Proteomics Uncovers Novel Components of an Interactive Protein Network Supporting RNA Export in Trypanosomes

Authors

Alexandre Haruo Inoue, Patricia Ferreira Domingues, Mariana Serpeloni, Priscila Mazzocchi Hiraiwa, Newton Medeiros Vidal, Erin R. Butterfield, Ricardo Canavate del Pino, Adriana Ludwig, Cordula Boehm, Mark C. Field, and Andréa Rodrigues Ávila

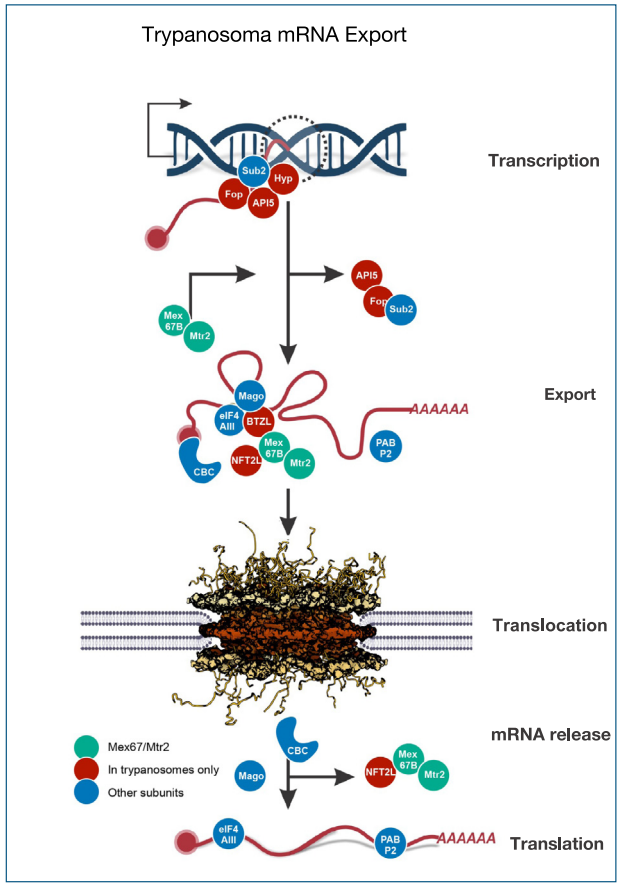
Correspondence

mfield@mac.com; andrea.avila@fiocruz.br

In Brief

This work was executed under full ethical compliance, and authorship is limited to those who have made significant contributions. The manuscript is original work, and works of others have been appropriately cited. We provide raw data for appropriate datasets at public databases. This work was performed standard compliant with community-acceptable guidelines and parameters. No known hazard was caused, nor any involvement of human subjects. Animal use for production of antibodies was performed under the institutional ethical guidelines.

Graphical Abstract



Highlights

- Gene expression in trypanosomes is mediated by noncanonical mechanisms.
- Trypanosome mRNA nuclear export system comprises unique proteins to kinetoplastids.
- The present work highlights an amalgam of kinetoplastid-specific and conserved components.
- Our data support a highly coupled mRNA maturation pathway.

2022, Mol Cell Proteomics *21*(*3*), 100208

https://doi.org/10.1016/j.mcpro.2022.100208

Proteomics Uncovers Novel Components of an Interactive Protein Network Supporting RNA Export in Trypanosomes

Alexandre Haruo Inoue[1,‡], Patricia Ferreira Domingues[1,‡], Mariana Serpeloni[1], Priscila Mazzocchi Hiraiwa[1], Newton Medeiros Vidal[2], Erin R. Butterfield[3], Ricardo Canavate del Pino[3], Adriana Ludwig[1], Cordula Boehm[3], Mark C. Field[3,4,*], and Andréa Rodrigues Ávila[1,*]

In trypanosomatids, transcription is polycistronic and all mRNAs are processed by *trans*-splicing, with export mediated by noncanonical mechanisms. Although mRNA export is central to gene regulation and expression, few orthologs of proteins involved in mRNA export in higher eukaryotes are detectable in trypanosome genomes, necessitating direct identification of protein components. We previously described conserved mRNA export pathway components in *Trypanosoma cruzi*, including orthologs of Sub2, a component of the TREX complex, and eIF4AIII (previously Hel45), a core component of the exon junction complex (EJC). Here, we searched for protein interactors of both proteins using cryomilling and mass spectrometry. Significant overlap between TcSub2 and TceIF4AIII-interacting protein cohorts suggests that both proteins associate with similar machinery. We identified several interactions with conserved core components of the EJC and multiple additional complexes, together with proteins specific to trypanosomatids. Additional immunoisolations of kinetoplastid-specific proteins both validated and extended the superinteractome, which is capable of supporting RNA processing from splicing through to nuclear export and cytoplasmic events. We also suggest that only proteomics is powerful enough to uncover the high connectivity between multiple aspects of mRNA metabolism and to uncover kinetoplastid-specific components that create a unique amalgam to support trypanosome mRNA maturation.

Trypanosoma cruzi is the causative agent of Chagas disease, a significant infection of humans and animals in Latin America where 10 million people are estimated to be infected (1). Climate change and human migration have both increased the range of the disease and risk for much of the Americas and Europe; recent global events will likely inflate this further, as well as compromise the ability of healthcare organizations to deliver therapeutics (2). Trypanosomatids diverged some 1 billion years ago from other eukaryotes and adapted to specific hosts and life cycle demands through multiple alterations in metabolic pathways, stage-specific gene expression, and other mechanisms (3, 4). Novel aspects of these processes in trypanosomes have attracted considerable interest, both through providing insight into diverse mechanisms of mRNA processing and representing possible therapeutic targets (5–7).

Unlike most eukaryotes, trypanosome genes are transcribed as polycistronic RNAs, frequently encompassing tens of open reading frames, while mature mRNAs (processed by *trans*-splicing and polyadenylation) exhibit differential abundance as a consequence of posttranscriptional regulation (8–11). The latter includes mRNA processing, differential *trans*-splice site utilization, mRNA stability as well as translation efficiency (12–14) and involves RNA-binding proteins and *cis*-regulatory elements within mature mRNA UTRs (15). Cytosolic RNA granules also impact mRNA abundance by interacting with the translation and stability machinery to regulate expression during differentiation processes, mainly in response to biological stimuli such as heat shock or starvation stress (16–21). Trypanosomes have multiple distinct RNA granules, and their composition is divergent from animals and fungi. For example, during translational repression, some populations of RNA granule are compositionally stable (22), while others are dynamic (23–25). RNA granules at the nuclear periphery may regulate mRNA fate by preventing misprocessed mRNAs from reaching the translation machinery (26, 27). However, it remains unclear how these complexes interact with the nuclear components involved in RNA export.

From the [1]Instituto Carlos Chagas, FIOCRUZ, Curitiba, Paraná, Brazil; [2]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA; [3]School of Life Sciences, University of Dundee, Dundee, Scotland, UK; [4]Biology Centre, University of South Bohemia, České Budějovice, Czech Republic

‡These authors contributed equally to this work.

*For correspondence: Andréa Rodrigues Ávila, andrea.avila@fiocruz.br; Mark C. Field, mfield@mac.com.

https://doi.org/10.1016/j.mcpro.2022.100208

In animals and fungi, mRNA export is initiated by cotranscriptional association of nascent mRNAs with the THO complex. THO is a conserved multimeric complex required for mRNP biogenesis consisting of at least four components: Tho2, Hpr1, Mft1, and Thp2 (28). THO also associates with two mRNA export factors, UAP56 and ALY (Sub2 and Yra1 in budding yeast), to form the TREX complex that couples transcription with mRNA nuclear export (29). In animals, recruitment of TREX to mRNA occurs by a splicing-coupled mechanism rather than a direct transcription-coupled mechanism as in yeast (30, 31). A second complex, TREX-2, links transcription and the nuclear pore complex (NPC) by facilitating transfer of mature mRNPs from the nuclear interior to the NPC (32). In yeast, TREX-2, composed of Sac3, Thp1, Sem1, Sus1 and Cdc3, regulates a surprisingly diverse number of chromatin-associated processes (33, 34).

The transport receptor, TAP (Mex67 in yeast), is recruited by TREX or TREX-2 *via* the UAP56/Sub2 subunits (35), to interact with the NPC and export mRNA to the cytoplasm. TAP/Mex67 share a modular domain organization that includes the aminoterminal RNA recognition motif (RRM) followed by Nuclear Transport Factor 2-like (NTF2L) and the ubiquitin associated (UBA) (36). The NTF2-like superfamily is a large group of related proteins that share a common fold, first observed in the structure of the rat NTF2 protein (37, 38).

Alternatively, UAP56/Aly can also be recruited by the exon-junction complex (EJC), which is composed of eIF4AIII, MLN51 (also known as BTZ/Barentsz), Y14 and Mago (or Magoh/Mago-nashi) in mammalian cells. The EJC is loaded onto immature mRNA molecules during *cis*-splicing at a conserved position upstream of the exon–exon junction (39) to form a stable mRNA export complex and is only removed from mRNP complexes when translation is initiated in the cytoplasm; hence, the EJC accompanies an mRNA molecule throughout nuclear export (40). The EJC also acts as a platform for transient binding of many proteins that define the fate of individual mRNAs *via* specific maturation pathways (41, 42). For example, the EJC can bind to the mRNA export factors REF/Aly and TAP/p15 to enhance the efficiency of transport of spliced mRNAs into the cytoplasm (43). EJC complex components have diversified in individual lineages, and not all components are retained universally, for example, MLN51 is not conserved in some trypanosomes (44, 45). Some EJC components have been identified in trypanosomes, but the direct association with the mRNA export machinery remains to be demonstrated.

Although many components of the RNA export machinery are known from animals, fungi, plants, and other organisms, *in silico* identification of components in trypanosomatid genomes is unreliable due to high divergence and few proteins appear conserved or recognizable (46). In fact, several nuclear complexes and nuclear-pore-associated proteins that play major roles in mRNA export in opisthokonts are lacking in trypanosomes, and a recent comparative review described the main differences in nuclear mRNA maturation and mRNA export (47). While UAP56/Sub2 (48) and TAP/Mex67 (49–51) are essential for mRNA transport in trypanosomes, little else has been identified with significant roles (49, 51–53). In *Trypanosoma brucei*, the general mRNA export factors Mex67 and Mtr2 interact with members of the 60S ribosomal subunit (54) and play distinct roles in tRNA export (55), confirming the presence of unique structures and potentially parasite-specific interactions for trypanosome RNA export. Further, Mex67 associates with Ran in trypanosomes, which is distinct from animals and fungi. Trypanosomes can initiate mRNA nuclear export cotranscriptionally, and initiation of export does not depend on completion of *trans*-splicing, suggesting that trypanosomes regulate completion as opposed to initiation of export. This is supported by the divergence of the cytoplasmic face composition of the NPC in trypanosomes (56) together with cytoplasmic nuclear peripheral granules (NPGs), but the components of both the NPC and export machinery responsible await identification (27).

Progress in characterizing the trypanosome NPC, kinetochores, and nuclear lamina has been achieved mainly through proteomics, essentially the only available approach due to high primary structure divergence (56–59). Here we have also explored proteomics with cryomilling and immunoisolation to identify proteins associated with validated components of the mRNA export pathway in *T. cruzi*, TcSub2, and TceIF4AIII. TcSub2 is a TREX complex protein involved in mRNA export (48), while TceIF4AIII is an RNA helicase that shuttles between the nucleus and cytoplasm *via* an Mex67-dependent route (60). We identify not only conserved components but also proteins specific to trypanosomatids. Among them, we identified a kinetoplastid-specific NTF2-like protein that interacts with Mex67, crucial for mRNA export. These observations indicate considerable divergence in the composition and function of trypanosome mRNA export factors.

EXPERIMENTAL PROCEDURES

Parasite Culture

T. cruzi Dm28c epimastigotes were maintained in axenic culture in liver infusion tryptose (LIT) medium at 28 °C (61). Procyclic forms of *T. brucei* Lister 427 and RNA interference *T. brucei* 29-13 line were maintained in SDM-79 medium supplemented with 10% fetal bovine serum at 28 °C (62). For *T. brucei* 29-13, medium was also supplemented with G418 (15 μg/ml) and hygromycin (50 μg/ml).

Affinity Purification of Protein Complexes

Anti-GFP nanobodies were expressed and purified (63). Antibodies (Sigma) or anti-GFP nanobodies were coupled to Dynabeads M-270 epoxy (Life technologies) (64). *T. cruzi* cell extracts were prepared by cryomilling with a Planetary Ball PM100 (RETSCH, UK) and affinity purification of GFP-tagged proteins (64). Pullout of TcSub2::GFP was achieved using a buffer containing 20 mM HEPES pH 7.4, 10 μM $CaCl_2$, 1 mM $MgCl_2$, 10 mM sodium citrate, 0.1% CHAPS (w/v), and protease inhibitors (COMPLETE Mini Protease inhibitor cocktail tablet, Roche); for TceIF4AIII, TcMago, TcNTF2L, and TcMex67 a buffer with

20 mM HEPES pH 7.4, 10 μM $CaCl_2$, 1 mM $MgCl_2$, 50 mM sodium citrate, 0.1% Triton X-100 (v/v), and protease inhibitors (COMPLETE Mini Protease inhibitor cocktail tablet, Roche); and for TcFOP and TcAPI5, a buffer with 20 mM HEPES pH 7.4, 10 μM $CaCl_2$, 1 mM $MgCl_2$, 50 mM sodium citrate, 0.1% Triton X-100 (v/v), and 10% glycerol (v/v). Aliquots of purified complexes were separated on 4 to 12% NuPAGE Novex Bis-Tris precast gels (Life technologies) and stained with SilverQuest Silver Staining (Life technologies).

For TcFOP, TcAPI5, TcNTF2L, and TcHYP, the complex was analyzed by mass spectrometry. *T. cruzi* cell extracts were prepared using 5×10^8 parasites lysed in 1 ml of citrate buffer (50 mM sodium citrate, 20 mM Hepes pH 7.4, 1 mM $MgCl_2$, 1 μM $CaCl_2$, 0,5% Triton X-100, protease inhibitor (COMPLETE Mini Protease inhibitor cocktail tablet, Roche) by sonication. Pullout of TcAPI5::GFP, TcFOP::GFP, TcNTF2L::GFP, and TcHYP::GFP was achieved by affinity purification of GFP-tagged proteins (64). For TcFOP and TcAPI5 pullout, 10% glycerol was added to citrate buffer. The complexes were eluted using 15 μl of elution buffer (2% SDS, 20 mM Tris-HCl pH 8.0) at 72 °C for 20 min. Aliquots of purified complexes were separated on 13% SDS-PAGE and stained with Silver Staining protocol.

Proteomics: Experimental Design and Statistical Rationale

Initially, TcSub2 and TceIF4AIII complexes were affinity purified and, for those analyses, three different technical replicates were performed for each TcSub2::GFP or TceIF4AIII::GFP complex. In each experiment, two samples were compared: TcSub2::GFP or TceIF4AIII::GFP pullout with a wild-type pullout (control). Next, samples were processed and identified by LC-MS/MS as previously described. Proteins were compared between TcSub2 and TceIF4AIII complex with the control samples, and only proteins identified in at least two replicates and with emPAI index >0.1 were considered for analyses.

