

# Genetic and pharmacological evidence for kinetic competition between alternative poly(A) sites in yeast

Rachael Emily Turner<sup>1</sup>, Paul F Harrison<sup>1,2</sup>, Angavai Swaminathan<sup>1</sup>, Calvin A Kraupner-Taylor<sup>1</sup>, Belinda J Goldie<sup>1</sup>, Michael See<sup>1,2</sup>, Amanda L Peterson<sup>3</sup>, Ralf B Schittenhelm<sup>4</sup>, David R Powell<sup>2</sup>, Darren J Creek<sup>3</sup>, Bernhard Dichtl<sup>5\*</sup>, Traude H Beilharz<sup>1\*</sup>

<sup>1</sup>Development and Stem Cells Program, Monash Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia; <sup>2</sup>Monash Bioinformatics Platform, Monash University, Melbourne, Australia; <sup>3</sup>Drug Delivery, Disposition and Dynamics, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Australia; <sup>4</sup>Monash Proteomics & Metabolomics Facility, Department of Biochemistry and Molecular Biology, Monash Biomedicine Discovery Institute, Monash University, Melbourne, Australia; <sup>5</sup>School of Life and Environmental Sciences, Deakin University, Geelong, Australia

**Abstract** Most eukaryotic mRNAs accommodate alternative sites of poly(A) addition in the 3' untranslated region in order to regulate mRNA function. Here, we present a systematic analysis of 3' end formation factors, which revealed 3'UTR lengthening in response to a loss of the core machinery, whereas a loss of the Sen1 helicase resulted in shorter 3'UTRs. We show that the anti-cancer drug cordycepin, 3' deoxyadenosine, caused nucleotide accumulation and the usage of distal poly(A) sites. Mycophenolic acid, a drug which reduces GTP levels and impairs RNA polymerase II (RNAP II) transcription elongation, promoted the usage of proximal sites and reversed the effects of cordycepin on alternative polyadenylation. Moreover, cordycepin-mediated usage of distal sites was associated with a permissive chromatin template and was suppressed in the presence of an *rpb1* mutation, which slows RNAP II elongation rate. We propose that alternative polyadenylation is governed by temporal coordination of RNAP II transcription and 3' end processing and controlled by the availability of 3' end factors, nucleotide levels and chromatin landscape.

**\*For correspondence:**

bernhard.dichtl@deakin.edu.au (BD);  
traude.beilharz@monash.edu (THB)

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 29

**Received:** 01 December 2020

**Accepted:** 22 June 2021

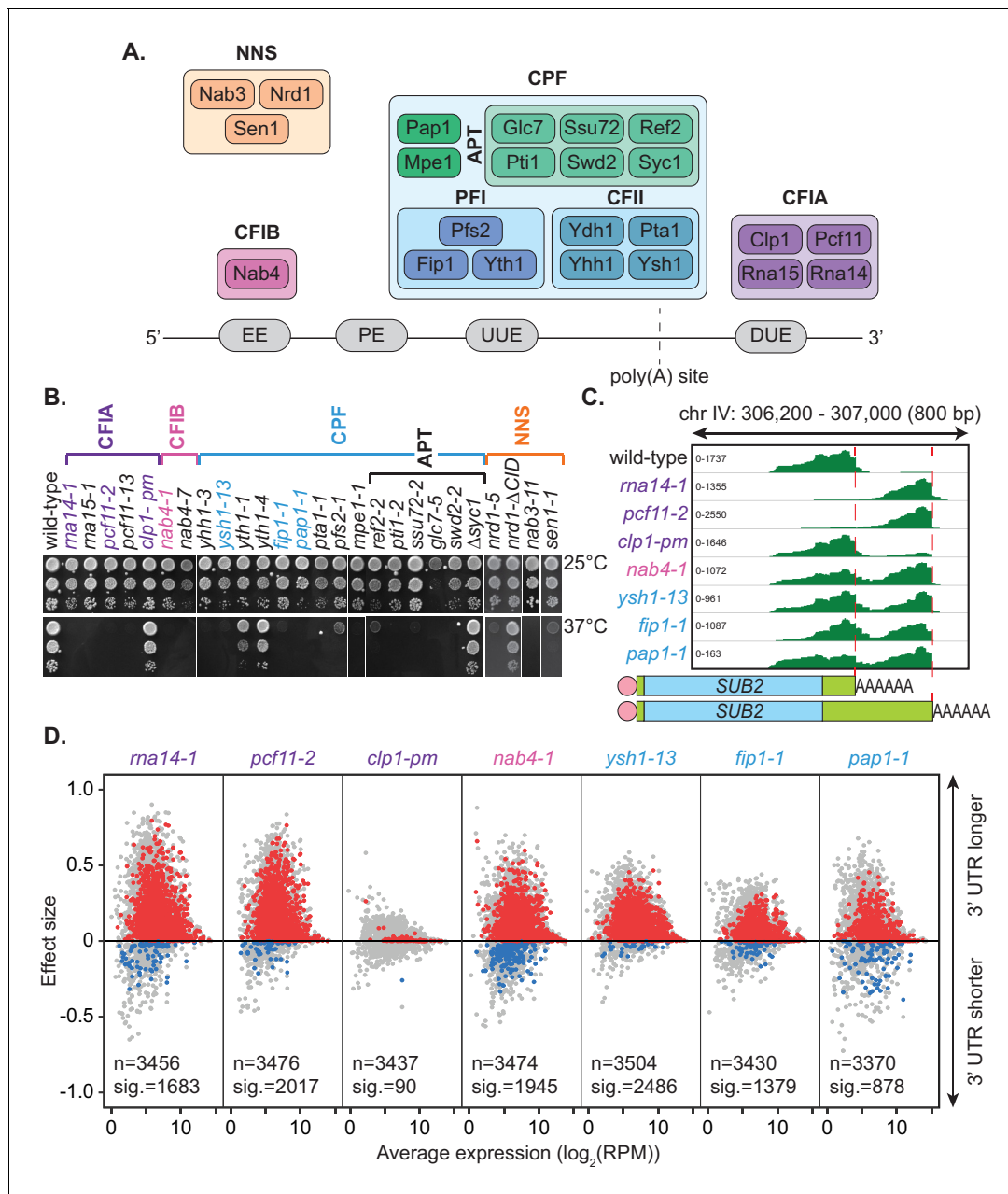
**Published:** 07 July 2021

**Reviewing editor:** Eric J Wagner, University of Texas Medical Branch at Galveston, United States

© Copyright Turner et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

3' end processing is a critical step in eukaryotic messenger RNA (mRNA) maturation. This two-step process involves co-transcriptional endonucleolytic cleavage of the pre-mRNA transcript and the subsequent addition of a non-templated polyadenosine (poly(A)) tail by poly(A) polymerase (*Colgan and Manley, 1997*). Cleavage site positioning is defined by a series of sequence elements within the pre-mRNA transcript that are recognised by the multi-subunit complexes that form the 3' end processing machinery. These cleavage and polyadenylation factors are highly conserved between eukaryotes; however, some differences in proteins and organisation exist between organisms. In the budding yeast *Saccharomyces cerevisiae*, three subcomplexes make up the core 3' end processing machinery. These include cleavage and polyadenylation factor (CPF), cleavage factor IA (CFIA) and cleavage factor IB (CFIB) (*Zhao et al., 1999; Mandel et al., 2008*).



**Figure 1.** Mutation of the yeast cleavage and polyadenylation machinery subunits confers changes in cleavage site choice. (A) Schematic of the yeast core cleavage and polyadenylation machinery and their interactions with the nascent mRNA strand. Cleavage and polyadenylation factor (CPF), Cleavage factor IA (CFIA) and Cleavage factor IB (CFIB) are involved in canonical 3' end formation whereas the Nrd1-Nab3-Sen1 (NNS) complex represents an alternative cleavage pathway utilised by snRNAs, snoRNAs, CUTs and some mRNAs. (B) Yeast strains, as indicated, were serially diluted 10-fold onto YPAD rich media plates and grown at 25°C or 37°C for 2 days. (C) Integrative Genomics Viewer (IGV) (Robinson et al., 2011) representation of PAT-seq reads aligned to the 3' end of the *SUB2* gene. Indicated yeast strains were grown at 25°C until OD<sub>600</sub> 0.6 then switched to 37°C for 1 hr. Peaks indicate nucleotide depth of coverage by sequencing reads and data ranges for each sample are shown on the left. The ratio of these peaks is used to calculate 3' end shifts. Schematic of 3' UTR length for each peak is shown below. (D) Global 3' end shift effects of genes observed in mutants of the indicated 3' end processing factor subunits. Values are relative to the wild-type strain W303 at 37°C using PAT-seq data. Each dot corresponds to one gene. A positive effect size indicates a general lengthening of the 3' UTR in the mutant relative to the wild-type, whereas a negative effect size indicates shifts towards shorter 3' UTRs. Grey dots are the estimated effect size. Where this is significant, a corresponding dot shows the 'confect', a confident inner bound of this effect size (FDR < 0.05) found using the Limma and TopConfacts R packages. Red indicates significant lengthening in the mutant whereas blue indicates significant shortening. The total number of detected genes (grey dots) and the number of significant shifts (red and blue dots) are indicated for each strain. Log<sub>2</sub>(RPM) is the logarithmic average number of reads per million. Data is representative of two biological replicates and is available in the **Figure 1—source data 1**.

Figure 1 continued on next page

Figure 1 continued

The online version of this article includes the following source data for figure 1:

**Source data 1.** 3' end processing mutant allele PAT-seq end shift values.

Together, these interact with the nascent transcript to promote both cleavage and polyadenylation (Figure 1A, Graber et al., 1999).

The 3' end processing signals involved in cleavage site recognition in yeast are generally more divergent than their mammalian counterparts and lack strong consensus sequences (Graber et al., 1999). However, in *S. cerevisiae*, four main sequence elements are associated with cleavage site positioning. Firstly, there is an A-rich positioning element (PE) located 10–30 nucleotides upstream of the cleavage site (Tian and Graber, 2012). This is the equivalent of the mammalian polyadenylation signal (PAS) and has the same optimal sequence of AAUAAA though this motif is less conserved in yeast (Zhao et al., 1999). Secondly, a UA-rich efficiency element (EE) occurs upstream of the PE sequence. The optimal sequence for the EE is UAUUA and is generally found 25–40 nucleotides away from the cleavage site (Tian and Graber, 2012). This sequence is not required for cleavage but improves the efficiency of the reaction and is important in site selection (Guo and Sherman, 1996). Lastly, U-rich elements exist either downstream (DUE) or upstream (UUE) of the cleavage site. Recombinant Yth1, Ydh1/Cft2, and Yhh1/Cft1 (all CPF subunits) as well as purified CPF interact with these U-rich elements (Barabino et al., 2000; Dichtl and Keller, 2001; Dichtl et al., 2002b) and mutation of these elements can result in reduced cleavage activity (Dichtl and Keller, 2001). Hrp1/Nab4, the sole CFIB subunit, interacts with the EE (Kessler et al., 1997; Chen and Hyman, 1998; Valentini et al., 1999). Rna15 (a CFIA subunit) has been found to interact with the PE motif in vitro (Gross and Moore, 2001), whereas in vivo mapping of factors on the nascent transcript suggested that binding of CFIA occurs downstream of the cleavage site (Baejen et al., 2014) similar to the CstF complex in the mammalian system (Chan et al., 2011; Mandel et al., 2008; Martin et al., 2012). Thus, the overall assembly of the 3' end formation machinery on the pre-mRNA substrate appears to be conserved between yeast and human (Baejen et al., 2014).

Pre-mRNA 3' end formation is physically and functionally coupled to RNAP II transcription. Assembly of 3' end factors on an emerging poly(A) site is accompanied by interactions with the carboxy terminal domain (CTD) of Rpb1, the largest subunit of RNAP II (McCracken et al., 1997). The CFIA subunit Pcf11 plays a central role in transmitting poly(A) site recognition to elongating RNAP II, triggering termination of transcription (Zhang et al., 2005; Sadowski et al., 2003; Barillà et al., 2001).

Many genes possess more than one potential site at which cleavage and polyadenylation may occur. This occurrence, termed alternative polyadenylation (APA), is pervasive in all studied eukaryotes and appears in up to 70% of human and yeast genes (Derti et al., 2012; Ozsolak et al., 2010). Use of such alternate cleavage sites has the potential to impact the stability, localisation and translational efficiency of an mRNA transcript. This has been linked to biological processes such as development and proliferation as well as diseases including cancer (Mayr and Bartel, 2009; Ji et al., 2009; Sandberg et al., 2008). However, the mechanisms that govern cleavage site choice are still under investigation.

Here, we used mutant alleles of yeast core cleavage and polyadenylation factors to probe their roles in alternative polyadenylation. Mutant phenotypes were linked to global preferred usage of distal cleavage sites and 3' UTR lengthening irrespective of subcomplex membership. Similarly, the adenosine analogue cordycepin (3' deoxyadenosine) was shown to interfere with correct 3' end processing and promote APA comparable to inactivation of the cleavage and polyadenylation machinery. This was linked to increased cellular nucleotides, and an enhanced transcription elongation rate in response to the drug treatment. In addition, the occurrence of APA following cordycepin treatment correlated with the size of intergenic regions and the size of nucleosome depleted regions after the stop codon in convergent genes. We propose that APA in yeast employs a balance between cleavage site strength, 3' end processing factor availability, and transcriptional rate to establish a kinetic control of poly(A) site choice.

## Results

### Mutations within the core cleavage and polyadenylation machinery promote lengthening of 3' UTRs

A compendium of *S. cerevisiae* mutants with defects in the cleavage and polyadenylation machinery has been collated from members of the yeast 3' end formation community (Table 1). This included strains with mutated subunits across the three core 3' end processing subcomplexes CFIA, CFIB and CPF, as well as the non-canonical Nrd1-Nab3-Sen1 (NNS) complex (Figure 1A). This repository combined with modern transcriptome-wide analyses has provided a unique opportunity to further understand how 3' end choices are globally regulated in yeast. We, therefore, sought to explore the role of individual cleavage factor subunits in cleavage site selection.

The majority of the cleavage and polyadenylation factors are essential for life necessitating the use of primarily temperature-sensitive alleles (Figure 1B). The phenotypes associated with mutant strains analysed here have been well described in the literature (Table 1). This includes primarily loss-of-function mutations, displaying distinct phenotypes in various steps of mRNA synthesis including pre-mRNA cleavage, pre-mRNA polyadenylation, cleavage site recognition, RNA binding, or transcription termination. For consistency, irrespective of the severity of the mutation, duplicates of each strain were grown in rich media at 25°C until an OD<sub>600</sub> of 0.6 was reached, then switched to a restrictive temperature (37°C) for 1 hr before harvesting.

A representative set of seven mutants was chosen for full transcriptome-wide APA analysis. This included three mutants from the CFIA complex, *rna14-1*, *pcf11-2*, and *clp1-pm*, a mutant of the single protein factor CFIB, *nab4-1*, and three mutants from the CPF complex, *ysh1-13*, *fip1-1*, and *pap1-1*. With the exception of *clp1-pm*, these were temperature-sensitive mutants (Figure 1B). Gene expression and cleavage site choice were analysed using the targeted poly(A) tail sequencing approach Poly(A)-Test Sequencing (PAT-seq) (Harrison et al., 2015). This is a 3' end-focused deep sequencing technique that involves biotin-labelled anchor oligonucleotide primers, limited digestion and gel size selection to enrich for the final 200 bp of polyadenylated transcripts as input into Illumina compatible sequencing libraries.

Sequencing data was processed using the tail-tools pipeline (<http://rnasystems.erc.monash.edu/software/>; Harrison et al., 2015). Sequencing reads were mapped to genes within the *S. cerevisiae* genome and the 3' end of transcripts were analysed for differences in positioning relative to the wild-type at 37°C. Figure 1C shows the 3' UTR shifts associated with inactivation of the 3' end processing machinery for the representative gene *SUB2*. Peak height reflects the number of transcripts ending at a certain position in the aligned sequence. Mutation of the *RNA14*, *PCF11*, *HRP1/NAB4*, *YSH1*, *FIP1*, and *PAP1* genes caused a switch to more distal cleavage site usage and longer 3' UTRs. This change occurred to the greatest extent in the *rna14-1* and *pcf11-2* mutants where reads mapping to the proximal site were rarely found. Little change was noticeable for the *clp1-pm* mutant, which contains point mutations in the P-loop motif (K136A and T137A) suggesting that ATP binding by Clp1 was dispensable for poly(A) site selection. Consistently, the mutations have previously been shown to have limited impact on pre-mRNA 3' end formation and termination in vitro (Holbein et al., 2011).

Of the 7091 genomic annotations identified by PAT-seq for our samples, 3511 were found to have two or more poly(A) sites within their 3' UTRs. Shift scores estimating the extent of 3' end shifting at 37°C in the seven mutants relative to the wild-type strain W303 were calculated for each of these genes. Effect size shift scores range from -1 to 1 with negative values indicating a shift to a more proximal cleavage site in the mutant and a positive shift score indicating a switch to a more distal cleavage site. An inner confidence bound, or 'confect', on these effect size scores was given for genes with significantly non-zero shifts (FDR < 0.05) using the topconfects R package (Harrison et al., 2019). Significant 3' UTR shifts were found in all seven mutants compared to the wild-type (Figure 1D). Excluding *clp1-pm*, 422 genes significantly changed cleavage sites in all mutants. The majority of observed shifts generated longer 3' UTRs upon mutation of the cleavage machinery indicating that functional attenuation of the individual subunits generally promoted an increase in 3' UTR length. *ysh1-13* had the highest number of significant shifts with 2486 genes significantly changing their cleavage site. Of these, 2457 switched to longer 3' UTRs compared to wild-type whilst only 29 became shorter. Ysh1, an endonuclease, is assumed to be the subunit responsible for cleaving the pre-mRNA transcript (Ryan et al., 2004; Mandel et al., 2006; Garas et al.,

**Table 1.** Yeast Strains.

Strain	Genotype	Relevant mutation if known	Reference
W303	MATa <i>leu2-3,112 his3-11,15 trp1-1 can1-100 ade2-1 ura3-1</i>	Wild type	<b>Thomas and Rothstein, 1989</b>
BY4741	MATa <i>leu2Δ0 his3Δ1 ura3Δ0 met15Δ0</i>	Wild type	<b>Brachmann et al., 1998</b>
<i>rna14-1</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 ade2-1 ura3-1 rna14-1</i>	T nucleotide insertion at 1986 causing a premature stop codon (replacement of last 16 amino acids with FK)	<b>Bloch et al., 1978</b> (named <i>cor1-1</i> ); <b>Minvielle-Sebastia et al., 1994</b> , <b>Rouillard et al., 2000</b>
<i>rna15-1</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 ade2-1 ura3-1 rna15-1</i>	L214P	<b>Bloch et al., 1978</b> (named <i>cor2-1</i> ); <b>Minvielle-Sebastia et al., 1994</b> , <b>Qu et al., 2007</b>
<i>pcf11-2</i>	MATa <i>leu2-3,112 his3-11,15 trp1Δ can1-100 ade2-1 ura3-1 pcf11-2</i>	E232G, D280G, C424R, S538G, F562S, S579P	<b>Amrani et al., 1997</b> , <b>Sadowski et al., 2003</b>
<i>pcf11-13</i>	MATa <i>leu2-3,112 his3-11,15 trp1Δ can1-100 ade2-1 ura3-1 pcf11-13</i>	D68A, S69A, I70A (CID region)	<b>Sadowski et al., 2003</b>
<i>clp1-pm</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 can1-100 ade2-1 ura3-1 clp1-pm</i>	K136A, T137A (P-loop motif)	<b>Ramirez et al., 2008</b>
<i>nab4-1</i>	MATa <i>ura3-52 nab4-1</i>	N167D, F179Y, P194H, Q265L	<b>Minvielle-Sebastia et al., 1998</b>
<i>nab4-7</i>	MATa <i>ura3-52 nab4-7</i>	S64R, K92M, T125A, I163T	<b>Minvielle-Sebastia et al., 1998</b>
<i>yhh1-3</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 can1-100 ade2-1 ura3-1 yhh1-3</i>	L302P, N716S, N762S, Y766C, K933R, D1070E, N1136D	<b>Dichtl et al., 2002b</b>
<i>ysh1-13</i>	MATa <i>leu2-3,112 his3-11,15 trp1Δ ade2-1 ura3-1 TRP1::ysh1 [ysh1-3-HIS3-CEN]</i>	V235A, N685H, D695V, E723V, R763G	<b>Garas et al., 2008</b>
<i>yth1-1</i>	MATa <i>leu2-3,112 his3-11,15 trp1Δ ade2-1 ura3-1 yth1::TRP1 [CEN4-ADE2-yth1-1]</i>	frameshift at 154 in causing a premature stop codon (deletion of last 55 amino acids)	<b>Barabino et al., 1997</b>
<i>yth1-4</i>	MATa <i>leu2-3,112 his3-11,15 trp1Δ ade2-1 ura3-1 yth1::TRP1 [CEN4-ADE2 -yth1-4]</i>	W70A	<b>Barabino et al., 2000</b>
<i>fip1-1</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 ura3-1 fip1-1</i>	L99F, Q216X (deletion of last 111 amino acids)	<b>Preker et al., 1995</b> , <b>Preker et al., 1997</b>
<i>pap1-1</i>	MATa <i>ade1 ade2 lys2 gal1? ura3-52 pap1-1</i>		<b>Patel and Butler, 1992</b>
<i>pta1-1</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 ade2-1 ura3-1 pta1-1</i>	Premature stop codon	<b>O'Connor and Peebles, 1992</b> , <b>Preker et al., 1997</b>
<i>pfs2-1</i>	MATa <i>leu2-3,112 his3-11,15 trp1Δ ade2-1 ura3-1 pfs2::TRP1 [pfs2-1-LEU2-CEN]</i>		<b>Ohnacker et al., 2000</b>
<i>mpe1-1</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 ade2-1 ura3-1 mpe1-1</i>	F9S, Q268K, K337F, K354X (deletion of last 87 amino acids)	<b>Vo et al., 2001</b>
<i>ref2-2</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 can1-100 ade2-1 ura3-1 ref2-2</i>		<b>Dheur et al., 2003</b>
<i>pti1-2</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 can1-100 ade2-1 ura3-1 pti1-2</i>		<b>Dheur et al., 2003</b>
<i>ssu72-2</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 ade2-1 ura3-1 ssu72-2</i>	R129A	<b>Pappas and Hampsey, 2000</b> , <b>Dichtl et al., 2002a</b>
<i>glc7-5</i>	MATa <i>leu2-3,112 his3-11,15 can1-100 ade2-1 ura3-1 glc7::LEU2 trp1:: glc7-5</i>	F226L	<b>Peggie et al., 2002</b>
<i>swd2-2</i>	MATa <i>leu2Δ0 his3Δ1 ura3Δ0 met15Δ0 swd2-2</i>	F14P, C27G, F41I, K185E, S253L, C257R	<b>Dichtl et al., 2004</b>
<i>Δsyc1</i>	MATa <i>leu2Δ0 his3Δ1 ura3Δ0 met15Δ0 syc1Δ</i>	Whole gene deletion	<b>Zhelkovsky et al., 2006</b>
<i>nrd1-5</i>	MATa <i>ura3-52 trp1-1 ade2-1 leu2-3,112 his3-11,15 lys2-Δ2 can1-100 met2-Δ1 nrd1-5</i>	G368V (RNA recognition motif)	<b>Steinmetz and Brow, 1996</b>
<i>nrd1-ΔCID</i>	MATa <i>ura3-52 trp1-1 ade2-1 leu2-3,112 his3-11,15 lys2-Δ2 can1-100 met2-Δ1 nrd1-ΔCID</i>	deletion of amino acids 39–169 (CTD region)	<b>Steinmetz and Brow, 1998</b>
<i>nab3-11</i>	MATa <i>leu2-3,112 his3-11,15 trp1-1 can1-100 ade2-1 ura3-1 nab3-11</i>	F371L, P374L (RNA recognition motif)	<b>Conrad et al., 2000</b>

