Recurrent rewiring and emergence of RNA regulatory networks

Daniel Wilinski, Natascha Buter, Andrew D. Klocko, Christopher P. Lapointe, Eric U. Selker, Audrey P. Gasch, and Marvin Wickens

*Department of Biochemistry, University of Wisconsin–Madison, Madison, WI 53706; †Institute of Molecular Biology, University of Oregon, Eugene, OR 97403; and ‡Laboratory of Genetics, University of Wisconsin–Madison, Madison, WI 53706

Edited by Robert H. Singer, Albert Einstein College of Medicine, Bronx, NY, and approved February 14, 2017 (received for review October 26, 2016)

Alterations in regulatory networks contribute to evolutionary change. Transcriptional networks are reconfigured by changes in the binding specificity of transcription factors and their cognate sites. The evolution of RNA–protein regulatory networks is far less understood. The PUF (Pumilio and FBF) family of RNA regulatory proteins controls the translation, stability, and movements of hundreds of mRNAs in a single species. We probe the evolution of PUF–RNA networks by direct identification of the mRNAs bound to PUF proteins in budding and filamentous fungi and by computational analyses of orthologous RNAs from 62 fungal species. Our findings reveal that PUF proteins gain and lose mRNAs with related and emergent biological functions during evolution. We demonstrate at least two independent rewiring events for PUF3 orthologs, independent but convergent evolution of PUF4/5 binding specificity and the rewiring of the PUF4/5 regulons in different fungal lineages. These findings demonstrate plasticity in RNA regulatory networks and suggest ways in which their rewiring occurs.

RNA regulation | PUF proteins | evolution | 3′ UTR elements

Coordinated regulation of genes and mRNAs pervades biology. Cells simplify the challenge by coregulating groups of functionally related genes via the same regulatory proteins. In this way, cells establish “networks” consisting of a site-specific regulatory protein, the downstream targets it controls, and the recognition sequences in DNA or RNA that link each target to its regulator. The discrimination of specific sequences in DNA and RNA is a cornerstone of biological regulation in living systems, as these binding sites determine which genes will be controlled.

Transcriptional networks display a remarkable capacity to evolve, as evidenced by comparative genomic studies. Networks can change at the individual gene level, in which specific transcriptional targets are gained or lost via changes in transcription factor binding sites within genomic DNA (1–4). Over extended periods of evolutionary time, the entire set of targets of a regulatory protein can change, until the network adapts and effectively controls a distinct set of genes. Alternatively, coregulated genes can fall under the control of different regulators. In this case, each individual target acquires DNA sites that are recognized by the alternate regulator. Ribosomal protein genes in fungi are a striking example: Their coregulation has been maintained for over a billion years of fungal evolution despite regulation by completely different systems in different species (5–7). This type of wholesale rewiring requires concerted evolution across the suite of targets. Although the mechanisms through which transcriptional networks evolve are unclear, it has been proposed to involve a period of “redundant” regulation, during which the same group of targets is temporarily controlled by two different regulatory systems, followed by loss of the ancestral connections (6, 8, 9). Rewiring often correlates with duplication and divergence of the responsible regulator (10–13).

The evolution of RNA regulatory networks is less well understood. A single RNA-binding protein (RBP) often controls groups of RNAs with common biological functions (14–17) and so governs their stability, translation, and localization. RBPs in the same protein family exhibit similar though not identical RNA-binding specificities (18, 19). Differences in family members generate divergence in their regulatory networks. 3′UTRs, like introns, are particularly favorable targets for network divergence, as they are relatively unconstrained in length and sequence. Alternative polyadenylation events, in which a single pre-mRNA species can give rise to two or more mRNAs that either contain or lack regulatory sites, provide a mechanism of network divergence unique to mRNAs (20–22). Other molecular mechanisms for the evolution of binding elements and mRNA–protein regulatory networks exist but are largely opaque.

Here we focus on the evolution of RNA–protein networks. We concentrate on the PUF (Pumilio and FBF) family of RBPs, because they are exemplary regulators of cytoplasmic mRNAs and provide a strong foundation based on genetics, biochemistry, and structural biology. PUF proteins generally bind single-stranded RNA elements in the 3′ UTR to trigger mRNA repression, activation, or decay and can influence localization as well (23–27). PUF family members are present in all eukaryotes and maintain stem cells and promote memory formation in metazoa (28, 29). RNA targets for the canonical Saccharomyces cerevisiae PUF proteins have been identified by immunoprecipitation methods, including RIP-chip, HITS-CLIP, and RNA tagging (14, 30–32). Each PUF protein binds a group of functionally related mRNAs, forming distinct networks that are each enriched for specific cellular processes. For example, of the 476 tagged RNAs bound to S. cerevisiae Puf5, 191 encode proteins that contribute to mitochondrial organization (32).

