

# A PUF family portrait: 3'UTR regulation as a way of life

Marvin Wickens, David S. Bernstein, Judith Kimble and Roy Parker

In eukaryotic cells, mRNAs are exquisitely controlled, often through regulatory elements in their 3' untranslated regions (3'UTRs). Proteins that bind to those sites are key players in controlling mRNA stability, translation and localization. One family of regulatory proteins – the PUF proteins – are not only structurally related, but also bind to 3'UTRs and modulate mRNA expression in a wide variety of eukaryotic species. They do so either by enhancing turnover or repressing translation, and act combinatorially with other regulatory proteins. Here, we discuss the evolution, biological function and mechanisms of action of the PUF protein family, and suggest that a primordial function of PUF proteins is to sustain mitotic proliferation of stem cells.

Gene expression in eukaryotic cells is regulated through an intricate web of interconnected steps. Many mRNAs contain elements in their 3' untranslated regions (3'UTRs) that bind regulatory proteins. These RNA–protein complexes control mRNA localization, translation and stability. In this article, we focus on one family of 3'UTR regulatory proteins. The first two members of the family to be analyzed in detail were *Drosophila melanogaster* Pumilio and *Caenorhabditis elegans* FBF; hence, the group is referred to as PUF [1] or Pum-HD [2] proteins.

Human families sometimes share superficial traits through many generations – a face, a stature, a voice. More rarely, family traditions persist as well. Molecular families are no different, and from that perspective, PUF proteins are remarkable. They not only look alike, but they practice the same trade, binding to 3'UTRs and repressing mRNAs. Some repress expression by promoting degradation, and others interfere with translation. An understanding of PUF protein function provides insights into the possible ancestral roles of these proteins and how 3'UTR regulators work more generally.

## Defining features

### Sequences and structures

PUF proteins are typically characterized by the presence of eight consecutive Puf repeats, each of approximately 40 amino acids, often flanked by two related sequences, Csp1 and Csp2 (Fig. 1a) [1–3]. Each repeat has a 'core consensus' containing aromatic and basic residues (Fig. 1b). Certain amino acids alternate in a pattern that suggests a two-repeat unit [1]. The entire cluster of Puf repeats is required for RNA binding [1,2,4]. Remarkably, this same region also interacts with protein co-regulators [5–8], and is sufficient to rescue, to a large extent, the defects of a PUF protein mutant [4].

The recently determined structure of the Puf repeat region of *Drosophila* Pumilio [9] and a closely related human PUF protein [10] reveals a striking extended crescent (Fig. 1c) [9,10]. Each Puf repeat folds into a three-helix domain; these domains stack on one another to form a crescent covering one-third of the circumference of a 42 Å radius circle (Fig. 1c). The core consensus sequences in each Puf repeat form helices that are arranged like rungs on a ladder on the inside face of the protein and that align conserved amino acids in adjacent Puf repeats (Fig. 1b,d). The Csp1 and Csp2 regions cap a hydrophobic core, and resemble Puf repeats. Interestingly, the repeated Puf repeat structures resemble those of Armadillo repeats in importin- $\alpha$  and  $\beta$ -catenin, suggesting that this protein-folding pattern could be ancestral [9,10].

The crescent provides two extended surfaces: one probably contacts RNA, and the other proteins. Mutations that disrupt RNA binding lie on the inner, concave face, where the conserved aromatic and basic amino acids suggest base stacking and backbone interactions with RNA (Fig. 1d) [9,10]. The inner surface of Pumilio is approximately 90 Å in length, corresponding well with the predicted length of the extended binding site of 20–30 nucleotides for a monomer [11]. Mutations that disrupt binding of two protein partners of Pumilio, Nanos and Brat, lie on the outside face [6,8–10]. In particular, a loop between repeats 7 and 8 decorates the otherwise very regular outside surface of Pumilio and is important for Nanos binding (asterisk in Fig. 1c) [6]. Other PUF proteins possess longer and more variable 'loop' regions between other repeats, and could dock other proteins all along their outer surface.

## Evolution of the family

The PUF protein family is widespread, with members found in many eukaryotes, including both plants and animals, and unicellular and metazoan organisms (Fig. 2) [1–3]. Several species, including *C. elegans*, *Saccharomyces cerevisiae* and *Arabidopsis thaliana*, possess multiple PUF proteins, whereas *Drosophila* possesses only one. Vertebrates appear to have a small number of PUF proteins – humans and mice each have two. To date, none has been detected in eubacteria or Archaea. In overall organization, Puf repeat regions are located near the C-terminus of the proteins in which they reside; typically, the N-terminal regions are more divergent. Within a

Marvin Wickens\*  
David S. Bernstein  
Dept. of Biochemistry,  
University of Wisconsin,  
Madison, WI 53706, USA.

Judith Kimble  
Howard Hughes Medical  
Institute, University of  
Wisconsin, Madison,  
WI 53706, USA.  
\*e-mail: wickens@  
biochem.wisc.edu

Roy Parker  
Howard Hughes Medical  
Institute, University of  
Arizona, Tucson,  
AZ 85721, USA.