Table 1 continued on next page

Table 1 continued

Strain	Genotype	Relevant mutation if known	Reference
sen1-1	MATa <i>leu2-3,112 his3-11,15 trp1-1 ade2-1 ura3-1 sen1-1</i>	G1747D (helicase domain)	Winey and Culbertson, 1988, Mischo et al., 2011
rbp1-1	MATa <i>ura3-52 leu2-3,112 rpb1-1</i>		Nonet et al., 1987

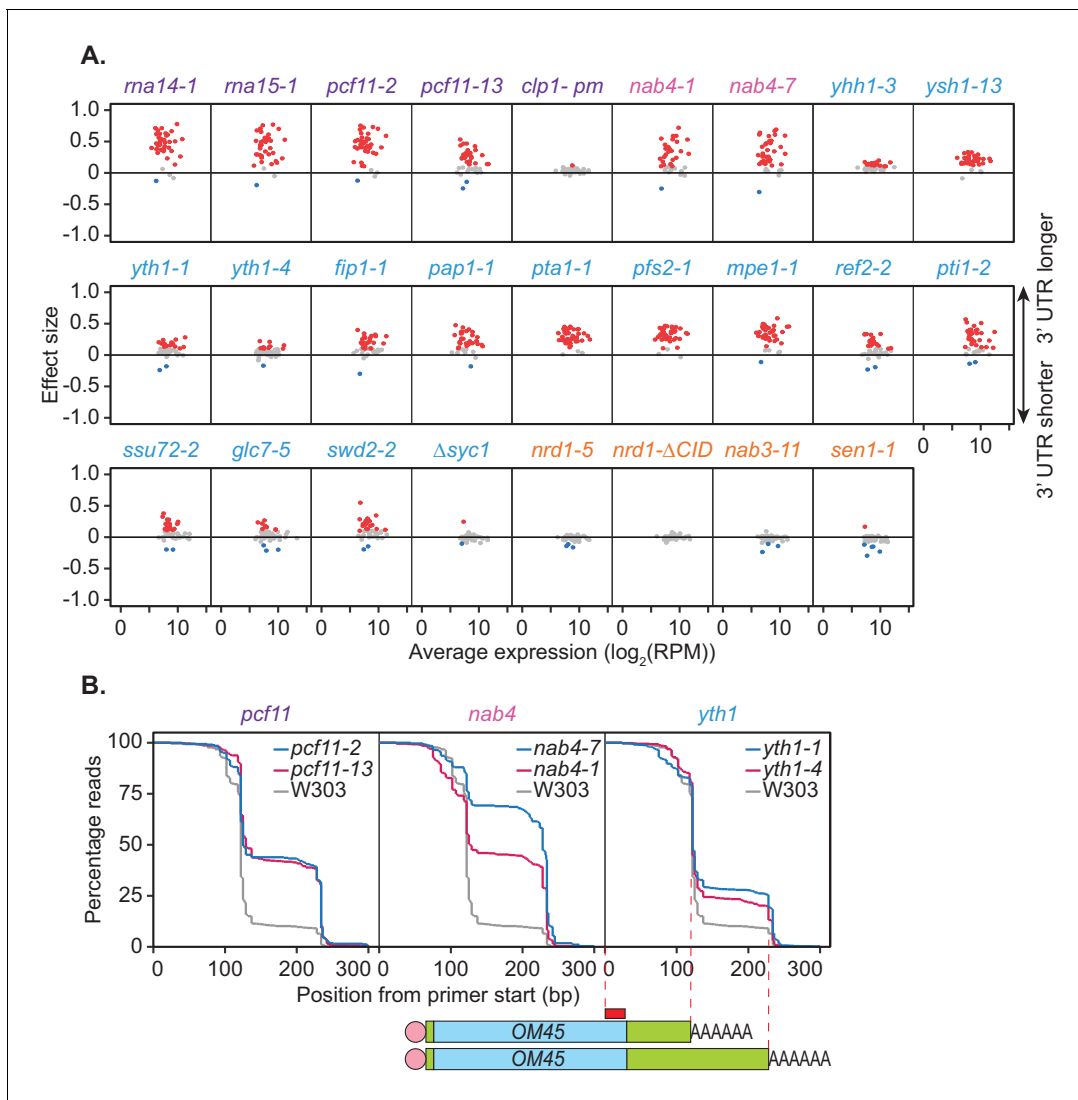
2008) and *ysh1* mutants may cause the proximal site to be skipped due to inefficiency of the actual cleavage reaction.

Despite more genes being affected in the *ysh1-13* strain, *rna14-1*, *pcf11-2*, and *nab4-1* strains all showed greater extents of shifting in impacted genes with absolute average confect values of 0.150, 0.125, and 0.093, respectively, compared to 0.086 in *ysh1-13*. The extent of the shift reflects the ratio of proximal to distal site usage. This suggests a key role for the CFIA and CFIB complexes in cleavage site selection. The role of these subunits in alternative polyadenylation may be attributed to a decrease in interactions with the poly(A) site sequences. It is also important to note that though the extent of switching varies in different mutant strains, this may be due to the severity of the mutation on protein function rather than an actual difference in the mechanisms of cleavage site recognition. Overall, inactivation of the yeast 3' end processing machinery appears to promote switching to more distal cleavage sites and higher expression of longer 3' UTR with little specificity for individual subunits.

To ensure the cleavage factor mutants examined were indeed representative of the general effect 3' end processing machinery impairment has on APA, a subset of shifts was analysed in the full panel of mutants. This was done using the multiplexed poly(A) test (mPAT) method (Beilharz et al., 2017), which is a targeted RNA sequencing adaptation of the extension poly(A) test (ePAT) assay (Jänicke et al., 2012) compatible with the Illumina MiSeq platform. The technique utilises a nested PCR to sequentially incorporate P5 and P7 extensions onto gene-specific PCR amplicons for sequencing on the Illumina flow-cell.

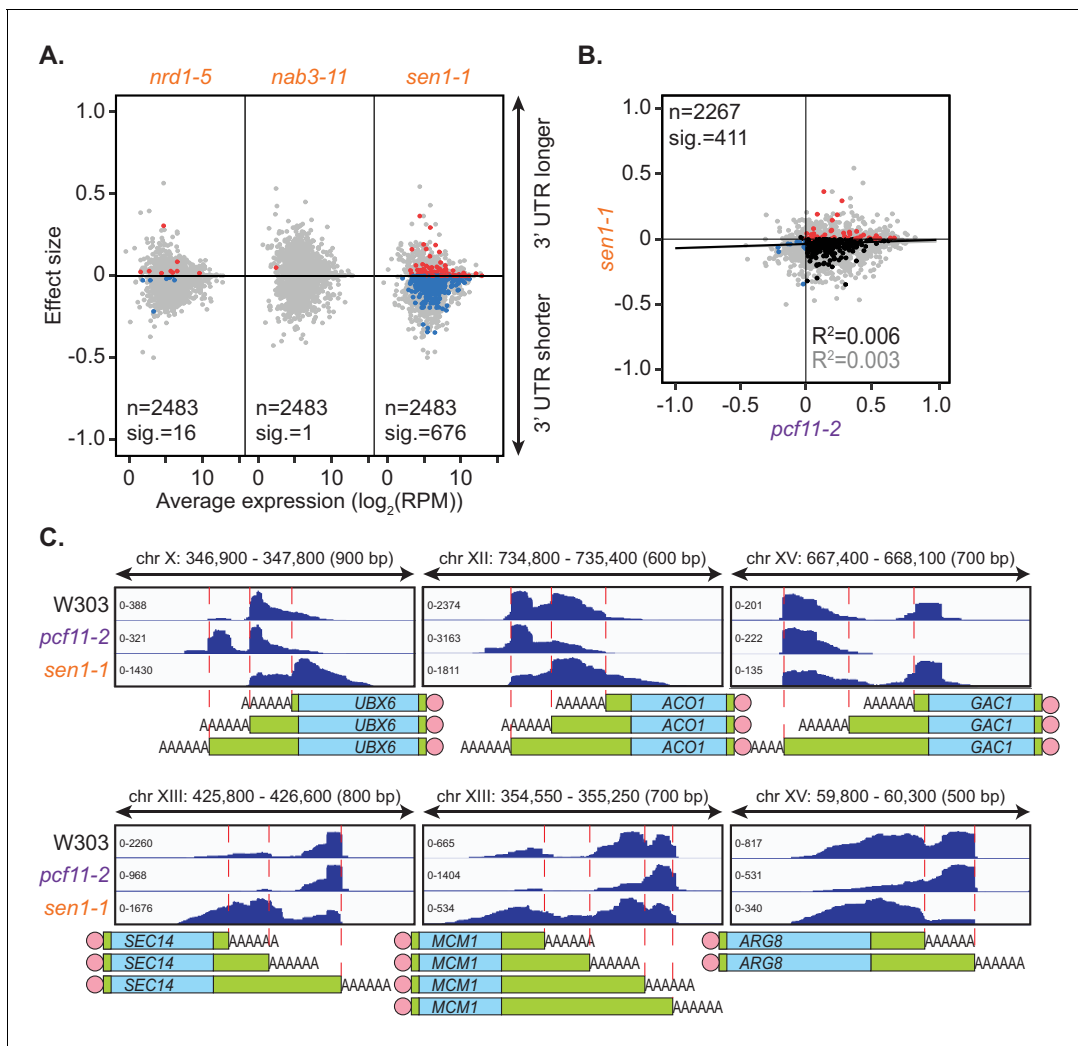
Forty-one genes were chosen for APA analysis, 39 of which were shown to significantly undergo alternative polyadenylation in the previous PAT-seq experiment. Cleavage site choice in each mutant strain was compared to that of the wild-type W303 at 37°C and shift scores calculated for each gene (Figure 2A). For the CFIA, CFIB, and CPF complexes, mutation of all but two subunits caused changes to cleavage site choice in at least a subset of the chosen genes. This was generally a switch to a more distal cleavage site and longer 3' UTR usage. However, very little shifting was seen for the *clp1-pm* and  $\Delta$ *syc1* mutants. *Syc1* is the only non-essential subunit of the core cleavage and polyadenylation complexes (Figure 1B, Zhelkovsky et al., 2006) indicating that its role in cleavage site choice may be negligible.

In an attempt to further dissect functional requirements in APA, we analysed additional mutant alleles of selected factors. Reverse cumulative distribution plots of the 3' ends of read alignments were used to visualise the percentage of reads terminating at each cleavage site of the gene *OM45* (Figure 2B). *nab4-1* and *nab4-7* mutants display defects in in vitro polyadenylation but function in in vitro cleavage (Minvielle-Sebastia et al., 1998). In our in vivo analyses both alleles were defective in *OM45* APA demonstrating a correlation with polyadenylation but not cleavage activity. The requirement for Hrp1/Nab4 in APA likely involves its role in poly(A) site selection (Minvielle-Sebastia et al., 1998). In vitro, the *yth1-1* and *yth1-4* alleles both have defects in the polyadenylation step of 3' end formation but only *yth1-4*, and not *yth1-1*, is defective in 3' end cleavage (Barabino et al., 2000). Additionally, in vivo APA defects did not correlate with cleavage activity, since both *YTH1* alleles showed comparable defects in APA (Figure 2B). The *pcf11-2* allele is defective in both cleavage and polyadenylation in vitro, while the *pcf11-13* allele is not (Sadowski et al., 2003). However, the latter contains mutations within the domain that interacts with the carboxyl-terminal domain (CTD) of Rpb1, the largest subunit of RNAP II (Barillà et al., 2001; Sadowski et al., 2003) and has formerly displayed defects in transcription termination (Sadowski et al., 2003). As such our APA results indicated that the interaction between Pcf11 and the transcriptional machinery was integral to APA in vivo. Overall, the results shown in Figure 2B suggested that defective in vitro cleavage activity was not indicative of defective APA in vivo. In contrast, defects in APA correlated with defective in vitro polyadenylation and defective interaction with the RNAP II CTD.



**Figure 2.** Mutations within the yeast cleavage and polyadenylation machinery subunits promote lengthening of mRNA 3' UTRs. (A) 3' end shift effects of genes observed upon mutation of the 3' end processing factor subunits. Values are relative to the wild-type strain W303 at 37°C using mPAT data. Each dot corresponds to a one gene as above. Grey dots indicate an effect size between  $-0.1$  and  $0.1$ , red dots indicate an effect size greater than  $0.1$  and blue dots indicate an effect size less than  $-0.1$ . Data is representative of two biological replicates. (B) Comparison of 3' end shifting in different alleles of the same 3' end processing factor. Reverse cumulative distribution plots show raw templated nucleotide length of reads aligned to the gene *OM45* from mPAT data, starting from the forward primer position as illustrated (red box) in the schematic. Mutant alleles for a single cleavage factor are shown within the same plot in comparison to the wild-type strain W303.

Impairment of the NNS complex appeared to have little impact on cleavage site selection in this analysis. However, as only a subset of protein-coding genes require the NNS complex for transcriptional termination (Steinmetz et al., 2006; Schulz et al., 2013; Steinmetz et al., 2001), genes in which this complex plays a role may have been excluded in the mPAT analysis. To address this possibility, global APA was analysed in one mutant strain for each of the NNS subunits using the PAT-seq method as above (Figure 3A). *nrd1-5* and *nab3-11* had little to no defect in cleavage site choice with only 16 and 1 genes, respectively, shifting significantly compared to the wild-type. Interestingly, reduced function of the helicase *Sen1* caused significant switching to shorter 3' UTRs for 676 genes. Of 411 genes that significantly switched cleavage site in both *sen1-1* and *pcf11-2*, 81.5% were in opposing directions (Figure 3B,C). Therefore, it appears that *Sen1* activity was required to antagonise cleavage site choice by the core 3' end processing factors at a subset of genes. This was consistent with the previous report that a *sen1-E1597K* mutant causes premature termination and



**Figure 3.** Sen1 mutation promotes 3' UTR shortening. (A) 3' end shift effects of genes observed upon mutation of the NNS subunits. Values are relative to the wild-type strain W303 at 37°C using PAT-seq data as in **Figure 1D**. Data is representative of two biological replicates and is available in the **Figure 3—source data 1**. (B) Comparison of 3' end shift effects of genes upon mutation of *PCF11* (x-axis) and *SEN1* (y-axis). *pcf11-2* and *sen1-1* shifts relative to their respective wild-type control at 37°C were compared. Grey dots indicate the estimated shift effects for all analysed features in common (n = 2267), black, red and blue dots indicate the significant shifts (n = 411, FDR < 0.05). Red indicates significant lengthening, blue indicates significant shortening and black indicates APA in opposite directions in both mutants. Lines show the linear regression for effect (grey) and confect (black) values respectively. (C) IGV representation of PAT-seq results aligned to the 3' end of the indicated genes in W303, *pcf11-2*, and *sen1-1* yeast strains at 37°C. Peaks indicate nucleotide depth of coverage by sequencing reads and data ranges for each sample are shown on the left. Schematic of 3' UTR length for each peak is shown below.

The online version of this article includes the following source data for figure 3:

**Source data 1.** NNS mutant allele PAT-seq end shift values.

shortening of the 3' UTR of the *NPL3* transcript (Steinmetz et al., 2006). We found that no significant changes in transcript levels were associated with the APA events in *sen1-1* strains (data not shown); however, a significant increase in gene coding region length was observed for *sen1-1* responsive genes (1270.5 bp in *sen1-1* APA genes and 986 bp in no APA genes, p-value=9.174×10<sup>-9</sup>). Overall, our observations are consistent with a wider role for this RNA helicase in antitermination of protein coding genes, contrasting its previously described roles in non-coding RNA termination (see Discussion).

In sum, these genetic data indicate that APA hinges on the level of functional machineries for cleavage and polyadenylation. However, given that each temperature sensitive allele has a unique



sensitivity and severity and therefore a different degree of apparent impact on 3' end choice, we next sought to explore the impact of cordycepin as a possible pharmacological driver of APA.

### Cordycepin induces lengthening of 3' UTRs

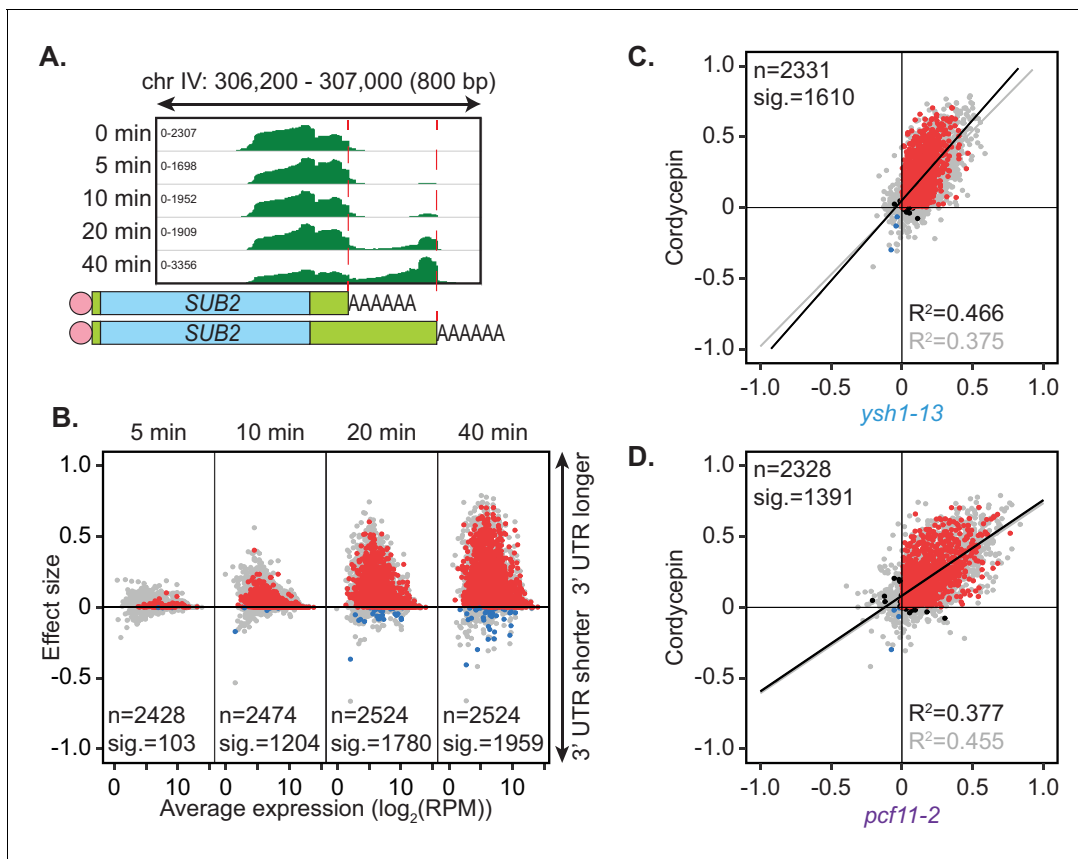
Cordycepin, or 3' deoxyadenosine, is an adenosine analogue that causes chain termination when incorporated into RNA during synthesis. We have previously demonstrated that independently of this role, cordycepin can interfere with normal 3' end formation for several genes in yeast (*Holbein et al., 2009*). We therefore further examined the effect of cordycepin on cleavage site choice genome-wide.

