Life and Death of mRNA Molecules in *Entamoeba histolytica*

Jesús Valdés-Flores¹, Itzel López-Rosas², César López-Camarillo³, Esther Ramirez-Moreno⁴, Juan D. Ospina-Villa⁵† and Laurence A. Marchat⁴*

¹ Departamento de Bioquímica, CINVESTAV, Ciudad de México, Mexico City, Mexico, ² CONACyT Research Fellow – Colegio de Postgraduados Campus Campeche, Campeche, Mexico, ³ Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, Ciudad de México, Mexico City, Mexico, ⁴ Escuela Nacional de Medicina y Homeopatía, Instituto Politécnico Nacional, Ciudad de México, Mexico City, Mexico

In eukaryotic cells, the life cycle of mRNA molecules is modulated in response to environmental signals and cell-cell communication in order to support cellular homeostasis. Capping, splicing and polyadenylation in the nucleus lead to the formation of transcripts that are suitable for translation in cytoplasm, until mRNA decay occurs in P-bodies. Although pre-mRNA processing and degradation mechanisms have usually been studied separately, they occur simultaneously and in a coordinated manner through protein-protein interactions, maintaining the integrity of gene expression. In the past few years, the availability of the genome sequence of *Entamoeba histolytica*, the protozoan parasite responsible for human amoebiasis, coupled to the development of the so-called “omics” technologies provided new opportunities for the study of mRNA processing and turnover in this pathogen. Here, we review the current knowledge about the molecular basis for splicing, 3′ end formation and mRNA degradation in amoeba, which suggest the conservation of events related to mRNA life throughout evolution. We also present the functional characterization of some key proteins and describe some interactions that indicate the relevance of cooperative regulatory events for gene expression in this human parasite.

**Keywords:** *Entamoeba*, mRNA decay, mRNA processing, P-bodies, polyadenylation, protozoan parasite, splicing

**INTRODUCTION**

The metabolism of messenger RNA (mRNA) is a complex process that is essential for gene expression regulation and mRNA turnover in response to environmental signals and cell-cell communication in eukaryotic cells. During pre-mRNA synthesis by RNA polymerase II (RNA Pol II) in the nucleus, they are modified to generate mature transcripts that can be exported to the cytoplasm and translated to proteins. First, the 5′ end of nascent mRNA is capped by a 7-methyl guanosine linked by a 5′-5′ triphosphate bridge to the first nucleoside of the transcript (capping). These reactions are catalyzed by three enzymes: RNA triphosphatase, guanylyltransferase, and RNA (guanine-7-)-methyltransferase (RNMT) (Cowling, 2010). Then, introns are removed and exons are ligated by the catalytic activity of the spliceosome components that include five small nuclear RNA (snRNAs), namely U1, U2, U4, U5, and U6, and small nuclear ribonucleic proteins (snRNPs) (splicing) (Shi, 2017). Finally, a phosphodiester bond is hydrolyzed at the 3′ end of mRNA and a poly(A) tail is added by the coordinated activity of a large set of polyadenylation factors that recognize specific motifs in RNA 3′ untranslated region (3′UTR) (cleavage/polyadenylation)
(Xiang et al., 2014). After translation, the elimination of mRNA molecules is necessary to ensure proper course of gene expression and prevent the accumulation of transcripts (Christie et al., 2011). Pathways of mRNA decay depend on the formation of RNA-protein complexes in microscopically detectable cytoplasmic structures, called processing bodies (P-bodies) (Sheth and Parker, 2003), in which mRNAs are translationally repressed (silenced) or degraded; their re-incorporation into ribosomes is also possible (Eulalio et al., 2007). Transcript decay involves 3′ end deadenylation by CAF1 and CCR4/NOT1–5 complex (or by PARN, PAN2 and PAN3 deadenylases) followed by 5′ end decapping by DCP1–DCP2 complex and Lsm1–7 proteins, and 5′-3′ digestion by exonuclease XRN1; alternatively, deadenylated transcripts can be degraded from the 5′ end by the exosome complex, while the scavenger-decapping DCP enzyme hydrolyzes the remaining cap structure (Labno et al., 2016). During translation, aberrant mRNAs with premature termination codons can be detected and eliminated through the nonsense-mediated decay (NMD) pathway (Rebbapragada and Lykke-Andersen, 2009).

Although pre-mRNA processing reactions have usually been studied separately, they occur co-transcriptionally, simultaneously and in a coordinated manner. Moreover, a large set of data has shown that they are interconnected with transcription, translation, and mRNA degradation; protein-protein interactions establish a functional link between the different molecular machineries and promote reciprocal regulation events to maintain the integrity of gene expression. Consequently, each of these processes plays a major role throughout the life cycle of mRNA. Thus, in addition to protect mRNA from 5′ to 3′ exonuclease cleavage, the m7G cap interacts with the cap-binding complex (CBC), which regulates spliceosome assembly, transcription termination, 3′ end processing, RNA export, and NMD in the nucleus. In the cytoplasm, CBC recruits eIF4G, RNA helicase eIF4A, and other proteins to promote translation initiation. Moreover, eIF4G interacts with poly(A) binding protein PABP1 bound to the poly(A) tail to create a mRNA pseudo-circularization and enhance the processivity of the ribosome. Furthermore, it has been recently demonstrated that 2′O methylated cap (cap 1) acts as a signature of self RNA molecules (Ramanathan et al., 2016). Several data indicate that U1 snRNP, the more abundant splicing factor, inhibits 3′ end processing. Notably, its interaction with PAP inhibits poly(A) tail synthesis and promotes degradation of U1A pre-mRNA (Gunderson et al., 1994, 1997). Moreover, its binding to the 5′ splice site (5′ss) of the terminal intron, avoids the use of premature cleavage and polyadenylation to protect the integrity of the transcriptome (Furth et al., 1994). Other data indicate that splicing and 3′ end processing factors may recruit each other and form a stabilized complex on the target pre-mRNA, resulting in reciprocal stimulation of efficiency. Thus, interactions between U2AF65 and CFIm59 (Millevoi et al., 2006), or U1A and CPSF160 (Lutz et al., 1996), enhance the polyadenylation reaction, while CPSF (Kyburz et al., 2006) and PAP (Vagner et al., 2000) stabilize U2AF65 to the terminal intron to stimulate splicing. On the other hand, exon–exon junction complexes (EJC) participate in mRNA degradation, as part of the CBC whose CBP80 component interacts directly with the NMD factor, up-frameshift 1 (UPF1), enhancing the efficiency of this process (Isken and Maquat, 2008). CFIm may bridge 3′ processing with capping through the binding of CFIm25 with CBP20 (Yang et al., 2011).

Until recently, little was known about mRNA metabolism in Entamoeba histolytica, the protozoan responsible for human amoebiasis. The availability of the E. histolytica genome sequence and the development of the so-called “omics” technologies have provided new opportunities for the study of mRNA processing and turnover in this parasite. To our knowledge, capping has not been described in E. histolytica, although preliminary searches in parasite genome database suggest the presence of genes that encode proteins with similarities to human capping enzymes. In this review, we focus on the current knowledge about the molecular basis for splicing, 3′ end formation and mRNA degradation, and describe some interactions between these events.