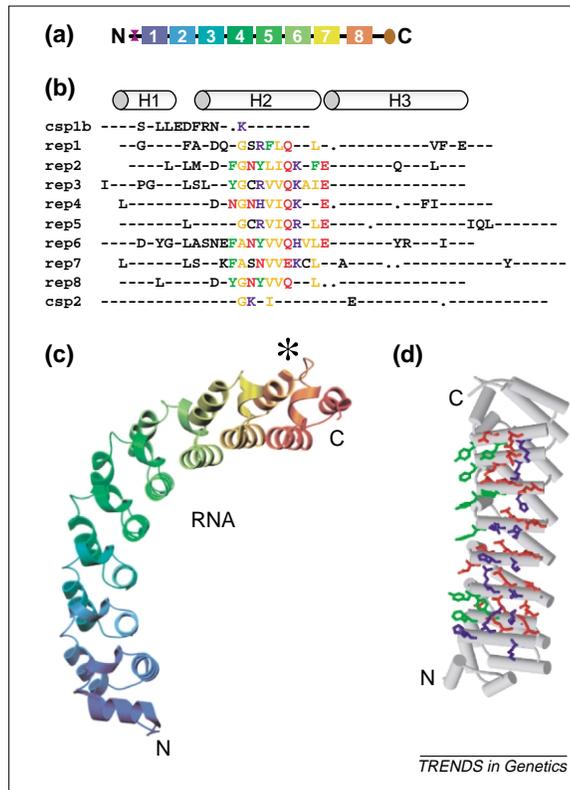


Fig. 1. Defining features: eight Puf repeats. (a) The RNA-binding domain of PUF proteins. The eight tandem repeats of approximately 40 amino acids are colored in boxes, and designated Puf repeats. These are flanked by two short sequences (anvil and oval) that form Puf repeat-like structures, named Csp1 (anvil) and Csp2 (oval). Supplementary information (including maps of other PUF proteins) is at <http://www.biochem.wisc.edu/wickens>. (b) Alignments and conserved sequences. A simplified version of a Hidden Markov Model of the entire Puf repeat region is presented. At the top are three helical units (H1-H3) present in each Puf repeat. Amino acids predicted to be present at a given position in at least 40% of all PUF proteins are shown; all other amino acids are indicated with dashes. Dots indicate spaces that are inserted for the alignment. The 'core consensus' comprises helix 2; conserved amino acids in that helix are colored: green, aromatic; blue, basic; red, acidic; yellow, aliphatic. Supplementary information (a more complete Markov model) is at <http://www.biochem.wisc.edu/wickens>. (c) Pumilio structure. Structure colored as in (a). The region and loop indicated with an \* is required for binding of a protein collaborator, Nanos [6,9,10]. RNA is thought to bind to the inside face of the crescent. Structure shown is that of *Drosophila* Pumilio, as presented in Ref. [9]. (d) The RNA-binding face of Pumilio. The orientation used in (c) is rotated 90 degrees, so that the conserved amino acids in helix 2 (the putative RNA-binding surface) face the reader. Conserved residues align in adjacent repeats, and are colored as in (b). Note the solvent exposed aromatic (green) and basic (blue) side chains, probably positioned to bind RNA. The structure shown is that of a human PUF, determined by Wang *et al.* [10]; we greatly appreciate their supplying pdb (protein database) coordinates before publication.

species, PUF proteins can form closely related subfamilies (e.g. *C. elegans* FBF-1 and FBF-2).

The dendrogram shown in Fig. 2 reveals two clusters of PUF proteins. One group – the 'Pumilio cluster' – includes *Drosophila* Pumilio, and PUFs from a wide range of other species. The second group contains nine of the eleven *C. elegans* PUF proteins, suggesting a burst of duplication and divergence in that species. Because FBF and Pumilio are only distantly related among the PUFs, their shared features could be common among family members [1].

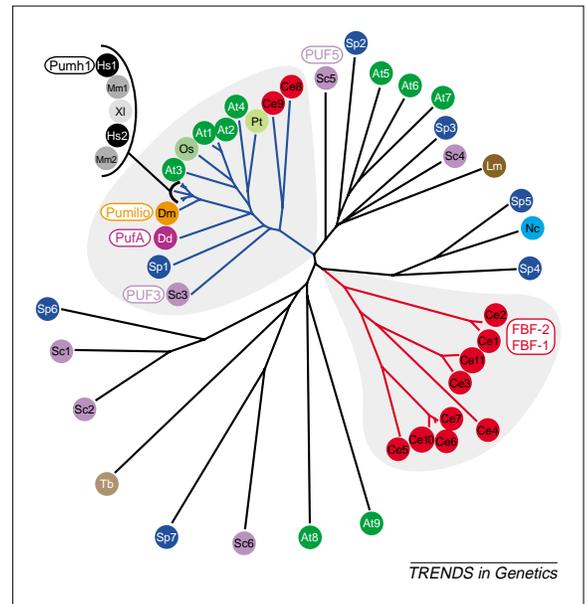


Fig. 2. PUF proteins throughout eukaryotes. An unrooted tree was derived by aligning only the Puf repeat regions using GCG. Partial expressed sequence tags (ESTs), which suggest many other PUF genes, were not included. PUF proteins specifically discussed in the text are highlighted. As much as possible, PUF proteins are named by their sequence relatedness within a species; thus PUF1 of yeast is more closely related to PUF2 of yeast than to the other yeast PUFs, but is not especially related to PUF1 of other organisms. Gray zones indicate the 'Pumilio cluster' (comprising *Drosophila* Pumilio, *C. elegans* PUF8 and 9, and several vertebrate PUFs) and the *C. elegans* cluster, containing nine of the 11 PUF proteins in that species (see text). Colors indicate different species. Blue to purple – fungi (*Saccharomyces cerevisiae*, Sc; *Schizosaccharomyces pombe*, Sp; *Neurospora crassa*, Nc; *Dictyostelium discoideum*, Dd). Gray to black – vertebrates (*Homo sapiens*, Hs; *Mus musculus*, Mm; *Xenopus laevis*, XI). Green – plants (*Arabidopsis thaliana*, At; *Oryza sativa*, Os; *Populus tremula* × *P. tremuloides* (poplar), Pt). Brown – trypanosomes (*Leishmania major*, Lm; *Trypanosoma brucei*, Tb). Red – *Caenorhabditis elegans*, Ce. Orange – *Drosophila melanogaster*, Dm. Accession numbers: Ce1 (FBF-1), AAF39879; Ce2 (FBF-2), Q09312; Ce3 (PUF-3), CAB63369; Ce4 (PUF-4), T33752; Ce5 (PUF-5), T22634; Ce6 (PUF-6), T21080; Ce7 (PUF-7), T32528; Ce8 (PUF-8), T15717; Ce9 (PUF-9), T26218; Ce10 (PUF-10), AAF60691; Ce11 (PUF-11), AAK68592; Dm (Pumilio), A46221; Lm, CAB62815; Tb, AAK62674; Sc1 (*JSN1*), P47135; Sc2 (PUF2), NP-015367; Sc3 (PUF3), NP-013088; Sc4 (PUF4), P25339; Sc5 (PUF5/MPT5/UTH4), P39016; Sc6 (PUF6), S69554; Sp1, CAB60694; Sp2, Q09829; Sp3, Q92359; Sp4, Q10238; Sp5, CAA20674; Sp6, CAA18887; Sp7, CAB54870; Nc, T49434; Dd (PufA), AAD39751; Hs1, AAG31807; Hs2, AAG31806; Mm1, AAG42319; Mm2, AAG31805; XI, BAB20864; At1, AAC95220; At2, AAC95216; At3, AAF02808; At4, AL049480; At5, BAA97177; At6, AAF87849; At7, CAB82120; At8, AAC28191; At9, AC007727; Os AAK73144; Pt, AAF71823.