Wild-type BY4741 yeast cells were treated with a non-lethal dose of cordycepin (20 µg/ml) for up to 40 min. Alternative polyadenylation was assessed using the PAT-seq method over a time course of drug treatment. Similar to depletion of the canonical cleavage and polyadenylation machinery, cordycepin treatment promoted use of a distal cleavage site for the gene *SUB2* (*Figure 4A*). Significant switching to alternative cleavage sites was observed for 1959 genes with the majority changing to longer 3' UTRs (*Figure 4B*). Therefore, cordycepin treatment generally caused a switch to more distal cleavage sites. Of note, any RNA with cordycepin incorporated will lack a 3' hydroxyl group and will therefore not be detectable by the PAT-seq assay. However, such transcripts are expected to be unstable and quickly targeted for degradation if premature termination occurs within the coding region of the gene (*Kamieniarz-Gdula and Proudfoot, 2019*).

How cordycepin interferes with cleavage site choice remains unknown. Genes affected by cordycepin treatment were therefore compared to those affected by mutations within the cleavage and polyadenylation machinery (*Figure 4—figure supplement 1*). The pattern of cleavage site switching in cordycepin treated cells was most similar to that of *pcf11-2* and *ysh1-13*, with 1391 and 1610 genes, respectively, being affected both by the drug and the mutation. However, there was also a group of 153 genes that had significant APA occurring only in the cordycepin-treated samples. This variation suggested that cordycepin did not alter cleavage site usage simply by interacting with or disabling the cleavage machinery. This is consistent with our observation above that mutants defective for in vitro cleavage retained proper APA in vivo (*Figure 2B*). Importantly, the genes that did significantly switch cleavage site in both cordycepin treated cells and the 3' end factor mutants, did so in the same direction with 1365 (98.1%) and 1592 (98.9%) of genes in quadrants 2 and 3 for cordycepin compared to *pcf11-2* and *ysh1-13*, respectively (*Figure 4C,D*). Overall, we conclude that cordycepin had a similar effect on cleavage site choice as mutations in the 3' end processing machinery but that it is unlikely to cause alternative polyadenylation through direct interaction with the 3' end formation factors.

Cordycepin treatment, however, had an impact on the expression of the cleavage and polyadenylation machinery (*Figure 4—figure supplement 2A*). The majority of the CPF complex subunits had decreased RNA expression after 40 min of drug treatment, whereas the CFIA and CFIB encoding genes *PCF11*, *RNA14*, and *HRP1/NAB4* showed increased expression. This may indicate the cell's attempt to return to wild-type cleavage positions, as these subunits were seen to play a key role in cleavage site selection. In addition, the NNS subunits *NRD1* and *SEN1* were upregulated. This might reflect a need to process the elevated level of cryptic unstable transcripts (CUTs) seen following cordycepin treatment (*Holbein et al., 2009*), or suggest a compensatory mechanism whereby cells under 3' end stress co-opt an alternative pathway to appropriately cleave mRNA transcripts (*Rondón et al., 2009*).

Cordycepin treatment also provided a convenient method to assess the implications of 3' UTR lengthening on translational efficiency within yeast. Polysome profiling of cells treated with cordycepin for 40 min demonstrated a decrease in polysome formation and, therefore, translation in response to cordycepin treatment (*Figure 4—figure supplement 2B*). This change could be due to stress from cordycepin's antifungal effect (*Sugar and McCaffrey, 1998*) or the global increase in 3' UTR length causing translation repression. Proteomic analysis of treated cells showed that this did not majorly transfer to global protein expression alterations with only five proteins significantly changing in expression (q-value 0.05, fold change > |1|) following cordycepin treatment when assessed using data-independent acquisition (DIA) mass spectrometry (*Figure 4—figure supplement 2C*). However, major changes in protein expression were not expected within this short time period of cordycepin treatment, and the lack of significant results may in part reflect the inherent inability of DIA mass spectrometry to assess newly translated proteins in isolation. Of note, the



**Figure 4.** Cordycepin induces lengthening of 3' UTRs in yeast cells. BY4741 cells were treated with 20  $\mu\text{g/ml}$  cordycepin for 0, 5, 10, 20, or 40 min and changes to 3' UTR length were observed. (A) IGV representation of PAT-seq reads aligned to the 3' end of the *SUB2* gene. Peaks indicate nucleotide depth of coverage by sequencing reads and data ranges for each sample are shown on the left. Schematic of 3' UTR length for each peak is shown below. (B) 3' end shift effects of genes observed upon cordycepin treatment. Values are relative to untreated cells using PAT-seq data as for **Figure 1D** with positive values indicating lengthening following drug treatment. Data is representative of two biological replicates and is available in the **Figure 4—source data 1**. (C) Comparison of 3' end shift effects of genes upon mutation of *YSH1* (x-axis) and after cordycepin treatment (y-axis). *ysh1-13* shifts were compared to W303 at 37°C and cordycepin treatment for 40 min was compared to untreated BY4741 cells at 30°C. Grey dots indicate the estimated shift effects for all analysed features in common ( $n = 2331$ ), black, red, and blue dots indicate the significant shifts ( $n = 1610$ ; FDR < 0.05). Red indicates significant lengthening, blue indicates significant shortening and black indicates APA in opposite directions for both the mutant and cordycepin treated cells. Lines show the linear regression for effect (grey) and confect (black) values, respectively. (D) Comparison of 3' end shift effects of genes upon mutation of *PCF11* (x-axis) and after cordycepin treatment (y-axis) as for C but with *pcf11-2* shifts compared to W303 at 37°C. The total number of analysed features in common was 2328 and the number of significant shifts was 1391 (FDR < 0.05).

The online version of this article includes the following source data and figure supplement(s) for figure 4:

**Source data 1.** Cordycepin treatment PAT-seq end shift values.

**Source data 2.** Cordycepin proteomics data.

**Figure supplement 1.** Cordycepin treatment and 3' end processing factor mutant alleles promote alternative polyadenylation in subsets of genes.

**Figure supplement 2.** Cordycepin treatment alters cleavage factor RNA expression and cellular translation.

cleavage factor Hrp1/Nab4 increased in abundance by 49% consistent with the upregulation of its mRNA and a cellular attempt to normalise 3' ends.

We previously showed that cordycepin-triphosphate (CoTP), rather than cordycepin, mediated the toxicity of the drug in yeast (**Holbein et al., 2009**). To address whether CoTP was also responsible for altered 3' end processing,  $\Delta\text{ado1}$  and  $\Delta\text{adk1}$  strains were subjected to cordycepin treatment (**Figure 4—figure supplement 2D**). Ado1, converts cordycepin to cordycepin monophosphate (CoMP), whereas Adk1 further converts CoMP to CoTP (**Lecoq et al., 2001; Konrad, 1988**). As  $\Delta\text{ado1}$  cells were resistant to cordycepin-induced alternative polyadenylation for the gene *OM14* and  $\Delta\text{adk1}$  cells were not, this suggested that CoMP, not CoTP, was the mediator of APA. This further separated the roles of cordycepin in RNA incorporation and toxicity from its effect on

alternative polyadenylation. In addition,  $\Delta nrt1$  cells conferred mild resistance to cordycepin-induced APA, suggesting that Nrt1, a high-affinity nicotinamide riboside transporter and low-affinity thiamine transporter (Enjo *et al.*, 1997; Belenky *et al.*, 2008), acted as a partially redundant transporter for cordycepin (Figure 4—figure supplement 2D).

Excess adenine has previously been shown to suppress cordycepin's effect on yeast cell growth (Holbein *et al.*, 2009). To determine whether adenine addition was also able to reduce cordycepin-induced APA, wild-type BY4741 cells were treated with cordycepin and adenine. However, adenine co-treatment did not return cells to wild-type cleavage site positions (Figure 4—figure supplement 2E). Furthermore, to ask whether other 3' deoxynucleosides can induce APA, cells were treated with the guanosine equivalent of cordycepin, 3' deoxyguanosine (Figure 4—figure supplement 2E). However, this analogue caused no changes to cleavage site choice further supporting the idea that cordycepin is unlikely to confer APA through RNA chain termination and indicates that the effect was specific to the adenosine analogue.

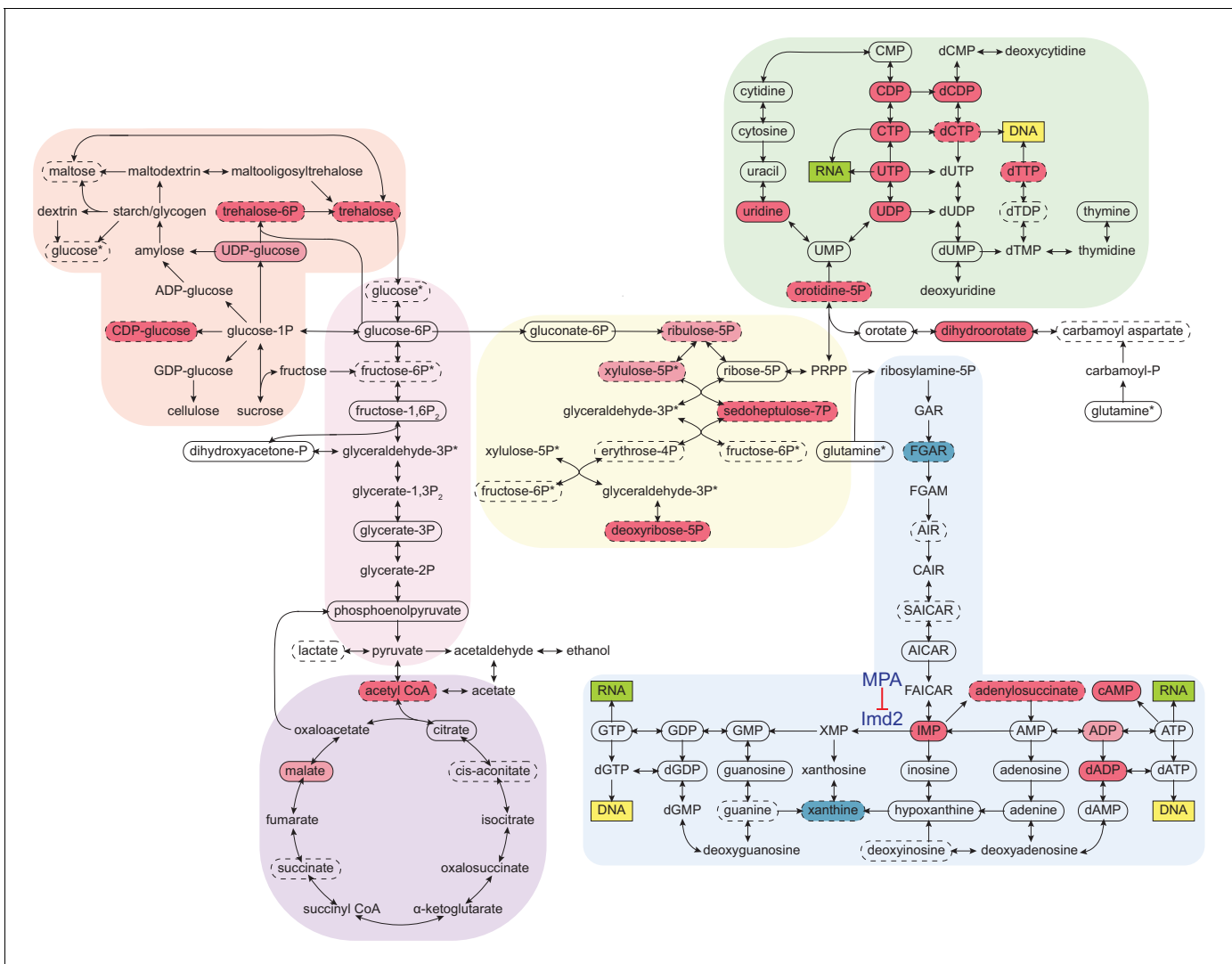
### Cordycepin alters yeast intermediate metabolism to promote nucleotide accumulation

The concentration of cordycepin used in our analyses did not significantly impair cell growth (Figure 4—figure supplement 2F). To further investigate the impact of the drug treatment, we analysed cellular metabolism. When grown in glucose-rich conditions, yeast preferentially ferments glucose to ethanol rather than perform aerobic respiration despite the presence of oxygen (Broach, 2012). To ask whether this was altered by cordycepin treatment, we first performed the Seahorse Energy Phenotyping assay on wild-type BY4741 yeast cells that were treated with cordycepin. This result was then compared with both DMSO treated and untreated cells. All three conditions responded to metabolic stress by increasing oxygen consumption, suggesting that cordycepin did not affect oxygen utilisation (Figure 5—figure supplement 1A). However, cordycepin-treated cells did not as efficiently undergo the increase in glycolytic output observed with both control groups (Figure 5—figure supplement 1B). In yeast, an increase in extracellular acidification is expected to report carbon dioxide production rather than lactic acid synthesis and indicates a decrease in overall ethanol synthesis in cordycepin-treated cells. As there was sufficient glucose in the media to support an increase in the control-treated cells, this suggests that metabolites of the glycolytic pathway of cordycepin-treated cells were diverted towards alternative pathways.

To probe the extent of metabolic rewiring by cordycepin, a metabolomic analysis was performed on cordycepin-treated cells. Several metabolites involved in glycolytically linked pathways were significantly upregulated following cordycepin treatment (Figure 5). The reserve carbohydrate, and stress protectant, trehalose was increased 3.8-fold, which correlated with drug-induced stress (Jain and Roy, 2009; Eleutherio *et al.*, 2015) and signifies amplified production of storage carbohydrates. This was accompanied by a 3.1-fold increase in acetyl CoA concentration and high levels of this metabolite are associated with increased acetylation of histones (Cai *et al.*, 2011; Donohoe *et al.*, 2012; Lee *et al.*, 2014) and therefore global transcriptional activity. Notably, we observed increased levels of key intermediates involved in nucleotide biosynthesis, including pentose phosphate pathway intermediates (sedoheptulose-7-phosphate and putative ribulose-5-phosphate and/or xylulose-5-phosphate), inosine monophosphate (IMP), adenylosuccinate, dihydroorotate, orotate monophosphate and several nucleotides and deoxynucleotides (CDP, CTP, UDP, UTP, dCDP, dCTP, dADP, and cAMP) indicating a major impact on nucleotide metabolism. Defective APA following cordycepin treatment was thus accompanied by a striking upsurge in nucleotide abundance, which is an important parameter for RNAP II transcription elongation rate *in vitro* (Uptain *et al.*, 1997).

### Slowing transcription attenuates cordycepin's effect on alternative polyadenylation

We and others have previously shown that the use of distal polyadenylation sites is linked to upregulation of transcriptional initiation (Swaminathan and Beilharz, 2016) and elongation rate (Pinto *et al.*, 2011). To test the idea that a change in nucleotide metabolism caused by cordycepin treatment leads to an increase in transcription rate and thus use of distal poly(A) sites, 4-thiouracil (4tU) labelling was used. This method monitors newly synthesised transcripts through the



**Figure 5.** Cordycepin alters yeast core metabolism to promote nucleotide biosynthesis. Yeast metabolomics pathway map showing metabolite fold changes following 40 min treatment with 20 µg/ml cordycepin. Solid circles indicate detected confirmed metabolites and dashed circles detected putative metabolites. Coloured circles show metabolites with significant fold change after cordycepin treatment with red indicating an increase in abundance and blue a decrease in abundance. Darker colour is used for metabolites with greater than 50% fold change. Pathways are based on the KEGG database and include starch and sucrose metabolism (orange), glycolysis/gluconeogenesis (pink), citrate cycle (TCA cycle) (purple), pentose phosphate pathway (yellow), pyrimidine metabolism (green), and purine metabolism (blue). Asterisk (\*) indicates metabolites that appear multiple times on the map. The pharmacological inhibition of Imd2 by mycophenolic acid (MPA) is indicated. Data is representative of six biological replicates and is available in the **Figure 5—source data 1**.

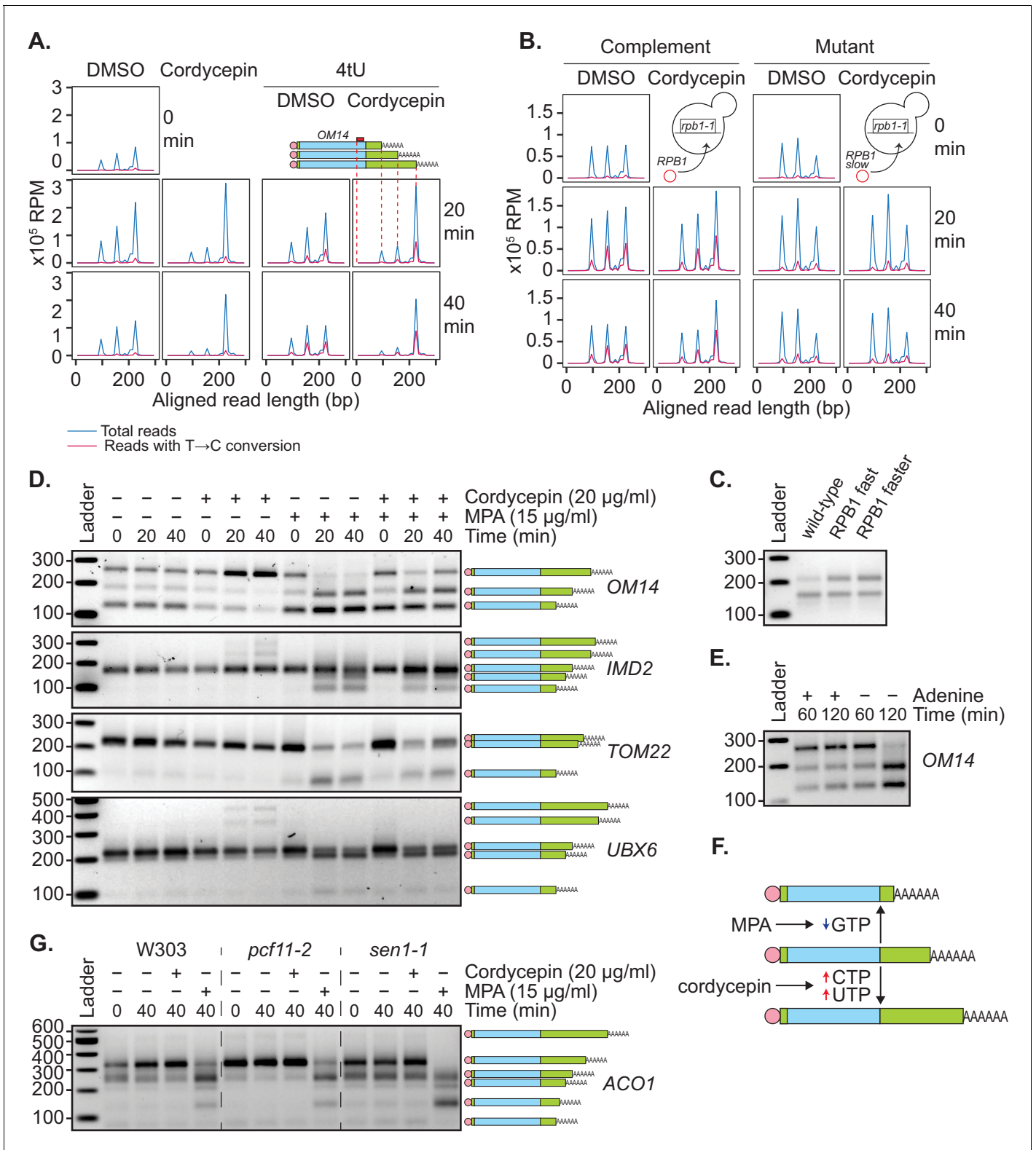
The online version of this article includes the following source data and figure supplement(s) for figure 5:

**Source data 1.** Cordycepin metabolomics data.

**Figure supplement 1.** Glycolytic output increase following mitochondrial stress is suppressed upon cordycepin treatment.

incorporation of 4tU, which subsequently appears as a T to C conversion when sequenced. Wild-type BY4741 cells were treated with cordycepin or DMSO and 4tU for 40 min. mRNA was then analysed using a modified version of the mPAT approach that used an anchored oligo-dT (TV12VN) primer for reverse transcription to focus on APA. For the gene *OM14*, three major cleavage sites were detected (**Figure 6A**).