**WHAT IS KNOWN ABOUT SPLICING IN E. HISTOLYTICA**

**Splicing Factors**

There are nearly four thousand introns in the 8333 annotated genes of *E. histolytica* (Weedall and Hall, 2011), most of them flanked by highly conserved 5′ and 3′ splice sites (ss), GUUUGU and UAG, respectively, but their branch point sequences (BS) lack such degree of conservation (Wilihoef et al., 2001; Hon et al., 2013). Whereas, no minor U12 introns have been identified in amoeba and most likely neither in the eukaryotic ancestor (Collins and Penny, 2005; Bartschat and Samuelsson, 2010), the majority of the main spliceosome components have been predicted and identified.

Molecular evidence and cloning confirmed the presence of U2, U4, U5, and U6 snRNAs (Miranda et al., 1996; Davis et al., 2007), however no significant homology with eukaryotic U1 snRNAs has led to the conclusion that such small nuclear RNA is absent in *Entamoeba* (Dávila et al., 2008). Nonetheless bioinformatic analyses predicted the presence of the three U1 snRNP U1-A, U1-C, and U1-70k factors, suggesting that activation of the 5′ss might be due to direct interaction of snRNP proteins or by U6 snRNA-5′ss complementarity substitution as demonstrated in other systems (Kandels-Lewis and Seraphin, 1993; Förch et al., 2002; Rhode et al., 2006; Huang et al., 2012). *In vivo* expression of tag-cloned U1-A and cross-linking immunoprecipitation (CLIP) assays of nuclear proteins coupled to mass spectroscopy allowed the identification of at least 32 splicing factors in trophozoites (Table 1), namely U2, U4, and U5 snRNPs, integral SmD1, SmD3, and SmF proteins; the U1 snRNP components and auxiliary factors U1-70k and TIA-1/TIAR; the U2 snRNP and related components U2-A, SF3a120, SF3a60/Prp9, SF3b1, SF3b3, U2AF65, and U2AF35; the U5 snRNP components Prp8 and Prp6 [which was previously identified and cloned (Hernandez-Rivas et al., 2000)]; the U6 snRNP integral components LSm2 and LSm5; two alleles
### TABLE 1 | Comparison of splicing factors in Entamoeba histolytica vs. human and yeast.

| Particle/class | Splicing factor | Remodeling | E. histolytica protein | Locus | UniProtKB |
|----------------|----------------|------------|------------------------|-------|-----------|
| SmNLSm snRNP   | SmD1           | snRNP Sm D1 | EHI_052090              | B1N466|           |
|                | SmD3           | snRNP F    | EHI_163710              | C4MDV1|           |
|                | SmF            | snRNP F    | EHI_060400              | C4M6J5|           |
|                | LSM2           | U6 snRNA-associated Sm-like LSM2 | EHI_088580 | C4LU49 |           |
|                | LSM5           | hypothetical protein | EHI_076840 | B1N3S3 |           |
| U1 snRNP U1-related No U1 snRNA | U1A, HA-tag | U1 snRNP-specific protein | EHI_050780 | C4LTU8 |           |
|                | U1-70K         | U1 snRNP subunit | EHI_153670 | C4LS9 |           |
|                | p68            | A-B (U1-5'ss) | EhDEAD20 | EHI_096390 | C4LWF2 |
|                | TIA-1/TIAR     | RNA-binding protein TIA-1 | EHI_056660 | C4M9T1 |           |
| U2 snRNP U2-related U2 snRNA | U2A'          | leucine rich repeat protein | EHI_167290 | C4MA8 |           |
|                | SF3a120        | splicing factor | EHI_058680 | C4LWT7 |           |
|                | SF3a50/Prp9PLC | splicing factor 3A subunit 3 | EHI_038600 | C4LZP4 |           |
|                | SF3b1PLC       | splicing factor 3B subunit 1 | EHI_049170 | C4MDA8 |           |
|                | SF3b3PLC       | splicing factor 3B subunit 3 | EHI_048160 | C4MA7 |           |
|                | U2AF65         | U2 snRNP aux. fact. large subunit | EHI_098300 | C4LX43 |           |
|                | U2AF35         | U2 snRNP auxiliary factor | EHI_192500 | C4M1H0 |           |
|                | Prp43          | ILS-disassembly | EhDEXh9 | EHI_184530 | C4MS9 |           |
|                | Prp43          | ILS-disassembly | EhDEXh13 | EHI_090040 | C4M27 |           |
|                | Prp43          | ILS-disassembly | EhDEXh7 | EHI_096230 | C4LWD6 |           |
|                | Prp5           | E-A (U2-3'ss) | EhDEAD3 | EHI_013960 | C4LX8 |           |
| U5 snRNP U5 snRNA | Snu114PLC     | U5 snRNP subunit | EHI_021380 | B1N3S3 |           |
|                | BrmPLC         | EhDEXh10/U5 snRNP-specific 200kd | EHI_045170 | C4LT0 |           |
|                | BrmPLC         | EhDEXh1 | EHI_131080 | C4LXH6 |           |
|                | Brr2PILC       | EhDEXh1 | EHI_060350 | C4MK6 |           |
|                | 220K/Prp8PILC  | pre-mRNA splicing factor | EHI_093960 | C4LY7 |           |
|                | 102K/Prp6PILC  | pre-mRNA splicing factor cwc2 | EHI_021440 | C4M056 |           |
|                | Prp28          | E-A (U2-3'ss) | EhDEAD4 | EHI_020340 | C4M7U6 |           |
| U4/U6 snRNP    | CPR6           | peptidyl-prolyl cis-trans isomerase | EHI_125840 | C4M7U6 |           |
| U4/U6 snRNA    | CPR6           | peptidyl-prolyl cis-trans isomerase | EHI_125840 | C4M7U6 |           |
| U4/U6-U5 tri-snRNP | 65K/SAD1      | ubiquitin C-term hydrolase | EHI_152110 | C4LS9 |           |
|                | Prp38          | PRP38 family protein | EHI_000490 | C4LZ60 |           |
| Prp19C/IBC     | Prp19PLC       | WD domain containing protein | EHI_130870 | C4LX5F |           |
|                | CDCs5PLC       | myb-like DNA-binding | EHI_000550 | C4LZ6A |           |
|                | Aquirius       | regulator of nonsense transcripts | EHI_193520 | C4MV4 |           |
|                | SyhPILC        | Hypothetical protein | EHI_073300 | C4LXJ5 |           |
|                | RBM22PILC      | pre-mRNA splicing factor cwc2 | EHI_126150 | C4LW00 |           |
| EJC            | CWGC22SPILC    | cell cycle control protein | EHI_093720 | C4LYG3 |           |
|                | Sub2p/UAP56SE | E-A (U2-3'ss) | EhDEAD18 | EHI_151600 | C4LSK1 |           |
| RES complex    | MGC13125PILC   | EF-hand calcium-binding | EHI_150550 | C4M234 |           |
| Complex B\(^\text{act}\) | Prp2         | B\(^\text{act}\)-B\(^*\) | EhDEXh4 | EHI_033720 | C4M435 |           |
| Complex C      | Abstrakt       | EhDEAD1 | EHI_175030 | C4LM8 |           |
| Step 2 factors | Prp22PLC       | Post splicing-ILS | EhDEXh3 | EHI_077640 | C4MBN6 |           |
|                | Prp16          | C-Post splicing | EhDEXh5 | EHI_122790 | C4MSM5 |           |

