

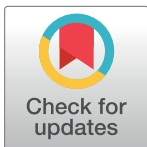
RESEARCH ARTICLE

Defining the role of the polyasparagine repeat domain of the *S. cerevisiae* transcription factor Azf1p

Taylor Stewart¹, Benjamin E. Wolfe¹, Stephen M. Fuchs^{1,2,3*}

1 Department of Biology, Tufts University, Medford, MA, United States of America, **2** Institute for Protein Innovation, Boston, MA, United States of America, **3** Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, United States of America

* stephen.fuchs@proteininnovation.org



Abstract

Across eukaryotes, homopolymeric repeats of amino acids are enriched in regulatory proteins such as transcription factors and chromatin remodelers. These domains play important roles in signaling, binding, prion formation, and functional phase separation. Azf1p is a prion-forming yeast transcription factor that contains two homorepeat domains, a polyglutamine and a polyasparagine domain. In this work, we report a new phenotype for Azf1p and identify a large set of genes that are regulated by Azf1p during growth in glucose. We show that the polyasparagine (polyN) domain plays a subtle role in transcription but is dispensable for Azf1p localization and prion formation. Genes upregulated upon deletion of the polyN domain are enriched in functions related to carbon metabolism and storage. This domain may therefore be a useful target for engineering yeast strains for fermentation applications and small molecule production. We also report that both the polyasparagine and polyglutamine domains vary in length across strains of *S. cerevisiae* and propose a model for how this variation may impact protein function.

OPEN ACCESS

Citation: Stewart T, Wolfe BE, Fuchs SM (2021) Defining the role of the polyasparagine repeat domain of the *S. cerevisiae* transcription factor Azf1p. PLoS ONE 16(5): e0247285. <https://doi.org/10.1371/journal.pone.0247285>

Editor: Alvaro Galli, CNR, ITALY

Received: January 31, 2021

Accepted: April 26, 2021

Published: May 21, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0247285>

Copyright: © 2021 Stewart et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting Information](#) files. Strains and plasmids are available upon request.

Introduction

Repeat domains are prevalent in the eukaryotic proteome and are found in 13% of proteins in yeast, 17% in humans, and 21% in *Drosophila*. Repeats are commonly found within intrinsically disordered regions (IDRs), which are protein domains that do not have a single native fold corresponding to a lowest energy state. Instead, these domains dynamically sample a number of conformations with similar energies [1]. Despite being historically dismissed as unimportant due to their lack of three-dimensional structure, repetitive protein domains are essential for binding, signaling, and functional phase separation. Across eukaryotes, repeats are enriched in cell-surface and regulatory proteins, such as transcription factors (TFs) and chromatin remodelers [2]. Repeats enriched in a single amino acid are deemed homorepeats. Polyglutamine (polyQ) is the most common type of homorepeat in eukaryotic proteins, followed by polyasparagine (polyN). While polyQ domains are found across eukaryotes, polyN domains are rare in vertebrates and found almost exclusively in more primitive invertebrates

Funding: This work was supported by the Army Research Office [W911NF-16-1-0175, W911NF-19-1-0299, and W911NF-20-1-0083 to S.M.F.]

Competing interests: The authors have declared that no competing interests exist.

[3]. Synthetic polyN peptides have been shown to mediate aggregation [4,5], but the role of natural polyN domains in protein function has only been explored in a few cases. For example, the polyN domain of Pdr1p has been shown to be important for transcriptional activation [6]. Whether or not this function expands to other polyN-containing TFs, such as *CCR4*, *MIT1*, *ASG1*, *HAP4*, *ABF1*, *MOT3*, *AZF1*, and *RLM1*, has not been investigated.

Several polyQ and polyN-containing TFs in yeast (*HAP4*, *AZF1*, *MOT3*, *RLM1*, *CYC8*, *MGA1*, and *SWI1*) have been demonstrated to form prions. Prions are proteins that have multiple conformations, at least one of which is self-templating and epigenetically heritable. Prion-forming domains (PrDs) are characterized by low-complexity amino acid sequences that are often rich in N and Q, insensitive to scrambling, and intrinsically disordered in their native state [7]. Many prion-forming proteins contain repeats that are important to the function of both the prion and the native protein. Although there is only one bona fide prion in mammals, PrP^{Sc}, screens to identify yeast prions have revealed that they are common in both lab and wild strains of yeast [8–11]. While the mammalian PrP^{Sc} and prion-like proteins are well documented to cause both infectious diseases and age-related neurodegeneration in mammals [12–18], there is mounting evidence that yeast prions act as adaptive strategies, especially during environmental stress. Yeast prions are well-suited to mediate adaptation due to their enrichment in TFs and other regulatory proteins [10,11,19–29]. A screen conducted by Chakrabortee and colleagues identified nearly 50 novel prion-like proteins in yeast that conferred epigenetically heritable phenotypes after transient overexpression. In many cases, the prion state improved yeast growth rate in different environmental stresses, including osmotic and acidic stress. Like previously identified prions, many of the proteins identified in this screen are TFs and RNA-binding proteins [10].

Prion-forming TFs have the capacity to impact yeast growth through multiple mechanisms. These proteins can alter gene expression in both their native and prion states, leading to changes cellular functions such as metabolism, surface display, and stress tolerance. These proteins are therefore important targets for yeast metabolic engineering. Yeast are used in a wide range of applications in research, food and drink production, and small molecule production. Because homorepeat domains are at the intersection of TFs and prions, these domains are particularly interesting targets for engineering. Target genes of Q-rich TFs in yeast have significantly higher levels of expression divergence, expression variability, mutational variance, and expression noise when compared to targets of non-repeat-containing TFs [30]. These results highlight the importance of repeat domains in modulating gene expression.

The yeast TF Azf1p was recently shown to form a prion [10]. Azf1p has a largely nonoverlapping set of target genes during growth in fermentable and non-fermentable carbon sources, and the null mutant has been shown to have a severe growth defect in glycerol. *AZF1* is also important for cell wall integrity [31]. The prion conformer, designated [*AZF1*⁺], confers resistance to the drug radicicol in a gain-of-function manner but decreases the expression of Azf1p's target genes [10]. Azf1p contains a polyN and polyQ domain, and the frequency of homorepeats in yeast TFs and PrDs predicts a role for these domains in transcription as well as prion formation. In this work, we investigated the function of the polyN domain of Azf1p and find that this domain plays a subtle role in transcription but is not required for Azf1p localization or prion formation.

Methods

Yeast strains and plasmids

Plasmids used in this work are summarized in [S1 Table](#). The ΔN mutant plasmid was constructed by QuikChange site-directed mutagenesis using pBY011-*AZF1* (Harvard PlasmID) as

a template with the primers described in [S2 Table](#). Yeast were grown on synthetic complete (SC) dropout medium, YPD (yeast extract, peptone, 2% dextrose), or YPG (yeast extract, peptone, 4% glycerol) at 30°C. Plasmids were freshly transformed into yeast and maintained on SC media lacking uracil (SC-Ura). Strains used in this study are summarized in [S3 Table](#) and were derived from BY4741 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*). The WT strain is from the yeast GFP collection in which the *AZF1* open reading frame is C-terminally tagged with GFP [[32,33](#)]. *AZF1ΔN* was constructed in this background using CRISPR [[34](#)]. A guide RNA sequence designed to target the polyQ domain of *AZF1* was cloned into the Cas9-expressing plasmid pML104. A ~1 kb fragment of *AZF1* containing the polyN deletion was amplified by PCR with 8 identical reactions using the *AZF1* F/R primers ([S2 Table](#)) and pBY011-ΔN as a template. Ethanol-precipitated PCR products were co-transformed with pML104-polyQ, and cells were plated on SC-Ura. After 5 days, individual transformants were selected and grown overnight in 1 mL YPD in a 96 deep-well plate to allow for loss of the Cas9 plasmid. Overnight cultures were diluted ten-fold four times and spotted on FOA. Fast-growing FOA-resistant colonies were selected and characterized by colony PCR with the *AZF1* F/R primers to identify yeast with the polyN deletion. This strain is referred to as ΔN in the text. *azf1-Δ1::URA3* was constructed in BY4741 using heterologous gene replacement with the *URA3* gene as a selectable marker. This strain was intended to be a full deletion of the *AZF1* open reading frame; however, *URA3* is instead integrated between *AZF1* and its promoter, abolishing *AZF1* expression ([S4 Table](#)). In the text, this strain is referred to as *azf1Δ* because it mimics the phenotypes associated with deletion of the *AZF1* open reading frame.

Spotting assays

For phenotypic growth assays, yeast were grown overnight in YPD. Saturated overnight cultures were used to start fresh cultures in the same medium at an OD₆₀₀ of 0.2. The cells were allowed to double at least two times before approximately 1.0 x 10⁷ cells were harvested and resuspended in sterile water in a 96-well plate. Cells were serially diluted five-fold five times and then spotted onto YPD or YPG plates using a 48-pin replicating tool. Plates were incubated at 30°C and imaged after three (YPD) to seven (YPG) days.

Growth curve assays

Each pBY011-*AZF1* plasmid (WT and ΔN) was transformed into the matched background and maintained on SC-Ura (i.e. the ΔN plasmid was transformed into *AZF1ΔN*). Three biological replicates of each strain were grown overnight in SC_{Raff}-Ura (2% raffinose). 2 μL of each culture were then transferred to 200 μL of four types of fresh media in a 96-well plate: SCGal-Ura (2% galactose), SCGal-Ura+Radicicol (75 μM), SC-URA, and SC-Ura+Radicicol. Growth was measured for 24 hours by OD₆₀₀ in a Molecular Devices SpectraMax M5 plate reader at 30°C with shaking prior to each reading. Readings were taken every 30 minutes. 10 μL of each SCGal-Ura and SC-Ura culture were then transferred to 1 mL fresh SC-URA and allowed to grow to saturation overnight. The plate reader assay was then repeated in SC-URA media with and without radicicol. We compared the growth of cells whose ancestors had experienced *AZF1* overexpression to those that did not in order to determine the role of the polyN domain in [*AZF1*⁺] prion formation and activity.