PUF proteins bind different RNAs due to variations in a conserved protein scaffold. The proteins consist of eight largely helical repeats (PUF repeats). Three amino acids within each repeat can potentially contact a single RNA base (23, 33). The sequence of the 5′-end of most PUF binding elements is UGUR (where R is U or C), and the 3′-end is AU or UA. Differences in RNA binding specificity among PUF proteins are due to changes in the curvature of a conserved, largely helical PUF protein scaffold, which imposes different limitations on the RNAs that can bind with high affinity.

*Significance*

Regulatory networks change during evolution. A protein that controls many genes in one species may control a different set of genes in another. We examined how mRNAs evolve, focusing on the PUF (Pumilio and FBF) family of RNA-binding proteins. These govern stability and translation of hundreds of mRNAs and enable coordinate regulation of discrete biological outcomes. To understand how RNA networks evolve, we used knowledge of the RNA specificity of each PUF protein to predict its mRNA targets and directly identified mRNAs bound to each protein in divergent fungi via biochemical methods. We find networks controlled by one protein switch during evolution to be controlled by another and that proteins with different specificities can share, gain, or lose batteries of mRNAs.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

To whom correspondence should be addressed. Email: wickens@biochem.wisc.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1617777114/-/DCSupplemental.
affinity (30). RNA nucleotides in the binding element contort to accommodate the protein scaffold (30, 34, 35). For example, in S. cerevisiae, Puf3 has high specificity for an 8-nucleotide (nt) PUF element, Puf4 instead recognizes a related 9-nt element, and Puf5 binds most tightly to 9-nt and 10-nt elements (14, 30, 34, 36). Each of their binding elements begins with UGUR and ends with UA. Their specificities diverge through the presence or absence of internal bases in the RNA binding element. The identities of these bases generally are unimportant.

Elegant computational studies of mRNAs across fungal species support a model in which the PuF5 regulon, which controls nuclear-encoded mitochondrial transcripts in S. cerevisiae, was rewired during evolution (37). Bioinformatics revealed that many of these mRNAs possessed 9-nt elements in filamentous fungi and 8-nt elements in S. cerevisiae (37, 38). However, PuF-RNA networks have not been identified experimentally in any fungal species besides S. cerevisiae.

To probe the evolution of RNA regulatory networks at a molecular level, we directly identified RNAs bound to multiple PuF proteins in budding and filamentous fungi. Our studies revealed multiple independent rewiring events in the Puf3, Puf4, and Puf5 networks during fungal evolution. We show that the binding specificities of PUF orthologs from the filamentous fungi Neurospora crassa and Aspergillus nidulans are similar to one another as well as to their S. cerevisiae counterparts. We find strong evidence for at least two independent rewiring events for Puf3 orthologs and demonstrate the independent but convergent evolution of PuF4/5 binding specificities and the rewiring of the PuF4/5 networks in different fungal lineages. Our findings demonstrate the plasticity of RNA regulatory networks, how these networks emerge, and ways in which they have been rewired.

**Results**

**Pattern of PUF Binding Elements Suggests Rewiring.** We identified matches to the known binding elements of S. cerevisiae PuF proteins in the 3′ UTRs of orthologous RNAs in 62 species of fungi to investigate the evolution of binding sites across orthologous gene sets (Fig. 1; full Phylogenetic tree in Fig. S1). The phylum represents at least 400 million years of evolution and includes species with strong foundations in genomics and functional information (39). We focused on species from two subphyla of Phylum Ascomycota: Saccharomycotina (budding yeasts) and Pezizomycotina (filamentous fungi) (40) (Fig. 1). We considered in our analysis only mRNAs with a clear ortholog in at least 50% of species—that is, mRNAs that encode orthologous proteins—as defined previously (41, 42).

We identified matches to PuF binding elements in the 3′ UTRs of orthologous mRNAs, comparing against sequence motifs that correspond to the major binding elements of the canonical S. cerevisiae PuF proteins (14). Because the S. cerevisiae PuF proteins bind different lengths of sequence elements, each bearing a UGU at one end and a UA or AU at the other (see the Introduction), we searched separately for matches to the 8-, 9-, and 10-nt PuF position-weight matrices (PWMs) defined by previous biochemical analysis of S. cerevisiae PuF5 (Sc_Puf5), PuF4 (Sc_Puf4), and PuF5 (Sc_Puf5) (14, 30, 34, 36). For each sequence in each 3′ UTR, we determined the log-likelihood ratio of the PWM model compared with a background genomic mononucleotide model (Fig. 1B). We assigned to each 3′ UTR the best log-likelihood ratio identified for each PWM. For each PWM comparison, we then clustered groups of orthologous genes based on the assigned log-likelihood score (Fig. 1 C–E; gene ontology (GO) enrichments for each cluster are found in Table S1).

Striking patterns of conservation and evolution of PuF binding-sequence enrichment emerged across groups of orthologous genes. However, we detected no clear clusters for which enrichment of a single PuF PWM was conserved in all fungi. Many clusters exhibit functional relatedness. A few appear to lack it (e.g., cluster 2 in the Eurotiomycetes (Aspergilli) (Fig. 1C)).

