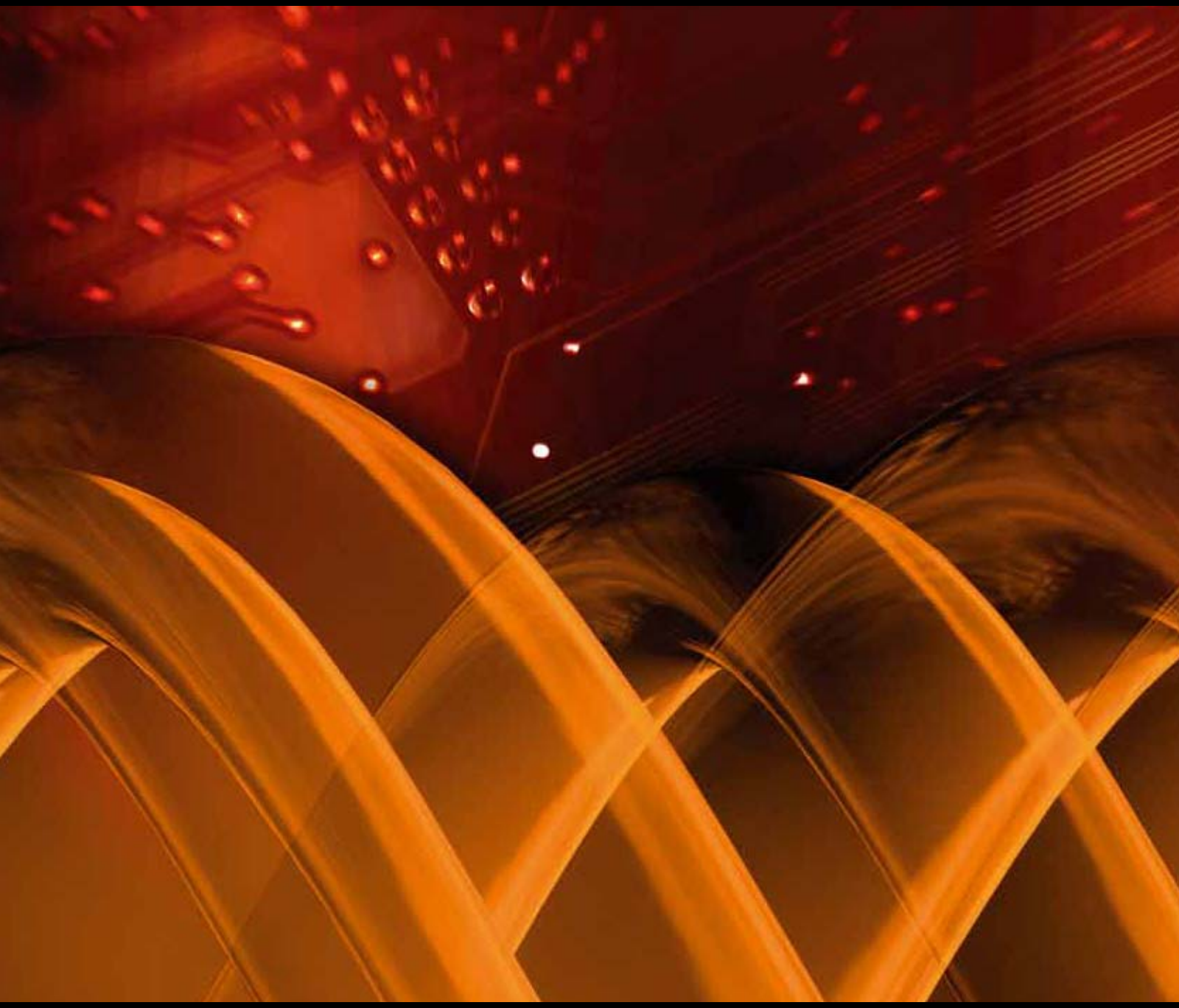
Comparative and Functional Genomics

TRANSLATIONAL CONTROL ACROSS EUKARYOTES

Guest Editors: Greco Hernández, Christopher Proud, Thomas Preiss, and Armen Parsyan



Translational Control across Eukaryotes

Comparative and Functional Genomics

Translational Control across Eukaryotes

Guest Editors: Greco Hernández, Christopher Proud, Thomas Preiss, and Armen Parsyan

Editorial Board

Contents

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 317690, 2 pages
doi:10.1155/2012/317690

Editorial

Translational Control across Eukaryotes

Greco Hernández

Division of Basic Research, National Institute for Cancer (INCan), Avenida San Fernando No. 22, Tlalpan, C.P. 14080, Mexico City, Mexico

Correspondence should be addressed to Greco Hernández, ghernandezr@incan.edu.mx

Received 27 May 2012; Accepted 27 May 2012

Diversity is one of the most significant features of life. The current assessment of biodiversity on Earth reaches several millions of living species, and an unimaginable number of species that have gone extinct. Further more, diversity is ingrained in all aspects of life, namely, genetics, metabolic needs, cell complexity, body plans and organismal morphology, developmental programs, behavioral patterns, and the occupancy of ecological niches. A deep understanding of all levels of organismal diversity is necessary to obtain a profound understanding of life. For this reason, the main goal of this special issue on translational control across eukaryotes is to acknowledge and discuss recent research showing that molecular and functional diversification also exists in the translation apparatus of eukaryotes. "Translation," that is, the process of decoding a messenger RNA to synthesize a protein by the ribosome, is a fundamental process for all forms of life.

The basic components and the global process of translation in eukaryotes have been largely established upon the study of traditional, so-called model organisms. However, this long-established overemphasis on model organisms has limited our knowledge of this process to few species. In recent years, the use of modern whole-genome sequencing and high-throughput technologies to study many nonmodel eukaryotes from disparate taxa has shown that the gene diversity in nature is far more expansive than we have ever imagined. These studies have uncovered a surprising diversity in the configuration of the translation apparatus across eukaryotes. This special issue acknowledges the diversity of translation factors in different lineages, the diversity found in key elements of mRNAs and RNA-binding proteins, and the diversity found in key RNA metabolism processes coupled to translational control, such as the RNA transport and the storage and degradation of mRNAs in cytoplasmic bodies. The first tutorial article, written by the editors of this installment, introduces the global topic as well as the spirit of the issue. Firstly, it introduces the general process of translation in eukaryotes and, secondly, it reviews the current knowledge on the many components of the translation machinery that have undergone diversification across phyla.

Diversification of initiation factors is perhaps the most remarkable feature of the evolution of the translation apparatus. The papers by R. Jagus's group (University of Maryland, Baltimore, USA) and by A. Zinoviev and M. Shapira (Ben-Gurion University of the Negev, Beer Sheva, Israel) review this topic in protists, the most diverse but less studied group of eukaryotes. The paper by R. M. Patrick and K. S. Browning (University of Texas at Austin, USA) covers the role of the key initiation factors eIF4F and eIFiso4F in plants. It also traces the evolutionary history of these two complexes across the kingdom of plants. A. Muñoz and M. M. Castellano (Polytechnic University of Madrid, Spain) analyze the different mechanisms regulating translation initiation in plants in response to abiotic stresses. They describe the mechanisms that plants share with other eukaryotes as well as the plant-specific ones. G. Hernández (National Institute for Cancer, Mexico City), in collaboration with P. Lasko and N. Sonenberg (McGill University, Montreal, Canada) investigate the evolutionary pattern that eIF4E has undergone in the class Insecta. Insects comprise the majority of extant animal species described and are the most diverse animal group on the planet. They have a huge impact on the biosphere as well as in all aspects of human life and economy.

In the last years, the role of the 5′-untranslated region of mRNA in translation regulation has been extensively studied. E. Martínez-Salas et al. (Autonomous University of Madrid, Spain) goes through internal ribosome entry sites (IRESs) as major elements regulating translation, specially

under stress conditions, in many phyla. L. O. F. Penalva's group (University of Texas, San Antonio, USA) explores the different, non-IRES, regulatory elements present at this region of the mRNA.

L. Wurth (Centre for Genomic Regulation, Barcelona, Spain) reviews the role and the conservation in different eukaryotes of key RNA-binding proteins (RBPs), which regulate translation to adjust cell proteome to different environmental conditions. As discussed by Wurth and by G. Grech and M. Lindern (University of Malta, Malta), some RBP are emerging as major players in the development of different tumors. C. Gamberi and P. Lasko (McGill University, Montreal, Canada) discuss the importance and evolutionary conservation of another RBP that plays crucial roles in development as well as in other cellular processes, that is, the Bic-C family of proteins and some of their protein partners.

Translational control in eukaryotes is tightly coupled to several components and processes of cell metabolism. One of them is the RNA transport, which establishes cellular asymmetries of protein synthesis to ensure different cell process or developmental programs to occur. The RNA transport machineries have also diverged in different phyla and, together with them, some components of the translation apparatus also diverged. This topic is reviewed by P. Vazquez-Pianzola and B. Suter (University of Bern, Switzerland), who analyze the conservation and divergence across eukaryotes of two major RNA transport machineries in unicellular and multicellular eukaryotes, namely, the yeast "Locasome" and the *Drosphila* Bic-D/Egl/Dyn complex, respectively. Another fundamental aspect of the translational control coupled to RNA metabolism is the storage and degradation of mRNAs in different cytoplasmic bodies, such as processing bodies and stress granules, which are collections of ribonucleoprotein complexes containing translation factors. Current research on the diversity of these foci in different phyla is analyzed by C. Layana, P. Ferreo and R. Rivera-Pomar from the National University of La Plata (Argentina).

The studies and reviews presented in this issue complement research published in other journals. Collectively, these studies show that after eukaryotes emerged, both components and regulatory mechanisms of the translation apparatus continued evolving during eukaryotes diversification. This supports the idea that this apparatus is far from being evolutionarily static.

Greco Hernández

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 813718, 10 pages
doi:10.1155/2012/813718

Review Article

Evolutionary Conservation and Diversification of the Translation Initiation Apparatus in Trypanosomatids

Alexandra Zinoviev and Michal Shapira

Department of Life Sciences, Ben-Gurion University of the Negev, P.O. Box 653, Beer Sheva 84105, Israel

Correspondence should be addressed to Michal Shapira, shapiram@bgu.ac.il

Received 31 January 2012; Accepted 12 March 2012

Academic Editor: Christopher Proud

Trypanosomatids are ancient eukaryotic parasites that migrate between insect vectors and mammalian hosts, causing a range of diseases in humans and domestic animals. Trypanosomatids feature a multitude of unusual molecular features, including polycistronic transcription and subsequent processing by *trans*-splicing and polyadenylation. Regulation of protein coding genes is posttranscriptional and thus, translation regulation is fundamental for activating the developmental program of gene expression. The spliced-leader RNA is attached to all mRNAs. It contains an unusual hypermethylated cap-4 structure in its 5′ end. The cap-binding complex, eIF4F, has gone through evolutionary changes in accordance with the requirement to bind cap-4. The eIF4F components in trypanosomatids are highly diverged from their orthologs in higher eukaryotes, and their potential functions are discussed. The cap-binding activity in all eukaryotes is a target for regulation and plays a similar role in trypanosomatids. Recent studies revealed a novel eIF4E-interacting protein, involved in directing stage-specific and stress-induced translation pathways. Translation regulation during stress also follows unusual regulatory cues, as the increased translation of Hsp83 following heat stress is driven by a defined element in the 3′ UTR, unlike higher eukaryotes. Overall, the environmental switches experienced by trypanosomatids during their life cycle seem to affect their translational machinery in unique ways.

1. Translation Initiation in Eukaryotes

Translation is a complex process that is controlled by a large number of proteins and factors, allowing the cell to generate a rapid response to external and internal signals. The initiation step of protein synthesis is predominantly viewed as the limiting step of this process, thus serving as a target for a multitude of regulatory cues.

During cap-dependent translation initiation, a preassembled 43S preinitiation complex (PIC) is targeted to the 5′ end of the mRNA through the cap-binding complex, eIF4F. The PIC subunits are comprised of the Met-tRNA along with eIF1, eIF1A, eIF2, and the multisubunit eIF3, which is responsible for recruiting the small ribosomal subunit (reviewed in [1, 2]). The eIF4F complex consists of the cap-binding protein, eIF4E, an RNA helicase, eIF4A, and the scaffold protein, eIF4G. The mammalian eIF4G holds together the eIF4F subunits and links them to the 43S PIC by interacting with eIF4E, eIF4A, and eIF3 [3, 4]; it is also responsible for the transient circularization of translating mRNAs, by interacting with the poly(A) binding protein (PABP) that is bound at their 3′ ends (reviewed in [5, 6]). The initiation complex scans the 5′ untranslated region (5′ UTR) until it reaches the initiator AUG codon, where the 60S ribosomal subunit joins to form the mature 80S ribosome. Two factors, eIF5B [7] and eIF6 [8], participate in this step, releasing most of the initiation factors. The scanning process is assisted by helicases which are associated with the initiation complex, with eIF4A serving as the main candidate, although it is not necessarily the only factor involved in this process [9, 10]. The activity of eIF4A is enhanced by an additional factor, eIF4B, which was shown to be essential for the helicase activity [11]. Once the ribosome advances along the mRNA and vacates the AUG start codon, a new initiation complex can be formed on the 5′ cap, leading to the formation of a polyribosome. Actively translating mRNAs are usually associated with the polysomal fraction of disrupted cells.

Exposure to a variety of environmental, nutritional, or virus induced stresses, causes a global inhibition of cap-dependent translation. However, certain mRNAs continue to be translated in a cap-independent manner, by which the small ribosomal subunit is targeted to a position that is adjacent to the initiator AUG, with the aid of a highly structured RNA element in the 5′ UTR. This element is denoted the internal ribosome entry site (IRES) [12]. The IRES assigns a functional secondary structure, usually with the aid of *trans*-activating factors, such as the poly-pyrimidine-tract-binding (PTB) protein, ITAF45 or the La antigen [13]. IRES elements were also reported for endogenous mRNAs [12, 14–16], however, they do not account for all cases of cap-independent translation, and other, less understood pathways, are probably also used by eukaryotic cells [17, 18].

Translation initiation is controlled by complex signaling networks, which integrate the internal and external conditions of living cells. It was reported that the arrest in cap-dependent translation observed under stress conditions is coordinated with the dephosphorylation and activation of the 4E-binding protein (4E-BP) The latter competes with eIF4G on binding to eIF4E, thus excluding assembly of the cap-binding complex [19–21]. The phosphorylation status of 4E-BP is mediated by the TOR-based complexes [22] that link between metabolism, protein synthesis and cell cycle [23, 24], which are controlled by a variety of cellular pathways [25]. While in most cases, blocking the dephosphorylation of 4E-BP led to inhibition of protein synthesis, exceptional cases have been reported where protein synthesis was not interrupted despite the dephosphorylation of 4E-BP, which was induced by specific mTOR inhibitors [26].

Another way by which translation can be globally stopped is through phosphorylation of eIF2α, which forms a ternary complex with Met-tRNA-GTP. Following the hydrolysis of GTP, activity of eIF2α is regenerated by the eIF2B-mediated exchange of GDP with GTP. It is commonly observed that phosphorylation of eIF2α at Ser 51 blocks this exchange, resulting in a global translational arrest [27].

Translation can be regulated also at a gene-specific level, without affecting the overall translation capacity of the cell. In many cases, a regulatory protein that is recruited by elements in the 3′ UTR binds eIF4E, in a manner that blocks the access to eIF4G and prevents assembly of the eIF4F complex. Such a pattern of regulation is observed during the development of *Drosophila* embryos, in which specific proteins and mRNAs are distributed unevenly along the body axis. For example, Maskin interacts with the eIF4E only on RNAs that contain a cytoplasmic polyadenylation element (CPE); disruption of the eIF4E-eIF4G complex by this protein is therefore specific to mRNAs that contain a CPE in their 3′ UTR (reviewed in [21]).

Exposure to extreme conditions such as temperature and pH alterations, osmotic switches, UV irradiation, nutrient starvation and oxidative stress is hazardous to all living cells, and is combatted by activation of a complex chaperone network. Extreme temperatures cause cellular damage at multiple levels, accompanied by a global arrest in synthesis of most cellular proteins, except for heat shock proteins (HSPs). In cases where the cell cannot overcome the inflicted damage, apoptosis may be induced [16]. The increased translation of HSPs in higher eukaryotes is controlled by the 5′ UTR, as shown for both HSP70 [28, 29] and HSP90 [30]. HSP90 and some variants of HSP70 are expressed at ambient temperatures, but their translation increases dramatically during heat shock. Despite extensive efforts, no functional motifs that could drive preferential translation at elevated temperatures were yet identified. It was also suggested that under normal temperatures, translation of Hsp90 occurs by a typical cap-dependent scanning mechanism, whereas during heat shock, translation shifts to a cap-independent mode.

2. Trypanosomatid Organisms

Trypanosomatids are unicellular, diploid protists from the Kinetoplastidæ order, which migrate between insect vectors and mammalian hosts, causing a variety of parasitic diseases in humans and their domestic animals. *Leishmania* parasites reside in the alimentary tract of female sandflies as extracellular flagellated promastigotes [31, 32]. Upon transmission to mammalian hosts, they differentiate into obligatory intracellular amastigotes within macrophages and dendritic cells [33, 34], causing a range of symptoms, that depend on the infecting species. *Trypanosoma cruzi* is the causative agent of Chagas disease in South America, and *T. brucei* spp. cause the African sleeping sickness. Altogether, trypanosomatid-borne diseases cause a great health threat and economical drawbacks for native populations living in endemic countries, as well as being a major problem for travelers visiting these regions. The divergence of most trypanosomatid species occurred very early in evolution, before the emergence of both vectors and hosts [35]. Thus, unique molecular properties are common in this family as compared to higher eukaryotes.

Protein coding genes in trypanosomatids are arranged as large, unidirectional transcription units, which are transcribed polycistronically. The resulting pre-mRNAs are matured by *trans*-splicing and polyadenylation into monocistronic mRNAs (reviewed in [37, 38]). Conventional RNA polymerase II promoters were not identified to date, and there is yet no evidence for the occurrence of regulated transcription activation processes for protein coding genes. Thus, mRNA processing, stability and translation serve as key mechanisms that direct differential program of gene expression throughout the pirasite life cycle (reviewed in [39]). During *trans*-splicing, a conserved mini-exon of 39 nucleotides is donated by the spliced leader RNA (SL RNA) to the 5′ end of all mRNAs. This mini-exon provides the 5′ cap structure, which in trypanosomatids is heavily methylated on the first four nucleotides and is thus denoted cap-4 [40]. In addition to the consensus m^7GTP, cap-4 consists of 2′-*O* ribose methylations on the first four nucleotides of the SL RNA and two base methylations on the first adenosine and fourth uridine (Figure 1). These base methylations are unique to trypanosomatids, and are not known in any other group of eukaryotes [36]. Substitutions

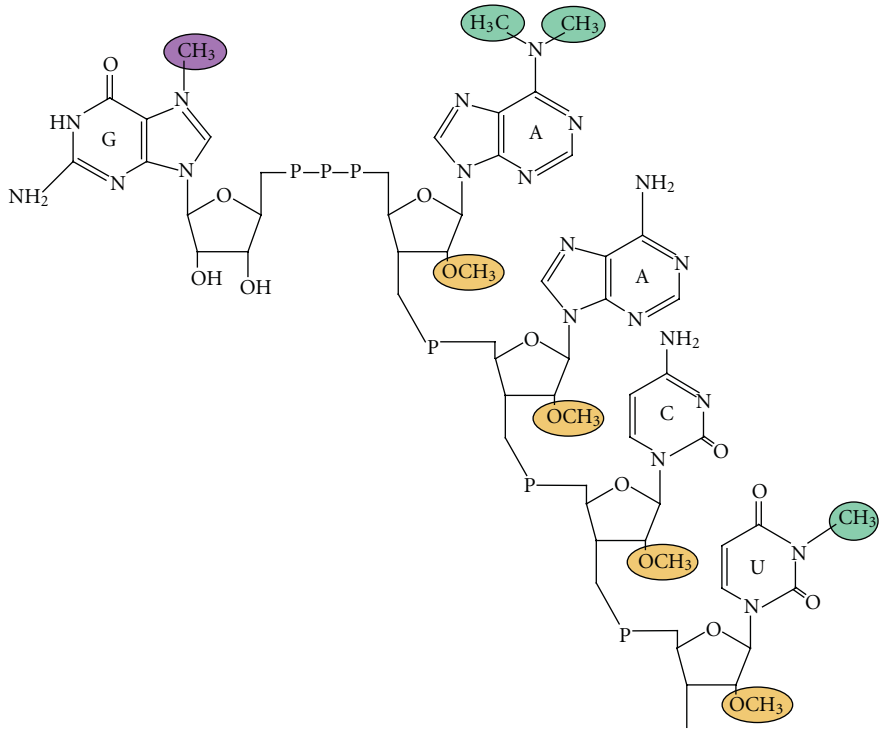


Figure 1: Trypanosomatid cap-4 structure, based on Bangs et al., 1992 [36].

of individual nucleotides in the cap-4 structure diminished the ability of the mutated SL RNA to be utilized in *trans*-splicing reactions [41] and thus it was not possible to evaluate the exact role of ribose methylation in translation. The enzymes that promote the 2′-*O*-methylations on the ribose moieties of the cap-4 nucleotides were identified [42, 43], and the effect of their elimination was tested. Single and double knock-out mutants of genes encoding one or two of the enzymes responsible for ribose 2′-*O*-methylation at position 1 (TbMTr1), 2 (TbMTr2), 3 and 4 (TbMTr3) were examined. Preventing the ribose methylations at positions 3 and 4, but not at position 1, reduced the translation rates; a further exacerbation was observed by additional loss of the methylation at position 2. Knock-out of TbMTr1 alone did not cause an inhibitory effect on translation, however, depletion of TbMTr2 or TbMTr3 on TbMTr1$^{-/-}$ background did not yield viable parasites [44, 45]. Therefore, it was proposed that only a minimal level of mRNA cap ribose methylation is essential for trypanosome viability. The role of cap-1 modification was shown to be related to the SL RNA biogenesis, as formerly shown for *Leptomonas collosoma* mutants [38].

3. eIF4E Isoforms of Trypanosomatids

Trypanosomatid genome annotation combined with a functional analysis approach, paved the way for studies of the translation apparatus in these organisms. Some of the factors were subject to a thorough biochemical and cellular analysis, whereas others were only identified based on sequence homology. Homology modeling of the four LeishIF4E paralogs suggested that the structure of the cap-binding pocket was conserved and maintained the basic features observed for the yeast and mammalian eIF4E [46]. It is of interest to note that the association constants for complexes of LeishIF4E homologues with m^{7}GTP as well as with cap-4 were two orders of magnitude lower than those of the mouse protein [46]. This could be related to the evolutionary changes that occurred in the parasite proteins to promote their interaction with cap-4. However, this has yet to be proven.

eIF4E is the eukaryote 5′ cap-binding translation initiation factor; its association with eIF4G is fundamental to the assembly of the cap-binding complex. The 3D structure of eIF4E from several organisms was formerly deciphered [3, 47–50]. Its N-terminal part is flexible and not conserved, while the C-terminus adopts a conserved structural signature. eIF4E acquires the shape of a baseball glove with a dedicated pocket for binding of the methylated guanine, which contains a sandwich of three Trp residues (W56, 102 and 166, according to the murine numbering). In addition, the basic residues Arg112, 157 and 162 interact with the negatively charged phosphate moieties of the cap structure. eIF4E binds eIF4G and several translation repressors, such as 4E-BP, through several conserved residues in its C-terminus. The high degree of structural conservation among orthologs of eIF4E enables homologs from different species to functionally replace each other in a yeast-based genetic complementation assay [51].

The trypanosomatid cap-4 binding proteins have gone through structural adaptations during their evolution to

adjust to binding of the highly modified cap-4 structure. Their genomes encode four isoforms of eIF4E [46, 52]. Although homology modeling of the four proteins supports a structural conservation of the cap-binding pocket, none of the four paralogs could complement the missing function of the yeast eIF4E, indicating high functional divergence from their higher eukaryotic orthologs [46]. Thus, deciphering the roles of the different isoforms is complex, especially since efficient cell-free systems for *in vitro* reconstitution of translation initiation using parasite mRNAs are not available for trypanosomatids. Biophysical assays using a chemically synthesized cap-4 and intermediate cap analogues [53], showed that the four eIF4E, denoted LeishIF4E-1 through LeishIF4E-4, vary in their cap-binding specificities. Trp fluorescence assays indicated that LeishIF4E-1 and LeishIF4E-4 bound m^7GTP and cap-4 with comparable affinities, LeishIF4E-3 bound mainly to m^7GTP and hardly to cap-4, whereas LeishIF4E-2 showed a great preference to cap-4, as compared to m^7GTP [46]. In agreement, the endogenous LeishIF4E-1 and LeishIF4E-4 were eluted from m^7GTP-Sepharose column [46, 54]. Despite the ability of LeishIF4E-3 to bind m^7GTP in the fluorescence titration assays, the recombinant and endogenous proteins showed very weak binding to m^7GTP-Sepharose column, for reasons not fully understood yet [46].

The subunits of the *Leishmania* eIF4F complex were examined by pulldown analysis and by monitoring their migration profile on sucrose gradients. These analyses led to the conclusion that LeishIF4E-4 is the most probable candidate to serve as the conventional eIF4E in promastigotes, and is part of the parasite LeishIF4F cap-4 binding complex. LeishIF4E-4 was shown to interact with LeishIF4G-3, a protein that contains a "middle of eIF4G" (MIF4G) domain [54, 55], which is responsible for recruiting LeishIF4A-1. A parallel interaction was shown for the *Trypanosoma brucei* ortholog [54]. The three eIF4F subunits comigrate on sucrose density gradients and are found in fractions that are expected to contain the pre-initiation complex [55].

Silencing experiments by RNAi were performed on the four *T. brucei* isoforms [54] in an attempt to reveal their function. Downregulation of TbIF4E-4 inhibited the growth of bloodstream form, but not procyclic cells. This result does not coincide with data obtained for the *Leishmania* paralog, since LeishIF4E-4 binds m^7GTP and LeishIF4G-3 mainly in promastigotes, and fails to perform these activities upon exposure to mammalian-like temperatures and in axenic amastigotes [56]. Furthermore, the migration pattern of LeishIF4E-4 in sucrose gradients shows that in heat-shocked cells this protein no longer forms large complexes. Altogether, the activity of LeishIF4E-4 is dramatically reduced at prolonged elevated temperatures [55], through a mechanism which is yet to be resolved. The differences observed between the stage-specific function of the LeishIF4E-4 and TbIF4E-4 described above could originate from a variable experimental setup, or alternatively, from differences between the species. It is also possible that after the final differentiation to the mammalian amastigote life-form, LeishlF4E-4 resumes its activity.

The dual silencing of TbIF4E-4 and TbIF4E-1 led to a growth arrest of procyclic parasites, suggesting that the two proteins may have a partially redundant function. Individual silencing of TbIF4E-4 or TbIF4E-1 was harmful only to the bloodstream life form. LeishIF4E-1, the *Leishmania* ortholog of TbIF4E-1, is the only eIF4E isoform that maintains its expression at elevated temperatures and in axenic amastigotes of *Leishmania*. The other three orthologs were downregulated under these conditions, excluding that they have a role in translation during heat shock [46, 56]. It is, therefore, possible that LeishIF4E-1 is associated with translation under stress and in amastigotes. TbIF4E-1, as well as its *Leishmania* ortholog LeishIF4E-1, do not interact efficiently with TbIF4G-3 or LeishIF4G-3, respectively [54, 56], suggesting that this isoform does not participate in building an eIF4F complex and therefore, could promote alternative pathways of translation, possibly in an eIF4G-independent manner. LeishIF4E-1 could either have an active role in translation which has not yet been resolved, or alternatively, it could passively protect the cap structure, if cap-independent mechanisms are practiced.

Elimination of TbIF4E-2, had no effect on *T. brucei* growth in both life forms [54]. The *Leishmana*, LeishIF4E-2, ortholog comigrated with heavy polysomes in sucrose gradients [46], unlike typical translation initiation factors that are found in lighter fractions [57, 58]. The RNAi experiments in *T. brucei* indicated that among the four paralogs, only TbIF4E-3 was essential for growth of both procyclic and bloodstream life forms. Its silencing caused a reduction in incorporation of a radiolabeled amino acid into newly synthesized polypeptides, and it was, therefore, suggested that TbIF4E-3 serves as a translation initiation factor [54]. However, the reduced translation rates occurred only after 72 hours, while the silencing was observed already after 24 hours. It could therefore be a downstream effect of other processes that inhibit cell growth in the absence of TbIF4E-3. TbIF4E-3 was reported to be part of RNA granules in *T. brucei* [59, 60] and could, therefore, be associated with trafficking of mRNAs either to storage bodies or to actively translating polysomes. The low affinity of the *Leishmania* ortholog LeishIF4E-3 to cap-4 contradicts its definition as a typical translation initiation factor. Thus, although TbIF4E-3 interacts with a MIF4G domain protein, TbIF4G-4, its precise role remains elusive.

4. eIF4E Binding Proteins of Trypanosomatids

Trypanosomatids encode several eIF4G candidates that contain the MIF4G domain. This element consists of several HEAT repeats [61], typical of all eIF4G proteins in higher eukaryotes [62]; it is responsible for binding to eIF4A as well as to RNA. The C-terminus of the human eIF4G contains a second site for interaction with eIF4A and its N-terminus carries binding sites for eIF4E and PABP. A consensus peptide motif, YXXXXLΦ, in eIF4G and in 4E-BP is responsible for binding to eIF4E. Substitution of the conserved Y or LΦ residues in the motif abrogates the eIF4E-eIF4G interaction [63].

The identity of the *Leishmania* eIF4G ortholog was deduced by its co-purification and interaction with LeishIF4E-4. This binding was monitored in yeast two-hybrid assays and pulldown experiments [52, 54–56]. LeishIF4G-3 and TbIF4G-3 contain a typical MIF4G domain [52, 55], but they vary from their mammalian counterpart in other parts of the protein. For example, the N-terminus of LeishIF4G-3 is short (50 amino acids) and contains the LeishIF4E-4 binding-peptide, but it cannot interact with PABP in a yeast two-hybrid assay. This peptide motif is only partially conserved with the YXXXXLΦ consensus, and binding to LeishIF4E-4 requires the presence of a Phe residue at position 4 [63], in addition to Tyr and Leu at positions 1 and 6 of the peptide. A partial requirement was observed for the Gly and Glu residues at positions 3 and 8 [55]. Altogether, this indicates a certain degree of variability from the eukaryote consensus sequence.

Other regions of the parasite eIF4G are also a source for variability, as compared to its human counterpart. eIF4G is a scaffold protein that links the 5′ and 3′ ends by interacting with both eIF4E and PABP. However, unlike in higher eukaryotes, mRNA circularization occurs through an interaction between the N-terminus of LeishIF4E-4 and LeishPABP-1. Elimination of this LeishIF4E-4 domain prevents its binding to LeishPABP-1 [56]. Therefore, LeishIF4E-4 along with LeishIF4G-3 and LeishIF4A-1, comprise a typical eIF4F complex in promastigotes [64], see Figure 2. The binding between LeishIF4G-3 and LeishIF4E-4 is eliminated at elevated temperatures and in axenic amastigotes, supporting that under these conditions translation may proceed through alternative pathways, possibly involving LeishIF4E-1.

An additional eIF4E-eIF4G pairing was reported for TbIF4E-3 and TbIF4G-4 [54]. The authors of this study propose that TbIF4E-3 is involved in translation initiation. This assumption is not supported by studies in *Leishmania* that excluded the copurification of these two proteins over a m^7GTP affinity column [54]. Furthermore, LeishIF4E-3 failed to bind cap-4 in fluorescence titration assays, thus making it an unlikely candidate to serve as a conventional translation initiation factor [46].

Another well-known eIF4E binding protein is 4E-BP of higher eukaryotes. It is a highly conserved protein that is expressed in most eukaryotes, except for *C. elegans*. The trypanosomatid genome database also does not contain any ortholog of the consensus 4E-BP (~10 kDa). However, affinity co-purification assays identified a novel LeishIF4E-interacting protein, denoted Leish4E-IP (~80 kDa) [56]. Leish4E-IP and the eukaryotic 4E-BP show no sequence homology, but share a predicted unstructured nature. Leish4E-IP binds mainly to LeishIF4E-1 with a tight requirement for the consensus YXXXXLΦ peptide [63]. Furthermore, although Leish4E-IP is expressed at all life stages, its binding to LeishIF4E-1 is observed only in promastigotes, suggesting that it participates in framing a stage-specific program of gene expression, via a mechanism that is yet to be resolved. Posttranslational modifications, such as phosphorylations, are likely to be involved. Data mining in the *Leishmania* genome revealed three orthologs of the TOR kinase. TOR 1 and 2 are essential, as null mutants could not be generated, whereas TOR3 could be eliminated. However, null mutants of TOR3 were nonvirulent and were impaired in their ability to survive within macrophages [65].

5. eIF4A Isoforms of Trypanosomatids

eIF4A is a member of the DEAD-box RNA helicases [66]. It promotes scanning of the 5′ UTR by the 43S PIC, via unwinding of the secondary RNA structures, until it arrives at the initiator AUG codon. DEAD-box proteins participate in a multitude of processes related to transcription, RNA processing, export and translation. The genome of most eukaryotes encodes for more than a single eIF4A isoform. In addition to the translation initiation factor eIF4A, another paralog was reported to be part of the exon junction complex (EJC) [67]. The genome of *T. brucei* encodes two isoforms of eIF4A. TbIF4AI and its *Leishmania* homolog, LeishIF4A-1, are part of the eIF4F cap-binding complex. Tb4AI is cytoplasmic and, as expected, interacts with TbIF4G-3 [52, 56]. Its downregulation caused a dramatic decrease in protein synthesis. TbIF4AIII is predominantly nuclear. It is essential, but its silencing hardly affects protein synthesis, and the resulting lethal effects are delayed, as compared to TbIF4AI. Thus TbIF4AI is the translation initiation factor that comprises the cap-4 binding complex, while the role of TbIF4AIII is not yet understood, although it may assign a similar role as in higher eukaryotes [64].

6. Poly(A) Binding Proteins of Trypanosomatids

Yeast encode a single PABP, whereas metazoans express several paralogs, with a tissue, or embryonic-specific pattern of expression [68]. The typical PABP are closely related, and consist of four conserved RRM domains. Their C-terminus promotes protein-protein interactions with their binding partners [69].

Elongated poly(A) tails are crucial for translation, but are also involved in a variety of processes related to RNA processing, export and stabilization. The mammalian PABP associates with eIF4G, resulting in transient circularization of the mRNA during translation initiation [70]. This interaction stabilizes the initiation complex and enhances translation. It is therefore targeted by endogenous regulators, such as the PABP-binding proteins PAIP1 and PAIP2, which can enhance or repress the activity of PABP, respectively [71]. PABP is also targeted by viral proteases, as part of their strategy to take over the cellular translational machinery [72].

T. brucei and *T. cruzi* genomes contain two PABP paralogs [73, 74]. Both are essential, indicating that they have different cellular functions. The leishmanias encode another unique isoform, PABP3 [73]. PABP1 was associated with components of the cap-binding complex [56, 73], suggesting that it is the cap-dependent translation initiation factor. To date, it is the only isoform that was shown to bind directly to a translation initiation factor. However, unlike in higher eukaryotes, it interacts with eIF4E instead of eIF4G. The higher eukaryote PABP shuttles between the nucleus

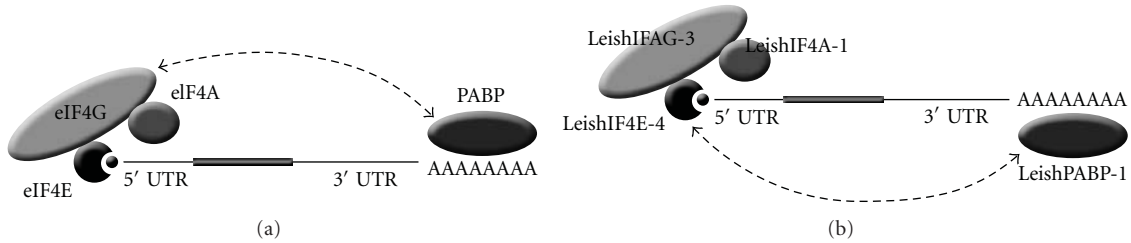


Figure 2: Interacting partners of cap-binding proteins in *Leishmania*. The typical eIF4F complex of higher eukaryotes (a) and *Leishmania* promastigotes (b) are shown. A table summarizing the known interacting partners of the *Leishmania* eIF4F complex are shown in Table 1. Accession numbers of the described proteins are: LeishIF4E-1—LmjF27.1620, LeishIF4E-3—LmjF28.2500, LeishIF4E-4—LmjF30.0450, LeishIF4G-3—LmjF16.1600, LeishIF4G-4—LmjF36.6060, LeishIF4A-1—LmjF01.0780 and LmjF01.0770, LeishPABP-1—LmjF35.5040, Leish4E-IP—LmjF35.3980.

Table 1

Protein	Interacts with	Source
LeishIF4E-1	Leish4E-IP	Zinoviev et al., 2011 [56]
LeishIF4E-3	LeishIF4G-4	Freire et al., 2011 [54]
LeishIF4E-4	LeishIF4G-3	Yoffe et al., 2009 [55]
LeishIF4A-1	LeishIF4G-3	Zinoviev et al., 2011 [56]
LeishPABP-1	LeishIF4G-3	Zinoviev et al., 2011 [56]

and the cytoplasm, but it is mostly cytoplasmic. All three *Leishmania* isoforms are indeed cytoplasmic, but inhibition of transcription causes PABP2 and PABP3 to accumulate in the nucleus. Their role is yet to be identified.

7. Translation under Stress Conditions in Trypanosomatids

The untranslated regions of trypanosomatid mRNAs, mostly those located downstream to the coding sequences, play a key role in differential expression of genes during the life cycle. This was established by the use of reporter systems for Hsp70 [75, 76] and Hsp83 [77, 78] of *Leishmania*. It was also established, that despite the key role of mRNA stability, elements that affected translation of Hsp83 alone could be identified. Large mutations that destabilized the Hsp83 mRNA at elevated temperatures did not interfere with its preferential translation, as long as the regulatory element for translation was included in the 3′ UTR [77]. Fine deletions finally identified a regulatory element of 30 nucleotides (positions 312–341), containing a stretch of polypyrimidines. This region was shown to be part of an RNA structure that was predicted with high probability [79], using the UNAfold algorithm [80]. A biophysical evaluation of the mRNA melting curves was performed to examine the role of secondary structures in the regulatory region. Incubation of the corresponding wild-type mRNA fragment (1–472) led to its melting at elevated temperatures (35°C). The mutant element, that did not induce preferential translation, failed to show a similar pattern. It was, therefore, assumed that preferential translation of *Leishmania* HSP83 during stress was promoted by melting of the regulatory region in the mRNA [79]. It is interesting to note that the element is not conserved throughout trypanosomatids, emphasizing the role of RNA structure in this mode of regulation.

An additional RNA element that confers stage-specific translation was identified in the the *Leishmania* Amastin genes. It was first mapped as an element of 450 nucleotides within their 3′ UTRs. Additional experiments narrowed this region down to 150 nucleotides that enhance translation of the harboring transcript in the amastigote life stage. The Amastin element was found in other amastigote-specific genes of *Leishmania*, such as Hsp100 [81, 82]. However, it bears no conservation to that of Hsp83. Furthermore, no defined secondary structure was reported for the Amastin RNA regulatory sequence. Therefore, the preferential translation conferred by the two elements is functionally distinct.

Trypanosomatids experience a broad range of environmental stresses during their complex life cycle. In addition to the switch between vector and host, they can also suffer from a shortage in sugars within the insect vector, due to its nutritional diet. Under these conditions, a transient arrest of translation is required until the stress is relieved. One common way to arrest global translation in higher eukaryotes is through inactivation of eIF2α, which can be achieved by phosphorylation of its conserved Ser51 residue. Indeed, phosphorylation of the *T. brucei* paralog occurs on the corresponding Thr169 [83] and on Thr 166 in the *Leishmania* eIF2α [84]. Since the trypanosomatid ortholog has a long non-conserved N-terminal extension, it shifts the position of the phosphorylated Thr. This phosphorylation of the *Leishmania* eIF2α is associated with stage transformation [84]. However, a mutant of *T. brucei* strain that encodes for a mutated eIF2α (T169A/−) and therefore cannot undergo phosphorylation, showed no effect on translation or on the ability to form heat-shock stress granules [59]. Thus, phosphorylation of Thr169 alone is most probably not sufficient to inhibit translation. Three potential eIF2α kinases (TbeIF2K1 to -3) were identified in the genome of *T. brucei* [83]. TbeIF2K2 was extensively studied, and shown to localize in the membrane of the flagellar pocket, a site that is known to promote exo- and endocytosis. This kinase phosphorylates the trypanosomatid eIF2α, as well as its mammalian counterpart. However, TbeIF2α is not a substrate for the mammalian GCN2 or PKR.

A second conserved mechanism for global translational arrest in higher eukaryotes involves the dephosphorylation of 4E-BP, although recent reports show numerous exceptions [26], indicating that this pathway is much more complex [23]. In most cases, the dephosphorylated 4E-BP binds to eIF4E and prevents assembly of the cap-dependent translation initiation complex. Trypanosomatid genomes lack the consensus 4E-BP and this pathway to achieve translational arrest is not functional as well. Thus, the mechanisms that confer a global decrease in translation in trypanosomatids remain unclear.

Differentiation of trypanosomatids from the insect-specific to the mammalian life-form is induced by extreme environmental switches. This raises the question of how these parasites deal with the damaging effects of the prolonged stresses, which would be deadly to other eukaryotes. In the absence of conventional mechanisms for transcriptional activation of protein coding genes [37], their differential pattern of expression along the life cycle is induced by post-transcriptional regulatory mechanisms [39], including splicing and mRNA stability, as well as translation [76, 77, 85]. Transcriptome [86, 87] and proteome [88] screens performed with *Leishmania* parasites during their axenic differentiation were published and are publicly available. In total, they reveal that differential gene expression occurs due to changes in mRNA levels at the early stages of differentiation. However, at later stages, translation and posttranslational regulation mechanisms are more influential [84]. It was also noted by the Zilberstein group that during the initial period at which signaling for differentiation takes place, translation decreased dramatically. Expression of ribosomal proteins was downregulated, eIF2α was phosphorylated and a general decrease in the amount of polysomes was observed. These effects were mainly transient, and translation resumed upon completion of the differentiation process [88].

8. What Is Next?

The research of trypanosomatids is rapidly advancing, and new tools are constantly being developed. The use of RNAi, which currently is restricted to *T. brucei*, is about to be developed for *Leishmania braziliensis* [88], the only *Leishmania* species that encodes all the components of this pathway. It should be interesting to see if silencing of the various factors described in this review confers similar effects in *Leishmania* as compared to *Trypanosoma*.

Another interesting point refers to the elegant strategy developed by *Leishmania* parasites to overtake their host cells. It was recently shown that the GP63 protease, which is secreted by the parasites into the infected macrophage, causes an efficient shutdown of the host protein synthesis. This is achieved by cleavage of mTOR, causing the constitutive dephosphorylation and activation of 4E-BP [89]. However, other targets of mTOR could presumably be affected, resulting in inhibitory effects on the macrophage metabolism.

Several recent publications describe the translational machinery of trypanosomatids and highlight its evolutionary variability, as compared to higher eukaryotes. The reported diversifications provide an exciting target for novel therapeutic approaches [90]. Potential drug targets that could be pursued include protein-RNA and protein-protein interactions that promote assembly of the translation initiation complex. For example, interactions between the cap-binding proteins and the unique cap-4 structure are of interest, as well as the binding between LeishIF4E-4 and LeishIF4G-3. TbIF4E-3, the ortholog of LeishIF4E-3, is also an interesting drug target, as it was shown to be essential in the blood stream form of *T. brucei*.

References

[1] A. C. Gingras, B. Raught, and N. Sonenberg, "eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation," *Annual Review of Biochemistry*, vol. 68, pp. 913–963, 1999.

[2] T. V. Pestova, J. R. Lorch, and C. H. T. Hellen, "The mechanism of translation initiation in Eukaryotes," in *Translation Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., pp. 87–128, Cold Spring Harbor, NY, USA, Cold Spring Harbor, 2007.

[3] J. D. Gross, N. J. Moerke, T. Von Der Haar et al., "Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E," *Cell*, vol. 115, no. 6, pp. 739–750, 2003.

[4] Y. Yamamoto, C. R. Singh, A. Marintchev et al., "The eukaryotic initiation factor (eIF) 5 HEAT domain mediates multifactor assembly and scanning with distinct to interacts to eIF1, eIF2, eIF3, and eIF4G," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 45, pp. 16164–16169, 2005.

[5] N. Amrani, S. Ghosh, D. A. Mangus, and A. Jacobson, "Translation factors promote the formation of two states of the closed-loop mRNP," *Nature*, vol. 453, no. 7199, pp. 1276–1280, 2008.

[6] S. Cheng and D. R. Gallie, "eIF4G, eIFiso4G, and eIF4B bind the poly(A)-binding protein through overlapping sites within the RNA recognition motif domains," *Journal of Biological Chemistry*, vol. 282, no. 35, pp. 25247–25258, 2007.

[7] A. Unbehaun, A. Marintchev, I. B. Lomakin et al., "Position of eukaryotic initiation factor eIF5B on the 80S ribosome mapped by directed hydroxyl radical probing," *EMBO Journal*, vol. 26, no. 13, pp. 3109–3123, 2007.

[8] V. Gandin, A. Miluzio, A. M. Barbieri et al., "Eukaryotic initiation factor 6 is rate-limiting in translation, growth and transformation," *Nature*, vol. 455, no. 7213, pp. 684–688, 2008.

[9] M. Oberer, A. Marintchev, and G. Wagner, "Structural basis for the enhancement of eIF4A helicase activity by eIF4G," *Genes and Development*, vol. 19, no. 18, pp. 2212–2223, 2005.

[10] H. Imataka and N. Sonenberg, "Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A," *Molecular and Cellular Biology*, vol. 17, no. 12, pp. 6940–6947, 1997.

[11] N. Methot, A. Pause, J. W. B. Hershey, and N. Sonenberg, "The translation initiation factor eIF-4B contains an RNA-binding region that is distinct and independent from its ribonucleoprotein consensus sequence," *Molecular and Cellular Biology*, vol. 14, no. 4, pp. 2307–2316, 1994.

[12] C. U. T. Hellen and P. Sarnow, "Internal ribosome entry sites in eukaryotic mRNA molecules," *Genes and Development*, vol. 15, no. 13, pp. 1593–1612, 2001.

[13] E. V. Pilipenko, T. V. Pestova, V. G. Kolupaeva et al., "A cell cycle-dependent protein serves as a template-specific translation initiation factor," *Genes and Development*, vol. 14, no. 16, pp. 2028–2045, 2000.

[14] A. G. Bert, R. Grépin, M. A. Vadas, and G. J. Goodall, "Assessing IRES activity in the HIF-1α and other cellular 5′ UTRs," *RNA*, vol. 12, no. 6, pp. 1074–1083, 2006.

[15] M. Holcik, N. Sonenberg, and R. G. Korneluk, "Internal ribosome initiation of translation and the control of cell death," *Trends in Genetics*, vol. 16, no. 10, pp. 469–473, 2000.

[16] M. Holcik and N. Sonenberg, "Translational control in stress and apoptosis," *Nature Reviews Molecular Cell Biology*, vol. 6, no. 4, pp. 318–327, 2005.

[17] D. E. Andreev, S. E. Dmitriev, I. M. Terenin, V. S. Prassolov, W. C. Merrick, and I. N. Shatsky, "Differential contribution of the m^7G-cap to the 5′ end-dependent translation initiation of mammalian mRNAs," *Nucleic Acids Research*, vol. 37, no. 18, pp. 6135–6147, 2009.

[18] S. F. Mitchell, S. E. Walker, M. A. Algire, E. H. Park, A. G. Hinnebusch, and J. R. Lorsch, "The 5'-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway," *Molecular Cell*, vol. 39, no. 6, pp. 950–962, 2010.

[19] A. Pause, G. J. Belsham, A. C. Gingras et al., "Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function," *Nature*, vol. 371, no. 6500, pp. 762–767, 1994.

[20] T. A. Lin, X. Kong, T. A. J. Haystead et al., "PHAS-I as a link between mitogen-activated protein kinase and translation initiation," *Science*, vol. 266, no. 5185, pp. 653–656, 1994.

[21] J. D. Richter and N. Sonenberg, "Regulation of cap-dependent translation by eIF4E inhibitory proteins," *Nature*, vol. 433, no. 7025, pp. 477–480, 2005.

[22] J. C. Lawrence and R. T. Abraham, "PHAS/4E-BPs as regulators of mRNA translation and cell proliferation," *Trends in Biochemical Sciences*, vol. 22, no. 9, pp. 345–349, 1997.

[23] M. Laplante and D. M. Sabatini, "mTOR signaling at a glance," *Journal of Cell Science*, vol. 122, no. 20, pp. 3589–3594, 2009.

[24] A. C. Gingras, S. P. Gygi, B. Raught et al., "Regulation of 4E-BP1 phosphorylation: a novel two step mechanism," *Genes and Development*, vol. 13, no. 11, pp. 1422–1437, 1999.

[25] M. Laplante and D. M. Sabatini, "An emerging role of mTOR in lipid biosynthesis," *Current Biology*, vol. 19, no. 22, pp. R1046–R1052, 2009.

[26] M. E. Feldman, B. Apsel, A. Uotila et al., "Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2," *PLoS Biology*, vol. 7, no. 2, article e38, 2009.

[27] T. E. Dever, "Gene-specific regulation by general translation factors," *Cell*, vol. 108, no. 4, pp. 545–556, 2002.

[28] T. J. McGarry and S. Lindquist, "The preferential translation of Drosophila hsp70 mRNA requires sequences in the untranslated leader," *Cell*, vol. 42, no. 3, pp. 903–911, 1985.

[29] R. Klemenz, D. Hultmark, and W. J. Gehring, "Selective translation of heat shock mRNA in Drosophila melanogaster depends on sequence information in the leader," *EMBO Journal*, vol. 4, no. 8, pp. 2053–2060, 1985.

[30] R. Ahmed and R. F. Duncan, "Translational regulation of Hsp90 mRNA: AUG-proximal 5′-untranslated region elements essential for preferential heat shock translation," *Journal of Biological Chemistry*, vol. 279, no. 48, pp. 49919–49930, 2004.

[31] D. L. Sacks and P. V. Perkins, "Identification of an infective stage of Leishmania promastigotes," *Science*, vol. 223, no. 4643, pp. 1417–1419, 1984.

[32] P. F. P. Pimenta, S. J. Turco, M. J. McConville, P. G. Lawyer, P. V. Perkins, and D. L. Sacks, "Stage-specific adhesion of Leishmania promastigotes to the sandfly midgut," *Science*, vol. 256, no. 5065, pp. 1812–1815, 1992.

[33] K. P. Chang and D. M. Dwyer, "Multiplication of a human parasite (Leishmania donovani) in phagolysosomes of hamster macrophages *in vitro*," *Science*, vol. 193, no. 4254, pp. 678–680, 1976.

[34] K. P. Chang, D. Fong, and R. S. Bray, "Biology of Leishmania and leishmaniasis," in *Leishmaniasis*, K. P. Chang and R. S. Bray, Eds., pp. 1–30, Elsevier, Amsterdam, The Netherlands, 1985.

[35] A. P. Fernandes, K. Nelson, and S. M. Beverley, "Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 24, pp. 11608–11612, 1993.

[36] J. D. Bangs, P. F. Crain, T. Hashizume, J. A. McCloskey, and J. C. Boothroyd, "Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides," *Journal of Biological Chemistry*, vol. 267, no. 14, pp. 9805–9815, 1992.

[37] C. E. Clayton, "Life without transcriptional control? From fly to man and back again," *EMBO Journal*, vol. 21, no. 8, pp. 1881–1888, 2002.

[38] S. Michaeli, "Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome," *Future Microbiology*, vol. 6, no. 4, pp. 459–474, 2011.

[39] C. Clayton and M. Shapira, "Post-transcriptional regulation of gene expression in trypanosomes and leishmanias," *Molecular and Biochemical Parasitology*, vol. 156, no. 2, pp. 93–101, 2007.

[40] X. H. Liang, A. Haritan, S. Uliel, and S. Michaeli, "trans and cis splicing in trypanosomatids: mechanism, factors, and regulation," *Eukaryotic Cell*, vol. 2, no. 5, pp. 830–840, 2003.

[41] M. Mandelboim, C. L. Estraño, C. Tschudi, E. Ullu, and S. Michaeli, "On the role of exon and intron sequences in trans-splicing utilization and cap 4 modification of the trypanosomatid Leptomonas collosoma SL RNA," *Journal of Biological Chemistry*, vol. 277, no. 38, pp. 35210–35218, 2002.

[42] J. P. Ruan, E. Ullu, and C. Tschudi, "Characterization of the Trypanosoma brucei cap hypermethylase Tgs1," *Molecular and Biochemical Parasitology*, vol. 155, no. 1, pp. 66–69, 2007.

[43] G. K. Arhin, E. Ullu, and C. Tschudi, "2′ -O-Methylation of position 2 of the trypanosome spliced leader cap 4 is mediated by a 48 kDa protein related to vaccinia virus VP39," *Molecular and Biochemical Parasitology*, vol. 147, no. 1, pp. 137–139, 2006.

[44] J. R. Zamudio, B. Mittra, D. A. Campbell, and N. R. Sturm, "Hypermethylated cap 4 maximizes Trypanosoma brucei translation," *Molecular Microbiology*, vol. 72, no. 5, pp. 1100–1110, 2009.

[45] J. R. Zamudio, B. Mittra, G. M. Zeiner et al., "Complete cap 4 formation is not required for viability in Trypanosoma brucei," *Eukaryotic Cell*, vol. 5, no. 6, pp. 905–915, 2006.

[46] Y. Yoffe, J. Zuberek, A. Lerer et al., "Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania," *Eukaryotic Cell*, vol. 5, no. 12, pp. 1969–1979, 2006.

[47] J. Marcotrigiano, A. C. Gingras, N. Sonenberg, and S. K. Burley, "Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP," *Cell*, vol. 89, no. 6, pp. 951–961, 1997.

[48] H. Matsuo, H. Li, A. M. McGuire et al., "Structure of translation factor elF4E bound to m7GDP and interaction with 4E-binding protein," *Nature Structural Biology*, vol. 4, no. 9, pp. 717–724, 1997.

[49] P. E. C. Hershey, S. M. McWhirter, J. D. Gross, G. Wagner, T. Alber, and A. B. Sachs, "The cap-binding protein eIF4E promotes folding of a functional domain of yeast translation initiation factor eIF4G1," *Journal of Biological Chemistry*, vol. 274, no. 30, pp. 21297–21304, 1999.

[50] K. Tomoo, X. Shen, K. Okabe et al., "Structural features of human initiation factor 4E, studied by X-ray crystal analyses and molecular dynamics simulations," *Journal of Molecular Biology*, vol. 328, no. 2, pp. 365–383, 2003.

[51] M. Altmann, P. P. Muller, J. Pelletier, N. Sonenberg, and H. Trachsel, "A mammalian translation initiation factor can substitute for its yeast homologue in vivo," *Journal of Biological Chemistry*, vol. 264, no. 21, pp. 12145–12147, 1989.

[52] R. Dhalia, C. R. S. Reis, E. R. Freire et al., "Translation initiation in Leishmania major: characterisation of multiple eIF4F subunit homologues," *Molecular and Biochemical Parasitology*, vol. 140, no. 1, pp. 23–41, 2005.

[53] M. Lewdorowicz, Y. Yoffe, J. Zuberek et al., "Chemical synthesis and binding activity of the trypanosomatid cap-4 structure," *RNA*, vol. 10, no. 9, pp. 1469–1478, 2004.

[54] E. R. Freire, R. Dhalia, D. M. N. Moura et al., "The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties," *Molecular and Biochemical Parasitology*, vol. 176, no. 1, pp. 25–36, 2011.

[55] Y. Yoffe, M. Léger, A. Zinoviev et al., "Evolutionary changes in the Leishmania eIF4F complex involve variations in the eIF4E-eIF4G interactions," *Nucleic Acids Research*, vol. 37, no. 10, pp. 3243–3253, 2009.

[56] A. Zinoviev, M. Leger, G. Wagner, and M. Shapira, "A novel 4E-interacting protein in Leishmania is involved in stage-specific translation pathways," *Nucleic Acids Research*, vol. 39, pp. 8404–8415, 2011.

[57] L. S. Hiremath, S. T. Hiremath, W. Rychlik, S. Joshi, L. L. Domier, and R. E. Rhoads, "*in vitro* synthesis, phosphorylation, and localization on 48 S initiation complexes of human protein synthesis initiation factor 4E," *Journal of Biological Chemistry*, vol. 264, no. 2, pp. 1132–1138, 1989.

[58] M. Rau, T. Ohlmann, S. J. Morley, and V. M. Pain, "A reevaluation of the Cap-binding protein, eIF4E, as a rate-limiting factor for initiation of translation in reticulocyte lysate," *Journal of Biological Chemistry*, vol. 271, no. 15, pp. 8983–8990, 1996.

[59] S. Kramer, R. Queiroz, L. Ellis et al., "Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2α phosphorylation at Thr169," *Journal of Cell Science*, vol. 121, no. 18, pp. 3002–3014, 2008.

[60] A. Cassola, J. G. De Gaudenzi, and A. C. Frasch, "Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes," *Molecular Microbiology*, vol. 65, no. 3, pp. 655–670, 2007.

[61] J. Marcotrigiano, I. B. Lomakin, N. Sonenberg, T. V. Pestova, C. U. T. Hellen, and S. K. Burley, "A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery," *Molecular Cell*, vol. 7, no. 1, pp. 193–203, 2001.

[62] A. Marintchev and G. Wagner, "eIF4G and CBP80 share a common origin and similar domain organization: implications for the structure and function of eIF4G," *Biochemistry*, vol. 44, no. 37, pp. 12265–12272, 2005.

[63] S. Mader, H. Lee, A. Pause, and N. Sonenberg, "The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4γ and the translational repressors 4E-binding proteins," *Molecular and Cellular Biology*, vol. 15, no. 9, pp. 4990–4997, 1995.

[64] R. Dhalia, N. Marinsek, C. R. S. Reis et al., "The two eIF4A helicases in Trypanosoma brucei are functionally distinct," *Nucleic Acids Research*, vol. 34, no. 9, pp. 2495–2507, 2006.

[65] L. M. D. Silva, K. L. Owens, S. M. F. Murta, and S. M. Beverley, "Regulated expression of the Leishmania major surface virulence factor lipophosphoglycan using conditionally destabilized fusion proteins," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 18, pp. 7583–7588, 2009.

[66] P. Linder, "Dead-box proteins: a family affair–active and passive players in RNP-remodeling," *Nucleic Acids Research*, vol. 34, no. 15, pp. 4168–4180, 2006.

[67] M. A. Ferraiuolo, C. S. Lee, L. W. Ler et al., "A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 12, pp. 4118–4123, 2004.

[68] A. Kahvejian, Y. V. Svitkin, R. Sukarieh, M. N. M'Boutchou, and N. Sonenberg, "Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms," *Genes and Development*, vol. 19, no. 1, pp. 104–113, 2005.

[69] R. C. Deo, J. B. Bonanno, N. Sonenberg, and S. K. Burley, "Recognition of polyadenylate RNA by the poly(A)-binding protein," *Cell*, vol. 98, no. 6, pp. 835–845, 1999.

[70] S. E. Wells, P. E. Hillner, R. D. Vale, and A. B. Sachs, "Circularization of mRNA by eukaryotic translation initiation factors," *Molecular Cell*, vol. 2, no. 1, pp. 135–140, 1998.

[71] M. C. Derry, A. Yanagiya, Y. Martineau, and N. Sonenberg, "Regulation of poly(A)-binding protein through PABP-interacting proteins," *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 71, pp. 537–543, 2006.

[72] R. J. Jackson, "Alternative mechanisms of initiating translation of mammalian mRNAs," *Biochemical Society Transactions*, vol. 33, no. 6, pp. 1231–1241, 2005.

[73] T. D. da Costa Lima, D. M. N. Moura, C. R. S. Reis et al., "Functional characterization of three Leishmania poly(A) binding protein homologues with distinct binding properties to RNA and protein partners," *Eukaryotic Cell*, vol. 9, no. 10, pp. 1484–1494, 2010.

[74] E. J. Bates, E. Knuepfer, and D. F. Smith, "Poly(A)-binding protein I of Leishmania: functional analysis and localisation in trypanosomatid parasites," *Nucleic Acids Research*, vol. 28, no. 5, pp. 1211–1220, 2000.

[75] L. Quijada, M. Soto, C. Alonso, and J. M. Requena, "Identification of a putative regulatory element in the 3′-untranslated region that controls expression of HSP70 in Leishmania infantum," *Molecular and Biochemical Parasitology*, vol. 110, no. 1, pp. 79–91, 2000.

[76] C. Folgueira, L. Quijada, M. Soto, D. R. Abanades, C. Alonso, and J. M. Requena, "The translational efficiencies of the two Leishmania infantum HSP70 mRNAs, differing in their 3′-untranslated regions, are affected by shifts in the temperature of growth through different mechanisms," *Journal of Biological Chemistry*, vol. 280, no. 42, pp. 35172–35183, 2005.

[77] A. Zilka, S. Garlapati, E. Dahan, V. Yaolsky, and M. Shapira, "Developmental regulation of heat shock protein 83 in Leishmania: 3′ processing and mRNA stability control transcript abundance, and translation is directed by a determinant in

the 3′-untranslated region," *Journal of Biological Chemistry*, vol. 276, no. 51, pp. 47922–47929, 2001.

[78] R. Larreta, M. Soto, L. Quijada et al., "The expression of HSP83 genes in Leishmania infantum is affected by temperature and by stage-differentiation and is regulated at the levels of mRNA stability and translation," *BMC Molecular Biology*, vol. 5, article 3, 2004.

[79] M. David, I. Gabdank, M. Ben-David et al., "Preferential translation of Hsp83 in Leishmania requires a thermosensitive polypyrimidine-rich element in the 3′ UTR and involves scanning of the 5′ UTR," *RNA*, vol. 16, no. 2, pp. 364–374, 2010.

[80] N. R. Markham and M. Zuker, "UNAFold: software for nucleic acid folding and hybridization," *Methods in Molecular Biology*, vol. 453, pp. 3–31, 2008.

[81] N. Boucher, Y. Wu, C. Dumas et al., "A common mechanism of stage-regulated gene expression in Leishmania mediated by a conserved 3′-untranslated region element," *Journal of Biological Chemistry*, vol. 277, no. 22, pp. 19511–19520, 2002.

[82] F. McNicoll, M. Müller, S. Cloutier et al., "Distinct 3′ - untranslated region elements regulate stage-specific mRNA accumulation and translation in Leishmania," *Journal of Biological Chemistry*, vol. 280, no. 42, pp. 35238–35246, 2005.

[83] M. C. S. Moraes, T. C. L. Jesus, N. N. Hashimoto et al., "Novel membrane-bound eIF2α kinase in the flagellar pocket of Trypanosoma brucei," *Eukaryotic Cell*, vol. 6, no. 11, pp. 1979–1991, 2007.

[84] T. Lahav, D. Sivam, H. Volpin et al., "Multiple levels of gene regulation mediate differentiation of the intracellular pathogen Leishmania," *FASEB Journal*, vol. 25, no. 2, pp. 515–525, 2011.

[85] M. Argaman, R. Aly, and M. Shapira, "Expression of heat shock protein 83 in Leishmania is regulated post-transcriptionally," *Molecular and Biochemical Parasitology*, vol. 64, no. 1, pp. 95–110, 1994.

[86] C. S. Peacock, K. Seeger, D. Harris et al., "Comparative genomic analysis of three Leishmania species that cause diverse human disease," *Nature Genetics*, vol. 39, no. 7, pp. 839–847, 2007.

[87] S. Haile and B. Papadopoulou, "Developmental regulation of gene expression in trypanosomatid parasitic protozoa," *Current Opinion in Microbiology*, vol. 10, no. 6, pp. 569–577, 2007.

[88] D. Rosenzweig, D. Smith, F. Opperdoes, S. Stern, R. W. Olafson, and D. Zilberstein, "Retooling Leishmania metabolismml: from sand fly gut to human macrophage," *FASEB Journal*, vol. 22, no. 2, pp. 590–602, 2008.

[89] M. Jaramillo, M. A. Gomez, O. Larson et al., "Translational control of Leishmania infection throught mTORC1 signaling," *Cell Host and Microbe*, vol. 9, pp. 331–341, 2011.

[90] N. J. Moerke, H. Aktas, H. Chen et al., "Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G," *Cell*, vol. 128, no. 2, pp. 257–267, 2007.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 134839, 21 pages
doi:10.1155/2012/134839

Review Article

Diversity of Eukaryotic Translational Initiation Factor eIF4E in Protists

Rosemary Jagus,[1] Tsvetan R. Bachvaroff,[2] Bhavesh Joshi,[3] and Allen R. Place[1]

[1] *Institute of Marine and Environmental Technology, University of Maryland Center for Environmental Science, 701 E. Pratt Street, Baltimore, MD 21202, USA*
[2] *Smithsonian Environmental Research Center, 647 Contees Wharf Road, Edgewater, MD 21037, USA*
[3] *BridgePath Scientific, 4841 International Boulevard, Suite 105, Frederick, MD 21703, USA*

Correspondence should be addressed to Rosemary Jagus, jagus@umces.edu

Received 26 January 2012; Accepted 9 April 2012

Academic Editor: Thomas Preiss

The greatest diversity of eukaryotic species is within the microbial eukaryotes, the protists, with plants and fungi/metazoa representing just two of the estimated seventy five lineages of eukaryotes. Protists are a diverse group characterized by unusual genome features and a wide range of genome sizes from 8.2 Mb in the apicomplexan parasite *Babesia bovis* to 112,000-220,050 Mb in the dinoflagellate *Prorocentrum micans*. Protists possess numerous cellular, molecular and biochemical traits not observed in "text-book" model organisms. These features challenge some of the concepts and assumptions about the regulation of gene expression in eukaryotes. Like multicellular eukaryotes, many protists encode multiple eIF4Es, but few functional studies have been undertaken except in parasitic species. An earlier phylogenetic analysis of protist eIF4Es indicated that they cannot be grouped within the three classes that describe eIF4E family members from multicellular organisms. Many more protist sequences are now available from which three clades can be recognized that are distinct from the plant/fungi/metazoan classes. Understanding of the protist eIF4Es will be facilitated as more sequences become available particularly for the under-represented opisthokonts and amoebozoa. Similarly, a better understanding of eIF4Es within each clade will develop as more functional studies of protist eIF4Es are completed.

1. Eukaryogenesis and Protein Synthesis

Protein synthesis is an ancient, conserved, complex multienzyme system, involving the participation of hundreds of macromolecules in which the mRNA template is decoded into a protein sequence on the ribosome. The ribosome, a complex and dynamic nucleoprotein machine, provides the platform for amino acid polymerization in all organisms [1, 2]. This process utilizes mRNAs, aminoacyl tRNAs, and a range of protein factors, as well as the inherent peptidyl-transferase activity of the ribosome itself. The common origin of protein synthesis in all domains of life is evident in the conservation of tRNA and ribosome structure, as well as some of the additional protein factors. Although the basic molecular mechanisms are conserved across the three domains of life, the Bacteria (eubacteria), Archaea (archaebacteria), and Eukarya (eukaryotes), important divergences have taken place as eukaryotic species have evolved. The origin of the eukaryotic cell is enigmatic. Eukaryotes are thought to have evolved from a fusion of a euryarchaeon with a deep-rooted Gram-positive proteobacteria, the phylum from which mitochondria are derived [3]. It is currently unclear whether the eubacterial fusion partner was distinct from the ancestor of mitochondria or identical to it. This view of the origins of eukaryotes is consistent with the observation that informational genes such as those involved in transcription, translation, and other related processes are most closely related to archaeal genes, whereas operational genes such as those involved in cellular metabolic processes including amino acid biosynthesis, cell envelope, and lipid synthesis are most closely related to eubacterial genes [4]. Such an origin is also consistent with the eukaryotic rooting implied by the presence of an insert within the elongation factor EF-1A that is found in all known eukaryotic and

eocytic (crenarchaeal) EF-1A sequences, but lacking in all paralogous EF-G sequences [3].

The mechanisms underlying protein synthesis in all organisms share common features and can be divided into three stages: *initiation*, *elongation*, *and termination*. During *initiation*, the ribosome is assembled at the initiation codon in the mRNA with a methionyl initiator tRNA bound in the peptidyl (P) site. During *elongation*, aminoacyl tRNAs enter the acceptor (A) site and the ribosome catalyzes the formation of a peptide bond. After the tRNAs and mRNA are translocated bringing the next codon into the A-site, the elongation process is repeated until a stop codon is encountered. During *termination*, the completed polypeptide is released from the ribosome, after which the ribosomal subunits are dissociated and the mRNA released for reuse. Different sets of protein accessory factors, the translation factors, assist the ribosome at each of these stages. These are referred to as initiation factors, elongation factors, and termination factors, respectively, to reflect the stage at which they are involved. The elongation process and machinery is well conserved from bacteria to eukaryotes, as is termination. However, the mechanisms of the initiation process, including recognition of the correct reading frame, differ, as do the mechanisms by which mRNA is recruited by the ribosome. Genomewide sequencing projects now allow us to assess the components of translational initiation in a wide range of organisms [5, 6].

Our view of protein synthesis is based mainly on information derived from *S. cerevisiae*, *Drosophila*, plant, and mammalian systems, with the translation components identified through sequencing projects. However, these are only narrow windows on the full diversity of extant eukaryotes. The greatest diversity of eukaryotic species is to be found within the protists, with plants and metazoans representing just two of the estimated 75 lineages of eukaryotes [7, 8]. We are only just beginning to uncover the vast diversity of bacterial-sized (pico- and nano-) eukaryotes, first discovered in clone libraries derived by PCR amplification of pooled "environmental" DNAs (culture-independent PCR) [9–11]. Microbial eukaryotes are a diverse group of organisms characterized by many unusual genome features. These features challenge some of the concepts and assumptions about regulation of gene expression in eukaryotes. In this paper, we will focus on a comparison of our current knowledge of the translation initiation factor eIF4E and its family members from protists. We will compare eIF4E in a range of protists and look at translational components in a simplified translation system found in an algal endosymbiont.

The control of gene expression is a complex process. Even after mRNA is transcribed from DNA, mRNAs can undergo many processing and regulatory steps that influence their expression [12]. Gene regulation at the translational level is widespread and significant. The extent of gene regulation at the translational level has been demonstrated during early *Drosophila* embryogenesis on a genomewide basis that was investigated by determining ribosomal density and ribosomal occupancy of over 10,000 transcripts during the first ten hours after egg laying in *Drosophila*. The diversity of the translation profiles indicates multiple mechanisms modulating transcript-specific translation with cluster analyses suggesting that the genes involved in some biological processes are coregulated at the translational level at certain developmental stages [13]. Similarly, protists have been shown to regulate translation over wide range of conditions and physiological changes, with groups like the dinoflagellates showing regulation of translation to be the predominant form of regulation of gene expression.

2. Origin of Eukaryotes

Eubacteria and Archaea show tremendous diversity in their metabolic capabilities but have limited morphological and behavioral diversity; conversely, eukaryotes share similar metabolic machinery but have tremendous morphological and behavioral diversity. Eukaryotes are thought to have evolved from the endosymbiosis of an α-proteobacteria and a phagotropic euryarchaeon approximately 2 billion years ago. The transition from prokaryotes to eukaryotes was the most radical change in cell organization since life began, with a burst of gene transfer, duplication, and the appearance of novel cell structures and processes such as the nucleus, the endomembrane system, actin-based cytoskeleton [14, 15], the spliceosome and splicing, nonsense-mediated decay of mRNA (NMD), and ubiquitin signaling [16, 17]. Although the deep phylogeny of eukaryotes currently should be considered unresolved, Koonin and his colleagues have postulated that the mitochondrial endosymbiont spawned an intron invasion which contributed to the emergence of these principal features of the eukaryotic cell [18–20]. Phagocytosis is thought to be central to the origin of the eukaryotic cell for the acquisition of the bacterial endosymbiont that became the ancestor of the mitochondrion. Findings suggest a hypothetical scenario of eukaryogenesis under which the archaeal ancestor of eukaryotes had no cell wall (like modern *Thermoplasma*) but had an actin-based cytoskeleton that allowed the euryarcheon to produce actin-supported membrane protrusions. These protrusions would enable accidental, occasional engulfment of bacteria, one of which would eventually became the mitochondrion. The acquisition of the endosymbiont triggered eukaryogenesis. From a fused cell with two independent prokaryotic gene expression systems, coordination of cell division developed and gene transfer took place through occasional membrane lysis. Some of eubacterial genes recombined into host chromosomes including group II introns [18]. Group II introns can be found among free-living α-proteobacteria, the ancestors of mitochondria [21]. They evolved specifically from group II introns that invaded the ancestrally intronless eukaryotic genome through the mitochondrial endosymbiont, thereby generating the prediction that group II introns should be found among free-living-proteobacteria, the ancestors of mitochondria [21]. This prediction was borne out supporting the idea that introns could originate from the mitochondrial endosymbiont. The mobility of group II introns in contemporary eubacteria [22] and their prevalence in α-proteobacteria [23] are consistent with such a view. The rapid, coincidental spread of introns following the origin of mitochondria is posited as the selective pressure

that forged nucleus-cytosol compartmentalization [18, 20]. The function of the nuclear envelope was to allow mRNA splicing, which is slow, to go to completion so that translation, which is fast, would occur only on mRNA with intact reading frames. The evolutionary relationships of proteins specific to the nuclear envelope and nuclear pore complex reveal that this protein set is a mix of proteins and domains of archaebacterial and eubacterial origins, along with some eukaryotic innovations, suggesting that the nucleus arose in a cell that already contained a mitochondrial endosymbiont [24].

3. Evolution of Translational Initiation and Eukaryogenesis

Eukaryotes inherited from their archaeal ancestor a core of translation initiation factors, which includes eukaryotic initiation factor (eIF)1, eIF1A, eIF2 (all three subunits), eIF2B (α, β, and δ subunits only) subunits), eIF4A, eIF5B, and eIF6 [25–27]. The establishment of the nuclear membrane resulted in the physical separation of transcription and translation and presented early eukaryotes with a different challenge; how to shuttle RNA from the nucleus to the site of protein synthesis in the cytoplasm. In prokaryotes, mRNA is translated as it is being synthesized, whereas in eukaryotes, mRNA is synthesized, and processed in the nucleus, and it is then exported to the cytoplasm. There is also a transition from uncapped and polycistronic mRNAs recognized by the ribosome through the Shine-Dalgarno sequence in the 5′-UTR to capped, polyadenylated, and, in most cases, monocistronic mRNAs and the evolution of the scanning process. The evolution of protein synthesis in the context of eukaryogenesis has been discussed previously by Hernández who proposed that recruitment of mRNAs in early eukaryotes was likely to have been through internal ribosome entry sites (IRESs) based on the functional similarity between IRESs and introns [28]. Although not universal, IRES transacting factors (ITAFs) are required for the proper functioning of most viral and cellular IRESs [29, 30]. ITAFs are predominantly nuclear proteins that also play key roles in pre-mRNA splicing and mRNA transport to the cytoplasm [31, 32]. Furthermore, polypyrimidine tracts, a hallmark of introns, are a common feature of cellular and some viral IRESs [33–35]. Hernández considers that the cellular IRESs are descendants of spliceosomal introns and that some of the ITAFs that existed as components of the splicing machinery (such as the ancestral PTB and hnRNPCs) were later incorporated into the nascent eukaryotic translational process. During this period, 5′-UTRs lacking Shine-Dalgarno motifs that were able to passively recruit the 40S ribosomal subunit would have been positively selected and could, therefore, have become the first examples of an IRES [28].

It also seems possible that capped spliced leader (SL) *trans*-spliced mRNAs may have arisen with eukaryogenesis and represent an early form of 5′ blocked mRNAs. In *trans*-splicing, a short SL exon is spliced from a capped small nuclear RNA and is transferred to pre-mRNA, thereby becoming the 5′-terminal end. The fully functional spliceosome is likely to have existed in the last eukaryote common ancestor, leading to splicing components and pre-mRNA signals that are found throughout eukaryotes and are similar among different eukaryotic lineages. It seems certain that SL *trans*-splicing arose through evolution from *cis*-splicing or *vice versa*. *Trans*-splicing shares the splicing signals and most of the components with *cis*-splicing, indicating a common relationship (reviewed [36]). Considering the similarities between the SL snRNP and the spliceosomal snRNPs, specialized *trans*-splicing SL RNAs could have arisen from a splicing U snRNP in ancestral *cis*-splicing early eukaryote and thus may be an ancient form of 5′-end blocking for emerging eukaryotes. SL *trans*-splicing is now found sporadically across the eukaryotic tree of life in a set of distantly related animal groups including urochordates, nematodes, flatworms, and hydra, as well as in the protist Euglenozoa and dinoflagellates, stimulating the argument that a common evolutionary origin seems unlikely. However, an attractive hypothesis to explain multiple evolutionary origins for the SL genes is that they have derived repeatedly from U-rich small nuclear RNAs (snRNAs) of the Sm-class involved in the nuclear spliceosome machinery [37]. In support of this, phylogenomic studies from *Hydra* indicate that SL genes can evolve rapidly in any organism because constraint on SL exon sequence evolution is low [38]. Furthermore, it has been reported that mammalian cells, which do not have SL *trans*-splicing, can SL *trans*-splice when supplied with the SL RNA of either nematodes or trypanosomes [39]. Duplications of the U1 snRNA gene followed by just a few mutations would be sufficient to lead to the acquisition of *trans*-splicing [39] suggesting that it could have happened in the emerging eukaryote as well as in more recent eukaryotic lines.

The separation of the nucleus from the cytoplasm led to the need for mechanisms to shuttle the transcripts into the cytoplasm and to provide for their protection against degradation. With the exception of eIF5, all the eukaryotic-specific initiation factors that evolved, eIF4E, eIF4G, eIF4B, eIF4H, and eIF3, are involved in the 5′-cap-binding and scanning processes. The 5′-cap structure provides stability from 5′ exonucleases and in extant eukaryotes is recognized by the small ribosomal subunit through the novel eukaryotic initiation factor eIF4E. eIF4E, a translational initiation factor found only in eukaryotes, has a unique alpha/beta fold that is considered to have no homologues outside the eukaryotes, as determined by sequence comparison or structural analyses [25]. Although in extant eukaryotes the main role of eIF4E is in translational initiation through cap recognition, it is possible that the cap structure and eIF4E emerged among the primary adaptive responses to the intron invasion and the need for nucleocytoplasmic RNA export, but initially had no role in translation [40]. For instance, it could have appeared in early eukaryotes either as a mediator of nuclear export of mRNAs, thus enhancing mRNA stability during nuclear export, or as a mediator of cytoplasmic storage of mRNAs. Consistent with this, one of the eIF4E proteins from the primitive eukaryote species *Giardia lamblia* binds only to nuclear noncoding small RNAs and has no function in translation [41]. eIF4E is found within different cytoplasmic bodies involved in such processes as mRNP remodeling, mRNA decay or storage [42–44]. In addition, a fraction of

this protein resides in the nucleus where it mediates the export of specific mRNAs to the cytoplasm [44, 45]. Since eIF4E has no ability to interact directly with the ribosome itself, the recruitment of eIF4E-bound mRNAs in emerging eukaryotes was likely to have been IRES-dependent.

4. Diversity of eIF4E Family Members

In eukaryotes, eIF4E is a central component in the initiation and regulation of translation in eukaryotic cells [46–49]. Through its interaction with the 5′-cap structure of mRNA and its translation partner, eIF4G, eIF4E functions to recruit mRNAs to the ribosome [46]. The interaction of eIF4E and eIF4G can be competed out by a family of 4E-binding proteins, the 4E-BPs, which are capable of repressing translation [46]. Three-dimensional structures of eIF4Es bound to cap-analogues resemble "cupped-hands" in which the cap-structure is sandwiched between two conserved Trp residues (W56 and W102 of *H. sapiens* eIF4E) [50–52]. A third conserved Trp residue (W166 of *H. sapiens* eIF4E) recognizes the 7-methyl moiety of the cap-structure. Aromatic residues Trp, Phe, and His show a distinctive pattern across from N- to C-terminus of the conserved core, containing eight similarly spaced tryptophans summarized by W(x2)W(x8–12)W(x17–20)W(x29–31)W(x9–12)W(x17)W(x32–36)W [6]. Multiple eIF4E family members have been identified in a wide range of organisms that includes plants, flies, mammals, frogs, birds, nematodes, fish, and various protists [53–55]. Evolutionarily, it seems that a single early eIF4E gene underwent a series of gene duplications, generating multiple structural classes and in some cases subclasses. Today, eIF4E and its relatives comprise a family of structurally related proteins within a given organism, although not all function as prototypical initiation factors. Sequence similarity is highest in a core region of 160 to 170 amino acid residues identified by evolutionary conservation and functional analyses [6]. Prototypical eIF4E is considered to be eIF4E-1 of mammals, eIF4E and eIF (iso)4E of plants, and eIF4E of *Saccharomyces cerevisiae*. With the exception of eIF4Es from protists, all eIF4Es can be grouped into one of three classes [6].

Class I members from Viridiplantae, Metazoa, and Fungi carry Trp residues equivalent to W43, W46, W56, W73, W102, W113, W130, and W166 of *H. sapiens* eIF4E-1 [6]. Prototypical eIF4Es bind the cap and eIF4G through the motif S/TVE/DE/DFW in which the Trp is W73. Substitution of a nonaromatic amino acid for W73 has been shown to disrupt the ability of eIF4E to interact with eIF4G and 4E-BPs [56, 57]. Substitution of a Gly residue in place of V69 creates an eIF4E variant that still binds 4E-BP1 but has a reduced capacity to interact with both eIF4G and 4E-BP2 [56]. A serine at residue equivalent to S209 in *H. sapiens* eIF4E-1 is the site of phosphorylation. Only Class I eIF4Es are known to function as translation factors. Genes, and cDNAs encoding members of Class I can be identified in species from plants/metazoans/fungi. As judged from completed genomes, many protists also encode Class I-like family members although these have proven hard to characterize and can show extension or compaction relative to prototypical eIF4E family members [6]. Evidence for gene duplication of Class I eIF4E family members can be found in certain plant species, as well as in nematodes, insects, chordates, and some fungi [53–55]. Class I members include the prototypical initiation factor but may also include eIF4Es that recognize alternative cap structures such as IFE-1, -2, and -5 of *Caenorhabditis elegans* [58, 59], or eIF4Es that fulfill regulatory functions such as the vertebrate eIF4E-1Bs [55, 60–62].

Class II members possess W → Y/F/L and W → Y/F substitutions relative to W43 and W56 of *H. sapiens* eIF4E. These substitions are absent from the model ascomycetes *S. cerevisiae* and *Schizosaccharomyces pombe*. Mammalian eIF4E-2 (Class II) binds only to cap and 4E-BPs [54]. They have been shown to regulate specific mRNA recruitment in *Drosophila* [63] and *C. elegans* [64].

Class III members possess a Trp residue equivalent to W43 of *H. sapiens* eIF4E but carry a W → C/Y substitution relative to *H. sapiens* W56. They have been identified primarily in chordates with rare examples in other Coelomata and in Cnidaria [6, 54]. Their biological function has not yet been determined, although mouse eIF4E-3 has been shown to bind both cap and eIF4G [54]. The protist eIF4Es do not fall into any of these three classes and by plant/metazoan/fungal standards appear to be compacted or possess extended sequences between the conserved tryptophans [6].

5. Diversity of Protists and Evolution of Eukaryotic Lineages

The greatest diversity of eukaryotic species is to be found within the protists. Eukaryotes appear to be monophyletic; all extant eukaryotes appear to postdate the acquisition of mitochondria. However, their phylogeny is currently not widely agreed upon. Molecular phylogenetics has the potential to resolve the systematics of eukaryotes. Sequence data continues to accumulate, but with few protists and fewer protist taxa and a distinct bias towards parasites infecting humans (and crop plants). There is increasing availability of multigene data from diverse lineages, although it seems likely that eukaryotic taxonomy will be further complicated by the discovery of ultrasmall eukaryotes. These are scattered across the eukaryotic tree and may include major new supergroups [9, 65]. The root of the eukaryotes remains open to debate, but recent analysis places the eukaryotic root between the monophyletic "unikonts" and "bikonts" [66].

The protists are defined loosely as unicellular eukaryotic organisms that are not plants, animals, or fungi. Eukaryotic features evolved within the protists that thrived for up to a billion years before they gave rise independently to multicellular eukaryotes, the familiar plants, animals, and fungi [67]. Extreme examples of genome sizes, both large and small, can be found among microbial eukaryotes from 8.2 Mb in the apicomplexan *Babesia bovis* to >200,000 Mb in certain dinoflagellates. Roughly forty sequenced genomes are available (depending on classification), some of which are multiple representatives of the same genus, for example, *Plasmodium*, *Leishmania*, and *Trypanosoma*. The last common ancestor of all eukaryotes is believed to have been

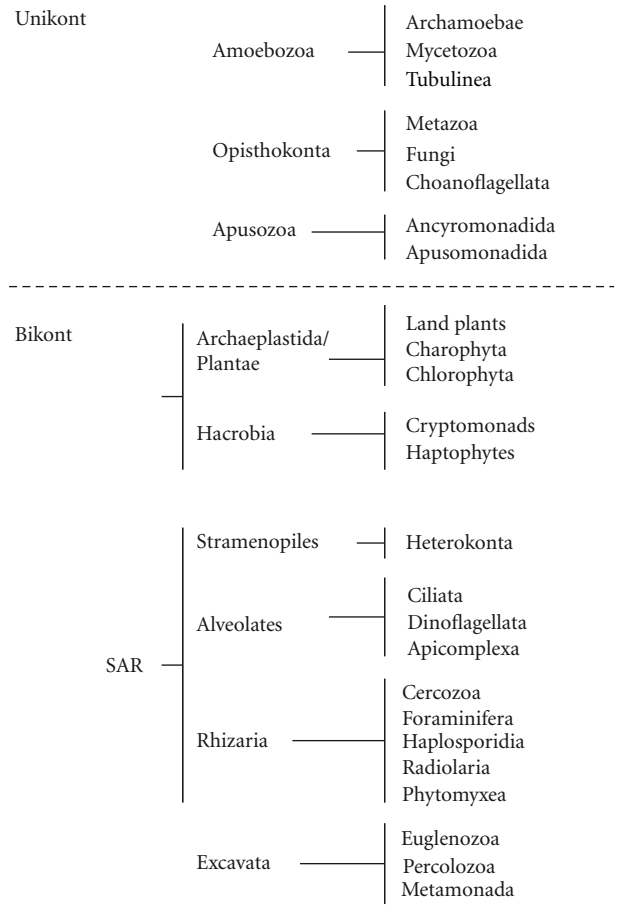


Figure 1: Relationships among major lineages of eukaryotes. Summary tree of eukaryotic relationships based on multigene analyses as outlined by Parfrey et al. [8].

a phagotrophic protist with a nucleus, at least one centriole and cilium, facultatively aerobic mitochondria, sex (meiosis and syngamy) and a dormant cyst with a cell wall of chitin and/or cellulose, and peroxisomes (based on a root along the lineage leading to Euglenozoa). Endosymbiosis led to the spread of plastids. Analyses of multigene genealogies have led to the conclusion that the acquisition of photosynthesis in eukaryotes arose from a primary endosymbiosis between a cyanobacterium and a eukaryotic host. This gave rise to glaucocystophytes (white lineage), red algae (red lineage), and green algae (green lineage, including plants) [7, 68–70]. Plastids spread by secondary endosymbiosis. Other photosynthetic eukaryotes such as cryptomonads, haptophytes, chlorarachniophytes (amoeboflagellate cercozoans), dinoflagellates, diatoms, brown algae, and euglenids are the result of secondary endosymbiosis, tertiary endosymbiosis, and, perhaps, even quaternary endosymbiosis in which a nonphotosynthetic eukaryotic ancestor engulfed a photosynthetic eukaryote [68, 71, 72]. Endosymbiosis resulted in the transfer of hundreds of genes to the host nucleus. Multiple gains and multiple losses of plastids are likely to have occurred, with plastids possibly lost in ciliates and remaining in relict form in apicomplexans [73] and *Perkinsus* [74]. Dinoflagellates have substituted the ancestral plastid several times by tertiary symbioses involving a diverse array of eukaryotes [71, 72].

There is no real consensus on eukaryotic phylogeny currently; part of the problem is that we are still very much in the discovery phase, and another is that some of the divisions are quite ancient. In recent years, eukaryotic taxonomy has shifted towards a new system of six supergroups that aims to portray evolutionary relationships between microbial and macrobial lineages [8, 75–77]. The six supergroups posited are the Amoebozoa, Opisthokonta, Apusozoa, the Archaeplastida/Plantae, SAR (Stramenopiles, Alveolates, and Rhizaria), and the Excavata (Table 1). These break down into two larger groups, those with a single flagellum (unikonts), which may or may not be retained, and those with two flagella (bikonts) (Table 1). A summary tree of eukaryotic relationships based on multigene analyses as outlined by Parfrey et al. [8] is shown in Figure 1.

The Amoebozoa includes a diversity of predominantly amoeboid members such as the tubulinid amoeba, *Amoeba* spp., *Dictyostelium discoideum* (cellular slime mold), and *Entamoeba* spp., which are secondarily amitochondriate. Opisthokonts include the metazoans, fungi, and the choanoflagellates such as *Monosiga brevicollis* that are the sister to the metazoans [78]. This is the best supported

Table 1: Eukaryotic groups and genera used for analysis of eIF4E family members. The six hypothesized supergroups of eukaryotes after Parfrey et al. [8]. Groups (pink) and genera (yellow) from which eIF4E sequences have been used to examine the relationship of protist eIF4E family members are highlighted.

Unranked	Super group	Group	Examples of Genera
Unikont	Amoebozoa	Archamoeba	*Entameoba*
		Mycetozoa	*Dictyostelium*
		Tubulinea	*Amoeba*, *Acanthamoeba*
Unikont	Opisthokonta	Metazoa	*Drosophila, Homo*
		Fungi	*Saccharomyces*
		Choanoflagellata	*Monosiga*
Unikont	Apusozoa	Ancyromonadida	*Ancyromonas*
		Apusomonadida	*Apusomona*
Bikont	Archaeplastida/Plantae (A/H):		
	Viridiplantae (true plants)	Land plants	*Arabidopsis*
		Charophyta	*Chara*
	Hacrobia	Chlorophyta	*Chlamydomonas,Volvox*
		Haptophytes	*Emiliania*
		Cryptomonads	*Guillardia*
	SAR:		
	Rhizaria	Cercozoa	*Bigelowiellia*
		Foraminifera	*Allogromia*
	Stramenopiles	Haplosporidia	*Bonamia*
	Alveolates	Radiolaria	*Collozoum*
		Phytomyxea	*Polymyxa*
		Heterokonta	*Phytopthora, Thalassiosira, Ectocarpus Phaeodactylum*
		Ciliata	*Tetrahymena, Paramecium*
		Apicomplexa	*Plasmodium, Eimeria, Babesia, Neospora, Theileria, Cryptosporidium, Toxoplasma*
		Dinoflagellata	*Karlodinium, Amoebophyra Amphidinium, Alexandrium*
Bikont	Excavata	Euglenozoa	*Euglena*, *Trypanosoma Leishmania*
		Percolozoa	*Naegleria*
		Metamonada	*Giardia, Trichomonas*

supergroup. The Apusozoa is a supergroup comprising flagellate protozoa, the apusomonads, and ancyromonads. On molecular trees, these two group together, but their relationship to other eukaryotes is uncertain [8]. The supergroup Archaeplastida/Plantae was posited to unite the three lineages with primary plastids: green algae (including land plants), rhodophytes, and glaucophytes with two other lineages, the cryptophytes and haptophytes, both of which have secondary plastids [79]. There is strong support for the SAR supergroup consisting of stramenopiles, alveolates, and plus rhizarians [8]. Within the SAR clade, each of the three members forms distinct lineages [68, 70]. For example, the Rhizaria emerged from molecular data to unite a heterogeneous group of flagellates and amoebae including cercomonads, foraminifera, diverse testate amoebae, and former members of the radiolaria [80] and represents an expansion of the Cercozoa to include foraminifera [81]. The Cercozoa was also recognized from molecular data [82]. Cercozoa and foraminifera appear to share a unique insertion in ubiquitin [83], although there is a paucity of nonmolecular characters uniting the members of this supergroup [8]. Within the alveolates, the Apicomplexa is a large monophyletic group many of which are parasites, including *Plasmodium*, the parasite responsible for malaria. The last supergroup is the Excavata, a supergroup composed predominately of heterotrophic flagellates, and includes many important parasites such as the trypanosomes, *Giardia*, and trichomonads. Within this supergroup, the "euglenozoa," the combination of eugleniids and trypanosomes is a grouping with good support.

6. Unusual Features of Protist eIF4Es

A previous phylogenetic analysis of eIF4E family members from protists indicated that they cannot be grouped with

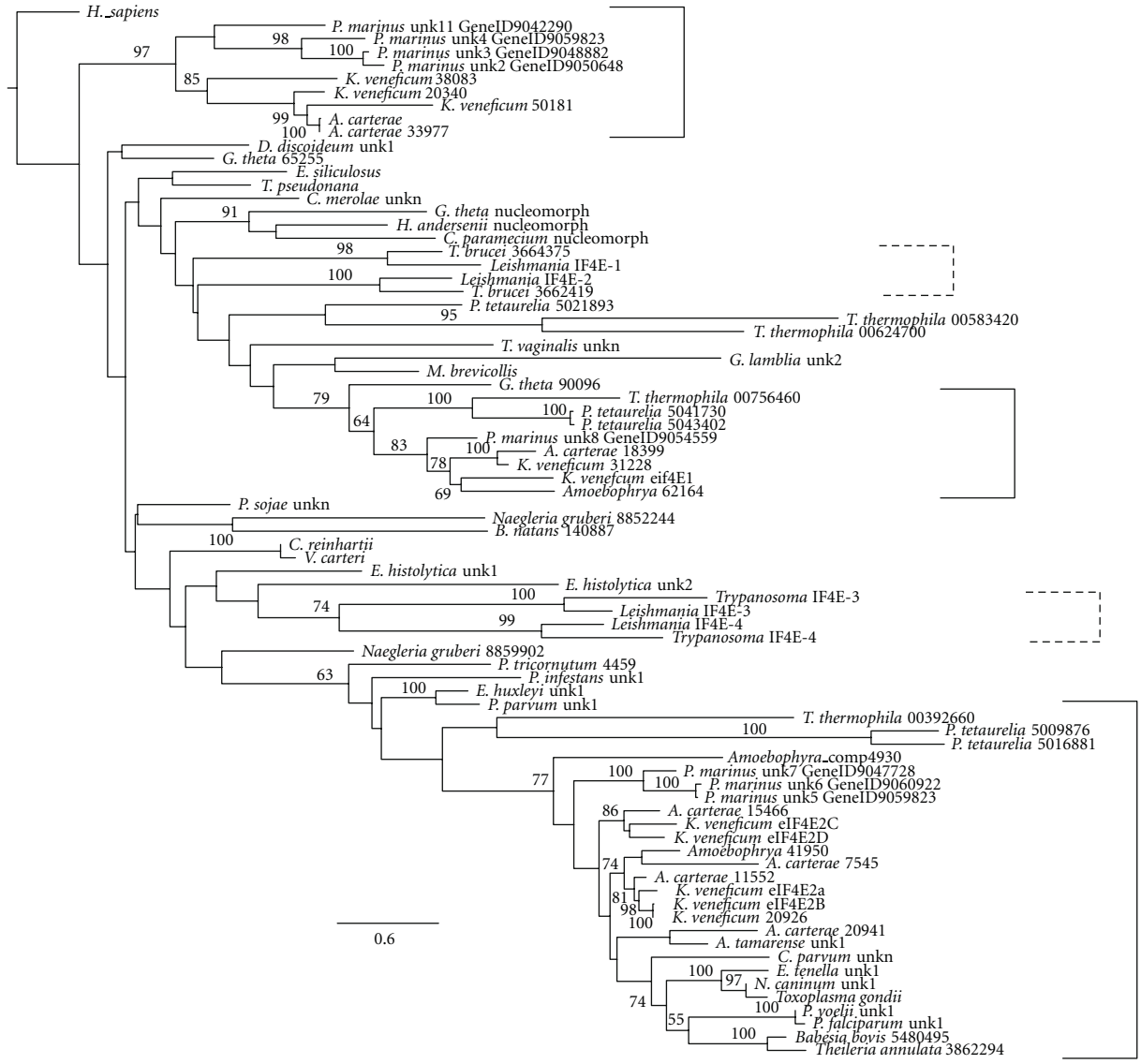


Figure 2: Relationship of selected eIF4E-family members from multiple protist species. Maximum likelihood phylogeny of eIF4E amino acid sequences aligned with T-coffee and trimmed to include only the core region of 453 aligned positions (corresponding to positions 30 to 203 of the human sequence). The tree was constructed using RAxML with the Jones Taylor Thornton gamma distributed model with 100 rapid bootstrap replicates. Bootstrap values above 50% are shown.

the three main classes that describe eIF4E family members from multicellular organisms [6]. At the time of the earlier analysis, very few sequences were available for protists. Many more are now available, though not all in publically available databases. Figure 2 shows a tree describing the overall relationships of selected eIF4E-family members from multiple protists species rooted with *H. sapiens* eIF4E-1. The tree shows maximum likelihood phylogeny of eIF4E amino acid sequences aligned with T-coffee and trimmed to include only the core regions corresponding to amino acids 30 to 203 of the human sequence). The tree was constructed using RAxML with the Jones Taylor Thornton gamma distributed model with 100 rapid bootstrap replicates. Bootstrap values above 50% are shown. Sequences derive predominantly from representatives of SAR mainly heterokonts, ciliates, apicomplexans, dinoflagellates, *Perkinsus*, along with Excavata representatives from Diplomonads (*Giardia*), Euglenozoa (*Trypanosoma*), and Parabasalids (*Trichomonas*).

Three clades stand out and are bracketed with solid lines. All three solid bracketed clades include eIF4Es from dinoflagellates and *Perkinsus* suggesting the possibility of three different classes. The bottom bracket shows a large clade (Clade 1) including eIF4Es from the ciliate, *Tetrahymena thermophila*; *Perkinsus marinus* eIF4E-5, -6, and -7; eIF4E-2a–d sequences from the dinoflagellate, *Karlodinium veneficum*, along with eIF4Es from the dinoflagellates *Amphidinium carterae* and *Amoebophrya*. This clade also includes eIF4Es from the closely related apicomplexans and is the only strong clade with apicomplexans in this tree. Clade 1 also includes "dotted line" clade eIF4E family members from the euglenozoan

Table 2: Summary of protist eIF4E family member characteristics. A selection of Clade 1 and Clade 2 protist eIF4E family members is shown, looking at the residue at positions equivalent to W46, W56, W73, W113 in human eIF4E-1; presence or absence of a Ser residue at the position equivalent to S209 in human eIF4E-1; presence or absence of insertions; the sequence of the sequence of the eIF4G-binding domain. Shading indicates Amoebozoa (Pale yellow); nucleomorph (Gray); haptophyte (Aqua); alveolate, apicomplexan (Yellow); alveolate dinoflagellate/Perkinsus (Pink); excavate (Aquamarine) eIF4Es.

Spp/form	Clade	W46	W56	W73	W113	S209	W46–W56	W73–W102	W113–W130	W130–W166	eIF4GBD
Plant/metazoan/fungi consensus	N/A	W	W	W	W	Y	N	N	N	N	S/TVxxFW
H. sapiens eIF4E-1	N/A	W	W	W	W	Y	N	N	N	N	TVEDFW
A. thaliana	N/A	W	W	W	W	Y	N	N	N	N	TVEDFW
E. histolytica, 180000	1	W	Y	W	W	Y	N	N	N	N	TVENFW
D. discoideum,	2	W	W	W	W	Y	N	N	N	N	SVEDFW
A. castellani eIF4E-1	2	W	W	W	W	Y	N	N	N	U	TVEDFW
G. theta nucleomorph	2	W	W	L	W	N	N	N	N	N	NLEDFL
H. andersonii nucleomorph	2	W	W	W	W	N	N	N	N	N	SIDNFW
C. paramecium nucleomorph	2	W	W	L	W	N	N	N	N	N	DVENFL
E. huxleyi, unk1	1	W	Y	W	W	Y	Ys	Y	N	N	TVEEFW
P. falciparum, unk1	1	W	Y	W	F	Y	N	Y	N	Ys	SVQKFW
N. caninum	1	W	Y	W	F	Y	N	Y	N	Ys	TVQKFW
E. tenella, unk1	1	W	Y	W	F	Y	N	Y	N	Ys	TVQTFW
T. gondii	1	W	Y	W	F	Y	N	Y	N	N	TVQKFW
B. bovis, 548495	1	W	Y	W	F	Y	Ys	N	N	N	SVQSFW
A. tamarense, unc 1	1	W	Y	W	F	Y	Y	Y	N	Y	SVEQFW
K. veneficum 2a	1	W	Y	W	F	Y	N	Y	N	Y	TVQEFW
K. veneficum 2b	1	W	Y	W	F	Y	N	Y	N	Y	TVQEFW
K. veneficum 2c	1	W	Y	W	F	N	N	Y	N	Y	TVQEFW
K. veneficum 2d	1	W	Y	W	F	N	N	Y	N	Y	TVKGFW
P. marinus 5	1	W	W	W	L	N	N	Y	N	Y	TVGEFW
P. marinus 6	1	W	W	W	L	N	N	Y	N	Y	TVGEFW
P. marinus 7	1	W	W	W	L	N	N	Y	N	Y	TVGEFW
L. major, EF4E3	1	Y	F	W	S	Y	N	N	Ys	N	DVESFW
T. brucei, EIF4E3	1	Y	Y	F	T	Y	N	N	Ys	N	DVECFW
N. gruberii, 8859902	1	W	Y	F	W	N	N	N	N	N	DVETFW
G. lamblia eIF4E2	U	F	F	F	K	N	Ys	N	N	N	SLKAFF

excavates, *Leishmania* and *Trypanosoma*, EIF4E3 and 4. The next bracketed clade (Clade 2) includes eIF4E family members from *K. veneficum* (eIF4E-1), *A. carterae* 18399, *P. marinus* eIF4E-8, *Amoebophrya* and the ciliate *T. thermophila* and "dotted line" clade that includes trypanosome sequences *Leishmania* EIF1 and 2. Characteristics of some Clade 1 and Clade 2 eIF4E family members are summarized in Table 2. The top bracketed clade (Clade 3) contains eIF4E family members from *P. marinus*, eIF4E-2, -3, -4, -11, *K. veneficum* eIF4E-1, and *A. carterae* 33977. eIF4Es from ciliates are absent from this top clade, and there is an "orphaned" clade of ciliate sequences. These results suggest gene duplication into three groups prior to divergence of the alveolates with the loss of one copy in *Amoebophrya* and the loss of two copies in apicomplexans. An alternate explanation could be that these copies are not apparent because they are so diverged, or, in the case of *Amoebophrya*, because of poor coverage.

7. eIF4E Family Members in *Giardia lamblia*

Giardia lamblia is an amitochondriate flagellated protozoan parasite that belongs to the diplomonad group (Excavata) that includes both parasitic and free living species [84]. Its genome is compact in structure and content (~11.7 Mb), contains few introns or mitochondrial relics, and has simplified machinery for DNA replication, transcription, RNA processing, and most metabolic pathways [85]. mRNA recruitment in these organisms is unusual in that their transcripts have exceedingly short 5′ untranslated regions

(5′-UTRs), ranging from 0 to 14 nucleotides, and similarly short 3′-UTRs of 10 to 30 nucleotides [86]. Extremely short 5′-UTRs are a highly conserved trait of transcripts from *Trichomonas*, *Entamoeba,* as well as *Giardia*. The precise cap structure in *Giardia* RNAs has not yet been determined, although native *Giardia* mRNAs have blocked 5′-ends and the genome encodes a yeast-like capping apparatus [87]. Furthermore, m^7GpppN-capped mRNA introduced into the cells is expressed well [87, 88]. Eight $m^{2,2,7}$GpppN-capped snRNA species have been identified in *Giardia* [89]. Experimentally, mRNA recruitment occurs efficiently in mRNAs that are capped and in which the first initiation codon is located only 1 nucleotide downstream from the m^7GpppN-cap structure. Recruitment can be decreased when the 5′-UTR between the cap and the initiation codon is lengthened beyond 9 nucleotides [88]. There are two eIF4E family members in *Giardia*, termed eIF4E1 and eIF4E2, which have distinct properties [41]. Of the two, eIF4E2 has been shown to be essential and binds to m^7GTP-Sepharose, suggesting that it functions in protein synthesis. The other, eIF4E1, is not essential and binds only to $m^{2,2,7}$GpppN-Sepharose. eIF4E1 is found concentrated and colocalized with the $m^{2,2,7}$GpppN cap, 16S-like rRNA, and fibrillarin in the nucleolus-like structure in the nucleus [41]. Of the eight conserved tryptophan residues typical of eIF4E Class I sequences, both forms have a Phe residue at the position equivalent to human W56. eIF4E1 has Leu at the position equivalent to human W73, and eIF4E2 has a Phe residue (Table 2). Both forms have poor consensus at the eIF4G binding site with substitutions of W113/Y and W113/I for eIF4E1 and eIF4E2, respectively (numbering as in human eIF4E), eIF4E1 has an insertion between residues 130–166.

8. eIF4E Family Members in Trypanosomatids

Trypanosomatids are a group of kinetoplast protozoa (Excavata/Euglenozoa) distinguished by having only a single flagellum. The haploid genome size in *Leishmania major* is ~36 Mb (haploid). mRNA maturation in trypanosomes differs from the process in most eukaryotes mainly because protein-coding genes are transcribed into polycistronic RNAs in this organism [36, 37, 90]. Transcription of protein coding genes occurs polycistronically, and processing to monocistronic mRNAs occurs through coupled splice leader (SL) *trans*-splicing and polyadenylation (reviewed [36]). The SL *trans*-splicing mechanism was once considered an anomaly of the kinetoplastids, but subsequent identification of *trans*-splicing in dinoflagellates, *Perkinsus*, euglenozoans, and several major invertebrate phyla suggests that this particular form of RNA processing may represent an evolutionarily important aspect of gene expression [36, 37, 91]. There are similarities, particularly in genomic arrangement of SL RNAs, between phyla known to exhibit *trans*-splicing and their mRNAs; however, there is little sequence similarity between the SLs of different organisms. In this RNA-mediated form of *trans*-splicing, a short SL exon is spliced from a capped small nuclear RNA and is transferred to pre-mRNA, thereby becoming the 5′-terminal end and providing an unusual cap structure to mature mRNAs. In *Euglena* (Excavata/Euglenozoa), the SL contribution results in trimethylguanosine, a so-called trimethyl cap, $m^{2,2,7}$GpppG (TMG), in which there are additional methylations to the prototypical monomethyl (m^7GpppN) cap structure found on most eukaryotic mRNAs [92]. In metazoans such as nematodes, where only a percentage of mRNAs are *trans*-spliced, the SL contribution results in a trimethyl cap [93]. In kinetoplastids, all of the mRNAs are *trans*-spliced and the SL contribution results in a highly unique cap structure where additional methylations are apparent. Whereas no more than three modified nucleotides have been described in any metazoan cap structure, the kinetoplastid cap has four consecutive modifiednucleotides (and thus by convention is referred to as a cap-4 structure) [94, 95]. This has been the most highly modified eukaryotic mRNA cap known to date. In trypanosomatids, mRNAs have a common 39-nt long spliced leader sequence at the distal end of the 5′-UTR, which is identical for all mRNAs of a given species. Regulation of gene expression in trypanosomatids is accomplished mainly through posttranscriptional mechanisms such as control of mRNA stability and translation [96–98].

Four eIF4E family members have been characterized from the trypanosomatids *Leishmania major* and *Trypanosoma brucei*, termed EIF4E1, 2, 3, and 4 [99, 100]. All four are expressed in both procyclic and bloodstream forms of the parasites. These four can be broadly classified into two groups (Figure 2). Sequence analysis has identified features that distinguish EIF4E1 and 2 from EIF4E3 and 4 in both *T. brucei* and *L. major*. Similarly, separation of the four eIF4Es into two distinct groups can be made on the basis of localization and function [100]. In *T. brucei,* EIF4E1 and 2 (Group 1, expanded Clade 2) localize both to the nucleus and the cytoplasm and do not seem to be directly involved in translation based on knockdown experiments, although they do perform functions essential for cellular viability [100]. The second group (Group 2, Clade 1) formed by EIF4E3 and 4 is more abundant, is strictly cytoplasmic, is required for translation, and interacts with *T. brucei* eIF4Gs [100].

Group 1 comprises the EIF4E1 and 2 sequences (expanded Clade 2), which are more similar in size to the human and yeast sequences, but show extensions between W102–W113. The function of this extension in Clade 2 eIF4Es in euglenozoans is not known, but the prolines suggest it is solvent exposed and thus could be involved in protein-protein interaction. eIF4E family members from Group 2 (expanded Clade 1), EIF4E3 and 4, share a few unusual features absent from the Group 1 members and distinct from plant, fungi, and metazoan eIF4Es. These include a long N-terminus of more than 150 amino acids which share extensive homology between different orthologues in the EIF4E3 sequences and also contain short segments of limited homology which seem to be conserved between the EIF4E3 and EIF4E4 sequences [100]. Of the eight conserved tryptophan residues typical of eIF4E Class I sequences, most are either conserved in the various trypanosomatid homologues or are replaced by other aromatic residues such as W56Y/F in the Group 2 eIF4Es (human eIF4E numbering) (Table 2) [100]. The only exception is W113, present in

the EIF4E1 and EIF4E2 sequences but which is replaced by nonaromatic hydrophilic residues in EIF4E3 and 4. Other substitutions in the trypanosomatid sequences are D104, next to the universally conserved W102/E103, involved in cap binding [50] which is replaced by a histidine in EIF4E2 and 3; V69/E70, part of the eIF4G-binding domain [101], which is missing in EIF42 and EIF4E4 [100].