Treating cells with cordycepin caused increased use of the most distal of these alternate cleavage sites. In 4tU-labelled samples, control cells had all three isoforms newly synthesised, whereas in cordycepin-treated cells, the majority of new transcripts were cleaved at the most distal cleavage site. Therefore, the increased use of longer 3' UTR isoforms was not a result of changed stability between



**Figure 6.** Genetic or pharmacological slowing of transcription attenuates cordycepin's effect on alternative polyadenylation. **(A and B)** Yeast cells treated with 20 µg/ml cordycepin or DMSO (control) and 4tU labelled for 20 or 40 min. mRNA was then analysed using the mPAT approach. Reads with T to C conversions represent transcripts synthesised during the 4tU labelling and thus newly produced mRNA. The 3' ends of reads aligned to the gene *OM14*, with the forward primer at 0, are shown in x10<sup>5</sup> reads per million. This was done for **A**. Wild-type BY4741 cells and **B**. *rpb1-1* mutant cells complemented with *RPB1* or an *rpb1* slow mutation. Data is representative of two biological replicates. **(C)** TVN-PAT analysis of the gene *OPI3* for cells **Figure 6 continued on next page**

Figure 6 continued

containing wild-type, L1101S 'fast' or E1103G 'faster' Rpb1. Samples were run relative to a 100 bp ladder. Bands correspond to the different APA isoforms of the gene with slower migrating bands indicating a more distal cleavage site and faster bands a more proximal cleavage. (D) TVN-PAT analysis of the genes *OM14*, *IMD2*, *TOM22*, and *UBX6* for yeast cells treated with 20 µg/ml cordycepin and 15 µg/ml MPA for 20 or 40 min. (E) TVN-PAT analysis of the gene *OM14* for wild-type W303 cells (adenine auxotroph) grown with or without adenine hemisulphate (100 µg/ml) for 60 or 120 min. (F) Schematic illustration of the pharmacological induction of APA. MPA treatment decreases GTP availability resulting in shorter 3'UTRs, whereas cordycepin treatment increases CTP and UTP concentrations and promotes 3'UTR lengthening. (G) TVN-PAT analysis of the gene *ACO1* for wild-type W303 and temperature-sensitive mutants *pcf11-2* and *sen1-1*. Cells were grown at 25°C then switched to the semi-restrictive temperature 30°C and treated with 20 µg/ml cordycepin or 15 µg/ml MPA for 40 min.

the different isoforms but was in fact due to the distal cleavage site being preferentially used in new transcripts. However, the total incorporation of 4tU did not significantly change between DMSO and cordycepin samples (mean incorporation of 1.011% in DMSO samples and 0.779% in cordycepin samples,  $p$ -value=0.227, 95% CI = -1.329 to 0.866). We argue that this was due to dilution of the label by the de novo nucleotide biosynthesis induced by cordycepin. As a result, an increased transcriptional rate by cordycepin could not be confirmed by an increased rate of 4tU incorporation.

To overcome this, 4tU labelling was repeated using a slow mutant of RNAP II (Kaplan et al., 2012). We reasoned that if the change in cleavage site use was caused by cordycepin increasing the transcription rate, a slow polymerase should be able to override this effect. When the *rpb1-1* mutant was complemented with a plasmid-borne copy of wild-type *RPB1* (Kaplan et al., 2012), cordycepin caused a shift toward the usage of distal cleavage sites (Figure 6B). However, *rpb1-1* complemented with a slow allele of *rpb1* (H1085Y mutation [Kaplan et al., 2012]), did not display this change. Therefore, slowing the global elongation rate ablated cordycepin's effect on 3' end processing. In contrast, cells containing fast *rpb1* alleles (Kaplan et al., 2012) displayed an increased use of the distal cleavage site for the gene *OPI3* (Figure 6C, Geisberg et al., 2020) similarly to cordycepin treatment. This observation was consistent with the idea that cordycepin did indeed promote alternative polyadenylation by altering the rate of transcription.

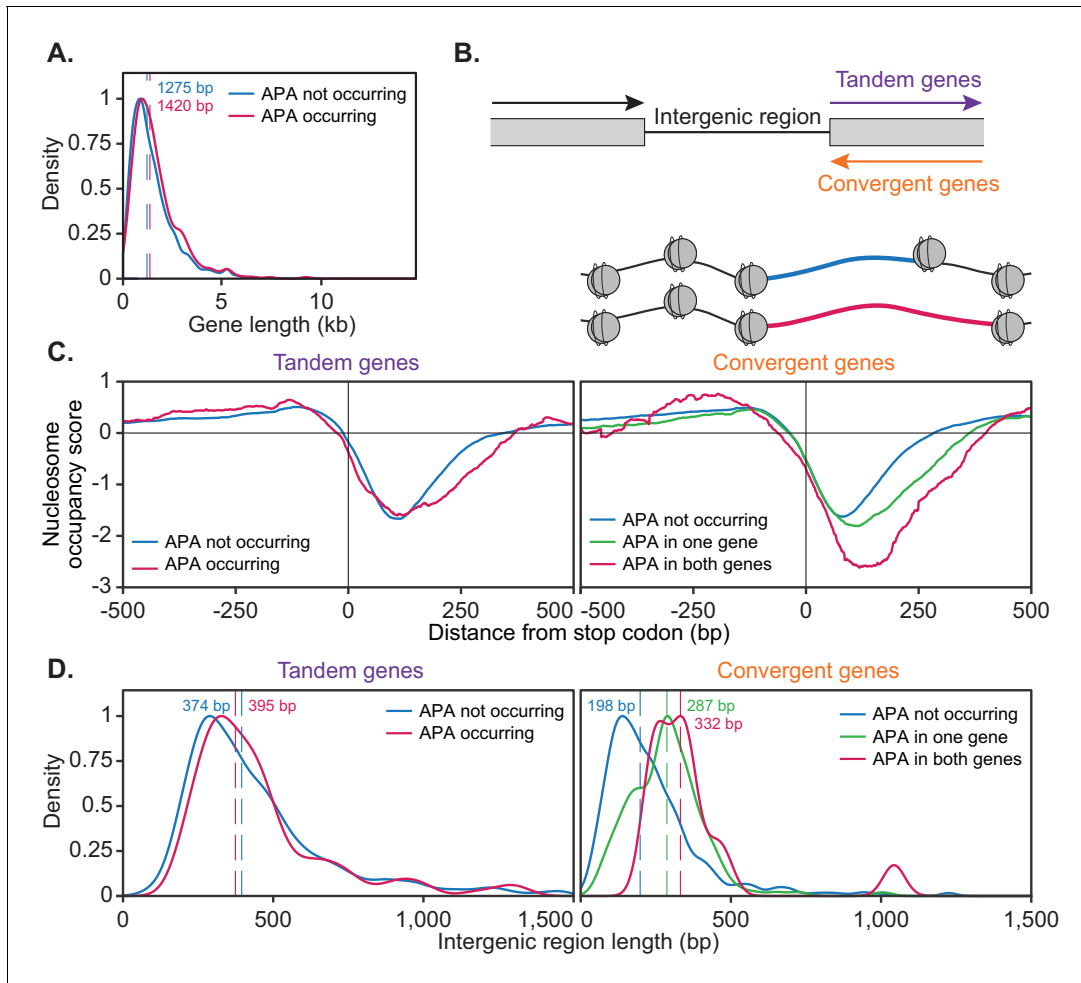
Mycophenolic acid (MPA) is an inhibitor of IMP dehydrogenase, the rate-limiting enzyme in de novo GTP synthesis. MPA treatment consequently interferes with transcription elongation due to substrate limitation (Mason and Struhl, 2005). To further demonstrate that cordycepin's APA effect was mediated by changes to nucleotide levels that in turn affected the transcriptional rate, cells were treated with 15 µg/ml MPA. Consistent with the idea that APA hinges on nucleotide levels, we found that MPA indeed caused increased usage of proximal cleavage sites (Figure 6D). This was similar to results seen upon elimination of adenine in W303 cells (Figure 6E), which require adenine supplementation for growth. Moreover, co-treatment of 20 µg/ml cordycepin with MPA led to a suppression of this shortening (Figure 6D). Collectively, these observations suggested that altered RNAP II elongation rate, as a consequence of nucleotide availability, was able to promote APA (Figure 6F).

To examine the interaction of nucleotide and cleavage factor levels in cleavage site selection, *pcf11-2* and *sen1-1* cells were treated with cordycepin or MPA (Figure 6G). Cordycepin treatment caused no further lengthening of the *ACO1* transcript in the 3'UTR lengthening mutant *pcf11-2* and was unable to suppress the shortening seen in the *sen1-1* mutant. In contrast, treatment with MPA prevented *pcf11-2* 3'UTR lengthening and caused increased shortening in *sen1-1*. This shows that 3'ends cannot extend past their most distal poly(A) site and produce a functional transcript and demonstrates that a reduced nucleotide availability has a stronger impact on cleavage site choice than the *pcf11-2* and *sen1-1* mutations.

### The ability of a gene to undergo alternative polyadenylation is dependent on the local nucleosome landscape

A large proportion of the yeast genome displayed changes in alternative polyadenylation following cordycepin treatment. But why did specific genes alter their 3' ends in response to both loss of the core cleavage machinery and the drug cordycepin whilst others did not? To address this question, we assessed the structural differences between genes that responded to cordycepin treatment and those that maintained their original cleavage sites.

For this analysis, genes were considered to be alternatively polyadenylated if more than 40% of reads switched to another cleavage site following cordycepin treatment and to not undergo alternative polyadenylation if the same cleavage site was primarily used in treated and untreated cells. This was a more stringent classification of APA than was utilised in the end-shift test in **Figure 3B** above. On average, genes with a high level of APA were significantly longer than those without APA ( $p$ -value=0.031) with median gene coding region lengths of 1420 bp and 1275 bp, respectively (**Figure 7A**). This suggested that the length of a gene may influence its ability to undergo APA.



**Figure 7.** The ability of a gene to undergo alternative polyadenylation is dependent on the local nucleosome landscape. Comparison of native genetic features in genes split based on APA following cordycepin treatment. APA was considered to occur if the major cleavage site (greater than 40% of reads) shifted to another site following 20  $\mu$ g/ml cordycepin treatment for 40 min and to not occur if the major cleavage site remains the same. Genes used in each comparison are included in the **Figure 7—source data 1**. (A) Scaled density plot comparing gene length (in kilobase pairs) for genes in which APA occurs (pink,  $n = 378$ ) and does not occur (blue,  $n = 4302$ ). Dashed lines and coloured values indicate the median gene length. T-test  $p$ -value=0.031. (B) Schematic of the two types of 3' end intergenic regions. The downstream gene can either be transcribed in the same direction as the upstream gene (tandem genes) or the opposite direction (convergent genes), in which case two genes will be assigned the same 3' end intergenic region. (C) Normalised nucleosome occupancy per base pair around the translation end site, averaged across all yeast genes, derived from **Kaplan et al., 2009**. For tandem genes, APA occurred in 128 genes (pink) and did not occur in 1719 genes (blue). T-test  $p$ -value=0.055. For convergent genes, APA occurred in both genes for 15 gene pairs (pink), in one gene for 155 pairs (green) and in neither gene for 746 pairs (blue). ANOVA  $p$ -value<0.001. Nucleosome occupancy values above zero indicate nucleosome enrichment compared to the genome-wide average. (D) Scaled density plot comparing 3' end intergenic region length (in base pairs) based on APA up to 1500 bp. Dashed lines and coloured values indicate the median intergenic region length. Tandem gene t-test  $p$ -value=0.646, convergent gene ANOVA  $p$ -value<0.001.

The online version of this article includes the following source data for figure 7:

**Source data 1.** Genes undergoing APA following cordycepin treatment.

Genes were then further classified based on the directionality of their downstream neighbouring gene (**Figure 7B**). If the neighbouring gene was transcribed in the same direction, the gene of interest was classified as 'tandem', whereas if the adjacent gene was transcribed in the opposite direction, this gene was termed 'convergent'. Convergent genes share a 3' end intergenic region and are referred to as 'gene pairs'. Gene pairs were divided based on whether both, one or neither genes undergo APA.

We next sought to determine whether the chromatin landscape of a gene correlated with alternative polyadenylation. Nucleosomes are positioned at defined locations across the genome. In particular, nucleosome depleted regions are associated with transcription start and stop sites (**Shivaswamy et al., 2008; Lee et al., 2004; Xu et al., 2009; Mavrich et al., 2008**). Focusing on the 3' end, we compared the native nucleosome occupancy of genes based on their responsiveness to cordycepin treatment (**Figure 7C**). This was done with a publicly available dataset of yeast nucleosome occupancy (**Kaplan et al., 2009**).

The nucleosome occupancy landscape of tandem genes did not significantly differ based on APA (p-value=0.055). However, convergent genes showed significant differences in nucleosome occupancy scores between gene pairs in which both, one and neither genes undergo APA (p-value<0.001). On average, the width of the nucleosome depleted region after the stop codon was wider when alternative polyadenylation did occur. This was enhanced when both genes in a pair displayed APA. As these cells were not treated with cordycepin, it appeared that responsive convergent genes had an innately wider nucleosome depleted region at the 3' end of the gene. Therefore, a gene's alternative polyadenylation potential may be linked to a pre-determined framework of nucleosome positioning and the directionality of their neighbouring downstream gene.

In addition to a wider nucleosome depleted region at the 3' end of the gene, cordycepin responsive convergent genes had on average a longer intergenic region between neighbouring genes (**Figure 7D**; median 332 bp for APA occurring in both genes and 198 bp for APA not occurring in either gene, p-value<0.001). Where one gene in a pair undergoes APA and the other did not the intergenic region length was intermediate (median 287 bp) compared to those in which APA occurs in both or neither gene. Increased intergenic region length may provide more room for additional cleavage sites to exist for a single gene and permit transcriptional extension without collision of transcriptional machinery between neighbouring genes. The length of the intergenic region for tandem genes did not significantly correlate with APA (median 397 bp for APA occurring and 380 bp for APA not occurring, p-value=0.856). It, therefore, appears that 3' end nucleosome positioning and the length of the 3' end intergenic region were determinants of APA potential for convergent but not tandem genes.

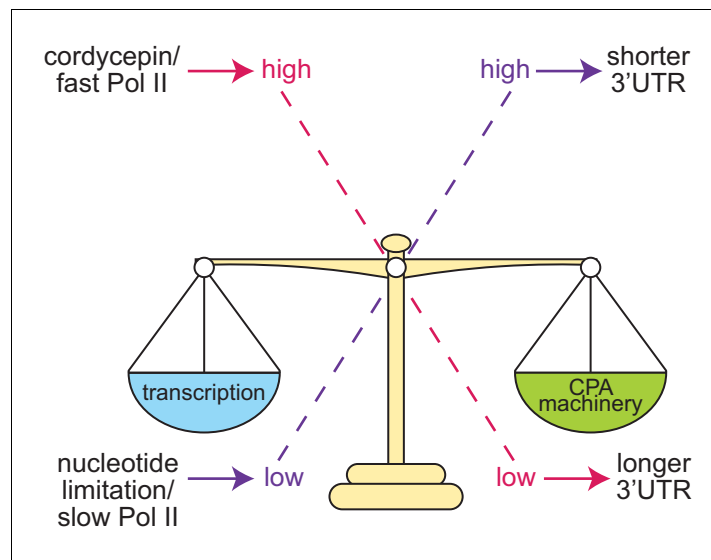
Based on the data presented here we propose the following model for alternative polyadenylation regulation in *S. cerevisiae* (**Figure 8**). Cleavage site choice exists as a balance between the availability of the cleavage and polyadenylation machinery, its affinity towards the sequence elements present on the pre-mRNA substrate and the transcriptional elongation rate. When this balance is tipped, it alters which cleavage site is preferentially utilised. High concentrations of the 3' end processing machinery ensure cleavage sites can be efficiently recognised promoting cleavage at more proximal cleavage sites due to their temporal advantage. Conversely, decreased concentrations, or cleavage factor mutations, will delay processing at proximal sites and increase distal cleavage site usage.

Changing the rate of transcription has a similar impact on cleavage site choice as this shifts the ratio between transcript elongation and the availability of the 3' end processing machinery. Faster transcription, as instigated by cordycepin treatment and facilitated by a permissive chromatin landscape, provides less time for the proximal site to be recognised by the cleavage factors before the distal form is also transcribed. The physical coupling of 3' end formation and transcription is likely to promote the preferential usage of the longer 3' UTR isoform once the distal cleavage site is available.

## Discussion

Despite its pervasive nature, the mechanisms of alternative polyadenylation remain poorly understood. The starting point of this study was a systematic analysis of 3' end formation factors, which uncovered that mutant alleles of most of the factors promoted global switching towards distal sites.





**Figure 8.** Proposed model of cleavage site selection in yeast. Schematic of cleavage site selection in yeast. We propose a kinetic model where balance exists between nucleotide levels, transcriptional rate, and the concentration of the cleavage and polyadenylation machinery. Tipping this scale by alteration of nucleotide concentration or mutation of cleavage factors and RNA polymerase II causes changes to cleavage site choice.

In contrast, a subset of genes shifted toward more proximal sites when the RNA helicase Sen1 was mutated. In addition, we demonstrate that cordycepin treatment caused a global shift to distal poly(A) sites, an effect that we found to be accompanied by an accumulation of nucleotides and their biosynthetic intermediates. In agreement with the idea that changes in cellular nucleotide concentrations impacted the usage of alternative poly(A) sites following cordycepin treatment, we show that mycophenolic acid, which is known to reduce GTP levels, caused a shift to more proximal sites and combined treatment reversed the shifts observed with cordycepin alone. Furthermore, suppression of cordycepin effects by a slow *rpb1* mutant of RNAP II functionally linked transcription elongation to alternative polyadenylation. Since transcription elongation is sensitive to nucleotide levels, we propose that nucleotide metabolism plays a key role in transmitting kinetic coordination of RNAP II elongation and coupled pre-mRNA 3' end processing in the regulation of alternative polyadenylation.

Employing global PAT-seq and targeted mPAT approaches we found that preferred poly(A) sites in wild-type cells shifted towards more distal sites in the majority of strains carrying mutations in subunits of the canonical cleavage and polyadenylation machinery. Overall, our results indicate that gene specific roles for individual 3' end factor subunits or their expression levels play a minor part in the regulation of poly(A) site recognition for the majority of genes. Mutations in these factors can compromise efficient recognition of the more proximal sites, which are then skipped in favour of a stronger more distal site that is more readily recognised by the impaired 3' end processing machinery (Shepard et al., 2011; Beudoing et al., 2000). However, as some genes are impacted upon mutation of individual subunits only, there also appears to be a degree of subunit specificity for alternative polyadenylation and some phenotypic variability likely can also be attributed to the specific alleles analysed in this study. Interestingly, changes in APA did not correlate with defective in vitro cleavage activity associated with mutant alleles (Figure 2B). A likely explanation for this is that reduced cleavage activity impacted the usage of alternative sites equally and thus there was no apparent shift in the cleavage site. In contrast, in vitro polyadenylation activity did correlate with defective APA in all analysed cases. The reasons for this remain unclear but the mutants analysed with defective in vitro polyadenylation also display RNA-binding activity, which may be compromised giving rise to the defect in APA. The requirement for a functional CTD-binding domain within Pcf11 highlighted the importance of co-transcriptional interactions for normal 3' end formation. Quantitative variations in the strength of 3' end shifts observed for different mutant strains thus likely

reflected distinct functional roles and deficiencies, which can impact on RNA binding, subcomplex interactions and complex stability, and coupling of 3' end formation to RNAP II transcription.

We observed a unique requirement for the Sen1 helicase for poly(A) site-choice on a subset of 676 genes (**Figure 3A**). Sen1 function has been well characterised in association with the NNS complex where it acts to promote transcription termination of non-coding RNA and of a small number of protein coding genes (**Steinmetz et al., 2006**). The role played by Sen1 in alternative polyadenylation described here is distinct from the previously described roles in RNA processing because the observed *sen1-1* mutant phenotype was not duplicated by *nrd1-5* and *nab3-11* mutations, and thus, Sen1 appeared to act independently of the RNA-binding proteins Nrd1 and Nab3. In contrast, deficiency of the Nrd1-like protein Seb1 in *Schizosaccharomyces pombe* resulted in widespread 3' UTR lengthening demonstrating that Nrd1's role in mRNA 3' end formation has not been conserved between budding and fission yeast (**Lemay et al., 2016**). Here, the *sen1-1* allele caused shifting toward sites proximal to those utilised in the wild-type. Therefore, functional Sen1 appeared to promote the usage of downstream poly(A) sites and thus suppressed usage of proximal poly(A) sites rather than promoting termination as was previously seen with non-coding RNA transcription. This may indicate that Sen1 antagonised the function of the core cleavage and polyadenylation machinery, since Pcf11 was inversely required for the usage of more proximal poly(A) sites on more than 80% of the identified Sen1 target genes. Interestingly, this activity was associated with genes that carried an above average length of the open reading frame suggesting that Sen1 may facilitate full-length expression of long genes. The mechanisms underlying Sen1 function in alternative polyadenylation remain unclear and need to be addressed in future studies. We speculated that defective termination of antisense non-coding RNA transcription in *sen1-1* mutants facilitated transcriptional interference that in turn caused an apparent suppression of the use of distal sites on the sense mRNA transcript. Such a role would, however, be expected to involve Nrd1 and Nab3 and therefore may not apply here. Alternatively, since Sen1 does interact with the RNAP II CTD (**Vasiljeva et al., 2008; Han et al., 2020**), it may impact on the interaction of other CTD binding factors, like Pcf11 (**Grzechnik et al., 2015**), and in this way interfere with efficient 3' end formation and/or termination of transcription that could shift poly(A) site usage from proximal to distal positions. Another possibility may be that Sen1 helicase activity was required to resolve RNA structures that form in proximity to the RNAP II exit channel, which may provoke transcriptional pausing and promote preferential usage of proximal sites if unresolved. Interestingly, an NNS-independent requirement for Sen1 has been reported previously in the protection of replication fork integrity that may be associated with pathology of Senataxin deficiencies (**Alzu et al., 2012**) but it remains unclear if this is related to Sen1 function in APA as described here.