A modified version of this plate reader assay was also performed with the WT, ΔN, and *azf1Δ* strains without the galactose-inducible plasmids. Three biological replicates were grown overnight in YPD. 2 μL of each saturated culture were then transferred to 200 μL of SC media, and growth was measured for 24 hours in the plate reader.

RNA-seq

Three biological replicates were grown overnight in either YPD or YPG. Saturated overnight cultures were diluted to OD₆₀₀ 0.2. Cells were grown to OD₆₀₀ 0.6–1.0 and collected by centrifugation. Total RNA was isolated using the GE Healthcare RNAspin Mini RNA Isolation Kit per the manufacturer's guidelines and was quantified by NanoDrop (Thermo Fisher Scientific). mRNA was isolated using a NEBNext poly(A) mRNA magnetic isolation module (New England Biolabs). mRNA was used to generate RNA-seq libraries using a NEBNext Ultra II RNA library prep kit for Illumina (New England Biolabs) following the manufacturer's recommended protocol with the following parameters: 15 minutes at 42°C for the RNA fragmentation time, a 1:25 dilution of the adapter, no size selection, and 15 cycles of denaturation/annealing/extension during the PCR enrichment step. After quantifying the library concentration with a Qubit HS DNA assay and determining the average peak size using a Fragment Analyzer, equimolar concentrations of each library were pooled together and sequenced on an Illumina NextSeq 550 at the Tufts University Genomics Core using a high output run with 150 cycles.

Data analysis was performed using Geneious Prime v. 2020.2.4 (Biomatters Ltd.). Sequencing reads were aligned to the s288c genome using the Geneious RNA Mapper with Medium-Low Sensitivity/Fast settings. Multiple best matches were mapped randomly. Expression levels were then calculated, and ambiguously mapped reads were counted as multiple full matches. To identify differentially expressed genes, expression levels were compared between strains in each carbon source using DEseq2 with a parametric fit type. Expression levels were also compared between carbon sources for each strain. Genes that are statistically significantly differentially expressed ($p < 0.001$) in *azf1Δ* compared to WT were selected for GO Enrichment Analysis. The genes in this subset all have log₂ ratios with a magnitude of 0.30 or greater. Because very few genes are statistically significantly differentially expressed in ΔN, genes with log₂ fold ratios of ± 0.30 or more were selected for analysis. GO Enrichment Analysis was performed using PANTHER v. 14 [35–37]. The PANTHER Overrepresentation Test (Released 20200728) employs a Fisher's exact test and reports significant enrichment in biological processes ($p < 0.05$).

Target gene expression measurements by RT-qPCR

RNA was isolated as described above, and cDNA synthesis was performed using the SuperScript™ First-Strand Synthesis System (Invitrogen) following the manufacturer's guidelines with an oligo-dT(20) primer and 300 ng of RNA per sample. cDNA was diluted 1:5 and stored at -20°C. Each 20 μL reaction was prepared using the Brilliant II SYBR Green QPCR system (Agilent) following the manufacturer's guidelines with the primers listed in S2 Table. Samples were contained in MicroAmp 96-Well Reaction Plates (Applied Biosystems) sealed with MicroAmp Optical Adhesive Film (Applied Biosystems). Amplification was performed using an Applied Biosystems 7300 Real-Time PCR System with the following cycling conditions: 5 minutes at 95°C and 40 cycles of 10 seconds at 95°C followed by 30 seconds at 55°C. The reference gene ACT1 was used for relative quantification. Gene expression changes were calculated using the $2^{-\Delta\Delta C_t}$ method. Two independent experiments were performed with three technical replicates for each biological replicate.

Microscopy

Cells in which *AZF1* is tagged with GFP were imaged on a Leica Thunder Imager in the log and stationary phases of growth with either glucose or glycerol as the carbon source. WT and ΔN cells were grown overnight in YPD or YPG. 100 μL of saturated YPD overnight cultures

were spun down and imaged to study cells in stationary phase with glucose as the carbon source. Saturated overnight YPD cultures were also diluted to an OD₆₀₀ of 0.2 and grown to mid-log phase. 1.0 mL of these cultures was spun down and imaged to study cells growing in log phase with glucose as the carbon source. YPG cultures reached mid-log phase after about 24 hours. 1.0 mL of these cultures were spun down and imaged to study cells growing in log phase with glycerol as the carbon source. YPG cultures were also allowed to grow for 48 hours, at which time 100 μ L were spun down and imaged to study cells growing in stationary phase with glycerol as the carbon source. Each condition was imaged two independent times, and at least 50 cells were imaged for each strain in each condition. Images were processed using Leica LAS X Instant Computational Clearing.

Evaluation of AZF1 repeat variation

Variation within *AZF1* was measured using the data available from 93 recently sequenced *S. cerevisiae* genomes [38] as described in Babokhov *et al.* 2018 [47].

Results

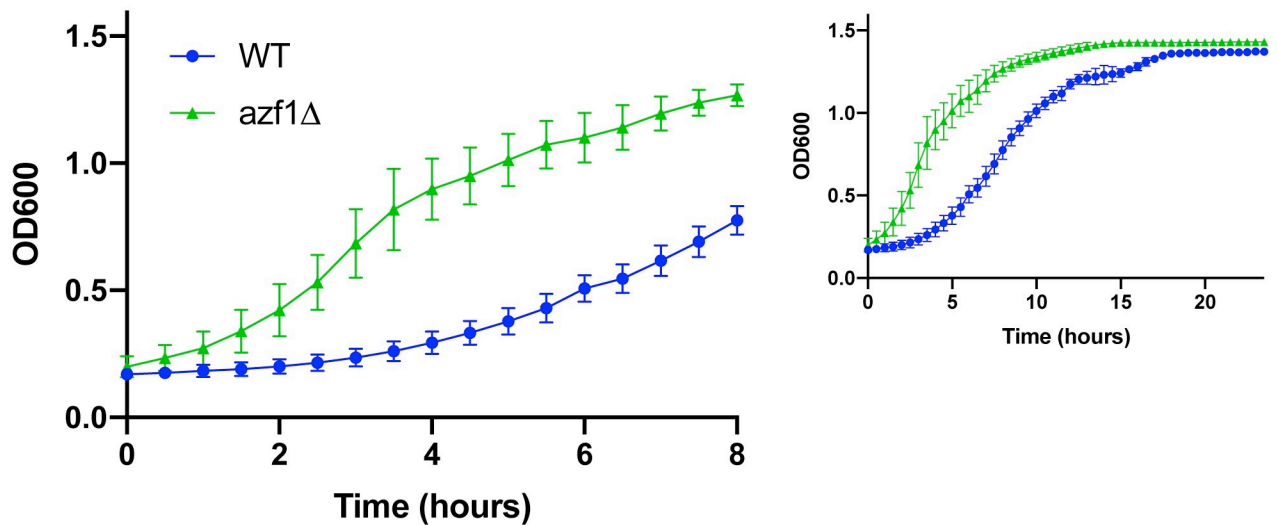
Azf1p is a prion-forming yeast TF that has been shown to activate genes involved in carbon metabolism and energy production during growth in glucose [10,31]. Previous work on Azf1p has been limited, so we began by monitoring the activity of *azf1 Δ* in different carbon sources. A null mutant has previously been shown by spotting assay to produce a severe growth defect on non-fermentable carbon sources [31]. Slattery and colleagues also investigated the impact of the null mutant on growth on glucose by spotting assay, which did not reveal any differences between *azf1 Δ* and the WT. We sought to further explore the role of Azf1p in glucose by monitoring growth in liquid media by optical density in a plate reader. We find that the null mutant has a shortened lag phase and reaches the log phase of growth more quickly when compared to the parent strain (Fig 1A). To further investigate the role of Azf1p during growth in glucose, we performed RNA-seq with RNA isolated from *azf1 Δ* and WT cells grown in glucose and found many genes that are differentially expressed in the null mutant (Fig 1B, S1 Fig, S5 Table).

Genes downregulated by *azf1 Δ* have not been previously reported, and we hypothesized that Azf1p may act as a co-repressor for these genes. The general co-repressor Cyc8p contains two Q-rich domains, one of which has been shown to mediate binding with cofactors [30]. Azf1p contains a polyQ and polyN domain (Fig 2A), and we hypothesized that one or both of these domains may be important for transcriptional repression by Azf1p. We first deleted the polyN domain using CRISPR (Fig 2A). We attempted to delete the polyQ domain using the same approach; however, deletion of the polyQ domain using this strategy was never successful. In an effort to probe the role of the polyQ domain, we deleted this domain from diploid BY4743. While the heterozygous mutant was viable, the strain did not sporulate. Null mutants are known to be sporulation deficient [31], and the sporulation deficiency observed in the heterozygous mutant suggests a fundamental role for the polyQ domain in Azf1p function.

Differential roles of the polyN domain during growth in different carbon sources

PolyQ domains have been extensively studied due to their involvement in a number of mammalian neurodegenerative diseases. Despite being the second most common type of homorepeat in the eukaryotic proteome, the function of polyN repeats has been underexplored. We were therefore interested in characterizing the role of the polyN domain of Azf1p. The severe growth defect in glycerol media produced by the null mutant is one of Azf1p's most well-

A.



B.