EIF4E3, the most abundant *Trypanosoma* and *Leishmania* eIF4E family member, is the only confirmed essential homologue in procyclic and bloodstream T. *brucei*. The similarities observed between *T. brucei* EIF4E3 and 4 at the sequence level, their similar subcellular localization, abundance, and their ability to bind to eIF4G partners are consistent with both performing related-roles in translational initiation. Interestingly, *T. brucei* EIF4E1, 2, and 4, but not *T. brucei* EIF4E3, can efficiently bind the m^7G cap. Nevertheless, when compared with EIF4E4 in *L. major*, it binds less efficiently to the trypanosomatid cap4 [99]. Although *T. brucei* EIF4E2 binds to the m^7G Sepharose in a similar manner to *T. brucei* EIF4E1 and 4, *L. major* EIF4E2 does not bind this cap [102], but rather, preferentially binds the methylated cap4 [99]. This difference, plus the existence of unusual insertions in the *L. major* EIF4E2 between W113–W130 that are missing from the *T. brucei* or *T. cruzi* orthologues, implies a divergence in function unique to the *L. major* protein. The earlier prediction [99] that this insertion might be related to the ability of *L. major* eIF4E to bind to the larger cap-4 seems therefore not to be a compelling argument.

9. eIF4E Family Members in Dinoflagellates

Dinoflagellates are alveolate unicellular protists and a sister group to the parasitic apicomplexans such as *Toxoplasma gondii* and *Plasmodium falciparum*. Dinoflagellates are a diversified group that exhibit a wide diversity in size, form, and lifestyle. They also show a wide spread of genome size, from 1500 to 4700 Mb in *Symbiodinium* sp to 112,000 to 220,050 Mb in *Prorocentrum micans* [103]. Ninety percent of all dinoflagellates are marine plankton with the remaining species being benthic, freshwater, or parasitic.

The free-living species are major primary producers, and several are known to produce harmful algal blooms that result in massive fish kills, human and marine mammal intoxications, as well as economic losses in fisheries and tourism. However, scientific interest with dinoflagellates extends beyond their ecological and economic importance. They possess numerous cellular, molecular, and biochemical traits not observed in "text-book" model organisms. It appears that the organization and regulation of genes in dinoflagellates is different from that of typical eukaryotes. DNA is in permanently condensed chromosomes not packaged in nucleosomes and DNA content ranging from 3 to 250 pg per cell (up to almost 60-fold larger than humans) [104]. Within the dinoflagellate genome, there appears to be a high degree of DNA redundancy, with multiple tandem copies (>20 in many cases) of protein coding genes to give complex gene families [103, 105] that are highly and coordinately expressed. Unlike trypanosomes, in which polycistronic mRNAs contain a series of different genes, the examples studied in dinoflagellates consist of tandemly arrayed copies of the same gene [105].

Recent studies find a predominance of posttranscriptional control of gene expression in dinoflagellate gene expression, including circadian controlled processes such as bioluminescence [106], carbohydrate metabolism [107], and the cell cycle [108], as well as a range of stressors [109–114]. The Van Dolah lab, at the NOAA Center for Coastal Environmental Health and Biomolecular Research, has developed an oligonucleotide microarray from 11,937 unique ESTs from the dinoflagellate *Karenia brevis* [115]. Following validation of the microarray, large-scale transcript profiling studies were performed examining diurnally regulated genes and genes involved in the acute stress response. These studies represent the largest transcript profiling experiments in a dinoflagellate species to date and showed only a small percentage of transcripts changing. None of the anticipated genes, under transcriptional control in other eukaryotes (e.g., cell cycle genes, heat shock, etc.), showed changes in mRNA abundance. Consistent with this, a massively parallel signature sequencing (MPSS) analysis of the transcriptome of the dinoflagellate *Alexandrium tamarense* has shown that of a total of 40,029, only 18, 2, and 12 signatures were found exclusively in the nutrient-replete, nitrogen-depleted, and phosphate-depleted cultures, respectively. The presence of bacteria had the most significant impact on the transcriptome, although the changes represented only ~1.0% of the total number of transcribed genes and a total of only ~1.3% signatures were transcriptionally regulated under any condition [116]. Since the levels of many proteins have been well documented to change in a variety of dinoflagellates, these large-scale studies point to translational regulation as a likely regulatory point in dinoflagellate gene expression. Currently, almost nothing is known about translational initiation or its regulation in these organisms.

Dinoflagellates have mRNAs with unique spliced leaders and cap structures: through analysis of sequences representing all major orders of dinoflagellates, nuclear mRNAs from fifteen species were recently found to be *trans*-spliced with the addition of a 22-nt conserved SL [117, 118]. SL *trans*-splicing has not been identified in a ciliate or apicomplexan to date; however, preliminary analysis using the 22-nt dinoflagellate SL revealed the usage of *trans*-splicing in *Perkinsus marinus* and *P. chesapeaki*, phylogenetic intermediates between apicomplexans and dinoflagellates [118]. Recently, SL *trans*-splicing has been identified in *Amoebophrya* sp, a member of the *Syndinales*, a dinoflagellate parasite of dinoflagellates, which represents a basal root of the dinoflagellates [119]. This suggests the SL machinery was present in an early ancestor of dinoflagellates. It is unclear whether all or only a subset of dinoflagellate genes are subject to SL *trans*-splicing, but, given the diversity of the cDNAs found in the full length libraries, a conservative estimate would be that greater than 90% of mRNAs are *trans*-spliced.

The 22-nt sequence found in dinoflagellate SL-RNA is 5′A(T)CCGTAGCCATTTTGGCTCAAG-3′ [118]. The identity of the cap structure for the SL-RNA needs to be verified, but preliminary analysis indicates only a monomethylated

5′ m^7G is present on mRNAs. Based on the SL-RNA sequence and LC-MS analysis, Place has proposed the following novel cap-4 structure for dinoflagellate mRNAs: $m^7GpppA(U)p^{m2'}Cp^{m2'}CpG$ with modifications to A (U) and G still needing to be established (unpublished results). There is no evidence for a trimethylguanosine or 2′O-methyl adenosine.

Dinoflagellates encode unusual eIF4E-family members. Two distinct eIF4E orthologues, eIF4E-1 and -2 have been partially characterized in *K. veneficum* (Jagus and Place, m/s in preparation) (Figure 3). To facilitate comparison of the sequences, the residues conserved in Class I eIF4Es in multicellular organisms are indicated and numbered as in human eIF4E-1: W43, W46, W56, W73, W102, W113, W130, W166 and S209. eIF4E-2 is represented by four distinct but closely related subtypes (eIF4E-2a–d) (Figure 3). Seven contigs encoding eIF4E-1 and 31 contigs encoding eIF4E-2 (approximately equivalent representation by the a–d subtypes) have been identified indicating the eIF4E-2 group is more highly expressed and may represent the dominant isoforms in the cell. A neighbor-joining tree predicts that the dinoflagellate eIF4E-2 is related to eIF4Es from the kinetoplasts *Leishmania* and *Trypanosoma* with 51% bootstrap support. RT-qPCR analysis for eIF4E transcript abundance is consistent with this assertion (Jagus and Place, m/s in preparation). The *K. veneficum* eIF4E sequences are aligned in Figure 3 with prototypical eIF4E-1 from human. Also included are the sequences for additional, as yet uncharacterized eIF4E family members. Additional sequences were uncovered after this paper was initiated and are shown as kv20926 and kv31228 in Figure 2; however, their sequences are not included in Figure 3. Kv20926 groups with *K. veneficum* Clade 1 eIF4E-2 subtypes and kv31228 with *K. veneficum* Clade 2 eIF4E-1. *K. veneficum* eIF4Es show a clear separation into two subclasses, based on an insert of 11 amino acids between W73 and W102 (numbering equivalent to human eIF4E-1) and distribute between three clades. *K. veneficum* eIF4E-1 and eIF4E 2a–d have a Tyr substitution at the position equivalent to human W56, one of the tryptophans involved in cap binding. This is also observed in eIF4Es from the dinoflagellate *Alexandrium tamarense*, but not from *Amphidinium carterae*. In addition, eIF4E-1 has glutamine instead of D/E in the eIFG/4E-BP-binding domain. The eIF4E-2 family members contain extended amino acid stretches between the structural units of the core, between residues equivalent to human W73 to W102, and W130 to W166. In addition, eIF4Es from several alveolate species have a Trp to Phe substitution at W113 [6], a characteristic shared by *K. veneficum* eIF4E-1. It is of interest that the different subtypes of *K. veneficum* eIF4E-2s show marked heterogeneity between W102–W113. The conserved phosphorylation site of eIF4E is only observed in eIF4E-2a and -2b of *K. veneficum*. eIF4E-2a and -2b share the TKS motif at the putative phosphorylation site in which the Lys residue is a sumoylation site in human eIF4E [120, 121]. The sumoylation site at the equivalent of human Lys35 is shared by eIF4E-2b, -2b, -2c, and -2d. eIF4E-1 contains the sumoylation site equivalent to human Lys210. eIF4E-2a, but not eIF4E-1, binds to m^7GTP-Sepharose *in vitro*, although neither interact with TMG. It is not known whether either form interacts with the unique cap-4 of dinoflagellates (Jagus/Place, m/s in preparation). These results are consistent with eIF4E-2a being a functional initiation factor, but not definitive. The *K. veneficum* eIF4E-2s fall into Clade 1 raising the possibility that other eIF4Es of Clade 1 bind to m^7GTP. The eIF4E-1s fall into Clade 2. Unlike *K. veneficum* eIF4E-1, some of the extended Clade 2 members like the *L. major* and *T. bruceii* eIF4E1 and 2 are known to bind m^7GTP but appear not to participate in protein synthesis [100], making it hard to predict function of the *K. veneficum* eIF4E-1s. Three of the *K. veneficum* eIF4Es fall into Clade 3. As with the *K. veneficum* Clade 2 representatives, these do not have the insert between W73 and W102.

10. eIF4E Family Members in *Perkinsus marinus*

Perkinsus marinus is an alveolate with a genome of 86 Mb and is closely related to the dinoflagellates [122]. Like the dinoflagellates, it also exhibits *trans*-splicing. Five different SLs of 21-22 nucleotides (nt) in length have been reported from *P. marinus* [123–125]. Variability at positions 1 and 2 between the different SLs suggests variability of cap structures. Overall these data suggest a complex gene regulatory system both at the level of mRNA generation and of translational control consistent with its complex life style. The *P. marinus* genome encodes eight eIF4E family members along with two very large (>600 amino acid) forms that contain only some of the typical eIF4E signatures. *P. marinus* eIF4E-5, -6, and -7 form a group that aligns most closely with the *K. veneficum* eIF4E-2s in Clade 1, suggesting they will bind m^7GTP caps (Figures 1 and 4 and Table 2). These share the insertions between W73 to W102 and W113 to W133. This group also has TVGEFW at the eIF4G binding domain. In addition, they each have a Trp to Leu substitution at W113. The *L. major* and *T. bruceii* also show a consistent substitution at this position, but to a hydrophilic amino acid. *P. marinus* eIF4E-2, -3, and -4 also form a group in Clade 3 with eIF4Es with two of the *K. veneficum* eIF4Es (Figure 2, Table 2). *P. marinus* eIF4E-8 groups with *K. veneficum* eIF4E-1.

11. Cryptomonads, *Guillardia theta*, and Nucleomorphs

Cryptomonads (Chromalveolata/Cryptophyta) are chimeras of two different eukaryotic cells; a flagellate host and a photosynthetic endosymbiont. These organisms are thought to have arisen by secondary symbiogenesis shortly after the origin of the common ancestor of green plants, red, and glaucophyte algae [126–128]. In the cryptomonad *Guillardia theta*, the flagellate host acquired a chloroplast by engulfing and retaining a red alga. In doing so, the host was able to convert from obligate heterotrophy to an autotrophic way of life [129–131]. In addition to the red algal chloroplast, cryptomonads have retained a vestigial red algal nuclear

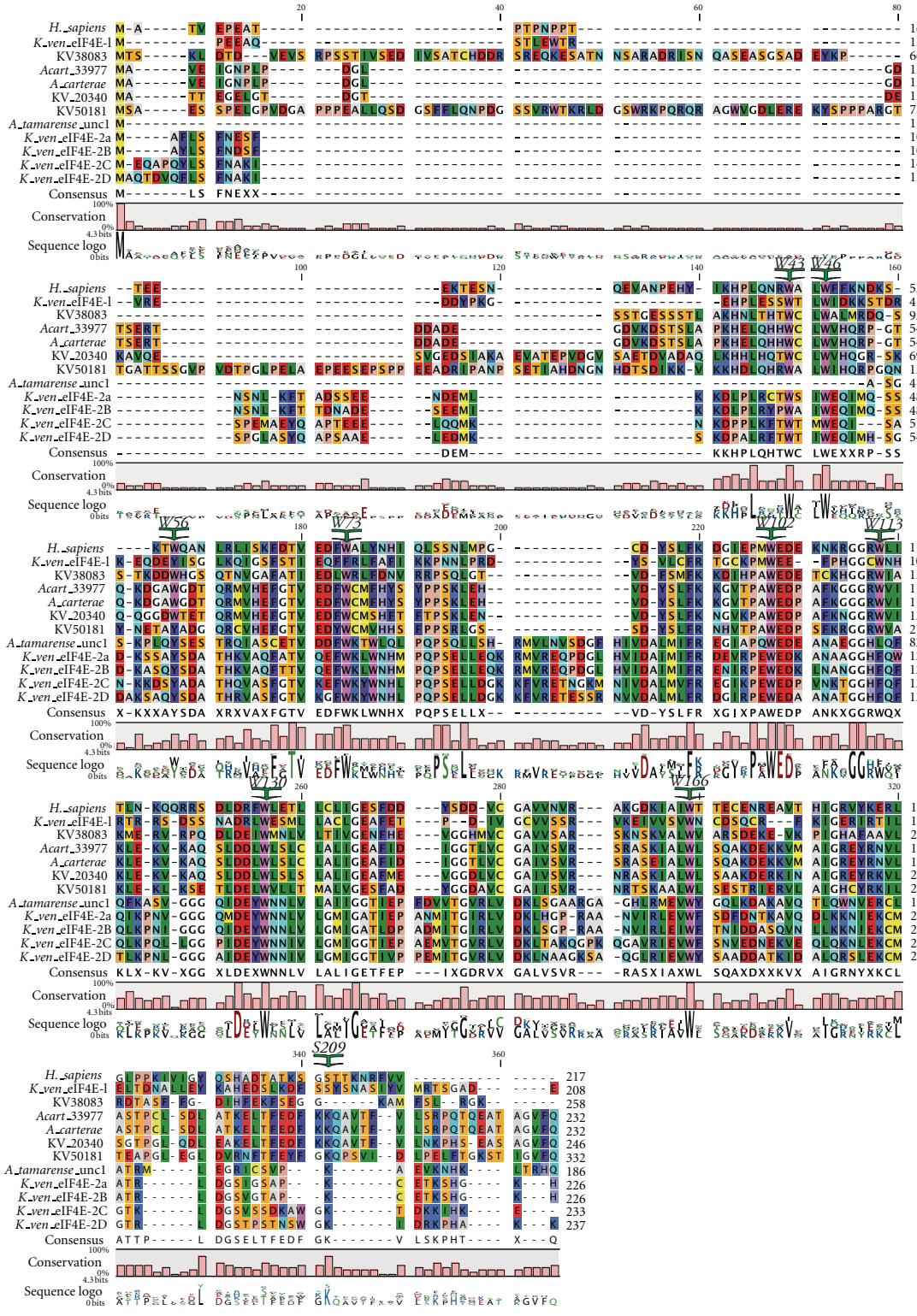


FIGURE 3: Comparison of the sequences of selected eIF4E-family members from *Karlodinium veneficum* and other dinoflagellates. Alignment of the amino acid sequences of selected established eIF4E-family members from *K. veneficum* and other dinoflagellates. Amino acid sequences were aligned with T-coffee using the BLOSUM62MT scoring matrix in CLC Main Workbench. To facilitate comparison of the sequences, the residues conserved in Class I eIF4Es in multicellular organisms are indicated and numbered as in human eIF4E-1: W43, W46, W56, W73, W102, W113, W130, W166, and S209.

FIGURE 4: Comparison of the sequences of selected eIF4E-family members from *Perkinsus marinus* with related species from *K. veneficum* and trypanosome species. Alignment of the amino acid sequences of selected established eIF4E-family members from *P. marinus* with related species from *K. veneficum*. Amino acid sequences were aligned with T-coffee using the BLOSUM62MT scoring matrix in CLC Main Workbench. To facilitate comparison of the sequences, the residues conserved in Class I eIF4Es in multicellular organisms are indicated and numbered as in human eIF4E-1: W43, W46, W56, W73, W102, W113, W130, W166, and S209.

genome as a minute nucleomorph with three chromosomes [132–134]. The nucleomorph resides in a cell compartment, the periplastid space, that also contains the chloroplast. The cellular organization of *Guillardia theta* is shown in Figure 5.

In the cartoon, former chloroplast genes now inserted in nucleomorph or nuclear chromosomes are indicated in green, and former red algal genes now in the host nucleus are indicated in red. The nucleomorph genome has been sequenced and shown to be 551 kbp with a gene density of 1 gene per 977 bp, encoding 464 putative protein coding genes [133]. This compact genome has infrequent overlapping genes, and short inverted repeats containing rRNA cistrons at its chromosome ends [132, 133, 135]. There is almost a total absence of spliceosomal introns which has facilitated gene annotation. Marked evolutionary compaction [126–128] has eliminated almost all the nucleomorph genes for metabolic functions, but left a few hundred housekeeping genes, and 30 genes encoding chloroplast-located proteins [133]. The housekeeping genes are limited to nuclear maintenance and transport, translation, protein degradation and folding, and microtubule/centrosome functions [133, 135]. More than 20% of the housekeeping genes encode components of the translational machinery. The nucleomorph and its periplastid space can be viewed as providing a minimum eukaryotic expression system for a small number of nucleomorph-encoded chloroplast proteins. The endosymbiont has been reduced to an organelle, equivalent to a "complex plastid." The relict, enslaved red alga is referred to here as the endosymbiont for convenience, although strictly speaking it should be considered an organelle.

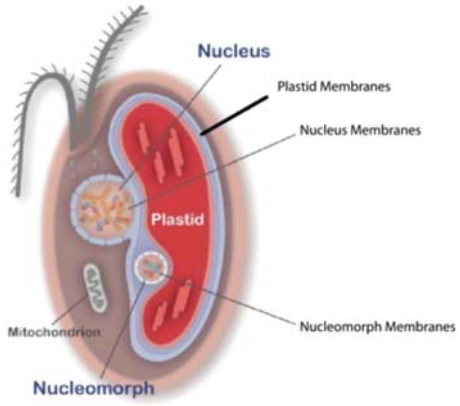


Figure 5: Cellular organization of *Guillardia theta*. Former chloroplast genes now inserted in nucleomorph or nuclear chromosomes are indicated in green, and former red algal genes now in the host nucleus are indicated in red. The four endosymbiont membranes are clearly represented.

12. The Translation Machinery of the *Guillardia theta* Nucleomorph

The endosymbiont encodes its own rRNA and 65 ribosomal proteins [133]. Functioning endosymbiont ribosomes have been demonstrated in the endosymbiont cytoplasm [136]. The endosymbiont has its own mRNAs with 5′-caps and poly(A) tails, elongation, and release factors, but only a subset of translational initiation factors. The nucleomorph encodes eIF1, eIF1A, eIF4A, eIF2 (all subunits, although the alpha subunit is truncated), eIF4E (truncated), eIF5B, eIF6, and poly(A) binding protein. It does not appear to encode any of the subunits of eIF2B, the factor that promotes guanine nucleotide exchange on eIF2. Furthermore, several initiation factors thought to be essential for eukaryotic initiation have not been identified; the nucleomorph does not encode eIF4B, eIF5, or the scaffold proteins eIF3 (any subunit) or eIF4G. All of these initiation factors have been shown to be essential in yeast (reviewed [137]). The nucleomorph is also without the eIF4E regulatory proteins, the 4E-BPs. Since the genome of the *G. theta* nucleomorph has been so severely compacted, it is hypothesized that the genes encoding complex cellular functions, such as protein synthesis, are limited to the minimal set needed to accomplish the function. Beyond the reduction in the number of initiation factors, several of the translational initiation factors encoded are truncated compared to their counterparts in nonprotist eukaryotes. This system can be considered to represent a natural experiment in deletion analysis and may tell us much about structure/function relationships in initiation factors, in addition to deepening our knowledge of this branch of the eukaryotic Tree of Life.

The factors eIF1, eIF1A, eIF2, eIF2B, eIF3, eIF4A, eIF4E, eIF4G, and eIF5 are all essential in yeast. eIF5B is not essential, although its deletion produces a severe slow growth phenotype [138]. The possibility that the lack of eIF2B, eIF3, eIF4G, and eIF5 in the *G. theta* endosymbiont reflects a primitive condition is unlikely since the deeply rooted, free-living red alga, *Cyanidioschyzon merolae*, encodes eIF4G, eIF5, and all the subunits of eIF2B and eIF3 [139]. *C. merolae* is considered to have the smallest genome of any free-living photosynthetic organism and molecular analyses support the primitiveness of this alga [140]. However, like the *Guillardia theta* nucleomorph, *C. merolae* does not appear to encode eIF4B, suggesting that eIF4B is a later evolutionary development [139]. Consistent with this, eIF4B is not essential in yeast, although its disruption results in a slow growth and cold-sensitive phenotype [141]. The endosymbiont has either evolved a minimal system of initiation through compaction of the genome, has made mechanistic adjustments to overcome factor deficiencies, or uses host factors. Use of host factors would require transport across the outer two membranes into the periplastidial compartment PPC and across all four membranes into the stroma [142].

The predicted eIF4E sequence of the *G. theta* nucleomorph is compacted, lacking extended amino-terminal and carboxy-terminal domains relative to the core of prototypical eIF4E (Figure 6) [6]. Although comparable forms from yeast, produced from deletion mutants, are still able to support life, they show considerably slower growth rates [143, 144]. This is likely to reflect a role of the N-terminal domain in enhancing stability. Scrutiny of the alignment also shows that the nucleomorph eIF4E has Leu at amino acid positions equivalent to V69 and W73 in human eIF4E-1. In human eIF4E-1, it is known that mutation to give a nonaromatic amino acid at position W73 disrupts the interaction with the adaptor protein, eIF4G, as does mutation of V69 to G [56, 57]. It is therefore unclear whether the nucleomorph eIF4E has the capacity to bind to eIF4G or indeed whether it needs to. It is possible that the nucleomorph eIF4E interacts with eIF4G imported from the host cytoplasm, although the sequence of the eIF4G-binding domain makes this unlikely. Alternatively, mRNA recruitment via an alternate interaction may be occurring. Interestingly, eIF4E sequences are available from additional nucleomorphs, those of another cryptophyte *Cryptomonas paramecium* and the heterokont *Haplogloia andersonii*. Both of these are truncated at the N-terminus, and both show substitutions in essential amino acids in the eIF4G binding domain.

13. Entamoeba and Mimivirus

Mimivirus is a double-stranded DNA virus isolated from amoebae [145]. It was first isolated from the water of a cooling tower in Bradford, England, during a study following a pneumonia outbreak in 1992 [146, 147]. Its name is derived from "mimicking microbe" because of the bacterium-like appearance of the particle and its Gram$^+$ staining. It has a cycle of viral transmission and replication that is typical of many dsDNA viruses. The study of mimivirus grown in *Acanthamoeba polyphaga* reveals a mature particle with the characteristic morphology of an icosahedral capsid with a diameter of at least 400 nm. At the beginning of the life cycle, the virus enters the amoeba and the viral genome is released. After expression of viral proteins and replication of the genome, the virus DNA is packaged into capsids and viral

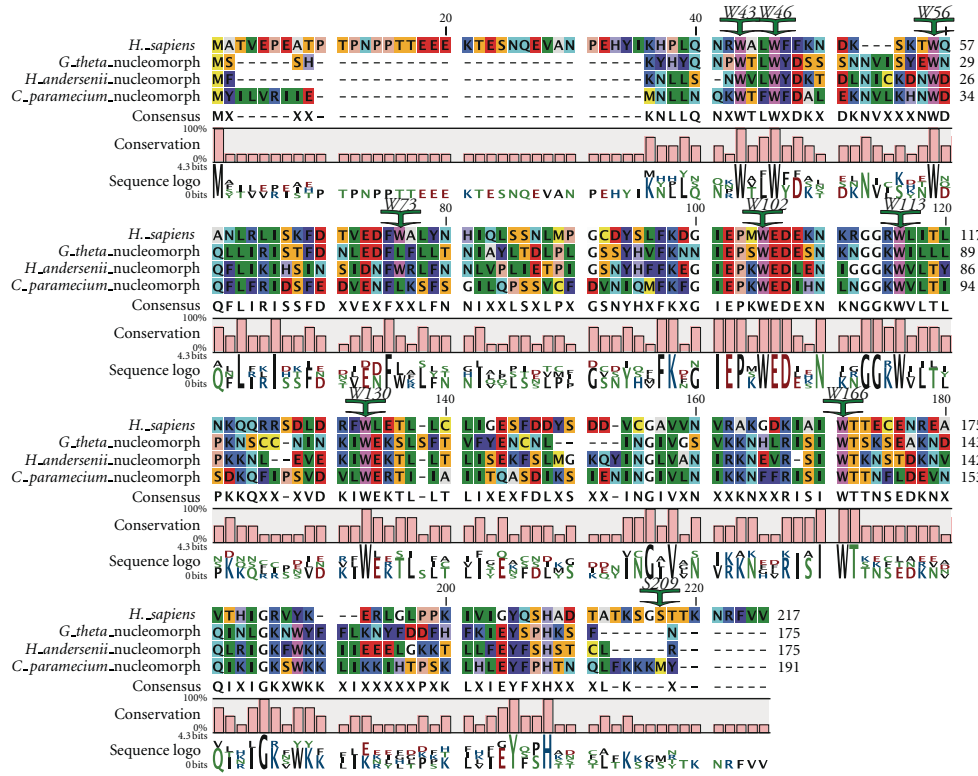


Figure 6: Comparison of the sequences of selected nucleomorph eIF4Es. Alignment of the amino acid sequences of the eIF4E from the nucleomorphs of *Guillardia theta*, *Haplogloia andersonii,* and *Cryptomonas paramecium*. Amino acid sequences were aligned with T-coffee using the BLOSUM62MT scoring matrix in CLC Main Workbench. To facilitate comparison of the sequences, the residues conserved in Class I eIF4Es in multicellular organisms are indicated and numbered as in human eIF4E-1: W43, W46, W56, W73, W102, W113, W130, W166, and S209.

particles are released from the amoeba [148]. Mimivirus has the largest known viral genome, 1.18 megabase pairs, and predicted to contain 1,262 genes, a very complex life cycle at the molecular level [146]. It encodes an unprecedented number of components of the transcriptional, translational, and replication machinery, many of which have not previously been described in viruses.

Although the mimivirus genome has more components resembling cellular genes than any other virus, it is still dependent on its host cell for the synthesis of proteins. Currently, the strategies by which mimivirus appropriates the host translation machinery have not been uncovered. Mimivirus exhibits many features that distinguish it from other nucleocytoplasmic large DNA viruses (NCLDVs). The most unexpected is the presence of numerous genes encoding central protein-translation components, encoding 10 proteins central to the translation apparatus: four aminoacy tRNA synthetases, eIF4E, ORF L496, eIF1A, eIF4A, eEF-1, and peptide chain release factor eRF1 [149, 150]. In addition, mimivirus encodes its own mRNA capping enzyme, and its own RNA cap guanine-N2 methyltransferase [151, 152]. Interestingly, mimivirus does not encode the mimic of the α-subunit of eIF2, found in many NCLDVs, that functions as a substrate to protect endogenous eIF2 from phosphorylation by an infection-activated kinase PKR. Finding these components of the translation apparatus in mimivirus calls into question the prevailing view that viruses rely entirely on the host translation machinery for protein synthesis [153]. Although the molecular mechanisms of its replicative cycle are yet to be uncovered, the detailed genome analysis has provided useful information on what viral genes may be involved in DNA replication and DNA repair, transcription, and protein folding, virion morphogenesis, and intracellular transport and suggests a complex life cycle.

The atypical eIF4E-family member of mimivirus is shown in Figure 7 aligned with the amino acid sequences of eIF4E-family members from *Acanthamoeba*. Mimivirus eIF4E has F49, W109, and E110, in positions equivalent to W56, W102, and E103 of human eIF4E-1, predicting that it should function in cap binding. However, mimivirus eIF4E, like the many protist eIF4E-family members, has extended stretches of amino acids between structural units of the core tryptophans. The positions of these stretches in mimivirus eIF4E resemble the extensions found in eIF4E-family members from Alveolata and Stramenopiles. However, the stretch of amino acids between residues equivalent to W102 and W166 of mouse/human eIF4E-1 are considerably longer in mimivirus eIF4E than those found in *P. falciparum* or other known stramenopile/alveolate eIF4E family members. Mimivirus eIF4E also differs from other eIF4E-family members in that it lacks a Trp residue equivalent to W73 of mouse

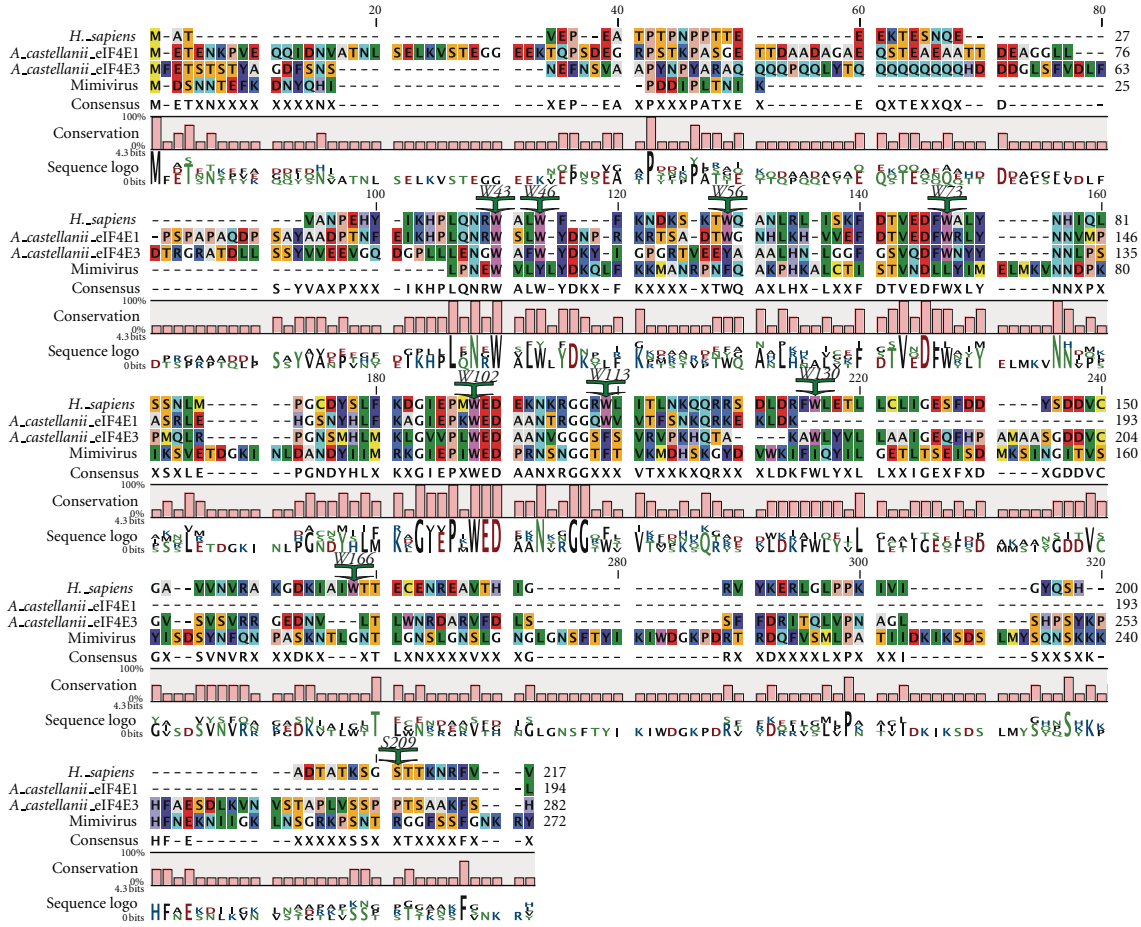


Figure 7: Comparison of the sequence of mimivirus eIF4E with those of an *Acanthamoeba* species. Alignment of the amino acid sequences of mimivirus eIF4E with eIF4Es from its host *Acanthamoeba castellani*. Amino acid sequences were aligned with T-coffee using the BLOSUM62MT scoring matrix in CLC Main Workbench. *A. castellani* sequences were derived from the Protist EST Program (PEP) in advance of scientific publication and acceptance by GenBank at http://amoebidia.bcm.umontreal.ca/public/pepdb/agrm.php. This site is no longer publically available. To facilitate comparison of the sequences, the residues conserved in Class I eIF4Es in multicellular organisms are indicated and numbered as in human eIF4E-1: W43, W46, W56, W73, W102, W113, W130, W166, and S209.

eIF4E-1 suggesting that the protein may not interact with eIF4G or 4E-BPs.

The host *A. castellanii* expresses at least five eIF4E-family members. None of the *A. castellanii* eIF4E family members shows extended stretches of amino acids in positions similar to those found in mimivirus eIF4E. Furthermore, *A. castellanii* eIF4E-family members possess conserved residues equivalent to V69 and W73 of human eIF4E-1 important for interaction with eIF4G and 4E-BPs, unlike mimivirus eIF4E. As a consequence of these differences in significant residues, it seems unlikely that mimivirus eIF4E has been acquired from the *Acanthamoeba* host. The sequence of mimivirus eIF4E predicts that it is likely to bind to 5′-cap structures but may not interact with eIF4G, suggesting that it could function as an inhibitor of cap-dependent translation. However, the mimivirus genome encodes genes for mRNA capping enzymes [151, 152], as do related NCLD viruses, suggesting that mimivirus mRNAs are capped and that the virus requires cap-dependent translation of its mRNAs. Since mimivirus can use both *A. polyphaga* and human as hosts, it will be of use to consider the role of its eIF4E in the context of mRNA recruitment in both environments.

14. Overview of Protist eIF4Es

Like multicellular eukaryotes, many protists encode multiple eIF4E family members. However, these do not fall into the eIF4E classes found in plants/metazoans/fungi. Of the eight conserved tryptophan residues typical of eIF4E Class I sequences, most are either conserved in protist eIF4E family members or are replaced by other aromatic residues. In many bikont protists, extensions are found between the conserved aromatic amino acids which vary with clade and phylogenetic grouping. Figure 2 shows the relationships of the protist eIF4Es and suggests that they fall into three clades. eIF4Es from dinoflagellates/*Perkinsus* and heterokonts can be found in all three clades. eIF4Es from ciliates and the parasitic dinoflagellate *Amoebophrya* are present in only two clades. Unfortunately, at the current time, there are many more eIF4E sequences available for alveolates and excavates

than for other protist groups, particularly the opisthokonts and amoebozoa. Furthermore, there is insufficient data on the functional characteristics of the eIF4Es in each of these clades to allow for any confident classification at this stage. Nevertheless, it is known that the *Leishmania* and *Trypanosoma* eIF4Es, EIF4E3 and 4 function as initiation factors and that the dinoflagellate eIF4E-2s from *K. veneficum* bind cap structures suggesting that this clade contains eIF4E family members that function as initiation factors. Table 2 shows the characteristics of some of the members from "Clades 1" and "2." As genome sequencing projects are completed, it is expected that the number of protist eIF4E family members available for scrutiny will increase dramatically in the near future. A wider representation of taxa will allow a more complete understanding of the relationships between these eIF4Es, as will a much needed expansion of functional studies particularly in the non-parasitic representatives.

Acknowledgments

This was supported by MCB no. 0626678, to A. R. Place and MCB no. 0134013 to R. Jagus. The authors are grateful to Drs. Terry Gaasterland, John Gill, Senjie Lin, Yu-Hui Rogers, and Huang Zhang, co-PIs with A. R. Place of the NSF Microbial Genome Sequencing Program Grant no. EF-0626678, "Dinoflagellate full-length cDNA sequencing," which made possible searching of the *K. veneficum* cDNA database. The authors would like to thank Charles Delwiche (supported by NSF DEB 0629624) for allowing them to use unpublished sequences for *Amoebophyra*. In addition, thanks are extended to Dr. Joseph Pitula, University of Maryland Eastern Shore, for useful discussions on the *P. marinus* eIF4Es and splice leaders and to Jorge Rodriguez, University of Maryland Eastern Shore, for initial alignments of the *P. marinus* eIF4Es. This paper represents contribution no. #12-234 from IMET and #4660 from UMCES.

References

[1] A. Ben-Shem, N. GarreaudeLoubresse, S. Melnikov, L. Jenner, G. Yusupova, and M. Yusupov, "The structure of the eukaryotic ribosome at 3.0 Å resolution," *Science*, vol. 334, pp. 1524–1529, 2011.

[2] A. Korostelev, D. N. Ermolenko, and H. F. Noller, "Structural dynamics of the ribosome," *Current Opinion in Chemical Biology*, vol. 12, no. 6, pp. 674–683, 2008.

[3] M. C. Rivera and J. A. Lake, "Evidence that eukaryotes and eocyte prokaryotes are immediate relatives," *Science*, vol. 257, no. 5066, pp. 74–76, 1992.

[4] M. C. Rivera, "Genomic analyses and the origin of the eukaryotes," *Chemistry and Biodiversity*, vol. 4, no. 11, pp. 2631–2638, 2007.

[5] G. Hernández and P. Vazquez-Pianzola, "Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families," *Mechanisms of Development*, vol. 122, no. 7-8, pp. 865–876, 2005.

[6] B. Joshi, K. Lee, D. L. Maeder, and R. Jagus, "Phylogenetic analysis of eIF4E-family members," *BMC Evolutionary Biology*, vol. 5, article 48, 2005.

[7] C. L. McGrath and L. A. Katz, "Genome diversity in microbial eukaryotes," *Trends in Ecology and Evolution*, vol. 19, no. 1, pp. 32–38, 2004.

[8] L. W. Parfrey, J. Grant, Y. I. Tekle et al., "Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life," *Systematic Biology*, vol. 59, no. 5, pp. 518–533, 2010.

[9] D. Moreira and P. López-García, "The molecular ecology of microbial eukaryotes unveils a hidden world," *Trends in Microbiology*, vol. 10, no. 1, pp. 31–38, 2002.

[10] J. D. Bangs, P. F. Crain, T. Hashizume, J. A. McCloskey, and J. C. Boothroyd, "Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides," *The Journal of Biological Chemistry*, vol. 267, no. 14, pp. 9805–9815, 1992.

[11] W. Marande, P. López-García, and D. Moreira, "Eukaryotic diversity and phylogeny using small- and large-subunit ribosomal RNA genes from environmental samples," *Environmental Microbiology*, vol. 11, no. 12, pp. 3179–3188, 2009.

[12] J. D. Keene, "Minireview: global regulation and dynamics of ribonucleic acid," *Endocrinology*, vol. 151, no. 4, pp. 1391–1397, 2010.

[13] X. Qin, S. Ahn, T. P. Speed, and G. M. Rubin, "Global analyses of mRNA translational control during early Drosophila embryogenesis," *Genome Biology*, vol. 8, no. 4, article R63, 2007.

[14] J. B. Dacks, A. A. Peden, and M. C. Field, "Evolution of specificity in the eukaryotic endomembrane system," *International Journal of Biochemistry and Cell Biology*, vol. 41, no. 2, pp. 330–340, 2009.

[15] M. C. Field and J. B. Dacks, "First and last ancestors: reconstructing evolution of the endomembrane system with ESCRTs, vesicle coat proteins, and nuclear pore complexes," *Current Opinion in Cell Biology*, vol. 21, no. 1, pp. 4–13, 2009.

[16] B. L. Semler and M. L. Waterman, "IRES-mediated pathways to polysomes: nuclear versus cytoplasmic routes," *Trends in Microbiology*, vol. 16, no. 1, pp. 1–5, 2008.

[17] A. Schröder-Lorenz and L. Rensing, "Circadian changes in protein-synthesis rate and protein phosphorylation in cell-free extracts of *Gonyaulax polyedra*," *Planta*, vol. 170, no. 1, pp. 7–13, 1987.

[18] E. V. Koonin, "The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate?" *Biology Direct*, vol. 1, article 22, 2006.

[19] E. V. Koonin, "The origin and early evolution of eukaryotes in the light of phylogenomics," *Genome Biology*, vol. 11, no. 5, article 209, 2010.

[20] W. Martin and E. V. Koonin, "Introns and the origin of nucleus-cytosol compartmentalization," *Nature*, vol. 440, no. 7080, pp. 41–45, 2006.

[21] T. Cavalier-Smith, "Intron phylogeny: a new hypothesis," *Trends in Genetics*, vol. 7, no. 5, pp. 145–148, 1991.

[22] B. Cousineau, S. Lawrence, D. Smith, and M. Belfort, "Retrotransposition of a bacterial group II intron," *Nature*, vol. 404, no. 6781, pp. 1018–1021, 2000.

[23] A. M. Lambowitz and S. Zimmerly, "Mobile group II introns," *Annual Review of Genetics*, vol. 38, pp. 1–35, 2004.

[24] B. J. Mans, V. Anantharaman, L. Aravind, and E. V. Koonin, "Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex," *Cell Cycle*, vol. 3, no. 12, pp. 1612–1637, 2004.

[25] L. Aravind and E. V. Koonin, "Eukaryote-specific domains in translation initiation factors: implications for translation regulation and evolution of the translation system," *Genome Research*, vol. 10, no. 8, pp. 1172–1184, 2000.

[26] D. Benelli and P. Londei, "Translation initiation in Archaea: conserved and domain-specific features," *Biochemical Society Transactions*, vol. 39, no. 1, pp. 89–93, 2011.

[27] P. Londei, "Evolution of translational initiation: new insights from the archaea," *FEMS Microbiology Reviews*, vol. 29, no. 2, pp. 185–200, 2005.

[28] G. Hernández, "Was the initiation of translation in early eukaryotes IRES-driven?" *Trends in Biochemical Sciences*, vol. 33, no. 2, pp. 58–64, 2008.

[29] K. Sawicka, M. Bushell, K. A. Spriggs, and A. E. Willis, "Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein," *Biochemical Society Transactions*, vol. 36, no. 4, pp. 641–647, 2008.

[30] S. M. Lewis and M. Holcik, "For IRES *trans*-acting factors, it is all about location," *Oncogene*, vol. 27, no. 8, pp. 1033–1035, 2008.

[31] B. L. Semler and M. L. Waterman, "IRES-mediated pathways to polysomes: nuclear versus cytoplasmic routes," *Trends in Microbiology*, vol. 16, no. 1, pp. 1–5, 2008.

[32] A. A. Komar and M. Hatzoglou, "Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states," *Cell Cycle*, vol. 10, no. 2, pp. 229–240, 2011.

[33] O. Elroy-Stein and W. C. Merrick, "Translation initiation via cellular internal ribosome entry sites," in *Translational Controlin Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., pp. 155–172, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[34] J. A. Doudna and P. Sarnow, "Translation initiation by viral internal ribosome entry sites," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., pp. 129–154, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[35] A. Pacheco and E. Martinez-Salas, "Insights into the biology of IRES elements through riboproteomic approaches," *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 458927, 12 pages, 2010.

[36] E. L. Lasda and T. Blumenthal, "*Trans*-splicing," *Wiley Interdisciplinary Reviews*, vol. 2, pp. 417–34, 2011.

[37] K. E. M. Hastings, "SL *trans*-splicing: easy come or easy go?" *Trends in Genetics*, vol. 21, no. 4, pp. 240–247, 2005.

[38] R. Derelle, T. Momose, M. Manuel, C. Da Silva, P. Wincker, and E. Houliston, "Convergent origins and rapid evolution of spliced leader *trans*-splicing in Metazoa: insights from the Ctenophora and Hydrozoa," *RNA*, vol. 16, no. 4, pp. 696–707, 2010.

[39] J. P. Bruzik and T. Maniatis, "Spliced leader RNAs from lower eukaryotes are *trans*-spliced in mammalian cells," *Nature*, vol. 360, no. 6405, pp. 692–695, 1992.

[40] G. Hernández, "On the origin of the cap-dependent initiation of translation in eukaryotes," *Trends in Biochemical Sciences*, vol. 34, no. 4, pp. 166–175, 2009.

[41] L. Li and C. C. Wang, "Identification in the ancient protist *Giardia lamblia* of two eukaryotic translation initiation factor 4E homologues with distinctive functions," *Eukaryotic Cell*, vol. 4, no. 5, pp. 948–959, 2005.

[42] M. A. Andrei, D. Ingelfinger, R. Heintzmann, T. Achsel, R. Rivera-Pomar, and R. Lührmann, "A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies," *RNA*, vol. 11, no. 5, pp. 717–727, 2005.

[43] N. P. Hoyle, L. M. Castelli, S. G. Campbell, L. E. A. Holmes, and M. P. Ashe, "Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies," *Journal of Cell Biology*, vol. 179, no. 1, pp. 65–74, 2007.

[44] L. Rong, M. Livingstone, R. Sukarieh et al., "Control of eIF4E cellular localization by eIF4E-binding proteins, 4E-BPs," *RNA*, vol. 14, no. 7, pp. 1318–1327, 2008.

[45] I. G. Goodfellow and L. O. Roberts, "Eukaryotic initiation factor 4E," *International Journal of Biochemistry and Cell Biology*, vol. 40, no. 12, pp. 2675–2680, 2008.

[46] A. C. Gingras, B. Raught, and N. Sonenberg, "eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation," *Annual Review of Biochemistry*, vol. 68, pp. 913–963, 1999.

[47] T. von der Haar, J. D. Gross, G. Wagner, and J. E. G. McCarthy, "The mRNA cap-binding protein eIF4E in post-transcriptional gene expression," *Nature Structural and Molecular Biology*, vol. 11, no. 6, pp. 503–511, 2004.

[48] N. Sonenberg, "eIF4E, the mRNA cap-binding protein: from basic discovery to translational research," *Biochemistry and Cell Biology*, vol. 86, no. 2, pp. 178–183, 2008.

[49] I. Topisirovic, Y. V. Svitkin, N. Sonenberg, and A. J. Shatkin, "Cap and cap-binding proteins in the control of gene expression," *Wiley Interdisciplinary Reviews*, pp. 277–98, 2011.

[50] J. Marcotrigiano, A. C. Gingras, N. Sonenberg, and S. K. Burley, "Co-crystal structure of the messenger RNA5′ cap-binding protein (elF4E) bound to 7-methyl-GDP," *Cell*, vol. 89, no. 6, pp. 951–961, 1997.

[51] H. Matsuo et al., "Structure of translation factor eIF4E bound to m^7GDP and interaction with 4E-binding protein," *Natural Structural Biology*, vol. 4, pp. 717–24.

[52] A. Niedzwiecka, J. Marcotrigiano, J. Stepinski et al., "Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5′ cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins," *Journal of Molecular Biology*, vol. 319, no. 3, pp. 615–635, 2002.

[53] G. Hernández, M. Altmann, J. M. Sierra et al., "Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in *Drosophila*," *Mechanisms of Development*, vol. 122, no. 4, pp. 529–543.

[54] B. Joshi, A. Cameron, and R. Jagus, "Characterization of mammalian eIF4E-family members," *European Journal of Biochemistry*, vol. 271, no. 11, pp. 2189–2203, 2004.

[55] J. Robalino, B. Joshi, S. C. Fahrenkrug, and R. Jagus, "Two zebrafish eIF4E family members are differentially expressed and functionally divergent," *The Journal of Biological Chemistry*, vol. 279, no. 11, pp. 10532–10541, 2004.

[56] M. Ptushkina, T. Von der Haar, S. Vasilescu, R. Frank, R. Birkenhäger, and J. E. G. McCarthy, "Cooperative modulation by eIF4G of eIF4E-binding to the mRNA5′ cap in yeast involves a site partially shared by p20," *The EMBO Journal*, vol. 17, no. 16, pp. 4798–4808, 1998.

[57] S. Pyronnet, H. Imataka, A. C. Gingras, R. Fukunaga, T. Hunter, and N. Sonenberg, "Human eukaryotic translation initiation factor 4G (eIF4G) recruits Mnk1 to phosphorylate eIF4E," *The EMBO Journal*, vol. 18, no. 1, pp. 270–279, 1999.

[58] M. Jankowska-Anyszka, B. J. Lamphear, E. J. Aamodt et al., "Multiple isoforms of eukaryotic protein synthesis initiation factor 4E in *Caenorhabditis elegans* can distinguish between mono- and trimethylated mRNA cap structures," *The Journal of Biological Chemistry*, vol. 273, no. 17, pp. 10538–10541, 1998.

[59] B. D. Keiper, B. J. Lamphear, A. M. Deshpande et al., "Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans*," *The Journal of Biological Chemistry*, vol. 275, no. 14, pp. 10590–10596, 2000.

[60] N. Minshall, M. H. Reiter, D. Weil, and N. Standart, "CPEB interacts with an ovary-specific eIF4E and 4E-T in early *Xenopus oocytes*," *The Journal of Biological Chemistry*, vol. 282, no. 52, pp. 37389–37401, 2007.

[61] N. Standart and N. Minshall, "Translational control in early development: CPEB, P-bodies and germinal granules," *Biochemical Society Transactions*, vol. 36, no. 4, pp. 671–676, 2008.

[62] A. V. Evsikov and C. Marín de Evsikova, "Evolutionary origin and phylogenetic analysis of the novel oocyte-specific eukaryotic translation initiation factor 4E in Tetrapoda," *Development Genes and Evolution*, vol. 219, no. 2, pp. 111–118, 2009.

[63] P. F. Cho, C. Gamberi, Y. Cho-Park, I. B. Cho-Park, P. Lasko, and N. Sonenberg, "Cap-dependent translational inhibition establishes two opposing morphogen gradients in *Drosophila* embryos," *Current Biology*, vol. 16, no. 20, pp. 2035–2041, 2006.

[64] T. D. Dinkova, B. D. Keiper, N. L. Korneeva, E. J. Aamodt, and R. E. Rhoads, "Translation of a small subset of *Caenorhabditis elegans* mRNAs is dependent on a specific eukaryotic translation initiation factor 4E isoform," *Molecular and Cellular Biology*, vol. 25, no. 1, pp. 100–113, 2005.

[65] S. L. Baldauf, "The deep roots of eukaryotes," *Science*, vol. 300, no. 5626, pp. 1703–1706, 2003.

[66] R. Derelle and B. F. Lang, "Rooting the eukaryotic tree with mitochondrial and bacterial proteins," *Molecular Biology and Evolution*, vol. 29, no. 4, pp. 1277–89, 2012.

[67] E. J. Javaux, A. H. Knoll, and M. R. Walter, "Morphological and ecological complexity in early eukaryotic ecosystems," *Nature*, vol. 412, no. 6842, pp. 66–69, 2001.

[68] P. J. Keeling, "Chromalveolates and the evolution of plastids by secondary endosymbiosis," *Journal of Eukaryotic Microbiology*, vol. 56, no. 1, pp. 1–8, 2009.

[69] P. J. Keeling, "Diversity and evolutionary history of plastids and their hosts," *American Journal of Botany*, vol. 91, no. 10, pp. 1481–1493, 2004.

[70] P. J. Keeling, "The endosymbiotic origin, diversification and fate of plastids," *Philosophical Transactions of the Royal Society B*, vol. 365, no. 1541, pp. 729–748, 2010.

[71] S. Y. Hwan, J. D. Hackett, F. M. Van Dolah, T. Nosenko, K. L. Lidie, and D. Bhattacharya, "Tertiary endosymbiosis driven genome evolution in dinoflagellate algae," *Molecular Biology and Evolution*, vol. 22, no. 5, pp. 1299–1308, 2005.

[72] C. F. Delwiche, "Tracing the thread of plastid diversity through the tapestry of life," *American Naturalist*, vol. 154, no. 4, pp. S164–S177, 1999.

[73] S. Sato, "The apicomplexan plastid and its evolution," *Cellular and Molecular Life Sciences*, vol. 68, no. 8, pp. 1285–1296, 2011.

[74] J. A. Fernández Robledo et al., "The search for the missing link: a relic plastid in *Perkinsus*?" *International Journal for Parasitology*, vol. 41, pp. 1217–1229, 2011.

[75] P. J. Keeling et al., "The tree of eukaryotes," *Trends in Ecology and Evolution*, vol. 20, pp. 670–676, 2005.

[76] S. M. Adl et al., "Diversity, nomenclature, and taxonomy of protists," *Systematics Biology*, pp. 684–689.

[77] C. E. Lane and J. M. Archibald, "The eukaryotic tree of life: endosymbiosis takes its TOL," *Trends in Ecology and Evolution*, vol. 23, no. 5, pp. 268–275, 2008.

[78] N. King, "The unicellular ancestry of animal development," *Developmental Cell*, vol. 7, no. 3, pp. 313–325, 2004.

[79] T. Cavalier-Smith, "Eukaryote kingdoms: seven or nine?" *BioSystems*, vol. 14, no. 3-4, pp. 461–481, 1981.

[80] T. Cavalier-Smith, "The phagotrophic origin of eukaryotes and phylogenetic classification on protozoa," *International Journal of Systematic and Evolutionary Microbiology*, vol. 52, no. 2, pp. 297–354, 2002.

[81] T. Cavalier-Smith, "A revised six-kingdom system of life," *Biological Reviews of the Cambridge Philosophical Society*, vol. 73, no. 3, pp. 203–266, 1998.

[82] T. Cavalier-Smith and E. E. Y. Chao, "Phylogeny and classification of phylum Cercozoa (Protozoa)," *Protist*, vol. 154, no. 3-4, pp. 341–358, 2003.

[83] J. M. Archibald and P. J. Keeling, "Actin and ubiquitin protein sequences support a cercozoan/foraminiferan ancestry for the plasmodiophorid plant pathogens," *Journal of Eukaryotic Microbiology*, vol. 51, no. 1, pp. 113–118, 2004.

[84] R. D. Adam, "Biology of *Giardia lamblia*," *Clinical Microbiology Reviews*, vol. 14, no. 3, pp. 447–475, 2001.

[85] J. Jerlström-Hultqvist, O. Franzén, J. Ankarklev et al., "Genome analysis and comparative genomics of a *Giardia intestinalis* assemblage E isolate," *BMC Genomics*, vol. 11, no. 1, article 543, 2010.

[86] R. D. Adam, "The *Giardia lamblia* genome," *International Journal for Parasitology*, vol. 30, no. 4, pp. 475–484, 2000.

[87] S. Hausmann, M. A. Altura, M. Witmer, S. M. Singer, H. G. Elmendorf, and S. Shuman, "Yeast-like mRNA capping apparatus in *Giardia lamblia*," *The Journal of Biological Chemistry*, vol. 280, no. 13, pp. 12077–12086, 2005.

[88] L. Li and C. C. Wang, "Capped mRNA with a single nucleotide leader is optimally translated in a primitive eukaryote, *Giardia lamblia*," *The Journal of Biological Chemistry*, vol. 279, no. 15, pp. 14656–14664, 2004.

[89] X. Niu, T. Hartshorne, X. Y. He, and N. Agabian, "Characterization of putative small nuclear RNAs from *Giardia lamblia*," *Molecular and Biochemical Parasitology*, vol. 66, no. 1, pp. 49–57, 1994.

[90] X. H. Liang, A. Haritan, S. Uliel, and S. Michaeli, "*Trans* and *cis* splicing in trypanosomatids: mechanism, factors, and regulation," *Eukaryotic Cell*, vol. 2, no. 5, pp. 830–840, 2003.

[91] V. Douris, M. J. Telford, and M. Averof, "Evidence for multiple independent origins of *trans*-splicing in Metazoa," *Molecular Biology and Evolution*, vol. 27, no. 3, pp. 684–693, 2010.

[92] M. Keller, L. H. Tessier, R. L. Chan, J. H. Weil, and P. Imbault, "In *Euglena*, spliced-leader RNA (SL-RNA) and 5S rRNA genes are tandemly repeated," *Nucleic Acids Research*, vol. 20, no. 7, pp. 1711–1715, 1992.

[93] T. Blumenthal, "*Trans*-splicing and polycistronic transcription in *Caenorhabditis elegans*," *Trends in Genetics*, vol. 11, no. 4, pp. 132–136, 1995.

[94] J. D. Bangs, P. F. Crain, T. Hashizume, J. A. McCloskey, and J. C. Boothroyd, "Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides," *The Journal of Biological Chemistry*, vol. 267, no. 14, pp. 9805–9815, 1992.

[95] K. L. Perry, K. P. Watkins, and N. Agabian, "Trypanosome mRNAs have unusual 'cap 4' structures acquired by addition of a spliced leader," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 84, no. 23, pp. 8190–8194, 1987.

[96] C. E. Clayton, "Life without transcriptional control? From fly to man and back again," *The EMBO Journal*, vol. 21, no. 8, pp. 1881–1888, 2002.

[97] S. Haile and B. Papadopoulou, "Developmental regulation of gene expression in trypanosomatid parasitic protozoa," *Current Opinion in Microbiology*, vol. 10, no. 6, pp. 569–577, 2007.

[98] R. Queiroz, C. Benz, K. Fellenberg, J. D. Hoheisel, and C. Clayton, "Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons," *BMC Genomics*, vol. 10, article 1471, p. 495, 2009.

[99] Y. Yoffe, J. Zuberek, A. Lerer et al., "Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in *Leishmania*," *Eukaryotic Cell*, vol. 5, no. 12, pp. 1969–1979, 2006.

[100] E. R. Freire, R. Dhalia, D. M. N. Moura et al., "The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties," *Molecular and Biochemical Parasitology*, vol. 176, no. 1, pp. 25–36, 2011.

[101] J. Marcotrigiano, A. C. Gingras, N. Sonenberg, and S. K. Burley, "Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G," *Molecular Cell*, vol. 3, no. 6, pp. 707–716, 1999.

[102] R. Dhalia, C. R. S. Reis, E. R. Freire et al., "Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues," *Molecular and Biochemical Parasitology*, vol. 140, no. 1, pp. 23–41, 2005.

[103] Y. Hou and S. Lin, "Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: gene content estimation for dinoflagellate genomes," *PLoS ONE*, vol. 4, no. 9, Article ID e6978, 2009.

[104] S. Moreno Díaz de la Espina, E. Alverca, A. Cuadrado, and S. Franca, "Organization of the genome and gene expression in a nuclear environment lacking histones and nucleosomes: the amazing dinoflagellates," *European Journal of Cell Biology*, vol. 84, no. 2-3, pp. 137–149, 2005.

[105] T. R. Bachvaroff and A. R. Place, "From stop to start: tandem gene arrangement, copy number and *Trans*-splicing sites in the dinoflagellate *Amphidinium carterae*," *PLoS ONE*, vol. 3, no. 8, Article ID e2929, 2008.

[106] M. Mittag, D. H. Lee, and J. W. Hastings, "Circadian expression of the luciferin-binding protein correlates with the binding of a protein to the 3′ untranslated region of its mRNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 12, pp. 5257–5261, 1994.

[107] T. Fagan, D. Morse, and J. W. Hastings, "Circadian synthesis of a nuclear-encoded chloroplast glyceraldehyde-3-phosphate dehydrogenase in the dinoflagellate *Gonyaulax polyedra* is translationally controlled," *Biochemistry*, vol. 38, no. 24, pp. 7689–7695, 1999.

[108] S. A. Brunelle and F. M. Van Dolah, "Post-transcriptional regulation of S-Phase genes in the dinoflagellate, *Karenia brevis*," *Journal of Eukaryotic Microbiology*, vol. 58, no. 4, pp. 373–382, 2011.

[109] A. Schröder-Lorenz and L. Rensing, "Circadian changes in protein-synthesis rate and protein phosphorylation in cell-free extracts of *Gonyaulax polyedra*," *Planta*, vol. 170, no. 1, pp. 7–13, 1987.

[110] M. R. Ten Lohuis and D. J. Miller, "Light-regulated transcription of genes encoding peridinin chlorophyll a proteins and the major intrinsic light-harvesting complex proteins in the dinoflagellate *Amphidinium carterae* hulburt (Dinophycae): changes in cytosine methylation accompany photoadaptation," *Plant Physiology*, vol. 117, no. 1, pp. 189–196, 1998.

[111] C. Rossini, W. Taylor, T. Fagan, and J. W. Hastings, "Lifetimes of mRNAs for clock-regulated proteins in a dinoflagellate," *Chronobiology International*, vol. 20, no. 6, pp. 963–976, 2003.

[112] F. W. F. Lee, D. Morse, and S. C. L. Lo, "Identification of two plastid proteins in the dinoflagellate *Alexandrium affine* that are substantially down-regulated by nitrogen-depletion," *Journal of Proteome Research*, vol. 8, no. 11, pp. 5080–5082, 2009.

[113] H. Akimoto, T. Kinumi, and Y. Ohmiya, "Circadian rhythm of a TCA cycle enzyme is apparently regulated at the translational level in the dinoflagellate *Lingulodinium polyedrum*," *Journal of Biological Rhythms*, vol. 20, no. 6, pp. 479–489, 2005.

[114] D. Morse, P. M. Milos, E. Roux, and J. W. Hastings, "Circadian regulation of bioluminescence in *Gonyaulax* involves translational control," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, no. 1, pp. 172–176, 1989.

[115] F. M. Van Dolah, K. B. Lidie, J. S. Morey et al., "Microarray analysis of diurnal- and circadian-regulated genes in the Florida red-tide dinoflagellate *Karenia brevis* (Dinophyceae)," *Journal of Phycology*, vol. 43, no. 4, pp. 741–752, 2007.

[116] A. Moustafa, A. N. Evans, D. M. Kulis et al., "Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence," *PLoS ONE*, vol. 5, no. 3, Article ID e9688, 2010.

[117] K. B. Lidie and F. M. van Dolah, "Spliced leader RNA-mediated *trans*-splicing in a dinoflagellate, *Karenia brevis*," *Journal of Eukaryotic Microbiology*, vol. 54, no. 5, pp. 427–435, 2007.

[118] H. Zhang, Y. Hou, L. Miranda et al., "Spliced leader RNA *trans*-splicing in dinoflagellates," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 11, pp. 4618–4623, 2007.

[119] T. R. Bachvaroff, A. R. Place, and D. W. Coats, "Expressed sequence tags from *Amoebophrya sp.* infecting *Karlodinium veneficum*: comparing host and parasite sequences," *Journal of Eukaryotic Microbiology*, vol. 56, no. 6, pp. 531–541, 2009.

[120] X. Xu, J. Vatsyayan, C. Gao, C. J. Bakkenist, and J. Hu, "Sumoylation of eIF4E activates mRNA translation," *EMBO Reports*, vol. 11, no. 4, pp. 299–304, 2010.

[121] X. Xu, J. Vatsyayan, C. Gao, C. J. Bakkenist, and J. Hu, "HDAC2 promotes eIF4E sumoylation and activates mRNA translation gene specifically," *The Journal of Biological Chemistry*, vol. 285, no. 24, pp. 18139–18143, 2010.

[122] T. R. Bachvaroff, S. M. Handy, A. R. Place, and C. F. Delwiche, "Alveolate phylogeny inferred using concatenated ribosomal proteins," *Journal of Eukaryotic Microbiology*, vol. 58, no. 3, pp. 223–233, 2011.

[123] J. A. F. Robledo, P. Courville, M. F. M. Cellier, and G. R. Vasta, "Gene organization and expression of the divalent cation transporter Nramp in the protistan parasite *Perkinsus marinus*," *Journal of Parasitology*, vol. 90, no. 5, pp. 1004–1014, 2004.

[124] J. L. Hearne and J. S. Pitula, "Identification of two spliced leader RNA transcripts from *Perkinsus marinus*," *Journal of Eukaryotic Microbiology*, vol. 58, no. 3, pp. 266–268, 2011.

[125] H. Zhang, D. A. Campbell, N. R. Sturm, C. F. Dungan, and S. Lin, "Spliced leader RNAs, mitochondrial gene frameshifts and multi-protein phylogeny expand support for the genus *Perkinsus* as a unique group of alveolates," *PLoS ONE*, vol. 6, no. 5, Article ID e19933, 2011.

[126] T. Cavalier-Smith, "Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree," *Journal of Eukaryotic Microbiology*, vol. 46, no. 4, pp. 347–366, 1999.

[127] T. Cavalier-Smith and M. J. Beaton, "The skeletal function of non-genic nuclear DNA: new evidence from ancient cell chimaeras," *Genetica*, vol. 106, no. 1-2, pp. 3–13, 1999.

[128] T. Cavalier-Smith, "Membrane heredity and early chloroplast evolution," *Trends in Plant Science*, vol. 5, no. 4, pp. 174–182, 2000.

[129] S. E. Douglas, C. A. Murphy, D. F. Spencer, and M. W. Grayt, "Cryptomonad algae are evolutionary chimaeras of two phylogenetically distinct unicellular eukaryotes," *Nature*, vol. 350, no. 6314, pp. 148–151, 1991.

[130] S. E. Douglas, "Eukaryote-eukaryote endosymbioses: insights from studies of a cryptomonad alga," *BioSystems*, vol. 28, no. 1–3, pp. 57–68, 1992.

[131] U. G. Maier, C. J. B. Hofmann, S. Eschbach, J. Wolters, and G. L. Igloi, "Demonstration of nucleomorph-encoded eukaryotic small subunit ribosomal RNA in cryptomonads," *Molecular and General Genetics*, vol. 230, no. 1-2, pp. 155–160, 1991.

[132] T. Cavalier-Smith, "Nucleomorphs: enslaved algal nuclei," *Current Opinion in Microbiology*, vol. 5, no. 6, pp. 612–619, 2002.

[133] S. Douglas, S. Zauner, M. Fraunholz et al., "The highly reduced genome of an enslaved algal nucleus," *Nature*, vol. 410, no. 6832, pp. 1091–1096, 2001.

[134] U. G. Maier, S. E. Douglas, and T. Cavalier-Smith, "The nucleomorph genomes of cryptophytes and chlorarachniophytes," *Protist*, vol. 151, no. 2, pp. 103–109, 2000.

[135] S. Zauner, M. Fraunholz, J. Wastl et al., "Chloroplast protein and centrosomal genes, a tRNA intron, and odd telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 1, pp. 200–205, 2000.

[136] G. I. McFadden, P. R. Gilson, and S. E. Douglas, "The photosynthetic endosymbiont in cryptomonad cells produces both chloroplast and cytoplasmic-type ribosomes," *Journal of Cell Science*, vol. 107, no. 2, pp. 649–657, 1994.

[137] L. D. Kapp and J. R. Lorsch, "The molecular mechanics of eukaryotic translation," *Annual Review of Biochemistry*, vol. 73, pp. 657–704, 2004.

[138] W. L. Zoll, L. E. Horton, A. A. Komar, J. O. Hensold, and W. C. Merrick, "Characterization of mammalian eIF2A and identification of the yeast homolog," *The Journal of Biological Chemistry*, vol. 277, no. 40, pp. 37079–37087, 2002.

[139] M. Matsuzaki, O. Misumi, T. Shin-I et al., "Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D," *Nature*, vol. 428, no. 6983, pp. 653–657, 2004.

[140] N. Ohta, N. Sato, H. Nozaki, and T. Kuroiwa, "Analysis of the cluster of ribosomal protein genes in the plastid genome of a unicellular red alga *Cyanidioschyzon merolae*: translocation of the *str* cluster as an early event in the rhodophyte-chromophyte lineage of plastid evolution," *Journal of Molecular Evolution*, vol. 45, no. 6, pp. 688–695, 1997.

[141] M. Altmann, P. P. Muller, B. Wittmer, F. Ruchti, S. Lanker, and H. Trachsel, "A *Saccharomyces cerevisiae* homologue of mammalian translation initiation factor 4B contributes to RNA helicase activity," *The EMBO Journal*, vol. 12, no. 10, pp. 3997–4003, 1993.

[142] S. B. Gould, "Ariadne's thread: guiding a protein across five membranes in cryptophytes," *Journal of Phycology*, vol. 44, no. 1, pp. 23–26, 2008.

[143] B. Joshi, J. Robalino, E. J. Schott, and R. Jagus, "Yeast "knock-out-and-rescue" system for identification of eIF4E-family members possessing eIF4E-activity," *BioTechniques*, vol. 33, no. 2, pp. 392–401, 2002.

[144] S. Vasilescu, M. Ptushkina, B. Linz, P. P. Müller, and J. E. G. McCarthy, "Mutants of eukaryotic initiation factor eIF-4E with altered mRNA cap binding specificity reprogram mRNA selection by ribosomes in *Saccharomyces cerevisiae*," *The Journal of Biological Chemistry*, vol. 271, no. 12, pp. 7030–7037, 1996.

[145] B. La Scola, T. J. Marrie, J. P. Auffray, and D. Raoult, "Mimivirus in pneumonia patients," *Emerging Infectious Diseases*, vol. 11, no. 3, pp. 449–452, 2005.

[146] D. Raoult, S. Audic, C. Robert et al., "The 1.2-megabase genome sequence of Mimivirus," *Science*, vol. 306, no. 5700, pp. 1344–1350, 2004.

[147] D. Raoult, B. La Scola, and R. Birtles, "The discovery and characterization of mimivirus, the largest known virus and putative pneumonia agent," *Clinical Infectious Diseases*, vol. 45, no. 1, pp. 95–102, 2007.

[148] M. Suzan-Monti, B. La Scola, and D. Raoult, "Genomic and evolutionary aspects of Mimivirus," *Virus Research*, vol. 117, no. 1, pp. 145–155, 2006.

[149] J. M. Claverie and C. Abergel, "Mimivirus and its virophage," *Annual Review of Genetics*, vol. 43, pp. 49–66, 2009.

[150] J. M. Claverie, C. Abergel, and H. Ogata, "Mimivirus," *Current Topics in Microbiology and Immunology*, vol. 328, pp. 89–121, 2009.

[151] D. Benarroch, P. Smith, and S. Shuman, "Characterization of a trifunctional mimivirus mRNA capping enzyme and crystal structure of the RNA triphosphatase domain," *Structure*, vol. 16, no. 4, pp. 501–512, 2008.

[152] D. Benarroch, Z. R. Qiu, B. Schwer, and S. Shuman, "Characterization of a mimivirus RNA cap guanine-N2 methyltransferase," *RNA*, vol. 15, no. 4, pp. 666–674, 2009.

[153] J. M. Claverie and C. Abergel, "Mimivirus: the emerging paradox of quasi-autonomous viruses," *Trends in Genetics*, vol. 26, no. 10, pp. 431–437, 2010.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 960420, 15 pages
doi:10.1155/2012/960420

Research Article

The Distribution of eIF4E-Family Members across Insecta

Gritta Tettweiler,[1,2] Michelle Kowanda,[1] Paul Lasko,[1] Nahum Sonenberg,[2] and Greco Hernández[3]

[1] *Department of Biology, McGill University, 1205 Dr. Penfield, Montreal, QC, Canada H3A 1B1*
[2] *Department of Biochemistry and Goodman Cancer Research Center, McGill University, Montreal, QC, Canada H3A 1A3*
[3] *Division of Basic Research, National Institute for Cancer (INCan), Avenida San Fernando No. 22, Tlalpan, 14080 Mexico City, DF, Mexico*

Correspondence should be addressed to Greco Hernández, ghernandezr@incan.edu.mx

Received 31 December 2011; Accepted 14 March 2012

Academic Editor: Armen Parsyan

Insects are part of the earliest faunas that invaded terrestrial environments and are the first organisms that evolved controlled flight. Nowadays, insects are the most diverse animal group on the planet and comprise the majority of extant animal species described. Moreover, they have a huge impact in the biosphere as well as in all aspects of human life and economy; therefore understanding all aspects of insect biology is of great importance. In insects, as in all cells, translation is a fundamental process for gene expression. However, translation in insects has been mostly studied only in the model organism *Drosophila melanogaster*. We used all publicly available genomic sequences to investigate in insects the distribution of the genes encoding the cap-binding protein eIF4E, a protein that plays a crucial role in eukaryotic translation. We found that there is a diversity of multiple ortholog genes encoding eIF4E isoforms within the genus *Drosophila*. In striking contrast, insects outside this genus contain only a single *eIF4E* gene, related to *D. melanogaster* eIF4E-1. We also found that all insect species here analyzed contain only one Class II gene, termed *4E-HP*. We discuss the possible evolutionary causes originating the multiplicity of eIF4E genes within the genus *Drosophila*.

1. Introduction

Insects are the most diverse animal group on Earth and comprise over half of all extant described species, dominating thus all terrestrial ecosystems [1–4]. Winged insects were the first organisms that evolved controlled flight, some 120, 200, and 300 million years (Myr) before flying reptiles, birds, and bats, respectively. Indeed, wings are believed to have led largely to the spectacular diversification of insects because they were able to explore and invade all terrestrial ecosystems, escape predators, and exploit scattered resources [2, 5]. Many studies show that insect diversity has been also strongly shaped by other evolutionary and ecological processes, including their relative ancient geological age, low extinction rate, ecological niches occupancy, sexual selection, and sexual conflict [1].

Insects originated 434–421 Myr ago during the Silurian Period, and it is suggested that earliest terrestrial faunas already included wingless insects [2, 5, 6]. Indeed, the aquatic-terrestrial transition of insect ancestors is associated with the earliest vascular land plants fossils. Thus, it is thought that true insects evolved from an aquatic arthropod that formed an ecological association with the earliest vascular plants and subsequently both lineages coevolved [2, 6]. By the Permian (299–251 Myr ago) nearly all extant insect orders already have emerged, and later a second spectacular radiation happened in the Jurassic. Insects have been diverging ever since [2, 7, 8]. Winged insects, which account for more than 98% of the class Insecta, emerged when early arborescent plants evolved (pteridophytes, mostly ferns, and horsetails) 380–354 Myr ago (during the Devonian). It is hypothesized that insect flight arose as an adaptation to the increasing height of trees, and that a number of highly successful insect species coevolved with flowering plants [2, 5, 6, 9].

Besides their crucial ecological importance in all terrestrial ecosystems, insects have a huge direct impact in all aspects of human life and economy. In agriculture, some

species cause huge damage to crops (e.g., aphids and weevil beetles), whilst others are of great benefit to flowering plants, which depend on pollinating species (e.g., bees, wasp, and butterflies). There are many species that can spread human pathogens (e.g., mosquitoes, fleas, and bed bugs) as well as key model organisms for basic research (*Drosophila*). Furthermore, several species serve as research objects for social behavior studies (e.g., bees and ants). Because of their overall significance, for many years immense efforts have been put forward to studying all aspects of insect biology. However, many biological processes, including translation, are still poorly studied at the molecular level. Therefore, further characterization of insect translation is necessary.

Most eukaryotic mRNAs are translated by a cap-dependent mechanism, whereby the mRNA is recruited to the ribosome through recognition of the 5′ cap structure (m^7GpppN, where N is any nucleotide) by the cap-binding protein eIF4E in complex with the scaffold protein eIF4G and the RNA helicase eIF4A [10, 11]. Three-dimensional studies demonstrated that eIF4E associated to cap-analogues resembles "cupped-hands" in which the cap structure is stacked between two highly conserved tryptophan residues (Trp-56 and Trp-102 of mouse eIF4E) through π bond interactions. A third conserved tryptophan residue (Trp-166 of mouse eIF4E) binds the N^7-methyl moiety of the cap structure [12–15]. Due to its pivotal role in translation, eIF4E activity is tightly regulated. Perhaps the most prominent regulatory mechanism is performed by eIF4E-binding proteins (4E-BPs), which bind eIF4E via an eIF4E-binding motif that is shared with eIF4G. 4E-BPs act as competitive inhibitors of eIF4E-eIF4G interaction and therefore of translation [10, 16, 17]. Another mechanism regulating eIF4E activity in some metazoans, including human, *Drosophila*, and *Aplysia*, is by phosphorylation of Ser-209 (mouse protein numbering; Ser251 in *Drosophila* eIF4E-1) [18–20].

Among insects, the unique translation initiation machinery that has been studied thus far is that from *D. melanogaster*. This species possesses seven genes encoding eight eIF4E cognates, one of them being 4E-HP (eIF4E-homolog protein) [19, 21–27]. All residues involved in 5′ cap structure binding are conserved in all eIF4Es [22–26], and experimental evidence confirmed their ability to bind this structure [21, 24, 26]. Likewise, most residues involved in eIF4G and 4E-BP binding are conserved and yeast two-hybrid experiments showed that all of them, except for eIF4E-6 and 4E-HP, interact with both proteins [26]. A functional assay showed that *D. melanogaster* eIF4E-1, eIF4E-2, eIF4E-4, eIF4E-5, and eIF4E-7, but not eIF4E-3 and eIF4E-6, are able to phenotypically rescue a lethal *eIF4E*-deficient yeast strain [26]. eIF4E-1 loss-of-function mutations cause growth arrest, severe embryonic defects, and lead to embryonic lethality [19, 28–30], and phosphorylation of eIF4E-1 at Ser251 is necessary for growth of the whole organisms [19]. Evidence supports the idea that there is physiological specialization of eIF4E cognates. While global translation is performed by eIF4E-1 [19, 28], eIF4E-3 is a testis-specific factor promoting translation in this tissue [31], eIF4E-5 might be involved in autophagy [32] and 4E-HP is a translational repressor [27, 33, 34]. Moreover, other activities have been reported for eIF4E-1, including a role in neurogenesis [35, 36] and a nuclear role in splicing [37]. Interactions of eIF4E-1 with different proteins, including 4E-BP [26, 38], Cup [39], Diap1 [40], and Ago2-Risc complex [41], have been described. Additionally, 4E-HP was found to interact with the RNA helicase Belle [42].

Recent advances in sequencing technology allow comparative analysis of multiple genomes across a wide range of evolutionarily related species. Thus, gene and protein annotation of twelve different *Drosophila* species [43] and from other insect species [44, 45] are now available. Here we investigated the distribution of the cap-binding proteins eIF4E and 4E-HP across the class Insecta.

2. Material and Methods

We compared annotated protein sequences of insects eIF4E-family members obtained from all publicly accessible databases, that is, http://umbicc3-215.umbi.umd.edu/ [45] and from several sequencing projects available in the NCBI GenBank NR, http://flybase.org/ and in http://www.butterflybase.org/ [44]. The genomes analyzed were from 12 *Drosophila* species [43], *Aedes aegypti*, *Anopheles gambiae* (all Diptera), *Camponotus floridanus, Harpegnathos saltator, Apis mellifera, Nasonia vitripennis* (all Hymenoptera), *Tribolium castaneum* (Coleoptera), *Manduca sexta*, *Spodoptera frugiperda*, *Heliconius melpomene*, *Bombix mori*, *Papilio xuthus* (all Lepidoptera), and *Acyrthosiphon pisum* (Hemiptera). Table 1 shows all annotated genes and the proteins they encode that were analyzed in this study. Incomplete sequences and sequences encoding partial putative proteins were excluded. Amino acid sequences were aligned using ClustalW [46, 47] with the Biology Workbench bioinformatics package and improved by eye. Phylograms were assembled by neighbor-joining using *MEGA5* program [48].

Jagus and colleagues proposed a classification of eIF4Es from 230 species into three classes according to variations in the residues Trp-43 and Trp-56 (human eIF4E numbering) [45, 49]. Class I members contain both Trp residues; Class II members contain Tyr, Phe, or Leu at the first position and Tyr or Phe at the second position; Class III proteins contain Trp at the first position and Cys or Tyr at the second position [45, 49]. In the present study we will follow this classification. Since *D. melanogaster* is one of the most characterized model organisms and thus the best-studied species of all insects (whose entire genome is available for over a decade now (http://flybase.org/ [50]), and because among insects only eIF4Es and 4E-HP from *D. melanogaster* have been characterized [19, 21–42], we chose *D. melanogaster* eIF4Es sequences, numbering and nomenclature (http://flybase.org/ [25, 26]) as a reference. To avoid misunderstanding with another nomenclature [45, 49], here we will keep the fly database (http://flybase.org/) nomenclature, referring when necessary, to the Class each eIF4E belongs to.

3. Results and Discussion

3.1. eIF4E Proteins across the Genus Drosophila. Gene duplication of eIF4E is particularly striking in *D. melanogaster*

Table 1: Overview of annotated genes analyzed in this study.

eIF4E paralogs within *D. melanogaster*.	Orthologs in other *Drosophila* species	Orthologs in other insects
eIF4E-1/2 (CG4035)	*D. simulans* GD12928	*A. aegypti* AAEL001916
	D. sechellia GM24878	*A. gambiae* AGAP007172
	D. erecta GG14044	*C. floridanus* EFN73765
	D. yakuba GE21247	*H. saltator* EFN83757
	D. ananassae GF23736	*A. mellifera* XP_624290.2
	D. willistoni GK20927	*T. castaneum* XP_973494
	D. pseudoobscura GA28658	*M. sexta* MSP00767
	D. persimilis GL12850	*S. frugiperda* AAK94897
	D. virilis GJ13832	*B. mori* BGIBMGA012674
	D. grimshawi GH16860	*H. melpomene* HMP00347
		A. pisum ACYPI001956
eIF4E-3 (CG8023)	*D. simulans* GD14067	
	D. sechellia GM25034	
	D. erecta GG14292	
	D. yakuba GE20721	
	D. ananassae GF25106	
	D. willistoni GK17185	
	D. pseudoobscura GA24628	
	D. persimilis GL26506	
	D. virilis GJ12520	
	D. grimshawi GH15498	
eIF4E-4 (CG10124)	*D. simulans* GD13118	
	D. sechellia GM13832	
	D. erecta GG15032	
	D. yakuba GE20475	
	D. ananassae GF10894	
	D. willistoni GK12583	
	D. pseudoobscura GA28599	
	D. persimilis GL13241	
	D. virilis GJ12668	
	D. grimshawi GH15637	
	D. mojavensis GI12684	
eIF4E-5 (CG8277)	*D. simulans* GD14038	
	D. sechellia GM25004	
	D. erecta GG14453	
	D. yakuba GE21642	
	D. ananassae GF10327	
	D. willistoni GK17737	
	D. pseudoobscura GA28380	
	D. persimilis GL18042	
	D. virilis GJ13889	
	D. grimshawi GH14978	
	D. mojavensis GI13141	
eIF4E-6 (CG1442)	*D. simulans* GD18002	
	D. sechellia GM12271	
	D. erecta GG12044	
	D. yakuba GE10483	

Table 1: Continued.

eIF4E paralogs within *D. melanogaster*.	Orthologs in other *Drosophila* species	Orthologs in other insects
eIF4E-7 (CG32859)	*D. simulans* GD16435	
	D. sechellia GM18997	
	D. erecta GG12718	
	D. yakuba GE16544	
	D. ananassae GF22031	
	D. willistoni GK16243	
	D. virilis GJ16331	
	D. mojavensis GI14982	
4E-HP (CG33100)	*D. simulans* GD18325	*A. aegypti* AAEL005796
	D. sechellia GM23515	*A. gambiae* AGAP002948
	D. erecta GG12377	*T. castaneum* XP_970157
	D. yakuba GE10831	*B. mori* NP_001091833
	D. ananassae GF23230	*P. xuthus* BAG30778
	D. willistoni GK11937	*A. pisum* ACYPI000423
	D. pseudoobscura GA26519	*C. floridanus* EFN65857
	D. persimilis GL24147	*N. vitripennis* XP_003426016
	D. virilis GJ14537	*A. mellifera* XP_623570
	D. grimshawi GH19097	
	D. mojavensis GI23433	

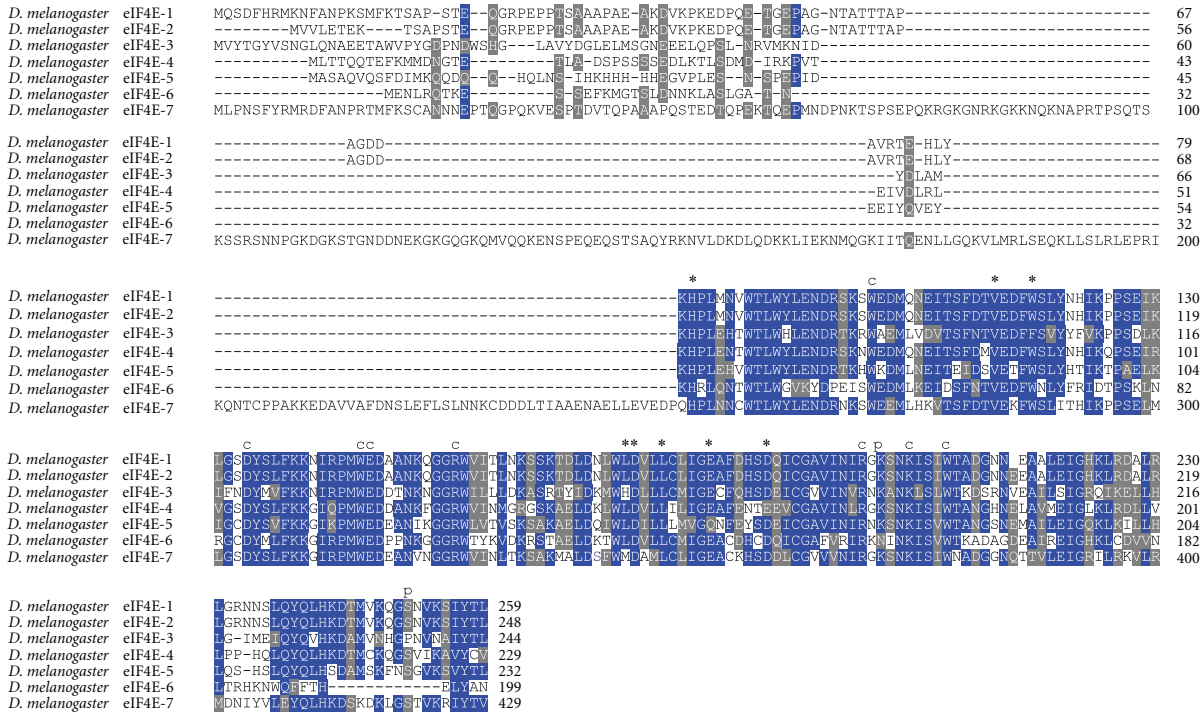


Figure 1: Seven eIF4E cognates in *D. melanogaster*. (A) ClustalW alignment of amino acid sequences representing Class I eIF4E family members from *D. melanogaster*. An extended version of eIF4E-6, not detected in previous studies, is included. Identical (blue) or conservative (gray) amino acid residues in at least 70% of sequences are highlighted. Conservative substitutions groups are STA, or NEQK, or NHQK, or NDEQ, or QHRK, or MILV, or MILF, or HY, or FYW, or GA. Residues essential for eIF4G- and 4E-BP binding are marked ∗; residues involved in cap binding are marked lower case *c*; phosphorylatable Ser, as well as Lys described to form a salt bridge with P-Ser, are marked lower case *p*.

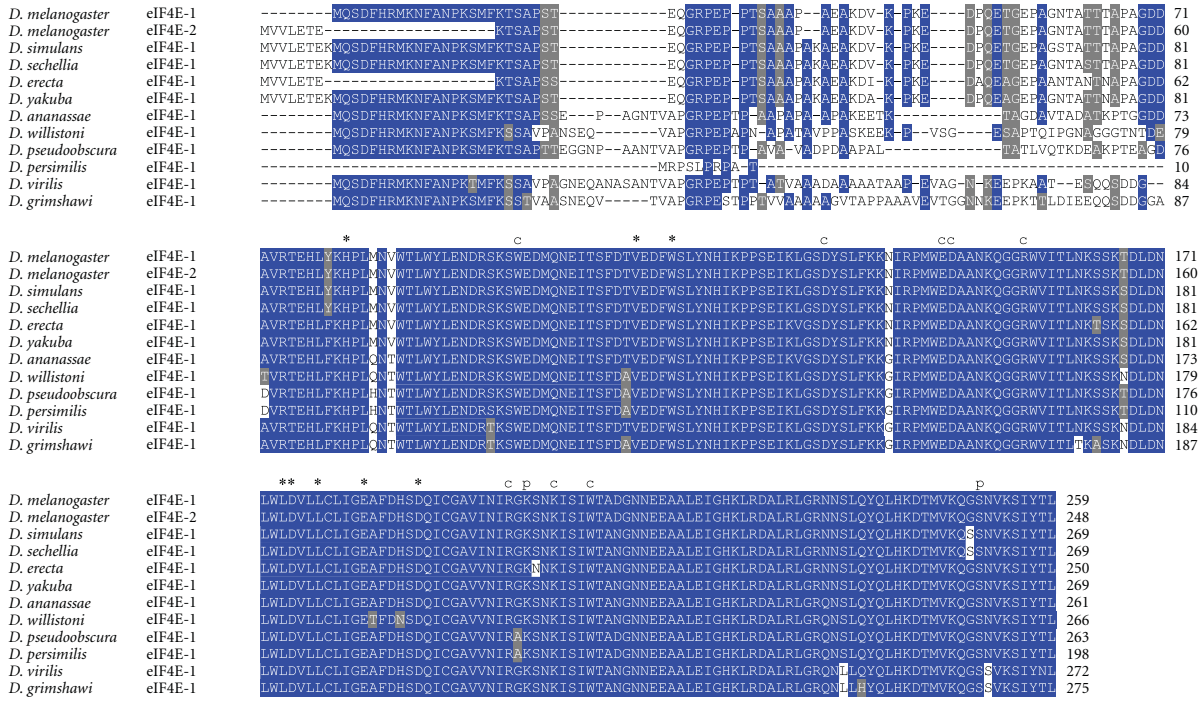


Figure 2: ClustalW alignment of amino acid sequences of eIF4E-1 orthologs from species of the genus *Drosophila*.

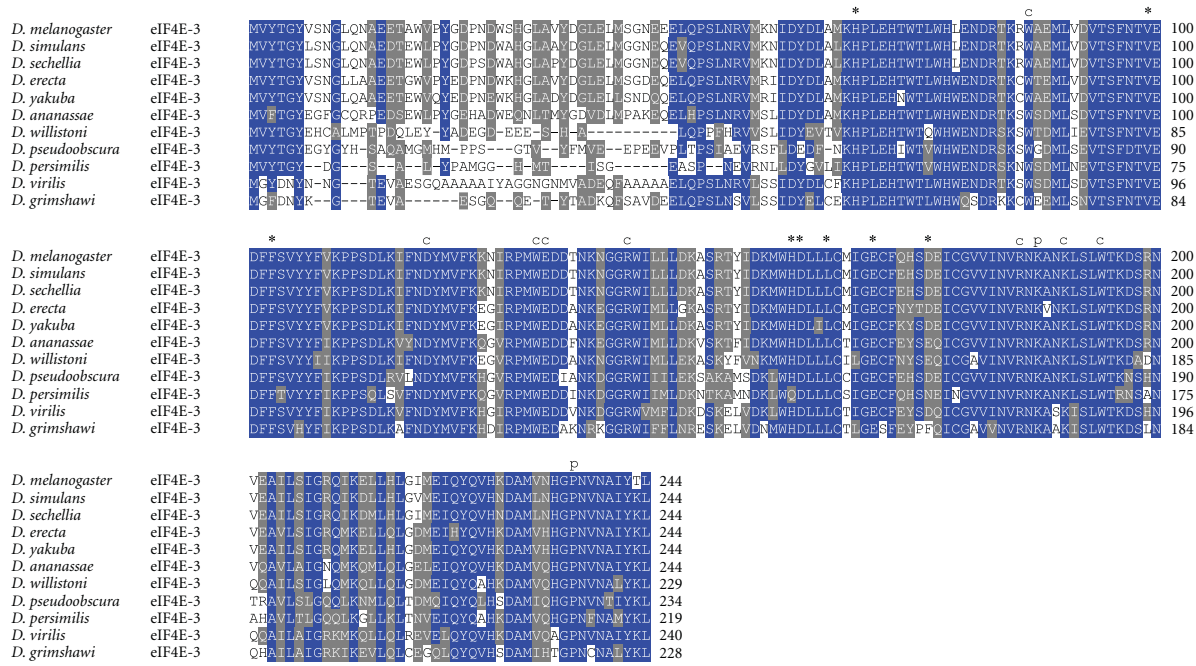


Figure 3: ClustalW alignment of amino acid sequences of eIF4E-3 orthologs from species of the genus *Drosophila*.

with seven different cognates of Class I eIF4Es (eIF4E-1 trough eIF4E-7) and one Class II gene, termed *4E-HP* [25, 26]. Although sequence comparisons of all *D. melanogaster* eIF4Es are shown elsewhere [25, 26], a comparison of these proteins including an extended version of eIF4E-6 (see below) is shown in Figure 1. Using BLAST searches, it became evident that gene duplication of eIF4E also happened across the entire genus *Drosophila*. Overall, 61 different Class I eIF4E-family members were identified in this genus. We found that *D. simulans*, *D. sechellia*, *D. erecta,* and *D. yakuba* contain each six *eIF4E* genes (*eIF4E-1*, *-3*, *-4*, *-5*, *-6*, and *-7*), *D. ananassae*, *D. willistoni,* and *D. virilis* contain each five (*eIF4E-1*, *-3*, *-4*, *-5*, and *-7*), *D. grimshawi*, *D. pseudoobscura,* and *D. persimilis* contain each four (*eIF4E-1*, *-3*, *-4*, and *-5*)

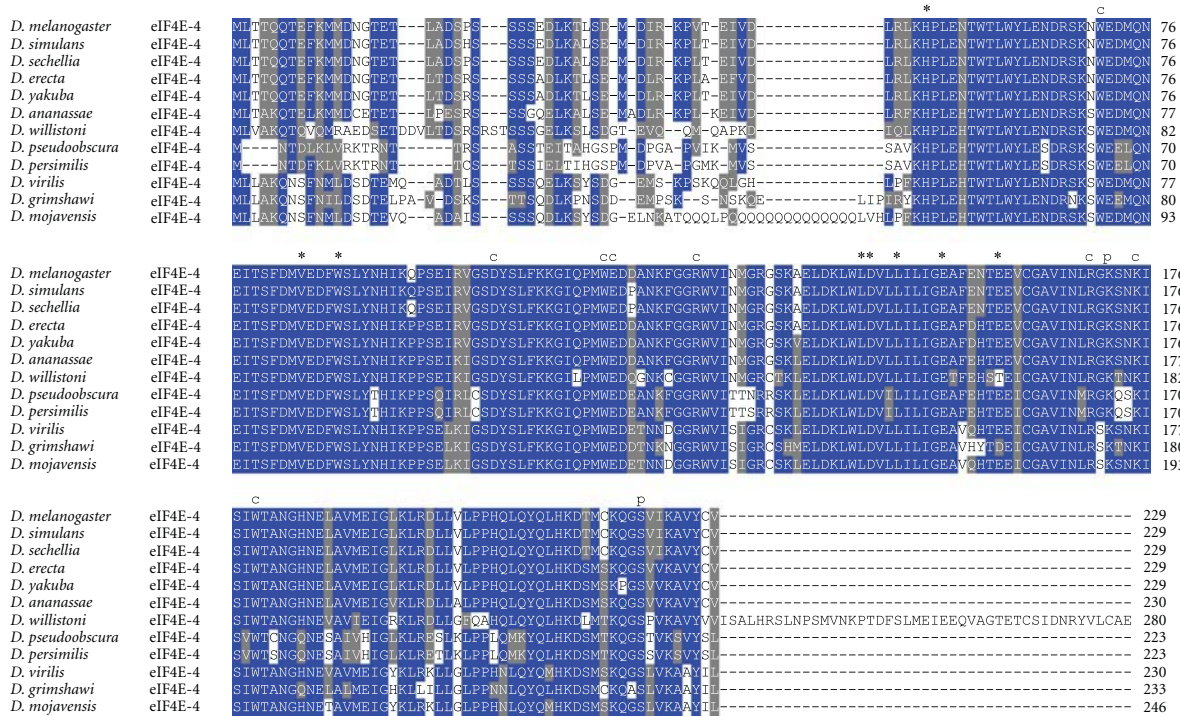


Figure 4: ClustalW alignment of amino acid sequences of eIF4E-4 orthologs from species of the genus *Drosophila*.

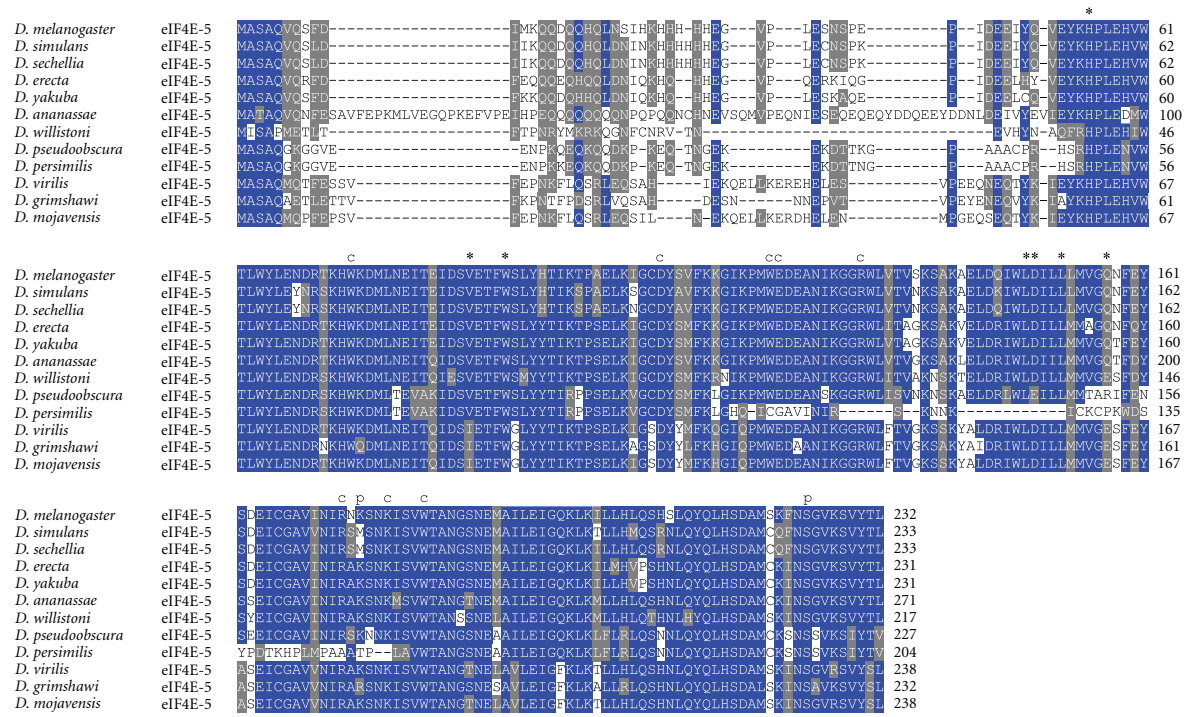


Figure 5: ClustalW alignment of amino acid sequences of eIF4E-5 orthologs from species of the genus *Drosophila*.

and *D. mojavensis* contains three cognates (*eIF4E-4*, *-5*, and *-7*) (Table 1).

It has been shown that *D. melanogaster* eIF4E-1 and eIF4E-2 arise by alternative splicing from the same gene (*eIF4E-1/2*), both proteins differing only in amino acids in the N-terminus. While eIF4E-1 contains the peptide sequence MQSDFHRMKNFANPKSMF, eIF4E-2 contains MVVLETE instead [23, 24] (Figure 1). BLAST searches showed that the gene *eIF4E-1/2* exists only in *D. melanogaster*. In all other species this gene encodes only

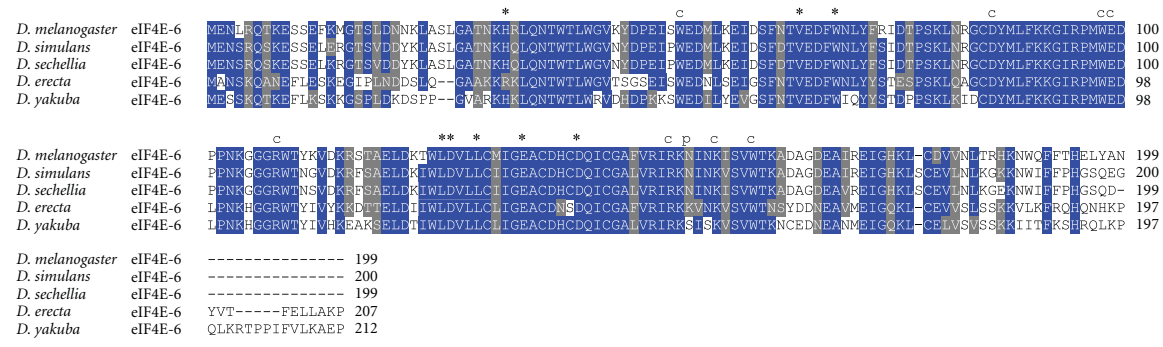


Figure 6: ClustalW alignment of amino acid sequences of eIF4E-6 orthologs from species of the genus *Drosophila*.

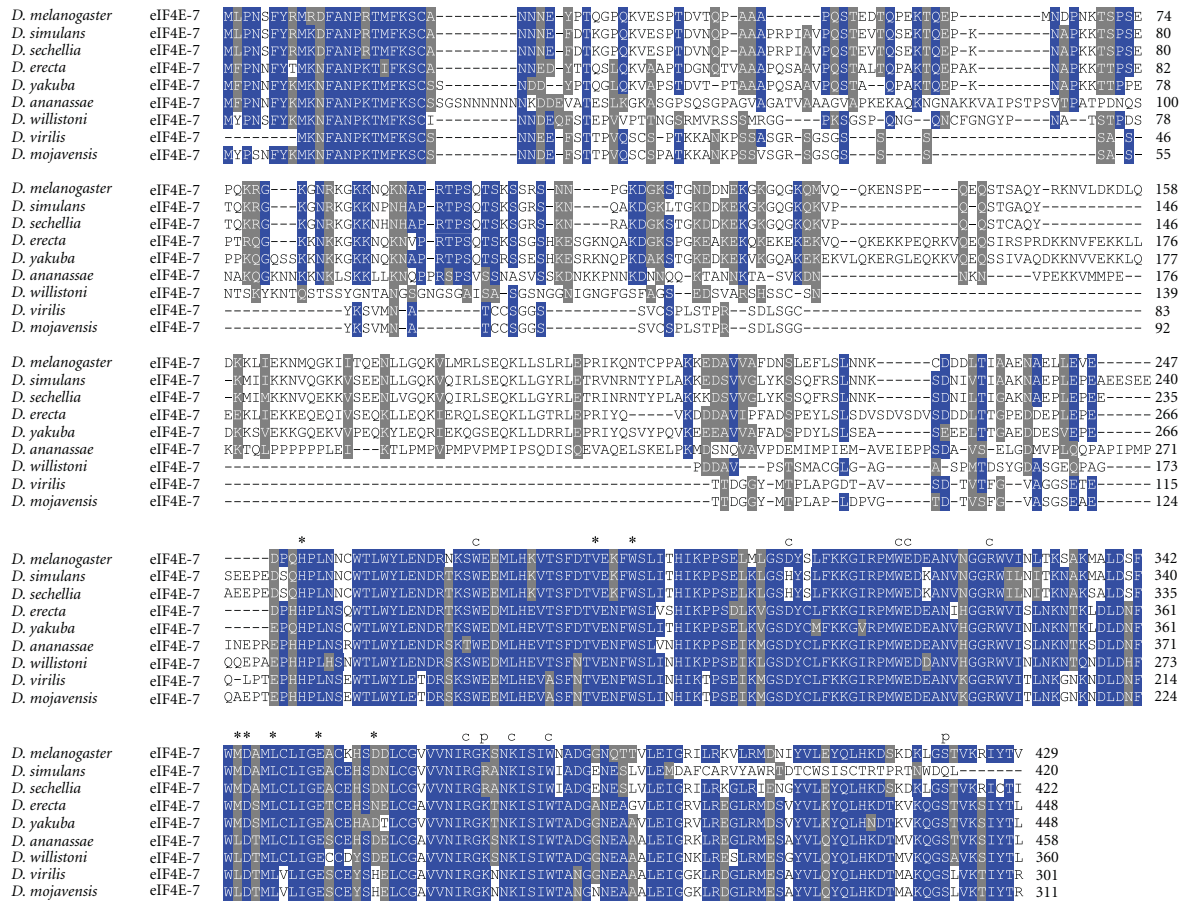


Figure 7: ClustalW alignment of amino acid sequences of eIF4E-7 orthologs from species of the genus *Drosophila*.

one protein, either eIF4E-1 (*D. ananassae*, *D. willistoni*, *D. pseudoobscura*, *D. virilis*, and *D. grimshawi*), eIF4E-2 (*D. erecta*), or a protein with both N-termini fused (*D. simulans*, *D. sechellia*, and *D. yakuba*). Interestingly, *D. persimilis* encodes an eIF4E-1 with a very short and divergent N-terminus (Figure 2). The high variability in eIF4E-1 N-terminus among *Drosophila* species suggests that this region of the protein has no biological relevance.

All residues involved in cap- and eIF4G/4E-BP-binding as well as for phosphorylation are conserved in eIF4E-1 from across the genus *Drosophila* (Figure 2). In eIF4E-3, residues involved in eIF4G/4E-BP binding are mutated in two positions, namely, Trp103>Phe, and Leu160>His (numbering according to *D. melanogaster* eIF4E-3; Figure 3). This significant alteration may explain the weak binding to eIF4G and 4E-BP shown in the yeast two-hybrid system [26]. Both changes are strongly conserved in eIF4E-3 across the genus *Drosophila*. Moreover, eIF4E-3 from all *Drosophila* species lack the counterpart of the phosphorylatable Ser251 of *D. melanogaster* eIF4E-1, possessing a proline instead [31]

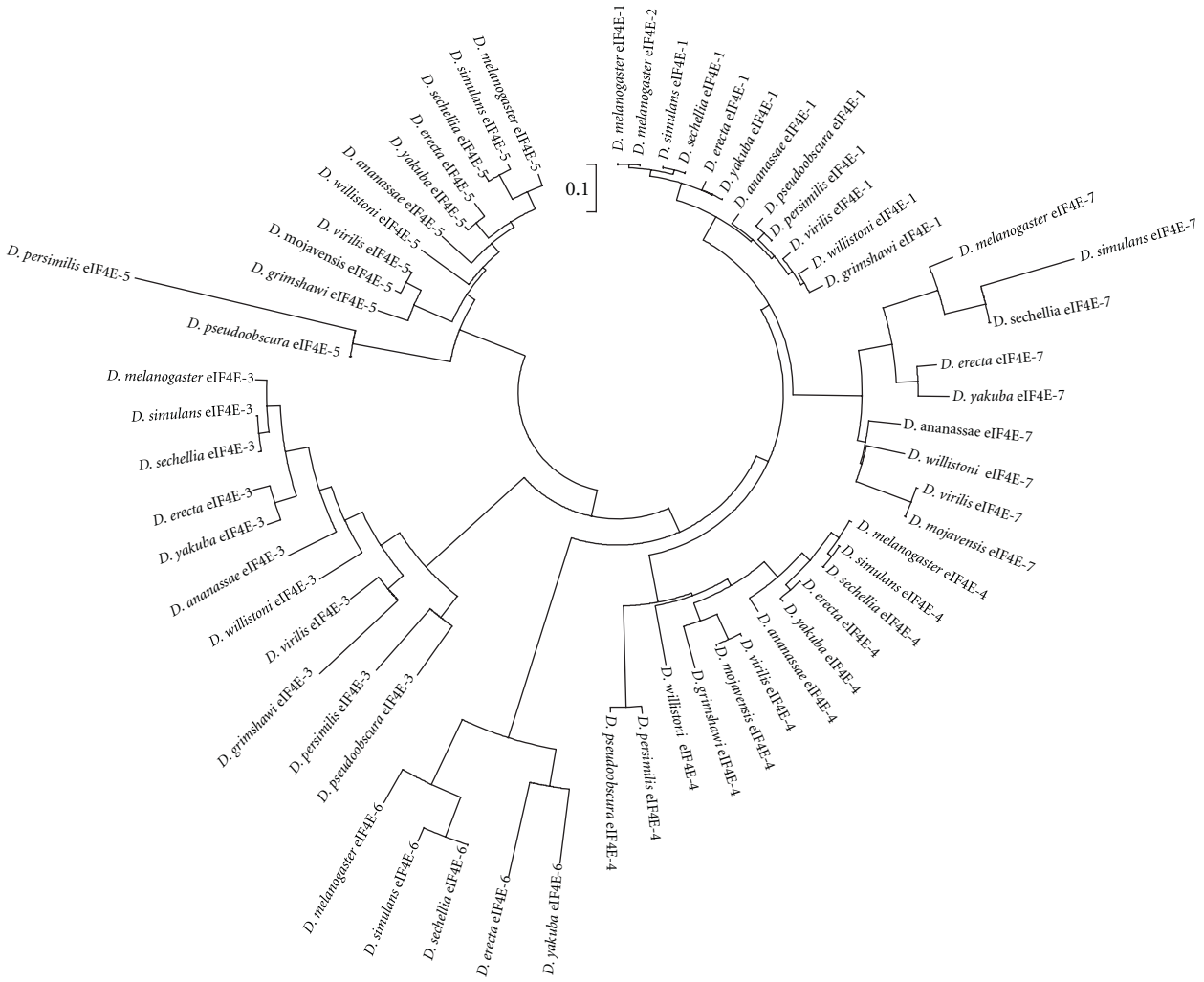


Figure 8: Radial phylogram computed from sequence alignments of eIF4Es from *Drosophila* species.

(Figure 3). *D. willistoni* eIF4E-4 displays the nonconservative amino acid exchange at position 160 Glu>Thr (numbering according to *D. melanogaster* eIF4E-4; Figure 4). eIF4E-5 varies considerably in length, ranging from 204 amino acids in *D. persimilis* to 271 amino acids in *D. ananassae*, and the N-terminus of eIF4E-5 (amino acids 1–53) is highly variable. However, eIF4E-5 is highly conserved from amino acid 54 on (Figure 5). *D. persimilis* eIF4E-5 also diverges from its orthologs in at least ten functionally important amino acids (Figure 5).