## Biological functions

### Functional diversity

Based on the biological functions of the five PUF proteins that have been analyzed (*C. elegans* FBF, *Dictyostelium* PufA, *Drosophila* Pumilio and *S. cerevisiae* PUF3 and PUF5), PUF proteins control remarkably diverse processes (Table 1). For example, *Drosophila* Pumilio governs anterior/posterior patterning of the early embryo by repression of hunchback mRNA [3,12]. *C. elegans* FBF regulates the germline switch from spermatogenesis to oogenesis by repressing *fem-3* [1]. Yeast PUF proteins regulate aging, mitochondrial function and

Table 1. PUF proteins and their functions

Organism	PUF family member	Biological function	Interacting proteins	Target mRNA	Refs <sup>a</sup>
<i>Caenorhabditis elegans</i>	FBF <sup>b</sup>	Sperm/oocyte switch	NOS <sup>e</sup>	<i>fem-3</i>	[1,5]
	FBF <sup>b</sup>	Maintenance of germline stem cells	?	?	
	FBF <sup>b</sup>	Spermatogenesis	CPB-1	?	[7]
	FBF/PUF-8/PUF-6,-7,-10 <sup>b,c</sup>	Mitotic arrest of primordial germ cells	NOS <sup>e</sup>	?	[17]
	FBF/PUF-8/PUF-6,-7,-10 <sup>b,c</sup>	Migration of primordial germ cells	NOS <sup>e</sup>	?	[17]
<i>Drosophila</i>	FBF/PUF-8/PUF-6,-7,-10 <sup>b,c</sup>	Germline survival	NOS <sup>e</sup>	?	[5,17]
	Pumilio	Anterior–posterior axis of embryo	Nos/Brat	<i>hunchback</i>	[3,6,8,56]
	Pumilio	Mitotic arrest of primordial germ cells	Nos	cyclin B	[25]
	Pumilio	Migration of primordial germ cells	Nos	?	[19,25,63]
<i>Saccharomyces cerevisiae</i>	Pumilio	Maintenance of germline stem cells	?	?	[18,19]
	Puf5p/Mpt5p/Uth4p <sup>d</sup> and Puf4p/Ygl014wp/Ygl023p	Lifespan (continued mother cell divisions)	?	?	[13]
	Puf5p/Mpt5p/Uth4p <sup>d</sup>	Mating-type switching	?	HO	[15]
	Puf5p/Mpt5p/Uth4p <sup>d</sup>	Pheromone response and G1 arrest	?	?	[59]
<i>Dictyostelium</i>	Puf3p	Mitochondrial function	?	COX17	[14]
	PufA	Continued mitotic divisions during vegetative growth	?	PKA-C	[16]
<i>Xenopus</i>	X-Puf1	Present in oocytes: function unknown	NOS/CPEB	cyclin B	[29]

<sup>a</sup>For brevity, references have been selected to provide an entry into the literature and are not comprehensive. Our apologies to those not included.  
<sup>b</sup>FBF comprises two proteins, FBF-1 and FBF-2, which are 93% identical [1]. Their functions have not been distinguished.  
<sup>c</sup>PUF-6, -7 and -10 are nearly identical; their functions have not been distinguished.  
<sup>d</sup>PUF5 was isolated as a multi-copy suppressor of *pop2* (hence designated *MPT5*) [47] and as a mutant decreasing yeast life span (hence *UTH4*) [13].  
<sup>e</sup>The three *C. elegans* NOS genes (*NOS-1*, *NOS-2* and *NOS-3*) are functionally redundant [5]. *NOS-3* protein interacts with FBF-1 and FBF-2 proteins, while *NOS-1* and *NOS-2* proteins interact with PUF-6, -7, -10 proteins (D. Bernstein *et al.*, unpublished).

mating-type switching [13–15], whereas *Dictyostelium* PufA promotes vegetative growth [16].

Within a species, a single PUF protein can have multiple functions, presumably by regulating multiple mRNAs. For example, *C. elegans* FBF is required not only for the sperm/oocyte switch, but also for maintenance of germline stem cells (S. Crittenden *et al.*, unpublished) and (together with four other PUF proteins) for germline survival, germline migration and germline mitotic arrest during embryogenesis [5,17]. Germline functions are probably ancestral, because *Drosophila* Pumilio also has key roles in germline migration and mitotic arrest during embryogenesis [18–22].

#### A common root: maintenance of stem cells

Despite the diversity of PUF functions (Table 1), PUF proteins share a common role in each species studied – maintenance of stem cells (Fig. 3). Thus, without Pumilio, female *Drosophila* germline stem cells differentiate prematurely as cystoblasts [18,19], and without FBF, *C. elegans* germline stem cells enter meiosis and undergo spermatogenesis (S. Crittenden *et al.*, unpublished; Ref. [1]) Therefore, in two distantly diverged metazoans, PUF proteins promote proliferation of germline stem cells and repress their differentiation.

What of organisms with no identifiable germ line? In this context, the role of PufA in the slime mold *Dictyostelium* is intriguing [16]. Normally, single *Dictyostelium* cells divide mitotically until, in response to nutrient deprivation, they aggregate and differentiate to form fruiting bodies containing stalk and spore cells. The PufA protein promotes continued vegetative divisions and inhibits differentiation by repressing an mRNA encoding the catalytic subunit

of cAMP-dependent protein kinase (PKA-C), which is required for differentiation [16]. In PufA mutants, cells precociously leave mitosis and differentiate [16]. Thus, PufA promotes proliferation in a manner analogous to Pumilio and FBF: proliferating single cells are viewed as stem cells that differentiate in response to an external cue.

A second provocative connection between stem cell controls and PUF proteins comes from studies of aging in *S. cerevisiae* [13]. Yeast aging is measured by counting the number of progeny generated by a single mother cell before dying [23]. Strains lacking *PUF5* (also called *MPT5* and *UTH4*) divide fewer times than wild type, whereas strains overexpressing *PUF5* divide more times and live longer [13]. In this case, one can view the mother cell as a stem cell, whose mitosis is not properly sustained in the absence of the PUF protein. Strikingly, the inhibition of PKA activity increases lifespan in yeast [24]. This leads to the suggestion that, like PufA in *Dictyostelium*, yeast *PUF5* might normally repress expression of PKA. An extension of this speculation is that PUF proteins might downregulate PKA in both *Drosophila* and *C. elegans*, thereby suggesting a crucial and ancestral role of cAMP-dependent protein kinase in controlling germline development.

