Rewiring of Posttranscriptional RNA Regulons: Puf4p in Fungi as an Example

Huifeng Jiang,1,2 Xiaoxian Guo,1,3 Lin Xu,1,4 and Zhenglong Gu*1

1Division of Nutritional Sciences, Cornell University
2Key Laboratory of Systems Microbial Biotechnology, Institute of Tianjin Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China
3College of Life Sciences, Zhejiang University, Hangzhou, People’s Republic of China
4Department of Molecular Biology and Genetics, Cornell University
*Corresponding author: E-mail: zg27@cornell.edu.

Abstract

It has been increasingly clear that changes in gene regulation play important roles in physiological and phenotypic evolution. Rewiring gene-regulatory networks, i.e., alteration of the gene-regulation system for different biological functions, has been demonstrated in various species. Posttranscriptional regulons have prominent roles in coordinating gene expression in a variety of eukaryotes. In this study, using Puf4p in fungi as an example, we demonstrate that postranscriptional regulatory networks can also be rewired during evolution. Although Puf4p is highly conserved in fungi, targets of the posttranscriptional regulon are functionally diverse among known fungal species. In the Saccharomycotina subdivision, target genes of Puf4p mostly conduct function in the nucleolus; however, in the Pezizomycotina subdivision, they are enriched in the mitochondria. Furthermore, we demonstrate different regulation efficiencies of mitochondrial function by PUF proteins in different fungal clades. Our results indicate that rewiring of posttranscriptional regulatory networks may be an important way of generating genetic novelties in gene regulation during evolution.

Key words: posttranscriptional regulon, evolution of gene regulation, yeast.

Introduction

Understanding how gene regulation evolves is a key area in research on evolution (King and Wilson 1975; Wray et al. 2003; Carroll 2005; Ihmels et al. 2005; Frankel et al. 2010; Kaplan et al. 2011; Rozpedowska et al. 2011; Shou et al. 2011). Gene regulation can occur at various levels. The concept of the posttranscriptional regulon was recently proposed to describe the phenomenon that mRNAs encoding functionally related proteins are coordinately regulated for stability, localization, and translation (Keene 2007). Previous works on the evolution of gene regulation have been conducted mostly at the gene-transcription level (Borneman et al. 2007; Wray 2007; Emerson et al. 2010), whereas the evolution of posttranscriptional regulation has received far less attention. Illustrating how changes occur at the level of posttranscriptional regulation is crucial for gaining better understanding of how gene regulation evolves in nature (Chen and Rajewsky 2007).

Posttranscriptional gene regulation plays an important role in determining biological complexity. Posttranscriptional regulons, consisting of RNA-binding proteins (RBPs) and their target genes, are essential for the coordination of gene expression (Mata et al. 2005; Keene 2007; Woodson and Chory 2008). Although extensive functional researches have been conducted on various RBPs and divergences of RBP targets in several species have been revealed (Keene 2007; Hogan et al. 2008), the evolutionary mode of posttranscriptional regulatory networks remains largely unknown. We have shown that a tinkering evolutionary mode of the PUF3 posttranscriptional regulon occurred during fungal evolution (Jiang et al. 2010). In this study, we focus on investigating the evolutionary mode of the PUF4 posttranscriptional regulon, which plays an important role in regulating genes that are associated with ribosome biogenesis in budding yeast, Saccharomyces cerevisiae (Gerber, Herschlag and Brown 2004; Grigull et al. 2004; Foat et al. 2005).

Puf3p and Puf4p belong to a group of well-characterized posttranscriptional regulators that contain a PUM-HD type (PUMilio homology domain, which contains eight helical repeats) RNA-binding domain to bind mRNAs on the 3′UTR (un-translated region). By binding with their target mRNAs, PUF proteins can recruit factors to transfer the target mRNAs into various subcellular locations, repress translation of mRNAs, and promote mRNA degradation (Chritton and Wickens 2010; Quenault et al. 2011). The degradation of mRNA by PUF proteins requires the recruitment of the Ccr4p–Pop2p–Not deadenylate complex (Goldstrohm et al. 2006). The mechanism of mRNA repression is likely conserved for all PUF proteins across species (Hook et al. 2007; Kadyrova et al. 2007; Lee et al. 2010).

