General amino acid control in fission yeast is regulated by a non-conserved transcription factor, with functions analogous to Gcn4/Atf4

Caia D. S. Duncan[1], Maria Rodriguez Lopez[2], Phil Ruis[1,3], Jürg Bähler[2] and Juan Mata[1,4]

1. Department of Biochemistry, University of Cambridge, UK

2. Department of Genetics, Evolution & Environment and Institute of Healthy Ageing, University College London, UK.

3. Current address: The Francis Crick Institute, London, UK

4. Corresponding author

Classification: Biological Sciences, Genetics

Corresponding author contact details:

Dr Juan Mata

Email: jm593@cam.ac.uk
Phone: +44 1223 760467
Address: Department of Biochemistry, Hopkins Building
University of Cambridge
Cambridge CB2 1QW
UK

Abstract

Eukaryotes respond to amino acid starvation by enhancing the translation of mRNAs encoding b-ZIP family transcription factors (*GCN4* in *Saccharomyces cerevisiae* and *ATF4* in mammals), which launch transcriptional programs to counter this stress. This pathway involves phosphorylation of the eIF2 translation factor by Gcn2-protein kinases, and is regulated by uORFs in the *GCN4*/*ATF4* 5'-leaders. Here we present evidence that the transcription factors that mediate this response are not evolutionary conserved. Although cells of the fission yeast *Schizosaccharomyces pombe* respond transcriptionally to amino acid starvation, they lack clear Gcn4 and Atf4 orthologs. We used ribosome profiling to identify mediators of this response in *S. pombe*, looking for transcription factor that behave like *GCN4*. We discovered a novel transcription factor (Fil1) translationally induced by amino acid starvation in a 5'-leader and Gcn2-dependent manner. Like Gcn4, Fil1 is required for the transcriptional response to amino acid starvation, and Gcn4 and Fil1 regulate similar genes. Despite their similarities in regulation, function, and targets, Fil1 and Gcn4 belong to different transcription factors families (GATA and b-ZIP, respectively). Thus, the same functions are performed by non-orthologous proteins under similar regulation. These results highlight the plasticity of transcriptional networks, which maintain conserved principles with non-conserved regulators.

Statement of significance

Eukaryotic cells respond to stress conditions by down-regulating general translation while selectively activating translation of genes required to cope with the stress (often encoding bZIP-family transcription factors, such as Gcn4 in *Saccharomyces cerevisiae* and Atf4 in mammals). Although the signal transduction pathways that mediate these responses are highly conserved, we report that the downstream transcriptional regulators are not: In the fission yeast *Schizosaccharomyces pombe*, this response is mediated by a GATA-type transcription factor (Fil1). Surprisingly, although Fil1 lacks any sequence homology to Atf4 and Gcn4, it regulates similar genes and is itself regulated in a similar manner. These results suggest that extensive rewiring has taken place during the evolution of this key response, and highlights the plasticity of transcriptional networks.

\body

Introduction

Cells respond to conditions of stress by implementing complex gene expression programs, both at the transcriptional and posttranscriptional levels (1). One of the best studied examples is the response to amino acid starvation. Amino acid starvation causes an accumulation of uncharged tRNAs, which activate a signalling pathway that leads to a reduction of the translation of the majority of cellular mRNAs. Simultaneously, the translation of specific mRNAs, some of them encoding key transcription factors, is induced. These transcription factors, in turn, launch a transcriptional program that promotes cellular survival under stress. This program is called the General Amino Acid Control (GAAC) in yeast (2), and the Amino Acid Response (AAR) in mammals (3).

This translational response to amino acid depletion is mediated by proteins of the Gcn2 protein kinase family, which are conserved throughout eukaryotes (4). Gcn2 is activated by deacylated tRNAs and phosphorylates the translation initiation factor eIF2, which is required to deliver the initiator tRNA to the ribosome, in its α subunit. eIF2 is a GTP-binding protein whose activation requires the activity of the GTP/GDP-exchange factor eIF2B, which facilitates the exchange of GDP with GTP. Phosphorylated eIF2 binds to eIF2B with high affinity, behaving as a competitive inhibitor. As eIF2 is more abundant than eIF2B, this binding leads to the rapid sequestration of all cellular eIF2B and thus triggers a global down-regulation of translation (5). This phosphorylation event occurs at a highly conserved serine residue.

In the budding yeast *Saccharomyces cerevisiae*, the key effector of the response to amino acid depletion is the Gcn4 transcription factor, which belongs to the basic Leucine ZIPper (bZIP) family. The translation of the *GCN4* mRNA is up-regulated upon amino acid starvation, in a process that is mediated by four upstream Open Reading Frames (uORFs) located at the 5'-leader sequence (2, 6). Ribosomes bind to the *GCN4* mRNA close to the 5' cap and scan the mRNA until they reach the AUG

of uORF1. The majority of ribosomes translate this uORF, which is permissive for reinitiation (uORF2 has similar properties to uORF1, and may function as a fail-safe mechanism). Thus, small subunits continue scanning the 5'-leader sequence until they translate uORF3 or uORF4, or until they reach the *GCN4* coding sequence. As uORF3/uORF4 are typically repressive for reinitiation, translation of uORF3/uORF4 and *GCN4* is mutually exclusive. Therefore, the decision by the reinitiating small subunit of whether to translate uORF3/4 determines the outcome of *GCN4* translation. Translation initiation requires the binding of the so called ternary complex (composed of the initiation factor eIF2, the initiator tRNA and a GTP molecule) to the small ribosomal subunit. Reinitiating small subunits lack a ternary complex, and must acquire it to recognise and translate a coding sequence. In the absence of stress, ternary complexes are abundant and the reinitiating subunits downstream of AUG1 can bind to one before they reach uORF3/4, thus allowing their translation and inhibiting *GCN4* translation. Under stress conditions, Gcn2 activation causes a reduction in the abundance of active ternary complexes, allowing the skipping of uORF3/4 and the translation of *GCN4* (2, 6). Thus, ribosome scanning through multiple inhibitory uORFs is required for preferential translation of *GCN4*. In mammals, AAR is mediated by Atf4, another b-ZIP transcription factor, whose translation is regulated similarly to that of *GCN4*. *ATF4* mRNA contains two uORFs: uORF1 is permissive for reinitiation, but uORF2 overlaps with the *ATF4* coding sequence and thus does not allow reinitiation. As in the case of *GCN4*, the abundance of ternary complex determines the translation of uORF2 or *ATF4* (6). In *Candida albicans*, a single uORF is necessary and sufficient to regulate translation of the *GCN4* homologue in response to amino acid deprivation (7).

All the known transcription factors that mediate the GAAC/AAR program belong to the family of b-ZIP transcription factors, although Gcn4 and Atf4 are not direct orthologs (8). Atf4 is well conserved among vertebrates, while *Drosophila* has a more distant homologue. Similarly, Gcn4 is conserved in related species of budding yeast such as *Candida* (7) and in the filamentous fungus *Aspergillus* (9). However other ascomycetes such as the fission yeasts *Schizosaccharomyces* (10) and many other fungi lack Gcn4-like transcription factors, and plants lack both Atf4 and Gcn4 orthologs

(11). Overall, the principle of translational control of b-ZIP transcription factor genes by Gcn2 in a uORF-regulated manner is widespread, and has been observed in mammals (8, 12), flies (13), and in several fungi (*Saccharomyces* (2), *Candida* (7) and *Aspergillus* (9)).

The fission yeast *Schizosaccharomyces pombe* displays a robust transcriptional response to amino acid depletion, which is dependent on the Gcn2-eIF2α signalling pathway (14). This program results in the elevated expression of multiple genes involved in amino acid biosynthesis, even though *S. pombe* lacks clear orthologs of Gcn4 or Atf4. Indeed, the transcription factor that implements this transcriptional program has not been identified (14).

Ribosome profiling (ribo-seq) provides a genome-wide, high-resolution view of translation (15). The approach is based on the treatment of translating ribosome-mRNA complexes with a ribonuclease (RNase), in such a way that only RNA fragments protected by bound ribosome survive the treatment. These fragments are then isolated and analysed by high-throughput sequencing. The number of sequence reads that map to a coding sequence, normalised by mRNA levels, provides an estimate of the efficiency of translation for every cellular mRNA (15).

Here we use ribosome profiling and parallel mRNA-sequencing to investigate the translational and transcriptional response of *S. pombe* cells to amino starvation. Genome-wide analyses identify a Gcn4 functional homologue: a novel transcription factor, Fil1, is essential for the transcriptional response to amino acid starvation and for normal growth in minimal medium lacking amino acids. We find that Fil1 binds to genes involved in amino acid biosynthetic pathways and causes their up-regulation. Importantly, a significant number of Fil1 targets are shared with Gcn4. In addition, *fil1* expression is controlled at the translational but not the transcriptional level, and reporter analyses suggest that this regulation is mediated by a 5'-leader sequence that contains multiple uORFs. Despite the conservation between Gcn4 and Fil1 at the target, functional and regulatory levels, the two proteins show no sequence homology, as Fil1 belongs to a different family of transcription factors

(GATA). These results provide a striking example of the plasticity of transcriptional gene expression programs, where function and regulation are maintained using non-conserved transcription factors.

Results

Modulation of the transcriptome in response to amino acid starvation

We used RNA-seq to investigate the response of *S. pombe* to amino acid starvation at the transcriptome level. *S. pombe* cells were treated with the histidine analogue 3-amino-1,2,4-triazol (3-AT), which inhibits histidine biosynthesis and thus mimics amino acid starvation. 573 genes were significantly up-regulated and 356 down-regulated (with conservative adjusted p-values < 10^{-3} and minimal changes of 1.5-fold, Figure 1A and Dataset S1). In addition, we identified 175 long non-coding RNAs significantly induced by 3-AT treatment (Dataset S1).

Both induced and repressed coding genes overlapped significantly with a previous microarray-based dataset of this process(14) (Figure S1), although the higher sensitivity of RNA-seq and the use of a large number of samples (7 biological replicates) allowed us to obtain a more complete view of this response. Up-regulated genes were enriched in GO categories related to amino acid biosynthesis (46 genes, GO:0008652, p-value = $2x10^{-8}$) and autophagy (20 genes, GO:0006914, p-value = $3x10^{-9}$,Figure S2A). Moreover, genes induced in response to all stress conditions (the so called core environmental stress response, or CESR (16)) were overrepresented and made over 50% of the induced genes (p-value=$3x10^{-121}$, Figure S2A). We divided the genes induced by 3-AT into those that were also part of the CESR and those that were not. The latter group was still enriched in genes related to amino acid biosynthesis and autophagy, while the CESR genes did not show this over-representation. This analysis is consistent with previous observations that the response to amino acid depletion involves a core response (CESR) together with a stress-specific program (equivalent to the GAAC) (14). Down-regulated genes were enriched in GO categories related to ribosome biogenesis (p-value = $4x10^{-24}$), cytoplasmic translation (p-value = $2x10^{-21}$, including genes encoding translation factors and ribosomal proteins), and ribonucleoside and glucose metabolism. About 45% of these genes are repressed in all stress situations, as part of the CESR.

The up-regulated group included several genes encoding transcription factors (the b-ZIP *atf1*, *pcr1* and *atf21*, the sporulation-induced *rsv2*, and two uncharacterised genes), suggesting that the gene expression program may involve a cascade of transcription factors that is activated by an unknown master regulator. All of these, with the exception of *atf21*, are also part of the CESR.

3-AT treatment leads to a robust phosphorylation of eIF2α (17, 18). Although *S. pombe* possesses three eIF2α kinases (Hri1, Hri2 and Gcn2) (10, 18, 19), 3-AT-dependent phosphorylation of eIF2α requires Gcn2 under the conditions used in this study (17, 18). Thus, we focused on the role of Gcn2. In the absence of stress, *gcn2Δ* cells showed only minor differences with respect to wild type (15 up- and 12 down-regulated genes, ~ 0.6% of all genes, Figure 1B, Figure S3A and Dataset S1). Upon 3-AT treatment of *gcn2Δ*, 112 genes were induced and 40 repressed (Figure 1C, Figure S2B and Dataset S1). The former group included some genes that are usually repressed as part of the CESR (namely genes required for ribosome biogenesis, but not genes encoding ribosomal proteins, GO: 0042254, p-value=$7x10^{-28}$). It also comprised numerous genes encoding heat-shock proteins as well as genes induced in response to oxidative stress (p-value = $9x10^{-15}$), as previously identified in a microarray study (20). A small number of CESR-induced genes were also induced by 3-AT treatment (although only ~10% of those induced in wild type cells, Figure S2B). By contrast, the expression of most genes induced by amino acid starvation in wild type cells was not up-regulated, confirming that Gcn2 is the major mediator of this response (Figures 1C-1D, Figure S2B and Dataset S1).

The translational response to amino acid starvation

To investigate the translational response that accompanies amino acid starvation, we applied ribosome profiling to cells treated with 3-AT as described above. Briefly, cell extracts were incubated with RNase I, and ribosome protected fragments (RPFs or ribosome footprints) were isolated and analysed by high-throughput sequencing. For every sample, mRNA levels were measured in parallel by RNA-seq. To obtain gene-specific estimates of translational efficiency we quantified the number

of RPFs that mapped to each annotated coding gene, and normalised this figure by the corresponding number of RNA-seq reads.