In Saccharomycotina, the 8-nt cluster 1 and the 9-nt cluster 5 have GO term enrichments for nuclear-encoded “mitochondrion” (P value 1.4 E-96) and “ribosome biogenesis” (P value 4.0E-21), respectively, as noted previously (14, 37, 38). Strikingly, the enrichment

---

**Fig. 1.** Conservation of putative PuF binding elements in Ascomycota fungi. (A) Phylogenetic tree of selected fungi (see full tree in Fig. S1). Colored branches represent each subphylum. Pink, Saccharomycotina (budding yeasts); orange, Pezizomycotina (filamentous fungi); blue, Taphrinomycotina (fission yeast), black, other fungi. (B) Representation of 8-, 9-, and 10-nt PWM models used to parameterize the log-likelihood function. Height of base represents probability of a base at each position in the binding element. (C–E) k-means clustering of orthologous transcripts based on log-likelihood scores for putative PuF binding elements in 3′ UTRs for species in A, as shown in the key. Each cluster contains the highest log-likelihood scores for an orthologous 3′ UTR, and each column represents a species (with the phylogenetic tree shown above). Clustering was done independently for each heat map. Only clusters with a GO term P value lower than E-9 were highlighted (except cluster 2). Shown are the log-likelihood scores based on the 8-nt binding elements (C) found in one or more of 4425 orthologous transcripts. (D) The 9-nt binding elements found in 4,898 transcripts. (E) The 10-nt binding elements found in 4,423 transcripts.
for similar-length PUF binding elements was not conserved across other fungi: The enrichment for 8-nt-long sequences in cluster 1 was only seen for budding yeasts, whereas enrichment for 9-nt elements in clusters 5 were observed for sublineages in filamentous fungi. Clusters of conservation restricted to the Pezizomycotina include 9-nt clusters 3 and 4 and 10-nt cluster 6 (Fig. 1D and E). All three of

Fig. 2. Evolution of PUF–mitochondrial RNA network. (A) Enrichment of PUF binding elements in transcripts from the mitochondrial cluster across Ascomycota fungi. Abbreviations used for fungal species are provided in Fig. S5. Ring 1 represents binding element enrichment orthologs in each species. The inner ring is the phylogenetic species tree. Black arrows mark the species used in C. (B) MEME-derived PWMs identified in 3′UTRs of mitochondrial cluster transcripts in each species. (C) Overlap of genes in mitochondrial cluster containing a putative PUF binding element for three species. (D) Relative luminescence values as a proxy for binding affinity for each PUF protein binding to each PUF site as assayed by yeast-three hybrid assays (48).
these clusters were enriched for mitochondrial function, but notably, none was enriched for the 8-nt PUF sequences known to mediate their control in *S. cerevisiae*. Interestingly, 9-nt cluster 3 was enriched for genes for both mitochondrion (3.1E-51) and ribosomal protein (4.1E-91). (A complete list of GO enrichments for each cluster is in Table S1.)

**Predicted Switch Between Budding and Filamentous Fungi.** The 309 transcripts in cluster 1 (Fig. 1C) are highly enriched for 8-nt elements, consistent with their binding to *Sc_PUF3* (14, 32). We refer to these transcripts as the “mitochondrial cluster.” In filamentous fungi, other clusters are enriched for mitochondria-related functions but contain transcripts with putative 9-nt or 10-nt binding elements (Fig. 1D and E, clusters 3, 4, and 6). Our data are consistent with bioinformatic analyses that suggested a “rewiring” event for this group (37, 38). They further suggest widespread evolution of enriched sequences among transcripts in the mitochondrial cluster. We determined the enrichment of each length of putative binding element for transcripts in the mitochondrial cluster across 62 fungi (Fig. 2, ring 1). All of the budding yeasts exhibit significant enrichment of 8-nt sequences in their 3′UTRs but not of 9-nt and 10-nt elements. Conversely, none of the filamentous fungi have significant enrichment for the 8-nt site. Instead, the filamentous fungi exhibit enrichment for putative binding elements 9 or 10 nt long (Fig. 2, ring 1). No other clade showed a significant enrichment of putative binding elements in these orthologs. These results were supported by unbiased motif discovery in the groups’ 3′UTRs (43), which identified PWMs matching the 9- and 10-nt binding elements seen in *A. nidulans* and *N. crassa* transcripts orthologous to the mitochondrial cluster (Fig. 2B). Most of the filamentous fungi transcripts orthologous to the mitochondrial cluster have a PUF-binding element, as is the case in *S. cerevisiae* (Fig. 2C). Thus, putative PUF binding elements in the mitochondrial cluster extend to filamentous fungi but have diverged so that they are of a different length, including 10-nt elements not previously detected.