On the basis of these observations, we suggest that an ancestral function of PUF proteins is to support the mitotic proliferation of stem cells. Anterior–posterior patterning, the sperm/oocyte switch, and regulation of mating-type switching are all later evolutionary accretions. In *Drosophila* embryos, Pumilio actually retards rather than promotes the cell cycle in pole cells (the presumptive germ line), probably by repressing cyclin B

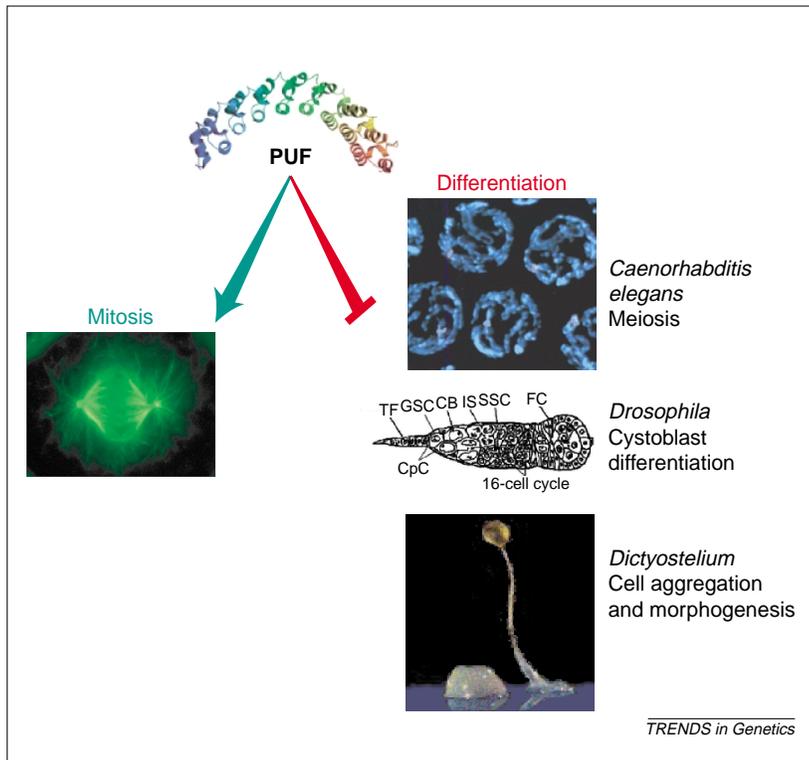


Fig. 3. PUF proteins promote mitotic proliferation of stem cells and suppress differentiation. In this model, PUF proteins promote mitosis of stem cells, and thereby repress differentiation. PUF proteins prevent meiosis in the *Caenorhabditis elegans* germ line, cystoblast differentiation in *Drosophila*, and fruiting body formation in *Dictyostelium*. These are viewed as alternatives to the continued proliferation of stem cells, and occur precociously in the absence of PUF proteins. See text for details. Images modified from originals by G. Borisy, A. Demberg, R. King and R. Kay.

mRNA [22,25,26]. It will be interesting to determine whether a PUF protein switches between positive and negative regulation of the cell cycle through different actions on a single mRNA, or by regulating multiple mRNA targets. In the many species in which PUF protein function has yet to be determined, we anticipate roles in other (perhaps unorthodox) stem cell populations.

#### Multiple targets and partners: combinatorial regulation mRNA targets

Individual PUF proteins regulate multiple mRNAs. This conclusion rests on two lines of evidence. First, Pumilio binds at least two mRNAs in *Drosophila* [8]: *hunchback* mRNA, which participates in embryonic patterning, and cyclin B mRNA, whose repression retards the cell cycle in pole cells [25,26]. Second, genetic elimination of a PUF protein causes a broader range of phenotypes than elimination of the binding site in a single mRNA target. For example, lack of an FBF-binding site in *fem-3* in *C. elegans* affects only the sperm/oocyte switch, but lack of FBF itself also perturbs mitosis (S. Crittenden *et al.*, unpublished) and spermatogenesis [7]. Similarly, yeast *puf5* mutants display a broad range of phenotypes, only one of which is accounted for by misregulation of *HO* mRNA, which is specifically involved in switching mating type (Table 1).

The mRNA regulated by a specific PUF protein will be determined in part by the binding specificity of that PUF protein. Individual PUF proteins bind selectively to different RNAs and are affected differentially by changes in the binding site (D. Bernstein *et al.*, unpublished; Ref. [2]). For example, human and *Drosophila* PUF proteins differ in their sensitivity to base changes in the RNA, despite the fact that the two proteins are 80% identical in their RNA-binding regions [2]. Thus, in organisms with multiple PUF proteins, different populations of transcripts might be recognized by subsets of the proteins. There are, however, common features among PUF protein binding sites [1,2]. In all cases analyzed so far, a UGUR tetranucleotide is essential for PUF proteins to bind to RNA [1,2,4,12,15]; flanking nucleotides determine specificity. The formation of PUF protein complexes on 3'UTRs can also require additional sequences, due to other RNA-binding proteins (see below) [6].

#### Protein partners

The specificity and consequence of a PUF–mRNA interaction will be determined partly by other proteins that bind to the PUF polypeptide. To date, PUF proteins are known to interact physically with members of three other protein families. Strikingly, these families consist of known 3'UTR regulators. The ability of PUF proteins to interact with additional 3'UTR-binding proteins allows the assembly of distinct PUF protein complexes, which might differentiate between different mRNA sequences and lead to different biological outcomes.

(1) Nanos (NOS) proteins. Nanos proteins are characterized by two distinctive CCHC zinc fingers, and bind RNA nonspecifically in the absence of a PUF protein [27,28]. FBF binds NOS-3, one of three *C. elegans* NOS proteins [5], and Pumilio interacts with the single NOS of *Drosophila* [6]. The significance of the interactions is corroborated by the overlapping phenotypes of mutants in the PUF and NOS proteins (Table 1), and, most compellingly, by the analysis of specific mutations in Pumilio that are defective in the interaction with NOS [6]. The other NOS proteins of *C. elegans*, NOS-1 and NOS-2, interact with different PUF proteins, emphasizing the commonality of PUF–NOS partnerships, as well as their specificity (D. Bernstein *et al.*, unpublished). Similarly, a *Xenopus* PUF binds to a *Xenopus* NOS homolog, XCAT-2, in frog oocytes [29].

FBF–NOS-3 and Pumilio–Nanos complexes probably differ in detail, as judged by both two- and three-hybrid experiments and biochemical analyses. Different regions of the *C. elegans* and *Drosophila* NOS proteins bind FBF and Pumilio [5,6]. Moreover, the balance between protein–protein and protein–RNA interactions apparently differs between the two complexes: the FBF–NOS-3 interaction is RNA-independent [5],

whereas the Pumilio–NOS interaction can only be detected in a ternary complex containing RNA [6].

(2) Cytoplasmic polyadenylation element-binding proteins (CPEB). CPEB proteins possess two RRM domains followed by CCCC and CCHC zinc fingers that are required for RNA binding [30,31]. Homologs are found among all metazoans examined, and bind to U-rich sequences in 3'UTRs, regulating polyadenylation and translation [31]. FBF physically interacts with a *C. elegans* CPEB homolog, CPB-1 [7]. Reduced activity in either partner cause spermatogenesis defects, though not at identical steps [7]. A *Xenopus* PUF protein binds to CPEB in oocytes, though the function of the complex has yet to be determined [29].

