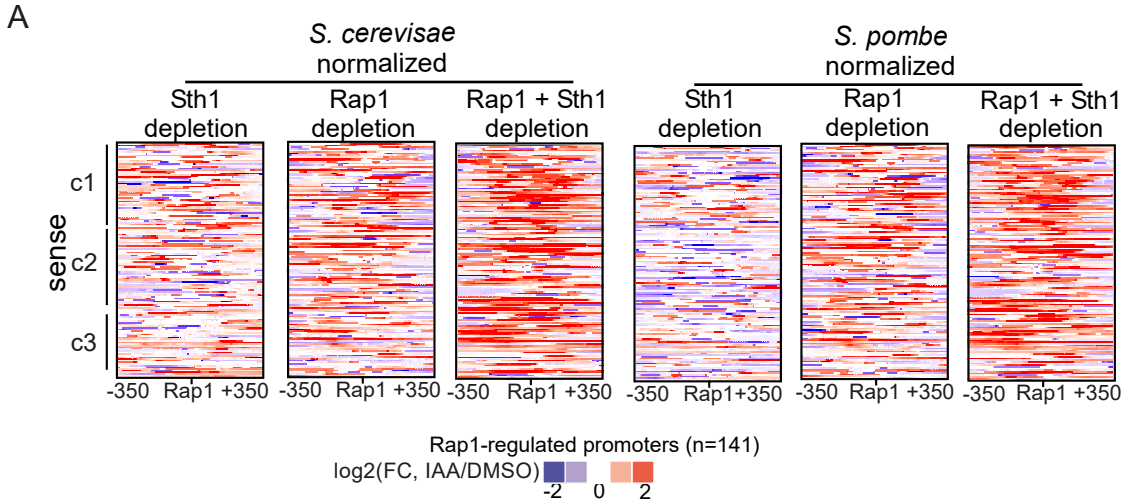

A
YFP
noncoding
mCherry
coding
transcription factor
binding site (bs)
20bp
B
Noncoding (YFP) signal
1x10^7
YFP signal (AU)
0
bs: -
Gal4
Gcn4
Cat8
Gcr1
Cbf1
Abf1
Reb1
wt
wt Δ
GAL4
GCN4
CAT8
GCR1
CBF1
- +
ABF1 -AID
REB1 -AID
IAA
C
Rap1
dCas9
coding
Pol II
divergent
(IRT2)
D
50 bp
TSS-seq (-)
Rap1 motifs
IRT2 TSS gRNA
E
EV
dCas9 + IRT2
IAA - + - +
Rap1-AID
dCas9
Hxk1
F
EV
dCas9 + IRT2
- + - +
IRT2
SNR190
G
IRT2 (RAP1-AID)
relative expr.
2
1
0
IAA -
gRNA
EV
+
IRT2
TEF1
H
Directional
RSC + GRFs
coding
TSS
Pol II
divergent
TSS
-1
FN
+1
Bi-directional
RSC or GRF depletion
Pol II
Pol II
-1
FN
+1