For TcMex67, TcAPI5, TcFOP, TcNTF2L, and TcHYP complexes, different replicate numbers were processed or reprocessed in different experiments to increase the reliability and the coverage. Three biological replicates were used for TcMex67 analyses. For TcAPI5, seven replicates were used: three biological replicates, but technical triplicates were performed for two biological replicates. For FOP, TcNTF2L, and TcHYP, five technical replicates were used for each protein: three biological replicates, but technical replicates were performed for one biological replicates. In each replicate, two samples were compared: target pullout and wild-type pullout (control). The TcMex67, TcAPI5, TcFOP, TcNTF2L, and TcHYP data analysis was performed based on the averages of all replicates and comparing the protein intensities between the target and control samples. The fold change ≥2 was considered to select the proteins.

Mass-Spectrometry-Based Proteomics

For TcSub2 and TceIF4AIII samples, mass spectrometry was performed at the FingerPrints Proteomics Facility at the University of Dundee (Scotland, UK). Samples were purified and desalted by electrophoresis into 4 to 12% NuPAGE Novex Bis-Tris precast gel (Life technologies), silver-stained, excised, and dehydrated in a SpeedVac (Thermo Scientific). The dehydrated gel was digested with trypsin (Modified Sequencing Grade, Roche) for 18 h at 30 °C. Tryptic peptides were extracted from the gel with 95% Acetonitrile/5% Formic acid (v/v) and captured with an Acclaim PepMap 100 system (C18, 100 μm × 2 cm) and fractionated on a C18 Easy-Spray PepMap RSLC (75 μm × 50 cm) (Thermo Scientific) columns with an Ultimate 3000 RS LCnano system (Thermo Scientific) coupled to an LTQ OrbiTrap Velos Pro (Thermo Scientific). For identification of proteins, the Mascot Search Daemon (Version 2.4.1) (Matrix Science) (http://www.matrixscience.com/) was used with a custom *T. cruzi* protein sequence library of 54,610 sequences from five different strains (CL Brener Esmeraldo-like, CL Brener non-Esmeraldo-like, Sylvio, *Dm*28c, and Marinkellei, downloaded on February 4, 2015 from Uniprot). This strategy increased coverage and hence the likelihood of identification of peptides. MS tolerance allowed up to two missed and/or nonspecific tryptic cleavages, and peptide mass tolerance and fragment mass tolerance were set as ±10 ppm and ±0.6 Da, respectively. Carbamidomethylation (C) was set as a fixed modification, and Acetyl (N-term), Dioxidation (M), Glu->pyro-Glu (N-term E), Oxidation (M) were included as variable modifications. FDR was calculated using a reverse translated proteome. This target-decoy strategy uses along with the target database a decoy database (randomized database, reverse target database used in maxquant). The number of matches to this decoy database equals the number of random matches (false positives) obtained in the target database and is used to calculate local or global FDRs. The relative abundance of proteins was estimated using the exponentially modified protein abundance index (emPAI score), determined in a peptide mixture based on the coverage of peptides from the identified protein (65) with a 1% FDR and significance p value ≤0.05.

TcMex67, TcAPI5, TcFOP, TcNTF2L, and TcHYP samples were processed at the mass spectrometry facility at the Carlos Chagas Institute/FIOCRUZ-PR (RPT02H PDTIS/Instituto Carlos Chagas—Fiocruz Paraná, Brazil). Samples were purified by electrophoresis into SDS–polyacrylamide gels; lanes were excised and dehydrated with 100% ethanol and dried in a vacuum centrifuge. After reduction with 10 mM DTT, 50 mM ammonium bicarbonate (ABC), samples were alkylated with 50 mM iodacetamide, 50 mM ABC. Gels were washed with 50 mM ABC and dehydrated with 100% ethanol, and this step was repeated. Then, gels were incubated with 12.5 ng/ml trypsin (Promega, V5113), 50 mM ABC at 37 °C for 18 h. After trypsinization, trifluoroacetic acid (TFA) was added to a final concentration of 0.5%. Peptides were extracted from the gel matrix through incubation twice with 30% MeCN, 3% TFA, and twice with 100% MeCN. The extracted peptide solution was concentrated in a vacuum centrifuge to remove MeCN and desalted with homemade C18 spin columns and analyzed by LC-MS/MS in a Thermo Scientific Easy-nLC 1000 system coupled to a LTQ Orbitrap XL ETD. Peptide separation was carried out in a 30 cm (75 μm inner diameter) fused silica in-house packed column with reverse-phase ReproSil-Pur C18-AQ 1.9 μm resin (Dr Maisch GmbH). Chromatography runs were performed with a flow rate of 250 nl/min from 5 to 40% MeCN in 0.1% formic acid, 5% DMSO in a 120 min gradient. The mass spectrometer operated in a data-dependent mode to automatically switch between MS and MS/MS (MS^2) acquisition. Survey full-scan MS spectra (at 300–2000 m/z range) were acquired in the Orbitrap analyzer with resolution of 60,000 at m/z 400 (after accumulation to a target value of 1,000,000 in the C-trap). The 12 most intense ions were sequentially isolated and fragmented in the linear ion trap using collision-induced dissociation at a target value of 30,000. The "lock mass" option was enabled in all full scans to improve mass accuracy of precursor ions (66). Peaklist picking, protein identification, quantification, and validation were done using the MaxQuant platform (version 1.5.5.1) (67), which includes the algorithm Andromeda (68) for database searching. Default parameters of the software were used for all analysis steps, unless stated otherwise. Proteins were searched against a "decoy database" prepared by reversing the sequence of each entry of the *T. cruzi* CL Brener protein sequence database (containing 19,242 protein sequences, downloaded on August 10, 2016 from Uniprot) and appending them to the forward sequences. This database was complemented with frequently observed contaminants (porcine trypsin, *Achromobacter lyticus* lysyl endopeptidase, and human keratins) and their reversed sequences. Search parameters specified an MS tolerance of 20 ppm, an MS/MS tolerance of 0.5 Da, and full trypsin specificity, allowing for up to two missed cleavages. Carbamidomethylation of cysteine was set as a

fixed modification, and oxidation of methionine and N-terminal acetylation (protein) were allowed as variable modifications.

For validation of the identifications, a minimum of seven amino acids for peptide length and two peptides per protein was required. In addition, an FDR threshold of 1% was applied at both peptide and protein levels. When transforming peptide identifications into protein identifications, similar protein sequences (*e.g.*, isoforms) present in the database that could not be distinguished by the experimentally detected peptides were grouped and referred to as protein groups. Protein quantification was performed using a label-free approach, where peptides eluting from each LC run are detected as three-dimensional features—retention time *versus* signal intensity (extracted ion chromatogram, XIC) *versus* mass/charge—aligned and compared across runs, as previously described (69).

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE (https://www.ebi.ac.uk/pride/) partner repository with the dataset identifier PXD028072.

Bioinformatic Analyses

For TceIF4AIII phylogenetic analysis, database searches were done using BLASTp (70) and ortholog proteins identified by reciprocal best hit (RBH). Protein sequences from representative species were obtained from the NCBI database and included proteomes from *Saccharomyces cerevisiae* (fungi), *Homo sapiens* (metazoa), *Dictyostelium discoideum* (amoebozoa), *Arabidopsis thaliana* (Viridiplantae), *Plasmodium falciparum* (Chromalveolata), *Toxoplasma gondii* (Chromalveolata), *T. cruzi* (Excavata), *T. brucei* (Excavata), and *Leishmania major* (Excavata). For calculating the similarity and identity among amino acid sequences (supplemental Table S1, sheet2), the Needle program from the EMBOSS package was used (71), which allows a global alignment between two protein sequences. Phylogenetic analysis was performed using maximum likelihood (ML) by FastTree 2.1 using WAG model (72). The phylogenetic tree was drawn using FigTree v1.4 program (http://tree.bio.ed.ac.uk/software/figtree/). TcHYP ortholog proteins were identified by RBH criteria and synteny evaluation. A total of 108 positions with more reliable alignment obtained by PSI-coffee (73) composed the final dataset. The tree was rooted using the *Bodo saltans* sequence. An evolutionary tree of TcHYP (TcCLB.506435.150) orthologs was inferred by using MrBayes 3.2.6 (74). Markov Chain Monte Carlo (MCMC) was run for 1,000,000 generations, sampling trees every 1000 generations. An ML tree (JTT +F + G + I) was also inferred using MegaX (75) with 1000 bootstrap replicates. For TcAPI5, TcFOP, and TcNTF2L, *T. cruzi* CL Brener Non-Esmeraldo-like sequences were obtained from TriTrypDB version 33 and used as queries for BLASTp against a database containing 121 eukaryotic predicted proteomes (supplemental Table S2) with an e-value threshold of 0.1. All hits identified were subject to reverse BLASTp using an e-value threshold of 0.1. Orthology was predicted based upon by RBH criteria being the original *T. cruzi* query sequence. Additional hits were identified using sequences obtained from *B. saltans* and *Dictyostelium purpureum*. Predicted orthologous sequences were aligned using MUSCLE (76), trimmed manually, and both ML and BA phylogenetic trees were constructed using PhyML (77) and MrBayes (74), respectively. PhyML 3.0 was used with default settings, and a bootstrap of 1000 and MrBayes 3.2.6 was used with an MCMC generation of 800,000, Γ shape rate variation with four discreet categories and a sampling frequency of 1000 with the first quarter as burn-in. MrBayes trees were run using the CIPRES Science Gateway portal (78). Proteins were identified by accession number in the TriTryp database (79). Protein domains were searched using InterProScan (80) or HMMscan from HMMER (81) looking for hidden Markov models (HMM) available in Pfam database 27.0 version (82).

Interactome visualization was performed using Cytoscape (available in https://cytoscape.org) from selected data from Tables 1 and 2. A static representation was generated from Cytoscape and processed using Affinity Designer (available in https://affinity.serif.com/en-gb/).

Polyclonal Antibody Production

The open reading frames of TcAPI5, TcFOP, TcNTF2L, and TcMago were amplified by PCR using specific oligonucleotide primers as shown in supplemental Table S3. *T. cruzi* Dm28c genomic DNA was used as template. The PCR products were cloned into the pDONR221 vector from Gateway technology (Invitrogen) and further recombination into the pDEST17 vector (Invitrogen) to produce his-tagged recombinant protein, according to the manufacturer's protocol. Expression of recombinant protein was induced in *Escherichia coli* BL21 (DE3) by addition of 1 mM IPTG and incubation for 3 h at 37 °C. His-tagged protein was purified by affinity chromatography on Ni-NTA resin (Qiagen) under denaturing conditions and used to inoculate mice to produce polyclonal antibodies, according to Inoue *et al.* (2014) (60). Animals were immunized by intraperitoneal injection with approximately 50 μg of the his6-tagged protein in Freund's complete adjuvant (Sigma) for the first inoculation. For the three booster injections, we used 20 μg of the recombinant protein in Alu-Gel (Serva) administered at 2-week intervals. Antiserum was obtained 5 days after the last booster injection and by blood collection *via* cardiac puncture. The protocol was approved by Ethics Commission on Animal Use (CEUA) from FICRUZ, protocol 47/12-3, license LW15/13; project title: "Characterization of trypanosome proteins."

Immunoblotting

Proteins were separated by gel electrophoresis (SDS-PAGE) and transferred to a nitrocellulose membrane (Hybond C, Amersham Biosciences). The membrane was blocked with 0.1% Tween 20 and 5% milk in phosphate-buffered saline (PBS). Primary antibodies were diluted in blocking solution at the following concentrations: mouse anti-TceIF4AIII (diluted 1:500) (60), rabbit anti-Protein A (Sigma-Aldrich, diluted 1:40,000), rabbit anti-TcSub2 (diluted 1:1000) (48), mouse anti-TcAPI5 (diluted 1:500), mouse anti-TcFOP (diluted 1:500), mouse anti-TcNTF2L (diluted 1:500), mouse anti-TcMago (diluted 1:500), rabbit anti-GFP (kindly provided by Dr Stenio Fragoso Perdigão, diluted 1:500), mouse anti-GFP (Roche, diluted 1:500), rat anti-HA (Roche, diluted 1:500). Antibodies were incubated with the membrane for 1 h. The membrane was washed three times in 0.1% Tween 20 in PBS. Bound antibodies were detected by fluorescence with IRDye whole IgG secondary antibody (LI-COR). Relative quantification of Western blots was performed using ImageJ software (available in https://imagej.nih.gov/ij/).

Protein Tagging

The coding sequences of TceIF4AIII, TcMago, TcAPI5, TcHYP, TcFOP, and TcFOPIΔFOP, a mutated version of TcFOP without the FoP domain, cloned previously into pDONR221 (Invitrogen) were recombined pTcGW (83) for N-terminal eGFP or PTP tagging. Similarly, the coding sequence of TcNTF2L was cloned into pTcGW1.1 (80) for C-terminal eGFP/PTP tagging. *T. cruzi* epimastigotes were transfected with these plasmids, and stable lines were selected by adding 500 μg/ml G418 to the culture medium. Cell lines carrying TcSub2 tagged at N-terminal with GFP and TcMex67 tagged at C-terminal with PTP were described in Serpeloni *et al.* (2011) (48) and Kugeratski *et al.* (2015) (84), respectively. In. *T. brucei*, cell lines expressing HA-tagged or GFP-tagged TbMago were generated by *in situ* tagging with the pMOTag4H or pMOTag4G plasmid, as described previously in Oberholzer *et al.* (2006) (85). The parasites were analyzed by