Recent analysis of 3' end factor interaction on nascent pre-mRNA suggested that the overall architecture of complex assembly across poly(A) sites is conserved between yeast and mammals (**Baejen et al., 2014**). In the mammalian system cleavage factor II subunits have been associated with the selection of proximal sites (**Kamieniarz-Gdula and Proudfoot, 2019; Ogorodnikov et al., 2018; Turner et al., 2020; Wang et al., 2019**), while cleavage factor I (CFIm) promotes the usage of distal sites suggesting that CFIm acts to represses cleavage at proximal sites (**Martin et al., 2012; Masamha et al., 2014**). As no CFIm homologues are known in yeast the repressive role of the core cleavage and polyadenylation machinery may not be conserved. Instead, the compact nature of the yeast genome with short intergenic regions may have favoured mechanisms that promote the preferential usage of proximal poly(A) site in order to avoid transcriptional interference between neighbouring genes (**Gullerova and Proudfoot, 2010**). The genes that require Sen1 for selection at distal sites may form a distinct group that was characterised by long open reading frames and it will be interesting to see whether the Sen1 homologue Senataxin plays a similar role in humans.

Cordycepin's ability to terminate RNA synthesis following its incorporation into the polynucleotide chain has been widely demonstrated (**Horowitz et al., 1976; Müller et al., 1977**). We have previously shown that cordycepin interferes with mRNA 3' end formation independently of RNA chain termination causing 3' UTR read-through in the genes *ACT1*, *ASC1* and *CYH2* (**Holbein et al., 2009**) opening the possibility that the drug may impact on alternative polyadenylation. Through PAT-seq analysis we have shown here that low-dose cordycepin treatment indeed stimulated alternative polyadenylation and distal cleavage site usage globally of similar extent to those observed with 3' end factor mutants. Cordycepin effects were most comparable to shifts in *pcf11-2* and *ysh1-13* yeast strains. However, as no alternative polyadenylation profile of a 3' end factor matched that of

cordycepin, the drug was unlikely to act through direct interaction with specific proteins. Furthermore, although cordycepin did provoke coordinated changes in 3' end factor RNA expression, only Hrp1/Nab4 showed significant changes in protein levels following a 40 min exposure to the drug. This indicated that changes in protein levels were unlikely to cause the cordycepin induced alternative polyadenylation, which was detected within 10 min of drug treatment. Our metabolomics analysis revealed accumulation of nucleotides and their biosynthetic intermediates in wild-type cells exposed to cordycepin. An increase in the pool of cellular nucleotides may reduce cordycepin toxicity by dilution. This idea is in line with our previous observation that poly(A) polymerase mutants displayed cordycepin resistance due to increased levels of free ATP (Holbein et al., 2009).

Importantly, nucleotide concentration is a key determinant of RNAP II processivity and elongation rate (Mason and Struhl, 2005; Uptain et al., 1997). Consistent with this idea drugs like mycophenolic acid and 6-azauracil that reduce GTP and UTP levels, respectively, act synergistically with defects in trans-acting factors required for efficient elongation causing a reduced rate of nucleotide incorporation, RNAP II pausing and dissociation of RNAP II from chromatin (Exinger and Lacroute, 1992; Shaw and Reines, 2000; Mason and Struhl, 2005). The increased dwelling time of RNAP II on the chromatin template is thought to enhance the probability of associated 3' end formation to take place and it was recently suggested that the rate of transcriptional elongation controls alternative polyadenylation (Geisberg et al., 2020) and Figure 6C. Consistent with these observations increased nucleotide levels observed following cordycepin treatment are predicted to enhance the rate of RNAP II elongation, which will reduce the window of opportunity for 3' end formation factors to associate with the transcript at proximal sites, and allowing for the transcription and use of distal poly(A) sites. The link between nucleotide levels and alternative polyadenylation is underscored by our observation that mycophenolic acid, which slows RNAP II transcription through reduced GTP levels, promoted usage of proximal sites and suppressed the effect of cordycepin. Similarly, a slow *rpb1* allele mirrored the effects of mycophenolic acid to suppress the effects of cordycepin on alternative polyadenylation. Importantly, distal site usage in *pcf11-2* was not further induced by cordycepin, indicating an upper limit to 3' lengthening, whereas depletion of GTP by MPA served to normalise 3' end choice.

In addition to nucleotide levels and trans-acting factors, the structure of the chromatin template is an important parameter for RNAP II elongation. Increased nucleosome occupancy at the 3' end of the gene reduces the rate of transcription due to the stalling of RNAP II (Hodges et al., 2009). This provides the 3' end formation machinery with more time to recognise this first site prior to the distal cleavage site being transcribed, meaning this position is nearly exclusively used. A stable nucleosome between potential cleavage positions is therefore predicted to provide a more proximal site with a kinetic advantage. Likewise, when the nucleosome depleted region is wider, the proximal site is expected to have less temporal advantage. Consistent with this idea we showed that genes that switch their poly(A) site in response to cordycepin treatment differ in their chromatin landscape compared to those that retain their original poly(A) sites. Convergent genes that were subject to alternative polyadenylation following cordycepin treatment tended to have wider intergenic regions downstream of the coding sequence, which likely allowed for transcriptional extension without collision of the transcriptional machinery between neighbouring genes. In contrast, no such widening of the intergenic region was observed for genes that are transcribed in tandem consistent with transcriptional interference being no issue at those genes. Moreover, we found that cordycepin-responsive convergent genes had innately wider nucleosome depleted regions at the 3' end of the gene consistent with the idea that nucleosome density restricts the occurrence of alternative polyadenylation.

Taken together, our results support a model where kinetic competition governs the selection of alternative poly(A) sites. On the one hand, proximal sites are transcribed first and thus carry a temporal advantage. The strength of a given poly(A) site and the availability of 3' end factors will determine the kinetics of cleavage and polyadenylation at those sites. Transcription of distal sites, on the other hand, will be achieved only under conditions which delay rapid processing at proximal sites and which favour rapid elongation by RNAP II. Once distal sites are available, however, they will compete with proximal sites for interaction with the 3' end formation machinery. The closer vicinity to elongating RNAP II may then favour the distal site due to the physical interaction of 3' end formation factors and the CTD of RNAP II. A more complete understanding of the mechanisms of alternative polyadenylation will thus require the analysis of a large number of nuclear proteins, which

directly impact on gene transcription by RNAP II, chromatin structure and its coupling to pre-mRNA 3' end formation.

## Materials and methods

### Yeast cell growth conditions

For cleavage factor mutant experiments, yeast strains were grown in YPAD rich media (1% yeast extract, 2% peptone, 2% glucose, 100 µg/ml adenine hemisulphate) at 25°C to an OD<sub>600</sub> of 0.6 then transferred to 37°C for 1 hr before cells were harvested.

For confirmation of temperature sensitivity, yeast strains were grown overnight in YPAD rich media at 25°C. Ten-fold serial dilutions were spotted on YPAD plates (1% yeast extract, 2% peptone, 2% glucose, 100 µg/ml adenine hemisulphate, 20 mg/ml agar) and grown for 2 days at either 25°C or 37°C.

For cordycepin treatment experiments, BY4741,  $\Delta nrt1$ ,  $\Delta ado1$ , and  $\Delta adk1$  yeast strains were grown in synthetic media (0.67% yeast nitrogen base [without amino acids or carbohydrates; with ammonium sulphate], 2% glucose, 20 µg/ml uracil, 20 µg/ml methionine, 20 µg/ml histidine, and 100 µg/ml leucine) at 30°C to an OD<sub>600</sub> of 0.6. W303, *pcf11-2* and *sen1-1* cells were grown in synthetic media at 25°C (0.67% yeast nitrogen base [without amino acids or carbohydrates; with ammonium sulphate], 2% glucose, 20 µg/ml uracil, 20 µg/ml tryptophan, 20 µg/ml histidine, 100 µg/ml leucine, 100 µg/ml adenine hemisulphate) to an OD<sub>600</sub> of 0.6 then moved to the semi-restrictive temperature 30°C. 20 µg/ml cordycepin, or an equal volume of DMSO (vehicle control) was then added and incubation was continued for the desired treatment lengths before cells were harvested. Where mycophenolic acid (MPA) was used, it was added at 15 µg/ml to cells grown exactly as for cordycepin treatment. All samples were collected immediately by centrifugation, washed with 1 mL ice cold milliQ H<sub>2</sub>O containing sodium azide (0.1% w/v) and pellets frozen at -80°C until RNA extraction.

For confirmation that a non-lethal dose of cordycepin was applied, BY4741 cells were grown overnight in YPAD rich media at 30°C. Ten-fold serial dilutions were spotted on SC plates (0.67% yeast nitrogen base (without amino acids or carbohydrates; with ammonium sulphate), 2% glucose, 20 µg/ml uracil, 20 µg/ml methionine, 20 µg/ml histidine and 100 µg/ml leucine, 20 mg/ml agar) and grown for 2 days at 30°C with 20 µg/ml or 120 µg/ml cordycepin. An equivalent volume of DMSO was used as a control.

For the cordycepin PAT-seq, metabolomics and proteomics datasets, two replicates of OD<sub>600</sub> 0.6 BY4741 cells treated with 20 µg/ml cordycepin for 0, 5, 10, 20, or 40 min were prepared over 3 separate days (six biological replicates total). Harvested cells were washed with ice cold water containing sodium azide (0.1% w/v) and frozen with liquid nitrogen. Two replicates were used for PAT-seq, six for metabolomics and five for proteomics.

For W303 growth with or without adenine, cells were grown in W303 background synthetic media supplemented with 100 µg/ml adenine hemisulphate at 30°C to an OD<sub>600</sub> of 0.6. Cells were then resuspended in the same media with or without adenine hemisulphate and grown for 1 or 2 hr.

Rpb1 fast mutants were created by transforming L1101S (fast) and E1103G (faster) Rpb1 mutant alleles (Kaplan *et al.*, 2012) on *LEU2* pRS315 based plasmids into CKY283 cells followed by replica plating on 5-fluoroototic acid to remove cells retaining the wild-type *RPB1 URA3* plasmid. Cells were then grown in YPAD media until an OD<sub>600</sub> of 0.6.

### Total RNA extraction

Total RNA extraction was carried out using the hot phenol method (Schmitt *et al.*, 1990) with slight modifications. Cell pellets were resuspended in 400 µl AE buffer (50 mM sodium acetate pH 5.2, 10 mM EDTA) on ice. A total of 33 µl of 25% SDS and 400 µl acid phenol was added then samples were vortexed and transferred to a 65°C water-bath for 20 min with intermittent vortexing. Samples were incubated on ice for 5 min to precipitate the SDS then centrifuged at top speed for 5 min at 4°C in a bench-top centrifuge. The supernatant (both phases) was transferred to a fresh tube and 400 µl of chloroform was added. Samples were vortexed thoroughly and centrifuged at top speed for 5 min at room temperature. The aqueous phase was then transferred to a fresh tube, avoiding the inter-phase and another 400 µl of chloroform was added. Following vortexing and centrifugation, the aqueous phase was transferred to a fresh tube containing 1/10th volume of 3 M sodium acetate pH

5.2 and an equal volume of isopropanol was added. Tubes were mixed by inversion then centrifuged at top speed for 20 min at 4°C. Supernatant was removed and pellets were washed with 80% ethanol and centrifuged at top speed for a further 10 min. Pellets were air-dried at room temperature and resuspended in RNase-free dH<sub>2</sub>O.

### PAT-seq

PAT-seq libraries were prepared as described previously (*Harrison et al., 2015*). A 5' biotinylated PAT-anchor primer, compatible with Illumina index primers, was used to create a 3' tag on adenylated RNA. This ensures that reverse transcription only occurs from true 3' ends rather than internal A repeat regions. For this 1 µg total RNA was combined with 1 µl Bio-EE primer and brought to a volume of 12 µl with dH<sub>2</sub>O in 200 µl PCR tubes. Mixtures were incubated at 80°C for 5 min then cooled to 37°C for 5 min. Once cooled, tubes were flash-centrifuged and 8 µl of master mix was added containing 4 µl 5x SuperScript III buffer, 1 µl 100 mM DTT, 1 µl 10 mM dNTPs, 1 µl RNaseOut and 1 µl Klenow polymerase per reaction. Samples were mixed thoroughly by inversion then flash-centrifuged and incubated at 37°C for 30 min. The temperature was then increased to 80°C for 5 min and 80 µl digest buffer (300 mM NaCl, 10 mM Tris-HCl pH 7.5) added before incubating for a further 10 min. Samples were then transferred to ice for 5 min.

RNA was then exposed to limited fragmentation by mixing with 100 µl 1:1000 diluted (with digest buffer) solution of RNase T1 for 2 min on ice in phase-lock tubes. This was followed by immediate phenol/chloroform RNA extraction with 200 µl 1:1 acid phenol: chloroform mix and vigorous shaking to stop the reaction. After centrifugation at top speed for 5 min in a benchtop centrifuge, 200 µl of chloroform was added, tubes were shaken and centrifuged again. The top layer was then transferred to new tubes containing 100 µl streptavidin magnetic beads (pre-washed in 55°C digest buffer) and incubated at 37°C with shaking for 15 min to allow binding of the DNA-RNA duplex to the beads. Extended 3' fragments were collected on the streptavidin magnetic beads and resuspended in 200 µl 1/10th strength 5x SuperScript III buffer at 55°C for 5 min. Fragments were collected on beads again then washed with 100 µl 1x T4 DNA ligase buffer (+ATP) at 37°C for 5 min with shaking.

5' phosphorylation of extended 3' fragments was carried out using T4 Polynucleotide Kinase. 1 µl T4 Polynucleotide kinase and 19 µl T4 DNA ligase buffer (+ATP) was added and samples were incubated at 37°C for 30 min with shaking. Tubes were then transferred to a 70°C heat block for 10 min then cooled slowly to room temperature to inactivate the enzyme. 200 µl of 1/10th strength 5x SuperScript III buffer was added and samples were transferred to 200 µl PCR tubes then concentrated using the beads. 7 µl of ligation master mix was added containing 2 µl dH<sub>2</sub>O, 2 µl stoichiometrically pre-annealed PAT-seq Splint A and Splint B, 3 µl 2x hybridisation buffer (40% PEG 8000, 1x SuperScript buffer) per reaction. Samples were then incubated at 65°C for 10 min then 16°C for 5 min. 8 µl 24% PEG 8000, 2 µl 10x RNL2 buffer and 2 µl RNA ligase two was added and samples incubated at 16°C for 16 hr.

Excess 5' splint was removed by adding 100 µl 1/10th strength 5x SuperScript buffer and concentrating on the magnetic beads. Reverse transcription was performed using the PAT-seq end-extend primer and SuperScript III on the magnetic matrix. A total of 18 µl master mix was added to each sample containing 11 µl dH<sub>2</sub>O, 4 µl 5x SuperScript III buffer, 1 µl 100 mM DTT, 1 µl 10 mM dNTPs and 1 µl RNaseOut per reaction. Samples were then incubated at 70°C for 5 min then cooled to 55°C. One µl SuperScript III was added to each tube, mixed by pipetting and incubated at 55°C for a further 60 min and heated to 70°C for 5 min.

The cDNA was size selected by elution from the beads in 2x formamide gel loading buffer and electrophoresis on a 6% urea-PAGE alongside a 25 bp DNA ladder. Fragments selected were between 125 bp and 300 bp in length.

Following gel excision, the cDNA was eluted using the 'crush and soak' method. The excised gel slices were placed inside 0.5 ml pierced lobind tubes in 2 ml lobind tubes. These were centrifuged at 13,000 g for 3 min and 300 µl dH<sub>2</sub>O added to the bottom tube which was incubated at 4°C with gentle shaking overnight. Eluted material and gel slurry were transferred to a corning spin-x column and centrifuged at top speed for 3 min. Supernatant was then transferred to a 1.5 ml lobind tube containing 2 µl glycoblue and 35 µl 5 M ammonium acetate and mixed thoroughly. A total of 350 µl isopropanol was added and tubes were incubated on ice for 5 min before being centrifuged at 13,000 g for 20 min at 4°C. Pellets were then washed with 80% ethanol and air-dried before being suspended in 30 µl dH<sub>2</sub>O. Ten µl of the purified cDNA was used as input for amplification with 1 µl

PAT-seq Universal forward primer, 2  $\mu$ l TruSeq Index PCR reverse primer, 50  $\mu$ l 2x Amplitaq Gold 360 Master Mix and 37  $\mu$ l dH<sub>2</sub>O. Amplification was performed with cycle conditions of 94°C for 10 min, then 16 cycles of 94°C for 20 s, 60°C for 20 s and 72°C for 30 s, followed by 1 min at 72°C. A different TruSeq primer was used for each experimental condition. PAT-seq libraries were sequenced on the Illumina HiSeq 1500 platform with 100 base rapid chemistry per the manufacturer's instructions at the Monash Health Translation Precinct medical genomic facility in Clayton, Victoria. mPAT mPAT libraries were prepared as previously described (Beilharz et al., 2017). To create complementary DNA (cDNA), 1  $\mu$ g of total RNA and 1  $\mu$ l of the anchor primer mPAT reverse were combined and brought to a volume of 11  $\mu$ l with dH<sub>2</sub>O in 200  $\mu$ l PCR tubes. Mixtures were incubated at 80°C for 5 min then cooled to 37°C for 1 min. Once cooled, tubes were flash-centrifuged and 8  $\mu$ l of master mix was added containing 4  $\mu$ l 5x SuperScript III buffer, 1  $\mu$ l 100 mM DTT, 1  $\mu$ l 10 mM dNTPs, 1  $\mu$ l RNaseOut and 1  $\mu$ l Klenow polymerase per reaction. Samples were mixed thoroughly by inversion then flash-centrifuged and incubated at 37°C for 30 min. The polymerase was then inactivated by increasing the temperature to 80°C for 10 min before cooling to 55°C for 1 min. At this temperature, 1  $\mu$ l of SuperScript III was added to each tube and a further 10 min of incubation at 55°C was carried out prior to mixing by inversion and flash-centrifuging. This is done to maintain the temperature during the initial stages of reverse transcription and prevent internal priming. Samples were then incubated at 55°C for 30 min followed by inactivation of reverse transcription at 80°C for 5 min and cooled to 12°C.

For PCR amplification, cDNA was diluted 1:5 with the addition of 100  $\mu$ l of dH<sub>2</sub>O. 25  $\mu$ l diluted cDNA was added to fresh 200  $\mu$ l PCR tubes with 75  $\mu$ l master mix containing 1  $\mu$ l mPAT reverse primer, 1  $\mu$ l pooled forward primer, 50  $\mu$ l Amplitaq Gold 360 Master Mix and 23  $\mu$ l dH<sub>2</sub>O per reaction. The pooled forward primer contained 46 gene-specific primers that each had a universal 5' extension for sequential addition of the 5' (P5) Illumina elements. Amplification was performed with cycle conditions of 95°C for 10 min, then 5 cycles of 95°C for 30 s, 60°C for 30 s and 72°C for 30 s, followed by 1 min at 72°C.

Amplicons were then purified using Macherey-Nagel NucleoSpin columns per manufacturer's instructions and NT1 buffer diluted 1:1 with dH<sub>2</sub>O. 49  $\mu$ l of 60°C dH<sub>2</sub>O was used for the final column elution. A second round of PCR amplification was carried out for each sample using 47  $\mu$ l purified amplicon, 1  $\mu$ l P5 universal forward primer, 2  $\mu$ l Illumina TruSeq indexed primer and 50  $\mu$ l Amplitaq Gold 360 Master Mix. A different TruSeq primer was used for each experimental condition and amplification as above was repeated for 10 cycles. All PCR reactions were pooled then purified using Nucleospin columns with 1:3 diluted NT1 buffer and a final elution volume of 5  $\mu$ l dH<sub>2</sub>O per original sample number. mPAT sequencing was done using MiSeq Reagent kit v2 with 300 cycles (300 bases of sequencing) on the Illumina MiSeq platform according to the manufacturer's specifications at Micromon Genomics, Monash University, Australia.

## Sequencing data processing and statistical tests

PAT-seq and mPAT reads were processed using the Tail Tools pipeline (Harrison et al., 2015). Read counts were produced at the level of whole genes and at the level of peaks detected by the pipeline. The reference annotation used was Ensembl *Saccharomyces cerevisiae* version 82 for the cordycepin PAT-seq, version 89 for the mPAT, cordycepin mTVN-PAT and NNS PAT-seq and version 93 for the *rpb1* mTVN-PAT and cleavage factor PAT-seq. These counts were then used to test for differential gene expression and APA. Raw sequencing and processed data has been uploaded to GEO, accession number GSE160539.

Differential gene expression was tested for between experimental groups, using the Fitnoise library (<https://github.com/pfh/fitnoise>; Harrison, 2018) as part of the Tail Tools pipeline, after log transforming and weighting peak counts with voom (Law et al., 2014) from the limma R package (RRID:SCR\_010943) after TMM library size normalisation (Robinson and Oshlack, 2010). Fitnoise is an implementation of Empirical Bayes moderated t-tests on weighted linear models as described in Smyth, 2004. Genes for which no relevant sample had at least 10 reads were removed before testing.

Shifts in APA usage was analysed as previously described (Turner et al., 2020). Peaks were first assigned to genes with up to 400 bp down-strand of the gene (but not proceeding into another gene on the same strand) were counted as belonging to that gene. Next, peak counts were log<sub>2</sub> transformed and linear models were fitted for each peak, using TMM normalisation, voom, and

limma, producing as coefficients estimates of the  $\log_2$  abundance of reads at each peak in each of two experimental conditions. For each gene, a 'shift score' was then estimated using these fitted coefficients, as follows.