Recent experimental evidence supports an extended C-terminus of eIF4E-6 (Tettweiler, Hernández, Sonenberg, and Lasko, unpublished), not detected in previous studies [25, 26]. This extended eIF4E-6 showed the highest similarity to eIF4E-3 and has functionally important residues diverged from eIF4E-1 (Figure 1). One of the differences is a lack of phosphorylatable Ser251 (numbering of eIF4E-1). Surprisingly, although extended eIF4E-6 possesses all amino acids involved in eIF4G/4E-BP binding, experimental evidence showed that it does not bind either of them (Tettweiler, Hernández, Sonenberg, and Lasko, unpublished). The extended eIF4E-6 could only be detected in five species, all of which contain conserved residues important for cap binding (Figure 6). In *D. erecta*, a conserved substitution His>Arg is observed in position 33 (numbering according to *D. melanogaster* eIF4E-6; Figure 6), a residue essential for eIF4G/4E-BP binding. Similar to eIF4E-3, no eIF4E-6 from any species has the counterpart of eIF4E-1 Ser251 (Figure 6).

eIF4E-7 is the longest protein from Class I family members with 301 amino acids in *D. virilis* to 458 amino acids in *D. ananassae* (Figure 7). The high degree of discrepancy in length is attributed to the variability in the N-terminal moiety of the protein (Figure 7). Although eIF4E-7 orthologs are most similar to eIF4E-1, eIF4E-7 from all *Drosophila* species cluster together in separate phylogram branches (Figure 8). Several species are lacking functionally important residues in the eIF4E-7 C-terminus. In particular, in *D. simulans* eIF4E-7 the eIF4E-1 Ser251 counterpart is substituted by a Gln, albeit it is conserved in other *Drosophila* species (Figure 7).

Overall, our analyses indicate that the seven eIF4E-cognates in the genus *Drosophila* form discrete clusters

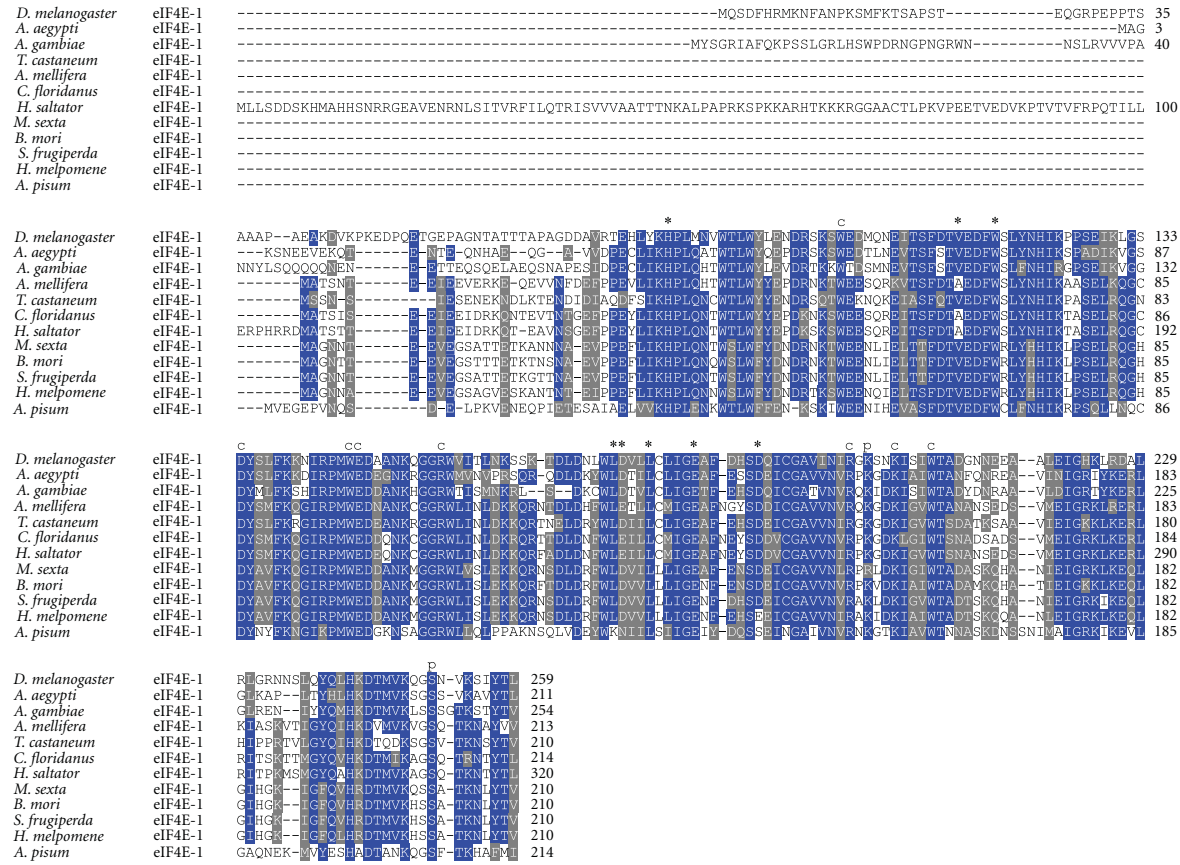


Figure 9: Orthologs of *D. melanogaster* eIF4E-1 in other insects. ClustalW alignment of amino acid sequences of eIF4E-1 orthologs from diverse insect species.

(Figure 8), indicating that they evolved separately from each other before the radiation of ancestral *Drosophila* into the current species.

3.2. eIF4E Proteins in Other Insects. We analyzed protein annotations from all insect genomes that are publicly available. These include species representing non-*Drosophila* Diptera, as well as Hymenoptera, Coleoptera, Lepidoptera, and Hemiptera. Outside of the genus *Drosophila*, eleven more Class I eIF4Es were identified in different insect species (Figures 9 and 10). In contrast to *Drosophila* species, which contain three to seven different Class I eIF4Es cognates, we identified only a single Class I *eIF4E* gene in each insect genome analyzed, all of them related to *D. melanogaster* eIF4E-1 and with a highly variable N-terminus moiety (Figure 9). All amino acids described to be involved in cap and eIF4G/4E-BP binding are conserved in all insect eIF4Es analyzed. The exception is Leu174 (numbering according to *D. melanogaster* eIF4E-1), which is exchanged to Lys in *A. pisum* eIF4E-1.

Several evolutionary forces could account for the multiplicity of eIF4E genes in *Drosophila* genus, as opposed to the other insect lineages containing only one *eIF4E* gene. Diptera experienced three episodes of explosive radiation, one of them happened during the emergence of Schizophora (close relatives of *D. melanogaster*) in the early Tertiary Period (65 MYA). The Schizophora radiation originated most of the family-level diversity in Diptera, accounting for more than a third of extant fly diversity [2, 51–53]. Interestingly, the temporal pattern of fruit flies speciation corresponds with the major periods of climate cooling and habitat fragmentation during the Cenozoic Era, which could be one of the causes for stimulating the rapid fruit flies speciation [52]. The vigorous burst of diversification of the Schizophora was also coincident with the emergence of some developmental novelties, including the ptilinal sac, an improved escape mechanism for the fly from its puparium [53]. Since flies originated in wet environments, it has been suggested that the emergence of an impervious pupation to their surrounding allowed flies to adapt to almost all substrates and to occupy a broad range of trophic niches [53]. The explosive diversification of schizophoran could have induced the repeated events of *eIF4E* duplication in *Drosophila* species. It is conceivable that specific modes of temporal and spatial regulation of protein synthesis driven by different eIF4E isoforms conferred an adaptive advantage to these environmental changes.

At the molecular level, genomic studies revealed that repeated tandem gene duplication has generated ~80% of the nascent genes during the *D. melanogaster* subgroup

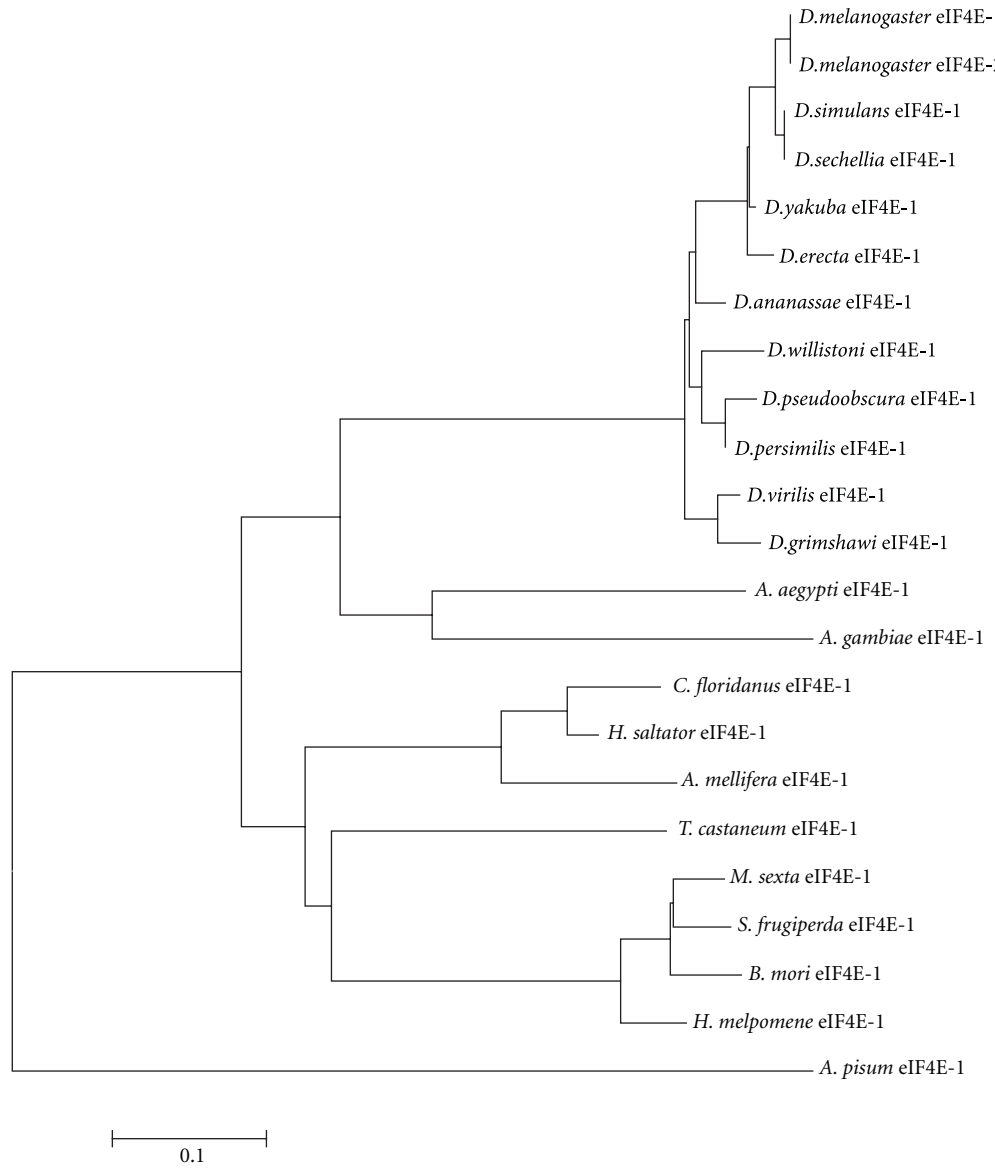


Figure 10: Phylogram computed from sequence alignments of eIF4E-1 from diverse insect species.

evolution, and that retroposition has generated ~10% of the new genes in these species [54, 55]. Five to eleven new functional genes per million years were originated during evolution of this lineage [54, 55]. These findings may explain that *D. melanogaster eIF4E-1/2*, *eIF4E-3*, *eIF4E-4,* and *eIF4E-5* genes lie within a narrow region of the chromosome 3L and share exon/intron genomic structure [23, 24, 26]. Thus, it is conceivable that these genes originated by tandem duplication of an original *eIF4E-1* gene. On the other hand, *eIF4E-6* and *eIF4E-7* genes, which lie in different chromosomes and contain no introns in the core region of the genes [26], could have originated by retroposition events from *eIF4E-3* and *eIF4E-1*, respectively. Noteworthy, *D. mojavensis* only encodes eIF4E-4, -5, and -7, but not eIF4E-1. Since eIF4E-7 appears to be an extended eIF4E-1, we speculate that eIF4E-7 functions for eIF4E-1 in this species, which at a certain point of evolution lost the original *eIF4E-1* gene. When available in the near future, the chromosomic location of *D. mojavensis eIF4E-7* gene could corroborate this hypothesis.

3.3. 4E-HP in the Genus Drosophila. We also analyzed Class II eIF4E, namely, 4E-HP, in species of the genus *Drosophila*. In a striking contrast to all eIF4Es, a single copy of the *4E-HP* gene was identified in each *Drosophila* species. Interestingly, 4E-HP displays an unusually strong conservation in the N-terminal moiety of the protein and residues important for eIF4G/4E-BP binding diverge considerably from eIF4E-1 in all *Drosophila* species (Figure 11). This is the case of Asn46, Gln82, Glu139, Asn140, and Met143 (positions refer to *D. melanogaster* 4E-HP), which are His, Glu, Leu,

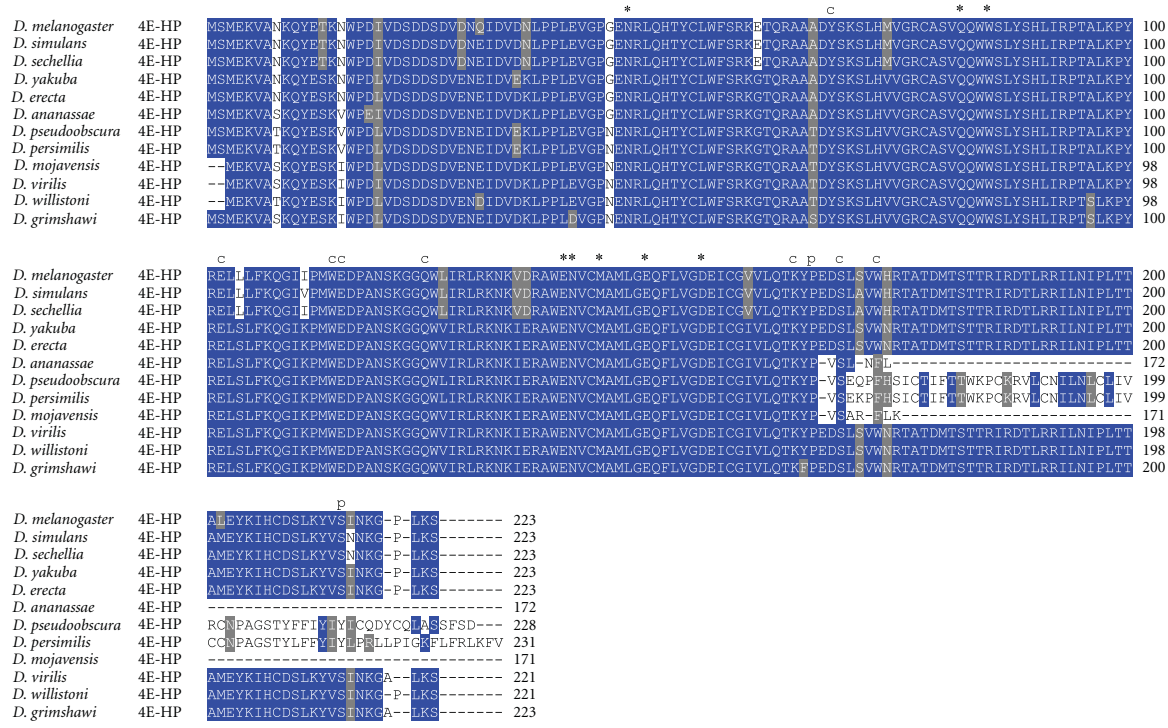


Figure 11: ClustalW alignment of amino acid sequences of 4E-HP orthologs from species of the genus *Drosophila*.

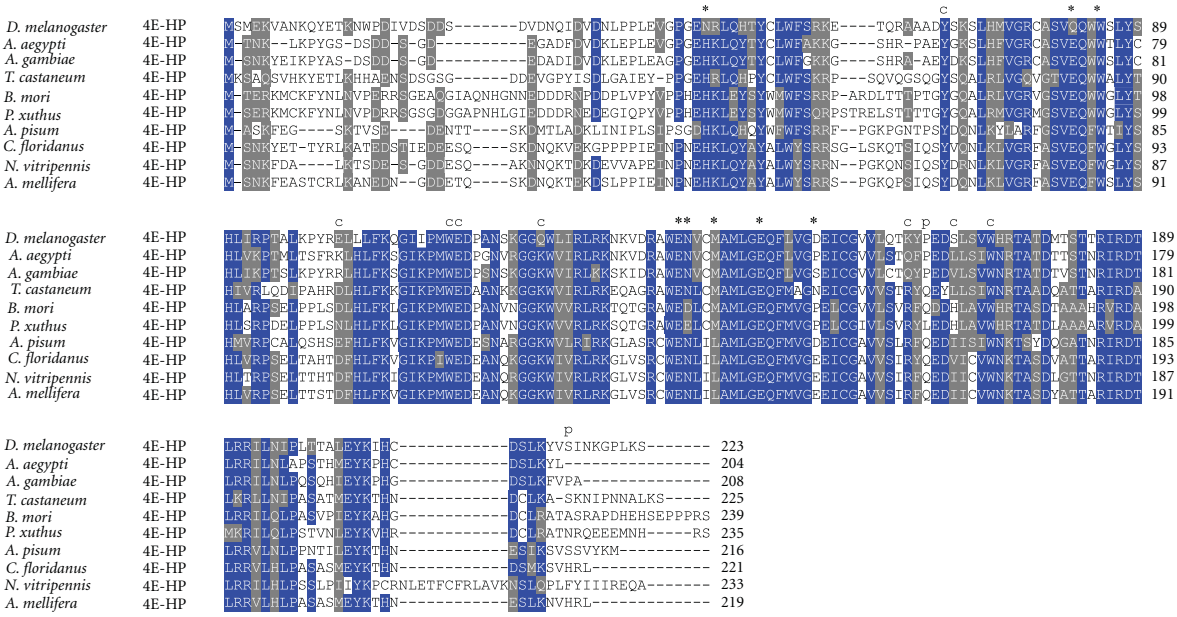


Figure 12: Orthologs of *D. melanogaster* 4E-HP in diverse insect species. ClustalW alignment of amino acid sequences of 4E-HP orthologs from diverse insect species.

Asp, and Leu residues in most *D. melanogaster* eIF4Es, respectively. Accordingly, *D. melanogaster* 4E-HP does not bind eIF4G [26] but it interacts with Bicoid (bcd) and Brain Tumor (Brat) instead [27, 33]. Many residues critical for cap binding underwent both conservative and nonconservative mutations in 4E-HP from all *Drosophila* species analyzed. Thus, Tyr68, Glu102, Gln124, Lys164, Pro166, and Ser169 are Trp, Asp, Arg, Arg, Lys, and Lys in all *D. melanogaster* eIF4Es, respectively. The counterpart of phosphorylatable Ser251 in *D. melanogaster* eIF4E-1 is conserved in most species of the genus *Drosophila*. Finally, *D. ananassae* and *D. mojavensis* 4E-HP is considerably shorter than 4E-HP in other species (Figure 11).

3.4. 4E-HP in Other Insects. Further BLAST searches identified again single-copy *4E-HP* genes in other insect species.

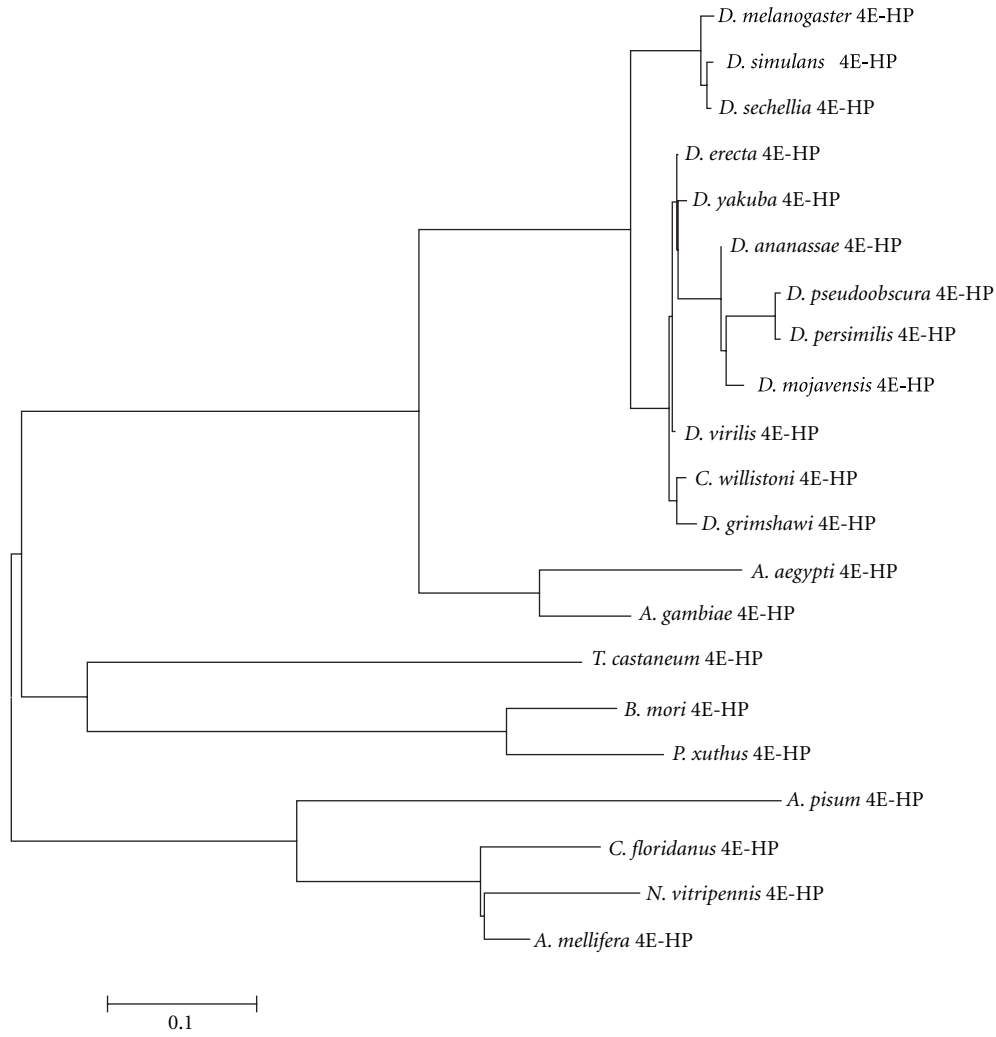


Figure 13: Phylogram computed from sequence alignments of 4E-HP from diverse insect species.

Sequence comparison showed a strong conservation in the core region of the protein, albeit N- and C-terminus are less conserved (Figure 12). In contrast to 4E-HP from *Drosophila* species, all residues important for eIF4G/4E-BP binding in eIF4Es are conserved in 4E-HP from all analyzed insects outside the genus *Drosophila*. This might suggest that 4E-HP in non-*Drosophila* insects do bind eIF4G/4E-BP. Similar to 4E-HP from all *Drosophila* species, most residues critical for cap-binding also show conservative changes in 4E-HP from Insecta species. The counterpart of phosphorylatable Ser251 of eIF4E-1 is only conserved in 4E-HP from *D. melanogaster*, *T. castaneum,* and *A. pisum* (Figure 12). A phylogram showing the relationships among 4E-HPs from all insects analyzed is shown in Figure 13.

Phylograms construction including all *Drosophila* 4E-HP and eIF4E sequences showed that all 4E-HPs cluster separately from all eIF4Es (not shown). Moreover, 4E-HP is widespread across metazoa, plants, and some fungi [45], and the *D. melanogaster* and human 4E-HP are able to bind the 5′ cap structure of the mRNA but not eIF4G [26, 56], thereby acting as a translational repressor of mRNAs associated to 4E-HP [27, 33, 34]. This, together with the findings that the *A. thaliana* [57] (termed nCBP) and the *C. elegans* [58] (termed IFE-4) orthologs can compete with reticulocyte eIF4E to reduce m^7GTP binding and can be found associated with small ribosomal subunits, respectively, which is consistent with a regulatory function, led to the suggestion that 4E-HP diverged from a widespread ancestral Class I eIF4E into a translational repressor in mammals and in *Drosophila* [59]. This is supported by the observation that all residues important for eIF4G/4E-BP binding in eIF4Es are highly conserved in 4E-HP from non-*Drosophila* insects, but not in *Drosophila* species (Figure 12). Thus, 4E-HP from insects outside the genus *Drosophila* should bind eIF4G and promote translation. It is important to experimentally analyze this controversial hypothesis.

3.5. Class III eIF4Es. Among insects, only two partial Class III eIF4Es were identified, one in *A. mellifera* and one in *H. coagulata*. Both are missing the start methionine and were therefore not further analyzed.

4. Concluding Remarks

Constant updating of genomic data and annotations as well as improved search algorithms provided a more comprehensive overview of insect eIF4E cognates than previously possible. Here we presented an updated analysis of eIF4Es and 4E-HP across Insecta. This analysis revealed an interesting observation, that is, that *eIF4E* is a single-copy gene in all insects analyzed, but in the genus *Drosophila* this gene underwent a striking multiplication along with the explosive radiation this lineage went through in the early Tertiary. eIF4E diversification led to variability of biochemical properties and physiological specialization, as documented for some *D. melanogaster* eIF4Es. It would be worthy to investigate whether this is also the case for other species with several eIF4E cognates, as sequence alignments showed how diverse this protein is in the genus *Drosophila*. It also would be interesting to search for novel, so far unknown, 4E-BPs in other *Drosophila* species. Moreover, it is possible that different eIF4Es could translate specific target mRNAs.

eIF4E from more insect species must be analyzed to obtain a better picture of the evolution and diversity of eIF4E in this group, and to see whether the rise of multiple *eIF4E* genes is found in other insect lineages too. If so, correlating eIF4E evolution with the natural history of those lineages might lead us to find general, underlying forces driving the translation apparatus evolution.

Acknowledgments

G. Hernández is supported by the National Institute for Cancer (Instituto Nacional de Cancerología, México). G. Tettweiler, P. Lasko, and N. Sonenberg were supported by a grant to P. Lasko and N. Sonenberg from the Canadian Institute of Health Research.

References

[1] P. J. Mayhew, "Why are there so many insect species? Perspectives from fossils and phylogenies," *Biological Reviews*, vol. 82, no. 3, pp. 425–454, 2007.

[2] D. A. Grimaldi and M. S. Engel, *Evolution of the Insects*, Cambridge University Press, New York, NY, USA, 2005.

[3] F. A. Bisby, Y. R. Roskov, T. M. Orrell, D. Nicolson, and L. E. Paglinawan, "Catalogue of life: 2010 annual checklist," *Species 2000*, vol. 23, no. 24, pp. 33–54, 2010.

[4] C. Mora, D. P. Titterson, S. Adl, A. G. B. Simpson, and B. Worm, "How many species are there on Earth and in the ocean," *PLoS Biology*, vol. 9, no. 8, Article ID e100112, 2011.

[5] M. S. Engel and D. A. Grimaldi, "New light shed on the oldest insect," *Nature*, vol. 427, no. 6975, pp. 627–630, 2004.

[6] M. W. Gaunt and M. A. Miles, "An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks," *Molecular Biology and Evolution*, vol. 19, no. 5, pp. 748–761, 2002.

[7] J. B. Whitfield and K. M. Kjer, "Ancient rapid radiations of insects: challenges for phylogenetic analysis," *Annual Review of Entomology*, vol. 53, pp. 449–472, 2008.

[8] E. A. Jarzembowski and A. J. Ross, "Insect origination and extinction in the Phanerozoic," in *Biotic Recovery from Mass Extinction Events*, M. B. Hart, Ed., pp. 65–78, Geological Society Special Publication, London, UK, 1996.

[9] P. G. Kevan, W. G. Chaloner, and D. B. O. Savile, "Interrelationships of early terrestrial arthropods and plants," *Paleontology*, vol. 18, pp. 391–417, 1975.

[10] N. Sonenberg and A. G. Hinnebusch, "Regulation of Translation Initiation in Eukaryotes: mechanisms and Biological Targets," *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[11] R. J. Jackson, C. U. T. Hellen, and T. V. Pestova, "The mechanism of eukaryotic translation initiation and principles of its regulation," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 2, pp. 113–127, 2010.

[12] J. Marcotrigiano, A. C. Gingras, N. Sonenberg, and S. K. Burley, "Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP," *Cell*, vol. 89, no. 6, pp. 951–961, 1997.

[13] H. Matsuo, H. Li, A. M. McGuire et al., "Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein," *Nature Structural Biology*, vol. 4, no. 9, pp. 717–724, 1997.

[14] K. Tomoo, X. Shen, K. Okabe et al., "Crystal structures of 7-methylguanosine 5′-triphosphate (m7GTP)- and P1-7-methylguanosine-P3-adenosine-5′, 5′-triphosphate (m7G-pppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region," *Biochemical Journal*, vol. 362, no. 3, pp. 539–544, 2002.

[15] K. Tomoo, X. Shen, K. Okabe et al., "Structural features of human initiation factor 4E, studied by X-ray crystal analyses and molecular dynamics simulations," *Journal of Molecular Biology*, vol. 328, no. 2, pp. 365–383, 2003.

[16] J. D. Richter and N. Sonenberg, "Regulation of cap-dependent translation by eIF4E inhibitory proteins," *Nature*, vol. 433, no. 7025, pp. 477–480, 2005.

[17] N. Sonenberg and A. G. Hinnebusch, "New modes of translational control in development, behavior, and disease," *Molecular Cell*, vol. 28, no. 5, pp. 721–729, 2007.

[18] L. Furic, L. Rong, O. Larsson et al., "EIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 32, pp. 14134–14139, 2010.

[19] P. E. D. Lachance, M. Miron, B. Raught, N. Sonenberg, and P. Lasko, "Phosphorylation of eukaryotic translation initiation factor 4E is critical for growth," *Molecular and Cellular Biology*, vol. 22, no. 6, pp. 1656–1663, 2002.

[20] G. Ross, J. R. Dyer, V. F. Castellucci, and W. S. Sossin, "Mnk is a negative regulator of cap-dependent translation in Aplysia neurons," *Journal of Neurochemistry*, vol. 97, no. 1, pp. 79–91, 2006.

[21] F. G. Maroto and J. M. Sierra, "Purification and characterization of mRNA cap-binding protein from *Drosophila melanogaster* embryos," *Molecular and Cellular Biology*, vol. 9, no. 5, pp. 2181–2190, 1989.

[22] G. Hernández and J. M. Sierra, "Translation initiation factor eIF-4E from *Drosophila*: cDNA sequence and expression of the gene," *Biochimica et Biophysica Acta*, vol. 1261, no. 3, pp. 427–431, 1995.

[23] G. Hernández, R. Diez Del Corral, J. Santoyo, S. Campuzano, and J. M. Sierra, "Localization, structure and expression of the gene for translation initiation factor eIF-4E from *Drosophila melanogaster*," *Molecular and General Genetics*, vol. 253, no. 5, pp. 624–633, 1997.

[24] C. A. Lavoie, P. E. D. Lachance, N. Sonenberg, and P. Lasko, "Alternatively spliced transcripts from the *Drosophila* eIF4E

gene produce two different cap-binding proteins," *The Journal of Biological Chemistry*, vol. 271, no. 27, pp. 16393–16398, 1996.

[25] P. Lasko, "The *Drosophila melanogaster* genome: translation factors and RNA binding proteins," *Journal of Cell Biology*, vol. 150, no. 2, pp. F51–F56, 2000.

[26] G. Hernández, M. Altmann, J. M. Sierra et al., "Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in *Drosophila*," *Mechanisms of Development*, vol. 122, no. 4, pp. 529–543, 2005.

[27] P. F. Cho, F. Poulin, Y. A. Cho-Park et al., "A new paradigm for translational control: inhibition via 5′-3′ mRNA tethering by Bicoid and the eIF4E cognate 4EHP," *Cell*, vol. 121, no. 3, pp. 411–423, 2005.

[28] G. Hernández, P. Vázquez-Pianzola, J. M. Sierra, and R. Rivera-Pomar, "Internal ribosome entry site drives cap-independent translation of reaper and heat shock protein 70 mRNAs in *Drosophila* embryos," *RNA*, vol. 10, no. 11, pp. 1783–1797, 2004.

[29] H. McNeill, G. M. Craig, and J. M. Bateman, "Regulation of neurogenesis and epidermal growth factor receptor signaling by the insulin receptor/target of rapamycin pathway in *Drosophila*," *Genetics*, vol. 179, no. 2, pp. 843–853, 2008.

[30] L. Gong, M. Puri, M. Ünlü et al., "*Drosophila* ventral furrow morphogenesis: a proteomic analysis," *Development*, vol. 131, no. 3, pp. 643–656, 2004.

[31] G. Hernández, H. Han, V. Gandin, T. Ferreira, N. Sonenberg, and P. Lasko, "Drosophila eukaryotic initiation factor 4E-3 is essential in post-meiotic stages of spermatogenesis," *Development*, In press.

[32] S. M. Gorski, S. Chittaranjan, E. D. Pleasance et al., "A SAGE approach to discovery of genes involved in autophagic cell death," *Current Biology*, vol. 13, no. 4, pp. 358–363, 2003.

[33] P. F. Cho, C. Gamberi, Y. Cho-Park, I. B. Cho-Park, P. Lasko, and N. Sonenberg, "Cap-dependent translational inhibition establishes two opposing morphogen gradients in *Drosophila* embryos," *Current Biology*, vol. 16, no. 20, pp. 2035–2041, 2006.

[34] J. C. Villaescusa, C. Buratti, D. Penkov et al., "Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation," *PLoS ONE*, vol. 4, no. 4, Article ID e5213, 2009.

[35] K. P. Menon, S. Sanyal, Y. Habara et al., "The translational repressor Pumilio regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E," *Neuron*, vol. 44, no. 4, pp. 663–676, 2004.

[36] S. J. Sigrist, P. R. Thiel, D. F. Reiff, P. E. D. Lachance, P. Lasko, and C. M. Schuster, "Postsynaptic translation affects the efficacy and morphology of neuromuscular junctions," *Nature*, vol. 405, no. 6790, pp. 1062–1065, 2000.

[37] P. L. Graham, J. L. Yanowitz, J. K. M. Penn, G. Deshpande, and P. Schedl, "The translation initiation factor eif4e regulates the Sex-Specific expression of the master switch gene Sxl in *Drosophila melanogaster*," *PLoS Genetics*, vol. 7, no. 7, Article ID e1002185, 2011.

[38] M. Miron, J. Verdú, P. E. D. Lachance, M. J. Birnbaum, P. F. Lasko, and N. Sonenberg, "The translational inhibitor 4E-BP is an effector of PI(3)K/Akt signalling and cell growth in *Drosophila*," *Nature Cell Biology*, vol. 3, no. 6, pp. 596–601, 2001.

[39] F. Piccioni, V. Zappavigna, and A. C. Verrotti, "A cup full of functions," *RNA Biology*, vol. 2, no. 4, pp. 125–128, 2005.

[40] S. K. Lee, J. S. Lee, K. S. Shin, and S. J. Yoo, "Translation initiation factor 4E (eIF4E) is regulated by cell death inhibitor, Diap 1," *Molecules and Cells*, vol. 24, no. 3, pp. 445–451, 2007.

[41] S. Iwasaki, T. Kawamata, and Y. Tomari, "*Drosophila* argonaute1 and argonaute2 employ distinct mechanisms for translational repression," *Molecular Cell*, vol. 34, no. 1, pp. 58–67, 2009.

[42] A. Yarunin, R. E. Harris, M. P. Ashe, and H. L. Ashe, "Patterning of the *Drosophila* oocyte by a sequential translation repression program involving the d4EHP and Belle translational repressors," *RNA Biology*, vol. 8, no. 5, pp. 904–912, 2011.

[43] A. G. Clark, M. B. Eisen, D. R. Smith et al., "Evolution of genes and genomes on the *Drosophila* phylogeny," *Nature*, vol. 450, no. 7167, pp. 203–218, 2007.

[44] A. Papanicolaou, S. Gebauer-Jung, M. L. Blaxter, W. Owen McMillan, and C. D. Jiggins, "ButterflyBase: a platform for lepidopteran genomics," *Nucleic Acids Research*, vol. 36, no. 1, supplement, pp. D582–D587, 2008.

[45] B. Joshi, K. Lee, D. L. Maeder, and R. Jagus, "Phylogenetic analysis of eIF4E-family members," *BMC Evolutionary Biology*, vol. 5, article 48, 2005.

[46] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice," *Nucleic Acids Research*, vol. 22, no. 22, pp. 4673–4680, 1994.

[47] D. G. Higgins, J. D. Thompson, and T. J. Gibson, "Using CLUSTAL for multiple sequence alignments," *Methods in Enzymology*, vol. 266, pp. 383–400, 1996.

[48] K. Tamura et al., "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods," *Molecular Biology and Evolution*, vol. 28, no. 10, pp. 2731–2739, 2011.

[49] R. E. Rhoads, T. D. Dinkova, and R. Jagus, "Approaches for analyzing the differential activities and functions of eIF4E family members," *Methods in Enzymology*, vol. 429, pp. 261–297, 2007.

[50] M. D. Adams, S. E. Celniker, R. A. Holt et al., "The genome sequence of *Drosophila melanogaster*," *Science*, vol. 287, no. 5461, pp. 2185–2195, 2000.

[51] D. K. Yeates, B. M. Wiegmann, G. W. Courtney, R. Meier, C. Lambkin, and T. Pape, "Phylogeny and systematics of Diptera: two decades of progress and prospects," *Zootaxa*, no. 1668, pp. 565–590, 2007.

[52] K. Tamura, S. Subramanian, and S. Kumar, "Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks," *Molecular Biology and Evolution*, vol. 21, no. 1, pp. 36–44, 2004.

[53] B. M. Wiegmann, M. D. Trautwein, I. S. Winkler et al., "Episodic radiations in the fly tree of life," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 14, pp. 5690–5695, 2011.

[54] Q. Zhou and W. Wang, "On the origin and evolution of new genes-a genomic and experimental perspective," *Journal of Genetics and Genomics*, vol. 35, no. 11, pp. 639–648, 2008.

[55] Q. Zhou, G. Zhang, Y. Zhang et al., "On the origin of new genes in *Drosophila*," *Genome Research*, vol. 18, no. 9, pp. 1446–1455, 2008.

[56] E. Rom et al., "Cloning and characterization of 4E-HP, a novel mammalian eIF4E-related cap-binding protein," *The Journal of Biological Chemistry*, vol. 273, no. 21, pp. 13104–13109, 1998.

[57] K. A. Ruud, C. Kuhlow, D. J. Goss, and K. S. Browning, "Identification and characterization of a novel cap-binding protein from *Arabidopsis thaliana*," *The Journal of Biological Chemistry*, vol. 273, no. 17, pp. 10325–10330, 1998.

[58] T. D. Dinkova, B. D. Keiper, N. L. Korneeva, E. J. Aamodt, and R. E. Rhoads, "Translation of a small subset of *Caenorhabditis elegans* mRNAs is dependent on a specific eukaryotic translation initiation factor 4E isoform," *Molecular and Cellular Biology*, vol. 25, no. 1, pp. 100–113, 2005.

[59] G. Hernández, M. Altmann, and P. Lasko, "Origins and evolution of the mechanisms regulating translation initiation in eukaryotes," *Trends in Biochemical Sciences*, vol. 35, no. 2, pp. 63–73, 2010.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 475731, 8 pages
doi:10.1155/2012/475731

Review Article

Before It Gets Started: Regulating Translation at the 5′ UTR

Patricia R. Araujo,[1] Kihoon Yoon,[1,2] Daijin Ko,[3] Andrew D. Smith,[4] Mei Qiao,[1] Uthra Suresh,[1] Suzanne C. Burns,[1] and Luiz O. F. Penalva[1,5]

[1] *Greehey Children's Cancer Research Institute, UTHSCSA, San Antonio, TX 78229-3900, USA*
[2] *Department of Epidemiology and Biostatistics, UTHSCSA, San Antonio, TX 78229-3900, USA*
[3] *Department of Management Science and Statistics, UTSA, San Antonio, TX 78249-0631, USA*
[4] *Molecular and Computational Biology, Department of Biological Sciences, USC, Los Angeles, CA 90089-2910, USA*
[5] *Department of Cellular and Structural Biology, UTHSCSA, San Antonio, TX 78229-3900, USA*

Correspondence should be addressed to Luiz O. F. Penalva, penalva@uthscsa.edu

Received 2 January 2012; Revised 22 February 2012; Accepted 11 March 2012

Academic Editor: Thomas Preiss

Translation regulation plays important roles in both normal physiological conditions and diseases states. This regulation requires cis-regulatory elements located mostly in 5′ and 3′ UTRs and trans-regulatory factors (e.g., RNA binding proteins (RBPs)) which recognize specific RNA features and interact with the translation machinery to modulate its activity. In this paper, we discuss important aspects of 5′ UTR-mediated regulation by providing an overview of the characteristics and the function of the main elements present in this region, like uORF (upstream open reading frame), secondary structures, and RBPs binding motifs and different mechanisms of translation regulation and the impact they have on gene expression and human health when deregulated.

1. Translation Regulation

Gene expression can be modulated at multiple levels from chromatin modification to mRNA translation. Despite the importance of transcriptional regulation, it is clear at this point that mRNA levels cannot be used as a sole parameter to justify the protein content of a cell. In fact, in a recent study from our lab, we determined that a direct correlation between mRNA and protein exists for less than a third of analyzed genes in a human cell line. Moreover, our analysis suggested that translation regulation contributes considerably to the protein variation as several parameters related to translation like 5′ UTR, 3′ UTR, coding sequence length, presence of uORFs and amino acid composition, and so forth showed good correlations with the obtained mRNA/protein ratios [1]. Translation regulation functions as an important switch when rapid changes in gene expression are required in reponse to internal and external stimuli (*PDGF2, VEGF, TGFβ* are examples of genes controlled in such way). Translation regulation also plays a significant role during development and cell differentiation by altering the levels of expression of specific mRNA subsets during a particular time window while the majority of transcripts remain unchanged (reviewed in [2–4]).

In this paper, we will focus on the importance of 5′ UTR mediated regulation and the different functional elements present in this region with the exception of IRES which is discussed in a different article of this issue. The main regulatory elements in 5′ UTR are secondary structures (including IRES), binding sites for RNA binding proteins, uAUGs and uORFs (Figure 1).

2. 5′ UTR

The average length of 5′ UTRs is ~100 to ~220 nucleotides across species [5]. In vertebrates, 5′ UTRs tend to be longer in transcripts encoding transcription factors, protooncogenes, growth factors and their receptors, and proteins that are poorly translated under normal conditions [6]. High GC content is also a conserved feature, with values surpassing 60% in the case of warm-blooded vertebrates. In the context of hairpin structures, GC content can affect protein translation efficiency independent of hairpin thermal stability and

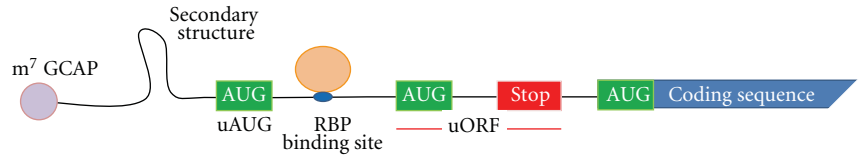


Figure 1: Regulatory elements present in 5′ UTR.

hairpin position [7]. UTRs of eukaryotic mRNAs also display a variety of repeats that include short and long interspersed elements (SINEs and LINEs, resp.), simple sequence repeats (SSRs), minisatellites, and macrosatellites [5].

Translation initiation in eukaryotes requires the recruitment of ribosomal subunits at either the 5′ m7G cap structure. The initiation codon is generally located far downstream, requiring ribosomal movement to this site. This movement appears to be nonlinear for some mRNAs (i.e., ribosomal subunits appear to bypass (shunt) segments of the 5′ UTR as they move in the direction of the AUG). Shunting could allow mRNAs containing uAUGs or hairpin structures to be translated efficiently. Important examples are provided by the cauliflower mosaic virus [8] and adenovirus [9] mRNAs. The mechanism of ribosomal shunting is rather complex requiring mRNA-rRNA base pairing [10].

Genes presenting differences in the 5′ UTR of their transcripts are relatively common. 10–18% of genes express alternative 5′ UTR by using multiple promoters [11, 12] while alternative splicing within UTRs is estimated to affect 13% of genes in the mammalian transcriptome [13]. These variations in 5′ UTR can function as important switches to regulate gene expression. Two important examples are provided by the cancer-related genes *BRCA1* (breast cancer 1) and *TGF-β* (transforming growth factor *β*). BRCA1 is a tumor suppressor, frequently mutated in breast cancer with functions in cell cycle, apoptosis, and DNA damage repair. BRAC1 produces two different transcripts that derive from two different promoters and therefore display differences in their 5′ UTR. A shorter transcript is expressed in cancerous as well as noncancerous breast tissue and efficiently translated, while a longer transcript is predominantly expressed in breast cancers. The presence of several uAUGs and a more complex structure dramatically affect the translation of this longer transcript. This causes an overall decrease in BRAC1 levels in tumor cells, leading to a relief in growth inhibition [14]. *TGF-β* is implicated in a large number of processes that include cell proliferation, migration, wound repair, development, tumorigenesis and immunosuppression. There are three known isoforms: *β*1, *β*2, and *β*3. *TGF-β3* produces two alternative transcripts: a 3.5 kb transcript with a very long 5′ UTR (1.1 kb) and a 2.6 kb transcript with a shorter 5′ UTR (0.23 kb). The presence of 11 uORFs in the longer transcript dramatically inhibits its translation while the shorter transcript is efficiently translated [15, 16].

3. Regulation by Secondary Structure

Secondary structures can function as major regulatory tools in 5′ UTRs. A correlation with gene function has been suggested; secondary structures have been determined to be particularly prevalent among mRNAs encoding transcription factors, protooncogenes, growth factors, and their receptors and proteins poorly translated under normal conditions. >90% of transcripts in these classes have 5′ UTRs containing stable secondary structures with average free energies less than −50 kcal/mol. 60% of these stable secondary structures are positioned very close to the cap structure [6]. These structures are very effective in inhibiting translation. In fact, a hairpin situated close to the cap with a free energy of −30 kcal/mol would be sufficient to block the access of the preinitiation complex to the mRNA. When located further away in the 5′ UTR, hairpins require a free energy stronger than −50 kcal/mol to be able to block translation [17, 18]. Stable secondary structure can resist the unwinding activity of the helicase elF4A. This effect can be overcome partially by the overexpression of elF4A in partnership with elF4B [19]. mRNAs with a highly structured 5′ UTR like proto-oncogenes and other growth factors use cap-dependent translation initiation. Not surprisingly, the overexpression of components of the translation initiation machinery including elf4E has been linked to tumorigenesis (reviewed in [18, 20]).

The gene *TGF-β1* provides a good example of translation inhibition mediated by secondary structure [21, 22]. An evolutionary conserved motif in the 5′ UTR forms a stable stem loop. However, this structure by itself is not sufficient to block translation. Translation repression of *TGF-β1* depends on increased binding of the RNA binding protein YB-1 to the *TGF-β1* transcript [23]. It was then proposed that YB-1 binds the 5′ UTR of *TGF-β1* with high affinity thanks to its GC content and cooperates with the stem loop to inhibit *TGF-β1* translation by facilitating duplex formation [24].

4. Regulation by RNA Binding Proteins

The human genome is predicted to encode circa 1,000 RNA binding proteins (RBPs) with a large percentage of them implicated in translation. They could be categorized into two main groups: RBPs that are part of the basic translation machinery and required for the translation of all expressed mRNAs (examples: PABPI, elf4E) and RBPs that function in a more selective way by controlling either positively or negatively the levels of translation of specific target mRNAs (examples: HuR, Musashi1). Regarding this later group, it has been observed that RBPs can use distinct mechanisms to increase or inhibit translation. Although several exceptions are known, it can be said that RBPs often recognize specific motifs in UTRs and interact with the translation machinery

to control expression. Interference with translation normally takes place during the initiation step (reviewed in [25]).

The best characterized example of RBP-mediated regulation involving 5′ UTRs is provided by the iron regulatory proteins (IRP 1 and 2). These proteins recognize a highly conserved stem loop structure with circa 30 nucleotides, known as the iron response element (IRE). The most important features include a hexanucleotide loop with the sequence CAGYCX (Y = U or A; X = U, C, or A) and a 5 bp upper stem that is separated from a lower stem of variable length by an unpaired cytosine. This regulation is crucial in maintaining cellular iron homeostasis as a large number of mRNAs connected to iron storage and metabolism including ferritin, mitochondrial aconitase, succinate dehydrogenase-iron protein, erythroid 5-aminolevulinate synthetase (eALAS), and an iron-exportin molecule named ferroportin (FPN1) have their expression modulated by this system. When cellular iron levels are low, IRP1 and IRP2 bind the IRE and block translation of the downstream ORF. When intracellular iron levels are high, the RNA binding activity of both IRPs is reduced (Figure 2(a)). IREs tend to be positioned close to the cap, which causes a steric inhibition of the binding of 40S ribosomal subunits to the transcript. When located distant to the cap, rather than affecting 40S recruitment, the IRE-IRP complex blocks ribosomal scanning (reviewed in [26]). An interesting bypass of the IRE/IRP mechanism can be observed in iron-starved duodenal and erythroid precursor cells. An upstream promoter is used to generate FPN1 pre-mRNAs containing one more exon that is connected by alternative splicing to a splice acceptor in the 3′ of the IRE. A mature FPN1 transcript containing the same open reading frame is generated; however, the 5′ UTR does not contain the IRE [27]. Therefore, these cells express the alternative FPN1 isoform in an iron-independent manner [27, 28]. Mutations affecting IREs can lead to diseases. This is the case of hereditary hyperferritinemia-cataract syndrome (HHCS), a genetic autosomal dominant disorder in which aggregation and crystallization of ferritin in the lens leads to bilateral cataracts [29].

RBP-mediated regulation can be very elaborate and involve multiple steps. One good example showing the crosstalk between factors and distinct regulatory processes is the male-specific-lethal 2 (*msl-2*) gene in *Drosophila*, a main player in dosage compensation. The female-specific RNA binding protein sex lethal (SXL) participates in multiple aspects of *msl-2* regulation where *msl-2* expression must be prevented (Figure 2(b)). Regulation starts at the splicing level; SXL binds to two polyU stretches located in an intron that is part of the 5′ UTR. This process causes intron retention and preserves critical sequences that later will be used in translation regulation [30, 31]. In the cytoplasm, the same SXL protein will function as a translation repressor of *msl-2* in two distinct mechanisms taking place at the 3′ and 5′ UTR [32]. SXL binds U-rich sequences in the 3′ UTR and recruits the corepressor protein UNR (upstream of N-ras) and PABP blocking the recruitment of the pre-initiation complex to the 5′ end of the mRNA [33–35]. To assure that *msl-2* gets fully repressed, a second regulatory step also mediated by SXL takes place at the 5′ UTR. This repression involves a novel regulatory mechanism where crosstalk between SXL and a uORF takes place to efficiently repress translation [36]. The 5′ UTR of *msl-2* contains 3 uORFs but only the 3rd one is involved in the repression. Interestingly, this repression is very weak in the absence of SXL (∼2-fold), but when present, SXL binds a poly U stretch a few nucleotides away from the uAUG and increases this repression to more than 14-fold. SXL acts by boosting translation initiation at the uAUG and not by acting as a simple steric arrest of scanning ribosomes. This effect may take place via an interaction between SXL and translation initiation factors; possibly members of elF3 component as indicated by a two-hybrid screening. This mechanism potentially affects a large number of mRNAs; 268 transcripts in *Drosophila* were determined to contain SXL binding motifs associated with uAUG spaced at an appropriate distance. For instance, a reporter construct containing the 5′UTR of the gene *Irr47* was repressed ∼4-fold by SXL protein [36].

RBPs can have antagonistic functions when regulating translation. An interesting example is the regulation of p21 in the context of replicative senescence, a cellular state where cells enter an irreversible growth arrest. Induction of p21 is required to initiate the process, and to inhibit cdk2-cyclin E complexes. The 5′ UTR of p21 contains a GC-rich sequence that forms a stem loop. This element is recognized by two RBPs with distinct properties: CUGBP1 and calreticulin (CRT). Competition between the two proteins determines final levels of p21 expression and establishes if cells will proliferate or undergo growth arrest and senescence. Binding of CUGBP1 to p21 mRNA is dramatically increased in senescence compared to young fibroblast cells. Protein levels do not change during the process and this increase in activity is due to phosphorylation. On the other hand, CRT IPs showed a four-to-fivefold reduction of activity in senescence cells due to a decrease in expression. Both proteins were shown to affect p21 translation. However, while CUGBP1 functions as an activator, CRT acts as a repressor. Since the two proteins have opposing activity in senescent cells, they were examined to see if they compete for interaction with p21 mRNA and to control its translation. Increasing amounts of one protein were able to reverse the binding of the other protein to p21 mRNA and its effect on translation; affinity to the binding site is rather different as CUGBP1 had to be present in the binding reactions at a four-to-eightfold molar excess to CRT to antagonize its binding to p21 mRNA and impact its translation [37].

5. Regulation by uORFs and Upstream AUGs

uORFs and uAUGs are major regulatory elements in 5′ UTRs. As their names suggest, uORFs are sequences defined by a start and stop codons upstream of the main coding region while uAUGs are start codons without an in-frame downstream stop codon located upstream of the main coding region. A large percentage of the human transcriptome contains uORF and/or uAUGs, with values ranging between 44 and 49% [38, 39]. Similar numbers are found in the

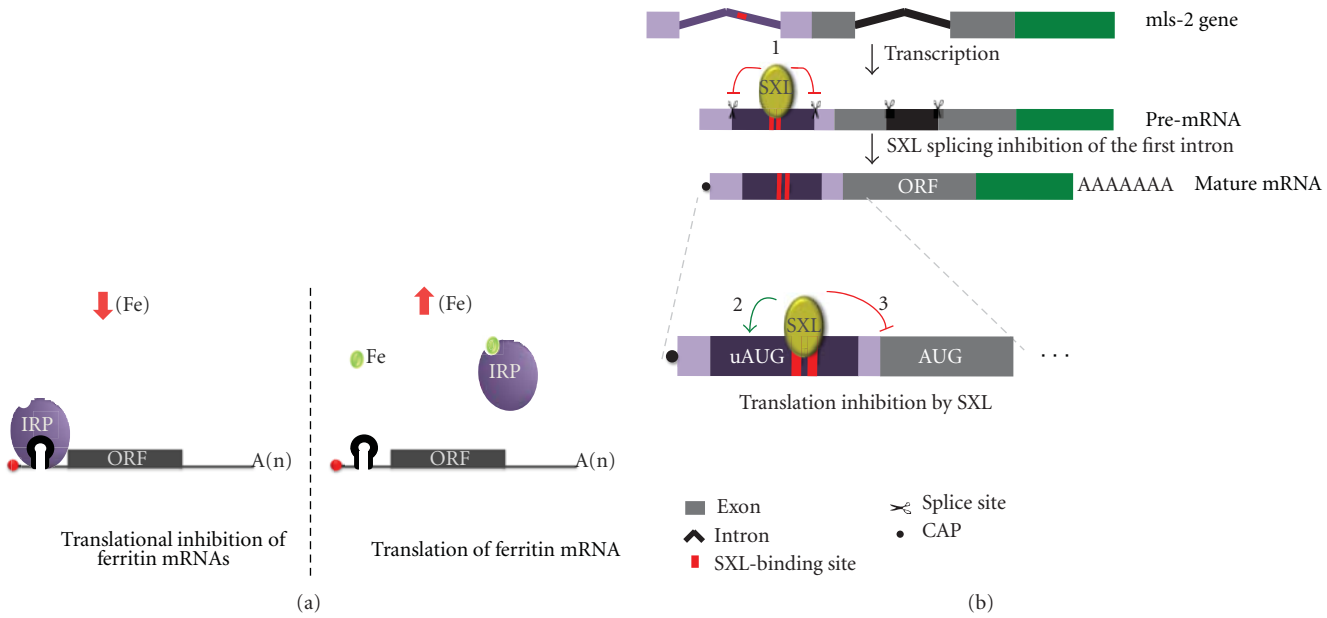


Figure 2: Translational regulation by RNA binding proteins. (a) In iron-deficient cells, IRPs bind to the IRE localized in the 5′ UTR of ferritin mRNA, blocking its translation. Once cellular iron levels increase, a complex containing Fe binds to IRPs. Thus, these proteins are allosterically modified, which reduces IRP-IRE binding and allows the translation of ferritin mRNAs. (b) *msl-2* gene regulation in females flies. After transcription in the nucleus, SXL specifically binds to intronic U-rich regions of msl-2 pre-mRNA and inhibits the intron removal (1). In the cytoplasm, SXL binds to the same elements localized now in the 5′ UTR of mature msl-2 mRNA, enhances the translation initiation of a upstream ORF (2), and prevents the main ORF translation (3). The regulatory elements in the 3′ UTR of msl-2 mRNA were not represented.

mouse transcriptome. Although these numbers might sound high, both uORFs and uAUGs are less frequent than expected by chance, suggesting that they are under selective pressure. uORFs and uAUG are overrepresented in particular subgroups like transcription factors, growth factors, and their receptors and proto-oncogenes [6]. Both uORFs and uAUGs are extremely diverse varying in position in relation to the cap and main AUG, number per transcript and length (in the case of uORFs) [38]. Supplementary Table 1 (in Supplementary Material available online at doi:10.1155/2012/475731) provides a comprehensive list of uORFs and uAUGs present in the human transcriptome. uORFs and uAUGs have not been extensively analyzed in terms of conservation. A pilot study done with a subset of human, mouse, and rat transcripts indicated that both elements are moderately conserved as 38% of uORFs and 24% of uAUGs were determined to be conserved among three species [39]. The modest conservation of uORFs combined with the fact that their average length (20 nucleotides) is expected by chance and uAUGs provide a stronger suppression in comparison to uORFs suggests that many uAUGs have been neutralized in the process of evolution by the acquisition of a downstream stop codon. It has been proposed then that only a few uORFs, very likely the conserved ones, have been recruited for expression regulation [39]. In yeast, it has been shown that uORFs are statistically underrepresented in 5′ UTRs and were removed by selective pressure, indicating similarly that the remaining uORFs may be implicated in translation regulation [40].

Although, overall it has been suggested that uORFs are negatively correlated with protein production [1, 38, 41] until now, functional activity has been demonstrated for only a limited number of uORFs and uAUGs. In Figure 3, we show examples of the impact uAUGs can have on translation efficiency. Among the most relevant features that can contribute to functionality are long 5′ cap-to-uORF distance, sequence conservation, context in which the AUG is located, strength of the initiation site for the ORF, length of the uORF, and number of AUGs in the 5′ UTR [38, 42]. Different outcomes have been observed when a ribosome encounters a uAUG or uORF [43]. Since the number of characterized events is still small, it is hard to define general mechanisms; we describe then a few well-characterized and relevant events. Leaky scanning is defined when a proportion of the scanning complexes bypass the uAUG or uORF and continue scanning for the next AUG. In this case, the upstream AUG acts as a "decoy" from the ORF AUG, functioning as a negative regulator of translation at least for some fraction of ribosomes. The production of cis-acting peptides by uORFs can reduce the initiation of translation of the downstream ORF by stalling the ribosome at the end of the uORF [44]. A classical example is provided by the evolutionarily conserved eukaryotic arginine attenuator peptide (AAP), that negatively controls the translation of proteins involved in the *de novo* fungal arginine biosynthesis in high arginine concentration [45]. In this scenario, arginine changes AAP conformation and/or P site environment causing ribosomal stalling at the termination codon of AAP uORF [46, 47]. AAP also reduces translation elongation

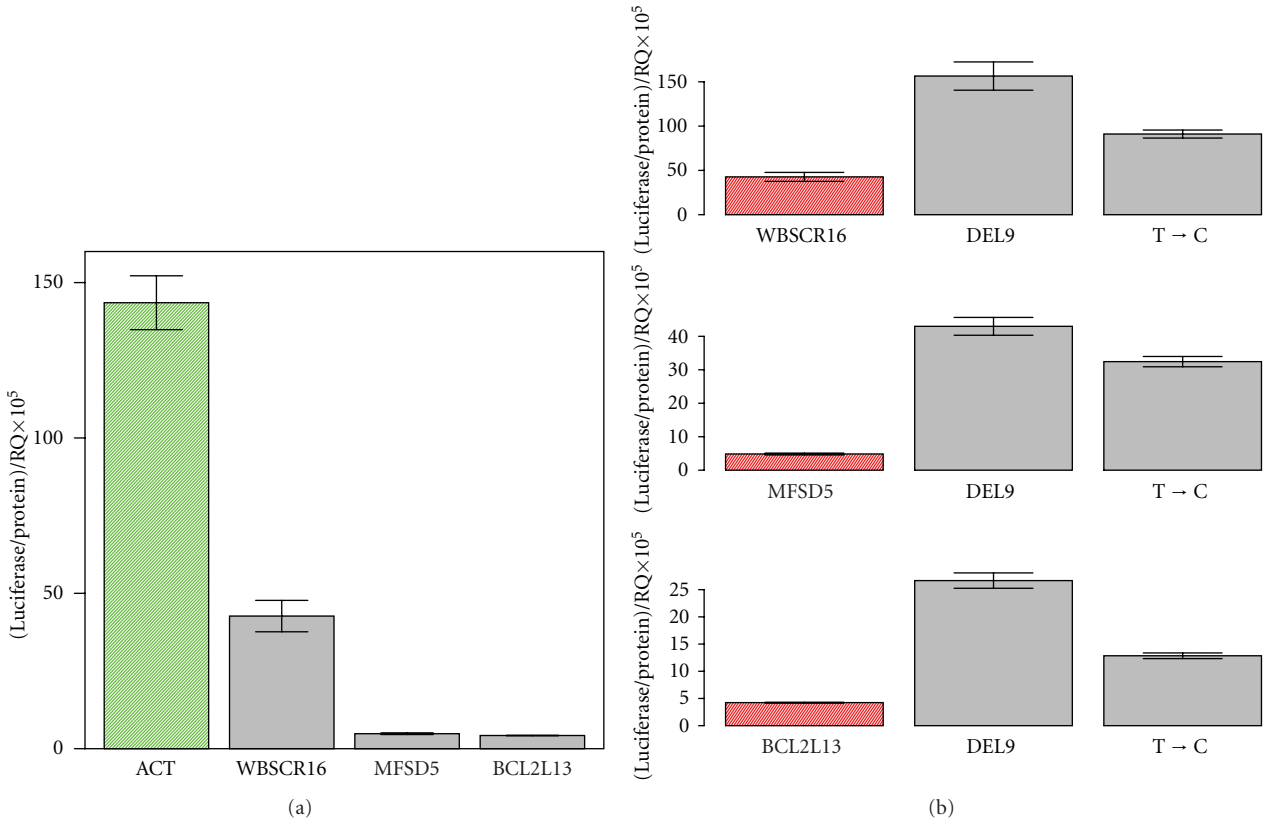


Figure 3: Impact of uAUG sequences on translation regulation. (a) Comparison of luciferase levels obtained for constructs having the 5′ UTR of the gene ACT (control) and genes containing uAUG: WBSCR16, MFSD5, and BCL2L13. (b) Deletion or mutation of uAUG sequence present in genes WBSCR16, MFSD5 and BCL2L13 reverts translation repression as seen as an increase in luciferase activity.

by ribosome stalling when the uORF is inserted within an encoding sequence [48]. Another classical example of uORF-mediated regulation comes from yeast. Four uORFs are present in the 5′ UTR of the transcription factor GCN4. The first of the four uORFs is always efficiently translated regardless of the nutritional conditions. In unperturbed cells, rapid reloading of ribosomes and initiation cofactors allow translation of uORFs 2–4 while inhibiting the translation of the main ORF. In situations of amino acid starvation, initiation factors are scarce, resulting in a decelerated reloading of ribosomes and scanning across the sequences containing the uORFs. A functional initiation complex is reassembled only at the main coding sequence and GCN4 expressed. This mechanism allows a fast response to nutritional stress [49, 50]. Another similar example of regulated expression via uORF is the Carnitine Palmitoyltransferase 1C (CPT1C) gene. CPT1C regulates metabolism in the brain in situations of energy surplus. The presence of uORF in the 5′ UTR represses the expression of the ORF. However, this repression is relieved in response to specific stress stimuli like glucose depravation and palmitate-BSA treatment [51]. It has been suggested that uORFs can also induce mRNA degradation. A series of 5′ UTR constructs containing as a reporter the cat gene from the bacterial transposon Tn9 was tested in yeast. A single nucleotide substitution was used to create a 7-codon ORF upstream of the cat gene. The uORF was translated efficiently and caused translation inhibition of the cat ORF and destabilization of the cat mRNA [52]. A connection between uORFs and mRNA decay was also suggested based on a comparison between average levels of expression of uORF-containing and non-uORF-containing transcripts [41].

Several mutations that eliminate or create uORFs that end up altering protein levels have been connected to human diseases. Their relevance was discussed recently [53]. Predisposition to melanoma can be caused by a mutation that introduces a uORF into the 5′ UTR of the gene cyclin-dependent-kinase inhibitor protein (CDKN2A) [54]. Hereditary thrombocythemia is caused by a mutation that creates a splicing variant that eliminates a uORF, leading to an increase in protein production of the gene thrombopoietin [55]. Marie Unna hereditary hypotrichosis derives from a mutation that disrupts a uORF present in the 5′ UTR of the gene hairless homolog and consequently increasing its expression [56]. A transition from G to A in one of the uORFs present in the 5′ UTR of TGF-β3 transcript was determined to be associated with arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC) [57]. Another group of five uORFs associated with diseases have been tested recently [58] using reporter assays; they include gonadal dysgenesis (*SRY*) [59], Van der Woude syndrome (*IRF6*) [60], Carney Complex Type 1 (*PRKAR1A*) [61], Hereditary pancreatitis (*SPINK1*) [62], and Thalassaemia-β (*HBB*) [63]. This list will certainly expand as more than

500 single-nucleotide polymorphisms (SNPs) creating or deleting uORFs have been reported.

6. Searching for Novel Regulatory Elements in the 5′ UTR

Only a small fraction of the posttranscriptional regulatory elements located in human 5′ UTRs have been characterized. Those identified UTR elements are catalogued in a web-resource maintained by Graziano Pesole's group called UTRdb (http://utrdb.ba.itb.cnr.it/) [49]. In vivo methods for the identification of posttranscriptional regulatory elements in UTRs, especially those associated with RBPs, have advanced dramatically in the last five years thanks to deep sequencing technology. CLIP and RIP-Seq are methods based on the isolation of RNA protein molecules (RNPs) via immunoprecipitation, followed by RNase digestion and precise identification of RBP binding sites with deep sequencing [64]. Although the number of RBPs analyzed so far by these methods is really small (reviewed in [65]), as deep sequencing technology becomes more accessible and the methods simplified, one could expect that very soon a large portion of the human RBP binding sites in UTRs will be mapped.

Another choice to map UTR elements regulating translation is to use purely computational methods based on analyzing the UTR sequences. These methods are based on identifying degenerate ribonucleotide patterns that have the expected properties of RBP binding sites. Similar methods have been applied for nearly 30 years to identify transcriptional regulatory in promoter sequences. These methods are reaching maturity, are very widely used, and have assisted greatly in compiling databases about transcriptional regulation (e.g., TRANSFAC) [66, 67]. Although much of the work directed towards designing and refining regulatory sequence analysis algorithms in context of transcriptional regulation can be adapted to corresponding analysis problems in the context of post-transcriptional regulatory elements, there are additional complications associated with RBP binding sites. The most obvious among these is that RBPs will have secondary structural preferences, and few existing analysis tools can incorporate information about RNA folding. Similarly, because of RNA folding regulatory elements can more easily function synergistically or display concerted binding to sequence elements that are distal in the primary sequence but very close in the folded molecule. Another difficulty is the lack of example translational regulatory elements for training the analysis. Based on a handful of well-studied examples, there is often a perception that RBP binding sites are on average shorter than transcription factors (TFs) binding sites, but this perception may be due to bias in the set of RBPs receiving the most research focus [65]. One of the most powerful methods for identifying regulatory elements is phylogenetic foot-printing, which takes advantage of locally elevated evolutionary conservation to reveal functional elements [5, 50, 51]. This logic works equally well for post-transcriptional regulatory elements. Unfortunately TF binding sites are also a major confound to direct application of computational sequence analysis for identifying 5′ UTR elements involved in translation. Elements involved in transcriptional regulation reside both up- and downstream of transcription start sites, and when 5′ UTRs are sufficiently short post-transcriptional regulatory elements are likely interleaved with TF binding sites.

Ultimately the best methods for identifying post-transcriptional regulatory elements will emerge from complementary application of experimental and computational techniques.

Acknowledgments

Research in Penalva's lab is supported by the Voelcker Foundation, Children's Brain Tumor foundation, and 5R21HG004664-02 and 1R01HG006015-01A1.

References

[1] C. Vogel, R. de Sousa Abreu, D. J. Ko et al., "Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line," *Molecular Systems Biology*, vol. 6, article 400, 2010.

[2] R. J. Jackson, C. U. T. Hellen, and T. V. Pestova, "The mechanism of eukaryotic translation initiation and principles of its regulation," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 2, pp. 113–127, 2010.

[3] N. Sonenberg and A. G. Hinnebusch, "Regulation of translation initiation in eukaryotes: mechanisms and biological targets," *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[4] F. Gebauer and M. W. Hentze, "Molecular mechanisms of translational control," *Nature Reviews Molecular Cell Biology*, vol. 5, no. 10, pp. 827–835, 2004.

[5] G. Pesole, F. Mignone, C. Gissi, G. Grillo, F. Licciulli, and S. Liuni, "Structural and functional features of eukaryotic mRNA untranslated regions," *Gene*, vol. 276, no. 1-2, pp. 73–81, 2001.

[6] R. V. Davuluri, Y. Suzuki, S. Sugano, and M. Q. Zhang, "CART classification of human 5′ UTR sequences," *Genome Research*, vol. 10, no. 11, pp. 1807–1816, 2000.

[7] J. R. Babendure, J. L. Babendure, J. H. Ding, and R. Y. Tsien, "Control of mammalian translation by mRNA structure near caps," *RNA*, vol. 12, no. 5, pp. 851–861, 2006.

[8] L. A. Ryabova, M. M. Pooggin, and T. Hohn, "Viral strategies of translation initiation: ribosomal shunt and reinitiation," *Progress in Nucleic Acid Research and Molecular Biology*, vol. 72, pp. 1–39, 2002.

[9] A. Yueh and R. J. Schneider, "Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to 18S rRNA," *Genes and Development*, vol. 14, no. 4, pp. 414–421, 2000.

[10] S. A. Chappell, J. Dresios, G. M. Edelman, and V. P. Mauro, "Ribosomal shunting mediated by a translational enhancer element that base pairs to 18S rRNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 25, pp. 9488–9493, 2006.

[11] N. D. Trinklein, S. J. F. Aldred, A. J. Saldanha, and R. M. Myers, "Identification and functional analysis of human transcriptional promoters," *Genome Research*, vol. 13, no. 2, pp. 308–312, 2003.

[12] T. Zhang, P. Haws, and Q. Wu, "Multiple variable first exons: a mechanism for cell—and tissue—specific gene regulation," *Genome Research*, vol. 14, no. 1, pp. 79–89, 2004.

[13] P. Carninci, T. Kasukawa, S. Katayama et al., "The transcriptional landscape of the mammalian genome," *Science*, vol. 309, no. 5740, pp. 1559–1563, 2005.

[14] K. Sobczak and W. J. Krzyzosiak, "Structural determinants of BRCA1 translational regulation," *The Journal of Biological Chemistry*, vol. 277, no. 19, pp. 17349–17358, 2002.

[15] B. A. Arrick, A. L. Lee, R. L. Grendell, and R. Derynck, "Inhibition of translation of transforming growth factor-β3 mRNA by its 5′ untranslated region," *Molecular and Cellular Biology*, vol. 11, no. 9, pp. 4306–4313, 1991.

[16] B. A. Arrick, R. L. Grendell, and L. A. Griffin, "Enhanced translational efficiency of a novel transforming growth factor β3 mRNA in human breast cancer cells," *Molecular and Cellular Biology*, vol. 14, no. 1, pp. 619–628, 1994.

[17] N. K. Gray and M. W. Hentze, "Regulation of protein synthesis by mRNA structure," *Molecular Biology Reports*, vol. 19, no. 3, pp. 195–200, 1994.

[18] B. M. Pickering and A. E. Willis, "The implications of structured 5′ untranslated regions on translation and disease," *Seminars in Cell and Developmental Biology*, vol. 16, no. 1, pp. 39–47, 2005.

[19] F. Rozen, I. Edery, K. Meerovitch, T. E. Dever, W. C. Merrick, and N. Sonenberg, "Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F," *Molecular and Cellular Biology*, vol. 10, no. 3, pp. 1134–1144, 1990.

[20] F. Robert and J. Pelletier, "Translation initiation: a critical signalling node in cancer," *Expert Opinion on Therapeutic Targets*, vol. 13, no. 11, pp. 1279–1293, 2009.

[21] S. J. Kim, K. Park, D. Koeller et al., "Post-transcriptional regulation of the human transforming growth factor-β1 gene," *The Journal of Biological Chemistry*, vol. 267, no. 19, pp. 13702–13707, 1992.

[22] D. S. Romeo, K. Park, A. B. Roberts, M. B. Sporn, and S. J. Kim, "An element of the transforming growth factor-β1 5′-untranslated region represses translation and specifically binds a cytosolic factor," *Molecular Endocrinology*, vol. 7, no. 6, pp. 759–766, 1993.

[23] D. J. Fraser, A. O. Phillips, X. Zhang et al., "Y-box protein-1 controls transforming growth factor-β1 translation in proximal tubular cells," *Kidney International*, vol. 73, no. 6, pp. 724–732, 2008.

[24] R. H. Jenkins, R. Bennagi, J. Martin, A. O. Phillips, J. E. Redman, and D. J. Fraser, "A conserved stem loop motif in the 59untranslated region regulates transforming growth factor-β1 translation," *PLoS one*, vol. 5, no. 8, Article ID e12283, 2010.

[25] I. Abaza and F. Gebauer, "Functional domains of Drosophila UNR in translational control," *RNA*, vol. 14, no. 3, pp. 482–490, 2008.

[26] D. J. Goss and E. C. Theil, "Iron responsive mRNAs: a family of Fe^{2+} sensitive riboregulators," *Accounts of Chemical Research*, vol. 44, no. 12, pp. 1320–1328, 2011.

[27] D. L. Zhang, R. M. Hughes, H. Ollivierre-Wilson, M. C. Ghosh, and T. A. Rouault, "A ferroportin transcript that lacks an iron-responsive element enables duodenal and erythroid precursor cells to evade translational repression," *Cell Metabolism*, vol. 9, no. 5, pp. 461–473, 2009.

[28] S. Abboud and D. J. Haile, "A novel mammalian iron-regulated protein involved in intracellular iron metabolism," *The Journal of Biological Chemistry*, vol. 275, no. 26, pp. 19906–19912, 2000.

[29] D. Girelli, R. Corrocher, L. Bisceglia et al., "Hereditary hyperferritinemia-cataract syndrome caused by a 29-base pair deletion in the iron responsive element of ferritin L-subunit gene," *Blood*, vol. 90, no. 5, pp. 2084–2088, 1997.

[30] L. Merendino, S. Guth, D. Bilbao, C. Martínez, and J. Valcárcel, "Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3′ splice site AG," *Nature*, vol. 402, no. 6763, pp. 838–841, 1999.

[31] P. Forch, L. Merendino, C. Martinez, and J. Valcarcel, "Modulation of msl-2 5′ splice site recognition by Sex-lethal," *RNA*, vol. 7, no. 9, pp. 1185–1191, 2001.

[32] K. Beckmann, M. Grskovic, F. Gebauer, and M. W. Hentze, "A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in drosophila," *Cell*, vol. 122, no. 4, pp. 529–540, 2005.

[33] K. Duncan, M. Grskovic, C. Strein et al., "Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2 mRNA 3′ UTR: translational repression for dosage compensation," *Genes & Development*, vol. 20, no. 3, pp. 368–379, 2006.

[34] I. Abaza, O. Coll, S. Patalano, and F. Gebauer, "Drosophila UNR is required for translational repression of male-specific lethal 2 mRNA during regulation of X-chromosome dosage compensation," *Genes & Development*, vol. 20, no. 3, pp. 380–389, 2006.

[35] K. E. Duncan, C. Strein, and M. W. Hentze, "The SXL-UNR corepressor complex uses a PABP-mediated mechanism to inhibit ribosome recruitment to msl-2 mRNA," *Molecular Cell*, vol. 36, no. 4, pp. 571–582, 2009.

[36] J. Medenbach, M. Seiler, and M. W. Hentze, "Translational control via protein-regulated upstream open reading frames," *Cell*, vol. 145, no. 6, pp. 902–913, 2011.

[37] P. Iakova, G. L. Wang, L. Timchenko et al., "Competition of CUGBP1 and calreticulin for the regulation of p21 translation determines cell fate," *The EMBO Journal*, vol. 23, no. 2, pp. 406–417, 2004.

[38] S. E. Calvo, D. J. Pagliarini, and V. K. Mootha, "Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 18, pp. 7507–7512, 2009.

[39] M. Iacono, F. Mignone, and G. Pesole, "UAUG and uORFs in human and rodent 5′ untranslated mRNAs," *Gene*, vol. 349, pp. 97–105, 2005.

[40] C. Lawless, R. D. Pearson, J. N. Selley et al., "Upstream sequence elements direct post-transcriptional regulation of gene expression under stress conditions in yeast," *BMC Genomics*, vol. 10, article 7, 2009.

[41] M. Matsui, N. Yachie, Y. Okada, R. Saito, and M. Tomita, "Bioinformatic analysis of post-transcriptional regulation by uORF in human and mouse," *The FEBS Letters*, vol. 581, no. 22, pp. 4184–4188, 2007.

[42] M. Kozak, "Constraints on reinitiation of translation in mammals," *Nucleic Acids Research*, vol. 29, no. 24, pp. 5226–5232, 2001.

[43] D. R. Morris and A. P. Geballe, "Upstream open reading frames as regulators of mRNA translation," *Molecular and Cellular Biology*, vol. 20, no. 23, pp. 8635–8642, 2000.

[44] M. Oyama, C. Itagaki, H. Hata et al., "Analysis of small human proteins reveals the translation of upstream open reading frames of mRNAs," *Genome Research*, vol. 14, no. 10, pp. 2048–2052, 2004.

[45] Z. Wang, A. Gaba, and M. S. Sachs, "A highly conserved mechanism of regulated ribosome stalling mediated by fungal arginine attenuator peptides that appears independent of the charging status of arginyl-tRNAs," *The Journal of Biological Chemistry*, vol. 274, no. 53, pp. 37565–37574, 1999.

[46] A. Gaba, Z. Wang, T. Krishnamoorthy, A. G. Hinnebusch, and M. S. Sachs, "Physical evidence for distinct mechanisms of translational control by upstream open reading frames," *The EMBO Journal*, vol. 20, no. 22, pp. 6453–6463, 2001.

[47] C. Wu, J. Wei, P. J. Lin et al., "Arginine changes the conformation of the arginine attenuator peptide relative to the ribosome tunnel," *Journal of Molecular Biology*, vol. 416, no. 4, pp. 518–533, 2012.

[48] P. Fang, C. C. Spevak, C. Wu, and M. S. Sachs, "A nascent polypeptide domain that can regulate translation elongation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 12, pp. 4059–4064, 2004.

[49] P. P. Mueller and A. G. Hinnebusch, "Multiple upstream AUG codons mediate translational control of GCN4," *Cell*, vol. 45, no. 2, pp. 201–207, 1986.

[50] A. G. Hinnebusch, "Translational regulation of GCN4 and the general amino acid control of yeast," *Annual Review of Microbiology*, vol. 59, no. 1, pp. 407–450, 2005.

[51] I. Lohse, P. Reilly, and K. Zaugg, "The CPT1C 5′UTR contains a repressing upstream open reading frame that is regulated by cellular energy availability and AMPK," *PLoS one*, vol. 6, no. 9, 2011.

[52] C. C. Oliveira and J. E. G. McCarthy, "The relationship between eukaryotic translation and mRNA stability. A short upstream open reading frame strongly inhibits translational initiation and greatly accelerates mRNA degradation in the yeast Saccharomyces cerevisiae," *The Journal of Biological Chemistry*, vol. 270, no. 15, pp. 8936–8943, 1995.

[53] K. Wethmar, J. J. Smink, and A. Leutz, "Upstream open reading frames: molecular switches in (patho)physiology," *Bioessays*, vol. 32, no. 10, pp. 885–893, 2010.

[54] L. Liu, D. Dilworth, L. Z. Gao et al., "Mutation of the CDKN2A 5′ UTR creates an aberrant initiation codon and predisposes to melanoma," *Nature Genetics*, vol. 21, no. 1, pp. 128–132, 1999.

[55] A. Wiestner, R. J. Schlemper, A. P. C. van der Maas, and R. C. Skoda, "An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia," *Nature Genetics*, vol. 18, no. 1, pp. 49–52, 1998.

[56] Y. R. Wen, Y. Liu, Y. M. Xu et al., "Erratum: 'loss-of-function mutations of an inhibitory upstream ORF in the human hairless transcript cause Marie Unna hereditary hypotrichosis (Nature Genetics (2009) 41 (228-233))'," *Nature Genetics*, vol. 41, no. 6, p. 762, 2009.

[57] G. Beffagna, G. Occhi, A. Nava et al., "Regulatory mutations in transforming growth factor-β3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1," *Cardiovascular Research*, vol. 65, no. 2, pp. 366–373, 2005.

[58] S. E. Calvo, D. J. Pagliarini, and V. K. Mootha, "Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 18, pp. 7507–7512, 2009.

[59] F. Poulat, M. Desclozeaux, S. Tuffery, P. Jay, B. Boizet, and P. Berta, "Mutation in the 5′ noncoding region of the SRY gene in an XY sex-reversed patient," *Human Mutation*, vol. 10, no. 1, pp. S192–S194, 1998.

[60] S. Kondo, B. C. Schutte, R. J. Richardson et al., "Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes," *Nature Genetics*, vol. 32, no. 2, pp. 285–289, 2002.

[61] L. Groussin, L. S. Kirschner, C. Vincent-Dejean et al., "Molecular analysis of the cyclic AMP-dependent protein kinase A (PKA) regulatory subunit 1A (PRKAR1A) gene in patients with carney complex and primary pigmented nodular adrenocortical disease (PPNAD) reveals novel mutations and clues for pathophysiology: augmented PKA signaling is associated with adrenal tumorigenesis in PPNAD," *American Journal of Human Genetics*, vol. 71, no. 6, pp. 1433–1442, 2002.

[62] H. Witt, W. Luck, H. C. Hennies et al., "Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis," *Nature Genetics*, vol. 25, no. 2, pp. 213–216, 2000.

[63] R. Oner, S. Agarwal, A. J. Dimovski et al., "The G → A mutation at position +22 3′ to the cap site of the β-globin gene as a possible cause for a β-thalassemia," *Hemoglobin*, vol. 15, no. 1-2, pp. 67–76, 1991.

[64] J. Konig, K. Zarnack, G. Rot et al., "ICLIP—transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution," *Journal of Visualized Experiments*, no. 50, article 2638, 2011.

[65] D. R. Morris and A. P. Geballe, "Upstream open reading frames as regulators of mRNA translation," *Molecular and Cellular Biology*, vol. 20, no. 23, pp. 8635–8642, 2000.

[66] V. Matys, O. V. Kel-Margoulis, E. Fricke et al., "TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes," *Nucleic Acids Research*, vol. 34, pp. D108–D110, 2006.

[67] R. J. Keyser, L. van der Merwe, M. Venter et al., "Identification of a novel functional deletion variant in the 5′-UTR of the DJ-1 gene," *BMC Medical Genetics*, vol. 10, article 105, 2009.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 504292, 7 pages
doi:10.1155/2012/504292

Review Article

Cytoplasmic Ribonucleoprotein Foci in Eukaryotes: Hotspots of Bio(chemical)Diversity

Carla Layana,[1,2] Paola Ferrero,[1,2] and Rolando Rivera-Pomar[1,2]

[1] *Centro Regional de Estudios Genómicos, Universidad Nacional de La Plata, CP 1888 Florencio Varela, Argentina*
[2] *Departamento de Ciencias Básicas y Experimentales, Universidad Nacional del Noroeste de Buenos Aires, Avenida Calchaquí 5900, CP 1888 Buenos Aires, Pergamino, Argentina*

Correspondence should be addressed to Rolando Rivera-Pomar, rrivera@unnoba.edu.ar

Received 1 February 2012; Accepted 22 March 2012

Academic Editor: Greco Hernández

The life of an mRNA from transcription to degradation offers multiple control check points that regulate gene expression. Transcription, splicing, and translation have been widely studied for many years; however, in recent years, new layers of posttranscriptional and posttranslational control have been uncovered. They involve the regulation of the metabolism of mRNA in cytoplasmic foci. They are collections of ribonucleoprotein complexes that, in most cases, remain still uncharacterized, except the processing bodies (PBs) and stress granules (SGs), which have been studied (and reviewed) in detail. A challenging prospective is to know how many different classes of foci exist, which functions they support, how are they formed, and how do they relate one to each other. Here, we present an update of the component of the different granules, a possible function, and hypothesis on their *in vivo* dynamics related to translational control.

1. Introduction

In recent years, several cytoplasmic foci/granules that contain proteins and RNA have been described. Two of them have been studied in more detail as they are related to mRNA silencing: stress granules (SG) and processing bodies (PB). SG are repressed mRNPs transiently induced in response to cellular stress. They range from 0,5 to 5 μm [1]. PB are discrete RNP cytoplasmic foci of 0,1-2 μm where the machinery of RNA interference, degradation and storage locates. In PB the mRNAs are forming mRNP complexes either repressing translation, in degradation complexes or stored for further use [2, 3]. SG and PB have been shown to share a growing number of proteins that are added in a day-to-day basis to the list of their components. SG, PB and other cytoplasmic foci are highly dynamic structures, although PB are quite stable over the time [4]; see also Supplementary Movie 1 available online at doi:10.1155/2012/504292. They are in a dynamic steady state with other mRNPs, such as polysomes in response to the translational state of the cell [5]. Although we do not intend to extensively review SG and PB, which have been matter of fine reviews in the last years [6–10], we will overview their functions before we address neglected issues and hypothesis.

2. Stress Granules

Translation initiation is the key regulatory step of translational control. Therefore, it is the most sensitive step to changes in the cellular environment, including stress. A key step in translation initiation inhibition is the phosphorylation of eIF2α, which results in an increase on the affinity of eIF2-GDP for eIF2B, sequestering this factor to prevent new round of translational initiation [11]. During this process, translation is inhibited and polysomes become released from the mRNA leading to the accumulation of inactive mRNPs in SG. The SG are in equilibrium with active polysomes. Protein elongation inhibitors, such as cycloheximide, prevent the assembly of SG by blocking the polysomes in an inactive state, while protein initiation inhibitors promote the formation of SG [8]. Table 1 shows the components of SG described up to now. They can be classified in three main groups as follows.

Table 1: Components of stress granules.

Protein	Function	Interacting proteins
Ago2	Cleaves interfered RNA	RISC, FXR1
APOBEC3G	Antiviral response	?
Ataxin-2	Translation	PABP-1
Caprin-1	Cell growth	G3BP
CPEB	mRNA repression	RCK, eIF4E, FXR1
DIS1	Unknown	eIF3h
eIF3	Translation	40S, eIF4G
eIF4E	Translation	CPEB, Smaug, eIF4G, 4ET
eIF4G	Translation	eIF4E, eIF3, PABP-1
FAST	Translation	TIA-1
FMRP, FXR1	Translation	Ago2, RISC
FBP, KSRP	mRNA degradation	TIA-1
FUS/TLS	Transcriptional control	Transcriptional machinery
G3BP	Ras signalling	Caprin
HuR	mRNA stabilization	?
IP5K	Signalling	?
Lin28	Developmental control	?
LINE 1 ORF1p	Transposon	?
MLN51	Splicing	Exon junction
PABP-1	Translation	eIF4G, eIF3, ataxina-2
RCK(p54)	mRNA degradation	GE-1, TTP
Plakophilin	Adhesion	G3BP, FXR1
PMR1	mRNA degradation	TIA-1
Pumilio 2	mRNA silencing	?
Rap 55	mRNA silencing	?
Rpb4	Transcription	?
SRC3	Transcription	TIA-1
Staufen	mRNA silencing	?
SMN	RNP assembly	Complejo SMN
TDP-43	Transcription and splicing regulator	eIF4G, eIF3, eIF2, ribosomal proteins, STAU-1, Xnr
TIA-1(rox-8), TIAR	mRNA silencing	FAST, SRC3, PMR1, FBP
TRAF2	Signalling	eIF4G
TTP, BRF-1	mRNA silencing	RCK (p54)
YB-1	Cold shock	?
ZBP1	Localization	?

(1) Core components: stalled initiation complexes (polyadenylated mRNAs and translation factors eIF4E, eIF4A, eIF4G, eIF3, eIF2, PABP, and proteins of the small ribosome subunit).

(2) RNA-binding proteins associated to silencing and transcript stability: TIA-1, TIAR [12], FAST, Argonaute [13], CPEB, smaug, DExD/H-box RCK/p54 (o Dhh1), XRN1 [5].

(3) RNA-binding proteins associated to mRNA metabolism either translation of degradation such as G3BP [14] and Staufen [15].

The key concept regarding SG is that they are responsible of protecting the mRNA during cell stress, altering the composition of the mRNPs in a reversible manner. As soon as the cell recovers, the mRNPs regain their translational capacity.

3. Processing Bodies

These structures have been described many times since 1997, when Bashkirov et al. observed that the exonuclease Xrn1 is located in small granular structures in the cytoplasm of mammalian cells and call them "Xrn1 foci" [17]. Later on, the decapping enzyme Dcp2 was also described to occur in cytoplasmic foci [18]. Contemporary, Eystathioy et al. have described that a protein associated to neuropathy named GW182 occurs in cytoplasmic speckles called GW bodies [19, 20]. Other RNA-related protein, the eIF4E-transporter, was also localized in discrete cytoplasmic foci [16, 21, 22]. Short after, a seminal work of Sheth and Parker established the functional bases of the now called PB that resulted in the same structures described many times before [23]. They

Table 2: Components of processing bodies.

Protein	Function	Organisms
XRN1, Sc Kem1	5′ → 3′ exonuclease	Human, mice, Sc
GW182, Ce AIN-1	miARN function	Human, Dm, Ce
DCP2, Ce DCAP2	Decapping	Human, Dm, Ce, Sc
DCP1, Ce DCAP1	Decapping	Human, Dm, Ce, Sc
Hedls, Ge-1	Decapping coactivator	Human, Dm
Dm CG5208, Pat1	Decapping coactivator	Dm, Sc
EDC3 (Lsm16)	Decapping coactivator	Human, Dm, Sc
Lsm1-7	Decapping coactivator	Human, Sc
RAP55	Putative decapping coactivator	Human
RCK/p54, Dm Me31B, Ce CGH-1, Sc Dhh1	Decapping coactivator, translational control	Human, Dm, Ce, Sc
eIF4E	Translation initiation	Human, rat, mouse, Dm, Sc
eIF4E-T	Translational repression	Human
SMG7	Nonsense mediated decay	Human
SNG5	Nonsense mediated decay	Human
UPF1, Sc Nam7	Nonsense mediated decay	Human, Sc
UPF2	Nonsense mediated decay	Human
UPF3	Nonsense mediated decay	Human
Argonaute	siRNA/miRNA pathways	Human, Dm, Ce
CCR4-CAF1-NOT complex	Deadenylation	Human, Dm, Sc
CPEB	Translational control	Human
FAST	S/T phosphoprotein activator of Fas	Human
TTP	ARE-mediated mRNA degradation	Human
Staufen	mRNA localization	Human, mice, Dm
Rbp1	Mitochondrial RNA degradation	Sc
Rbp4	RNA pol II subunit	Sc
Sbp1	Suppressor of deccaping	Sc
Germin 5	Part of small nuclear RNPs	Human
Dcs2	Stress-induced regulator	Sc
APOBEC3G, APOBEC3F	Antiviral activity	Human

demonstrated that PB contains enzymes involved in the degradation of the mRNA. Later one further studies showed that they are also related to miRNA metabolism and can store mRNAs to bring them back to polysomes (reviewed in [3, 24]. They include, different than SG, neither ribosomal proteins nor translation factors, except eIF4E. They do not present either the exosome components [25]. eIF4G and PABP were found in yeast PB, although at low level and in stress conditions resulting on glucose deprivation [26]. Proteins and mRNA can reversely go in and out of PB [25]. The relationship of PB and polysomes is demonstrated by the blocking of PB formation by cycloheximide. A summary of the components in different organisms is shown in Table 2. The occurrence of such large and diverse set of proteins (and the list continuously grows up) suggests that PBs are involved in a plethora of posttranslational processes regulating gene expression, such as mRNA degradation and silencing. mRNA degradation starts with the shortening of the poly-A tail—the deadenylation. In eukaryotes, there are several complexes involved in the process: PARN2-PARN3 initiates the process, which continues with the action of the CAF1-CCR4-NOT complex. Later on, mRNA degradation continues by nucleolytic cleavage on both ends. 3′ → 5′ degradation is catalyzed by the exosome and the SKI complex, while 5′ → 3′ degradation requires previous decapping by DCP2 and the coactivator DCP1 and the action of the exonuclease XRN1. All these enzyme localize in PB. There are several evidences indicating that mRNA degradation occurs in PB.

(i) The assembly of PB depends on mRNA, as RNase treatment of the cells induces the disappearing of PB [27, 28].

(ii) Inhibition or removal of the deadenylase Ccr4 reduces the number and size of PB, while the removal of the downstream-acting enzymes Xrn1 and Dcp1 does not affect the stability of PB [21].

(iii) mRNA degradation intermediates are present in PB [23].

Therefore, one can conclude that mRNA degradation occurs in PB and depends on the existence of degradation enzymes and mRNA degradation intermediates [21, 23, 25, 29]. Many of the PB components are not restricted to the foci and also are present in the soluble cytoplasm and nuclei, suggesting that the different processes might start before the mRNAs entry into PB. PBs are also related to mRNA quality control mechanisms, such as nonsense-mediated

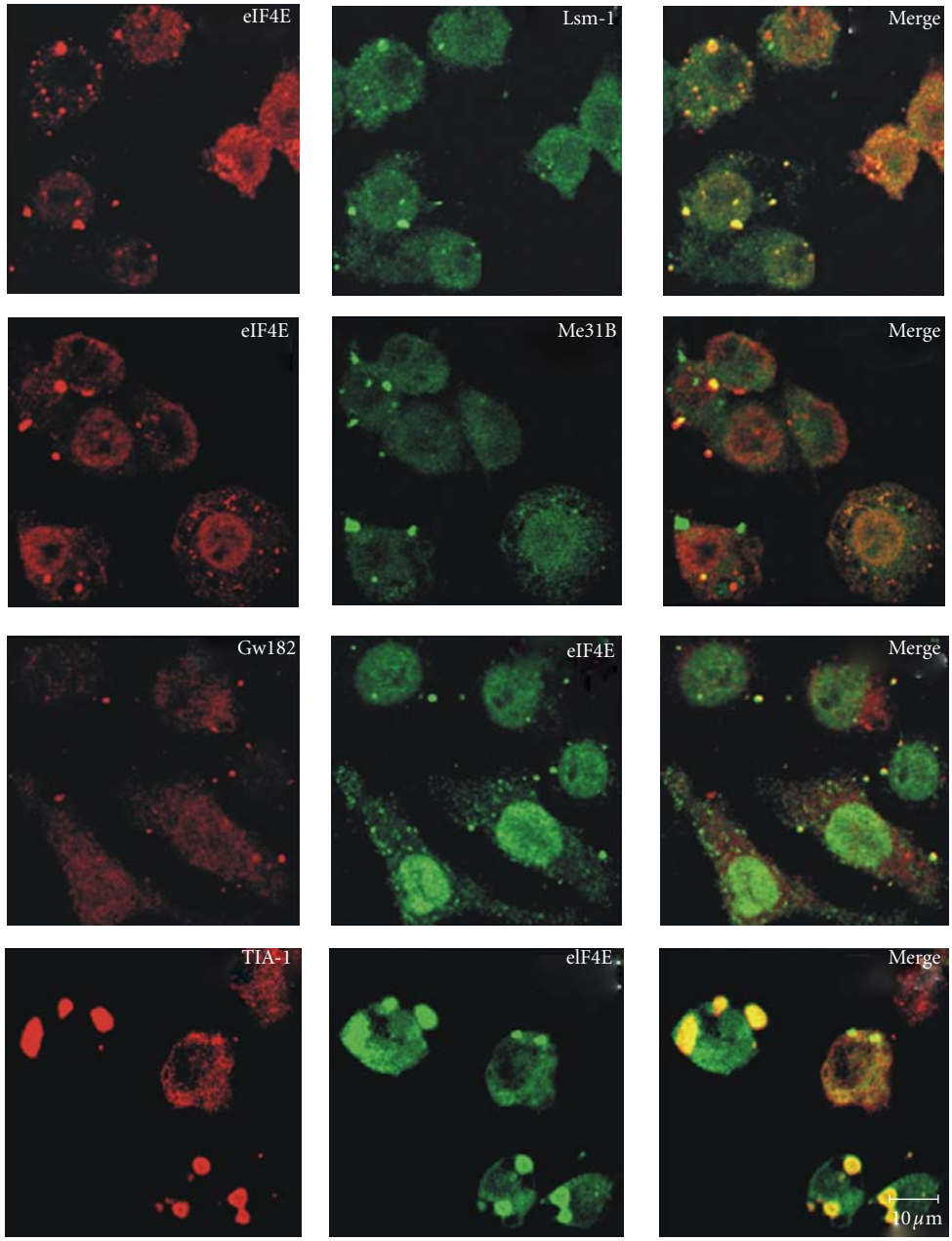


Figure 1: Colocalization of eIF4E with components of PBs and SGs shows a diversity of cytoplasmic foci quality. The experiment shows that, in every case, the granules contain both components or either one or the other in different quantities. This would represent intermediates or different forms of eIF4E-containing foci. *Drosophila melanogaster* S2 cells were transfected with proteins fusion CFP-Lsm-1 or CFP-Me31B. GW182, eIF4E, and Rox8 (the TIA-1 ortholog in *Drosophila*) were revealed using antibodies against GW182, anti-V5, and anti-TIA-1, respectively. In the bottom panel (row 4), the cells were prestressed with arsenite for 30 minutes.

decay (NMD). The detection of premature termination in the cells by spotting an mRNA with an abnormal stop codon is mediated by a surveillance complex composed by UPF1, UPF2 y UPF3, additional proteins, namely, SMG1, and SMG5-7 [30–32]. As soon as the surveillance complex is assembled, the degradation enzymes (Dcp1, Xrn1) are recruited to the mRNA in PB. Although the degrading enzymes are located in PB, the mechanism of recruitment is unknown. In silencing, there are two types of small mRNAs that regulate posttranscriptional gene expression: siRNAs and miRNAs. Despite the different mechanism of silencing, in both cases participate the protein Argonaute (Ago) and the RISC (RNA-induced silencing complex). In the case of siRNAs, Ago produces an endonucleolytic cleavage of the mRNA to promote degradation by the $3' \rightarrow 5'$ and $5' \rightarrow 3'$ decay machinery in PB. In the case of the miRNAs, they

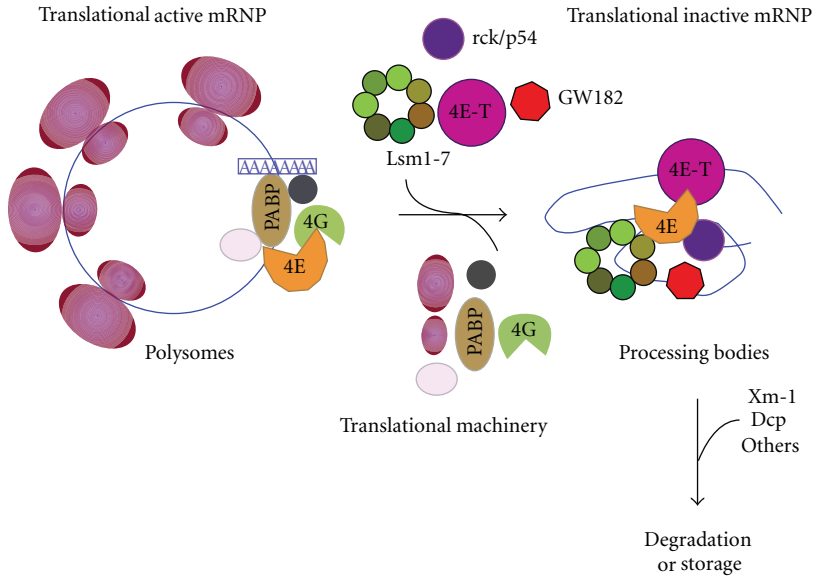


Figure 2: Relationship among active polysomes and PBs. The recruitment of active polysomes to PB implies the removal from the mRNA of the translation factors by translational repressors. Some of them have been demonstrated to interact with eIF4E *in vivo* (rck/p54 and eIF4E-T, [16]). They further interact and/or recruit the enhancers of decapping Lsm1-7 or the miRNA-related protein GW182 to form the PB. Later on, they assemble the decapping and degradation enzymes and/or the proteins required for silencing and storage into PB. All the intermediate steps of this process can represent different populations of granules coexisting in the cell and visible with different morphology that might reflect a variety of components and/or diverse stoichiometry.

recruit Ago to direct the repressed mRNA to degradation mediated by the PB proteins GW182, CCR4-CAF1-NOT1, DCP2, DCP1, and XRN1.

4. The Cycle of an mRNA in the Cell: SG, PB, Polysomes, and the Unknown Intermediates

From the previous analysis, one can establish many unsolved aspects on cytoplasmic foci function. One of them is the dynamic of the mRNP remodeling. The current model suggests an active movement of mRNPs from and to polysomes and from and to SG and PB [33]. However, how does it happen and the factors involved are not known. Translationally active mRNAs can interact, in response to errors in translational initiation or to specific recruitment of regulatory proteins, with translational repressors such as Dhh1, Pat 1, Lsm1-7, eIF4E-T. Those factors would promote the replacement of the translational machinery from the mRNA, promote the cap removal and determine degradation [33] or the accumulation of silenced mRNA in PB. Within PB, mRNPs could undergo further remodeling and define a path to follow, including their return to polysomes. In addition, PBs have been shown to interact and exchange components or their own nature with SG (reviewed in [6, 7]) in a process that may result in mRNPs intermediates of unknown nature. Evidence for the diversity of cytoplasmic foci and their components results from immunocytochemistry and colocalization studies. A common factor present in most cytoplasmic mRNPs is the cap-binding protein eIF4E. eIF4E occurs in active polysomes as a translation initiation factor, in SG as part of the stalled initiation complex, and in PB as the only translation factor present there in multicellular eukaryotes. We observed in *Drosophila* S2 cells that eIF4E colocalizes with different pairs of markers, either for PB (GW182, Lsm1, Me31B—an ortholog of the helicase rck/p54) or SG (TIA-1) and that the colocalization does not occur in all foci in the same way (PVF, CL, and RRP, unpublished data and Figure 1). In some cases, the foci contain one, the other, or both components. In the foci that show colocalization of both factors, the relative amount of each component may vary from foci to foci, as judged by confocal microscopy quantification of the colocalized factors (PVF, CL, RRP, unpublished observation). This implies that there are a diversity of granules. An appealing hypothesis is that eIF4E is a common link among different mRNPs, playing different roles depending on their interactors. One plausible function could be that the accumulation of mRNPs in eIF4E-containing foci is a way to regulate the rate of translation in different physiological states (cell cycle phases, developmental stages, circadian rhythms). Moreover, it has been reported that, in mammalian cells, eIF4E interacts in PB with at least two factors, rck/p54 and eIF4E-T [21]. These are simultaneous interactions within the PB and imply that both proteins could contact different domains of the same eIF4E molecule or that they would represent different populations of mRNPs or different functions within the same PB. In either cases, the complexity of the interactions *in vivo* is more diverse than it has been expected. A model for the remodeling of active mRNPs to silence and degradation based on Andrei et al. [21] is depicted in Figure 2. This might require several intermediate states that can be the maturation steps of a mRNA in the way of a PB or within a PB. This

would correlate with the large diversity of components and interactions within a cytoplasmic foci and the diversity of the foci within a cell. The understanding of the dynamics of mRNP is far from clear and unpredictable paths remain to be discovered. They will need further research and more sophisticated methods for *in vivo* studies.

Acknowledgments

This work was supported by grants from CONICET (PIP 00318 to P. Ferrero) and ANPCyT (PICT-2008-1237 to RRP and P. Ferrero). P. Ferrero and R. Rivera-Pomar are investigators and CL doctoral fellow of the CONICET.

References

[1] M. G. Thomas, L. J. Martinez Tosar, M. A. Desbats, C. C. Leishman, and G. L. Boccaccio, "Mammalian staufen 1 is recruited to stress granules and impairs their assembly," *Journal of Cell Science*, vol. 122, no. 4, pp. 563–573, 2009.

[2] S. P. Chan and F. J. Slack, "MicroRNA-mediated silencing inside P-bodies," *RNA Biology*, vol. 3, no. 3, pp. 97–100, 2006.

[3] A. Eulalio, I. Behm-Ansmant, D. Schweizer, and E. Izaurralde, "P-body formation is a consequence, not the cause, of RNA-mediated gene silencing," *Molecular and Cellular Biology*, vol. 27, no. 11, pp. 3970–3981, 2007.

[4] A. Aizer and Y. Shav-Tal, "Intracellular trafficking and dynamics of P bodies," *Prion*, vol. 2, no. 4, pp. 131–134, 2008.

[5] N. Kedersha, G. Stoecklin, M. Ayodele et al., "Stress granules and processing bodies are dynamically linked sites of mRNP remodeling," *Journal of Cell Biology*, vol. 169, no. 6, pp. 871–884, 2005.

[6] M. G. Thomas, M. Loschi, M. A. Desbats, and G. L. Boccaccio, "RNA granules: the good, the bad and the ugly," *Cellular Signalling*, vol. 23, no. 2, pp. 324–334, 2011.

[7] J. R. Buchan and R. Parker, "Eukaryotic stress granules: the ins and outs of translation," *Molecular Cell*, vol. 36, no. 6, pp. 932–941, 2009.

[8] P. Anderson and N. Kedersha, "Stress granules: the Tao of RNA triage," *Trends in Biochemical Sciences*, vol. 33, no. 3, pp. 141–150, 2008.

[9] A. Eulalio, I. Behm-Ansmant, and E. Izaurralde, "P bodies: at the crossroads of post-transcriptional pathways," *Nature Reviews Molecular Cell Biology*, vol. 8, no. 1, pp. 9–22, 2007.

[10] R. Parker and U. Sheth, "P bodies and the control of mRNA translation and degradation," *Molecular Cell*, vol. 25, no. 5, pp. 635–646, 2007.

[11] S. Yamasaki and P. Anderson, "Reprogramming mRNA translation during stress," *Current Opinion in Cell Biology*, vol. 20, no. 2, pp. 222–226, 2008.

[12] N. L. Kedersha, M. Gupta, W. Li, I. Miller, and P. Anderson, "RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2α to the assembly of mammalian stress granules," *Journal of Cell Biology*, vol. 147, no. 7, pp. 1431–1442, 1999.

[13] A. K. Leung, J. M. Calabrese, and P. A. Sharp, "Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 48, pp. 18125–18130, 2006.

[14] H. Tourrière, K. Chebli, L. Zekri et al., "The RasGAP-associated endoribonuclease G3BP assembles stress granules," *Journal of Cell Biology*, vol. 160, no. 6, pp. 823–831, 2003.

[15] M. G. Thomas, L. J. Martinez Tosar, M. Loschi et al., "Staufen recruitment into stress granules does not affect early mRNA transport in oligodendrocytes," *Molecular Biology of the Cell*, vol. 16, no. 1, pp. 405–420, 2005.

[16] J. Dostie, F. Lejbkowicz, and N. Sonenberg, "Nuclear eukaryotic initiation factor 4E (eIF4E) colocalizes with splicing factors in speckles," *Journal of Cell Biology*, vol. 148, no. 2, pp. 239–246, 2000.

[17] V. I. Bashkirov, H. Scherthan, J. A. Solinger, J. M. Buerstedde, and W. D. Heyer, "A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates," *Journal of Cell Biology*, vol. 136, no. 4, pp. 761–773, 1997.

[18] E. van Dijk, N. Cougot, S. Meyer, S. Babajko, E. Wahle, and B. Séraphin, "Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures," *The European Molecular Biology Organization Journal*, vol. 21, no. 24, pp. 6915–6924, 2002.

[19] T. Eystathioy, E. K. Chan, S. A. Tenenbaum, J. D. Keene, K. Griffith, and M. J. Fritzler, "A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles," *Molecular Biology of the Cell*, vol. 13, no. 4, pp. 1338–1351, 2002.

[20] T. Eystathioy, A. Jakymiw, E. K. Chan, B. Séraphin, N. Cougot, and M. J. Fritzler, "The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies," *RNA*, vol. 9, no. 10, pp. 1171–1173, 2003.

[21] M. A. Andrei, D. Ingelfinger, R. Heintzmann, T. Achsel, R. Rivera-Pomar, and R. Lührmann, "A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies," *RNA*, vol. 11, no. 5, pp. 717–727, 2005.

[22] M. A. Ferraiuolo, S. Basak, J. Dostie, E. L. Murray, D. R. Schoenberg, and N. Sonenberg, "A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay," *Journal of Cell Biology*, vol. 170, no. 6, pp. 913–924, 2005.

[23] U. Sheth and R. Parker, "Decapping and decay of messenger RNA occur in cytoplasmic processing bodies," *Science*, vol. 300, no. 5620, pp. 805–808, 2003.

[24] R. Parker and U. Sheth, "P bodies and the control of mRNA translation and degradation," *Molecular Cell*, vol. 25, no. 5, pp. 635–646, 2007.

[25] M. Brengues, D. Teixeira, and R. Parker, "Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies," *Science*, vol. 310, no. 5747, pp. 486–489, 2005.

[26] M. Brengues and R. Parker, "Accumulation of polyadenylated mRNA, Pab1p, eIF4E, and eIF4G with P-bodies in *Saccharomyces cerevisiae*," *Molecular Biology of the Cell*, vol. 18, no. 7, pp. 2592–2602, 2007.