TABLE 1
Lists of filtered proteins identified in TceIF4AIII and TcSub2 isolates

eIF4III isolation only			Sub2 isolation only			Both Sub2 and eIF4III isolation		
Designation[a]	Location, annotation[b]	Accession	Designation	Location, annotation	Accession	Designation	Location, annotation	Accession
NTF2L[c]	Cytoplasmic, RNA binding, N-terminal NTF2 domain	TcCLB.511367.220	*BTZL*[c]	Nuclear, RRM superfamily	TcCLB.508879.80	Sub2[c]	Nuclear, ATP-dependent RNA helicase SUB2	TcCLB.508319.40
PABP2[c]	Cytoplasmic (granules), polyadenylate-binding protein	TcCLB.508461.140	Y14[c]	Nuclear, RNA-binding protein Y14/8A	TcCLB.508213.40	eIF4AIII[c]	Nuclear/Nucleolar ATP-dependent RNA helicase FAL1	TcCLB.506587.40
Ce1[d]	Nucleolar, RNA 3′-terminal phosphate cyclase	TcCLB.510761.30	TSR1[c]	Nuclear, splicing factor TSR1, SR type	TcCLB.503715.10	**API5**[c]	Nuclear, mRNA binding, ARM repeats	TcCLB.511807.280
Lark-like[c]	Cytoplasmic, dual RRM RNA-binding protein	TcCLB.511727.270	TSR1IP[c]	Nuclear, splicing factor ptsr1 interacting protein, SR type	TcCLB.504105.160	**FOP**[c]	Nuclear, Zn finger CCCH	TcCLB.509033.80
U3 snRAP6[e]	Nucleolar, U3 small nucleolar RNA-associated protein 6	TcCLB.510769.30	U2AF26[c]	Nuclear, U2 splicing auxiliary factor	TcCLB.503577.20	Mago[c]	Nuclear, Mago nashi	TcCLB.506945.200
Dhh1[c]	Cytoplasmic, ATP-dependent DEAD/H RNA helicase	TcCLB.506959.30	RRM1[c]	Nuclear, RNA-binding protein SR type	TcCLB.511621.50	PRP19[c]	Nuclear, PRP19 domain, WD40 repeats	TcCLB.509103.10
ZC3H41[c]	Cytoplasmic, Zn-finger and RNA helicase, ZC3H41	TcCLB.508355.330	RBSR2/SR34[c]	Nuclear, spicing factor, SR type	TcCLB.503919.30	*Hel67/DDX3/DBP1*[f]	Nuclear/cytoplasmic, ATP-dependent RNA helicase Hel67/DDX3	TcCLB.506213.120
ZC3H39[c]	Cytoplasmic (granules), RING and RNA-binding protein	TcCLB.506211.70	RBP33[c]	Nuclear, RNA-binding protein 33, spicing factor	TcCLB.508569.90	*DRBD2*[c]	Nuclear, Double RNA-binding domain protein 2	TcCLB.510755.120
NOP5[e]	Nucleolar, pre-snoRNA splicing protein, NOP domain	TcCLB.508277.230	DRBD18[c]	Nuclear, double RBD, extensive interactions	TcCLB.511727.190	DRBD3[c]	Nuclear/cytoplasmic, double RNA-binding domain protein 3	TcCLB.506649.80
NOP56[e]	Nucleolar, pre-snoRNA splicing protein, NOP domain	TcCLB.506189.10	*ZC3H40*[c]	Cytoplasmic, mRNA binding and stabilizing, Zn finger type	TcCLB.506211.60	ALBA1[g]	ALBA1	TcCLB.504001.10

TABLE 1—*Continued*

eIF4III isolation only			Sub2 isolation only			Both Sub2 and eIF4III isolation		
Designation[a]	Location, annotation[b]	Accession	Designation	Location, annotation	Accession	Designation	Location, annotation	Accession
NOP136/ BMS1[d]	Nucleolar, ribosomal biogenesis	TcCLB.510899.59	UBP-2[c]	U-rich RNA-binding protein, mRNA destabilizer	TcCLB.507093.229	*ALBA2*[g]	ALBA2	TcCLB.504001.20
Tc38[c]	Nucleolar/ cytoplasmic, RNA-binding	TcCLB.503833.50	Hel64/ DBP2B[c]	Nuclear/nucleolar, ATP-dependent RNA helicase DDX3X domain	TcCLB.508973.50	ALBA3[g]	ALBA3	TcCLB.510877.30
			Mle[c]	Cytoplasmic, dsRNA unwindling Mle RNA helicase	TcCLB.505123.4	Nucleolar RNA helicase II[d]	Nucleolar, DEAD box helicase, GUCT domain	TcCLB.508205.20
			eEF1a[f]	Cytoplasmic, elongation factor 1a	TcCLB.511369.10	*SCD6*[f]	Cytoplasmic, LSm domain, Trailer hitch protein	TcCLB.507093.300
			Mex67B[f]	Cytoplasmic/ nuclear Mex67B	TcCLB.506127.50	Histone 3[g]	Nuclear, histone H3	TcCLB.505931.50
						Hypothetical	Unknown, low complexity regions	TcCLB.506435.150

[a]Protein name according to conserved and specific domain identified using INTERPRO, HMMER, PFAM, or ortholog in NCBI nr database. Bold represents kinetoplastid-specific granules, and italic represents nuclear peripheral granules.

[b]Location based on TrypTag, mass spectrometry, or direct experimental evidence. Restricted to either *T. cruzi* or *T. brucei* evidence, except in case of highly conserved ortholog, where evidence from animal or fungal cells is accepted.

[c]mRNA, underline trans-splicing factor.

[d]rRNA/nucleolar.

[e]sn/snoRNA.

[f]Translation.

[g]Chromatin.

TABLE 2
Lists of filtered proteins identified in TcHyp, TcAPI5, TcFOP, and TcNTF2L isolates

TcCLB.506435.150 isolation			API5 isolation			FOP isolation			NTF2L isolation		
Designation[a]	Location, annotation[b]	Accession	Designation[a]	Location, annotation[b]	Accession	Designation[a]	Location, annotation[b]	Acession	Designation[a]	Location, annotation[b]	Acession
Hyp[c]	Unknown, low complexity regions	TcCLB.506435.150	**API5**[c]	Nuclear, mRNA binding, ARM repeats	TcCLB.511807.280	**FOP**[c]	Nuclear, Zn finger CCCH	TcCLB.509033.80	**NTF2L**[c]	Cytoplasmic, RNA binding, N-terminal NTF2 domain	TcCLB.511367.220
Mex67B[c]	Cytoplasmic/nuclear Mex67B	TcCLB.506127.50	**FOP**[c]	Nuclear, Zn finger CCCH	TcCLB.509033.80	Importin - beta[c]	importin beta-1 subunit, putative[c]	TcCLB.504105.150	*PABP2*[c]	Cytoplasmic (granules), polyadenylate-binding protein	TcCLB.508461.140
S27a[d]	ubiquitin/ribosomal protein S27a, putative	TcCLB.510409.39	Mex67B[c]	Cytoplasmic/nuclear Mex67B	TcCLB.506127.50	eiF4AIII[c]	Nuclear/Nucleolar ATP-dependent RNA helicase FAL1	TcCLB.506587.40	*DRBD2*[c]	Nuclear, Double RNA-binding domain protein 2	TcCLB.510755.120;
PABP1[e]	polyadenylate-binding protein 1, putative	TcCLB.506885.70	eiF4AIII[c]	Nuclear/Nucleolar ATP-dependent RNA helicase FAL1	TcCLB.506587.40	Mago[c]	Nuclear, Mago nashi	TcCLB.506945.200	Tc38[c]	Nucleolar/ cytoplasmic, RNA-binding	TcCLB.503833.50
DED1[c]	ATP-dependent RNA helicase, putative	TcCLB.510661.90	DRBD4[c]	polypyrimidine tract-binding protein, putative	TcCLB.511727.160	NAP[f]	nucleosome assembly protein (NAP), putative	TcCLB.509003.10	ALBA4[e]	DNA/RNA-binding protein Alba 4	TcCLB.510877.40
SUB2[c]	Nuclear, ATP-dependent RNA helicase SUB2	TcCLB.508319.40	**Hyp**[c]	Unknown, low complexity regions	TcCLB.506435.150	Importin alfa[c]	Importin subunit alpha	TcCLB.509057.20	Hel67/ DDX3/ DBP1[e]	Nuclear/cytoplasmic, ATP-dependent RNA helicase Hel67/DDX3	TcCLB.511285.120
hnRNPH[c]	heterogeneous nuclear ribonucleoprotein H/F, putative	TcCLB.511109.130	Mago[c]	Nuclear, Mago nashi	TcCLB.506945.200	DRBD4[c]	polypyrimidine tract-binding protein, putative	TcCLB.511727.160	ZC3H41[c]	Cytoplasmic, Zn-finger and RNA helicase, ZC3H41	TcCLB.508355.330
ZC3H40[c]	RNA-binding protein, putative	TcCLB.508895.60	Fibrillarin[d]	nucleolar RNA binding protein, putative	TcCLB.508277.230	hnRNPH[c]	Heterogeneous nuclear ribonucleoprotein H/F, putative	TcCLB.511109.130	DRBD4[c]	polypyrimidine tract-binding protein, putative	TcCLB.511727.160
eiF4AIII[c]	Nuclear/Nucleolar ATP-dependent RNA helicase FAL1	TcCLB.506587.40	DRBD3[c]	Nuclear/cytoplasmic, double RNA-binding domain protein 3	TcCLB.508349.39	NRBD[c]	nuclear RNA binding domain	TcCLB.511727.290	RBP[c]	RNA-binding protein, putative	TcCLB.509317.60
RanBP1[c]	Ran-binding protein 1, putative	TcCLB.507099.30	SMD3[c]	small nuclear ribonucleoprotein sm d3	TcCLB.508257.150	*Dhh1*[c]	Cytoplasmic, ATP-dependent DEAD/ H RNA helicase	TcCLB.506959.30	eEF1a[e]	Cytoplasmic, elongation factor 1a	TcCLB.511369.10;
DBP2A[c]	ATP-dependent RNA helicase DBP2A,	TcCLB.510187.290	UBP-2[c]	U-rich RNA-binding protein, mRNA destabilizer	TcCLB.507093.229	Hel67/ DDX3/ DBP1[e]	Nuclear/cytoplasmic, ATP-dependent RNA helicase Hel67/DDX3	TcCLB.506213.120	Nonsense Reg[c]	regulator of nonsense transcripts 1	TcCLB.511317.30
MTR2[c]	mRNA transport regulator MTR2, putative	TcCLB.508173.180	SUB2[c]	Nuclear, ATP-dependent RNA helicase SUB2	TcCLB.508319.40	NOP5[d]	Nucleolar, pre-snoRNA splicing protein, NOP domain	TcCLB.508277.230	ZC3H39[c]	RNA-binding protein, putative	TcCLB.508895.50
RiboHII[c]	ribonuclease HII, putative	TcCLB.510287.60	*SCD6*[e]	Cytoplasmic, LSm domain, Trailer hitch protein	TcCLB.507093.300	ZC3H40[c]	RNA-binding protein, putative	TcCLB.506211.60	ZC3H40[c]	RNA-binding protein, putative	TcCLB.508895.60
eIF5A[e]	eukaryotic translation initiation factor 5A	TcCLB.506925.130	**NTF2L**[c]	Cytoplasmic, RNA binding, N-terminal NTF2 domain	TcCLB.511367.220	ZC3H39[c]	RNA-binding protein, putative	TcCLB.508895.50	*PABP1*[e]	polyadenylate-binding protein 1, putative	TcCLB.506885.70

TABLE 2—*Continued*

TcCLB.506435.150 isolation			API5 isolation			FOP isolation			NTF2L isolation		
Designation[a]	Location, annotation[b]	Accession	Designation[a]	Location, annotation[b]	Accession	Designation[a]	Location, annotation[b]	Acession	Designation[a]	Location, annotation[b]	Acession
snoRBP[d]	nucleolar RNA-binding protein, putative	TcCLB.507649.80	ZC3H40[c]	RNA-binding protein, putative	TcCLB.508895.60	DRBD3[c]	Nuclear/cytoplasmic, double RNA-binding domain protein 3	TcCLB.506649.80	*Hel67/ DDX3/ DBP1*[e]	Nuclear/cytoplasmic, ATP-dependent RNA helicase Hel67/DDX3	TcCLB.506213.120
Fibrillarin[d]	Nucleolar RNA binding protein, putative	TcCLB.508277.230	*Dhh1*[c]	ATP-dependent DEAD/H RNA helicase, putative	TcCLB.506959.30	*SCD6*[e]	Cytoplasmic, LSm domain, Trailer hitch protein	TcCLB.507093.300	PUM[c]	pumilio protein, putative	TcCLB.507049.199
FOP[c]	Nuclear, Zn finger CCCH	TcCLB.509033.80	*PABP2*[c]	Cytoplasmic (granules), polyadenylate-binding protein	TcCLB.508461.140	RBP42[c]	RNA-binding protein 42 (RNA-binding motif protein 42), putative	TcCLB.509167.140	Ran[c]	GTP-binding nuclear protein rtb2, putative	TcCLB.509455.80
NOP56[g]	Nucleolar, ribosomal biogenesis	TcCLB.506189.10	PRP19[c]	Nuclear, PRP19 domain, WD40 repeats	TcCLB.509103.10	**Hyp**[c]	Unknown, low complexity regions	TcCLB.506435.150	PUF6[c]	pumilio/PUF RNA binding protein 6, putative	TcCLB.510125.10
RNA Hel[c]	ATP-dependent DEAD/H RNA helicase, putative	TcCLB.507641.120	Histone H3[f]	Nuclear, histone H3	TcCLB.505931.50	snoRBP[d]	Nucleolar RNA-binding protein, putative	TcCLB.507649.80	snoRBP[d]	nucleolar RNA-binding protein, putative	TcCLB.507649.80
DRBD2[c]	RNA-binding protein, putative	TcCLB.510755.120				RBP2[c]	DNA-directed RNA polymerase III subunit, putative	TcCLB.505997.210	ZFP2[c]	zinc finger protein 2, putative	TcCLB.503989.10
eEF1-beta[e]	translation elongation factor 1-beta, putative	TcCLB.509733.100				RNA Hel II[c]	nucleolar RNA helicase II, putative	TcCLB.506123.40	CAP - methyltrans-ferase[e]	mRNA cap guanine-N7 methyltransferase, putative	TcCLB.508799.80
snoRBP[d]	nucleolar RNA-binding protein, putative	TcCLB.510859.17				RBP42[c]	RNA-binding protein 42 (RNA-binding motif protein 42),	TcCLB.509167.140	RBP31[c]	RNA-binding protein, putative	TcCLB.510007.30
NTF2L[c]	Cytoplasmic, RNA binding, N-terminal NTF2 domain	TcCLB.511367.220				DBP2A[c]	ATP-dependent RNA helicase DBP2A, putative	TcCLB.510187.290	u2af26[c]	U2 splicing auxiliary factor, putative	TcCLB.510943.60;
PABP2[c]	Cytoplasmic (granules), polyadenylate-binding protein	TcCLB.508461.140				SmD2[c]	small nuclear ribonucleoprotein SmD2, putative	TcCLB.511189.80	UBP-2[c]	U-rich RNA-binding protein, mRNA destabilizer	TcCLB.507093.229;
eIF4A1[e]	Eukaryotic initiation factor 4A-1	TcCLB.511585.190				*DRBD2*[c]	Nuclear, Double RNA-binding domain protein 2	TcCLB.510755.120	Histone H2B[f]	histone H2B, putative	TcCLB.511635.20
eEF1-gamma[e]	elongation factor 1-gamma (EF-1-gamma, pseudogene), putative	TcCLB.506459.290				ALBA3[f]	ALBA3	TcCLB.510877.30	RNA Hel II[c]	nucleolar RNA helicase II, putative	TcCLB.506123.40
DRBD2[c]	Nuclear, Double RNA-binding domain protein 2	TcCLB.510755.120				**NTF2L**[c]	Cytoplasmic, RNA binding, N-terminal NTF2 domain	TcCLB.511367.220	PUF2[c]	pumilio/PUF RNA binding protein 2, putative	TcCLB.511261.120
Ran[c]	GTP-binding nuclear protein rtb2, putative	TcCLB.509455.80				SUB2[c]	Nuclear, ATP-dependent RNA helicase SUB2	TcCLB.508319.40	eiF4AIII[c]	Nuclear/Nucleolar ATP-dependent RNA helicase FAL1	TcCLB.506587.40

TABLE 2—*Continued*

TcCLB.506435.150 isolation			API5 isolation			FOP isolation			NTF2L isolation		
Designation[a]	Location, annotation[b]	Accession	Designation[a]	Location, annotation[b]	Accession	Designation[a]	Location, annotation[b]	Acession	Designation[a]	Location, annotation[b]	Acession
RRM1[c]	RNA-binding protein, putative	TcCLB.511621.50				*PABP2*[c]	Cytoplasmic (granules), polyadenylate-binding protein	TcCLB.508461.140	RRM1[c]	RNA-binding protein, putative	TcCLB.511621.50
Hel67/ DDX3/ DBP1[e]	ATP-dependent RNA helicase HEL67	TcCLB.511285.120				NOP5[g]	Nucleolar, ribosomal biogenesis	TcCLB.506189.10	DED1[c]	ATP-dependent RNA helicase, putative	TcCLB.510661.90
eEF1-alfa[e]	elongation factor 1-alpha, putative	TcCLB.510119.20				RRM1[c]	RNA-binding protein, putative	TcCLB.511621.50	Histone H2A[f]	histone H2A, putative (fragment)	TcCLB.508321.11
Histone H4[f]	histone H4, putative	TcCLB.508203.29				**API5**[c]	Nuclear, mRNA binding, ARM repeats	TcCLB.511807.280	PBP1[e]	PAB1-binding protein	TcCLB.511409.10
Histone H3[f]	Nuclear, histone H3	TcCLB.505931.50				kZFP2[f]	poly-zinc finger protein 2, putative	TcCLB.509731.10			
Histone H2B[f]	histone H2B, putative	TcCLB.511635.20									
SCD6[e]	Cytoplasmic, LSm domain, Trailer hitch protein	TcCLB.507093.300									
DUF 1126[c]	Repeat of unknown function (DUF1126)/EF-hand domain pair, putative	TcCLB.510797.30									
NUP59[c]	Nucleoporin NUP59	TcCLB.506301.30									

[a]Protein name according to conserved and specific domain identified using INTERPRO, HMMER, PFAM, or ortholog in NCBI nr database. Bold represents kinetoplastid-specific granules, and italic represents nuclear peripheral granules.