For PAT-seq experiments, considering a single gene, for two conditions  $i = A, B$ , and  $n$  peaks  $j = 1 \dots n$ , there is a fitted coefficient  $\beta_{i,j}$ , and the proportion of reads for each peak within a condition is

$$p_{i,j} = \frac{2^{\beta_{i,j}}}{\sum_{k=1}^n 2^{\beta_{i,k}}}$$

A shift score  $s$  is then calculated, ranging from  $-1$  to  $1$  where  $-1$  indicates all reads in condition B are up-strand of condition A and one indicates all reads in condition B are down-strand of condition A.

$$s = \sum_{i=1}^n \sum_{j=1}^n \text{sgn}(j-i) p_{A,i} p_{B,j}$$

where  $\text{sgn}$  is the sign function. In the simplest case of two peaks,  $s$  is equal to the proportion distal peak usage in B minus the proportion distal peak usage in A.

The fitted coefficients  $\beta_{i,j}$  have an associated error covariance matrix, which can be propagated through these steps by the delta method to obtain an approximate standard error associated with  $s$ . Shift scores and associated standard errors are calculated for each gene with two or more peaks, then the topconfacts R package (Harrison et al., 2019) is used to provide 'confect' inner confidence bounds on the shift scores with correction for multiple testing. Confect values are only given for significantly non-zero shifts, with a False Discovery Rate of 0.05. Ordering results by absolute confect values and reading down this list only as far as is desired, a False Coverage-statement Rate of 0.05 is guaranteed for these confidence bounds.

For mPAT experiments, we use the same  $s$  statistic as above, but with a simplified calculation method. Comparing two samples, A and B, with  $c_{A,j}$  and  $c_{B,j}$  read count at each peak  $j$  respectively we calculate  $p_{i,j}$  as  $p_{i,j} = \frac{c_{i,j}}{\sum_{k=1}^n c_{i,k}}$  and calculate  $s$  as above.

## Metabolomics

Cell pellets obtained from 10 ml culture at  $\text{OD}_{600}$  of 0.6 were resuspended in equal volume (10  $\mu\text{l}$ ) H<sub>2</sub>O and cells broken open with cryo-pulverisation using a Retsch CryoMill (50 small ball bearings per sample,  $4 \times 45$  s at a frequency of 30 Hz with intermittent cooling in liquid nitrogen for 1 min). 140  $\mu\text{l}$  CMW extraction solvent (1:3:0.8 chloroform:methanol:H<sub>2</sub>O v/v/v containing the internal standards 1  $\mu\text{M}$  CHAPS, 1  $\mu\text{M}$  CAPS, 1  $\mu\text{M}$  PIPES and 1  $\mu\text{M}$  TRIS) was added and samples vortexed intermittently with ball bearings for 5 min (20 s vortexing, 40 s on ice). Supernatant and cell debris were moved to a fresh tube and intermittent vortexing continued for 15 min. Samples were then centrifuged at top speed for 10 min at 4°C. Supernatant was extracted and stored at  $-80^\circ\text{C}$  until analysis by liquid chromatography-mass spectrometry.

Chromatography was performed using a Dionex Ultimate 3000 UHPLC system (Thermo) and a ZIC-pHILIC column (Merck) with ammonium carbonate and acetonitrile in the mobile phase (Creek et al., 2013) coupled to a Q-Exactive high-resolution mass spectrometer (Thermo). The instrument was operated in both positive and negative ion mode. Parameters for the high-performance liquid chromatography and mass spectrometry analysis were as previously described (Stoessel et al., 2016).

Untargeted metabolomics data analysis was performed using the freely available software packages mzMatch (Scheltema et al., 2011) and IDEOM (Creek et al., 2012) as previously described (Trochine et al., 2014; Stoessel et al., 2016). A total of 226 authentic metabolites were used for verification of retention times and to aid metabolite identification. Identification of metabolites with these standards have high confidence (MSI level 1) and are highlighted yellow in **Figure 5—source data 1**. Other metabolites were putatively identified (MSI level 2–3) using exact mass and predicted retention times from KEGG, Lipidmaps and MetaCyc databases (Creek et al., 2011). Mean peak height was used for relative quantification and unpaired Welch's T test used for statistical analysis. Data was not normalised, with analysis of the four spiked internal standards (CHAPS, CAPS, PIPES,

and TRIS), total ion current chromatograms, median peak heights and pooled quality control samples analysed throughout the batch used to determine signal reproducibility. It is noted that one replicate from each time point was excluded during analysis due to the observation of unwanted phase separation in the samples (likely due to excess residual water in the pellet).

### Gene and intergenic region length and nucleosome occupancy analysis

Alternative polyadenylation was based on the cordycepin PAT-seq experiment. Reads per million for each peak and sample were calculated using edgeR's TMM library size correction. A primary peak was assigned for each gene for untreated BY4741 cells (t0) and cells treated with 20 µg/ml cordycepin for 40 min (t40). The primary peak must have at least RPM  $\geq 5$  and account for at least 40% of reads for that gene at that time point. Where no primary peak was assigned, the gene was discarded. 4410 genes had a primary peak at both t0 and t40. If the primary peak differed at t0 and t40, the gene was considered to undergo APA whereas if the primary peak remained the same after cordycepin treatment, no APA has occurred.

The reference annotation used was Ensembl *Saccharomyces cerevisiae* version 82. Gene and intergenic region width were plotted as scaled density plots so that density peaks for APA and non-APA genes were comparable despite different gene numbers. For gene width, genes were split based on only APA with 378 genes undergoing APA and 4032 genes not undergoing APA. To test whether there was a significant difference between gene length in APA vs. non-APA genes an unpaired two-sided t-test was used. This returned a value of  $p=0.031$  indicating that the difference was significant.

Using the GenomicRanges package (Lawrence *et al.*, 2013, RRID:SCR\_000025), intergenic regions annotations were created by finding the difference between unannotated chromosomes and the ENSEMBL gene annotations. This was not strand-specific so that intergenic regions are based on genes being on either strand. These were then assigned to the upstream gene so that each allocated intergenic region is at the 3' end of its gene.

Intergenic regions were further characterised into tandem and convergent based on the directionality of their 3' end neighbour gene. The gene on the other side of the intergenic region was classified as either going in the same direction as the primary gene ('tandem') or in the opposite direction ('convergent'). Tandem gene intergenic regions are only assigned to one primary gene as they are on the 3' end of one gene and the 5' end of the neighbouring gene, whereas convergent gene intergenic regions are on the 3' end of two genes and are therefore assigned to two genes, or a 'gene pair'. 1847 genes were characterised as tandem and 2405 genes as convergent.

For tandem genes, 128 genes undergo APA and 1719 genes do not. When limited to an intergenic region width of 1500 bp (due to outliers of up to 4670 bp), 125 genes undergo APA and 1677 do not. To test whether there was a significant difference between tandem gene intergenic region length in APA vs. non-APA genes an unpaired two-sided t-test was used. This returned a value of  $p=0.856$  when considering all genes and  $p=0.646$  when including only intergenic regions up to 1500 bp indicating that the difference was not significant.

Convergent gene pairs were split based on whether both genes undergo APA, one gene undergoes APA or neither gene undergoes APA. Pairs for which APA information was not available for both genes were discarded. APA occurred in both genes for 15 gene pairs (30 genes), in one gene for 155 gene pairs (310 genes) and in neither gene for 746 gene pairs (1492 genes). To test whether there was a significant difference between convergent gene intergenic region length based on APA, analysis of variance (ANOVA) was used. This returned a value of  $p<0.001$  indicating that the difference was significant.

Nucleosome positioning data was obtained from Kaplan *et al.*, 2009. Yeast nucleosome DNA was prepared from log-phase cells grown in YPD treated with micrococcal nuclease (MNase). Log<sub>2</sub> normalised nucleosome occupancy per base pair values were used. Genes were characterised as above as tandem or convergent and then split based on APA. Nucleosome positioning data for 500 bp either side of the translational end site for each gene was then averaged for each category. A nucleosome occupancy value of 0 represents the genome-wide average. Values above zero indicate nucleosome enrichment relative to the genome-wide average and values below zero a relatively nucleosome depleted region. To test whether there was significant difference in 3' end nucleosome occupancy based on APA, an unpaired two-sided t-test was used for tandem genes and ANOVA



was used for convergent genes. These returned values of  $p=0.055$  and  $p<0.001$ , respectively indicating that the difference was significant convergent genes but not tandem genes.

## 4tu labelling

BY4741 cells were grown in synthetic media (0.67% yeast nitrogen base without amino acids, 2% glucose, 20  $\mu\text{g/ml}$  uracil, leucine 100  $\mu\text{g/ml}$ , histidine 20  $\mu\text{g/ml}$ , methionine 20  $\mu\text{g/ml}$ ) at 30°C. The Pol II mutant *rpb1-1* cells (Olivas and Parker, 2000), carrying plasmids either with either *RPB1* wild type (pCK859) or the *rpb1* H1085Y slow mutation (pCK870) (Kaplan et al., 2012; Malik et al., 2017) were grown in synthetic media (0.67% yeast nitrogen base without amino acids, 2% glucose, 20  $\mu\text{g/ml}$  uracil) at 30°C.

4tU labelling was carried out according to the timelapse-seq method (Schofield et al., 2018) with modifications for the yeast model system. This involves the use of 4-thiouracil (4tU) rather than 4-thiouridine (4sU), which is readily taken up by yeast cells (Munchel et al., 2011) without expressing a nucleoside transporter (Miller et al., 2011). Once cell cultures reached an  $\text{OD}_{600}$  of 0.6, an untreated sample was harvested, and the remaining cells were split into four flasks. Twenty  $\mu\text{g/ml}$  cordycepin or an equivalent volume of DMSO (vehicle control) was added to each flask along with 500  $\mu\text{M}$  4tU or an equivalent volume of DMSO (unlabelled vehicle control). Samples were taken after 20 and 40 min. Due to the light-sensitive nature of 4tU, samples were treated and total RNA extracted in a light-proof setting.

Two  $\mu\text{g}$  total RNA was combined with 0.84  $\mu\text{l}$  3 M pH 5.2 sodium acetate, 0.2  $\mu\text{l}$  0.5 M EDTA, 1.3  $\mu\text{l}$  TFEA and 1.3  $\mu\text{l}$  192 mM  $\text{NaIO}_4$  and made up to 25  $\mu\text{l}$  with RNase-free  $\text{dH}_2\text{O}$ . Samples were incubated at 45°C for 1 hr prior to cooling on ice.  $\text{NaIO}_4$  was precipitated with 300 mM sodium acetate pH 5.2 and 300 mM potassium chloride for 10 min on ice then centrifuged at top speed for 30 min at 4°C in a bench-top centrifuge. The supernatant was transferred to a fresh tube and RNA was precipitated with 2.5 volumes of 100% ethanol and 0.5  $\mu\text{l}$  5 mg/mL glycogen for 1 hr on ice. These were centrifuged at top speed for 30 min at 4°C. Supernatant was removed and pellets were washed with 80% ethanol and centrifuged at top speed for a further 10 min. Pellets were air-dried at room temperature and resuspended in 18  $\mu\text{l}$  RNase-free  $\text{dH}_2\text{O}$ .

RNA samples were combined with 2  $\mu\text{l}$  10x reducing mix (0.2  $\mu\text{l}$  1 M Tris-HCl pH 7.5, 0.2  $\mu\text{l}$  1 M DTT, 0.4  $\mu\text{l}$  5 M NaCl, 0.04  $\mu\text{l}$  0.5 M EDTA pH 8.0) and incubated at 37°C for 30 min. These were then cleaned using Zymo RNA Clean and Concentrator-5 kit per manufacturer's instructions for total RNA and eluted in 15  $\mu\text{l}$  of  $\text{dH}_2\text{O}$ .

An altered version of the mPAT assay involving a mTVN-PAT reverse primer rather than the mPAT reverse was used. This primer has an identical primer sequence to the mPAT reverse except for the addition of two 3' variable bases, V (A, G or C) and N (any base). These variable bases lock the primer to the polyadenylation site for reverse transcription.

To create mTVN-PAT cDNA, 400 ng of total RNA and 1  $\mu\text{l}$  of the mTVN-PAT reverse were combined and brought to a volume of 12  $\mu\text{l}$  with  $\text{dH}_2\text{O}$  in 200  $\mu\text{l}$  PCR tubes. Mixtures were incubated at 80°C for 5 min then cooled to 42°C for 1 min. Once cooled, tubes were flash-centrifuged and 7  $\mu\text{l}$  of master mix was added containing 4  $\mu\text{l}$  5x SuperScript III buffer, 1  $\mu\text{l}$  100 mM DTT, 1  $\mu\text{l}$  10 mM dNTPs, and 1  $\mu\text{l}$  RNaseOut per reaction. Samples were mixed thoroughly by inversion then flash-centrifuged and held at 42°C for 1 min. At this temperature, 1  $\mu\text{l}$  of SuperScript III was added followed by mixing by inversion and flash-centrifuging. Samples were then further incubated at 42°C for 15 min, 48°C for 15 min and 55°C for 15 min followed by inactivation of reverse transcription at 80°C for 5 min and cooling to 12°C.

PCR amplification, sequencing and data processing were carried out as for the mPAT with 21 and 12 gene-specific forward primers for the BY4741 and *rpb1-1* experiments respectively and primer sequences are indicated in Table 2.

For each read aligned to the target genes, every mismatch along this sequence compared to the reference was found and a list of SNPs created. From this, the total number of T to C nucleotide conversions, indicative of 4tU labelling, was calculated. The total reads (in reads per million) and those with T to C conversions for each read length was compared for each sample condition. Reads aligned to the *OM14* gene were plotted with values representative of two replicates.

**Table 2.** Oligonucleotide Primers.

Primer	Sequence (5' – 3')
PAT-seq biotin end-extend	Biotin-CAGACGTGTGCTCTTCCGATCTTTTTTTTTTTTTTTTT
PAT-seq splint A (200 μM)	CCCTACACGACGCTCTTCCG(rA)(rT)(rC)(rT)
PAT-seq Splint B (200 μM)	NNNNAGATCGGAAGAGCGTCGTGTAGGG
Illumina universal Rd1 forward (50 μM for mPAT)	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCG
PAT anchor reverse	GCGAGCTCCGCGCCGCGTTTTTTTTTTTT
TVN-PAT anchor reverse	GCGAGCTCCGCGCCGCGTTTTTTTTTTTTTVN
OM14 PAT forward	GGGTCTTTTGACGCTGGAC
mPAT reverse	CAGACGTGTGCTCTTCCGATCTTTTTTTTTTTTT
mPAT-TVN reverse	CAGACGTGTGCTCTTCCGATCTTTTTTTTTTTTTTVN
AAD16 mPAT	CCTACACGACGCTCTTCCGATCTTGCGTTTGTGTAAGAAATATGC
ADE2 mPAT	CCTACACGACGCTCTTCCGATCTAGAAACTGTCGGTTACGAAGC
APQ12 mPAT	CCTACACGACGCTCTTCCGATCTGAAACGCCTCTGCTTACTCGG
ARG8 mPAT	CCTACACGACGCTCTTCCGATCTGGCTATTGAAGCGTTTACG
ARP5 mPAT	CCTACACGACGCTCTTCCGATCTGAGACAGCAAATGAAACGC
CHD1 mPAT	CCTACACGACGCTCTTCCGATCTGCTGATGGCAATGTACGAC
COX17 mPAT	CCTACACGACGCTCTTCCGATCTTGACAGTCTGCCGACAACCA
CRN1 mPAT	CCTACACGACGCTCTTCCGATCTCGGCGGCATAATAATGC
CST6 mPAT	CCTACACGACGCTCTTCCGATCTCTCGAGCTGCATCCTTTCTT
DBF2 mPAT	CCTACACGACGCTCTTCCGATCTTCAACTAGCACCTATGAACGC
ECM16 mPAT	CCTACACGACGCTCTTCCGATCTGCTTCCAGACCATCACAGG
ECM25 mPAT	CCTACACGACGCTCTTCCGATCTGCATATACGACAACAAAATACCC
END3 mPAT	CCTACACGACGCTCTTCCGATCTGAGAAATCAATTGACACCGA
ENT1 mPAT	CCTACACGACGCTCTTCCGATCTGTGATTCTGTCTTCCAGTCCG
ERG8 mPAT	CCTACACGACGCTCTTCCGATCTGGTAGATAATAGTGGTCCATGTGA
GFD1 mPAT	CCTACACGACGCTCTTCCGATCTCACATGGACACTTTTAAACAGC
HSP26 mPAT	CCTACACGACGCTCTTCCGATCTGGTTTCTTCTCAAGAATCGTG
IMP2 mPAT	CCTACACGACGCTCTTCCGATCTGAGCCATTTTAGAATGAAAATCAGC
LOS1 mPAT	CCTACACGACGCTCTTCCGATCTGCAAGGTCAATAGCTTTTCCAGG
MRPL19 mPAT	CCTACACGACGCTCTTCCGATCTCGAGAAATTTTCAACAGACCTTCC
MRPL22 mPAT	CCTACACGACGCTCTTCCGATCTGCTGAGAAAGATGAACTGCTACTC
MSA1 mPAT	CCTACACGACGCTCTTCCGATCTGCATGTGAATGGAGTTGACCTTC
NOP16 mPAT	CCTACACGACGCTCTTCCGATCTGGCAACTACCAATTGATTACCA
NOT3 mPAT	CCTACACGACGCTCTTCCGATCTCAGTGCTAATGGCAGTATAATTTG
NUP159 mPAT	CCTACACGACGCTCTTCCGATCTGCTATATGTACGTTGTTAGTGCCG
OM14 mPAT	CCTACACGACGCTCTTCCGATCTGGTCTTTTGACGCTGGACGG
OM45 mPAT	CCTACACGACGCTCTTCCGATCTCTGGAGCTCGAAAAAGGAC
PDE2 mPAT	CCTACACGACGCTCTTCCGATCTTTCCTTTTGTGAAGTATTTGTGC
PNT1 mPAT	CCTACACGACGCTCTTCCGATCTCATGACATTATCTATGCTGTACATATTG
PRY2 mPAT	CCTACACGACGCTCTTCCGATCTGGATTCTTCTTTTCTAGGGTACGC
RCL1 mPAT	CCTACACGACGCTCTTCCGATCTGGTGAACCTCACGGACAACCT
RER2 mPAT	CCTACACGACGCTCTTCCGATCTGAGCAAGATAAATGAGTTCCG
RHO1 mPAT	CCTACACGACGCTCTTCCGATCTCAATCCCATTCTTTTCTCA
RPF1 mPAT	CCTACACGACGCTCTTCCGATCTCCGTAAGAACCCTGGTCCG
RSC4 mPAT	CCTACACGACGCTCTTCCGATCTGTCTTCCCATCATATGCATGT

Table 2 continued on next page

Table 2 continued

Primer	Sequence (5' – 3')
SLF1 mPAT	CCTACACGACGCTCTCCGATCTGGTGAAATTAGCAGGCAGTTTG
SNF2 mPAT	CCTACACGACGCTCTCCGATCTGCATGACAGAAGCGAGTGTATAG
SRP68 mPAT	CCTACACGACGCTCTCCGATCTGGTTTCTGGGCCTATTTGG
TIM54 mPAT	CCTACACGACGCTCTCCGATCTCCAAGGAAGAGCCAGAATCAG
TOM70 mPAT	CCTACACGACGCTCTCCGATCTGCAGCAATGACATTGACATCTCAC
TRP2 mPAT	CCTACACGACGCTCTCCGATCTAATGATGTATAGCAGGATCCTGA
UBC9 mPAT	CCTACACGACGCTCTCCGATCTGAATCCATCTTTCCATTCTTCC
VTS1 mPAT	CCTACACGACGCTCTCCGATCTGCATATACCAACACAGGGAACA
YET3 mPAT	CCTACACGACGCTCTCCGATCTGTGCGATGTGCAAAGCCTACA
YRA1 mPAT	CCTACACGACGCTCTCCGATCTACCGCCACTAGGTGACGC
YSC84 mPAT	CCTACACGACGCTCTCCGATCTGGATGGGTTCTTATTCAGC
MID2 mPAT	CCTACACGACGCTCTCCGATCTCAAGGTAACGAATTATCACCACG

All primers have a concentration of 100  $\mu$ M unless stated.

### Polysome profiling

Prior to cell harvesting, 100  $\mu$ g/ml cycloheximide was added and incubation continued for another minute to stall ribosomes on the mRNA. After removal of supernatant, pellets were resuspended in 5 ml breaking buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 0.5% NP40, 2 mM MgCl<sub>2</sub>, 200  $\mu$ g/ml heparin, 100  $\mu$ g/ml cycloheximide, 2 mM DTT, 0.5 mM PMSF) on ice and transferred to corex tubes. Pellets were then resuspended in 1 ml breaking buffer and 4  $\mu$ l RNaseOut and approximately 500  $\mu$ l worth of glass beads added. Cells were lysed by vortexing corex tubes 10 times for 1 min with 1 min on ice in between. Supernatant was then transferred to 2 ml tubes and stored at  $-80^{\circ}$ C until ready for further use.