By altering the binding elements in genes, functions of targets for different PUF regulons vary in different species (Crittenden et al. 2002; Gerber et al. 2004, 2006; Galgano et al. 2008; Tam et al. 2010). In budding yeast, different PUF proteins recognize different elements and regulate genes in different functional categories by recruiting the
RNA decay complex. For example, Puf3p and Puf4p mostly regulate genes that function in mitochondria and the nucleolus, respectively (Gerber et al. 2004). Only one PUF protein in fruit flies is essential for early embryogenesis and cell development (Murata and Wharton 1995; Lin and Spradling 1997; Gerber et al. 2006). One of the PUF proteins in worms determines the proliferation of germline stem cells (Crittenden et al. 2002). In humans, two PUF proteins interact with the microRNA system to regulate genes in transcription, the cell cycle, and cell proliferation (Galgano et al. 2008; Morris et al. 2008). The dynamic functions of PUF regulons indicate rich evolutionary history of these posttranscriptional regulatory networks. Our results from this study indeed indicate that the functions of genes that are regulated by the PUF-4 posttranscriptional regulon were reprogrammed in different fungal clades during evolution. In the *Saccharomycotina* subdivision, Puf4p regulates genes in the nucleolus, whereas in the *Pezizomycotina* subdivision, it regulates genes that function in mitochondria. Together with our previous study on the evolution of Puf3p targets in fungi (Jiang et al. 2010), our results reveal new principles on the evolution of novel posttranscriptional regulation and call for more efforts to investigate the evolution at this critical level of gene regulation.

**Materials and Methods**

**Sequence Data**

Sequences for the studied fungal species were from Fitzpatrick et al. (2006) and the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). The human, mouse, worm, and fly sequence data were downloaded from the Ensembl database (http://asia.ensembl.org/index.html). The sequence data for Arabidopsis *thaliana* was downloaded from http://www.plantgdb.org/ . Mitochondrial genes for human, mouse, and *A. thaliana* were downloaded from the MitoP2 database (http://www.mitop.de ) (Prokisch et al. 2006). The mitochondrial genes for worm and fly were identified based on their orthology with human mitochondrial genes. The nucleolar genes for human were defined based on the Gene Ontology annotation (GO:0005730). The nucleolar genes for mouse, worm, fly, and *A. thaliana* were detected by their orthology with nucleolar genes in the human genome.

**Definition of Orthologous Genes**

Using the InParanoid software package (Remm et al. 2001), orthologs between the budding yeast, *S. cerevisiae*, and other fungal species were identified. The default parameters were used in detection of PUM-HD repeats detected by HMMER3 (http://hmmr.org/ ) based on the Hidden Markov Model profile for PUF domain (PF00806) from the Pfam database (http://pfam.sanger.ac.uk/).

**Reconstruction of the Fungal Phylogenetic Tree**

Using genes in *S. cerevisiae* as queries, we obtained the ten most conserved orthologs between *S. cerevisiae* and *Cryptococcus neoformans* (the most remotely related fungal species to *S. cerevisiae*) and used these genes to reconstruct a phylogenetic tree for all the studied fungal species. Protein sequences were aligned using MUSCLE (Edgar 2004). The phylogenetic tree was constructed by ClustalW2.0 using distance matrix method and Neighbour-Joining (NJ) clustering method. Finally, the tree was drawn by MEGA5 (Larkin et al. 2007; Tamura et al. 2011).

**Detection of Target Genes of PUF4 Posttranscriptional Regulon**

As in our previous study (Jiang et al. 2010), we used a Perl script to detect the target motif (P4E) of Puf4p. As shown in Gerber et al. (2004), we allowed flexibility (among A, C, and T) in the fifth and seventh nucleotides and required exact match for the rest nucleotides of the P4E motif during motif detection. For each species, we scanned 250 bp of DNA sequence downstream of all annotated genes to discern the occurrence of the motifs. Genes with the P4E motif in their 250-bp 3′ downstream regions were defined as the target genes of Puf4p.