[b]Location based on TrypTag, mass spectrometry, or direct experimental evidence. Restricted to either *T. cruzi* or *T. brucei* evidence, except in case of highly conserved ortholog, where evidence from animal or fungal cells is accepted.

[c]mRNA, underline trans-splicing factor.

[d]rRNA/nucleolar.

[e]Translation.

[f]Chromatin.

[g]sn/snoRNA.

FACSAria II (BD) in the flow cytometry facility at Carlos Chagas Institute–Fiocruz/PR. GFP-positive parasites were cloned using single cell precision mode into 96-well plate containing 100 μl of LIT medium/well and the selection antibiotic.

The endogenous and tagged proteins were localized by indirect immunofluorescence assays, as described by Serpeloni *et al.* (2011) (48). The parasites were incubated with specific antibodies for 1 h at 37 °C and then washed with PBS and incubated with Alexa Fluor 488-conjugated goat anti-mouse IgG or Alexa Fluor 594-conjugated goat anti-rabbit IgG antibodies (Invitrogen, 1:600 dilution), as appropriate, for 1 h. DNA was stained by incubation with 5 μg/ml DAPI for 15 min. Slides were analyzed by fluorescence microscopy (Nikon 80i), and images were captured with a CoolSnap PROcf (Media Cybernetics) camera and analyzed with Image Pro-Plus v. 4.5.1.22 (Media Cybernetics). Images were also obtained by inverted microscopy (Leica DMI6000B) associated with deconvolution software Leica AF6000 (microscope facility RPT07C PDTIS/Carlos Chagas Institute–Fiocruz Paraná).

RNA Interference

The coding sequences of orthologs of TcFOP (TbFOP, Gene ID Tb927.6.1470), TcAPI5 (TbAPI5, Gene ID Tb927.7.2240), TcNTF2L (TbNTF2L, Gene ID Tb927.10.2240), and TcHYP (TbHYP, Gene ID Tb927.04.3060) from *T. brucei* were analyzed by RNAit (86) to identify specific target nucleotide sequences for RNA interference. The oligonucleotides predicted by RNAit were used to amplify a DNA fragment (supplemental Table S3) for cloning into the p2T7-177 vector (87). A total of 10 μg insert-containing vector was linearized with NotI and transfected into procyclic forms of *T. brucei* 29-13. Transfected parasites were selected by the addition of 5 μg/ml phleomycin to the medium, and RNAi was induced by adding 2 μg/ml tetracycline to log phase parasite cultures.

mRNA in Situ *Hybridization*

To analyze the distribution of poly(A)$^+$ RNA in *T. brucei*, a digoxigenin-conjugated oligo(dT) probe was used as previously protocol described in Inoue *et al.* (2014) (60). Probe binding was detected by indirect immunofluorescence analysis, as described above, with mouse monoclonal anti-digoxigenin antibody (Sigma-Aldrich, 1:300 dilution) and Alexa Fluor 488-conjugated goat anti-mouse IgG secondary antibody (Invitrogen, 1:600 dilution).

Quantitative PCR Analysis

T. brucei total RNA was extracted using TRIzol Reagent (Invitrogen). To remove residual DNA, the DNase treatment was performed with RQ1 (RNA Qualified) RNase-Free DNase (Promega) according to the manufacturer's instruction. RNA quantification was performed with NanoDrop One Spectrophotometer (Thermo Fisher Scientific), and 1 μg of RNA was run on a formaldehyde gel to check the integrity of RNA. To prepare templates for quantitative PCR, cDNA was synthesized from 1 μg of RNA by ImProm-II Reverse Transcription System (Promega), using random primers (Invitrogen. Real-time PCR was performed in triplicate with the SYBERSelect Master Mix (Applied Biosystems) on LightCycle 96 system (Roche).

qPCR analysis for TbAPI5, TbFOP, TbNTF2L, and TbHYP was normalized using the data obtained by amplification of 7SL RNA. supplemental Table S3 lists the primers used for qPCR, PCR conditions were as follows: preincubation at 95 °C for 10 min, followed by 65 cycles of denaturation at 95 °C for 20 s, annealing at 55 °C for 20 s, and extension at 72 °C for 20 s, one cycle of melting at 95 °C for 60 s, 40 °C for 60 s, 65 °C for 1 s, and 97 °C for 1 s. Expression levels were calculated by Pfaffl's method (88).

Transcriptome Analysis

The TbSub2 RNAi lineage, previously obtained in Serpeloni *et al.* (2011) (48), was used to isolate RNA samples. The induction of RNAi was performed in triplicate by adding 2 μg/ml tetracycline to log phase parasite culture. Total RNA from 1×10^9 cells was isolated using RNeasy Mini Kit (Qiagen) at 24- and 48-h postinduction. RNA was treated with DNAse1 on-column using a RNase-Free DNase Set (Qiagen) following the manufacturer protocol. The integrity was determined by nanoelectrophoresis using an Agilent 2100 Bioanalyzer (Agilent) and quantified by spectrophotometry using a NanoDrop One Spectrophotometer (Thermo Scientific). Approximately 5 μg RNA was sent to BGI services (China) for transcriptome sequencing using Illumina HiSeq 2000 technology (Agilent). The transcriptome data have been submitted to SRA at NCBI with the accession number SUB10190938. We used https://www.bioinformatics.babraham.ac.uk/projects/fastqc and MultiQC for quality control, and analysis was done with Snakemake script. We performed differential expression between each pair of conditions, using edgeR (specifically, exactTest), selecting only coding sequences with at least ten counts in at least one condition (mean across replicates). We used sleuth for differential expression of Salmon data, which simulates technical replicates and improves accuracy. There are 3736 transcripts "differentially expressed" across all three conditions (that is changed in at least one condition) out of 8485 total transcripts analyzed.

RESULTS

TcSub2 and TcIF4AIII Share Kinetoplastid-Specific Protein Interactions

To identify components of the *T. cruzi* mRNA maturation machinery, we used *T. cruzi* Dm28c ectopically expressing N-terminal GFP fusions of TcSub2 (48) and TcIF4AIII (60) as bait. We used cryomilling in liquid nitrogen to preserve dynamic and low-affinity protein–protein interactions (64). Complexes were affinity-captured using anti-GFP nanobodies conjugated to magnetic beads (63). Cell grindates were directly thawed into buffers to determine optimum conditions for preservation and isolation of complexes. Following analysis by SDS-PAGE (Fig. 1*A*), where we confirmed the enrichment of the tagged-proteins, TcSub2 and TcIF4AIII isolates were analyzed by mass spectrometry and protein identifications filtered according to exponentially modified protein abundance index (emPAI). Proteins from three independent experiments with emPAI >0.1, and identified in at least two of the three replicates, are listed in the supplemental material (supplemental Table S4). In addition to proteins known to be involved with RNA metabolism, many additional proteins were recovered among the significant identifications (Table 1).

There was considerable coincidence in the proteins interacting with TcSub2 and TcIF4AIII, together with specific interactions (Table 1). The coincident identifications suggest that TcSub2 and TcIF4AIII associate as a supercomplex, further supported by copurification of TcSub2 and TcIF4AIII. The TcSub2/TcIF4AIII cohort includes orthologs of PRP19-like, Dbp1, Alba1-3, and orthologs of DRBD2 and 3, all of which have defined functions in mRNA metabolism (89–93). Significantly, components of TREX or TREX-2 were not

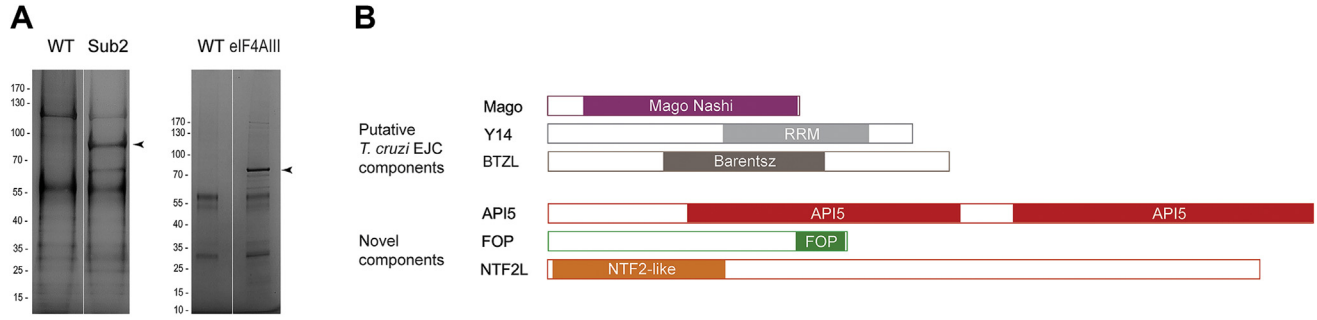


FIG. 1. **Identification of proteins copurified with TcSub2 and TceIF4AIII.** *A*, affinity purified TcSub2 (Sub2) and TceIF4AIII complexes were analyzed on 4 to 12% NuPAGE Novex Bis-Tris precast gel (Life technologies) and stained with SilverQuest Silver Stain (Life technologies). As a negative control, cryomilled cell extracts from wild-type (WT) parasites were treated in the same conditions. *Arrowheads* indicate the bands corresponding to TcSub2 and TceIF4AIII. *B*, schematic representation of selected proteins identified by mass spectrometry: Trypanosome EJC proteins (TcMago, TcY14 and TcBTZL) and three novel proteins (TcAPI5, TcFOP, and TcNTF2L). The motifs and domains were predicted using INTERPRO, HMMER, and PFAM databases and are shown in colors.

identified, in accordance with their apparent absence from trypanosomatid genomes (46). However, proteins associated with chromatin and splicing were identified, including TSR1, TSR1IP, U2AF26, and SR34 (94–97), which interact with quality control (QC) components and the NPC.

In addition, proteins involved in ribosome biogenesis, SDA1 (98) and NOP5 (99), as well as RNA granule associators, SCD6 (22, 100) and TcDHH1 (26, 101), were also detected, indicating an interplay of transcription/mRNA export with other RNA metabolism pathways. The domain architectures of several hypothetical proteins were predicted and include representatives containing an FOP domain (TcFOP), a dual API5 domain (TcAPI5), and an NTF2-like domain (TcNTF2L), while others were identified as core components of the EJC (Fig. 1*B*). Phylogenetic reconstructions demonstrated that these proteins are restricted to the trypanosomatids and eubodonids and hence are likely kinetoplastid-specific mRNA export factors (Fig. 2 and supplemental Fig. S1). Among them, only NTF2L has been already described in trypanosomes as a component of mRNA export, both associated to EJC complex (44) and to nuclear peripherical granules (27).

For TcNTF2L, while the N-terminal NTF2-like domain is broadly conserved, the overall architecture is clearly kinetoplastid-specific. Other proteins correspond to the human Ran GTPase-activating protein-binding protein 1 and 2 (Q13283 and Q9UN86) and contain an RNA recognition motif (RRM) in the C-terminal region absent from the kinetoplastid NTF2L proteins (supplemental Table S2, sheet 3). Phylogenetic reconstructions for NTF2L (Fig. 2) recovered two major clades, one containing only kinetoplastid NTF2L proteins and the other all other sequences, consistent with unique pathways for mRNA export (102).

To provide orthogonal evidence for the identified interactions between TcAPI5, TcFOP, and TcNTF2L with TcSub2 and TceIF4AIII, we raised antibodies to TcAPI5, TcFOP, and TcNTF2L. Western blots confirmed TcAPI5 and TcFOP as present in the TcSub2 isolates, while TcNTF2L was present in TceIF4AIII pullouts, confirming the assignments (supplemental Fig. S2).

EJC Proteins Interact With the T. cruzi *mRNA Export Pathway*

We previously described the *T. cruzi* ortholog of eIF4AIII (60), and phylogenetic tree indicated that TceIF4AIII grouped into the eIF4AIII clade (supplemental Fig. S3). eIF4AIII is a core component of the metazoan EJC and functions in multiple pathways of mRNA metabolism (41), while the *S. cerevisiae* ortholog, Fal1, is involved in 40S-ribosomal-subunit biogenesis (103). Additional proteins related to the mammalian EJC core were identified here; TcMago is represented in the TceIF4AIII/TcSub2 cohort, while TcY14 was identified only in TcSub2 isolations. A protein containing a Barentsz/BTZ domain, named as TcBarentsz-like (TcBTZL), was also present among the TcSub2 interactors.

To confirm the interaction between TcMago and TcSub2/TceIF4AIII, we created *T. cruzi* parasites expressing TcMago::GFP. The fusion protein localized mainly in the nucleus (Fig. 3*A*), and the *T. brucei* ortholog tagged with HA or GFP displayed a nuclear localization (supplemental Fig. S4), as described in Tryp Tag database (http://tryptag.org). TcMago localizes with TcSub2 in the nucleus (Fig. 3, *A* and *B*), and Western blot analysis of TcMago::GFP immunoprecipitations (Fig. 3*D*) identified both TcSub2 and TceIF4AIII (Fig. 3*E*), robustly confirming that conserved core EJC proteins interact with components of the trypanosome mRNA export pathway.