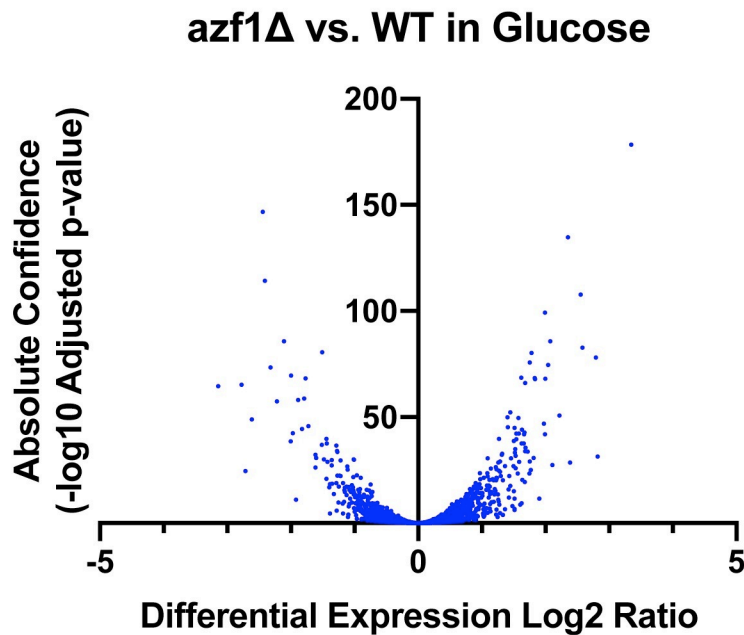


Fig 1. *azf1Δ* shortens lag phase during growth in glucose. A) Growth of WT and *azf1Δ* with glucose as the carbon source. Growth was measured by OD₆₀₀ in a plate reader for 24 hours. The first 8 hours are highlighted to show the differences in growth more clearly and due to the reduced accuracy of the plate reader at high optical densities. *azf1Δ* has a dramatically shorter lag phase than WT. Growth curves are the average of three biological replicates, and error bars indicate standard deviation. B) Volcano plot showing genes that are differentially expressed in *azf1Δ* compared to WT.

<https://doi.org/10.1371/journal.pone.0247285.g001>

characterized phenotypes. We therefore measured the impact of deleting the polyN domain on growth in glycerol by spotting assay (Fig 2B). Deletion of the polyN domain produces a moderate growth defect on glycerol when compared to the parent strain. WT, ΔN, and *azf1Δ*

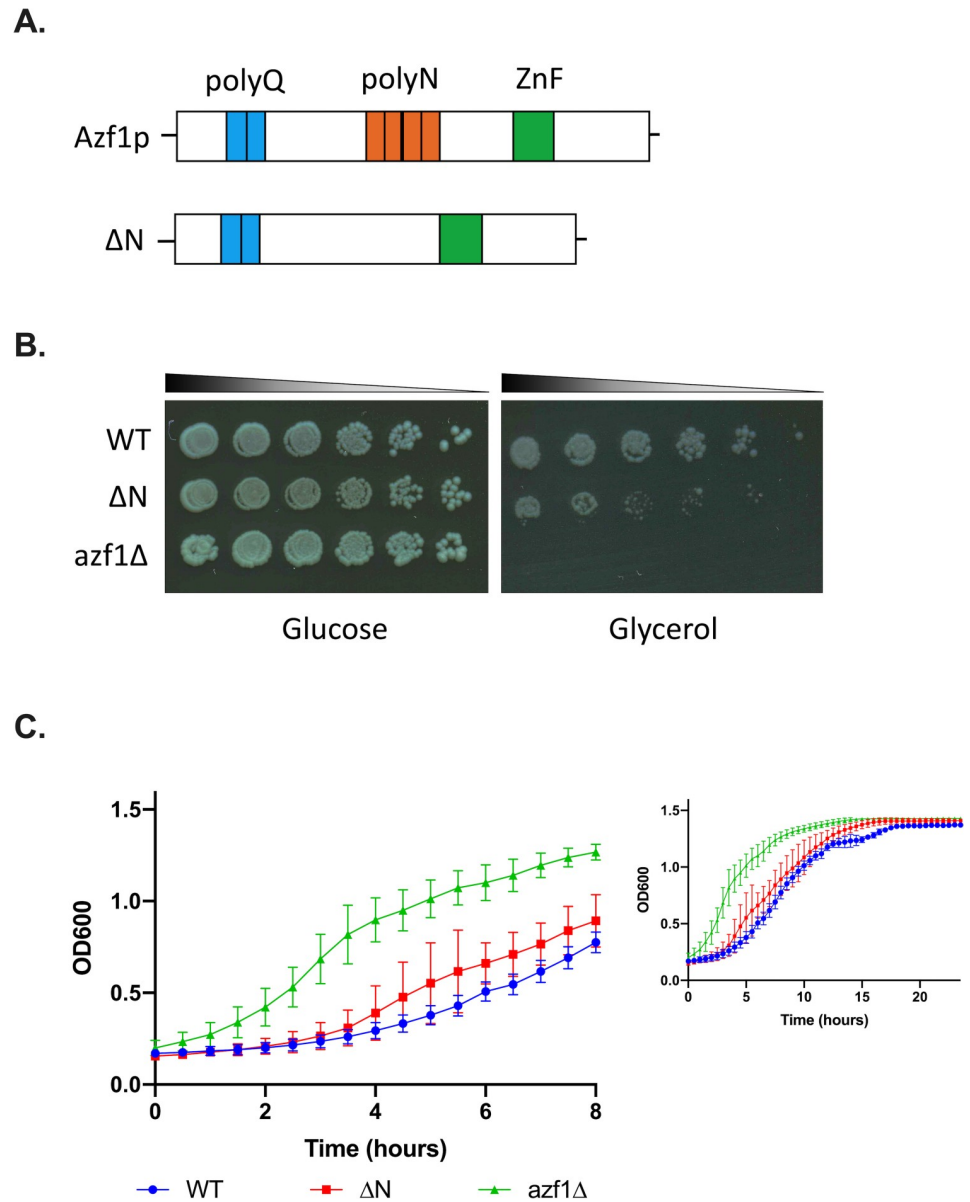


Fig 2. The polyN domain of Azf1p impacts growth in different carbon sources. A) Schematic of WT and ΔN Azf1p. Azf1p contains a polyQ (teal), polyN (orange), and a zinc-finger DNA-binding domain (ZnF) (green). To probe the role of the polyN domain, we deleted this domain in *AZF1-GFP* [32,33]. B) Spotting assay measuring the growth of WT, ΔN , and *azf1* Δ on glucose and glycerol. On glycerol, deletion of the polyN domain produces a mild growth defect, while *azf1* Δ produces a severe growth defect. Images are representative of 3 independent experiments. C) Growth of WT, ΔN , and *azf1* Δ with glucose as the carbon source. Growth was measured by OD₆₀₀ in a plate reader for 24 hours. The first 8 hours are highlighted to show the differences in growth more clearly and due to the reduced accuracy of the plate reader at high optical densities. Deletion of the polyN domain moderately shortens the lag phase, and *azf1* Δ dramatically shortens the lag phase compared to WT. Growth curves are the average of three biological replicates, and error bars indicate standard deviation.

<https://doi.org/10.1371/journal.pone.0247285.g002>

appear to grow similarly on glucose, which is consistent with the finding by Slattery and colleagues that *AZF1* is not essential for growth in fermentable carbon sources. In order to detect any more subtle differences in growth between these strains in glucose, we monitored growth for 24 hours by optical density in a plate reader (Fig 1C). We find that deletion of the polyN

domain also reduces the length of the lag phase, although not to the same extent as the null mutant. This suggests that the polyN domain plays a fundamental role in Azf1p function, although deletion of this domain does not completely abolish Azf1p activity. To demonstrate that the growth phenotype observed in glucose is due to the deletion of the polyN domain, we repeated the plate reader assay with the WT and ΔN strains expressing a wildtype copy of *AZF1* from the pBY011-*AZF1* plasmid. Expression of wildtype *AZF1* goes beyond rescuing the shortened lag phase phenotype and produces a growth defect in ΔN cells compared to WT (S3A Fig).

PolyN domains are prevalent in yeast TFs, but their function has only been explored in a few cases. In order to evaluate the role of the polyN domain in transcription, we performed RNA-seq using RNA isolated from WT and ΔN cells grown in both glucose and glycerol. We find that deletion of the polyN domain alters gene expression in both glucose (Fig 3A, S6 Table) and glycerol (Fig 3B, S7 Table) compared to WT. Notably, there are more statistically significantly differentially expressed genes in ΔN during growth in glucose compared to glycerol. The number of differentially expressed genes and the magnitudes of these changes, however, are much less than those observed in the null mutant, which indicates that the polyN domain plays a subtle Azf1p function.

In order to test the validity of the RNA-seq experiment, we measured relative gene expression of several of the most differentially expressed genes from each carbon source, as well as two known Azf1p target genes (*MDH2* and *GAS1*), by qPCR. Overall, the relative gene expression trends observed in ΔN and *azf1* Δ compared to WT are consistent with those measured by RNA-seq (S2 Fig).

In order to gain a better understanding of how the transcriptional changes in ΔN underlie the growth phenotypes observed in Fig 2B and 2C, we characterized differentially expressed genes with a log₂ ratio of 0.30 or higher for upregulated genes and -0.30 or lower for downregulated genes by GO Enrichment Analysis [35–37]. Genes upregulated by ΔN during growth in glucose are enriched in two main biological processes: environmental stress tolerance and carbon metabolism and storage (Fig 3C). There are only 3 genes downregulated -0.30 -fold or more during growth in glucose. As a result, there is no statistically significant biological process enrichment. Out of 119 genes upregulated 0.30-fold or higher in ΔN during growth in glycerol, 5 are involved in gluconeogenesis (17-fold enrichment), 8 are involved in glycolysis (15-fold enrichment), and 13 are involved in cell wall organization (3-fold enrichment). Out of 56 genes downregulated -0.30 -fold or more in glycerol, 12 are involved in rRNA processing (4-fold enrichment) (Fig 3E).