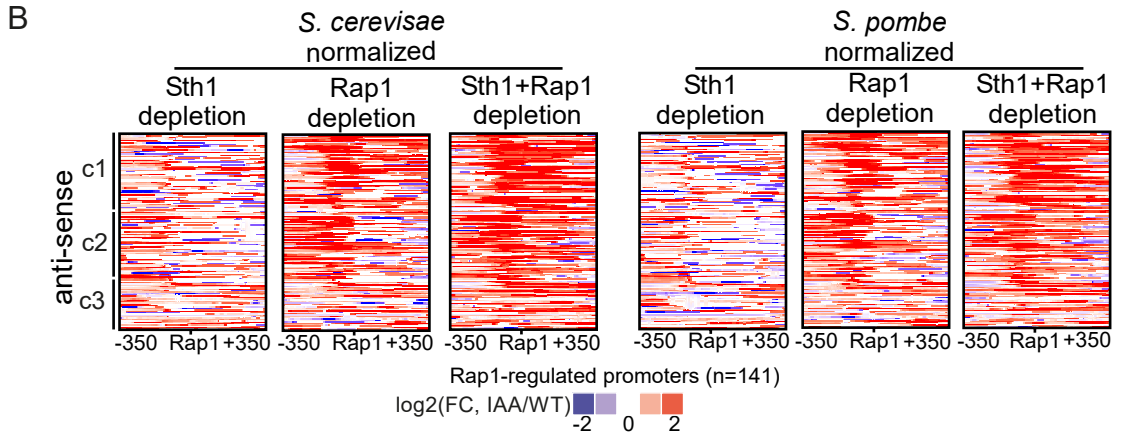
A

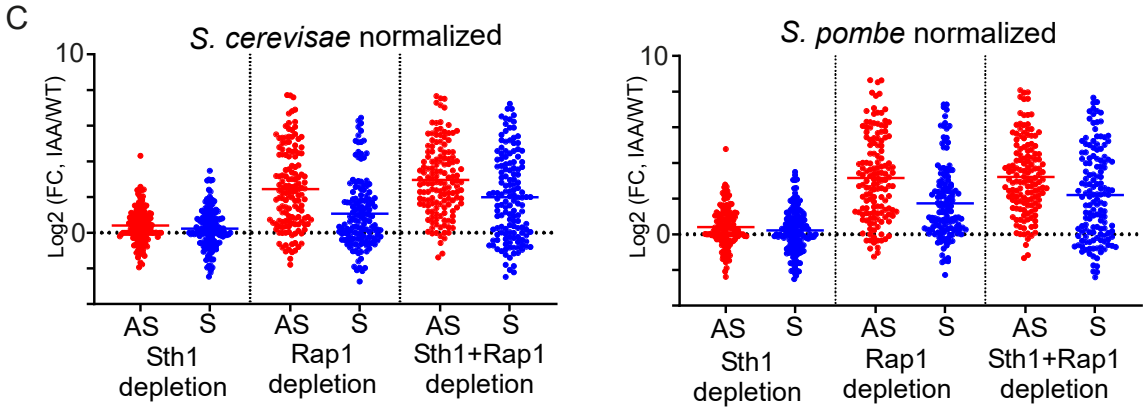

Rap1-regulated genes (n=141)
log2 (FC, IAA/WT)
4
2
0
-2
-4
-6
sc
sp
Sth1 dpl.
sc
sp
Rap1 dpl.
sc
sp
Rap1 & Sth1 dpl.

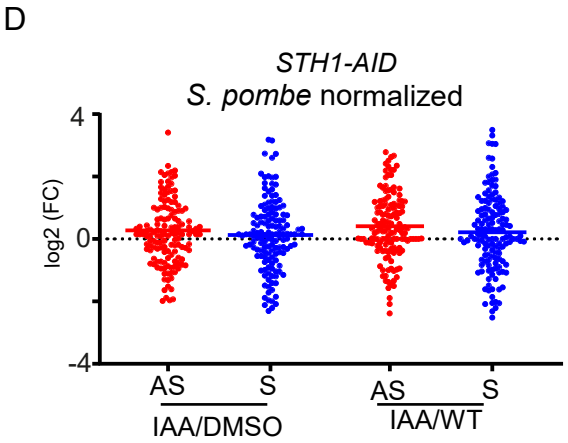
B


0-7500
norm.
WT
Rap1 dpl.
Sth1 dpl.
Rap1 + Sth1 dpl.
- strand
0-200
+ strand
sc
sp
RPL40B
iMLP1
(div. RNA)
MLP1

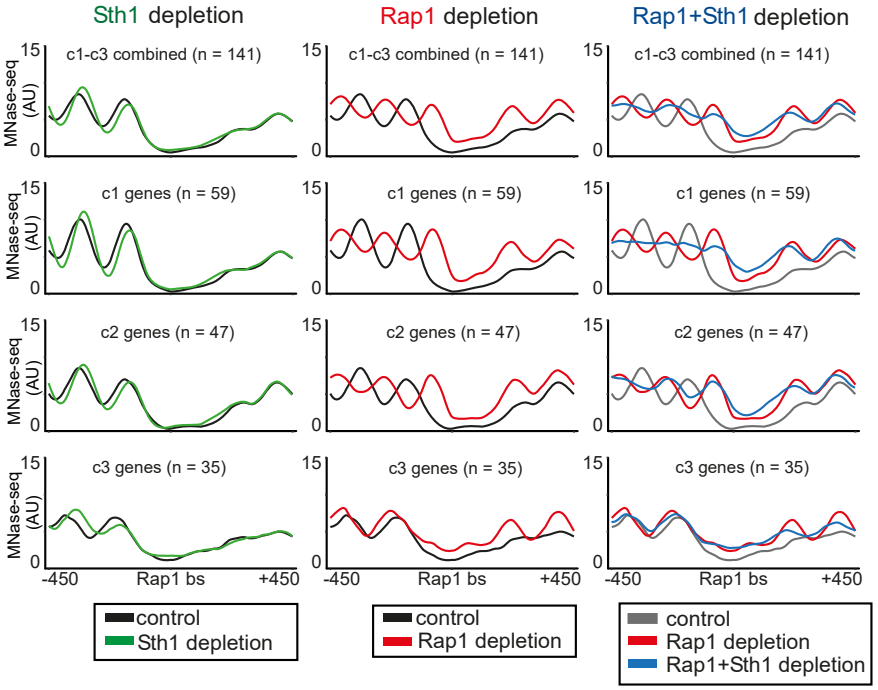

A
S. cerevisae normalized
S. pombe normalized
Sth1 depletion
Rap1 depletion
Rap1 + Sth1 depletion
sense
c1
c2
c3
-350 Rap1 +350
Rap1-regulated promoters (n=141)
log2(FC, IAA/DMSO)
-2 0 2


B
S. cerevisae normalized
S. pombe normalized
Sth1 depletion
Rap1 depletion
Sth1+Rap1 depletion
anti-sense
c1
c2
c3
-350 Rap1 +350
Rap1-regulated promoters (n=141)
log2(FC, IAA/WT)
-2 0 2

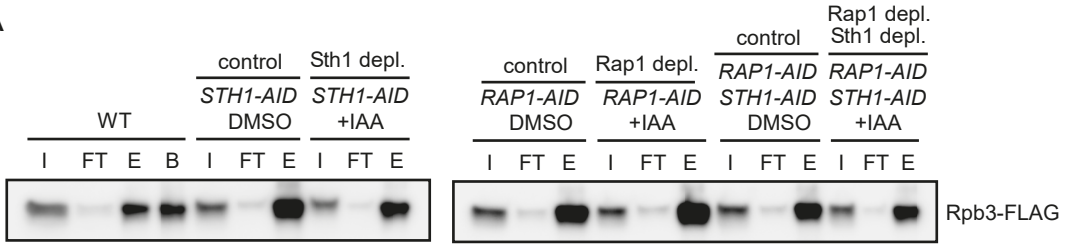

C
S. cerevisae normalized
S. pombe normalized
Log2 (FC, IAA/WT)
10
0
AS S
Sth1 depletion
Rap1 depletion
Sth1+Rap1 depletion