As 3-AT causes a general down-regulation of translation, our translation efficiency measurements (TEs) are likely to be overestimates. However, the data reflect relative changes between conditions and identify genes that behave differently from the majority of transcripts. Relative changes in TE were much less widespread than those in the transcriptome. Upon 3-AT treatment, only 19 genes were consistently up-regulated at the translational level (at least 1.5-fold in seven independent replicates), while 11 were down-regulated (Figure 2A and Dataset S1). None of the two groups showed any specific enrichments in GO or expression categories, although they included potential regulators of gene expression (see below). This response appears weaker than that of *S. cerevisiae*, where 251 genes were either induced or repressed at least 2-fold at the TE level (15). This difference might be partly technical, due to our stringent selection conditions over seven biological replicates, or reflect biological features of both systems. We then examined whether this translational response was dependent on Gcn2 presence. Ribosome profiling and matching RNA-seq in *gcn2Δ* cells treated or untreated with 3-AT revealed that the majority of the translationally-induced genes did not respond to amino acid starvation (Figure 2B). Indeed, the translational efficiency of several of them was even reduced in the mutant upon 3-AT treatment (Figure 2B).

As translation is a highly dynamic process, many early studies (including our initial work) used an incubation with inhibitors of translation elongation, namely cycloheximide (CHX), to 'freeze' ribosomes on their *in vivo* localization during cell collection. While we performed these experiments, several studies reported that the results with ribosome profiling experiments may be affected by pre-treatment with CHX, especially under conditions of stress (21-24). To investigate if our experiments were affected by the use of CHX in media, we performed two additional ribosome profiling experiments in cells treated and untreated with CHX (both in the absence and the presence of 3-AT). Consistent of a recent study of the response to nitrogen starvation, we observed an increase in

ribosome density of ribosomal protein mRNAs upon CHX treatment, but no changes for most other mRNAs (Figures 2C and 2D, yellow dots) (21).

Codon-specific ribosomal occupancies and histidine starvation

Ribosome profiling allows the determination of codon-specific ribosomal occupancies, by measuring the fraction of ribosomes that are translating each of the 61 amino acid coding codons. If these data are normalized by the abundance of each codon in the genome, they can be used to generate a 'relative codon enrichment', which is related to the average time spent by the ribosome on each codon (Figure S4). We used this property to validate the quality of the ribosome profiling dataset: as 3-AT treatment interferes with histidine synthesis, it is expected to cause a decrease in the levels of activated histidine-tRNA, and thus to raise the time that the ribosome spends decoding histidine codons. Consistently, relative ribosomal occupancies of both histidine codons were strongly increased in 3-AT treated cells, while those of the other 59 codons were unaffected (Figure S4).

The Fil1 transcription factor is a master regulator of amino acid biosynthesis genes

The translation of *S. cerevisiae GCN4*, but not its transcription, is induced upon 3-AT treatment (15). Thus, we reasoned that this might be a property of other master regulators of the amino acid starvation response. To explore this idea we mined our dataset for genes whose TE was induced by histidine starvation, while maintaining constant levels of mRNA. Only seven genes fulfilled these criteria (Dataset S1), including one encoding a predicted transcription factor (*SPCC1393.08*). *SPCC1393.08* TE was reproducibly induced upon 3-AT treatment, with an average increase of 3.8-fold over 7 biological replicates.

SPCC1393.08 gene (hereafter referred to as *fil1*, for *gcn **f**our* induction like) encodes an uncharacterised protein of 557 amino acids containing two tandem GATA-type zinc fingers in its C-

terminus. The N-terminal part is conserved in *Schizosaccharomyces octosporus* and *cryophilus*, but not in *japonicus*. By contrast, the zinc fingers are well conserved in some fungi (such as *Pneumocystis*) and in animals (especially vertebrates). Note that all *S. cerevisiae* members of the GATA family have a single zinc finger domain. Importantly, we confirmed that 3-AT-dependent increase in translation efficiency of *fil1* (and the lack of changes in mRNA levels) also occurred in the absence of CHX incubation (Figures 2C-2D, arrows).

fil1Δ cells were viable and behaved as wild type in rich medium, but grew very slowly in minimal medium (Figure 3A). This phenotype is very similar to that of *S. cerevisiae GCN4* mutants (25, 26). Addition of all 20 amino acids to the medium improved *fil1Δ* growth, but not enough to reach wild type rates (note wild type cells also grew faster in the supplemented medium, Figure 3A). Expression of a *fil1* transgene containing the *fil1* endogenous promoter completely rescued the phenotype of *fil1Δ* cells, confirming that the effects were not due to secondary mutations (Figure 3A). We tested whether *fil1Δ* cells are hypersensitive to 3-AT, but were unable to obtain a clear answer due to their very poor growth in minimal medium.

We then used RNA-seq to investigate the effects of *fil1* deletion on gene expression. A total of 165 genes were expressed at low levels in *fil1Δ* cells compared to wild type. These genes showed a very strong enrichment in genes involved in amino acid biosynthesis (GO: 0006520, p-value = 10^{-39}), but not in autophagy (GO: 0006914, p-value = 0.78, Figure S3B). This is different from mammalian Atf4, which directly regulates the expression of a number of autophagy genes (27). Fil1-dependent genes also displayed a small but significant overlap with genes that are induced as part of the CESR (p-value = $3x10^{-6}$, Figure S3B). By contrast, only 39 genes were up-regulated in the mutant, which were enriched in genes induced by nitrogen starvation (including those involved in pheromone responses, p-value = $2x10^{-15}$). Fil1-dependent genes showed a large overlap with those genes induced by 3-AT treatment (Figure 3B). Together with the phenotypic data, these results suggest that Fil1 activates

amino acid biosynthesis (and probably that of other metabolites) in unstressed cells, and that many of Fil1 targets are further induced in response to amino acid starvation (Figure 3B).

We examined whether Fil1 has a role in the response to 3-AT by performing RNA-seq of *fil1Δ* cells after 3-AT exposure. Strikingly, only 10 genes showed significant changes in the *fil1Δ* mutant (Figure 3C and Figure S2C). These genes did not include amino acid metabolism genes, autophagy genes, or those encoding the transcription factors *atf1*, *atf21*, *pcr1*, and *rsv2*. This lack of response was not due to 3-AT-dependent genes being constitutively expressed in the *fil1Δ* mutant, as these genes were not expressed at high levels in untreated mutant cells (Figure 3D). Thus, these data indicate that *fil1Δ* cells are unable to respond transcriptionally to 3-AT.

These data predict that Fil1 overexpression should mimic the response to amino acid starvation. To test this idea we expressed *fil1* under the control of the regulatable promoter *nmt1* (note that this construct did not contain the endogenous 5'-leader sequences of *fil1*). Upon promoter derepression, 3-AT responsive genes were strongly upregulated, consistent with the idea that elevated Fil1 expression is responsible for the transcriptional response to amino acid depletion (Figure 3E).

We then investigated whether Fil1 regulates gene expression directly. Cells expressing Fil1-TAP from its endogenous locus were used for ChIP-seq experiments, both in the absence and the presence of 3-AT treatment. Figures 4A-4B and Figure S5A-S5D show examples of three direct targets of Fil1, which have high levels of Fil1 on their promoters, are induced by 3-AT treatment (albeit to different extents), and require Fil1 for normal levels of expression. In untreated cells, we detected 352 peaks of Fil1 enrichment, which were associated with 240 promoters. About 30% of genes under-expressed in *fil1Δ* were bound by Fil1 (Figure 4C), and were strongly enriched in GO categories related to amino acid biosynthesis. The lack of binding to the remaining genes may be due to ChIP-seq being less sensitive than RNA-seq, or to indirect effects of the *fil1* deletion. By contrast, genes bound by Fil1 but whose expression was not reduced by *fil1* deletion did not show any enrichment in genes related

to amino acid production pathways. Finally, most genes overexpressed in the *fil1Δ* mutant were not bound by Fil1 (Figure 4D), indicating that Fil1 does not function as a repressor. We also identified an enriched sequence in Fil1-bound peaks (Figure S6). This motif is related to the 'GATA' sequence, and may represent the Fil1-binding site. Importantly, the motif is not similar to the Gcn4 consensus binding site (TGACTC) (28). Together, these results are consistent with the poor growth in minimal medium phenotype of *fil1Δ* mutants, and suggests that Fil1 regulates amino acid biosynthesis even in unstressed cells by directly stimulating the transcription of key genes.

Fil1 was present in 232 peaks after 3-AT treatment, which were linked to 170 promoter regions (Figures 4E and 4F). Genes bound by Fil1 and those induced by 3-AT overlapped significantly, but those repressed did not. This is consistent with Fil1 functioning as an activator of transcription. Genes bound by Fil1 before and after 3-AT incubation also overlapped extensively (Figure 4G). However, the GATA-related sequence was identified with much lower significance and other unrelated sequences were more strongly enriched (Figure S6). Moreover, we were not able to detect a general rise in Fil1 binding upon 3-AT treatment. This lack of increased binding may simply reflect the lack of quantitative behaviour of ChIP-seq experiments. To address this issue we used quantitative PCR to measure the enrichment of Fil1 on promoters upon 3-AT treatment. For this purpose, we focused on three Fil1 target genes: one that shows very weak increases in mRNA levels (*leu2*, 1.3-fold, Figure S5D), and two that display larger increases (*arg1* and *asn1*, 1.9 and 2.6-fold increases, respectively, Figures 4B and S5B). Both *arg1* and *asn1* showed a clear rise in Fil1 binding after 3-AT exposure, which was not observed for *leu2* (Figure S5F). These results are consistent with the idea that elevated Fil1 levels, and thus increased Fil1 protein on promoters, activates transcription of Fil1 targets.

Overall, Fil1 bound to only ~10% of all genes induced by 3-AT (Figure 4E). Given that the whole transcriptional response is dependent on Fil1 (Figure 3C and 3D), this suggests that Fil1 may drive part of the response indirectly through the up-regulation of other transcription factors. Consistent with this idea, the *atf1* gene is both induced by 3-AT and bound by Fil1, and has been reported to be

sensitive to 3-AT (14). Moreover, other transcription factor genes that are induced by 3-AT (*rsv2*, *pcr1* and *atf21*) showed peaks of Fil1 binding, although they did not pass the significance threshold of the peak-finding algorithm. Taken together, our data suggest that Fil1 regulates cellular metabolism in unstressed cells by directly binding to and activating genes involved in amino acid biosynthesis. In addition, Fil1 drives the response to amino acid starvation both directly and through the action of downstream transcription factors. The transcriptional response to stress in mammalian cells involves a similar mechanism, where Atf4 directly activates the expression of the transcription factors c/EBPβ, Atf3, Atf5, and CHOP (3).

Finally, we investigated whether Fil1 and Gcn4 regulate orthologous genes. Comparisons of ChIP-seq results for both transcription factors (29) using tables of orthologs identified a highly significant (p-value = $9x10^{-11}$) 'core' set of 26 genes that are directly bound by Gcn4 and Fil1 (Figure 4H). This group was strongly enriched in genes with functions in amino acid biosynthesis. Thus, despite the complete lack of sequence similarity between Gcn4 and Fil1, both transcription factors bind to the promoters of orthologous genes.

Regulation of Fil1 expression

The 5'-leader sequence of *fil1* mRNA is very long (962 nucleotides), and contains five AUG-starting uORFs. Of these, four showed clear translation in ribosome profiling experiments (Figure 5A). In addition, a CUG-starting uORF is also translated (Figure 5A). This suggests that *fil1* translation might be regulated by translation of some of these uORFs. Consistent with this idea, ribosomal density in the coding sequence of *fil1* was strongly increased upon amino acid depletion, while it remained unchanged in the 5'-leader sequence (Figure 5B). This was a highly specific effect, as the relative alteration in occupancy of the *fil1* CDS compared to its 5'-leader sequence was the highest in the transcriptome. By contrast, 3-AT treatment in a *gcn2Δ* background resulted in a decrease in ribosomal density in the coding sequence, but not in the 5'-leader (Figure S7).

The ribosome profiling data indicates that *fil1* mRNA translation is enhanced in response to 3-AT, suggesting that protein levels are also increased. To confirm this prediction we followed Fil1 protein by Western blotting using the Fil1-TAP strain described above. As expected, Fil1 protein accumulated upon 3-AT treatment (Figure 5C). Moreover, addition of histidine together with 3-AT, which is expected to suppress the effect of the drug, prevented the increase of Fil1 protein (Figure 5C). Finally, 3-AT did not cause an increase in Fil1 protein in a *gcn2Δ* strain (Figure 5C). All together, these results confirm that elevated translation of the *fil1* mRNA results in increased Fil1 protein levels.

To investigate the nature of *fil1* translational induction, we built a reporter system containing a constitutive promoter (*adh1*), the 5'-leader sequence of *fil1*, one copy of the mCherry fluorescence protein gene (mCherry) and the 3' untranslated region (3'-UTR) of *adh1*. We generated a similar construct with the 5'-leader of the *adh1* gene to be used as a control (Figure 5D). We monitored mCherry protein accumulation by flow cytometry, and mCherry mRNA levels by quantitative PCR. In the absence of stress, the mRNA levels of both reporters were similar to each other, but protein accumulation was ~56 times lower in constructs containing the 5'-leader of *fil1*, indicating that this region exerts a very strong repressive effect on translation (Figure 5D). Importantly, the protein expression from the *adh1* 5'-leader reporter was unaffected by 3-AT treatment, while that of *fil1* displayed a reproducible increase (Figure 5E, left). As expected, this increase was suppressed by addition of histidine to the culture and was dependent on Gcn2 (Figure 5E, left). By contrast, mRNA reporter levels decreased slightly in all treated wild type cells (Figure 5E, right), and remained unchanged in wild type cells treated with histidine and in *gcn2Δ* mutants, possibly suggesting a stress-dependent inhibition of the *adh1* promoter. Overall, these results demonstrate that the *fil1* 5'-leader sequence is sufficient to confer 3-AT-responsive translation to a downstream coding sequence.