The polyN domain is dispensable for Azf1p localization

Azf1p has previously been shown to be localized to the nucleus during growth in glucose in both log and stationary phase, consistent with its function as a TF. Azf1p is also nuclear during log phase growth in glycerol, but relocalizes to cytoplasmic foci during stationary phase growth in glycerol [31]. Because polyN domains have previously been shown to promote protein aggregation [4,5], we hypothesized that the polyN domain is important for the formation of these foci. During growth in glucose, we find that both WT and ΔN localize to the nucleus (Fig 4A). In glycerol, both WT and ΔN localize to the nucleus during log phase, and both form cytoplasmic foci during stationary phase growth (Fig 4B). We therefore conclude that the polyN domain is not required for Azf1p relocalization during stationary phase growth in glycerol.

[*AZF1*⁺] prion formation does not require the polyN domain

Azf1p was recently shown to form a prion, designated [*AZF1*⁺], that confers resistance to the drug radicicol. Both the polyQ and polyN domains are predicted to be capable of promoting

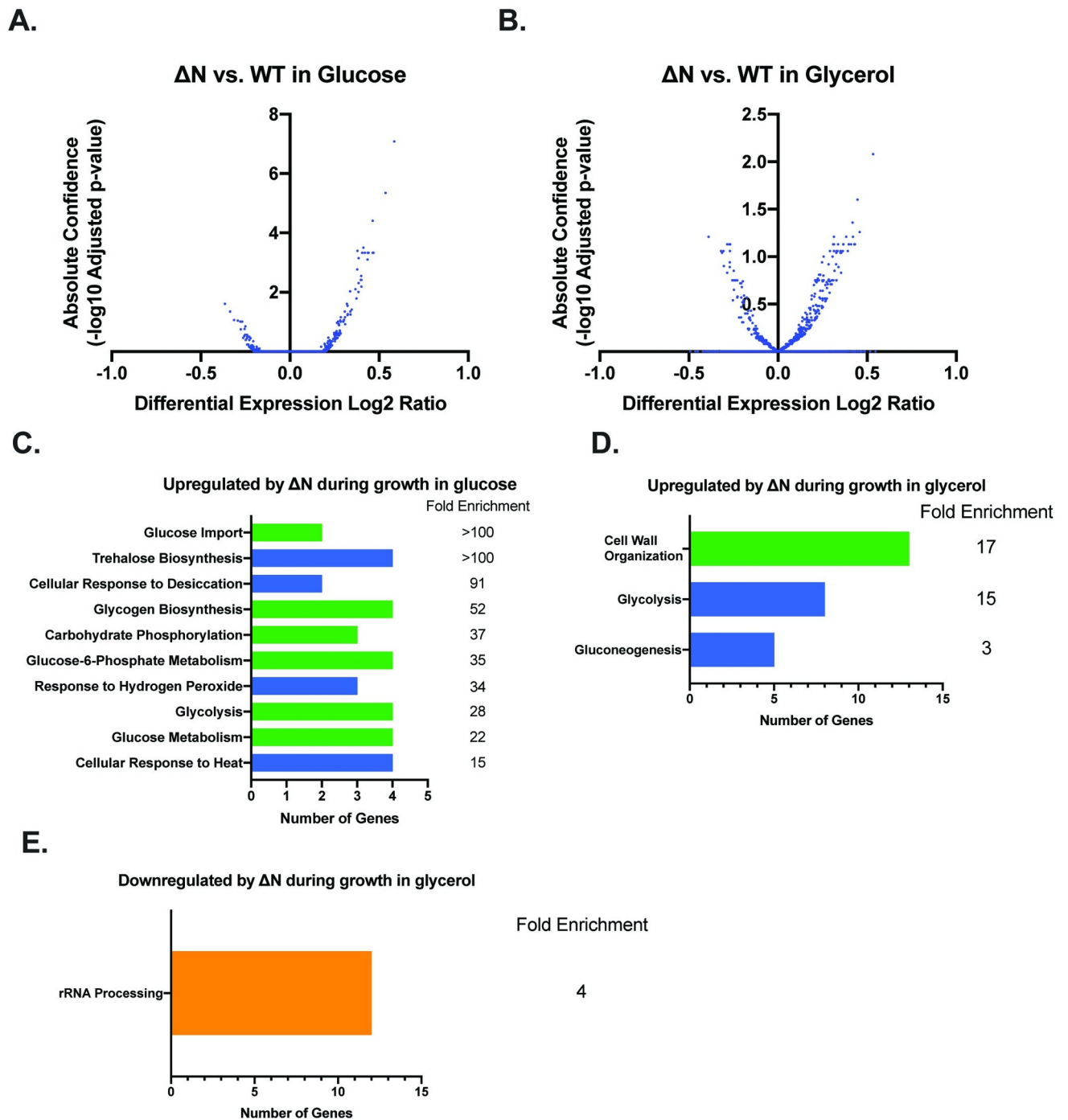


Fig 3. The polyN domain of Azf1p plays a subtle role in transcription. A) Volcano plot showing genes that are differentially expressed in ΔN compared to WT in glucose. B) Volcano plot showing genes that are differentially expressed in ΔN compared to WT in glycerol. C) GO Enrichment Analysis of upregulated genes with a log₂ ratio of 0.30 or more in ΔN in glucose compared to WT. Fold enrichment is indicated to the right of each bar. Processes related to environmental stress tolerance are shown in blue, and processes related to carbon metabolism and storage are shown in green. D) GO Enrichment Analysis of upregulated genes with a log₂ ratio 0.30 or higher in ΔN in glycerol compared to WT. Fold enrichment is indicated to the right of each bar. Processes related to environmental stress tolerance are shown in blue, and processes related to carbon metabolism and storage are shown in green. E) GO Enrichment Analysis of upregulated genes with a log₂ ratio ≤ -0.30 or lower in ΔN in glycerol compared to WT. rRNA processing is the only process for which this set of genes is statistically significantly enriched (4-fold enrichment).

<https://doi.org/10.1371/journal.pone.0247285.g003>

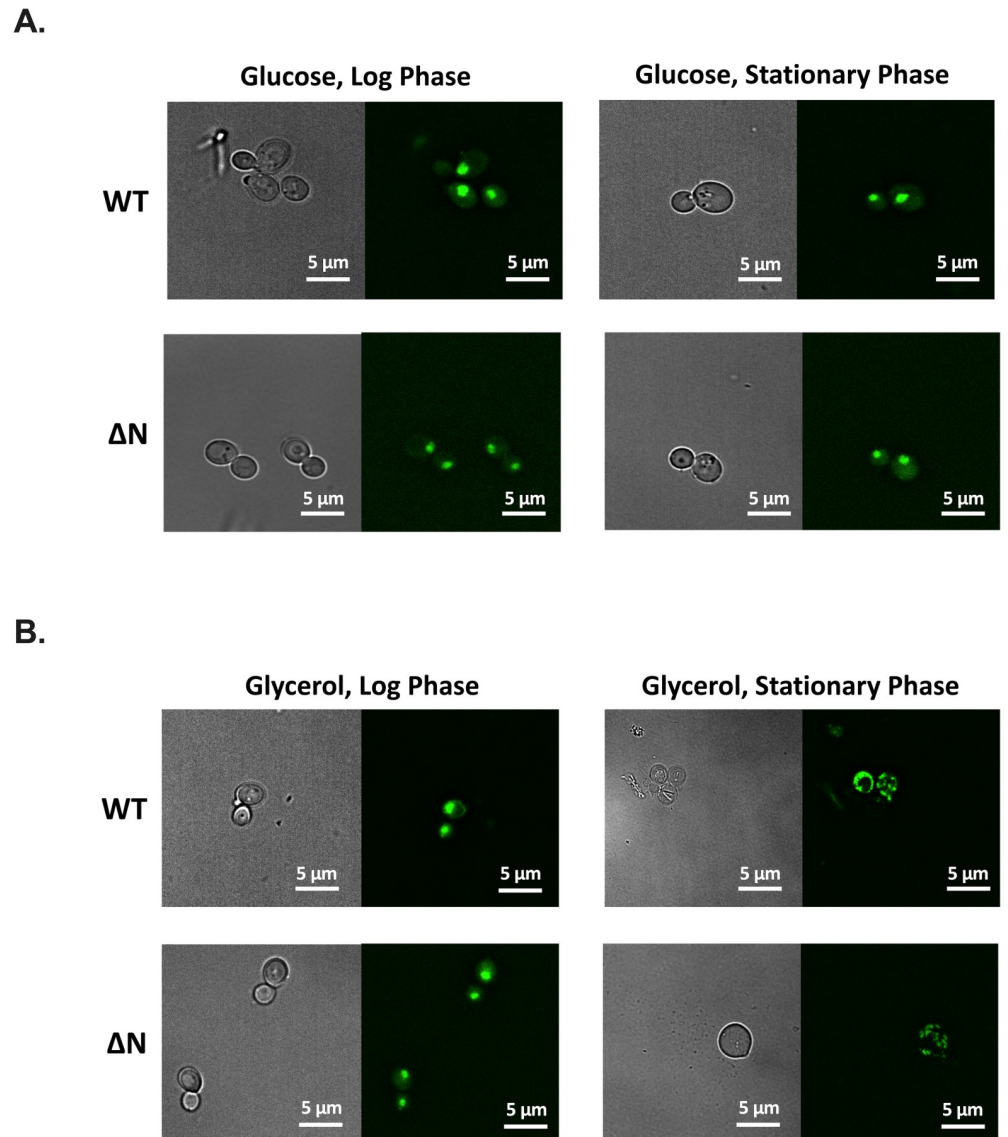


Fig 4. Role of the polyN domain in localization of Azf1p. A) Representative images of localization of Azf1p in glucose during growth in log and stationary phase. During growth in glucose, Azf1p is localized to the nucleus. B) Representative images of localization of Azf1p during growth in glycerol in log and stationary phase. During log phase, Azf1p is localized to the nucleus. During stationary phase, both Azf1p and Azf1-ΔNp form cytoplasmic foci. At least 50 cells were imaged for each strain in each condition.