D
STH1-AID
S. pombe normalized
log2 (FC)
4
0
-4
AS S
IAA/DMSO
IAA/WT

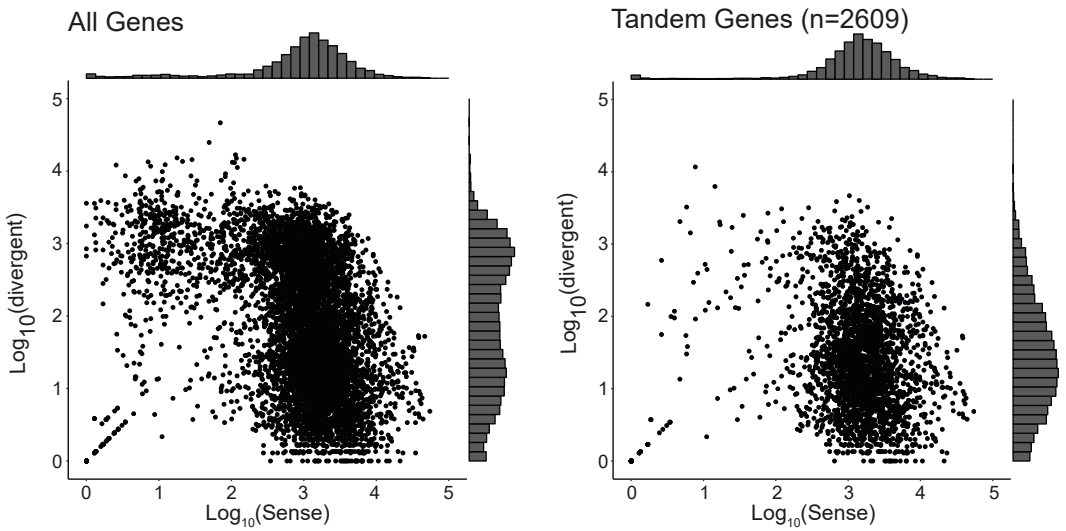
Wu *et al.* Figure S3



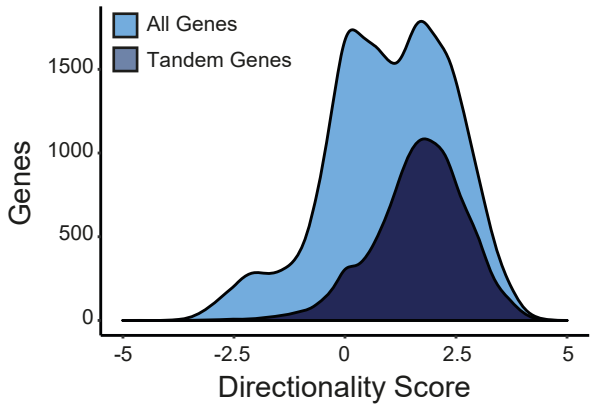
A


WT
control
STH1-AID
DMSO
Sth1 depl.
STH1-AID
+IAA
control
RAP1-AID
DMSO
Rap1 depl.
RAP1-AID
+IAA
control
RAP1-AID
STH1-AID
DMSO
Rap1 depl.
Sth1 depl.
RAP1-AID
STH1-AID
+IAA
I FT E B I FT E I FT E
I FT E I FT E I FT E I FT E
Rpb3-FLAG