*AmoebaDB. E. histolytica splicing factors, U snRNAs and components of the post-catalytic/intron lariat spliceosome complexes (PILS in superscript) were described by Miranda et al. (1996), Hernandez-Rivas et al. (2000), Davis et al. (2007), Dávila et al. (2008), Fourmann et al. (2013), and Valdés et al. (2014). Previously undetected additional Entamoeba PILS components (shaded) were identified in the ProteomeXchange repository PXD001080. # indicates that CWC22 is also part of the Prp19C. Transitions of spliceosome complexes remodeling by the respective DExH/D-box helicases are indicated (Liu, 2002; Marchat et al., 2008; Hahn et al., 2012; Wahl and Luhrmann, 2015).
of the U4/U6 di-snRNP component CP6; the U4/U6.U5 tri-snRNP components SAD1 and Prp38; and the nineteen complex (NTC) components Prp19, KIAA0560/Aquarius intron-binding spliceosomal factor, DDX5, and Abstrakt/DDX41 (Valdés et al., 2014) (Figure 1).

Splicing E (early) complex formation involves 5’ss recognition by the U1 snRNP (Larson and Hoskins, 2017). However, the less conserved and poorly recognized (weak) 5′ss are activated by 5ss-U1-C interactions or when TIA-1/TIAR binds to U-tracts localized in front of the 5′ss (Förch et al., 2002). Only U1-70k and TIA-1/TIAR were detected in the U1-A CLIP assays, therefore the most likely scenario for Entamoeba 5′ss activation involves direct interaction of U1-A/U1-70k with the 5′ss with the participation of TIA-1/TIAR bound to the U-rich most often spliced *Entamoeba* 5′ss (GUUUGUUU) (Hon et al., 2013) as described for weak 5′ss.

Because cross-linking was carried out with UV, the number of factors identified is limited but it represents all complexes formed during spliceosome assembly, first and second steps of splicing, disassembly, turnover, exon junction complex, and mRNA transport. Moreover, the presence of the core protein of the NTC, Prp19, and U2AF65, which interact with the PSer2 CTD of the large subunit of RNA pol II, ensure proper co-transcriptional activation of the spliceosome, splicing catalysis, termination factors recruitment, and extranuclear mRNA transport factors (David et al., 2011; Gu et al., 2013).

![FIGURE 1 | Co-transcriptional pre-mRNA processing in *E. histolytica*: focus on splicing factors. The model summarizes the data available to date (Valdés et al., 2014). During transcript elongation by RNA polymerase II (RNAPII; purple), Ser2 residues of the few heptapeptide repeats of its carboxy-terminal domain become phosphorylated (PSer2-CTD; maroon circles) and apt to recruit the spliceosomal (salmon triangle) and polyadenylation machineries. The large subunit of the U2 Auxiliary Splicing Factor U2AF65 (of 84 kDa in *E. histolytica*; orange oval) is a major player in pre-mRNA processing by tethering the spliceosome and the pre-mRNA (light green boxes) to RNAPII. U2AF65 interacts with the RNAPII-PSer2-CTD and with splicing factors conforming the Prp19 Complex (NTC; blue circle). The NTC regulates the formation and progression of essential spliceosome conformations required for the two steps of spacing. Splicing complex E formation occurs when the snRNP U1-A (yellow oval) binds to the 5′ss (splice site) and the splicing factor TIA-1/TIAR (yellow box) binds to the U-rich sequence just downstream the 5′ss. Splicing complex E also involves the 3′ss definition (not shown). When RNAPII releases the 3′ss from the transcription site, splicing factor 1 binds the branch site at the same time that U2AF65 binds the intron’s polypyrimidine tract located between the branch site and the 3′ss; also simultaneously, the small subunit of U2AF (U2AF35, of 29 kDa in *E. histolytica*; pale orange circle) recognizes the 3′ss. The interaction of U2AF65 with splicing factor 1 and U2AF65 at the 3′ss and with the CTD of RNAPII ensures that U2AF65 also tethers the pre-mRNA to RNAPII. Finally, in addition to the previously reported interactions of RNAPII with the polyadenylation complex (vide infra), U1-A directly or indirectly interacts with the splicing complexes B-C, and more importantly with CstF77 (dark green oval), a member of the polyadenylation machinery.](image-url)
Intron Lariat Debranching Enzyme

Intron lariat debranching enzyme, or Dbr1, is a member of the calcineurin-like metallophosphoesterases (MPEs) superfamily of binuclear metal-ion-containing enzymes that hydrolyze phosphomonoesters, phosphodiesterases, and phosphotriesters in a metal-dependent manner. From bacteriophages to humans, the MPE domain is found in Mre11/SbcD DNA-repair enzymes, mammalian phosphoprotein phosphatases, acid sphingomyelinases, purple acid phosphatases, nucleotidases, and bacterial cyclic nucleotide phosphodiesterases. Despite this functional diversity, MPEs show a remarkably similar structural fold and active-site architecture composed of five sequence blocks that allow metal coordination in the conserved motif D[X]H[x]nGD[x]nGHNH[D/E] [x]nH[x]nGH[X]H (Matange et al., 2015). Alanine scanning assays identified the yeast Dbr1 active site (Valdés et al., 2014). Interactions between EhCFIm25 and EhPAP have been experimentally confirmed, although our previous findings indicated the presence of the pocket could accommodate pyrimidines. In addition, there are few sequence-specific interactions at the BS, confirming recognition of atypical BS. Finally, the interactions between RNA and the LRL are stabilized by secondary contacts between residues 141–144 of the LRL and residues Phe292, Pro293, and Phe337 of the carboxy-terminal domain (CTD) of Dbr1; and intricate hydrogen bonds centered in Arg158 aid to stabilize further the conformation of the LRL (Montemayor et al., 2014). The structural data was confirmed by Dbr1 activity in vivo. Whereas, E. histolytica wild type Dbr1 was able to complement Saccharomyces cerevisiae Dbr1-deletant strains relieving intron lariat accumulation, none of the constructs carrying Cys14Ala/Ser substitutions or 141-146Ala substitutions, or CTD or LRL deletions relieved intron lariat accumulation (Montemayor et al., 2014). The presence of E. histolytica Dbr1 in intron large post-spliceosomal complexes along with U2, U5 and U6 snRNPs, and the proteins Ntr1/TFIP11 and Prp43 (Yoshimoto et al., 2009) or the Drn1/Ygr093w protein that transiently binds Dbr1 to post-spliceosomal complexes is still unproven (Garrey et al., 2014), although our previous findings point to this possibility (Valdés et al., 2014).