<https://doi.org/10.1371/journal.pone.0247285.g004>

prion formation by the Prion-Like Amino Acid Composition (PLAAC) algorithm [39] (Fig 5A). Prion formation can be driven by overexpression of the constituent protein, and phenotypes conferred by the prion are stable over many generations after transient overexpression [10]. In order to measure the impact of deleting the polyN domain on prion formation, we measured resistance to radicicol of strains that experienced previous overexpression of *AZF1*. Consistent with previous reports, overexpression of WT improves growth in radicicol compared to cells that have never experienced overexpression (Fig 5B). Unexpectedly, overexpression of ΔN also confers increased resistance to radicicol compared to naïve cells (Fig 5C). Yeast prions have been referred to as a “bet-hedging” strategy due to their ability to confer beneficial phenotypes under some conditions and be deleterious in others [40]. We find that

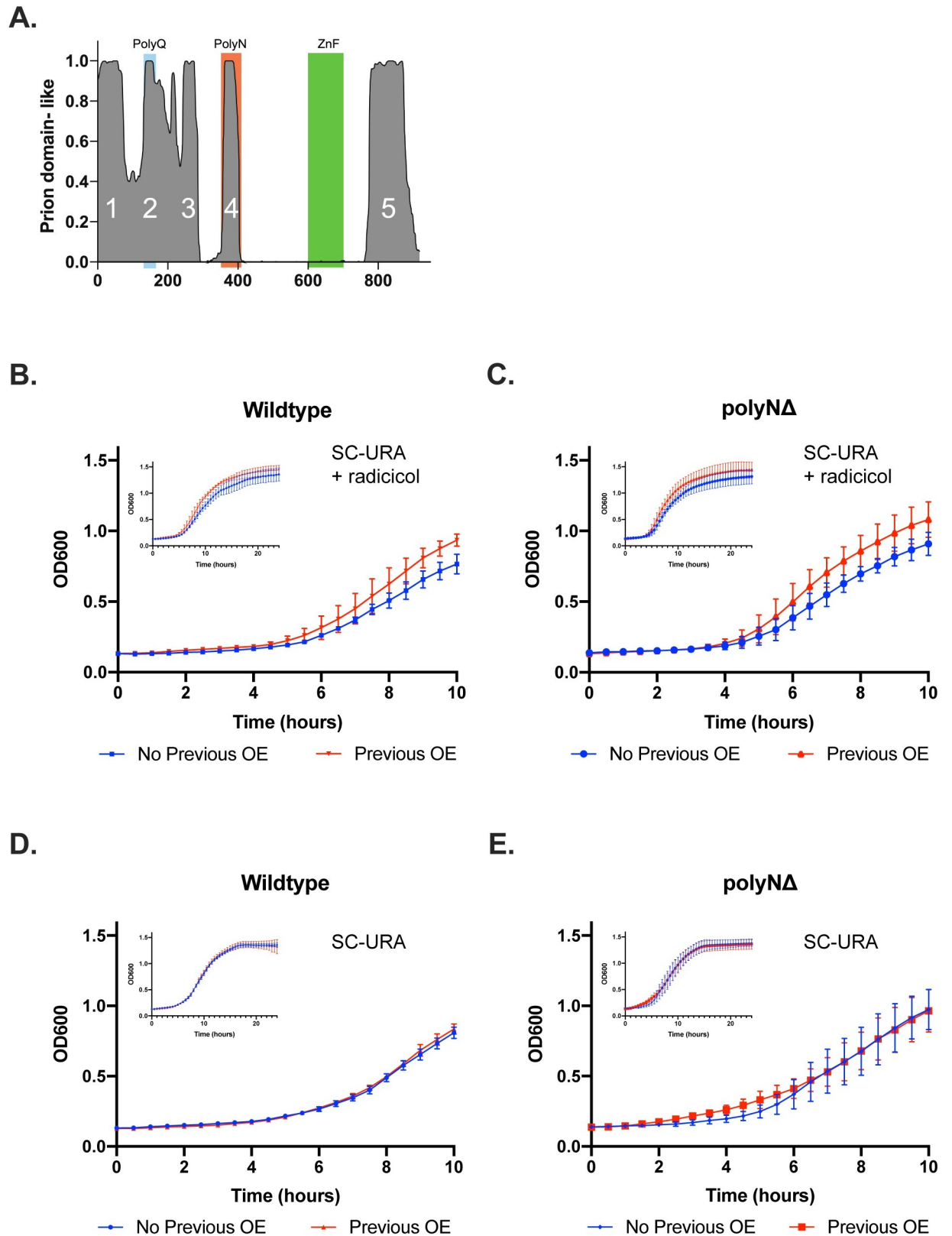


Fig 5. Role of the polyN domain in prion formation and activity. A) Azf1p domains predicted to be capable of promoting prion formation by PLAAC. The polyQ and polyN domains of Azf1p are both predicted to have prion-like properties. B) Growth of WT cells that did (red) and did

not (blue) experience previous overexpression (OE) of *AZF1* in SC-Ura + 75 μ M radicicol measured by OD₆₀₀ in a plate reader. C) Growth of ΔN cells that did (red) and did not (blue) experience previous overexpression of ΔN in SC-Ura + 75 μ M radicicol measured by OD₆₀₀ in a plate reader. Previous overexpression of both WT and ΔN confer resistance to radicicol. D) Growth of WT cells in SC-Ura with and without previous overexpression of *AZF1* measured by OD₆₀₀ in a plate reader. E) Growth ΔN cells in SC-Ura with and without previous overexpression of ΔN measured by OD₆₀₀ in a plate reader. Overexpression of *AZF1* does not impact growth under normal conditions. All growth curves are the average of three biological replicates. The first 10 hours are highlighted to show the differences in growth more clearly and due to the reduced accuracy of the plate reader at high optical densities.

<https://doi.org/10.1371/journal.pone.0247285.g005>

transient overexpression of both WT and ΔN do not impact growth in the absence of radicicol (Fig 5D and 5E), indicating that [*AZF1*⁺] is not toxic to cells under normal conditions.

Azf1p exhibits repeat-length variation across strains of yeast

Repeat domains are well known to be genetically unstable, and changes in repeat copy number can impact protein function. This variation can enable organisms to adapt to different environmental conditions and drive evolution [30,41–46]. We therefore examined the homorepeat domains of Azf1p for length variation. Through a bioinformatic analysis of the genomes of 93 lab and wild isolates of *S. cerevisiae* [38,47], we determined that the polyN and polyQ domains of Azf1p vary in repeat copy number across strains of yeast. The polyN domain has 8 alleles ranging from 17 to 25 N residues (Fig 6A), and the polyQ domain has 3 alleles, 9Q, 12Q, and 17Q (Fig 6B).

Azf1p is a prion-forming yeast TF that contains a polyQ and a polyN domain. To begin to elucidate how repetitive regions modulate Azf1p function, in this work we sought to investigate the role of the polyN domain in transcription and [*AZF1*⁺] prion formation. We find that the polyN domain plays a subtle role in transcription but is not required for prion formation. We also report that both the polyQ and polyN domains vary in copy number across strains of yeast.

Discussion

Azf1p has previously been reported to be a transcriptional activator that is important for growth on glycerol and other non-fermentable carbon sources [31]. In this work, we report a new phenotype for *AZF1* during growth in glucose and identify a larger set of genes that are regulated by Azf1p. We monitored growth of WT and *azf1* Δ cells in glucose by optical density and find that the null mutant has a shortened lag phase compared to WT (Fig 1A). To further investigate the role of Azf1p during growth in glucose, we measured differential gene expression in the null mutant compared to WT by RNA-seq (Fig 1B and S5 Table). To our knowledge, *azf1* Δ has not been previously shown to upregulate gene expression. Our data suggests that Azf1p has a repressive function.

Despite being historically dismissed as functionally unimportant, homorepeats, particularly polyQ and polyN, are conserved across eukaryotic TFs and have been shown to play important roles in signaling, binding, and protein solubility. N-rich and Q-rich sequences are also characteristic of the prion-forming domains of yeast prions, and many repeat-containing TFs form prions. Azf1p contains a polyQ and a polyN repeat domain, and we sought to investigate the roles of these domains in transcription and prion activity. We attempted to delete each of these domains in the *AZF1-GFP* strain from the yeast GFP collection [32,33]. While deletion of the polyN domain was successful (Fig 2A), the polyQ domain could not be deleted using multiple strategies. This predicts a fundamental role for the polyQ domain of Azf1p in yeast but requires significant additional exploration that is beyond the scope of this manuscript. Here we continued the interrogation of the polyN domain as they are thought to serve similar functions to polyQ domains but are not nearly as well studied.