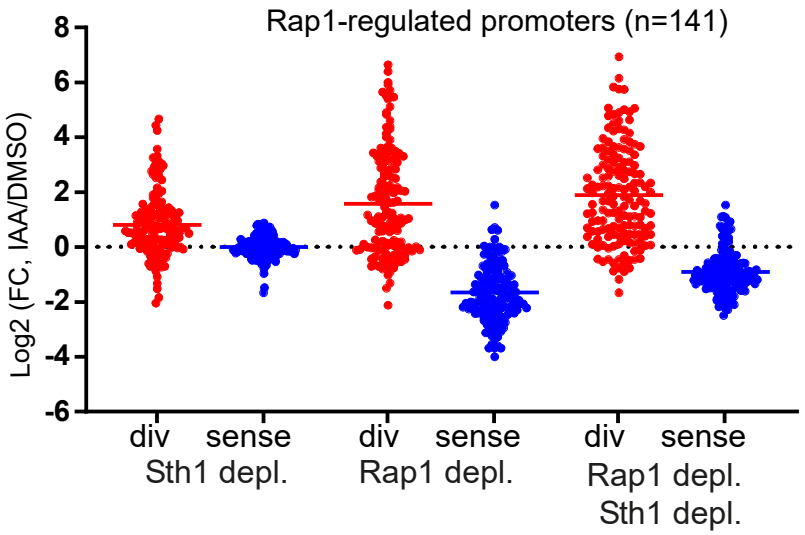
B


All Genes
Tandem Genes (n=2609)
Log10(divergent)
Log10(Sense)
0
1
2
3
4
5

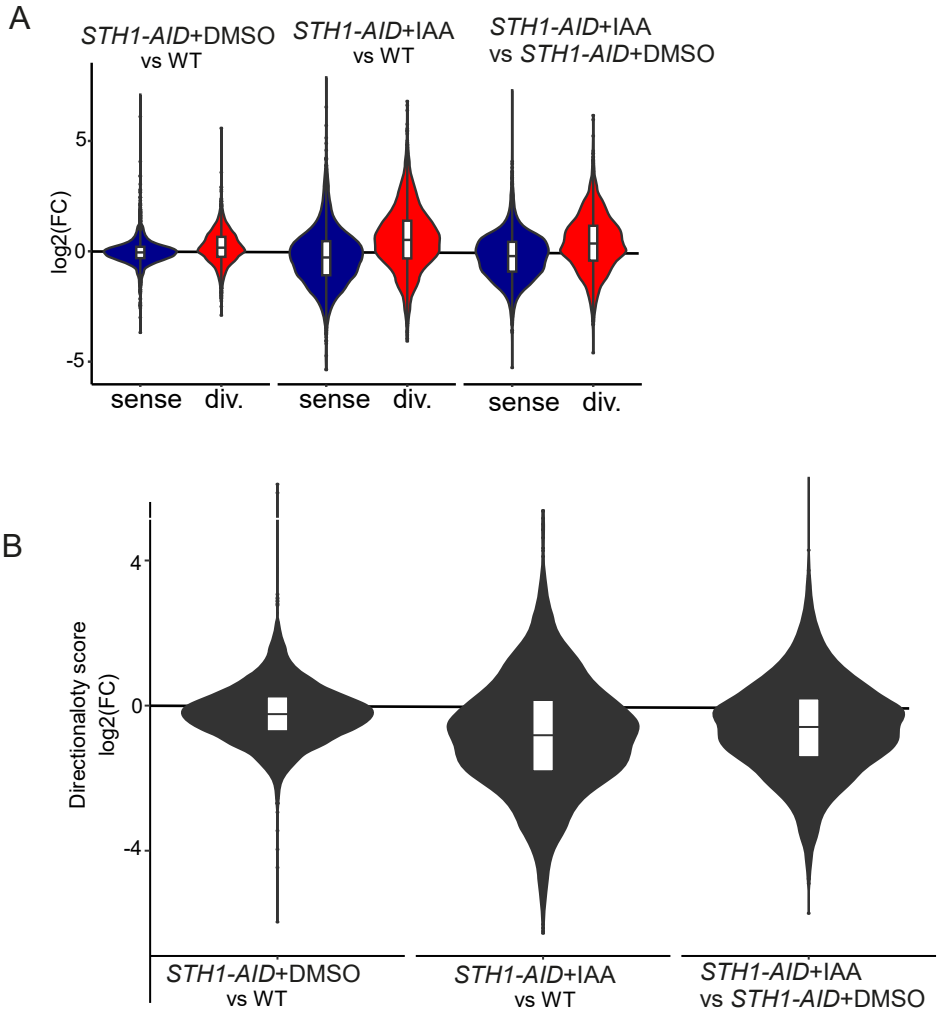
C

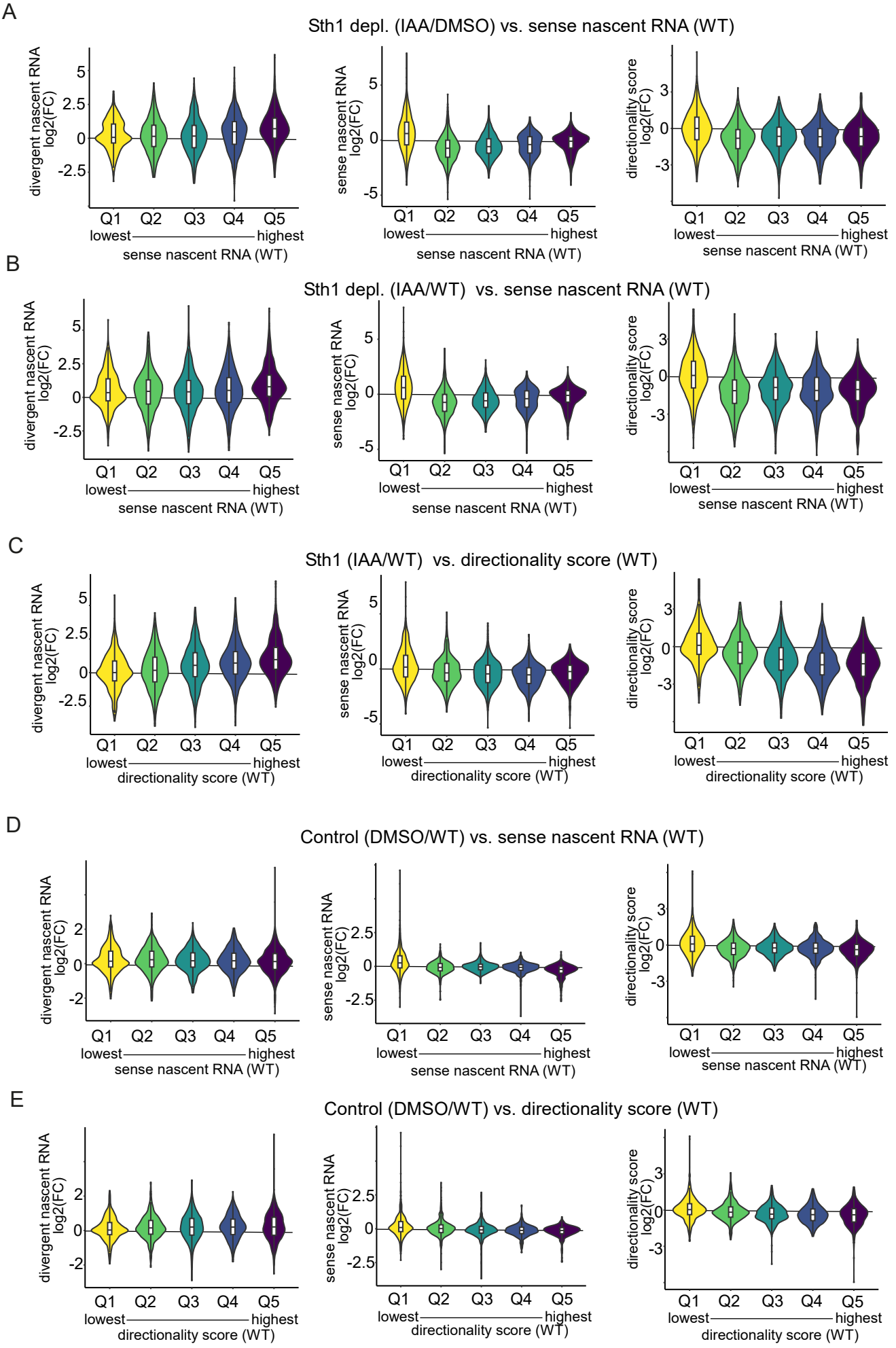

All Genes
Tandem Genes
Genes
1500
1000
500
0
-5
-2.5
0
2.5
5
Directionality Score

D

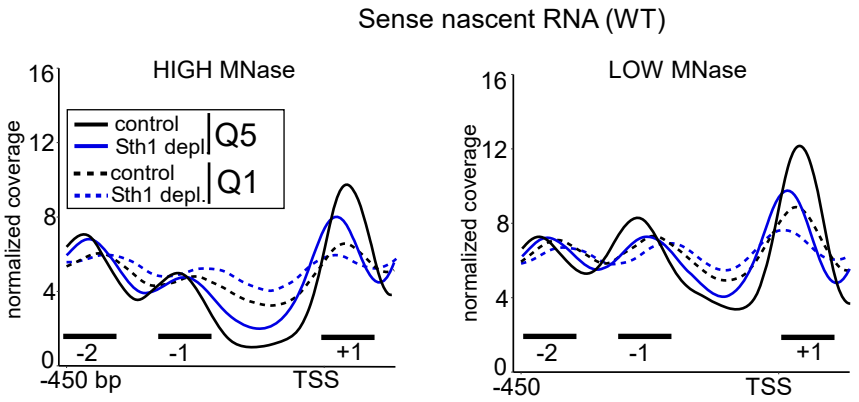

Rap1-regulated promoters (n=141)
Log2 (FC, IAA/DMSO)
8
6
4
2
0
-2
-4
-6
div
sense
Sth1 depl.
div
sense
Rap1 depl.
div
sense
Rap1 depl.
Sth1 depl.