UNDERSTANDING POLYADENYLATION IN E. HISTOLYTICA

Cis-Elements for Pre-mRNA 3′ End Formation

One of the first reports about mRNA polyadenylation in E. histolytica was published in 1993 and describes the existence of a putative polyadenylation motif TAATT and a 12 pyrimidine stretch in the 3′ UTR of parasite genes (Bruchhaus et al., 1993). Then, other groups showed that alternative polyadenylation sites and poly(A) tail size represent efficient posttranscriptional and post-spliceosomal mechanisms for gene expression regulation (Urban et al., 1996; López-Camarillo et al., 2003). But the publication of the first version of the E. histolytica genome sequence in 2005 (assembly of

---

**FIGURE 2** | Model of the pre-mRNA 3′ processing complex in E. histolytica. The core 3′ processing complex is composed of four main complexes: CPSF (green), CstF (yellow), CFiIm (orange), and CFiIm (purple) that bind cis-elements within pre-mRNA 3′ UTR. Based on the current knowledge in other eukaryotic cells, we hypothesize that the WDR33 subunit of CPSF complex recognizes the polyadenylation signal (polyA signal) located upstream of the cleavage site (polyA site marked here with a black arrow), while CstF64 recognizes the U-rich downstream element. We also propose that the binding of CFiIm25 to the upstream U-rich motif promotes the recruitment of CFiIm subunits (CIP1 and PCF11) and interactions between CPSF and CstF, allowing the RNA cleavage by CPSF37 and the poly(A) tail synthesis by PAP. Additional factors, such as FIP1, PC4, RBBP6, SSU72, and PAPB2 would also contribute to the regulation of cleavage/polyadenylation reaction (López-Camarillo et al., 2005). Interactions between EhCFIm25 and EhPAP have been experimentally confirmed in vitro (Pezet-Valdez et al., 2013). Interestingly, the interaction of EhCstF77 with U1-A provided the first evidence for a link between mRNA polyadenylation and splicing in Entamoeba (Valdés et al., 2014).
~23 Mb that predicted 9938 coding genes comprising 49% region of the genome) (Loftus et al., 2005) represented the critical step to identify motifs in the mRNA 3′UTR and the polyadenylation machinery in this parasite.

A small-scale in silico analysis of cDNA and genomic sequences revealed that E. histolytica 3′ UTRs contain three conserved motifs: (i) the consensus UA(A/U)UU polyadenylation signal or variants located 10–30 nt upstream the poly(A) site, (ii) the U-rich tract located 1–30 nt upstream the poly(A) site, and (iii) a U-rich element located 3–30 nt downstream the poly(A) site (López-Camarillo et al., 2005). Computational examination of a larger number of cDNA and genomic sequences confirmed this molecular array and suggested the presence of an additional distal A-rich element (Figure 2) (Zamorano et al., 2008). Study of the alternative usage of poly(A) sites using RNA-Seq indicated that microheterogeneity in poly(A) sites is likely to be stochastic in E. histolytica and only a small fraction of alternative polyadenylation isoforms appeared to be genuine (Hon et al., 2013). Interestingly, genes with alternative poly(A) sites may have a large impact on global gene expression in E. histolytica since most of them participate in DNA condensation, DNA binding, translation, splicing, mRNA binding, protein folding and protein transport; other genes are related to signaling, oxidation/reduction, calcium ion binding, cell cycle, and intracellular transport. Indeed, the upstream shift in poly(A) site selection resulting from the silencing of the polyadenylation factor EhCFIm25, was confirmed for thioredoxin and 60S ribosomal protein L7 transcripts and related to specific phenotypical changes and parasite death (Ospina-Villa et al., 2017).

**Polyadenylation Factors**

Analyses of the 9938 coding genes predicted in the first version of the genome indicated that E. histolytica has genes that encode proteins with homology to the majority of polyadenylation factors described in human and yeast (Table 2): the cleavage and specificity factor (CPSF160, 100, 73, and 30), the cleavage stimulating factor (CstF77, 64, and 50), the 25 kDa subunit of the cleavage factor Im (CfIm) and both CIP1 and PCF11 subunits of CFIm, as well as FIP1, poly(A) polymerase (PAP), poly(A) binding protein (PABP), RBBP6 (Mpe1 in yeast), WDR33 (Ps2 in yeast), PNAS-120 (Su72 in yeast), and PC4 (Sub1 in yeast) (Figure 2) (López-Camarillo et al., 2005, 2014). In human cells, WDR33 and CPSF30 are the CPSF subunits that binds the polyadenylation signal (Chan et al., 2014; Schonemann et al., 2014). CPSF73 is the endonuclease responsible for RNA cleavage (Mandel et al., 2006). CstF77 interacts with CPSF160, promoting their cooperative RNA binding during the assembly process (Murthy and Manley, 1995). CFIm25 regulates the selection of distal poly(A) sites, contributes to the recruitment of polyadenylation factors and is necessary for poly(A) tail synthesis (Brown and Gilmartin, 2003; Kubo et al., 2006). Although CFIm subunits are the less characterized polyadenylation factors, evidence suggest that Pcf11 is required for degradation of the 3′ product following cleavage (West and Proudfoot, 2008), while Clp1 interacts with both CPSF and CFIm and likely tethers them to CFIm (de Vries et al., 2000). PAP is responsible for the addition of the polyadenosine tail (Raabe et al., 1991) and its activity is accelerated by PABP (Wahle, 1991). RBBP6 associates with other core polyadenylation factors through its unusual ubiquitin-like domain and modulates expression of mRNAs with AU-rich 3′ UTR (Di Giammartino et al., 2014). PC4 (Sub1 in yeast) associates with the polyadenylation factor CstF64 to modulate transcription termination and polyadenylation initiation (Calvo and Manley, 2001, 2003).

The presence of these proteins in E. histolytica suggests that 3′ end processing of parasite mRNA could be performed as it has been described in other eukaryotic cells. Accordingly, EhPAP has the conserved PAP central catalytic domain with the three invariant aspartate residues involved in nucleotide transfer and the F/YGS motif responsible for ATP binding, confirming that it belongs to the polymerases-like superfamily of nucleotidyl transferases (García-Vivas et al., 2005). In the predicted three-dimensional model, both functional domains fold as a U-shaped structure that likely orients the incoming ATP and RNA molecules for poly(A) tail extension. In agreement with its expected role in the poly(A) tail synthesis, EhPAP was.