Kinetoplastid-Specific Proteins Are Associated With mRNA Export Pathways

We selected four proteins identified here and which are restricted to kinetoplastids for further characterization. TcFOP, TcAPI5, and TcHYP are also localized to the nucleus, like TcSub2 (Fig. 4, *A*and *B* and supplemental Fig. S7*A*) while TcNTF2L is cytoplasmic, like TceIF4AIII (Fig. 4*C*). To

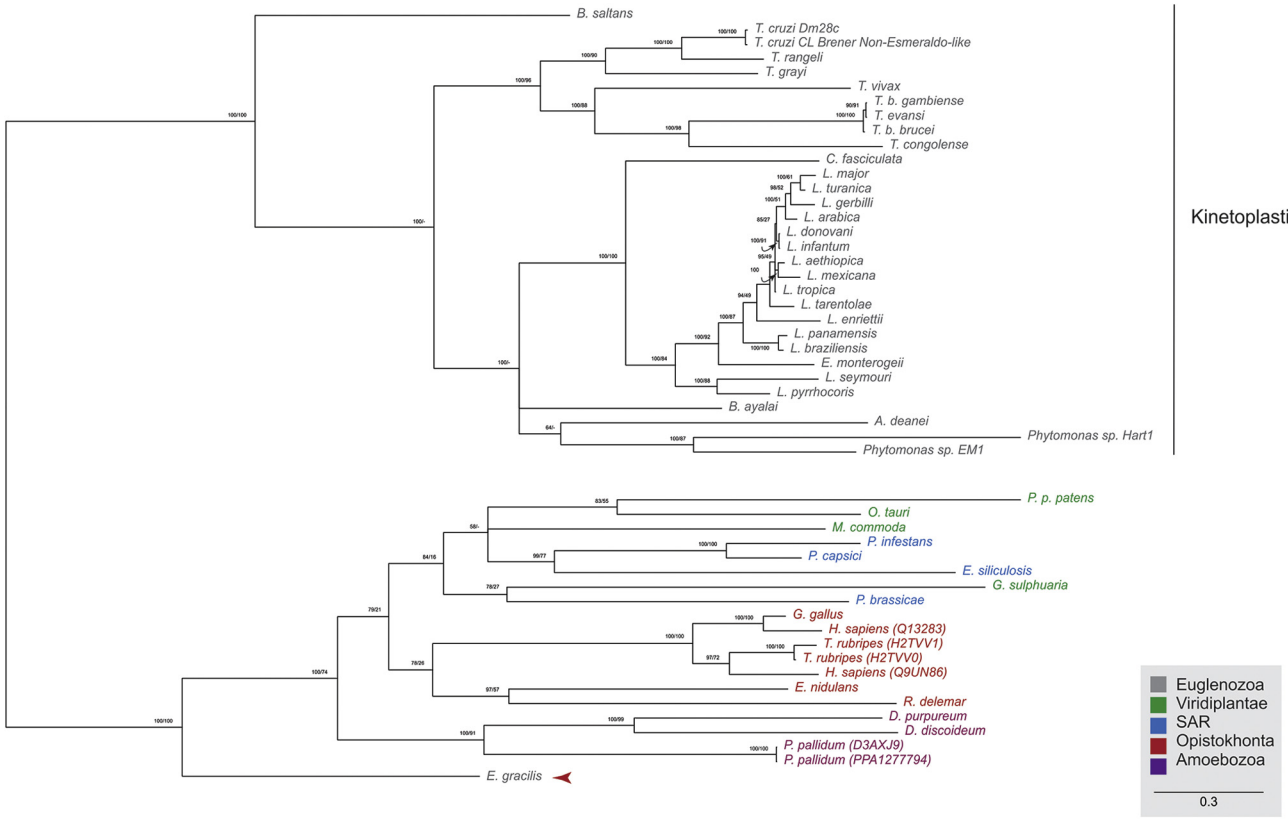


FIG. 2. **TcNTF2L orthologs were found in trypanosomatids and Eubodonida.** A list of TcNTF2L homologs and the organisms searched are shown in supplemental Table S2. *Red arrow head* indicates Euglena, which is outside of the Kinetoplastoda lineage but a relative. Colors represent homolog proteins from representative species of distinct eukaryotic groups, which grouped into a distinct clade from TcNTF2L. Near the nodes are Bayesian posterior probabilities and ML bootstrap values. Bar indicates the estimated number of changed sites.

investigate the physical interactions of these proteins, we created lines expressing TcFOP, TcAPI5, TcNTF2L, and TcHYP tagged with GFP (supplemental Figs. S5 and S7*A*). We confirmed that TcSub2 is present in immunoprecipitates from both TcFOP::GFP (Fig. 5*B*) and TcAPI5::GFP, while TceIF4AIII is indeed present in TcNTF2L::GFP pullouts (Fig. 6*B*). To ensure that the GFP tag was not affecting TcNTF2L location, we used antibodies to TcNTF2L and observed a similar cytoplasmic localization to TcNTF2L::GFP (supplemental Fig. S5*E*). Altogether, these data also provide robust support for the presence of kinetoplastid-specific proteins TcFOP, TcAPI5, TcNTF2L, and TcHYP in association with conserved components within trypanosomatid mRNA export pathways.

The FOP domain of TcFOP is divergent from canonical domains (supplemental Fig. S6). In metazoa, FOP contains a conserved LDXXLDAYM UAP56-binding motif (UBM) that is duplicated in the C-terminal region (X = any amino acid) (104, 105), but in the trypanosomatid FOP, the UBMs are separated by an RGG motif. We asked if these are functional for binding TcSub2, the trypanosome ortholog of the binding partner UAP56. Deletion of the C-terminal UBM from TcFOP (FOPΔC) altered interactions with TcSub2 (Fig. 5*B*), decreasing TcSub2 recovery in pullouts by ~50%, indicating that despite divergence, the FOP domain does mediate interaction with TcSub2 (Fig. 5*C*).

A Superinteractome Connecting Splicing, Translation, and mRNA Export

Mex67, a major mRNA export factor, has been partially characterized in *T. brucei* (49–51). Using *T. cruzi* cells expressing TcMex67 tagged with PTP (84) and the same conditions as the TceIF4AIII pullout (Fig. 7*A*), we confirmed interactions with several proteins including Mtr2 and Ran previously identified in *T. brucei* (51, 56) (supplemental Table S5). Although TceIF4AIII itself was not detected, several proteins common to the TceIF4AIII and TcSub2 interactomes were found, including PABP2, RNA helicase Hel67, and DRBD2, together with TcFOP and TcNTF2L (Fig. 7*B*). This further supports the presence of kinetoplastid-specific proteins as core components of the trypanosome mRNA export machinery.

Analysis of the kinetoplastid-specific protein interactions (supplemental Table S6) confirmed the presence of several conserved components of the processing, export, quality control, and translation systems (Table 2). These include TcSub2, ALBA, DRBD2, NOP, Mago, TceIF4AIII, TceEF1a,

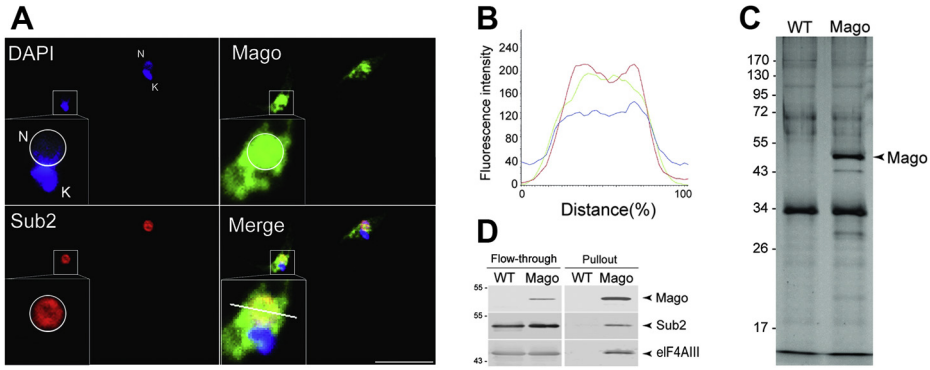


FIG. 3. **Localization and protein interactions of TcMago.** *A*, cellular localization of TcSub2 (Sub2) and TcMago tagged with eGFP (Mago) by confocal microscopy using anti-GFP antibodies. Nucleus is circled in white. DAPI = DNA stained with DAPI. MERGE = merged images. N = nucleus. K = kinetoplast. The kinetoplast DNA usually displays stronger DAPI signal than nucleus. *B*, graphs show the quantification of fluorescence intensity detected across the *dotted line* in merge images. DAPI in *blue*, TcMago::GFP in *green* and TcSub2 in *red*. *C*, TcMago::GFP pullout (Mago) was analyzed in silver-stained SDS–polyacrylamide gel. As a negative control, cryomilled cell extracts from wild-type (WT) parasites were treated in the same conditions. *D*, TcMago::GFP isolates were analyzed by Western blot using the following antibodies: mouse anti-GFP, mouse anti- TceIF4AIII, and rabbit anti-TcSub2 and detection occurred with IRDye whole IgG secondary antibody (LI-COR).

and TcPABP2. Most proteins identified with TcFOP, TcAPI5, TcNTF2L, and TcHYP were also identified as TcSub2 and TceIF4AIII interactors. RNA granule proteins were also identified, including TcDhh1, TcSCD6, and TcPABP2, reinforcing connectivity between mRNA export and the cytoplasmic processing machinery. Further, we identified PUMILIO proteins, RNA-binding proteins involved in controlling gene expression (106–108). The intersection of the interacting proteins recovered with the six isolated proteins was processed using a Venn Diagram tool (https://bioinformatics.psb.ugent.be/cgi-bin/liste/Venn/calculate_venn.htpl), and those proteins recovered with three or more protein complexes plotted (Fig. 8). Overall, these data indicate an unexpected high level of connectivity between multiple aspects of mRNA processing and transport.

Phenotypic Screening by RNAi in T. brucei *of New mRNA Export Factors*

We analyze the roles of TcFOP, TcAPI5, and TcNTF2L by silencing their *T. brucei* orthologs. Quantitative PCR confirmed the efficiency of RNAi against all targets (Fig. 9*A*). Knockdown of TbAPI5 and TbNTF2L affected proliferation, while knockdown of TbFOP had little proliferative impact (Fig. 9*B*). None of the proteins appear involved in global export of mRNAs since

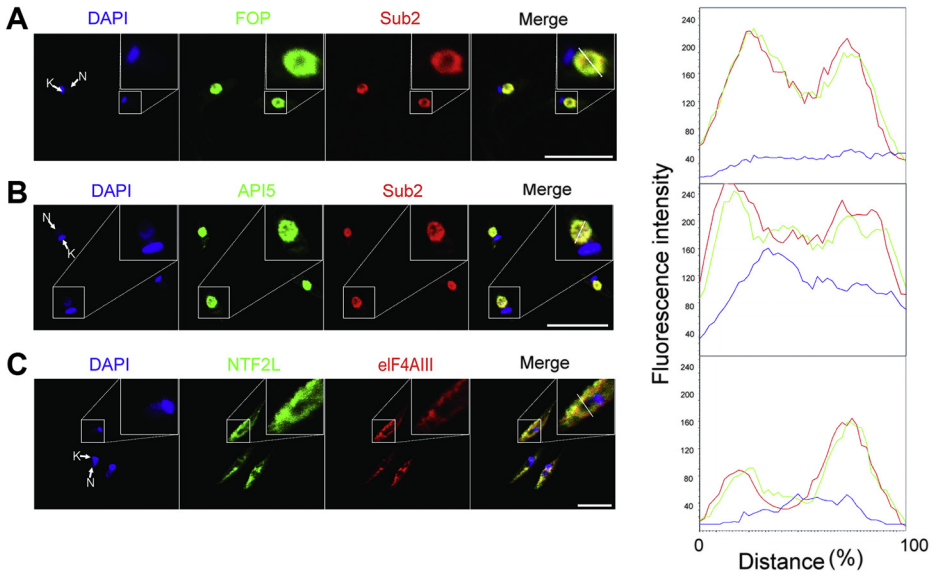


FIG. 4. **Localizations of TcFOP, TcAPI5, and TcNTF2L.** Cellular localization of (*A*) TcSub2 (Sub2) and TcFOP (FOP); (*B*) TcSub2 and TcAPI5 (API5); (*C*) TceIF4AIII (eIF4AIII) and TcNTF2L::GFP (NTF2L). DAPI = DNA stained with DAPI. MERGE = merged images. N = nucleus. K = kinetoplast. Bar = 10 μm (*A* and *B*) and 7.5 μm (*C*). At *right*, graphs show the quantification of fluorescence intensity detected across the dotted line in merge images.

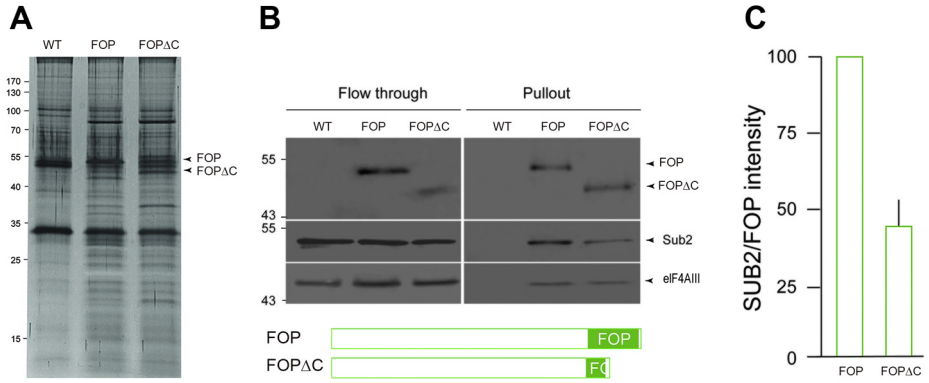


FIG. 5. **Interactions of TcFOP depend on the FOP domain.** *A*, TcFOP (FOP) and mutated TcFOP (FOPΔC) tagged with GFP were pulled out and the copurified proteins were analyzed on silver-stained SDS-polyacrylamide gels. As a negative control, cryomilled cell extracts from wild-type (WT) parasites were treated in the same conditions. *B*, the interaction with TcSub2 (Sub2) and TceIF4AIII was analyzed by Western blot. Isolates were analyzed by Western blot using the following antibodies: mouse anti-TceIF4AIII, rabbit anti-TcSub2, or mouse anti-TcFOP and detection were carried out with IRDye whole IgG secondary antibody (LI-COR). *C*, relative quantification by densitometry of TcSub2 signal, normalized against FOP and ΔFOP signal. Data are expressed as means and standard deviation of three independent experiments.

accumulation of mRNA in the nucleus by polyA FISH was not seen (Fig. 9*C*). Two different controls (without probe and RNAse treatment) confirmed the specificity of probe signal (supplemental Fig. S8). Additionally, we also analyzed TbHYP knockdown and obtained a similar phenotype to TbAPI5, TbFOP, and TbNTF2L with no obvious inhibition of global mRNA export (supplemental Fig. S7*G*). Hence, these new proteins may be mediating the processing of a subset of RNA molecules or that their functions are redundant.