The above data predict that loss of translational control of *fil1 in vivo* should cause constitutive activation of the amino acid starvation transcriptional response. To explore this possibility, we constructed a strain in which the six initiation codons of the *fil1* mRNA were inactivated. Cells carrying

this mutation did not show strong gene expression changes in response to 3-AT treatment (1.1 median fold-induction, compared to 1.74 of wild type cells, Figure S8A-B). This lack of response was due to constitutive expression of 3-AT-dependent genes, as these genes were overexpressed in the mutant even in the absence of 3-AT treatment (1.73-fold induction, Figure S8C). Moreover, a reporter containing the mutated *fil1* 5'-leader was expressed at very high levels (20-fold higher than the wild type, Figure 5D) but insensitive to 3-AT treatment (Figure 5E). Taken together, our results suggest that uORF-mediated translational control of *fil1,* leading to increased Fil1 protein levels, is directly responsible for the activation of the transcriptional response to amino acid starvation.

Discussion / conclusions

We have systematically examined the response of *S. pombe* to amino acid starvation, both at the transcriptome and translational levels, and identified the key transcriptional effector of the program. We report that this response is mediated by the eIF2α kinase Gcn2, which is required for the uORF-mediated translational induction of the Fil1 transcription factor (Figure 2B). Fil1 directs the transcriptional program, probably by both directly activating transcriptional targets and through the transcriptional up-regulation of other transcription factors. Fil1 also has essential roles in unstressed cells to maintain normal levels of amino acid biosynthesis genes.

The role of Gcn2 in the regulation of the CESR is surprising, as a previous microarray study suggested that the 3-AT induction of the CESR was Gcn2-independent (14). One possibility is that the CESR is only induced directly with the higher concentrations of 3-AT used in the microarray study (30 mM), and that the lower concentrations employed in this work only activate the CESR through the Gcn2 – Fil1 pathway, which would activate the transcription of the genes encoding the *atf1* and *pcr1* transcription factors.

The biological function, targets and regulation of Fil1 suggest that it is a functional homologue of the *S. cerevisiae* Gcn4 transcription factor. Indeed, a highly significant group of orthologous genes are directly bound by both transcription factors (Figure 4H). Interestingly, different elements of the response to amino acid starvation display strikingly different levels of conservation (Figure S9). The eIF2α/Gcn2 signalling pathway is extremely conserved, and Gcn2 protein kinases phosphorylate eIF2α across eukaryotes (including fungi (4), mammals (4), plants (11), *Leishmania* (30) and the Apicomplexans *Plasmodium* (31) and *Toxoplasma* (32)). The next layer, the translational up-regulation of an mRNA encoding a transcription factor, employs a common general mechanism (uORFs that are differentially used during starvation), but the details are different (see Introduction). By contrast, our results demonstrate complete divergence in the nature of the transcription factors that directly activates the response. Given the lack of conservation of Gcn4 homologues (even within fungi) and of Atf4 (not conserved beyond metazoans), this may turn out to be a general phenomenon. Indeed, the downstream transcription factor has not been identified in many organisms that display Gcn2-mediated stress responses (such as many fungi, plants (11), *Leishmania* (30), *Plasmodium* (32) and *Toxoplasma* (31)). Our results, together with published ribosome profiling of *S. cerevisiae* (15) and mammals (33), establish ribosome profiling as a powerful approach to identify these key transcriptional regulators.

The interactions between transcription factors and their sets of targets genes (regulons) can be flexible across evolution, and large-scale rewiring may occur (34). For instance, the expression of genes encoding ribosomal proteins is controlled by different, unrelated regulators in *S. cerevisiae* and *Candida albicans* (Yap1 and Tbf1 / Cbf1, respectively) (35). Another example is the expression of sterol biosynthesis genes, which in most eukaryotes is performed by basic helix-loop-helix transcription activators, whereas in *S. cerevisiae* and *C. albicans* is regulated by a Gal4-type zinc finger protein (36). The behaviour of Gcn4 and Fil1, two unrelated transcription factors with highly similar regulons and biological functions, and under the control of an exceptionally conserved signal transduction pathway, is a striking example of the plasticity of transcriptional circuits.

Methods

Strains, growth conditions and experimental design

Standard methods and media were used for *S. pombe* (37). For all genome-wide experiments *S. pombe* cells were grown in Edinburgh Minimal Medium 2 (EMM2) without additional amino acids at 32 °C. Histidine starvation was induced by incubating cells with 3-amino-1,2,4-triazol (3-AT) at 10 mM for 60 minutes (genome-wide experiments), 5 hours (reporters) or as indicated in the figures (time courses). For measurements of growth rates (Figure 3A) *fil1Δ* and wild type cells were grown in yeast extract medium with supplements (YES), washed three times with water and resuspended in EMM2. When histidine was used as a control, cells were grown in EMM2 containing 75 mg/L histidine, and an extra 75 mg/L histidine was added together with the 3-AT. For *fil1* overexpression from the *nmt1* promoter, *nmt1-fil1* cells were grown in EMM2 containing 15 µM thiamine, washed three times with EMM2, resuspended in EMM2 and incubated for 18 hours at 32 °C.

Table S1 presents a full list of strains. All strains used were prototrophic. Deletions of *fil1* and *gcn2* were confirmed by diagnostic PCR and by examination of the RNA-seq data. A *fil1Δ* strain with a copy of *fil1* integrated at the *leu1* locus behaved as wild type, confirming that the deletion was the cause of the observed phenotypes (Figure 3A). The C-terminal tagged Fil1-tap strain containing the endogenous 3' UTR was constructed using CRISPR/Cas9 (38). The gRNA-encoding sequence AGAAATAGAGAATAAATTTT was cloned into the CRISPR-Cas9 plasmid pMZ374 (38) using Gibson assembly. A repair fragment was constructed containing the last 700 nucleotides of the *fil1* coding sequence, a copy of the TAP-tag and 430 nucleotides of *fil1* 3' UTR and cloned into pJET2.1 (Thermo Scientific) by Gibson assembly. The final construct was PCR-amplified with Phusion (Thermo). 10 µg of the PCR repair fragment and 1 µg of the CRISPR plasmid were transformed into a *ura4-D18* strain. *ura*$^{+}$ colonies were selected for and checked by colony PCR for the correct integration. The *ura4-D18* marker was removed by crossing.

Reporters were constructed by removing the *nmt1* promoter and His6-Flag-GFP tag from pDUAL-His6-Flag-GFP with SphI and NdeI and replacing them with the mCherry coding sequence. A PCR product containing the genomic *adh1* promoter and 5'-leader was inserted using Gibson Assembly. The *fil1* 5'-leader reporter was built by inserting a PCR product containing the *adh1* promoter and a PCR product with the *fil1* 5'-leader sequence in the vector above using Gibson assembly. The *fil1* transcriptional start side was identified from published CAGE mapping (39) and is located 3 nucleotides upstream of the annotated 5'-leader. Both reporters were integrated into the *leu1* locus.

The mutant 5'-leader was synthesised by GeneArt (Thermo Fisher Scientific) and contains mutations in six initiation codes (GUG1, and AUG1 to AUG5), which were mutated to CAA (GUG1), AAG (AUG1), and AAA (AUG2, AUG3 and AUG5). uORF4 contains two consecutive AUG codons, which were mutated to AAAAAA. . AUG1 was mutated to AAG (and not AAA, as the others) to avoid creating an AUG codon. This sequence was used to replace the wild type sequence using Gibson Assembly.

All repeats of genome-wide experiments are independent biological replicates carried out on separate days (see ArrayExpress submission for a complete list). The following experiments were performed: 1] Ribosome profiling and matching RNA-seq of wild type cells in plus/minus 3AT (3 repeats); 2] Ribosome profiling and matching RNA-seq of wild type cells in plus/minus 3AT with plus/minus cycloheximide treatment (2 repeats); 3] Ribosome profiling and matching RNA-seq of *gcn2Δ* in plus/minus 3-AT (2 repeats); 4] RNA-seq of *fil1Δ* and matching wild type control, plus/minus 3-AT (4 repeats); 5] Chip-seq with untagged strain, Fil1-TAP minus 3-AT, and Fil1-TAP plus 3-AT (2 repeats).

Protein analyses

For Fil1-TAP detection, cells were harvested by filtration, washed with water and frozen as a dry pellet. Cell pellets were resuspended in 100 µl 20% TCA, and lysed with 1 ml of acid treated glass beads in a bead beater (FastPrep-5, MP Biomedicals) at level 7.5 for 15 seconds. 150 µl of 10% TCA was added before eluting from the glass beads. Lysates were frozen on dry ice and spun at 18,000 RCF for 15 min at 4 °C. Pellets were washed 4 times with cold acetone, dried and resuspended in 2x Laemmli buffer (4% SDS, 20% glycerol, 120 mM Tris-Cl [pH 6.8], 0.02% [w/v] bromophenol blue, 1% βME), and boiled at 100 °C for 5 min. TAP tag was detected with PAP (peroxide-anti-peroxide, Sigma P1291) and histone H3 with the #9715 polyclonal from Cell Signalling Technology.

Quantitative PCR to measure RNA levels

RNA was extracted as described (40). 0.1 µg of total RNA was digested with RQ1 RNase-free DNase (Promega) and cDNA was generated GoScript™ Reverse Transcriptase mix with random primers (Promega) following the manufacturer's protocol. qPCR reactions were performed in triplicate using PowerUP Sybr Mix (Applied Biosystems) and primers within the mCherry coding sequence and control primers in the genomic *myo1* gene. Reactions were analysed on a Rotor-Gene Q (Qiagen).

Quantitative PCR analysis of chromatin immunoprecipitation

Three independent biological replicates were performed, and for each of them two technical repeats were carried out (independent qPCRs performed on different days from the original immunopurified DNA). Immunoprecipitated chromatin was subjected to qPCR analysis. Peak region enrichment the tested genes (*leu2*, *arg1*, *asn1*) was normalized to a control gene (*cdc2*) using the ΔCt method, with fold enrichments calculated as 2^- (Ct of peak region – Ct of control region) (41). Enrichments of technical replicates were then averaged. Enrichment values were calculated for both +/- 3-AT cells, and the ratio of +3-AT to -3-AT calculated for each sample. The normalised ratios for the three

biological replicates were plotted (Figure S5) and used for statistical analysis. Significance was determined using a one-sample, one-sided Student's t-test.

Library preparation and sequencing

For RPF analyses, preparation of cell extracts, RNase treatment, separation of samples by centrifugation through sucrose gradients, and isolation of protected RNA fragments were performed as previously described (40). For samples wt.noAT.ribo.2 and wt.AT.ribo.2 (see ArrayExpress submission), libraries were prepared using a polyadenylation protocol as described (42). For all RPF samples, gel-purified RNA fragments were treated with 10 units of T4 PNK (Thermo Fisher) in a low pH buffer (700 mM Tris pH 7, 50 mM DTT, 100 mM MgCl2) for 30 min at 37 °C. ATP and buffer A (Thermo Fisher) were then added for an additional 30 min incubation. RNA fragments were column-purified (PureLink RNA micro-columns, Life Technologies). 100 ng were used as input for the NEXTflex Small RNA Sequencing Kit v2 (Bioo Scientific), and libraries were generated following manufacturer's protocol. For mRNA analyses, total RNA was isolated as previously described (40). Total RNA was then depleted from rRNA using Ribo-Zero Gold rRNA Removal Kit Yeast (Illumina) with 4 µg as input. Finally, 30 ng of ribo-depleted RNA was used as starting material for the NEXTflex Rapid Directional qRNA-Seq Kit (Bioo Scientific). Libraries were sequenced in an Illumina HiSeq 2000 or NextSeq 500 as indicated (ArrayExpress submission). ChIP-seq experiments were performed exactly as described (43).

Data pre-processing and read alignment

For ribosome-profiling samples wt.noAT.ribo.2 and wt.AT.ribo.2 (see ArrayExpress submission), the structure of the reads is the following: RRRRRRRRRRRR (NNNN….NNNN) BBB, where R represents a random nucleotide, N denotes the sequence of the protected RNA fragment, and BBB is a multiplexing barcode. For all other ribosome profiling samples, the structure of the reads is as follows: RRRR

(NNNN…NNNN) RRRR-adaptor-, where R represents random nucleotides, N corresponds to the sequence of the RNA protected fragment, and the adaptor sequence is TGGAATTCTCGGGTGCCAAGG. In both cases, random nucleotides serve as unique molecular identifiers (UMIs) (40) that allow the removal of PCR duplicates and the generation of a non-redundant dataset. To prepare reads for mapping we first removed partial adaptor sequences from the 3' end of the read. Duplicate reads were then discarded, followed by removal of UMIs.

For all RNA-seq experiments, the structure of reads is the following: RRRRRRRRA (NNNN….NNNN), where R(8) corresponds to a UMI, A to an adenosine residue, and N to the sequence of the RNA fragment. Duplicated reads were discarded, the RRRRRRRRA sequence removed from the reads, and reads were reverse complemented before mapping.