### TABLE 2 | Polyadenylation factors in human and E. histolytica.

| Protein | Access number<sup>a</sup> | Locus<sup>b</sup> | Protein | Access number<sup>a</sup> |
|---------|---------------------------|------------------|---------|---------------------------|
| **CLEAVAGE AND POLYADENYLATION SPECIFIC FACTORS** |
| CPSF160 | Q10570 | EHI_160110 | CPSF160 | C4M386 |
| CPSF100 | Q9P210 | EHI_033130 | CPSF100 | C4M6Y0 |
| CPSF73 | Q9UKF6 | EHI_136700 | CPSF73 | C4M297 |
| CPSF30 | Q95639 | EHI_067580 | CPSF30 | C4M9G4 |
| FIP1 | Q6UN15 | EHI_052180 | FIP1 | C4M765 |
| WDR33 | Q9C0J8 | EHI_170080 | WDR33 | C4M1D0 |
| **CLEAVAGE STIMULATING FACTORS** |
| CatF77 | Q12996 | EHI_098370 | CatF-77 | C4L2W3 |
| CatF64 | P33240 | EHI_151900 | CatF-64 | C4LSN8 |
| CatF50 | Q05048 | EHI_152770 | CatF-50 | C4LSW0 |
| **CLEAVAGE FACTORS IM** |
| CFIm68 | Q16630 | – | – | – |
| CFIm59 | Q8NE64 | – | – | – |
| CFIm25 | O43809 | EHI_077110 | CFIm25 | C4M2T1 |
| **CLEAVAGE FACTORS IIM** |
| CIP1 | Q92989 | EHI_008100 | CIP1 | C4LYE5 |
| PCF11 | Q94913 | EHI_045130 | PCF11 | C4LTO6 |
| **OTHER POLYADENYLATION FACTORS** |
| PAP-B | Q9NFJ5 | EHI_012040 | PAP | Q51D88 |
| Ssu72 | Q98Z56 | EHI_027340 | Ssu72 | C4M1T3 |
| PC4 | Q59999 | EHI_192520 | PC4 | C4M1H2 |
| PABP1 | P11940 | EHI_198750 | PABP1 | C4LSW1 |
| RBBP6 | Q7Z6E9 | EHI_014000 | RBBP6 | C4LPX2 |
| Symplekin | Q92797 | – | – | – |

<sup>a</sup>UhrProtKB; <sup>b</sup>AmoebaDB. Data about parasite proteins were obtained from (López-Camarillo et al., 2005, 2014).
EhCFIm25 is the homolog of the human positive coactivator 4, a multifunctional protein that establishes an important link between transcription and polyadenylation. On one hand, its binding to promoters facilitates the recruitment of transcription factors to stimulate the pre-initiation complex assembly (Conesa and Acker, 2010); on the other hand, its interaction with EhCstf64 avoids premature transcription termination and polyadenylation initiation until the polyadenylation motifs have been transcribed (Calvo and Manley, 2001). Moreover, it mediates chromatin organization and heterochromatin gene silencing by interacting with histones H3 and H2B (Das et al., 2006, 2010). As homologous proteins, the EhPC4 protein contains a single-strand DNA (ssDNA) binding region whose residue K127 is required for DNA interaction, and a dimerization domain in the so-called PC4 domain at the C-terminus (Hernandez de la Cruz et al., 2014). Interestingly EhPC4 and its potential partner, EhCstf-64, were significantly up-regulated in virulent trophozoites (Santi-Rocca et al., 2008). Consistently,
the overexpression of EhPC4 induced the modulation of proteins with key functions in cytoskeleton dynamics, cell migration and invasion in trophozoites. Among them, the up-regulation of a 16-kDa actin-binding protein (EhABP16) which is a putative member of the coflin/tropomyosin family involved in actin polymerization was associated with an increase in parasite migration of trophozoites and destruction of human SW480 colon cells, confirming that EhPC4 has an impact on parasite virulence (Hernández de la Cruz et al., 2014). On the other hand, the overexpression of EhPC4 significantly increased cell proliferation, DNA replication and DNA content of trophozoites, promoting the formation of giant multinucleated trophozoites. EhPC4 modulates the expression of genes involved in carbohydrate and nucleic acid metabolism, chromosome segregation and cytokinesis, evidencing the relevance of this factor in polyploidy and genome stability in E. histolytica (Hernández de la Cruz et al., 2016). The role of EhPC4 in mRNA 3′ end formation and its relevance for the events mentioned above remain to be investigated.

**MOLECULAR EVENTS FOR mRNA DECAY IN E. HISTOLYTICA**

**mRNA Degradation Machineries**

E. histolytica has most of the factors that are involved in mRNA degradation in eukaryotic cells, including proteins involved in deadenylation, decapping, and exonuclease activity, but it lacks several components (Table 3) (López-Rosas et al., 2012). The reduced mRNA deadenylation machinery includes the CAF1/NOT complex with the five NOT proteins and the poly-A specific ribonucleases CAF1 and CAF1-like, but the carbon catabolite repressor 4 (CCR4) described in yeast and human, as well as PAN2, PAN3, and PARN deadenylases, are missing (Figure 3). EhCAF1 is a ribonuclease D family member, having the CAF1 nuclease domain and the conserved DEDD residues (D81E86D208D276) that are important for 3′ to 5′ exonuclease activity in homologous proteins (Dauergon et al., 2001). Consequently, EhCAF1 is a functional deadenylase that binds 3′ UTR and degrades the poly(A) tail of parasite transcripts in *in vitro* assays (López-Rosas et al., 2014).

Although capping has not been described in E. histolytica, bioinformatics analyses revealed the existence of a decapping complex that is formed by the catalytic subunit EhDCP2, and EhXRN2, EhLSM1–6, EhEDC3, and EhDHH1 as decapping associated proteins, whereas it also includes DCP1, SCD6, PAT1, and LSM7 in yeast and animals (Table 3; Figure 3) (López-Rosas et al., 2012). EhXRN2 and EhDCP2 have the typical architecture of homologous proteins. Notably, EhXRN2 has the XRN_N nuclease domain and the internal tower domain with the active site motif KX5QQX5RR, which is critical for ribonuclease function (Xiang et al., 2009). EhDCP2 has the conserved DCP2 box A domain and the conserved nudix Box (GX5EX5REUXEXXGU) that are both responsible for cap structure removal (She et al., 2008). In eukaryotic cells, the elimination of the 5′ cap compromises mRNA to 5′ to 3′ exonucleolytic decay, apparently in an irreversible way, hence, decapping activity is tightly regulated (Li and Kiledjian, 2010). The heptameric Lsm1–7 complex associates with the 3′ end of deadenylated mRNAs and promotes decapping (Tharun et al., 2000; Tharun and Parker, 2001). The activity of the decapping enzyme is stimulated by accessory proteins, such as Edc proteins.