**Evolution of Putative Binding Elements Despite Conservation of PUF Binding Specificities.** Two models could explain the differences in binding site lengths: Transcripts in the mitochondrial cluster could be regulated by the same PUF ortholog whose binding specificity changed during evolution; alternatively, the sequence of the binding elements could have changed, while the protein specificity did not. To address this, we investigated the specificities of PUF proteins between species. All proteins containing a Pumilio domain for all 62 species were aligned, and then a protein tree was generated by maximum likelihood (Fig. S2).

Nearly all species have a single ortholog of Puf3 (Fig. 2A, ring 3). In contrast, Saccharomyces and Schizosaccharomyces have two paralogs that belong to a Puf4/Puf5 clade, whereas Pezizomycotina and other fungi have only one copy of the Puf4/5 protein (Fig. 2A, ring 4). This suggests there was one ancestral gene, which we refer to as Puf4/5, that was duplicated in the Saccharomycotina (37) and then again in the Schizosaccharomyces. This is particularly intriguing because *S. cerevisiae* Puf4 and *S. cerevisiae* Puf5 have different binding specificities, suggesting that their binding preferences diverged after duplication (14, 30).

To distinguish between changes in protein specificity or binding site sequence, we assayed the RNA binding specificities of *A. nidulans* Puf4/5 (AnPuf4/5; AN10071) and *N. crassa* Puf4/5 (Nc_Puf4/5; NCU16560.1) using a series of RNAs with binding elements ranging in length from 8 to 10 nt, using the three-hybrid system (47). In this assay, levels of LacZ expression are directly related to binding affinity (48). *A. nidulans* Puf4/5 bound best to sequences matching the 9-nt PUF binding element, similar to the known binding preference of Sc_Puf4, but also bound weakly to a 10-nt element (Fig. 2D). In contrast, *N. crassa* PUF4/5 bound well to both 9- and 10-nt-long binding elements (Fig. 2D). The broader specificity of Nc_PUF4/5 echoes the expanded binding specificity of Sc_Puf5 (30). Together, these data suggest that the ancestral PUF4/5 protein could bind both 9-nt and 10-nt elements and that the known specificities for the *S. cerevisiae* proteins emerged after the PUF4/5 duplication (see Discussion).

The binding specificities of the filamentous fungi PUF4/5 orthologs and the enrichment for putative 9- and 10-nt elements in the orthologs of the mitochondrial cluster suggest that there was a “regulator switch,” in which *A. nidulans* and *N. crassa* underwent a change in the protein that mediates control of orthodox mitochondrial RNAs. In this hypothesis, mRNAs that possess 8-nt elements in *S. cerevisiae*, which are recognized by Sc_Puf3, instead possess 9- or 10-nt elements in *Neurospora* and *Aspergillus* and are likely bound by PUF4/5 in those species.

**In Vivo Targets of *S. cerevisiae* Puf3 and *N. crassa* Puf4/5 Demonstrate a Regulatory Switch.** To test the regulator switch hypothesis, we identified mRNAs bound in vivo to *Sc_Puf3*, *Nc_Puf3* (Nc_PUF3), and Nc_PUF4/5. We cross-linked RNA–protein complexes in vivo with UV irradiation of intact cells, immunopurified these protein–RNA complexes, and identified the RNAs by high throughput sequencing (“HITS-CLIP”) (49). We recently analyzed Sc_Puf5 using the same methods and included those data in the meta-analysis (30). The *S. cerevisiae* PUF genes were TAP-tagged and the *N. crassa* PUF genes were FLAG-tagged; all genes were integrated by homologous recombination at the endogenous chromosomal locus under control of the natural promoter. The three biological replicates for each protein displayed excellent reproducibility (Fig. 3A–C).

We identified 500, 835, and 746 significant CLIP peaks for Sc_Puf3, Nc_Puf3, and Nc_PUF4/5, respectively. These arise from 7.6, 8.4, and 14.3 million uniquely assignable reads (for complete mapping statistics, see Table S2). Data quality in HITS-CLIP is often assessed by the presence of binding elements in peaks of read enrichment (50–52). In our data, 90%, 68%, and 99% of Sc_Puf3, Nc_Puf3, and Nc_PUF4/5 binding peaks, respectively, contain one or more sequences matching the PWMs derived for each set of targets, suggesting our sequencing data identify genuine PUF protein targets.

The peaks, predominantly found in the 3′UTRs of mRNAs, represent 465, 803, and 686 transcripts bound by Sc_Puf3, Nc_Puf3, and Nc_PUF4/5, respectively (Fig. S3 and Dataset S1). Representative CLIP peaks are shown in Fig. 3 D–G for COX17 orthologs, a well-characterized target of Sc_Puf3 (Fig. 3D) that is not efficiently bound by Sc_Puf5 (Fig. 3E). Strikingly, the Neurospora COX17 ortholog was bound by Nc_Puf5 but not Nc_Puf3 (Fig. 3 F and G). In all cases, binding peaks were centered over the known binding elements in the *S. cerevisiae* COX17 mRNA or the predicted elements in the *N. crassa* ortholog. This trend was seen for the profiles of the five loci with the largest binding signal in the *N. crassa* and *S. cerevisiae* Puf3 and PUF4/5 proteins (Fig. S4 A–C).