Wu *et al.* Figure S5

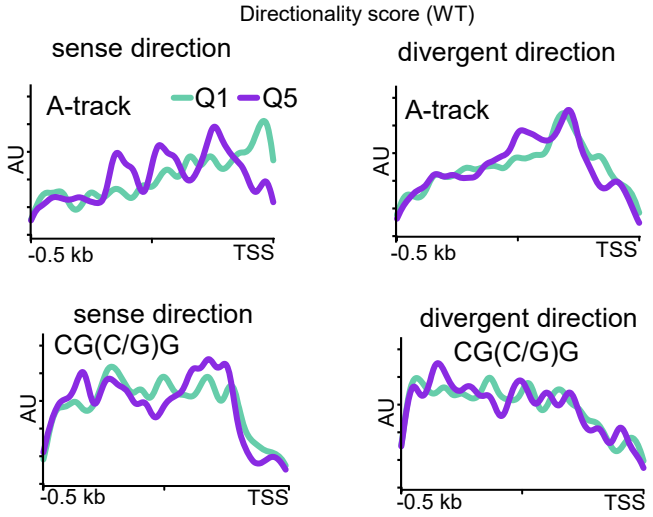



A
Sth1 depl. (IAA/DMSO) vs. sense nascent RNA (WT)
divergent nascent RNA log2(FC)
sense nascent RNA log2(FC)
directionality score log2(FC)
Q1 Q2 Q3 Q4 Q5
lowest highest
sense nascent RNA (WT)
B
Sth1 depl. (IAA/WT) vs. sense nascent RNA (WT)
C
Sth1 (IAA/WT) vs. directionality score (WT)
directionality score (WT)
D
Control (DMSO/WT) vs. sense nascent RNA (WT)
E
Control (DMSO/WT) vs. directionality score (WT)