(3) Brat. *Drosophila* Brat is a member of the NHL protein family, characterized by two to six repeats of a 44-residue sequence rich in glycine and hydrophobic residues plus a cluster of charged amino acids [32]. Brat is recruited to *hunchback* mRNA through a ternary complex of Pumilio, Nanos and the mRNA [8]. Single amino acid substitutions in Brat compromise quaternary complex formation and disrupt regulation of *hunchback in vivo*. Two well-characterized genes in *C. elegans*, *ncl-1* and *lin-41*, also contain NHL domains: *ncl-1* regulates nucleolar size and rRNA abundance [33]; *lin-41* represses differentiation and promotes mitosis of specific somatic cells [34], a role that echoes PUF functions in stem cells. However, no genetic or physical interaction between *C. elegans* PUF and NHL proteins has been observed, nor are NHL proteins other than Brat known to participate in 3'UTR regulation.

The multiplicity of partners interact with the Puf repeat region of the protein. The Puf repeat regions of *Drosophila* Pumilio and *C. elegans* FBF mediate interactions with Nanos proteins [5,6], and the Puf repeat region of FBF also is sufficient for interaction with CPEB [7]. The interaction region has been mapped most precisely with Pumilio: small perturbations in a loop near the eighth Puf repeat disrupt its interaction with Nanos (Fig. 1) [6]. Other protein collaborators probably interact with the PUF protein elsewhere on its extended outer surface (Fig. 1; see above). In addition, many PUF proteins contain substantial regions outside the Puf repeats, toward the N-terminus, which might have additional functions.

#### Models of mechanism

In the cases tested thus far, PUFs decrease expression of target mRNAs, either by repressing their translation or enhancing their decay. We suggest that PUFs pull the same molecular trigger in both cases, and that the difference is the particular cell's balance between translation and turnover pathways.

In oocytes and early embryos, most mRNAs with short (or no) poly(A) tails are stable, but translationally inactive [31,35]; mRNA decay turns

on later during development. By contrast, in yeast, mRNAs with little or no poly(A) are rapidly degraded. This decay after deadenylation is due to enzymatic cleavage of the 7mGpppG cap ('decapping') and degradation of the mRNA by a 5' to 3' exonuclease [36]. In one view, the difference between the behavior of deadenylated mRNAs in embryos and yeast arises simply because the activity of the deadenylation-dependent decapping pathway is higher in yeast [37]. For example, if cap-binding protein (eIF-4E) were prevented from binding to the cap, this might expose the cap and cause decay in yeast and repression in the embryo. Thus, turnover and translational repression could result from the same, single trigger. The focus then becomes its identity.

We consider two separate issues. First, we discuss the evidence that PUF proteins can decrease expression both by promoting deadenylation and by other mechanisms. Second, we consider how PUF proteins accelerate removal of poly(A), and suggest that PUF proteins might mediate transitions from active to inactive mRNA–protein (messenger ribonucleoprotein; mRNP) states.

#### Deadenylation-dependent and deadenylation-independent mechanisms

Decreases in expression by PUF proteins are correlated with removal of the poly(A) tail. For example, *fem-3* and *hunchback* mRNAs carrying mutations that disrupt PUF protein binding have longer tails than their wild-type counterparts [38,39]. Conversely, the poly(A) tail of an mRNA (*bicoid*) carrying a PUF-binding site is shorter when exposed to ectopic Nanos *in vivo* [40]. The rate of deadenylation of target mRNAs is enhanced by Pumilio and Nanos in *Drosophila* [39], and by Puf3p in yeast [14]. In all these cases, the PUF protein could either facilitate deadenylation or repress cytoplasmic polyadenylation [41]. Because cytoplasmic polyadenylation is not known to occur in yeast, the finger points at deadenylation.

An important distinction is whether PUF proteins directly promote deadenylation, which then leads to repression of translation in certain conditions, or whether PUF proteins inhibit translation, which then leads to enhanced deadenylation rates [42,43]. In yeast, Puf3p accelerates deadenylation of *COX17* mRNA without affecting its polysome association, suggesting a direct effect on deadenylation [14]. In *Drosophila* embryos, Pumilio can promote deadenylation of mRNA fragments that are not expected to be translated [39]. Thus, PUF proteins appear to enhance deadenylation independent of any effect on translation. This is an appealing molecular trigger for both turnover and translational repression.

PUF proteins also act through other mechanisms, however. For example, yeast PUF3 enhances decapping after deadenylation [14], suggesting that PUF proteins can affect events at the 5' end

independent of the poly(A) tail. Similarly, Pumilio can repress *hunchback* mRNAs *in vivo* through a mechanism that does not require a poly(A) tail [44]. These observations indicate that PUF proteins either have multiple functions, or that the mechanism by which they enhance deadenylation rates might also enhance decapping rates and repress translation (see below).

#### Accelerating deadenylation and changing mRNP organization

Recent work suggests that the same few deadenylases appear to act on most mRNAs. For example, Ccr4p/Caf1p, a major deadenylase in yeast, is required for the slow deadenylation of stable mRNAs, and for the rapid deadenylation of unstable mRNAs, including reactions enhanced by PUF proteins [45]. How do PUF proteins make these deadenylases act preferentially on specific mRNAs? In principle, the PUFs could attract the deadenylase (Fig. 4a), or could alter the mRNP to make it a good substrate for the enzyme (Fig. 4b). In either case, the ultimate effect on the mRNA would be the same: the PUF protein would change an mRNP organization that was translationally competent and inhibited for deadenylation, into one in which the mRNA was translationally repressed and accessible to deadenylation (Fig. 4).

In the first hypothesis (Fig. 4a), the PUF protein bound to the 3'UTR recruits a general deadenylase with little specificity for any particular mRNA. The rate of deadenylation is enhanced on a particular mRNA, much as the rate of polyadenylation by poly(A) polymerase is enhanced by RNA-binding specificity factors [46]. The PUF could bind the deadenylase directly, or be linked through intermediates. Removal of poly(A) effectively removes poly(A)-binding protein (PAB), and thereby disrupts the mRNP. Interestingly, overexpression of yeast *PUF5* suppresses some of the phenotypes of a deletion of a deadenylase component (*CAF1/POP2*) [47]. The large network of interactions among deadenylation and turnover components (e.g. see Ref. [48]) suggests that a PUF could grab the entire machinery by clutching any one of many handles. In this model, additional mechanisms must be invoked to explain the effects of PUF proteins on translation of nonpolyadenylated mRNAs [44], and the ability of PUF proteins to stimulate decapping [14].