To assess whether targets were shared between Sc_Puf3 and *N. crassa* PUFs, we determined their overlap across those species (Fig. 3H). The overlap in transcripts bound by Sc_Puf3 and Nc_PUF4/5 was much more significant (6.1E-41, hypergeometric distribution test) than the overlap between Sc_Puf3 and Nc_PUF3 (5.0E-3, hypergeometric distribution test) (Fig. 3H). Orthologs of 42% of Sc_Puf3 targets were bound by Nc_PUF4/5, whereas only 17% were bound by Nc_Puf3. We conclude that Sc_Puf3 and Nc_PUF4/5 proteins bind a similar set of RNAs, even more than the two PUF3 orthologs.

**PUF Binding Specificities Across Species.** To determine the in vivo sequence preference of each PUF protein, we identified PWMs based on overrepresented sequences in targets from the HITS-CLIP database (43). The binding specificities of the *N. crassa* PUFs were strikingly similar to those of their orthologs in *S. cerevisiae* (Fig. 4 A and C). Sc_Puf3-bound sites contained sequences matching the known 9-nt binding consensus comprised of a 5′ UGU and a 3′ AUA (Fig. 4A). In addition to the core PUF binding element, ∼50% of the sequences associated with Sc_Puf3 contain a cytosine at the minus 2 position, which enhances binding in vitro and biological activity in vivo (32, 36). Nc_PUF3-bound regions contained 8-nt sequences similar to the known Sc_Puf3 binding element, but the PWM lacked

Wilinski et al.  PNAS | Published online March 20, 2017 | E8219
the –2C enrichment seen in *S. cerevisiae* (Fig. 4B) (14, 32, 36). The PWM identified for targets of Sc_Puf5 contains the canonical 5′-UGUA motif with apparent degeneracy in the 3′-end; we previously showed that the single matrix includes discrete 8–12-nt-long binding elements that differ in spacer length (30). When training a single PWM on *Ne_PUF4/5* targets, we identify the same type of apparent degeneracy (Fig. 4D) that can be explained as a combination of 9-nt- and 10-nt-long elements (Fig. 4E). The PWMs derived from in vivo targets agree well with our bioinformatic predictions of sequence enrichment (Fig. 2B) and the binding specificities measured ex vivo (Fig. 2D). We conclude that the in vivo binding preferences of PUF3 orthologs is highly conserved and that the binding preference of the single *Ne_PUF4/5* is similar to that of Sc_Puf5.

**Evolution of PUF Target Networks in *N. crassa***. To determine the functional enrichment for the targets of each PUF protein, we performed GO analyses. As expected, the Sc_Puf3 targets exhibited a strong enrichment for nuclear-encoded mitochondrial proteins (*P* value 1E-103) and translation factors (*P* value 6.5E-37) (Fig. 4F and Dataset S2). The targets of *Ne_PUF4/5* were dramatically enriched for those encoding nuclear-encoded mitochondrial proteins (*P* value 1E-25); Sc_Puf3 targets also were enriched for this annotation, although much more modestly (*P* value 1E-5) (Fig. 4F and Dataset S2). In agreement with these enrichments, we found that 50% (155/309) of orthologous transcripts in the mitochondrial cluster (Fig. 1C) were bound by Sc_Puf3 in *S. cerevisiae*, whereas 44% (137/309) were bound by Ne_PUF4/5 and 13% (39/309) by Ne_Puf3 in *N. crassa* (Fig. 4G). Interestingly, of the 39 targets of Ne_Puf3 in the mitochondrial cluster, 71% were also bound by Ne_PUF4/5 (Fig. 4G). These transcripts often contained both 8-nt- and 9/10-nt-long putative PUF binding elements. Two examples of cotargets of Ne_PUF3 and Ne_PUF4/5 are transcripts CBP1 and NCU07386, which harbor both binding elements, one under each respective peak (Fig. 4H). Thus, a subset of mitochondrial-related transcripts are bound by both PUF4/5 and PUF3 through their respective recognition elements. *Ne_PUF4/5* targets were enriched for other functional annotations beyond mitochondrial transcripts. The group of targets was also enriched for those encoding cytosolic ribosomal proteins (*P* value 3E-16), which is intriguing, as previously identified Sc_Puf4 targets are enriched for transcripts encoding ribosome biogenesis proteins (14) (see Discussion).