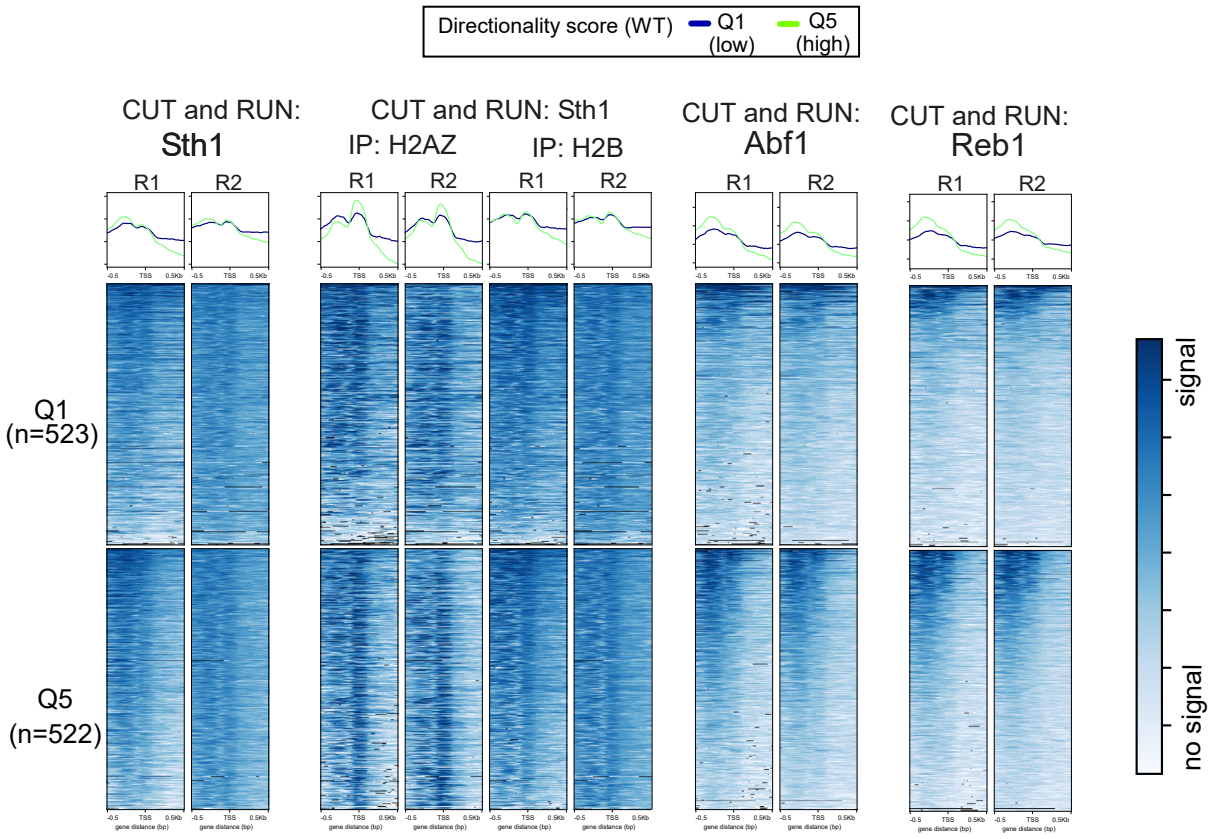
A


Sense nascent RNA (WT)
HIGH MNase
LOW MNase
normalized coverage
control
Sth1 depl
Q5
control
Sth1 depl.
Q1
16
12
8
4
0
-2
-1
+1
-450 bp
TSS
-450
TSS

B


Directionality score (WT)
sense direction
divergent direction
A-track
Q1
Q5
AU
-0.5 kb
TSS
CG(C/G)G