**Enrichment Analysis in the Localization Categories**

Sublocalization information for genes in budding yeast was downloaded from the Saccharomyces Genome Database (http://www.yeastgenome.org/) (Huh et al. 2003). The localization of genes in other species was defined based on the localization of their orthologs in *S. cerevisiae*. If a gene from one species has two orthologs in *S. cerevisiae*, its localization was defined as the combination of localization for both copies. The hypergeometric test was used to test the enrichment of genes with P4E in each localization category. For multiple comparisons, the false discovery rates are no larger than 0.01 (Benjamini and Hochberg 1995).

**Results**

**The Evolution of PUF Proteins in Fungi**

The PUF protein family is conserved throughout the eukaryotic kingdom (Wickens et al. 2002; Spassov and Jurecic 2003). However, the number and function of PUF proteins vary among species (Quenault et al. 2011). In *S. cerevisiae*, seven proteins with PUM-HD repeats were identified as PUF proteins (Supplementary table 1, Supplementary Material online). In order to investigate the evolution of these PUF proteins in fungi, we identified their orthologous genes in the sequenced fungal species (Supplementary Table 2, Supplementary Material online). In figure 1, we show the distribution of orthologous genes for these seven PUF proteins in different fungal species. Figure 1 shows that PUF2 originated during a whole genome duplication (WGD) event because its orthologs are only present in the post-WGD yeast species (Byrne and Wolfe 2005). The orthologs of PUF5 are present only in the *Saccharomyces* subdivision except *Yarrowia lipolytica*, which is the furthest species from budding yeast in the subdivision. Therefore, both PUF2 and PUF5 genes exist only in some species of the *Saccharomyces* subdivision.

Since the PUM-HD domain in PUF proteins is essential for binding mRNA, we examined the number of PUM-HD...
Fig. 1. The evolution of Puf proteins in fungi. The phylogenetic tree for species in fungi was reconstructed based on ten conserved proteins (the bootstrap scores are displayed in Supplementary fig. 1, Supplementary Material online, and the abbreviations for the studied species are shown in Supplementary table 2, Supplementary Material online). The tree was consistent with the results of a previous study (Fitzpatrick et al. 2006). The color in each cell denotes the number of PUM-HD repeats for that gene in the respective species. Gray means that there was no detectable orthologous gene. The arrows indicate the WGD event and the saccharomyces subdivision, respectively.

Repeats for all of the PUF orthologs we studied. As shown in figure 1, although orthologs of PUF1, PUF6, and NOP9 can be identified in all studied fungal species, the number of PUM-HD repeats in these genes is not conserved, indicating that the functions of Puf1p, Puf6p, and Nop9p may be dramatically changed in different fungal species.

In contrast to the above PUF proteins, most orthologs of Puf3p and Puf4p consist of a canonical PUM-HD domain with eight repeats. Targets of these two PUF proteins also have significant functional enrichment in the budding yeast, warranting further investigation on how their targets evolved in different fungal species. Our previous work has shown that the PUF3 posttranscriptional regulon is highly conserved in fungi, but the dedicated regulation of genes that function in the mitochondria occurs only in the Saccharomyces subdivision (Jiang et al. 2010). In the following study, we focused on the functional evolutionary process of the PUF4 posttranscriptional regulon in fungal species.

Functional Rewiring of PUF4 Posttranscriptional Regulon in Fungal Species

Previous researches have shown that Puf4p plays an important role in ribosome biogenesis (Gerber et al. 2004; Foat et al. 2005). This motivated us to investigate whether the function of the PUF4 posttranscriptional regulon is conserved among fungal species. Accordingly, we identified all orthologous genes between each studied fungal species and budding yeast, S. cerevisiae. Genes in each species were categorized based on the subcellular localization of their orthologs in the budding yeast (Huh et al. 2003). In each localization category, we defined a gene as a target of the PUF4 regulon if its 3’ UTR had the recognized binding element of the PUF4 regulon (P4E). We further estimated the enrichment of target genes in each species (the phylogenetic relationship among species is displayed in Supplementary fig. 1, Supplementary Material online, and the total number of genes in each category are listed in Supplementary table 2, Supplementary Material online). As shown in figure 2A, we confirmed previous experimental results that genes functioning in the nucleolus have enriched Puf4p binding. Indeed, this pattern was observed in most species in the Saccharomyces subdivision. However, to our surprise, the studied species in the Pezizomycotina subdivision showed significant enrichment of P4E only in genes that are encoded by the nuclear genome but function in mitochondria (therefore called mitochondrial genes for simplicity hereafter), indicating a possible functional rewiring of the PUF4 posttranscriptional regulon during fungal evolution.