Mapping was performed using TopHat2 version 2.1.1 and Bowtie2 version 2.2.8 (44, 45). For ribosome profiling experiments, processed reads were first mapped to the *S. pombe* rDNA genome using the following parameters: --read-mismatches 2 --no-coverage-search --min-intron-length 29 --max-intron-length 819 -z 0 -g 1. Unmapped reads were then aligned to the full *S. pombe* genome with the same settings and with a gff3 file (Schizosaccharomyces_pombe_ASM294v2.28.gff3, downloaded from Ensembl) as a source of information on exon-intron junctions. For RNA-seq data, reads were directly mapped to the *S. pombe* genome using the parameters detailed above.

For ChIP-seq experiments, reads were aligned to the *S. pombe* genome using Bowtie2 2.2.8 (45) with the following non-default parameter: -k 2. Reads that map to repetitive sequences were removed from the analysis.

Data analysis

Data quantification (number of reads per coding sequence) was carried out using in-house Perl scripts. All statistical analyses were performed using R.

Differential expression analysis was performed using the Bioconductor DESeq2 package (46). Raw counts were directly fed to the program, and no filtering was applied. Unless otherwise indicated, a threshold of 10^{-3} was chosen for the adjusted p-value and a cut-off of 1.5-fold minimal change for the change in RNA levels.

For codon usage analyses RPF reads were aligned to nucleotide 16 (corresponding to position 1 of the codon in the ribosome A site). Only codons after 90 were used. For each coding sequence the following calculations were performed: 1] determination of the fraction of RPFs that occupy each codon (RPFs in a given codon divided by total RPFs); 2] quantification of the relative abundance of each codon on the coding sequence (number of times each codon is present divided by total codon number); 3] definition of the normalized codon occupancy by dividing parameter 1 by parameter 2. The average codon enrichments (Figure S4) were then calculated with data from all coding sequences.

For the analysis of TE, we used two different methods. In the first one, we required a threshold of 1.5-fold increase or reduction over the median of all genes in all seven ribosome profiling experiments (five biological replicates performed in the presence of CHX and 2 in its absence). This analysis identified 19 up-regulated and 11 down-regulated genes. In a second, more stringent approach, a z-score higher than 2 was required in all seven experiments. This identified 8 up-regulated and 5 down-regulated genes. Both approaches produced very similar results, including the identification of *fil1* (see Dataset S1 for complete lists of genes). Only genes with at least 20 counts in 75% or more relevant samples (RPF and corresponding RNA-seq in all ribosome profiling experiments) were used for TE analysis (91.8% of all genes). The data plotted in Figures 5A, and S7 were obtained from the CHX-treated samples.

Gene set enrichment was performed with AnGeLi (47). Lists of orthologous genes between *S. cerevisiae* and *S. pombe* were generated using YeastMine (48), and are based on a manually curated set prepared by Pombase (49). The significance of the overlap between gene lists was calculated using Fisher's exact test. The list of Gcn4 direct targets was obtained from a ChIP-microarray study (29).

For ChIP-seq, peaks were called with GPS/GEM (50), using the untagged strain experiment as background and default parameters (which include a minimal fold-difference between IP and control of 3, and a q-value threshold of 0.01). Potential binding motifs were searched using GPS/GEM with default parameters and with motif sizes restricted to between 6 and 8 nucleotides (--kmin 6 --kmax 8). Peaks were assigned to the closest gene promoter(s) using an in-house Perl script. Note that peaks located between divergent genes could not be assigned unambiguously to either gene, and thus were allocated to both. For the comparison of Fil1 binding between control and 3-AT-treated cells we used Homer with a false discovery rate of 0.001 (51).

Data deposition

All sequencing data has been deposited in the ArrayExpress database (5252), with the following accession numbers: *fil1*Δ/wild type RNA-seq (E-MTAB-5601), ChIP-seq experiments (E-MTAB-5580), ribosome profiling and parallel RNA-seq (E-MTAB-5810) and *fil1* overexpression (E-MTAB-6226)

Acknowledgements

This work was supported by Biotechnology and Biological Sciences Research Council (BBSRC) grants to J. M. [BB/N007697/1] and [BB/M021483/1], and by a Wellcome Trust Senior Investigator Award to J.B. [095598/Z/11/Z]. We thank Sanjay Ghosh and George Wood for help with experiments, and Sanjay Ghosh for comments on the manuscript.

Figure legends

Figure 1. Transcriptomic response to amino acid starvation.

A. Scatter plot comparing mRNA levels of wild type cells before and after 3-AT treatment (RNA-seq). All cells were grown in EMM2 without amino acids and 3-AT was added as indicated. All data have been normalized to RPKMs (Reads Per Kilobase per Million mapped reads). The dashed lines correspond to 2-fold differences. The results of a single experiment are shown. Genes in green have been selected as significantly up-regulated by 3-AT over multiple independent biological replicates (see Methods for details). **B.** As in A, comparison between wild type cells and *gcn2Δ* mutants in the absence of 3-AT. **C.** As in A, comparison of *gcn2Δ* mutants before and after 3-AT treatment. **D.** As in A, comparison of wild type and *gcn2Δ* cells after exposure to 3-AT.

Figure 2. Translational responses to amino acid starvation.

Scatter plots comparing $\log_2$ changes in mRNA levels and translation efficiencies between 3-AT-treated and untreated cells. All cells were grown in EMM2 without amino acids, and 3-AT was added as indicated. The *fil1* gene is plotted in black and highlighted by the arrows. **A.** Wild type cells. Genes whose TE is reproducibly induced upon 3-AT treatment in wild type cells are shown in red (at least 1.5-fold induction in seven out of seven experiments, including both CHX-treated and untreated cells). Cells from this experiment were pre-treated with CHX. **B.** As in A, data for *gcn2Δ* mutants. **C.** Wild type cells incubated with CHX. Genes encoding ribosomal proteins are displayed in yellow. **D.** As in C, wild type cells not incubated with CHX prior to collection.

Figure 3. Characterisation of *fil1Δ* mutants.

A. Relative growth rates of wild type cells (WT), *fil1Δ*, and *fil1Δ* expressing the *fil1* gene from the *leu1* locus. Cells were grown in rich medium (Rich), minimal medium with the addition of 20 amino acids (EMM+aa), and standard minimal medium (EMM). Data are normalised to the rich medium samples. Each dot corresponds to an independent biological replicate (n=3 or 6 as seen in the figure) and horizontal lines indicate the mean. Significance was determined using two-sample two-sided Student's t-tests (** P<0.01). Only significant comparisons are shown. **B.** Venn diagram showing the overlap between genes induced by 3-AT treatment in wild type cells and those expressed at low levels in *fil1Δ* mutants (no 3-AT). All cells were grown in EMM2 without amino acids, and 3-AT was added as indicated. The p-value of the observed overlap is shown. **C.** Scatter plot comparing mRNA levels of *fil1Δ* cells before and after 3-AT treatment. All data have been normalized to RPKMs (Reads Per Kilobase per Million mapped reads). The dashed lines correspond to 2-fold differences. Genes in green are significantly up-regulated by 3-AT in wild type cells (see Methods). **D.** As in C, but comparison of wild type and *fil1Δ* cells in the absence of 3-AT treatment. **E.** As in C, cells expressing the *fil1* coding sequence from the *nmt1* promoter.

Figure 4. ChIP-seq analysis of Fil1 binding.

A. Enrichment of Fil1-TAP in the *asn1* locus. Cells were grown in EMM2 without amino acids, and 3-AT was added as indicated. The arrow corresponds to the *asn1* gene and the box to the coding sequence. Enrichment is shown for two independent biological replicates in the absence of 3-AT exposure (red and black lines). The x axis shows the chromosomal coordinates (chromosome 2). **B.** Changes in mRNA levels of the *asn1* gene upon 3-AT treatment of wild type cells (left), in *fil1* mutants compared to wild type cells without3-AT exposure (middle), or in *fil1* mutants compared to wild type cells after 3-AT-treatment (right). Data are from RNA-seq experiments. Each point correspond to an independent biological replicate and the horizontal lines show the mean (n=3). **C.** Venn diagram showing the overlap between genes bound by Fil1 (without 3-AT treatment) and those expressed at

low levels in *fil1Δ* mutants. The p-value was calculated as described in Methods. **D.** As in C, showing the overlap between genes bound by Fil1 (no 3-AT) and those expressed at increased levels in *fil1Δ* mutants. **E.** As in C, displaying the overlap between genes bound by Fil1 (upon 3-AT exposure) and those induced by 3-AT treatment of wild type cells. **F.** As in C, comparing genes bound by Fil1 (with 3-AT) and those repressed by 3-AT treatment of wild type cells. **G.** As in C, comparing genes bound by Fil1 in cells untreated or treated with 3-AT. **I.** As in C, comparing genes bound by *S. pombe* Fil1 and *S. cerevisiae* Gcn4.

Figure 5. Translational control of the *fil1* mRNA.

A. Structure of the *fil1* transcript and distribution of RPFs. Lines represent the location of the untranslated regions and the box that of the coding sequence (CDS). The position of 6 uORFs (five AUG and one CUG) are indicated. Cells were grown in EMM2 without amino acids and incubated with 3-AT for 1 hour. The panels underneath display the density of RPFs along the transcript for untreated (top) and 3-AT treated cells (bottom). **B.** Quantification of RPF read density for the *fil1* coding sequence (CDS) and 5'-leader sequences. Data are presented for control and 3-AT treated cells. Each dot corresponds to an independent biological replicate (n=4) and the horizontal lines indicate the mean. **C.** Western blots to measure Fil1-TAP protein levels. Cells were incubated in EMM2 without amino acids and containing 3-AT for the indicated times. One sample (His) was incubated with both histidine and 3-AT for 180 minutes (histidine is expected to prevent the effects of 3-AT). The experiment was performed with wild type cells (right) and *gcn2Δ* mutants (left). Histone H3 was used as a loading control. The bottom panel shows quantification of three independent biological replicates of the experiment above, with data normalised to the levels of untreated cells of the corresponding genotypes. **D.** Expression of a mCherry fluorescent reporter in wild type cells containing the 5'-leader of *adh1,* or *fil1* with the six mutated uORFs (*fil1-6x*). Data are presented for fluorescence (protein) or RNA levels, and normalised to the levels of the wild type *fil1* reporter. Each dot corresponds to an

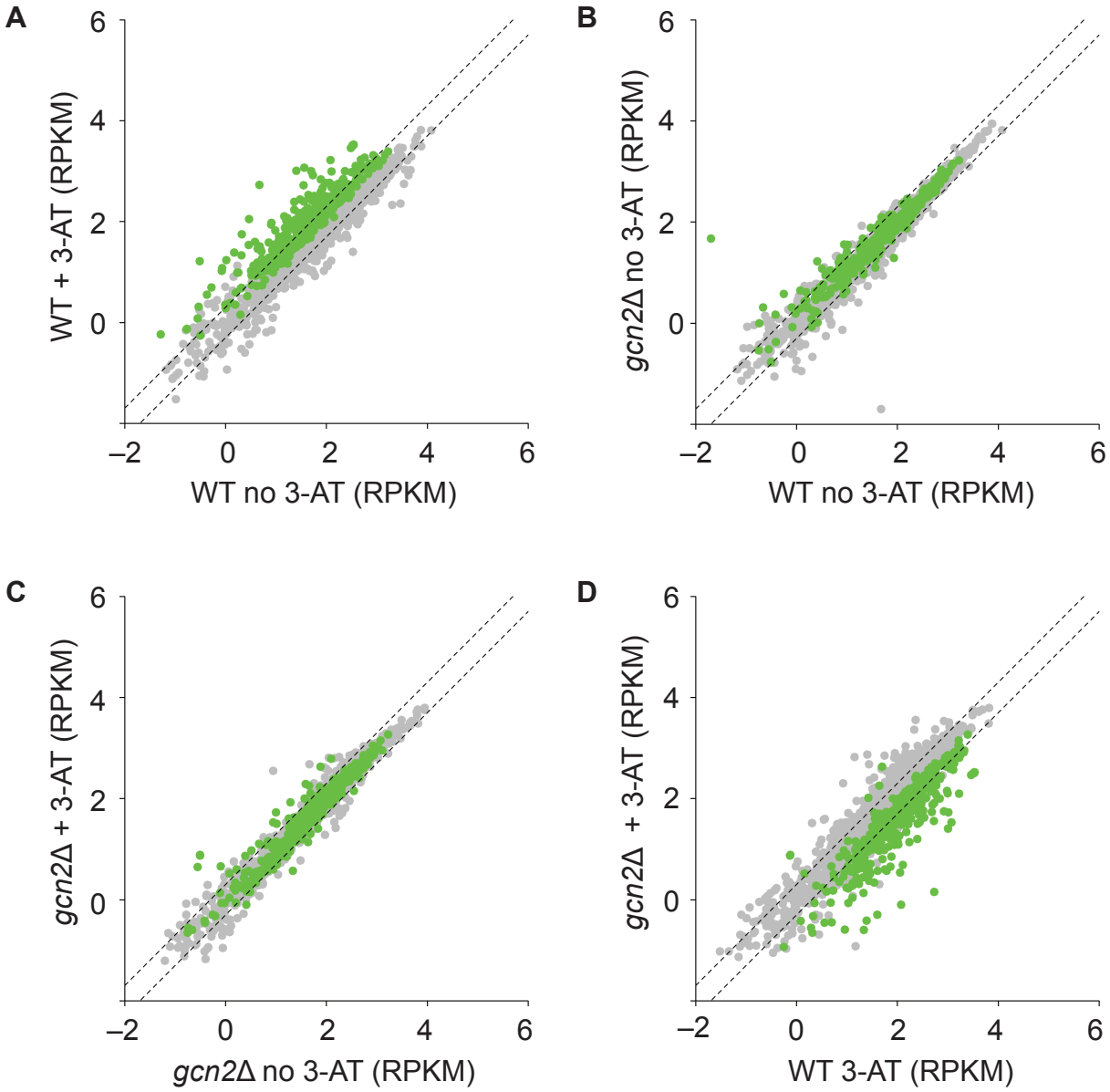
independent biological replicate (n=3) and horizontal lines indicate the mean **E.** Expression of an mCherry fluorescent reporter in wild type cells containing the 5'-leader sequences of *adh1*, wild type *fil1* (*fil1-wt*), or *fil1* with the 6 mutated uORFs (*fil1-6x*), or in wild type with *fil1* 5'-leader incubated in the presence of histidine (His), or with the *fil1* 5'-leader in *gcn2Δ* cells. Cell treatment as in C, except that the cells were incubated with 3-AT for 5 hours. All data are normalised to the levels of the corresponding reporter in untreated cells. Each dot corresponds to an independent biological replicate (n=3) and horizontal lines indicate the mean. Significance was determined using one-sample one-sided Student's t-tests. Only significant comparisons are shown (** p value<0.01). The data are shown for fluorescence levels estimated by flow cytometry (left) and for mRNA abundance quantified by qPCR (right).