Wu *et al.* Figure S8



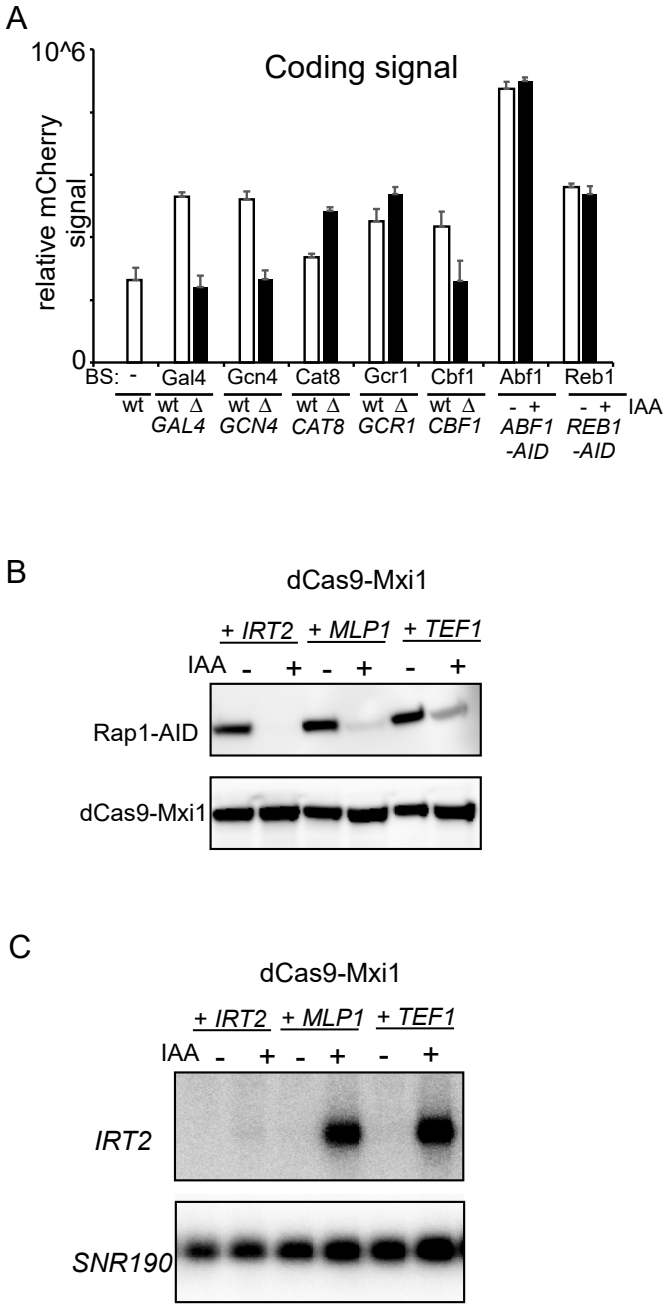

A
Coding signal
10^6
0
relative mCherry signal
BS: -
Gal4
Gcn4
Cat8
Gcr1
Cbf1
Abf1
Reb1
wt
wt Δ
GAL4
GCN4
CAT8
GCR1
CBF1
- +
ABF1 -AID
REB1 -AID
IAA
B
dCas9-Mxi1
+ IRT2
+ MLP1
+ TEF1
IAA - + - + - +
Rap1-AID
dCas9-Mxi1
C
dCas9-Mxi1
+ IRT2
+ MLP1
+ TEF1
IAA - + - + - +
IRT2
SNR190

Table S3. Genotypes of strains used in this study

Strain Number	Genotype
FW7228	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rpb3::RPB3-3xFLAG::NATMX*
FW7220	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rpb3::RPB3-3xFLAG::NATMX leu2::pGPD1-osTIR1::LEU2 sth1::STH1-V5-AID::KANMX6*
FW7232	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rpb3::RPB3-3xFLAG::NATMX leu2::pGPD1-osTIR1::LEU2 rap1::RAP1-V5-AID::KANMX6 sth1::STH1-V5-AID::KANMX6*
FW7238	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rpb3::RPB3-3xFLAG::NATMX leu2::pGPD1-osTIR1::LEU2 rap1::RAP1-V5-AID::KANMX6*
FW6404	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ppt1::p615-YFP-Cbf1_bs-pPPT1-mCherry::NATMX6*
FW6306	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 cbf1::KANMX ppt1::p615-YFP-Cbf1_bs-pPPT1-mCherry::NATMX6*
FW6401	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ppt1::p612-YFP-Gcn4_bs-pPPT1-mCherry::NATMX6*
FW6300	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gcn4::KANMX ppt1::p612-YFP-Gcn4_bs-pPPT1-mCherry::NATMX6*
FW6402	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ppt1::p613-YFP-Cat8_bs-pPPT1-mCherry::NATMX6*
FW6302	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 cat8::KANMX ppt1::p613-YFP-Cat8_bs-pPPT1-mCherry::NATMX6*
FW6403	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ppt1::p614-YFP-Gal4_bs-pPPT1-mCherry::NATMX6*
FW6424	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gal4::KANMX ppt1::p614-YFP-Gal4_bs-pPPT1-mCherry::NATMX6*
FW6405	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ppt1::p616-YFP-Gcr1_bs-pPPT1-mCherry::NATMX6*
FW6315	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gcr1::KANMX ppt1::p616-YFP-Gcr1_bs-pPPT1-mCherry::NATMX6*
FW6415	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 abf1::ABF1-V5-IAA7::KANMX6 leu2::pGPD1-osTIR1::LEU2 ppt1::p619-YFP-Abf1_bs-pPPT1-mCherry::NATMX6*

FW6411	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 reb1::REB1-V5-IAA7::KANMX6 leu2::pGPD1-osTIR1::LEU2 ppt1::p620-YFP-Reb1_bs-pPPT1-mCherry::NATMX6*
FW6407	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ppt1::p592-YFP-pPPT1-mCherry::NATMX6*
FW629	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*
FW8477	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rap1::RAP1-V5-IAA7::KANMX6 his3::pGPD1-osTIR1::HIS3 ura3::p375::URA3*
FW8523	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rap1::RAP1-V5-IAA7::KANMX6 his3::pGPD1-osTIR1::HIS3 ura3::p704::pTDH3-dCas9-3xFLAG-ADH1term::URA3 p692::pRS305-pSNR52-sgIRT2*
FW8531	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rap1::RAP1-V5-IAA7::KANMX6 his3::pGPD1-osTIR1::HIS3 ura3::p703::pTDH3-dCas9-Mxi1-3xFLAG-ADH1term::URA3 p692::pRS305-pSNR52-sgIRT2*
FW8527	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rap1::RAP1-V5-IAA7::KANMX6 his3::pGPD1-osTIR1::HIS3 ura3::p703::pTDH3-dCas9-3xFLAG-ADH1term::URA3 p705::pRS305-pSNR52-sgTEFI*
FW8535	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rap1::RAP1-V5-IAA7::KANMX6 his3::pGPD1-osTIR1::HIS3 ura3::p703::pTDH3-dCas9-Mxi1-3xFLAG-ADH1term::URA3 p352::pRS305-pSNR52-sgTEF1*
FW8529	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rap1::RAP1-V5-IAA7::KANMX6 his3::pGPD1-osTIR1::HIS3 ura3::p703::pTDH3-dCas9-Mxi1-3xFLAG-ADH1term::URA3 p706::pRS305-pSNR52-sgMLP1*