Regulation of nucleolar genes by Puf4p also showed clade-specific dynamics in the Saccharomyces subdivision. All of the species in one of the clades (Group 1) showed functional enrichment of Puf4p binding for the nuclear genes, whereas only some of the species in the other clade (Group 2) showed a similar pattern (Supplementary fig. 2, Supplementary Material online). As shown in figure 2B, on average, only 2% and 4% of genes in the nucleolus had P4E for species in Group 3 and 4, respectively. Many more genes in the nucleolus had P4E in Group 1 (~30%, two-tailed t-test P-value = 2 × 10⁻⁹ comparing with the Groups 3 and 4) and Group 2 species (~17%, two-tailed t-test P-value = 5 × 10⁻⁴ comparing with the Groups 3 and 4). The percentage of genes in the nucleolus with P4E in the Group 1 species was also significantly higher than that for the genes in the Group 2 species (two-tailed t-test, P value = 0.001).

Deep Functional Divergence of PUF3 and PUF4 Regulons in Fungal Species

We have shown that mitochondrial genes in the Saccharomyces subdivision are regulated by the PUF3 posttranscriptional regulon (Jiang et al. 2010). Interestingly, the above results indicate that the mitochondrial genes may be regulated by the PUF4 posttranscriptional regulon for species in the Pezizomycotina subdivision. In order to further understand the evolutionary relationship between these two posttranscriptional regulons, we reconstructed the phylogenetic tree for the orthologs of two regulators, Puf3p and Puf4p, in the studied fungal species. The PUM-HD repeats, which is the functional domain for RNA binding, was used in tree reconstruction. The ortholog of PUF protein in Gardia lamblia, which is considered as the sister group to all other eukaryotes (Sogin et al. 1989; Sun et al.
and colleagues using protein structures of Puf3p and Puf4p and genetics study showed that two amino acids changes, C ↔ T and R ↔ C, at the third and fifth PUM-HD repeats, respectively, are essential for the recognition difference between Puf3p and Puf4p (Miller et al. 2008). Comparing these two repeats in all orthologs of Puf3p and Puf4p, we found that these two pairs of different amino acids were conserved in all Puf3p and Puf4p orthologs (fig. 3B), indicating that the recognition element by PUM-HD repeats in Puf4p might be conserved and distinct from the recognition element of Puf3p for all the studied fungal species. Genes with P3E or P4E in their 3′ UTR regions were further compared in each species. As shown in figure 3C, the possible target genes of PUF3 and PUF4 regulons (genes with P3E and P4E, respectively) has very little overlap (less than 4%, on average), further supporting the conclusion that these two proteins may have different functions in all fungal species.

Posttranscriptional Regulation of Mitochondrial Genes during Fungal Evolution

We further compared the regulatory efficiency of two posttranscriptional regulons for mitochondrial genes in their respective clades. On average, about 12% of the mitochondrial genes are bound by PUF4 genes in the Pezizomycotina subdivision, which is significantly lower than for PUF3 genes in the Saccharomycotina subdivision (~45%, two-tailed t-test P < 2 × 10^{-16}). Because occurrence of the binding motif is correlated with GC content at the genome level (Jiang et al. 2010), we normalized the frequencies of P3E and P4E in mitochondrial genes by their occurrences in the genome. As shown in figure 4A, on average ~15% P4E in the whole genome appears in the mitochondrial genes for species in the Pezizomycotina subdivision. This is still significantly lower than that for P3E in the Saccharomycotina subdivision (~29%, two-tailed t-test P = 4 × 10^{-8}; Supplementary table 3, Supplementary Material online). Since genes that are involved in mitochondrial translation are essential for mitochondrial biogenesis, we investigated the occurrence of PUF motif in mitochondrial ribosomal protein genes and found that about 21% mitochondrial ribosomal protein genes are bound by Puf4p in the Pezizomycotina species, but ~90% mitochondrial ribosomal protein genes are bound by Puf3p in the Saccharomycotina species (Supplementary fig. 4, Supplementary Material online). Therefore the regulation of mitochondrial genes by the PUF4 regulon in the Pezizomycotina subdivision may be less efficient than that of the PUF3 regulon in the Saccharomycotina subdivision.