[27] V. I. Bashkirov, H. Scherthan, J. A. Solinger, J.-M. Buerstedde, and W.-D. Heyer, "A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates," *The Journal of Cell Biology*, vol. 136, no. 4, pp. 761–773, 1997.

[28] D. Ingelfinger, D. J. Arndt-Jovin, R. Lührmann, and T. Achsel, "The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrn1 in distinct cytoplasmic foci," *RNA*, vol. 8, no. 12, pp. 1489–1501, 2002.

[29] N. Cougot, S. Babajko, and B. Séraphin, "Cytoplasmic foci are sites of mRNA decay in human cells," *Journal of Cell Biology*, vol. 165, no. 1, pp. 31–40, 2004.

[30] E. Conti and E. Izaurralde, "Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species,"

Current Opinion in Cell Biology, vol. 17, no. 3, pp. 316–325, 2005.

[31] N. Amrani, M. S. Sachs, and A. Jacobson, "Early nonsense: mRNA decay solves a translational problem," *Nature Reviews Molecular Cell Biology*, vol. 7, no. 6, pp. 415–425, 2006.

[32] F. Lejeune and L. E. Maquat, "Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells," *Current Opinion in Cell Biology*, vol. 17, no. 3, pp. 309–315, 2005.

[33] V. Balagopal and R. Parker, "Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs," *Current Opinion in Cell Biology*, vol. 21, no. 3, pp. 403–408, 2009.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 287852, 13 pages
doi:10.1155/2012/287852

Review Article

Conservation of the RNA Transport Machineries and Their Coupling to Translation Control across Eukaryotes

Paula Vazquez-Pianzola and Beat Suter

Institute of Cell Biology, University of Bern, Baltzerstrasse 4, 3012 Bern, Switzerland

Correspondence should be addressed to Paula Vazquez-Pianzola, paula.vazquez@izb.unibe.ch

Received 28 December 2011; Accepted 9 February 2012

Academic Editor: Greco Hernández

Restriction of proteins to discrete subcellular regions is a common mechanism to establish cellular asymmetries and depends on a coordinated program of mRNA localization and translation control. Many processes from the budding of a yeast to the establishment of metazoan embryonic axes and the migration of human neurons, depend on this type of cell polarization. How factors controlling transport and translation assemble to regulate at the same time the movement and translation of transported mRNAs, and whether these mechanisms are conserved across kingdoms is not yet entirely understood. In this review we will focus on some of the best characterized examples of mRNA transport machineries, the "yeast locasome" as an example of RNA transport and translation control in unicellular eukaryotes, and on the *Drosophila* Bic-D/Egl/Dyn RNA localization machinery as an example of RNA transport in higher eukaryotes. This focus is motivated by the relatively advanced knowledge about the proteins that connect the localizing mRNAs to the transport motors and the many well studied proteins involved in translational control of specific transcripts that are moved by these machineries. We will also discuss whether the core of these RNA transport machineries and factors regulating mRNA localization and translation are conserved across eukaryotes.

1. Introduction

RNA transport coupled with translation control is a crucial mechanism to target protein expression to specific regions of a cell or an organism. During transport, mRNAs associate with proteins that control every step in the mRNA life cycle. Together, mRNAs and proteins form large ribonucleoprotein (RNP) complexes in which different factors control assembly, stability, translation, and transport of localized mRNAs. Microtubules, microfilaments, and their motors then transport these complexes to their final destination. To achieve local protein synthesis at the final target site, translation of transported mRNAs must be repressed during their journey and then activated only once the mRNAs reach their final destination.

Although several proteins involved in translation control of localized transcripts have been described, how translation repression during transport occurs and how local protein synthesis is activated at the final destination of a given mRNA are only partially known for a few mRNAs. In this paper we will focus on some of the best-characterized examples of translational regulation of localized transcripts and we will analyze whether the complexes regulating localization and translation are conserved in other eukaryotes. We will also attempt to shed light on the conservation of the coupling between mRNA localization and translational control across eukaryotes.

2. RNA Localization Supports Local Protein Synthesis from Bacteria to Metazoans

Mechanisms to target mRNAs to discrete subcellular locations, where their protein products are expressed locally, were traditionally thought to be a hallmark of eukaryotes, which synthesize and translate mRNAs in different cellular compartments, namely, the nucleus and the cytoplasm [1–4]. However, recent findings indicate that even in bacteria some mRNAs move from the nucleoid to particular regions of the cell before they get translated [5]. Most interestingly,

a *Drosophila* transcript encoding a membrane protein was recently reported to localize to the membrane in *E. coli,* too [5]. This would imply that recognition of localizing signals within the transcripts have been conserved during evolution and with this probably large parts of the RNA transport machinery.

In unicellular eukaryotes such as yeast, many transcripts are actively transported to the cell bud. This is the case for *ASH1* (asymmetric synthesis of HO) mRNA, which is localized to the bud of daughter cells and is essential for the mating-type switch (see what follows). RNA transport phenomenons have also been described in plants. In addition to cellular localization, transport of viral genomes, cellular mRNAs, and small RNAs (miRNAs and siRNAs) between cells through plasmodesmata and through the phloem is a common process in higher plants [6, 7]. Although the mechanisms regulating these processes are not well studied, it seems that cell to cell RNA movement is mediated by plant factors and that plasmodesmal transport is a highly regulated process. As an example, the homeodomain protein KNOTTED1 facilitates the transport of its own mRNA from cell-to-cell and this RNA is translated after its translocation [8, 9].

Some of the first examples of regulation of gene expression involving translational control of localized RNAs were described while studying embryonic development in metazoans. During this stage, maternal mRNAs accumulate in specific regions of *Xenopus* and *Drosophila* embryos, and translational control of these localized mRNAs is essential for embryogenesis. Examples of such process in *Drosophila* are the localization of the mRNA encoding the maternal determinant Bicoid (Bcd) to the anterior cortex of the oocyte, and of *oskar* (*osk*) and *nanos* (*nos*) mRNAs to the posterior cortex [10]. Their proper localization and translational control are essential for specifying the anteroposterior axis of the embryo. Similarly, localization of *gurken* (*grk*) mRNA to the dorsoanterior corner of the oocyte, next to the oocyte nucleus, is essential for the specification of the dorsoventral axis of the egg chamber and of the embryo [10]. In *Xenopus*, the mRNAs encoding the T-box transcription factor VegT and a member of the transforming growth factor-beta (TGF-β) family, Vg1, localize to the vegetal cortex of *Xenopus* oocytes and play roles in endodermal and mesodermal specification during early embryogenesis [11]. Importantly, a growing number of other mRNAs have also been reported to be localized in oocytes, eggs, and cleaving embryos of diverse organisms including the wasp *Nasonia vitripennis*, the cnidarian *Clytia hemisphaerica*, zebrafish, and several ascidian species, highlighting the importance of the RNA localization process across eukaryotes [12]. Other examples involve the accumulation and local synthesis of RNAs in the protruding edges of polarized cells, like fibroblasts and neurons [13, 14]. β-actin mRNA targeting to lamellipodia of chicken fibroblasts combined with its local translation at this site produces an enrichment of actin at the leading edge of these cells, a process that is required for cell motility [14]. β-actin mRNA is also localized in dendrites, where it is needed for ligand-dependent filopodial growth of rat hippocampal neurons [15, 16]. Some mRNAs are also localized and locally translated in axonal growth cones [13]. For example, local translation of β-actin mRNA facilitates Ca^{2+}- and netrin-1-dependent growth cone guidance in *Xenopus* [17, 18].

Surprisingly, a systematic study of 3370 transcripts expressed during embryonic development in *Drosophila* showed that 71% of the analyzed mRNAs exhibited clear subcellular distribution patterns, suggesting that virtually all aspects of cellular function are impacted by RNA localization pathways [19]. Interestingly, many of these mRNAs showed novel patterns of localization, which suggest the existence of so-far unknown subcellular structures where these mRNAs and their protein products might play specific local functions. The cited study was only taking into account ~25% of the *Drosophila* genome, leaving considerable room to discover additional localizing transcripts and novel spatially restricted subcellular locations, which could unveil the existence of unknown subcellular compartments. Importantly, there was also a high correlation between the RNA localization sites and the localization of the proteins they encode, confirming that translation control is tightly regulated during RNA transport [19].

3. Localization and Translation of RNAs in Non-Polarized Cells

Recently, mRNA transport and localized translation have been found to occur in very specific regions within non-polarized cells as well. In yeast, 423 mRNAs were found to localized to mitochondrion-bound polysomes [20]. In this case, about half of them encode putative nuclear-encoded mitochondrial proteins, suggesting that this serves to locally translate them in the vicinity of mitochondria [20]. Interestingly, the 3′-UTR of some of these mRNAs is sufficient to target the mRNAs to the vicinity of the mitochondria in a translation-independent manner [20, 21]. In another study, some mRNAs were found localized to the endoplasmic reticulum (ER), and further studies demonstrated that this localization can happen in a translation- and Signal Recognition Particle- (SRP-) independent manner [22]. In yeast, many mRNAs encoding peroxin or matrix proteins also colocalize with peroxisomes. For example, *PEX14* mRNA seems to localize to the peroxisomes and its 3′-UTR plays a role in this localization [23]. In conclusion, although localization of mRNAs around the mitochondria, ER, and peroxisomes was first thought to take place cotranslationally by the presence of specific import signals in the nascent proteins, it is becoming now clear that mRNAs have intrinsic targeting information for localization to the vicinity of these compartments.

Centrosomes and spindles also contain RNAs that either have structural functions or are carried along for asymmetric distribution during cell division. Ribosomes are also associated with spindles in frog egg extracts. Recently, Sharp and colleagues used high throughput sequencing to identify ~450 mRNAs significantly enriched on microtubules (MT-RNAs) [24]. They found an overrepresentation of transcripts involved in regulation of mitosis or playing roles in cell division, spindle formation, and chromosome function.

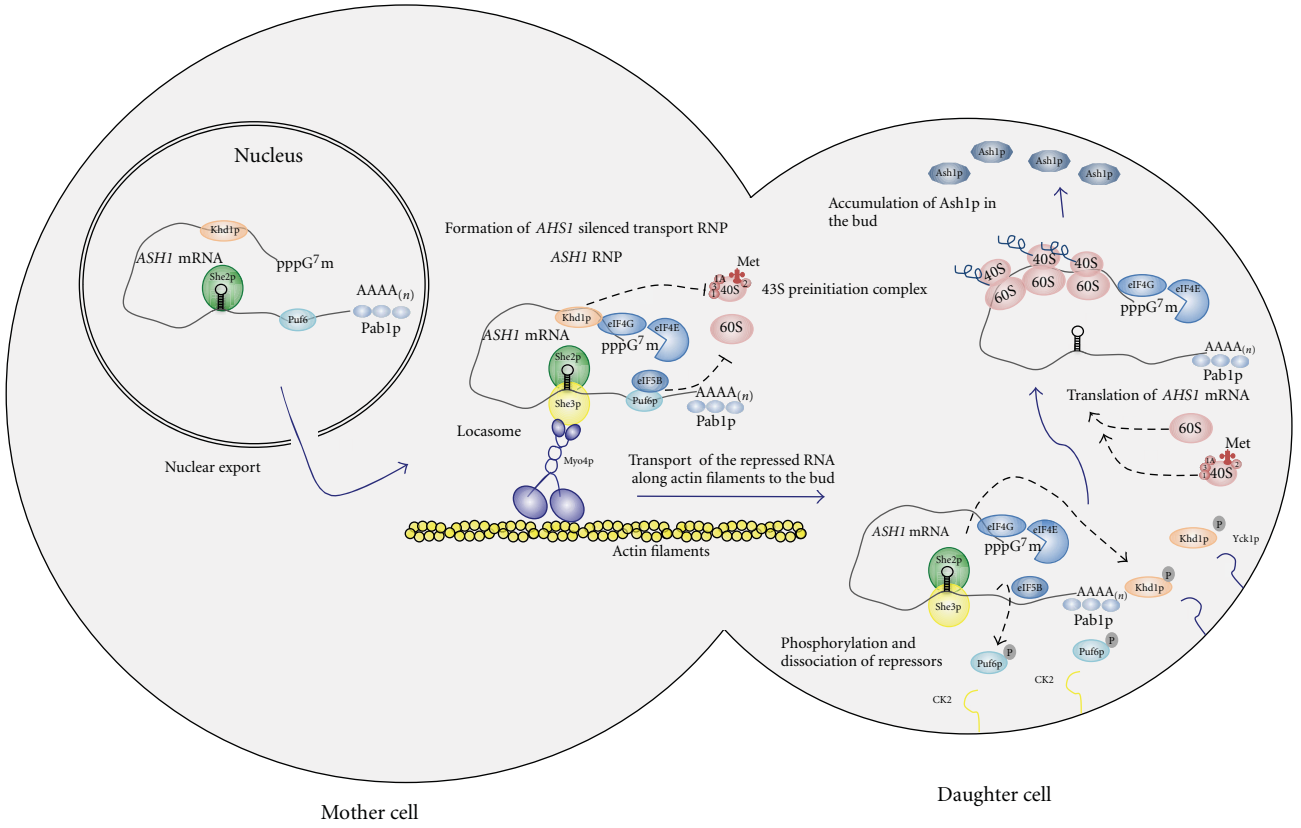


Figure 1: Transport and translation repression of *ASH1* mRNA in *S. cerevisiae*. *ASH1* mRNA is synthesized in the nucleus of the mother cell. The She2p protein is loaded onto *ASH1* mRNA in the nucleus. Once in the cytoplasm, the *ASH1*-She2p complex binds to She3p which associates with Myo4p to form the transport machinery called the "locasome". The translation repressors Puf6p and Khd1p and Pabp1 (which is needed for localization) are thought to be also loaded onto *ASH1* mRNA before nuclear export. The locasome then transports silenced *ASH1* RNPs to the bud through the actin filaments. Puf6p and Khd1p block *AHS1* mRNA translation during transport by different mechanisms. One of them is through the interaction of Puf6p with eIF5B and further inhibition of the recruitment of the 60S ribosomal to the mRNA. Khd1p binds eIF4G. This interaction might prevent the recruitment of the 43S pre-initiation complex (consisting of the 40S subunit, the stabilizing factors eIF3, eIF1 and eIF1A and a ternary complex composed of eIF2 bound to an initiator Met-tRNA and GTP) to the mRNA, thereby blocking translation initiation. However, the exact mechanism is not clearly understood. Once the complex is localized to the bud tip, membrane associated kinases, CK2 and Yck1p, phosphorylate Puf6p and Khd1p respectively. This produces the dissociation of the repressors from *ASH1* mRNA allowing thus translation activation. Ash1p then inhibits mating-type switching only in the daughter cell.

This supports the notion that association of mRNAs with microtubules is a mechanism used to compartmentalize functionally related mRNAs also within the nucleocytoplasmic space of mitotic cells, where MT-RNAs are likely to contribute to spindle-localized mitotic translation.

4. Localization Coupled to Translational Control in Unicellular Organisms: The "Locasome" and the Ash1 Paradigm

One of the best-characterized examples of RNA transport coupled to translation control is the localization of *Saccharomyces cerevisiae ASH1* mRNA [25]. This mRNA is transported to the distal tip of the bud, resulting in the asymmetric sorting of the transcriptional repressor Ash1p into the daughter cell nucleus. In the daughter cell Ash1p represses transcription of the HO endonuclease, inhibiting mating-type switching in the daughter cell [26, 27]. Trans-acting factors Myo4p, She3p, and She2p drive *ASH1* mRNA localization and form a complex known as the "locasome," which is also essential for the localization of many other bud-localized mRNAs [28–30]. She2p is an RNA-binding protein that directly interacts with the *ASH1* mRNA *cis*-acting elements, and Myo4p is a type V myosin motor that functions to directly transport *ASH1* mRNA to the bud along the actin cytoskeleton. She3p was initially suggested to act as an adaptor protein because it can simultaneously associate with Myo4p and She2p. However, recent data also suggested that She3p interacts directly with *ASH1* mRNA [31], suggesting that it stabilizes RNP assembly through different interactions (Figure 1).

Silencing *ASH1* mRNA before it is localized at the bud cortex in late anaphase is critical for asymmetric segregation of Ash1p to the daughter cell nucleus. Puf proteins are known to be versatile posttranscriptional repressors that can

bind different transcripts with diverse cellular functions [32]. In yeast, Puf6p binds *ASH1* mRNA and it is involved in translational repression of this mRNA and in its localization [33]. Deng et al. reported that Puf6p interferes with the conversion of the 48S preinitiation complex to the 80S initiation complex during translation initiation, and this repression is mediated through an interaction of Puf6p with the general translation factor eIF5B [34]. When the mRNA reaches the bud tip, protein kinase CK2 (casein kinase II) phosphorylates the N-terminal region of Puf6p and the repression is then relieved [34]. Khd1p is another protein that interacts with *ASH1* mRNA and reduces translation initiation of the *ASH1* mRNA [35, 36]. Several translation factors have been found to associate with Khd1p, including eIF4G1, eIF4G2, eIF4E, and PABP. Interestingly, Khd1p has been found to interact directly with the C-terminal domain of eIF4G1 to regulate the translation of *ASH1* transcripts. Again, a phosphorylation step seems to trigger translational derepression at its final destination. At the bud plasma membrane, the type I Casein kinase (Yck1p) phosphorylates Khd1p. This leads to the dissociation of Khd1p from the *ASH1* mRNA, releasing its translational repression [35] (Figure 1).

4.1. The Puf Family of Proteins, but Not the Locasome, Is Conserved. The adaptor proteins She2p and She3p link mRNAs to the myosin motor. They are only present in fungi, indicating that the main core of the "locasome" either evolved only in this lineage or was lost and further replaced by other machineries in other eukaryotes. Interestingly, members of the Puf family of proteins are present across kingdoms. *Drosophila melanogaster* has two Puf orthologs, vertebrates have three, yeast six, *Caenorhabditis elegans* 12, rice 19, and *Arabidopsis* 26 [32, 37]. Besides the aforementioned translation repression of *ASH1* mRNA by Puf6, several other mechanisms of translation repression involving Puf members have been described. For example, yeast Puf5, Puf4, and Puf3, *D. melanogaster* Pumilio (Pum), *C. elegans* FBF and human Pum1 interact with the Ccr4-Pop-Not deadenylase complex, indicating that they influence translation and stability of their target mRNAs by controlling poly(A) tail length [38–42]. *Drosophila* Pum recruits the translation inhibitor 4E-HP to *hunchback* mRNA via the protein Brain tumor (Brat), thereby inhibiting translation initiation [43]. In *Xenopus,* Pum2 competes with eIF4E for cap structure binding and this also inhibits translation initiation [44]. Yeast Pufs function in mRNA localization; Puf5 is involved in the localization of *PEX14* mRNA to the peroxisomes, and Puf3 drives mRNA localization to mitochondria [23, 45, 46]. Different classes of mRNAs have been found to be associated with different yeast Puf proteins. Puf3 binds mainly to nuclear mRNAs that encode mitochondrial proteins, Puf1 and Puf2 bind preferentially mRNAs encoding nucleolar ribosomal RNA-processing factors, and Puf5 associates with mRNAs that encode components of the spindle pole and chromatin modifiers [47]. This specificity of the interaction of a Puf family with subsets of functionally related mRNAs seems to indicate that different Puf families may regulate translation and localization of specific subsets of mRNAs. Mammalian Pum2 (mPum2) forms discrete RNA-containing particles in the somatodendritic compartment of polarized neurons, suggesting a role in localization of RNPs in dendrites and in the formation of stress granules [48]. In hippocampal neurons mPum2 is involved in translation repression of the mRNA encoding the translation initiation factor eIF4E and, interestingly, postsynaptic Pum also negatively regulates the expression eIF4E at the *Drosophila* neuromuscular junction (NMJ) [49, 50]. This suggests a conserved role of Pum proteins in regulating local translation at the synapses by controlling the local levels of eIF4E and thus general translation initiation on localizing mRNAs. Altogether, these observations support the notion that Pufs are conserved proteins that regulate localization and local translation of different mRNAs.

5. Localization-Coupled Translational Control in Multicellular Organisms: The Bic-D/Egl mRNA Localization Machinery and the *Osk* Paradigm

Is There a General Drosophila RNA Localization Machinery? In *Drosophila,* an RNA transport machinery plays a key role in oogenesis by localizing into the oocyte RNAs required for oocyte determination, differentiation, and formation of anterior-posterior and dorsal-ventral polarity. This machinery is composed of Bicaudal-D (Bic-D) and Egalitarian (Egl) proteins, which interact with the cytoplasmic microtubule motors Dynein(Dyn)/Dynactin to move the mRNA cargo on microtubules (MTs) to distinct cellular compartments [51, 52]. During oocyte determination, a single cell among an interconnected cyst of sixteen germline cells differentiates into an oocyte, and this process involves the preferential accumulation of specific messenger RNAs and proteins in this cell. The other fifteen cells adopt a nurse cell fate and provide the oocyte with the materials required for oocyte growth. *Bic-D* loss-of-function mutant females produce egg chambers composed of 16 polyploid cells with nurse cell appearance, indicating that the oocyte fails to differentiate. Since *Bic-D* mutant egg chambers fail to accumulate oocyte-specific mRNAs (such as *osk*, *orb, Bic-D,* and *fs(1)K10*) in the future oocyte, it is suggested that the loss of oocyte differentiation is due to a failure in the transport of oocyte-specific proteins and mRNAs from the nurse cells into the oocyte [53, 54]. Ovaries mutant for *egl* as well as wild-type ovaries treated with microtubule-depolymerising drugs show the same 16 nurse cell phenotype as *Bic-D* mutants [55, 56]. Studies using fluorescently labelled mRNAs injected into the nurse cells have shown that Bic-D and Egl are recruited to injected *grk* and *bcd* mRNAs in the nurse cells, and these proteins are required for *grk* transport into the oocyte [57]. The same studies found that transport along MTs via Dyn is also required for the efficient transport of *grk*, *bcd*, and *osk* RNA from the nurse cells to the oocyte [57]. Moreover, the Bic-D/Egl/Dyn machinery is also used for the apical localization of *inscuteable* mRNA in neuroblasts [58] and pair rule and

wingless segmentation mRNAs in the blastoderm embryos [59].

The formation of the Bic-D/Egl/Dyn complex has been studied in *Drosophila* and in mammals. While *Drosophila* Egl interacts directly with Bic-D and also binds the Dyn light chain (Dlc), mammalian orthologues of Bic-D bind *in vitro* directly to components of the Dyn and Dynactin complexes and they also associate *in vivo* with them [60, 61]. Therefore, it is suggested that the Bic-D/Egl complex acts as a link between a microtubule-dependent Dyn motor and the mRNAs. Dienstbier et al. showed that Egl binds directly to mRNAs that localize in the oocyte and apically in the embryos, suggesting that Egl is the factor that links the molecular motors and Bic-D with the transported mRNAs [62]. However, it is still not clear whether Egl is a general link for all mRNAs transported by this machinery or whether additional proteins are required for the specificity of the interaction since, so far, only a specific direct link between the complex Bic-D/Egl and the localization signals of *grk, K10,* and *I factor* mRNAs have been demonstrated. Moreover, Egl alone seems to have an inherent degree of mRNA promiscuity *in vivo* and *in vitro* [62, 63].

The current model proposes that all maternally localized mRNAs are transported by the Bic-D/Egl/Dyn localization machinery from nurse cells to the oocyte [10]. The current data also suggest that Bic-D and Egl form part of a general mRNA transport machinery used repeatedly throughout *Drosophila* development. Although many mRNAs are transported by this machinery, studies of the proteins controlling their translation while transported are still missing for most of the localized transcripts. Nevertheless, in a similar way to the yeast locasome, the Bic-D/Egl/Dyn machinery must also be part of a bigger RNP complex that contains proteins involved in translation control of the transported mRNAs, ensuring that protein synthesis is only activated once the mRNAs reach their final destination.

Control of Translation of RNAs Transported by the Bic-D/Egl Machinery: The Osk Paradigm. Drosophila osk mRNA is transported by the Bic-D/Egl/Dyn transport machinery from nurse cells to the oocyte [57, 59] (Figure 2). Within the oocyte *osk* mRNA switches to a kinesin-based motor that transports it to the posterior cortex. However, only kinesin heavy chain (KHC), but not the kinesin light chain (KLC), is required for this movement, and the KLC-likeprotein PAT1 functions as a positive regulator of KHC during posterior localization of *osk* mRNA [64–66]. Although the mechanism of localization to the posterior has been controversial, based on recent studies that followed the movement of *osk* mRNA particles *in vivo* Zimyanin et al. proposed a new model where *osk* mRNA is localized by random walk on microtubules. Each particle undergoes large numbers of active movements in different directions, but shows an excess of movements towards the posterior which is sufficient to produce the strong posterior localization seen by stage 9 [67]. While kinesin is involved in this long-range MT-based transport of *osk* mRNA throughout the oocyte and into the posterior cytoplasm, recent results indicate that this movement is followed by short-range actomyosinV-dependent translocation or entrapment of *osk* mRNA at the posterior cortex [68].

osk mRNA is one of the most studied models for translation control during transport, and the aforementioned transport machineries must associate with different factors that control translation of the mRNA during transport (Figure 2). During its extended journey, *osk* mRNA translation must be repressed since Osk protein is only observed once the mRNA reaches its final destination at the posterior cortex of the oocyte after stage 8 of oogenesis. Mutants in *armitage* (*armi*), *aubergine* (*aub*, also known as *sting*), *spindle-E* (*spn-E*, also known as *homeless* (*hls*)), *maelstrom* (*mael*) [69], *zucchini* (*zuc*), *squah* (*squ*) [70], and *krimper* (*krimp*) [71] show premature translation of *osk* mRNA in the oocyte during the first part of oogenesis up to stage 6. Interestingly, *spn-E*, *armi*, *aub*, *zuc,* and *squ* are also needed for silencing of the *Stellate* locus, a gene regulated endogenously by small RNAs [70, 72–75]. *spn-E*, *zuc, aub*, *squ* and *krimp* are additionally needed for silencing of retrotransposons in the *Drosophila* germline [70, 71, 76, 77]. Silencing of the transposable elements and the *Stellate* locus is achieved by a mechanism that uses a class of small RNAs called repeat-associated small interfering RNA (rasiRNAs), subsequently renamed Piwi-interacting RNAs (piRNAs) [78–80]. The *mael, krimp, spn-E* [71], *zuc, squ, aub,* and *spn-E* [70] genes have been implicated in the production of these piRNAs. This pathway is germline-specific and depends on the Piwi subfamily of argonaute proteins, which include Aub, Piwi, and Ago3 (reviews in [79, 80]). It is therefore possible that translational silencing of *osk* mRNA during early oogenesis is driven by piRNA-Piwi-Argonaute complexes interacting with *osk* mRNA. However, whether piRNAs play a direct or indirect role in translation control and which piRNAs are involved in translation repression of *osk* mRNA should still be studied. Egg chambers mutants in the *Maternal expression at 31B* (*Me31B*) gene show ectopic Osk accumulation in the nurse cells rather than in the oocyte during early oogenesis indicating that Me31B repress *osk* translation during its transport through the nurse cell into the oocyte [81]. Since Me31B egg chambers degenerate around stage 6, a role in translation repression in later stages could not be ruled out. *Drosophila* PTB (polypyrimidine tract-binding protein) is also involved in translational repression during early (starting at stage 5-6) and late oogenesis of the localizing *osk* mRNA by binding along the *osk* 3′-UTR and mediating assembly of high-order complexes containing multiple *osk* RNAs that produce translational silencing [82]. Interestingly, a 50 kDa pumpkin phloem RNA-binding protein (RBP50), which is evolutionarily related to animal PTBs, seems to be part of the core of an RNP that contains proteins and RNAs transported in the phloem [83]. A complex made up by Bruno (Bru) and Cup represses *osk* mRNA cap-dependent translation from stages 5-6 on [84]. Bru binds simultaneously to Bruno-response elements (BREs) in *osk* 3′-UTR and to Cup. Cup is an eIF4E-binding protein that competes with eIF4G for the binding to eIF4E, thereby inhibiting recruitment of the small ribosomal subunit to *osk* mRNA [84]. Egg chambers expressing a Cup mutant protein that cannot bind eIF4E

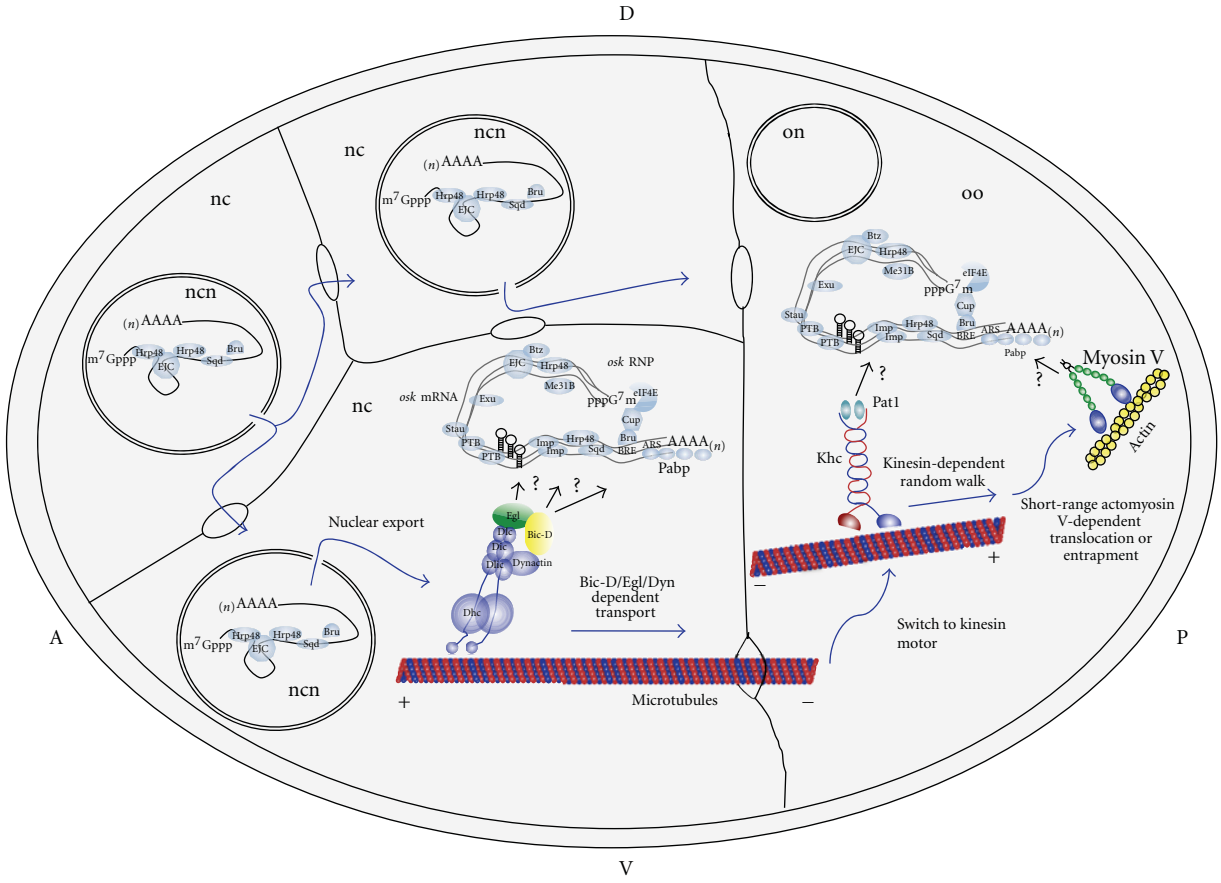


Figure 2: Transport and translation repression of *osk* mRNA during *Drosophila* oogenesis. *osk* mRNA is synthesized in the nucleus (ncn) of the nurse cells (nc) and exported already as a complex with several factors controlling its transport and/or translation (light blue circles), like the exon junction complex (EJC, composed of Mago-Nashi/Y14/eIF4AIII), Hrp48, Bru and Sqd. In the nc cytoplasm more factors controlling translation (Me31B, Cup, Bru, PTB, Imp), localization (Stau, Exu, Sqd, Btz, Pabp) or stability (Pabp) (light blue circles) associate with *osk* mRNA to form a big RNP complex (light blue circles). This RNP contains many *osk* mRNA molecules and multiple factors that repress translation of *osk* by several different mechanisms (see text for details). This big silenced *osk* RNP is recruited by the Bic-D/Egl/Dyn localization machinery which directs its minus end directed microtubule transport in the nurse cell cytoplasm and through the ring canals into the oocyte (oo). Factors linking *osk* RNPs to the transport machinery are not known. Since Egl binds directly some other localized mRNAs, Egl may be the linking factor. Other proteins in complex with Bic-D, such as Pabp, (which binds directly to *osk* mRNA through adenine rich sequences (ARS) and the poly-(A) tail) could also be involved. Within the oocyte the silenced *osk* RNP is then transported by a kinesin motor probably by a random walk process in a poorly polarized microtubule network with a net movement toward the posterior cortex. This movement is followed by a short-range actomyosin-dependent transport or entrapment of *osk* mRNA to the posterior cortex. During its journey *osk* mRNA associates with the factors repressing its translation. Although different proteins may associate with *osk* during different stages of oogenesis, most of them are probably associated with it during its all trip to the posterior. When *osk* mRNA reaches the posterior cortex at stage 9 of oogenesis, translation repression is relieve and the mRNA gets translated (not shown in this figure).

show precocious expression of *osk* mRNA in stages 6 to 9 and also increased expression in stage 9 oocytes. Another mechanism, independent of the Cup-eIF4E interaction but dependent on Bru, also drives translation repression during mid oogenesis. This mechanism involves the formation of *osk* mRNA oligomers by binding of Bru that produces the formation of large (50S-80S) silencing particles that cannot be accessed by ribosomes [85]. Hrp48 binds sequences in the *osk* 5′ and 3′UTRs and has also been involved in localization and translational repression of *osk* mRNA after stage 9 of oogenesis, although how Hrp48 regulates translation is still not known [86].

Interestingly, Cup was shown to be also involved in translational repression of *grk* mRNA, which is also transported by the Bic-D/Egl complex. Based on genetic and biochemical interactions studies, Clouse et al. proposed a model for translation regulation of *grk* mRNA [87]. In their model, Cup and Bruno also function in complex with Sqd, Otu and Hrb27C/Hrp48 in repressing translation of *grk* mRNA before it is localized. While this is not proven yet, this repression would also appear to act at the level of translation initiation. They also showed that before the RNA has reached its final destination in the future dorsal-anterior region of the oocyte, a well-established translation factor, the poly(A)-binding

protein (PABP), functions with Encore (Enc) to facilitate the translational activation of *grk* mRNA [87].

Our group recently reported that *Drosophila pabp* genetically interacts with *Bic-D* and that the two proteins form an RNA-dependent complex. *pabp* mutants show reduced *osk* mRNA stability and display defects in *osk* mRNA localization during early oogenesis. These findings demonstrated that PABP plays a key role in *osk* mRNA localization and is also essential in the germline for oocyte growth [88]. The recent finding that mammalian PABP can bind to microtubules [89] also hints that PABP links *osk* mRNA to the transport machinery in addition to controlling its RNA stability during transport. Although it seems that PABP is not involved in controlling translation during early oogenesis, a study of the role of PABP in activating translation of *osk* mRNA after it has reached its final destination is still missing due to the lack of *pabp* mutants that specifically affect late oogenesis. Another factor that may play a role in controlling translation and localization of both *grk* and *osk* mRNAs is the insulin-like growth factor II mRNA-binding protein (IMP) [90, 91]. However, genetic studies so far failed to reveal such a requirement for IMP, indicating that its function is at best a redundant one [90, 91]. In summary, Cup, PABP, IMP, Bruno, and Hrp48 are factors that can associate with the Bic-D/Egl/Dyn transport machinery to regulate the fate and translation of specific transported mRNAs.

How Conserved Is the Bic-D/Egl Complex across Eukaryotes? Studies on the functional role of Bic-D homologs in different species suggest that Bic-D proteins are coiled-coil proteins that function as factors linking the Dyn/Dynactin minus-end-directed motor complex with different cargos [52]. Besides its role in the aforementioned Bic-D/Egl/Dyn RNA transport machinery, *Drosophila* Bic-D is also involved in lipid droplet transport [92], migration of photoreceptor cell nuclei [93], movement of the oocyte nucleus [94], transport of Chc and synaptic vesicle recycling at the neuromuscular junction [95]. *Drosophila* Bic-D also binds an RNA binding protein, the mental retardation protein (FMRP), and both are required for efficient branching of the dendritic arbour [96]. In mammals, Bic-D proteins are required for anchoring the centrosomes to the microtubules [97]. Mammalian Bic-D2 associates with RanBP2, a component of the nuclear pore complex, and is needed to regulate centrosome and nuclear positioning during mitotic entry [98]. By binding Rab6, mammalian Bic-D also controls COPI-independent Golgi-ER transport [99], and Rab6B-Bic-D1 interaction regulates retrograde membrane transport in neurites of human neuronal cells [100]. Like *D. melanogaster* Bic-D, *C. elegans* Bic-D is also involved in nuclear migration [101] and in dendritic branching [102]. Altogether, these studies show that Bic-D acts as a modulator of the Dyn transport complex in different organisms, linking different, but sometimes conserved cargos, such as mRNAs, nuclei, and vesicles.

The *Bic-D* gene is conserved throughout the animal kingdom, but is not present in plants and fungi. While there is only one gene encoding Bic-D in insects, *C. elegans* and the ascidians *Ciona intestinalis* and *Ciona savignyi*, the gene is duplicated in higher vertebrate lineages, including mammals (human, mouse, and gorilla) and birds (chicken). Accordingly, the two homologs of *Drosophila Bic*-D were named *Bic-D1* and *Bic-D2*. In the amphibian *Xenopus,* one *Bic-D1* and two *Bic-D2* homologs are found. Interestingly, fishes (*Danio rerio, Gasterosteus aculeatus, Oryzias latipes, takifugu rubripes, Tetraodon nigroviridis*) have two homologues of the *Bic-D1* gene and two homologs of *Bic-D2*. In addition, in fishes there is also a third, deeply divergent gene, probably representing an ancestral version of the *Bic-D* gene. In the sea lamprey *Petromyzon marinus* there are also two *Bic-D* genes, one *Bic-D1* ortholog and one that also seems to be close to the original ancestor *Bic-D* gene (taken from http://cegg.unige.ch/orthodb [103]).

Recent studies in the wasp *Nasonia vitripennis* point to a conserved role for Bic-D in mRNA localization in nondipteran insects [104]. Knocking down *Bic-D* by RNAi in the *Nasonia* germ line produced oogenesis phenotypes similar to the ones observed in *Drosophila Bic-D* mutants. More importantly, mRNAs that localize to the *Nasonia* oocyte also fail to localize to their normal destination in *Bic-D* loss-of-function animals. These studies strongly suggest that the role of Bic-D in the localization of oocyte determinants, which also involved the organization of a polarized microtubule network, is conserved between *Nasonia* and *Drosophila.* Even though these insects share a similar germ line development, evolutionary they diverged over 200 million years ago. Thus, although no other examples of Bic-D-dependent mRNA transport in other species have been investigated, the high conservation of Bic-D proteins in the animal kingdom suggests that Bic-D proteins have played a conserved role in mRNA transport during evolution. The study of the biological roles of Bic-D in different eukaryotes is an interesting field that deserves further investigation.

In contrast to *Bic-D, egl* is not present in mammals, and only one homolog is found in *D. melanogaster* and *C. elegans.* So far, functional studies on *egl* have been restricted to *D. melanogaster,* but recent studies on the giant shrimp (*Penaeus monodon) egl* ortholog suggested an involvement in ovary development as well [105]. Since *Drosophila* Egl is an adaptor that binds directly to localization signals in mRNAs, most likely Bic-D/Egl complexes function in mRNA localization only in Arthropoda and Nematoda, and it appears that other adaptor proteins not related or only distantly related to Egl may link the Bic-D/Dyn localization machinery to localizing mRNAs in other phyla.

Bic-D is highly conserved across the animal kingdom and other highly conserved RNA-binding proteins that play roles in RNA localization, such as PABP [88], FMRP [96] and other proteins (Vazquez and Suter, unpublished results) are also present in Bic-D complexes. This suggests that different adaptor proteins may be linking the transported mRNAs to the transport machinery in a species-specific manner, as well as in a mRNA- or tissue-specific manner. One of these proteins might be PABP. Cytoplasmic PABPs are general translation factors and are conserved throughout eukaryotes [106]. One cytoplasmic *pabp* gene is present in the unicellular fungi *Candida albicans, Saccharomyces cerevisiae,* and *Schizosaccharomyces pombe*, as well as in

D. melanogaster. In contrast, vertebrates contain multiple cytoplasmic PABPs. They include PABP1 (also known as PABP, PAB1, PAB, and PABPC1), PABP4 (also called PABPC4, iPABP or APP-1), ePABP (embryonic PABP), ePABP2, and the mammalian-specific tPABP (testis-specific PABP, also called PABPC2 in mouse and PABPC3 in humans) [107, 108]. A shorter version of these PABPs called PABP5 or PABP5C is also present in higher eukaryotes, and it is highly conserved in primates, rodents, and humans [109]. Counting all the family members, eight genes are present in *Arabidopsis*, four in *Zebrafish*, Chicken, and *Xenopus Tropicalis,* seven in humans, and six in mouse (taken from http://cegg.unige.ch/orthodb [103, 110]). To date, most of the functional studies have been focused on the prototype PABP1; however, the versatility and the high number of genes encoding PABP family members open the possibility that different PABPs may regulate localization and/or translation of different transported mRNAs. A conserved role for PABP in RNA localization is supported by the recent finding that PABP binds directly or in a complex with other proteins to non-poly(A) sequences in the *osk*, *bcd,* and *Vasopressin* mRNAs, which are essential for correct localization of these transcripts in *Drosophila* oocytes and mammalian dendrites, respectively [88, 111–114]. Furthermore, yeast Pab1p is required to restrict *ASH1* mRNA to the bud tip, indicating that the role of these proteins in RNA localization is also conserved in unicellular eukaryotes [115].

FMRP is also a highly conserved protein, displaying 92% amino acid identity between humans and chicken [116]. In humans and mouse there are three paralogous proteins (namely, FMRP, FXR1 and FXR2) [117–120]. The three genes have a conserved gene structure suggesting they may be derived from a common ancestor [120]. Zebrafish possess also three FMR1-related genes that are orthologous to the human and murine ones [121]. In *Drosophila* there is only one single orthologous gene that has higher overall similarity to human FXR2 than to FMR1 or FXR1 [122]. Several lines of evidence prove that FMRP orthologs are involved in RNA localization and translation control [123]. FMRP colocalizes and immunoprecipitates with several dendritically localized mRNAs in mammalian neurons [123]. FMRP knock down mice show an excess of protein synthesis and loss of stimulus-induced translation of some localized mRNAs as well as a failure to augment trafficking of certain mRNAs in neurons upon mGluR activation, indicating that FMRP is crucial for transport and regulation of local translation of certain mRNAs at the synapses [123]. The recent finding that *Drosophila* FMRP binds to Bic-D and that both cooperate to control dendrite morphogenesis [96] and that FMRP controls RNA transport in neurons [124] suggests that the role of FMRP as a link between the transport machinery and localizing neuronal mRNAs may be conserved between mammals and *Drosophila*.

Another protein that is required for translational control of localized *osk* and *grk* mRNAs is the insulin-like growth factor II mRNA-binding protein (IMP) [90, 91]. Preliminary results from our lab indicate that *Drosophila* IMP is a component of the Bic-D/Egl complex (Vazquez-Pianzola, Bullock and Suter, unpublished). IMP is highly conserved in the animal kingdom. Most likely the vertebrate IMP family originated by repeated gene duplications shortly after the divergence of vertebrates from other major metazoan clades. This is supported by the finding that *D. melanogaster, C. elegans* and the ascidians *C. intestinalis, and C. savignyi* have only one gene, whereas most vertebrates possess more than one ortholog. In most mammals (i.e., human, rat, mice), birds (Chicken), and reptiles (*Anolis Carolinensis*), three IMPs (namely, IMP1, IMP2 and IMP3) are present. Interestingly, *Gorilla* and the fish *D. Rerio* have four orthologous genes, the additional one being most closely related to mammalian IMP2. Mammalian IMP1 is most closely related to *Drosophila* IMP. The amphibian *Xenopus Tropicalis* contains only one IMP gene, homologous to mammalian IMP3, which was originally named Vg1 RNA-binding protein (Vg1RBP/Vera) (Taken from http://cegg.unige.ch/orthodb [103, 125]). These proteins are paradigms of RNA binding proteins required for transport and local translation of RNAs. The chicken IMP1, also known as the zipcode-binding protein (ZBP-1), is required for beta-actin mRNA localization and translational repression during transport to the leading edge of motile fibroblasts and neurons, while the Vg1RBP/Vera is required for Vg1 mRNAs localization to the vegetal Pole of the *xenopus* oocytes during maturation [28, 125]. These observations show that IMP proteins play a function that has been conserved during animal evolution.

6. Concluding Remarks

Many studies have been performed on the factors regulating translation of specific mRNAs while transported to their destination. One conclusion from these reports is that many of them, such as Pufs, FMRPs, IMP and PABP proteins, have been highly conserved during evolution and that their roles in translation also seem to be conserved across eukaryotes. The "locasome" in yeast and the Bic-D/Egl localization machinery in *Drosophila* seem to be general links between the RNA-transporting molecular motors and the translation machinery, acting either via myosin or Dyn/Dynactin motors. The "locasome" seems to be only present in unicellular fungi while Bic-D proteins are conserved in the animal kingdom. Thus, further studies of Bic-D proteins in RNA transport in other animals, including humans, will shed light on the question if the mechanisms of RNA transport are indeed conserved over the entire animal kingdom.

The reports that even in non-polarized cells mRNAs are localized to different compartments, such as the vicinity of mitochondria, peroxisomes, spindles or ER, raise the question of how these mRNAs are transported. Regarding this, it is known that some mRNAs are localized independent of translation and that 3′-UTR regions of many mRNAs are involved in their localization to these subcellular compartments. However, only few studies of the factors involved in these processes have been performed. Another intriguing question that remains to be investigated is whether mRNA localization to the vicinity of these organelles involves an active transport or just a diffusion mechanism followed

by anchoring of the mRNAs through factors, such as specific RNA binding proteins localized to these structures. Thus, high-throughput *in situ* hybridization screens and proteomics approaches of different subcellular fractions are needed to shed light on the existence of new subcellular compartments and the common features of the RNAs targeted to them. In plants, some RNAs travel between cells and in the phloem, but the study of the factors controlling their transport and translation is still scarce. Neither the locasome nor the Bic-D transport machineries are conserved in plants, pointing to novel, so-far unknown RNA transport players awaiting discovery in these organisms. Extensive studies on mRNA translation should also be done in non-model organisms. To our knowledge, nothing is known about subcellular localization of RNAs and translational control in protists, even though examples of localized mRNAs have been described in bacteria, animals, fungi and plants. This strongly suggests that subcellular localization of mRNAs is an essential process that most likely is required for most forms of life, and that the mechanisms of subcellular localization of RNAs were conserved during evolution. It is worth testing whether this process also functions in archeal lineages, since some of them are believed to be current representatives of the eukaryotic ancestors.

Many of the proteins controlling translation during transport of their target mRNAs seem to repress translation at the level of translation initiation either competing for the formation of the eIF4E complex or inhibiting 60S subunit joining. This makes sense, since translation initiation is the limiting step in all the translation process indicating that it must be tightly regulated. Phosphorylation of the yeast translational repressors Puf6 and Khd1p and the chicken ZBP by specific kinases localized at the mRNA final destination is involved in local translational de-repression of their targets [34, 35, 126]. The presence of similar mechanisms of translation control of localized mRNAs in unicellular and higher eukaryotes, may indicate that the control of translation initiation, the presence of locally expressed kinases and the phosphorylation status of the RNA-binding proteins are conserved features used for the RNA localization machineries during evolution to control translation of localized mRNAs.

Elucidating the global composition of different RNP-complexes and identifying the factors that are common and the ones that are specific to sort individual mRNAs to the different subcellular compartments is an interesting and important question for future research in the field. Similarly, elucidating in detail the mechanisms that are in place to couple mRNA localization to local protein synthesis across different eukaryotes is another fascinating question to tackle.

Acknowledgments

The specially thank Greco Hernández for valuable comments on the paper and useful discussions. This work was supported by the Swiss National Science Foundation and the Canton of Bern.

References

[1] L. Vardy and T. L. Orr-Weaver, "Regulating translation of maternal messages: multiple repression mechanisms," *Trends in Cell Biology*, vol. 17, no. 11, pp. 547–554, 2007.

[2] T. G. Du, M. Schmid, and R. P. Jansen, "Why cells move messages: the biological functions of mRNA localization," *Seminars in Cell and Developmental Biology*, vol. 18, no. 2, pp. 171–177, 2007.

[3] C. E. Holt and S. L. Bullock, "Subcellular mRNA localization in animal cells and why it matters," *Science*, vol. 326, no. 5957, pp. 1212–1216, 2009.

[4] K. C. Martin and A. Ephrussi, "mRNA localization: gene expression in the spatial dimension," *Cell*, vol. 136, no. 4, pp. 719–730, 2009.

[5] K. Nevo-Dinur, A. Nussbaum-Shochat, S. Ben-Yehuda, and O. Amster-Choder, "Translation-independent localization of mRNA in *E. coli*," *Science*, vol. 331, no. 6020, pp. 1081–1084, 2011.

[6] T. K. Hyun, M. N. Uddin, Y. Rim, and J. Y. Kim, "Cell-to-cell trafficking of RNA and RNA silencing through plasmodesmata," *Protoplasma*, vol. 248, no. 1, pp. 101–116, 2011.

[7] J. Kehr and A. Buhtz, "Long distance transport and movement of RNA through the phloem," *Journal of Experimental Botany*, vol. 59, no. 1, pp. 85–92, 2008.

[8] W. J. Lucas, S. Bouché-Pillon, D. P. Jackson et al., "Selective trafficking of KNOTTED1 homeodomain protein and its mRNA through plasmodesmata," *Science*, vol. 270, no. 5244, pp. 1980–1983, 1995.

[9] J. Y. Kim, Y. Rim, J. Wang, and D. Jackson, "A novel cell-to-cell trafficking assay indicates that the KNOX homeodomain is necessary and sufficient for intercellular protein and mRNA trafficking," *Genes and Development*, vol. 19, no. 7, pp. 788–793, 2005.

[10] J. M. Kugler and P. Lasko, "Localization, anchoring and translational control of *oskar, gurken, bicoid* and *nanos* mRNA during *Drosophila* oogenesis," *Fly*, vol. 3, no. 1, pp. 15–28, 2009.

[11] M. L. King, T. J. Messitt, and K. L. Mowry, "Putting RNAs in the right place at the right time: RNA localization in the frog oocyte," *Biology of the Cell*, vol. 97, no. 1, pp. 19–33, 2005.

[12] G. Kumano, "Polarizing animal cells via mRNA localization in oogenesis and early development," *Development, Growth and Differentiation*, vol. 54, no. 1, pp. 1–18, 2011.

[13] C. J. Donnelly, M. Fainzilber, and J. L. Twiss, "Subcellular communication through RNA transport and localized protein synthesis," *Traffic*, vol. 11, no. 12, pp. 1498–1505, 2010.

[14] J. Condeelis and R. H. Singer, "How and why does *β-actin* mRNA target?" *Biology of the Cell*, vol. 97, no. 1, pp. 97–110, 2005.

[15] T. Eom, L. N. Antar, R. H. Singer, and G. J. Bassell, "Localization of a *β-actin* messenger ribonucleoprotein complex with zipcode-binding protein modulates the density of dendritic filopodia and filopodial synapses," *Journal of Neuroscience*, vol. 23, no. 32, pp. 10433–10444, 2003.

[16] H. L. Zhang, T. Eom, Y. Oleynikov et al., "Neurotrophin-induced transport of a *β-actin* mRNP complex increases *β-actin* levels and stimulates growth cone motility," *Neuron*, vol. 31, no. 2, pp. 261–275, 2001.

[17] J. Yao, Y. Sasaki, Z. Wen, G. J. Bassell, and J. Q. Zheng, "An essential role for *β-actin* mRNA localization and translation in Ca^{2+}-dependent growth cone guidance," *Nature Neuroscience*, vol. 9, no. 10, pp. 1265–1273, 2006.

[18] K. M. Leung, F. P. Van Horck, A. C. Lin, R. Allison, N. Standart, and C. E. Holt, "Asymmetrical *β-actin* mRNA translation in growth cones mediates attractive turning to netrin-1," *Nature Neuroscience*, vol. 9, no. 10, pp. 1247–1256, 2006.

[19] E. Lécuyer, H. Yoshida, N. Parthasarathy et al., "Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function," *Cell*, vol. 131, no. 1, pp. 174–187, 2007.

[20] P. Marc, A. Margeot, F. Devaux, C. Blugeon, M. Corral-Debrinski, and C. Jacq, "Genome-wide analysis of mRNAs targeted to yeast mitochondria," *EMBO Reports*, vol. 3, no. 2, pp. 159–164, 2002.

[21] M. Corral-Debrinski, C. Blugeon, and C. Jacq, "In yeast, the 3′ untranslated region or the presequence of *ATM1* is required for the exclusive localization of its mRNA to the vicinity of mitochondria," *Molecular and Cellular Biology*, vol. 20, no. 21, pp. 7881–7892, 2000.

[22] B. Pyhtila, T. Zheng, P. J. Lager, J. D. Keene, M. C. Reedy, and C. V. Nicchitta, "Signal sequence- and translation-independent mRNA localization to the endoplasmic reticulum," *RNA*, vol. 14, no. 3, pp. 445–453, 2008.

[23] G. Zipor, L. Haim-Vilmovsky, R. Gelin-Licht, N. Gadir, C. Brocard, and J. E. Gerst, "Localization of mRNAs coding for peroxisomal proteins in the yeast, *Saccharomyces cerevisiae*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 47, pp. 19848–19853, 2009.

[24] J. A. Sharp, J. J. Plant, T. K. Ohsumi, M. Borowsky, and M. D. Blower, "Functional analysis of the microtubule-interacting transcriptome," *Molecular Biology of the Cell*, vol. 22, no. 22, pp. 4312–4323, 2011.

[25] R. G. Heym and D. Niessing, "Principles of mRNA transport in yeast," *Cellular and Molecular Life Sciences*, vol. 69, pp. 1843–1853, 2012.

[26] P. A. Takizawa, A. Sil, J. R. Swedlow, I. Herskowitz, and R. D. Vale, "Actin-dependent localization of an RNA encoding a cell-fate determinant in the yeast," *Nature*, vol. 389, no. 6646, pp. 90–93, 1997.

[27] R. M. Long, R. H. Singer, X. Meng, I. Gonzalez, K. Nasmyth, and R. P. Jansen, "Mating type switching in yeast controlled by asymmetric localization of *ASH1* mRNA," *Science*, vol. 277, no. 5324, pp. 383–387, 1997.

[28] N. Paquin and P. Chartrand, "Local regulation of mRNA translation: new insights from the bud," *Trends in Cell Biology*, vol. 18, no. 3, pp. 105–111, 2008.

[29] K. A. Shepard, A. P. Gerber, A. Jambhekar et al., "Widespread cytoplasmic mRNA transport in yeast: identification of 22 bud-localized transcripts using DNA microarray analysis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 20, pp. 11429–11434, 2003.

[30] M. Oeffinger, K. E. Wei, R. Rogers et al., "Comprehensive analysis of diverse ribonucleoprotein complexes," *Nature Methods*, vol. 4, no. 11, pp. 951–956, 2007.

[31] S. M. Landers, M. R. Gallas, J. Little, and R. M. Long, "She3p possesses a novel activity required for *ASH1* mRNA localization in *Saccharomyces cerevisiae*," *Eukaryotic Cell*, vol. 8, no. 7, pp. 1072–1083, 2009.

[32] T. Quenault, T. Lithgow, and A. Traven, "PUF proteins: repression, activation and mRNA localization," *Trends in Cell Biology*, vol. 21, no. 2, pp. 104–112, 2011.

[33] W. Gu, Y. Deng, D. Zenklusen, and R. H. Singer, "A new yeast PUF family protein, Puf6p, represses *ASH1* mRNA translation and is required for its localization," *Genes and Development*, vol. 18, no. 12, pp. 1452–1465, 2004.

[34] Y. Deng, R. H. Singer, and W. Gu, "Translation of *ASH1* mRNA is repressed by Puf6p-Fun12p/eIF5B interaction and released by CK2 phosphorylation," *Genes and Development*, vol. 22, no. 8, pp. 1037–1050, 2008.

[35] N. Paquin, M. Ménade, G. Poirier, D. Donato, E. Drouet, and P. Chartrand, "Local Activation of Yeast *ASH1* mRNA Translation through Phosphorylation of Khd1p by the Casein Kinase Yck1p," *Molecular Cell*, vol. 26, no. 6, pp. 795–809, 2007.

[36] K. Irie, T. Tadauchi, P. A. Takizawa, R. D. Vale, K. Matsumoto, and I. Herskowitz, "The Khd1 protein, which has three KH RNA-binding motifs, is required for proper localization of *ASH1* mRNA in yeast," *The EMBO Journal*, vol. 21, no. 5, pp. 1158–1167, 2002.

[37] P. P. C. Tam, I. H. Barrette-Ng, D. M. Simon, M. W. C. Tam, A. L. Ang, and D. G. Muench, "The Puf family of RNA-binding proteins in plants: phylogeny, structural modeling, activity and subcellular localization," *BMC Plant Biology*, vol. 10, article 44, 2010.

[38] A. C. Goldstrohm, D. J. Seay, B. A. Hook, and M. Wickens, "PUF protein-mediated deadenylation is catalyzed by Ccr4p," *The Journal of Biological Chemistry*, vol. 282, no. 1, pp. 109–114, 2007.

[39] B. A. Hook, A. C. Goldstrohm, D. J. Seay, and M. Wickens, "Two yeast PUF proteins negatively regulate a single mRNA," *The Journal of Biological Chemistry*, vol. 282, no. 21, pp. 15430–15438, 2007.

[40] N. Suh, S. L. Crittenden, A. Goldstrohm et al., "FBF and its dual control of *gld-1* expression in the *Caenorhabditis elegans* germline," *Genetics*, vol. 181, no. 4, pp. 1249–1260, 2009.

[41] L. Y. Kadyrova, Y. Habara, T. H. Lee, and R. P. Wharton, "Translational control of maternal *Cyclin B* mRNA by Nanos in the *Drosophila* germline," *Development*, vol. 134, no. 8, pp. 1519–1527, 2007.

[42] A. C. Goldstrohm, B. A. Hook, D. J. Seay, and M. Wickens, "PUF proteins bind Pop2p to regulate messenger RNAs," *Nature Structural and Molecular Biology*, vol. 13, no. 6, pp. 533–539, 2006.

[43] P. F. Cho, C. Gamberi, Y. Cho-Park, I. B. Cho-Park, P. Lasko, and N. Sonenberg, "Cap-dependent translational inhibition establishes two opposing morphogen gradients in *Drosophila* embryos," *Current Biology*, vol. 16, no. 20, pp. 2035–2041, 2006.

[44] Q. Cao, K. Padmanabhan, and J. D. Richter, "Pumilio 2 controls translation by competing with eIF4E for 7-methyl guanosine cap recognition," *RNA*, vol. 16, no. 1, pp. 221–227, 2010.

[45] N. Gadir, L. Haim-Vilmovsky, J. Kraut-Cohen, and J. E. Gerst, "Localization of mRNAs coding for mitochondrial proteins in the yeast *Saccharomyces cerevisiae*," *RNA*, vol. 17, no. 8, pp. 1551–1565, 2011.

[46] Y. Saint-Georges, M. Garcia, T. Delaveau et al., "Yeast mitochondrial biogenesis: a role for the PUF RNA-binding protein puf3p in mRNA localization," *PLoS One*, vol. 3, no. 6, Article ID e2293, 2008.

[47] A. P. Gerber, D. Herschlag, and P. O. Brown, "Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast," *PLoS Biology*, vol. 2, no. 3, article E79, 2004.

[48] J. P. Vessey, A. Vaccani, Y. Xie et al., "Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules," *Journal of Neuroscience*, vol. 26, no. 24, pp. 6496–6508, 2006.

[49] J. P. Vessey, L. Schoderboeck, E. Gingl et al., "Mammalian Pumilio 2 regulates dendrite morphogenesis and synaptic function," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 7, pp. 3222–3227, 2010.

[50] K. P. Menon, S. Sanyal, Y. Habara et al., "The translational repressor Pumilio regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E," *Neuron*, vol. 44, no. 4, pp. 663–676, 2004.

[51] S. L. Bullock, "Translocation of mRNAs by molecular motors: think complex?" *Seminars in Cell and Developmental Biology*, vol. 18, no. 2, pp. 194–201, 2007.

[52] M. Claußen and B. Suter, "BicD-dependent localization processes: from Drosophilia development to human cell biology," *Annals of Anatomy*, vol. 187, no. 5-6, pp. 539–553, 2005.

[53] B. Ran, R. Bopp, and B. Suter, "Null alleles reveal novel requirements for Bic-D during *Drosophila* oogenesis and zygotic development," *Development*, vol. 120, no. 5, pp. 1233–1242, 1994.

[54] B. Suter and R. Steward, "Requirement for phosphorylation and localization of the Bicaudal-D protein in *Drosophila* oocyte differentiation," *Cell*, vol. 67, no. 5, pp. 917–926, 1991.

[55] T. Schupbach and E. Wieschaus, "Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology," *Genetics*, vol. 129, no. 4, pp. 1119–1136, 1991.

[56] W. E. Theurkauf, B. M. Alberts, Y. N. Jan, and T. A. Jongens, "A central role for microtubules in the differentiation of *Drosophila* oocytes," *Development*, vol. 118, no. 4, pp. 1169–1180, 1993.

[57] A. Clark, C. Meignin, and I. Davis, "A Dynein-dependent shortcut rapidly delivers axis determination transcripts into the *Drosophila* oocyte," *Development*, vol. 134, no. 10, pp. 1955–1965, 2007.

[58] J. R. Hughes, S. L. Bullock, and D. Ish-Horowicz, "Inscuteable mRNA localization is dynein-dependent and regulates apicobasal polarity and spindle length in *Drosophila* neuroblasts," *Current Biology*, vol. 14, no. 21, pp. 1950–1956, 2004.

[59] S. L. Bullock and D. Ish-Horowicz, "Conserved signals and machinery for RNA transport in *Drosophila* oogenesis and embryogenesis," *Nature*, vol. 414, no. 6864, pp. 611–616, 2001.

[60] C. C. Hoogenraad, A. Akhmanova, S. A. Howell et al., "Mammalian golgi-associated Bicaudal-D2 functions in the dynein-dynactin pathway by interacting with these complexes," *The EMBO Journal*, vol. 20, no. 15, pp. 4041–4054, 2001.

[61] C. Navarro, H. Puthalakath, J. M. Adams, A. Strasser, and R. Lehmann, "Egalitarian binds dynein light chain to establish oocyte polarity and maintain oocyte fate," *Nature Cell Biology*, vol. 6, no. 5, pp. 427–435, 2004.

[62] M. Dienstbier, F. Boehl, X. Li, and S. L. Bullock, "Egalitarian is a selective RNA-binding protein linking mRNA localization signals to the dynein motor," *Genes and Development*, vol. 23, no. 13, pp. 1546–1558, 2009.

[63] S. L. Bullock, A. Nicol, S. P. Gross, and D. Zicha, "Guidance of bidirectional motor complexes by mRNA cargoes through control of dynein number and activity," *Current Biology*, vol. 16, no. 14, pp. 1447–1452, 2006.

[64] R. P. Brendza, L. R. Serbus, J. B. Duffy, and W. M. Saxton, "A function for kinesin I in the posterior transport of *oskar* mRNA and staufen protein," *Science*, vol. 289, no. 5487, pp. 2120–2122, 2000.

[65] I. M. Palacios and D. St Johnston, "Kinesin light chain-independent function of the Kinesin heavy chain in cytoplasmic streaming and posterior localisation in the *Drosophila* oocyte," *Development*, vol. 129, no. 23, pp. 5473–5485, 2002.

[66] P. Loiseau, T. Davies, L. S. Williams, M. Mishima, and I. M. Palacios, "*Drosophila* PAT1 is required for Kinesin-1 to transport cargo and to maximize its motility," *Development*, vol. 137, no. 16, pp. 2763–2772, 2010.

[67] V. L. Zimyanin, K. Belaya, J. Pecreaux et al., "In vivo imaging of *oskar* mRNA transport reveals the mechanism of posterior localization," *Cell*, vol. 134, no. 5, pp. 843–853, 2008.

[68] J. Krauss, S. López de Quinto, C. Nüsslein-Volhard, and A. Ephrussi, "Myosin-V Regulates oskar mRNA Localization in the *Drosophila* Oocyte," *Current Biology*, vol. 19, no. 12, pp. 1058–1063, 2009.

[69] H. A. Cook, B. S. Koppetsch, J. Wu, and W. E. Theurkauf, "The *Drosophila* SDE3 homolog armitage is required for oskar mRNA silencing and embryonic axis specification," *Cell*, vol. 116, no. 6, pp. 817–829, 2004.

[70] A. Pane, K. Wehr, and T. Schüpbach, "*zucchini* and *squash* encode two putative nucleases required for rasiRNA production in the *Drosophila* germline," *Developmental Cell*, vol. 12, no. 6, pp. 851–862, 2007.

[71] A. K. Lim and T. Kai, "Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila* melanogaster," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 16, pp. 6714–6719, 2007.

[72] Y. Tomari, T. Du, B. Haley et al., "RISC assembly defects in the *Drosophila* RNAi mutant armitage," *Cell*, vol. 116, no. 6, pp. 831–841, 2004.

[73] A. A. Aravin, N. M. Naumova, A. V. Tulin, V. V. Vagin, Y. M. Rozovsky, and V. A. Gvozdev, "Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. *melanogaster* germline," *Current Biology*, vol. 11, no. 13, pp. 1017–1027, 2001.

[74] W. Stapleton, S. Das, and B. D. McKee, "A role of the *Drosophila homeless* gene in repression of Stellate in male meiosis," *Chromosoma*, vol. 110, no. 3, pp. 228–240, 2001.

[75] A. Schmidt, G. Palumbo, M. P. Bozzetti, P. Tritto, S. Pimpinelli, and U. Schäfer, "Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of *Drosophila melanogaster*," *Genetics*, vol. 151, no. 2, pp. 749–760, 1999.

[76] M. Savitsky, D. Kwon, P. Georgiev, A. Kalmykova, and V. Gvozdev, "Telomere elongation is under the control of the RNAi-based mechanism in the *Drosophila* germline," *Genes and Development*, vol. 20, no. 3, pp. 345–354, 2006.

[77] Y. Chen, A. Pane, and T. Schüpbach, "*cutoff* and *aubergine* mutations result in retrotransposon upregulation and checkpoint activation in *Drosophila*," *Current Biology*, vol. 17, no. 7, pp. 637–642, 2007.

[78] V. V. Vagin, A. Sigova, C. Li, H. Seitz, V. Gvozdev, and P. D. Zamore, "A distinct small RNA pathway silences selfish genetic elements in the germline," *Science*, vol. 313, no. 5785, pp. 320–324, 2006.

[79] M. Ghildiyal and P. D. Zamore, "Small silencing RNAs: an expanding universe," *Nature Reviews Genetics*, vol. 10, no. 2, pp. 94–108, 2009.

[80] C. Klattenhoff and W. Theurkauf, "Biogenesis and germline functions of piRNAs," *Development*, vol. 135, no. 1, pp. 3–9, 2008.

[81] A. Nakamura, R. Amikura, K. Hanyu, and S. Kobayashi, "Me31B silences translation of oocyte-localizing RNAs

through the formation of cytoplasmic RNP complex during *Drosophila* oogenesis," *Development*, vol. 128, no. 17, pp. 3233–3242, 2001.

[82] F. Besse, S. López De Quinto, V. Marchand, A. Trucco, and A. Ephrussi, "Drosophila PTB promotes formation of high-order RNP particles and represses *oskar* translation," *Genes and Development*, vol. 23, no. 2, pp. 195–207, 2009.

[83] B. K. Ham, J. L. Brandom, B. Xoconostle-Cázares, V. Ringgold, T. J. Lough, and W. J. Lucas, "A polypyrimidine tract binding protein, pumpkin RBP50, forms the basis of a phloem-mobile ribonucleoprotein complex," *Plant Cell*, vol. 21, no. 1, pp. 197–215, 2009.

[84] A. Nakamura, K. Sato, and K. Hanyu-Nakamura, "*Drosophila* Cup is an eIF4E binding protein that associates with Bruno and regulates *oskar* mRNA Translation in Oogenesis," *Developmental Cell*, vol. 6, no. 1, pp. 69–78, 2004.

[85] M. Chekulaeva, M. W. Hentze, and A. Ephrussi, "Bruno acts as a dual repressor of *oskar* translation, promoting mRNA oligomerization and formation of silencing particles," *Cell*, vol. 124, no. 3, pp. 521–533, 2006.

[86] T. Yano, S. López de Quinto, Y. Matsui, A. Shevchenko, A. Shevchenko, and A. Ephrussi, "Hrp48, a *Drosophila* hnRNPA/B homolog, binds and regulates translation of *oskar* mRNA," *Developmental Cell*, vol. 6, no. 5, pp. 637–648, 2004.

[87] K. N. Clouse, S. B. Ferguson, and T. Schüpbach, "Squid, Cup, and PABP55B function together to regulate *gurken* translation in *Drosophila*," *Developmental Biology*, vol. 313, no. 2, pp. 713–724, 2008.

[88] P. Vazquez-Pianzola, H. Urlaub, and B. Suter, "Pabp binds to the osk 3′UTR and specifically contributes to *osk* mRNA stability and oocyte accumulation," *Developmental Biology*, vol. 357, no. 2, pp. 404–418, 2011.

[89] K. G. Chernov, P. A. Curmi, L. Hamon, A. Mechulam, L. P. Ovchinnikov, and D. Pastré, "Atomic force microscopy reveals binding of mRNA to microtubules mediated by two major mRNP proteins YB-1 and PABP," *FEBS Letters*, vol. 582, no. 19, pp. 2875–2881, 2008.

[90] C. Geng and P. M. Macdonald, "Imp associates with squid and Hrp48 and contributes to localized expression of *gurken* in the oocyte," *Molecular and Cellular Biology*, vol. 26, no. 24, pp. 9508–9516, 2006.

[91] T. P. Munro, S. Kwon, B. J. Schnapp, and D. St. Johnston, "A repeated IMP-binding motif controls *oskar* mRNA translation and anchoring independently of *Drosophila melanogaster* IMP," *Journal of Cell Biology*, vol. 172, no. 4, pp. 577–588, 2006.

[92] K. S. Larsen, J. Xu, S. Cermelli, Z. Shu, and S. P. Gross, "BicaudalD actively regulates microtubule motor activity in lipid droplet transport," *PLoS One*, vol. 3, no. 11, Article ID e3763, 2008.

[93] A. Swan, T. Nguyen, and B. Suter, "*Drosophila Lissencephaly*-1 functions with Bic-D and *dynein* in oocyte determination and nuclear positioning," *Nature Cell Biology*, vol. 1, no. 7, pp. 444–449, 1999.

[94] A. Swan and B. Suter, "Role of *Bicaudal*-D in patterning the *Drosophila* egg chamber in mid-oogenesis," *Development*, vol. 122, no. 11, pp. 3577–3586, 1996.

[95] X. Li, H. Kuromi, L. Briggs et al., "Bicaudal-D binds clathrin heavy chain to promote its transport and augments synaptic vesicle recycling," *The EMBO Journal*, vol. 29, no. 5, pp. 992–1006, 2010.

[96] A. Bianco, M. Dienstbier, H. K. Salter, G. Gatto, and S. L. Bullock, "Bicaudal-D regulates fragile X mental retardation protein levels, motility, and function during neuronal morphogenesis," *Current Biology*, vol. 20, no. 16, pp. 1487–1492, 2010.

[97] K. Fumoto, C. C. Hoogenraad, and A. Kikuchi, "GSK-3β-regulated interaction of BICD with dynein is involved in microtubule anchorage at centrosome," *The EMBO Journal*, vol. 25, no. 24, pp. 5670–5682, 2006.

[98] D. Splinter, M. E. Tanenbaum, A. Lindqvist et al., "Bicaudal D2, dynein, and kinesin-1 associate with nuclear pore complexes and regulate centrosome and nuclear positioning during mitotic entry," *PLoS Biology*, vol. 8, no. 4, Article ID e1000350, 2010.

[99] T. Matanis, A. Akhmanova, P. Wulf et al., "Bicaudal-D regulates COPI-independent Golgi-ER transport by recruiting the dynein-dynactin motor complex," *Nature Cell Biology*, vol. 4, no. 12, pp. 986–992, 2002.

[100] B. F. J. Wanschers, R. van de Vorstenbosch, M. A. Schlager et al., "A role for the Rab6B Bicaudal-D1 interaction in retrograde transport in neuronal cells," *Experimental Cell Research*, vol. 313, no. 16, pp. 3408–3420, 2007.

[101] H. N. Fridolfsson, N. Ly, M. Meyerzon, and D. A. Starr, "UNC-83 coordinates kinesin-1 and dynein activities at the nuclear envelope during nuclear migration," *Developmental Biology*, vol. 338, no. 2, pp. 237–250, 2010.

[102] C. Aguirre-Chen, H. E. Bülow, and Z. Kaprielian, "*C. elegans* bicd-1, homolog of the *Drosophila* dynein accessory factor Bicaudal D, regulates the branching of PVD sensory neuron dendrites," *Development*, vol. 138, no. 3, pp. 507–518, 2011.

[103] R. M. Waterhouse, E. M. Zdobnov, F. Tegenfeldt, J. Li, and E. V. Kriventseva, "OrthoDB: the hierarchical catalog of eukaryotic orthologs in 2011," *Nucleic Acids Research*, vol. 39, pp. D283–D288, 2011.

[104] E. C. Olesnicky and C. Desplan, "Distinct mechanisms for mRNA localization during embryonic axis specification in the wasp *Nasonia*," *Developmental Biology*, vol. 306, no. 1, pp. 134–142, 2007.

[105] R. Preechaphol, S. Klinbunga, B. Khamnamtong, and P. Menasveta, "Isolation and characterization of genes functionally involved in ovarian development of the giant tiger shrimp *Penaeus monodon* by suppression subtractive hybridization (SSH)," *Genetics and Molecular Biology*, vol. 33, no. 4, pp. 676–685, 2010.

[106] G. Hernández, M. Altmann, and P. Lasko, "Origins and evolution of the mechanisms regulating translation initiation in eukaryotes," *Trends in Biochemical Sciences*, vol. 35, no. 2, pp. 63–73, 2010.

[107] H. M. Burgess and N. K. Gray, "mRNA-specific regulation of translation by poly(A)-binding proteins," *Biochemical Society Transactions*, vol. 38, no. 6, pp. 1517–1522, 2010.

[108] B. Gorgoni and N. K. Gray, "The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: a developmental perspective," *Brief Funct Genomic Proteomic*, vol. 3, no. 2, pp. 125–141, 2004.

[109] P. Blanco, C. A. Sargent, C. A. Boucher, G. Howell, M. Ross, and N. A. Affara, "A novel poly(A)-binding protein gene (PABPC5) maps to an X-specific subinterval in the Xq21.3/Yp11.2 Homology block of the human sex chromosomes," *Genomics*, vol. 74, no. 1, pp. 1–11, 2001.

[110] D. A. Mangus, M. C. Evans, and A. Jacobson, "Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression," *Genome Biology*, vol. 4, no. 7, article 223, 2003.

[111] E. A. Arn, B. J. Cha, W. E. Theurkauf, and P. M. Macdonald, "Recognition of a *bicoid* mRNA localization signal by a

protein complex containing Swallow, Nod, and RNA binding proteins," *Developmental Cell*, vol. 4, no. 1, pp. 41–51, 2003.

[112] E. Mohr, N. Prakash, K. Vieluf, C. Fuhrmann, F. Buck, and D. Richter, "*Vasopressin* mRNA localization in nerve cells: characterization of cis-acting elements and trans-acting factors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 13, pp. 7072–7079, 2001.

[113] E. Mohr and D. Richter, "Subcellular *vasopressin* mRNA trafficking and local translation in dendrites," *Journal of Neuroendocrinology*, vol. 16, no. 4, pp. 333–339, 2004.

[114] E. Mohr, C. Fuhrmann, and D. Richter, "VP-RBP, a protein enriched in brain tissue, specifically interacts with the dendritic localizer sequence of rat *vasopressin* mRNA," *European Journal of Neuroscience*, vol. 13, no. 6, pp. 1107–1112, 2001.

[115] M. Trautwein, J. Dengjel, M. Schirle, and A. Spang, "Arf1p provides an unexpected link between COPI vesicles and mRNA in *Saccharomyces cerevisiae*," *Molecular Biology of the Cell*, vol. 15, no. 11, pp. 5021–5037, 2004.

[116] D. K. Price, F. Zhang, C. T. Ashley, and S. T. Warren, "The chicken FMR1 gene is highly conserved with a CCT 5′-untranslated repeat and encodes an RNA-binding protein," *Genomics*, vol. 31, no. 1, pp. 3–12, 1996.

[117] J. F. Coy, Z. Sedlacek, D. Bachner et al., "Highly conserved 3′ UTR and expression pattern of *FXR1* points to a divergent gene regulation of *FXR1* and *FMR1*," *Human Molecular Genetics*, vol. 4, no. 12, pp. 2209–2218, 1995.

[118] M. C. Siomi, H. Siomi, W. H. Sauer, S. Srinivasan, R. L. Nussbaum, and G. Dreyfuss, "*FXR1*, an autosomal homolog of the fragile X mental retardation gene," *The EMBO Journal*, vol. 14, no. 11, pp. 2401–2408, 1995.

[119] Y. Zhang, J. P. O'Connor, M. C. Siomi et al., "The Fragile X Mental Retardation Syndrome protein interacts with novel homologs FXR1 and FXR2," *The EMBO Journal*, vol. 14, no. 21, pp. 5358–5366, 1995.

[120] L. L. Kirkpatrick, K. A. McIlwain, and D. L. Nelson, "Comparative genomic sequence analysis of the *FXR* gene family: *FMR1*, *FXR1*, and *FXR2*," *Genomics*, vol. 78, no. 3, pp. 169–177, 2001.

[121] B. Tucker, R. Richards, and M. Lardelli, "Expression of three zebrafish orthologs of human *FMR1*-related genes and their phylogenetic relationships," *Development Genes and Evolution*, vol. 214, no. 11, pp. 567–574, 2004.

[122] J. Morales, P. R. Hiesinger, A. J. Schroeder et al., "*Drosophila* fragile X protein, DFXR, regulates neuronal morphology and function in the brain," *Neuron*, vol. 34, no. 6, pp. 961–972, 2002.

[123] G. J. Bassell and S. T. Warren, "Fragile X Syndrome: loss of local mRNA regulation alters aynaptic development and function," *Neuron*, vol. 60, no. 2, pp. 201–214, 2008.

[124] P. S. Estes, M. O'Shea, S. Clasen, and D. C. Zarnescu, "Fragile X protein controls the efficacy of mRNA transport in *Drosophila* neurons," *Molecular and Cellular Neuroscience*, vol. 39, no. 2, pp. 170–179, 2008.

[125] F. C. Nielsen, J. Nielsen, and J. Christiansen, "A family of IGF-II mRNA binding proteins (IMP) involved in RNA trafficking," *Scandinavian Journal of Clinical and Laboratory Investigation, Supplement*, vol. 61, no. 234, pp. 93–99, 2001.

[126] S. Hüttelmaier, D. Zenklusen, M. Lederer et al., "Spatial regulation of *β-actin* translation by Src-dependent phosphorylation of ZBP1," *Nature*, vol. 438, no. 7067, pp. 512–515, 2005.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 178525, 11 pages
doi:10.1155/2012/178525

Review Article

Versatility of RNA-Binding Proteins in Cancer

Laurence Wurth

Gene Regulation Programme, Center for Genomic Regulation (CRG) and UPF, 08003 Barcelona, Spain

Correspondence should be addressed to Laurence Wurth, laurence.wurth@crg.es

Received 20 December 2011; Accepted 28 February 2012

Academic Editor: Armen Parsyan

Posttranscriptional gene regulation is a rapid and efficient process to adjust the proteome of a cell to a changing environment. RNA-binding proteins (RBPs) are the master regulators of mRNA processing and translation and are often aberrantly expressed in cancer. In addition to well-studied transcription factors, RBPs are emerging as fundamental players in tumor development. RBPs and their mRNA targets form a complex network that plays a crucial role in tumorigenesis. This paper describes mechanisms by which RBPs influence the expression of well-known oncogenes, focusing on precise examples that illustrate the versatility of RBPs in posttranscriptional control of cancer development. RBPs appeared very early in evolution, and new RNA-binding domains and combinations of them were generated in more complex organisms. The identification of RBPs, their mRNA targets, and their mechanism of action have provided novel potential targets for cancer therapy.

1. Introduction

Traditionally, it has been well accepted that cancer development is dictated in part by aberrant transcriptional events and signaling pathways. More recently, it has become clear that posttranscriptional regulation of gene expression also controls cell proliferation, differentiation, invasion, metastasis, apoptosis, and angiogenesis which influence initiation and progression of cancer [1–4]. Regulation of already transcribed messenger RNAs (mRNAs) is an efficient and rapid way to alter gene expression and plays a crucial role in tumorigenesis.

After transcription, nascent mRNAs undergo several processing steps including splicing, capping, 3′ end formation, surveillance, nucleocytoplasmic transport, and, for many transcripts, localization before being translated and finally degraded [5, 6]. The mRNA does not exist alone in the cell, and its metabolism is largely defined by bound RNA-binding proteins (RBPs). RBPs, which regulate all steps of RNA biogenesis, form dynamic units with the RNA, called ribonucleoprotein complexes (RNPs) [7]. Different sets of RBPs are associated to the mRNA at different time points and in different compartments, thereby regulating the fate of their target in a time- and space-dependent way. RBPs often provide a landing platform for the recruitment of additional factors and enzymes to the mRNA. RBPs are the master regulators of post-transcriptional gene expression and, thus, are expected to play important roles in cancer development [1]. Besides RBPs, the discovery of microRNAs (miRNA) was of great inspiration for the RNA field and provided a new powerful tool to regulate gene expression. miRNAs associate with RBPs to form microRNPs (miRNP) which regulate translation and RNA stability by binding to complementary sequences in target mRNAs. miRNPs have been found to regulate expression of factors implicated in tumorigenesis, but we will not discuss this mechanism here (for recent reviews see [8, 9]).

RBPs bind to specific sequences or secondary structures typically found in the untranslated regions (UTRs) but also in the open reading frame (ORF) of target mRNAs [10, 11]. UTRs in particular have offered more flexibility to evolution, as the constraints of encoding a protein product have not been imposed upon them. As a consequence, diverse and often conserved regulatory elements are present in the UTRs [12]. In the 5′UTR, ribose methylation of the cap structure as well as 5′ terminal polypyrimidine sequences or secondary structures such as internal ribosome entry sites (IRESs) control protein expression. Sequence elements in the 3′UTR regulate the stability of the mRNA, its translational efficiency and localization. Specific binding of regulatory proteins to

these elements is achieved through RNA-binding domains (RBDs). More than 40 RBDs have been identified. Among them, the most prominent are the RNA recognition motif (RRM), K-homology domain (KH), double stranded RNA-binding domain (dsRBD), zinc finger, Arginine-rich domain, cold-shock domain (CSD), and the PAZ and PIWI domains [13]. An RNA-binding protein can contain combinations of different RBDs, which allow a high flexibility for interaction with different targets. RBP purification techniques followed by high throughput proteomics will hopefully allow us in the near future to identify new RNA-binding proteins as well as new RNA-binding domains. Powerful techniques like CLIP-seq (UV cross-linking and immunoprecipitation followed by high throughput sequencing) are helping to identify new RBP targets in a genome wide scale, as well as new RBP binding sites [14–16]. The list of RBPs, RBDs and their targets is far from being complete. New technology is proving helpful to unravel the complexity of post-transcriptional gene regulation.

In cancer cells, expression of numerous oncoproteins or tumor suppressors is under the control of specific RBPs. Splicing, stability, localization as well as translation of these mRNAs are highly regulated, often in a tissue-specific manner [6]. Many RBPs are aberrantly expressed in cancer cells and have thus a cancer-specific regulatory activity [1, 17, 18]. Deregulation of RBP expression in cancer may have its origin on epigenetic events or on miRNA-dependent controls, although the detailed molecular mechanisms are often obscure [19–21]. An additional layer of regulation is provided by signaling: the phosphorylation status of some RBPs is defined by signaling pathways that are deregulated in cancer, and this phosphorylation controls RBP activity and subsequently the expression of its target mRNAs [22, 23]. Signaling pathway alterations occur in different stages of tumor formation and are often correlated with tumor grade.

In this paper, we will summarize the different functions of RBPs in post-transcriptional gene regulation and the impact of aberrant regulation on tumorigenesis. In addition, we will discuss the conservation of specific RBPs across eukaryotes, which may yield hints on how diversity has been generated.

2. RNA-binding Proteins Implicated in Cancer Development

Post-transcriptional gene regulation implies factors which act at different levels of mRNA metabolism, including alternative splicing, localization, stability of the mRNA or cap-dependent and -independent translation. In this section I will introduce a subset of RBPs involved in cancer development which play key roles in each of the steps of RNA regulation, namely, Sam68, eIF4E, La, and HuR to illustrate the powerful RBP regulatory capacity in cancer.

2.1. Sam68 Regulates Alternative Splicing of Cancer-Related mRNAs. Sam68 belongs to the evolutionarily conserved signal transduction and activation of RNA (STAR) family of RBPs [4, 24, 25]. Sam68 is predominately nuclear but has also been detected in the cytoplasm and exerts multiple activities

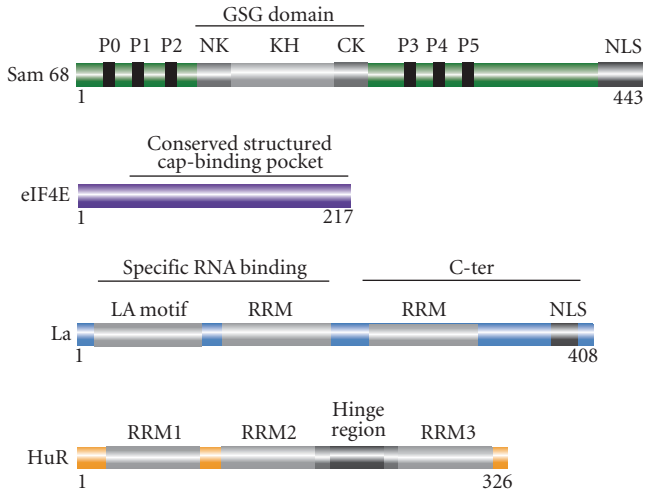


Figure 1: Schematic representation of the 4 RBPs discussed in this paper: Sam68, eIF4E, La, and HuR; RBDs are depicted in light gray. RRM: RNA recognition motif; GSG: GRP33/SAM68/GLD-1 domain, composed of a KH domain (KH) flanked by N-terminal (NK) and C-terminal (CK) extensions; LA: La motif. Phosphorylation sites of La are indicated in black (P0–P5). Nuclear localization signals (NLSs) are represented in dark gray. The number of amino acids of each protein is indicated.

in gene expression, from transcription and signaling to splicing regulation [4, 26]. RNA binding is achieved by a KH domain embedded in a highly conserved region called GSG (GRP/Sam68/GLD1) domain [27] (Figure 1). RNA binding is used for splicing regulation and is modulated by posttranslational modifications, such as phosphorylation or acetylation [22, 25, 28] (Figure 2).

The role of Sam68 in alternative splicing seems directly related to its oncogenic properties. Alternative splicing (AS) allows the majority of human genes to encode for multiple protein isoforms, which often play different or even opposite roles [29]. In addition to the spliceosome, a set of RBPs are necessary to control alternative splicing [7]. Aberrant expression of RBPs in cancer can lead to deregulation of splicing, and subsequent changes in the proteome [30]. The splicing targets of Sam68 support its involvement in tumor progression [4, 31]. Furthermore, the function of Sam68 in AS is regulated by signaling pathways which are often deregulated in cancer cells, establishing a link between signal transduction, alternative splicing, and gene expression during tumorigenesis [22, 32, 33] (Figure 2).

Sam68 is overexpressed in breast, prostate, renal, and cervical cancer cells [26, 34–36] and is also frequently upregulated in tumors [34, 37].

The first hard evidence that Sam68 is involved in regulation of alternative splicing with an impact on tumorigenesis was provided by the demonstration that it promotes inclusion of exon v5 in the CD44 pre-mRNA [33]. CD44 encodes a cell surface molecule involved in cancer cell proliferation. CD44 transcript isoforms are alternatively generated by the inclusion of 10 variant exons, which are decisive in tumor

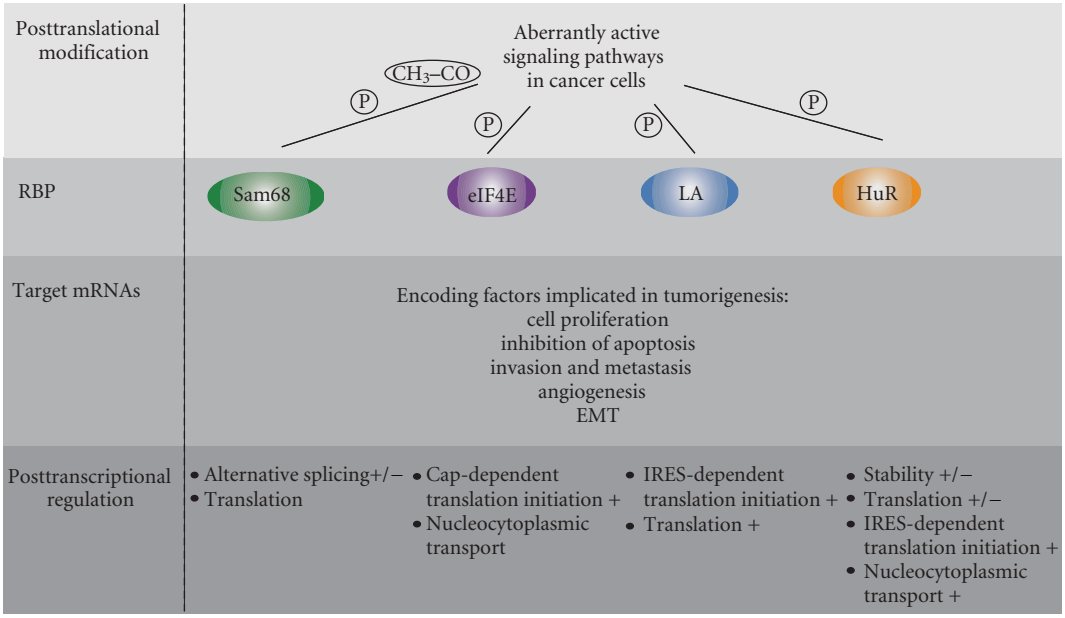


Figure 2: Overview of posttranscriptional gene regulation by Sam68, eIF4E, La, and HuR in tumorigenesis. In cancer cells, RBPs are posttranslationally modified by aberrantly active signaling pathways that activate their binding to targets encoding proteins implicated in tumorigenesis. The steps of mRNA metabolism regulated by RBPs are indicated. (+) and (–) specify up- or downregulation. P and CH_3–CO indicate phosphorylation and acetylation of RBPs. EMT: epithelial to mesenchymal transition.

progression [38]. Depletion of Sam68 strongly reduces the inclusion of several variable exons. Interestingly, Sam68 activity is controlled by the Ras signaling pathway, and Sam68 phosphorylation by ERK is needed to promote v5 inclusion [33].

Sam68 also regulates AS of cyclin D1, a protooncogene frequently deregulated in cancer cells [39, 40]. In addition, Sam68 promotes the generation of a stable SF2/ASF, isoform through regulation of splicing. The protooncogene SF2/ASF, also a splicing factor, is in turn responsible for processing of ΔRon pre-mRNA, which encodes a factor involved in EMT in colon cancer cells [41].

Another connection of Sam68 with cancer could be provided by the control of AS of the Bcl-x transcript. The Bcl-x gene can yield the antiapoptotic Bcl-x(L) factor or the proapoptotic Bcl-x(S) [22, 42]. Some studies have reported that Sam68 overexpression causes the accumulation of proapoptotic Bcl-x(s) in a manner that depends on the RNA-binding activity of Sam68 [22, 42]. However, the observation that Sam68 and the antiapoptotic Bcl-x(L) are upregulated in prostate cancer cells is at odds with a proposed activity of Sam68 in Bclx(S) upregulation [34, 43]. This apparent contradiction was resolved by the finding that the activity of Sam68 on Bcl-x AS depends on its phosphorylation status, which can switch Sam68 function from proapototic to anti-apoptotic in cancer cells. Indeed, Src-like kinase, which is often activated in cancer, phosphorylates Sam68 and thereby promotes splicing of the antiapoptotic Bcl-x(l) variant which inhibits cell death [22].

Intriguingly, in advanced breast and renal tumors, Sam68 was found to localize in the cytoplasm [26, 35]. These observations suggest a potential function of Sam68 in translational control in advanced stages of tumorigenesis. In accordance with a potential role of Sam68 in translation, it was previously proposed to regulate the translation of selected mRNAs in male germ cells and neurons [44, 45].

Other RBPs regulating splicing in cancer cells are hnRNPs (A/B) H, SR proteins (ASF/SF2), RBM5, HuR, and PTB. The interested reader can refer to the following reviews and articles [30, 46].

2.2. eIF4E Overexpression in Cancer Enhances Translation Initiation of Specific mRNAs. Translation initiation is a critical step of protein synthesis and is highly regulated [47]. One of the most crucial regulators is the cap-binding protein eIF4E (eukaryotic initiation factor 4E) [48]. In the cytoplasm, eIF4E binds directly to the m^7GTP-cap structure present at the 5′end of all mRNAs and interacts with eIF4G, which in turn recruits the 43S ribosomal complex during initiation of translation. eIF4E and eIF4G together with the RNA helicase eIF4A form the eIF4F complex, which is often targeted for translational regulation [47].

Early findings indicated that eIF4E overexpression leads to malignant transformation of fibroblasts [49, 50]. Since then, numerous studies have reported overexpression of eIF4E in different tumor types (e.g., breast, prostate, gastric colon, lung, skin, and lymphomas) [51]. Elevated expression of eIF4E often correlates with malignancy and poor prognosis [52, 53]. Surprisingly, overexpression of eIF4E does not induce a global increase in protein synthesis but augments translation of a subset of mRNAs encoding mostly prooncogenic proteins [2, 54] (Figure 2). mRNAs regulated

by eIF4E overexpression include those encoding components of the cell cycle machinery (cyclin D1, CDK2, c-myc, RNR2, ODC, surviving, Mcl-1, Bcl-2) or factors implicated in angiogenesis (VEGF, FGF-2, PDGF) and invasion (MMP9) [2, 51, 55, 56].

It has been proposed that mRNAs coding for proteins upregulated in oncogenesis contain long and highly structured 5′UTRs [10]. mRNAs bearing stable secondary structures in the 5′UTR are poorly translated in normal conditions and may be particularly dependent on the eIF4F complex and the unwinding capacity of the eIF4A helicase to initiate translation. Thus eIF4E overexpression may lead to enhanced translation of otherwise inefficiently translated transcripts involved in tumorigenesis [2]. Interestingly, eIF4E seems to be implicated in nucleocytoplasmic transport of mRNAs (e.g., cyclin D) and thus may regulate expression of some genes in an initiation-independent way [57].

eIF4E activity is regulated by signaling pathways amplified in human cancers (Figure 2). The protein kinase mTor phosphorylates eIF4E-binding proteins (4E-BP). In their unphosphorylated state, 4E-BPs bind to eIF4E on the same site recognized by eIF4G, blocking the formation of the cap-binding complex. Phosphorylation of 4E-BP leads to loss of affinity for eIF4E and increases translation [2]. In addition, eIF4E phosphorylation by MAPK-integrating kinases MNK1 and MNK2 enhances cap-dependent initiation [47, 54].

Given the important role of eIF4E in tumorigenesis, reducing either eIF4E activity or levels in cancer cells has become an attractive anticancer strategy [51, 58]. Many compounds inhibiting mTor kinase activity have proven to be efficient. For example, PP242, Tonin1, and INK128 are ATP active site inhibitors of mTOR and block the phosphorylation of all mTor targets including 4E-BP [51]. Unfortunately, cells of some cancer types are insensitive to treatment with mTor inhibitors [59]. As an alternative strategy, inhibiting eIF4E expression with antisense oligonucleotides (AON) has given promising results in suppressing tumor growth in vivo [60].

2.3. La Is an ITAF Implicated in Cancer. The multifunctional RNA-binding protein La is primarily nuclear but can shuttle between the nucleus and the cytoplasm [61, 62]. According to its localization, La functions in small RNA processing [63] and in translation of mRNAs [64–66]. La can be divided into three regions: the N-terminus, which contains the conserved La motif; a less conserved RNA recognition motif (RRM); and a weakly conserved C terminus, which contains an RRM, and a nuclear localization signal (NLS) [67] (Figure 1). The La motif folds into an RRM and its high conservation suggests that it carries out a specific function [68, 69]. La interacts with cellular and viral mRNAs and regulates IRES and cap-dependent translation initiation [64, 66, 70–73]. An IRES is a nucleotide sequence folding in a specific secondary structure that recruits ribosomes independently of the cap structure [74]. During cellular stress, cap-dependent translation is downregulated, and IRES-dependent translation of many mRNAs is favored [75]. For example, under the hypoxic conditions usually found in the interior of a tumor, IRES-mediated translation of the angiogenic factor VEGF is favored leading to vascularization of the tumor [76]. Specific RNA-binding proteins termed IRES transacting factors (ITAFS) are required to regulate IRES-dependent translation in cancer development [74]. La is an ITAF that regulates the IRES-dependent translation of mRNAs involved in cell proliferation, angiogenesis and apoptosis [64, 77, 78] (Figure 2).

As an ITAF, La interacts directly with the IRES of the mRNA encoding the proapoptotic factor XIAP [64]. In addition, La regulates IRES-dependent translation of LamB1, a factor that drives invasion, angiogenesis and metastasis [79, 80]. La also binds to the IRES of cyclin D1 (CCND1) in cervical cancer tissues, and its overexpression correlates with upregulation of cyclin D1 while its depletion leads to a reduction of cyclin D1 levels and a defect in cell proliferation [77].

La is overexpressed in chronic myeloid leukemia, cervical cancer tissues, oral squamous cell carcinoma (SCC), and in a number of cancer cell lines compared to nontumorigenic cells [66, 77, 78, 81]. In SCC, La is required for expression of β-catenin and MMP-2, proteins implicated in cell-cell adhesion and cell motility, respectively [78]. In leukemia, increased levels of La correlate with upregulation of MDM2 (an oncogenic tyrosine kinase). La interacts directly with the 5′UTR of mdm2 mRNA and enhances its translation [66].

Using mouse glial progenitor cells, Brennet proposed that La functions as a translational regulator during KRas/Akt oncogenic signaling [62]. Ras and Akt pathways are aberrantly active in cancer cells and play a pivotal role in the formation and regulation of glioblastoma [82]. In this tumor type La is phosphorylated by Akt, and this changes its distribution from the nucleus to the cytoplasm leading to association of a subset of La-bound mRNAs to polysomes. Many of these mRNAs encode factors implicated in oncogenesis such as Cyclin G2, Bcl2, and PDGFA [62].

The number of known La mRNA targets is still limited and further studies are necessary to understand its function in tumorigenesis. However, La already represents a promising target for cancer therapy. As an example, La activity has been efficiently blocked by a synthetic peptide corresponding to amino acids 11 to 28 of La. By competition, the peptide inhibits IRES-driven translation of Hepatitis C without affecting cap-dependent translation of cellular mRNAs [83]. This peptide could also be used to block expression of cancer related mRNA targets of La.

Other ITAFs implicated in cancer are PTB, hnRNP A1, hnRNP E1, hnRNP E2, and YB1. The interested reader can refer to the following reviews and articles [84, 85].

2.4. HuR Regulates the Stability and Translation of Cancer-Related Transcripts. The human antigen R (HuR) is the most prominent RBP known to be implicated in tumorigenesis [3]. Overexpression of HuR has been observed in lymphomas, gastric, breast, pancreatic, prostate, oral, colon, skin, lung, ovarian, and brain cancers [86–91]. Elevated cytoplasmic accumulation of HuR correlates with high-grade malignancy and serves as a prognostic factor of poor clinical outcome in some cancer types [92–95]. Localized in the nucleus of normal cells, HuR often translocates to

the cytoplasm in transformed cells [96, 97]. HuR's subcellular localization is regulated by posttranslational modifications, and the enzymes modifying HuR are all implicated in cancer [97] (Figure 2). In the cytoplasm, HuR binds to adenine- and uridine-rich elements (AU-rich elements or AREs) located in 3′UTR of target mRNAs [98]. AU-rich elements serve as binding sites for a variety of RBPs that modulate mRNA half-life [11]. An estimated 10% of all mRNAs bear AU-rich sequences [99]. The minimal functional ARE sequence is a nonamer UUAUUUAWW [100]. Most RBPs binding to AREs promote rapid deadenylation and degradation of substrate mRNAs by targeting them to the exosome (e.g., TTP, AUF1, CUGBP2) [101]. On the contrary, HuR most often enhances the stability of its target mRNAs [3]. In addition, HuR can also regulate the splicing of a certain number of targets [102].

HuR is a member of the embryonic lethal abnormal vision (ELAV) family of proteins and contains three RRMs that provide high-affinity RNA binding [103] (Figure 1). HuR target mRNAs encode products that promote proliferation, inhibit apoptosis, increase angiogenesis, and facilitate invasion and metastasis. For an extensive list of HuR targets, see [3]. Below I will give an overview of HuR targets and will summarize the different mechanisms by which HuR regulates their expression.

Upon binding to the 3′UTR, HuR stabilizes the mRNAs coding for cyclins (cyclin D1, E1, A2, B1), favoring cell cycle progression and promoting proliferation of cancer cells [104–106]. HuR also promotes cancer cell survival by stabilizing transcripts encoding antiapoptotic factors like Bcl-2, Mcl-1, SIRT1, and p21 [90, 107–110]. mRNAs coding for proteins implicated in invasion and metastasis (MMP-9) [111, 112], cell migration and adhesion (Urokinase A and uPA receptor) [113] or EMT (snail) are also stabilized by HuR [114]. Expression of the proangiogenic factors VEGF and HIF-1α is controlled by HuR. Regulation of HIF-1α mRNA is interesting, as HuR binds to both the 5′ and 3′UTRs and promotes translation and stability [115, 116]. The mechanism by which HuR stabilizes its targets is still unclear, but recent studies have proposed an interplay between HuR and miRNAs [117]. HuR is able to suppress activity of miRNAs, by inhibiting their recruitment to the mRNA or even by promoting their downregulation. Some examples of cross-talk between HuR and miRNAs will be given in the next paragraph.

ERBB-2 overexpression is associated with development and progression of prostate cancer. HuR enhances ERBB-2 expression using a miRNA-dependent mechanism. HuR binds to a uridinerich element (URE) in the 3′UTR of ERBB-2 and inhibits action of miR-331-3p to a nearby site [118]. The presence of HuR on the mRNA does not alter miR-331-3p binding, which leads to the hypothesis that HuR may rather reduce association between ERBB-2 mRNA and the RNA silencing complex [118]. In colorectal cancer, HuR overexpression and localization in the cytoplasm correlate with decreased levels of miR-16, a miRNA that binds to the 3′UTR of COX-2 mRNA and inhibits its expression by mRNA decay [119]. Intriguingly, HuR interacts with miR-16 and promotes its downregulation in an mRNA ligand-dependent manner. Thus, HuR stabilizes COX-2 mRNA by binding to the ARE and by downregulating miR-16 [119].

Interestingly, HuR is able to repress the translation of the proapoptotic factor c-Myc by recruiting the let-7 miRNP to the 3′UTR [120]. HuR is not the only RBP which assists in targeting miRNPs to the 3′UTR of mRNAs, as was shown with the example of TTP [121].

HuR also represses the translation of some of its targets by binding to the 5′UTR. This is the case for p27, which prevents cell proliferation [122].

It has been recently shown that HuR can act as an ITAF binding to the IRES of XIAP mRNA, which encodes an antiapoptotic factor [123]. HuR stimulates the translation of XIAP mRNA by binding to XIAP IRES and enhancing its recruitment into polysomes.

Interestingly in the case of the antiapoptotic factor prothymosin alpha (ProTα), HuR binding to its 3′UTR enhances nuclear export of the mRNA followed by induced translation upon UV irradiation [124].

In summary, the majority of HuR mRNA targets are stabilized upon binding, and translation is enhanced. As an ITAF, HuR binds to IRES structures and enhances translation. HuR is also able to inhibit translation by binding to 5′UTR or by recruiting miRNPs to the 3′UTR. On the other hand, HuR also inhibits miRNA binding to the 3′UTR of its target mRNAs. Finally, HuR is increasing cytoplasmic abundance of target mRNAs probably via enhanced mRNA nulear export. These examples illustrate the complexity of HuR regulatory activity.

The large spectrum of mRNA targets regulated by HuR confirms its potential to coordinate nearly all steps of tumorigenesis. Overexpressed in a high number of cancer types, HuR provides a good candidate for therapy design. Surprisingly, however, a recent study showed that elevated levels of HuR may be advantageous for cancer therapy. In pancreatic ductal adenocarcinoma, HuR levels modulate the therapeutic activity of gemcitabine (GEM), a common chemotherapeutic agent [125]. GEM exposure to cancer cells increases the amount of cytoplasmic HuR and promotes its association with dCK mRNA, which encodes the enzyme that activates GEM, establishing a positive feedback loop that improves its therapeutic efficacy. This example shows that therapies that reduce the level of HuR have to be designed carefully, and perhaps in a tumor type-dependent manner [126].

Besides HuR, a number of other factors can regulate the stability and expression of mRNAs bearing AREs [127]. The TIS11 family of RBPs composed of Tristetraprolin (TTP) and butyrate response factors 1 and 2 (BRF-1 and-2) bind and target ARE-containing mRNAs for rapid degradation [101]. AUF1 is able to stabilize or destabilize ARE-containing mRNAs [128]. The CELF family of RNA-binding proteins is composed of 6 members, which promote either mRNA decay or translation of its target mRNAs [129, 130]. For example CUGBP2 binds COX-2 mRNA which is then stabilized but translationally repressed [131]. T-cell intracellular antigen-1 (TIA-1) and TIA-1-related (TIAR) proteins are translational silencers [132]. Some of these factors have common targets and compete for binding depending on cellular conditions.

3. Conservation of RBPs across Eukaryotes

Post-transcriptional gene regulation is a coordinated, efficient, rapid and flexible mechanism to control the proteome of the cell in response to different physiological conditions. It is thus not surprising that some organisms have become highly dependent on post-transcriptional mechanisms to regulate gene expression, like, for example, the protozoan parasite, trypanosome [133–135]. The trypanosome genome encodes very few potential regulatory transcription factors, and gene regulation relies mostly on RNA-binding proteins [136]. It has been proposed that 3–11% of the proteome in bacteria, archea and eukaryotes are putative RNA-binding proteins [137]. The large number of RBPs suggests that RNA metabolism may be a central and evolutionarily conserved contributor to cell physiology. Most of the RNA-binding domains known today are present in early stages of evolution. Interestingly, several new eukaryotic-specific RNA-binding domains have emerged, like the RRM, which suggests that post-transcriptional gene regulation became more complex with evolution [137].

The RNA-binding proteins described in this paper are widely conserved across eukaryotes (Figure 3). We could detect homologues of HuR only in metazoa and not in fungi and plants. Human HuR is the most divergent family member of the ELAV proteins. While the other members, HuD, HuC, and Hel-N1, present a neuron- and brain-specific expression, where they are mostly implicated in alternative splicing, HuR is ubiquitously expressed and fulfills numerous functions [138, 139].

Sam68 homologues exist in all eukaryotes except fungi (Figure 3). In the STAR protein family, the Sam68 subfamily is composed of Sam68 (SRC-associated in mitosis, 68 kd) and the Sam68-like mammalian proteins 1 and 2 (SLM-1 and SLM-2, also named T-STAR in humans) [140–143]. As in the case of HuR, Sam68 is ubiquitously expressed, whereas SLM-1 and SLM-2 expression is restricted to few cell types or tissues [144]. In humans, Sam68 has acquired a larger spectrum of functions and plays a major role in signaling and splicing in different tissues.

Contrary to HuR and Sam68, La homologues can be identified in all three phyla: metazoa, fungi, and plants (Figure 3). La was first characterized as a human protein, and homologues have been identified in a wide variety of other eukaryotes [63]. The N-terminal part containing the La motif is highly conserved, in contrast to the C-terminal domain which varies both in size and sequence between species, ranging from 70 amino acids in the yeasts *S. cerevisiae* and *S. pombe* to more than 220 amino acids in vertebrates. Human La is phosphorylated at different sites, all located in the C terminus [63] (Figure 2). Interestingly these sites are only conserved in vertebrate La proteins. The presence of an additional C-terminal region including different functional domains and phosphorylation sites shows that La has evolved to a highly regulated and multifunctional factor in vertebrates.

The translation initiation factor eIF4E is highly conserved across eukaryotes. Sequence comparisons revealed a phylogenetically conserved 182 amino acid C-terminal region [145, 146]. In contrast, the N-terminal region is poorly conserved and is not required for cap-dependent translation [145]. Functional conservation has also been demonstrated, as mammalian eIF4E can rescue the lethality caused by disruption of the yeast eIF4E gene [147]. The crystallographic structure of eIF4E in mouse, yeast, human, and wheat has been solved [145, 148–150]. The three-dimensional structure of the C-terminal part of murine eIF4E demonstrates that the surface of the molecule resembles a cupped hand that contains a narrow cap-binding slot. The remarkable level of sequence identity across phylogeny suggests that all known eIF4Es share the same structure in their conserved C-terminal region [145]. eIF4E thus does not contain a canonical RBD but adopts a conserved three-dimensional structure which interacts with the cap.