**Evidence for Multiple Rewiring Events and the Emergence of New Networks**. Our data suggest that orthologous transcripts have evolved different lengths of PUF binding elements in their 3′UTRs and hence are regulated by different PUF proteins. To further investigate this, we assembled a combined set of all orthologous transcripts bound by Sc_Puf3, Ne_Puf3, and Ne_PUF4/5, or Sc_Puf4 (identified in a previous study by a different method) (14). We clustered this combined set of transcripts based on putative

---

**Fig. 3.** HITS-CLIP data. (A–C) Reproducibility of biological replicates. Height of CLIP peaks were log-transformed and then plotted on each axis. Data were colored based on the precipitated protein (Sc_PUF3, blue; Ne_PUF3, red; Ne_PUF4/5, green). Spearman’s correlation coefficients (*r*), associated *P* values (*P*), and Pearson’s correlation coefficient (*r*) are indicated in A (*r* = 0.65, 0.77; *P* = 0; *r* = 0.77, 0.96; *P* = 500), B (*r* = 0.64, 0.85, 0.76; *P* = 0; *r* = 0.99, 0.99, 0.99; *n* = 835), and C (*r* = 0.85, 0.92, 0.79; *P* = 0, 0, 0; *r* = 0.71, 0.97, 0.67; *n* = 746). (D–G) Examples of binding peaks (read depth) for COX17, a canonical target of Sc_Puf3. YLL009C (COX17), NCU0058 (al-2), and NCU02530 (cox17) are depicted, with ORFs annotated in blue and the likely binding elements displayed beneath. (H) Overlapping CLIP targets for each protein. Only genes with orthologs in *S. cerevisiae* and *N. crassa* are included. RPM, reads per million mapped reads.
Fig. 4. In vivo binding elements of PUF proteins. MEME-derived PWMs from CLIP data. (A) S. cerevisiae PUF3, (B) N. crassa PUF3, (C) S. cerevisiae PUF5 (30), (D) N. crassa PUF4/5. (E) The previously defined Par-CLIP of Homo sapiens PUF3 ortholog, PUM2, defined by PhyloGibbs (52, 59). (F) Deconvolution of Nc_PUF4/5 composite binding element (D) into binding elements of 9 nt and 10 nt in length. (G) Mitochondrion GO term enrichment for all mRNAs targets of each PUF protein. (H) Venn diagram depicting the overlap between PUF targets from the mitochondrial cluster. (H) Nc_PUF3 and Nc_PUF4/5 interaction peaks identify separate binding elements for each PUF protein. One biological replicate for Nc_PUF3 and Nc_PUF4/5 are plotted in red and brown, respectively. Likely PUF binding elements are highlighted. RPM, reads per million mapped reads.

PUF binding elements identified in 3’UTRs of each species’ orthologs across 62 fungal species (Fig. 4A). This analysis captured differential enrichment of putative PUF elements in the orthologous mRNAs, all of which are targets identified in our CLIP studies. Several striking patterns emerged that implicate additional PUF regulatory rewiring and the emergence of networks. As reported above, transcripts encoding mitochondrial proteins (Fig. 5, cluster A) are enriched for 9-nt elements in S. cerevisiae and, to some extent in N. crassa, but are strongly enriched for 9-nt and 10-nt elements in filamentous fungi. Interestingly, many transcripts linked to mitochondrial envelope and organization (Fig. 5, cluster B) were specifically depleted of PUF binding elements in the Neurospora clade. In contrast, cluster C (Fig. 5) was bound by Nc_PUF3 and enriched for the 8-nt PUF element—strikingly, this group was enriched for transcripts encoding hydrolases and cellulose-binding proteins, functions that may be important for the optimal growth of this filamentous fungus. This cluster may represent the emergence of a new network in that species. Cluster D reveals ribosomal protein mRNAs with 9-nt elements in Neurospora but not in S. cerevisiae. Cluster E reveals mRNAs with 9-nt elements in S. cerevisiae (and recognized by its Puf4 and Puf5 proteins), which again may imply the emergence of a new network during evolution. Finally, we observe an enrichment of PUF targets in Neurospora that are responsive to light (Nc_PUF3: 233, P value 1.7E-13, Nc_PUF4/5: 257, P value 1.1E-22; hypergeometric distribution test). Thus, this regulatory system is also impacted by PUF regulation.

To visualize the overlaps, we depicted the connection between every orthologous target and its cognate binding protein(s) (Fig. 6). Each target of the four proteins (Sc_Puf3, Sc_Puf5, Nc_Puf3, or Nc_PUF4/5) is represented as a node (small circle). Many of the transcripts with mitochondrial functions are bound by Sc_Puf3 and Nc_PUF4/5 in the respective species, whereas others
are uniquely associated with one protein or another (Fig. 6 A). The fact that many connections are unique, but with targets that encode proteins in the same biological process, implies that PUF regulation of this biological function is conserved, whereas the specific mRNAs that are regulated diverged (Fig. 6 B–D).