Sub2 Knockdown Leads to Restricted Impact on the Transcriptome

TbSub2 is the sole protein analyzed whose absence clearly blocked mRNA export (48). The impact of TbSub2 knockdown on the transcriptome was restricted and given a supposed central role in mRNA maturation, surprising. Only seven genes demonstrated highly significant increases in abundance (Fig. 10), of which three correspond to 28S and 16S rRNAs (Fig. 11) and the remaining four are protein-encoding. The top protein hits Tb927.7.6960, Tb927.5.680, and Tb927.4.910, with the fourth (Tb927.9.7290) corresponding to a VSG pseudogene. Tb927.7.6960 is an E2 ubiquitin ligase with a Ubc2 domain, while Tb927.5.680 and Tb927.4.910 are hypothetical proteins, both kinetoplast-restricted; based on localization at TrypTag Tb927.5.680 is likely endolysosomal. Similarly, few genes were found to have decreased associated transcript levels, and there was no obvious functional or other feature uniting the cohort (supplemental Table S7). The possibility that TbSub2 is involved in rRNA processing is intriguing as it is associated with mRNA export and quality control in

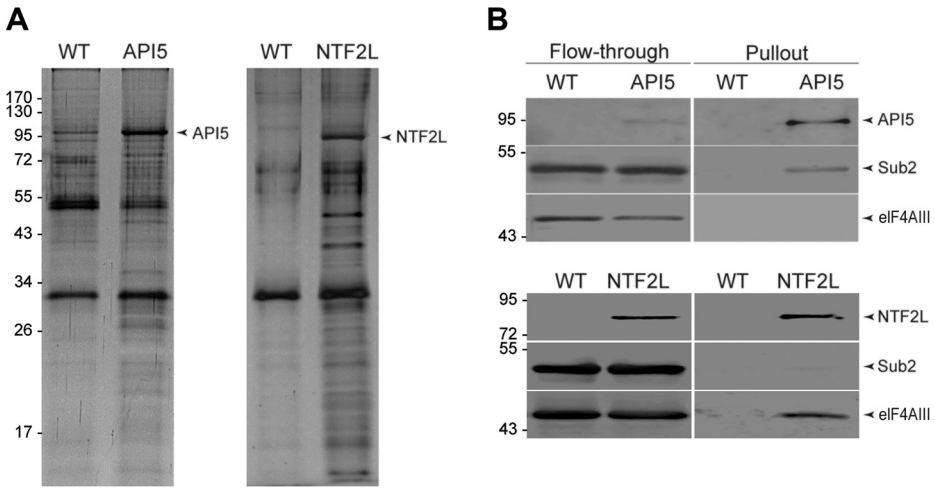


FIG. 6. **Protein interaction of TcAPI5 and TcNTF2L.** *A*, copurified proteins with TcAPI5 (API5) and TcNTF2L (NTF2L) were analyzed on silver-stained SDS–polyacrylamide gels. As a negative control, cryomilled cell extracts from wild-type (WT) parasites were treated in the same conditions. *B*, TcAPI5 and TcNTF2L isolates were analyzed by Western blot using the following antibodies: mouse anti-TceIF4AIII, rabbit anti-TcSub2, mouse anti-TcAPI5, or mouse anti-TcNTF2L. Detection was carried out with IRDye whole IgG secondary antibody (LI-COR).

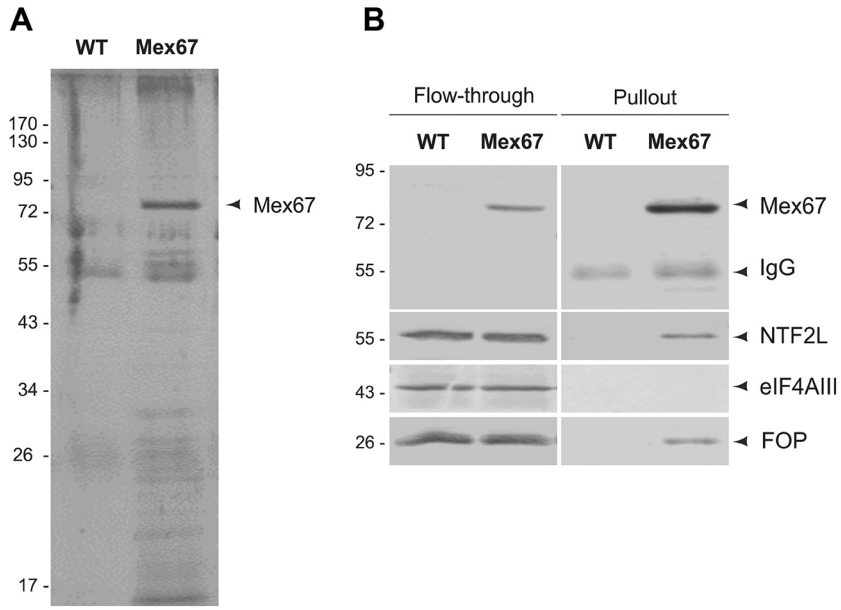


FIG. 7. **Interactions of TcMex67.** *A*, proteins copurified with TcMex67 (Mex67) were analyzed on silver-stained SDS–polyacrylamide gels. As a negative control, cryomilled cell extracts from wild-type (WT) parasites were treated in the same conditions. *B*, TcMex67 isolates were analyzed using the following antibodies: rabbit anti-Protein A, mouse anti-TcNTF2L, mouse anti-TceIF4AIII, mouse anti-TcFOP and detection were carried out with IRDye whole IgG secondary antibody (LI-COR).

other organisms, but the presence of several nucleolar proteins in our interactome would be consistent with an expanded role or close associations between proteins of the mRNA and rRNA export pathways.

DISCUSSION

Messenger RNA maturation requires successful completion of multiple processes, including splicing, polyadenylation, 5′-cap modification, and export to the cytoplasm prior to translation. These steps ensure that only mRNA molecules encoding native proteins are translated, preventing accumulation of unfolded proteins and also supporting differential expression (12–15). In animals and fungi, TREX/TREX-2, the EJC, and the integrator complex (INT) are known players in mRNA processing (109, 110). NPC components, both at the nuclear basket and the cytoplasmic face, together with the Mex67/Mtr2 transport receptor and a vast repertoire of ATP-dependent RNA helicases are critically important. Many factors are widely conserved, but distinct features between lineages are clear. In *Arabidopsis*, several INT complex proteins have divergent architectures compared with their animal orthologs (111) and multiple differences between animals and fungi are known (109, 110, 112). Trypanosomes exhibit even greater divergence as orthologs of many mRNA processing proteins are undetectable *in silico*.

Established differences in mRNA processing between trypanosomatids and other eukaryotes include cotranscriptional nuclear export and polycistronic transcription, placing additional burdens on posttranscriptional systems. Additionally, trypanosomes have two PABP paralogs with distinct functions (21), three Mex67 paralogs (50) and major mRNA processing activity within nuclear periphery granules (NPGs) proximal to the NPC cytoplasmic face (26, 27). Deferment of mRNA QC, at least in part to the cytoplasm, is in agreement with the absence of several proteins involved in animal/yeast mRNA

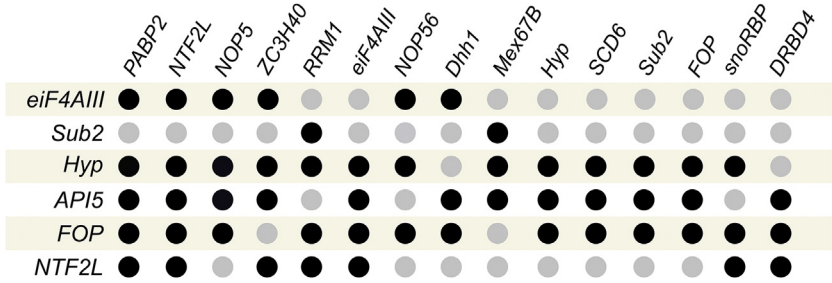


FIG. 8. **Summary of coincident interacting proteins with three or more isolated proteins.** The six isolated proteins are represented in *lines* and interacting proteins are in columns. *Black circles* indicate the presence of the protein in the isolated complex.

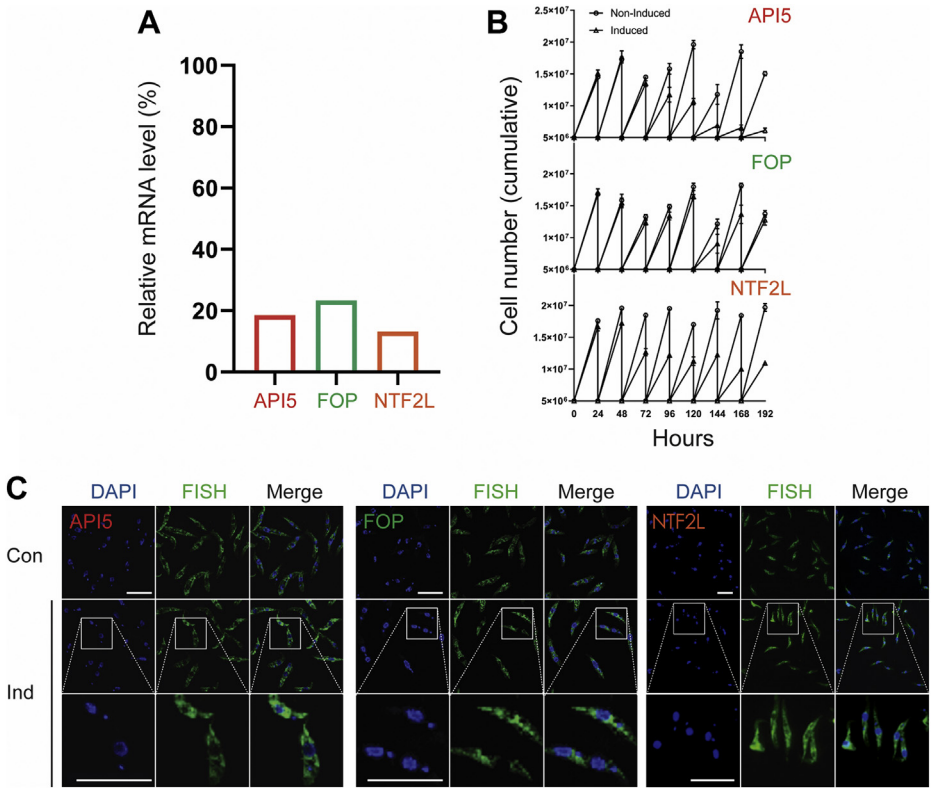


FIG. 9. **TbFOP, TbAPI5, and TbNTF2L are nonessential proteins.** *A*, real-time PCR analysis using noninduced (N) and induced parasites (I) for RNAi of TbFOP (FOP), TbAPI5 (API5), and TbNTF2L (NTF2L). The relative mRNA levels were quantified at 72 h postinduction and normalized by detection of 7SL RNA. *B*, growth curve of cell lines after RNAi induction for TbFOP, TbAPI5 and TbNTF2L. The density of cells in the culture was determined by triplicate counting in a particle counter (Beckman Coulter, Z Series Coulter Counter Cell). Both noninduced and induced cultures were diluted daily to 5×10^5 cells/ml for all RNAi inductions. *C*, cellular localization of mRNA by fluorescence *in situ* hybridization (FISH) in noninduced (Con) and induced (Ind) parasites. Digoxigenin-conjugated oligo(dT) was used as probe. DAPI = DNA stained with DAPI. N = nucleus. K = kinetoplast. Bar = 10 μm.

maturation (27, 56). Remarkably, using TcSub2 and TceIF4AIII as affinity handles, we uncovered an interactome encompassing chromatin-associated proteins through to NPG components and several Zn^{2+}-finger RNA-binding proteins (Tables 1 and 2, Fig. 12). One-third of identifications were coincident between both TcSub2 and TceIF4AIII, and neither PABP1 nor any PABP1 interactor was found, indicating our data are robust and specific (21). Identification of components across the entire journey of mRNA processing suggests a highly integrated system.

There are several functional cohorts within this TcSub2/eIF4AIII interactome, including proteins associated with chromatin, the *trans*-splicing machinery, and the EJC. Only a single Mex67 paralog (Mex67B) was recovered, suggesting specificity between mRNA export interactors. With TcMex67 as the affinity handle, we identified Ran, Mtr2 and PABP2 as expected, in addition to the kinetoplastid-specific proteins TcFOP and TcNTF2L, themselves also within the TcSub2/eIF4AIII interactome. A further cohort, preferentially identified with TceIF4AIII is likely involved in rRNA processing and/or ribosomal biogenesis and suggests a close connection between mRNA and rRNA pathways. Significantly, neither TcSub2 nor TceIF4AIII retrieved proteins that would constitute a canonical TREX or EJC.

In metazoa, the EJC is deposited upstream of exon–exon boundaries during *cis*-splicing (39) and contributes to multiple activities including NMD (113, 114), export (43, 114–116), and translation (117). Trypanosomes have both components of the EJC and NMD-related pathway (44, 118) and associations between the EJC (Mago, eIF4AIII) and Sub2 suggest a platform linking splicing and export (29, 119). This is further supported by identification of TSR1, TSR1IP, SR34, and additional splicing machinery proteins. Another SR protein, RRM1, is associated with the core spliceosome and modulates chromatin structure (120). Interestingly, we also observed RBSR2 that is associated with RRM1, RBSR1, and other RNA-binding proteins that were also identified, such as UBP-1, UBP-2, and ALBA3 (121). Additionally, DRBD18, identified in TcSub2 and TcMex67 isolates, is involved in mRNA export in association with Mex67/Mtr2 by regulating the mobilization of mRNPs through the nuclear pore complex (122).