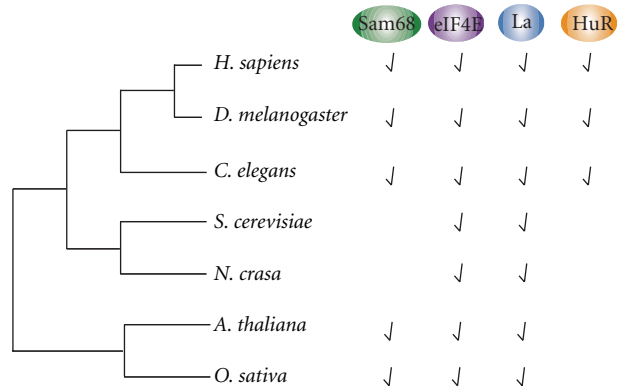


Figure 3: Conservation of Sam68, eIF4E, La, and HuR in different phyla. Phylogenetic tree of the RBPs described in this paper. The presence of homologues is indicated.

Interestingly most eukaryotic organisms express multiple eIF4E family members, and it has been proposed that a ubiquitously expressed member of the family may be implicated in general translation initiation while others could be involved in specialized functions [151, 152]. eIF4E family members may provide an additional layer of control in translation and may regulate specific subsets of mRNAs, which could be linked to cancer development.

4. Concluding Remarks

In cancer research, the impact of post-transcriptional gene regulation has been considered only since a few years. Today, it is well established that a subset of RBPs are key regulators of processes involved in tumorigenesis. The genome wide analysis of RBPs and their RNA targets has allowed a better understanding of the complex world of mRNA metabolism and the connections existing between different RBPs. According to the "RNA operon" concept, mRNAs encoding functionally related proteins are coregulated by specific RBPs, ensuring an efficient, flexible, and coordinated response to cellular need [144, 153, 154]. RNA operons can be interconnected. HuR and eIF4E for example, share common mRNA targets like c-myc, cyclin D1 and VEGF, suggesting an orchestrated regulation of the expression of genes implicated

in tumorigenesis [155, 156]. In addition, HuR regulates expression of eIF4E in cancer cells [156]. These observations show that post-transcriptional regulation events are highly linked and provide a powerful mechanism to control the fate of a cell.

RBPs are highly versatile factors that can bind to multiple RNA targets and regulate their fate by a variety of mechanisms. The fact that every step of the mRNA life cycle is narrowly controlled allows RBPs to fine tune expression in a very precise manner. The conservation of RBPs across eukaryotes and the emergence of more complexity along evolution also point to an essential role of RBPs. Post-transcriptional gene regulation is a central mechanism of emerging importance in cancer research which is expected to provide novel targets for therapy design.

Acknowledgments

The author would like to thank Fátima Gebauer for suggestions and critical reading of the paper. He also thanks Margarita Meer and Fyodor Kondrashov for performing phylogenetic analysis. L. Wurth is supported by the National Research Fund, Luxembourg, and cofounded under the Marie Curie Actions of European Commission (FP7-COFUND) and work in Fátima Gebauer's lab is supported by grants BFU2009-08243 and Consolider CSD2009-00080 from MICINN.

References

[1] M. Y. Kim, J. Hur, and S. Jeong, "Emerging roles of RNA and RNA-binding protein network in cancer cells," *BMB Reports*, vol. 42, no. 3, pp. 125–130, 2009.

[2] D. Silvera, S. C. Formenti, and R. J. Schneider, "Translational control in cancer," *Nature Reviews Cancer*, vol. 10, no. 4, pp. 254–266, 2010.

[3] K. Abdelmohsen and M. Gorospe, "Posttranscriptional regulation of cancer traits by HuR," *Wiley Interdisciplinary Reviews*, vol. 1, no. 2, pp. 214–229, 2011.

[4] P. Bielli, R. Busa, M. P. Paronetto, and C. Sette, "The RNA-binding protein Sam68 is a multifunctional player in human cancer," *Endocrine-Related Cancers*, vol. 18, no. 4, pp. R91–R102, 2011.

[5] J. D. Keene, "Ribonucleoprotein infrastructure regulating the flow of genetic information between the genome and the proteome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 13, pp. 7018–7024, 2001.

[6] M. J. Moore, "From birth to death: the complex lives of eukaryotic mRNAs," *Science*, vol. 309, no. 5740, pp. 1514–1518, 2005.

[7] M. C. Wahl, C. L. Will, and R. Lührmann, "The Spliceosome: design Principles of a Dynamic RNP Machine," *Cell*, vol. 136, no. 4, pp. 701–718, 2009.

[8] M. van Kouwenhove, M. Kedde, and R. Agami, "MicroRNA regulation by RNA-binding proteins and its implications for cancer," *Nature Reviews Cancer*, vol. 11, pp. 644–656, 2011.

[9] A. L. Kasinski and F. J. Slack, "MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy," *Nature Reviews Cancer*, vol. 11, no. 12, pp. 849–864, 2011.

[10] B. M. Pickering and A. E. Willis, "The implications of structured 5′ untranslated regions on translation and disease," *Seminars in Cell and Developmental Biology*, vol. 16, no. 1, pp. 39–47, 2005.

[11] Y. Audic and R. S. Hartley, "Post-transcriptional regulation in cancer," *Biology of the Cell*, vol. 96, no. 7, pp. 479–498, 2004.

[12] L. Duret, F. Dorkeld, and C. Gautier, "Strong conservation of non-coding sequences during vertebrates evolution: potential involvement in post-transcriptional regulation of gene expression," *Nucleic Acids Research*, vol. 21, no. 10, pp. 2315–2322, 1993.

[13] B. M. Lunde, C. Moore, and G. Varani, "RNA-binding proteins: modular design for efficient function," *Nature Reviews Molecular Cell Biology*, vol. 8, no. 6, pp. 479–490, 2007.

[14] J. Ule, K. Jensen, A. Mele, and R. B. Darnell, "CLIP: a method for identifying protein-RNA interaction sites in living cells," *Methods*, vol. 37, no. 4, pp. 376–386, 2005.

[15] D. D. Licatalosi, A. Mele, J. J. Fak et al., "HITS-CLIP yields genome-wide insights into brain alternative RNA processing," *Nature*, vol. 456, no. 7221, pp. 464–469, 2008.

[16] S. Lebedeva, M. Jens, K. Theil et al., "Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR," *Molecular Cell*, vol. 43, no. 3, pp. 340–352, 2011.

[17] K. E. Lukong, K. W. Chang, E. W. Khandjian, and S. Richard, "RNA-binding proteins in human genetic disease," *Trends in Genetics*, vol. 24, no. 8, pp. 416–425, 2008.

[18] E. Ortiz-Zapater, D. Pineda, N. Martinez-Bosch et al., "Key contribution of CPEB4-mediated translational control to cancer progression," *Nature Medicine*, vol. 18, pp. 83–90, 2012.

[19] F. Xu, X. Zhang, Y. Lei et al., "Loss of repression of HuR translation by miR-16 may be responsible for the elevation of HuR in human breast carcinoma," *Journal of Cellular Biochemistry*, vol. 111, no. 3, pp. 727–734, 2010.

[20] B. H. Sohn, I. Y. Park, J. J. Lee et al., "Functional switching of TGF-beta1 signaling in liver cancer via epigenetic modulation of a single CpG site in TTP promoter," *Gastroenterology*, vol. 138, no. 5, pp. 1898–e12, 2010.

[21] C. A. Gebeshuber, K. Zatloukal, and J. Martinez, "miR-29a suppresses tristetraprolin, which is a regulator of epithelial polarity and metastasis," *EMBO Reports*, vol. 10, no. 4, pp. 400–405, 2009.

[22] M. P. Paronetto, T. Achsel, A. Massiello, C. E. Chalfant, and C. Sette, "The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x," *Journal of Cell Biology*, vol. 176, no. 7, pp. 929–939, 2007.

[23] H. H. Kim, K. Abdelmohsen, A. Lal et al., "Nuclear HuR accumulation through phosphorylation by Cdk1," *Genes and Development*, vol. 22, no. 13, pp. 1804–1815, 2008.

[24] C. Vernet and K. Artzt, "STAR, a gene family involved in signal transduction and activation of RNA," *Trends in Genetics*, vol. 13, no. 12, pp. 479–484, 1997.

[25] K. E. Lukong and S. Richard, "Sam68, the KH domain-containing superSTAR," *Biochimica et Biophysica Acta*, vol. 1653, no. 2, pp. 73–86, 2003.

[26] Z. Zhang, J. Li, H. Zheng et al., "Expression and cytoplasmic localization of SAM68 is a significant and independent prognostic marker for renal cell carcinoma," *Cancer Epidemiology Biomarkers and Prevention*, vol. 18, no. 10, pp. 2685–2693, 2009.

[27] Q. Lin, S. J. Taylor, and D. Shalloway, "Specificity and determinants of Sam68 RNA binding. Implications for the biological function of K homology domains," *The Journal of Biological Chemistry*, vol. 272, no. 43, pp. 27274–27280, 1997.

[28] I. Babic, A. Jakymiw, and D. J. Fujita, "The RNA binding protein Sam68 is acetylated in tumor cell lines, and its acetylation correlates with enhanced RNA binding activity," *Oncogene*, vol. 23, no. 21, pp. 3781–3789, 2004.

[29] B. Hartmann and J. Valcárcel, "Decrypting the genome's alternative messages," *Current Opinion in Cell Biology*, vol. 21, no. 3, pp. 377–386, 2009.

[30] C. J. David and J. L. Manley, "Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged," *Genes and Development*, vol. 24, no. 21, pp. 2343–2364, 2010.

[31] K. E. Lukong and S. Richard, "Targeting the RNA-binding protein Sam68 as a treatment for cancer?" *Future Oncology*, vol. 3, no. 5, pp. 539–544, 2007.

[32] R. B. Irby and T. J. Yeatman, "Role of Src expression and activation in human cancer," *Oncogene*, vol. 19, no. 49, pp. 5636–5642, 2000.

[33] N. Matter, P. Herrlich, and H. König, "Signal-dependent regulation of splicing via phosphorylation of Sam68," *Nature*, vol. 420, no. 6916, pp. 691–695, 2002.

[34] R. Busà, M. P. Paronetto, D. Farini et al., "The RNA-binding protein Sam68 contributes to proliferation and survival of human prostate cancer cells," *Oncogene*, vol. 26, no. 30, pp. 4372–4382, 2007.

[35] L. Song, L. Wang, Y. Li et al., "Sam68 up-regulation correlates with, and its down-regulation inhibits, proliferation and tumourigenicity of breast cancer cells," *Journal of Pathology*, vol. 222, no. 3, pp. 227–237, 2010.

[36] Z. Li, C. P. Yu, Y. Zhong et al., "Sam68 expression and cytoplasmic localization is correlated with lymph node metastasis as well as prognosis in patients with early-stage cervical cancer," *Annals of Oncology*, vol. 23, no. 3, pp. 638–646, 2012.

[37] P. Rajan, L. Gaughan, C. Dalgliesh et al., "Regulation of gene expression by the RNA-binding protein Sam68 in cancer," *Biochemical Society Transactions*, vol. 36, no. 3, pp. 505–507, 2008.

[38] D. L. Cooper, "Retention of CD44 introns in bladder cancer: understanding the alternative splicing of pre-mRNA opens new insights into the pathogenesis of human cancers," *Journal of Pathology*, vol. 177, no. 1, pp. 1–3, 1995.

[39] M. P. Paronetto, M. Cappellari, R. Busà et al., "Alternative splicing of the cyclin D1 proto-oncogene is regulated by the RNA-binding protein Sam68," *Cancer Research*, vol. 70, no. 1, pp. 229–239, 2010.

[40] E. S. Knudsen and K. E. Knudsen, "Retinoblastoma tumor suppressor: where cancer meets the cell cycle," *Experimental Biology and Medicine*, vol. 231, no. 7, pp. 1271–1281, 2006.

[41] C. Valacca, S. Bonomi, E. Buratti et al., "Sam68 regulates EMT through alternative splicing-activated nonsense-mediated mRNA decay of the SF2/ASF proto-oncogene," *Journal of Cell Biology*, vol. 191, no. 1, pp. 87–99, 2010.

[42] L. H. Boise, M. Gonzalez-Garcia, C. E. Postema et al., "bcl-x, A bcl-2-related gene that functions as a dominant regulator of apoptotic cell death," *Cell*, vol. 74, no. 4, pp. 597–608, 1993.

[43] D. R. Mercatante, J. L. Mohler, and R. Kole, "Cellular response to an antisense-mediated shift of Bcl-x pre-mRNA splicing and antineoplastic agents," *The Journal of Biological Chemistry*, vol. 277, no. 51, pp. 49374–49382, 2002.

[44] M. P. Paronetto, F. Zalfa, F. Botti, R. Geremia, C. Bagni, and C. Sette, "The nuclear RNA-binding protein Sam68 translocates to the cytoplasm and associates with the polysemes in mouse spermatocytes," *Molecular Biology of the Cell*, vol. 17, no. 1, pp. 14–24, 2006.

[45] J. Grange, V. Boyer, R. Fabian-Fine, N. B. Fredj, R. Sadoul, and Y. Goldberg, "Somatodendritic Localization and mRNA Association of the Splicing Regulatory Protein Sam68 in the Hippocampus and Cortex," *Journal of Neuroscience Research*, vol. 75, no. 5, pp. 654–666, 2004.

[46] J. M. Izquierdo, N. Majos, and S. Bonnal, "Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition," *Molecular Cell*, vol. 19, no. 4, pp. 475–484, 2005.

[47] N. Sonenberg and A. G. Hinnebusch, "Regulation of translation initiation in eukaryotes: mechanisms and biological targets," *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[48] I. Topisirovic, Y. V. Svitkin, N. Sonenberg, and A. J. Shatkin, "Cap and cap-binding proteins in the control of gene expression," *Wiley Interdisciplinary Reviews*, vol. 2, no. 2, pp. 277–298, 2011.

[49] A. Lazaris-Karatzas, K. S. Montine, and N. Sonenberg, "Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5′ cap," *Nature*, vol. 345, no. 6275, pp. 544–547, 1990.

[50] S. G. Zimmer, A. Debenedetti, and J. R. Graff, "Translational control of malignancy: the mRNA cap-binding protein, eIF-4E, as a central regulator of tumor formation, growth, invasion and metastasis," *Anticancer Research*, vol. 20, no. 3, pp. 1343–1351, 2000.

[51] A. C. Hsieh and D. Ruggero, "Targeting eukaryotic translation initiation factor 4E (eIF4E) in cancer," *Clinical Cancer Research*, vol. 16, no. 20, pp. 4914–4920, 2010.

[52] J. R. Graff, B. W. Konicek, R. L. Lynch et al., "eIF4E activation is commonly Elevated in advanced human prostate cancers and significantly related to reduced patient survival," *Cancer Research*, vol. 69, no. 9, pp. 3866–3873, 2009.

[53] L. J. Coleman, M. B. Peter, T. J. Teall et al., "Combined analysis of eIF4E and 4E-binding protein expression predicts breast cancer survival and estimates eIF4E activity," *British Journal of Cancer*, vol. 100, no. 9, pp. 1393–1399, 2009.

[54] H. G. Wendel, R. L. A. Silva, A. Malina et al., "Dissecting eIF4E action in tumorigenesis," *Genes and Development*, vol. 21, no. 24, pp. 3232–3237, 2007.

[55] Y. Mamane, E. Petroulakis, L. Rong, K. Yoshida, L. W. Ler, and N. Sonenberg, "eIF4E—from translation to transformation," *Oncogene*, vol. 23, no. 18, pp. 3172–3179, 2004.

[56] A. De Benedetti and J. R. Graff, "eIF-4E expression and its role in malignancies and metastases," *Oncogene*, vol. 23, no. 18, pp. 3189–3199, 2004.

[57] B. Culjkovic, I. Topisirovic, L. Skrabanek, M. Ruiz-Gutierrez, and K. L. B. Borden, "eIF4E promotes nuclear export of cyclin D1 mRNAs via an element in the 3′UTR," *Journal of Cell Biology*, vol. 169, no. 2, pp. 245–256, 2005.

[58] A. Malina, R. Cencic, and J. Pelletier, "Targeting translation dependence in cancer.," *Oncotarget*, vol. 2, no. 1-2, pp. 76–88, 2011.

[59] M. A. Bjornsti and P. J. Houghton, "The TOR pathway: a target for cancer therapy," *Nature Reviews Cancer*, vol. 4, no. 5, pp. 335–348, 2004.

[60] J. R. Graff, B. W. Konicek, T. M. Vincent et al., "Therapeutic suppression of translation initiation factor eIF4E expression reduces tumor growth without toxicity," *Journal of Clinical Investigation*, vol. 117, no. 9, pp. 2638–2648, 2007.

[61] S. A. Rutjes, P. J. Utz, A. Van Der Heijden, C. Broekhuis, W. J. Van Venrooij, and G. J. M. Pruijn, "The La (SS-B) autoantigen, a key protein in RNA biogenesis, is dephosphorylated and cleaved early during apoptosis," *Cell Death and Differentiation*, vol. 6, no. 10, pp. 976–986, 1999.

[62] F. Brenet, N. D. Socci, N. Sonenberg, and E. C. Holland, "Akt phosphorylation of La regulates specific mRNA translation in glial progenitors," *Oncogene*, vol. 28, no. 1, pp. 128–139, 2009.

[63] S. L. Wolin and T. Cedervall, "The La protein," *Annual Review of Biochemistry*, vol. 71, pp. 375–403, 2002.

[64] M. Holcik and R. G. Korneluk, "Functional characterization of the X-linked inhibitor of apoptosis (XIAP) internal ribosome entry site element: role of la autoantigen in XIAP translation," *Molecular and Cellular Biology*, vol. 20, no. 13, pp. 4648–4657, 2000.

[65] C. Crosio, P. P. Boyl, F. Loreni, P. Pierandrei-Amaldi, and F. Amaldi, "La protein has a positive effect on the translation of TOP mRNAs in vivo," *Nucleic Acids Research*, vol. 28, no. 15, pp. 2927–2934, 2000.

[66] R. Trotta, T. Vignudelli, O. Candini et al., "BCR/ABL activates mdm2 mRNA translation via the La antigen," *Cancer Cell*, vol. 3, no. 2, pp. 145–160, 2003.

[67] D. J. Van Horn, C. J. Yoo, D. Xue, H. Shi, and S. L. Wolin, "The La protein in Schizosaccharomyces pombe: a conserved yet dispensable phosphoprotein that functions in tRNA maturation," *RNA*, vol. 3, no. 12, pp. 1434–1443, 1997.

[68] R. J. Maraia and R. V. A. Intine, "Recognition of nascent RNA by the human La antigen: conserved and divergent features of structure and function," *Molecular and Cellular Biology*, vol. 21, no. 2, pp. 367–379, 2001.

[69] D. J. Kenan and J. D. Keene, "La gets its wings," *Nature Structural and Molecular Biology*, vol. 11, no. 4, pp. 303–305, 2004.

[70] R. S. McLaren, N. Caruccio, and J. Ross, "Human la protein: a stabilizer of histone mRNA," *Molecular and Cellular Biology*, vol. 17, no. 6, pp. 3028–3036, 1997.

[71] F. Brenet, N. Dussault, J. Borch et al., "Mammalian peptidylglycine α-amidating monooxygenase mRNA expression can be modulated by the La autoantigen," *Molecular and Cellular Biology*, vol. 25, no. 17, pp. 7505–7521, 2005.

[72] K. Spångberg, L. Wiklund, and S. Schwartz, "Binding of the La autoantigen to the hepatitis C virus 3′ untranslated region protects the RNA from rapid degradation in vitro," *Journal of General Virology*, vol. 82, no. 1, pp. 113–120, 2001.

[73] Y. K. Kim, S. H. Back, J. Rho, S. H. Lee, and S. K. Jang, "La autoantigen enhances translation of Bip mRNA," *Nucleic Acids Research*, vol. 29, no. 24, pp. 5009–5016, 2001.

[74] A. A. Komar and M. Hatzoglou, "Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states," *Cell Cycle*, vol. 10, no. 2, pp. 229–240, 2011.

[75] K. A. Spriggs, M. Bushell, and A. E. Willis, "Translational Regulation of Gene Expression during Conditions of Cell Stress," *Molecular Cell*, vol. 40, no. 2, pp. 228–237, 2010.

[76] I. Stein, A. Itin, P. Einat, R. Skaliter, Z. Grossman, and E. Keshet, "Translation of vascular endothelial growth factor mRNA by internal ribosome entry: implications for translation under hypoxia," *Molecular and Cellular Biology*, vol. 18, no. 6, pp. 3112–3119, 1998.

[77] G. Sommer, C. Rossa, A. C. Chi, B. W. Neville, and T. Heise, "Implication of RNA-binding protein La in proliferation, migration and invasion of lymph node-metastasized hypopharyngeal SCC cells," *PLoS ONE*, vol. 6, no. 10, Article ID e25402, 2011.

[78] G. Sommer, J. Dittmann, J. Kuehnert et al., "The RNA-binding protein la contributes to cell proliferation and CCND1 expression," *Oncogene*, vol. 30, no. 4, pp. 434–444, 2011.

[79] M. Petz, N. Them, H. Huber, H. Beug, and W. Mikulits, "La enhances IRES-mediated translation of laminin B1 during malignant epithelial to mesenchymal transition," *Nucleic Acids Research*, vol. 40, no. 1, pp. 290–302, 2012.

[80] X. Sanjuan, P. L. Fernandez, R. Miquel et al., "Overexpression of the 67-kD laminin receptor correlates with tumour progression in human colorectal carcinoma," *The Journal of Pathology*, vol. 179, no. 4, pp. 376–380, 1996.

[81] F. Al-Ejeh, J. M. Darby, and M. P. Brown, "The La autoantigen is a malignancy-associated cell death target that is induced by DNA-damaging drugs," *Clinical Cancer Research*, vol. 13, no. 18, pp. 5509s–5518s, 2007.

[82] E. C. Holland, J. Celestino, C. Dai, L. Schaefer, R. E. Sawaya, and G. N. Fuller, "Combined activation of Ras and Akt in neural progenitors induces glioblastoma formation in mice," *Nature Genetics*, vol. 25, no. 1, pp. 55–57, 2000.

[83] R. E. Izumi, S. Das, B. Barat, S. Raychaudhuri, and A. Dasgupta, "A peptide from autoantigen La blocks poliovirus and hepatitis C virus cap-independent translation and reveals a single tyrosine critical for La RNA binding and translation stimulation," *Journal of Virology*, vol. 78, no. 7, pp. 3763–3776, 2004.

[84] L. C. Cobbold, L. A. Wilson, K. Sawicka et al., "Upregulated c-myc expression in multiple myeloma by internal ribosome entry results from increased interactions with and expression of PTB-1 and YB-1," *Oncogene*, vol. 29, no. 19, pp. 2884–2891, 2010.

[85] Y. Shi, P. J. Frost, B. Q. Hoang et al., "IL-6-induced stimulation of c-Myc translation in multiple myeloma cells is mediated by myc internal ribosome entry site function and the RNA-binding protein, hnRNP A1," *Cancer Research*, vol. 68, no. 24, pp. 10215–10222, 2008.

[86] J. Bergalet, M. Fawal, C. Lopez et al., "HuR-Mediated control of C/EBPβ mRNA stability and translation in ALK-Positive anaplastic large cell lymphomas," *Molecular Cancer Research*, vol. 9, no. 4, pp. 485–496, 2011.

[87] W. Kakuguchi, T. Kitamura, T. Kuroshima et al., "HuR knockdown changes the oncogenic potential of oral cancer cells," *Molecular Cancer Research*, vol. 8, no. 4, pp. 520–528, 2010.

[88] S. L. Nowotarski and L. M. Shantz, "Cytoplasmic accumulation of the RNA-binding protein HuR stabilizes the ornithine decarboxylase transcript in a murine nonmelanoma skin cancer model," *The Journal of Biological Chemistry*, vol. 285, no. 41, pp. 31885–31894, 2010.

[89] J. Wang, B. Wang, J. Bi, and C. Zhang, "Cytoplasmic HuR expression correlates with angiogenesis, lymphangiogenesis, and poor outcome in lung cancer," *Medical Oncology*, vol. 28, pp. 577–585, 2010.

[90] N. Filippova, X. Yang, Y. Wang et al., "The RNA-binding protein HuR promotes glioma growth and treatment resistance," *Molecular Cancer Research*, vol. 9, no. 5, pp. 648–659, 2011.

[91] F. Bolognani, A. I. Gallani, and L. Sokol, "mRNA stability alterations mediated by HuR are necessary to sustain the fast growth of glioma cells," *Journal of Neuro-Oncology*, vol. 106, no. 3, pp. 531–542, 2012.

[92] P. S. Yoo, C. A. W. Sullivan, S. Kiang et al., "Tissue microarray analysis of 560 patients with colorectal adenocarcinoma: high expression of HuR predicts poor survival," *Annals of Surgical Oncology*, vol. 16, no. 1, pp. 200–207, 2009.

[93] M. Heinonen, R. Fagerholm, K. Aaltonen et al., "Prognostic role of HuR in hereditary breast cancer," *Clinical Cancer Research*, vol. 13, no. 23, pp. 6959–6963, 2007.

[94] C. Denkert, W. Weichert, K. J. Winzer et al., "Expression of the ELAV-like protein HuR is associated with higher tumor grade and increased cyclooxygenase-2 expression in human

breast carcinoma," *Clinical Cancer Research*, vol. 10, no. 16, pp. 5580–5586, 2004.

[95] C. Denkert, W. Weichert, S. Pest et al., "Overexpression of the embryonic-lethal abnormal vision-like protein HuR in ovarian carcinoma is a prognostic factor and is associated with increased cyclooxygenase 2 expression," *Cancer Research*, vol. 64, no. 1, pp. 189–195, 2004.

[96] H. Hasegawa, W. Kakuguchi, T. Kuroshima et al., "HuR is exported to the cytoplasm in oral cancer cells in a different manner from that of normal cells," *British Journal of Cancer*, vol. 100, no. 12, pp. 1943–1948, 2009.

[97] A. Doller, J. Pfeilschifter, and W. Eberhardt, "Signalling pathways regulating nucleo-cytoplasmic shuttling of the mRNA-binding protein HuR," *Cellular Signalling*, vol. 20, no. 12, pp. 2165–2173, 2008.

[98] C. M. Brennan and J. A. Steitz, "HuR and mRNA stability," *Cellular and Molecular Life Sciences*, vol. 58, no. 2, pp. 266–277, 2001.

[99] A. S. Halees, R. El-badrawi, and K. S. A. Khabar, "ARED Organism: expansion of ARED reveals AU-rich element cluster variations between human and mouse," *Nucleic Acids Research*, vol. 36, no. 1, pp. D137–D140, 2008.

[100] C. A. Lagnado, C. Y. Brown, and G. J. Goodali, "AUUUA is not sufficient to promote poly(A) shortening and degradation of an mRNA: the functional sequence within AU-rich elements may be UUAUUUA(U/A)(U/A)," *Molecular and Cellular Biology*, vol. 14, no. 12, pp. 7984–7995, 1994.

[101] S. Sanduja and D. A. Dixon, "Tristetraprolin and E6-AP: killing the messenger in cervical cancer," *Cell Cycle*, vol. 9, no. 16, pp. 3135–3136, 2010.

[102] J. M. Izquierdo, "Hu antigen R (HuR) functions as an alternative pre-mRNA splicing regulator of Fas apoptosis-promoting receptor on exon definition," *The Journal of Biological Chemistry*, vol. 283, no. 27, pp. 19077–19084, 2008.

[103] T. D. Levine, F. Gao, P. H. King, L. G. Andrews, and J. D. Keene, "Hel-N1: an autoimmune RNA-binding protein with specificity for 3' uridylate-rich untranslated regions of growth factor mRNAs," *Molecular and Cellular Biology*, vol. 13, no. 6, pp. 3494–3504, 1993.

[104] A. Lal, K. Mazan-Mamczarz, T. Kawai, X. Yang, J. L. Martindale, and M. Gorospe, "Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs," *EMBO Journal*, vol. 23, no. 15, pp. 3092–3102, 2004.

[105] X. Guo and R. S. Hartley, "HuR contributes to cyclin E1 deregulation in MCF-7 breast cancer cells," *Cancer Research*, vol. 66, no. 16, pp. 7948–7956, 2006.

[106] W. Wang, M. C. Caldwell, S. Lin, H. Furneaux, and M. Gorospe, "HuR regulates cyclin A and cyclin B1 mRNA stability during cell proliferation," *EMBO Journal*, vol. 19, no. 10, pp. 2340–2350, 2000.

[107] D. Ishimaru, S. Ramalingam, T. K. Sengupta et al., "Regulation of Bcl-2 expression by HuR in HL60 leukemia cells and A431 carcinoma cells," *Molecular Cancer Research*, vol. 7, no. 8, pp. 1354–1366, 2009.

[108] K. Abdelmohsen, A. Lal, H. K. Hyeon, and M. Gorospe, "Posttranscriptional orchestration of an anti-apoptotic program by HuR," *Cell Cycle*, vol. 6, no. 11, pp. 1288–1292, 2007.

[109] K. Abdelmohsen, R. Pullmann Jr., A. Lal et al., "Phosphorylation of HuR by Chk2 Regulates SIRT1 Expression," *Molecular Cell*, vol. 25, no. 4, pp. 543–557, 2007.

[110] S. J. Cho, J. Zhang, and X. Chen, "RNPC1 modulates the RNA-binding activity of, and cooperates with, HuR to regulate p21 mRNA stability," *Nucleic Acids Research*, vol. 38, no. 7, Article ID gkp1229, pp. 2256–2267, 2010.

[111] E. S. Akool, H. Kleinert, F. M. A. Hamada et al., "Nitric oxide increases the decay of matrix metalloproteinase 9 mRNA by inhibiting the expression of mRNA-stabilizing factor HuR," *Molecular and Cellular Biology*, vol. 23, no. 14, pp. 4901–4916, 2003.

[112] A. Huwiler, E. S. Akool, A. Aschrafi, F. M. A. Hamada, J. Pfeilschifter, and W. Eberhardt, "ATP Potentiates Interleukin-1β-induced MMP-9 Expression in Mesangial Cells via Recruitment of the ELAV Protein HuR," *The Journal of Biological Chemistry*, vol. 278, no. 51, pp. 51758–51769, 2003.

[113] H. Tran, F. Maurer, and Y. Nagamine, "Stabilization of urokinase and urokinase receptor mRNAs by HuR is linked to its cytoplasmic accumulation induced by activated mitogen-activated protein kinase-activated protein kinase 2," *Molecular and Cellular Biology*, vol. 23, no. 20, pp. 7177–7188, 2003.

[114] R. Dong, J. G. Lu, Q. Wang et al., "Stabilization of Snail by HuR in the process of hydrogen peroxide induced cell migration," *Biochemical and Biophysical Research Communications*, vol. 356, no. 1, pp. 318–321, 2007.

[115] S. Galban, Y. Kuwano, and R. Pullmann Jr., "RNA-binding proteins HuR and PTB promote the translation of hypoxia-inducible factor 1alpha," *Molecular and Cellular Biology*, vol. 28, no. 1, pp. 93–107, 2008.

[116] L. G. Sheflin, A. P. Zou, and S. W. Spaulding, "Androgens regulate the binding of endogenous HuR to the AU-rich 3′UTRs of HIF-1alpha and EGF mRNA," *Biochemical and Biophysical Research Communications* , vol. 322, no. 2, pp. 644–651, 2004.

[117] N. C. Meisner and W. Filipowicz, "Properties of the regulatory RNA-binding protein HuR and its role in controlling miRNA repression," *Advances in experimental medicine and biology*, vol. 700, pp. 106–123, 2010.

[118] M. R. Epis, A. Barker, and K. M. Giles, "The RNA-binding protein HuR opposes the repression of ERBB-2 gene expression by microRNA miR-331-3p in prostate cancer cells," *The Journal of Biological Chemistry*, vol. 286, no. 48, pp. 41442–41454, 2011.

[119] L. E. Young, A. E. Moore, L. Sokol, N. Meisner-Kober, and D. A. Dixon, "The mRNA stability factor HuR inhibits microRNA-16 targeting of cyclooxygenase-2," *Molecular Cancer Research*, vol. 10, no. 1, pp. 167–180, 2012.

[120] H. K. Hyeon, Y. Kuwano, S. Srikantan, K. L. Eun, J. L. Martindale, and M. Gorospe, "HuR recruits let-7/RISC to repress c-Myc expression," *Genes and Development*, vol. 23, no. 15, pp. 1743–1748, 2009.

[121] Q. Jing, S. Huang, S. Guth et al., "Involvement of microRNA in AU-Rich element-mediated mRNA instability," *Cell*, vol. 120, no. 5, pp. 623–634, 2005.

[122] M. Kullmann, U. Göpfert, B. Siewe, and L. Hengst, "ELAV/Hu proteins inhibit p27 translation via an IRES element in the p27 5′UTR," *Genes and Development*, vol. 16, no. 23, pp. 3087–3099, 2002.

[123] D. Durie, S. M. Lewis, U. Liwak, M. Kisilewicz, M. Gorospe, and M. Holcik, "RNA-binding protein HuR mediates cytoprotection through stimulation of XIAP translation," *Oncogene*, vol. 30, no. 12, pp. 1460–1469, 2011.

[124] A. Lal, T. Kawai, X. Yang, K. Mazan-Mamczarz, and M. Gorospe, "Antiapoptotic function of RNA-binding protein HuR effected through prothymosin α," *EMBO Journal*, vol. 24, no. 10, pp. 1852–1862, 2005.

[125] C. L. Costantino, A. K. Witkiewicz, Y. Kuwano et al., "The role of HuR in gemcitabine efficacy in pancreatic cancer: HuR up-regulates the expression of the gemcitabine metabolizing

enzyme deoxycytidine kinase," *Cancer Research*, vol. 69, no. 11, pp. 4567–4572, 2009.

[126] J. R. Brody and G. E. Gonye, "HuR's role in gemcitabine efficacy: an exception or opportunity?" *Wiley Interdisciplinary Reviews*, vol. 2, no. 3, pp. 435–444, 2011.

[127] I. Lopez de Silanes, M. P. Quesada, and M. Esteller, "Aberrant regulation of messenger RNA 3′-untranslated region in human cancer," *Cellular Oncology*, vol. 29, no. 1, pp. 1–17, 2007.

[128] F. M. Gratacos and G. Brewer, "The role of AUF1 in regulated mRNA decay," *Wiley Interdisciplinary Reviews*, vol. 1, no. 3, pp. 457–473, 2011.

[129] C. Barreau, L. Paillard, A. Méreau, and H. B. Osborne, "Mammalian CELF/Bruno-like RNA-binding proteins: molecular characteristics and biological functions," *Biochimie*, vol. 88, no. 5, pp. 515–525, 2006.

[130] I. A. Vlasova and P. R. Bohjanen, "Posttranscriptional regulation of gene networks by GU-rich elements and CELF proteins," *RNA Biology*, vol. 5, no. 4, pp. 201–207, 2008.

[131] D. Mukhopadhyay, C. W. Houchen, S. Kennedy, B. K. Dieckgraefe, and S. Anant, "Coupled mRNA stabilization and translational silencing of cyclooxygenase-2 by a novel RNA binding protein, CUGBP2," *Molecular Cell*, vol. 11, no. 1, pp. 113–126, 2003.

[132] D. A. Dixon, G. C. Balch, N. Kedersha et al., "Regulation of cyclooxygenase-2 expression by the translational silencer TIA-1," *Journal of Experimental Medicine*, vol. 198, no. 3, pp. 475–481, 2003.

[133] M. Ouellette and B. Papadopoulou, "Coordinated gene expression by post-transcriptional regulons in African trypanosomes," *Journal of Biology*, vol. 8, no. 11, article 100, 2009.

[134] C. Clayton and M. Shapira, "Post-transcriptional regulation of gene expression in trypanosomes and leishmanias," *Molecular and Biochemical Parasitology*, vol. 156, no. 2, pp. 93–101, 2007.

[135] S. Haile and B. Papadopoulou, "Developmental regulation of gene expression in trypanosomatid parasitic protozoa," *Current Opinion in Microbiology*, vol. 10, no. 6, pp. 569–577, 2007.

[136] J. B. Palenchar and V. Bellofatto, "Gene transcription in trypanosomes," *Molecular and Biochemical Parasitology*, vol. 146, no. 2, pp. 135–141, 2006.

[137] V. Anantharaman, E. V. Koonin, and L. Aravind, "Comparative genomics and evolution of proteins involved in RNA metabolism," *Nucleic Acids Research*, vol. 30, no. 7, pp. 1427–1464, 2002.

[138] W. J. Ma, S. Cheng, C. Campbell, A. Wright, and H. Furneaux, "Cloning and characterization of HuR, a ubiquitously expressed Elav-like protein," *The Journal of Biological Chemistry*, vol. 271, no. 14, pp. 8144–8151, 1996.

[139] D. Antic and J. D. Keene, "Embryonic lethal abnormal visual RNA-binding proteins involved in growth, differentiation, and posttranscriptional gene expression," *American Journal of Human Genetics*, vol. 61, no. 2, pp. 273–278, 1997.

[140] S. Fumagalli, N. F. Totty, J. J. Hsuan, and S. A. Courtneidge, "A target for Src in mitosis," *Nature*, vol. 368, no. 6474, pp. 871–874, 1994.

[141] S. J. Taylor and D. Shalloway, "An RNA-binding protein associated with Src through its SH2 and SH3 domains in mitosis," *Nature*, vol. 368, no. 6474, pp. 867–871, 1994.

[142] M. Di Fruscio, T. Chen, and S. Richard, "Characterization of Sam68-like mammalian proteins SLM-1 and SLM-2: SLM-1 is a Src substrate during mitosis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 6, pp. 2710–2715, 1999.

[143] J. P. Venables, C. Vernet, S. L. Chew et al., "T-STAR/ETOILE: a novel relative of SAM68 that interacts with an RNA-binding protein implicated in spermatogenesis," *Human Molecular Genetics*, vol. 8, no. 6, pp. 959–969, 1999.

[144] M. P. Paronetto and C. Sette, "Role of RNA-binding proteins in mammalian spermatogenesis," *International Journal of Andrology*, vol. 33, no. 1, pp. 2–12, 2009.

[145] J. Marcotrigiano, A. C. Gingras, N. Sonenberg, and S. K. Burley, "Cocrystal structure of the messenger RNA 5′ cap-binding protein (eIF4E) bound to 7-methyl-GDP," *Cell*, vol. 89, no. 6, pp. 951–961, 1997.

[146] A. C. Gingras, B. Raught, and N. Sonenberg, "eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation," *Annual Review of Biochemistry*, vol. 68, pp. 913–963, 1999.

[147] M. Altmann, P. P. Muller, J. Pelletier, N. Sonenberg, and H. Trachsel, "A mammalian translation initiation factor can substitute for its yeast homologue in vivo," *The Journal of Biological Chemistry*, vol. 264, no. 21, pp. 12145–12147, 1989.

[148] H. Matsuo, H. Li, A. M. McGuire et al., "Structure of translation factor elF4E bound to m7GDP and interaction with 4E-binding protein," *Nature Structural Biology*, vol. 4, no. 9, pp. 717–724, 1997.

[149] K. Tomoo, X. Shen, K. Okabe et al., "Crystal structures of 7-methylguanosine 5′-triphosphate (m7GTP)- and P1-7-methylguanosine-P3-adenosine-5′, 5′-triphosphate (m7GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region," *Biochemical Journal*, vol. 362, no. 3, pp. 539–544, 2002.

[150] A. F. Monzingo, S. Dhaliwal, A. Dutt-Chaudhuri et al., "The structure of eukaryotic translation initiation factor-4E from wheat reveals a novel disulfide bond," *Plant Physiology*, vol. 143, no. 4, pp. 1504–1518, 2007.

[151] R. E. Rhoads, "EIF4E: new family members, new binding partners, new roles," *The Journal of Biological Chemistry*, vol. 284, no. 25, pp. 16711–16715, 2009.

[152] G. Hernández and P. Vazquez-Pianzola, "Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families," *Mechanisms of Development*, vol. 122, no. 7-8, pp. 865–876, 2005.

[153] A. R. Morris, N. Mukherjee, and J. D. Keene, "Systematic analysis of posttranscriptional gene expression," *Wiley Interdisciplinary Reviews*, vol. 2, no. 2, pp. 162–180, 2010.

[154] S. C. Janga and N. Mittal, "Construction, structure and dynamics of post-transcriptional regulatory network directed by RNA-binding proteins," *Advances in Experimental Medicine and Biology*, vol. 722, pp. 103–117, 2011.

[155] I. Topisirovic, N. Siddiqui, and K. L. B. Borden, "The eukaryotic translation initiation factor 4E (eIF4E) and HuR RNA operons collaboratively regulate the expression of survival and proliferative genes," *Cell Cycle*, vol. 8, no. 7, pp. 960–961, 2009.

[156] I. Topisirovic, N. Siddiqui, S. Orolicki et al., "Stability of eukaryotic translation initiation factor 4E mRNA is regulated by HuR, and this activity is dysregulated in cancer," *Molecular and Cellular Biology*, vol. 29, no. 5, pp. 1152–1162, 2009.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 256848, 14 pages
doi:10.1155/2012/256848

Review Article

On the Diversification of the Translation Apparatus across Eukaryotes

Greco Hernández,[1] Christopher G. Proud,[2] Thomas Preiss,[3] and Armen Parsyan[4,5]

[1] *Division of Basic Research, National Institute for Cancer (INCan), Avenida San Fernando No. 22, Col. Sección XVI, Tlalpan, 14080 Mexico City, Mexico*
[2] *Centre for Biological Sciences, University of Southampton, Life Sciences Building (B85), Southampton SO17 1BJ, UK*
[3] *Genome Biology Department, The John Curtin School of Medical Research, The Australian National University, Building 131, Garran Road, Acton, Canberra, ACT 0200, Australia*
[4] *Goodman Cancer Centre and Department of Biochemistry, Faculty of Medicine, McGill University, 1160 Pine Avenue West, Montreal, QC, Canada H3A 1A3*
[5] *Division of General Surgery, Department of Surgery, Faculty of Medicine, McGill University Health Centre, Royal Victoria Hospital, 687 Pine Avenue West, Montreal, QC, Canada H3A 1A1*

Correspondence should be addressed to Greco Hernández, ghernandezr@incan.edu.mx

Received 11 December 2011; Accepted 7 March 2012

Academic Editor: Brian Wigdahl

Diversity is one of the most remarkable features of living organisms. Current assessments of eukaryote biodiversity reaches 1.5 million species, but the true figure could be several times that number. Diversity is ingrained in all stages and echelons of life, namely, the occupancy of ecological niches, behavioral patterns, body plans and organismal complexity, as well as metabolic needs and genetics. In this review, we will discuss that diversity also exists in a key biochemical process, translation, across eukaryotes. Translation is a fundamental process for all forms of life, and the basic components and mechanisms of translation in eukaryotes have been largely established upon the study of traditional, so-called model organisms. By using modern genome-wide, high-throughput technologies, recent studies of many nonmodel eukaryotes have unveiled a surprising diversity in the configuration of the translation apparatus across eukaryotes, showing that this apparatus is far from being evolutionarily static. For some of the components of this machinery, functional differences between different species have also been found. The recent research reviewed in this article highlights the molecular and functional diversification the translational machinery has undergone during eukaryotic evolution. A better understanding of all aspects of organismal diversity is key to a more profound knowledge of life.

1. Protein Synthesis Is a Fundamental Process of Life

Proteins are one of the elementary components of life and account for a large fraction of mass in the biosphere. They catalyze most reactions that sustain life and play structural, transport, and regulatory roles in all living organisms. Hence, "translation," that is, the synthesis of proteins by the ribosome using messenger (m)RNA as the template, is a fundamental process for all forms of life, and a large proportion of an organism's energy is committed to translation [1, 2]. Accordingly, regulating protein synthesis is crucial for all organisms. Indeed, many mechanisms to control gene expression at the translational level have evolved in eukaryotes [3]. These mechanisms have endowed eukaryotes with the potential to rapidly and reversibly respond to stress or sudden environmental changes [1, 2, 4]. Translational control also plays a crucial role in tissues and developmental processes where transcription is quiescent, or where asymmetric spatial localization of proteins is required, such as early embryogenesis, learning and memory, neurogenesis, and gametogenesis [5–10]. Moreover, recent global

gene expression measurements have shown that the cellular abundance of proteins in mammalian cells is predominantly controlled at the level of translation [11, 12].

Eukaryotic translation is a sophisticated, tightly regulated, multistep process, the basic steps of which are conserved in all eukaryotes. It is performed by the ribosome together with multiple auxiliary "translation" factors (proteins) and is divided into four steps: initiation, elongation, termination, and recycling. These basic processes of translation were established experimentally in eukaryotes some decades ago, and many regulatory mechanisms have been subsequently elucidated [13, 14]. However, it was only recently that, with the use of powerful genome-wide sequencing, proteomics and bioinformatics-based technologies, a surprising diversity in components of the translation apparatus across eukaryotes was unveiled. In some cases, even functional differences between same molecules from different species have also been identified. Additionally, there is evidence that even the genetic code itself has continued to evolve in some phyla. These findings indicate that after eukaryotes emerged, the translational apparatus further evolved during eukaryotic diversification. In this article, we will review recent research revealing the diversification that the genetic code and many components of the translational machinery have undergone across eukaryotes.

2. Overview of the Translation Process in Eukaryotes

2.1. Initiation. The aim of the initiation step is both to ensure the recruitment of the mRNA to the ribosome and the positioning the ribosome in the proper frame at the start codon, which is achieved in a set of steps mediated by eukaryotic initiation factors (eIF). For most eukaryotic mRNAs, this happens by the so-called cap-dependent mechanism (Figure 1) [15–18]. It begins with the dissociation of the ribosome into its 60S and 40S subunits by eIF6. Free 40S subunit, which is stabilized by eIF3, eIF1, and eIF1A, binds to a ternary complex (consisting of eIF2 bound to an initiator Met-tRNA$_i^{Met}$ and GTP) to form a 43S preinitiation complex. On the other hand, the cap structure (m^{7}GpppN, where N is any nucleotide) of the mRNA is recognized by eIF4E in complex with the scaffold protein eIF4G. Then, the 43S preinitiation complex is recruited to the 5′ end of the mRNA, a process that is coordinated by eIF4E through its interactions with eIF4G and the 40S ribosomal subunit-associated eIF3. The ribosomal complex then scans in a 5′ $\longrightarrow$ 3′ direction along the 5′-untranslated region (UTR) through interactions with the eIF4G-bound RNA helicase eIF4A and eIF4B to reach the start codon, usually an AUG. During scanning, eIF4B stimulates the activity of eIF4A which unwinds secondary RNA structures in the mRNA. eIF1, eIF1A, and eIF5 assist in the positioning of the 40S ribosomal subunit at the correct start codon so that eIF2 can deliver the anti-codon of the initiator Met-tRNA$_i^{Met}$ as the cognate partner for the start codon, directly to the peptidyl (P)-site of the 40S ribosomal subunit. Once the ribosomal subunit is placed on the start codon, a 48S pre-initiation complex is formed. Then, eIF5 promotes GTP hydrolysis by eIF2 to release the eIF proteins. Finally, the GTPase eIF5B is required for the joining of the 60S ribosomal subunit to the 40S subunit to form an 80S initiation complex. The poly A-binding protein (PABP) is able to interact with the 3′-poly(A) tail and eIF4G promoting circularization of the mRNA and increasing the efficiency of subsequent rounds of initiation (Figure 1) [15–20].

In the case of some viral and cellular mRNAs, 5′-UTR recognition by the 40S ribosomal subunit happens without involvement of eIF4E and is, instead, driven by RNA structures located in *cis* within the mRNA itself. Such structures are operationally defined as internal ribosome entry site (IRES) and are located in the proximity of the start codon ([21–23]; Martinez-Salas et al. this issue).

2.2. Elongation. After initiation, the 80S ribosome is assembled at the start codon of the mRNA containing a Met-tRNA$_i^{Met}$ in the P-site. Then, elongation takes place (Figure 1); this is the process of decoding codons and formation of peptide bonds sequentially to add amino acid residues to the carboxy-terminal end of the nascent peptide [16, 24–26]. This process is assisted by elongation factors (eEF) and involves four major steps. (1) Formation of the ternary complex eEF1A·GTP·aminoacyl-tRNA and delivery of the first elongator aminoacyl-tRNAs to an empty ribosomal tRNA-binding site called the A-(acceptor) site. It is in the A-site where codon/anticodon decoding takes place. (2) Interaction of the ribosome with the mRNA-tRNA. This duplex activates eEF1A·GTP hydrolysis and guanine nucleotide exchange on eEF1A. (3) Peptide bond formation then occurs between the P-site peptidyl-tRNA and the incoming aminoacyl moiety of an A-site aminoacyl-tRNA. This reaction is catalyzed by the peptidyl transferase center of the 60S ribosomal subunit, and the products comprise of a new peptidyl-tRNA that is one amino acid residue longer and a deacylated (discharged) tRNA. (4) Binding of eEF2·GTP and GTP hydrolysis promotes the translocation of the mRNA such that the deacylated tRNA moves to the E-(exit) site, the peptidyl-tRNA is in the P-site, and the mRNA moves by three nucleotides to place the next mRNA codon into the A-site. The deacylated tRNA in E-site is then ejected from the ribosome. The whole process is repeated along the mRNA sequence until a stop codon is reached and the process of termination is initiated [16, 24–26].

2.3. Termination. Translation termination is mediated by two polypeptide chain-release factors, eRF1 and eRF3 (Figure 1). When any of the termination codons (UAA, UAG, and UGA) is exposed in the A-site, eRF1 recognizes the codon, binds the A-site, and triggers the release of the nascent polypeptide from the ribosome by hydrolysing the ester bond linking the polypeptide chain to the P-site tRNA. This reaction leaves the P-site tRNA in a deacylated state, leaving it to be catalyzed by the peptidyl transferase center of the ribosome. eRF1 recognizes stop signals and functionally acts as a tRNA-mimic, whereas eRF3 is a ribosome- and eRF1-dependent GTPase that, by forming a stable complex with eRF1, stimulates the termination process [16, 27, 28].

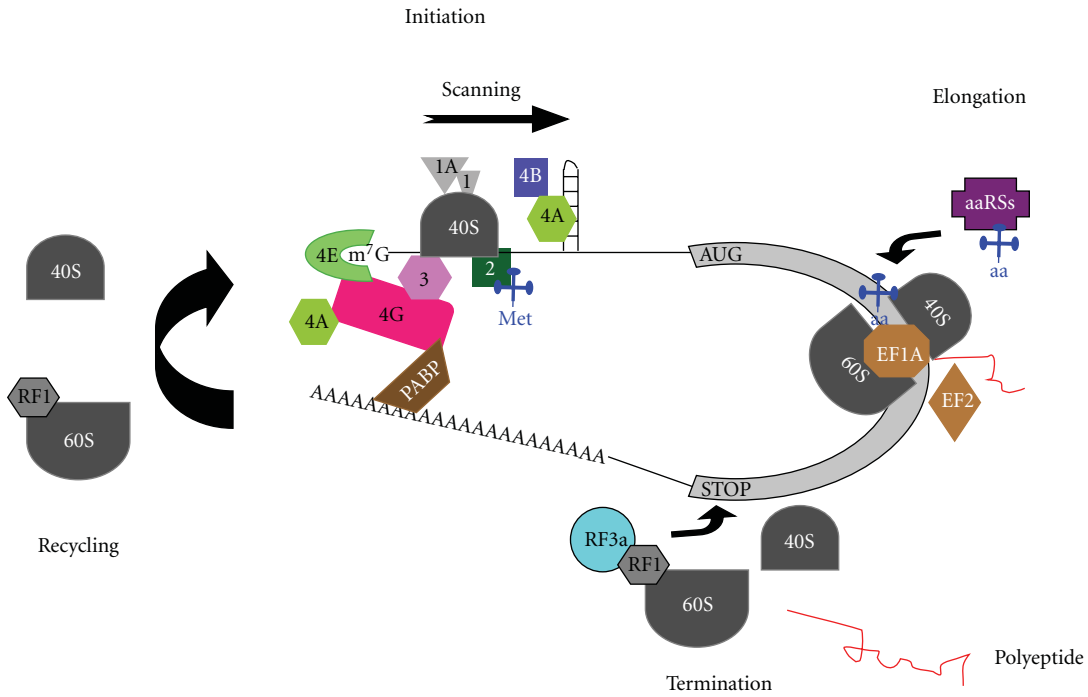


FIGURE 1: The general process of translation in eukaryotes. A typical eukaryotic mRNA is represented. The cap structure (m^7G), the open reading frame (light gray box) and the poly(A) tail are depicted. During *Initiation*, most eukaryotic mRNAs are translated by the cap-dependent mechanism, which requires recognition by eIF4E (green crescent) complexed with eIF4G (red) and eIF4A (light green)—the so-called eIF4F complex—of the cap structure at the 5′ end. A 43S preinitiation complex (consisting in a 40S ribosomal subunit (dark gray) loaded with eIF3 (pink), eIF1 and eIF1A (light grey), initiator Met-$tRNA_i^{Met}$ (blue clover), eIF2 (dark green), and GTP binds the eIF4F-mRNA complex and scans along the 5′-UTR of the mRNA to reach the start codon (usually an *AUG* triplet). During the scanning eIF4A, stimulated by eIF4B (dark blue), unwinds secondary RNA structure in an ATP-dependent manner. The poly A-binding protein (PABP, dark brown) binds both the poly(A) tail and eIF4G promoting mRNA circularization. *Elongation* is assisted by elongation factors eEF1A and eEF2 (light brown). During this step, aminoacyl-tRNA synthetases (aaRSs, purple) catalyze the binding of amino acids (aa) to cognate tRNAs. *Termination* is mediated by the release factors eRF1 (gray) and eRF3 (light blue) and happens when a termination codon (*STOP*) of the mRNA is exposed in the A-site of the ribosome. In this step, the completed polypeptide (red) is released. During *Recycling*, which is required to allow further rounds of translation, both ribosomal subunits dissociate from the mRNA. eRF1 remains associated with the posttermination complexes after polypeptide release.

2.4. Recycling. In the recycling step, both ribosomal subunits are dissociated, releasing the mRNA and deacetylated tRNA, so that both ribosomal subunits can be used for another round of initiation [16, 27, 28] (Figure 1). The closed-loop model proposes that, during translation, cross-talk occurs between both ends due to the circular conformation of the mRNA. According to this model, termination and recycling may not release the 40S ribosomal subunit back into the cytoplasm. Instead, this subunit may be passed from the poly(A) tail back to the 5′-end of the mRNA, so that a new round of initiation can be started [16, 27].

3. Divergence in the Genetic Code

The deciphering of the genetic code in the early 1960's established one of the basic foundations of modern biology. Soon after, the essential universality of the genetic code was recognized, that is, the assignment of 20 amino acids to 64 codons and two punctuation marks (start and stop signals) is substantially the same for all extant forms of life on earth [29]. Nevertheless, variations to the "universal" genetic code, wherein the meaning of a "universal" codon is changed to a different one, have recently been uncovered in a wide range of bacteria, organelles, and the nuclear genome of eukaryotes, revealing that the genetic code is still evolving in some lineages [30–33]. In eukaryotes, deviations from the standard nuclear genetic code have arisen independently multiple times in unicellular organisms of five lineages, namely, ciliates, Diplomonads, fungi (in the genus *Candida* and some ascomycetes), polymastigid oxymonads, and green algae (in Dasycladales and Cladophorales) [30, 31, 33–39]. Most codon variations in eukaryotes are found to be the reassignment of the stop codons UAG and UAA to glutamine, and the stop codon UGA to tryptophan or cysteine (Figure 2). All reported code variations in ciliates, Diplomonads, and green algae belong to this kind. In contrast, *Candida* ambiguously utilizes the codon CUG (universally used for leucine) for both serine and leucine. The underlying mechanisms of codon reassignment are mutations in tRNA genes that affect decoding, RNA editing, or mutations in eRF1 [30, 31, 34–39].

The observation that the same codon reassignments have occurred independently in closely related species (within the yeasts, green algae, and ciliate taxa) supports the notion that these changes provide a selective advantage in similar ecological niches [30]. Whether there is a restriction for the genetic code to change in multicellular organisms is not known.

4. Diversity in the Initiation Step

4.1. Functional Divergence of eIF Proteins. While the fundamental principles of translation are well conserved across all forms of life, in eukaryotes the initiation step has undergone substantial increase in complexity as compared to prokaryotes [3, 22, 40–44]. Most evidence for molecular and functional diversification among the translation components has been found in the eIF4 proteins (Figure 2). Most eukaryotic phyla possess several paralog genes for members of the eIF4 families, with well-documented differential expression patterns and variable biochemical properties among paralogs of the same organism [45–72]. For eIF4E and eIF4G cognates, even evidence of physiological specialization has been found among both unicellular and multicellular organisms (Table 1). These findings support the hypothesis that in organisms with several paralogs, an ubiquitous set of eIF4 factors supports global translation initiation whereas other paralogs perform their activity in specific cellular processes [45]. In some cases, eIF4E cognates have evolved towards translational repressors. Class 2 eIF4Es are exemplified by eIF4E-homolog protein (4E-HP) in human, eIF4E-2 in mouse [63], eIF4E-8 in *Drosophila* [52, 58, 73], IF4 in *C. elegans* [74, 75], and nCBP in *A. thaliana* [76], and they can bind the 5′ cap structure of mRNA but do not bind eIF4G [58, 77], thereby acting as a translational repressors of mRNAs associated with it [73, 78]. Class 2 eIF4Es are widespread across metazoa, plants, and some fungi although absent in the model ascomycetes *S. cerevisiae* and *S. pombe* [46]. Since the *Arabidopsis* [76] and *Caenorhabditis* [74] orthologs promote translation of some mRNAs, it seems most likely 4E-HP diverged from a widespread ancestral eIF4E to form a translational repressor in metazoa [3]. A similar example is eIF4E-1B, which emerged only in vertebrates as a translational repressor of a subset of oocyte mRNAs [57, 59, 79], and *Leishmania* eIF4E-1, which under heat shock conditions binds to a *Leishmania*-specific 4E-BP and becomes translationally inactive [71]. In other cases, eIF4E cognates have evolved towards a new molecular function not related to translation. This is the case with *Trypanosoma* eIF4E-1 and eIF4E-2, which are essential nuclear and cytoplasmic proteins, respectively [49], and *Giardia* (eIF4E-2), which binds only to nuclear noncoding small RNAs [64]. However, it is also possible that this was an ancestral function of eIF4E [22, 40].

Whereas the need for distinct eIF4 proteins in different tissues may have been the driving force behind the evolution of various paralogs in multicellular organisms, in unicellular eukaryotes different paralogs may be differentially needed during distinct life stages [49]. Specific features of mRNA metabolism in some phyla also might have driven the evolution of eIF4Es in specific organisms, such as the use of different cap structures (usually mono- and trimethylated) in mRNAs from worms of the phylum Nematoda [50, 51, 54, 80], and flagellate protists of the order Kinetoplastida [49, 65, 66]. These mRNAs result from the *trans*-splicing process to produce mature mRNAs.

Other eIFs have also undergone molecular diversification across eukaryotes, including the multisubunit eIF3 whose subunit composition ranges from 5 to 13 nonidentical polypeptides in different phyla [99], and eIF6 that is duplicated into two or three paralogs in plants [100]. However, the functional relevance of these phenomena (if any) is not known.

4.2. Multiple RNA Helicases for Translation Initiation. The evolution of cap-dependent translation has led to a dependency on RNA helicase activity to unwind the 5′-UTR secondary structure during the scanning [22, 40]. The DEAD-box RNA helicase/ATPase eIF4A is the main helicase thought to perform this activity. Recently, other RNA helicases from diverse organisms have also been found to facilitate translation of specific mRNAs with structured 5′-UTRs (Figure 2). Such is the case of the mammalian, *Drosophila* and yeast DEAD-box helicases DDX3 and Ded1, as well as the human DExH-box helicases RHA and DHX29 [101–103]. In *Drosophila*, the DEAD-box helicase Vasa interacts with eIF5B and regulates the translation of *gurken* and *mei-P26* mRNAs. Evidence supports the idea that Vasa is a translational activator of specific mRNAs involved in germline development [6, 7]. In contrast, orthologs of the *Xenopus* helicase Xp54 (DEAD-box, DDX6-like helicases) in a spectrum of organisms, including *Drosophila* Me31B, *Saccharomyces* Dhh1, human rck/p54, and *Caenorhabditis* CGH-1 have been found to repress translation of stored mRNAs and promote aggregation into germplasm-containing structures [104].

Most RNA helicases involved in translation also play a variety of roles in other processes of RNA metabolism, including mRNA RNP assembly, RNA degradation, RNA export, and splicing [103]. This functional versatility of RNA helicases leads us to speculate that a wider diversity of other, yet unidentified, helicases might be involved in translation in all eukaryotes. This could be the case of the *Arabidopsis* eIF4F complex, which contains eIF4A in proliferating cells but different RNA helicases in quiescent cells [105]. Whether these helicases play a role in translation is not known.

4.3. Divergence in the Regulation of Initiation: Diversity of eIF4E-Binding Proteins. Almost twenty years ago, it was discovered that eIF4E is negatively regulated in mammalian cells by three related proteins, the eIF4E-binding proteins (4E-BPs) 1, 2, and 3. These proteins share with eIF4G the motif YXXXXLϕ (where X is any amino acid and ϕ is a hydrophobic residue) that interacts with the convex dorsal surface of eIF4E, so binding of 4E-BPs to eIF4E precludes its association with eIF4G and represses cap-dependent

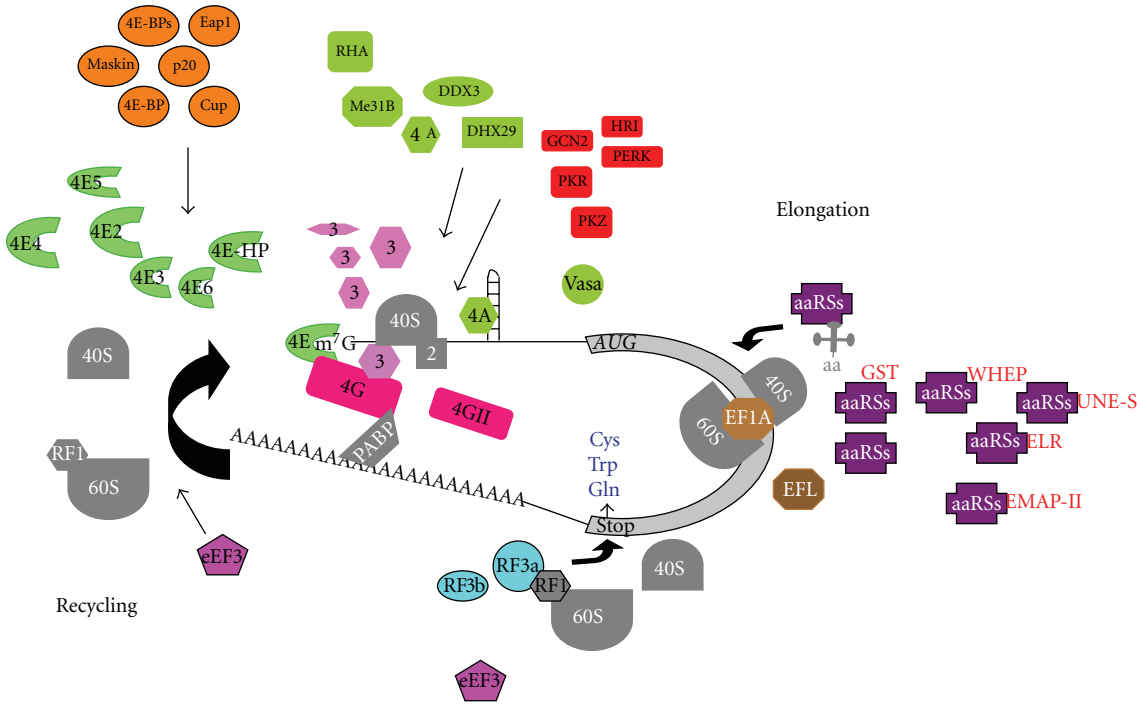


FIGURE 2: Diversity in the configuration of the translation apparatus across eukaryotes. The different components of the translation machinery that show diversity in different phyla are shown in colors. Components with some diversity that is not discussed here are depicted in gray. Several copies of eIF4E (green crescent) and eIF4G (red) have been found in plants, metazoan, and protists. In some cases, eIF4E cognates have evolved towards translational repressors (4E-HP is an example). Many 4E-binding proteins (orange) have been discovered in species from metazoan, fungi and protists. The subunit composition of eIF3 (pink) ranges from 5 to 13 nonidentical polypeptides in different phyla. There is, however, a core of five homolog subunits shared by most eukaryotes. Several RNA helicases (light green) from diverse organisms have been found to be involved in *Initiation*. A family of five kinases (*HRI*, *PERK*, *GCN2*, *PKR*, and *PKZ*, red) phosphorylate the alpha subunit of eIF2 to inhibit global translation under stress conditions. The presence of eIF2alpha kinases varies in different lineages. Different domains (red), such as *WHEP*, *EMAPII*, *ELR*, *GST*, and *UNE-S*, have been added to different aminoacyl-tRNA synthetases (*aaRSs*, purple) in distinct phyla of multicellular species. For *Elongation* to happen, a number of protist, algae and fungi (most of them unicellular organisms) lack eEF1A (light brown) and instead possess the related factor elongation factor-like (EFL, dark brown). For *Termination*, most organisms only contain a single eRF3 (light blue). In contrast, mammalian species express two eRF3s (viz. eRF3a and eRF3b). Ribosomes from all eukaryotes perform *Elongation* with eEF1A and eEF2. However, the yeast *S. cerevisiae* requires an additional essential factor, eEF3 (light purple), for *Elongation* to proceed. Genes encoding eEF3 have been found exclusively in many species of fungi. Evidence supports the notion that eEF3 activity promotes ribosome recycling. Variations to the "universal" genetic code, wherein the meaning of a "universal" codon is changed to a different one, exist in several species of in unicellular eukaryotes. Most codon variations are the reassignment of the stop codons UAG and UAA to glutamine, and the stop codon UGA to tryptophan or cysteine.

translation [8, 106]. In the last years, a myriad of 4E-binding proteins has been discovered in species from distantly related taxa, including mammals, plants, *Drosophila*, *Caenorhabditis*, yeast [3, 8, 106], and *Leishmania* [71] (Figure 2). Interestingly, most 4E-BPs are phylogenetically unrelated to each other and control translation in disparate, species-specific processes, such as embryogenesis in *Drosophila*, neurogenesis in mammals, or pseudohyphal growth in yeast. Moreover, some 4E-BPs utilize non-canonical motifs to bind eIF4E. These observations support the idea that binding to eIF4E evolved independently in multiple taxonomic groups [3].

4.4. Divergence in the Regulation of Initiation: The Case of eIF4E Phosphorylation. In mammalian cells, the kinases ERK or p38MAPK phosphorylate and activate the MAPK-interacting kinases (Mnk1/2). Mnk interacts with the carboxy-terminal part of eIF4G to directly phosphorylate eIF4E on Ser-209. This phosphorylation appears to regulate the function of eIF4E although the precise consequences are unclear [107–110]. Mammals possess two Mnk genes (*MKNK1/2*) which in humans, but not mice, give rise to four Mnk isoforms by alternative splicing; these isoforms have distinct properties in terms of activity, regulation, and subcellular localization [111]. In *Drosophila*, the single Mnk orthologue, LK6, also phosphorylates eIF4E-1 at a serine residue corresponding to mammalian Ser-209, a phosphorylation that is critical for development and cell growth [112–115]. However, the effects of phosphorylation on eIF4E activity and its physiological relevance are different across eukaryotes. Indeed, a residue equivalent to Ser-209 is present in metazoan eIF4Es but is absent in different fungi, protists and plants ([67];R. Jagus et al., this issue). Accordingly, Mnk is conserved among metazoans, but no Mnk ortholog exists in *S. cerevisiae* or plants, whose eIF4Gs

Table 1: Specialized activities of eIF4 proteins.

Protein[a]	Activity	Reference
eIF4E cognates		
Dm eIF4E-1, M eIF4E-1, Ce IFE-3, Sp eIF4E-1, Sc eIF4E, Plant eIF4E and eIF(iso)4E, Z eIF4E-1A, Gl eIF4E-2; Tb eIF4E-3 and eIF4E-4; Lm eIF4E-1 and eIF4E-4	Supports general cap-dependent initiation of translation. Essential gene.	[49, 54, 55, 57, 58, 62, 64, 65, 67, 72, 81–84]
M eIF4E-1	mRNA nucleocytoplasm transport.	[85]
Dm eIF4E-1	Involved in *sex-lethal* (*Sxl*)-dependent female-specific alternative splicing of male specific lethal-2 (msl-2) mRNA and *Sxl* pre-mRNAs.	[86]
Sp eIF4E-2	Supports cap-dependent translation initiation during stress response.	[62]
Ce IFE-1	Required for gametogenesis.	[87–89]
Ce IFE-2	Involved in chromosome segregation at meiosis at elevated temperatures.	[90]
Ce IFE-4	Promotes expression of specific mRNAs involved in egg lying. Nonessential gene.	[74]
Dm eIF4E-3	Testis-specific protein, essential for spermatogenesis.	[91]
La eF4E-4	Supports translation in promastigotes stage.	[71]
Dm 4E-HP, M 4E-HP	Negative regulator of translation.	[58, 73, 77, 78]
Xl eIF4E-1B	Negative regulator of translation.	[57, 79]
La eIF4E-1	Represses translation under heat shock conditions.	[71]
Gl eIF4E-1	Involved in nuclear snRNAs metabolism and play no role in translation.	[64]
Tb eIF4E-1 and eIF4E-2	Essential genes. Play no role in translation.	[49]
eIF4G cognates		
M eIF4G-I and eIF4G-II, Dm eIF4G, Sc eIF4G-I and eIF4G-II, plant eIF4G and Plant eIF(iso)4G, Ce p170 of IFG-1,	Scaffold protein. Supports general cap- and IRES-dependent initiation of translation.	[55, 60, 67, 71, 92–96]
Dm eIF4G-2	Support translation initiation in testis.	[47, 48]
M eIF4G-2	Involved in hematopoietic cell differentiation.	[97]
M eIF4G-3	Essential for spermatogenesis.	[98]
Ce IFG-1	p130 of *ifg-1* gene is involved in mitotic and early meiotic germ cell development.	[93]
La eIF4G-3	Supports translation in promastigotes stage.	[71]

[a] At, *Arabidopsis thaliana*; Ce, *Caenorhabditis elagans*; Dm, *Drosophila melanogaster*; Lm, *Leishmania major*; La, *Leishmania amazonensis*; M, *mammalian*; Nt, *N. tabacum*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; W, *wheat germ*; Xl, *Xenopus laevis*; Z, *zebra fish*; Gl, *Giardia lamblia*; Tb, *Trypanosoma brusei*.

lack a Mnk-binding domain ([67]; R. M. Patrick and K. S. Browning, this issue). Moreover, *Trypanosoma* eIF4E-3 [116] and *S. cerevisiae* eIF4E [117] are phosphorylated on residues which are not equivalent to mammalian Ser-209, and *S. cerevisiae* cells expressing a nonphosphorylatable version as sole source of eIF4E do not display any evident defect on global protein synthesis or cell growth [117]. These observations support the idea that eIF4E phosphorylation at Ser-209 by the MAPK-Mnk signaling pathway evolved only in metazoans and that, perhaps, alternative mechanisms regulate eIF4E in nonmetazoan eukaryotes [3].

4.5. Diversity in the Regulation of Initiation: The Case of eIF2alpha Phosphorylation. Under different stress conditions, general protein synthesis is inhibited through phosphorylation of the alpha subunit of eIF2 at Ser-51 by a family of kinases that are present in widely scattered lineages. They include the double-stranded RNA protein

kinase (PKR) that is activated during viral infection, the heme-regulated inhibitor kinase (HRI) that is activated under heme deprivation or arsenite exposure, the PKR-like endoplasmatic reticulum kinase (PERK) that is activated by unfolded proteins in the lumen of the endoplasmic reticulum, and the general control nonderepressible 2 (GCN2) that is activated by uncharged tRNA and thus senses amino acid starvation [107, 118] (Figure 2). The presence of eIF2alpha kinases varies in different lineages; while GCN2 is present in all eukaryotes; PERK is found in only metazoans; HRI is found in vertebrates, the dipteran *Anopheles*, the fungi *Schizosaccharomyces,* and the echinoderm *Strongylocentrotus*; PKR is only found in vertebrates [53, 118]. Interestingly, in some teleost fishes, PKR has undergone further duplication into PKR and PKZ, which perhaps led teleost fishes to respond to an extended range of viral infections [119].

5. Diversity in the Elongation Step

5.1. Divergence in the Aminoacyl-tRNA Synthetases. The process of elongation is highly conserved among all forms of life [16, 24, 25]. Key molecules for elongation are aminoacyl-tRNA synthetases (aaRSs), which catalyze the aminoacylation reaction whereby an amino acid is attached to the cognate tRNA. aaRSs are the only components of the gene expression machinery that function at the interface between nucleic acids and proteins. Thus, by performing their activity, aaRSs establishes the fundamental rules of the universal genetic code and, thus, of translation. aaRSs constitute a family of 20 essential cellular enzymes that are grouped into two classes: class I, in which the aminoacylation domain has a Rossmann nucleotide-binding fold, and class II, in which this domain is a seven-stranded beta-sheet with flanking alpha-helices. The conservation of the genetic code suggests that aaRSs evolved very early before the emergence of the last universal common ancestor [120, 121].

Throughout evolution of multicellularity, different domains, such as the WHEP domain, the oligonucleotide binding fold-containing EMAPII domain, the tripeptide ELR (Glu-Leu-Arg), the glutathione S-transferase (GST) domain and a specialized amino-terminal helix (N-helix), have been progressively added to different aaRSs in distinct phyla (Figure 2). The tripeptide ELR and the EMAPII domain were incorporated simultaneously to TyrRSs in metazoans starting from insects; the WHEP domain is present in TrpRS only in chordates; a unique sequence motif, UNE-S, became fused to the C-terminal of SerRS of all vertebrates [120, 121]. In bilaterian animals, the glutamylRS and prolylRS were linked via WHEP domains giving rise to a bifunctional glutamyl-prolylRS [120, 121]. It was recently found that this fused enzyme is also present in the cnidarian *Nematostella*, which pushes the origin of glutamyl-prolylRS back to the cnidaria-bilaterian ancestor [122], and suggests that this enzyme further underwent fission in the nematode *C. elegans* where glutamylRS and prolylRS enzymes are separated. GlutamylRS and prolylRS are also separate in plants and fungi [120–122].

It has been found that the function of the aaRSs was either increased or impaired by the addition of the new domains. Whereas the WHEP domain regulates interaction of TrpRS with its cognate receptor, with MetRS this domain plays a tRNA-sequestering function. The Leu-zipper motif in ArgRS is important for the formation of multi aaRSs complex (MSC), which enhances channeling of tRNA to the ribosome. Moreover, different aaRSs play diverse roles in cellular activities beyond translation, such as stress response, plant and animal embryogenesis, cell death, immune responses, transcriptional regulation, and RNA splicing [120, 121, 123]. It was found that the incorporation of domains to aaRSs correlates positively with the increase in organism's complexity. For example, the number of aaRSs carrying the GST domain increases from two in fungi to four in insects, to five in fish, and six in humans [121]. Thus, it has been proposed that the newly fused aaRSs domains triggered the appearance of new biological functions for these proteins in different lineages and that the fusion of domains to aaRSs could have played an important part in expanding the complexity of newly emerging metazoan phyla [121].

5.2. Divergence in Elongation Factors. eEF1A plays a critical role in translation. It binds and delivers aa-tRNAs to the A-site of ribosomes during the elongation step. Because homologs of this essential protein occur in all domains of life, it was thought to exist in all eukaryotes. Strikingly, a recent genome-wide survey revealed that a number of lineages lack eEF1A and instead possess a related factor called elongation factor-like (EFL) protein that retains the residues critical for eEF1A function [124] (Figure 2). It was later found that EFL-encoding species are scattered widely across eukaryotes and that *eEF1A* and *EFL* genes display mutually exclusive phylogenetic distributions. Thus, it is assumed that eEF1A and EFL are functionally equivalent [124–132]. Since EFL is present only in eukaryotes, it is thought that eEF1A is ancestral to all extant eukaryotes and that a single duplication event in a specific lineage gave rise to EFL. EFL genes were then spread to other lineages via multiple independent lateral gene transfer events, where EFL took over the original eEF1A function resulting in secondary loss of the endogenous eEF1A. It is thought that both genes coexisted for some time before one or the other was lost. Indeed, the diatom *Thalassiosira* bears both *EFL* and *eEF1A* genes [129] and might be an example of this situation. It is also possible that there was a single gain of EFL early in evolution followed by differential loss of it [124, 128, 129, 131, 132]. So far, EFL genes have been identified in widespread taxa, including diatoms, green and red algae, fungi, euglenozoans, foraminiferans, cryptophytes, goniomonads, katablepharid, chlorarachniophytes, oomycetes, dinoflagellates, choanozoans, centrohelids, and haptophytes [124–132]. Most of them are unicellular organisms. In contrast, eEF1A is found in most eukaryotes, and multiple copies of this gene have been found in some insect orders, including Coleoptera, Hymenoptera, Diptera, Thysanoptera, and Hemiptera [133].

The eEF1A activity is modulated by diverse post-translational modifications, including phosphorylation, lysine methylation, and methyl-esterification. eEF1A also undergoes modification by covalent binding of ethanolamine phosphoglycerol (EPG), whose function is not known and for whom the number of moieties attached varies in different eukaryotes [134]. Moreover, in addition to its role in translation, eEF1A has been reported to play several "moonlighting" functions, including binding to cytoskeletal proteins, signal transduction, protein nuclear export and import of tRNAs into mitochondria [134]. It is not known whether EFL undergoes the same posttranslational modifications as eEF1A does and whether it also displays non-translational activities.

6. Divergence in the Termination Step

The termination of protein synthesis is governed by eRF1, which is a monophyletic and highly conserved protein that is universally present in eukaryotes. Comprehensive analyses of genomic datasets show that eRF1 was inherited by eukaryotes from archaeal ancestors and that most eukaryotes encode only one eRF1. Known exceptions are *Arabidopsis thaliana*, which possesses three *eRF1* genes, and the ciliates *Tetrahymena, Oxytricha, Nyctotherus, Oxytricha*, *Euplotes,* and *Paramecium* which have two *eRF1* genes [135–138]. Interestingly, unusually high rates of eRF1 evolution have been found in organisms with variant genetic codes, especially in the N-terminal domain, which is responsible for stop-codon recognition [30, 34, 135, 136, 138, 139]. eRF1 displays structural similarity to tRNA molecules and mimics its activity during binding of ribosomal A-site during recognition of a stop codon [34, 139–141]. Since mutations in eRF1 N-terminal domain switch from omnipotent to bipotent mode for stop-codon specificity [35–38, 141], most likely the accelerated evolution of eRF1 in organisms with variations to the nuclear genetic code has been driven mainly to accommodate these variations [30, 31, 34–38, 135, 138–141].

eRF3 is a GTPase that stimulates the activity of eRF1 during the translation termination process. eRF3 arose in early eukaryotes by the duplication of the GTPase eEF1A. Consistent with this, eRF3 binds and transports eRF1, a structural mimic of tRNA, to the ribosomal A-site, similar to the role of eEF1A in binding and delivering aminoacyl-tRNAs to the same site during translation elongation [142, 143]. eRF3 is much more divergent than eRF1, especially in its N-terminal domain. In addition, *eRF3* is universal among eukaryotes, and most organisms only contain single-copies of this gene [137, 143]. In contrast, mammalian species express two eRF3s (viz. eRF3a and eRF3b; Figure 2). They possess different N regions and display drastically different tissue distribution and expression profiles during the cell cycle [143, 144]. Moreover, eRF3b but not eRF3a can substitute for yeast eRF3 in translation termination [145]. These observations indicate duplication and further functional divergence of eRF3 proteins in this lineage.

7. Divergence in the Recycling Step

Ribosomes from all eukaryotes perform elongation with eEF1A and eEF2. Interestingly, it has been known for some time that the yeast *Saccharomyces cerevisiae* requires an additional essential factor, eEF3, for the elongation cycle to proceed [146]. Genes encoding eEF3 were subsequently identified exclusively in other fungi (both yeasts and filamentous), including *Candida*, *Pneumocystis*, *Neurospora*, *Aspergillus,* and *Mucor* [147–150] (Figure 2). eEF3 is an ATPase that interacts with both ribosomal subunits and stimulates binding of aminoacyl-tRNA to the ribosomal A-site by enhancing the rate of deacylated tRNA dissociation from the E-site. Because E-site release is needed for efficient A-site binding of aminoacyl-tRNA, it was thought that eEF3 functions as a so-called "E-site" factor [16, 151]. Most recently, it was shown that post-termination complex, consisting of a ribosome, mRNA, and tRNA, is disassembled into single components by ATP and eEF3. Because the release of mRNA and deacylated tRNA and ribosome dissociation takes place simultaneously and no 40S—mRNA complexes remain, it is proposed that eEF3 activity promotes ribosome recycling [152]. "What were the evolutionary forces that led to the emergence of eEF3 exclusively in fungi?" is a very interesting, still open question.

8. Concluding Remarks

One of the most conspicuous features of life is its prominent ability to diversify. Current assessments of the biodiversity on Earth reaches 2 million species, although the true number of living organisms could easily be four times that number and likely much higher [153, 154]. The diversification of life has occurred at different levels, including the occupancy of ecological niches, behavioral patterns, body plans, and organismal complexity, and metabolic needs and capabilities. More recently, intensive whole-genome shotgun sequencing of microbial communities from different environments has unveiled a vast profusion of diversification also at the genetic level [155–157]. We have discussed that diversity also exists in the machinery that performs a fundamental process, translation, across eukaryotes. We speculate that the molecular diversification of the translation apparatus is among the basis that provided to early eukaryotes the scope to invade new ecological niches and overcome the different environmental and biological challenges this represented. Different evolutionary mechanisms might have been the driving forces leading to this molecular diversification in different lineages, including natural selection, sexual selection, genetic drift and neutral evolution. However, at this point, we can be nothing but speculative on the biological meaning of the molecular diversification reviewed here.

Traditional studies on so-called model organisms have taught us the global processes of eukaryotic translation. In the last years, the use of modern genome-wide, high-throughput technologies to study many non-model eukaryotes from different taxa has unveiled that diversification of the translation machinery configuration is far more expansive than previously thought. Collectively, these studies

show that the translation apparatus in eukaryotes is far from being evolutionarily static. Therefore, we anticipate that, as more organisms are studied, additional diversification of components of the translation apparatus will be revealed. We believe that a better understanding of the diversity of all levels of organism will provide us a more profound understanding of Life.

Acknowledgments

The authors are very thankful to Rosemary Jagus for critical review of the manuscript, and to Michelle Kowanda for proofreading the manuscript. G. Hernández is supported by the National Institute for Cancer (Instituto Nacional de Cancerología, México) and Consejo Nacional de Ciencia y Tecnología (CONACyT). C. G. Proud is supported by AstraZeneca, The Biotechnology and Biological Sciences Research Council, the British Heart Foundation, Cancer Research UK, Janssen, the Kerkut Trust, the Medical Research Council, The Royal Society, and The Wellcome Trust. The work in the laboratory of T. Preiss is supported by grants from the Australian Research Council and the National Health and Medical Research Council of Australia. In many cases, the authors cite reviews, not original research articles. They apologise to many authors for not citing their original literature.

References

[1] M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, "Origins and principles of translational control," in *Translational Control of Gene Expression*, N. Sonenberg, J. W. B. Hershey, and M. B. Mathews, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2000.

[2] M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, "Origins and principles of translational control," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[3] G. Hernández, M. Altmann, and P. Lasko, "Origins and evolution of the mechanisms regulating translation initiation in eukaryotes," *Trends in Biochemical Sciences*, vol. 35, no. 2, pp. 63–73, 2010.

[4] B. Mazumder, V. Seshadri, and P. L. Fox, "Translational control by the 3′-UTR: the ends specify the means," *Trends in Biochemical Sciences*, vol. 28, no. 2, pp. 91–98, 2003.

[5] R. Renkawitz-Pohl, L. Hempel, M. Hollman, and M. A. Schafer, "Spermatogenesis," in *Comprehensive Molecular Insect Science*, L. I. Gilbert, K. Iatrou, and S. S. Gill, Eds., Elsevier Pergamon, 2005.

[6] J. D. Richter and P. Lasko, "Translational control in oocyte development," *Cold Spring Harbor Perspectives in Biology*, vol. 3, no. 9, Article ID a002758, 2011.

[7] P. Lasko, "Translational control during early development," in *Progress in Molecular Biology and Translational Science*, J. W. B. Hershey, Ed., Academic Press, Burlington, Vt, USA, 2009.

[8] N. Sonenberg and A. G. Hinnebusch, "New modes of translation control in development, behavior, and disease," *Molecular Cell*, vol. 28, no. 5, pp. 721–729, 2007.

[9] B. Thompson, M. Wickens, and J. Kimble, "Translational control in development," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[10] M. Costa-Mattioli, W. S. Sossin, E. Klann, and N. Sonenberg, "Translational control of long-lasting synaptic plasticity and memory," *Neuron*, vol. 61, no. 1, pp. 10–26, 2009.

[11] B. Schwanhüusser, D. Busse, N. Li et al., "Global quantification of mammalian gene expression control," *Nature*, vol. 473, no. 7347, pp. 337–342, 2011.

[12] C. Vogel, R. De Sousa Abreu, D. Ko et al., "Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line," *Molecular Systems Biology*, vol. 6, article no. 400, 2010.

[13] M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., *Translational Bontrol in Biology and Medicine*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[14] N. Sonenberg, J. W. B. Hershey, and M. B. Mathews, Eds., *Translational Control of Gene Expression*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.

[15] J. W. B. Hershey and W. C. Merrick, "Pathway and mechanism of initiation of protein synthesis," in *Translational Control of Gene Expression*, N. Sonenberg, J. W. B. Hershey, and M.B. Mathews, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2000.

[16] L. D. Kapp and J. R. Lorsch, "The molecular mechanics of eukaryotic translation," *Annual Review of Biochemistry*, vol. 73, pp. 657–704, 2004.

[17] N. Sonenberg and A. G. Hinnebusch, "Regulation of translation initiation in eukaryotes: mechanisms and biological targets," *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[18] R. J. Jackson, C. U. T. Hellen, and T. V. Pestova, "The mechanism of eukaryotic translation initiation and principles of its regulation," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 2, pp. 113–127, 2010.

[19] T. Preiss and M. W. Hentze, "Starting the protein synthesis machine: eukaryotic translation initiation," *BioEssays*, vol. 25, no. 12, pp. 1201–1211, 2003.

[20] F. Gebauer and M. W. Hentze, "Molecular mechanisms of translational control," *Nature Reviews Molecular Cell Biology*, vol. 5, no. 10, pp. 827–835, 2004.

[21] J. A. Dounda and P. Sarnow, "Translation initiation by viral internal ribosome entry sites," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[22] G. Hernández, "Was the initiation of translation in early eukaryotes IRES-driven?" *Trends in Biochemical Sciences*, vol. 33, no. 2, pp. 58–64, 2008.

[23] A. Pacheco and E. Martinez-Salas, "Insights into the biology of IRES elements through riboproteomic approaches.," *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 458927, 12 pages, 2010.

[24] D. J. Taylor, J. Frank, and T. G. Kinzy, "Structure and function of the eukaryotic ribosome and elongation factors," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[25] G. R. Andersen, P. Nissen, and J. Nyborg, "Elongation factors in protein biosynthesis," *Trends in Biochemical Sciences*, vol. 28, no. 8, pp. 434–441, 2003.

[26] T. P. Herbert and C. G. Proud, "Regulation of translation elongation and the cotranslational protein targeting pathway," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[27] M. Ehrenberg, V. Hauryliuk, C. G. Crist, and Y. Nakamura, "Translation termination, the prion [PSI+], and ribosomal recycling," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[28] R. J. Jackson, C. U. T. Hellen, and T. V. Pestova, "Termination and post-termination events in eukaryotic translation," *Advances in Protein Chemistry and Structural Biology*, vol. 86, pp. 45–93, 2012.

[29] M. Szymański and J. Barciszewski, "The genetic code—40 years on," *Acta Biochimica Polonica*, vol. 54, no. 1, pp. 51–54, 2007.

[30] R. D. Knight, S. J. Freeland, and L. F. Landweber, "Rewiring the keyboard: evolvability of the genetic code," *Nature Reviews Genetics*, vol. 2, no. 1, pp. 49–58, 2001.

[31] A. V. Lobanov, A. A. Turanov, D. L. Hatfield, and V. N. Gladyshev, "Dual functions of codons in the genetic code," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 45, no. 4, pp. 257–265, 2010.

[32] E. V. Koonin and A. S. Novozhilov, "Origin and evolution of the genetic code: the universal enigma," *IUBMB Life*, vol. 61, no. 2, pp. 99–111, 2009.

[33] M. A. S. Santos, G. Moura, S. E. Massey, and M. F. Tuite, "Driving change: the evolution of alternative genetic codes," *Trends in Genetics*, vol. 20, no. 2, pp. 95–102, 2004.

[34] C. A. Lozupone, R. D. Knight, and L. F. Landweber, "The molecular basis of nuclear genetic code change in ciliates," *Current Biology*, vol. 11, no. 2, pp. 65–74, 2001.

[35] B. Eliseev, P. Kryuchkova, E. Alkalaeva, and L. Frolova, "A single amino acid change of translation termination factor eRF1 switches between bipotent and omnipotent stop-codon specificity," *Nucleic Acids Research*, vol. 39, no. 2, pp. 599–608, 2011.

[36] S. Lekomtsev, P. Kolosov, L. Bidou, L. Frolova, J. P. Rousset, and L. Kisselev, "Different modes of stop codon restriction by the Stylonychia and Paramecium eRF1 translation termination factors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 26, pp. 10824–10829, 2007.

[37] Y. Inagaki, C. Blouin, W. F. Doolittle, and A. J. Roger, "Convergence and constraint in eukaryotic release factor 1 (eRF1) domain 1: the evolution of stop codon specificity," *Nucleic Acids Research*, vol. 30, no. 2, pp. 532–544, 2002.

[38] A. Seit-Nebi, L. Frolova, and L. Kisselev, "Conversion of omnipotent translation termination factor eRF1 into ciliate-like UGA-only unipotent eRF1," *EMBO Reports*, vol. 3, no. 9, pp. 881–886, 2002.

[39] E. Cocquyt, G. H. Gile, F. Leliaert, H. Verbruggen, P. J. Keeling, and O. De Clerck, "Complex phylogenetic distribution of a non-canonical genetic code in green algae," *BMC Evolutionary Biology*, vol. 10, no. 1, article no. 327, 2010.

[40] G. Hernández, "On the origin of the cap-dependent initiation of translation in eukaryotes," *Trends in Biochemical Sciences*, vol. 34, no. 4, pp. 166–175, 2009.

[41] N. C. Kyrpides and C. R. Woese, "Universally conserved translation initiation factors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 1, pp. 224–228, 1998.

[42] P. Londei, "Evolution of translational initiation: new insights from the archaea," *FEMS Microbiology Reviews*, vol. 29, no. 2, pp. 185–200, 2005.

[43] L. Aravind and E. V. Koonin, "Eukaryote-specific domains in translation initiation factors: implications for translation regulation and evolution of the translation system," *Genome Research*, vol. 10, no. 8, pp. 1172–1184, 2000.

[44] D. Benelli and P. Londei, "Begin at the beginning: evolution of translational initiation," *Research in Microbiology*, vol. 160, no. 7, pp. 493–501, 2009.

[45] G. Hernández and P. Vazquez-Pianzola, "Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families," *Mechanisms of Development*, vol. 122, no. 7-8, pp. 865–876, 2005.

[46] B. Joshi, K. Lee, D. L. Maeder, and R. Jagus, "Phylogenetic analysis of eIF4E-family members," *BMC Evolutionary Biology*, vol. 5, article no. 48, 2005.

[47] C. C. Baker and M. T. Fuller, "Translational control of meiotic cell cycle progression and spermatid differentiation in male germ cells by a novel eIF4G homolog," *Development*, vol. 134, no. 15, pp. 2863–2869, 2007.

[48] T. M. Franklin-Dumont, C. Chatterjee, S. A. Wasserman, and S. DiNardo, "A novel eIF4G homolog, off-schedule, couples translational control to meiosis and differentiation in Drosophila spermatocytes," *Development*, vol. 134, no. 15, pp. 2851–2861, 2007.

[49] E. R. Freire, R. Dhalia, D. M. N. Moura et al., "The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties," *Molecular and Biochemical Parasitology*, vol. 176, no. 1, pp. 25–36, 2011.

[50] B. D. Keiper, B. J. Lamphear, A. M. Deshpande et al., "Functional characterization of five eIF4E isoforms in Caenorhabditis elegans," *Journal of Biological Chemistry*, vol. 275, no. 14, pp. 10590–10596, 2000.

[51] H. Miyoshi, D. S. Dwyer, B. D. Keiper, M. Jankowska-Anyszka, E. Darzynkiewicz, and R. E. Rhoads, "Discrimination between mono- and trimethylated cap structures by two isoforms of Caenorhabditis elegans eIF4E," *EMBO Journal*, vol. 21, no. 17, pp. 4680–4690, 2002.

[52] P. Lasko, "The Drosophila melanogaster genome: translation factors and RNA binding proteins," *Journal of Cell Biology*, vol. 150, no. 2, pp. F51–F56, 2000.

[53] J. Morales, O. Mulner-Lorillon, B. Cosson et al., "Translational control genes in the sea urchin genome," *Developmental Biology*, vol. 300, no. 1, pp. 293–307, 2006.

[54] M. Jankowska-Anyszka, B. J. Lamphear, E. J. Aamodt et al., "Multiple isoforms of eukaryotic protein synthesis initiation factor 4E in Caenorhabditis elegans can distinguish between mono- and trimethylated mRNA cap structures," *Journal of Biological Chemistry*, vol. 273, no. 17, pp. 10538–10541, 1998.

[55] K. S. Browning, "Plant translation initiation factors: it is not easy to be green," *Biochemical Society Transactions*, vol. 32, no. 4, pp. 589–591, 2004.

[56] L. K. Mayberry, M. Leah Allen, M. D. Dennis, and K. S. Browning, "Evidence for variation in the optimal translation initiation complex: plant eIF4B, eIF4F, and eIF(iso)4F differentially promote translation of mRNAs," *Plant Physiology*, vol. 150, no. 4, pp. 1844–1854, 2009.

[57] J. Robalino, B. Joshi, S. C. Fahrenkrug, and R. Jagus, "Two zebrafish eIF4E family members are differentially expressed

and functionally divergent," *Journal of Biological Chemistry*, vol. 279, no. 11, pp. 10532–10541, 2004.

[58] G. Hernández, M. Altmann, J. M. Sierra et al., "Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in Drosophila," *Mechanisms of Development*, vol. 122, no. 4, pp. 529–543, 2005.

[59] A. V. Evsikov and C. Marín de Evsikova, "Evolutionary origin and phylogenetic analysis of the novel oocyte-specific eukaryotic translation initiation factor 4E in Tetrapoda," *Development Genes and Evolution*, vol. 219, no. 2, pp. 111–118, 2009.

[60] C. Goyer, M. Altmann, H. S. Lee et al., "TIF4631 and TIF4632: two yeast genes encoding the high-molecular-weight subunits of the cap-binding protein complex (eukaryotic initiation factor 4F) contain an RNA recognition motif-like sequence and carry out an essential function," *Molecular and Cellular Biology*, vol. 13, no. 8, pp. 4860–4874, 1993.

[61] M. Wakiyama, A. Suzuki, M. Saigoh et al., "Analysis of the isoform of Xenopus euakryotic translation initiation factor 4E," *Bioscience, Biotechnology and Biochemistry*, vol. 65, no. 1, pp. 232–235, 2001.

[62] M. Ptushkina, K. Berthelot, T. Von der Haar, L. Geffers, J. Warwicker, and J. E. G. McCarthy, "A second elF4E protein in Schizosaccharomyces pombe has distinct elF4G-binding properties," *Nucleic Acids Research*, vol. 29, no. 22, pp. 4561–4569, 2001.

[63] B. Joshi, A. Cameron, and R. Jagus, "Characterization of mammalian eIF4E-family members," *European Journal of Biochemistry*, vol. 271, no. 11, pp. 2189–2203, 2004.

[64] L. Li and C. C. Wang, "Identification in the ancient protist Giardia lamblia of two eukaryotic translation initiation factor 4E homologues with distinctive functions," *Eukaryotic Cell*, vol. 4, no. 5, pp. 948–959, 2005.

[65] Y. Yoffe, J. Zuberek, A. Lerer et al., "Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania," *Eukaryotic Cell*, vol. 5, no. 12, pp. 1969–1979, 2006.

[66] R. Dhalia, C. R. S. Reis, E. R. Freire et al., "Translation initiation in Leishmania major: characterisation of multiple eIF4F subunit homologues," *Molecular and Biochemical Parasitology*, vol. 140, no. 1, pp. 23–41, 2005.

[67] D. R. Gallie, "Translational control in plants and chloroplasts," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[68] G. Hernández, P. Vázques-Pianzola, A. Zurbriggen, M. Altmann, J. M. Sierra, and R. Rivera-Pomar, "Two functionally redundant isoforms of Drosophila melanogaster eukaryotic initiation factor 4B are involved in cap-dependent translation, cell survival, and proliferation," *European Journal of Biochemistry*, vol. 271, no. 14, pp. 2923–2936, 2004.

[69] G. Hernández, M. Del Mar Castellano, M. Agudo, and J. M. Sierra, "Isolation and characterization of the cDNA and the gene for eukaryotic translation initiation factor 4G from Drosophila melanogaster," *European Journal of Biochemistry*, vol. 253, no. 1, pp. 27–35, 1998.

[70] G. W. Owttrim, T. Mandel, H. Trachsel, A. A. Thomas, and C. Kuhlemeier, "Characterization of the tobacco eIF-4A gene family," *Plant Molecular Biology*, vol. 26, no. 6, pp. 1747–1757, 1994.

[71] A. Zinoviev, M. Leger, G. Wagner, and M. Shapira, "A novel 4E-interacting protein in Leishmania is involved in stage-specific translation pathways," *Nucleic Acids Research*, vol. 39, no. 19, pp. 8404–8415, 2011.

[72] Y. Yoffe, M. Léger, A. Zinoviev et al., "Evolutionary changes in the Leishmania eIF4F complex involve variations in the eIF4E-eIF4G interactions," *Nucleic Acids Research*, vol. 37, no. 10, pp. 3243–3253, 2009.

[73] P. F. Cho, F. Poulin, Y. A. Cho-Park et al., "A new paradigm for translational control: inhibition via 5′-3′ mRNA tethering by Bicoid and the eIF4E cognate 4EHP," *Cell*, vol. 121, no. 3, pp. 411–423, 2005.

[74] T. D. Dinkova, B. D. Keiper, N. L. Korneeva, E. J. Aamodt, and R. E. Rhoads, "Translation of a small subset of Caenorhabditis elegans mRNAs is dependent on a specific eukaryotic translation initiation factor 4E isoform," *Molecular and Cellular Biology*, vol. 25, no. 1, pp. 100–113, 2005.

[75] R. E. Rhoads, "EIF4E: new family members, new binding partners, new roles," *Journal of Biological Chemistry*, vol. 284, no. 25, pp. 16711–16715, 2009.

[76] K. A. Ruud, C. Kuhlow, D. J. Goss, and K. S. Browning, "Identification and characterization of a novel cap-binding protein from Arabidopsis thaliana," *Journal of Biological Chemistry*, vol. 273, no. 17, pp. 10325–10330, 1998.

[77] E. Rom, H. C. Kim, A. C. Gingras et al., "Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein," *Journal of Biological Chemistry*, vol. 273, no. 21, pp. 13104–13109, 1998.

[78] J. C. Villaescusa, C. Buratti, D. Penkov et al., "Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation," *PLoS One*, vol. 4, no. 4, Article ID e5213, 2009.

[79] N. Minshall, M. H. Reiter, D. Weil, and N. Standart, "CPEB interacts with an ovary-specific eIF4E and 4E-T in early Xenopus oocytes," *Journal of Biological Chemistry*, vol. 282, no. 52, pp. 37389–37401, 2007.

[80] S. Lall, C. C. Friedman, M. Jankowska-Anyszka, J. Stepinski, E. Darzynkiewicz, and R. E. Davis, "Contribution of trans-splicing, 5′-leader length, cap-poly(A) synergism, and initiation factors to nematode translation in an Ascaris suum embryo cell-free system," *Journal of Biological Chemistry*, vol. 279, no. 44, pp. 45573–45585, 2004.

[81] M. Altmann, C. Handschin, and H. Trachsel, "mRNA cap-binding protein: cloning of the gene encoding protein synthesis initiation factor eIF-4E from Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 7, no. 3, pp. 998–1003, 1987.

[82] F. G. Maroto and J. M. Sierra, "Purification and characterization of mRNA cap-binding protein from Drosophila melanogaster embryos," *Molecular and Cellular Biology*, vol. 9, no. 5, pp. 2181–2190, 1989.

[83] C. M. Rodriguez, M. A. Freire, C. Camilleri, and C. Robaglia, "The Arabidosis thaliana cDNAs encoding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development," *Plant Journal*, vol. 13, no. 4, pp. 465–473, 1998.

[84] N. Sonenberg, "eIF4E, the mRNA cap-binding protein: from basic discovery to translational research," *Biochemistry and Cell Biology*, vol. 86, no. 2, pp. 178–183, 2008.

[85] L. Rong, M. Livingstone, R. Sukarieh et al., "Control of eIF4E cellular localization by eIF4E-binding proteins, 4E-BPs," *RNA*, vol. 14, no. 7, pp. 1318–1327, 2008.

[86] P. L. Graham, J. L. Yanowitz, J. K. M. Penn, G. Deshpande, and P. Schedl, "The translation initiation factor eif4e regulates the Sex-Specific expression of the master switch gene

Sxl in Drosophila melanogaster," *PLoS Genetics*, vol. 7, no. 7, Article ID e1002185, 2011.

[87] A. Amiri, B. D. Keiper, I. Kawasaki et al., "An isoform of eIF4E is a component of germ granules and is required for spermatogenesis in *C. elegans*," *Development*, vol. 128, no. 20, pp. 3899–3912, 2001.

[88] M. A. Henderson, E. Croniand, S. Dunkelbarger, V. Contreras, S. Strome, and B. D. Keiper, "A germline-specific isoform of eIF4E (IFE-1) is required for efficient translation of stored mRNAs and maturation of both oocytes and sperm," *Journal of Cell Science*, vol. 122, no. 10, pp. 1529–1539, 2009.

[89] I. Kawasaki, M. H. Jeong, and Y. H. Shim, "Regulation of sperm-specific proteins by IFE-1, a germline-specific homolog of eIF4E, in C. elegans," *Molecules and Cells*, vol. 31, no. 2, pp. 191–197, 2011.

[90] A. Song, S. Labella, N. L. Korneeva et al., "A *C. elegans* eIF4E-family member upregulates translation at elevated temperatures of mRNAs encoding MSH-5 and other meiotic crossover proteins," *Journal of Cell Science*, vol. 123, no. 13, pp. 2228–2237, 2010.

[91] G. Hernández, V. Gandin, H. Han, T. Ferreira, N. Sonenberg, and P. Lasko, "Translational control by Drosophila eIF4E-3 is essential for cell differentiation during spermiogenesis," *Development*. In press.

[92] J. M. Zapata, M. A. Martinez, and J. M. Sierra, "Purification and characterization of eukaryotic polypeptide chain initiation factor 4F from Drosophila melanogaster embryos," *Journal of Biological Chemistry*, vol. 269, no. 27, pp. 18047–18052, 1994.

[93] V. Contreras, M. A. Richardson, E. Hao, and B. D. Keiper, "Depletion of the cap-associated isoform of translation factor eIF4G induces germline apoptosis in *C. elegans*," *Cell Death and Differentiation*, vol. 15, no. 8, pp. 1232–1242, 2008.

[94] D. Prévôt, J. L. Darlix, and T. Ohlmann, "Conducting the initiation of protein synthesis: the role of eIF4G," *Biology of the Cell*, vol. 95, no. 3-4, pp. 141–156, 2003.

[95] T. V. Pestova, I. N. Shatsky, and C. U. T. Hellen, "Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes," *Molecular and Cellular Biology*, vol. 16, no. 12, pp. 6870–6878, 1996.

[96] A. Gradi, H. Imataka, Y. V. Svitkin et al., "A novel functional human eukaryotic translation initiation factor 4G," *Molecular and Cellular Biology*, vol. 18, no. 1, pp. 334–342, 1998.

[97] S. Caron, M. Charon, E. Cramer, N. Sonenberg, and I. Dusanter-Fourt, "Selective modification of eukaryotic initiation factor 4F (eIF4F) at the onset of cell differentiation: recruitment of eIF4GII and long-lasting phosphorylation of eIF4E," *Molecular and Cellular Biology*, vol. 24, no. 11, pp. 4920–4928, 2004.

[98] F. Sun, K. Palmer, and M. A. Handel, "Mutation of Eif4g3, encoding a eukaryotic translation initiation factor, causes male infertility and meiotic arrest of mouse spermatocytes," *Development*, vol. 137, no. 10, pp. 1699–1707, 2010.

[99] A. G. Hinnebusch, "eIF3: a versatile scaffold for translation initiation complexes," *Trends in Biochemical Sciences*, vol. 31, no. 10, pp. 553–562, 2006.

[100] J. Guo, Z. Jin, X. Yang, J. F. Li, and J. G. Chen, "Eukaryotic initiation factor 6, an evolutionarily conserved regulator of ribosome biogenesis and protein translation," *Plant Signaling and Behavior*, vol. 6, no. 5, pp. 766–771, 2011.

[101] A. L. Stevenson and J. E. G. McCarthy, "Found in translation: another RNA helicase function," *Molecular Cell*, vol. 32, no. 6, pp. 755–756, 2008.

[102] A. Parsyan, Y. Svitkin, D. Shahbazian et al., "MRNA helicases: the tacticians of translational control," *Nature Reviews Molecular Cell Biology*, vol. 12, no. 4, pp. 235–245, 2011.

[103] P. Linder and E. Jankowsky, "From unwinding to clamping—the DEAD box RNA helicase family," *Nature Reviews Molecular Cell Biology*, vol. 12, no. 8, pp. 505–516, 2011.

[104] A. Weston and J. Sommerville, "Xp54 and related (DDX6-like) RNA helicases: roles in messenger RNP assembly, translation regulation and RNA degradation," *Nucleic Acids Research*, vol. 34, no. 10, pp. 3082–3094, 2006.

[105] M. S. Bush, A. P. Hutchins, A. M. E. Jones et al., "Selective recruitment of proteins to 5′ cap complexes during the growth cycle in Arabidopsis," *Plant Journal*, vol. 59, no. 3, pp. 400–412, 2009.

[106] J. D. Richter and N. Sonenberg, "Regulation of cap-dependent translation by eIF4E inhibitory proteins," *Nature*, vol. 433, no. 7025, pp. 477–480, 2005.

[107] C. G. Proud, "Signalling to translation: how signal transduction pathways control the protein synthetic machinery," *Biochemical Journal*, vol. 403, no. 2, pp. 217–234, 2007.

[108] B. Raught and A. C. Gingras, "Signaling to translation initiation," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[109] L. Furic, L. Rong, O. Larsson et al., "EIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 32, pp. 14134–14139, 2010.

[110] G. C. Scheper and C. G. Proud, "Does phosphorylation of the cap-binding protein eIF4E play a role in translation initiation?" *European Journal of Biochemistry*, vol. 269, no. 22, pp. 5350–5359, 2002.

[111] M. Buxade, J. L. Parra-Palau, and C. G. Proud, "The Mnks: MAP kinase-interacting kinases (MAP kinase signal-integrating kinases).," *Frontiers in Bioscience*, vol. 13, pp. 5359–5373, 2008.

[112] J. H. Reiling, K. T. Doepfner, E. Hafen, and H. Stocker, "Diet-dependent effects of the Drosophila Mnk1/Mnk2 homolog Lk6 on growth via eIF4E," *Current Biology*, vol. 15, no. 1, pp. 24–30, 2005.

[113] P. E. D. Lachance, M. Miron, B. Raught, N. Sonenberg, and P. Lasko, "Phosphorylation of eukaryotic translation initiation factor 4E is critical for growth," *Molecular and Cellular Biology*, vol. 22, no. 6, pp. 1656–1663, 2002.

[114] N. Arquier, M. Bourouis, J. Colombani, and P. Léopold, "Drosophila Lk6 kinase controls phosphorylation of eukaryotic translation initiation factor 4E and promotes normal growth and development," *Current Biology*, vol. 15, no. 1, pp. 19–23, 2005.

[115] J. L. Parra-Palau, G. C. Scheper, D. E. Harper, and C. G. Proud, "The Drosophila protein kinase LK6 is regulated by ERK and phosphorylates the eukaryotic initiation factor eIF4E in vivo," *Biochemical Journal*, vol. 385, no. 3, pp. 695–702, 2005.

[116] I. R. E. Nett, D. M. A. Martin, D. Miranda-Saavedra et al., "The phosphoproteome of bloodstream form Trypanosoma brucei, causative agent of African sleeping sickness," *Molecular and Cellular Proteomics*, vol. 8, no. 7, pp. 1527–1538, 2009.

[117] N. I. T. Zanchin and J. E. G. McCarthy, "Characterization of the in vivo phosphorylation sites of the mRNA·Cap- binding complex proteins eukaryotic initiation factor-4E and p20 in Saccharomyces cerevisiae," *Journal of Biological Chemistry*, vol. 270, no. 44, pp. 26505–26510, 1995.

[118] T. E. Dever, A. C. Dar, and F. Sicheri, "The eIF2a kinases," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[119] S. Rothenburg, N. Deigendesch, M. Dey, T. E. Dever, and L. Tazi, "Double-stranded RNA-activated protein kinase PKR of fishes and amphibians: varying the number of double-stranded RNA binding domains and lineage-specific duplications.," *BMC biology*, vol. 6, article 12, 2008.

[120] M. Szymański, M. Deniziak, and J. Barciszewski, "The new aspects of aminoacyl-tRNA synthetases," *Acta Biochimica Polonica*, vol. 47, no. 3, pp. 821–834, 2000.

[121] M. Guo, X. L. Yang, and P. Schimmel, "New functions of aminoacyl-tRNA synthetases beyond translation," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 9, pp. 668–674, 2010.