In the second hypothesis (Fig. 4b), PUFs affect the state of the mRNP, and thereby cause changes in poly(A) length. In one incarnation of this model, PUF proteins promote dissociation of PAB from poly(A), exposing it to deadenylases. Several other 3'UTR-binding proteins bind PAB (e.g. see Refs [49,50]). More generally, because the two ends of the mRNA can interact through a PAB-eIF-4G-eIF-4E bridge [51], perturbations at either end of the mRNA could affect mRNP structure globally. Interestingly, the Dhh1p protein, a DEAD-box helicase, functions in

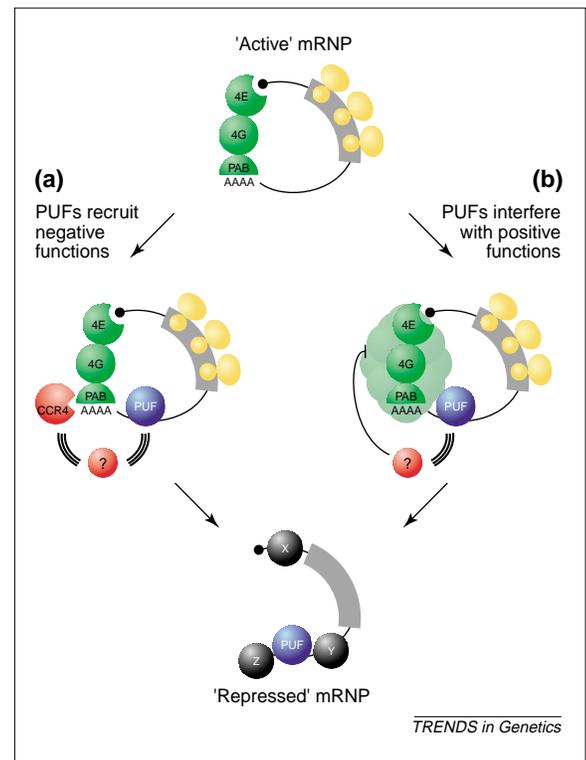


Fig. 4. Models for PUF protein-mediated repression of mRNAs. PUF proteins convert an active mRNA-protein particle – ‘active mRNP’ (messenger ribonucleoprotein), top – into a ‘repressed mRNP’ (bottom). mRNAs in the active complex are translated, stable and deadenylated only slowly; mRNAs in the repressed complex are translationally silent, unstable and deadenylated rapidly. The transition from one form to another can occur through either of two paths, (a) and (b). Path (a): PUFs recruit the deadenylase. The number and nature of the interactions is not known, and recruitment of the deadenylase could simultaneously recruit components of the turnover machinery. Path (b): PUF proteins directly trigger a change in mRNP organization. The key feature of this model is that PUF proteins inhibit the action of positive-acting components. Although the eIF-4E–eIF-4G–poly(A)-binding protein (PAB) linkage (and associated translational components) is depicted (green), other candidates exist. Blue, PUF protein; green, positive-acting factors, including the translation initiation factors, eIF-4E and eIF-4G; red, negative-acting factors including the deadenylase (Ccr4p); black, components of the repressed mRNP that are allowed access to the mRNA after the transition.

translational repression of mRNA during early metazoan development and promotes decapping in yeast [48,52–54], and interacts with the yeast general turnover machinery. Its involvement in both turnover and translational repression, and the fact that such helicases can mediate changes in RNA-protein complexes, suggests that Dhh1p could be critical in the mRNP transitions depicted in Fig. 4b. An appealing aspect of the model in which PUFs trigger a global change in mRNP structure is that one event can explain stimulation of both deadenylation and decapping. The two types of model are not mutually exclusive: the deadenylase could be part of a complex that also triggers an mRNP transition.

Physiological transitions mediated by PUF proteins, as suggested by Fig. 4, have not been demonstrated directly. Rather, most of the evidence on how PUFs work comes from comparing mRNAs in cells with or without a PUF. Thus PUF proteins could

drive an mRNA into the 'repressed' mRNP state without its ever having been translated.

### Regulating the regulator

Understanding how a PUF protein works or interacts with its partners is only the first step. How are PUFs controlled? Thus far, the mere abundance of the PUF protein appears not to be the key determinant. Instead, regulation of the partners could be critical. For example, the localized repression of *hunchback* mRNA in the presumptive posterior of a fly embryo requires Pumilio and Nanos, yet Pumilio is present throughout the syncytial embryo [55]. Nanos protein is present at high levels only in the posterior [56,57], and that spatial asymmetry underlies the localized repression of *hunchback* mRNA. The appearance or distribution of partners could control the activity of PUFs in many other contexts, making the dynamics of interactions a key issue. How many distinct complexes are there *in vivo*, and of what composition? Are the interactions of different partners with one PUF compatible or mutually exclusive? When and where do the interactions occur?

Cells might also regulate PUF proteins through a variety of cellular kinases and signaling pathways. In yeast, two MAP kinases (Fus3p and Kss1p), the cell cycle kinase Cdc28p, and protein kinase C (PKC) each interact genetically with *PUF5*, and bind to Puf5p in two-hybrid experiments; interactions with the MAP kinases and Cdc28p have been confirmed by co-immunoprecipitation [47,58,59]. A YAK kinase regulates slime mold PufA, which in turn controls cAMP-dependent protein kinase mRNA [16]. Mitosis in the *C. elegans* germ line is regulated by the Notch signaling pathway [60], which must feed into FBF's role in proliferation. How these signaling pathways impinge on the PUFs is unclear, but could hold important clues to the regulation of PUF action more generally.

### Prospects and implications

PUF proteins share a remarkable array of features, aside from their anatomical similarities. They repress expression by affecting translation and turnover, and interact with related (albeit distinct) sequences. These features are probably not the whole story: PUF proteins might regulate other steps in mRNA function and might sometimes activate rather than repress. Nevertheless, the commonality of the connections to repression and deadenylation crystallize key issues. How do PUFs promote deadenylation, and does that cause repression?

How many layers of interactors are there between the PUFs and the basal machineries, or the mRNP's structure? How do the many interactors with a single PUF protein influence one another, and RNA-binding specificity? There are echoes here of prominent questions in transcriptional control, and solutions probably will come from an analogous mingling of genetic, biochemical and structural analysis focused on PUF protein complexes.