Linear sucrose gradients were prepared in Beckman Coulter 14  $\times$  96 mm centrifuge tubes with sequential loading of sucrose stocks in gradient buffer (50 mM Tris-HCl pH 7.0, 50 mM NH<sub>4</sub>Cl, 0.5% NP40, 4 mM MgCl<sub>2</sub>, 1 mM DTT and 100  $\mu$ g/ml cycloheximide). From top to bottom, tubes contained 2.5 ml 17.5%, 2.5 ml 25%, 2 ml 33%, 2 ml 41% and 2.5 ml 50% sucrose solutions. Lysate was loaded to the top of the sucrose gradients and spun in a swinging bucket rotor Beckman Coulter SW40TI in a L-90K Preparative Ultracentrifuge at 38000 rpm for 2.5 hr at 4 $^{\circ}$ C, max acceleration, no break. Gradients were fractionated by upward displacement with 60% sucrose at 3 ml/min and absorbance was monitored at 254 nm.

### Proteomics

Cell pellets were resuspended in 25  $\mu$ l 1% SDC (sodium deoxycholate), 100 mM Tris and broken open with cryo-pulverisation using a cryogenic mixer mill (50 small ball bearings per sample, 4  $\times$  45 s at a frequency of 30 Hz with intermittent cooling in liquid nitrogen for 1 min). The protein content was estimated using the bicinchoninic acid method. Samples were denatured and alkylated for 5 min at 95 $^{\circ}$ C with TCEP (Tris(2-carboxyethyl phosphine hydrochloride)) and CAA (2-chloroacetamide) at final concentrations of 10 mM TCEP and 40 mM CAA. Sequencing-grade trypsin was added at an enzyme:protein ratio of 1:100 and incubated overnight at 37 $^{\circ}$ C. Protein digestion was arrested through the addition of formic acid to a final concentration of 1% and two-phase extraction with water-saturated ethyl acetate was used to remove SDC. The aqueous phase containing the peptides was collected, concentrated in a vacuum concentrator and reconstituted in buffer A (0.1% formic acid).

Data-independent acquisition (DIA) mass spectrometry was performed on an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific) coupled to a Dionex UltiMate 3000 RSLCnano system and a Dionex UltiMate 3000 RS autosampler. The samples were loaded onto an Acclaim PepMap 100 trap column (100  $\mu$ m  $\times$  2 cm; nanoViper; C18; 5  $\mu$ m; 100  $\text{\AA}$ ) and separated on an Acclaim PepMap RSLC analytical column (75  $\mu$ m  $\times$  5 cm; nanoViper; C18; 2  $\mu$ m; 100  $\text{\AA}$ ). The parameters for the

mass spectrometric DIA acquisition were described previously (Deo *et al.*, 2020). The acquired DIA data was analysed in Spectronaut 8 Cassini (Biognosys) using an in-house generated spectral library established by acquiring the same samples with the same set-up in data-dependent acquisition mode.

### TVN-PAT

TVN-PAT was performed as previously described (Jänicke *et al.*, 2012). To create TVN-PAT cDNA, 1 µg of total RNA and 1 µl TVN primer were combined and brought to a volume of 12 µl with dH<sub>2</sub>O in 200 µl PCR tubes. Mixtures were incubated at 80°C for 5 min then cooled to 42°C for 1 min. Once cooled, tubes were flash-centrifuged and 7 µl of master mix was added containing 4 µl 5x SuperScript III buffer, 1 µl 100 mM DTT, 1 µl 10 mM dNTPs and 1 µl RNaseOut per reaction. Samples were mixed thoroughly by inversion then flash-centrifuged and held at 42°C for 1 min. At this temperature, 1 µl of SuperScript III was added followed by mixing by inversion and flash-centrifuging. Samples were then further incubated at 42°C for 15 min, 48°C for 15 min and 55°C for 15 min followed by inactivation of reverse transcription at 80°C for 5 min and cooling to 12°C.

For PCR amplification, cDNA was diluted 1:10 with the addition of 200 µl of dH<sub>2</sub>O. 5 µl diluted cDNA was added to fresh 200 µl PCR tubes with 15 µl master mix containing 0.2 µl PAT assay reverse primer, 0.2 µl *OM14* gene-specific forward primer, 10 µl Amplitaq Gold 360 Master Mix and 4.6 µl dH<sub>2</sub>O per reaction. Amplification was performed with cycle conditions of 95°C for 10 min, then 28 cycles of 95°C for 20 s, 60°C for 20 s and 72°C for 30 s, followed by 1 min at 72°C.

Four µl 6x Orange G loading dye was added to each PCR sample. 12 µl was run on a 2% high resolution agarose gel pre-stained with SYBR safe. To estimate PCR product sizes, band migration was determined relative to a 100 bp ladder.

### Seahorse metabolic analysis

BY4741 cells were subjected to Seahorse XF Cell Energy Phenotype assay per the manufacturer's instructions (Agilent) with modifications as follows. Cells were grown at 30°C to OD<sub>600</sub> 0.6 in synthetic media then resuspended in assay medium (Seahorse XF media, 20 µg/ml uracil, 5 mg/ml glucose, 5 µg/ml ammonium sulphate) and plated (5 × 10<sup>4</sup> per well) in CellTak-coated XF96 plates via centrifugation (1000 g, 5 min, RT). After 1 hr recovery at 30°C, Oxygen Consumption Rate (OCR) and Extracellular Acidification Rate (ECAR) were determined using an XF96 extracellular flux analyser (Seahorse Biosciences) at 30°C. Using cycling parameters 1 min mix, 1 min wait, 2 min measure, basal activity was measured for four cycles. After addition of 20 µg/ml cordycepin, or an equivalent volume of DMSO or assay media (controls), a further 10 cycles were measured before uncoupling of mitochondrial function via stressor cocktail (3 µM FCCP, 1 µM oligomycin), followed by a further six measurement cycles. Data are representative of three independent experimental repeats, with 12 wells per condition in each. Graphs were prepared using Prism seven software.

### Acknowledgements

We thank the following researchers for strain and plasmid resources: David Brow, Michael Culbertson, Walter Keller, Lionel Minvielle-Sebastia, Beate Schwer, Mike Stark, Maurice Swanson, Françoise Wyers, Craig Kaplan and Wendy Olivas. Melissa J Curtis is acknowledged for her technical assistance in the experimental data collection early in the project, and members of the Beilharz laboratory provided critical feedback. The MHTP Medical Genomics Facility, the Monash Proteomics and Metabolomics Facility, Micromon, and the Monash Bioinformatics Platform are thanked for the provision of technical support and infrastructure. Monash research technology platforms are enabled by Bioplatforms Australia (BPA) and the National Collaborative Research Infrastructure Strategy (NCRIS). RET was supported by an Australian Postgraduate Research award. THB was supported by a Monash Bio Discovery Fellowship and grants from the Australian Research Council (ARC: DP170100569 and FT180100049).

## Additional information

### Funding

Funder	Grant reference number	Author
Australian Research Council	DP170100569	Traude H Beilharz
Australian Research Council	FT180100049	Traude H Beilharz

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

Rachael Emily Turner, Data curation, Formal analysis, Validation, Visualization, Methodology, Writing - original draft; Paul F Harrison, Data curation, Formal analysis, Supervision, Visualization, Writing - review and editing; Angavai Swaminathan, Supervision, Investigation, Methodology, Writing - review and editing; Calvin A Kraupner-Taylor, Data curation, Formal analysis, Investigation, Methodology, Writing - review and editing; Belinda J Goldie, Amanda L Peterson, Data curation, Formal analysis, Methodology, Writing - review and editing; Michael See, Data curation, Formal analysis, Supervision, Writing - review and editing; Ralf B Schittenhelm, Resources, Data curation, Formal analysis, Visualization, Methodology, Writing - review and editing; David R Powell, Resources, Data curation, Supervision, Writing - review and editing; Darren J Creek, Resources, Data curation, Formal analysis, Methodology, Writing - review and editing; Bernhard Dichtl, Conceptualization, Resources, Formal analysis, Writing - original draft, Writing - review and editing; Traude H Beilharz, Conceptualization, Resources, Supervision, Funding acquisition, Investigation, Project administration, Writing - review and editing

### Author ORCIDs

Rachael Emily Turner  <https://orcid.org/0000-0001-9319-4825>

Paul F Harrison  <http://orcid.org/0000-0002-3980-268X>

Michael See  <http://orcid.org/0000-0002-7231-3896>

Traude H Beilharz  <https://orcid.org/0000-0002-8942-9502>

### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.65331.sa1>

Author response <https://doi.org/10.7554/eLife.65331.sa2>

## Additional files

### Supplementary files

- Transparent reporting form

### Data availability

Sequencing data have been deposited in GEO under accession code GSE160539.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Turner RE, Harrison PF, Swaminathan A, Kraupner-Taylor CA, Goldie BJ, See M, Peterson AL, Schittenhelm RB, Powell DR, Creek DJ, Dichtl B, Beilharz TH	2021	Genetic and pharmacological evidence for kinetic competition between alternative poly(A) sites in yeast	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?&amp;acc=GSE160539">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?&amp;acc=GSE160539</a>	NCBI Gene Expression Omnibus, GSE160539

## References

- Alzu A**, Bermejo R, Begnis M, Lucca C, Piccini D, Carotenuto W, Saponaro M, Brambati A, Cocito A, Foiani M, Liberi G. 2012. Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* **151**:835–846. DOI: <https://doi.org/10.1016/j.cell.2012.09.041>, PMID: 23141540
- Amrani N**, Minet M, Wyers F, Dufour ME, Aggerbeck LP, Lacroute F. 1997. PCF11 encodes a third protein component of yeast cleavage and polyadenylation factor I. *Molecular and Cellular Biology* **17**:1102–1109. DOI: <https://doi.org/10.1128/MCB.17.3.1102>, PMID: 9032237
- Baejen C**, Torkler P, Gressel S, Essig K, Söding J, Cramer P. 2014. Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition. *Molecular Cell* **55**:745–757. DOI: <https://doi.org/10.1016/j.molcel.2014.08.005>, PMID: 25192364
- Barabino SM**, Hübner W, Jenny A, Minvielle-Sebastia L, Keller W. 1997. The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes & Development* **11**:1703–1716. DOI: <https://doi.org/10.1101/gad.11.13.1703>, PMID: 9224719
- Barabino SM**, Ohnacker M, Keller W. 2000. Distinct roles of two Yth1p domains in 3'-end cleavage and polyadenylation of yeast pre-mRNAs. *The EMBO Journal* **19**:3778–3787. DOI: <https://doi.org/10.1093/emboj/19.14.3778>, PMID: 10899131
- Barillà D**, Lee BA, Proudfoot NJ. 2001. Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in *Saccharomyces cerevisiae*. *PNAS* **98**:445–450. DOI: <https://doi.org/10.1073/pnas.98.2.445>, PMID: 11149954
- Beaudoing E**, Freier S, Wyatt JR, Claverie JM, Gautheret D. 2000. Patterns of variant polyadenylation signal usage in human genes. *Genome Research* **10**:1001–1010. DOI: <https://doi.org/10.1101/gr.10.7.1001>, PMID: 10899149
- Beilharz TH**, Harrison PF, Miles DM, See MM, Le UM, Kalanon M, Curtis MJ, Hasan Q, Saksouk J, Margaritis T, Holstege F, Geli V, Dichtl B. 2017. Coordination of cell cycle progression and mitotic spindle assembly involves histone H3 lysine 4 methylation by Set1/COMPASS. *Genetics* **205**:185–199. DOI: <https://doi.org/10.1534/genetics.116.194852>, PMID: 28049706
- Belenky PA**, Moga TG, Brenner C. 2008. *Saccharomyces cerevisiae* YOR071C encodes the high affinity nicotinamide riboside transporter Nrt1. *Journal of Biological Chemistry* **283**:8075–8079. DOI: <https://doi.org/10.1074/jbc.C800021200>, PMID: 18258590
- Bloch JC**, Perrin F, Lacroute F. 1978. Yeast temperature-sensitive mutants specifically impaired in processing of poly(A)-containing RNAs. *Molecular and General Genetics MGG* **165**:123–127. DOI: <https://doi.org/10.1007/BF00269900>, PMID: 366372
- Brachmann CB**, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD. 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**:115–132. DOI: [https://doi.org/10.1002/\(SICI\)1097-0061\(19980130\)14:2<115::AID-YEA204>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2), PMID: 9483801
- Broach JR**. 2012. Nutritional control of growth and development in yeast. *Genetics* **192**:73–105. DOI: <https://doi.org/10.1534/genetics.111.135731>, PMID: 22964838
- Cai L**, Sutter BM, Li B, Tu BP. 2011. Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Molecular Cell* **42**:426–437. DOI: <https://doi.org/10.1016/j.molcel.2011.05.004>, PMID: 21596309
- Chan S**, Choi EA, Shi Y. 2011. Pre-mRNA 3'-end processing complex assembly and function. *Wiley Interdisciplinary Reviews: RNA* **2**:321–335. DOI: <https://doi.org/10.1002/wrna.54>, PMID: 21957020
- Chen S**, Hyman LE. 1998. A specific RNA-protein interaction at yeast polyadenylation efficiency elements. *Nucleic Acids Research* **26**:4965–4974. DOI: <https://doi.org/10.1093/nar/26.21.4965>, PMID: 9776761
- Colgan DF**, Manley JL. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes & Development* **11**:2755–2766. DOI: <https://doi.org/10.1101/gad.11.21.2755>, PMID: 9353246
- Conrad NK**, Wilson SM, Steinmetz EJ, Patturajan M, Brow DA, Swanson MS, Corden JL. 2000. A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics* **154**:557–571. DOI: <https://doi.org/10.1093/genetics/154.2.557>, PMID: 10655211
- Creek DJ**, Jankevics A, Breitling R, Watson DG, Barrett MP, Burgess KE. 2011. Toward global metabolomics analysis with hydrophilic interaction liquid chromatography-mass spectrometry: improved metabolite identification by retention time prediction. *Analytical Chemistry* **83**:8703–8710. DOI: <https://doi.org/10.1021/ac2021823>, PMID: 21928819
- Creek DJ**, Jankevics A, Burgess KE, Breitling R, Barrett MP. 2012. IDEOM: an excel interface for analysis of LC-MS-based metabolomics data. *Bioinformatics* **28**:1048–1049. DOI: <https://doi.org/10.1093/bioinformatics/bts069>, PMID: 22308147
- Creek DJ**, Nijagal B, Kim DH, Rojas F, Matthews KR, Barrett MP. 2013. Metabolomics guides rational development of a simplified cell culture medium for drug screening against *Trypanosoma brucei*. *Antimicrobial Agents and Chemotherapy* **57**:2768–2779. DOI: <https://doi.org/10.1128/AAC.00044-13>, PMID: 23571546
- Deo P**, Chow SH, Han ML, Speir M, Huang C, Schittenhelm RB, Dhital S, Emery J, Li J, Kile BT, Vince JE, Lawlor KE, Naderer T. 2020. Mitochondrial dysfunction caused by outer membrane vesicles from Gram-negative *Bacteria* activates intrinsic apoptosis and inflammation. *Nature Microbiology* **5**:1418–1427. DOI: <https://doi.org/10.1038/s41564-020-0773-2>, PMID: 32807891

- Derti A**, Garrett-Engle P, Macisaac KD, Stevens RC, Sriram S, Chen R, Rohl CA, Johnson JM, Babak T. 2012. A quantitative atlas of polyadenylation in five mammals. *Genome Research* **22**:1173–1183. DOI: <https://doi.org/10.1101/gr.132563.111>, PMID: 22454233
- Dheur S**, Vo leTA, Voisin-Hakil F, Minet M, Schmitter JM, Lacroute F, Wyers F, Minvielle-Sebastia L. 2003. Pti1p and Ref2p found in association with the mRNA 3' end formation complex direct snoRNA maturation. *The EMBO Journal* **22**:2831–2840. DOI: <https://doi.org/10.1093/emboj/cdg253>, PMID: 12773397
- Dichtl B**, Blank D, Ohnacker M, Friedlein A, Roeder D, Langen H, Keller W. 2002a. A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. *Molecular Cell* **10**:1139–1150. DOI: [https://doi.org/10.1016/S1097-2765\(02\)00707-4](https://doi.org/10.1016/S1097-2765(02)00707-4), PMID: 12453421
- Dichtl B**, Blank D, Sadowski M, Hübner W, Weiser S, Keller W. 2002b. Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. *The EMBO Journal* **21**:4125–4135. DOI: <https://doi.org/10.1093/emboj/cdf390>, PMID: 12145212
- Dichtl B**, Aasland R, Keller W. 2004. Functions for *S. cerevisiae* Swd2p in 3' end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. *RNA* **10**:965–977. DOI: <https://doi.org/10.1261/rna.7090104>, PMID: 15146080
- Dichtl B**, Keller W. 2001. Recognition of polyadenylation sites in yeast pre-mRNAs by cleavage and polyadenylation factor. *The EMBO Journal* **20**:3197–3209. DOI: <https://doi.org/10.1093/emboj/20.12.3197>, PMID: 11406596
- Donohoe DR**, Collins LB, Wali A, Bigler R, Sun W, Bultman SJ. 2012. The warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Molecular Cell* **48**:612–626. DOI: <https://doi.org/10.1016/j.molcel.2012.08.033>, PMID: 23063526
- Eleuthero E**, Panek A, De Mesquita JF, Trevisol E, Magalhães R. 2015. Revisiting yeast trehalose metabolism. *Current Genetics* **61**:263–274. DOI: <https://doi.org/10.1007/s00294-014-0450-1>, PMID: 25209979
- Enjo F**, Nosaka K, Ogata M, Iwashima A, Nishimura H. 1997. Isolation and characterization of a thiamin transport gene, THI10, from *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **272**:19165–19170. DOI: <https://doi.org/10.1074/jbc.272.31.19165>, PMID: 9235906
- Exinger F**, Lacroute F. 1992. 6-Azauracil inhibition of GTP biosynthesis in *Saccharomyces cerevisiae*. *Current Genetics* **22**:9–11. DOI: <https://doi.org/10.1007/BF00351735>, PMID: 1611672
- Garas M**, Dichtl B, Keller W. 2008. The role of the putative 3' end processing endonuclease Ysh1p in mRNA and snoRNA synthesis. *RNA* **14**:2671–2684. DOI: <https://doi.org/10.1261/rna.1293008>, PMID: 18971324
- Geisberg JV**, Moqtaderi Z, Struhl K. 2020. The transcriptional elongation rate regulates alternative polyadenylation in yeast. *eLife* **9**:e59810. DOI: <https://doi.org/10.7554/eLife.59810>, PMID: 32845240
- Graber JH**, Cantor CR, Mohr SC, Smith TF. 1999. In silico detection of control signals: mrna 3'-end-processing sequences in diverse species. *PNAS* **96**:14055–14060. DOI: <https://doi.org/10.1073/pnas.96.24.14055>, PMID: 10570197
- Gross S**, Moore CL. 2001. Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation. *Molecular and Cellular Biology* **21**:8045–8055. DOI: <https://doi.org/10.1128/MCB.21.23.8045-8055.2001>, PMID: 11689695
- Grzechnik P**, Gdula MR, Proudfoot NJ. 2015. Pcf11 orchestrates transcription termination pathways in yeast. *Genes & Development* **29**:849–861. DOI: <https://doi.org/10.1101/gad.251470.114>, PMID: 25877920
- Gullerova M**, Proudfoot NJ. 2010. Transcriptional interference and gene orientation in yeast: noncoding RNA connections. Cold Spring Harbor Symposia on Quantitative Biology 299–311. DOI: <https://doi.org/10.1101/sqb.2010.75.048>
- Guo Z**, Sherman F. 1996. 3'-end-forming signals of yeast mRNA. *Trends in Biochemical Sciences* **21**:477–481. DOI: [https://doi.org/10.1016/S0968-0004\(96\)10057-8](https://doi.org/10.1016/S0968-0004(96)10057-8), PMID: 9009831
- Han Z**, Jasnovidova O, Haidara N, Tudek A, Kubicek K, Libri D, Stefl R, Porrua O. 2020. Termination of non-coding transcription in yeast relies on both an RNA pol II CTD interaction domain and a CTD-mimicking region in Sen1. *The EMBO Journal* **39**:e101548. DOI: <https://doi.org/10.15252/emboj.2019101548>, PMID: 32107786
- Harrison PF**, Powell DR, Clancy JL, Preiss T, Boag PR, Traven A, Seemann T, Beilharz TH. 2015. PAT-seq: a method to study the integration of 3'-UTR dynamics with gene expression in the eukaryotic transcriptome. *RNA* **21**:1502–1510. DOI: <https://doi.org/10.1261/rna.048355.114>, PMID: 26092945
- Harrison P**. 2018. Fitnoise. *GitHub*. <https://github.com/pfh/fitnoise>
- Harrison PF**, Pattison AD, Powell DR, Beilharz TH. 2019. Topconfects: a package for confident effect sizes in differential expression analysis provides a more biologically useful ranked gene list. *Genome Biology* **20**:67. DOI: <https://doi.org/10.1186/s13059-019-1674-7>, PMID: 30922379
- Hodges C**, Bintu L, Lubkowska L, Kashlev M, Bustamante C. 2009. Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science* **325**:626–628. DOI: <https://doi.org/10.1126/science.1172926>, PMID: 19644123
- Holbein S**, Wengi A, Decourty L, Freimoser FM, Jacquier A, Dichtl B. 2009. Cordycepin interferes with 3' end formation in yeast independently of its potential to terminate RNA chain elongation. *RNA* **15**:837–849. DOI: <https://doi.org/10.1261/rna.1458909>, PMID: 19324962
- Holbein S**, Scola S, Loll B, Dichtl BS, Hübner W, Meinhart A, Dichtl B. 2011. The P-loop domain of yeast Clp1 mediates interactions between CF IA and CPF factors in pre-mRNA 3' end formation. *PLOS ONE* **6**:e29139. DOI: <https://doi.org/10.1371/journal.pone.0029139>, PMID: 22216186
- Horowitz B**, Goldfinger BA, Marmur J. 1976. Effect of cordycepin triphosphate on the nuclear DNA-dependent RNA polymerases and poly(A) polymerase from the yeast, *Saccharomyces cerevisiae*. *Archives of Biochemistry and Biophysics* **172**:143–148. DOI: [https://doi.org/10.1016/0003-9861\(76\)90059-X](https://doi.org/10.1016/0003-9861(76)90059-X), PMID: 766704