[122] P. S. Ray, J. C. Sullivan, J. Jia, J. Francis, J. R. Finnerty, and P. L. Fox, "Evolution of function of a fused metazoan tRNA synthetase," *Molecular Biology and Evolution*, vol. 28, no. 1, pp. 437–447, 2011.

[123] P. L. Fox, P. S. Ray, A. Arif, and J. Jia, "Noncanonical functions of aminoacyl-tRNA synthetases in translational control," in *Translational Control in Biology and Medicine*, M. B. Mathews, N. Sonenberg, and J. W. B. Hershey, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2007.

[124] P. J. Keeling and Y. Inagaki, "A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1α," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 43, pp. 15380–15385, 2004.

[125] M. Sakaguchi, K. Takishita, T. Matsumoto, T. Hashimoto, and Y. Inagaki, "Tracing back EFL gene evolution in the cryptomonads-haptophytes assemblage: separate origins of EFL genes in haptophytes, photosynthetic cryptomonads, and goniomonads," *Gene*, vol. 441, no. 1-2, pp. 126–131, 2009.

[126] G. H. Gile, P. M. Novis, D. S. Cragg, G. C. Zuccarello, and P. J. Keeling, "The distribution of elongation factor-1 alpha (EF-1α), elongation factor-like (EFL), and a non-canonical genetic code in the ulvophyceae: Discrete genetic characters support a consistent phylogenetic framework," *Journal of Eukaryotic Microbiology*, vol. 56, no. 4, pp. 367–372, 2009.

[127] E. Cocquyt, H. Verbruggen, F. Leliaert, F. W. Zechman, K. Sabbe, and O. De Clerck, "Gain and loss of elongation factor genes in green algae," *BMC Evolutionary Biology*, vol. 9, no. 1, article no. 39, 2009.

[128] G. P. Noble, M. B. Rogers, and P. J. Keeling, "Complex distribution of EFL and EF-1α proteins in the green algal lineage," *BMC Evolutionary Biology*, vol. 7, article no. 82, 2007.

[129] R. Kamikawa, Y. Inagaki, and Y. Sako, "Direct phylogenetic evidence for lateral transfer of elongation factor-like gene," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 19, pp. 6965–6969, 2008.

[130] R. Kamikawa, A. Yabuki, T. Nakayama, K. I. Ishida, T. Hashimoto, and Y. Inagaki, "Cercozoa comprises both EF-1α-containing and EFL-containing members," *European Journal of Protistology*, vol. 47, no. 1, pp. 24–28, 2011.

[131] R. Kamikawa, M. Sakaguchi, T. Matsumoto, T. Hashimoto, and Y. Inagaki, "Rooting for the root of elongation factor-like protein phylogeny," *Molecular Phylogenetics and Evolution*, vol. 56, no. 3, pp. 1082–1088, 2010.

[132] G. H. Gile, D. Faktorová, C. A. Castlejohn et al., "Distribution and phylogeny of EFL and EF1alpha in Euglenozoa suggest ancestral co-occurrence followed by differential loss," *PLoS One*, vol. 4, no. 4, Article ID e5162, 2009.

[133] M. Djernaes and J. Damgaard, "Exon-intron structure, paralogy and sequenced regions of elongation factor-1 alpha in Hexapoda," *Arthropod Systematics and Phylogeny*, vol. 64, no. 1, pp. 45–52, 2006.

[134] E. Greganova, M. Altmann, and P. Bütikofer, "Unique modifications of translation elongation factors," *FEBS Journal*, vol. 278, no. 15, pp. 2613–2624, 2011.

[135] O. T. P. Kim, K. Yura, N. Go, and T. Harumoto, "Newly sequenced eRF1s from ciliates: the diversity of stop codon usage and the molecular surfaces that are important for stop codon interactions," *Gene*, vol. 346, pp. 277–286, 2005.

[136] D. Moreira, S. Kervestin, O. Jean-Jean, and H. Philippe, "Evolution of eukaryotic translation elongation and termination factors: variations of evolutionary rate and genetic code deviations," *Molecular Biology and Evolution*, vol. 19, no. 2, pp. 189–200, 2002.

[137] G. C. Atkinson, S. L. Baldauf, and V. Hauryliuk, "Evolution of nonstop, no-go and nonsense-mediated mRNA decay and their termination factor-derived components," *BMC Evolutionary Biology*, vol. 8, no. 1, article no. 290, 2008.

[138] Y. Inagaki and W. F. Doolittle, "Class I release factors in ciliates with variant genetic codes," *Nucleic Acids Research*, vol. 29, no. 4, pp. 921–927, 2001.

[139] H. Song, P. Mugnier, A. K. Das et al., "The crystal structure of human eukaryotic release factor eRF1—mechanism of stop codon recognition and peptidyl-tRNA hydrolysis," *Cell*, vol. 100, no. 3, pp. 311–321, 2000.

[140] P. Kolosov, L. Frolova, A. Seit-Nebi et al., "Invariant amino acids essential for decoding function of polypeptide release factor eRF1," *Nucleic Acids Research*, vol. 33, no. 19, pp. 6418–6425, 2005.

[141] K. Ito, L. Frolova, A. Seit-Nebi, A. Karamyshev, L. Kisselev, and Y. Nakamura, "Omnipotent decoding potential resides in eukaryotic translation termination factor eRF1 of variant-code organisms and is modulated by the interactions of amino acid sequences within domain 1," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 13, pp. 8494–8499, 2002.

[142] Y. Inagaki and W. F. Doolittle, "Evolution of the eukaryotic translation termination system: origins of release factors," *Molecular Biology and Evolution*, vol. 17, no. 6, pp. 882–889, 2000.

[143] G. Zhouravleva, V. Schepachev, A. Petrova, O. Tarasov, and S. Inge-Vechtomov, "Evolution of translation termination factor eRF3: is GSPT2 generated by retrotransposition of GSPT1's mRNA?" *IUBMB Life*, vol. 58, no. 4, pp. 199–202, 2006.

[144] S. I. Hoshino, M. Imai, M. Mizutani et al., "Molecular cloning of a novel member of the eukaryotic polypeptide chain-releasing factors (eRF): its identification as eRF3 interacting with eRF1," *Journal of Biological Chemistry*, vol. 273, no. 35, pp. 22254–22259, 1998.

[145] C. L. Goff, O. Zemlyanko, S. Moskalenko et al., "Mouse GSPT2, but not GSPT1, can substitute for yeast eRF3 in vivo," *Genes to Cells*, vol. 7, no. 10, pp. 1043–1057, 2002.

[146] L. Skogerson and E. Wakatama, "A ribosome dependent GTPase from yeast distinct from elongation factor 2," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 73, no. 1, pp. 73–76, 1976.

[147] M. F. Ypma-Wong, W. A. Fonzi, and P. S. Sypherd, "Fungus-specific translation elongation factor 3 gene present in *Pneumocystis carinii*," *Infection and Immunity*, vol. 60, no. 10, pp. 4140–4145, 1992.

[148] S. Qin, A. Xie, M. C. M. Bonato, and C. S. McLaughlin, "Sequence analysis of the translational elongation factor 3 from Saccharomyces cerevisiae," *Journal of Biological Chemistry*, vol. 265, no. 4, pp. 1903–1912, 1990.

[149] B. J. Di Domenico, J. Lupisella, M. Sandbaken, and K. Chakraburtty, "Isolation and sequence analysis of the gene encoding translation elongation factor 3 from Candida albicans," *Yeast*, vol. 8, no. 5, pp. 337–352, 1992.

[150] L. Skogerson, "Separation and characterization of yeast elongation factors," *Methods in Enzymology*, vol. 60, no. C, pp. 676–685, 1979.

[151] K. Chakraburtty and F. J. Triana-Alonso, "Yeast elongation factor 3: structure and function," *Biological Chemistry*, vol. 379, no. 7, pp. 831–840, 1998.

[152] S. Kurata, K. H. Nielsen, S. F. Mitchell, J. R. Lorsch, A. Kaji, and H. Kaji, "Ribosome recycling step in yeast cytoplasmic protein synthesis is catalyzed by eEF3 and ATP," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 24, pp. 10854–10859, 2010.

[153] F. A. Bisby, Y. R. Roskov, T. M. Orrell, D. Nicolson, L. E. Paglinawan et al., "Species 2000 & ITIS Catalogue of Life: 2010 Annual Checklist," 2010, http://www.catalogueoflife.org/annual-checklist/2010.

[154] C. Mora, D. P. Titterson, S. Adl, A. G. B. Simpson, and B. Worm, "How many species are there on Earth and in the ocean," *PLoS Biology*, vol. 9, no. 8, Article ID e1001127, 2011.

[155] S. G. Tringe, C. Von Mering, A. Kobayashi et al., "Comparative metagenomics of microbial communities," *Science*, vol. 308, no. 5721, pp. 554–557, 2005.

[156] A. C. McHardy and I. Rigoutsos, "What's in the mix: phylogenetic classification of metagenome sequence samples," *Current Opinion in Microbiology*, vol. 10, no. 5, pp. 499–503, 2007.

[157] J. C. Venter, K. Remington, J. F. Heidelberg et al., "Environmental genome shotgun sequencing of the Sargasso Sea," *Science*, vol. 304, no. 5667, pp. 66–74, 2004.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 576540, 10 pages
doi:10.1155/2012/576540

Review Article

The Role of Translation Initiation Regulation in Haematopoiesis

Godfrey Grech[1] and Marieke von Lindern[2]

[1] *Department of Pathology, Medical School, University of Malta, Msida MSD 2090, Malta*
[2] *Department of Hematopoiesis, Sanquin Research and Landsteiner Laboratory, AMC/UvA, 1066 CX Amsterdam, The Netherlands*

Correspondence should be addressed to Godfrey Grech, godfrey.grech@um.edu.mt

Received 20 January 2012; Accepted 25 February 2012

Academic Editor: Greco Hernández

Organisation of RNAs into functional subgroups that are translated in response to extrinsic and intrinsic factors underlines a relatively unexplored gene expression modulation that drives cell fate in the same manner as regulation of the transcriptome by transcription factors. Recent studies on the molecular mechanisms of inflammatory responses and haematological disorders indicate clearly that the regulation of mRNA translation at the level of translation initiation, mRNA stability, and protein isoform synthesis is implicated in the tight regulation of gene expression. This paper outlines how these posttranscriptional control mechanisms, including control at the level of translation initiation factors and the role of RNA binding proteins, affect hematopoiesis. The clinical relevance of these mechanisms in haematological disorders indicates clearly the potential therapeutic implications and the need of molecular tools that allow measurement at the level of translational control. Although the importance of miRNAs in translation control is well recognised and studied extensively, this paper will exclude detailed account of this level of control.

1. Introduction

Hematopoietic stem cells (HSCs) have a life-long capacity to replenish the stem cell compartment and give rise to multipotent progenitors. These progenitors expand to maintain the hematopoietic compartment and differentiate into various blood lineage progenitors. Lineage positive progenitors are committed for differentiation into mature blood cells.

Transcription factors have a pivotal role in hematopoiesis to maintain a gene expression program that endows self-renewal properties to HSCs and enables commitment and differentiation into different blood cell lineages [1]. The upregulation of both PU.1 and Gata-1 reprograms HSC to become common myeloid progenitors (CMPs) [2]. The CMPs undergo further lineage divergence into megakaryocyte/erythroid progenitors (MEPs) and granulocyte/monocyte progenitors (GMPs) upon Gata-1 and PU.1 mutual exclusive expression, respectively. Commitment to the erythroid lineage is characterized by the expression of erythroid-specific transcription factors Gata-1, Eklf, and Nfe2 determining the erythroid program.

Upon commitment, the balance between proliferation and differentiation of lineage-specific progenitors is under tight control, to maintain the progenitor pool and ensure maturation in response to physiological demand. The production of increased numbers of mature blood cells during stress situations such as inflammation or hypoxia requires higher progenitor proliferation rates. Concurrently, feedback mechanisms must be closely coordinated to repress progenitor proliferation when the stress is over [3]. The human bone marrow must replace 10^{11} erythrocytes daily under normal physiological erythropoiesis. Gene expression is regulated at the transcription level, producing a cell-specific mRNA pool. Subsequent control of mRNA translation enables the cells to further adapt to environmental and developmental cues. Translation regulation (i) permits fast cellular responses to growth factors, inducing specific proteins to be expressed, (ii) selective expression of different protein isoforms from a given transcript, and (iii) induction of expression of pro-apoptotic proteins when the transcription program is inhibited. Translation Initiation is an important level of translation control. Cap-dependent translation initiation depends on two major

limiting steps: (i) the formation of the initiation complex by release of the cap-binding initiation factor eIF4E from its binding factor 4E-BP and (ii) binding of the ternary complex (TC) consisting of GTP-loaded translation initiation factor 2 (eIF2:GTP) plus a methionine-loaded initiator tRNA ($tRNA_i^{met}$) to the 40S ribosome subunit. Cap-independent translation initiation depends on an internal ribosomal entry site (IRES) in the transcript that has to bind IRES transactivating factors (ITAFs). The biochemistry of translation initiation has been extensively reviewed [4, 5]. We will focus on the importance of translation initiation for haematopoiesis.

2. Growth Factor-Dependent Proliferation of Hematopoietic Progenitors

The main regulator of erythropoiesis is the glycoprotein hormone Erythropoietin (Epo), produced in the kidney in response to oxygen tension in the blood. The function of Epo initiates from the specific interaction to its cell surface receptor (EpoR). In stress erythropoiesis, stem cell factor (cKit ligand) and glucocorticoids (GR) work in concert with Epo to induce expansion of progenitors in the mouse spleen [6, 7]. The requirement for SCF in acute erythroid expansion was demonstrated by the observation that inhibiting c-Kit antibodies abolished splenic hematopoiesis upon induction of haemolytic anaemia in mice, while the antibodies had no effect on steady-state erythropoiesis [6].

Epo and SCF transduce signals via multiple cooperating pathways in erythroid progenitors [7–10], among which the activation of the PI3K pathway. Although both Epo and SCF activate PI3K in erythroid progenitors, the efficiency with which downstream signalling pathways are activated shows large differences [11, 12], suggesting differential susceptibility to feedback pathways. Activation of PI3K results in phosphorylation and activation of PKB and subsequently of mTOR (Figure 1). In turn, mTOR phosphorylates and activates S6K (Rps6kb1; p70S6Kinase) and 4EBP (4E-Binding Protein) [13]. In erythroblasts, only SCF can induce full phosphorylation of 4EBP [14]. PP2A is the main phosphatase acting on S6K and 4EBP1 and thereby the main antagonist of mTOR function in erythroblasts (Figure 1).

4EBP hyperphosphorylation results in the release of the mRNA cap-binding factor eIF4E (eukaryotic Initiation Factor 4E) [15]. Subsequently, eIF4E can bind the scaffold protein eIF4G, which enables the formation of an eIF4F scanning complex containing eIF4E, eIF4G, and the RNA helicase eIF4A. eIF4F associates with several other translation factors including the 40S small subunit of the ribosome and the associated ternary complex consisting of eIF2:$tRNA_i^{met}$ [5]. This preinitiation complex scans the 5′UTR for the first AUG codon in an appropriate sequence context [16]. There the ribosome associates and methionine is deposited at the P-site and eIF2 is released [5, 17]. The cap-binding eIF4E protein is a rate limiting factor in the formation of the preinitiation complex [18] and therefore its release upon phosphorylation of 4EBP is a crucial control mechanism in polysome recruitment of mRNAs. In addition, eIF4E is phosphorylated by MAP-kinase signal-integrating kinases

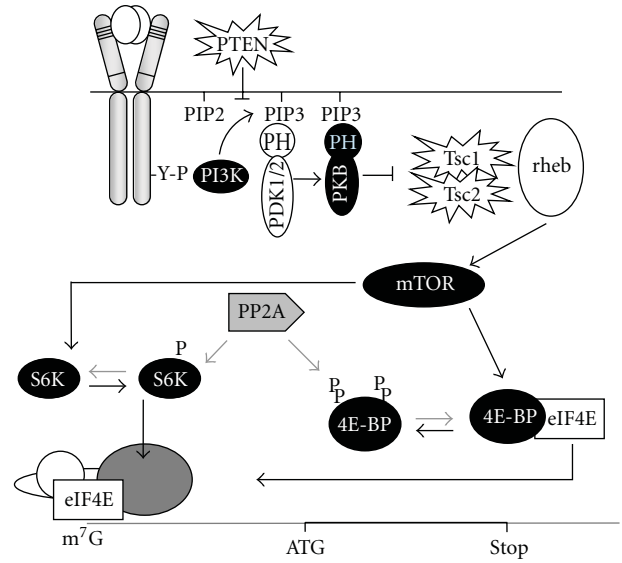


Figure 1: The PI3K/PKB/mTOR pathway controls mRNA translation. SCF-receptor activation results in recruitment of PI3K to the receptor, which generates phosphorylates membrane lipids (PIP3) that form an anchor for the PH-domain containing kinases PDK1 and PKB. PIP3 is dephosphorylated by the tumour suppressor PTEN, which silences the PI3K-pathway. At the membrane PDK1 phosphorylates PKB, which phosphorylates the tuberous sclerosis tumour suppressor genes Tsc1 and Tsc2. Upon phosphorylation these genes release the GTPase Rheb to activate mTOR. Activation of mTOR results in phosphorylation of p70S6kinase (S6K) and eIF4E-binding protein (4E-BP). Upon phosphorylation, 4E-BP releases the cap-binding translation initiation factor 4E (eIF4E), which allows for association of eIF4E with the proteins that form the eIF4F scanning complex and with the 40S ribosomal subunit [4].

Mnk1 and Mnk2 [19, 20] in response to insulin and stress [21]. The role of eIF4E phosphorylation is still controversial.

2.1. Translation Initiation Sensitivity to eIFs: Regulation of eIF4F-Sensitive Transcripts. Although the preinitiation complex consists of general translation factors, it scans mRNAs with a short and simple 5′ Untranslated Region (UTR) much more efficiently than mRNAs with a long and structured 5′UTR [22]. Structured mRNAs require a higher density of preinitiation complexes to maintain an open structure, which renders them more sensitive to the concentration of eIF4F complexes in the cell. Well-known examples are growth promoting proteins such as VEGF, MYC, and ODC [22]. PI3K-dependent, selective polysome recruitment of mRNA is important in v-Ras/v-Akt-transformed glioblastoma cells [23] and in metastasis of human epithelial cells [24]. Transcripts that are specifically recruited to polysomes upon overexpression of eIF4E have been identified in mouse and human cells [25, 26].

Proliferation of erythroid progenitors under conditions that mimic stress erythropoiesis is strictly dependent on PI3K activity [11, 27]. Overexpression of eIF4E abrogated the requirement for SCF-induced PI3K activation, which

suggested a role of selective polysome recruitment of transcript with a complex RNA structure [12]. Genomewide profiling of both total and polysome-associated mRNAs in an erythroblast cell line, I/11, cultured in presence or absence of growth factors identified a large number of constitutively expressed transcripts that are selectively translated in response to PI3K activation, or upon overexpression of eIF4E [14]. This list of eIF4E-sensitive transcripts included alpha4, a subunit of protein phosphatase 2a (PP2a). PP2A exists in various complexes that shift target specificity depending on the binding of regulatory components. mTOR modulates the formation of the PP2A-α4 complex sequestrating the phosphatase activity away from its own downstream targets 4EBP and S6K [28–31]. Constitutive expression of alpha4 in erythroid progenitors completely blocked erythroid differentiation and endowed erythroblast with long-term proliferation in the absence of SCF. It maintained the activation state of the mTOR targets 4EBP and p70S6k in the presence of Epo alone and enhanced polysome recruitment of other eIF4E-senstive transcripts [14]. This further underlines the importance of eIF4E-dependent translation and of the proteins controlled by this mechanism. Among them are indeed several proteins that are essential for erythropoiesis such as Nme2 [14, 32, 33] and the SNARE protein Use1 that is essential for retrograde transport of vesicles from the Golgi to the ER [34].

Mutations that enhance translation initiation efficiency have been implicated in the aggressiveness of various human cancers including Acute Myeloid Leukemia (AML) [35, 36]. Constitutive activation of PI3K and mTOR occurs at high frequency in AML [37]. Mostly the cause of constitutive PI3K activation is unknown, but mutations in the receptor kinases cKIT and FLT3 are candidates. The D816V mutation in the kinase domain of cKit activates the PI3K/PKB/mTOR pathway and confers sensitivity to rapamycin [38]. Interestingly, rapamycin induces cell cycle arrest and apoptosis in patient-derived neoplastic mast cells harbouring the D816V *cKIT*, but not in normal human cord-blood-derived mast cells. This implies that inhibitors targeting translation initiation regulators are therapeutic candidates in the treatment of aggressive systemic mastocytosis (associated with *cKIT* D816V) and in AML harbouring the D816V *cKIT* mutant (present in 10 to 40% of core-binding factor leukaemia [39]).

Enhanced eIF4E-dependent translation of transcripts with a structured 5′UTR also contributes to chronic myelogenous leukaemia (CML). Leukemic transformation of hematopoietic progenitors by the BCR/ABL fusion protein depends on PI3K activation, which will enhance 4EBP phosphorylation and eIF4E release. In addition, BCR/ABL induces expression of SET, which subsequently acts as an inhibitory regulatory subunit of PP2A similar to α4 [14, 40]. Notably, the leukaemic potential of BCR/ABL-expressing cells can be inhibited by pharmacological activation of the phosphatase pp2a [41].

Pharmacological inactivation of mTOR with rapamycin reduces neoplastic proliferation in PTEN deficient mice, and reverses tumour growth in cancer cells characterised by activated PKB [42]. In AML, however, trials with rapamycin have been put on halt because rapamycin induced a feedback pathway resulting in further PKB activation [43]. Given the role of the Pp2a regulatory subunit α4, reactivation of the phosphatase, pp2a, offers a potential alternative treatment to therapy-resistant patients [44].

2.2. Translation Initiation Sensitivity to eIFs: eIF2α Phosphorylation and AUG Selection. The second limiting translation factor next to eIF4E is eIF2, a GTPase that is associated with methionine-loaded initiator tRNA only in its GTP-bound state. This complex of eIF2:GTP/$tRNA_i^{met}$ is known as the ternary complex (TC) [5]. The GTPase activity of eIF2 is activated by recognition of an AUG codon and depends on the sequence context. The better an initiation codon resembles the consensus sequence (A/GnnAUGG), the higher the chance that the eIF2 GTPase activity is triggered and methionine is delivered to the P-site of the ribosome [45]. This then leads to release of eIF2:GDP. In addition to availability, also the activity of eIF2 is regulated. The initiation factor eIF5 increases GTPase activity and increases the probability that methionine is deposited at the P-site of the ribosome at a start codon in a less optimal context [17]. The eIF2 is a trimeric protein that is regulated by phosphorylation of the eIF2α subunit. Phosphorylation inhibits the recovery of eIF2:GTP from eIF2:GDP by protein eIF2B, and thereby the reassociation with $tRNA_i^{met}$. The kinases involved in eIF2 phosphorylation are activated by lack of haem (HRI), unfolded proteins in the ER (PERK), double strand RNA (RNA viruses; PKR), or lack of amino acids (GCN2) [46]. In the hematopoietic system, HRI is extremely important in erythropoiesis because it coordinates haemoglobin synthesis. It links iron availability and haem synthesis to the translation of globin chains [47]. Phosphorylation of eIF2 results in translation of Atf4 (activating transcription factor 4) and subsequent transcription of Gadd34. Gadd34 is an activating regulatory subunit of phosphatase Pp1 that is able to dephosphorylate eIF2-P. Mice lacking this feedback regulation develop severe anemia [48].

Similar to eIF4E, also eIF2 availability has consequences for both the overall protein synthesis rate and the translation of transcripts that carry regulatory sequences in their 5′UTR. The phosphorylation level of eIF2 specifically controls translation of transcripts with upstream open reading frames (uORFs). Following the translation of a short uORF, the preinitiation complex lacking eIF2 continues scanning. During scanning the TC will reassociate to enable translation initiation at the next AUG startcodon in an appropriate context. The availability of TC will determine how fast translation can reinitiate. Thus, eIF2 phosphorylation renders translation dependent on the distance between an uORF and the AUG start codon. Notably, the majority of transcripts that are dependent on SCF-induced PI3K activity and on availability of eIF4E in erythroblasts also contain several uORFs, including the previously mentioned transcripts encoding Nme2, Use1, and α4. Translation of these transcripts is hypersensitive not only to growth factor signalling but also to inhibition by oxidative stress, or iron availability.

The transcript encoding thrombopoietin (TPO) contains 7 uORFs of which the last uORF overlaps with the TPO start codon [49]. In hereditary thrombocytosis (HT) mutations

that deregulate translation of some of these uORFs cause high levels of TPO expression and thrombosis.

Notably, the transcripts of several Ets-family members, Scl/Tal, and C/EBPα and β contain an upstream ORF (uORF) that overlaps and is out of frame of the proper AUG start codon of the full-length isoform of the transcription factor [50, 51]. The uORF starts with a uAUG in a suboptimal Kozak consensus and hence is only translated at enhanced activity and availability of eIF2:GTP and eIF4E. When the uORF is translated, the initiation codon of the full-length protein is skipped and reinitiation at a downstream AUG codon results in synthesis of a shorter protein isoform. In the case of C/EBPα translation initiation at a downstream AUG results in a truncated transcription factor protein that acts as a dominant negative isoform [52].

The importance of the relative abundance of C/EBPα isoforms is evidenced by the occurrence of mutations in acute myeloid leukaemia (AML) cases that inhibit translation of the full-length C/EBPα protein [52]. The functional 30 kDa-truncated protein expressed in these patients was shown to inhibit G-CSF receptor in 32Dcl3 cells induced to differentiate into neutrophils [53]. In addition, C/EBPα is required for the generation of the GMP compartment and its expression also denotes selectivity in differential commitment to monocytic lineage [54]. In a subset of AML C/EBPα is mutated [55]. Strikingly, two mutations are combined: an N-terminal mutation on 1 allele and a C-terminal mutation on the other allele. The N-terminal mutation abrogates the long isoform of C/EBPα but allows for translation from the downstream AUG start codon, which contributes to the phenotype of the AML [56].

Increasing evidence supports the importance of expressing different isoforms that modulate lineage commitment in hematopoietic cells. The expression of truncated forms of the transcription factor, Stem Cell Leukaemia (Scl), is regulated by differential initiation of translation [51] and results in erythroid lineage differentiation. Expression of functional isoforms due to differential translation has been described for other hematopoietic transcription factors [50, 57] and disruption of isoform ratios is implicated in disease [52, 53].

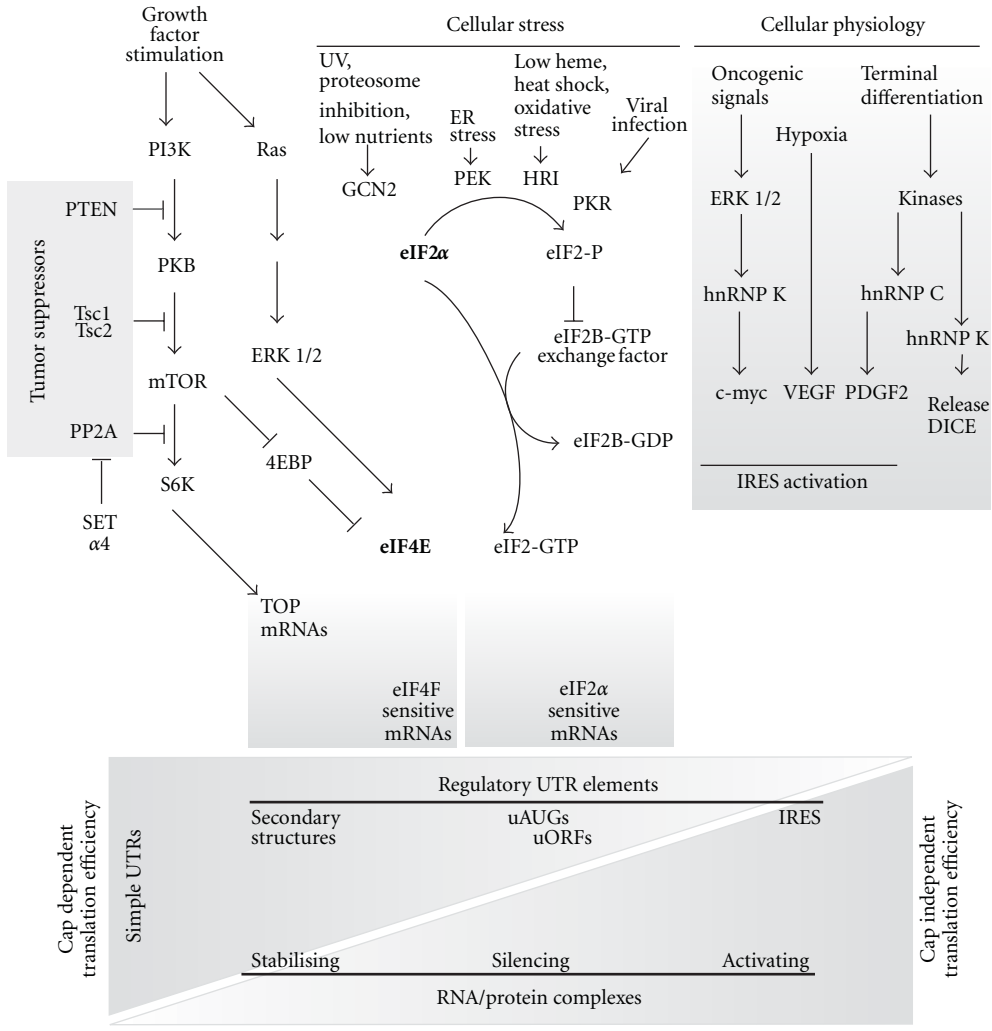
The list of transcripts regulated at the level of alternative AUG usage in hematopoietic regulation is increasing. A novel technique to identify translation start codons indicated that 65% of all transcripts expressed in mouse ES cells are translated from at least 2 alternative start sites [58].

Translation initiation inhibition of growth regulatory proteins (growth factors, cytokines, oncogenes, repressors of tumour suppressor inhibitors, and others) is a known phenomenon [59] and this may be extended to regulatory proteins that attenuate cellular terminal differentiation.

2.3. Translation Initiation Sensitivity to eIFs: Cap-Independent Translation. Regulatory elements at the 5′UTR of mature transcripts render translation dependent on signalling or other environmental conditions such as iron availability. Some transcripts however have a highly structured 5′UTR meant to completely block cap-dependent translation initiation. Under the conditions that these proteins need to be synthesized, cap-dependent translation initiation is strongly decreased and translation initiation reverts to internal ribosome entry sites (IRESs). The structural complexity of IRES elements argue in favour of their role as translation inhibitors, although it is more correct to define these structures as modulators of translation. For instance, although the 5′UTR of platelet-derived growth factor (*PDGF2*) is long (1022 bp), is structured and contains upstream ORFs, it is efficiently translated during megakaryocytic differentiation via binding and activation of the IRES by hnRNP C (Figure 2) [60]. Hence, specificity of IRES-mediated gene expression is determined by IRES trans-acting factors (ITAFs), representing a particular cellular state. In addition to differentiation, translation initiation starting from an IRES is found in transcripts encoding proapoptotic proteins or proteins required specifically during G2/M when transcription is silenced by compaction of the DNA [61, 62].

IRES-dependent translation is less competitive for polysome recruitment than cap-dependent translation. Therefore, IRES-dependent translation may be preferentially impaired when ribosome subunits are reduced, a situation that is typical for Diamond Blackfan Anemia (DBA). Expression profiling of polyribosome-bound mRNAs from erythroblasts identified a specific set of transcripts that are selectively lost from polyribosomes upon reduced expression of Rps19 [63]. Among these mRNAs were transcripts encoding the Hsp70/Hsc70 cochaperone Bag1 (Bcl-2-associated athanogene 1), and the RNA binding protein Csde1 (cold shock domain containing E1), both requiring an IRES for translation initiation. Importantly, expression of BAG1 and CSDE1 was also reduced in human erythroblasts cultured from peripheral blood of DBA patients, whereas *BAG1* and *CSDE1* mRNA level was constant or even elevated. Interestingly, Csde1 binds the IRES of several transcripts and controls IRES-mediated translation [64].

2.4. RNA-Binding Proteins: Transcript Stability and Translation Control. mRNA translation is regulated by activities of the translation machinery components as described previously but also via regulation of proteins that bind to specific mRNAs. RNA binding protein complexes attenuate expression by modulating stability, degradation, cellular localisation, and silencing of specific RNAs or subgroups of mRNAs. The most studied RNA-binding proteins present in ribonucleoprotein (RNP) particles are the heterogeneous nuclear ribonucleoproteins (hnRNPs) that recognize AU-rich elements (ARE) and coordinate expression of mRNAs at the level of nuclear-cytoplasmic shuttling [65], cytoplasmic mRNA turnover [66], and silencing of cell state- and type-specific mRNAs [67, 68]. The ARE is located in the 3′ untranslated region of many short-lived transcripts from cytokines, proto-oncogenes, growth factors, or cell cycle regulators [69]. Interestingly, tristetraprolin (TTP) and butyrate response factor (BRF1) belong to the same protein family and both promote ARE-dependent decay [70]. Interferon γ (IFNγ) suppresses the survival and expansion of T-helper 17 (Th17) cells by inducing expression of TTP, resulting in the destabilisation of the p19 mRNA, coding for a subunit of IL23 [71, 72]. The tight regulation of proinflammatory cytokines through mRNA stability is required to suppress



In bold: translation initiation factors

Figure 2: Translation initiation control during growth factor stimulation, cellular stress, and cellular physiology. Growth factor addition activates the PI3K/PKB/mTOR pathway releasing the limiting translation initiation factor 4E (eIF4E) from a repression complex with 4EBP and activating S6K resulting in enhanced cap-dependent translation efficiency of structured mRNAs and ribogenesis. Interestingly, the tumour suppressor proteins PTEN, Tsc1/2, and Pp2a are involved in attenuating this pathway. Another limiting initiation factor, eIF2α, is involved in providing methionine-tRNA in a complex with the 60S ribosome subunit to start peptide synthesis once the proper AUG is recognised. eIF2 is phosphorylated by GCN2, PEK, HRI, or PRK in response to various stress conditions. Low levels of eIF4E and eIF2-GTP as a result of 4EBP repression or stress-induced eIF2 phosphorylation, respectively, repress cap-dependent translation. These conditions are optimal for translation initiation from Internal Ribosomal Entry Sites (IRESs). The levels of eIFs modulate translation initiation and this depends on the codes offered by the transcripts. Some transcripts are ideal to be translated under stress conditions having IRES structures in their 5′UTRs; others have secondary structures that are difficult to melt and hence hinder the scanning process. The presence of uORFs, attenuates translation initiation and also has a role in protein isoform formation. RNA-binding proteins modulate specific mRNAs by stabilising, silencing, or activating the transcripts. These RNA/protein complexes (RNPs) have an important role in cellular physiology. Some RNPs respond to oncogenic signals, while others are covalently modified and drive translation in response to terminal differentiation signals as in the case of the DICE elements (translation initiation factors in bold).

inflammatory responses. In fact, TTP deficient mice have an overproduction of cytokines resulting in a systemic inflammatory response with clinical outcomes including arthritis and autoimmunity [73].

In contrast to TTP and Brf1, members of the ELAV family of RNA-binding proteins (e.g., HuR) bind and stabilize ARE-containing transcripts [74]. Interestingly, in the context of a closed loop model of translated eukaryotic mRNAs, the recruitment of HuR and other RNA binding proteins to 3′UTR elements results in the formation of complexes between HuR and the scanning ribosome at the 5′UTR. This stabilises the transcript and facilitates translation initiation at the proper AUG in transcripts with a structured 5′UTR [75]. Upregulation of TNFalpha in a mouse model for autoimmune disease depends on binding of ELAV-like protein to the AU-rich elements in its 3′UTR [76].

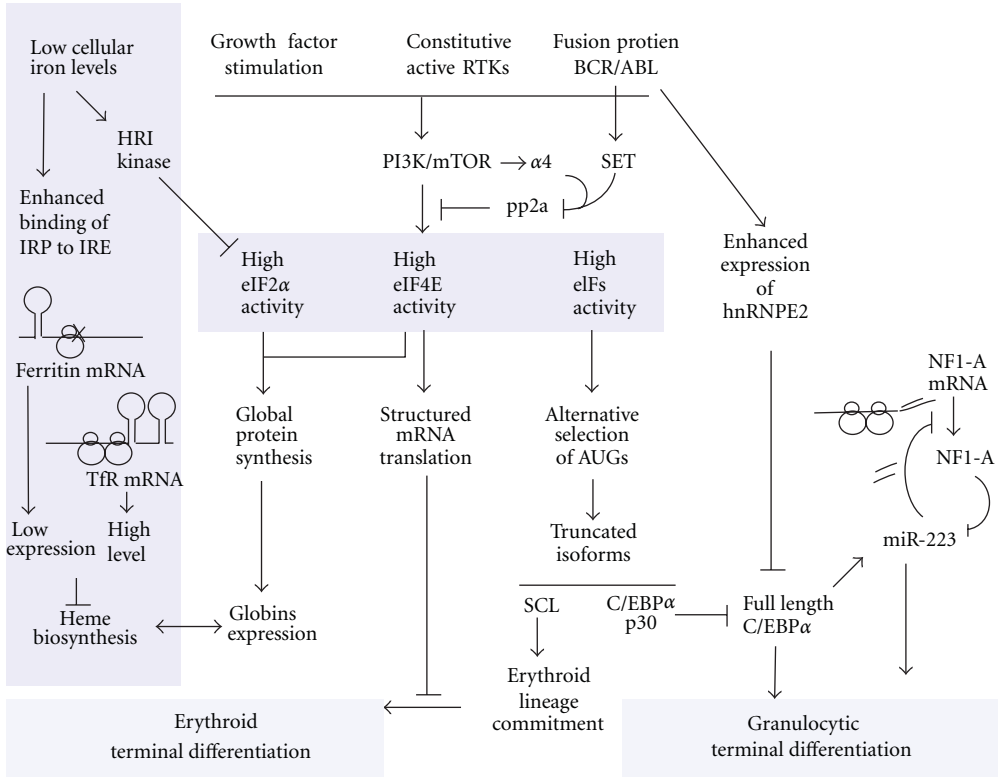


Figure 3: Translation initiation control relays signals to erythroid and granulocytic differentiation. SCF binds c-kit, a Receptor Tyrosine Kinases (RTKs), activating PI3K/mTOR pathway in the same way as constitutive active mutant RTKs and kinase active fusion protein, BCR/ABL. mTOR downstream effector proteins are maintained active by attenuating the phosphatase Pp2a which is inhibited by SCF-driven alpha4 expression and enhanced expression of SET in response to BCR/ABL. High activity of translation initiation factors enhances polysome recruitment of structured mRNAs and delays erythroid terminal differentiation. During erythroid terminal differentiation the balance between globin synthesis and haeme biosynthesis is under the tight control of translation initiation. Iron Responsive Element (IRE) in the UTRs of ferritin and transferrin modulates iron uptake and storage in accordance to demand of haeme. Low cellular iron levels trigger phosphorylation of eIF2α to reduce the production of globin proteins. High eIFs levels also regulate commitment to the erythroid or megakaryocytic lineage by selective usage of AUGs in the SCL transcript driving different isoform production. The same mechanism is used to produce truncated isoforms of the transcription factor C/EBPα that acts as a dominant negative form of the full length and hence inhibits granulocytic terminal differentiation. Another form of translation control is involved in regulation of C/EBPα transcription activity. Full-length C/EBPα enhances transcription of micro RNA 223 (miRNA-223), an inhibitor of NFI-A translation. NFI-A is a competitor for binding C/EBPα DNA sites and hence its inhibition results in a positive feedback loop driving granulocytic differentiation. In addition to transcription inhibition of full-length C/EBPα driven by selective AUG usage or translation silencing of competitors, the role of RNA-binding proteins is important in modulating terminal differentiation. BCR/ABL enhances the expression of hnRNPE2 that binds the UTR of C/EBPα transcript and inhibits translation.

Modification of RNA binding-proteins by signalling can be another level of regulating translation efficiency at the proper AUG. Silencing of transcripts by binding of hnRNP K and hnRNP E1 to differentiation control element (DICE) in the 3′ UTR of 15-lipoxygenase (LOX) mRNA transcript [67] can be released during terminal erythroid differentiation by phosphorylation of hnRNP K (Figure 2) [77].

Another prototype example in which RNA-binding proteins regulate translation in hematopoietic cells is the iron-response element (IRE) [78]. It is a stem-loop structure that binds iron regulatory proteins IRP1 and IRP2 dependent on their association with iron [79, 80]. Whereas the free IRE can easily be unwound by the eIF4A helicase during scanning of the preinitiation complex, IRP binding stabilises the stem-loop structure and impairs continued scanning and translation. Initially characterised for ferritin mRNA, IRE elements are found in a number of transcripts encoding proteins involved in iron metabolism or hemoglobin synthesis. An example of the latter is alpha-hemoglobin-stabilizing protein (AHSP) mRNA. Because alpha globin is synthesised ahead of beta globin, it has to be stabilised until its incorporation into haemoglobin. Regulation of AHSP by iron implies regulation of the stability of alpha globin [81].

Ceruloplasmin is involved in iron metabolism and is translationally regulated by interferon gamma [82]. Phosphorylation of ribosomal protein L13a by interferon-gamma results in dissociation of L13a from the 60S ribosome subunit and recruitment of an Rpl13-containing protein complex to a structural element in the 3′UTR of ceruloplasmin [83] resulting in translation repression. This mechanism incorporates 2 novel issues. First, the ribosome is able to present signalling sensitive factors that can be released to

attenuate translation of specific transcripts without affecting global synthesis rates. Second, regulatory elements in the 3′UTR recruit protein complexes within the circular mature transcripts and interact with scanning complexes in the 5′UTR, hence modulating translation initiation efficiency.

The expression of RNA-binding proteins that attenuate translation of specific subsets of mRNAs has been implicated in the transition from chronic CML to blast crisis events [84, 85]. For instance, ectopic expression of hnRNP E2, an RNA-binding protein upregulated during blast crisis of CML, resulted in downregulation of C/EBPα (Figure 3) and G-CSFR in myeloid progenitor cells, inhibiting granulocytic differentiation [86].

3. Conclusion and Future Perspectives

Regulation of gene expression has been studied extensively in disease models and patient groups giving detailed annotations of the differential expression at the level of transcription regulation. Differential expression between the transcriptome and the proteome supports the importance of post-transcriptional regulation. Increasing evidence supports the prominent and so far underestimated regulation of mRNA translation, which depends on the availability and activity of the translation machinery, the structure of the transcript, and the expression of RNA binding proteins. Growth factor signaling enhances polysome recruitment of specific RNA transcripts, with marginal effect on the transcription regulation [23]. Similarly, increased availability of the translation initiation factor, eIF4E, through growth factor signaling, overexpression studies, and suppressions of the attenuation mechanism of mTOR signaling results in proliferative advantage of erythroid progenitors and block of differentiation [12, 14]. Increasingly, differential or conditional expression of transcripts are being associated with haematological malignancies, including expression of dominant negative isoforms involved in progenitor differentiation [52], expression of RNA binding proteins involved in inflammatory responses [71], and suppressed expression of phosphatase subunits resulting in progression of disease [40]. Interestingly, deregulation of negative feedback mechanisms due to low activity of phosphatases is potentially targeted using mTOR inhibitors or phosphatase activators. Hence understanding translational control mechanisms and its deregulation in disease will impact on patient stratification into therapeutic groups.

References

[1] R. L. Phillips, R. E. Ernst, B. Brunk et al., "The genetic program of hematopoietic stem cells," *Science*, vol. 288, no. 5471, pp. 1635–1640, 2000.

[2] E. W. Scott, M. C. Simon, J. Anastasi, and H. Singh, "Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages," *Science*, vol. 265, no. 5178, pp. 1573–1577, 1994.

[3] K. M. Vattem and R. C. Wek, "Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 31, pp. 11269–11274, 2004.

[4] A. C. Gingras, B. Raught, and N. Sonenberg, "Regulation of translation initiation by FRAP/mTOR," *Genes and Development*, vol. 15, no. 7, pp. 807–826, 2001.

[5] N. Sonenberg and A. G. Hinnebusch, "Regulation of translation initiation in Eukaryotes: mechanisms and biological targets," *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[6] V. C. Broudy, N. L. Lin, G. V. Priestley, K. Nocka, and N. S. Wolf, "Interaction of stem cell factor and its receptor c-kit mediates lodgment and acute expansion of hematopoietic cells in the murine spleen," *Blood*, vol. 88, no. 1, pp. 75–81, 1996.

[7] A. Bauer, F. Tronche, O. Wessely et al., "The glucocorticoid receptor is required for stress erythropoiesis," *Genes and Development*, vol. 13, no. 22, pp. 2996–3002, 1999.

[8] M. von Lindern, W. Zauner, G. Mellitzer et al., "The glucocorticoid receptor cooperates with the erythropoietin receptor and c-Kit to enhance and sustain proliferation of erythroid progenitors in vitro," *Blood*, vol. 94, no. 2, pp. 550–559, 1999.

[9] H. Dolznig, B. Habermann, K. Stangl et al., "Apoptosis protection by the Epo target Bcl-XL allows factor-independent differentiation of primary erythroblasts," *Current Biology*, vol. 12, no. 13, pp. 1076–1085, 2002.

[10] M. von Lindern, U. Schmidt, and H. Beug, "Control of erythropoiesis by erythropoietin and stem cell factor: a novel role for Bruton's tyrosine kinase," *Cell Cycle*, vol. 3, no. 7, pp. 876–879, 2004.

[11] W. J. Bakker, M. Blázquez-Domingo, A. Kolbus et al., "FoxO3a regulates erythroid differentiation and induces BTG1, an activator of protein arginine methyl transferase 1," *Journal of Cell Biology*, vol. 164, no. 2, pp. 175–184, 2004.

[12] M. Blázquez-Domingo, G. Grech, and M. von Lindern, "Translation initiation factor 4E inhibits differentiation of erythroid progenitors," *Molecular and Cellular Biology*, vol. 25, no. 19, pp. 8496–8506, 2005.

[13] X. Wang, A. Beugnet, M. Murakami, S. Yamanaka, and C. G. Proud, "Distinct signaling events downstream of mTOR cooperate to mediate the effects of amino acids and insulin on initiation factor 4E-binding proteins," *Molecular and Cellular Biology*, vol. 25, no. 7, pp. 2558–2572, 2005.

[14] G. Grech, M. Blazquez-Domingo, A. Kolbus et al., "Igbp1 is part of a positive feedback loop in stem cell factor-dependent, selective mRNA translation initiation inhibiting erythroid differentiation," *Blood*, vol. 112, no. 7, pp. 2750–2760, 2008.

[15] M. Murakami, T. Ichisaka, M. Maeda et al., "mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells," *Molecular and Cellular Biology*, vol. 24, no. 15, pp. 6710–6718, 2004.

[16] M. Kozak, "A second look at cellular mRNA sequences said to function as internal ribosome entry sites," *Nucleic Acids Research*, vol. 33, no. 20, pp. 6593–6602, 2005.

[17] T. V. Pestova, V. G. Kolupaeva, I. B. Lomakin et al., "Molecular mechanisms of translation initiation in eukaryotes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 13, pp. 7029–7036, 2001.

[18] N. Sonenberg and A. C. Gingras, "The mRNA 5′ cap-binding protein eIF4E and control of cell growth," *Current Opinion in Cell Biology*, vol. 10, no. 2, pp. 268–275, 1998.

[19] A. Flynn and C. G. Proud, "Serine 209, not serine 53, is the major site of phosphorylation in initiation factor eIF-4E in serum-treated Chinese hamster ovary cells," *Journal of Biological Chemistry*, vol. 270, no. 37, pp. 21684–21688, 1995.

[20] B. Joshi, A. L. Cai, B. D. Keiper et al., "Phosphorylation of eukaryotic protein synthesis initiation factor 4E at Ser-209," *Journal of Biological Chemistry*, vol. 270, no. 24, pp. 14597–14603, 1995.

[21] A. J. Waskiewicz, A. Flynn, C. G. Proud, and J. A. Cooper, "Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2," *EMBO Journal*, vol. 16, no. 8, pp. 1909–1920, 1997.

[22] A. De Benedetti and J. R. Graff, "eIF-4E expression and its role in malignancies and metastases," *Oncogene*, vol. 23, no. 18, pp. 3189–3199, 2004.

[23] V. K. Rajasekhar, A. Viale, N. D. Socci, M. Wiedmann, X. Hu, and E. C. Holland, "Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes," *Molecular Cell*, vol. 12, no. 4, pp. 889–901, 2003.

[24] T. Waerner, M. Alacakaptan, I. Tamir et al., "ILEI: a cytokine essential for EMT, tumor formation, and late events in metastasis in epithelial cells," *Cancer Cell*, vol. 10, no. 3, pp. 227–239, 2006.

[25] Y. Mamane, E. Petroulakis, Y. Martineau et al., "Epigenetic activation of a subset of mRNAs by elF4E explains its effects on cell proliferation," *PLoS One*, vol. 2, no. 2, article e242, 2007.

[26] O. Larsson, S. Li, O. A. Issaenko et al., "Eukaryotic translation initiation factor 4E-induced progression of primary human mammary epithelial cells along the cancer pathway is associated with targeted translational deregulation of oncogenic drivers and inhibitors," *Cancer Research*, vol. 67, no. 14, pp. 6814–6824, 2007.

[27] M. von Lindern, E. M. Deiner, H. Dolznig et al., "Leukemic transformation of normal murine erythroid progenitors: v- and c-ErbB act through signaling pathways activated by the EpoR and c-Kit in stress erythropoiesis," *Oncogene*, vol. 20, no. 28, pp. 3651–3664, 2001.

[28] S. Inui, H. Sanjo, K. Maeda, H. Yamamoto, E. Miyamoto, and N. Sakaguchi, "Ig receptor binding protein 1 (α4) is associated with a rapamycin- sensitive signal transduction in lymphocytes through direct binding to the catalytic subunit of protein phosphatase 2A," *Blood*, vol. 92, no. 2, pp. 539–546, 1998.

[29] M. Kong, C. J. Fox, J. Mu et al., "The PP2A-associated protein α4 is an essential inhibitor of apoptosis," *Science*, vol. 306, no. 5696, pp. 695–698, 2004.

[30] H. Chung, A. C. Nairn, K. Murata, and D. L. Brautigan, "Mutation of Tyr307 and Leu309 in the protein phosphatase 2A catalytic subunit favors association with the α4 subunit which promotes dephosphorylation of elongation factor-2," *Biochemistry*, vol. 38, no. 32, pp. 10371–10376, 1999.

[31] T. D. Prickett and D. L. Brautigan, "The α4 regulatory subunit exerts opposing allosteric effects on protein phosphatases PP6 and PP2A," *Journal of Biological Chemistry*, vol. 281, no. 41, pp. 30503–30511, 2006.

[32] M. Joosten, M. Blazquez-Domingo, F. Lindeboom et al., "Translational control of putative protooncogene Nm23-M2 by cytokines via phosphoinositide 3-kinase signaling," *The Journal of Biological Chemistry*, vol. 279, no. 37, pp. 38169–38176, 2004.

[33] E. H. Postel, X. Zou, D. A. Notterman, and K. M. D. La Perle, "Double knockout Nme1/Nme2 mouse model suggests a critical role for NDP kinases in erythroid development," *Molecular and Cellular Biochemistry*, vol. 329, no. 1-2, pp. 45–50, 2009.

[34] M. Dilcher, B. Veith, S. Chidambaram, E. Hartmann, H. D. Schmitt, and G. F. Von Mollard, "Use1p is a yeast SNARE protein required for retrograde traffic to the ER," *EMBO Journal*, vol. 22, no. 14, pp. 3664–3674, 2003.

[35] B. J. Longley, M. J. Reguera, and Y. Ma, "Classes of c-KIT activating mutations: proposed mechanisms of action and implications for disease classification and therapy," *Leukemia Research*, vol. 25, no. 7, pp. 571–576, 2001.

[36] Y. Li, S. Fan, J. Koo et al., "Elevated expression of eukaryotic translation initiation factor 4E is associated with proliferation, invasion and acquired resistance to erlotinib in lung cancer," *Cancer Biology & Therapy*, vol. 13, no. 5, 2012.

[37] P. Cornillet-Lefebvre, W. Cuccuini, V. Bardet et al., "Constitutive phosphoinositide 3-kinase activation in acute myeloid leukemia is not due to p110δ mutations," *Leukemia*, vol. 20, no. 2, pp. 374–376, 2006.

[38] M. Gabillot-Carré, Y. Lepelletier, M. Humbert et al., "Rapamycin inhibits growth and survival of D816V-mutated c-kit mast cells," *Blood*, vol. 108, no. 3, pp. 1065–1072, 2006.

[39] A. Beghini, P. Peterlongo, C. B. Ripamonti et al., "C-kit mutations in core binding factor leukemias [3]," *Blood*, vol. 95, no. 2, pp. 726–727, 2000.

[40] P. Neviani, R. Santhanam, J. J. Oaks et al., "FTY720, a new alternative for treating blast crisis chronic myelogenous leukemia and Philadelphia chromosome-positive acute lymphocytic leukemia," *Journal of Clinical Investigation*, vol. 117, no. 9, pp. 2408–2421, 2007.

[41] D. Perrotti, F. Turturro, and P. Neviani, "BCR/ABL, mRNA translation and apoptosis," *Cell Death and Differentiation*, vol. 12, no. 6, pp. 534–540, 2005.

[42] M. S. Neshat, I. K. Mellinghoff, C. Tran et al., "Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 18, pp. 10314–10319, 2001.

[43] J. Tamburini, N. Chapuis, V. Bardet et al., "Mammalian target of rapamycin (mTOR) inhibition activates phosphatidylinositol 3-kinase/Akt by up-regulating insulin-like growth factor-1 receptor signaling in acute myeloid leukemia: rationale for therapeutic inhibition of both pathways," *Blood*, vol. 111, no. 1, pp. 379–382, 2008.

[44] K. G. Roberts, A. M. Smith, F. McDougall et al., "Essential requirement for PP2A inhibition by the oncogenic receptor c-KIT suggests PP2A reactivation as a strategy to treat c-KIT+ cancers," *Cancer Research*, vol. 70, no. 13, pp. 5438–5447, 2010.

[45] M. Kozak, "The scanning model for translation: an update," *Journal of Cell Biology*, vol. 108, no. 2, pp. 229–241, 1989.

[46] R. C. Wek, H. Y. Jiang, and T. G. Anthony, "Coping with stress: EIF2 kinases and translational control," *Biochemical Society Transactions*, vol. 34, no. 1, pp. 7–11, 2006.

[47] J. J. Chen, "Regulation of protein synthesis by the heme-regulated eIF2α kinase: relevance to anemias," *Blood*, vol. 109, no. 7, pp. 2693–2699, 2007.

[48] A. D. Patterson, M. C. Hollander, G. F. Miller, and A. J. Fornace Jr., "Gadd34 requirement for normal hemoglobin synthesis," *Molecular and Cellular Biology*, vol. 26, no. 5, pp. 1644–1653, 2006.

[49] M. Cazzola and R. C. Skoda, "Translational pathophysiology: a novel molecular mechanism of human disease," *Blood*, vol. 95, no. 11, pp. 3280–3288, 2000.

[50] C. F. Calkhoven, C. Muller, and A. Leutz, "Translational control of C/EBPα and C/EBPβ isoform expression," *Genes and Development*, vol. 14, no. 15, pp. 1920–1932, 2000.

[51] C. F. Calkhoven, C. Müller, R. Martin, G. Krosl, T. Hoang, and A. Leutz, "Translational control of SCL-isoform expression in hematopoietic lineage choice," *Genes and Development*, vol. 17, no. 8, pp. 959–964, 2003.

[52] T. Pabst, B. U. Mueller, P. Zhang et al., "Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding

protein-α (C/EBPα), in acute myeloid leukemia," *Nature Genetics*, vol. 27, no. 3, pp. 263–270, 2001.

[53] R. Cleaves, Q. F. Wang, and A. D. Friedman, "C/EBPαp30, a myeloid leukemia oncoprotein, limits G-CSF receptor expression but not terminal granulopoiesis via site-selective inhibition of C/EBP DNa binding," *Oncogene*, vol. 23, no. 3, pp. 716–725, 2004.

[54] D. Wang, J. D'Costa, C. I. Civin, and A. D. Friedman, "C/EBPα directs monocytic commitment of primary myeloid progenitors," *Blood*, vol. 108, no. 4, pp. 1223–1229, 2006.

[55] S. B. van Waalwijk van Doorn-Khosrovani, C. Erpelinck, J. Meijer et al., "Biallelic mutations in the CEBPA gene low CEBPA expression levels as prognostic markers in intermediate-risk AML," *Hematology Journal*, vol. 4, no. 1, pp. 31–40, 2003.

[56] B. J. Wouters, B. Löwenberg, C. A. J. Erpelinck-Verschueren, W. L. J. van Putten, P. J. M. Valk, and R. Delwel, "Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome," *Blood*, vol. 113, no. 13, pp. 3088–3091, 2009.

[57] R. Calligaris, S. Bottardi, S. Cogoi, I. Apezteguia, and C. Santoro, "Alternative translation initiation site usage results in two functionally distinct forms of the GATA-1 transcription factor," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 25, pp. 11598–11602, 1995.

[58] N. T. Ingolia, L. F. Lareau, and J. S. Weissman, "Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes," *Cell*, vol. 147, no. 4, pp. 789–802, 2011.

[59] D. Bergeron, L. Poliquin, C. A. Kozak, and E. Rassart, "Identification of a common viral integration region in Cas-Br-E murine leukemia virus-induced non-T-, non-B-cell lymphomas," *Journal of Virology*, vol. 65, no. 1, pp. 7–15, 1991.

[60] J. Bernstein, I. Shefler, and O. Elroy-Stein, "The translational repression mediated by the platelet-derived growth factor 2/c-sis mRNA leader is relieved during megakaryocytic differentiation," *Journal of Biological Chemistry*, vol. 270, no. 18, pp. 10559–10565, 1995.

[61] S. Pyronnet, L. Pradayrol, and N. Sonenberg, "A cell cycle-dependent internal ribosome entry site," *Molecular Cell*, vol. 5, no. 4, pp. 607–616, 2000.

[62] K. A. Spriggs, M. Bushell, S. A. Mitchell, and A. E. Willis, "Internal ribosome entry segment-mediated translation during apoptosis: the role of IRES-trans-acting factors," *Cell Death and Differentiation*, vol. 12, no. 6, pp. 585–591, 2005.

[63] R. Horos, H. Ijspeert, D. Pospisilova et al., "Ribosomal deficiencies in Diamond-Blackfan anemia impair translation of transcripts essential for differentiation of murine and human erythroblasts," *Blood*, vol. 119, no. 1, pp. 262–272, 2012.

[64] B. Schepens, S. A. Tinton, Y. Bruynooghe et al., "A role for hnRNP C1/C2 and Unr in internal initiation of translation during mitosis," *EMBO Journal*, vol. 26, no. 1, pp. 158–169, 2007.

[65] J. L. Veyrune, G. P. Campbell, J. Wiseman, J. M. Blanchard, and J. E. Hesketh, "A localisation signal in the 3′ untranslated region of c-myc mRNA targets c-myc mRNA and β-globin reporter sequences to the perinuclear cytoplasm and cytoskeletal-bound polysomes," *Journal of Cell Science*, vol. 109, no. 6, pp. 1185–1194, 1996.

[66] G. Shaw and R. Kamen, "A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation," *Cell*, vol. 46, no. 5, pp. 659–667, 1986.

[67] D. H. Ostareck, A. Ostareck-Lederer, I. N. Shatsky, and M. W. Hentze, "Lipoxygenase mRNA silencing in erythroid differentiation: the 3′UTR regulatory complex controls 60S ribosomal subunit joining," *Cell*, vol. 104, no. 2, pp. 281–290, 2001.

[68] M. Notari, P. Neviani, R. Santhanam et al., "A MAPK/HNRPK pathway controls BCR/ABL oncogenic potential by regulating MYC mRNA translation," *Blood*, vol. 107, no. 6, pp. 2507–2516, 2006.

[69] T. Bakheet, M. Frevel, B. R. G. Williams, W. Greer, and K. S. A. Khabar, "ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins," *Nucleic Acids Research*, vol. 29, no. 1, pp. 246–254, 2001.

[70] P. J. Blackshear, "Tristetraprolin and other CCCH tandem zinc-finger proteins in the regulation of mRNA turnover," *Biochemical Society Transactions*, vol. 30, no. 6, pp. 945–952, 2002.

[71] X. Qian, H. Ning, J. Zhang et al., "Posttranscriptional regulation of IL-23 expression by IFN-γ through tristetraprolin," *Journal of Immunology*, vol. 186, no. 11, pp. 6454–6464, 2011.

[72] I. Sauer, B. Schaljo, C. Vogl et al., "Interferons limit inflammatory responses by induction of tristetraprolin," *Blood*, vol. 107, no. 12, pp. 4790–4797, 2006.

[73] G. A. Taylor, E. Carballo, D. M. Lee et al., "A pathogenetic role for TNFα in the syndrome of cachexia, arthritis, and autoimmunity resulting from tristetraprolin (TTP) deficiency," *Immunity*, vol. 4, no. 5, pp. 445–454, 1996.

[74] L. P. Ford, J. Watson, J. D. Keene, and J. Wilusz, "ELAV proteins stabilize deadenylated intermediates in a novel in vitro mRNA deadenylation/degradation system," *Genes and Development*, vol. 13, no. 2, pp. 188–201, 1999.

[75] A. Mehta, C. R. Trotta, and S. W. Peltz, "Derepression of the Her-2 uORF is mediated by a novel post-transcriptional control mechanism in cancer cells," *Genes and Development*, vol. 20, no. 8, pp. 939–953, 2006.

[76] K. Sakai, Y. Kitagawa, M. Saiki, S. Saiki, and G. Hirose, "Binding of the ELAV-like protein in murine autoimmune T-cells to the nonameric AU-rich element in the 3′ untranslated region of CD154 mRNA," *Molecular Immunology*, vol. 39, no. 14, pp. 879–883, 2003.

[77] H. Habelhah, K. Shah, L. Huang et al., "ERK phosphorylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation," *Nature Cell Biology*, vol. 3, no. 3, pp. 325–330, 2001.

[78] A. M. Thomson, J. T. Rogers, and P. J. Leedman, "Iron-regulatory proteins, iron-responsive elements and ferritin mRNA translation," *International Journal of Biochemistry and Cell Biology*, vol. 31, no. 10, pp. 1139–1152, 1999.

[79] T. A. Rouault, "The role of iron regulatory proteins in mammalian iron homeostasis and disease," *Nature Chemical Biology*, vol. 2, no. 8, pp. 406–414, 2006.

[80] M. U. Muckenthaler, B. Galy, and M. W. Hentze, "Systemic iron homeostasis and the iron-responsive element/iron-regulatory protein (IRE/IRP) regulatory network," *Annual Review of Nutrition*, vol. 28, pp. 197–213, 2008.

[81] C. O. Dos Santos, L. C. Dore, E. Valentine et al., "An iron responsive element-like stem-loop regulates α-hemoglobin-stabilizing protein mRNA," *Journal of Biological Chemistry*, vol. 283, no. 40, pp. 26956–26964, 2008.

[82] B. Mazumder, P. Sampath, and P. L. Fox, "Translational control of ceruloplasmin gene expression: beyond the IRE," *Biological Research*, vol. 39, no. 1, pp. 59–66, 2006.

[83] B. Mazumder, P. Sampath, V. Seshadri, R. K. Maitra, P. E. DiCorleto, and P. L. Fox, "Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control," *Cell*, vol. 115, no. 2, pp. 187–198, 2003.

[84] D. Perrotti, R. Trotta, and B. Calabretta, "Altered mRNA translation: possible mechanism for CML disease progression.," *Cell Cycle*, vol. 2, no. 3, pp. 177–180, 2003.

[85] D. Perrotti and B. Calabretta, "Translational regulation by the p210 BCR/ABL oncoprotein," *Oncogene*, vol. 23, no. 18, pp. 3222–3229, 2004.

[86] D. Perrotti, V. Cesi, R. Trotta et al., "BCR-ABL suppresses C/EBPα expression through inhibitory action of hnRNP E2," *Nature Genetics*, vol. 30, no. 1, pp. 48–58, 2002.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 141386, 23 pages
doi:10.1155/2012/141386

Review Article

The Bic-C Family of Developmental Translational Regulators

Chiara Gamberi and Paul Lasko

Department of Biology, McGill University, 3649 Promenade Sir William Osler, Montréal, QC, Canada H3G 0B1

Correspondence should be addressed to Chiara Gamberi, chiara.gamberi@mcgill.ca

Received 24 January 2012; Accepted 18 February 2012

Academic Editor: Greco Hernández

Regulation of mRNA translation is especially important during cellular and developmental processes. Many evolutionarily conserved proteins act in the context of multiprotein complexes and modulate protein translation both at the spatial and the temporal levels. Among these, Bicaudal C constitutes a family of RNA binding proteins whose founding member was first identified in *Drosophila* and contains orthologs in vertebrates. We discuss recent advances towards understanding the functions of these proteins in the context of the cellular and developmental biology of many model organisms and their connection to human disease.

1. Introduction

Translational regulation of mRNA distributed asymmetrically in the early *Drosophila* embryo underlies pattern formation and germ cell specification. Furthermore, expression of certain proteins occurs only at definite stages of development. Exquisite, often partially redundant mechanisms of control ensure the coordination of the spatial and temporal expression of proteins with morphogenetic potential. These mechanisms have been reviewed recently [6]. Here we will discuss the case of one of such translational regulators, Bicaudal C (Bic-C), which is evolutionarily conserved, and for which there is recent accumulating functional evidence from both invertebrate and vertebrate model organisms suggesting that Bic-C is a fundamental regulator of cellular processes and an outstanding example of the fascinating complexity of the developmental mechanisms.

2. Materials and Methods

The sequences shown in this paper are listed in Table 1, and they were recovered by running BLAST [7] with the *Drosophila* sequence and the NCBI sequence database, using the Homologene feature at the NCBI. The sequences for the different *Drosophila* species were retrieved from FlyBase [8]. Sequences were aligned with Clustal W [1, 2].

3. Results and Discussion

3.1. Bic-C . The *Bic-C* gene was originally identified during a *Drosophila* screen for maternal genes affecting embryonic polarity [9]. In fact, adult females bearing *Bic-C* mutations in one of their second chromosomes produce embryos exhibiting anterior-posterior defects of severity ranging from anterior defects, to the development of bicaudal embryos composed of as few as four segments arranged as two, mirror-image posterior ends, to embryos that fail to cellularize [3]. This pleiotropy indicates that Bic-C participates in (or influences) many different pathways.

Early work demonstrated that Bic-C is required during oogenesis to establish anterior-posterior polarity in the oocyte [3, 5, 9, 10]. It encodes a 905-amino-acid (aa) RNA binding protein containing two canonical and three non-canonical KH RNA binding domains (KH2, 4 and KH 1, 3, 5, resp., aa 56–524) [3, 11, 12], a C-terminal Sterile Alpha Motif domain (SAM domain, aa 805–868, Prosite) [13], and a region rich in serine and glycine (aa 598–693). In the Bic-C protein, both the region containing the KH domains and the full-length, recombinant protein possess affinity for RNA [14, 15] with the full-length protein exhibiting more selective binding of synthetic probes *in vitro*. RNA binding is likely important to Bic-C function in fruit flies, as a spontaneous mutation (G296R) that affects the third KH domain, decreases RNA affinity *in vitro,* and exhibits

Table 1: Sequences used in this study.

Sequences	Species
Bic-C	
Gene Bank ID	
gi|24584539	*D. melanogaster* B isoform
gi|158300058	*A. gambiae*
gi|13994223	*M. musculus*
gi|109509376	*R. norvegicus*
gi|122937472	*H. sapiens*
gi|114631037	*P. troglodytes*
gi|73953060	*C. familiaris*
gi|194679417	*B. taurus*
gi|292623098	*D. rerio*
gi|212646112	*C. elegans*
gi|118092391	*G. gallus*
FlyBase ID	
FBpp0080362	*D. melanogaster* B isoform
FBpp0080363	*D. melanogaster* D isoform
FBpp0080361	*D. melanogaster* A isoform
FBpp0118127	*D. ananassae*
FBpp0143734	*D. erecta*
FBpp0144300	*D. grimshawi*
FBpp0166588	*D. mojavensis*
FBpp0179414	*D. persimilis*
FBpp0287937	*D. pseudobscura*
FBpp0200128	*D. sechellia*
FBpp0222439	*D. simulans*
FBpp0232468	*D. virilis*
FBpp0253912	*D. willistoni*
FBpp0266309	*D. yakuba*
Not3/5	
Gene Bank ID	
gi|39945962	*Magnaportae oryzae*
gi|85075997	*Neurospora crassa*
gi|19115701	*S. pombe*
gi|19921660	*D. melanogaster*
gi|158299738	*A. gambiae*
gi|22122717	*M. musculus*
gi|34854462	*R. norvegicus*
gi|7657387	*H. sapiens*
gi|114678945	*P. troglodytes*
gi|73946891	*C. familiaris*
gi|119911200	*B. taurus*
gi|53933228	*D. rerio*
gi|133901756	*C. elegans*
gi|238481292	*A. thaliana*
gi|115454389	*O. sativa japonica*
FlyBase ID	
FBpp0085398	*D. melanogaster*
FBpp0125948	*D. ananassae*
FBpp0129398	*D. erecta*
FBpp0147530	*D. grimshawi*
FBpp0160933	*D. mojavensis*
FBpp01852	*D. persimilis*
FBpp0288020	*D. pseudobscura*
FBpp0197981	*D. sechellia*
FBpp0208756	*D. simulans*
Bpp0227498	*D. virilis*
FBpp0243918	*D. willistoni*
FBpp0264455	*D. yakuba*

a strong phenotype *in vivo* [3]. However, this mutation may be affecting more than RNA binding of the whole protein, for example, by perturbing secondary structure in its neighbourhood, as it may be the case for a similar mutation occurring in another KH domain [12]. If this were the case, the severity of the phenotype may be due to the combination of lack of RNA interaction and other defective pathways under Bic-C control in the wild type. The region containing the KH domains in two Bic-C orthologs shows conserved RNA binding capability in the mouse Bicc1 [16] and, surprisingly, not in the *C. elegans* GLD-3 [12].

SAM domains are ancient modules present in most species that are commonly engaged in mediating protein-protein interaction [13, 17] and can multimerize [18, 19]. Multimerization of RNA binding proteins and RNA is most likely the basis for building RNP particles and a target of regulation. Interestingly, the SAM domain of the human BICC1 can form polymers *in vitro* [20] and some KH domains can mediate interactions between proteins [21, 22]. This is also the case for the C. elegans GLD-3 that interacts with the GLD-2 polymerase via its first KH domain [23] therefore it is likely that Bic-C is part of multiprotein complexes such as cellular RNPs. Certain SAM domains have also been implicated in RNA binding, as the case of *Drosophila* Smaug and *S. cerevisiae* Vts1 [24]. Interestingly, among all the *Drosophila* SAM domains, Bic-C contains the one most similar to Smaug's, which includes the critical residues for RNA interaction [25], suggesting the possibility that it may contribute to the Bic-C RNA binding capacity in the cell [17]. Studies of the vertebrate Bic-C homologs, whose targets are largely unknown, have suggested that presence of the SAM domain may mediate association with the P-bodies [26, 27]. Another interesting possibility is that the putative RNA binding and protein-protein interaction capabilities of the SAM domain may be regulated, possibly via posttranslational modifications. In this scenario protein modification in this domain may change the specificity and/or affinity of Bic-C for RNA to switch between protein and RNA binding activities in certain cellular or developmental contexts. Interestingly, a tyrosine residue in position 822 that can be phosphorylated in other SAM domains to regulate their activity is also conserved [28] (Figure 1).

3.2. Evolutionary Conservation of the Bic-C Protein. Bic-C is found in all the sequenced *Drosophila* species and its homologs are virtually identical to each other, except for regions of

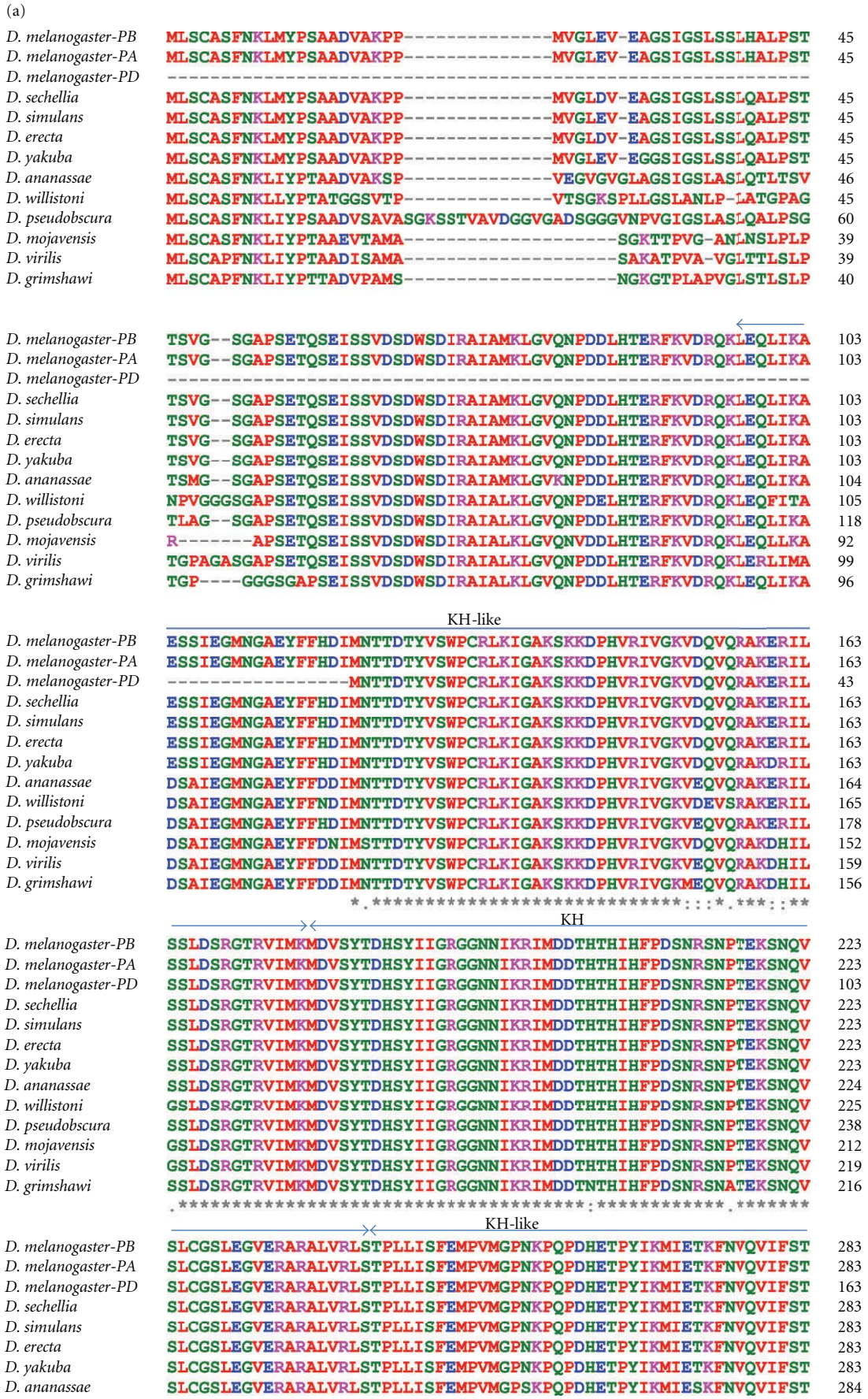


Figure 1: Continued.

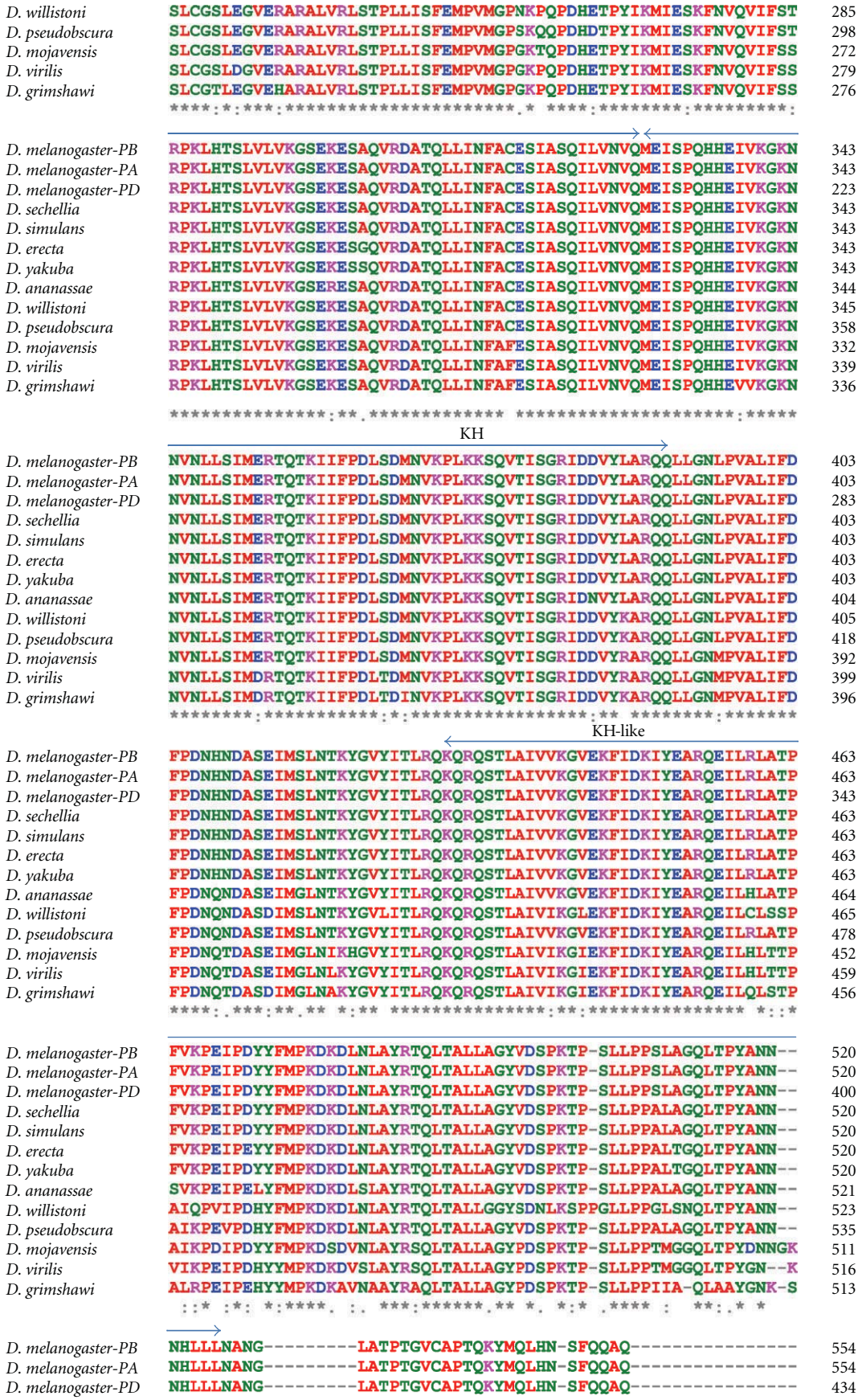


Figure 1: Continued.

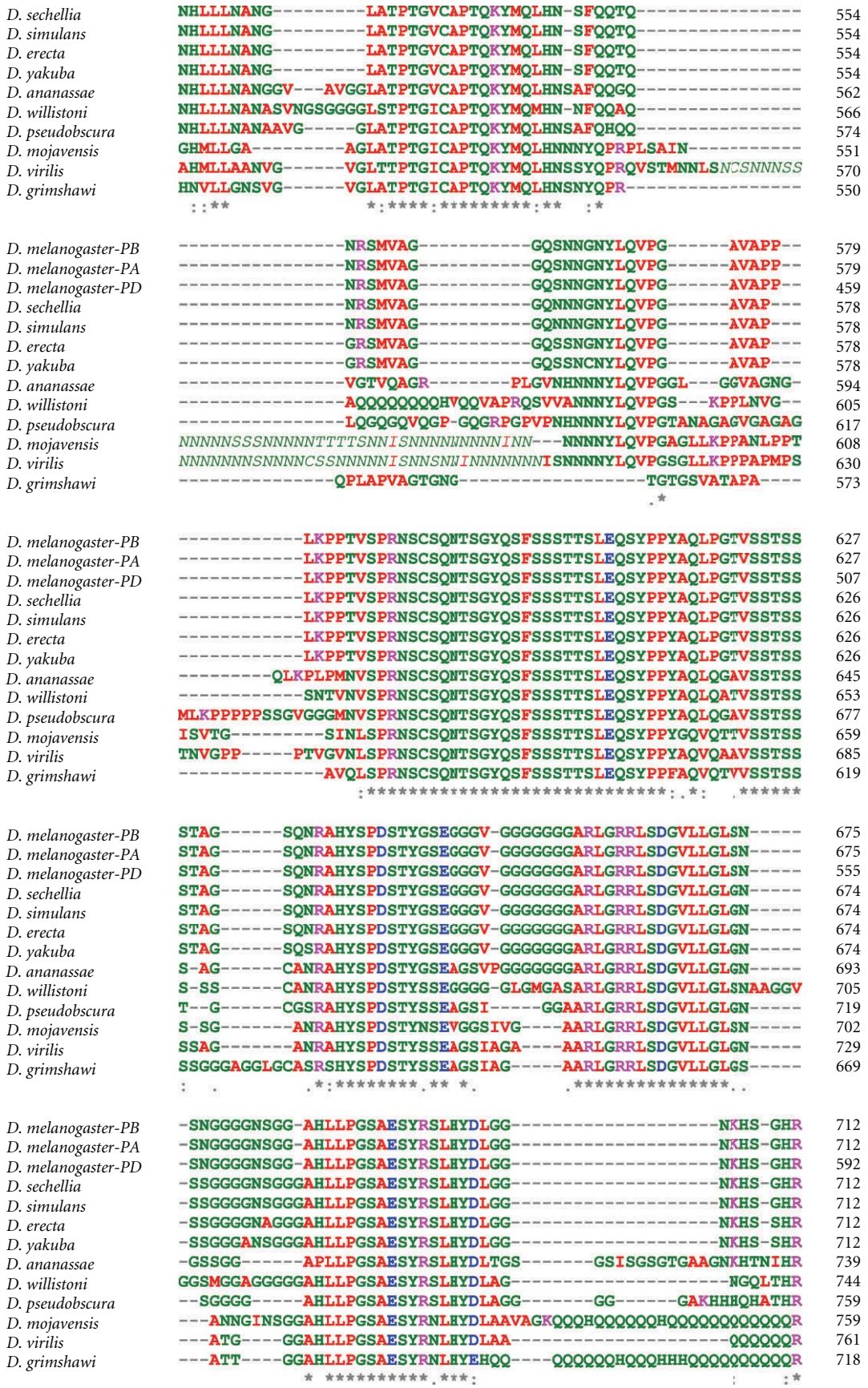


Figure 1: Continued.

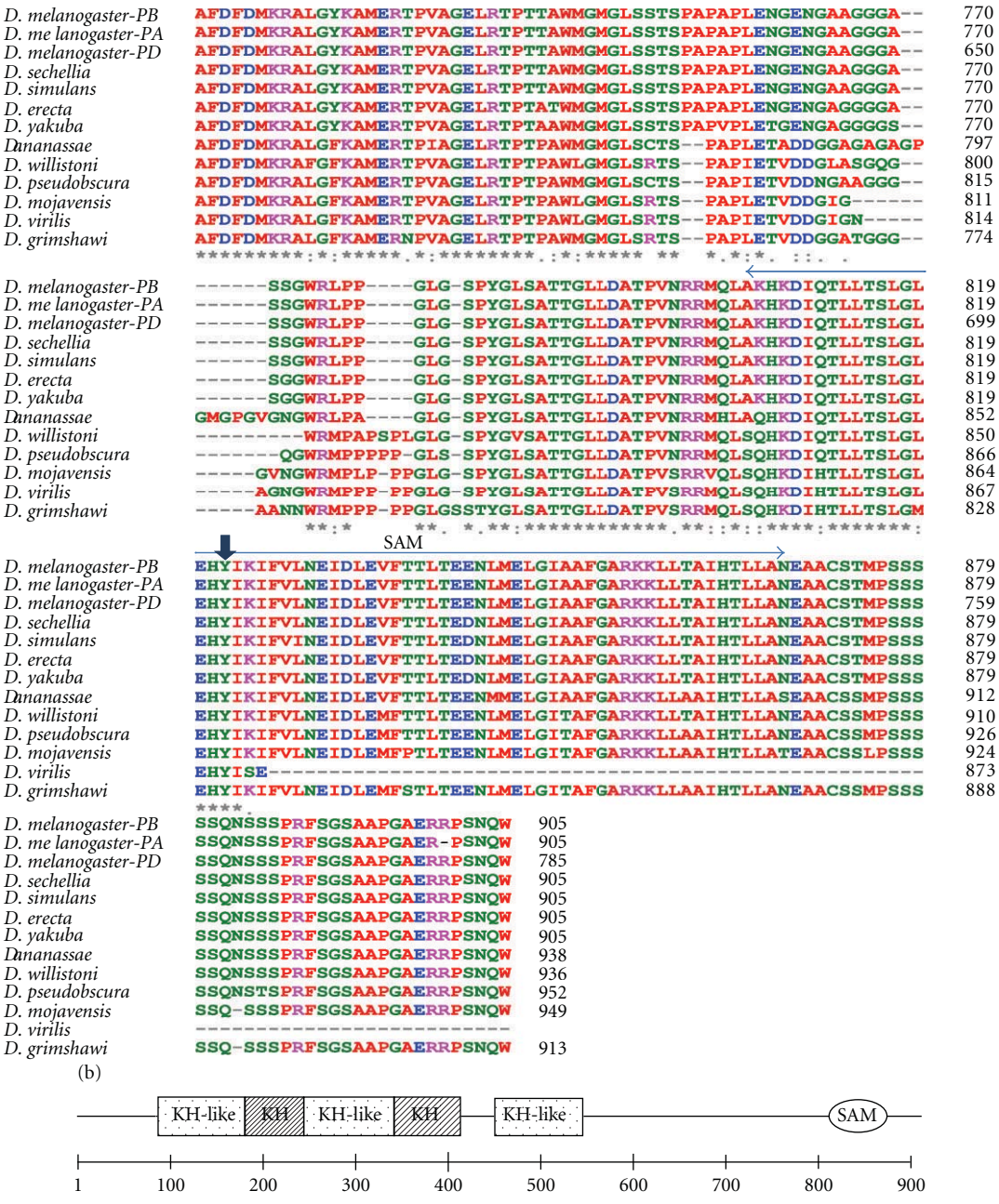


Figure 1: (a) Alignment of Bic-C sequences from 11 *Drosophila* species. Clustal W [1, 2] was used to align sequences extracted from FlyBase. Amino acid (aa) color coding is from Clustal W: red, small aliphatic, hydrophobic, and aromatics; blue, acidic; magenta, basic; green, hydroxyl, sulphydryl, amine, and glycine; grey, unusual aa. Symbols for aa conservation are from Clustal W: (asterisk ∗): positions with a single, fully conserved residue. (Colon :): conservation between groups of strongly similar properties scoring >0.5 in the Gonnet PAM 250 matrix. (Period .): conservation between groups of weakly similar properties scoring ≤0.5 in the Gonnet PAM 250 matrix. All three *D. melanogaster* Bic-C isoforms are shown (*PA*, *PB*, *PD*). The two canonical (KH) and three noncanonical (KH-like) KH RNA-binding modules are indicated (arrows, top). Domain assignment is as in [3] except for the fourth KH-related motif and the SAM domains, that are labelled according to the Pfam database [4]. A conserved, potentially phosphorylated, tyrosine is also indicated (arrowhead, top). Divergence occurs in regions of low complexity in the encoding DNA. Relative to the numbering of the *Drosophila* sequence: insertion at 555, variable length of the serine stretches around aa 623, and between aa 647–658 in the serine-glycine rich region. Further, after aa 715 there seems to be insertions of glutamine stretches of various lengths in *D. mojavensis*, *D. virilise*, and *D. grimshawi*. Finally, *D. ananassae* shows a short insertion at aa 770. The *D. virilise,* sequence results truncated. A TBLASTn search with the C-terminal region of Bic-C from *D. melanogaster* reveals many ESTs with similarity to the *D. melanogaster* sequence, suggesting a possible misannotation (not shown). Another region of possible sequencing misannotation in the *D. virilis* and the *D. mojavensis* Bic-C is italicized and not in bold type. Note that the *Bic-C* gene in *D. melanogaster* has nine mapped introns [5], and there is the possibility that the sequence was misannotated with this respect. (b) Block structure of the *D. melanogaster* Bic-C highlighting the protein motifs described in the text.

low complexity, where there are stretches of adjacent identical amino acids whose number varies in different species, the possible result of evolutionary mechanisms acting on triplet repeats or of stuttering sequencing polymerases (Figure 1).

An alignment of Bic-C orthologs from different animals reveals extensive sequence conservation from aa 83 to 268, (referring to the *Drosophila* sequence). Between aa 269 and 303, the vertebrate proteins lack the acidic residues present in the two Dipterans (*D. melanogaster and Anopheles gambiae*) while the basic residues between aa 281 and 286 are conserved (Figure 2).

Similarly, between aa 417 and 423 the acidic residues are exchanged with a basic (K) or a neutral (G) residues, while the adjacent phenylalanine 424 is changed conservatively into a tyrosine, suggesting that the overall protein folding may be preserved and that the electrostatic environment may be different between the insect and the vertebrate proteins. Since this region contains possible KH-domain-like modules, this may influence their ability to interact with RNA by contributing positive charges that might help retain or stabilize the interaction with RNA. At aa 458, vertebrate sequences diverge from those of *Drosophila*, *Anopheles*, and *Caenorhabditis elegans*. These sequences show blocks of conservation (aa 712–737 and 815–863) interspersed with regions of divergence and one insertion of 38 residues at aa 778. The SAM domain is one such block of conservation, with its phosphorylatable tyrosine [28] that is invariant in all the sequences analysed and the identity (or conservative substitution) of most of the amino acids that contribute to create an environment conducive to RNA binding in the case of Smaug [24].

3.3. Bic-C and Translational Regulation. Evidence that Bic-C was involved in control of mRNA translation came first from studies in *Drosophila* where it was observed that Oskar, a well-studied morphogen, was upregulated in ovaries from Bic-C mutated females [14]. The identification of other mRNA targets coimmunoprecipitated with Bic-C yielded the *Bic-C* mRNA itself and several mRNAs encoding factors involved in the Wnt pathway, vesicular trafficking, and organization of the actin cytoskeleton [15]. Bic-C interacts directly with the Not3/5 subunit of the CCR4 deadenylase complex, and it is believed that, when bound to its target RNA, it is able to recruit the deadenylase. This shifts the cellular balance between polyadenylation and deadenylation towards the latter, impairing translation [15]. Since Not3/5 is also evolutionarily conserved, it is discussed below in the perspective of its contribution to the Bic-C complexes.

The other invertebrate family member for which there is substantial functional information is the *C. elegans* GLD-3. GLD-3 is involved in germline development and embryogenesis by regulating the time of expression of developmental factors [23, 29, 30]. GLD-3, via its first KH domain, interacts with GLD-2, a noncanonical polyA polymerase devoid of an RNA interaction domain of its own [23, 30]. Although it was expected that GLD-3 may tether GLD-2 to the RNA, a recent structural study could not find any RNA binding activity for the GLD-3 KH region [12]; therefore further studies are needed to elucidate how GLD-3 participates to *C. elegans* development.

In the *Drosophila* ovary Bic-C is present in cytoplasmic granules enriched for Trailer Hitch (Tral) and Me31B [31, 32], two proteins marking sponge bodies, ovarian organelles related to the repression of mRNA translation [33–35]. Mouse and *Xenopus* Bicc1 in cultured cells are also found within subcellular structures associated with mRNA silencing, the processing granules (P granules, [26, 27, 36]), strongly suggesting that the members of the Bic-C protein family may share a conserved function in translational control. For example, P bodies may destabilize mRNAs via the action of decapping enzymes such as Dcp1 in many tissues undergoing rapid mRNA turnover, while certain yeast mRNAs can be reversibly associated with P-bodies [37]. Further, in metazoans, deadenylation is often the rate-limiting, first step of mRNA decay [38]. While in the kidney, high turnover of certain mRNA may be instrumental to rapidly adapt organ function to the environmental changes, in tissues with a strong "anabolic" activity such as the ovary it would not be surprising to find that some maternal mRNAs are silenced and stored in cellular compartments refractory to translation during oogenesis, to be deployed later in the early embryo. Consistent with the possibility that Bic-C may not function by destabilizing its mRNA targets, no global changes in *Bic-C* mRNA stability were observed in the *Drosophila* ovary, neither by quantitative RT-PCR of ovarian total mRNA nor by *in situ* hybridization (Bic-C negatively regulates its own mRNA) [15]. While there seems to be a mild effect on stability of the polycystic kidney disease 2 (*Pkd2*) mRNA in the kidneys of the *Bicc1*$^{-/-}$ KO mice, in this case, no direct association of this mRNA with the Bicc1 protein was formally demonstrated [27]. It is also possible that only a fraction of the cellular Bic-C pool is involved in destabilization and degradation of mRNA targets, possibly constituting a distinct compartment. This scenario would have escaped detection via traditional biochemical methods because they cannot preserve the integrity of the tissues analyzed. Until more regulatory targets for the Bic-C family members will be identified, validated, and characterized functionally, this current puzzle will remain unanswered.

3.4. Not3/5: An Evolutionarily Conserved Bic-C Partner Affecting mRNA Translation. Not3 is one of the subunits of the CCR4-NOT deadenylase, which is the predominant deadenylase, at least in the yeast *S. cerevisiae* [39–41]. Other subunits include CCR4, CAF1, NOT1-5 [40–44]. In *Drosophila* homologous genes are present for each of these subunits, with the exception of NOT3 and NOT5, for which there is only one gene displaying homology to both proteins [45]. Interestingly, Not3/5 does not contain any known protein domain, as identified via Prosite [46].

Drosophila Not3/5 proteins are virtually identical in 12 species, the differences being concentrated in areas of low-sequence complexity (Figure 3). A BLAST search [7] reveals that besides insects and vertebrates, there are Not3/5 orthologs, in *fungi* (*S. cerevisiae*, *Schizosaccharomyces pombe*, as well as the mushrooms *Laccaria bicolor*, *Coprinopsis*

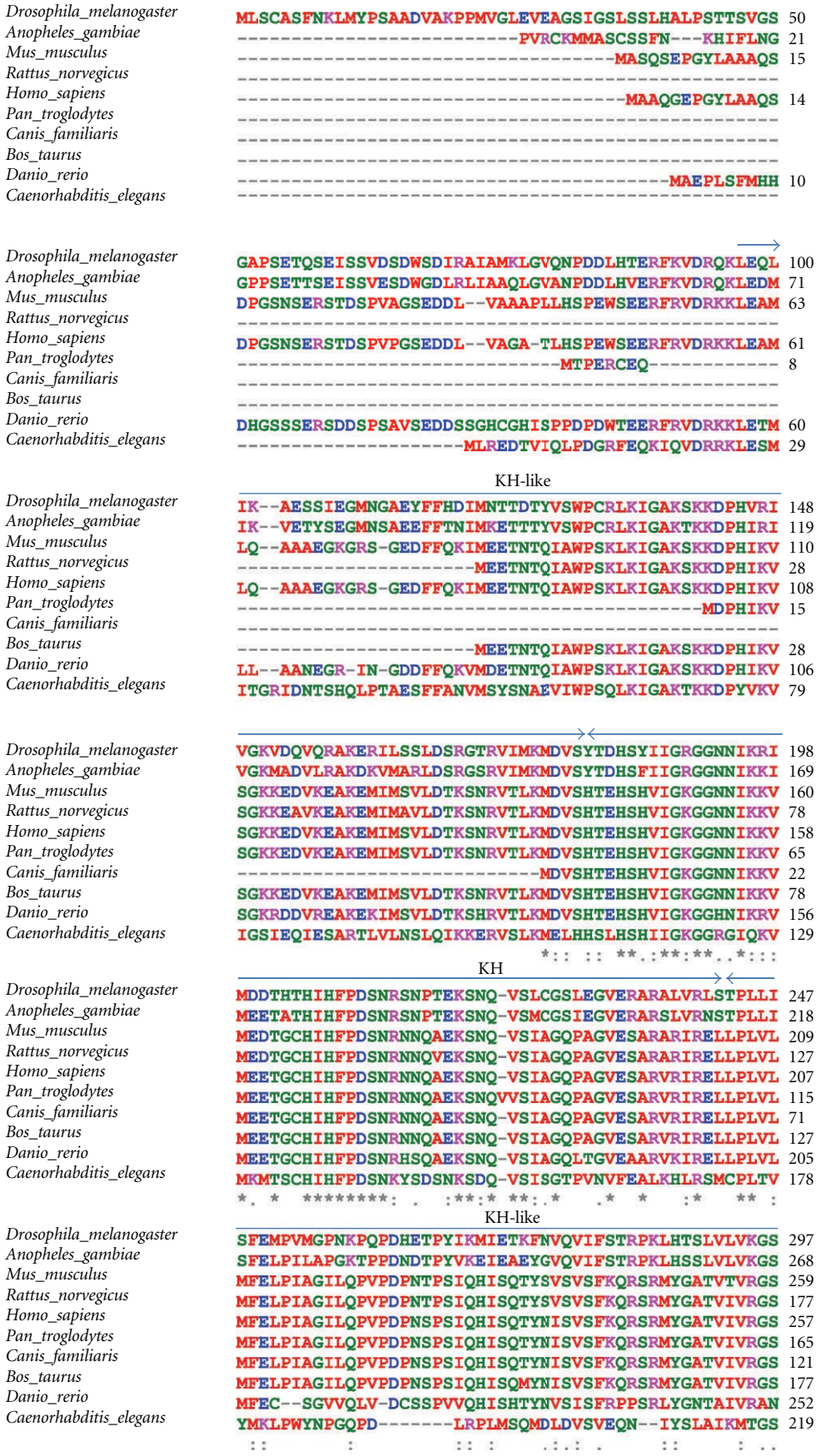


Figure 2: Continued.

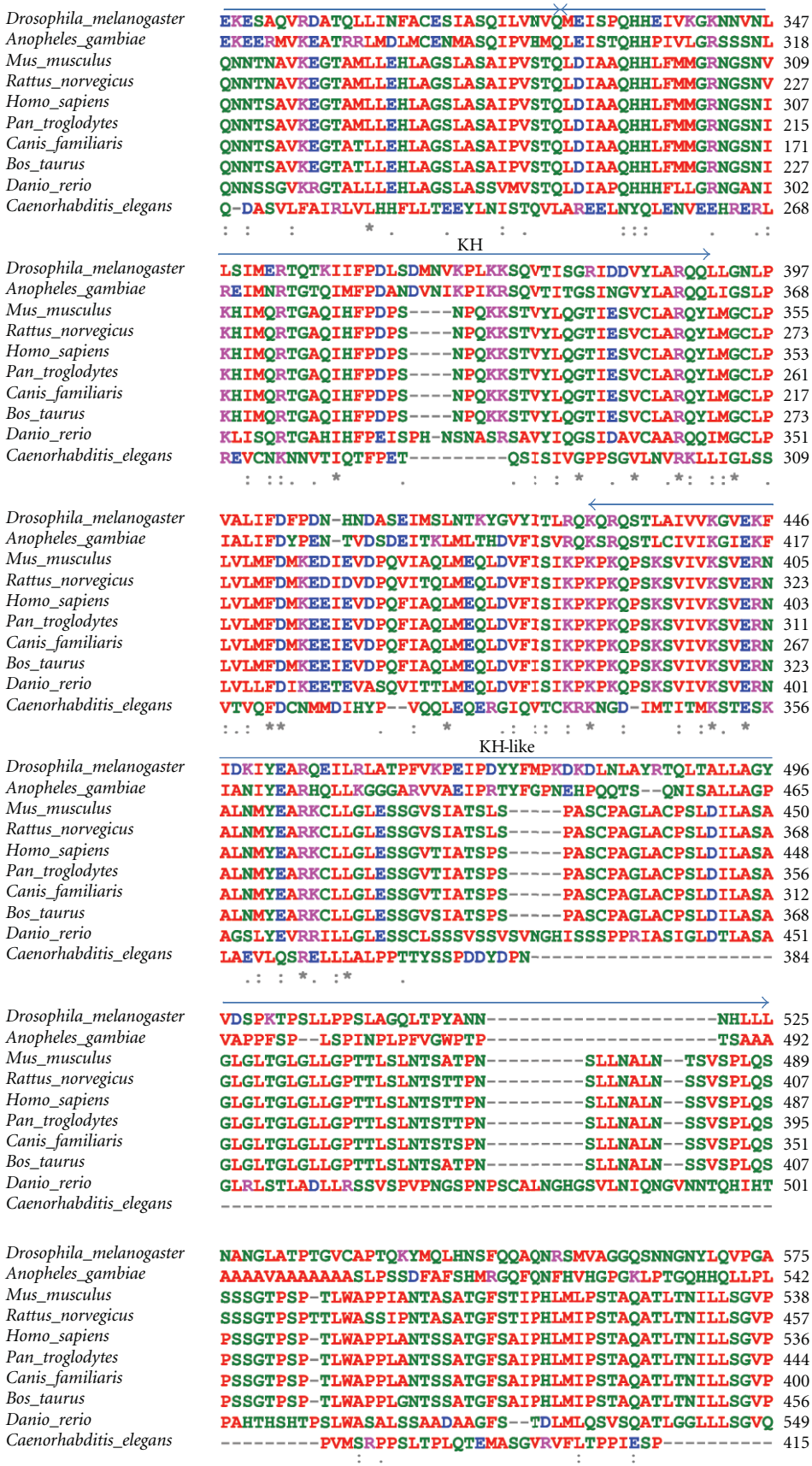


Figure 2: Continued.

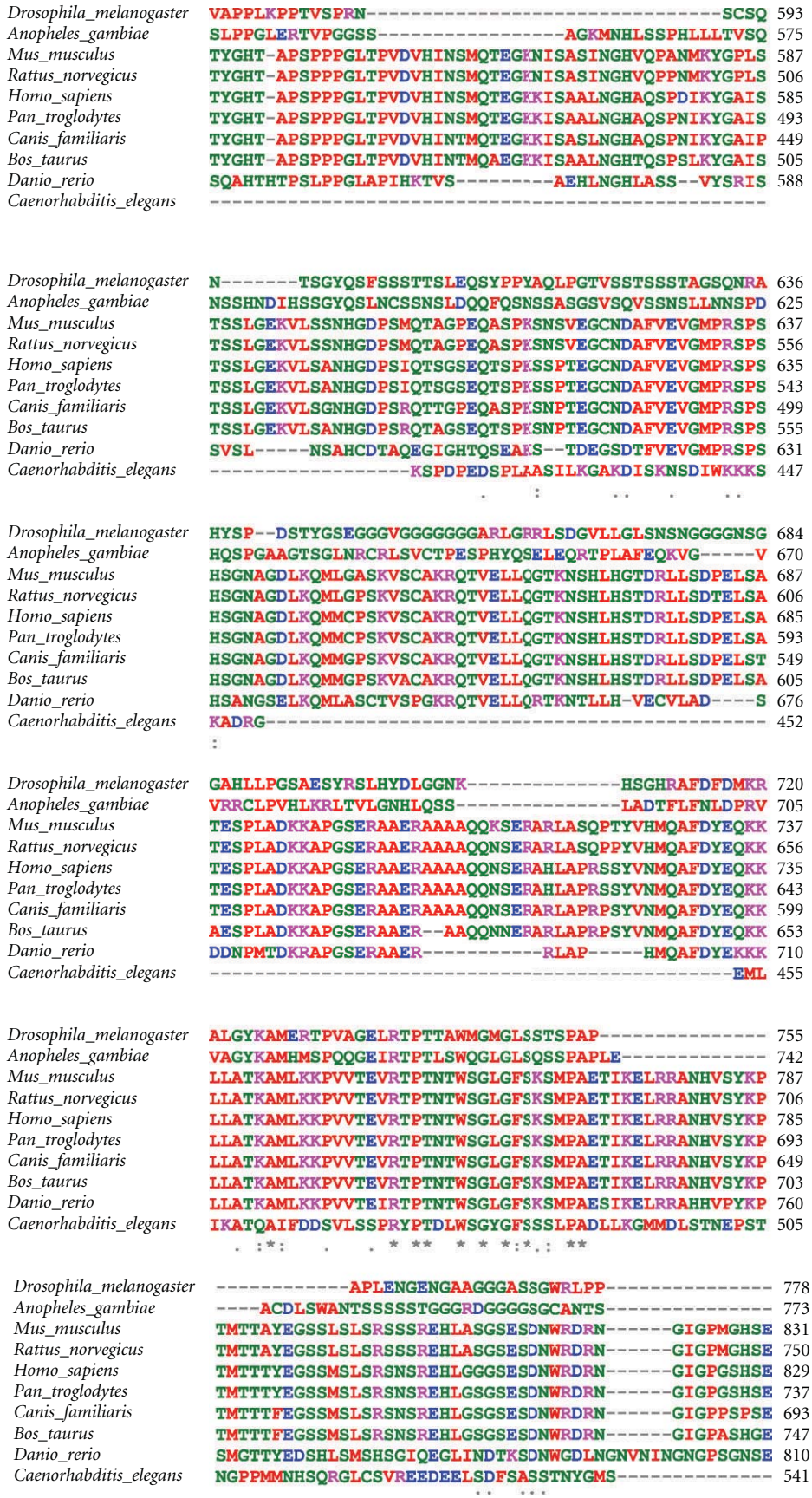


Figure 2: Continued.