Because mitochondrial genes in the Pezizomycotina and Saccharomycotina species are regulated by different posttranscriptional regulons, we investigated relationships between these two regulons. We grouped orthologous mitochondrial genes in the Pezizomycotina subdivision based on the number of species in the Pezizomycotina subdivision that have P4E in their 3′ UTR regions. For each orthologous gene category, we
calculated the percentage of species in the *Saccharomycotina* subdivision that have PUF3 regulation element. We found that the more species have P4E in mitochondrial orthologs in the *Pezizomycotina* subdivision, the more species have P3E in the *Saccharomycotina* subdivision (fig. 4B, analysis of variance, $F = 6.5$, $P$ value $= 3 \times 10^{-4}$).

**Fig. 3.** Deep functional divergence between Puf3p and Puf4p. (A) The phylogenetic tree of the PUM-HD domains from the Puf3p and Puf4p orthologs (bootstrap scores are presented in Supplementary fig. 3, Supplementary Material online). Glam-Puf protein is used as the out-group. The branches to Puf3p and Puf4p in each species were coded in red and blue, respectively. (B) Functional conservation of Puf3p and Puf4p orthologs in fungi. The conserved amino acids for the repeat #3 and #5 of the PUM-HD domain are displayed using ClustalX (Thompson et al. 1997). The red (C $\rightarrow$ T) and blue (R $\rightarrow$ C) arrows denote the conserved amino acids in Puf3p and Puf4p that are critical for different motif recognition between these two proteins. (C) Little functional overlapping between Puf3p and Puf4p. The $y$ axis denotes the ratio of genes having both P3E and P4E over the total number of genes with P3E and/or P4E in each species ($x$ axis).

**Fig. 4.** Posttranscriptional regulation of mitochondrial genes in fungi. (A) The occurrence of P3E and P4E in mitochondrial genes. The $x$ axis denotes each studied fungal species (the Group 4 species, which include the *Saccharomycotina* and *Pezizomycotina* subdivisions, were put in the middle of fig. 4). The $y$ axis denotes the percentage of all Puf3p or Puf4p target genes in each genome that are dedicated to mitochondrial regulation. (B) The correlation of Puf3p and Puf4p mitochondrial regulation in fungi. All mitochondrial genes were divided into four groups based on the number of species in the *Pezizomycotina* subdivision that have the P4E motif. The $x$ axis represents the group of mitochondrial genes. The $y$ axis is the average percentage of species in the *Saccharomycotina* subdivision that have P3E for each orthologous mitochondrial gene group. The mean and standard deviation for each category are shown.
Discussion

Functional Innovation of Posttranscriptional Regulons

Our results show that all fungal species in the Pezizomycotina, Schizosaccharomyces, and Basidiomycota subdivisions do not have functional enrichment of PUF4 posttranscriptional regulation for the nucleolar genes. We conducted further analysis for the nucleolar genes in five out-group species, including human, mouse, Caenorhabditis elegans, Drosophila melanogaster, and A. thaliana (fig. 5). Our results show that nucleolar genes do not have an enriched PUF4 regulatory motif in their 3′ UTR (fig. 5A). The results for the PUF3 motif are the same (fig. 5B). The parsimony principle suggests that the PUF4 posttranscriptional regulation for the nucleolar genes might be a novel feature that evolved only in the Saccharomycotina species. Based on a similar parsimonious argument, we also concluded that the enriched posttranscriptional regulation by PUF proteins for the mitochondrial genes may also represent evolutionary innovations in the Saccharomycotina and Pezizomycotina subdivisions because mitochondrial genes in the species of Schizosaccharomyces and Basidiomycota subdivisions, and the above five out-group species had no enrichment for either PUF3 or PUF4 posttranscriptional regulation (fig. 5C and D).