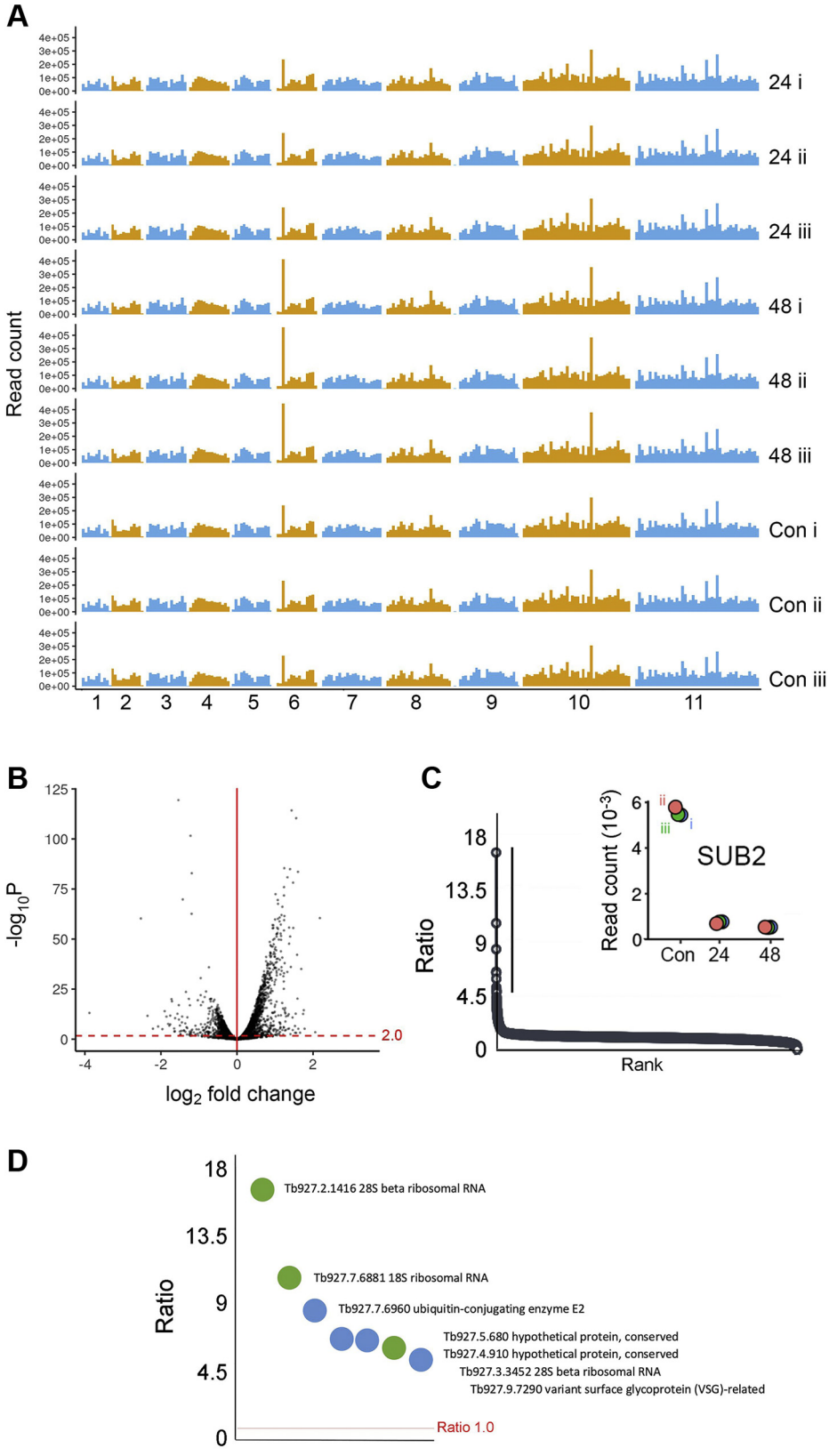


FIG. 10. **TbSub2 is involved in a small number of RNA processing events.** *A*, Manhattan plots for all RBNAseq datasets mapped against the *T. brucei* TREU927 genome. Reads are indicated as the *y-axis* and chromosome numbers as the *x-axis*. Note that overall, there is little major variation in the overall transcriptional pattern. *B*, Volcano plot of the same dataset. *Solid red line* indicates a log fold change of zero, and the

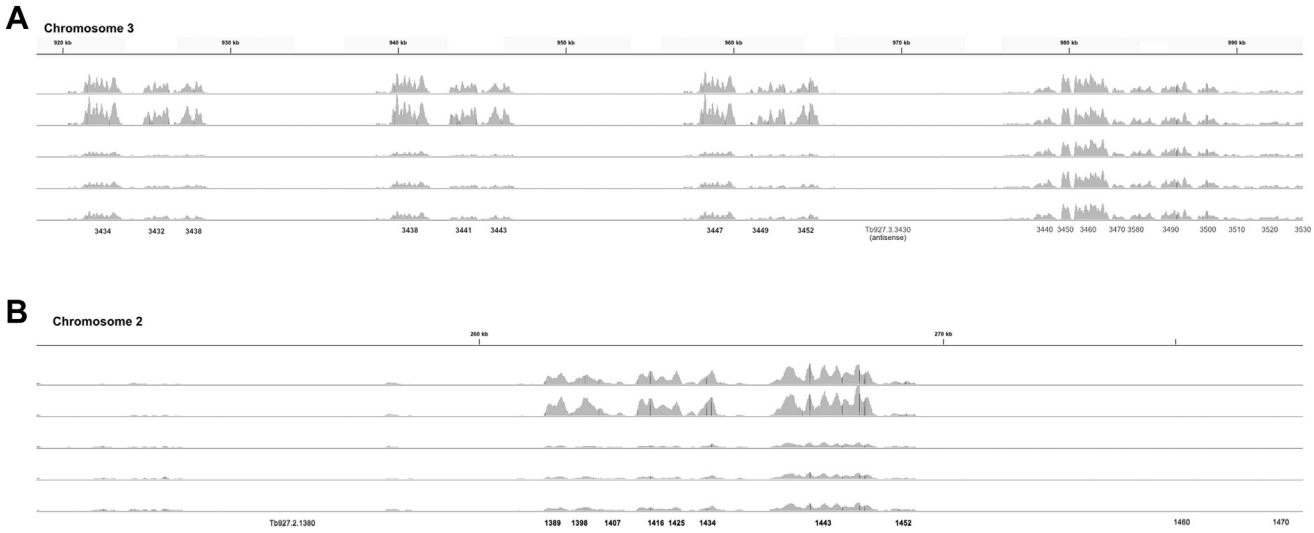


FIG. 11. **28S and 16S RNA levels dramatically increase on Sub2 silencing.** Screengrabs from IG Viewer showing two regions of the 28S rRNA loci reads mapped to the *T. brucei* TREU927 genome. Top two systems represent induced samples in each panel, and lower three are unindexed control cells. Genes in *bold* are rRNA genes, and those in regular font are protein coding genes. Note that the protein coding genes are unchanged in terms of reads mapped. *A,* region of chromosome 3. *B,* region of chromosome 2. Annotation taken rom TritrypDB.

TcFOP, TcAPI5, TcNTF2L, and TcHYP possess lineage-specific features, indicating mechanistic divergence (123). TcFOP contains an FOP-like domain initially described in mammalian CHTOP, a TREX component (105). CHTOP binds Sub2 and activates mRNA delivery to Mex67 (104). The trypanosome FOP domain is distinct from metazoan FOP but is required for interaction with TcSub2. TcFOP interacts with TcMex67, supporting a role in mRNA export. However, knockdown of TbFOP did not drastically affect mRNA export, although this is also noted for CHTOP protein, where mRNA export is only affected by a dual knockdown of CHTOP and partner ALYREF (104). Since ALYREF orthologs are absent from trypanosome genomes (28) and was not identified here by proteomics, TcFOP may have both divergent architecture and distinct mechanisms in trypanosomes. Knockdown of API5 also did not significantly impair mRNA export, despite interacting with mRNA export factors. Moreover, API5 interacts directly with Sub2 in mammalian cells and depletion results in accumulation of poly(A)$^+$ RNA in the nucleus, which was not found here (119). TcNTF2L has a domain common to several transport receptors, including NTF2 (124) and Mex67 (125). Although TcNTF2L RNAi did not cause clear nuclear mRNA accumulation, this result cannot exclude the involvement of TcNTF2L in mRNA export pathway since the *T. brucei* ortholog copurifies with NPG proteins (26, 27). Furthermore, TbNTF2L is a chromatin-associated protein (126) and copurifies with EJC components (44). These results and the very mild impact on mRNA abundance from Sub2 silencing may indicate a level of redundancy and/or additional mechanisms remaining to be uncovered.

Cytoplasmic proteins PABP2, DHH1, and multiple Zn-finger RNA helicases that associate with NPGs were identified (26). PABP2 regulates bulk mRNA directed for translation (21) while DHH1 is localized in cytoplasmic granules (23, 24, 101). Although DHH1 removes polysome mRNAs and interacts with decapping and deadenylase complexes in yeast (127), DHH1 acts only as a translational repressor in trypanosomes (128). In yeast, DHH1 is associated with SCD6, a decapping activator that binds eIF4G and inhibits translation initiation (100, 129). Unlike yeast, SCD6 in *T. brucei* does not form a complex with DHH1 (100), even though it also represses translation and significantly we identified TcSCD6 with both TcSub2 and TceIF4AIII.

In conclusion, we report both novel and conserved components of the mRNA export pathway in trypanosomes identified through immunoisolation and proteomics (Fig. 12). Among many interactions identified, we highlight the kinetoplastid-specific proteins TcFOP, TcAPI5, and TcHYP that interact directly with TcSub2, while TcNTF2L, a shuttling protein, interacts with the export receptor TcMex67. TcNTFL2 and TcFOP are kinetoplastid-specific but connect transcription/splicing components with the conserved export factor Mex67 and other components of mRNA export. We illustrated the comparison between major proteins described in metazoa and those identified in *T. cruzi* (Fig. 13). Overall, we propose that trypanosome mRNA export comprises a core of pan-

dotted red line a significance of 2.0, above which data are considered robust. *C*, ratio of fold change (linear) against rank order; only seven transcripts demonstrate a greater than fivefold increase. Inset data for Sub2, the knockdown target. *D*, plots for the seven most increased transcripts, where *green* dots are rRNA and *blue* correspond to protein-coding transcripts. Identities of the transcripts are indicated, and data are available in full as supplemental Table S7. A ratio of 1.0 indicates no change.

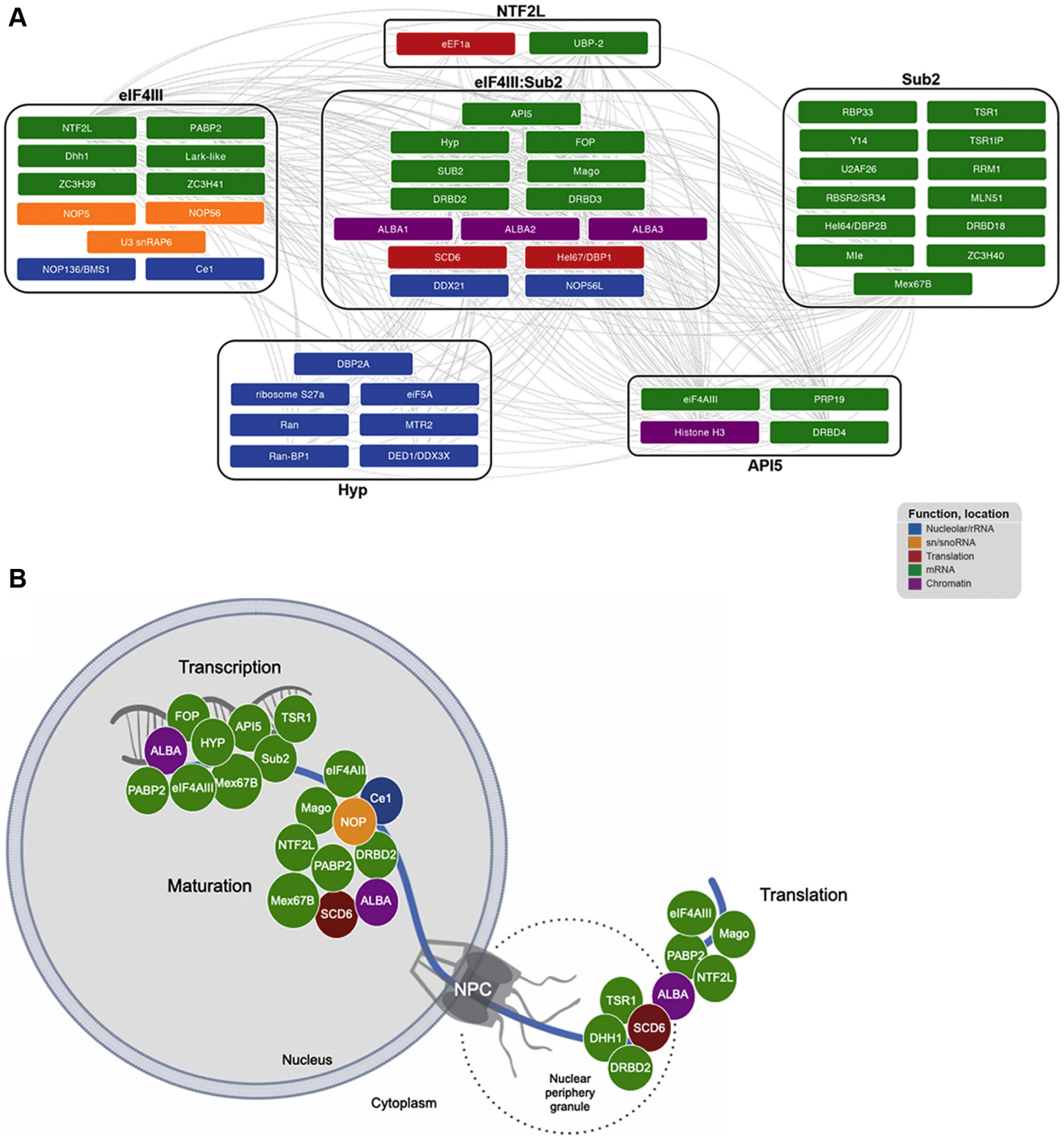


FIG. 12. **Schematic representation of *T. cruzi* mRNA export machinery.** *A*, using a cohort of six affinity handles, we identified a large interactome that illustrates multiple interactions as well as the presence of conserved and divergent proteins. Furthermore, these data validate the absence of many components that have failed to be identified *in silico*, including proteins of the TREX and TREX2 complexes. Cohorts are arranged as identified in the pullouts and colored by predicted function using the scheme as in Tables 1 and 2. *Large boxes* represent a single interactome, with the affinity handle indicated. *Dotted lines* between entries indicate a detected interaction. Many potential interactions and the TcMex67 isolation have been omitted for clarity. Conserved proteins include TcSub2, TceIF4AIII and multiple RNA binding proteins and EJC components. kinetoplastid-specific proteins include TcAPI5, TcFOP and TcNTF2L, and TcFOP and TcAPI5 interact directly with TcSub2. Core components of the EJC are directly associated with the mRNA machinery. TcNTF2L seems to be a shuttling protein that interacts with TceIF4AIII and the export receptor TcMex67. Connectivity data are also available as a cytoscape file for readers who wish to explore further. *B*, model for interactome and interactions with the NPC and export systems. Conserved proteins, such as TcSub2, TceIF4AIII, RNA binding proteins, components of the EJC, and kinetoplastid-specific proteins, such as TcAPI5, TcFOP, and TcNTF2L are shown. The nuclear proteins TcFOP and TcAPI5 interact directly with TcSub2. Core components of EJC are directly associated to mRNA machinery. TcNTF2L seems to be a shuttling protein that interacts with TceIF4AIII and the export receptor TcMex67.