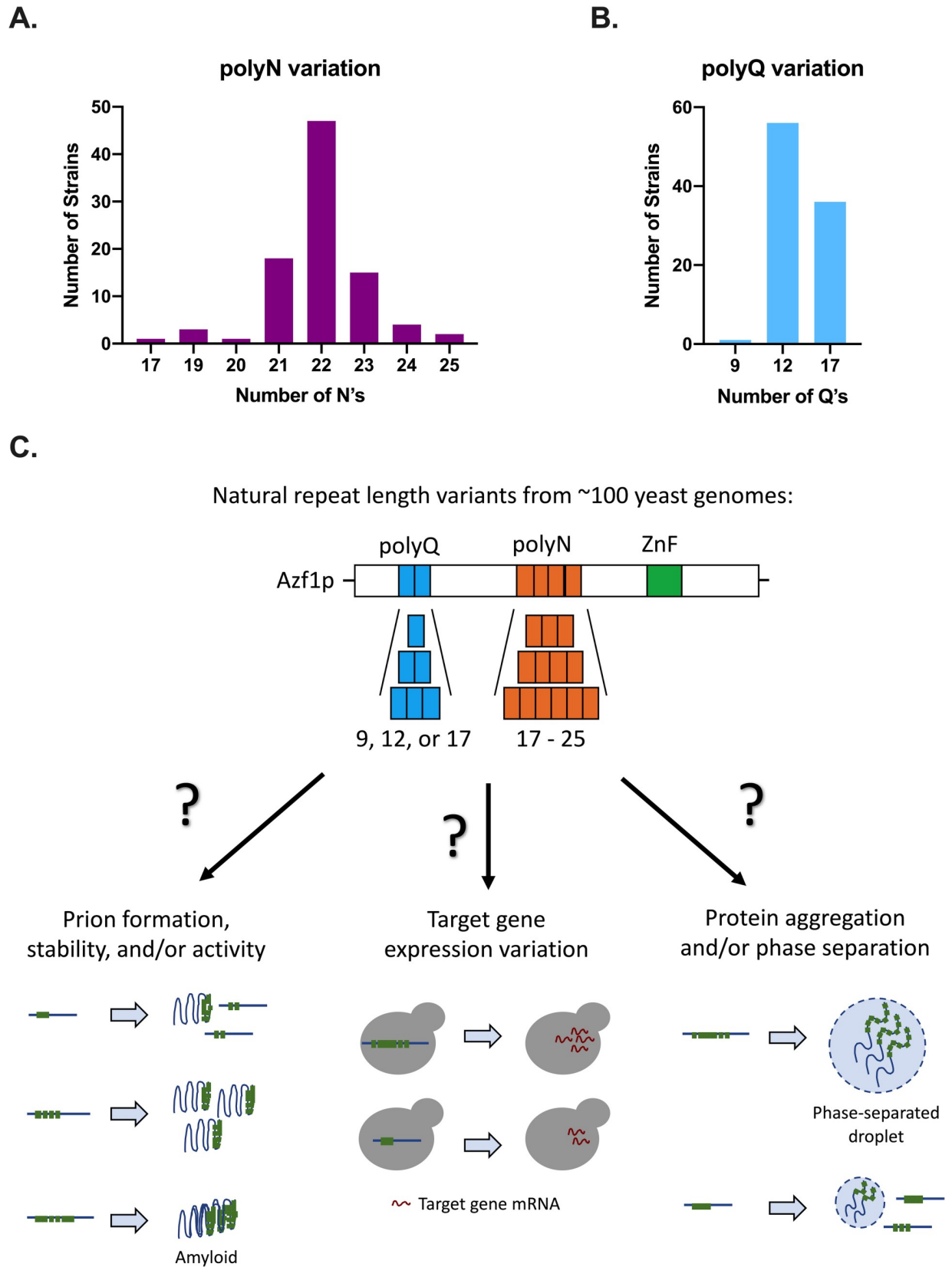


Fig 6. Repeat variation in the polyQ and polyN domains of Azf1p may impact protein function. A) The distribution of polyN domain variants across 93 strains of *S. cerevisiae*. The polyN domain has a range of alleles from 17 to 25 N residues. B) The distribution of polyQ domain variants across 93 strains of *S. cerevisiae*. The polyQ domain has 3 alleles: 9Q, 12Q, and 17Q. C) Model for ways that repeat copy number variation in Azf1p may impact protein function. Repeat variation may impact prion formation and/or stability, target gene expression, or protein aggregation and/or phase separation.

<https://doi.org/10.1371/journal.pone.0247285.g006>

Due to the known involvement of a homorepeat domain in the repressive activity of Cyc8p, we hypothesized that the polyN domain may be important for this function. In order to address this hypothesis, we first examined the role of the polyN domain in growth in different carbon sources. We find that deletion of the polyN domain produces a mild growth defect on glycerol and moderately reduces the length of lag phase during growth in glucose compared to WT (Fig 2B and 2C). Because each of these phenotypes is less pronounced than the corresponding phenotype observed in the null mutant, we conclude that the polyN domain has a minor but not essential role in Azf1p function. In order to verify that the shortened lag phase observed in glucose is due to the deletion of the polyN domain, we attempted to rescue this phenotype by expressing a wildtype copy of *AZF1* from the pBY011-*AZF1* plasmid. While this plasmid encodes *AZF1* under the control of a galactose-inducible promoter, we have determined by qPCR that *AZF1* is expressed from this plasmid at a low level in glucose. In the WT strain, we compared the relative expression of *AZF1* in cells transformed with pBY011-*AZF1* to cells transformed with pBY011-GFP, which encodes GFP under the galactose-inducible promoter instead of *AZF1*. We find that during growth in glucose, pBY011-*AZF1* results in significantly increased expression of *AZF1* (S3B Fig). Even with this low level of leaky expression, we find that the plasmid copy of wildtype *AZF1* goes beyond rescuing the shortened lag phase phenotype and produces a growth defect in ΔN cells compared to WT (S3A Fig). While we recognize that rescuing the observed ΔN growth phenotypes with a wildtype copy of *AZF1* under the native promoter would be more appropriate, the constraints of the COVID-19 pandemic prevented us from performing this experiment.

In order to better understand the observed growth phenotypes, we compared gene expression levels in ΔN to WT as measured by RNA-seq. Consistent with our phenotypic growth data, we find that deletion of the polyN domain impacts gene expression in both glucose (Fig 3A) and glycerol (Fig 3B), suggesting that this domain is involved in Azf1p's TF activity. During growth in both carbon sources, more genes are upregulated than downregulated upon deletion of the polyN domain. This result is consistent with our hypothesis that the polyN domain plays a role in Azf1p's repressive function.

In order to explore how the changes in gene expression measured by RNA-seq give rise to the observed growth patterns in each carbon source, we performed GO Enrichment Analysis on genes differentially expressed between ΔN and WT. In glycerol, genes upregulated by ΔN are enriched in gluconeogenesis (17-fold enrichment), glycolysis (15-fold enrichment), and cell wall organization (3-fold enrichment) (Fig 3D). These processes are consistent with the those previously reported to be regulated by Azf1p [31]. Genes upregulated by ΔN during growth in glucose fit into two main categories: environmental stress tolerance and carbon metabolism and storage (Fig 3C). Upregulation of carbon metabolism genes may underlie the shortened lag phase observed in ΔN , as cells that can import and utilize glucose more rapidly or efficiently may be able to enter the log phase of growth more quickly. This result may have implications for fermentation applications and small molecule production. Yeast are used in the production of a variety of foods and drinks, as well as natural and non-natural small molecules for pharmaceuticals, fuels, and even perfumes [48–50]. Metabolic engineering is a crucial technology geared towards improving the efficiency and output of these processes, and the new phenotypes we report for a null mutant and a ΔN mutant indicate that Azf1p may be a useful target.

We hypothesized that the transcriptional changes observed in ΔN may result from altered Azf1p localization. During both log and stationary phase growth, we find that ΔN exhibits the same localization patterns as WT regardless of the carbon source (Fig 4). Deletion of the polyN domain therefore does not appear to impact Azf1p localization. This observation is consistent with the result that deletion of the polyN domain impacts the expression of only a subset of Azf1p's target genes, as altered localization would likely affect all of Azf1p's targets. The polyN domain may instead act as a binding domain with transcriptional cofactors. This hypothesis is supported by the different growth and gene expression patterns observed between ΔN and *azf1 Δ* . Different target genes have different combinations of transcriptional regulators. The polyN domain may be important for binding with only a subset of Azf1p's cofactors, causing deletion of this domain to impact expression of only a subset of Azf1p's target genes.

In a large-scale screen for yeast prions conducted by Chakrabortee and colleagues, Azf1p was found to form a prion, [AZF1⁺], that confers resistance to the drug radicicol. Overexpression of the constituent protein drives prion formation, and phenotypes conferred by the prion are stable over many generations after transient overexpression [10]. The PrD of Azf1p has not been characterized experimentally, but PLAAC predicts that the polyN domain is one of several sequences capable of promoting prion formation [39] (Fig 5A). In order to determine if the polyN domain plays a role in [AZF1⁺] prion formation or activity, we monitored growth in radicicol of WT and ΔN cells that previously overexpressed WT and ΔN , respectively. Consistent with Chakrabortee and colleagues' results, transient overexpression of WT improves growth in radicicol compared to cells that have never experienced overexpression (Fig 5B). Surprisingly, overexpression of ΔN also confers resistance to radicicol compared to naïve cells (Fig 5C), suggesting that the polyN domain is dispensable for prion formation. The polyQ domain is within a large N-terminal IDR predicted by PLAAC to be PrD-like [39] (Fig 5A). This domain may therefore be a more relevant candidate for future studies aimed at determining the PrD of Azf1p.

Finally, we report that the polyN and polyQ domains of Azf1p vary in repeat copy number across strains of yeast. The polyN domain has 8 alleles ranging from 17 to 25 N residues (Fig 6A), and the polyQ domain has 3 alleles, 9Q, 12Q, and 17Q (Fig 6B). We propose a model outlining several ways that repeat-length variation may impact Azf1p function (Fig 6C). First, repeat variation could alter prion formation and/or stability. In yeast, expanded repeats in PrDs have been shown to increase prion formation [51–53]. While we find that the polyN domain does not play a significant role in prion formation, PLAAC predicts that the polyQ domain is PrD-like (Fig 5A). If the polyQ domain is in fact important for prionogenesis, short polyQ alleles may reduce prion formation and/or stability of the prion conformer, while long alleles or abnormally expanded repeats could result in the formation of an irreversible amyloid fiber.