**TABLE 3** | Comparison of mRNA decay machineries between human and E. histolytica.

| Protein | Access number | Locus | Protein | Access number |
|---------|---------------|-------|---------|--------------|
| **DECAPPING FACTORS** | | | | |
| DCP2 | Q8U60 | EHI_058810 | EhDCP2 | C4M5G6 | López-Rosas et al., 2012 |
| Lsm1 | O15116 | EHI_188020 | EhLsm1 | B1N3A8 | | |
| Lsm2 | Q9Y333 | EHI_068580 | EhLsm2 | C4LU49 | | |
| Lsm3 | P62310 | EHI_151310 | EhLsm3 | C4L5H4 | | |
| Lsm4 | Q9Y420 | EHI_049370 | EhLsm4 | C4LU9D | | |
| Lsm5 | Q9Y49Y | EHI_078480 | EhLsm5 | B1N33 | | |
| Lsm6 | P62312 | EHI_188130 | EhLsm6 | C4M187 | | |
| Lsm7 | Q9UK45 | EHI_025840 | EhLsm7 | C4M99 | | |
| Edc3 | Q9E6F6 | EHI_198940 | EhEcd3 | C4LWU0 | | |
| Drhl | P26196 | EHI_093900 | EhDrdh1 | C4L911 | | |
| **DEADENYLATION FACTORS** | | | | |
| CAIF | Q9UIJ1 | EHI_048150 | EhCAIF | Q56AY2 | López-Rosas et al., 2012, 2014 |
| CALIF | Q9UIJ1 | EHI_039000 | EhCAIF-1 | C4LJ21 | | |
| NOT1 | Q8Y8AX | EHI_008810 | EhNOT1 | C4MY89 | | |
| NOT2 | Q9NZN8 | EHI_041180 | EhNOT2 | C4MV90 | | |
| NOT3 | Q7157 | EHI_119650 | EhNOT3 | C4MY73 | | |
| NOT4 | Q9X652 | EHI_080710 | EhNOT4 | C4MN99 | | |
| **EXOSOME FACTORS** | | | | |
| RRP4 | Q13868 | EHI_163510 | EhRRP4 | C4M372 | López-Camarillo et al., 2014 |
| RRP6 | Q01780 | EHI_021400 | EhRRP6 | C4M504 | | |
| RRP40 | Q8Y375 | EHI_004770 | EhRRP40 | C4M673 | | |
| RRP41 | Q9NP33 | EHI_040230 | EhRRP41 | C4M4G1 | | |
| RRP42 | Q75024 | EHI_005680 | EhRRP42 | C4LZ49 | | |
| RRP43 | Q9E626 | EHI_180880 | EhRRP43 | C4M982 | | |
| RRP46 | Q9NOT4 | EHI_088520 | EhRRP46 | C4M89Y | | |
| MTR3 | Q5PKV6 | EHI_126330 | EhMTR3 | C4LW17 | | |
| DIS3 | Q8Y2L | EHI_160720 | EhDIS3 | C4MAJ9 | | |
| **NON-SENSE MEDIATED DECAY FACTOR** | | | | |
| UPF1 | Q92900 | EHI_035550 | EhUPF1 | C4LY5X | López-Rosas et al., 2012 |
| **RNA INTERFERENCE FACTORS** | | | | |
| AGO2 | Q9JKV8 | EHI_188850 | EhAGO2-1 | C4LVQ2 | Zhang et al., 2008 |
| – | – | EHI_125680 | EhAGO2-2 | C4LVV2 | | |
| – | – | EHI_177170 | EhAGO2-3 | C4LY31 | | |
| – | – | EHI_139420 | EhRlfP | C4M6W7 | | |
| – | – | EHI_179800 | EhRlfP | C4MB51 | | |
| DICER | Q9UPY3 | EHI_068740 | EhRNaseIII | C4M6U4 | Abed and Ankit, 2005; Zhang et al., 2008 |

*a*UhrProtKB; *b*AmoebaDB.
the DExD/H-box RNA helicases Dhh1 and Pat1 (Bonnerot et al., 2000; Schwartz et al., 2003).

*Entamoeba histolytica* contains seven exosome encoding genes including Rrp41, Rrp43, Rrp46, Mtr3-Rrp42 and the catalytic subunit Dis3, as well as accessory stabilizing Rrp4, and Rrp40 proteins; but it lacks Rrp45 and Cal4 genes (Table 3; Figure 3) (López-Rosas et al., 2012). The EhRRP41 protein colocalizes and physically interacts with EhL-PSP, which also interacts and colocalizes with the EhCAF1 deadenylase. But the fact that EhRRP41 did not coimmunoprecipitate with EhCAF1, suggests the existence of two EhL-PSP-containing complexes. The colocalization of exosome factors (EhRrp41) with EhCAF1 and EhL-PSP in trophozoites showed novel interactions between mRNA degradation protein and suggests the existence of cooperative interactions between mRNA decay machineries in *E. histolytica* (López-Rosas et al., 2014). In yeast and human, the nine subunits of the exosome complex form a ring structure in which Dis3 (RRP44) is the key player in mRNA turnover being the catalytic subunit responsible for exonucleolytic and endonucleolytic activities in the 3′-5′ decay of deadenylated transcripts in cytoplasm (Ibrahim et al., 2008). Additionally, the nuclear exosome is involved in 3′-end trimming of rRNA, snRNA, and snoRNA, as well as mRNA surveillance and degradation of cryptic unstable transcripts (Parker and Song, 2004).

*E. histolytica* genome also contains genes for components of the NMD and RNA interference (RNAi) pathways, namely three Ehupf genes (López-Rosas et al., 2012), as well as two EhRdRP, one EhRNaseIII and three EhAGO2 proteins (Abed and Ankri, 2005; Zhang et al., 2008, 2011), respectively (Table 3). The absence of DICER and GW182 homologs suggests that RNA interference may use DICER-independent mechanisms in *E. histolytica* (Zhang et al., 2011). Pompey et al. (2015) recently showed that EhRNaseIII is able to cleave dsRNA to generate shorter fragments in a heterologous system. This suggests that EhRNaseIII in conjunction with other amoebic factors might reconstitute an active DICER-like complex. Congruently, numerous reports involving gene-silencing assays confirmed the functionality of the RNAi pathway in *E. histolytica*.

**P-Body-Like Structures**

Several experiments suggest that mRNA decay reactions, namely deadenylation, decapping, and 5′-exonucleolytic decay, take place in microscopically detectable cytoplasmic P-bodies like structures in *E. histolytica* (López-Rosas et al., 2012), as it has been described in other eukaryotic cells (Sheh and Parker, 2003). The EhCAF1 deadenylase, EhXRN2 exoribonuclease and EhDCP2 decapping proteins, as well as the EhAGO2-2 protein, were detected in cytoplasmic foci in immunofluorescence and confocal microscopy experiments (López-Rosas et al., 2012). Additionally biochemical analysis revealed that EhCAF1 co-immunoprecipitated with EhXRN2, thus linking deadenylation to 5′-to-3′ mRNA degradation. Interestingly, these cytoplasmic structures also contain polyadenylated transcripts and dsRNA, which is congruent with their role in RNA decay. Moreover their formation depends on the presence of active transcription and translation (López-Rosas et al., 2012), as well as cellular stress, such as DNA damage, heat shock, and nitric oxide (López-Rosas et al., 2014), which make them bona fide P-body structures (Figure 3). Altogether, these data suggest that, as in human cells, the accumulation of transcripts in cytoplasmic P-bodies like structures for silencing or decay, represents a key regulatory process for gene expression regulation in response to specific conditions or signals in *E. histolytica*.

**CONCLUSION**

Besides the evolutionary distance between *E. histolytica* and its human host, the screening of parasite genome sequences and the functional characterization of specific factors, revealed that molecular mechanisms regulating mRNA processing and degradation seem to be roughly similar in both organisms. Several subtle differences exist, but canonical factors involved in splicing, polyadenylation and decay are generally conserved in this primitive eukaryote, which highlights that these events are key players for gene expression regulation in eukaryotic cells. The study of a larger number of factors involved in splicing, polyadenylation or mRNA degradation remains to be addressed to elucidate all the relationship among these reactions. In addition to contribute to the better understanding of posttranscriptional regulation in *E. histolytica*, the characterization of these factors and events may also lead to the identification of a biochemical target involved in various mRNA processing pathways, whose inhibition would have a massive impact on parasite survival. On the other hand, recent data indicated the potential of factors involved in polyadenylation as biochemical targets for parasite control, which may open the way for the design of new molecules for the control of this parasitic disease. In this context, the results of the proof-of-concept study in *E. histolytica* may promote the use of aptamers to control *E. histolytica* during the development of amoebiosis or to eradicate residual trophozoites during antibiotic treatment. Further experiments are required to confirm their affinity, evaluate their effect in vivo and improve their bioavailability.