The nucleolus is the location for ribosome synthesis. It has been shown that despite extreme functional conservation, ribosomal protein genes had a very dynamic evolutionary process for regulation at the transcriptional level in various fungal species (Tanay et al. 2005). Interestingly, a novel gene transcriptional regulatory network, including a new protein domain in the transcription factor, Rap1p, and a new promoter motif for ribosomal protein genes, evolved in the Group 1 fungal species. Our results show that posttranscriptional regulation of the ribosomal protein genes in the nucleolus also evolved as a novel regulatory module in the same species group. The underlying biological cause for these concurrent regulatory innovations of ribosome biogenesis and regulation during evolution remains to be further investigated.

Functional Rewiring of Posttranscriptional Regulation in Fungi

Our results indicate that the function of the PUF4 posttranscriptional regulon might be rewired from binding mitochondrial genes in the Pezizomycotina clade (Group 3,
fig. 2) to binding nucleolar genes in the *Saccharomycotina* clade (Groups 1 and 2, fig. 2). We also hypothesized that the mitochondrial genes were rewired from being posttranscriptionally regulated by Puf4p in the *Pezizomycotina* clade to Puf3p in the *Saccharomycotina* clade because genes functioning in mitochondria are enriched for their respective target 3′ UTR motifs in these two clades. The switch of the binding motif in the 3′ UTR should have occurred individually in their target mitochondrial genes. We cannot directly rule out the possibility that two regulators, Puf3p and Puf4p, have functional redundancy in the common ancestor of the *Pezizomycotina* and *Saccharomycotina* clades. However, several lines of evidence argue against this possibility: First, PUF3 and PUF4 genes were formed by an ancient duplication event, either at or before the common ancestor of all fungal species. Second, two orthologs show clear divergence of amino acids that are functionally critical for different motif recognition in all fungal species. Third, target genes identified by RIP-Chip experiments for Puf3p and Puf4p overlap very little (<1%) in *S. cerevisiae* (Gerber et al. 2004). And fourth, the target genes with P3E or P4E also overlap very little in all studied fungal species (fig. 3C). We envisioned that it might be adaptive to rewire the mitochondrial posttranscriptional gene regulation in the *Saccharomycotina* clade for higher efficiency, which may fit the clades, respectively. We are currently conducting further experiments in different fungal species to test these hypotheses.

How genetic novelties are created in nature is a fundamental issue in evolutionary biology. Functional rewiring is one principal mechanism to generate genetic novelties by reusing preexisting modules (Jacob 1977; Jacob 2001; Long 2001; Carroll 2005). In nature, functional rewiring has been frequently used to reshape transcriptional networks (Rodriguez-Caso et al. 2005; Lavioie et al. 2010). Similar evolutionary processes also occurred in reorganizing protein domains (Long and Langley 1993; Bockaert and Pin 1999; Long et al. 2003; Kleine and Leister 2007), reprogramming DNA methylation (Mandrioli 2004), and rewiring gene function by reusing transposable elements (Jordan 2006). Our results show another way to rewire regulation at the posttranscriptional level. Together with our previous study on the evolution of Puf3p targets in fungi (Jiang et al. 2010), our results reveal new principles governing the evolution of novel posttranscriptional regulation and provide a new perspective on understanding how gene regulation evolves in nature.

**Supplementary Material**

Supplementary figures 1–4 and tables 1–3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

**Acknowledgments**

We thank Jian Lu and Hong Chen for discussions, and Jun Sun and Brandon Barker for reading the manuscript and their helpful comments. We are also very grateful to the editor and three reviewers for their valuable comments for the earlier version of this paper. This work was supported by startup funds from Cornell University, National Science Foundation grant DEB-0949556, and National Institutes of Health 1R01AI085286 awarded to Z.G.

**References**