Repeat variation could also impact expression of Azf1p's target genes. Another prion-forming yeast TF, Cyc8p, has a naturally variable polyQ domain. Variation in this domain impacts Cyc8p solubility and interaction with its binding partners, which in turn alters target gene expression and the associated phenotypes [30]. We find that the polyN domain of Azf1p plays a subtle role in transcription, possibly by mediating binding to transcriptional cofactors. The impact of repeat variation in the polyN domain of Azf1p, as well as the variable repeat domains of many other TFs across eukaryotes, on target gene expression warrants further study. The gene expression and resulting metabolic changes we observed in ΔN may have implications in fermentation applications and small molecule production. Deletion of the polyN domain moderately shortens the lag phase in glucose (Fig 1C), a phenotype that can potentially be explained by the upregulation of genes involved in carbon metabolism and storage (Fig 3C). Repeat variation may serve as a mechanism to further tune this phenotype by altering gene expression.

One possible mechanism by which repeat variation could impact Azf1p's function in transcription is phase separation. Several eukaryotic TFs have been shown to phase separate with Mediator via their IDRs, and the formation of these droplets was demonstrated to be required for transcription [54,55]. The polyQ and polyN domains of Azf1p are both intrinsically disordered [47], and length variation within these domains could impact the extent to which Azf1p can be incorporated into phase separated droplets.

Repetitive domains were historically dismissed as functionally unimportant due to their lack of three-dimensional structure. These domains are now known to be important for signaling, binding, prion formation, and phase separation. Homorepeats are particularly common in eukaryotic TFs and often exhibit length variation across organisms. The function of these domains and the impact of repeat variation, however, has only been minimally explored. We report a new role for the variable polyN domain of the yeast TF Azf1p in transcription and highlight repeat copy number variation as a possible driver of target gene expression variance in this and other eukaryotic TFs. We also report metabolic changes that may be useful for fermentation applications and small molecule production.

Supporting information

S1 Fig. GO Enrichment Analysis of genes differentially expressed in *azf1Δ* during growth in glucose. A) GO Enrichment Analysis of genes that are statistically significantly upregulated by *azf1Δ* ($p < 0.001$) compared to WT as measured by RNA-seq with RNA isolated from cells grown in glucose. B) GO Enrichment Analysis of genes that are statistically significantly downregulated by *azf1Δ* compared to WT ($p < 0.001$) as measured by RNA-seq using RNA isolated from cells grown in glucose.

(TIF)

S2 Fig. Relative gene expression analysis of Azf1p target genes by qPCR. A) Fold change in expression of Azf1p target genes in *AZF1ΔN* and *azf1Δ* compared to WT during growth in glucose. B) Fold change in expression of Azf1p target genes in *AZF1ΔN* compared to WT during growth in glycerol. Error bars represent standard deviation. Statistical significance was calculated using an unpaired student's t test. * indicates $p < 0.05$ and ** indicates $p < 0.001$.

(TIF)

S3 Fig. Expression of wildtype *AZF1* abolishes the shortened lag phase observed in ΔN cells during growth in glucose. A) Growth of WT (blue) and ΔN (red) cells with glucose as the carbon source. Growth was measured by OD_{600} in a plate reader for 24 hours. Growth curves are the average of three biological replicates, and error bars indicate standard deviation. B) Fold change in expression of *AZF1* in WT cells expressing pBY011-*AZF1* compared to pBY011-GFP during growth in glucose. Error bars represent standard deviation. Statistical significance was calculated using an unpaired student's t test ($p = 0.0028$).

(TIF)

S1 Table. Plasmids used in this work.

(PDF)

S2 Table. Primers used in this work.

(PDF)

S3 Table. Yeast strains used in this work.

(PDF)

S4 Table. Fold change in *AZF1* expression in 3 biological replicates of *azf1Δ* compared to WT.

(PDF)

S5 Table. Differentially expressed genes in *azf1Δ* compared to WT during growth in glucose.

(CSV)

S6 Table. Differentially expressed genes in ΔN compared to WT during growth in glucose.

(CSV)

S7 Table. Differentially expressed genes in ΔN compared to WT during growth in glycerol.

(CSV)

Acknowledgments

The authors would like to acknowledge Mackenzie Parmenter and Mike Babokhov for early contributions to this project. We thank M. McVey, K.A. McElroy, and members of the McVey Lab and Tufts Biology department for their helpful discussions during the preparation of this manuscript.

Author Contributions

Conceptualization: Taylor Stewart, Stephen M. Fuchs.

Data curation: Taylor Stewart, Benjamin E. Wolfe.

Funding acquisition: Stephen M. Fuchs.

Investigation: Taylor Stewart, Stephen M. Fuchs.

Methodology: Taylor Stewart, Benjamin E. Wolfe, Stephen M. Fuchs.

Project administration: Stephen M. Fuchs.

Resources: Stephen M. Fuchs.

Supervision: Stephen M. Fuchs.

Writing – original draft: Taylor Stewart, Stephen M. Fuchs.

Writing – review & editing: Taylor Stewart, Benjamin E. Wolfe, Stephen M. Fuchs.

References

1. Van Der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, et al. Classification of intrinsically disordered regions and proteins. *Chemical Reviews*. 2014. pp. 6589–6631. <https://doi.org/10.1021/cr400525m> PMID: 24773235
2. Gemayel R, Vincens MD, Legendre M, Verstrepen KJ. Variable Tandem Repeats Accelerate Evolution of Coding and Regulatory Sequences. *Annu Rev Genet*. 2010; 44: 445–477. <https://doi.org/10.1146/annurev-genet-072610-155046> PMID: 20809801
3. Faux NG, Bottomley SP, Lesk AM, Irving JA, Morrison JR, De La Banda MG, et al. Functional insights from the distribution and role of homopeptide repeat-containing proteins. *Genome Res*. 2005; 15: 537–551. <https://doi.org/10.1101/gr.3096505> PMID: 15805494
4. Lu X, Murphy RM. Synthesis and disaggregation of asparagine repeat-containing peptides. *J Pept Sci*. 2014; 20: 860–867. <https://doi.org/10.1002/psc.2677> PMID: 25044797
5. Peters TW, Huang M. Protein aggregation and polyasparagine-mediated cellular toxicity in *Saccharomyces cerevisiae*. *Prion*. 2007; 1: 144–153. <https://doi.org/10.4161/pri.1.2.4630> PMID: 19164913

6. Kolaczowska A, Kolaczowski M, Delahodde A, Goffeau A. Functional dissection of Pdr1p, a regulator of multidrug resistance in *Saccharomyces cerevisiae*. *Mol Genet Genomics*. 2002; 267: 96–106. <https://doi.org/10.1007/s00438-002-0642-0> PMID: 11919720
7. Cascarina SM, Ross ED. Yeast prions and human prion-like proteins: Sequence features and prediction methods. *Cell Mol Life Sci*. 2014; 71: 2047–2063. <https://doi.org/10.1007/s00018-013-1543-6> PMID: 24390581
8. Alberti S, Halfmann R, King O, Kapila A, Lindquist S. A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins. *Cell*. 2009; 137: 146–158. <https://doi.org/10.1016/j.cell.2009.02.044> PMID: 19345193
9. Halfmann R, Jarosz DF, Jones SK, Chang A, Lancaster AK, Lindquist S. Prions are a common mechanism for phenotypic inheritance in wild yeasts. *Nature*. 2012; 482: 363–368. <https://doi.org/10.1038/nature10875> PMID: 22337056
10. Chakrabortee S, Byers JS, Jones S, Garcia DM, Bhullar B, Chang A, et al. Intrinsically Disordered Proteins Drive Emergence and Inheritance of Biological Traits. *Cell*. 2016; 167: 369–381.e12. <https://doi.org/10.1016/j.cell.2016.09.017> PMID: 27693355
11. Itakura AK, Chakravarty AK, Jakobson CM, Jarosz DF. Widespread Prion-Based Control of Growth and Differentiation Strategies in *Saccharomyces cerevisiae*. *Mol Cell*. 2020; 77: 266–278. <https://doi.org/10.1016/j.molcel.2019.10.027> PMID: 31757756
12. Collinge J. Mammalian prions and their wider relevance in neurodegenerative diseases. *Nature*. 2016; 539: 217–226. <https://doi.org/10.1038/nature20415> PMID: 27830781
13. Prusiner SB. Novel proteinaceous infectious particles cause scrapie. *Science* (80-). 1982; 216: 136–144. <https://doi.org/10.1126/science.6801762> PMID: 6801762
14. Murakami T, Qamar S, Lin JQ, Schierle GSK, Rees E, Miyashita A, et al. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron*. 2015; 88: 670–690. <https://doi.org/10.1016/j.neuron.2015.10.030> PMID: 26526393
15. Patel A, Lee HO, Jawerth L, Maharana S, Jahnel M, Hein MY, et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*. 2015; 162: 1066–1077. <https://doi.org/10.1016/j.cell.2015.07.047> PMID: 26317470
16. Harrison AF, Shorter J. RNA-binding proteins with prion-like domains in health and disease. *Biochem J*. 2017; 474: 1417–1438. <https://doi.org/10.1042/BCJ20160499> PMID: 28389532
17. Paul KR, Mollieux A, Cascarina S, Boncella AE, Taylor JP, Ross ED. Effects of Mutations on the Aggregation Propensity of the Human Prion-Like Protein hnRNPA2B1. *Mol Cell Biol*. 2017; 37: e00652–16. <https://doi.org/10.1128/MCB.00652-16> PMID: 28137911
18. Guo L, Kim HJ, Wang H, Monaghan J, Freyermuth F, Sung JC, et al. Nuclear-Import Receptors Reverse Aberrant Phase Transitions of RNA-Binding Proteins with Prion-like Domains. *Cell*. 2018; 173: 677–692. <https://doi.org/10.1016/j.cell.2018.03.002> PMID: 29677512
19. Tuite MF, Serio TR. The prion hypothesis: From biological anomaly to basic regulatory mechanism. *Nat Rev Mol Cell Biol*. 2010; 11: 823–833. <https://doi.org/10.1038/nrm3007> PMID: 21081963
20. Newby GA, Lindquist S. Blessings in disguise: Biological benefits of prion-like mechanisms. *Trends Cell Biol*. 2013; 283: 251–259. <https://doi.org/10.1016/j.tcb.2013.01.007> PMID: 23485338
21. Chernova TA, Chernoff YO, Wilkinson KD. Prion-based memory of heat stress in yeast. *Prion*. 2017; 11: 151–161. <https://doi.org/10.1080/19336896.2017.1328342> PMID: 28521568
22. Harvey ZH, Chen Y, Jarosz DF. Protein-Based Inheritance: Epigenetics beyond the Chromosome. *Molecular Cell*. 2018. pp. 195–202. <https://doi.org/10.1016/j.molcel.2017.10.030> PMID: 29153393
23. Jarosz DF, Brown JCS, Walker GA, Datta MS, Ung WL, Lancaster AK, et al. Cross-kingdom chemical communication drives a heritable, mutually beneficial prion-based transformation of metabolism. *Cell*. 2014; 158: 1083–1093. <https://doi.org/10.1016/j.cell.2014.07.025> PMID: 25171409
24. Jarosz DF, Lancaster AK, Brown JCS, Lindquist S. An evolutionarily conserved prion-like element converts wild fungi from metabolic specialists to generalists. *Cell*. 2014; 158: 1072–1082. <https://doi.org/10.1016/j.cell.2014.07.024> PMID: 25171408
25. Halfmann R, Lindquist S. Epigenetics in the extreme: Prions and the inheritance of environmentally acquired traits. *Science* (80-). 2010; 330: 629–632. <https://doi.org/10.1126/science.1191081> PMID: 21030648
26. March ZM, King OD, Shorter J. Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain Res*. 2016; 1647: 9–18. <https://doi.org/10.1016/j.brainres.2016.02.037> PMID: 26996412