PUF proteins embody three features that might apply to 3'UTR regulators more generally. First, they act with partners, some of which are already known to be 3'UTR regulators in their own right. Although PUF proteins are present throughout the eukaryotes, their collaborators have been characterized thus far only in metazoans. One suspects that similar collaborators hide in the genomes of unicellular organisms and plants, and have simply not been spotted yet. Alternatively, collaborators could be a metazoan invention. Second, regulation of poly(A) length might be a common mode of PUF action. The connection between regulated changes in poly(A) length and translational activity has been discussed elsewhere [31,35], and will not be recapitulated here. However, an important emerging conclusion is that the PUF proteins can operate through both poly(A)-dependent and poly(A)-independent mechanisms. Third, combinatorial regulation is probably common. In particular, multiple mRNAs will be targets of a single regulator that engages in multiple partnerships. Moreover, the RNA with which a PUF protein interacts could dictate the identity of the protein complex. For example, Brat is recruited by Pumilio and NOS bound to *hunchback*, but not cyclin B mRNA [8]. The implied allosteric change has been observed in other RNA-binding proteins [61] and DNA-binding transcription factors [62]. In transcription, and perhaps with PUFs, the nucleic acid ligand can determine whether the complex activates or represses. Regardless, PUF-mediated controls parallel well-documented principles of transcriptional regulation, in which distinct protein-protein interactions discriminate among promoters and yield unique biological outcomes.

The study of PUF proteins has been invigorated by the unexpected convergence of many approaches and species. Genetics, biochemistry and structural biology will continue to provide a rich picture of the central molecules; cell and developmental biology will enrich the biological perspective. The outline of a new PUF family portrait has been drawn, but much blank canvas remains.

### Acknowledgements

We thank Phil Zamore and Kathy Tanaka-Hall and their colleagues for the communication of results before publication, including protein database (pdb) coordinates of Pumilio. We are also grateful to Liaoteng Wang for pointing out *C. elegans* PUF11 and relevant literature. We also thank several groups working on PUF proteins, including Drs Gray, Herskowitz, Irie, Lehmann, Wharton and Williamson, for discussions and communication of results before publication. We thank members of the Wickens, Kimble and Parker labs for comments on the manuscript and discussions of many topics covered here. We are grateful to Carol Pfeffer for preparing the manuscript, and to the UW Biochemistry Media Lab for the figures. Work in the Wickens laboratory is supported by the NIH (GM31892 and GM50942). J.K. and R.P. are investigators of the Howard Hughes Medical Institute.

### References

- Zhang, B. *et al.* (1997) A conserved RNA binding protein that regulates patterning of sexual fates in the *C. elegans* hermaphrodite germ line. *Nature* 390, 477–484
- Zamore, P.D. *et al.* (1997) The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* 3, 1421–1433
- Barker, D. *et al.* (1992) *Pumilio* is essential for function but for distribution of the *Drosophila* abdominal determinant nanos. *Genes Dev.* 6, 2312–2326
- Wharton, R.P. *et al.* (1998) The Pumilio RNA-binding domain is also a translational regulator. *Mol. Cell* 1, 863–872
- Kraemer, B. *et al.* (1999) NANOS-3 and FBF proteins physically interact to control the sperm/oocyte switch in *C. elegans*. *Curr. Biol.* 9, 1009–1018
- Sonoda, J. *et al.* (1999) Recruitment of Nanos to *hunchback* mRNA by Pumilio. *Genes Dev.* 13, 2704–2712
- Luitjens, C. *et al.* (2000) CPEB proteins control two key steps in spermatogenesis in *C. elegans*. *Genes Dev.* 14, 2596–2609
- Sonoda, J. *et al.* (2001) *Drosophila* Brain Tumor is a translational repressor. *Genes Dev.* 15, 762–773