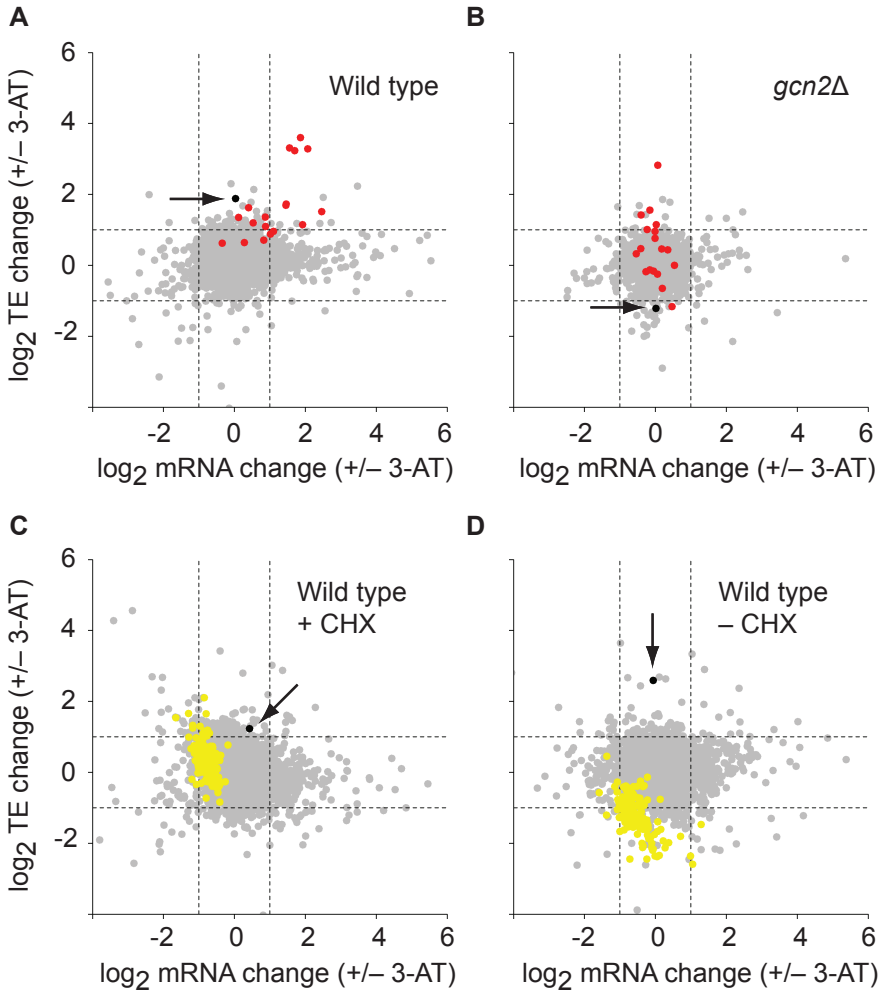
References

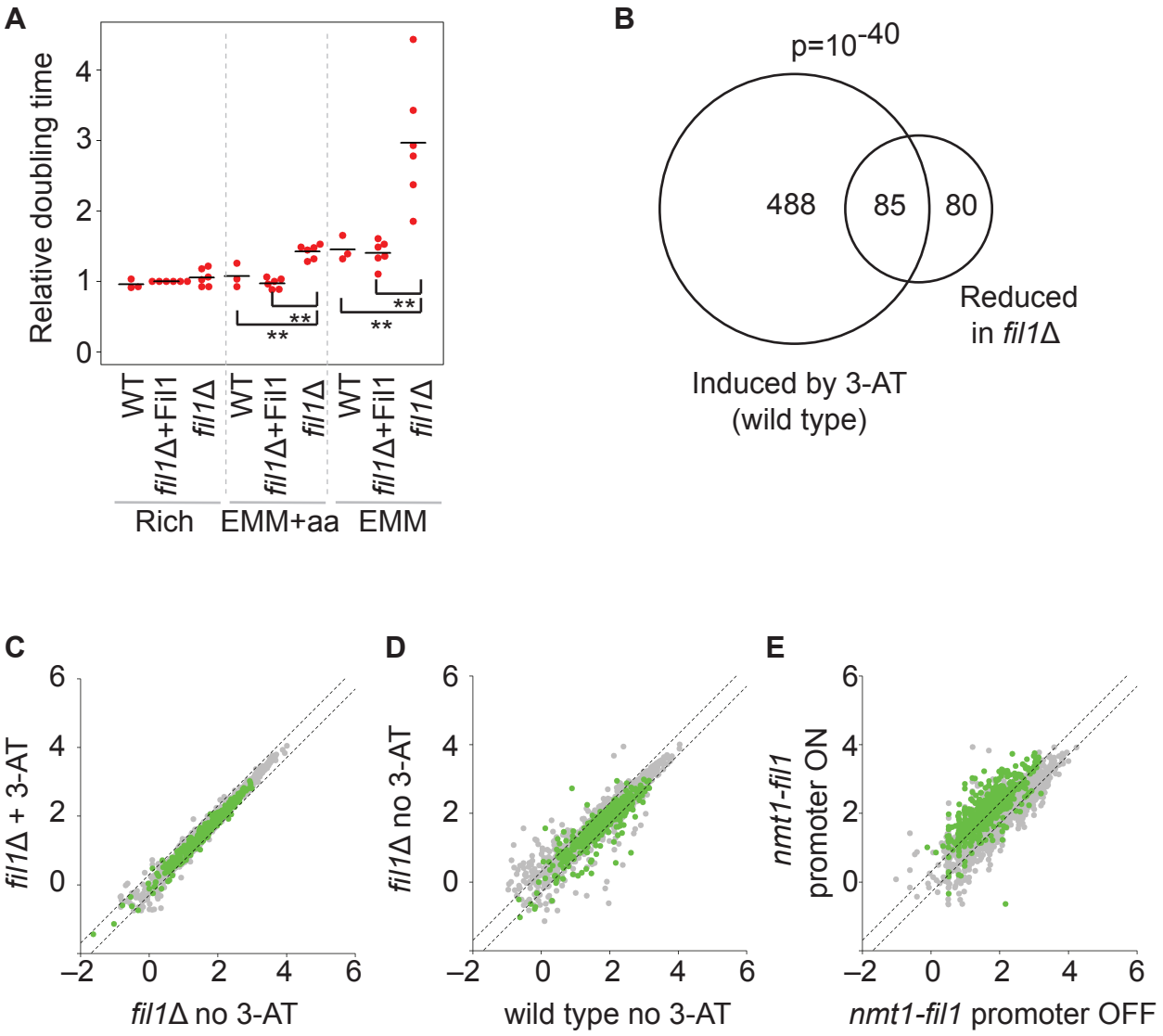
1. Spriggs KA, Bushell M, & Willis AE (2010) Translational regulation of gene expression during conditions of cell stress. *Mol Cell* 40(2):228-237.
2. Hinnebusch AG (2005) Translational regulation of *GCN4* and the general amino acid control of yeast. *Annu Rev Microbiol* 59:407-450.
3. Kilberg MS, Balasubramanian M, Fu L, & Shan J (2012) The transcription factor network associated with the amino acid response in mammalian cells. *Adv Nutr* 3(3):295-306.
4. Castilho BA*, et al.* (2014) Keeping the eIF2 alpha kinase Gcn2 in check. *Biochimica et biophysica acta* 1843(9):1948-1968.
5. Dever TE, Kinzy TG, & Pavitt GD (2016) Mechanism and Regulation of Protein Synthesis in *Saccharomyces cerevisiae*. *Genetics* 203(1):65-107.
6. Jackson RJ, Hellen CU, & Pestova TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11(2):113-127.
7. Sundaram A & Grant CM (2014) A single inhibitory upstream open reading frame (uORF) is sufficient to regulate *Candida albicans GCN4* translation in response to amino acid starvation conditions. *RNA* 20(4):559-567.
8. Vattem KM & Wek RC (2004) Reinitiation involving upstream ORFs regulates *ATF4* mRNA translation in mammalian cells. *PNAS* 101(31):11269-11274.
9. Hoffmann B, Valerius O, Andermann M, & Braus GH (2001) Transcriptional autoregulation and inhibition of mRNA translation of amino acid regulator gene *cpcA* of filamentous fungus *Aspergillus nidulans*. *Mol Biol Cell* 12(9):2846-2857.
10. Berlanga JJ*, et al.* (2010) Role of mitogen-activated protein kinase Sty1 in regulation of eukaryotic initiation factor 2alpha kinases in response to environmental stress in *Schizosaccharomyces pombe*. *Eukaryot Cell* 9(1):194-207.