27. Holmes DL, Lancaster AK, Lindquist S, Halfmann R. Heritable remodeling of yeast multicellularity by an environmentally responsive prion. *Cell*. 2013; 153: 153–165. <https://doi.org/10.1016/j.cell.2013.02.026> PMID: 23540696
28. Chakravarty AK, Smejkal T, Itakura AK, Garcia DM, Jarosz DF. A Non-amyloid Prion Particle that Activates a Heritable Gene Expression Program. *Mol Cell*. 2020; 77: 251–265. <https://doi.org/10.1016/j.molcel.2019.10.028> PMID: 31757755
29. Franzmann TM, Alberti S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J Biol Chem*. 2019; 294: 7128–7136. <https://doi.org/10.1074/jbc.TM118.001190> PMID: 29921587
30. Gemayel R, Chavali S, Pougach K, Legendre M, Zhu B, Boeynaems S, et al. Variable Glutamine-Rich Repeats Modulate Transcription Factor Activity. *Mol Cell*. 2015; 59: 615–627. <https://doi.org/10.1016/j.molcel.2015.07.003> PMID: 26257283
31. Slattery MG, Liko D, Heideman W. The Function and Properties of the Azf1 Transcriptional Regulator Change with Growth Conditions in *Saccharomyces cerevisiae*. *Eukaryot Cell*. 2006; 5: 313–320. <https://doi.org/10.1128/EC.5.2.313-320.2006> PMID: 16467472
32. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, et al. Global analysis of protein localization in budding yeast. *Nature*. 2003. <https://doi.org/10.1038/nature02026> PMID: 14562095
33. Howson R, Huh WK, Ghaemmaghami S, Falvo JV., Bower K, Belle A, et al. Construction, verification and experimental use of two epitope-tagged collections of budding yeast strains. *Comp Funct Genomics*. 2005; 6: 2–16. <https://doi.org/10.1002/cfg.449> PMID: 18629296
34. Laughery MF, Hunter T, Brown A, Hoopes J, Ostbye T, Shumaker T, et al. New vectors for simple and streamlined CRISPR-Cas9 genome editing in *Saccharomyces cerevisiae*. *Yeast*. 2015; 32: 711–720. <https://doi.org/10.1002/yea.3098> PMID: 26305040
35. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019; 47: D419–D426. <https://doi.org/10.1093/nar/gky1038> PMID: 30407594
36. Carbon S, Douglass E, Dunn N, Good B, Harris NL, Lewis SE, et al. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res*. 2019; 47: D330–D338. <https://doi.org/10.1093/nar/gky1055> PMID: 30395331
37. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: Tool for the unification of biology. *Nat Genet*. 2000; 25: 25–29. <https://doi.org/10.1038/75556> PMID: 10802651
38. Strobe PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Magwene PM, et al. The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Res*. 2015; 25: 762–774. <https://doi.org/10.1101/gr.185538.114> PMID: 25840857
39. Lancaster AK, Nutter-Upham A, Lindquist S, King OD. PLAAC: A web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics*. 2014; 30: 2501–2502. <https://doi.org/10.1093/bioinformatics/btu310> PMID: 24825614
40. Halfmann R, Alberti S, Lindquist S. Prions, protein homeostasis, and phenotypic diversity. *Trends Cell Biol*. 2010; 20: 125–133. <https://doi.org/10.1016/j.tcb.2009.12.003> PMID: 20071174
41. Hoyer LL. The ALS gene family of *Candida albicans*. *Trends Microbiol*. 2001; 9: 176–180. [https://doi.org/10.1016/s0966-842x\(01\)01984-9](https://doi.org/10.1016/s0966-842x(01)01984-9) PMID: 11286882
42. Fidalgo M, Barrales RR, Jimenez J. Coding repeat instability in the FLO11 gene of *Saccharomyces yeasts*. *Yeast*. 2008; 25: 879–889. <https://doi.org/10.1002/yea.1642> PMID: 19160455
43. Press MO, Queitsch C. Variability in a Short Tandem Repeat Mediates Complex Epistatic Interactions in *Arabidopsis thaliana*. *Genetics*. 2017; 205: 455–464. <https://doi.org/10.1534/genetics.116.193359> PMID: 27866166
44. Sawyer LA, Hennessy JM, Peixoto AA, Rosato E, Parkinson H, Costa R, et al. Natural variation in a *Drosophila* clock gene and temperature compensation. *Science (80-)*. 1997; 278: 2117–2120. <https://doi.org/10.1126/science.278.5346.2117> PMID: 9405346
45. Ebisawa T, Uchiyama M, Kajimura N, Mishima K, Kamei Y, Katoh M, et al. Association of structural polymorphisms in the human period3 gene with delayed sleep phase syndrome. *EMBO Rep*. 2001; 2: 342–346. <https://doi.org/10.1093/embo-reports/kve070> PMID: 11306557
46. Johnsen A, Fidler AE, Kuhn S, Carter KL, Hoffmann A, Barr IR, et al. Avian Clock gene polymorphism: Evidence for a latitudinal cline in allele frequencies. *Mol Ecol*. 2007; 16: 4867–4880. <https://doi.org/10.1111/j.1365-294X.2007.03552.x> PMID: 17927702
47. Babokhov M, Reinfeld BI, Hackbarth K, Bentov Y, Fuchs SM. Tandem repeats drive variation of intrinsically disordered regions in budding yeast. *bioRxiv*. 2018. <https://doi.org/10.1101/339663>

48. Lee JW, Na D, Park JM, Lee J, Choi S, Lee SY. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat Chem Biol.* 2012; 8: 536–546. <https://doi.org/10.1038/nchembio.970> PMID: 22596205
49. Zhang Y, Nielsen J, Liu Z. Engineering yeast metabolism for production of terpenoids for use as perfume ingredients, pharmaceuticals and biofuels. *FEMS Yeast Res.* 2017;17. <https://doi.org/10.1093/femsyr/fox080> PMID: 29096021
50. Chen R, Yang S, Zhang L, Zhou YJ. Advanced Strategies for Production of Natural Products in Yeast. *iScience.* 2020; 23: 100879. <https://doi.org/10.1016/j.isci.2020.100879> PMID: 32087574
51. Tank EMH, Harris DA, Desai AA, True HL. Prion Protein Repeat Expansion Results in Increased Aggregation and Reveals Phenotypic Variability. *Mol Cell Biol.* 2007; 27: 5445–5455. <https://doi.org/10.1128/MCB.02127-06> PMID: 17548473
52. Liu JJ, Lindquist S. Oligopeptide-repeat expansions modulate “protein-only” inheritance in yeast. *Nature.* 1999; 400: 576. <https://doi.org/10.1038/23054> PMID: 10448861
53. Parham SN, Resende CG, Tuite MF. Oligopeptide repeats in the yeast protein Sup35p stabilize intermolecular prion interactions. *EMBO J.* 2001; 20: 2111–2119. <https://doi.org/10.1093/emboj/20.9.2111> PMID: 11331577
54. Sabari BR, Dall’Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science (80-).* 2018; 361: eaar3958. <https://doi.org/10.1126/science.aar3958> PMID: 29930091
55. Boija A, Klein IA, Sabari BR, Dall’Agnese A, Coffey EL, Zamudio AV., et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell.* 2018; 175: 1842–1855. <https://doi.org/10.1016/j.cell.2018.10.042> PMID: 30449618