- 9 Edwards, T.A. *et al.* (2001) Structure of pumilio reveals similarity between RNA and peptide binding motifs. *Cell* 105, 281–289
- 10 Wang, X. *et al.* (2001) Crystal structure of a Pumilio homology domain. *Mol. Cell* 7, 855–865
- 11 Zamore, P.D. *et al.* (1999) The PUMILIO–RNA interaction: a single RNA-binding domain monomer recognizes a bipartite target sequence. *Biochemistry* 38, 596–604
- 12 Murata, Y. *et al.* (1995) Binding of pumilio to maternal hunchback mRNA is required for posterior patterning in *Drosophila* embryos. *Cell* 80, 747–756
- 13 Kennedy, B.K. *et al.* (1997) Redistribution of silencing proteins from telomeres to the nucleolus is associated with extension of life span in *S. cerevisiae*. *Cell* 89, 381–391
- 14 Olivas, W. *et al.* (2000) The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast. *EMBO J.* 19, 6602–6611
- 15 Tadauchi, T. *et al.* (2001) Post-transcriptional regulation through the HO 3'-UTR by Mpt5, a yeast homolog of Pumilio and BFB. *EMBO J.* 20, 552–561
- 16 Souza, G.M. *et al.* (1999) Starvation promotes *Dictyostelium* development by relieving PufA inhibition of PKA translation through the Yaka kinase pathway. *Development* 126, 3263–3274
- 17 Subramaniam, K. *et al.* (1999) *nos-1* and *nos-2*, two genes related to *Drosophila nanos*, regulate primordial germ cell development and survival in *Caenorhabditis elegans*. *Development* 126, 4861–4871
- 18 Lin, H. *et al.* (1997) A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. *Development* 124, 2463–2476
- 19 Forbes, A. *et al.* (1998) Nanos and Pumilio have critical roles in the development and function of *Drosophila* germline stem cells. *Development* 125, 679–690
- 20 Bhat, K.M. (1999) The posterior determinant gene *nanos* is required for the maintenance of the adult germline stem cells during *Drosophila* oogenesis. *Genetics* 151, 1479–1492
- 21 Parisi, M. *et al.* (1999) The *Drosophila pumilio* gene encodes two functional isoforms that play multiple roles in germline development, gonadogenesis, oogenesis and embryogenesis. *Genetics* 153, 235–250
- 22 Parisi, M. *et al.* (2000) Translational repression: a duet of Nanos and Pumilio. *Curr. Biol.* 10, R81–R83
- 23 Mortimer, R.K. *et al.* (1959) Life span of individual yeast cells. *Nature* 183, 1751–1752
- 24 Lin, S.J. *et al.* (2000) Requirement of NAD and SIR2 for life-span extension by calorie restriction in *Saccharomyces cerevisiae*. *Science* 289, 2126–2128
- 25 Asaoki-Taguchi, M. *et al.* (1999) Maternal Pumilio acts together with Nanos in germline development in *Drosophila* embryos. *Nat. Cell Biol.* 1, 431–437
- 26 Deshpande, G. *et al.* (1999) Novel functions of *nanos* in downregulating mitosis and transcription during the development of the *Drosophila* germline. *Cell* 99, 271–281
- 27 Curtis, D. *et al.* (1995) Nanos is an evolutionarily conserved organizer of anterior–posterior polarity. *Development* 121, 1899–1910
- 28 Curtis, D. *et al.* (1997) A CCHC metal-binding domain in Nanos is essential for translational regulation. *EMBO J.* 16, 834–843
- 29 Nakahata, S. *et al.* (2001) Biochemical identification of *Xenopus pumilio* as a sequence-specific cyclin B1 mRNA-binding protein that physically interacts with a Nanos homolog, Xcat-2, and a cytoplasmic polyadenylation element-binding protein. *J. Biol. Chem.* 276, 20945–20953
- 30 Hake, L.E. *et al.* (1998) Specificity of RNA binding by CPEB: requirement for RNA recognition motifs and a novel zinc finger. *Mol. Cell Biol.* 18, 685–693
- 31 Richter, J.D. (2000) Influence of polyadenylation-induced translation on metazoan development and neuronal synaptic function. In *Translational Control of Gene Expression* (Hershey, J.W.B. *et al.*, eds), pp. 785–805, Cold Spring Harbor Laboratory Press
- 32 Slack, F.J. *et al.* (1998) A novel repeat domain that is often associated with RING finger and B box motifs. *Trends Biochem. Sci.* 23, 474–475
- 33 Frank, D.J. *et al.* (1998) *ncl-1* is required for the regulation of cell size and ribosomal RNA synthesis in *Caenorhabditis elegans*. *J. Cell Biol.* 140, 1321–1329
- 34 Slack, F.J. *et al.* (2000) The *lin-41* RBCC gene acts in the *C. elegans* heterochronic pathway between the *let-7* regulatory RNA and the *lin-29* transcription factor. *Mol. Cell* 5, 659–669
- 35 Wickens, M. *et al.* (2000) Translational control of developmental decisions. In *Translational Control of Gene Expression* (Hershey, J.W.B. *et al.*, eds), pp. 295–370, Cold Spring Harbor Laboratory Press
- 36 Tucker, M. *et al.* (2000) Mechanisms and control of mRNA decapping in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.* 69, 571–595
- 37 Wickens, M. *et al.* (1997) Life and death in the cytoplasm: messages from the 3' end. *Curr. Opin. Genet. Dev.* 7, 220–232
- 38 Ahringer, J. *et al.* (1991) Control of the sperm–oocyte switch in *Caenorhabditis elegans* hermaphrodites by the *fem-3*' untranslated region. *Nature* 349, 346–348
- 39 Wreden, C. *et al.* (1997) Nanos and pumilio establish embryonic polarity in *Drosophila* by promoting posterior deadenylation of hunchback mRNA. *Development* 124, 3015–3023
- 40 Wharton, R. *et al.* (1991) RNA regulatory elements mediate control of *Drosophila* body pattern by the posterior morphogen *nanos*. *Cell* 67, 955–967
- 41 Wickens, M. (1992) Forward, backward, how much, when: mechanisms of poly(A) addition and removal and their role in early development. *Semin. Dev. Biol.* 3, 399–412
- 42 Muhlrud, D. *et al.* (1995) Turnover mechanisms of the stable yeast *PGK1* mRNA. *Mol. Cell Biol.* 15, 2145–2156
- 43 Muckenthaler, M. *et al.* (1997) Regulated poly(A) tail shortening in somatic cells mediated by cap-proximal translational repressor proteins and ribosome association. *RNA* 3, 983–995
- 44 Chagnovich, D. *et al.* (2001) Poly(A)-independent regulation of maternal hunchback translation in the *Drosophila* embryo. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11359–11364
- 45 Tucker, M. *et al.* Ccr4p is the catalytic subunit of the Ccr4p/Not complex mRNA deadenylase in *Saccharomyces cerevisiae*. *EMBO J.* (in press)
- 46 Wahle, E. *et al.* (1996) The biochemistry of polyadenylation. *Trends Biochem. Sci.* 21, 247–250
- 47 Hata, H. *et al.* (1998) Dhh1p, a putative RNA helicase, associates with the general transcription factors Pop2p and Ccr4p from *Saccharomyces cerevisiae*. *Genetics* 148, 571–579
- 48 Collier, J. *et al.* (2001) The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* 7, 1717–1727
- 49 Wang, Z. *et al.* (2000) The poly(A)-binding protein and an mRNA stability protein jointly regulate an endoribonuclease activity. *Mol. Cell Biol.* 20, 6334–6341
- 50 Laroia, G. *et al.* (1999) Control of mRNA decay by heat shock-ubiquitin-proteasome pathway. *Science* 284, 499–502
- 51 Sachs, A.B. (2000) Cell cycle-dependent translation initiation: IRES elements prevail. *Cell* 101, 243–245
- 52 Nakamura, A. *et al.* Me31B silences translation of oocyte-localising RNAs through the formation of cytoplasmic RNP complex during *Drosophila* oogenesis. *Development* (in press)
- 53 Navarro, R.E. *et al.* (2001) *cgh-1* a conserved predicted RNA helicase required for gametogenesis and protection from physiological germline apoptosis in *C. elegans*. *Development* 128, 3221–3232
- 54 Minshall, N. *et al.* (2001) A conserved role of a DEAD-box helicase in mRNA masking. *RNA* 7, 1728–1742
- 55 Curtis, D. *et al.* (1995) Translational regulation in development. *Cell* 81, 171–178
- 56 Wang, C. *et al.* (1991) nanos is the localized posterior determinant in *Drosophila*. *Cell* 66, 637–647 [erratum appears in *Cell* (1992) 68, 1177]
- 57 Wang, C. *et al.* (1994) Genetics of *nanos* localization in *Drosophila*. *Dev. Dyn.* 199, 103–115
- 58 Gray, J.V. *et al.* (1997) A role for the Pkc-1 MAP kinase pathway of *S. cerevisiae* in bud emergence and identification of putative upstream regulator. *EMBO J.* 16, 4924–4937
- 59 Chen, T. *et al.* (1997) *Saccharomyces cerevisiae* Mpt5p interacts with Sst2p and plays roles in pheromone sensitivity and recovery from pheromone arrest. *Mol. Cell Biol.* 17, 3429–3439
- 60 Kimble, J. *et al.* (1997) The LIN-12/Notch signaling pathway and its regulation. *Annu. Rev. Cell. Dev. Biol.* 13, 333–361
- 61 Williamson, J.R. (2000) Induced fit in RNA-protein recognition. *Nat. Struct. Biol.* 7, 834–837
- 62 Lefstin, J.A. *et al.* (1998) Allosteric effects of DNA on transcriptional regulators. *Nature* 392, 885–888
- 63 Kobayashi, S. *et al.* (1996) Essential role of the posterior morphogen *nanos* for germline development in *Drosophila*. *Nature* 380, 708–711

## Trends in Genetics editorial policy

For information on our editorial policy and how to submit papers to Trends in Genetics see <http://www.trends.com/tig/about.htm>