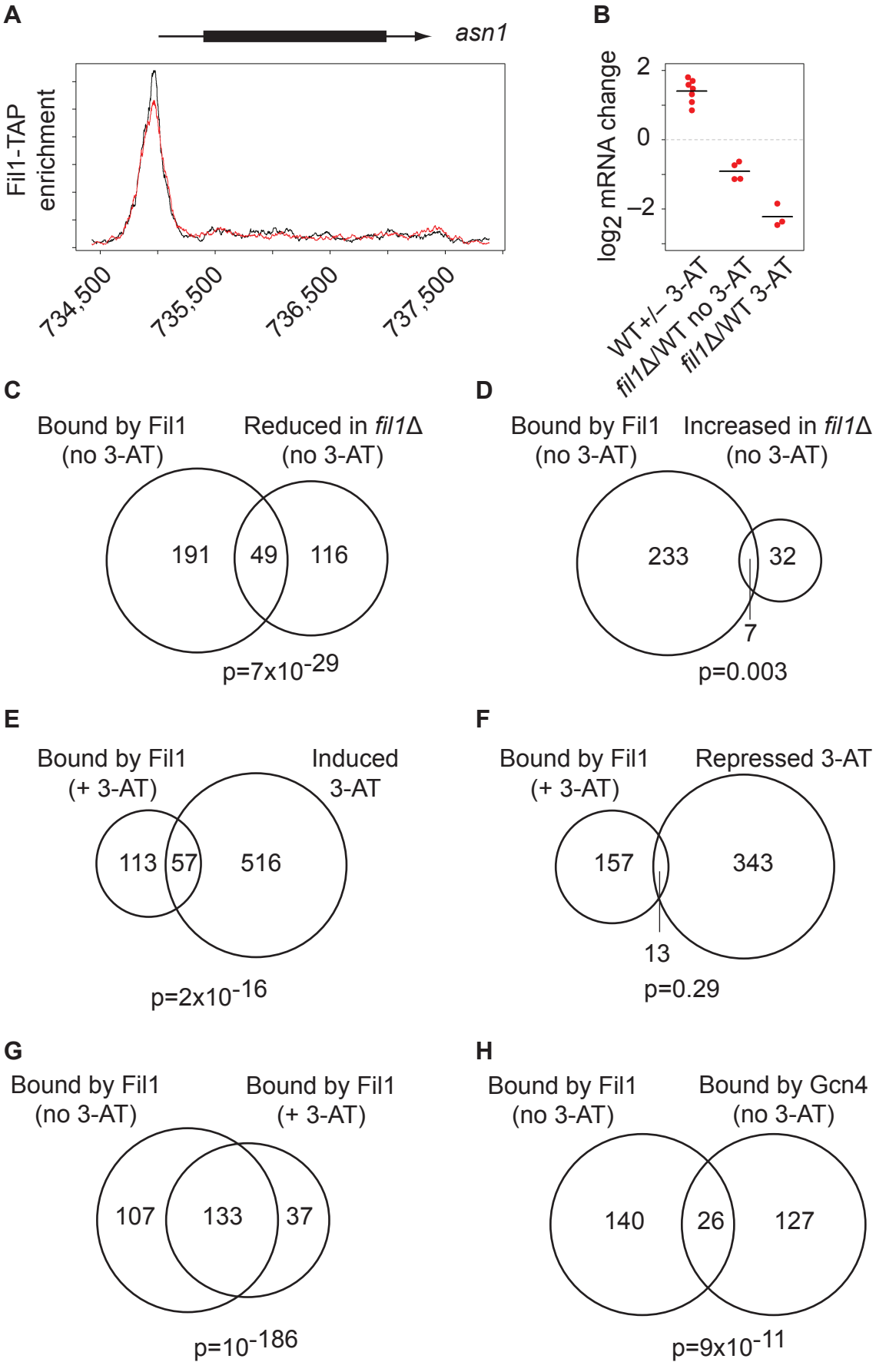
11. Li MW, AuYeung WK, & Lam HM (2013) The GCN2 homologue in *Arabidopsis thaliana* interacts with uncharged tRNA and uses *Arabidopsis* eIF2alpha molecules as direct substrates. *Plant Biol (Stuttg)* 15(1):13-18.
12. Harding HP*, et al.* (2000) Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol Cell* 6(5):1099-1108.
13. Kang K, Ryoo HD, Park JE, Yoon JH, & Kang MJ (2015) A *Drosophila* Reporter for the Translational Activation of *ATF4* Marks Stressed Cells during Development. *PLoS ONE* 10(5):e0126795.
14. Udagawa T*, et al.* (2008) Int6/eIF3e promotes general translation and Atf1 abundance to modulate Sty1 MAPK-dependent stress response in fission yeast. *J Biol Chem* 283(32):22063-22075.
15. Ingolia NT, Ghaemmaghami S, Newman JR, & Weissman JS (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* 324(5924):218-223.
16. Chen D*, et al.* (2003) Global transcriptional responses of fission yeast to environmental stress. *Mol Biol Cell* 14(1):214-229.
17. Tarumoto Y, Kanoh J, & Ishikawa F (2013) Receptor for activated C-kinase (RACK1) homolog Cpc2 facilitates the general amino acid control response through Gcn2 kinase in fission yeast. *J Biol Chem* 288(26):19260-19268.
18. Zhan K, Narasimhan J, & Wek RC (2004) Differential activation of eIF2 kinases in response to cellular stresses in *Schizosaccharomyces pombe*. *Genetics* 168(4):1867-1875.
19. Martin R, Berlanga JJ, & de Haro C (2013) New roles of the fission yeast eIF2alpha kinases Hri1 and Gcn2 in response to nutritional stress. *J Cell Sci* 126(Pt 14):3010-3020.
20. Nemoto N*, et al.* (2010) The roles of stress-activated Sty1 and Gcn2 kinases and of the protooncoprotein homologue Int6/eIF3e in responses to endogenous oxidative stress during histidine starvation. *J Mol Biol* 404(2):183-201.
21. Duncan CDS & Mata J (2017) Effects of cycloheximide on the interpretation of ribosome profiling experiments in Schizosaccharomyces pombe. *Scientific reports* 7(1):10331.
22. Gerashchenko MV & Gladyshev VN (2014) Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res* 42(17):e134.
23. Hussmann JA, Patchett S, Johnson A, Sawyer S, & Press WH (2015) Understanding Biases in Ribosome Profiling Experiments Reveals Signatures of Translation Dynamics in Yeast. *PLoS Genet* 11(12):e1005732.
24. Lareau LF, Hite DH, Hogan GJ, & Brown PO (2014) Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *eLife* 3:e01257.
25. Hueso G*, et al.* (2012) A novel role for protein kinase Gcn2 in yeast tolerance to intracellular acid stress. *Biochem J* 441(1):255-264.
26. Penn MD, Galgoci B, & Greer H (1983) Identification of AAS genes and their regulatory role in general control of amino acid biosynthesis in yeast. *PNAS* 80(9):2704-2708.
27. B'Chir W*, et al.* (2013) The eIF2alpha/ATF4 pathway is essential for stress-induced autophagy gene expression. *Nucleic Acids Res* 41(16):7683-7699.
28. Arndt K & Fink GR (1986) GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5' TGACTC 3' sequences. *PNAS* 83(22):8516-8520.
29. Harbison CT*, et al.* (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature* 431(7004):99-104.
30. Rao SJ, Meleppattu S, & Pal JK (2016) A GCN2-Like eIF2alpha Kinase (LdeK1) of *Leishmania donovani* and Its Possible Role in Stress Response. *PLoS ONE* 11(6):e0156032.
31. Fennell C*, et al.* (2009) PfeIK1, a eukaryotic initiation factor 2alpha kinase of the human malaria parasite *Plasmodium falciparum*, regulates stress-response to amino-acid starvation. *Malaria journal* 8:99.

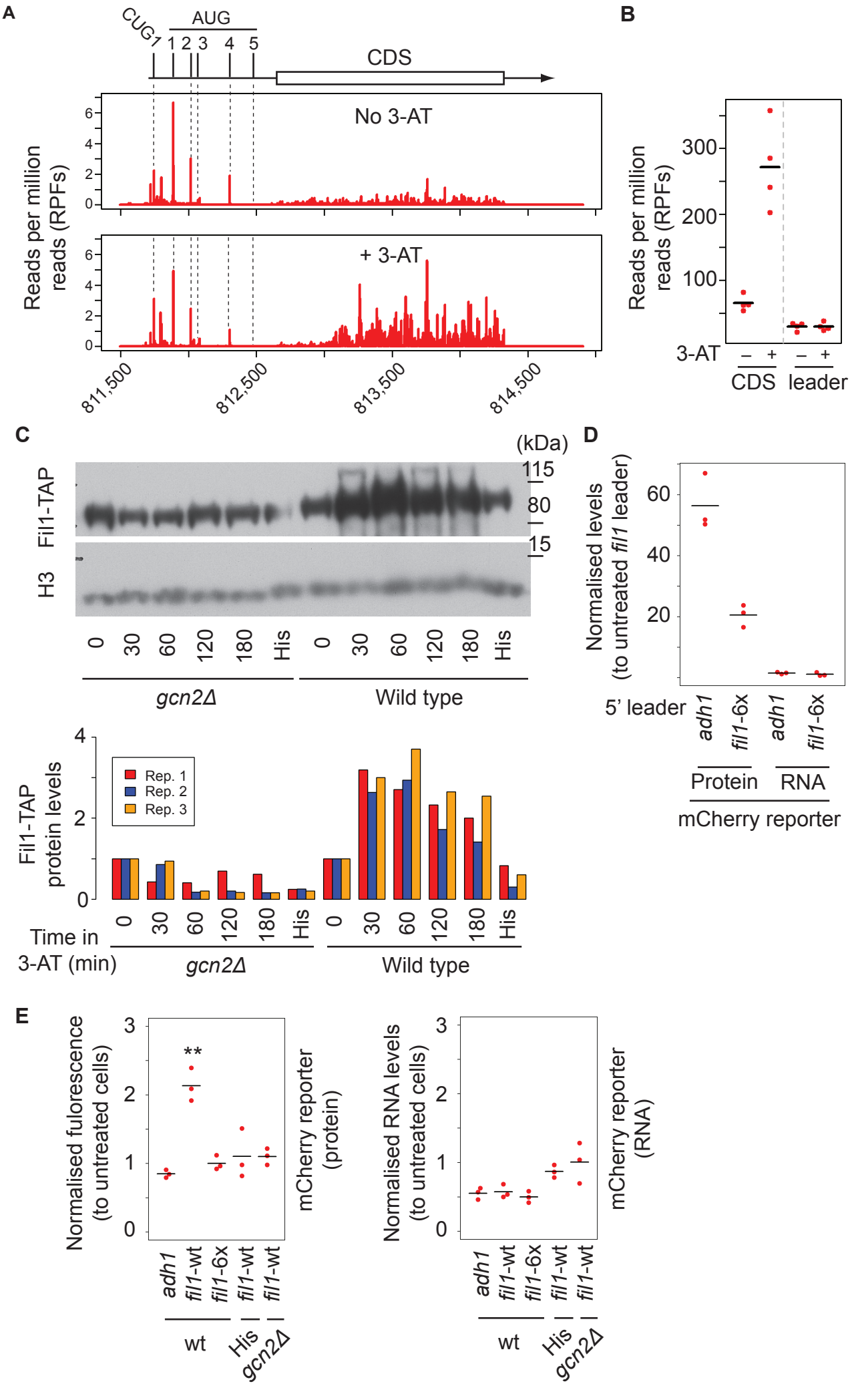
32. Konrad C, Wek RC, & Sullivan WJ, Jr. (2014) GCN2-like eIF2alpha kinase manages the amino acid starvation response in *Toxoplasma gondii*. *Int J Parasitol* 44(2):139-146.
33. Sidrauski C, McGeachy AM, Ingolia NT, & Walter P (2015) The small molecule ISRIB reverses the effects of eIF2alpha phosphorylation on translation and stress granule assembly. *eLife* 4(e05033).
34. Li H & Johnson AD (2010) Evolution of transcription networks--lessons from yeasts. *Curr Biol* 20(17):R746-753.
35. Hogues H*, et al.* (2008) Transcription factor substitution during the evolution of fungal ribosome regulation. *Mol Cell* 29(5):552-562.
36. Maguire SL*, et al.* (2014) Zinc finger transcription factors displaced SREBP proteins as the major Sterol regulators during *Saccharomycotina* evolution. *PLoS Genet* 10(1):e1004076.
37. Moreno S, Klar A, & Nurse P (1991) Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*. *Methods Enzymol* 194:795-823.
38. Jacobs JZ, Ciccaglione KM, Tournier V, & Zaratiegui M (2014) Implementation of the CRISPR-Cas9 system in fission yeast. *Nature communications* 5:5344.
39. Li H*, et al.* (2015) Genome-wide analysis of core promoter structures in Schizosaccharomyces pombe with DeepCAGE. *RNA Biol* 12(5):525-537.
40. Duncan CD & Mata J (2014) The translational landscape of fission-yeast meiosis and sporulation. *Nat Struct Mol Biol* 21(7):641-647.
41. Livak KJ & Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25(4):402-408.
42. Duncan C & Mata J (2017) Ribosome Profiling for the Analysis of Translation During Yeast Meiosis. *Methods Mol Biol* 1471:99-122.
43. Cotobal C*, et al.* (2015) Role of Ccr4-Not complex in heterochromatin formation at meiotic genes and subtelomeres in fission yeast. *Epigenetics & chromatin* 8:28.
44. Kim D*, et al.* (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14(4):R36.
45. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357-359.
46. Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.
47. Bitton DA*, et al.* (2015) AnGeLi: A Tool for the Analysis of Gene Lists from Fission Yeast. *Frontiers in genetics* 6:330.
48. Balakrishnan R*, et al.* (2012) YeastMine--an integrated data warehouse for *Saccharomyces cerevisiae* data as a multipurpose tool-kit. *Database : the journal of biological databases and curation* 2012:bar062.
49. Wood V*, et al.* (2012) PomBase: a comprehensive online resource for fission yeast. *Nucleic Acids Res* 40(Database issue):D695-699.
50. Guo Y, Mahony S, & Gifford DK (2012) High resolution genome wide binding event finding and motif discovery reveals transcription factor spatial binding constraints. *PLoS computational biology* 8(8):e1002638.
51. Heinz S*, et al.* (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38(4):576-589.
52. Rustici G*, et al.* (2013) ArrayExpress update--trends in database growth and links to data analysis tools. *Nucleic Acids Res* 41(Database issue):D987-990.