- Jain NK, Roy I. 2009. Effect of trehalose on protein structure. *Protein Science: A Publication of the Protein Society* **18**:24–36. DOI: <https://doi.org/10.1002/pro.3>, PMID: 19177348
- Jänicke A, Vancuylenberg J, Boag PR, Traven A, Beilharz TH. 2012. ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. *RNA* **18**:1289–1295. DOI: <https://doi.org/10.1261/rna.031898.111>, PMID: 22543866
- Ji Z, Lee JY, Pan Z, Jiang B, Tian B. 2009. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *PNAS* **106**:7028–7033. DOI: <https://doi.org/10.1073/pnas.0900028106>, PMID: 19372383
- Kamieniarz-Gdula K, Proudfoot NJ. 2019. Transcriptional control by premature termination: a forgotten mechanism. *Trends in Genetics* **35**:553–564. DOI: <https://doi.org/10.1016/j.tig.2019.05.005>, PMID: 31213387
- Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, Segal E. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**:362–366. DOI: <https://doi.org/10.1038/nature07667>, PMID: 19092803
- Kaplan CD, Jin H, Zhang IL, Belyanin A. 2012. Dissection of pol II trigger loop function and pol II activity-dependent control of start site selection in vivo. *PLOS Genetics* **8**:e1002627. DOI: <https://doi.org/10.1371/journal.pgen.1002627>, PMID: 22511879
- Kessler MM, Henry MF, Shen E, Zhao J, Gross S, Silver PA, Moore CL. 1997. Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes & Development* **11**:2545–2556. DOI: <https://doi.org/10.1101/gad.11.19.2545>, PMID: 9334319
- Konrad M. 1988. Analysis and in vivo disruption of the gene coding for adenylate kinase (ADK1) in the yeast *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **263**:19468–19474. DOI: [https://doi.org/10.1016/S0021-9258\(19\)77657-1](https://doi.org/10.1016/S0021-9258(19)77657-1), PMID: 2848829
- Law CW, Chen Y, Shi W, Smyth GK. 2014. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* **15**:R29. DOI: <https://doi.org/10.1186/gb-2014-15-2-r29>, PMID: 24485249
- Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. 2013. Software for computing and annotating genomic ranges. *PLOS Computational Biology* **9**:e1003118. DOI: <https://doi.org/10.1371/journal.pcbi.1003118>, PMID: 23950696
- Lecoq K, Belloc I, Desgranges C, Daignan-Fornier B. 2001. Role of Adenosine kinase in *Saccharomyces cerevisiae*: identification of the ADO1 gene and study of the mutant phenotypes. *Yeast* **18**:335–342. DOI: [https://doi.org/10.1002/1097-0061\(20010315\)18:4<335::AID-YEA674>3.0.CO;2-X](https://doi.org/10.1002/1097-0061(20010315)18:4<335::AID-YEA674>3.0.CO;2-X), PMID: 11223943
- Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD. 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nature Genetics* **36**:900–905. DOI: <https://doi.org/10.1038/ng1400>, PMID: 15247917
- Lee JV, Carrer A, Shah S, Snyder NW, Wei S, Venneti S, Worth AJ, Yuan ZF, Lim HW, Liu S, Jackson E, Aiello NM, Haas NB, Rebbeck TR, Judkins A, Won KJ, Chodosh LA, Garcia BA, Stanger BZ, Feldman MD, et al. 2014. Akt-dependent metabolic reprogramming regulates tumor cell histone acetylation. *Cell Metabolism* **20**:306–319. DOI: <https://doi.org/10.1016/j.cmet.2014.06.004>, PMID: 24998913
- Lemay JF, Marguerat S, Larochelle M, Liu X, van Nues R, Hunyadkúrti J, Hoque M, Tian B, Granneman S, Bähler J, Bachand F. 2016. The Nrd1-like protein Seb1 coordinates cotranscriptional 3' end processing and polyadenylation site selection. *Genes & Development* **30**:1558–1572. DOI: <https://doi.org/10.1101/gad.280222.116>, PMID: 27401558
- Malik I, Qiu C, Snavely T, Kaplan CD. 2017. Wide-ranging and unexpected consequences of altered pol II catalytic activity in vivo. *Nucleic Acids Research* **45**:4431–4451. DOI: <https://doi.org/10.1093/nar/gkx037>, PMID: 28119420
- Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, Tong L. 2006. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **444**:953–956. DOI: <https://doi.org/10.1038/nature05363>, PMID: 17128255
- Mandel CR, Bai Y, Tong L. 2008. Protein factors in pre-mRNA 3'-end processing. *Cellular and Molecular Life Sciences* **65**:1099–1122. DOI: <https://doi.org/10.1007/s00018-007-7474-3>, PMID: 18158581
- Martin G, Gruber AR, Keller W, Zavolan M. 2012. Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell Reports* **1**:753–763. DOI: <https://doi.org/10.1016/j.celrep.2012.05.003>, PMID: 22813749
- Masamha CP, Xia Z, Yang J, Albrecht TR, Li M, Shyu AB, Li W, Wagner EJ. 2014. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* **510**:412–416. DOI: <https://doi.org/10.1038/nature13261>, PMID: 24814343
- Mason PB, Struhl K. 2005. Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. *Molecular Cell* **17**:831–840. DOI: <https://doi.org/10.1016/j.molcel.2005.02.017>, PMID: 15780939
- Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh BF. 2008. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Research* **18**:1073–1083. DOI: <https://doi.org/10.1101/gr.078261.108>, PMID: 18550805
- Mayr C, Bartel DP. 2009. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**:673–684. DOI: <https://doi.org/10.1016/j.cell.2009.06.016>, PMID: 19703394
- McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**:357–361. DOI: <https://doi.org/10.1038/385357a0>, PMID: 9002523



- Miller C, Schwalb B, Maier K, Schulz D, Dümcke S, Zacher B, Mayer A, Sydow J, Marcinowski L, Dölken L, Martin DE, Tresch A, Cramer P. 2011. Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Molecular Systems Biology* **7**:458. DOI: <https://doi.org/10.1038/msb.2010.112>, PMID: 21206491
- Minvielle-Sebastia L, Preker PJ, Keller W. 1994. RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3'-end processing factor. *Science* **266**:1702–1705. DOI: <https://doi.org/10.1126/science.7992054>, PMID: 7992054
- Minvielle-Sebastia L, Beyer K, Krecic AM, Hector RE, Swanson MS, Keller W. 1998. Control of cleavage site selection during mRNA 3' end formation by a yeast hnRNP. *The EMBO Journal* **17**:7454–7468. DOI: <https://doi.org/10.1093/emboj/17.24.7454>, PMID: 9857200
- Mischo HE, Gómez-González B, Grzechnik P, Rondón AG, Wei W, Steinmetz L, Aguilera A, Proudfoot NJ. 2011. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Molecular Cell* **41**:21–32. DOI: <https://doi.org/10.1016/j.molcel.2010.12.007>, PMID: 21211720
- Müller WE, Seibert G, Beyer R, Breter HJ, Maidhof A, Zahn RK. 1977. Effect of cordycepin on nucleic acid metabolism in L5178Y cells and on nucleic acid-synthesizing enzyme systems. *Cancer Research* **37**:3824–3833.
- Munchel SE, Shultzaberger RK, Takizawa N, Weis K. 2011. Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay. *Molecular Biology of the Cell* **22**:2787–2795. DOI: <https://doi.org/10.1091/mbc.e11-01-0028>, PMID: 21680716
- Nonet M, Scafe C, Sexton J, Young R. 1987. Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. *Molecular and Cellular Biology* **7**:1602–1611. DOI: <https://doi.org/10.1128/MCB.7.5.1602>, PMID: 3299050
- O'Connor JP, Peebles CL. 1992. PTA1, an essential gene of *Saccharomyces cerevisiae* affecting pre-tRNA processing. *Molecular and Cellular Biology* **12**:3843–3856. DOI: <https://doi.org/10.1128/MCB.12.9.3843>, PMID: 1508188
- Ogorodnikov A, Levin M, Tattikota S, Tokalov S, Hoque M, Scherzinger D, Marini F, Poetsch A, Binder H, Macher-Göppinger S, Probst HC, Tian B, Schaefer M, Lackner KJ, Westermann F, Danckwardt S. 2018. Transcriptome 3' end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nature Communications* **9**:5331. DOI: <https://doi.org/10.1038/s41467-018-07580-5>, PMID: 30552333
- Ohnacker M, Barabino SM, Preker PJ, Keller W. 2000. The WD-repeat protein pfs2p bridges two essential factors within the yeast pre-mRNA 3'-end-processing complex. *The EMBO Journal* **19**:37–47. DOI: <https://doi.org/10.1093/emboj/19.1.37>, PMID: 10619842
- Olivas W, Parker R. 2000. The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast. *The EMBO Journal* **19**:6602–6611. DOI: <https://doi.org/10.1093/emboj/19.23.6602>, PMID: 11101532
- Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, John B, Milos PM. 2010. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* **143**:1018–1029. DOI: <https://doi.org/10.1016/j.cell.2010.11.020>
- Pappas DL, Hampsey M. 2000. Functional interaction between Ssu72 and the Rpb2 subunit of RNA polymerase II in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **20**:8343–8351. DOI: <https://doi.org/10.1128/MCB.20.22.8343-8351.2000>, PMID: 11046131
- Patel D, Butler JS. 1992. Conditional defect in mRNA 3' end processing caused by a mutation in the gene for poly(A) polymerase. *Molecular and Cellular Biology* **12**:3297–3304. DOI: <https://doi.org/10.1128/MCB.12.7.3297>, PMID: 1620131
- Peggie MW, MacKelvie SH, Bloecher A, Knatko EV, Tatchell K, Stark MJ. 2002. Essential functions of Sds22p in chromosome stability and nuclear localization of PP1. *Journal of Cell Science* **115**:195–206. DOI: <https://doi.org/10.1242/jcs.115.1.195>, PMID: 11801737
- Pinto PA, Henriques T, Freitas MO, Martins T, Domingues RG, Wyrzykowska PS, Coelho PA, Carmo AM, Sunkel CE, Proudfoot NJ, Moreira A. 2011. RNA polymerase II kinetics in polo polyadenylation signal selection. *The EMBO Journal* **30**:2431–2444. DOI: <https://doi.org/10.1038/emboj.2011.156>, PMID: 21602789
- Preker PJ, Lingner J, Minvielle-Sebastia L, Keller W. 1995. The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell* **81**:379–389. DOI: [https://doi.org/10.1016/0092-8674\(95\)90391-7](https://doi.org/10.1016/0092-8674(95)90391-7), PMID: 7736590
- Preker PJ, Ohnacker M, Minvielle-Sebastia L, Keller W. 1997. A multisubunit 3' end processing factor from yeast containing poly(A) polymerase and homologues of the subunits of mammalian cleavage and polyadenylation specificity factor. *The EMBO Journal* **16**:4727–4737. DOI: <https://doi.org/10.1093/emboj/16.15.4727>, PMID: 9303317
- Qu X, Perez-Canadillas JM, Agrawal S, De Baecke J, Cheng H, Varani G, Moore C. 2007. The C-terminal domains of vertebrate CstF-64 and its yeast orthologue Rna15 form a new structure critical for mRNA 3'-end processing. *Journal of Biological Chemistry* **282**:2101–2115. DOI: <https://doi.org/10.1074/jbc.M609981200>, PMID: 17116658
- Ramirez A, Shuman S, Schwer B. 2008. Human RNA 5'-kinase (hClp1) can function as a tRNA splicing enzyme in vivo. *RNA* **14**:1737–1745. DOI: <https://doi.org/10.1261/rna.1142908>, PMID: 18648070
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nature Biotechnology* **29**:24–26. DOI: <https://doi.org/10.1038/nbt.1754>, PMID: 21221095
- Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11**:R25. DOI: <https://doi.org/10.1186/gb-2010-11-3-r25>, PMID: 20196867

- Rondón AG**, Mischo HE, Kawauchi J, Proudfoot NJ. 2009. Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. *Molecular Cell* **36**:88–98. DOI: <https://doi.org/10.1016/j.molcel.2009.07.028>, PMID: 19818712
- Rouillard JM**, Brendolise C, Lacroute F. 2000. Rna14p, a component of the yeast nuclear cleavage/polyadenylation factor I, is also localised in mitochondria. *Molecular and General Genetics MGG* **262**:1103–1112. DOI: <https://doi.org/10.1007/PL00008653>, PMID: 10660071
- Ryan K**, Calvo O, Manley JL. 2004. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA* **10**:565–573. DOI: <https://doi.org/10.1261/rna.5214404>, PMID: 15037765
- Sadowski M**, Dichtl B, Hübner W, Keller W. 2003. Independent functions of yeast Pcf1p in pre-mRNA 3' end processing and in transcription termination. *The EMBO Journal* **22**:2167–2177. DOI: <https://doi.org/10.1093/emboj/cdg200>, PMID: 12727883
- Sandberg R**, Neilson JR, Sarma A, Sharp PA, Burge CB. 2008. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**:1643–1647. DOI: <https://doi.org/10.1126/science.1155390>, PMID: 18566288
- Scheltema RA**, Jankevics A, Jansen RC, Swertz MA, Breitling R. 2011. PeakML/mzMatch: a file format, Java library, R library, and tool-chain for mass spectrometry data analysis. *Analytical Chemistry* **83**:2786–2793. DOI: <https://doi.org/10.1021/ac2000994>, PMID: 21401061
- Schmitt ME**, Brown TA, Trumpower BL. 1990. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Research* **18**:3091–3092. DOI: <https://doi.org/10.1093/nar/18.10.3091>, PMID: 2190191
- Schofield JA**, Duffy EE, Kiefer L, Sullivan MC, Simon MD. 2018. TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding. *Nature Methods* **15**:221–225. DOI: <https://doi.org/10.1038/nmeth.4582>, PMID: 29355846
- Schulz D**, Schwab B, Kiesel A, Baejen C, Torkler P, Gagneur J, Soeding J, Cramer P. 2013. Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell* **155**:1075–1087. DOI: <https://doi.org/10.1016/j.cell.2013.10.024>, PMID: 24210918
- Shaw RJ**, Reines D. 2000. *Saccharomyces cerevisiae* transcription elongation mutants are defective in PUR5 induction in response to nucleotide depletion. *Molecular and Cellular Biology* **20**:7427–7437. DOI: <https://doi.org/10.1128/MCB.20.20.7427-7437.2000>, PMID: 11003640
- Shepard PJ**, Choi E-A, Lu J, Flanagan LA, Hertel KJ, Shi Y. 2011. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA* **17**:761–772. DOI: <https://doi.org/10.1261/rna.2581711>
- Shivaswamy S**, Bhinge A, Zhao Y, Jones S, Hirst M, Iyer VR. 2008. Dynamic Remodeling of Individual Nucleosomes Across a Eukaryotic Genome in Response to Transcriptional Perturbation. *PLOS Biology* **6**:e65. DOI: <https://doi.org/10.1371/journal.pbio.0060065>
- Smyth GK**. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* **3**:1–25. DOI: <https://doi.org/10.2202/1544-6115.1027>
- Steinmetz EJ**, Conrad NK, Brow DA, Corden JL. 2001. RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature* **413**:327–331. DOI: <https://doi.org/10.1038/35095090>, PMID: 11565036
- Steinmetz EJ**, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. 2006. Genome-Wide Distribution of Yeast RNA Polymerase II and Its Control by Sen1 Helicase. *Molecular Cell* **24**:735–746. DOI: <https://doi.org/10.1016/j.molcel.2006.10.023>
- Steinmetz EJ**, Brow DA. 1996. Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. *Molecular and Cellular Biology* **16**:6993–7003. DOI: <https://doi.org/10.1128/MCB.16.12.6993>
- Steinmetz EJ**, Brow DA. 1998. Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association. *PNAS* **95**:6699–6704. DOI: <https://doi.org/10.1073/pnas.95.12.6699>
- Stoessel D**, Nowell CJ, Jones AJ, Ferrins L, Ellis KM, Riley J, Rahmani R, Read KD, McConville MJ, Avery VM, Baell JB, Creek DJ. 2016. Metabolomics and lipidomics reveal perturbation of sphingolipid metabolism by a novel anti-trypanosomal 3-(oxazololo[4,5-b]pyridine-2-yl)anilide. *Metabolomics* **12**:12. DOI: <https://doi.org/10.1007/s11306-016-1062-1>
- Sugar AM**, McCaffrey RP. 1998. Antifungal Activity of 3-Deoxyadenosine (Cordycepin). *Antimicrobial Agents and Chemotherapy* **42**:1424–1427. DOI: <https://doi.org/10.1128/AAC.42.6.1424>
- Swaminathan A**, Beilharz TH. 2016. Epitope-tagged yeast strains reveal promoter driven changes to 3'-end formation and convergent antisense-transcription from common 3' UTRs. *Nucleic Acids Research* **44**:377–386. DOI: <https://doi.org/10.1093/nar/gkv1022>, PMID: 26481348
- Thomas BJ**, Rothstein R. 1989. Elevated recombination rates in transcriptionally active DNA. *Cell* **56**:619–630. DOI: [https://doi.org/10.1016/0092-8674\(89\)90584-9](https://doi.org/10.1016/0092-8674(89)90584-9)
- Tian B**, Graber JH. 2012. Signals for pre-mRNA cleavage and polyadenylation. *Wiley Interdisciplinary Reviews: RNA* **3**:385–396. DOI: <https://doi.org/10.1002/wrna.116>
- Trochine A**, Creek DJ, Faral-Tello P, Barrett MP, Robello C. 2014. Benzimidazole Biotransformation and Multiple Targets in *Trypanosoma cruzi* Revealed by Metabolomics. *PLOS Neglected Tropical Diseases* **8**:e2844. DOI: <https://doi.org/10.1371/journal.pntd.0002844>

- Turner RE**, Henneken LM, Liem-Weits M, Harrison PF, Swaminathan A, Vary R, Nikolic I, Simpson KJ, Powell DR, Beilharz TH, Dichtl B. 2020. Requirement for cleavage factor II<sub>m</sub> in the control of alternative polyadenylation in breast Cancer cells. *RNA* **26**:969–981. DOI: <https://doi.org/10.1261/rna.075226.120>, PMID: 32295865
- Uptain SM**, Kane CM, Chamberlin MJ. 1997. Basic mechanisms of transcript elongation and its regulation. *Annual Review of Biochemistry* **66**:117–172. DOI: <https://doi.org/10.1146/annurev.biochem.66.1.117>, PMID: 9242904
- Valentini SR**, Weiss VH, Silver PA. 1999. Arginine methylation and binding of Hrp1p to the efficiency element for mRNA 3'-end formation. *RNA* **5**:272–280. DOI: <https://doi.org/10.1017/S1355838299981633>, PMID: 10024178
- Vasiljeva L**, Kim M, Mutschler H, Buratowski S, Meinhart A. 2008. The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nature Structural & Molecular Biology* **15**:795–804. DOI: <https://doi.org/10.1038/nsmb.1468>, PMID: 18660819
- Vo LT**, Minet M, Schmitter JM, Lacroute F, Wyers F. 2001. Mpe1, a zinc knuckle protein, is an essential component of yeast cleavage and polyadenylation factor required for the cleavage and polyadenylation of mRNA. *Molecular and Cellular Biology* **21**:8346–8356. DOI: <https://doi.org/10.1128/MCB.21.24.8346-8356.2001>, PMID: 11713271
- Wang R**, Zheng D, Wei L, Ding Q, Tian B. 2019. Regulation of intronic polyadenylation by PCF11 impacts mRNA expression of long genes. *Cell Reports* **26**:2766–2778. DOI: <https://doi.org/10.1016/j.celrep.2019.02.049>
- Winey M**, Culbertson MR. 1988. Mutations affecting the tRNA-splicing endonuclease activity of *Saccharomyces cerevisiae*. *Genetics* **118**:609–617. DOI: <https://doi.org/10.1093/genetics/118.4.609>
- Xu Z**, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM. 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**:1033–1037. DOI: <https://doi.org/10.1038/nature07728>, PMID: 19169243
- Zhang Z**, Fu J, Gilmour DS. 2005. CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes & Development* **19**:1572–1580. DOI: <https://doi.org/10.1101/gad.1296305>
- Zhao J**, Hyman L, Moore C. 1999. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiology and Molecular Biology Reviews* **63**:405–445. DOI: <https://doi.org/10.1128/MMBR.63.2.405-445.1999>, PMID: 10357856
- Zhelkovsky A**, Tacahashi Y, Nasser T, He X, Sterzer U, Jensen TH, Domdey H, Moore C. 2006. The role of the Brr5/Ysh1 C-terminal domain and its homolog Syc1 in mRNA 3'-end processing in *Saccharomyces cerevisiae*. *RNA* **12**:435–445. DOI: <https://doi.org/10.1261/rna.2267606>