**The Emergence of New Sites Underlies Rewiring.** Two evolutionary routes could produce an apparent switch from 8-nt PUF3 regulation to 9/10-nt PUF4/5 regulation. An element of one length could change into a different length simply by gain or loss of nucleotides within the element; such changes are particularly facile, as the identity of bases in the central region of the site is unimportant in binding (see the Introduction). Alternatively, a new binding element could arise elsewhere in the same 3′ UTR before the original motif degenerates. We find a preponderance of evidence for the latter model. *Nc*_PUF4/5 targets in the mitochondrial cluster are significantly enriched for transcripts with multiple binding elements per mRNA UTR (average of 2.04 elements per 3′ UTR) compared with all *Nc*_PUF4/5 targets (average of 1.67) (*P* value 6.5E-4, two-tailed *t* test). Furthermore, transcripts within a species can be bound by both proteins, as previously detailed for the cotargeting of CBP3 and *NCU07386* by *Nc*_PUF-3 and *Nc*_PUF-4/5 (Fig. 4 H). Although element-length switching likely exists, the presence of targets with two sites in a single 3′ UTR suggests a transient intermediate may often exist in which both proteins act on the same mRNAs.

**Discussion**

Our work provides compelling experimental support for extensive rewiring of PUF-dependent regulatory networks in fungi and insights into how that occurred. Computational analyses prompted models of the rewiring of the nuclear-encoded mitochondrial network (5, 37, 38). Our biochemical data provide molecular evidence for such a regulator switch, reveal multiple rewiring events, and support fluidity in the emergence of new PUF–RNA networks.

The most parsimonious model for the evolution of the mitochondrial networks posits that mitochondrial transcripts (along with cytosolic ribosomal protein mRNAs and other transcripts) were regulated by both PUF3 and PUF4/5 in the common ancestor (Fig. 7). The relative prominence of PUF3 versus PUF4/5 control of mitochondrial transcripts evolved in the modern-day lineages, with
PUF3 playing the predominant role in Sacharamycotina and PUF4/5 the major role in filamentous fungi. Orthologous PUF3 proteins have stringent specificity for 8-nt-long elements in virtually every organism tested to date, including mammalian PUM (14, 30, 52, 53). We show that, in contrast, Nc_PUF4/5 and An_PUF4/5 can bind 9 nt, and even 10 nt in the case of Nc_PUF4/5 long binding elements (Fig 2D), as does Sc_Puf5 (30). Together, this suggests that the ancestral mitochondrial transcripts were regulated by PUF3 and PUF4/5 bound to their 8-nt and 9/10-nt 3′ UTR elements, respectively.

Several events produced specialization of the PUF networks along the branch leading to budding yeasts. Our phylogenetic analysis indicates that PUF4/5 duplicated after the split of budding and filamentous fungi and then diverged in binding preference; broad binding specificity was likely ancestral (37). We propose that Sc_Puf5 retained the ancestral specificity whereas Sc_Puf4 specialized to bind a 9-nt element. These changes likely involved alterations in the topology of the PUF scaffold (30). As protein binding specificity evolved, so too did the elements of the downstream targets: Mitochondrial transcripts bound by Sc_Puf3 evolved toward 8-nt elements, whereas transcripts bound by Sc_Puf4 and Sc_Puf5 evolved 9-nt or maintained mixed-length PUF sequences, respectively. It is intriguing that Sc_Puf4 targets are enriched for ribosome-biogenesis transcripts, which are functionally related to the ribosomal protein transcripts bound by the orthologous Nc_PUF4/5 in N. crassa. We propose that the ancestral PUF4/5 protein regulated both mitochondrial and cytosolic ribosome-biogenesis functions but that this coregulation was lost in budding yeasts when the two networks specialized toward PUF3- and PUF4-dependent control, respectively. This may have facilitated a further decoupling of mitochondrial and cytosolic ribosomal protein gene expression in respiro-fermentative yeasts, as previously proposed to have occurred through upstream transcription factor rewiring (54, 55). The acquisition of sugar-responsive PUF3 regulation could have provided selective pressure to evolve and maintain the PUF3 network in respiro-fermentative

---

**Fig. 6.** Functional relatedness of all mRNAs bound by PUF3 and PUF4/5 proteins of *S. cerevisiae* and *N. crassa*. All orthologous targets of PUF proteins defined by CLIP are plotted as nodes (gray balls). Edges (lines) link the PUF to the RNAs it binds. Targets are colored based on their annotated function: (A) mitochondria (purple), (B) ribosome (red), (C) hydrolase (black), and (D) light responsive (blue).
yeasts. *S. cerevisiae* is extensively phosphorylated when *S. cerevisiae* is depleted of glucose and switches from a repressor to an activator of mitochondrial mRNAs (24). Acquisition of regulated activation via phosphorylation may have committed the budding yeast lineage to using PuF3 to control mitochondrial RNAs.