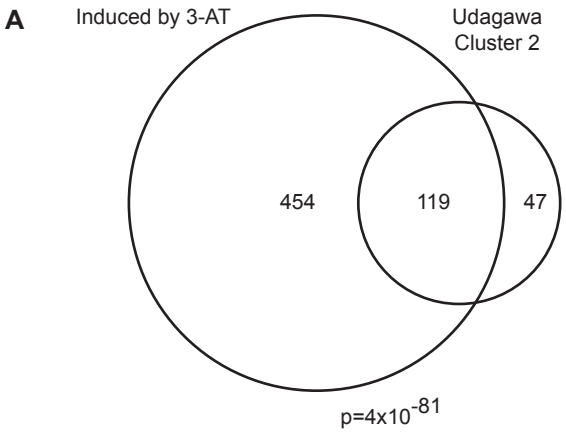

A
WT + 3-AT (RPKM)
WT no 3-AT (RPKM)
B
gcn2Δ no 3-AT (RPKM)
WT no 3-AT (RPKM)
C
gcn2Δ + 3-AT (RPKM)
gcn2Δ no 3-AT (RPKM)
D
gcn2Δ + 3-AT (RPKM)
WT 3-AT (RPKM)

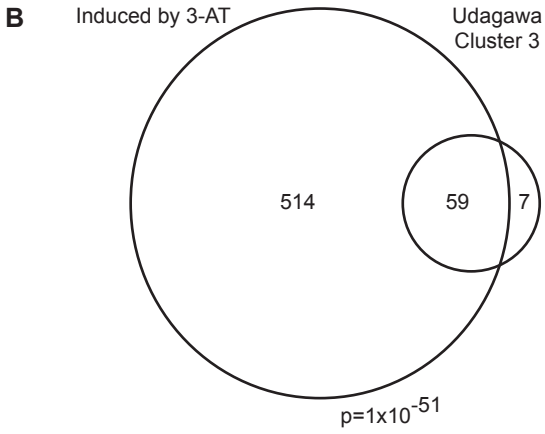

A
B
C
D
Wild type
gcn2Δ
Wild type
+ CHX
Wild type
– CHX
log2 TE change (+/– 3-AT)
log2 mRNA change (+/– 3-AT)
6
4
2
0
-2
-2
0
2
4
6

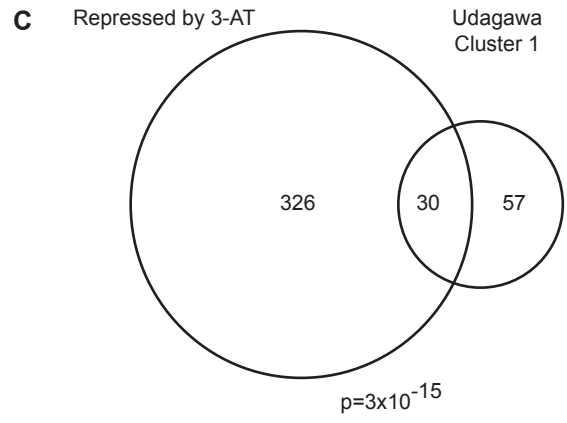

A
Relative doubling time
0
1
2
3
4
**
**
**
**
WT
fil1Δ+Fil1
fil1Δ
WT
fil1Δ+Fil1
fil1Δ
WT
fil1Δ+Fil1
fil1Δ
Rich
EMM+aa
EMM
B
p=10-40
488
85
80
Reduced in fil1Δ
Induced by 3-AT (wild type)
C
fil1Δ + 3-AT
fil1Δ no 3-AT
D
fil1Δ no 3-AT
wild type no 3-AT
E
nmt1-fil1 promoter ON
nmt1-fil1 promoter OFF

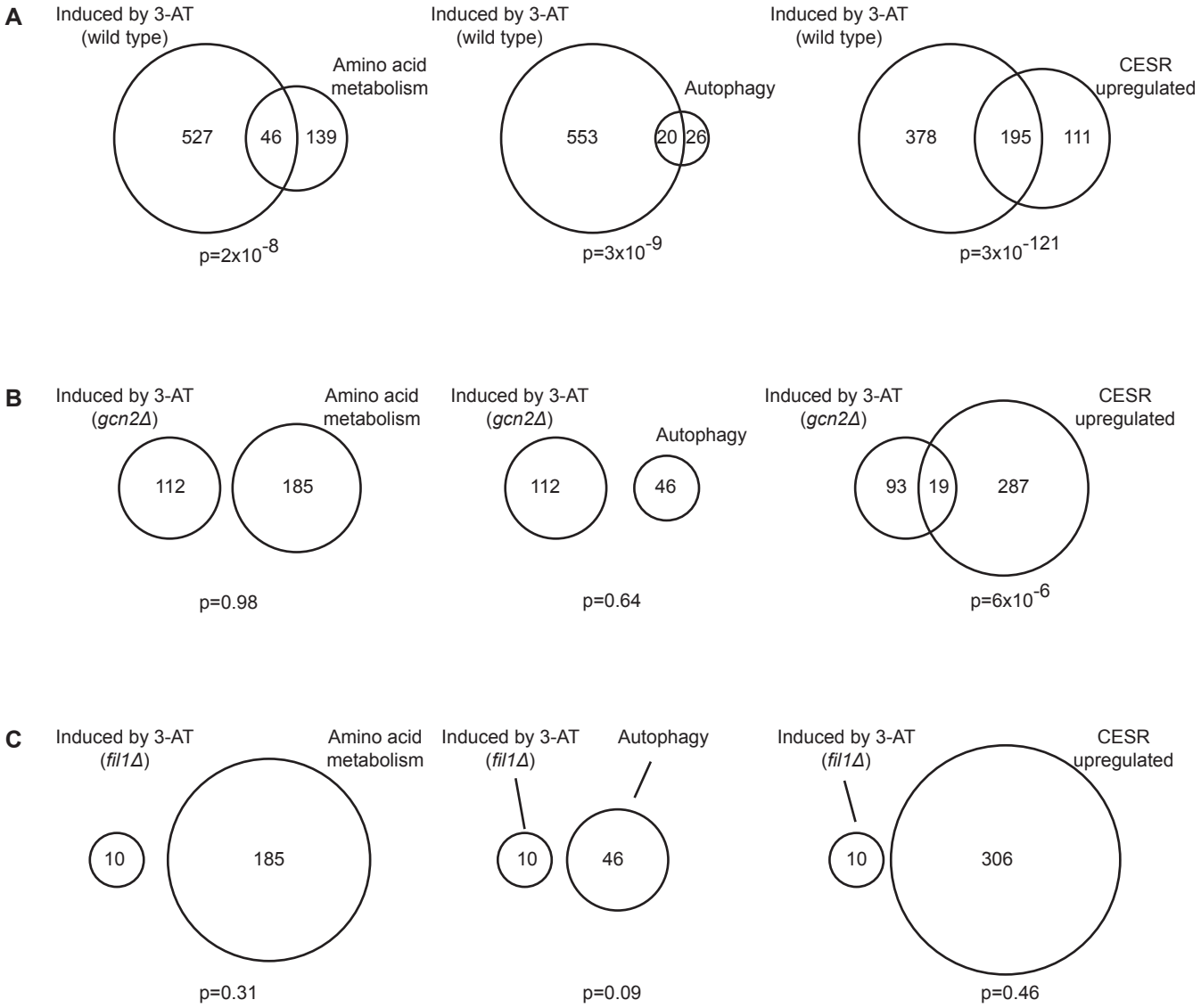

A
asn1
Fil1-TAP enrichment
734,500
735,500
736,500
737,500
B
log2 mRNA change
2
0
−2
WT+/− 3-AT
fil1Δ/WT no 3-AT
fil1Δ/WT 3-AT
C
Bound by Fil1 (no 3-AT)
Reduced in fil1Δ (no 3-AT)
191
49
116
p=7x10-29
D
Bound by Fil1 (no 3-AT)
Increased in fil1Δ (no 3-AT)
233
32
7
p=0.003
E
Bound by Fil1 (+ 3-AT)
Induced 3-AT
113
57
516
p=2x10-16
F
Bound by Fil1 (+ 3-AT)
Repressed 3-AT
157
343
13
p=0.29
G
Bound by Fil1 (no 3-AT)
Bound by Fil1 (+ 3-AT)
107
133
37
p=10-186
H
Bound by Fil1 (no 3-AT)
Bound by Gcn4 (no 3-AT)
140
26
127
p=9x10-11

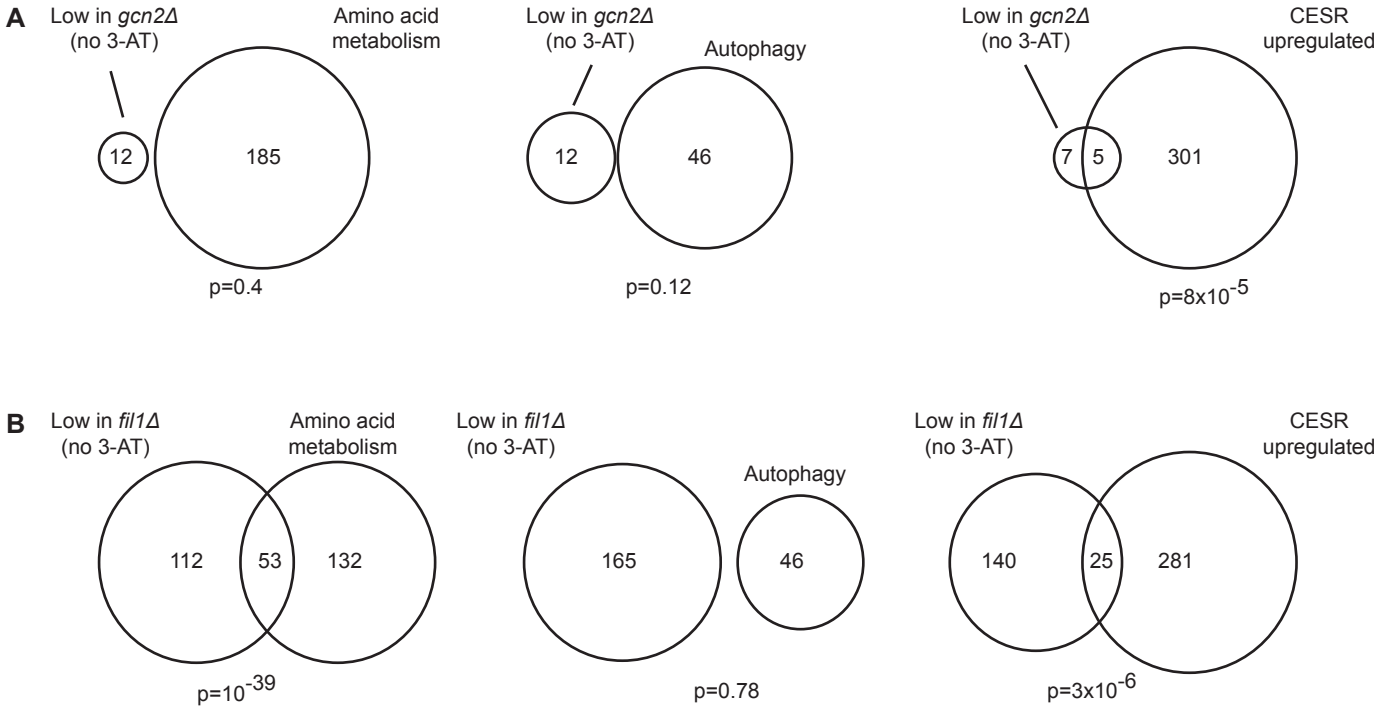

A
CUG1
AUG
1 2 3 4 5
CDS
No 3-AT
+ 3-AT
Reads per million reads (RPFs)
811,500
812,500
813,500
814,500
B
Reads per million reads (RPFs)
3-AT
CDS
leader
C
(kDa)
Fil1-TAP
H3
gcn2Δ
Wild type
Fil1-TAP protein levels
Rep. 1
Rep. 2
Rep. 3
Time in 3-AT (min)
D
Normalised levels (to untreated fil1 leader)
5' leader
adh1
fil1-6x
Protein
RNA
mCherry reporter
E
Normalised fulorescence (to untreated cells)
mCherry reporter (protein)
Normalised RNA levels (to untreated cells)
mCherry reporter (RNA)
wt
His
gcn2Δ