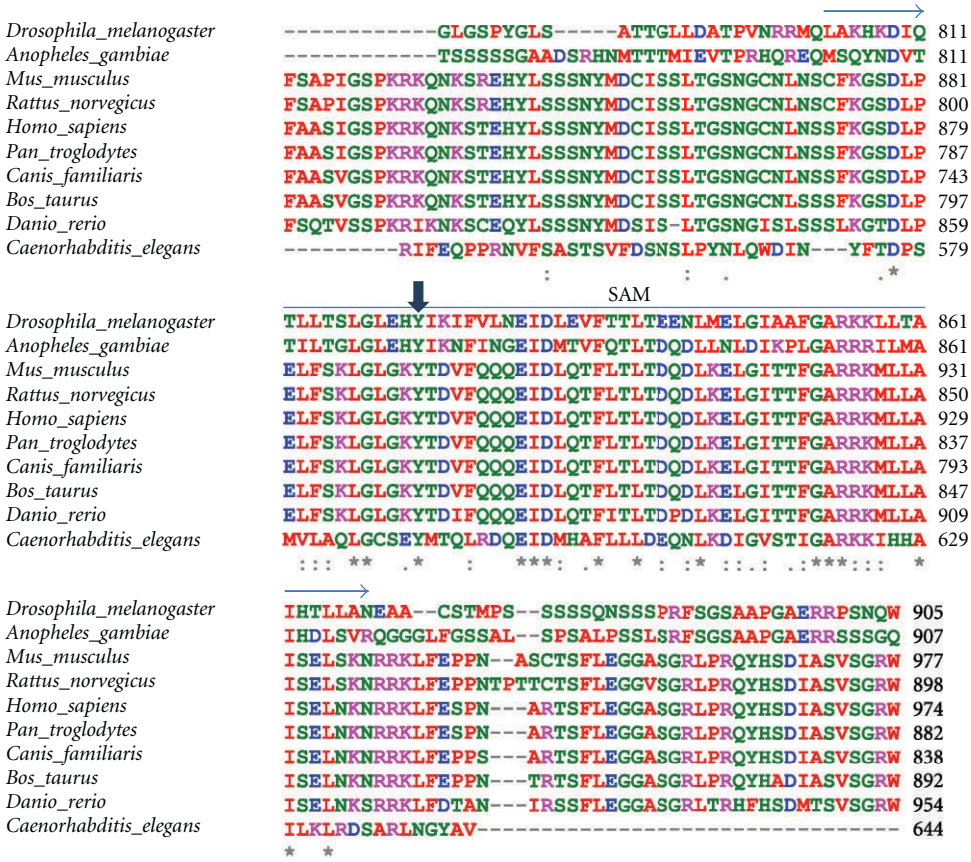


Figure 2: Bic-C orthologs. Clustal W [1, 2] was used to align sequences extracted from the NCBI sequence database. As in Figure 1, the two canonical (KH) and three noncanonical (KH-like) KH RNA-binding modules are indicated (arrows, top). Domain assignment is as in [3] except for the fourth KH-related motif and the SAM domains, that are labelled according to the Pfam database [4]. A conserved, potentially phosphorylated, tyrosine is also indicated (arrowhead, top). Amino acid (aa) color-coding is from Clustal W: red, small aliphatic, hydrophobic and aromatics; blue, acidic; magenta, basic; green, hydroxyl, sulphydryl, amine, and glycine; grey, unusual aa. Symbols for aa conservation are from Clustal W: (asterisk $*$): positions with a single, fully conserved residue. (Colon :): conservation between groups of strongly similar properties-scoring >0.5 in the Gonnet PAM 250 matrix. (Period .): conservation between groups of weakly similar properties-scoring ≤0.5 in the Gonnet PAM 250 matrix. Highlighted yellow: residues that contribute to RNA binding in the Smaug protein. Grey highlight denotes mild (versus strong) basic charges. Light blue highlights a charged aa in a conserved position, but an opposite electrical charge. The *Gallus gallus* genome also contains a predicted sequence with extensive homology to Bic-C (Table 1) and with a long extension at the N terminal end. Since there is no experimental evidence of the true starting methionine we did not include it in this alignment.

cinerea, and *Schizophyllum commune*), vascular plants (e.g., *Arabidopsis thaliana*, *Oryza sativa*, *Vitis vinifera*, and *Ustilago maydis*), and mosses (*Physcomitrella patens*). An alignment of complete sequences is shown in Figure 3. The conservation is highest at the N- and C-terminus of the protein (aa 1–238 and 680–844, with reference to the *Drosophila* sequence) where all the family members show extensive identity. Between aa 330 and 679 the sequences diverge with the orthologues from the two insects (*D. melanogaster* and *Anopheles gambiae*), the *fungi*, the higher eukaryotes, and the plants being more similar with each other than with members of a different group. Notably, the vertebrate sequences, with the exception of zebrafish that contains various small deletions, have blocks of almost complete identity in this region (Figure 4). The partial divergence in the central region of Not3/5 is likely due to the fact that the *Drosophila* gene is homologous to both the NOT3 and NOT5 genes and likely plays the functional roles of both yeast proteins, [45] a seemingly unique feature of *Drosophila [47]*. Not3/5 was recovered in a two-hybrid screen for proteins interacting with *Drosophila* Bic-C, and multiple pieces of evidence support the existence of this interaction *in vivo*: there is genetic interaction between *Bic-C* and *twin*, the *Drosophila* gene encoding for CCR4; other subunits of the CCR4-NOT complex can be coimmunoprecipitated with Bic-C from ovary extracts and the Bic-C target mRNAs that were tested were found with longer polyA tails in *Bic-C* mutants [15]. Although one study of vertebrate models could not detect differences in polyadenylation in a presumptive Bic-C target [27], due to the high homology of the Bic-C and NOT orthologs it is possible that Bic-C from other species can interact with NOT homologs and, possibly, other subunits of the deadenylase complex. These may contribute to the interaction only in the context of the assembled complex and

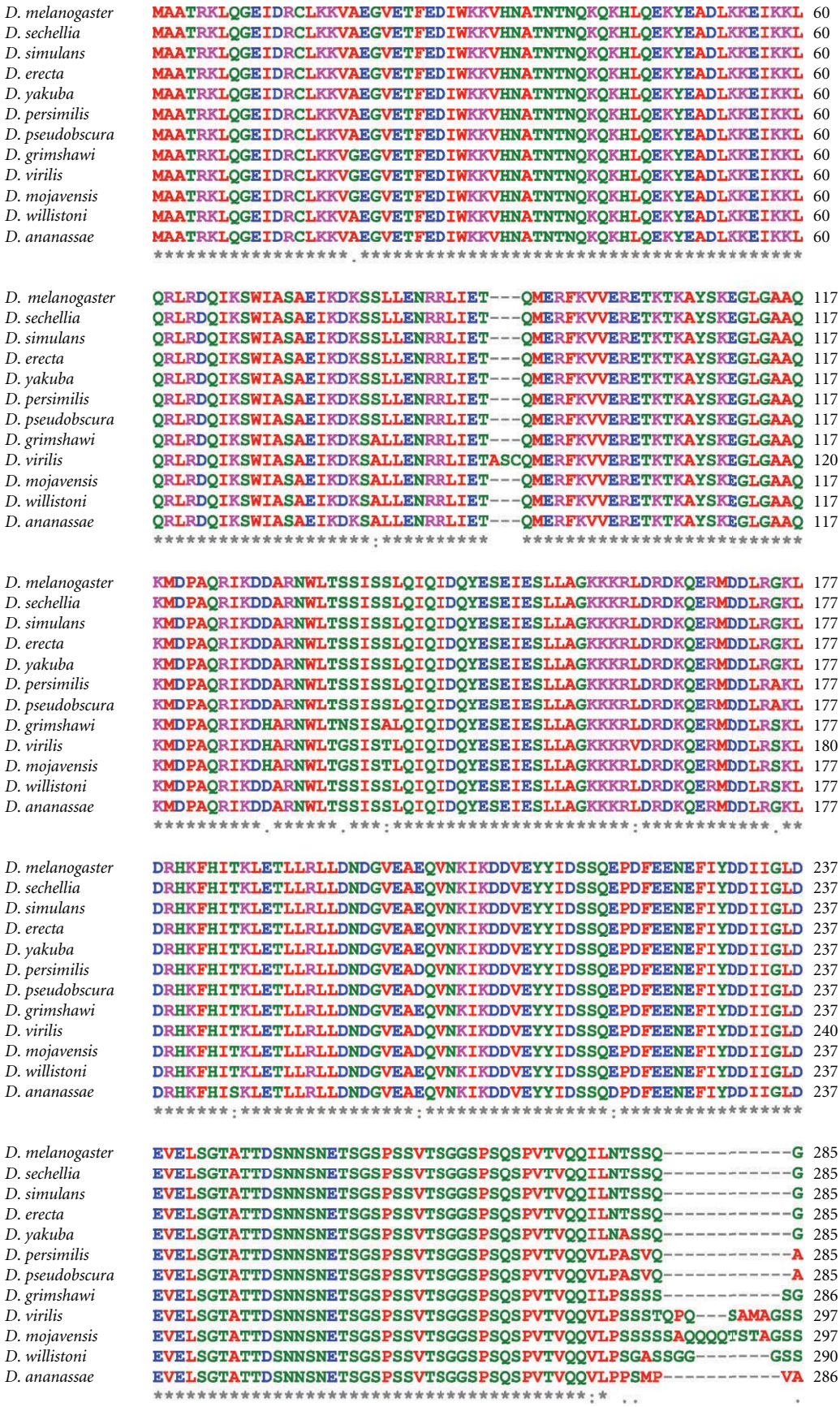


Figure 3: Continued.

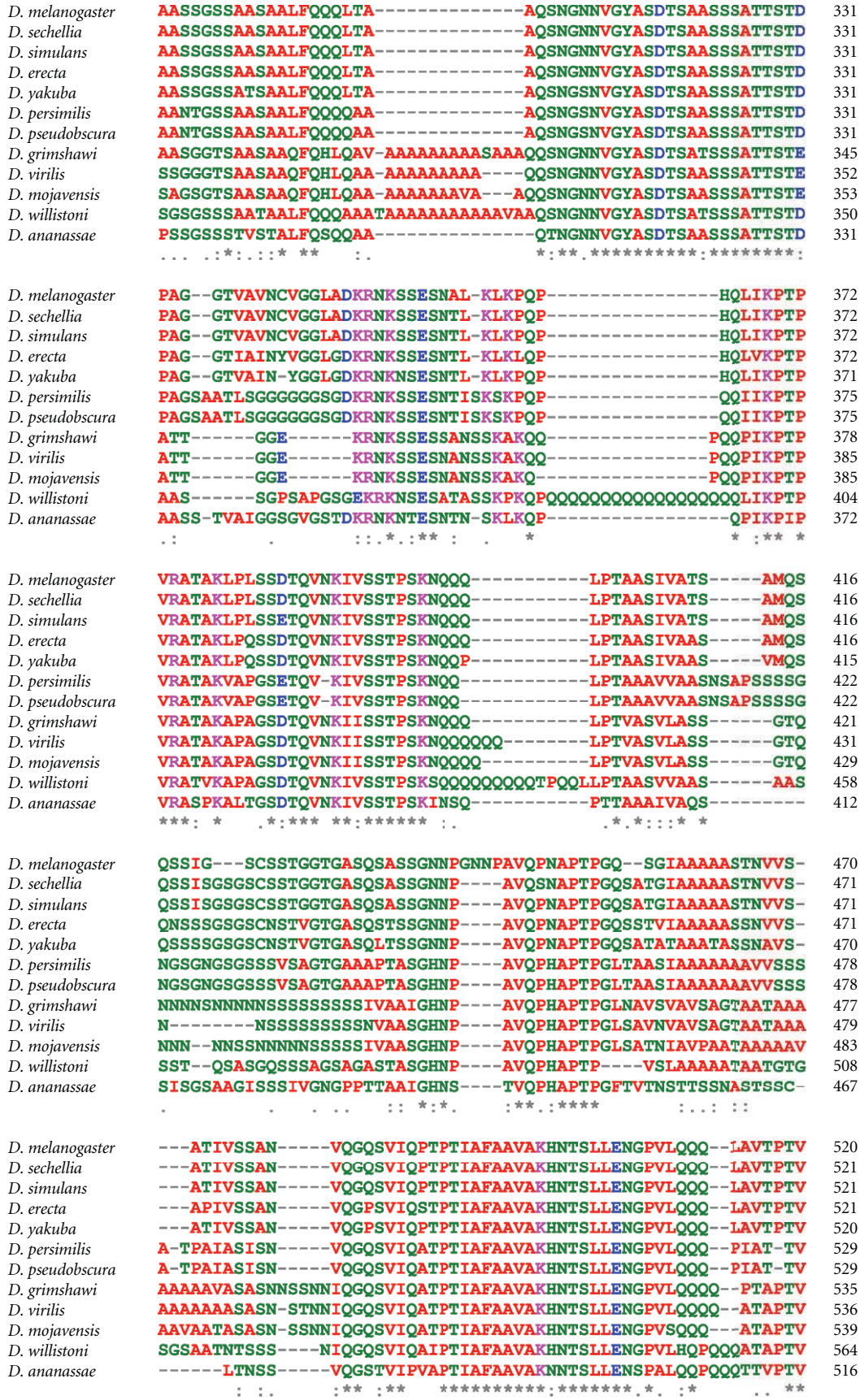


Figure 3: Continued.

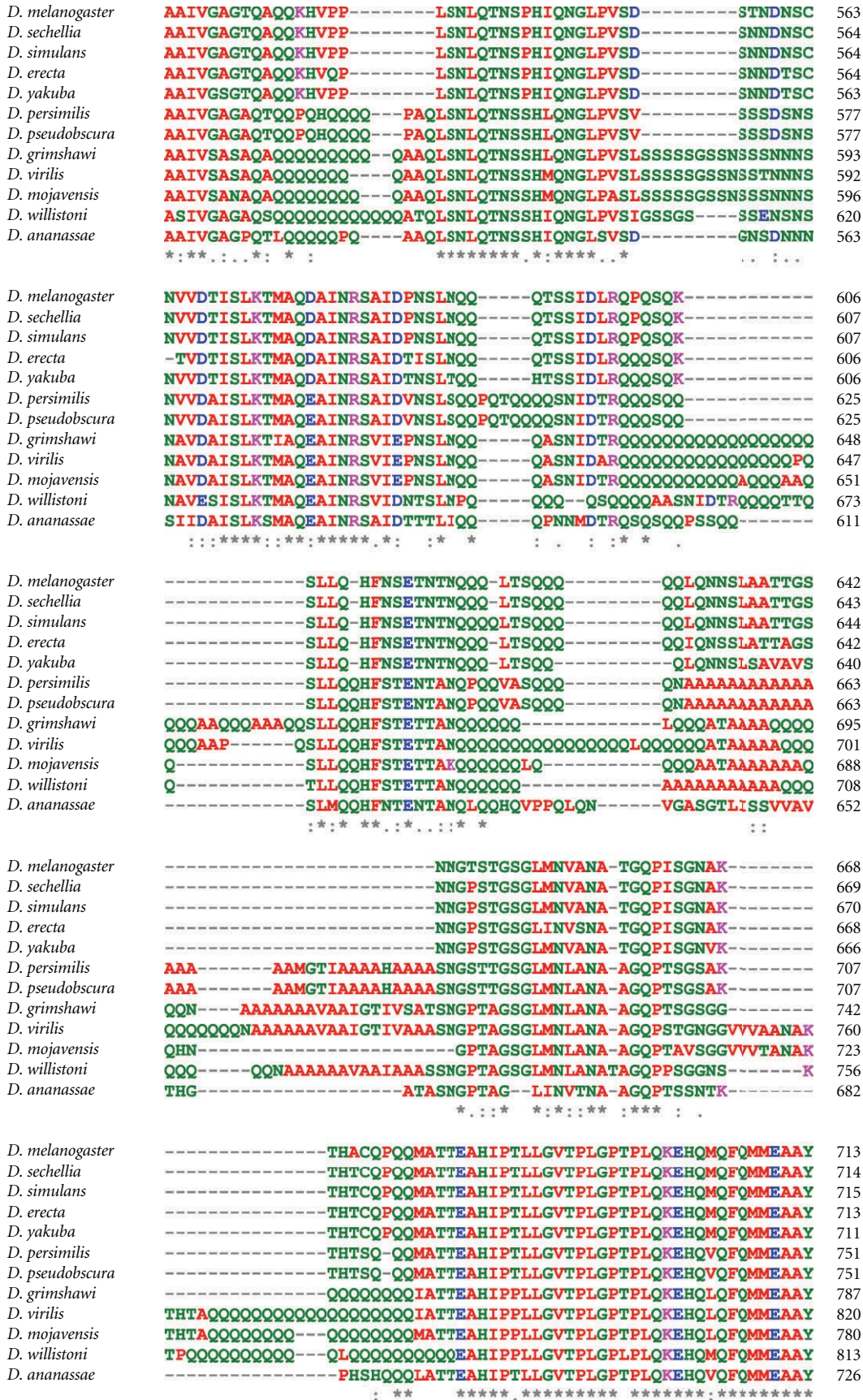


Figure 3: Continued.

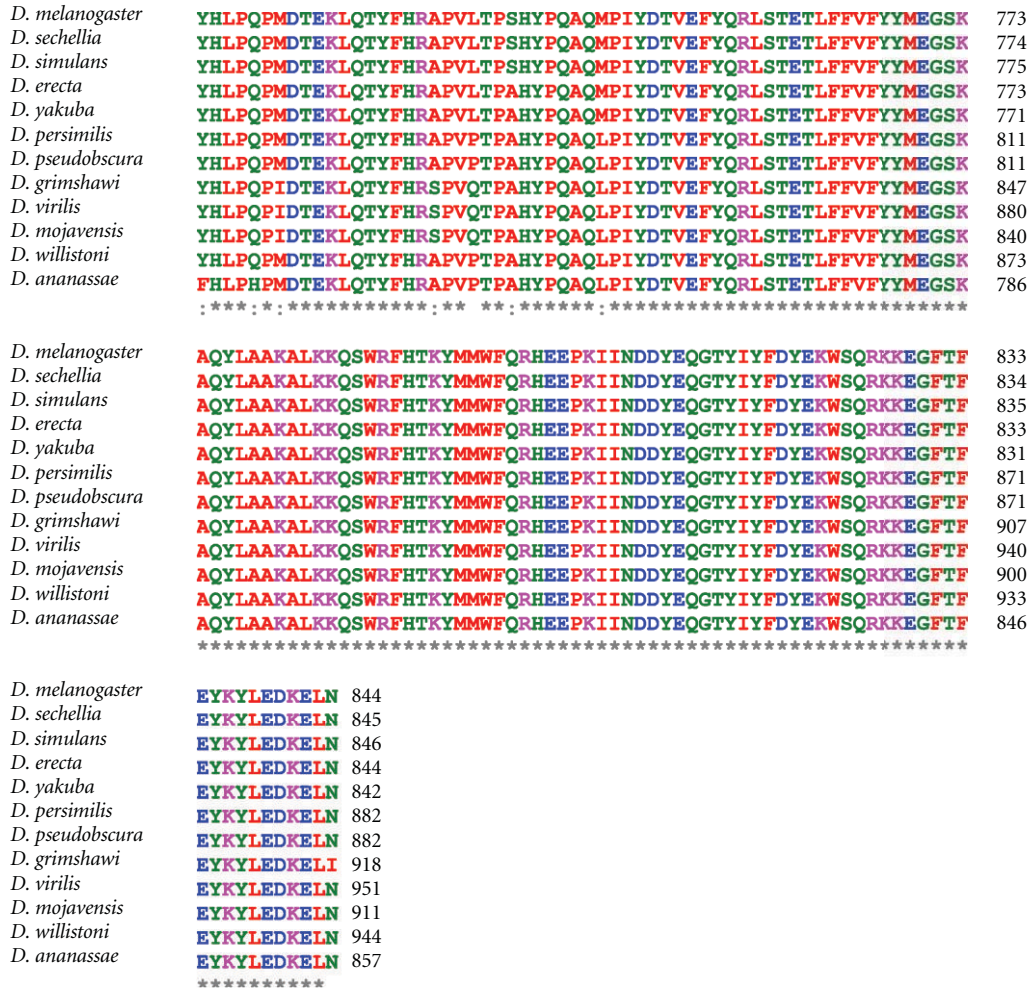


Figure 3: Not3/5 homologs from 12 *Drosophila* species. Clustal W [1, 2] was used to align sequences extracted from FlyBase. Amino acid (aa) color-coding is from Clustal W: red, small aliphatic, hydrophobic, and aromatics; blue, acidic; magenta, basic; green, hydroxyl, sulphydryl, amine, and glycine; grey, unusual aa. Symbols for aa conservation are from Clustal W: (asterisk $*$): positions with a single, fully conserved residue. (Colon :): conservation between groups of strongly similar properties-scoring >0.5 in the Gonnet PAM 250 matrix. (Period .): conservation between groups of weakly similar properties-scoring ≤0.5 in the Gonnet PAM 250 matrix.

may have therefore escaped detection in the *Drosophila* two-hybrid screen. Coimmunoprecipitation studies from tissue extracts and the precise mapping of the interaction domains on both proteins will be required to resolve this issue.

3.5. Multiple Bic-C Isoforms. *Drosophila* Bic-C has three predicted mRNA isoforms, RA, RB, and RD, that encode two identical (RA and RB) and one shorter (RD) proteins lacking the first 120 aa (Figure 1). These mRNA isoforms are expressed at different times during development (FlyBase): *Bic-C-RA* is expressed in the early embryo (0–6 hrs old) and in the adult female (i.e., most likely in the ovary), and *Bic-C-RB* is found mostly in late embryogenesis (7–22 hrs old). This is also consistent with our earlier report of multiple protein isoforms [14]. During the larval phases *Bic-C* is undetectable, and during pupation *Bic-C* expression is resumed, with its *RD* isoform being the most abundant and remaining prominent in adult males (FlyBase). The presence of two distinct mRNAs encoding the same amino acid sequence at definite developmental stages also suggests the possibility that they may be subjected to distinct regulation(s) in different tissues or at different developmental times and that the Bic-C activity may be required in specific time windows. This is consistent with a report that *Bic-C* function is especially needed at embryonic day (*E*) 18.5 during mouse development [27].

Interestingly, the mouse *Bicc1* gene and human *BICC1* also produce two distinct mRNAs by alternative splicing, which differ for the presence of exon 21 [26, 48] although no further functional information is known to date, so it is difficult to speculate if the presence of multiple Bic-C isoforms has conserved functional roles.

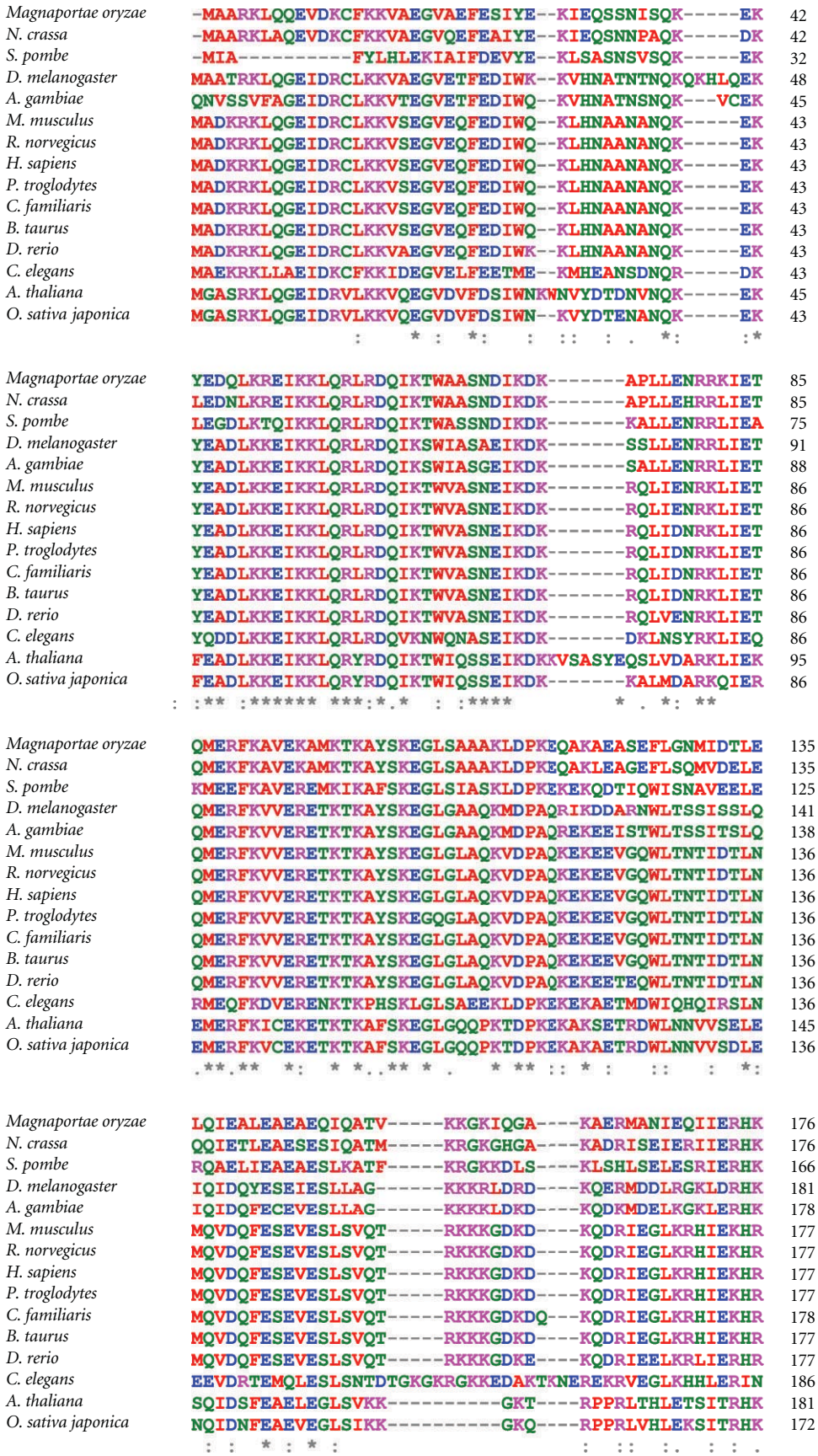


Figure 4: Continued.

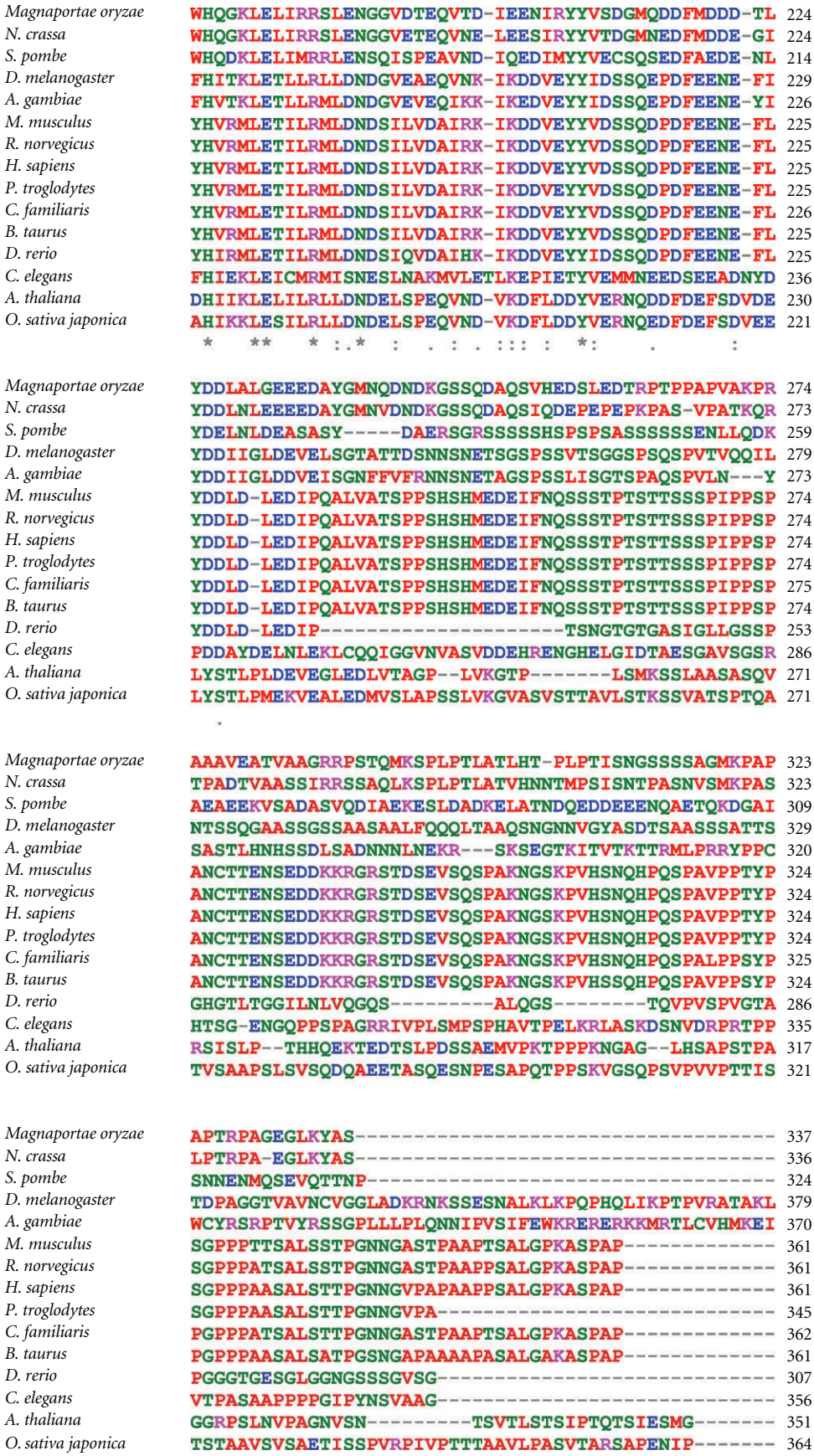


Figure 4: Continued.

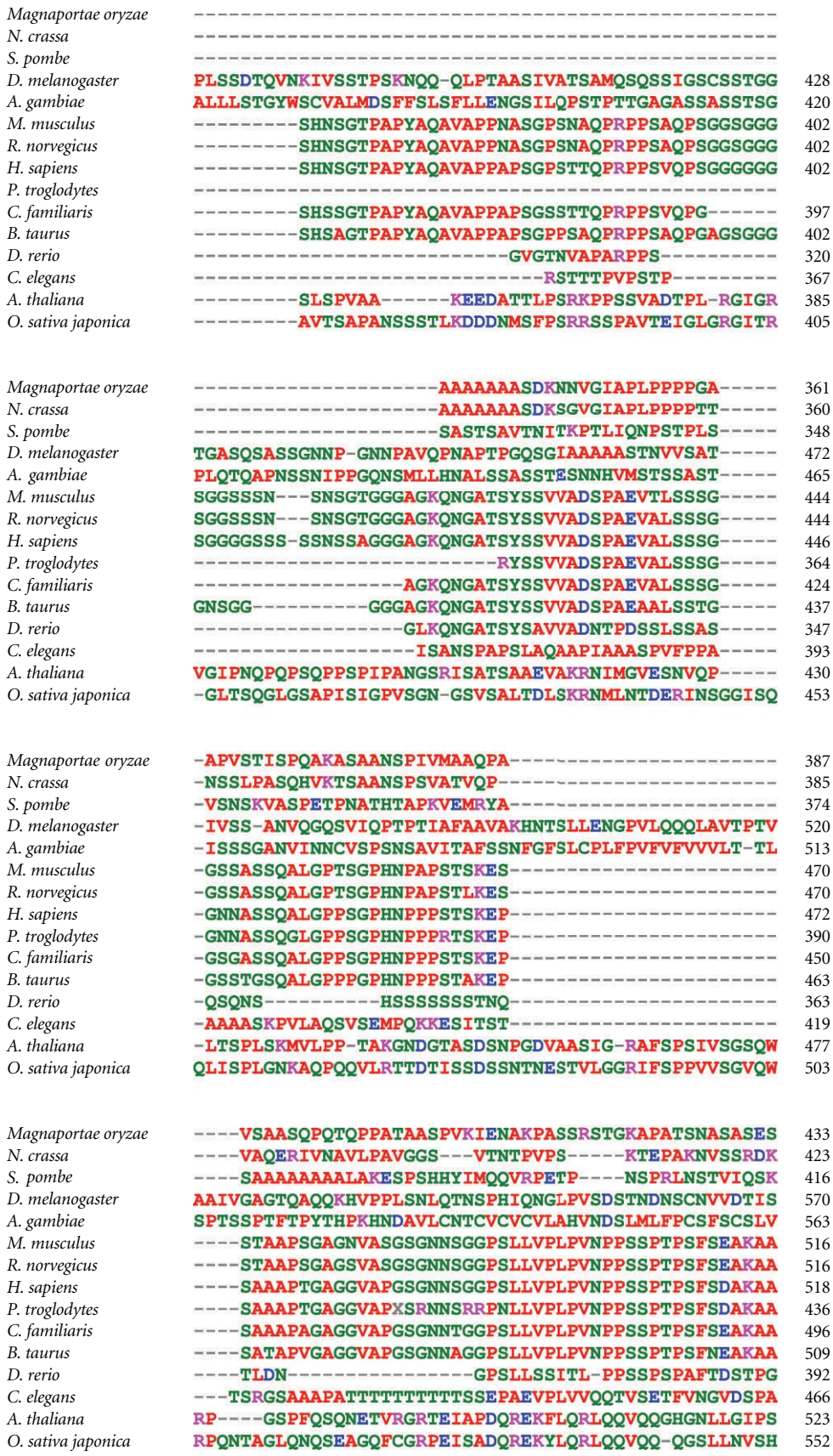


Figure 4: Continued.

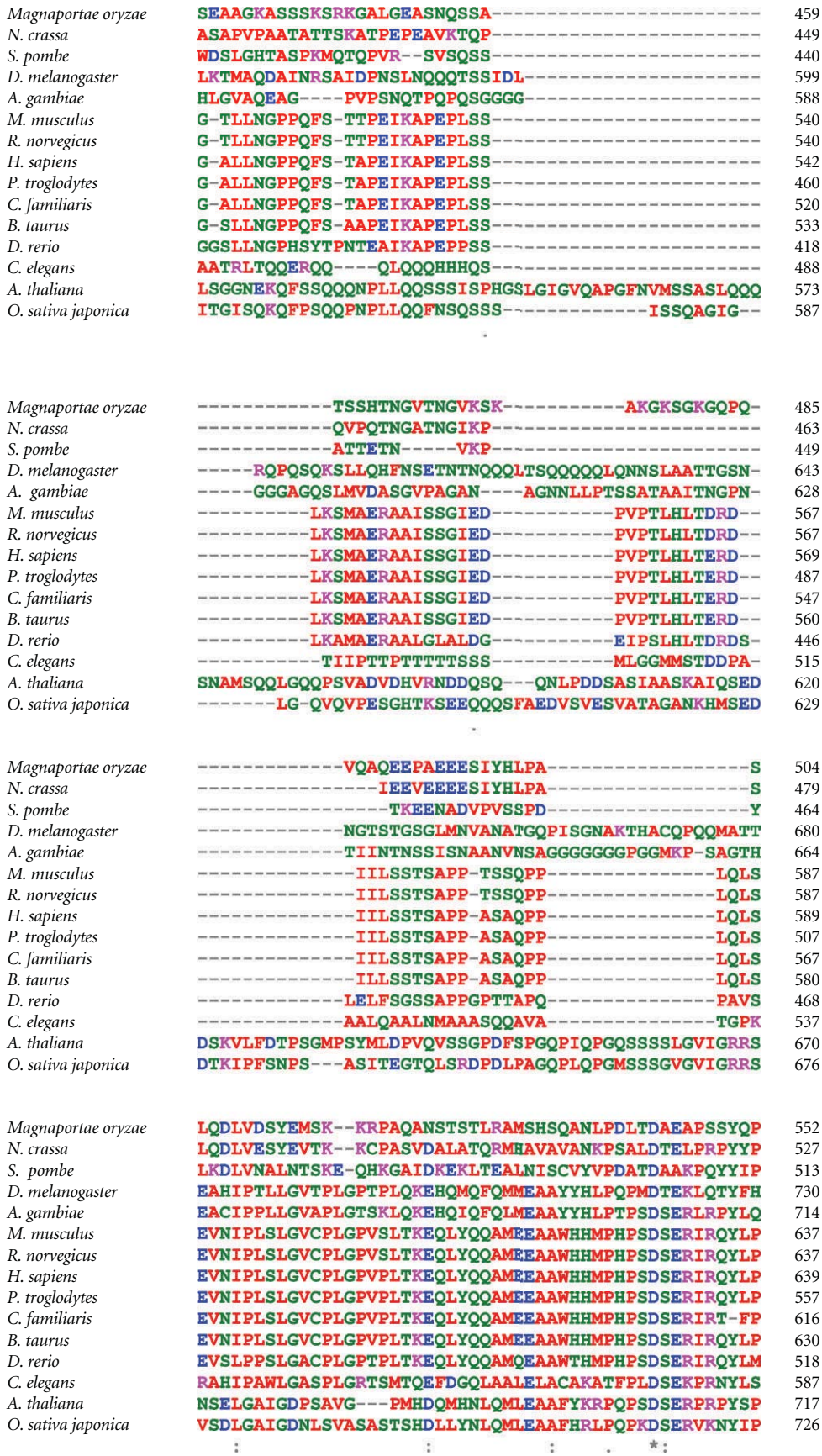


Figure 4: Continued.

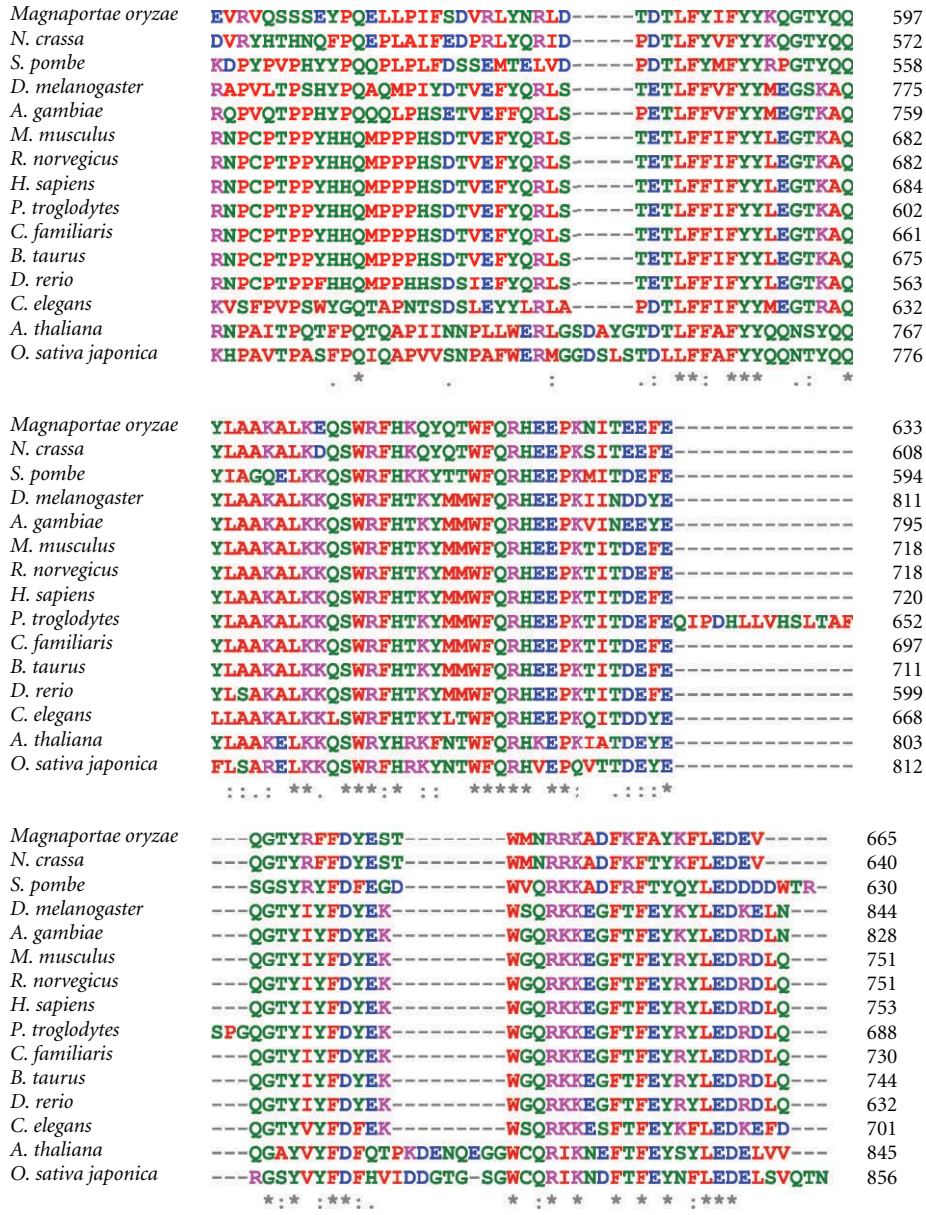


Figure 4: Not3/5 orthologs. Clustal W [1, 2] was used to align sequences extracted from the NCBI database. Amino acid (aa) color-coding is from Clustal W: red, small aliphatic, hydrophobic, and aromatics; blue, acidic; magenta, basic; green, hydroxyl, sulphydryl, amine, and glycine; grey, unusual aa. Symbols for aa conservation are from Clustal W: (asterisk $*$): positions with a single, fully conserved residue. (Colon :): conservation between groups of strongly similar properties-scoring >0.5 in the Gonnet PAM 250 matrix. (Period .): conservation between groups of weakly similar properties-scoring $\leq$0.5 in the Gonnet PAM 250 matrix.

3.6. Bic-C and Polycystic Kidney Disease. In humans, two polycystic kidney disease (PKD) forms are caused by mutations in the *PKD1* and *PKD2* genes (autosomal, dominant [49–55]) or in PKHD1 (autosomal recessive, [52–55]). The link between Bic-C malfunction and PKD is compelling: two mouse models developing polycystic kidneys harbor mutations of the *Bicc1* gene [56]; Bic-C inactivation in *Xenopus* induces cystic kidneys [27, 57]; recently, a zebrafish model of PKD was validated that inhibits the Bicc1 function [58]. Finally, human studies on patients with renal disorders identified two mutations associated with the *BICC1* gene: one affecting the first KH domain and the other affecting the SAM domain [48], proving the relevance of the *Bic-C* animal models for understanding the etiology of this incurable disease.

In 3D cultures of mouse IMCD cells, depleting Bicc1 disrupts cadherin-mediated cell adhesion, normal epithelial polarization, proliferation, and apoptosis that prevent tubulomorphogenesis *in vitro* [59]. Interestingly, aspects of the *Drosophila* phenotype also affect cell migration and may influence cell-cell interaction and polarization. For example, migration of the follicle cells (FCs) in the ovary is defective

in Bic-C mutant [3], resulting in eggs that remain open at the anterior end. This defect may occur because of inefficient communication between germ line and somatic cells, although to date we do not know the molecular pathway underlying this phenomenon (for an alternate possibility, see also Section 3.7).

In a recent paper [27] Tran and colleagues report that in a novel *Bicc1*$^{-/-}$ mutant mice and in *Xenopus* depleted for Bicc1 the *Pkd2* mRNA and its cognate protein are downregulated (29 and 54%, resp.), while both *Pkd1* and *Pdhd1* levels are unaffected. In the mouse these effects are clearest specifically at stage E18.5. The regulation appears to be mediated via a cellular microRNA, *miR-17* [27] that is also amplified in certain cancers [60]. Here Bicc1 may relieve the *miR-17*-mediated repression via a mechanism that does not involve regulation of the polyadenylation state of at least the mRNAs tested and may mildly impact mRNA stability [27]. The fact that the Bicc1 protein may bind multiple mRNAs and that it may be involved in the possible antagonistic regulation of the *miR-17* complexes, also assembled on multiple mRNAs, reinforces the view that the Bic-C orthologs are central to the regulation of many cellular processes and that many more aspects of their function await elucidation.

3.7. Other Bic-C Functions. Another hint to Bic-C function comes again from *Drosophila*, where the *Bic-C* mutants exhibit disrupted pattern of the cortical filamentous actin in the growing oocyte and abnormal actin-containing structures in the ooplasm that trap both the dorsal fate determinant Gurken [61–63] and other proteins that would normally be secreted [31, 32]. This function requires Trailer hitch, a protein originally identified in a screen for mutants for axial polarity that may regulate expression of endoplasmic reticulum (ER) exit site components on the ER surface. A malfunctioning secretory pathway could affect communication between the oocyte and the overlying FC and may affect their migration. Since many mRNAs involved in vesicular trafficking and/or organization of the actin cytoskeleton were also recovered in Bic-C immunoprecipitates [15], it is possible that their posttranscriptional control may contribute to the observed Bic-C defects. Lastly, and not mutually exclusive, the altered actin dynamics exhibited by the *Bic-C* and *Tral* mutants must also add to the observed inhibition of the normal dumping of nurse cell contents into the nascent oocyte during late oogenesis.

4. Concluding Remarks

Bic-C is an ancient protein conserved from *Drosophila* to man. Its mutation induces a pleiotropic phenotype. In fruit flies the Bic-C protein binds to RNAs involved in establishing the embryonic polarity, the Wnt pathway, actin dynamics and results in many observed defects, including abnormal development. In the vertebrates the better characterized aspect of lack of Bic-C function is the induction of cystic kidneys and the alteration of cell proliferation and three dimensional organization; however, defects in pancreatic and liver function and heterotaxia (i.e., randomization of the left-right symmetry) of the visceral organs have also been observed [26, 27]. Further, effects on the Wnt pathway have also been reported in human patients with renal displasia [48], as well as in mice and frogs [26]. Bicc1 is also expressed in the nervous system [58] which suggests that there may be novel aspects of its function ready to be discovered and that Bic-C homologs may be involved in fundamental, evolutionarily conserved mechanisms of determination of polarity, from establishment of the body axes to planar cell polarity.

The experimental evidence so far also suggests that Bic-C function may also be required at specific times of development in many species. Since Bic-C is a negative regulator of translation, we can expect at least part of the mutant phenotypes to be linked with inappropriate spatial and/or temporal regulation of gene expression. Further, Bic-C has multiple mRNA targets, and it exists in multiple isoforms in many organisms. At least in the case of one of the Bic-C interacting partners, the CCR4 deadenylase, it is proposed that multiple forms of this complex exist in higher vertebrates [47], as there are documented isoforms for a few of the complex subunits. Therefore, it is possible that the Bic-C-CCR4-dependent regulation acts via and is regulated by combinatorial mechanisms, with variant complexes having partially redundant function. This could also explain why all the individual molecular effects/phenotypes described for Bic-C tend to be mild and why years of concerted experimental efforts have yielded only a few proven targets for this gene, since many of the real targets would presumably not have been highly enriched compared to the controls.

References

[1] M. A. Larkin, G. Blackshields, N. P. Brown et al., "Clustal W and Clustal X version 2.0," *Bioinformatics*, vol. 23, no. 21, pp. 2947–2948, 2007.

[2] M. Goujon, H. McWilliam, W. Li et al., "A new bioinformatics analysis tools framework at EMBL-EBI," *Nucleic Acids Research*, vol. 38, pp. W695–W699, 2010.

[3] M. Mahone, E. E. Saffman, and P. F. Lasko, "Localized Bicaudal-C RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1," *The Embo Journal*, vol. 14, no. 9, pp. 2043–2055, 1995.

[4] R. D. Finn, J. Mistry, J. Tate et al., "The Pfam protein families database," *Nucleic Acids Research*, vol. 38, supplement 1, pp. D211–D222, 2010.

[5] M. Ashburner, P. Thompson, J. Roote et al., "The genetics of a small autosomal region of *Drosophila* melanogaster containing the structural gene for alcohol dehydrogenase—VII. Characterization of the region around the snail and cactus loci," *Genetics*, vol. 126, no. 3, pp. 679–694, 1990.

[6] P. Lasko, "Posttranscriptional regulation in *Drosophila* oocytes and early embryos," *Wiley Interdisciplinary Reviews: RNA*, vol. 2, no. 3, pp. 408–416, 2011.

[7] S. F. Altschul, T. L. Madden, A. A. Schäffer et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Research*, vol. 25, no. 17, pp. 3389–3402, 1997.

[8] P. McQuilton, S. E. St Pierre, and J. Thurmond, "FlyBase 101—the basics of navigating flyBase," *Nucleic Acids Research*, vol. 40, no. 1, pp. D706–D714, 2012.

[9] J. Mohler and E. F. Wieschaus, "Dominant maternal-effect mutations of *Drosophila* melanogaster causing the production of double-abdomen embryos," *Genetics*, vol. 112, no. 4, pp. 803–822, 1986.

[10] T. Schupbach and E. Wieschaus, "Germline autonomy of maternal-effect mutations altering the embryonic body pattern of *Drosophila*," *Developmental Biology*, vol. 113, no. 2, pp. 443–448, 1986.

[11] N. V. Grishin, "KH domain: one motif, two folds," *Nucleic Acids Research*, vol. 29, no. 3, pp. 638–643, 2001.

[12] K. Nakel, S. A. Hartung, F. Bonneau, C. R. Eckmann, and E. Conti, "Four KH domains of the C. elegans Bicaudal-C ortholog GLD-3 form a globular structural platform," *RNA*, vol. 16, no. 11, pp. 2058–2067, 2010.

[13] C. A. Kim and J. U. Bowie, "SAM domains: uniform structure, diversity of function," *Trends in Biochemical Sciences*, vol. 28, no. 12, pp. 625–628, 2003.

[14] E. E. Saffman, S. Styhler, K. Rother, W. Li, S. Richard, and P. Lasko, "Premature translation of *oskar* in oocytes lacking the RNA-binding protein Bicaudal-C," *Molecular and Cellular Biology*, vol. 18, no. 8, pp. 4855–4862, 1998.

[15] J. Chicoine, P. Benoit, C. Gamberi, M. Paliouras, M. Simonelig, and P. Lasko, "Bicaudal-C Recruits CCR4-NOT deadenylase to target mRNAs and regulates Oogenesis, Cytoskeletal organization, and its own expression," *Developmental Cell*, vol. 13, no. 5, pp. 691–704, 2007.

[16] D. J. Bouvrette, S. J. Price, and E. C. Bryda, "K homology domains of the mouse polycystic kidney disease-related protein, Bicaudal-C (Bicc1), mediate RNA binding *in vitro*," *Nephron Experimental Nephrology*, vol. 108, no. 1, pp. e27–e34, 2008.

[17] B. M. Lunde, C. Moore, and G. Varani, "RNA-binding proteins: modular design for efficient function," *Nature Reviews*, vol. 8, no. 6, pp. 479–490, 2007.

[18] A. J. Peterson, M. Kyba, D. Bornemann, K. Morgan, H. W. Brock, and J. Simon, "A domain shared by the Polycomb group proteins Scm and ph mediates heterotypic and homotypic interactions," *Molecular and Cellular Biology*, vol. 17, no. 11, pp. 6683–6692, 1997.

[19] F. Qiao and J. U. Bowie, "The many faces of SAM," *Science's STKE*, vol. 2005, no. 286, article re7, 2005.

[20] M. J. Knight, C. Leettola, M. Gingery, H. Li, and J. U. Bowie, "A human sterile alpha motif domain polymerizome," *Protein Science*, vol. 20, no. 10, pp. 1697–1706, 2011.

[21] T. Chen, B. B. Damaj, C. Herrera, P. Lasko, and S. Richard, "Self-association of the single-KH-domain family members Sam68, GRP33, GLD-1, and Qk1: role of the KH domain," *Molecular and Cellular Biology*, vol. 17, no. 10, pp. 5707–5718, 1997.

[22] M. Di Fruscio, T. Chen, S. Bonyadi, P. Lasko, and S. Richard, "The identification of two *Drosophila* K homology domain proteins: KEP1 and SAM are members of the Sam68 family of GSG domain proteins," *Journal of Biological Chemistry*, vol. 273, no. 46, pp. 30122–30130, 1998.

[23] C. R. Eckmann, S. L. Crittenden, N. Suh, and J. Kimble, "GLD-3 and control of the mitosis/meiosis decision in the germline of *Caenorhabditis elegans*," *Genetics*, vol. 168, no. 1, pp. 147–160, 2004.

[24] T. Aviv, Z. Lin, S. Lau, L. M. Rendl, F. Sicheri, and C. A. Smibert, "The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators," *Nature Structural Biology*, vol. 10, no. 8, pp. 614–621, 2003.

[25] C. Gamberi, O. Johnstone, and P. Lasko, "*Drosophila* RNA binding proteins," *International Review of Cytology*, vol. 248, pp. 43–139, 2006.

[26] C. Maisonneuve, I. Guilleret, P. Vick et al., "Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow," *Development*, vol. 136, no. 17, pp. 3019–3030, 2009.

[27] U. Tran, L. Zakin, A. Schweickert et al., "The RNA-binding protein bicaudal C regulates *polycystin* 2 in the kidney by antagonizing miR-17 activity," *Development*, vol. 137, no. 7, pp. 1107–1116, 2010.

[28] J. Schultz, C. P. Ponting, K. Hofmann, and P. Bork, "SAM as a protein interaction domain involved in developmental regulation," *Protein Science*, vol. 6, no. 1, pp. 249–253, 1997.

[29] C. R. Eckmann, B. Kraemer, M. Wickens, and J. Kimble, "GLD-3, a bicaudal-C homolog that inhibits FBF to control germline sex determination in *C. elegans*," *Developmental Cell*, vol. 3, no. 5, pp. 697–710, 2002.

[30] L. Wang, C. R. Eckmann, L. C. Kadyk, M. Wickens, and J. Kimble, "A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*," *Nature*, vol. 419, no. 6904, pp. 312–316, 2002.

[31] J. M. Kugler, J. Chicoine, and P. Lasko, "Bicaudal-C associates with a trailer hitch/Me31B complex and is required for efficient gurken secretion," *Developmental Biology*, vol. 328, no. 1, pp. 160–172, 2009.

[32] M. J. Snee and P. M. Macdonald, "Bicaudal C and trailer hitch have similar roles in *gurken* mRNA localization and cytoskeletal organization," *Developmental Biology*, vol. 328, no. 2, pp. 434–444, 2009.

[33] M. K. Jaglarz, M. Kloc, W. Jankowska, B. Szymanska, and S. M. Bilinski, "Nuage morphogenesis becomes more complex: two translocation pathways and two forms of nuage coexist in *Drosophila* germline syncytia," *Cell and Tissue Research*, vol. 344, no. 1, pp. 169–181, 2011.

[34] M. J. Snee and P. M. Macdonald, "Dynamic organization and plasticity of sponge bodies," *Developmental Dynamics*, vol. 238, no. 4, pp. 918–930, 2009.

[35] M. Wilsch-Bräuninger, H. Schwarz, and C. Nüsslein-Volhard, "A sponge-like structure involved in the association and transport of maternal products during *Drosophila* oogenesis," *Journal of Cell Biology*, vol. 139, no. 3, pp. 817–829, 1997.

[36] M. Olszewska, J. J. Bujarski, and M. Kurpisz, "P-bodies and their functions during mRNA cell cycle: mini-review," *Cell Biochemistry and Function*, vol. 30, no. 3, pp. 177–182, 2012.

[37] R. Parker and U. Sheth, "P bodies and the control of mRNA translation and degradation," *Molecular Cell*, vol. 25, no. 5, pp. 635–646, 2007.

[38] C. Y. Chen and A. B. Shyu, "AU-rich elements: characterization and importance in mRNA degradation," *Trends in Biochemical Sciences*, vol. 20, no. 11, pp. 465–470, 1995.

[39] M. Tucker, R. R. Staples, M. A. Valencia-Sanchez, D. Muhlrad, and R. Parker, "Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae," *The Embo Journal*, vol. 21, no. 6, pp. 1427–1436, 2002.

[40] M. Tucker, M. A. Valencia-Sanchez, R. R. Staples, J. Chen, C. L. Denis, and R. Parker, "The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*," *Cell*, vol. 104, no. 3, pp. 377–386, 2001.

[41] J. Chen, Y. C. Chiang, and C. L. Denis, "CCR4, a 3′-5′ poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase," *The Embo Journal*, vol. 21, no. 6, pp. 1414–1426, 2002.

[42] S. Thore, F. Mauxion, B. Séraphin, and D. Suck, "X-ray structure and activity of the yeast Pop2 protein: A nuclease subunit of the mRNA deadenylase complex," *EMBO Reports*, vol. 4, no. 12, pp. 1150–1155, 2003.

[43] M. C. Daugeron, F. Mauxion, and B. Séraphin, "The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation," *Nucleic Acids Research*, vol. 29, no. 12, pp. 2448–2455, 2001.

[44] J. Chen, J. Rappsilber, Y. C. Chiang, P. Russell, M. Mann, and C. L. Denis, "Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex," *Journal of Molecular Biology*, vol. 314, no. 4, pp. 683–694, 2001.

[45] U. Oberholzer and M. A. Collart, "Characterization of NOT5 that encodes a new component of the Not protein complex," *Gene*, vol. 207, no. 1, pp. 61–69, 1998.

[46] C. J. A. Sigrist, L. Cerutti, E. de Castro et al., "PROSITE, a protein domain database for functional characterization and annotation," *Nucleic Acids Research*, vol. 38, pp. D161–D166, 2010.

[47] C. Temme, S. Zaessinger, S. Meyer, M. Simonelig, and E. Wahle, "A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in *Drosophila*," *The Embo Journal*, vol. 23, no. 14, pp. 2862–2871, 2004.

[48] M. R. Kraus, S. Clauin, Y. Pfister et al., "Two mutations in human BICC1 resulting in Wnt pathway hyperactivity associated with cystic renal dysplasia," *Human Mutation*, vol. 33, no. 1, pp. 86–90, 2012.

[49] T. C. Burn, T. D. Connors, W. R. Dackowski et al., "Analysis of the genomic sequence for the autosomal dominant polycystic kidney disease (PKD1) gene predicts the presence of a leucine-rich repeat: the American PKD1 consortium (APKD1 Consortium)," *Human Molecular Genetics*, vol. 4, no. 4, pp. 575–582, 1995.

[50] J. Hughes, C. J. Ward, B. Peral et al., "*The polycystic kidney disease* 1 (PKD1) gene encodes a novel protein with multiple cell recognition domains," *Nature Genetics*, vol. 10, no. 2, pp. 151–160, 1995.

[51] T. Mochizuki, G. Wu, T. Hayashi et al., "PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein," *Science*, vol. 272, no. 5266, pp. 1339–1342, 1996.

[52] F. Hildebrandt, E. Otto, C. Rensing et al., "A novel gene encoding an SH3 domain protein is mutated in nephronophthisis type 1," *Nature Genetics*, vol. 17, no. 2, pp. 149–153, 1997.

[53] L. F. Onuchic, L. Furu, Y. Nagasawa et al., "*PKHD1*, the polycystic kidney and hepatic disease 1 gene, encodes a novel large protein containing multiple immunoglobulin-like plexin-transcription-factor domains and parallel beta-helix 1 repeats," *American Journal of Human Genetics*, vol. 70, no. 5, pp. 1305–1317, 2002.

[54] C. J. Ward, M. C. Hogan, S. Rossetti et al., "The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein," *Nature Genetics*, vol. 30, no. 3, pp. 259–269, 2002.

[55] H. Xiong, Y. Chen, Y. Yi et al., "A novel gene encoding a TIG multiple domain protein is a positional candidate for autosomal recessive polycystic kidney disease," *Genomics*, vol. 80, no. 1, pp. 96–104, 2002.

[56] C. Cogswell, S. J. Price, X. Hou, L. M. Guay-Woodford, L. Flaherty, and E. C. Bryda, "Positional cloning of jcpk/bpk locus of the mouse," *Mammalian Genome*, vol. 14, no. 4, pp. 242–249, 2003.

[57] U. Tran, L. M. Pickney, B. D. Özpolat, and O. Wessely, "Xenopus bicaudal-C is required for the differentiation of the amphibian pronephros," *Developmental Biology*, vol. 307, no. 1, pp. 152–164, 2007.

[58] D. J. Bouvrette, V. Sittaramane, J. R. Heidel, A. Chandrasekhar, and E. C. Bryda, "Knockdown of bicaudal C in zebrafish (*Danio rerio*) causes cystic kidneys: a nonmammalian model of polycystic kidney disease," *Comparative Medicine*, vol. 60, no. 2, pp. 96–106, 2010.

[59] Y. Fu, I. Kim, P. Lian et al., "Loss of Bicc1 impairs tubulomorphogenesis of cultured IMCD cells by disrupting E-cadherin-based cell-cell adhesion," *European Journal of Cell Biology*, vol. 89, no. 6, pp. 428–436, 2010.

[60] M. Jovanovic and M. O. Hengartner, "miRNAs and apoptosis: RNAs to die for," *Oncogene*, vol. 25, no. 46, pp. 6176–6187, 2006.

[61] F. S. Neuman-Silberberg and T. Schupbach, "The *Drosophila* dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGFα-like protein," *Cell*, vol. 75, no. 1, pp. 165–174, 1993.

[62] R. P. Ray and T. Schüpbach, "Intercellular signaling and the polarization of body axes during *Drosophila* oogenesis," *Genes & Development*, vol. 10, no. 14, pp. 1711–1723, 1996.

[63] J. E. Wilhelm, M. Buszczak, and S. Sayles, "Efficient protein trafficking requires trailer hitch, a component of a ribonucleoprotein complex localized to the ER in *Drosophila*," *Developmental Cell*, vol. 9, no. 5, pp. 675–685, 2005.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 287814, 12 pages
doi:10.1155/2012/287814

Review Article

The eIF4F and eIFiso4F Complexes of Plants: An Evolutionary Perspective

Ryan M. Patrick and Karen S. Browning

Department of Chemistry and Biochemistry and the Institute for Cell and Molecular Biology, The University of Texas at Austin, Austin, TX 78712, USA

Correspondence should be addressed to Karen S. Browning, kbrowning@mail.utexas.edu

Received 12 January 2012; Accepted 16 February 2012

Academic Editor: Thomas Preiss

Translation initiation in eukaryotes requires a number of initiation factors to recruit the assembled ribosome to mRNA. The eIF4F complex plays a key role in initiation and is a common target point for regulation of protein synthesis. Most work on the translation machinery of plants to date has focused on flowering plants, which have both the eIF4F complex (eIF4E and eIF4G) as well as the plant-specific eIFiso4F complex (eIFiso4E and eIFiso4G). The increasing availability of plant genome sequence data has made it possible to trace the evolutionary history of these two complexes in plants, leading to several interesting discoveries. eIFiso4G is conserved throughout plants, while eIFiso4E only appears with the evolution of flowering plants. The eIF4G N-terminus, which has been difficult to annotate, appears to be well conserved throughout the plant lineage and contains two motifs of unknown function. Comparison of eIFiso4G and eIF4G sequence data suggests conserved features unique to eIFiso4G and eIF4G proteins. These findings have answered some questions about the evolutionary history of the two eIF4F complexes of plants, while raising new ones.

1. Introduction

In eukaryotes, posttranscriptional gene regulation at the level of translation initiation is an important mechanism [1]. The process of translation initiation begins with the eIF4F complex, made up of the subunits eIF4E, which recognizes the 7-methylguanosine (m^7G) cap on the 5′ end of mRNA, and eIF4G, which binds to eIF4E and serves as a scaffold for other initiation factors [2]. eIF4G has sites for binding poly(A)-binding proteins (PABPs), which bind to the poly(A)-tail at the 3′ end of the mRNA, effectively allowing the eIF4F complex to circularize the mRNA molecule [3]. eIF4G also has RNA binding activity which may promote association with mRNA and improve eIF4E cap recognition [4]. eIF4G additionally binds the RNA helicase eIF4A [5], which promotes ATP-dependent unwinding of RNA secondary structure in a manner promoted by eIF4G and eIF4B [6]. The 43S preinitiation complex, made up of the 40S ribosomal subunit, eIF2 bound to GTP and Met-$tRNA^{Met}_i$, eIF3, eIF1, eIF1a, and eIF5 [2], is recruited to the mRNA by eIF4G through contacts with eIF3 [7] as well as eIF5 and eIF1 [8]. The docking of the 43S preinitiation complex is followed by scanning for the AUG start codon and joining of the 60S ribosomal subunit to begin translation [2]. The placement of the eIF4F complex at the beginning of this process makes it a key point for regulation of protein synthesis [9].

Flowering plants have two distinct isoforms of the eIF4F complex. In addition to the evolutionarily conserved eIF4F complex made up of eIF4E and eIF4G, they also have a plant-specific eIFiso4F complex made up of eIFiso4E and eIFiso4G [10, 11]. Wheat eIF4F and eIFiso4F have been shown to have differential effects on translation of various RNAs [12]. It has been reported that eIF4E-binding to eIF4G is very tight (0.18 nM K_D) and eIFiso4E-binding to eIFiso4G is similarly tight (0.08 nM K_D), while mixed complexes of eIF4E to eIFiso4G and eIFiso4E to eIF4G have ~80–100-fold less tight binding than their preferred partner; however, the mixed complexes retain activity *in vitro* [13]. *Arabidopsis thaliana* mutant plants with only a mixed complex of eIFiso4G and eIF4E are able to survive; but, those plants with only

Table 1: Distribution of eIF4F subunit genes in Viridiplantae. Nonflowering plants and green algae are bold.

	eIF4G	eIFiso4G	eIF4E	eIFiso4E	4EHP	eIF4E1b
Arabidopsis thaliana	1	2	1	1	1	2
Arabidopsis lyrata	1	3	1	1	1	2
Thellungiella halophila	1	2	1	1	1	1
Carica papaya	1	1	1	1	1	0
Theobroma cacao	1	2	1	1	1	0
Citrus clementina	1	2	1	1	1	0
Citrus sinensis	1	2	1	1	1	0
Eucalyptus grandis	1	2	1	1	1	1
Solanum tuberosum	2	2	1	1	1	0
Prunus persica	1	2	1	1	1	0
Fragaria vesca	1	2	1	1	1	1
Cucumis sativus	2	2	1	1	1	0
Glycine max	4	4	2	2	2	0
Medicago truncatula	1	1	1	1	1	0
Populus trichocarpa	2	4	1	2	1	0
Ricinus communis	1	1	1	1	1	0
Manihot esculenta	2	2	1	2	2	0
Vitis vinifera	1	2	1	1	1	0
Mimulus guttatus	2	2	1	2	1	0
Aquilegia coerulea	1	2	2	1	1	0
Sorghum bicolor	2	1	1	1	1	0
Zea mays	3	2	2	2	1	0
Setaria italica	2	2	1	1	1	0
Oryza sativa	1	2	1	1	1	0
Brachypodium distachyon	2	1	1	1	1	0
Selaginella moellendorffii	**2**	**2**	**4**	**0**	**1**	**0**
Physcomitrella patens	**2**	**5**	**4**	**0**	**1**	**0**
Chlamydomonas reinhardtii	**1**	**1**	**1**	**0**	**0**	**0**
Volvox carteri	**1**	**1**	**1**	**0**	**0**	**0**
Micromonas pusilla	**1**	**1**	**1**	**0**	**0**	**0**

eIF4G and eIFiso4E do not appear to be able to progress through a normal developmental program (Mayberry and Browning, unpublished observations). These results suggest that unique properties are associated with the two cap-binding complexes and their subunits in plants.

The increasing amount of sequence data from Viridiplantae (the monophyletic group of green plants, including the green algae and land plants) has made it possible to ask questions about the evolutionary history of the eIF4F and eIFiso4F complexes. Essentially all work to date on the translation machinery of Viridiplantae has been done in flowering plants. This work seeks to clarify the distribution of eIF4F and eIFiso4F subunit genes through Viridiplantae and identifying sequence traits in order to better understand the evolutionary significance of these complexes.

2. Materials and Methods

Plant eIF4F/eIFiso4F subunit protein sequences were obtained by BLAST of genome databases including NCBI [14], Joint Genome Institute [15], Phytozome [16], Sol Genomics Network [17], the Strawberry Genome [18], and Cacao Genome Database (http://www.cacaogenomedb.org/). Upstream genomic regions were translated using the ExPASy Translate tool [19] and were in some cases used where annotated eIF4G protein sequences may be incomplete. eIF4G and eIFiso4G alignments were performed by ClustalW2 [20] with manual adjustments (see Supplementary Table 1 in Supplementary Material available online at doi:10.1155/2012/287814 for a list of genes/loci used). eIF4E and eIFiso4E alignment and phylogeny were generated by MAFFT [21].

3. Results and Discussion

3.1. eIFiso4E Appears in Flowering Plants. All flowering plants with available completed genome sequences encode eIF4E and eIFiso4E proteins (Table 1). Most Viridiplantae also encode the conserved additional eIF4E family member 4EHP (also known as nCBP in plants) [22], though it is lost

in green algae. Additionally, some plants, like *A. thaliana*, encode eIF4E-like genes with divergence from the canonical plant eIF4E sequence which we term eIF4E1b genes (Patrick and Browning, manuscript in preparation). To address the lineage of eIF4E and eIFiso4E, a phylogeny of eIF4E genes from Viridiplantae was constructed (Figure 1).

To our knowledge, it has not been previously noted that eIFiso4E first appears at the emergence of flowering plants; it is not present in the genomes of the bryophyte *Physcomitrella patens*, the lycophyte *Selaginella moellendorffii*, or green algae, and there is no expressed sequence tag (EST) support for eIFiso4E before angiosperms evolved. *Amborella trichopoda*, the earliest diverging angiosperm known [23], has EST support for both eIF4E and an early eIFiso4E, and ESTs from other early angiosperms (such as the aquatic flowering plant *Cabomba aquatica*, see Figure 1) support a fully developed flowering plant eIFiso4E.

We have also found that gymnosperms have two forms of eIF4E, with one resembling the more conserved plant eIF4E and one being a divergent form of eIF4E that is distinct from eIFiso4E, which we term $eIF4E_{gs}$ (eIF4E Gymnosperm). There is currently good EST support for $eIF4E_{gs}$ within conifers, as well as evidence of its presence in the cycad *Cycas rumphii* (Figure 1). Research of the translation machinery in conifers would be needed to address whether $eIF4E_{gs}$ has a preferred binding partner in eIFiso4G or eIF4G, creating a parallel form of eIFiso4F in gymnosperms. It is unclear whether gene duplication happened in the common ancestor of gymnosperms and angiosperms, with the duplicated eIF4E diverging to $eIF4E_{gs}$ in gymnosperms and to eIFiso4E in angiosperms, or whether parallel gene duplication and divergence happened in each lineage; it is interesting in either case that the development of a second distinct eIF4E in plants seems coincident with transition to seed-based reproduction.

3.2. Distribution of eIF4G and eIFiso4G in Viridiplantae. The domain structure of eIF4G in plants is organized similarly to mammals, with a shared core structure of an eIF4E-binding site, the HEAT-1/MIF4G and HEAT-2/MA3 domains which bind eIF4A and contribute to mRNA scanning [24], and a long N-terminus with little identified structure [25]. Plant eIF4G differs from mammalian eIF4G in that it lacks the C-terminal HEAT-3/W2 domain. Plant eIFiso4G is similar in structure to eIF4G, but lacks the long N-terminus (see Figure 2).

One of the most interesting questions regarding the translation machinery of plants is why they contain both eIF4G and the plant-specific isoform eIFiso4G. In flowering plants, these proteins form distinct eIF4F (eIF4G with eIF4E) and eIFiso4F (eIFiso4G with eIFiso4E) complexes, that differ in their ability to promote translation of structured mRNAs *in vitro* [26]. Plant viruses often require one of these complexes for replication, but not the other, and the genes for the subunits of eIF4F or eIFiso4F have been identified as virus resistance genes for many types of plant viruses [27]. Most flowering plants with completed genomes available have more than one eIFiso4G gene (Table 1); *A. thaliana* has two, with the eIFiso4G1 gene being more highly expressed than eIFiso4G2. They appear to have overlapping functions, since deletion of either eIFiso4G subunit has little effect, but simultaneous deletion leads to a severe phenotype [28].

Flowering plants with completed genomes are about evenly divided between those that have a single copy of eIF4G and those that have two or more, but it is more common for the eIFiso4G copy number to be higher than eIF4G than vice versa (Table 1). *A. thaliana* has one eIF4G gene, and interestingly deletion of eIF4G has little effect (Mayberry and Browning, unpublished observations), in contrast to the severe growth phenotype of the eIFiso4G double mutant [28]. Nearly all Viridiplantae species which currently have sequenced genomes available contain genes for both eIF4G and eIFiso4G (*Chlorella variabilis* is a possible exception, as it appears to encode only eIFiso4G). This evolutionary conservation suggests that, while the genes have overlapping functions in translation initiation, each may have important specific roles in gene regulation as well.

As there was no eIFiso4E present before the evolution of angiosperms, it is unclear whether the binding partner of eIFiso4G at the conserved 4E-binding site (see below) was eIF4E or 4EHP in earlier Viridiplantae. Wheat eIFiso4G can form a complex with 4EHP that has some capacity to enhance translation initiation [22]; however, in *A. thaliana,* 4EHP does not appear to form a complex with eIF4G (Patrick and Browning, unpublished observations). 4EHP does not appear to be present in green algae (Table 1), leaving eIF4E the most likely option to form a complex with eIFiso4G in that lineage. As the function of eIFiso4G has only been studied in flowering plants that express eIFiso4E and form the eIFiso4F complex, research would be necessary to confirm that eIFiso4G has similar roles in translation initiation in nonflowering plants.

3.3. The N-Terminus of Plant eIF4G. Due to poor sequence conservation in the N-terminus, there is often difficulty annotating the eIF4G start site, especially outside of angiosperms. Based on available genomic information from flowering plants, we have been able to identify two conserved motifs in the N-terminal region, referred to here as the 4G-PN1 and 4G-PN2 sites (plant eIF4G N-terminal motif 1 and 2). 4G-PN1 is 17 amino acids long, with the consensus sequence PARTSAPPNxDEQKRxQ (Figure 3(a)), and appears 180 amino acids into *A. thaliana* eIF4G. 4G-PN2 is 15 amino acids long, with the consensus sequence VKITxPxTHEELxLD (Figure 3(b)), and appears 375 amino acids into the *A. thaliana* eIF4G. The region N-terminal of 4G-PN1 and between 4G-PN1 and 4G-PN2 is poorly conserved at a sequence level in plants but the positions of the two motifs and length of the intervening sequence are maintained. The 4G-PN2 motif is followed by a long poorly conserved region leading into the 4E-binding site and HEAT-1 domain. The role of these motifs, whether structural or supporting protein-protein interactions, is not known.

Though the 4G-PN1 and 4G-PN2 motifs are present upstream of the eIF4G HEAT-1 in almost all available Viridiplantae genome sequences, they are sometimes not included in the predicted protein coding sequence. They are

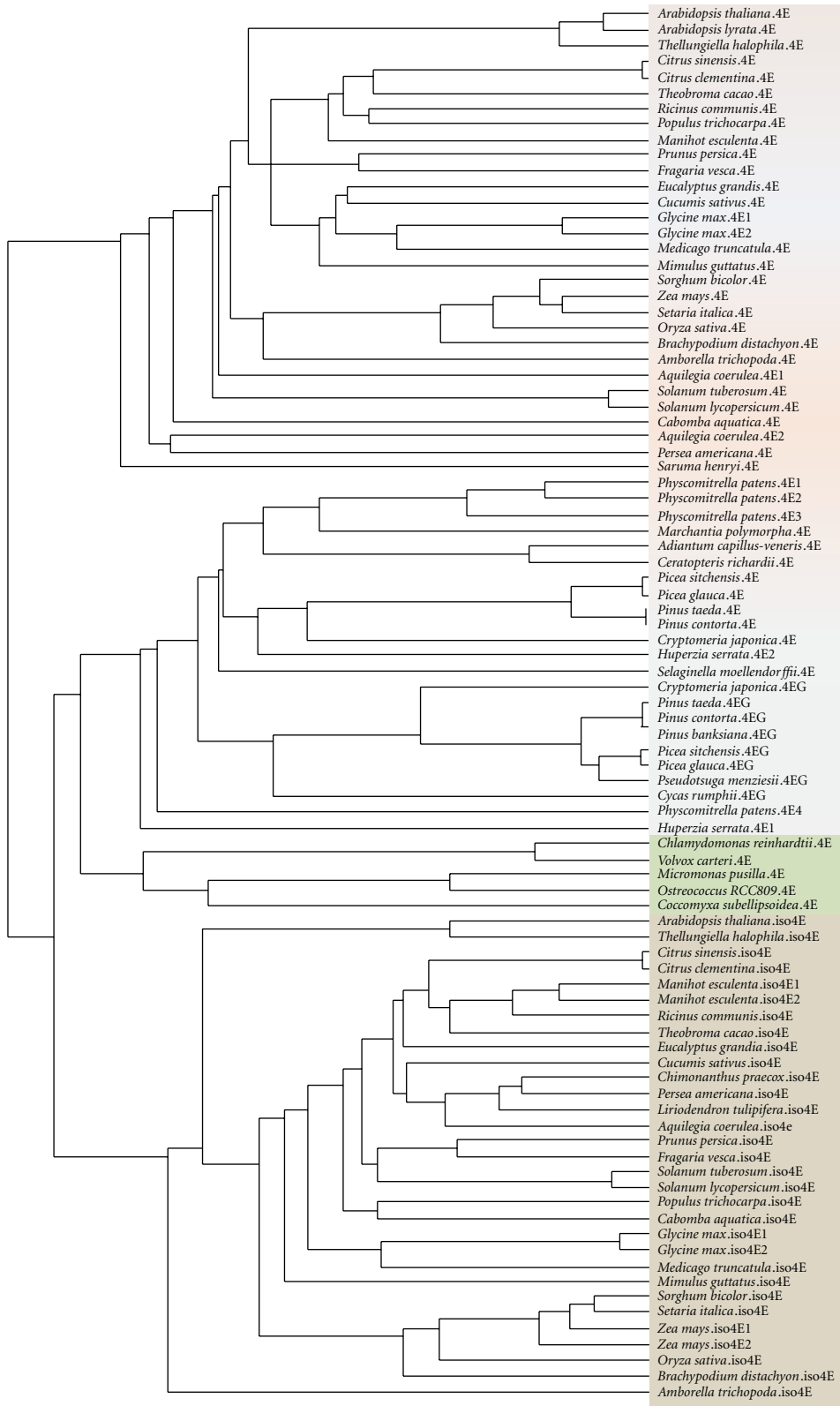


Figure 1: Phylogeny of Viridiplantae eIF4E and eIFiso4E. $eIF4E_{gs}$ genes of gymnosperms are labeled eIF4EG. Phylogeny generated by alignment of eIF4E, eIFiso4E, and $eIF4E_{gs}$ genes using MAFFT version 6 [21].

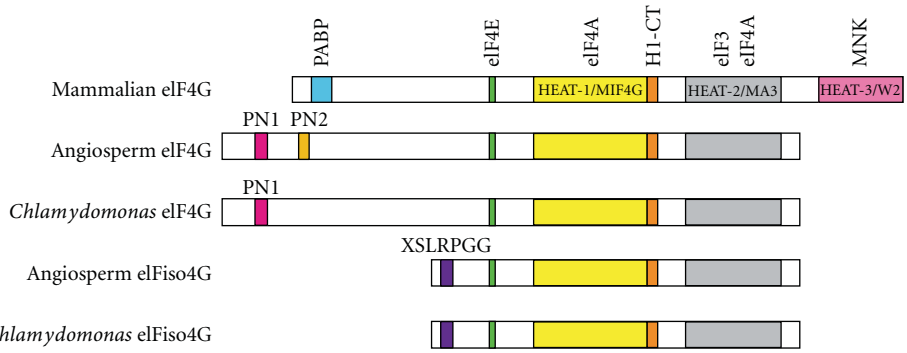


Figure 2: Domain organization of eIF4G and eIFiso4G from mammals, angiosperms, and the green algae *Chlamydomonas reinhardtii*. Plant eIF4G and eIFiso4G share the core organization of the eIF4E-binding site, HEAT-1/MIF4G domain, H1-CT motif, and HEAT-2/MA3 domain with mammals, but do not have the C-terminal HEAT-3 domain. The eIF3- and eIF4A-binding regions are thought to be maintained between all shown isoforms. Plant eIF4G has a longer N-terminus than mammals and contains the plant-specific 4G-PN1 and 4G-PN2 motifs as shown. *Chlamydomonas* eIF4G has a 4G-PN1-like sequence but no 4G-PN2 motif, while other green algae may have a 4G-PN2 motif but no 4G-PN1 motif. eIFiso4G is remarkably well conserved across plants, with the N-terminal XSLRPGG motif maintained from green algae to angiosperms.

present in the genome of *P. patens* and *S. moellendorffii*, as well as EST evidence supporting their existence in the conifer *Picea glauca*, which supports a conserved long N-terminus for eIF4G at least back to the emergence of land plants. Further investigation will be needed to determine if there are alternative splicing and translation initiation sites giving rise to multiple forms of eIF4G in plants. Supporting proteomic data is needed as well to fully understand the role of these motifs.

3.4. eIF4G of Green Algae. Green algae genomes currently annotate eIF4G as several different lengths, with *Chlamydomonas reinhardtii* being predicted as the same length as vascular plant eIF4G, but the close relative *Volvox carteri* being annotated without the N-terminus though its sequence is present in the genome. These green algae encode a 4G-PN1-like motif at the proper location (Figure 3(a)), but do not appear to have a PN2-like motif. *Ostreococcus* and *Micromonas* species have their eIF4G annotated as severely truncated, to the point where the 4E-binding sequence is not included, though it is encoded in the genome. Assuming the annotations are erroneously short, a 4G-PN2-like motif is encoded at the proper location upstream of the eIF4G HEAT-1 domain (Figure 3(b)); however, no 4G-PN1-like motif can be found.

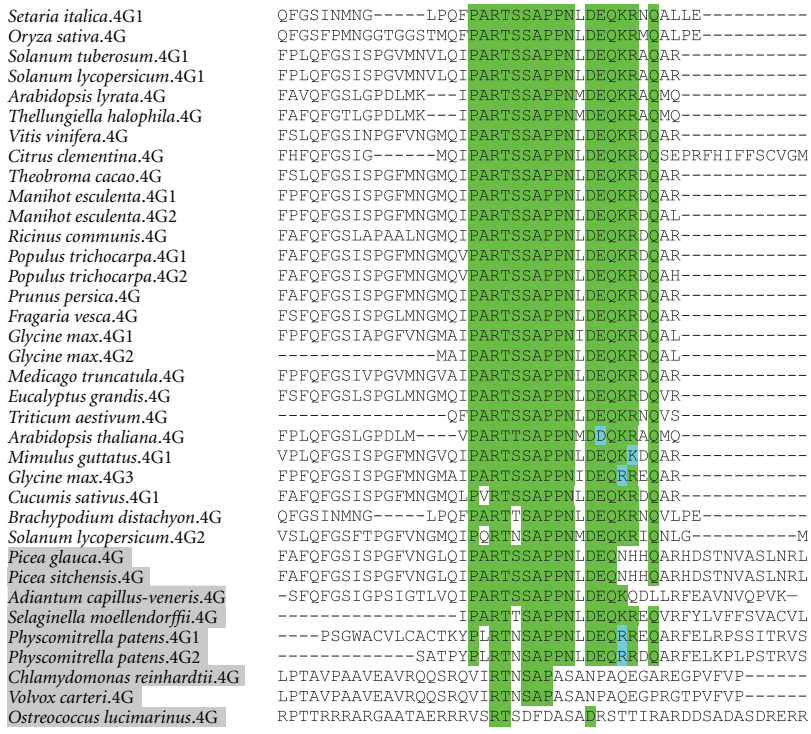
These lines of evidence support the possibility of a common Viridiplantae ancentral eIF4G with a full length N-terminus containing the 4G-PN1 and 4G-PN2 motifs. If this is the case, either motif may have been lost in some algae lineages, while both were maintained in the land plant lineage.

3.5. The H1-CT Site in Plants. The *cum2* mutation in *A. thaliana* was identified as a point mutation of a proline residue in eIF4G that inhibits replication of *Cucumber mosaic virus* [29]. Interestingly, this mutation occurs at a motif that is well conserved in eukaryotes, with the proline at this location conserved in animals and fungi. The motif, found between the end of the HEAT-1 domain and the predicted eIF3 binding site, has previously been identified as the H1-CT motif [25], conserved in fungi and animals, and here we provide evidence that this motif is conserved in most eukaryotic eIF4G proteins (Figure 4).

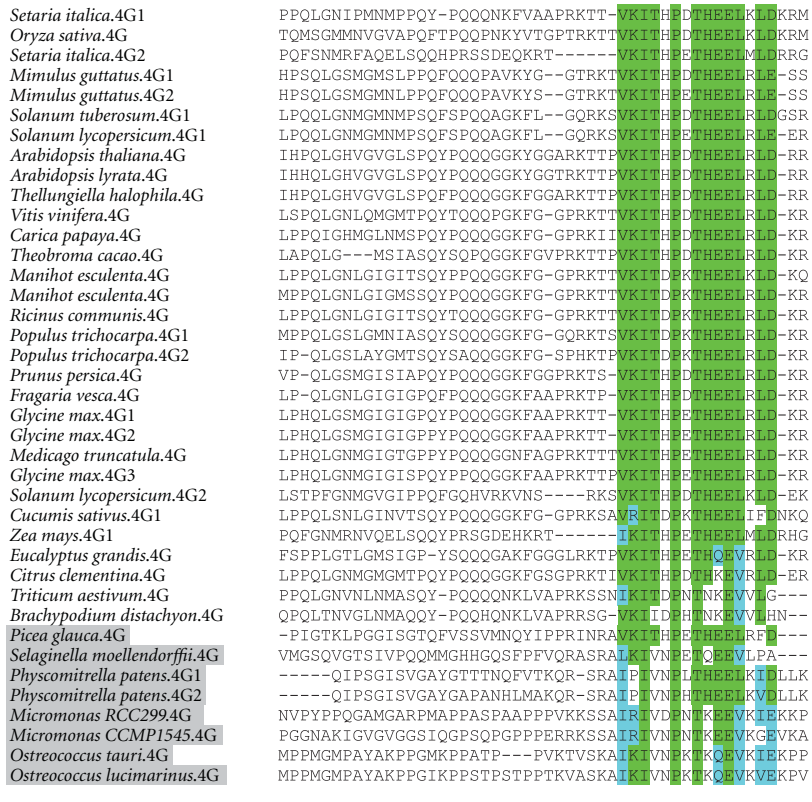
The core shared motif of the H1-CT region in plant eIF4G and eIFiso4G, which is also well conserved in other eukaryotes, is $RRx_5KxIxExHxxA$ (Figure 4). The residues around this core are divergent in eIF4G and eIFiso4G, the eIF4G motif at the site being RRVEGPKKI(D/E)EVHRDA (Figure 4(a)) and for eIFiso4G being PRREExKAKTIxEHx-EAExxLG (Figure 4(b)). The H1-CT motif in mammals and yeast shares similarities with both the eIF4G and eIFiso4G motifs (Figure 4(c)). The reason for the difference at this motif in the two plant isoforms is not clear, but it is useful for differentiation between divergent eIF4G and eIFiso4G genes.

3.6. Is the Origin of eIFiso4G Outside Viridiplantae? A second site useful for identification of eIFiso4G genes is a conserved N-terminal sequence of XSLRPGG (Figure 5), with X being a hydrophobic amino acid (I, V, or L). This sequence is conserved in eIFiso4G throughout the Viridiplantae lineage, but is not present in eIF4G. The purpose of this conserved motif is unknown, as N-terminal truncations of eIFiso4G lacking this sequence were found to bind eIFiso4E, eIF4A, synthesize polypeptides, and hydrolyze ATP at wild-type levels [30].

While eIFiso4G is present in all Viridiplantae, it is not clear whether the origin of the plant-specific isoform of eIF4G was before or after the divergence of Viridiplantae. Interestingly, two heterokonts, the brown algae *Ectocarpus siliculosus* and the marine diatom *Thalassiosira pseudonana,* encode a sequence similar to the eIFiso4G XSLRPGG motif at the correct position upstream of an eIF4G HEAT-1 domain. The *E. siliculosus* gene also bears more similarity to eIFiso4G than eIF4G at the H1-CT motif, while the *T. pseudonana* has similarities to both (Figure 4(c)).



(a)



(b)

Figure 3: The N-terminal motifs of eIF4G. Residues highlighted in green have identity to the consensus sequence, and residues highlighted in blue have similarity. Genes of nonflowering plants and green algae are shaded grey. (a) The PG-N1 motif with consensus sequence PARTSAPPNxDEQKRxQ. (b) The PGN-2 motif with consensus sequence VKITxPxTHEELxLD.

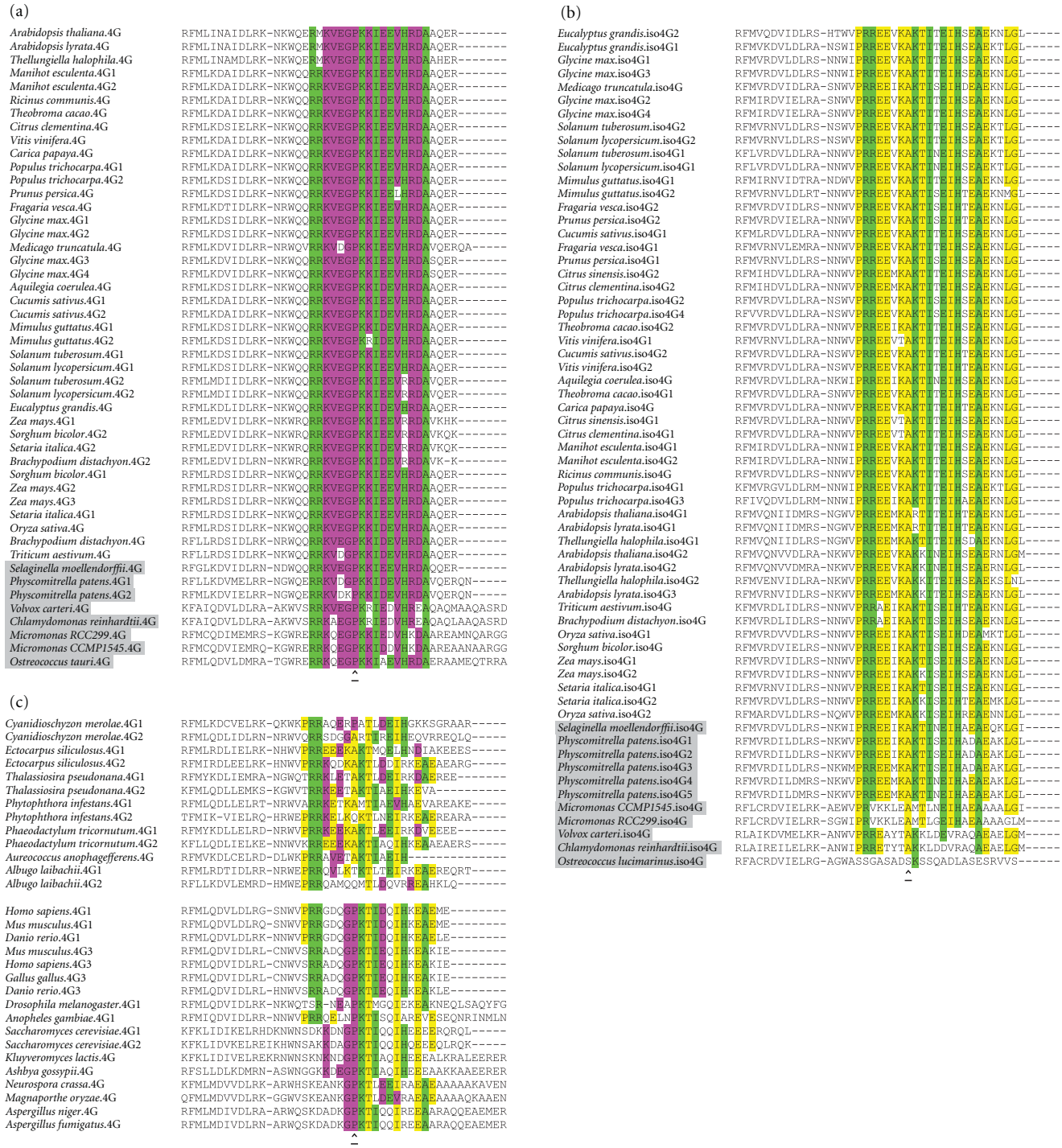


Figure 4: The H1-CT motif of eIF4G and eIFiso4G. Residues highlighted in green have identity to the shared core sequence RRx$_5$KxIxExHxxA. The arrow identifies the site of the *cum2* mutation in eIF4G. (a) The H1-CT motif of eIF4G. Residues highlighted in purple have identity to the unique residues of the eIF4G H1-CT motif RRVEGPKKI(D/E)EVHRDA. Genes of nonflowering plants and green algae are shaded grey. (b) The H1-CT motif of eIFiso4G. Residues highlighted in yellow have identity to the unique residues of the eIFiso4G H1-CT motif PRREExKAKTIxEHxEAExxLG. Genes of nonflowering plants and green algae are shaded grey. (c) The H1-CT motif of eIF4G genes of heterokonts, animals, and fungi. Residues are highlighted according to their identity to the shared core motif (green), the motif of plant eIF4G (purple), or the motif of eIFiso4G (yellow).

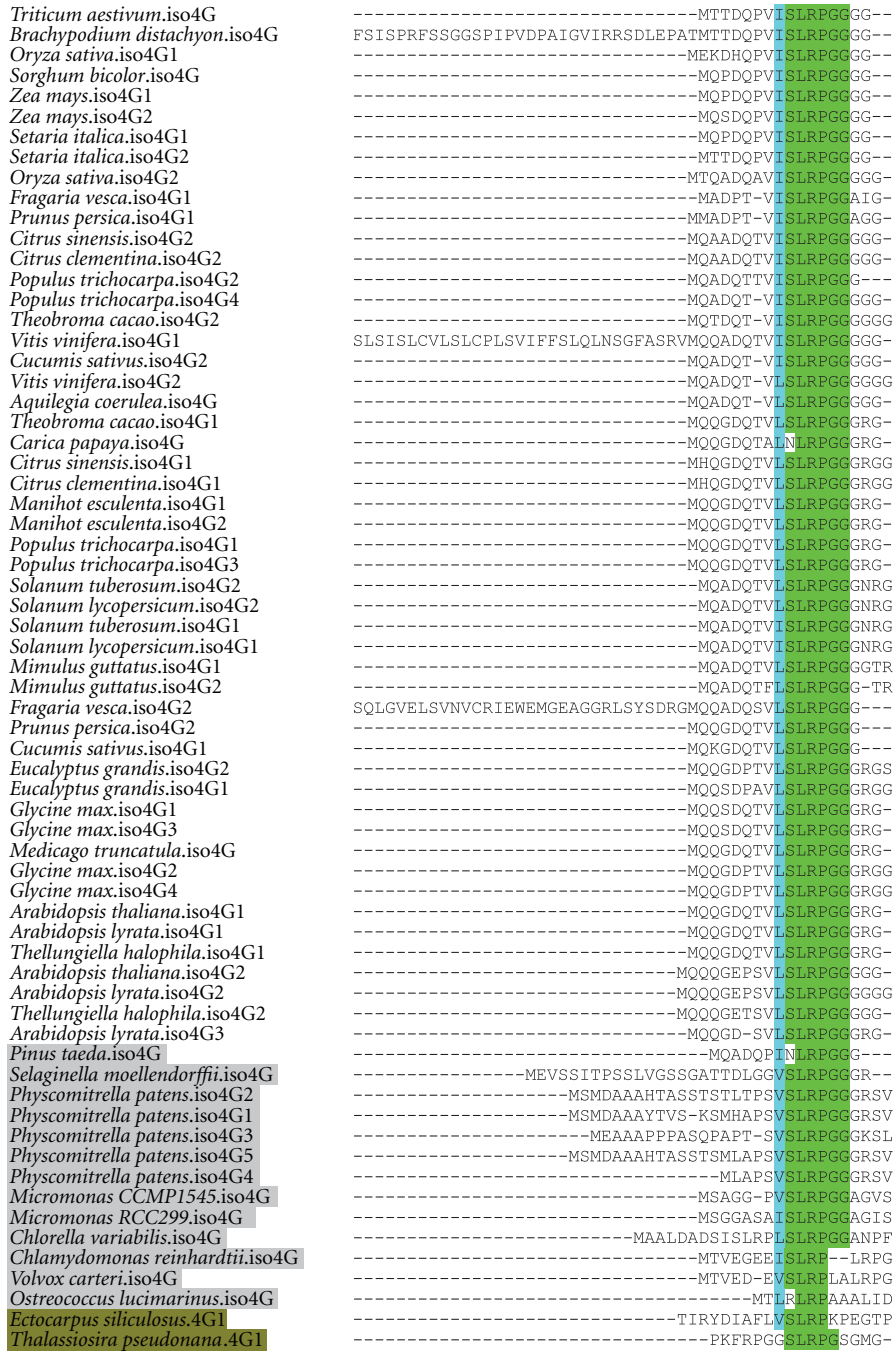


Figure 5: The N-terminal XSLRPGG motif of eIFiso4G. Residues highlighted in green have identity to the consensus sequence, and the variable hydrophobic residue is highlighted in blue. Genes of nonflowering plants and green algae are shaded grey. Genes of the heterokont eIF4G sequences containing this motif are shaded in brown.

The red algae *Cyanidioschyzon merolae*, more closely related to Viridiplantae [31], encodes two eIF4G genes, but they are divergent to the point it is not possible to identify them as either eIF4G or eIFiso4G homologs. The *E. siliculosus* gene may contribute evidence of a conserved eIFiso4G outside of Viridiplantae, but there is not enough support at this time to definitively state that the origin of eIFiso4G predates Viridiplantae.

3.7. The 4E-Binding Site of eIF4G and eIFiso4G. As eIF4G and eIFiso4G prefer to form discrete complexes with eIF4E and eIFiso4E, respectively [6], we used alignment of known sequences for angiosperm eIF4G and eIFiso4G to find if they have distinct 4E-binding motifs and whether the 4E-binding site in these proteins changed after the evolution of eIFiso4E. eIF4G has a well-conserved 4E-binding site sequence of KKYSRDFLLx$_8$LPxxF, which appears in its flowering plant

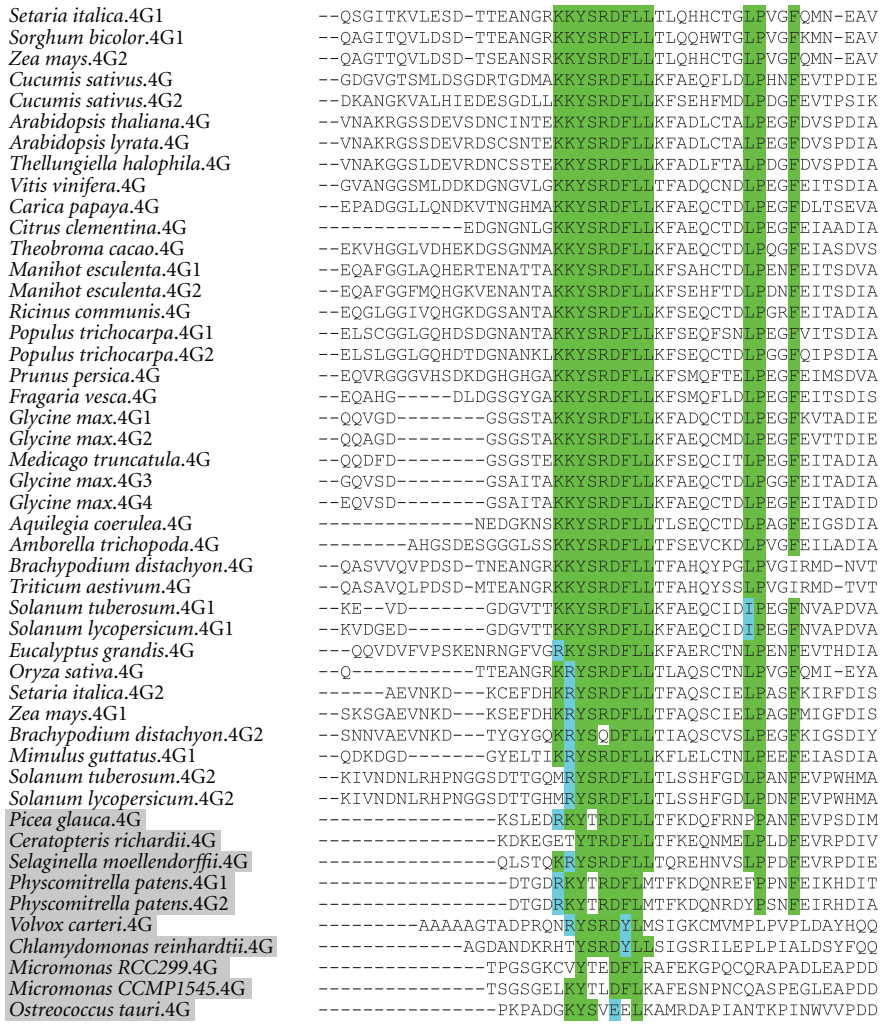


Figure 6: The 4E-binding site of plant eIF4G. Residues highlighted in green have identity to the consensus sequence KKYSRDFLLx$_8$LPxxF, and residues highlighted blue have similarity. Genes of nonflowering plants and green algae are shaded grey.

form as early as the lycophyte *S. moellendorffii* (Figure 6). The eIFiso4G site for 4E binding is ERVRYTR(D/E)QLLZLRE (Z being Glu or Gln) (Figure 7). Interestingly, it seems common for plants to have one eIFiso4G copy closely matching this consensus sequence, while other copies may diverge from this sequence. For example, *A. thaliana* eIFiso4G1 is close to the consensus sequence, while eIFiso4G2 diverges at several residues. eIFiso4G2 copurifies with eIFiso4E and has similar activity to eIFiso4G1 *in vitro* [12], so it is unclear at this time whether these differences are meaningful.

The flowering plant 4E-binding sequence of eIFiso4G seems nearly fully formed in the bryophyte *P. patens*, and the sequence in green algae eIFiso4G is roughly as similar to its angiosperm counterpart as the green algae eIF4G 4E-binding site is to its angiosperm version. One might expect the 4E-binding sites to have evolved after the emergence of eIFiso4E to each bind their preferred partner and discriminate against the other, but it seems in both cases the 4E-binding site was well formed before eIFiso4E evolved and has changed little since. The discrimination may therefore be at a site on the large subunit away from the identified 4E-binding site, or it may have evolved on the 4G-binding interface of eIF4E and eIFiso4E.

4. Conclusions

The increasing availability of genomic sequences from Viridiplantae has helped clarify the evolutionary history of the flowering plant eIF4F and eIFiso4F complexes, but has also raised many new questions. The discovery that evolution of eIFiso4G occurred long before eIFiso4E is surprising; *in vitro* observations on the eIFiso4F complex of wheat [13–26] and *Arabidopsis* [12] as well as the ability for either eIFiso4E or eIFiso4G gene disruptions to confer resistance to *Lettuce mosaic virus*, *Plum pox virus*, and *Turnip mosaic virus* in *A. thaliana* [32] point to a strongly intertwined role for eIFiso4E and eIFiso4G. This opens up several questions. Before the evolution of eIFiso4E, was eIF4E shared between eIF4G and eIFiso4G, or was 4EHP involved? Does eIFiso4G promote translation in green algae and early land plants, as it seems to in flowering plants, or did it have a different role altogether?

Figure 7: The 4E-binding site of plant eIFiso4G. Residues highlighted in green have identity to the consensus sequence ERVRYTR(D/E)QLLZLRE, and residues highlighted blue have similarity. Genes of nonflowering plants and green algae are shaded grey. Plants generally have one copy of eIFiso4G that closely resembles the consensus sequence; this primary copy is highlighted in yellow. Secondary copies, which are unhighlighted, may diverge from this sequence.

What is the relationship between the evolution of flowering plants and the coincident appearance of eIFiso4E, which appears conserved in all available angiosperm sequences? Future work will hopefully begin to answer these questions and should build toward an understanding of the function in flowering plants of the eIF4F and eIFiso4F complexes.

While mutational and deletion studies have been performed on eIFiso4G [30, 33], less analysis has been published on the activity of different domains of plant eIF4G, and the role of the N-terminal region remains mysterious. Deletion of a significant portion of the eIF4G N-terminus has little effect *in vitro* on translational activity ([34] and Mayberry and Browning, unpublished observations) suggesting the N-terminus may have a regulatory or unknown function. The identification of two N-terminal motifs in the plant eIF4G conserved back to at least the evolution of land plants and possibly as far back as the root of Viridiplantae implies that the N-terminal region does have some important function.

Future studies will be necessary to determine whether these motifs are involved in interactions with other proteins (possibly to PABP, the binding site of which has not been identified in plant eIF4G) and to discover whether the N-terminus contributes to translation initiation or to some other as yet unrecognized function(s) of eIF4G.

Acknowledgment

This work was supported by the National Science Foundation to K. S. Browning (MCB1052530 and Arabidopsis 2010 S-0000335).

References

[1] W. C. Merrick, "Eukaryotic protein synthesis: still a mystery," *Journal of Biological Chemistry*, vol. 285, no. 28, pp. 21197–21201, 2010.

[2] R. J. Jackson, C. U. T. Hellen, and T. V. Pestova, "The mechanism of eukaryotic translation initiation and principles of its regulation," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 2, pp. 113–127, 2010.

[3] S. E. Wells, P. E. Hillner, R. D. Vale, and A. B. Sachs, "Circularization of mRNA by eukaryotic translation initiation factors," *Molecular Cell*, vol. 2, no. 1, pp. 135–140, 1998.

[4] A. Yanagiya, Y. V. Svitkin, S. Shibata, S. Mikami, H. Imataka, and N. Sonenberg, "Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4GI (eIF4GI) for efficient interaction of eIF4E with the mRNA cap," *Molecular and Cellular Biology*, vol. 29, no. 6, pp. 1661–1669, 2009.

[5] M. Oberer, A. Marintchev, and G. Wagner, "Structural basis for the enhancement of eIF4A helicase activity by eIF4G," *Genes and Development*, vol. 19, no. 18, pp. 2212–2223, 2005.

[6] A. R. Ozes, K. Feoktistova, B. C. Avanzino et al., "Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B," *Journal of Molecular Biology*, vol. 412, no. 4, pp. 674–687, 2011.

[7] A. K. Lefebvre, N. L. Korneeva, M. Trutschl et al., "Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit," *Journal of Biological Chemistry*, vol. 281, no. 32, pp. 22917–22932, 2006.

[8] H. He, T. Von der Haar, C. R. Singh et al., "The yeast eukaryotic initiation factor 4G (eIF4G) HEAT domain interacts with eIF1 and eIF5 and is involved in stringent AUG selection," *Molecular and Cellular Biology*, vol. 23, no. 15, pp. 5431–5445, 2003.

[9] N. Sonenberg and A. G. Hinnebusch, "Regulation of translation initiation in eukaryotes: mechanisms and biological targets," *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[10] K. S. Browning, "Plant translation initiation factors: it is not easy to be green," *Biochemical Society Transactions*, vol. 32, no. 4, pp. 589–591, 2004.

[11] K. S. Browning, "The plant translational apparatus," *Plant Molecular Biology*, vol. 32, no. 1-2, pp. 107–144, 1996.

[12] L. K. Mayberry, M. Leah Allen, M. D. Dennis, and K. S. Browning, "Evidence for variation in the optimal translation initiation complex: plant eIF4B, eIF4F, and eIF(iso)4F differentially promote translation of mRNAs," *Plant Physiology*, vol. 150, no. 4, pp. 1844–1854, 2009.

[13] L. K. Mayberry, M. L. Allen, K. R. Nitka et al., "Plant cap-binding complexes eukaryotic initiation factors eIF4F and eIFISO4F: molecular specificity of subunit binding," *Journal of Biological Chemistry*, vol. 286, no. 49, pp. 42566–42574, 2011.

[14] S. F. Altschul, J. C. Wootton, E. Zaslavsky, and Y. K. Yu, "The construction and use of log-odds substitution scores for multiple sequence alignment," *PLoS Computational Biology*, vol. 6, no. 7, Article ID e1000852, 2010.

[15] I. V. Grigoriev, H. Nordberg, I. Shabalov et al., "The genome portal of the department of energy joint genome institute," *Nucleic Acids Research*, vol. 40, no. 1, pp. D26–D32, 2012.

[16] D. M. Goodstein, S. Shu, R. Howson et al., "Phytozome: a comparative platform for green plant genomics," *Nucleic Acids Research*, vol. 40, no. 1, pp. D1178–D1186, 2012.

[17] A. Bombarely, N. Menda, I. Y. Tecle et al., "The Sol Genomics Network (solgenomics.net): growing tomatoes using Perl," *Nucleic Acids Research*, vol. 39, pp. D1149–D1155, 2011.

[18] V. Shulaev, D. J. Sargent, R. N. Crowhurst et al., "The genome of woodland strawberry (*Fragaria vesca*)," *Nature Genetics*, vol. 43, no. 2, pp. 109–116, 2010.

[19] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel, and A. Bairoch, "ExPASy: the proteomics server for in-depth protein knowledge and analysis," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3784–3788, 2003.

[20] M. A. Larkin, G. Blackshields, N. P. Brown et al., "Clustal W and clustal X version 2.0," *Bioinformatics*, vol. 23, no. 21, pp. 2947–2948, 2007.

[21] K. Katoh and H. Toh, "Recent developments in the MAFFT multiple sequence alignment program," *Briefings in Bioinformatics*, vol. 9, no. 4, pp. 286–298, 2008.

[22] K. A. Ruud, C. Kuhlow, D. J. Goss, and K. S. Browning, "Identification and characterization of a novel cap-binding protein from *Arabidopsis thaliana*," *Journal of Biological Chemistry*, vol. 273, no. 17, pp. 10325–10330, 1998.

[23] M. J. Moore, C. D. Bell, P. S. Soltis, and D. E. Soltis, "Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 49, pp. 19363–19368, 2007.

[24] A. Marintchev, K. A. Edmonds, B. Marintcheva et al., "Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation," *Cell*, vol. 136, no. 3, pp. 447–460, 2009.

[25] A. Marintchev and G. Wagner, "eIF4G and CBP80 share a common origin and similar domain organization: implications for the structure and function of eIF4G," *Biochemistry*, vol. 44, no. 37, pp. 12265–12272, 2005.

[26] D. R. Gallie and K. S. Browning, "eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs," *Journal of Biological Chemistry*, vol. 276, no. 40, pp. 36951–36960, 2001.

[27] C. Robaglia and C. Caranta, "Translation initiation factors: a weak link in plant RNA virus infection," *Trends in Plant Science*, vol. 11, no. 1, pp. 40–45, 2006.

[28] A. D. Lellis, M. L. Allen, A. W. Aertker et al., "Deletion of the eIFiso4G subunit of the *Arabidopsis* eIFiso4F translation initiation complex impairs health and viability," *Plant Molecular Biology*, vol. 74, no. 3, pp. 249–263, 2010.

[29] M. Yoshii, M. Nishikiori, K. Tomita et al., "The *Arabidopsis* cucumovirus multiplication 1 and 2 loci encode translation initiation factors 4E and 4G," *Journal of Virology*, vol. 78, no. 12, pp. 6102–6111, 2004.

[30] A. M. Metz and K. S. Browning, "Mutational analysis of the functional domains of the large subunit of the isozyme form of

wheat initiation factor eIF4F," *Journal of Biological Chemistry*, vol. 271, no. 49, pp. 31033–31036, 1996.

[31] T. Cavalier-Smith, "Kingdoms Protozoa and Chromista and the eozoan root of the eukaryotic tree," *Biology Letters*, vol. 6, no. 3, pp. 342–345, 2010.

[32] V. Nicaise, J. L. Gallois, F. Chafiai et al., "Coordinated and selective recruitment of eIF4E and eIF4G factors for potyvirus infection in *Arabidopsis thaliana*," *The FEBS Letters*, vol. 581, no. 5, pp. 1041–1046, 2007.

[33] S. Cheng and D. R. Gallie, "Competitive and noncompetitive binding of eIF4B, eIF4A, and the poly(A) binding protein to wheat translation initiation factor eIFiso4G," *Biochemistry*, vol. 49, no. 38, pp. 8251–8265, 2010.

[34] S. Lax, W. Fritz, K. Browning, and J. Ravel, "Isolation and characterization of factors from wheat germ that exhibit eukaryotic initiation factor 4B activity and overcome 7-methylguanosine 5'-triphosphate inhibition of polypeptide synthesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 82, no. 2, pp. 330–333, 1985.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 406357, 8 pages
doi:10.1155/2012/406357

Review Article

Regulation of Translation Initiation under Abiotic Stress Conditions in Plants: Is It a Conserved or Not so Conserved Process among Eukaryotes?

Alfonso Muñoz[1] and M. Mar Castellano[2]

[1] *Departamento de Genética Molecular de Plantas, Centro Nacional de Biotecnología-CSIC, 28049 Madrid, Spain*
[2] *Centro de Biotecnología y Genómica de Plantas, INIA-UPM, Campus de Montegancedo, 28223 Madrid, Spain*

Correspondence should be addressed to M. Mar Castellano, castellano.mar@inia.es

Received 23 December 2011; Accepted 8 February 2012

Academic Editor: Greco Hernández

For years, the study of gene expression regulation of plants in response to stress conditions has been focused mainly on the analysis of transcriptional changes. However, the knowledge on translational regulation is very scarce in these organisms, despite in plants, as in the rest of the eukaryotes, translational regulation has been proven to play a pivotal role in the response to different stresses. Regulation of protein synthesis under abiotic stress was thought to be a conserved process, since, in general, both the translation factors and the translation process are basically similar in eukaryotes. However, this conservation is not so clear in plants as the knowledge of the mechanisms that control translation is very poor. Indeed, some of the basic regulators of translation initiation, well characterised in other systems, are still to be identified in plants. In this paper we will focus on both the regulation of different initiation factors and the mechanisms that cellular mRNAs use to bypass the translational repression established under abiotic stresses. For this purpose, we will review the knowledge from different eukaryotes but paying special attention to the information that has been recently published in plants.

1. Introduction

One of the main responses of cells to stress conditions involves partial or virtually total cessation of energetically consumptive processes normally vital to homeostasis, including transcription and protein synthesis. Translation consumes a substantial amount of cellular energy and, therefore, it is one of the main targets to be inhibited in response to most, if not all, types of cellular stresses. However, under conditions where global protein synthesis is severely compromised, some proteins are still synthesised as part of the mechanisms of cell survival, as these proteins are able to mitigate the damage caused by the stress and enable cells to tolerate the stressful conditions more effectively [1]. Appearance of abiotic stresses, as environmental conditions, is in many cases sudden. Therefore, a quick response to stress should be established to assure cell survival. In such a context, translational regulation of preexisting mRNAs provides a prompt and alternative way to control gene expression, as compared to other slower cellular processes such as mRNA transcription, processing, and transport to cytoplasm [2].