**FUNDING**

This study was supported by the Mexico-France program grants ECOS NORD (M14S02) and SEP-CONACYT-ANUIES (249554), and Mexican grants from SIP-IPN (20170969) and CONACyT (178550), Mexico. JO-V was a scholarship recipient of Mexican BEIFI-IPN and CONACyT programs. LM is supported by COFAA-IPN.

**AUTHOR CONTRIBUTIONS**

JV-F, IL-R, CL-C, ER-M, and JO-V reviewed data about mRNA splicing, polyadenylation and decay, respectively. JO-V designed the figures. CL-C and LM designed the review organization, revised and integrated the different parts of the manuscript.
Kandelis-Lewis, S., and Séraphin, B. (1993). Involvement of U6 snRNA in 5’ splice site selection. Science 262, 2035–2039. doi: 10.1126/science.262.5140.1598
Kubo, T., Wada, T., Yamaguchi, Y., Shimizu, A., and Handa, H. (2006). Knock-down of 25 kDa subunit of cleavage factor Im in Hela cells alters alternative polyadenylation within 30-UTRs. Nucleic Acids Res. 34, 6264–6271. doi: 10.1093/nar/gkl273
Kyburz, A., Friedlein, A., Langen, H., and Keller, W. (2006). Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3’ end processing and splicing. Mol. Cell 23, 195–205. doi: 10.1016/j.molcel.2006.05.037
Labno, A., Tomecki, R., and Dziembowski, A. (2016). Cytoplasmic RNA decay: what defines a substrate? Curr. Opin. Cell Biol. 46, 107–114. doi: 10.1016/j.ceb.2016.09.023
Larson, J. D., and Hoskins, A. A. (2017). Dynamics and consequences of spliceosome E complex formation. Elife 6:e27592. doi: 10.7554/eLife.27592
Li, Y., and Kiledjian, M. (2010). Regulation of mRNA decapping. Wiley Interdiscip. Rev. RNA 1, 253–265. doi: 10.1002/wrna.15
Liu, Z. R. (2002). 68S RNA helicase is an essential human splicing factor that acts at the U1 snRNA-S 5’ splice site duplex. Mol. Cell. Biol. 22, 5443–5450. doi: 10.1128/MCB.22.15.5443-5450.2002
Loftus, B., Anderson, I., Davies, R., Alsmark, U. C., Samuelson, J., Amedeo, P., et al. (2006). An interaction between U2AF 65 and CFIm links pre-mRNA 3’-end processing to mRNA export. Nucleic Acids Res. 34, 591, 2003–2010. doi: 10.1002/1873-3468.12677
Millevoi, S., Loulergue, C., Dettwyler, S., Karaa, S. Z., Keller, W., Antoniou, M., et al. (2006). An interaction between U2AF65 and CFIm links the splicing and 3’ end processing machineries. EMBO J. 25, 4854–4864. doi: 10.1038/sj.emboj.7601331
Miranda, R., Salgado, L. M., Sánchez-López, R., Alagón, A., and Lizardi, P. M. (1996). Identification and analysis of the u6 small nuclear RNA gene from Entamoeba histolytica. Gene 180, 37–42. doi: 10.1016/S0378-1119(96)00397-6
Mortenmeyer, E. J., Katalik, A., Clark, N. E., Taylor, A. B., Schueriemann, J.P., Combis, D. J., et al. (2014). Structural basis of lariat RNA recognition by the intron debranching enzyme Dbr1. Nucleic Acids Res. 42, 10845–10855. doi: 10.1093/nar/gku725
Murthy, K. G., and Manley, J. L. (1995). The 160-kb subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3’-end formation. Genes Dev. 9, 2672–2683. doi: 10.1101/gad.9.21.2672
Osypina-Villa, J. D., Dufour, A., Weber, C., Ramirez-Moreno, E., Zamaro-Carrillo, E., Guillen, N., et al. (2018). Targeting the polyadenylation factor EhpCfm25 with RNA aptamers controls survival in Entamoeba histolytica. Sci. Rep. 8:5270. doi: 10.1038/s41598-018-23997-w
Osypina-Villa, J. D., Guillen, N., Lopez-Camarillo, C., Soto-Sanchez, J., Ramirez-Moreno, E., Garcia-Vazquez, R., et al. (2017). Silencing the cleavage factor CFIm25 as a new strategy to control Entamoeba histolytica parasite. J. Microbiol. 55, 783–791. doi: 10.1007/s12725-017-7299-9
Osypina-Villa, J. D., Zamaro-Carrillo, A., Lopez-Camarillo, C., Castro-Sanchez, C. A., Soto-Sanchez, J., Ramirez-Moreno, E., et al. (2015). Amino acid residues Leu135 and Tyr236 are required for RNA binding activity of CFIm25 in Entamoeba histolytica. Biochimie 115, 44–51. doi: 10.1016/j.biochi.2015.04.017
Palencia, A., Boudgour, A., Brenier-Pichart, M. P., Touquet, B., Bertini, R. L., Sensi, C., et al. (2017). Targeting Toxoplasma gondii CPSF3 as a new approach to control toxoplasmosis. EMBO Mol. Med. 9, 385–394. doi: 10.15252/emmm.201607370
Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. Nat. Struct. Mol. Biol. 11, 121–127. doi: 10.1038/nsmb724
Pezet-Valdez, M., Fernandez-Retana, J., Osypina-Villa, J. D., Ramirez-Moreno, M. E., Orozco, E., Charcas-Lopez, S., et al. (2013). The 25 kDa subunit of cleavage factor Im Is a RNA-binding protein that interacts with the poly(A) polymerase in Entamoeba histolytica. PLoS ONE 8:e69777. doi: 10.1371/journal.pone.0069777
Pompey, J. M., Foda, B., and Singh, U. (2015). A Single RNaseII Domain Protein from Entamoeba histolytica Has dsRNA cleavage activity and can help mediate RNAi gene silencing in a heterologous system. PLoS ONE 10:e0133740. doi: 10.1371/journal.pone.0133740
Raae, T., Bollum, F. J., and Manley, J. L. (1991). Primary structure and expression of bovine poly(A) polymerase. Nature 353, 229–234.
Ramanathan, A., Robb, G. B., and Chan, S. H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res. 44, 7511–7526. doi: 10.1093/nar/gkw551
Ransey, E., Paredes, E., Dey, S. K., Das, S. R., Heroux, A., and Macbeth, M. R. (2017). Crystal structure of the Entamoeba histolytica RNA lariat debranching enzyme EhDbr1 reveals a catalytic Zn(2+) /Mn(2+) heterobinucleation. FEBS Lett. 591, 2003–2010. doi: 10.1002/1873-3468.12677
Rebbapragada, I., and Lykke-Andersen, J. (2009). Execution of nonsense-mediated mRNA decay: what defines a substrate? Curr. Opin. Cell Biol. 21, 394–402. doi: 10.1016/j.ceb.2009.02.007
Rhode, B. M., Hartmuth, K., Westhof, E., and Lührmann, R. (2006). Proximity of conserved U6 and U2 snRNA elements to the 5’ splice site region in activated spliceosomes. EMBO J. 25, 2475–2486. doi: 10.1038/sj.emboj.7601134
Santi-Rocca, J., Weber, C., Guigon, G., Sismeiro, O., Coppée, J. Y., and Guillén, N. (2008). The lysine- and glutamic acid-rich protein KerPI plays a role in Entamoeba histolytica liver abscess pathogenesis. Cell. Microbiol. 10,202–217. doi: 10.1111/j.1462-5822.2007.01030.x
Schönemann, L., Kuhn, U., Martin, G., Schafer, P., Gruber, A. R., et al. (2014). Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. Genes Dev. 28, 2381–2393. doi: 10.1101/gad.250985.114
Schwartz, D., Decker, C. J., and Parker, R. (2003). The enhancer of decapping, Edc1p and Edc2p, bind RNA and stimulate the activity of the decapping enzyme. RNA 9, 239–251. doi: 10.1261/rna.2171203
Schwer, B., Khalid, F., and Shuman, S. (2016). Mechanistic insights into the manganese-dependent phosphodiesterase activity of yeast Dbr1 with bis-p-nitrophenylphosphate and branched RNA substrates. RNA 22, 1819–1827. doi: 10.1261/rna.058555.2.116