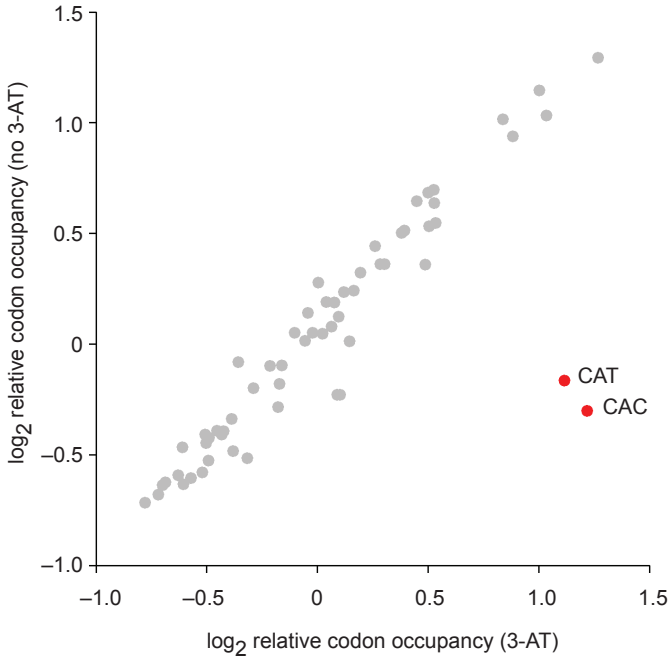

A
Induced by 3-AT
Udagawa
Cluster 2
454
119
47
p=4x10-81

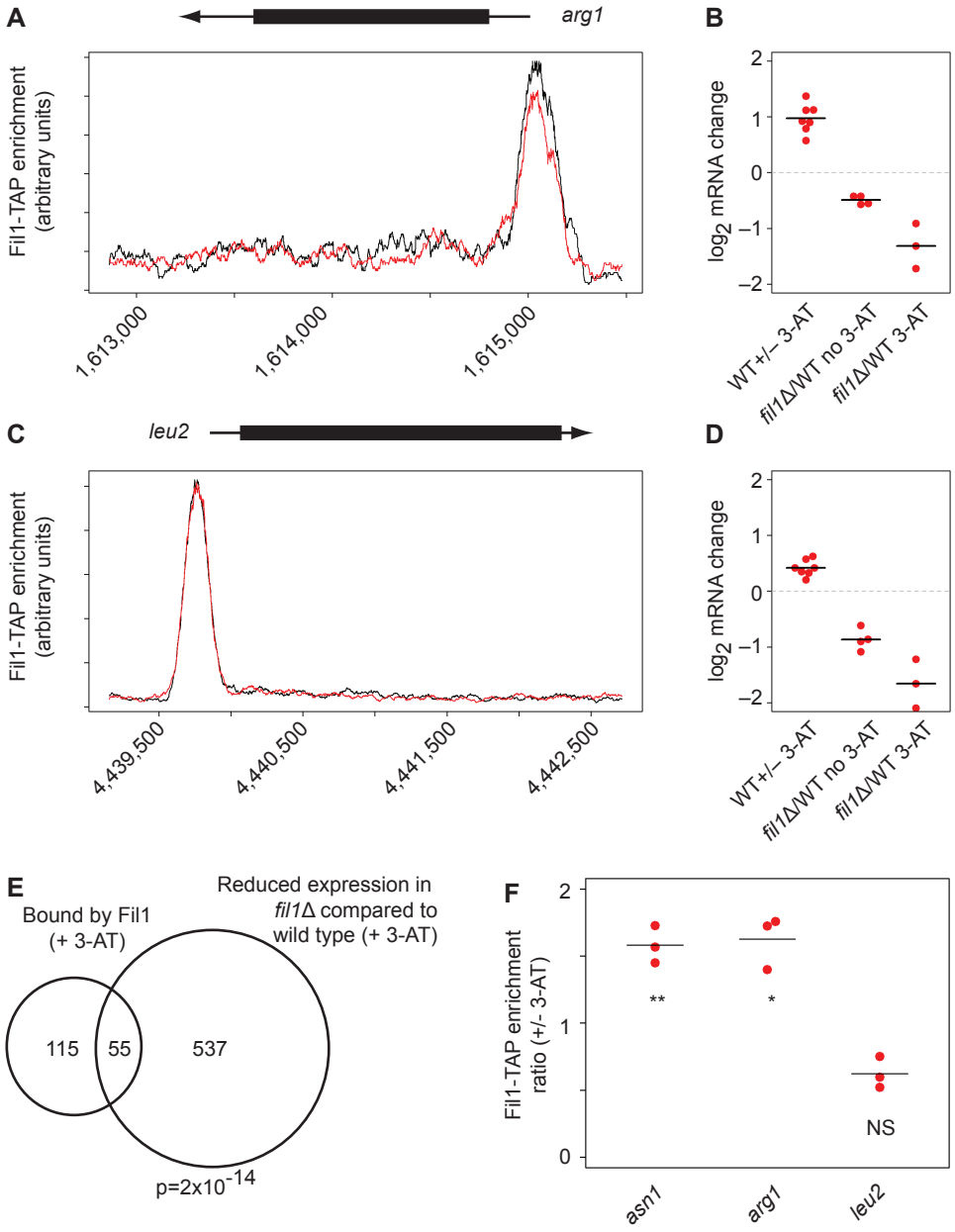

B
Induced by 3-AT
Udagawa
Cluster 3
514
59
7
p=1x10-51

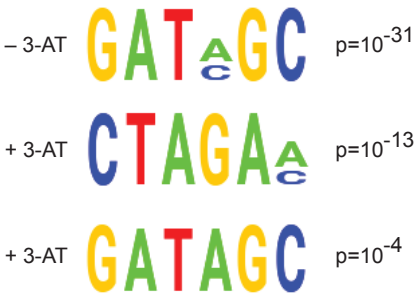

C
Repressed by 3-AT
Udagawa
Cluster 1
326
30
57
p=3x10-15

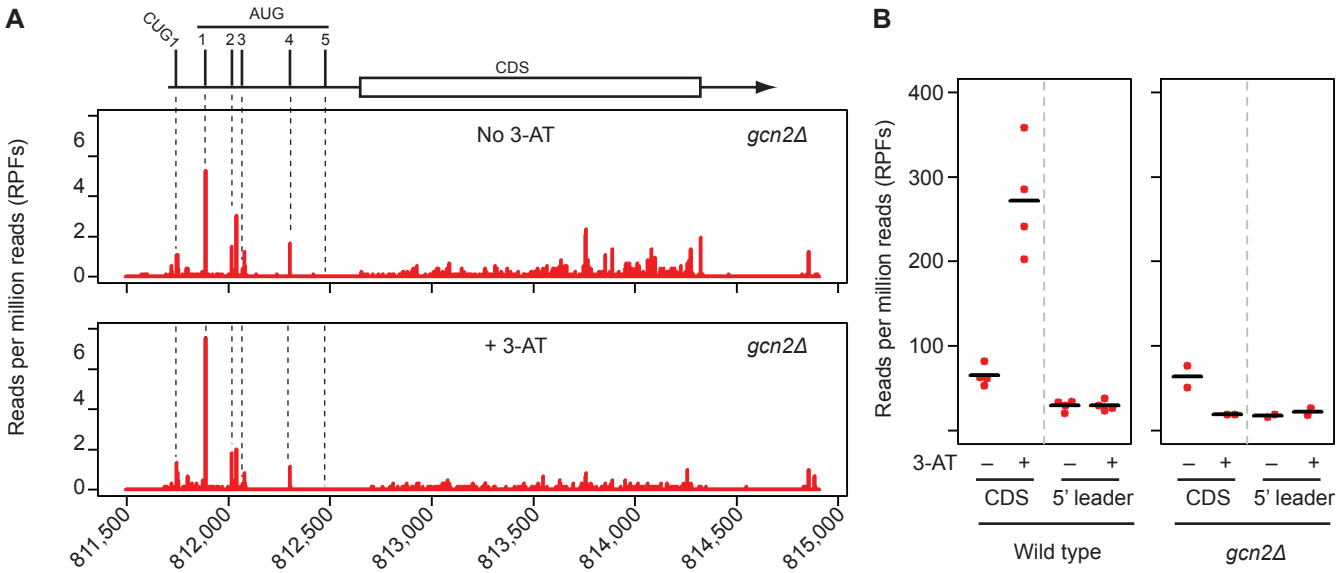

A
Induced by 3-AT
(wild type)
Amino acid
metabolism
527
46
139
p=2x10-8
Induced by 3-AT
(wild type)
Autophagy
553
20
26
p=3x10-9
Induced by 3-AT
(wild type)
CESR
upregulated
378
195
111
p=3x10-121
B
Induced by 3-AT
(gcn2Δ)
Amino acid
metabolism
112
185
p=0.98
Induced by 3-AT
(gcn2Δ)
Autophagy
112
46
p=0.64
Induced by 3-AT
(gcn2Δ)
CESR
upregulated
93
19
287
p=6x10-6
C
Induced by 3-AT
(fil1Δ)
Amino acid
metabolism
10
185
p=0.31
Induced by 3-AT
(fil1Δ)
Autophagy
10
46
p=0.09
Induced by 3-AT
(fil1Δ)
CESR
upregulated
10
306
p=0.46

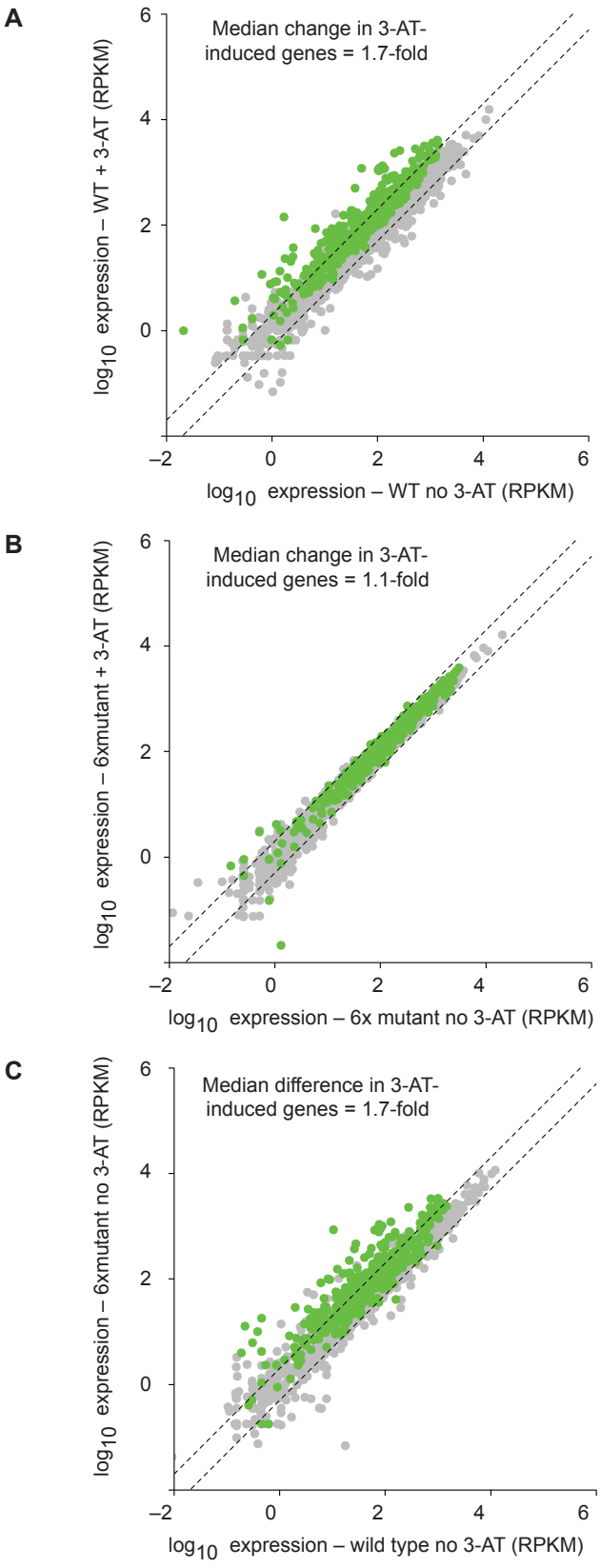

A
Low in gcn2Δ
(no 3-AT)
Amino acid
metabolism
12
185
p=0.4
Low in gcn2Δ
(no 3-AT)
Autophagy
12
46
p=0.12
Low in gcn2Δ
(no 3-AT)
CESR
upregulated
7
5
301
p=8x10-5
B
Low in fil1Δ
(no 3-AT)
Amino acid
metabolism
112
53
132
p=10-39
Low in fil1Δ
(no 3-AT)
Autophagy
165
46
p=0.78
Low in fil1Δ
(no 3-AT)
CESR
upregulated
140
25
281
p=3x10-6

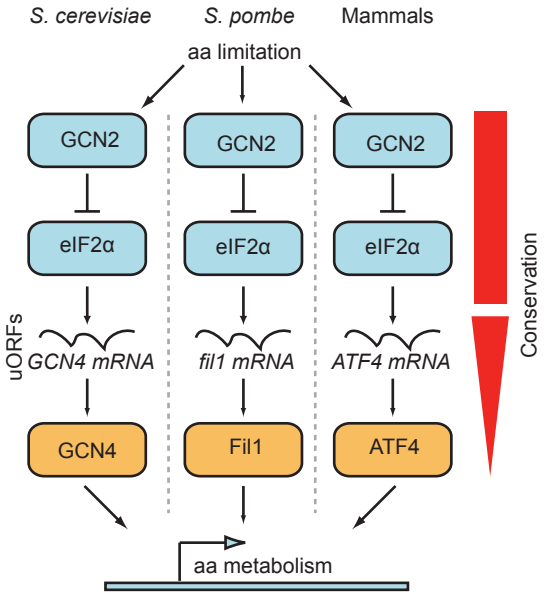

1.5
1.0
0.5
0
–0.5
–1.0
log2 relative codon occupancy (no 3-AT)
–1.0
–0.5
0
0.5
1.0
1.5
log2 relative codon occupancy (3-AT)
CAT
CAC


A
arg1
Fil1-TAP enrichment (arbitrary units)
1,613,000
1,614,000
1,615,000
B
log2 mRNA change
2
1
0
−1
−2
WT+/− 3-AT
fil1Δ/WT no 3-AT
fil1Δ/WT 3-AT
C
leu2
Fil1-TAP enrichment (arbitrary units)
4,439,500
4,440,500
4,441,500
4,442,500
D
log2 mRNA change
2
1
0
−1
−2
WT+/− 3-AT
fil1Δ/WT no 3-AT
fil1Δ/WT 3-AT
E
Bound by Fil1 (+ 3-AT)
Reduced expression in fil1Δ compared to wild type (+ 3-AT)
115
55
537
p=2x10-14
F
Fil1-TAP enrichment ratio (+/- 3-AT)
2
1
0
**
*
NS
asn1
arg1
leu2


– 3-AT
p=10-31
+ 3-AT
p=10-13
+ 3-AT
p=10-4


A
CUG1
AUG
1
2
3
4
5
CDS
No 3-AT
gcn2Δ
+ 3-AT
gcn2Δ
Reads per million reads (RPFs)
0
2
4
6
811,500
812,000
812,500
813,000
813,500
814,000
814,500
815,000
B
Reads per million reads (RPFs)
100
200
300
400
3-AT
–
+
CDS
5' leader
Wild type
gcn2Δ


A
Median change in 3-AT-induced genes = 1.7-fold
log10 expression – WT + 3-AT (RPKM)
log10 expression – WT no 3-AT (RPKM)
B
Median change in 3-AT-induced genes = 1.1-fold
log10 expression – 6xmutant + 3-AT (RPKM)
log10 expression – 6x mutant no 3-AT (RPKM)
C
Median difference in 3-AT-induced genes = 1.7-fold
log10 expression – 6xmutant no 3-AT (RPKM)
log10 expression – wild type no 3-AT (RPKM)
6
4
2
0
–2
0
2
4
6


S. cerevisiae
S. pombe
Mammals
aa limitation
GCN2
GCN2
GCN2
eIF2α
eIF2α
eIF2α
uORFs
GCN4 mRNA
fil1 mRNA
ATF4 mRNA
GCN4
Fil1
ATF4
aa metabolism
Conservation