In animals and yeast, there are many known examples of global translational inhibition and preferential production of key proteins critical for survival under different abiotic insults [3–8]. This scenario also begins to be envisioned in plants where several studies demonstrate that general mRNA translation inhibition and selective translation of some mRNAs are key points in the adaptation process of plants to different abiotic stresses, including hypoxia, heat shock, water deficit, sucrose starvation, and saline stress [9]. Thus, in Arabidopsis seedlings subjected to oxygen deprivation, mRNAs coding proteins involved in glycolysis and alcoholic fermentation are efficiently translated, meanwhile the translation of other constitutively synthesised proteins is inhibited [10]. In a similar way, a decrease in the *de novo*

protein synthesis has been demonstrated in *Brassica napus* seedlings after being subjected to heat shock for several hours. Under these conditions, in an opposite way to the proteins synthesised under normal conditions, only the translation of heat shock proteins is observed [11]. Furthermore, a reduction of protein synthesis with an increase in the synthesis of membrane proteins and of sulphur assimilation enzymes and transporters has been described in Arabidopsis-cultured cells subjected to sublethal cadmium stress [12]. In addition, the translational repression of specific components of the translation machinery and cell cycle-related mRNAs has been observed during sucrose starvation using the same system [13]. Other examples of rapid impairment of *de novo* protein synthesis by osmotic stress in Arabidopsis and rice have recently been published [14].

2. Initiation of Translation: Main Target of the Translation Regulation in Response to Abiotic Stress

To date, the different experiments carried out to unravel the translational phase regulated under stress conditions point to a regulation mainly at the initiation step. In eukaryotes, under physiological conditions, the vast majority of mRNAs initiate translation via a canonical cap-dependent mechanism that begins with the recognition by the eIF4E factor of the cap structure (7-methyl guanosine) placed at the 5′ end of the mRNAs to be translated. eIF4E interacts with eIF4G and with eIF4A, forming the cap binding complex called eIF4F. This complex allows the further recruitment of the preinitiation complex 43S, which consists of the small ribosomal subunit 40S, the ternary complex eIF2/GTP/$tRNA_i^{met}$ and the factors eIF3, eIF1 and eIF1A. The resulting preinitiation complex scans the mRNAs in the 5′ → 3′ direction until an initiation codon is found. There the ribosomal subunit 60S is loaded and the elongation phase begins [15]. However, under abiotic stress conditions this canonical translation initiation is impeded by different mechanisms that affect mainly the activity of the initiation factors eIF2α, eIF4E, and eIF4A [1, 2, 5, 16–18].

3. Regulation of Translational Initiation Factors under Abiotic Stress

3.1. Translation Regulation by eIF2α Phosphorylation. In eukaryotes, one of the main mechanisms of translation inhibition in response to stress is the regulation of the subunit α of the eIF2 factor by phosphorylation. eIF2α phosphorylation is mediated by different kinases that are specifically activated in response to different stresses promoting the inhibition of translation by hindering the formation of the eIF2/GTP/$tRNA_i^{met}$ ternary complex [17]. eIF2α kinases and their activation by stress conditions are different among different eukaryotes. In vertebrates four different eIF2α-kinases, namely, GCN2, PERK, PKR and HRI that are activated by nutrient limitation [19], protein misfolding in the endoplasmic reticulum (ER) [20], virus infection [21], and heme group availability [22], respectively, have been described (Figure 1(a)). However, other eukaryotes have a different number of these enzymes. For instance, *Schizosaccharomyces pombe* has three eIF2α kinases (two distinct HRI and a GCN2), *Drosophila melanogaster* and *Caenorhabditis elegans* have only two (PERK and GCN2), and *Saccharomyces cerevisiae* has only one (GCN2) [23].

A strong inhibition of protein synthesis by eIF2α phosphorylation under different stress conditions has also been reported in plants, demonstrating that this mechanism of regulation of translation is conserved in these organisms [24]. Genome-wide searches for the presence of eIF2α kinases in Arabidopsis and rice suggest that higher plants only contain a GCN2-like eIF2α kinase [24]. In agreement with these *in silico* searches, so far only the eIF2α kinase GCN2 has been characterized in plants [24, 25], although some reports also suggest the controversial existence in plants of an eIF2α kinase with the biochemical properties of the mammalian PKR [26–29]. Arabidopsis GCN2 is activated under different stress conditions including amino acid and purine deprivation, cadmium, UV, cold shock, and wounding (Figure 1(a)), or in response to different hormones involved in the activation of defence response to insect herbivores [24, 25]. Although AtGCN2 activity is linked to a strong reduction in global protein synthesis under the aforementioned conditions, the activity of this enzyme does not account for the general inhibition of translation under all stresses in plants, as treatments using NaCl or H_2O_2 do not promote actively the phosphorylation of eIF2α. Moreover, results in Arabidopsis demonstrate that heat shock does not lead to eIF2α phosphorylation either, confirming previous results obtained in wheat [30]. Interestingly, heat shock causes a striking inhibition of protein synthesis in plants, suggesting that different mechanisms might be involved in the global protein synthesis inhibition observed under these conditions.

3.2. Translation Regulation by the Association of eIF4E with Interacting Proteins. The regulation of mammalian eIF4E under abiotic stress conditions is by far the mechanism that has been better studied. This regulation in mammals involves the interaction of eIF4E with the 4E-binding proteins (4E-BPs). 4E-BPs show the same conserved eIF4E-binding domain as eIF4G, so their action mechanism is based on their capability to compete out the eIF4G-eIF4E interaction, thereby inhibiting further recruitment of the ribosome to the mRNA "cap" structure. This mechanism is regulated by the phosphorylation status of 4E-BPs. Under physiological conditions, the TOR (target of rapamycin) kinase phosphorylates 4E-BPs, which turns 4E-BPs unable to interact with eIF4E. In response to different stresses, TOR is inhibited and 4E-BPs become dephosphorylated. This hypophosphorylation state increases the affinity of 4E-BPs for eIF4E, inhibiting cap-dependent translation and setting up a switch in the translational initiation mechanism from cap-dependent to cap-independent [18] (Figure 1(b)).

Regulation of eIF4E activity in budding yeast *S. cerevisiae* shares common features with that of mammals. In *S. cerevisiae* two functional homologs of the mammaliam

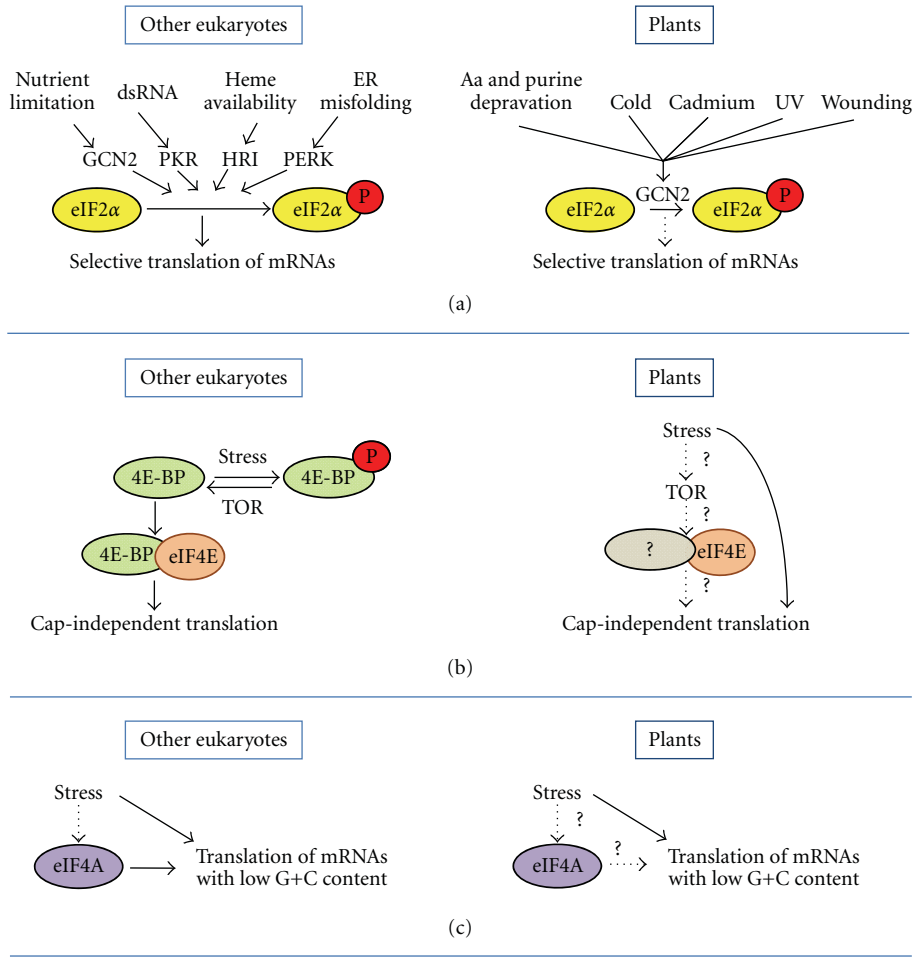


Figure 1: Regulation of translational initiation factors and transcript-differential translation under abiotic stress conditions. (a) Protein synthesis inhibition has been observed upon eIF2α phosphorylation both in plants and in other eukaryotes subjected to different abiotic stress conditions. In plants (right panel), in contrast to vertebrates (the case illustrated) (left panel), only the eIF2α kinase GCN2 has been described. In yeast and mammals eIF2α phosphorylation mediated by GCN2 promotes the selective translation of some mRNAs as *GCN4* or *ATF4*, respectively; however, whether eIF2α phosphorylation leads to the stimulation of translation of specific mRNAs is unknown in plants. (b) In mammals the activity of eIF4E under abiotic stress is regulated by the eIF4E binding to hypophosphorylated forms of the 4E-BPs (left panel). Such binding promotes cap-dependent translation inhibition and the observation of cap-independent translation. Different evidences point out that plants can support cap-independent translation under abiotic stress conditions (right panel). However, the role of eIF4E and TOR in this process has to be elucidated. (c) Some abiotic stress conditions promote the selective translation of mRNAs with low G+C content in yeast and plants. In yeast (left panel) the involvement of eIF4A in this regulation has been described, although the exact mechanism regulating the activity of eIF4A is still unclear. In plants (right panel) the role of eIF4A in this process has to be determined.

4E-BPs, p20 and EAP1, have been described [31, 32]. Both proteins block cap-dependent translation by interfering with the interaction of eIF4E with eIF4G, a mechanism analogous to that of the mammaliam 4E-BPs [31, 32]. In addition, TOR signalling pathway also plays a critical role in yeast, as in higher eukaryotes, in the modulation of translation initiation via regulation of eIF4E activity. Indeed, disruption of the *EAP1* gene confers partial resistance to the growth-inhibitory properties of rapamycin, implicating EAP1 in the TOR signaling pathway controlling cap-dependent translation in *S. cerevisiae* [32].

Cap-independent translation has also been observed in plants subjected to both abiotic and biotic stresses (Figure 1(b)). In maize, two cellular mRNAs, the alcohol dehydrogenase *ADH1* and the heat shock protein *HSP101*, are translated in a cap-independent manner in oxygen-deprived roots [33] and during heat stress [34], respectively. These data, together with the fact that plant viruses use a cap-independent translation strategy to translate their mRNAs lacking the cap structure in the host cells [35], demonstrate that plant translational apparatus is able to support cap-independent translation under stress conditions. In addition, TOR also plays an important role in the regulation of protein synthesis in plants as RNAi reduction of *TOR* results in a strong inhibition of translation initiation in Arabidopsis, while *TOR*-overexpressing lines show an increase in translation initiation efficiency [36]. Moreover, in these lines the expression levels of *AtTOR* are correlated to the tolerance

of Arabidopsis to osmotic stress indicating that AtTOR, possibly by its role in protein synthesis, modulates the response to abiotic stress conditions [36].

Regardless these striking parallelisms, the link between the role of TOR and the regulation of the eIF4E activity under abiotic stress in plants, if it exists, is far from being understood (Figure 1(b)). Indeed, no homolog of the 4E-BPs has been found in the plant genomes available to date. In spite of that, it has been described that the β subunit of the nascent polypeptide-associated complex (NAC) and the plant lipoxygenase 2 (AtLOX2) could putatively act as 4E-BP analogs since they interact with the Arabidopsis eIFiso4E in yeast two hybrid assays and these interactions can be displaced by the addition of AtIF4G *in vitro* [37, 38]. Moreover, AteIF4E has been proven to coimmunoprecipate with AtLOX2 from Arabidopsis extracts [38]. However, their role in the regulation of protein translation has not been demonstrated, as no evidences for changes in translation mediated by these proteins or for the regulation of their activities by TOR have been described either *in vitro* or *in vivo*.

3.3. Translation Regulation by eIF4A. Recently, new alternative mechanisms for the regulation of translation initiation under stress conditions which involve the regulation of the eIF4A RNA helicase have been discovered. A clear example is shown in yeast [5], where the authors demonstrated that glucose depletion causes a global translation inhibition due to a reduction in the amount of eIF4A bound to eIF4G. Concomitant with this reduction, changes in the levels of eIF3 associated to eIF4G are observed indicating that eIF4A could be required for the turnover in the association of eIF4G-eIF3 complex in a way that modulates translation initiation. Furthermore, the involvement of the regulation of eIF4A in translation in the response to lithium stress in *S. cerevisiae* has also been described [39] (Figure 1(c)).

As shown for the yeast eIF4A, plant eIF4A activity seems to be involved in the regulation of translation under abiotic stress in these organisms, as the overexpression of the pea DNA helicase 45, which seems to be the eIF4A ortholog, has been proven to confer high salinity tolerance in tobacco [40]. However, this observation should be further studied as the exact mechanism underlying this stress tolerance is not currently completely understood (Figure 1(c)).

4. Differential Translation of mRNAs in Response to Abiotic Stress Conditions

Under general translational inhibition conditions induced by abiotic stresses, some mRNAs involved in triggering stress responses are able to be selectively and efficiently translated. These transcripts have special characteristics that allow them to bypass specifically the different regulation points of translational inhibition. In this section we will focus on understanding the features that allow these mRNAs to circumvent downregulation of translation and we will deepen our knowledge in the information available in plants.

4.1. Differential Translation Mediated by eIF2α Regulation. Specific examples of mRNAs immune to eIF2α regulation under a variety of stress conditions as *GCN4* and *ATF4* have been characterized in yeast [41] and mammals [42] (Figure 1(a)). Both mRNAs are able to be translated by a complex mechanism based on the fact that when eIF2α is phosphorylated and, therefore, the ternary complex is scarce, the scanning ribosome fails to initiate translation at upstream reading frames (uORFs), which are terminated in premature stop codons. In this case, scanning continues downstream towards the functional initiation codon allowing, with this long scanning, the enough time for ternary complex recruitment and, therefore, to promote the subsequent translation of the functional peptide [16, 41].

In plants, eIF2α phosphorylation causes a drastic inhibition of protein synthesis during amino acid starvation that is correlated with a partial inhibition of mRNA association to polysomes [24], demonstrating that, under eIF2α phosphorylation, there are some transcripts still able to be translated. However, at the moment, it is not known whether or not eIF2α phosphorylation leads to stimulation of translation of specific mRNAs, as reported for other systems (Figure 1(a)). In plants, no homolog to GCN4 transcription factor has been characterized and there is no evidence for the involvement of GCN2 in the transcriptional activation of Arabidopsis genes homologous to those regulated by GCN4 in yeast [25].

4.2. Differential Translation Mediated by IRESs and CITEs. In the late 1980s, the study of viral gene expression led to the discovery of the most studied alternative mode of translation initiation, the IRES-driven initiation. This mechanism allows the 40S ribosome to be directly recruited to sequences located within the 5′-UTR of viral RNAs called Internal Ribome Entry Sites (IRES) without the need of cap-recognition by eIF4E [43–45]. Since then, IRES activity has been described in an increasingly number of cellular transcripts including those coding for translation initiation factors, transcription factors, oncoproteins, growth factors, and homeotic and survival proteins. The presence of these cellular IRESs (cIRESs) allows the efficient translation of mRNAs under conditions, where cap-dependent initiation is inhibited or seriously compromised, as it is the case of abiotic stress or during physiological processes as mitosis, apoptosis, or cell differentiation [46, 47].

In plants, three cIRESs have been characterized to support cap-independent translation *in vitro*. These cIRESs have been found within the 5′-leader sequences of the mRNAs coding for the Arabidopsis ribosomal protein S18 subunit C (RPS18C) [48], the maize heat shock protein 101 (HSP101) [34], and the maize alcohol dehydrogenase (ADH1) [33]. Two of these mRNAs, the *HSP101* and the *ADH1* mRNAs, are efficiently translated under heat shock and under hypoxia, respectively [33, 34], suggesting an important role of cIRESs in the mechanism for selective translation under abiotic stress in plants. Indeed, the 5′-leader of *ADH1* was able to provide efficient translation of a reporter gene *in vivo* in *Nicotiana benthamiana* cells both under oxygen shortage and heat shock, while translation of

the same construct lacking this sequence was significantly reduced [33]. Although promising, the examples of known plant cIRESs are scarce and, therefore, whether the use of cIRESs as translational enhancers of specific cellular mRNAs under abiotic stress is a generalized mechanism in these organisms remains still an open question.

For years the presence of cIRESs has been considered the only possible mechanism underlying cap-independent translation of cellular mRNAs. Interestingly, new mechanisms of cap-independent translation have been proposed to explain the translation observed under conditions where eIF4E activity is reduced [49, 50]. One of them is the translation of the mouse *HSP70* mRNA under heat stress conditions [4]. In this paper, Sun and collaborators demonstrate that the *HSP70* 5′-UTR is able to drive the translation of reporter genes under cap-independent conditions. However, the same sequence is unable to maintain cap-independent translation when placed in the intercistronic region of a bicistronic construct, ruling out the presence of an IRES within the sequence. Examples of such sequences have been described within plant viral mRNAs. The mRNAs of a large portion of all plant viruses lack the cap structure and, therefore, are forced to be translated in a cap-independent manner. To do so, in addition to viral IRESs, they use special elements termed cap-independent translational enhancers (CITEs). CITEs are able to recruit eIF4E and eIF4G cognates, or directly the 40S ribosomal subunit to the proximity to the AUG initiation codon, licensing in such a way the mRNA to initiate translation in a cap-independent manner [35, 51]. Although the existence of CITE-like elements is still considered exclusive of plant viral mRNAs, it would not be surprising if such elements are also discovered in plant cellular mRNAs. Cellular CITE-like elements, if present, might provide an alternative to cIRESs to drive translation of plant mRNAs [33].

Differential translation of some mRNAs under certain abiotic conditions could also be explained by the binding of specific RNA binding proteins to certain sequences within the mRNAs, acting as cap-dependent translational enhancing factors and cap-dependent enhancers, respectively. Most abiotic stress conditions reduce cap-dependent initiation and, therefore, enhancers acting synergistically with the cap could increase selectively the translational rate of those transcripts containing them. A good example of cap-dependent enhancing factors is the protein disulfide isomerase (PDI) that is a key regulator of insulin translation in response to glucose in mammals [52]. PDI is able to bind specifically to glucose responsive mRNAs under glucose stimulation and recruits the poly(A)-binding protein (PABP) to unknown enhancer elements in their 5′-UTR. Although how PABP binding could increase translation of such mRNAs is still unknown, it is reasonable to think that it is by the interaction of PABP with eIF4G. Cap-dependent enhancers of translation in plant viruses have also been described [53–55], being one of the better known examples the Ω sequence found in plant tobacco mosaic virus (TMV) [56]. This sequence is recognized by the HSP101 that, in turn and through its interaction with the Ω sequence, recruits eIF4G subunit to the 5′-UTR of the viral RNA [55]. The Ω sequence has been used to promote translation of cellular mRNAs enhancing both cap-dependent and cap-independent translation of the downstream gene by 2–10-fold. Therefore, these enhancers of cap-dependent translation could facilitate cap-dependent translation and even sustain some cap-independent translation under low eIF4E activity. If these kind of enhancers are also found in plant cellular mRNAs is a question that remains unanswered but that should be studied.

4.3. Differential Translation Mediated by eIF4A Regulation. Sequence analysis of polysome-bound mRNAs during glucose starvation in yeast, where a reduction of eIF4A association within the initiation complexes was observed, demonstrates that a common feature of these mRNAs is the low G+C content immediately upstream of the AUG [5]. These results suggest that the specific translation of mRNAs with low secondary structure could be selectively promoted under low eIF4A activity (Figure 1(c)). However, other alternative explanations cannot be fully excluded as, for example, the activation of IRES-driven translation of unstructured mRNAs by low level of helicase activity [6] or the possibility that other RNA helicases, with substrate preference for poorly structured mRNAs, may substitute the function of eIF4A. In a similar way, a study in Arabidopsis demonstrated that ribosome loading of mRNAs with high G+C content is differentially reduced under mild dehydration conditions [57]. These results may reflect, as in the previous case, a higher requirement for RNA helicase activity to initiate translation under stress in plants and may point to a low mRNA G+C content as a mechanism to bypass the restrain in eIF4A activity under abiotic stress (Figure 1(c)).

5. Unique Features of Regulation of Translation Initiation in Plants

It is well known that plants have unique translational characteristics as the existence, in addition to the canonical eIF4E and eIF4G factors, of IF(iso)4E and eIF(iso)4G isoforms. In Arabidopsis, one eIF(iso)4E and two eIF(iso)4Gs have been described; however, the number of these isoforms varies between plant species. eIF(iso)4E and eIF(iso)4G isoforms interact specifically between them to form eIF(iso)4F complexes [58]. The ability of the eIF(iso)4F complexes to support translation initiation of specific mRNAs has been proven different to that of the eIF4F complexes, suggesting that certain mRNA features allow different transcripts to interact preferentially with either complexes [59, 60]. Indeed, Lellis and coworkers have recently demonstrated that the double-mutant in the two Arabidopsis eIF(iso)4G factors displays strong phenotypes in growth and development, which, in the apparent absence of general protein synthesis inhibition, could be caused by the selective translation of specific genes [61]. Moreover, in maize it has been demonstrated that eIF(iso)4E is particularly required for the translation of stored mRNAs from dry seeds, and that eIF4E is unable to fully replace this eIF(iso)4E function [62]. If eIF4F and eIF(iso)4F complexes regulate translation of

different sets of mRNAs, this would mean a plant-specific layer of gene expression regulation that is worth studying in depth.

6. Conclusion

The conservation of mechanisms to globally inhibit protein synthesis concomitant to mRNA translation reprogramming under different stresses points out to the fundamental importance of translation regulation during the response to abiotic stresses in all eukaryotes. Although we already know that there are multiple parallel mechanisms across eukaryotes that modulate translation under abiotic stresses, we are still far away from understanding completely this regulation, as new alternative mechanisms taking part in this regulation are still being described. In plants, the study of translational regulation under stress is still in its infancy, and some of the most conserved regulators have not been found in these organisms yet. A considerable effort should be done in this respect, since understanding how plants respond to environmental conditions can only be fulfilled by a complete knowledge of how translation is regulated.

Acknowledgments

This work was supported by the ERC Starting Grant 260468 and by the Spanish Ministerio de Ciencia e Innovación (BIO2010-15751) to M. Mar Castellano and by a JAE-Doc Contract from CSIC to A. Muñoz.

References

[1] M. Holcik and N. Sonenberg, "Translational control in stress and apoptosis," *Nature Reviews Molecular Cell Biology*, vol. 6, no. 4, pp. 318–327, 2005.

[2] T. E. Graber and M. Holcik, "Cap-independent regulation of gene expression in apoptosis," *Molecular BioSystems*, vol. 3, no. 12, pp. 825–834, 2007.

[3] M. B. Al-Fageeh and C. M. Smales, "Control and regulation of the cellular responses to cold shock: the responses in yeast and mammalian systems," *Biochemical Journal*, vol. 397, no. 2, pp. 247–259, 2006.

[4] J. Sun, C. S. Conn, Y. Han, V. Yeung, and S.-B. Qian, "PI3K-mTORC1 attenuates stress response by inhibiting cap-independent Hsp70 translation," *Journal of Biological Chemistry*, vol. 286, no. 8, pp. 6791–6800, 2011.

[5] L. M. Castelli, J. Lui, S. G. Campbell et al., "Glucose depletion inhibits translation initiation via eIF4A loss and subsequent 48S preinitiation complex accumulation, while the pentose phosphate pathway is coordinately up-regulated," *Molecular Biology of the Cell*, vol. 22, no. 18, pp. 3379–3393, 2011.

[6] W. V. Gilbert, K. Zhou, T. K. Butler, and J. A. Doudna, "Cap-independent translation is required for starvation-induced differentiation in yeast," *Science*, vol. 317, no. 5842, pp. 1224–1227, 2007.

[7] L. Liu and M. C. Simon, "Regulation of transcription and translation by hypoxia," *Cancer Biology and Therapy*, vol. 3, no. 6, pp. 492–497, 2004.

[8] S. Braunstein, K. Karpisheva, C. Pola et al., "A Hypoxia-controlled cap-dependent to cap-independent translation switch in breast cancer," *Molecular Cell*, vol. 28, no. 3, pp. 501–512, 2007.

[9] M. Floris, H. Mahgoub, E. Lanet, C. Robaglia, and B. Menand, "Post-transcriptional regulation of gene expression in plants during abiotic stress," *International Journal of Molecular Sciences*, vol. 10, no. 7, pp. 3168–3185, 2009.

[10] C. Branco-Price, K. A. Kaiser, C. J. H. Jang, C. K. Larive, and J. Bailey-Serres, "Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*," *Plant Journal*, vol. 56, no. 5, pp. 743–755, 2008.

[11] S. Dhaubhadel, K. S. Browning, D. R. Gallie, and P. Krishna, "Brassinosteroid functions to protect the translational machinery and heat-shock protein synthesis following thermal stress," *Plant Journal*, vol. 29, no. 6, pp. 681–691, 2002.

[12] R. Sormani, E. Delannoy, S. Lageix et al., "Sublethal cadmium intoxication in *Arabidopsis thaliana* impacts translation at multiple levels," *Plant and Cell Physiology*, vol. 52, no. 2, pp. 436–447, 2011.

[13] M. Nicolaï, M. A. Roncato, A. S. Canoy et al., "Large-scale analysis of mRNA translation states during sucrose starvation in Arabidopsis cells identifies cell proliferation and chromatin structure as targets of translational control," *Plant Physiology*, vol. 141, no. 2, pp. 663–673, 2006.

[14] H. Matsuura, U. Kiyotaka, Y. Ishibashi et al., "A short period of mannitol stress but not LiCl stress led to global translational repression in plants," *Bioscience, Biotechnology and Biochemistry*, vol. 74, no. 10, pp. 2110–2112, 2010.

[15] R. J. Jackson, C. U. T. Hellen, and T. V. Pestova, "The mechanism of eukaryotic translation initiation and principles of its regulation," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 2, pp. 113–127, 2010.

[16] R. C. Wek, H. Y. Jiang, and T. G. Anthony, "Coping with stress: EIF2 kinases and translational control," *Biochemical Society Transactions*, vol. 34, no. 1, pp. 7–11, 2006.

[17] M. J. Clemens, "Initiation factor eIF2 alpha phosphorylation in stress responses and apoptosis," *Progress in Molecular and Subcellular Biology*, vol. 27, pp. 57–89, 2001.

[18] M. J. Clemens, "Translational regulation in cell stress and apoptosis. Roles of the eIF4E binding proteins," *Journal of Cellular and Molecular Medicine*, vol. 5, no. 3, pp. 221–239, 2001.

[19] P. Zhang, B. C. McGrath, J. Reinert et al., "The GCN2 eIF2α kinase is required for adaptation to amino acid deprivation in mice," *Molecular and Cellular Biology*, vol. 22, no. 19, pp. 6681–6688, 2002.

[20] H. P. Harding, M. Calfon, F. Urano, I. Novoa, and D. Ron, "Transcriptional and translational control in the mammalian unfolded protein response," *Annual Review of Cell and Developmental Biology*, vol. 18, pp. 575–599, 2002.

[21] M. J. Clemens, "PKR—a protein kinase regulated by double-stranded RNA," *International Journal of Biochemistry and Cell Biology*, vol. 29, no. 7, pp. 945–949, 1997.

[22] J. J. Chen, "Regulation of protein synthesis by the heme-regulated eIF2α kinase: relevance to anemias," *Blood*, vol. 109, no. 7, pp. 2693–2699, 2007.

[23] K. Zhan, J. Narasimhan, and R. C. Wek, "Differential activation of eIF2 kinases in response to cellular stresses in Schizosaccharomyces pombe," *Genetics*, vol. 168, no. 4, pp. 1867–1875, 2004.

[24] S. Lageix, E. Lanet, M. N. Pouch-Pélissier et al., "Arabidopsis eIF2α kinase GCN2 is essential for growth in stress conditions and is activated by wounding," *BMC Plant Biology*, vol. 8, article no. 134, 2008.

[25] Y. Zhang, Y. Wang, K. Kanyuka, M. A. J. Parry, S. J. Powers, and N. G. Halford, "GCN2-dependent phosphorylation of eukaryotic translation initiation factor-2α in Arabidopsis," *Journal of Experimental Botany*, vol. 59, no. 11, pp. 3131–3141, 2008.

[26] H. J. Hiddinga, C. J. Crum, Jie Hu, and D. A. Roth, "Viroid-induced phosphorylation of a host protein related to a dsRNA-dependent protein kinase," *Science*, vol. 241, no. 4864, pp. 451–453, 1988.

[27] J. O. Langland, L. A. Langland, K. S. Browning, and D. A. Roth, "Phosphorylation of plant eukaryotic initiation factor-2 by the plant-encoded double-stranded RNA-dependent protein kinase, pPKR, and inhibition of protein synthesis in vitro," *Journal of Biological Chemistry*, vol. 271, no. 8, pp. 4539–4544, 1996.

[28] J. O. Langland, Jin Song, B. L. Jacobs, and D. A. Roth, "Identification of a plant-encoded analog of PKR, the mammalian double-stranded RNA-dependent protein kinase," *Plant Physiology*, vol. 108, no. 3, pp. 1259–1267, 1995.

[29] C. J. Crum, J. Hu, H. J. Hiddinga, and D. A. Roth, "Tobacco mosaic virus infection stimulates the phosphorylation of a plant protein associated with double-stranded RNA-dependent protein kinase activity," *Journal of Biological Chemistry*, vol. 263, no. 26, pp. 13440–13443, 1988.

[30] D. R. Gallie, H. Le, C. Caldwell, R. L. Tanguay, N. X. Hoang, and K. S. Browning, "The phosphorylation state of translation initiation factors is regulated developmentally and following heat shock in wheat," *Journal of Biological Chemistry*, vol. 272, no. 2, pp. 1046–1053, 1997.

[31] M. Altmann, N. Schmitz, C. Berset, and H. Trachsel, "A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E," *EMBO Journal*, vol. 16, no. 5, pp. 1114–1121, 1997.

[32] G. P. Cosentino, T. Schmelzle, A. Haghighat, S. B. Helliwell, M. N. Hall, and N. Sonenberg, "Eap1p, a novel eukaryotic translation initiation factor 4E-associated protein in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 20, no. 13, pp. 4604–4613, 2000.

[33] E. S. Mardanova, L. A. Zamchuk, M. V. Skulachev, and N. V. Ravin, "The 5′ untranslated region of the maize alcohol dehydrogenase gene contains an internal ribosome entry site," *Gene*, vol. 420, no. 1, pp. 11–16, 2008.

[34] T. D. Dinkova, H. Zepeda, E. Martínez-Salas, L. M. Martínez, J. Nieto-Sotelo, and E. Sánchez De Jiménez, "Cap-independent translation of maize Hsp101," *Plant Journal*, vol. 41, no. 5, pp. 722–731, 2005.

[35] E. L. P. Kneller, A. M. Rakotondrafara, and W. A. Miller, "Cap-independent translation of plant viral RNAs," *Virus Research*, vol. 119, no. 1, pp. 63–75, 2006.

[36] D. Deprost, L. Yao, R. Sormani et al., "The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation," *EMBO Reports*, vol. 8, no. 9, pp. 864–870, 2007.

[37] M. A. Freire, "Translation initiation factor (iso) 4E interacts with BTF3, the β subunit of the nascent polypeptide-associated complex," *Gene*, vol. 345, no. 2, pp. 271–277, 2005.

[38] M. A. Freire, C. Tourneur, F. Granier et al., "Plant lipoxygenase 2 is a translation initiation factor-4E-binding protein," *Plant Molecular Biology*, vol. 44, no. 2, pp. 129–140, 2000.

[39] M. Montero-Lomelí, B. L. B. Morais, D. L. Figueiredo, D. C. S. Neto, J. R. P. Martins, and C. A. Masuda, "The initiation factor eIF4A is involved in the response to lithium stress in Saccharomyces cerevisiae," *Journal of Biological Chemistry*, vol. 277, no. 24, pp. 21542–21548, 2002.

[40] N. Sanan-Mishra, X. H. Pham, S. K. Sopory, and N. Tuteja, "Pea DNA helicase 45 overexpression in tobacco confers high salinity tolerance without affecting yield," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 2, pp. 509–514, 2005.

[41] A. G. Hinnebusch, "Translational regulation of GCN4 and the general amino acid control of yeast," *Annual Review of Microbiology*, vol. 59, pp. 407–450, 2005.

[42] H. P. Harding, I. Novoa, Y. Zhang et al., "Regulated translation initiation controls stress-induced gene expression in mammalian cells," *Molecular Cell*, vol. 6, no. 5, pp. 1099–1108, 2000.

[43] J. Pelletier and N. Sonenberg, "Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA," *Nature*, vol. 334, no. 6180, pp. 320–325, 1988.

[44] J. Pelletier, G. Kaplan, V. R. Racaniello, and N. Sonenberg, "Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5' noncoding region," *Molecular and Cellular Biology*, vol. 8, no. 3, pp. 1103–1112, 1988.

[45] S. K. Jang, H. G. Krausslich, M. J. H. Nicklin, G. M. Duke, A. C. Palmenberg, and E. Wimmer, "A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation," *Journal of Virology*, vol. 62, no. 8, pp. 2636–2643, 1988.

[46] K. A. Spriggs, M. Stoneley, M. Bushell, and A. E. Willis, "Re-programming of translation following cell stress allows IRES-mediated translation to predominate," *Biology of the Cell*, vol. 100, no. 1, pp. 27–38, 2008.

[47] A. A. Komar and M. Hatzoglou, "Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states," *Cell Cycle*, vol. 10, no. 2, pp. 229–240, 2011.

[48] R. Vanderhaeghen, R. De Clercq, M. Karimi, M. Van Montagu, P. Hilson, and M. Van Lijsebettens, "Leader sequence of a plant ribosomal protein gene with complementarity to the 18S rRNA triggers in vitro cap-independent translation," *FEBS Letters*, vol. 580, no. 11, pp. 2630–2636, 2006.

[49] I. N. Shatsky, S. E. Dmitriev, I. M. Terenin, and D. E. Andreev, "Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs," *Molecules and Cells*, vol. 30, no. 4, pp. 285–293, 2010.

[50] W. V. Gilbert, "Alternative ways to think about cellular internal ribosome entry," *Journal of Biological Chemistry*, vol. 285, no. 38, pp. 29033–29038, 2010.

[51] W. A. Miller, Z. Wang, and K. Treder, "The amazing diversity of cap-independent translation elements in the 3′-untranslated regions of plant viral RNAs," *Biochemical Society Transactions*, vol. 35, no. 6, pp. 1629–1633, 2007.

[52] S. D. Kulkarni, B. Muralidharan, A. C. Panda, B. Bakthavachalu, A. Vindu, and V. Seshadri, "Glucose-stimulated translation regulation of insulin by the 5′ UTR-binding proteins," *Journal of Biological Chemistry*, vol. 286, no. 16, pp. 14146–14156, 2011.

[53] D. A. Zelenina, O. I. Kulaeva, E. V. Smirnyagina et al., "Translation enhancing properties of the 5'-leader of potato virus X genomic RNA," *FEBS Letters*, vol. 296, no. 3, pp. 267–270, 1992.

[54] L. Neeleman, R. C. L. Olsthoorn, H. J. M. Linthorst, and J. F. Bol, "Translation of a nonpolyadenylated viral RNA is enhanced by binding of viral coat protein or polyadenylation of the RNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 25, pp. 14286–14291, 2001.

[55] D. R. Gallie, "The 5′-leader of tobacco mosaic virus promotes translation through enhanced recruitment of elF4F," *Nucleic Acids Research*, vol. 30, no. 15, pp. 3401–3411, 2002.

[56] D. R. Gallie and V. Walbot, "Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation," *Nucleic Acids Research*, vol. 20, no. 17, pp. 4631–4638, 1992.

[57] R. Kawaguchi and J. Bailey-Serres, "mRNA sequence features that contribute to translational regulation in Arabidopsis," *Nucleic Acids Research*, vol. 33, no. 3, pp. 955–965, 2005.

[58] L. K. Mayberry, M. L. Allen, K. R. Nitka, L. Campbell, P. A. Murphy, and K. S. Browning, "Plant cap-binding complexes eukaryotic initiation factors eIF4F and eIFISO4F: molecular specificity of subunit binding," *Journal of Biological Chemistry*, vol. 286, no. 49, pp. 42566–42574, 2011.

[59] D. R. Gallie and K. S. Browning, "eIF4G Functionally Differs from eIFiso4G in Promoting Internal Initiation, Cap-independent Translation, and Translation of Structured mRNAs," *Journal of Biological Chemistry*, vol. 276, no. 40, pp. 36951–36960, 2001.

[60] L. K. Mayberry, M. Leah Allen, M. D. Dennis, and K. S. Browning, "Evidence for variation in the optimal translation initiation complex: plant eIF4B, eIF4F, and eIF(iso)4F differentially promote translation of mRNAs," *Plant Physiology*, vol. 150, no. 4, pp. 1844–1854, 2009.

[61] A. D. Lellis, M. L. Allen, A. W. Aertker et al., "Deletion of the eIFiso4G subunit of the Arabidopsis eIFiso4F translation initiation complex impairs health and viability," *Plant Molecular Biology*, vol. 74, no. 3, pp. 249–263, 2010.

[62] T. D. Dinkova, "Tight translational control by the initiation factors eIF4E and eIF(iso)4E is required for maize seed germination," *Seed Science Research*, vol. 21, no. 2, pp. 85–93, 2011.

Hindawi Publishing Corporation
Comparative and Functional Genomics
Volume 2012, Article ID 391546, 12 pages
doi:10.1155/2012/391546

Review Article

Alternative Mechanisms to Initiate Translation in Eukaryotic mRNAs

Encarnación Martínez-Salas, David Piñeiro, and Noemí Fernández

Centro de Biología Molecular Severo Ochoa, Consejo Superior de Investigaciones Científicas, Universidad Autónoma de Madrid, Nicolás Cabrera, 1, Cantoblanco, 28049 Madrid, Spain

Correspondence should be addressed to Encarnación Martínez-Salas, emartinez@cbm.uam.es

Received 4 November 2011; Accepted 20 January 2012

Academic Editor: Greco Hernández

The composition of the cellular proteome is under the control of multiple processes, one of the most important being translation initiation. The majority of eukaryotic cellular mRNAs initiates translation by the cap-dependent or scanning mode of translation initiation, a mechanism that depends on the recognition of the $m^7G(5')ppp(5')N$, known as the cap. However, mRNAs encoding proteins required for cell survival under stress bypass conditions inhibitory to cap-dependent translation; these mRNAs often harbor internal ribosome entry site (IRES) elements in their 5′UTRs that mediate internal initiation of translation. This mechanism is also exploited by mRNAs expressed from the genome of viruses infecting eukaryotic cells. In this paper we discuss recent advances in understanding alternative ways to initiate translation across eukaryotic organisms.

1. Alternative Translation Initiation Mechanisms: An Important Layer of Gene Expression Control

The coding capacity of eukaryotic genomes is much larger than anticipated. Many layers of gene expression control operate at the posttranscriptional level, as illustrated by the RNA splicing process, the noncoding RNAs regulatory elements, and the large repertoire of factors that contribute to control mRNA transport, localization, stability, and translation. Translation control is one of the posttranscriptional cellular processes that exert a profound impact on the composition of the cellular proteome. This is particularly relevant to maintain homeostasis in response to stress induced by a large variety of environmental factors, as well as during development or disease [1]. In addition, these layers of gene expression control contribute to increase the coding capacity of the genome by generating different polypeptides from the same transcriptional unit.

The majority of cellular mRNAs initiate translation by a mechanism that depends on the recognition of the $m^7G(5')ppp(5')N$ structure (termed cap) located at the 5′end of most mRNAs (Figure 1(a)). This manner of initiating translation involves a large number of auxiliary proteins termed eukaryotic initiation factors (eIFs) [1]. The 5′cap structure is recognized by eIF4E that, in turn, is bound to the scaffold protein eIF4G and the RNA helicase eIF4A (within a trimeric complex termed eIF4F). Additionally, eIF4G further interacts with eIF3 and the poly(A)-binding protein (PABP) that is bound to the poly(A) tail of the mRNA. Separately, the 40S ribosomal subunit associates with the ternary complex (TC) consisting of the initiator methionyl-$tRNA_i$ and eIF2-GTP, leading to the formation of the 43S complex that is stabilized by eIF1A and eIF3. Following assembly of the 43S complex into the eIF4F-bound mRNA, scanning of the 5′UTR region commences until the first AUG codon is encountered, leading to the formation of the 48S initiation complex. At this step, eIF1 is displaced and eIF5 mediates the hydrolysis of eIF2-bound GTP; joining of the 60S subunit is then mediated by eIF5B yielding the 80S ribosome that gives rise to the start of polypeptide synthesis. For a review on the translation initiation process, see [1] and references therein.

Various alternative mechanisms to initiate translation are, however, operative during cellular stress [1, 2]. Furthermore, atypical mRNAs that lack the cap structure at the 5′end or the poly(A) tail at the 3′end employ specific mechanisms to initiate translation. Histones are among the most abundant proteins in eukaryotic cells, despite having mRNAs with an organization that is incompatible with the conventional scanning initiation model. Peculiar features of metazoan histone mRNAs are that they harbor a short 5′UTR and lack a poly(A) tail. Instead, these mRNAs harbor a conserved stem loop near the 3′end (Figure 1(b)) that is recognized by the stem-loop-binding protein (SLBP). In addition, the open reading frame of mouse histone mRNA contains two structural elements critical for translation initiation. One of them binds to eIF4E without the need of the cap, such that the 43S complex is recruited to this site and loaded near the AUG start codon [3]. This process is assisted by a second structural element located downstream of the initiator triplet that sequesters the cap, facilitating the direct loading of the ribosome on the cognate codon.

A different example of unconventional RNA organization is presented by a plant viral RNA, the pea enation mosaic virus, that initiates translation using a cap-independent mechanism. This mRNA does not possess a cap at the 5′end even though it requires eIF4E for translation initiation. In this particular case, the RNA structure adopts a pseudoknot structure that projects a guanosine residue into the cap-binding pocket of eIF4E [4].

In addition to the 5′-cap and 3′-poly(A) tail, eukaryotic mRNAs can contain upstream open reading frames (uORFs) (Figure 1(c)), 3′cap-independent translation elements (3′CITEs) (Figure 1(d)), or internal ribosome entry site (IRES) elements (Figures 1(e), 1(f), and 1(g)). These types of structural elements can act as sensors of environmental factors, mediate efficient translation of some viral mRNAs, or control translation of mRNAs that encode proteins performing critical roles during cell death, DNA-damage response, or in the differentiation process of higher eukaryotes and algae [5–10]. In general, these structural elements act as strong barriers to scanning ribosomes in the 5′UTR of mRNAs. Hence, their presence is incompatible with the scanning model of translation initiation, and the corresponding mRNAs have evolved different manners to initiate translation using nonconventional mechanisms.

One extensively analyzed uORF-containing mRNA is that encoding the yeast transcription factor GCN4. Translation of this particular mRNA is strongly enhanced during nutrient deprivation, an event that induces eIF2α-phosphorylation leading to low levels of TC and, thus, inhibiting cap-dependent translation. However, GCN4 mRNA overcomes the translation inhibitory effects of four uORFs under low TC levels by allowing a fraction of posttermination 40S subunits to reinitiate at the authentic GCN4 start codon [1]. In mRNAs subjected to polyamine-responsive translation regulation, uORFs act as sensors of polyamine levels [10]. Finally, in many other mRNAs, uORFs are obstacles that block or delay scanning ribosomes causing a strong reduction of protein synthesis, as exemplified in the

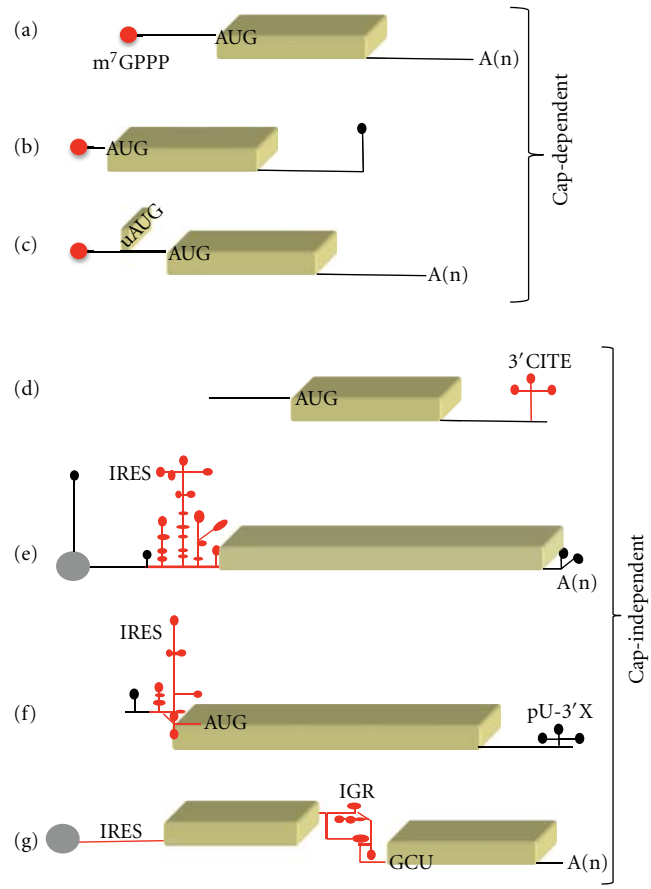


Figure 1: Schematic representation of eukaryotic mRNAs. (a) Features of a conventional mRNA. The red circle at the 5′end depicts the cap (m^7Gppp); A(n) depicts the poly(A) tail at the 3′end. (b) (c) Schematic of atypical RNA structures, stem loops (black hairpin), or uAUGs, respectively, located in mRNAs translated via cap-dependent initiation. (d) RNA structural elements located in the 3′untranslated region of viral RNAs mediating cap-independent translation (3′CITE). Different types of IRES elements found in the viral RNA of picornaviruses (e), hepatitis C (f), and dicistroviruses (g) are schematically depicted in red.

5′ terminal region of p53 mRNA or the embryonic form of the chicken proinsulin mRNA [5, 6].

Consistent with the fact that mRNA translation operates on functionally circularized molecules, regulatory elements also are located at the 3′UTR; this is the case of 3′CITE elements which are particularly abundant in plant RNA viruses [7, 11, 12] and promote cap-independent translation by means of complex RNA structures that generate functional bridges between the 5′ and 3′UTRs of the mRNA.

In summary, translation initiation mechanisms affecting the efficiency of protein synthesis of a given mRNA are diverse and, importantly, more frequent than anticipated, sometimes giving rise to the expression of different polypeptides from a single transcriptional unit. Therefore, presence of any of these regulatory elements can seriously complicate efforts to accurately define the sites of translation initiation at the genomewide scale.

2. IRES Elements in Eukaryotic Organisms and Their Viruses

IRES elements are specialized RNA regulatory sequences governing cap-independent translation initiation in eukaryotic mRNAs that are translated during cellular stress, that is, when cap-dependent translation is compromised [2, 13, 14]. IRES elements, initially reported in the genomic RNA of two picornaviruses (namely, poliovirus (PV) and encephalomyocarditis virus (EMCV)), drive internal initiation of translation in the mRNA of all members of the *Picornaviridae* family [15–20] (Table 1). Soon after their discovery, IRES elements were also found in other RNA viruses infecting mammals, such as hepatitis C (HCV), pestiviruses [21, 22], or retroviruses [23–27], as well as in RNA viruses infecting invertebrates [28–33], plants [34–39], and protozoa [40, 41]. Recently, IRES-dependent translation in mRNAs transcribed from DNA viruses belonging to the *Herpesviridae* family has been reported [42–45] (Table 1).

As expected from the fact that viruses require components of the host machinery to translate their genome, IRES elements have been found in mRNAs encoded in the genome of the host (Table 2). Thus, IRES-dependent translation initiation has been described in mRNAs expressed in animal cells, both vertebrate and invertebrate [46–69], plants, and yeasts [70–73]. Not surprisingly, various examples of IRES elements reported in animal cells, plants, or yeasts drive internal initiation of translation in mRNAs that encode proteins performing similar functions or belonging to the same regulatory pathway, for example, nutrient deprivation, apoptosis, or heat-shock (see Table 2). Published IRES elements are available at (http://iresite.org/) (http://140.135.61.9/ires/) [74, 75].

The ability of being translated under conditions inhibitory to cap-dependent initiation, though with different efficiency, is a general feature of all IRES-containing mRNAs. With notable exceptions [23, 37], IRES elements are located in the 5′UTR of mRNAs upstream of the initiator codon. Other features such as long length of the 5′UTR (200 to 500 nucleotides), heavy RNA structure, high GC content, initiation at non-AUG codons, ignored AUGs upstream of the functional start codon are often but not universally found [46, 53, 76, 77]. For instance, some IRES elements found in plant RNA viruses, *D. melanogaster* and *S. cerevisiae,* have a high AU content [37, 62, 63, 73]. In this regard, the 5′UTR sequences of mRNAs are highly divergent [78]. This is consistent with the great sequence diversity of the currently known IRES elements (http://rfam.sanger.ac.uk/) found in mRNAs of different organisms, including viruses, protozoa, yeast, plants, and animals. Furthermore, this lack of sequence conservation creates serious problems to predict the presence of IRES elements in eukaryotic mRNAs using computational methods. Therefore, defining critical hints (short primary sequences, tertiary structure elements, unique RNA-binding protein motifs, etc.) of model IRES elements is crucial to accurately predict putative IRES elements at the genomic level.

3. Types of IRES Elements

Two picornavirus RNAs, PV and EMCV, contain the first reported IRES elements [15, 16]. This property was later extended to all picornavirus RNAs as well as to several positive-strand RNA viruses, such as HCV, pestiviruses, and dicistroviruses [79]. Nevertheless, it is remarkable that despite performing the same function, viral IRES elements differ in nucleotide sequence, RNA secondary structure, and trans-acting factors requirement. A distinctive feature of the picornavirus IRES is their long length that varies between 350 to 450 nucleotides, depending on the virus genera. Furthermore, picornavirus IRES elements are classified in four types (termed I, II, III, and HCV-like) according to their RNA structure organization. The genome of picornaviruses consists of a single-stranded RNA of positive polarity that harbors a short poly(A) tail at the 3′end (Figure 1(e)). However, picornavirus RNAs differ from cellular mRNAs in having a long, heavily structured 5′UTR and a viral-encoded protein (VPg) covalently linked to the 5′end instead of cap, hence, incompatible with the cap-dependent mechanism of translation initiation. Not surprisingly, translation of the viral genome is governed by the IRES element using a cap-independent mechanism that is resistant to the action of viral proteases. Picornavirus-encoded proteases execute the processing of the viral polyprotein but also recognize as substrates several host factors. Among the host factors proteolyzed during infection are eIF4G and PABP [80, 81], which are key components of the cap-dependent translation initiation machinery. Thereby, cleavage of host factors induces the shut-off of cap-dependent translation in infected cells.

Picornavirus elements IRES belonging to types I and II require the C-terminal end of eIF4G, eIF4A, and eIF3 to assemble 48S initiation complexes [82–84]. Type III IRES require intact eIF4G, and, in contrast, the HCV-like IRES does not need eIF4G to assemble 48S complexes [18]. In addition to eIFs, auxiliary factors termed IRES transacting factors (ITAFs) contribute to modulate (either stimulate or repress) picornavirus IRES activity. In support of the relevance of factors different than eIFs for internal initiation, transcripts encompassing the region interacting with eIFs do not possess IRES activity [85], indicating that interaction with eIFs is necessary but not sufficient for IRES function.

The HCV viral RNA does not possess poly(A) tail; instead, a poly(U) tract and a complex RNA structure are located near the 3′end (Figure 1(f)). The HCV IRES element is located close to the 5′end of the viral genome and differs profoundly in RNA structure organization from picornavirus IRES belonging to types I, II, and III. Specifically, the HCV IRES (340 nucleotides) is arranged in structural domains II, III, and IV, including a pseudoknot upstream of the AUG start codon [86] which is conserved with pestivirus IRES elements [22]. Domain III participates in the interaction with eIF3 and the 40S subunit, while domain II helps to accommodate the mRNA in the tRNA-exit site of the ribosome and mediates eIF2 release during 80S assembly [87, 88]. Interestingly, HCV-like IRES elements with similar eIF requirement have been found in some picornavirus genera [18, 89, 90], which presumably arose by recombination events.

Table 1: Distribution of IRES elements in viral mRNAs.

Host	*Virus family*/genus	Virus/IRES name	Reference
Mammals	*Picornaviridae*/Enterovirus	Poliovirus (PV)	[15]
	Picornaviridae/Cardiovirus	Encephalomyelitis virus (EMCV)	[16]
	Picornaviridae/Aphthovirus	Foot-and-mouth disease virus (FMDV)	[17]
	Picornaviridae/Teschovirus	Porcine teschovirus-1 (PTV-1)	[18]
	Picornaviridae/Kobuvirus	Aichivirus (AiV)	[19]
	Picornaviridae/Senecavirus	Seneca Valley virus (SVV)	[20]
	Flaviviridae/Hepacivirus	Hepatitis C virus (HCV)	[21]
	Flaviviridae /Pestivirus	Classical swine fever virus (CSFV)	[22]
	Retroviridae/Lentivirus	Human immunodeficiency virus-2 (HIV-2)	[23]
	Retroviridae/Lentivirus	Human immunodeficiency virus-1 (HIV-1)	[24]
	Retroviridae/Retrovirus	Moloney murine leukemia virus (MoMLV)	[25]
	Retroviridae/Lentivirus	Feline immunodeficiency virus (FIV)	[26]
	Retroviridae/Retrovirus	Mouse mammary tumor virus (MMTV)	[27]
	Herpesviridae/Cytomegalovirus	Human cytomegalovirus latency (pUL138)	[42]
	Herpesv./Lymphocryptovirus	Epstein-Barr virus (EBNA-1)	[44]
	Herpesv./Mardivirus	Herpes virus Marek's disease (MDV RLORF9)	[45]
	Papovaviridae/Polyomavirus	SV40 polycistronic 19S (SV40 19S)	[43]
Insects	*Dicistroviridae*/Cripavirus	Rhopalosiphum padi virus (RhPV)	[28]
	Dicistroviridae/Cripavirus	Cricket paralysis virus (CrPV)	[29]
	Dicistroviridae/Cripavirus	Ectropis obliqua picorna-like virus (EoPV)	[31]
	Dicistroviridae/Cripavirus	Plautia stali intestine virus (PSIV)	[32]
	Dicistroviridae/Cripavirus	Triatoma virus (TrV)	[33]
	Dicistroviridae/Aparavirus	Bee paralysis dicistrovirus (IAPV, KBV)	[30]
Plants	*Comoviridae*/Nepovirus	Black currant reversion virus (BRV)	[34]
	Tombusviridae/Carmovirus	Pelargonium flower break virus (PFBV)	[35]
	Tombusviridae/Carmovirus	Hibiscus chlorotic ringspot virus (HCRSV)	[38]
	Tobamovirus	Crucifer-infecting tobamovirus (CrTMV)	[36]
	Luteoviridae/Polerovirus	Potato leaf roll polerovirus (PLRV)	[37]
	Potyviridae/Potyvirus	Tobacco etch virus (TEV)	[39]
Protozoa	*Totiviridae*/Giardiavirus	Giardiavirus (GLV)	[40]
	Totiviridae/Leishmaniavirus	Leishmania RNA virus-1 (LRV-1)	[41]

A unique type of IRES element is located in the intergenic region (IGR) of the genome of dicistroviruses (Figure 1(g)). This RNA region spans about 200 nucleotides and adopts a tertiary structure including three pseudoknots that functionally substitute the initiator tRNA during internal initiation [29, 91]. The IGR mimics a tRNA anticodon loop base-paired to mRNA and a translation elongation factor, facilitating initiation without the help of eIFs. A unique feature of these IRES elements is to initiate protein synthesis at non-AUG codons (CUU, GCU, CCU, CUC, depending on the dicistrovirus genus) [92, 93], with preference for alanine-coding triplets.

IRES elements were reported in various cellular mRNAs that remained attached to polysomes under conditions inhibitory to cap-dependent translation [51, 94]. These mRNAs contain a cap at the 5′end although they are translated at very low levels and have the capacity to switch to an IRES-dependent mechanism when cap-dependent initiation is impaired. This process is assisted by ITAFs, a group of RNA-binding proteins that are thought to help in the proper folding of the IRES region facilitating the mRNA recruitment to the translation machinery.

Thus, attending to the essential requirements for internal initiation, IRES elements can be grouped in two main categories: (a) those that do not need proteins to assemble the initiation complex (e.g., the IGR of dicistroviruses that adopts a docking structure capable of fitting in the ribosomal subunit [95]) and (b) those that do need factors to recruit the ribosome (typically, picornaviruses, HCV, and cellular IRES elements [60, 96, 97]). Within the second category, distinct groups can be made depending on the RNA structural motifs and proteins required for activity.

Table 2: Distribution of IRES elements in cellular mRNAs.

Organism	Protein function	IRES name	Reference
Mammals	Apoptotic proteins	Apaf-1	[60]
		XIAP	[56]
		HIAP2/c-IAP1	[57]
		DAP5	[58]
		Bcl-2	[61]
	Oncogene	c-myc	[77]
	Amino acid starvation	CAT-1	[46]
	Nutrient signaling	INR	[59]
	Differentiation	LEF-1	[65]
		PDGF2	[66]
	Hypoxia	HIF-1a	[50]
		VEGF	[55]
		FGF2	[53]
	Heat shock	BiP	[51]
		BAG-1	[49]
	Cold shock	CIRP	[48]
	DNA damage response	p53	[47]
		SHMT1	[54]
	Mitosis	PITSLREp58	[50]
		CDK1	[61]
Insects	Apoptotic proteins,	Rpr, hid	[62]
	Heat shock	hsp70	[63]
		grim, skl	
	Homeotic protein	Antennapedia	[64]
	Insulin signaling	dFoxO	[67]
		dInR	[68]
	Alcohol dehydrogenase	Adh-Adhr	[69]
Plants	Heat shock	HSP101	[70]
	Alcohol dehydrogenase	ADH	[71]
Yeast	Nitrogen assimilation	URE-2	[72]
	glucose starvation	GPR1, NCE102	[73]
		YMR181a, MSN1	
		BOI1, FLO8, GIC1	

4. RNA Structural Motifs Found in IRES Elements

RNA structure plays a fundamental role in viral IRES-dependent translation initiation [13]. In support of this, mutations leading to the disruption of specific RNA structure motifs impaired IRES activity while the corresponding compensatory mutations restored IRES function [21, 98]. Furthermore, RNA structure of viral IRES elements is organized in modules which are phylogenetically conserved [99–102], providing evidence in favour of a distribution of functions among the different RNA domains [103, 104].

Examples of structural motifs found in IRES elements are the pseudoknots (Pks). These are tertiary motifs that play important role in the IRES of HCV, bovine viral diarrhea virus (BVDV), and classical swine fever virus (CSFV) [22]. Three different Pks conform the IGR of cricket paralysis virus (CrPV) and plautia stali intestine virus (PSIV), as well as other dicistroviruses [91, 93]. Giardiavirus (GLV) and tobacco etch virus (TEV) IRES elements also contain Pk structures [39, 40]. While these IRES elements are located in genetically distant RNA viruses, the Pk structure is conserved, indicating that the RNA organization is biologically relevant for internal initiation. Some cellular IRES elements were reported to contain Pk structures, as illustrated by c-myc and L-myc [77, 105]. In further support of the role of RNA structure for IRES-dependent translation, the zipper model proposed for the cationic amino acid transporter CAT-1 mRNA suggested that RNA structure modification via translation of an upstream uORF induces the formation of the active IRES [46].

Evidence for tight links between RNA structure and biological function are provided by the conservation of structural motifs within IRES elements of highly variable genomes. Specifically, the purine-rich GNRA and RAAA motifs (N stands for any nucleotide, and R, purine), as well as the G : C-rich stems that hold these motifs are conserved between aphthovirus and cardiovirus [106, 107]. In the case of foot-and-mouth disease virus (FMDV), the prototype of the aphthovirus genus, the central IRES domain is a self-folding region that has been proposed to instruct the functional conformation of the whole IRES element [85, 108, 109]. Along this idea, RNA structural analysis provided evidence for stem loops whose structural conformation depends on distant interactions within this domain, involving residues of the GNRA motif [110]. Thus, it is likely that RNA structural motifs located in the apical region of the central domain could constitute a signature of picornavirus type II IRES elements.

The GNRA motif of picornavirus IRES adopts a tetraloop conformation at the tip of a stem loop [111–113]. This motif is essential for IRES activity in FMDV and EMCV [114, 115], showing a strong preference for GUAA in the case of the FMDV IRES [108]. This observation, together with the lack of genetic variability within the apical stem and the covariation observed in the adjacent stems, pointed towards their joint contribution to IRES activity. Importantly, mutational analysis of the invariant apical stem revealed a better performance of G : C than C : G base-pairs, demonstrating the relevance of the three-dimensional RNA conformation for IRES activity [106].

Although related RNA viruses share the overall organization of their genomic RNA, viral IRES is organized in high-order structures that differ between distant families. Cryo-electron microscopy studies of the HCV IRES and the IGR provided information on the capacity of these RNAs to be accommodated in the interface of the ribosomal subunits [88, 95]. Even though the IGR of CrPV and the HCV IRES exhibit different structural organization, they interact with the ribosomal protein RpS25 [116] and induce similar conformational changes in the 40S ribosomal subunit. This finding opens the possibility that IRES elements could possess a universal structural motif mediating its direct interaction with the 40S subunit. This putative universal RNA

motif still remains elusive, but it could be a promising tool to search for unidentified IRES elements at the genomic level.

Concerning the identification of structural motifs conserved between genetically distant RNAs, the IRES region of FMDV, EMCV, CrPV, HCV, CSFV, and BVDV contains a structural element recognized as substrate of the RNase P ribozyme [117–119]. RNase P is a nuclear structure-dependent endonuclease involved in the processing of the tRNA precursor, that also recognizes as substrate viral RNAs containing tRNA-like structures at the 3′end. The RNase P cleavage site in the FMDV IRES maps within an internal region that is involved in tertiary interactions; in addition, defective IRES mutants bearing modified RNA structures exhibited a differential response to ribozyme cleavage both *in vitro* and in transfected cells [119, 120]. The significance of the RNase P recognition motif in IRES elements is unknown since there is no proof for its direct involvement in the translation process. However, it does not constitute an RNA processing motif in transfected cells [121], consistent with the fact that the picornavirus infection cycle, as well as that of HCV and pestivirus, occurs in the cytoplasm of infected cells; therefore, the viral RNA has no access to RNase P.

The possibility that this structural motif constitutes a remnant of an ancient tRNA-like structure, similar to that found in the IGR IRES, is open to further investigations. Indeed, the evolutionary origin of IRES elements is unknown, but it has been proposed that this mode of initiating protein synthesis could be operating earlier than the cap-dependent [122]. In keeping with this hypothesis, the IRES property of self-interacting with the ribosome is a very attractive idea, also consistent with the finding that IRES activity is sensitive to changes in ribosome composition [123–126].

Another possibility to explain the presence of tRNA-like motifs within viral IRES elements is that they were inherited from RNA replication signals accommodated to assist in the translation process. In plant RNA viruses, tRNA-like structures located at the 3′end of the viral genome control cap-independent translation initiation and viral RNA replication [11, 127]. RNA-RNA interactions between the 3′and the 5′UTR of the viral genome assist in these processes. In this regard, long-range RNA-RNA interactions between the 5′ and the 3′end of some viral genomes have been observed [128, 129]. Consistent with a functional link between the ends of the viral RNA, IRES activity is stimulated by the 3′UTR [130, 131]. In picornavirus RNAs, the 3′UTR is composed of two stem loops and a short poly(A) tail that are required for replication and infectivity [132]. Furthermore, the insulin-like growth factor II mRNA-binding protein 1 (IGF2BP1) was identified among the proteins identified in complexes assembled with RNAs that contained the HCV IRES and the 3′UTR. This protein coimmunoprecipitates with eIF3 and the 40S subunit [133], suggesting that it enhances HCV IRES activity by recruiting the ribosomal subunits to a pseudo circularized RNA. Thus, bridging of 5′ and 3′ends involves direct RNA-RNA contacts and RNA-protein interactions. These results provide a mechanistic basis for translation stimulation and replication of the viral RNA resembling the synergistic stimulation of cap-dependent translation.

5. RNA-Protein Interactions Controlling IRES Activity

The lack of conserved features among distantly related IRES elements has led to the view that different IRES elements could recruit the ribosomal subunits assisted by unique sets of ITAFs. Along this idea, riboproteomic approaches have facilitated the identification of various proteins interacting with different IRES elements [134]. Most ITAFs are RNA-binding proteins previously identified as transcription regulators, splicing factors, RNA transport, RNA stability, or translation control proteins [133, 135, 136]. Typical examples of multifunctional proteins that act as ITAFs are the polypyrimidine tract-binding protein (PTB), the poly-r(C) binding protein (PCBP2), the SR splicing factor (SRp20), the far upstream element binding protein 2 (FBP2), the lupus La antigen (La), or Gemin5, among others [136–140].

The IRES elements of HCV and HIV-1 differ from those of picornaviruses not only in RNA structure but also in some factor requirement. Assembly of the HCV IRES-48S initiation complex requires eIF3, but not eIF4G [96]. Consistent with this, eIF3 has been identified by mass spectrometry of IRES-bound protein complexes [135, 141]. Other proteins bound to HCV and picornavirus IRES are PTB, PCBP 2, nucleolin, Gemin5, upstream of n-ras (unr), heterogeneous nuclear RNA-binding protein (hnRNP) A1/A2, La autoantigen (La), NS1-associated protein, as well as several RNA helicases DEAH-box polypeptide 9 (DHX9) [134]. Gemin5 binds directly to FMDV and HCV IRES regions and down regulates translation efficiency. Additionally, Gemin5 binds m^7GTP [142], explaining its down regulation of cap-dependent translation [140]. In contrast, the HIV-1 IRES is stimulated by hnRNP A1/2, the RNA helicase DEAD/H Box 3 (DDX3), the human Rev-interacting protein (hRIP), and the nuclear RNA-binding protein Src-associated in mitosis (Sam68) [143].

Cellular IRES elements are typically present in mRNAs encoding stress response proteins, such as those needed during nutrient deprivation, temperature shock, hibernation, hypoxia, cell cycle arrest, or apoptosis (Table 2) [2, 14, 57, 144]. However, with the exception of polypyrimidine tracts, conservation of primary sequence is not readily detected between viral and cellular IRES elements [145]. The observation that cellular IRES elements do not share overall structural similarity [146] has led to the view that short motifs may control the interaction with transacting factors needed to recruit the mRNA to 40S subunits. In favor of this hypothesis, PTB stimulates the IRES of the apoptotic protease-activating factor 1 (apaf-1), BCL2-associated athanogene (BAG)-1, and the hypoxia-inducible factor (HIF1a) [50, 147], allowing the synthesis of proteins that mediate cell survival under apoptosis, hypoxia, nutrient deprivation, or cell growth dysregulation. Proteins interacting with the lymphoid enhancer factor (LEF-1) IRES recently identified using biotin-tagged RNAs combined with stable isotope labeling

with amino acids in cell culture (SILAC)-based quantitative mass spectrometry [65] include the splicing-related protein proline and glutamine-rich SFPQ/PSF, the non-POU domain-containing octamer-binding nuclear RNA-binding protein (nonO/p54nrb), PCBP 2, HuR, and the oncoprotein DEK (named by the initials of a patient affected of acute myeloid leukemia). Investigating whether the proteins identified in intracellular IRES-ribonucleoprotein complexes perform the same or different functions from those found in IRES complexes assembled *in vitro* requires further work.

Since ITAFs usually act in large complexes with various factors within the cellular compartments, proteins interacting with different targets may lead to distinct effects depending on the target RNA and the other partners of the complex. Thus, changes in the abundance, posttranslational modifications, or subcellular location of ITAFs could be responsible for the distinct IRES response to stress conditions. For example, the kinase PITSLREp58 IRES is specifically activated during mitosis, while mPer 1 translation oscillates during circadian rhythmic period [148, 149]; the Apaf IRES is activated during apoptosis, while the X-linked inhibitor of apoptosis protein (XIAP) is inhibited [150]. Relocalization of hnRNP A1 mediates internal initiation of c-myc, unr, cyclin D1, vascular endothelial growth factor (VEGF), fibroblast growth factor (FGF-2), Apaf-1, and XIAP mRNAs [151]. In contrast, death-associated protein (DAP 5), nuclear factor NF45, G-rich RNA sequence binding factor (GRSF-1), fragile-X mental retardation protein (FMRP), dyskeratosis congenita (DKC1), heterogeneous nuclear ribonucleoprotein D-like protein (JKTBP1), or zinc-finger protein (ZNF9) are IRES-specific [57, 61, 152–154]. Hence, individual mRNAs seem to use different mechanisms to evade the global repression of protein synthesis.

6. Perspectives towards the Identification of IRES Elements at Genomewide Scale in Eukaryotes

A functional assay testing the cap-independent capacity and the ability to resist cap-inhibitory conditions is the usual way to identify IRES elements in mRNAs. This is a cumbersome task in terms of genomewide scale identification of IRES elements in eukaryotic genomes. To facilitate this task, short conserved structural motifs identified in model IRES elements could provide a tool to search for putative IRES at a genomic level. In considering critical features of IRES elements, signals that may be suggestive of functional IRES could be the presence of polypyrimidine tracts, pseudoknots near the start codon, or hairpin-loops mimicking those present in the IRES of picornavirus, HCV, or the IGR of dicistrovirus RNAs. Although none of these features individually are sufficient to define a functional IRES element, the presence of one (or more) of these motifs may provide hints to select potential IRES in mRNAs.

The question that remains unresolved is what are the distinctive features of IRES elements that may allow their accurate prediction at the genomewide scale, even though computer programs have being designed to predict IRES elements using bioinformatics tools (http://140.135.61.9/ires/) [75]. From our point of view this question is still far from being answered; however, the detailed molecular, biochemical, and structural characterization of model IRES elements will provide critical hints to reveal the presence of similar elements within eukaryotic genomes.

Acknowledgment

This work was supported by Grants BFU2008-02159, CSD2009-00080, and by an Institutional grant from Fundación Ramón Areces.

References

[1] N. Sonenberg and A. G. Hinnebusch, "Regulation of translation initiation in eukaryotes: mechanisms and biological targets," *Cell*, vol. 136, no. 4, pp. 731–745, 2009.

[2] K. A. Spriggs, M. Bushell, and A. E. Willis, "Translational regulation of gene expression during conditions of cell stress," *Molecular Cell*, vol. 40, no. 2, pp. 228–237, 2010.

[3] F. Martin, S. Barends, S. Jaeger, L. Schaeffer, L. Prongidi-Fix, and G. Eriani, "Cap-assisted internal initiation of translation of histone H4," *Molecular Cell*, vol. 41, pp. 197–209, 2011.

[4] Z. Wang, M. Parisien, K. Scheets, and W. A. Miller, "The cap-binding translation initiation factor, eIF4E, binds a pseudoknot in a viral cap-independent translation element," *Structure*, vol. 19, pp. 868–880, 2011.

[5] C. Hernandez-Sanchez, A. Mansilla, E. J. de la Rosa, G. E. Pollerberg, E. Martínez-Salas, and F. de Pablo, "Upstream AUGs in embryonic proinsulin mRNA control its low translation level," *The EMBO Journal*, vol. 22, no. 20, pp. 5582–5592, 2003.

[6] L. Blaszczyk and J. Ciesiolka, "Secondary structure and the role in translation initiation of the 5′-terminal region of p53 mRNA," *Biochemistry*, vol. 50, pp. 7080–7092, 2011.

[7] A. M. Rakotondrafara, C. Polacek, E. Harris, and W. A. Miller, "Oscillating kissing stem-loop interactions mediate 5′ scanning-dependent translation by a viral 3′-cap-independent translation element," *RNA*, vol. 12, no. 10, pp. 1893–1906, 2006.

[8] I. R. Powley, A. Kondrashov, L. A. Young et al., "Translational reprogramming following UVB irradiation is mediated by DNA-PKcs and allows selective recruitment to the polysomes of mRNAs encoding DNA repair enzymes," *Genes and Development*, vol. 23, no. 10, pp. 1207–1220, 2009.

[9] K. Babinger, A. Hallmann, and R. Schmitt, "Translational control of regA, a key gene controlling cell differentiation in Volvox carteri," *Development*, vol. 133, no. 20, pp. 4045–4051, 2006.

[10] I. P. Ivanov, J. F. Atkins, and A. J. Michael, "A profusion of upstream open reading frame mechanisms in polyamine-responsive translational regulation," *Nucleic Acids Research*, vol. 38, no. 2, pp. 353–359, 2009.

[11] V. A. Stupina, X. Yuan, A. Meskauskas, J. D. Dinman, and A. E. Simon, "Ribosome binding to a 5′ translational enhancer is altered in the presence of the 3′ untranslated region in cap-independent translation of turnip crinkle virus," *Journal of Virology*, vol. 85, pp. 4638–4653, 2011.

[12] X. Yuan, K. Shi, A. Meskauskas, and A. E. Simon, "The 3′ end of Turnip crinkle virus contains a highly interactive

structure including a translational enhancer that is disrupted by binding to the RNA-dependent RNA polymerase," *RNA*, vol. 15, no. 10, pp. 1849–1864, 2009.

[13] E. Martínez-Salas, "The impact of RNA structure on picornavirus IRES activity," *Trends in Microbiology*, vol. 16, no. 5, pp. 230–237, 2008.

[14] A. A. Komar and M. Hatzoglou, "Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states," *Cell Cycle*, vol. 10, pp. 229–240, 2011.

[15] J. Pelletier and N. Sonenberg, "Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA," *Nature*, vol. 334, no. 6180, pp. 320–325, 1988.

[16] S. K. Jang, H. G. Krausslich, M. J. H. Nicklin, G. M. Duke, A. C. Palmenberg, and E. Wimmer, "A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation," *Journal of Virology*, vol. 62, no. 8, pp. 2636–2643, 1988.

[17] E. Martinez-Salas, J. C. Saiz, M. Davila, G. J. Belsham, and E. Domingo, "A single nucleotide substitution in the internal ribosome entry site of foot-and-mouth disease virus leads to enhanced cap-independent translation in vivo," *Journal of Virology*, vol. 67, no. 7, pp. 3748–3755, 1993.

[18] A. V. Pisarev, L. S. Chard, Y. Kaku, H. L. Johns, I. N. Shatsky, and G. J. Belsham, "Functional and structural similarities between the internal ribosome entry sites of hepatitis C virus and porcine teschovirus, a picornavirus," *Journal of Virology*, vol. 78, no. 9, pp. 4487–4497, 2004.

[19] Y. Yu, T. R. Sweeney, P. Kafasla, R. J. Jackson, T. V. Pestova, and C. U. Hellen, "The mechanism of translation initiation on Aichivirus RNA mediated by a novel type of picornavirus IRES," *The EMBO Journal*, vol. 30, pp. 4423–4436, 2011.

[20] M. M. Willcocks, N. Locker, Z. Gomwalk et al., "Structural features of the Seneca valley virus internal ribosome entry site (IRES) element: a picornavirus with a pestivirus-like IRES," *Journal of Virology*, vol. 85, pp. 4452–4461, 2011.

[21] M. Honda, L. H. Ping, R. C. A. Rijnbrand et al., "Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA," *Virology*, vol. 222, no. 1, pp. 31–42, 1996.

[22] R. Rijnbrand, T. van der Straaten, P. A. van Rijn, W. J. M. Spaan, and P. J. Bredenbeek, "Internal entry of ribosomes is directed by the 5′ noncoding region of classical swine fever virus and is dependent on the presence of an RNA pseudoknot upstream of the initiation codon," *Journal of Virology*, vol. 71, no. 1, pp. 451–457, 1997.

[23] N. Locker, N. Chamond, and B. Sargueil, "A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3," *Nucleic Acids Research*, vol. 39, no. 6, pp. 2367–2377, 2011.

[24] M. Vallejos, J. Deforges, T. D. Plank et al., "Activity of the human immunodeficiency virus type 1 cell cycle-dependent internal ribosomal entry site is modulated by IRES trans-acting factors," *Nucleic Acids Research*, vol. 39, pp. 6186–6200, 2011.

[25] S. Vagner, A. Waysbort, M. Marenda, M. C. Gensac, F. Amalric, and A. C. Prats, "Alternative translation initiation of the Moloney murine leukemia virus mRNA controlled by internal ribosome entry involving the p57/PTB splicing factor," *The Journal of Biological Chemistry*, vol. 270, no. 35, pp. 20376–20383, 1995.

[26] V. Camerini, D. Decimo, L. Balvay et al., "A dormant internal ribosome entry site controls translation of feline immunodeficiency virus," *Journal of Virology*, vol. 82, no. 7, pp. 3574–3583, 2008.

[27] M. Vallejos, P. Ramdohr, F. Valiente-Echeverría et al., "The 5′-untranslated region of the mouse mammary tumor virus mRNA exhibits cap-independent translation initiation," *Nucleic Acids Research*, vol. 38, no. 2, pp. 618–632, 2009.

[28] K. E. Woolaway, K. Lazaridis, G. J. Belsham, M. J. Carter, and L. O. Roberts, "The 5′ untranslated region of Rhopalosiphum padi virus contains an internal ribosome entry site which functions efficiently in mammalian, plant, and insect translation systems," *Journal of Virology*, vol. 75, no. 21, pp. 10244–10249, 2001.

[29] J. E. Wilson, T. V. Pestova, C. U. T. Hellen, and P. Sarnow, "Initiation of protein synthesis from the A site of the ribosome," *Cell*, vol. 102, no. 4, pp. 511–520, 2000.

[30] A. E. Firth, Q. S. Wang, E. Jan, and J. F. Atkins, "Bioinformatic evidence for a stem-loop structure 5′-adjacent to the IGR-IRES and for an overlapping gene in the bee paralysis dicistroviruses," *Virology Journal*, vol. 6, article 193, 2009.

[31] J. Lu, Y. Hu, L. Hu et al., "Ectropis obliqua picorna-like virus IRES-driven internal initiation of translation in cell systems derived from different origins," *Journal of General Virology*, vol. 88, no. 10, pp. 2834–2838, 2007.

[32] N. Shibuya and N. Nakashima, "Characterization of the 5′ internal ribosome entry site of Plautia stali intestine virus," *Journal of General Virology*, vol. 87, no. 12, pp. 3679–3686, 2006.

[33] C. Czibener, D. Alvarez, E. Scodeller, and A. V. Gamarnik, "Characterization of internal ribosomal entry sites of Triatoma virus," *Journal of General Virology*, vol. 86, no. 8, pp. 2275–2280, 2005.

[34] A. Karetnikov and K. Lehto, "The RNA2 5′ leader of Blackcurrant reversion virus mediates efficient in vivo translation through an internal ribosomal entry site mechanism," *Journal of General Virology*, vol. 88, pp. 286–297, 2007.

[35] O. Fernandez-Miragall and C. Hernandez, "An internal ribosome entry site directs translation of the 3′-gene from pelargonium flower break virus genomic RNA: implications for infectivity," *PLoS ONE*, vol. 6, article e22617, 2011.

[36] Y. L. Dorokhov, P. A. Ivanov, T. V. Komarova, M. V. Skulachev, and J. G. Atabekov, "An internal ribosome entry site located upstream of the crucifer-infecting tobamovirus coat protein (CP) gene can be used for CP synthesis in vivo," *Journal of General Virology*, vol. 87, no. 9, pp. 2693–2697, 2006.

[37] H. M. Jaag, L. Kawchuk, W. Rohde, R. Fischer, N. Emans, and D. Prüfer, "An unusual internal ribosomal entry site of inverted symmetry directs expression of a potato leafroll polerovirus replication-associated protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 15, pp. 8939–8944, 2003.

[38] D. C. Y. Koh, S. M. Wong, and D. X. Liu, "Synergism of the 3′-untranslated region and an internal ribosome entry site differentially enhances the translation of a plant virus coat protein," *The Journal of Biological Chemistry*, vol. 278, no. 23, pp. 20565–20573, 2003.

[39] V. Zeenko and D. R. Gallie, "Cap-independent translation of tobacco etch virus is conferred by an RNA pseudoknot in the 5′-leader," *The Journal of Biological Chemistry*, vol. 280, no. 29, pp. 26813–26824, 2005.

[40] S. Garlapati and C. C. Wang, "Structural elements in the 5′-untranslated region of giardiavirus transcript essential for

internal ribosome entry site-mediated translation initiation," *Eukaryotic Cell*, vol. 4, no. 4, pp. 742–754, 2005.

[41] J. A. Maga, G. Widmer, and J. H. LeBowitz, "Leishmania RNA virus 1-mediated cap-independent translation," *Molecular and Cellular Biology*, vol. 15, no. 9, pp. 4884–4889, 1995.

[42] L. Grainger, L. Cicchini, M. Rak et al., "Stress-inducible alternative translation initiation of human cytomegalovirus latency protein pUL138," *Journal of Virology*, vol. 84, no. 18, pp. 9472–9486, 2010.

[43] Y. Yu and J. C. Alwine, "19S late mRNAs of simian virus 40 have an internal ribosome entry site upstream of the virion structural protein 3 coding sequence," *Journal of Virology*, vol. 80, no. 13, pp. 6553–6558, 2006.

[44] A. Isaksson, M. Berggren, K. Ekeland-Sjoberg, T. Samuelsson, and A. Ricksten, "Cell specific internal translation efficiency of Epstein-Barr virus present in solid organ transplant patients," *Journal of Medical Virology*, vol. 79, no. 5, pp. 573–581, 2007.

[45] A. Tahiri-Alaoui, L. P. Smith, S. Baigent et al., "Identification of an intercistronic internal ribosome entry site in a Marek's disease virus immediate-early gene," *Journal of Virology*, vol. 83, no. 11, pp. 5846–5853, 2009.

[46] J. Fernandez, I. Yaman, C. Huang et al., "Ribosome stalling regulates ires-mediated translation in eukaryotes, a parallel to prokaryotic attenuation," *Molecular Cell*, vol. 17, no. 3, pp. 405–416, 2005.

[47] D. Q. Yang, M. J. Halaby, and Y. Zhang, "The identification of an internal ribosomal entry site in the 5′-untranslated region of p53 mRNA provides a novel mechanism for the regulation of its translation following DNA damage," *Oncogene*, vol. 25, no. 33, pp. 4613–4619, 2006.

[48] M. B. Al-Fageeh and C. M. Smales, "Cold-inducible RNA binding protein (CIRP) expression is modulated by alternative mRNAs," *RNA*, vol. 15, no. 6, pp. 1165–1176, 2009.

[49] B. M. Pickering, S. A. Mitchell, K. A. Spriggs, M. Stoneley, and A. E. Willis, "Bag-1 internal ribosome entry segment activity is promoted by structural changes mediated by poly(rC) binding protein 1 and recruitment of polypyrimidine tract binding protein 1," *Molecular and Cellular Biology*, vol. 24, no. 12, pp. 5595–5605, 2004.

[50] B. Schepens, S. A. Tinton, Y. Bruynooghe, R. Beyaert, and S. Cornelis, "The polypyrimidine tract-binding protein stimulates HIF-1α IRES-mediated translation during hypoxia," *Nucleic Acids Research*, vol. 33, no. 21, pp. 6884–6894, 2005.

[51] D. G. Macejak and P. Sarnow, "Internal initiation of translation mediated by the 5′ leader of a cellular mRNA," *Nature*, vol. 353, no. 6339, pp. 90–94, 1991.

[52] M. Stoneley, T. Subkhankulova, J. P. C. Le Quesne et al., "Analysis of the c-myc IRES; A potential role for cell-type specific trans-acting factors and the nuclear compartment," *Nucleic Acids Research*, vol. 28, no. 3, pp. 687–694, 2000.

[53] S. Bonnal, C. Schaeffer, L. Créancier et al., "A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons," *The Journal of Biological Chemistry*, vol. 278, no. 41, pp. 39330–39336, 2003.

[54] J. T. Fox, W. K. Shin, M. A. Caudill, and P. J. Stover, "A UV-responsive internal ribosome entry site enhances serine hydroxymethyltransferase 1 expression for DNA damage repair," *The Journal of Biological Chemistry*, vol. 284, no. 45, pp. 31097–31108, 2009.

[55] M. J. Morris, Y. Negishi, C. Pazsint, J. D. Schonhoft, and S. Basu, "An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES," *Journal of the American Chemical Society*, vol. 132, no. 50, pp. 17831–17839, 2010.

[56] A. Riley, L. E. Jordan, and M. Holcik, "Distinct 5′ UTRs regulate XIAP expression under normal growth conditions and during cellular stress," *Nucleic Acids Research*, vol. 38, no. 14, pp. 4665–4674, 2010.

[57] T. E. Graber, S. D. Baird, P. N. Kao, M. B. Mathews, and M. Holcik, "NF45 functions as an IRES transacting factor that is required for translation of cIAP1 during the unfolded protein response," *Cell Death and Differentiation*, vol. 17, no. 4, pp. 719–729, 2010.

[58] S. Henis-Korenblit, G. Shani, T. Sines, L. Marash, G. Shohat, and A. Kimchi, "The caspase-cleaved DAP5 protein supports internal ribosome entry site-mediated translation of death proteins," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 8, pp. 5400–5405, 2002.

[59] K. A. Spriggs, L. C. Cobbold, S. H. Ridley et al., "The human insulin receptor mRNA contains a functional internal ribosome entry segment," *Nucleic Acids Research*, vol. 37, no. 17, pp. 5881–5893, 2009.

[60] S. A. Mitchell, K. A. Spriggs, M. J. Coldwell, R. J. Jackson, and A. E. Willis, "The Apaf-1 internal ribosome entry segment attains the correct structural conformation for function via interactions with PTB and unr," *Molecular Cell*, vol. 11, no. 3, pp. 757–771, 2003.

[61] L. Marash, N. Liberman, S. Henis-Korenblit et al., "DAP5 promotes cap-independent translation of Bcl-2 and CDK1 to facilitate cell survival during mitosis," *Molecular Cell*, vol. 30, no. 4, pp. 447–459, 2008.

[62] G. Hernández, P. Vázquez-Pianzola, J. M. Sierra, and R. Rivera-Pomar, "Internal ribosome entry site drives cap-independent translation of reaper and heat shock protein 70 mRNAs in Drosophila embryos," *RNA*, vol. 10, no. 11, pp. 1783–1797, 2004.

[63] P. Vazquez-Pianzola, G. Hernández, B. Suter, and R. Rivera-Pomar, "Different modes of translation for hid, grim and sickle mRNAs in Drosophila," *Cell Death and Differentiation*, vol. 14, no. 2, pp. 286–295, 2007.

[64] S. K. Oh, M. P. Scott, and P. Sarnow, "Homeotic gene Antennapedia mRNA contains 5′-noncoding sequences that confer translational initiation by internal ribosome binding," *Genes and Development*, vol. 6, no. 9, pp. 1643–1653, 1992.

[65] B. P. Tsai, X. Wang, L. Huang, and M. L. Waterman, "Quantitative profiling of in vivo-assembled RNA-protein complexes using a novel integrated proteomic approach," *Molecular & Cellular Proteomics*, vol. 10, no. 4, article M110 007385, 2011.

[66] O. Sella, G. Gerlitz, S. Y. Le, and O. Elroy-Stein, "Differentiation-induced internal translation of c-sis mRNA: analysis of the cis elements and their differentiation-linked binding to the hnRNP C protein," *Molecular and Cellular Biology*, vol. 19, no. 8, pp. 5429–5440, 1999.

[67] E. Villa-Cuesta, B. T. Sage, and M. Tatar, "A role for drosophila dFoxO and dFoxO 5′UTR internal ribosomal entry sites during fasting," *PLoS ONE*, vol. 5, no. 7, article e11521, 2010.

[68] M. T. Marr, J. A. D'Alessio, O. Puig, and R. Tjian, "IRES-mediated functional coupling of transcription and translation amplifies insulin receptor feedback," *Genes and Development*, vol. 21, no. 2, pp. 175–183, 2007.

[69] P. Ramanathan, J. Guo, R. N. Whitehead, and S. Brogna, "The intergenic spacer of the Drosophila Adh-Adhr dicistronic

mRNA stimulates internal translation initiation," *RNA Biology*, vol. 5, no. 3, pp. 149–156, 2008.

[70] T. D. Dinkova, H. Zepeda, E. Martinez-Salas, L. M. Martinez, J. Nieto-Sotelo, and S. E. de Jimenez, "Cap-independent translation of maize Hsp101," *Plant Journal*, vol. 41, no. 5, pp. 722–731, 2005.

[71] E. S. Mardanova, L. A. Zamchuk, M. V. Skulachev, and N. V. Ravin, "The 5′ untranslated region of the maize alcohol dehydrogenase gene contains an internal ribosome entry site," *Gene*, vol. 420, no. 1, pp. 11–16, 2008.

[72] L. C. Reineke and W. C. Merrick, "Characterization of the functional role of nucleotides within the URE2 IRES element and the requirements for eIF2A-mediated repression," *RNA*, vol. 15, no. 12, pp. 2264–2277, 2009.

[73] W. V. Gilbert, K. Zhou, T. K. Butler, and J. A. Doudna, "Cap-independent translation is required for starvation-induced differentiation in yeast," *Science*, vol. 317, no. 5842, pp. 1224–1227, 2007.

[74] M. Mokrejs, V. Vopalensky, O. Kolenaty et al., "IRESite: the database of experimentally verified IRES structures (http://www.iresite.org/)," *Nucleic Acids Research*, vol. 34, pp. D125–D130, 2006.

[75] T. Y. Wu, C. C. Hsieh, J. J. Hong, C. Y. Chen, and Y. S. Tsai, "IRSS: a web-based tool for automatic layout and analysis of IRES secondary structure prediction and searching system in silico," *BMC Bioinformatics*, vol. 10, article 160, 2009.

[76] S. López de Quinto and E. Martínez-Salas, "Involvement of the aphthovirus RNA region located between the two functional AUGs in start codon selection," *Virology*, vol. 255, no. 2, pp. 324–336, 1999.

[77] J. P. C. Le Quesne, M. Stoneley, G. A. Fraser, and A. E. Willis, "Derivation of a structural model for the c-myc IRES," *Journal of Molecular Biology*, vol. 310, no. 1, pp. 111–126, 2001.

[78] G. Grillo, A. Turi, F. Licciulli et al., "UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs," *Nucleic Acids Research*, vol. 38, no. 1, pp. D75–D80, 2009.

[79] E. Martínez-Salas, A. Pacheco, P. Serrano, and N. Fernandez, "New insights into internal ribosome entry site elements relevant for viral gene expression," *Journal of General Virology*, vol. 89, no. 3, pp. 611–626, 2008.

[80] M. Rodríguez Pulido, P. Serrano, M. Sáiz, and E. Martínez-Salas, "Foot-and-mouth disease virus infection induces proteolytic cleavage of PTB, eIF3a,b, and PABP RNA-binding proteins," *Virology*, vol. 364, no. 2, pp. 466–474, 2007.

[81] E. Martinez-Salas and M. Ryan, "Translation and protein processing," in *Picornaviruses*, E. Ehrenfeld, E. Domingo, and R. Roos, Eds., pp. 141–161, ASM Press, 2010.

[82] D. E. Andreev, O. Fernandez-Miragall, J. Ramajo et al., "Differential factor requirement to assemble translation initiation complexes at the alternative start codons of foot-and-mouth disease virus RNA," *RNA*, vol. 13, no. 8, pp. 1366–1374, 2007.

[83] S. de Breyne, Y. Yu, A. Unbehaun, T. V. Pestova, and C. U. T. Hellen, "Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 23, pp. 9197–9202, 2009.

[84] S. López de Quinto, E. Lafuente, and E. Martínez-Salas, "IRES interaction with translation initiation factors: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII," *RNA*, vol. 7, no. 9, pp. 1213–1226, 2001.

[85] O. Fernandez-Miragall, S. Lopez de Quinto, and E. Martinez-Salas, "Relevance of RNA structure for the activity of picornavirus IRES elements," *Virus Research*, vol. 139, pp. 172–182, 2009.

[86] K. E. Berry, S. Waghray, S. A. Mortimer, Y. Bai, and J. A. Doudna, "Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning," *Structure*, vol. 19, pp. 1456–1466, 2011.

[87] N. Locker, L. E. Easton, and P. J. Lukavsky, "HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly," *The EMBO Journal*, vol. 26, no. 3, pp. 795–805, 2007.

[88] C. M. T. Spahn, J. S. Kieft, R. A. Grassucci et al., "Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit," *Science*, vol. 291, no. 5510, pp. 1959–1962, 2001.

[89] S. de Breyne, Y. Yu, T. V. Pestova, and C. U. T. Hellen, "Factor requirements for translation initiation on the Simian picornavirus internal ribosomal entry site," *RNA*, vol. 14, no. 2, pp. 367–380, 2008.

[90] L. E. Easton, N. Locker, and P. J. Lukavsky, "Conserved functional domains and a novel tertiary interaction near the pseudoknot drive translational activity of hepatitis C virus and hepatitis C virus-like internal ribosome entry sites," *Nucleic Acids Research*, vol. 37, no. 16, pp. 5537–5549, 2009.

[91] D. A. Costantino, J. S. Pfingsten, R. P. Rambo, and J. S. Kieft, "tRNA-mRNA mimicry drives translation initiation from a viral IRES," *Nature Structural and Molecular Biology*, vol. 15, no. 1, pp. 57–64, 2008.

[92] J. Sasaki and N. Nakashima, "Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus in vitro," *Journal of Virology*, vol. 73, no. 2, pp. 1219–1226, 1999.

[93] N. Nakashima and T. Uchiumi, "Functional analysis of structural motifs in dicistroviruses," *Virus Research*, vol. 139, no. 2, pp. 137–147, 2009.

[94] G. Johannes, M. S. Carter, M. B. Eisen, P. O. Brown, and P. Sarnow, "Identification of eukaryotic mRNAs that are translated at reduced cap binding complex elF4F concentrations using a cDNA microarray," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 23, pp. 13118–13123, 1999.

[95] C. M. T. Spahn, E. Jan, A. Mulder, R. A. Grassucci, P. Sarnow, and J. Frank, "Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor," *Cell*, vol. 118, no. 4, pp. 465–475, 2004.

[96] T. V. Pestova, I. N. Shatsky, S. P. Fletcher, R. J. Jackson, and C. U. T. Hellen, "A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initation of hepatitis C and classical swine fever virus RNAs," *Genes and Development*, vol. 12, no. 1, pp. 67–83, 1998.

[97] S. Lopez de Quinto and E. Martinez-Salas, "Interaction of the elF4G initiation factor with the aphthovirus IRES is essential for internal translation initiation in vivo," *RNA*, vol. 6, no. 10, pp. 1380–1392, 2000.

[98] E. Martinez-Salas, M. P. Regalado, and E. Domingo, "Identification of an essential region for internal initiation of translation in the aphthovirus internal ribosome entry site and implications for viral evolution," *Journal of Virology*, vol. 70, no. 2, pp. 992–998, 1996.

[99] E. Martinez-Salas and O. Fernandez-Miragall, "Picornavirus IRES: structure function relationship," *Current Pharmaceutical Design*, vol. 10, pp. 3757–3767, 2004.

[100] M. Honda, R. Rijnbrand, G. Abell, D. Kim, and S. M. Lemon, "Natural variation in translational activities of the 5′ nontranslated RNAs of hepatitis C virus genotypes 1a and 1b: evidence for a long-range RNA- RNA interaction outside of the internal ribosomal entry site," *Journal of Virology*, vol. 73, no. 6, pp. 4941–4951, 1999.

[101] J. C. Saiz, S. Lopez de Quinto, N. Ibarrola et al., "Internal initiation of translation efficiency in different hepatitis C genotypes isolated from interferon treated patients," *Archives of Virology*, vol. 144, pp. 215–229, 1999.

[102] M. I. Barria, A. Gonzalez, J. Vera-Otarola et al., "Analysis of natural variants of the hepatitis C virus internal ribosome entry site reveals that primary sequence plays a key role in cap-independent translation," *Nucleic Acids Research*, vol. 37, pp. 957–971, 2009.

[103] P. Serrano, J. Ramajo, and E. Martínez-Salas, "Rescue of internal initiation of translation by RNA complementation provides evidence for a distribution of functions between individual IRES domains," *Virology*, vol. 388, no. 1, pp. 221–229, 2009.

[104] C. J. Jang and E. Jan, "Modular domains of the Dicistroviridae intergenic internal ribosome entry site," *RNA*, vol. 16, no. 6, pp. 1182–1195, 2010.

[105] C. L. Jopling, K. A. Spriggs, S. A. Mitchell, M. Stoneley, and A. E. Willis, "L-Myc protein synthesis is initiated by internal ribosome entry," *RNA*, vol. 10, no. 2, pp. 287–298, 2004.

[106] N. Fernandez, O. Fernandez-Miragall, J. Ramajo et al., "Structural basis for the biological relevance of the invariant apical stem in IRES-mediated translation," *Nucleic Acids Research*, vol. 39, pp. 8572–8585, 2011.

[107] Y. Yu, I. S. Abaeva, A. Marintchev, T. V. Pestova, and C. U. Hellen, "Common conformational changes induced in type 2 picornavirus IRESs by cognate trans-acting factors," *Nucleic Acids Research*, vol. 39, pp. 4851–4865, 2011.

[108] N. Fernandez, A. Garcia-Sacristan, J. Ramajo, C. Briones, and E. Martinez-Salas, "Structural analysis provides insights into the modular organization of picornavirus IRES," *Virology*, vol. 409, pp. 251–261, 2011.

[109] R. Ramos and E. Martínez-Salas, "Long-range RNA interactions between structural domains of the aphthovirus internal ribosome entry site (IRES)," *RNA*, vol. 5, no. 10, pp. 1374–1383, 1999.

[110] O. Fernandez-Miragall, R. Ramos, J. Ramajo, and E. Martinez-Salas, "Evidence of reciprocal tertiary interactions between conserved motifs involved in organizing RNA structure essential for internal initiation of translation," *RNA*, vol. 12, pp. 223–234, 2006.

[111] O. Fernandez-Miragall and E. Martinez-Salas, "Structural organization of a viral IRES depends on the integrity of the GNRA motif," *RNA*, vol. 9, pp. 1333–1344, 2003.

[112] M. Phelan, R. J. Banks, G. Conn, and V. Ramesh, "NMR studies of the structure and Mg^{2+} binding properties of a conserved RNA motif of EMCV picornavirus IRES element," *Nucleic Acids Research*, vol. 32, no. 16, pp. 4715–4724, 2004.

[113] Z. Du, N. B. Ulyanov, J. Yu, R. Andino, and T. L. James, "NMR structures of loop B RNAs from the stem-loop IV domain of the enterovirus internal ribosome entry site: a single C to U substitution drastically changes the shape and flexibility of RNA," *Biochemistry*, vol. 43, no. 19, pp. 5757–5771, 2004.

[114] S. Lopez de Quinto and E. Martinez-Salas, "Conserved structural motifs located in distal loops of aphthovirus internal ribosome entry site domain 3 are required for internal initiation of translation," *Journal of Virology*, vol. 71, pp. 4171–4175, 1997.

[115] M. E. M. Robertson, R. A. Seamons, and G. J. Belsham, "A selection system for functional internal ribosome entry site (IRES) elements: analysis of the requirement for a conserved GNRA tetraloop in the encephalomyocarditis virus IRES," *RNA*, vol. 5, no. 9, pp. 1167–1179, 1999.

[116] D. M. Landry, M. I. Hertz, and S. R. Thompson, "RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs," *Genes and Development*, vol. 23, no. 23, pp. 2753–2764, 2009.

[117] A. Nadal, M. Martell, J. R. Lytle et al., "Specific cleavage of hepatitis C virus RNA genome by human RNase P," *The Journal of Biological Chemistry*, vol. 277, no. 34, pp. 30606–30613, 2002.

[118] A. J. Lyons and H. D. Robertson, "Detection of tRNA-like structure through RNase P cleavage of viral internal ribosome entry site RNAs near the AUG start triplet," *The Journal of Biological Chemistry*, vol. 278, no. 29, pp. 26844–26850, 2003.

[119] P. Serrano, J. Gomez, and E. Martínez-Salas, "Characterization of a cyanobacterial RNase P ribozyme recognition motif in the IRES of foot-and-mouth disease virus reveals a unique structural element," *RNA*, vol. 13, no. 6, pp. 849–859, 2007.

[120] N. Fernandez and E. Martinez-Salas, "Tailoring the switch from IRES-dependent to 5′-end-dependent translation with the RNase P ribozyme," *RNA*, vol. 16, pp. 852–862.

[121] M. Piron, N. Beguiristain, A. Nadal, E. Martinez-Salas, and J. Gomez, "Characterizing the function and structural organization of the 5′ tRNA-like motif within the hepatitis C virus quasispecies," *Nucleic Acids Research*, vol. 33, pp. 1487–1502, 2005.

[122] G. Hernández, "Was the initiation of translation in early eukaryotes IRES-driven?" *Trends in Biochemical Sciences*, vol. 33, no. 2, pp. 58–64, 2008.

[123] F. Martínez-Azorín, M. Remacha, E. Martínez-Salas, and J. P. G. Ballesta, "Internal translation initiation on the foot-and-mouth disease virus IRES is affected by ribosomal stalk conformation," *FEBS Letters*, vol. 582, no. 20, pp. 3029–3032, 2008.

[124] C. Yang, C. Zhang, J. D. Dittman, and S. A. Whitham, "Differential requirement of ribosomal protein S6 by plant RNA viruses with different translation initiation strategies," *Virology*, vol. 390, no. 2, pp. 163–173, 2009.

[125] A. Basu, P. Das, S. Chaudhuri et al., "Requirement of rRNA methylation for 80S ribosome assembly on a cohort of cellular Internal Ribosome Entry Sites," *Molecular and Cellular Biology*, vol. 31, pp. 4482–4499, 2011.

[126] R. Horos, H. Ijspeert, D. Pospisilova et al., "Ribosomal deficiencies in Diamond-Blackfan anemia impair translation of transcripts essential for differentiation of murine and human erythroblasts," *Blood*, vol. 119, pp. 262–272, 2012.

[127] X. Yuan, K. Shi, M. Y. L. Young, and A. E. Simon, "The terminal loop of a 3′ proximal hairpin plays a critical role in replication and the structure of the 3′ region of Turnip crinkle virus," *Virology*, vol. 402, no. 2, pp. 271–280, 2010.

[128] P. Serrano, M. R. Pulido, M. Saiz, and E. Martinez-Salas, "The 3′ end of the foot-and-mouth disease virus genome establishes two distinct long-range RNA-RNA interactions with the 5′ end region," *Journal of General Virology*, vol. 87, pp. 3013–3022, 2006.

[129] C. Romero-López and A. Berzal-Herranz, "A long-range RNA-RNA interaction between the 5′ and 3′ ends of the HCV genome," *RNA*, vol. 15, no. 9, pp. 1740–1752, 2009.

[130] S. Lopez de Quinto, M. Saiz, D. de la Morena, F. Sobrino, and E. Martinez-Salas, "IRES-driven translation is stimulated separately by the FMDV 3′-NCR and poly(A) sequences," *Nucleic Acids Research*, vol. 30, pp. 4398–4405, 2002.

[131] E. Dobrikova, P. Florez, S. Bradrick, and M. Gromeier, "Activity of a type 1 picornavirus internal ribosomal entry site is determined by sequences within the 3′ nontranslated region," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 25, pp. 15125–15130, 2003.

[132] M. Saiz, S. Gomez, E. Martinez-Salas, and F. Sobrino, "Deletion or substitution of the aphthovirus 3′ NCR abrogates infectivity and virus replication," *Journal of General Virology*, vol. 82, pp. 93–101, 2001.

[133] S. Weinlich, S. Hüttelmaier, A. Schierhorn, S. E. Behrens, A. Ostareck-Lederer, and D. H. Ostareck, "IGF2BP1 enhances HCV IRES-mediated translation initiation via the 3′UTR," *RNA*, vol. 15, no. 8, pp. 1528–1542, 2009.

[134] A. Pacheco and E. Martinez-Salas, "Insights into the biology of IRES elements through riboproteomic approaches," *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 458927, 12 pages, 2010.

[135] A. Pacheco, S. Reigadas, and E. Martínez-Salas, "Riboproteomic analysis of polypeptides interacting with the internal ribosome-entry site element of foot-and-mouth disease viral RNA," *Proteomics*, vol. 8, no. 22, pp. 4782–4790, 2008.

[136] P. Vazquez-Pianzola, H. Urlaub, and R. Rivera-Pomar, "Proteomic analysis of reaper 5′ untranslated region-interacting factors isolated by tobramycin affinity-selection reveals a role for La antigen in reaper mRNA translation," *Proteomics*, vol. 5, no. 6, pp. 1645–1655, 2005.

[137] S. K. Jang and E. Wimmer, "Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein," *Genes and Development*, vol. 4, no. 9, pp. 1560–1572, 1990.

[138] K. M. Bedard, S. Daijogo, and B. L. Semler, "A nucleocytoplasmic SR protein functions in viral IRES-mediated translation initiation," *The EMBO Journal*, vol. 26, no. 2, pp. 459–467, 2007.

[139] J. Y. Lin, M. L. Li, and S. R. Shih, "Far upstream element binding protein 2 interacts with enterovirus 71 internal ribosomal entry site and negatively regulates viral translation," *Nucleic Acids Research*, vol. 37, no. 1, pp. 47–59, 2009.

[140] A. Pacheco, S. Lopez de Quinto, J. Ramajo, N. Fernandez, and E. Martinez-Salas, "A novel role for Gemin5 in mRNA translation," *Nucleic Acids Research*, vol. 37, pp. 582–590, 2009.

[141] Y. Yu, H. Ji, J. A. Doudna, and J. A. Leary, "Mass spectrometric analysis of the human 40S ribosomal subunit: native and HCV IRES-bound complexes," *Protein Science*, vol. 14, no. 6, pp. 1438–1446, 2005.

[142] S. S. Bradrick and M. Gromeier, "Identification of gemin5 as a novel 7-methylguanosine cap-binding protein," *PLoS ONE*, vol. 4, no. 9, article e7030, 2009.

[143] J. Liu, J. Henao-Mejia, H. Liu, Y. Zhao, and J. J. He, "Translational regulation of HIV-1 replication by HIV-1 Rev cellular cofactors Sam68, eIF5A, hRIP, and DDX3," *Journal of NeuroImmune Pharmacology*, vol. 6, pp. 308–321, 2011.

[144] P. Pan and F. van Breukelen, "Preference of IRES-mediated initiation of translation during hibernation in golden-mantled ground squirrels, Spermophilus lateralis," *American Journal of Physiology*, vol. 301, pp. R370–R377, 2011.

[145] M. Bushell, M. Stoneley, Y. W. Kong et al., "Polypyrimidine tract binding protein regulates IRES-mediated gene expression during apoptosis," *Molecular Cell*, vol. 23, no. 3, pp. 401–412, 2006.

[146] X. Xia and M. Holcik, "Strong eukaryotic IRESs have weak secondary structure," *PLoS ONE*, vol. 4, no. 1, article e4136, 2009.

[147] H. C. Dobbyn, K. Hill, T. L. Hamilton et al., "Regulation of BAG-1 IRES-mediated translation following chemotoxic stress," *Oncogene*, vol. 27, no. 8, pp. 1167–1174, 2008.

[148] B. Schepens, S. A. Tinton, Y. Bruynooghe et al., "A role for hnRNP C1/C2 and Unr in internal initiation of translation during mitosis," *The EMBO Journal*, vol. 26, no. 1, pp. 158–169, 2007.

[149] K. H. Lee, K. C. Woo, D. Y. Kim et al., "Rhythmic interaction between period1 mRNA and hnRNP Q leads to circadian time-dependent translation," *Molecular and Cellular Biology*, vol. 32, pp. 717–728, 2012.

[150] N. H. Ungureanu, M. Cloutier, S. M. Lewis et al., "Internal ribosome entry site-mediated translation of Apaf-1, but not XIAP, is regulated during UV-induced cell death," *The Journal of Biological Chemistry*, vol. 281, no. 22, pp. 15155–15163, 2006.

[151] A. Cammas, F. Pileur, S. Bonnal et al., "Cytoplasmic relocalization of heterogeneous nuclear ribonucleoprotein A1 controls translation initiation of specific mRNAs," *Molecular Biology of the Cell*, vol. 18, no. 12, pp. 5048–5059, 2007.

[152] S. M. Lewis, S. Cerquozzi, T. E. Graber, N. H. Ungureanu, M. Andrews, and M. Holcik, "The eIF4G homolog DAP5/p97 supports the translation of select mRNAs during endoplasmic reticulum stress," *Nucleic Acids Research*, vol. 36, no. 1, pp. 168–178, 2008.

[153] M. A. Sammons, A. K. Antons, M. Bendjennat, B. Udd, R. Krahe, and A. J. Link, "ZNF9 activation of IRES-mediated translation of the human ODC mRNA is decreased in myotonic dystrophy type 2," *PLoS ONE*, vol. 5, no. 2, article e9301, 2010.

[154] C. Bellodi, N. Kopmar, and D. Ruggero, "Deregulation of oncogene-induced senescence and p53 translational control in X-linked dyskeratosis congenita," *The EMBO Journal*, vol. 29, no. 11, pp. 1865–1876, 2010.