A different story emerges for the evolutionary trajectory in the filamentous fungi. Our experimental data demonstrate PuF rewiring of the mitochondrial targets. Orthologs of *S. cerevisiae* PuF3-bound targets display strong enrichment for 9-nt or 10-nt sites in filamentous species. We show here that these orthologs of *S. cerevisiae* PuF3 targets are bound by *Neurospora* Puf4/5 in *N. crassa* and to a lesser extent by *Neurospora* Puf3. We find one set of transcripts that is strongly enriched for 9-nt PuF elements in the *Neurospora* species and, to a lesser extent, in the budding yeasts; however, orthologs in the Aspergilli are enriched primarily for 8-nt PuF elements (Fig. 5, cluster D, and Fig. 7, purple box). This observation suggests a second hand-off of targets potentially regulated by PUF3 in Aspergilli. Given the enrichment in this group for ribosome-related functions and for 9-nt elements in budding-yeast transcripts, we propose that a subset of these mRNAs are regulated by PuF4 in *S. cerevisiae* (see Fig. 5, cluster D *S. cerevisiae* binding profiles). This model suggests a complex interplay between three different PuF regulators in three branches of fungi and implicates a second independent rewiring of PuF3 regulation.

Our data also demonstrate an additional, independent evolutionary pattern of PuF4/5 binding specificity in Aspergilli. Whereas the single PuF4/5 protein in *N. crassa* binds 9-nt and 10-nt sequences equally well, an PuF4/5 strongly prefers the 9-nt-long element (Fig. 2D). Concordantly, the mitochondrial transcripts from *Neurospora* species are nearly equally enriched for 9-nt and 10-nt elements, whereas those from *Aspergilli* are significantly biased toward 9-nt elements (Fig. 2A, arrows). Together, these data indicate that *Aspergilli* have evolved both in terms of PuF4/5 specificity and in the 3′UTR sequences found in hundreds of its transcripts. It also confirms that there have been two independent origins of PuF4/5 binding preference: once after the split of *Aspergilli* and *Neurospora* species and again after PuF4/5 duplication in the *Saccharomyces* lineage. It is plausible that the determinants for 9-nt versus 10-nt site preference in the PuF4/5 proteins may be relatively simple to change; modest changes in the curvature of the PuF scaffold may be sufficient (30, 35, 56). The related binding specificities of PuF proteins and the relaxed binding preferences of the ancestral PuF4/5 protein likely both contributed.

Regulatory redundancy is a recurring theme in transcription-factor network rewiring. Ancestral and newly derived regulators are proposed to cofunction until ancestral regulation is lost, completing the “handoff” of targets. In most examples, that redundancy is provided by the co-occurrence of distinct transcription factor binding sites (6, 8, 57). This model of multiple elements likely pertains to rewiring of the PuF regulons. For example, many mitochondrial transcripts bound by *Neurospora* PuF3 are also bound by *Neurospora* PuF4/5—in these cases, binding appears to occur through 8-nt or 9/10-nt elements, respectively. The co-occurrence of multiple elements may have facilitated evolution toward predominantly PuF3-dependent regulation in the budding yeasts versus primarily PuF4/5-dependent regulation in the filamentous fungi, simply through loss of one class of binding site. Once initiated, this situation could have amplified the selective pressure for other functionally related targets to acquire 8-nt or 9/10-nt elements in the respective lineages to maintain coregulation under the new regulatory system.

PuF proteins bind a set of related binding elements—this, combined with the relaxed specificity of PuF4/5, could provide an alternate model of redundant regulatory systems. PuF elements in the target transcripts could evolve by de novo creation of new binding sequences and loss of ancestral sites from elsewhere in the 3′UTRs. Alternatively, an existing PuF element could rapidly switch into a distinct-length PuF sequence through simpler changes. Insertion or deletion of bases into the PuF-element spacer region could instantly change PuF element length and thus change which PuF protein binds that element. Alternatively, addition of a simple UA dinucleotide downstream of an existing 8-nt element (e.g., UGUAN-

---

**Fig. 7.** Evolutionary model of PuF network rewiring. Colored boxes represent orthologous sets of mRNA targets. Solid arrows indicate PuF binding of the denoted binding element enriched in the target mRNA UTRs; dashed arrows represent weak binding or secondary modes of regulation (see Discussion for details). Dashed colored boxes represent the acquisition or loss of transcripts from the denoted regulon. Mt, mitochondrial; Ribo, ribosomal.

---

**Materials and Methods**

Detailed methods and descriptions of yeast three-hybrid assay, HITS-CLIP, *N. crassa* manipulations, and accession numbers can be found in SI Materials and Methods. Ascomycota fungi were chosen based on previously determined orthology (41). Phylogenetic tree shown in Fig. 1A and Fig. S1 was calculated using ribosomal RNA...
remaining genes were then k-means clustered with k = 13. We chose 13 clusters based on the “elbow” method, which considers the percentage of variance as a function of the number of clusters. Additional descriptions of bioinformatic analyses are provided in SI Materials and Methods.

ACKNOWLEDGMENTS. We thank the M.W. and Kimble laboratories, and Elena Sorokin, in particular, for helpful discussions of this work. We also thank Laura Vanderploeg of the UW Bioinformatics Lab for help with figures. This work was supported by NIH Grants R01 GM035690 and R01 GM03061 (to E.S.), F32 GM079721 (to A.D.K.), R01 GM038989 (to A.P.G.), and R01 GM50942 (to M.W.).