She, M., Decker, C. J., Svergun, D. I., Round, A., Chen, N., Muhlrad D., et al. (2008). Structural basis of dcp2 recognition and activation by dcp1. Mol. Cell. 29, 337–349. doi: 10.1016/j.molcel.2008.01.002

Sheth, U., and Parker, R. (2003). Decapping and decay of RNAs occur in cytoplasmic processing bodies. Science 300, 805–808. doi: 10.1126/science.1082320

Shi, Y. (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. Nat. Rev. Mol. Cell Biol. 18, 655–670. doi: 10.1038/nrm.2017.66

Sidik, S. M., Huet, D., Ganesan, S. M., Huynh, M. H., Wang, T., Nasamu, A. S., et al. (2016). A Genome-wide CRISPR Screen in Toxoplasma identifies essential apicomplexan genes. Cell 166, 1423–1435. doi: 10.1016/j.cell.2016.08.019

Sonoki, E., Ng, C. L., Lee, M. C., Guo, D., Zhang, Y. K., Zhou, Y., et al. (2017). A potent antimalarial benzoxaborole targets a Plasmodium falciparum cleavage and polyadenylation specificity factor homologue. Nat. Commun. 8:14574. doi: 10.1038/ncomms14574

Tharun, S., He, W., Mayes, A. E., Lennertz, P., Beggs, J. D., and Parker, R. (2000). Yeast Sm-like proteins function in mRNA decapping and decay. Nature 404, 515–518. doi: 10.1038/35066676

Tharun, S., and Parker, R. (2001). Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p–7p complex on deadenylated yeast mRNAs. Mol. Cell. 8, 1075–1083. doi: 10.1016/S1097-2765(01)00395-1

Urban, B., Blasic, G., Förster, B., Hamelmann, C., and Horstmann, R.D. (1996). Putative serine/threonine protein kinase expressed in complement-resistant forms of Entamoeba histolytica. Mol. Biochem. Parasitol. 80, 171–178. doi: 10.1016/0166-8851(96)00264-9

Vagner, S., Vagner, C., and Mattaj, I. W. (2000). The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF65 to couple 3’-end processing and splicing. Genes Dev. 14, 403–413.

Valdés, J., Nozaki, T., Sato, E., Chiba, Y., Nakada-Tsukui, K., Villegas-Sepúlveda, N., et al. (2014). Proteomic analysis of Entamoeba histolytica in vivo assembled pre-mRNA splicing complexes. J. Proteomics 111, 30–45. doi: 10.1016/j.jprot.2014.07.027

Wahl, M. C., and Luhrmann, R. (2015). SnapShot: spliceosome dynamics. I. Cell 162:690.e1. doi: 10.1016/j.cell.2015.07.033

Wahlé, E. (1991). A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. Cell 66, 759–768. doi: 10.1016/0092-8674(91)90119-1

Weedall, G. D., and Hall, N. (2011). Evolutionary genomics of Entamoeba. Res. Microbiol. 162, 637–645. doi: 10.1016/j.resmic.2011.01.007

West, S., and Proudfoot, N. J. (2008). Human Pcf11 enhances degradation of RNA polymerase II-associated nascent RNA and transcriptional termination. Nucleic Acids Res. 36, 905–914. doi: 10.1093/nar/gkn1112

Willoeht, U., Campos-Góngora, E., Touzni, S., Bruchhauser, L., and Tannich, E. (2001). Introns of Entamoeba histolytica and Entamoeba dispar. Protist 152, 149–156. doi: 10.1078/1434-4610-00053

Xiang, K., Tong, L., and Manley, J. L. (2014). Delineating the structural blueprint of the pre-mRNA 3’-end processing machinery. Mol. Cell. Biol. J. 34, 1894–1910. doi: 10.1128/MCB.00884-14

Xiang, S., Cooper-Morgan, A., Jiao, X., Kiledjian, M., Manley, J. L., Tong, L. J. et al. (2009). Structure and function of the 5’ → 3’ exoribonuclease Rat1 and its activating partner Rai1. Nature 458, 784–788. doi: 10.1038/nature07731

Yang, Q., Gilmartin, G. M., and Doublé, S. (2011). The structure of human cleavage factor I(m) hints at functions beyond UGUA-specific RNA binding: a role in alternative polyadenylation and a potential link to 5’ capping and splicing. RNA Biol. 8, 748–753. doi: 10.4161/tra.8.5.16040

Yoshimoto, R., Kataoka, N., Okawa, K., and Ohno, M. (2009). Isolation and characterization of post-splicing lariat-intron complexes. Nucleic Acids Res. 37, 891–902. doi: 10.1093/nar/gkn1002

Zamorano, A., López-Camarillo, C., Orozco, E., Weber, C., Guillen, N., and Marchat, I. A. (2008). In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for Entamoeba histolytica mRNA polyadenylation. Comput. Biol. Chem. 32, 256–263. doi: 10.1016/j.compbiolchem.2008.03.019

Zhang, H., Ehrenkaufer, G. M., Pompey, J. M., Hackney, J. A., and Singh, U (2008). Small RNAs with 5’-polysphosphate termini associate with a piwi-related protein and regulate gene expression in the single-celled eukaryote Entamoeba histolytica. PLoS Pathog. 4:e1000219. doi: 10.1371/journal.ppat.1000219

Zhang, H., Pompey, J. M., and Singh, U (2011). RNA interference in Entamoeba histolytica: implications for parasite biology and gene silencing. Future Microbiol. 61, 103–117. doi: 10.2217/fmb.10.154

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with one of the authors JV-F.

Copyright © 2018 Valdés-Flores, López-Rosas, López-Camarillo, Ramirez-Moreno, Opinia-Villa and Marchat. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.