Table S4. Plasmids used in this study.

plasmid number	plasmid name
592	*YFP-pPPT1-mCherry::NATMX6*
615	*YFP-Cbf1_bs-pPPT1-mCherry::NATMX6*
612	*YFP-Gcn4_bs-pPPT1-mCherry::NATMX6*
613	*YFP-Cat8_bs-pPPT1-mCherry::NATMX6*
614	*YFP-Gal4_bs-pPPT1-mCherry::NATMX6*
616	*YFP-Gcr1_bs-pPPT1-mCherry::NATMX6*
619	*YFP-Abf1_bs-pPPT1-mCherry::NATMX6*
620	*YFP-Reb1_bs-pPPT1-mCherry::NATMX6*
247	*pNH605 pGPD1-osTIR1 LEU2*
250	*pNH603 pGPD1-osTIR1 HIS3*
227	*pWG444 (plasmid for making NAT deletion), NAT marker*
627	*3xFLAG-NATMX*
704	*pTDH3-dCas9-3xFLAG-ADH1term*
703	*pTDH3-dCas9-Mxi1-3xFLAG-ADH1term*
692	*pRS305-pSNR52-sgIRT2*
706	*pRS305-pSNR52-sgMLP1*
352	*pRS305-pSNR52-sgTEF1*
252	*pFA6A-V5-IAA7::KANMX6*

Table S5. Oligonucleotide sequences used in this study.

IME1_NB_probe_R	ATGCAACGCCTACTTGTTTT	*IRT2* northern blot probe
IME1_NB_probe_F	GATGGAGGGTTGGCATAAAA	*IRT2* northern blot probe
SNR190_NB_probe_F	GGCCCTGATGATAATG	*SNR190* northern blot probe
SNR190_NB_probe_R	GGCTCAGATCTGCATG	*SNR190* northern blot probe
IRT2_RT_F	GACATCCGCATTCTTGCAGC	*IRT2* for RT-qPCR
IRT2_RT_R	CATGCTGTTCTTTCCGCCAC	*IRT2* for RT-qPCR
ACT1_RT_F	GTACCACCATGTTCCCAGGTATT	*ACT1* RT-qPCR
ACT1_RT_R	AGATGGACCACTTTCGTCGT	*ACT1* RT-qPCR
IRT2_gRNA_F	**ATGATATGTAG**TTTTAGAGCTAGAAATAGCAAGTTAAA	*IRT2* gRNA primer for cloning, highlighted is the gRNA sequence used
IRT2_gRNA_R	**TTAGATTATT**GATCATTTATCTTTCACTGCGGA	*IRT2* gRNA primer for cloning, highlighted is the gRNA sequence used
MLP1_gRNA_F	**GAGGCACGGAG**TTTTAGAGCTAGAAATAGCAAGTTAAA	*MLP1* gRNA primer for cloning, highlighted is the gRNA sequence used
MLP1_gRNA_R	**TAATAGATTT**GATCATTTATCTTTCACTGCGGA	*MLP1* gRNA primer for cloning, highlighted is the gRNA sequence used
Reb1bs_top	TCGGGTAACTTATCGGGTAACAT	Oligo used for cloning Reb1 binding site
Reb1bs_bottom	ATGTTACCCGATAAGTTACCCGA	Oligo used for cloning Reb1 binding site
Abf1bs_top	TTATCACTTCCCACGATTATCACTTCCCACGATT	Oligo used for cloning Abf1 binding site
Abf1bs_bottom	AATCGTGGGAAGTGATAATCGTGGGAAGTGATAA	Oligo used for cloning Abf1 binding site
Gal4bs_top	ACGGATTAGAAGCCGCCGAGCGGGCGACAGCCCTCCGACGGAAGACTCTCCTCCGT	Oligo used for cloning Gal4 binding site
Gal4bs_bottom	ACGGAGGAGAGTCTTCCGTCGGAGGGCTGTCGCCCGCTCGGCGGCTTCTAATCCGT	Oligo used for cloning Gal4 binding site

Gcn4bs_top	TATGACTCATTCTATGACTCATTC	Oligo used for cloning Gcn4 binding site
Gcn4bs_bottom	GAATGAGTCATAGAATGAGTCATA	Oligo used for cloning Gcn4 binding site
Cbf1bs_top	T CACGTGA TCACGTGA T CACGTGA	Oligo used for cloning Cbf1 binding site
Cbf1bs_bottom	TCACGTG ATCACGTGA TCACGTG A	Oligo used for cloning Cbf1 binding site
Gcr1bs_top	TGGAAGCCTTGGAAGCCTTGGAAGCCT	Oligo used for cloning Gcr1 binding site
Gcr1bs_bottom	AGGCTTCCAAGGCTTCCAAGGCTTCCA	Oligo used for cloning Gcr1 binding site
Cat8bs_top	CCTTTAAGCCGACCTTTAAGCCGACCTT TAAGCCGA	Oligo used for cloning Cat8 binding site
Cat8bs_bottom	TCGGCTTAAAGGTCGGCTTAAAGGTCG GCTTAAAGG	Oligo used for cloning Cat8 binding site