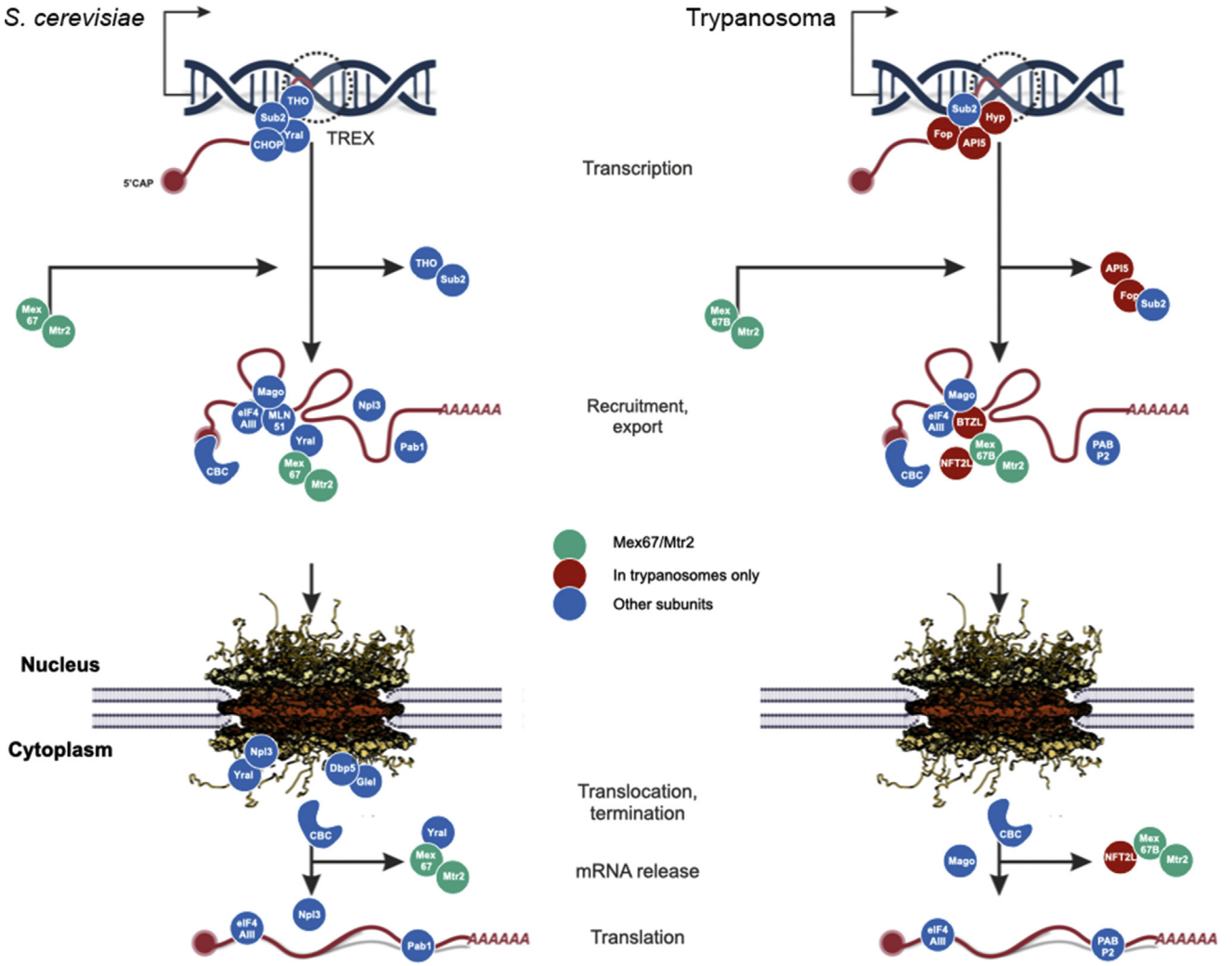


FIG. 13. **Comparision between *S. cerevisae* and *T. cruzi* components of mRNA export coupled to transcription/splicing.** The figure contains the major components of *S. cerevisiae* as described on Kohler and Hurt (2007) (130) and Kramer (2021) (47). To simplify, components of Peripheral Nuclear Granules and Nuclear Pore Complex are omitted. In *T. cruzi*, the specific proteins are in *red*.

eukaryotic proteins together with a significant number of proteins unique to kinetoplastids and a further example of the power of proteomics to uncover evolutionary divergent mechanisms.

DATA AVAILABILITY

All mass spectrometry data have been deposited with the ProteomeXchange Consortium *via* the PRIDE partner repository (https://www.ebi.ac.uk/pride/) with the dataset identifier PXD028072. The TbSub2 RNAi transcriptome data have been submitted to SRA at NCBI with the accession number SUB10190938.

Supplemental data—This article contains supplemental data.

Acknowledgments—We thank the Program for Technological Development in Tools for Health–RPT FIOCRUZ for Technological Platforms of Microscopy, Flow Cytometry and Mass spectrometry, and the FingerPrints proteomics facility at the University of Dundee, supported by the "Wellcome Trust Technology Platform" award [097945/B/11/Z]. We would like to thank Wagner Nagib de Souza Birbeire for assistance in producing Figures 11*B* and 13 and James Abbott for assistance in producing Figure 11*A*.

Funding and additional information—This work was supported by the Wellcome Trust (204697/Z/16/Z to M. C. F.), the Medical Research Council (MR/N010558/1 to M. C. F.), Coordenação De Aperfeiçoamento De Pessoal De Nível Superior (CAPES to A. R. Á., A. H. I., P. F. D., P. M. H. and M. S.), Fundação Oswaldo Cruz (FIOCRUZ to A. R. Á.), and Conselho

Nacional de Desenvolvimento Científico e Tecnológico (421990/2017-1 CNPq to A. R. Á.).

Author contributions—A. H. I., M. C. F., and A. R. Á. conceptualization; M. S., P. M. H., N. M. V., E. R. B., R. C. d. P., and A. L. data curation; M. C. F. and A. R. Á. funding acquisition; A. H. I. and P. F. D. investigation; A. H. I. and P. F. D. methodology; A. R. Á. project administration; P. M. H., R. C. d. P., and C. B. resources; C. B. and M. C. F. supervision; M. S., N. M. V., and A. L. validation; M. C. F. and A. R. Á. visualization; A. H. I. and P. F. D. writing—original draft; M. C. F. and A. R. Á. writing—review and editing.

Conflict of interest—The authors declare no competing interests.

Abbreviations—The abbreviations used are: EJC, exon junction complex; INT, integrator complex; mRNP, messenger ribonucleoprotein; NMD, nonsense-mediated *mRNA* decay; NPC, nuclear pore complex; NPGs, nuclear peripheral granules; TREX, TRanscription-EXport; UTR, mRNA untranslated region.

Received August 30, 2021, and in revised from, January 15, 2022
Published, MCPRO Papers in Press, January 26, 2022, https://doi.org/10.1016/j.mcpro.2022.100208

REFERENCES

1. WHO. (2010) *First WHO Report on Neglected Tropical Diseases: Working to Overcome the Global Impact of Neglected Tropical Diseases*. World Health Organization, Geneva: 1–184
2. De Souza, R. D. C. M., Gorla, D. E., Chame, M., Jaramillo, N., Monroy, C., and Diotaiuti, L. (2021) Chagas disease in the context of the 2030 agenda: Global warming and vectors. *Mem. Inst. Oswaldo Cruz* **116**, 1–14
3. Kramer, S., and Carrington, M. (2011) Trans-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. *Trends Parasitol.* **27**, 23–30
4. Bartholomeu, D. C., de Paiva, R. M. C., Mendes, T. A. O., DaRocha, W. D., and Teixeira, S. M. R. (2014) Unveiling the intracellular survival gene Kit of trypanosomatid parasites. *PLoS Pathog.* **10**, e1004399
5. Kamina, A. D., and Williams, N. (2017) Ribosome assembly in trypanosomatids: A novel therapeutic target. *Trends Parasitol.* **33**, 256–257
6. Begolo, D., Vincent, I. M., Giordani, F., Pöhner, I., Witty, M. J., Rowan, T. G., Bengaly, Z., Gillingwater, K., Freund, Y., Wade, R. C., Barrett, M. P., and Clayton, C. (2018) The trypanocidal benzoxaborole AN7973 inhibits trypanosome mRNA processing. *PLoS Pathog.* **14**, e1007315
7. Gosavi, U., Srivastava, A., Badjatia, N., and Günzl, A. (2020) Rapid block of pre-mRNA splicing by chemical inhibition of analog-sensitive CRK9 in Trypanosoma brucei. *Mol. Microbiol.* **113**, 1225–1239
8. Johnson, P. J., Kooter, J. M., Borst, P., and Van Leeuwenhoekhuis, A. (1987) Inactivation of transcription by UV irradiation of T. brucei provides evidence for a multicistronic transcription unit including a VSG gene. *Cell* **51**, 273–281
9. Agabian, N. (1990) Trans splicing of nuclear pre-mRNAs. *Cell* **61**, 1157–1160
10. Martínez-Calvillo, S., Yan, S., Nguyen, D., Fox, M., Stuart, K., Myler, P. J., and Street, N. (2003) Transcription of Leishmania major friedlin chromosome 1 initiates in both. *Mol. Cell* **11**, 1291–1299
11. Preußer, C., Jaé, N., and Bindereif, A. (2012) mRNA splicing in trypanosomes. *Int. J. Med. Microbiol.* **302**, 221–224
12. Siegel, T. N., Gunasekera, K., Cross, G. A., and Ochsenreiter, T. (2011) Gene expression in Trypanosoma brucei: Lessons from high throughput RNA sequencing studies. *Trends Parasitol.* **27**, 434–441
13. Kramer, S. (2012) Developmental regulation of gene expression in the absence of transcriptional control: The case of kinetoplastids. *Mol. Biochem. Parasitol.* **181**, 61–72
14. Smircich, P., Eastman, G., Bispo, S., Duhagon, M. A., Guerra-slompo, E. P., Garat, B., Goldenberg, S., Munroe, D. J., Dallagiovanna, B., Holetz, F., and Sotelo-silveira, J. R. (2015) Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in Trypanosoma cruzi. *BMC Genomics* **16**, 1–14
15. Clayton, C. E. (2016) Gene expression in kinetoplastids. *Curr. Opin. Microbiol.* **32**, 46–51
16. Kramer, S. (2014) RNA in development: How ribonucleoprotein granules regulate the life cycles of pathogenic protozoa. *Wiley Interdiscip. Rev. RNA* **5**, 263–284
17. Minia, I., Merce, C., Terrao, M., and Clayton, C. (2016) Translation regulation and RNA granule formation after heat shock of procyclic form Trypanosoma brucei: Many heat-induced mRNAs are also increased during differentiation to mammalian-infective forms. *PLoS Negl. Trop. Dis.* **10**, e0004982
18. Klein, C., Terrao, M., and Clayton, C. (2017) The role of the zinc finger protein ZC3H32 in bloodstream-form Trypanosoma brucei. *PLoS One* **12**, e0177901
19. Chakraborty, C., and Clayton, C. (2018) Stress susceptibility in Trypanosoma brucei lacking the RNA-binding protein ZC3H30. *PLoS Negl. Trop. Dis.* **12**, e0006835
20. Romaniuk, M. A., Frasch, A. C., and Cassola, A. (2018) Translational repression by an RNA-binding protein promotes differentiation to infective forms in Trypanosoma cruzi. *PLoS Pathog.* **14**, e1007059
21. Zoltner, M., Krienitz, N., Field, M. C., and Kramer, S. (2018) Comparative proteomics of the two T. brucei PABPs suggests that PABP2 controls bulk mRNA. *PLoS Negl. Trop. Dis.* **12**, e0006679
22. Krüger, T., Hofweber, M., Kramer, S., and Strome, S. (2013) SCD6 induces ribonucleoprotein granule formation in trypanosomes in a translation-independent manner, regulated by its Lsm and RGG domains. *Mol. Biol. Cell* **24**, 2098–2111
23. Holetz, F. B., Alves, L. R., Probst, C. M., Dallagiovanna, B., Marchini, F. K., Manque, P., Buck, G., Krieger, M. A., Correa, A., and Goldenberg, S. (2010) Protein and mRNA content of TcDHH1-containing mRNPs in Trypanosoma cruzi. *FEBS J.* **277**, 3415–3426
24. Alves, L. R., Ávila, A. R., Correa, A., Holetz, F. B., Mansur, F. C. B., Manque, P. A., De Menezes, J. P. B., Buck, G. A., Krieger, M. A., and Goldenberg, S. (2010) Proteomic analysis reveals the dynamic association of proteins with translated mRNAs in Trypanosoma cruzi. *Gene* **452**, 72–78
25. Ferreira, J., Ferrarini, M. G., Nardelli, S. C., Goldenberg, S., Ávila, A. R., and Holetz, F. B. (2018) Trypanosoma cruzi XRNA granules colocalise with distinct mRNP granules at the nuclear periphery. *Mem. Inst. Oswaldo Cruz* **113**, e170531
26. Kramer, S., Marnef, A., Standart, N., and Carrington, M. (2012) Inhibition of mRNA maturation in trypanosomes causes the formation of novel foci at the nuclear periphery containing cytoplasmic regulators of mRNA fate. *J. Cell Sci.* **125**, 2896–2909
27. Goos, C., Dejung, M., Wehman, A. M., M-natus, E., Schmidt, J., Sunter, J., Engstler, M., Butter, F., and Kramer, S. (2019) Trypanosomes can initiate nuclear export co-transcriptionally. *Nucleic Acids Res.* **47**, 266–282
28. Jimeno, S., Rondón, A. G., Luna, R., and Aguilera, A. (2002) The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *EMBO J.* **21**, 3526–3535
29. Strässer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondón, A. G., Aguilera, A., Struhl, K., Reed, R., and Hurt, E. (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417**, 304–308
30. Chi, B., Wang, Q., Wu, G., Tan, M., Wang, L., Shi, M., Chang, X., and Cheng, H. (2013) Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Res.* **41**, 1294–1306
31. Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005) Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* **19**, 1512–1517
32. Jani, D., Lutz, S., Hurt, E., Laskey, R. A., Stewart, M., and Wickramasinghe, V. O. (2012) Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. *Nucleic Acids Res.* **40**, 4562–4573

33. Jani, D., Valkov, E., and Stewart, M. (2014) Structural basis for binding the TREX2 complex to nuclear pores, GAL1 localisation and mRNA export. *Nucleic Acids Res.* **42**, 6686–6697
34. Schubert, T., and Köhler, A. (2016) Mediator and TREX-2: Emerging links between transcription initiation and mRNA export. *Nucleus* **7**, 126–131
35. Wickramasinghe, V. O., and Laskey, R. A. (2015) Control of mammalian gene expression by selective mRNA export. *Nat. Rev. Mol. Cell Biol.* **16**, 431–442
36. Katahira, J., Dimitrova, L., Imai, Y., and Hurt, E. (2015) NTF2-like domain of Tap plays a critical role in cargo mRNA recognition and export. *Nucleic Acids Res.* **43**, 1894–1904
37. Bullock, T. L., Clarkson, W. D., Kent, H. M., and Stewart, M. (1996) The 1.6 Å resolution crystal structure of nuclear transport factor 2 (NTF2). *J. Mol. Biol.* **260**, 422–431
38. Eberhardt, R. Y., Chang, Y., Bateman, A., Murzin, A. G., Axelrod, H. L., Hwang, W. C., and Aravind, L. (2013) Filling out the structural map of the NTF2-like superfamily. *BMC Bioinformatics* **14**, 1–11
39. Hir, H. Le, Izaurralde, E., Maquat, L. E., and Moore, M. J. (2000) The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19**, 6860–6869
40. Gehring, N. H., Lamprinaki, S., Kulozik, A. E., and Hentze, M. W. (2009) Disassembly of exon junction complexes by PYM. *Cell* **137**, 536–548
41. Tange, T.Ø., Shibuya, T., Jurica, M. S., and Moore, M. J. (2005) Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA* **11**, 1869–1883
42. Andreou, A. Z., and Klostermeier, D. (2013) The DEAD-box helicase eIF4A paradigm or the odd one out? *RNA Biol.* **10**, 19–32
43. Hir, H. Le, Gatfield, D., Izaurralde, E., and Moore, M. J. (2001) The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**, 4987–4997
44. Bercovich, N., Levin, M. J., Clayton, C., and Vazquez, M. P. (2009) Identification of core components of the exon junction complex in trypanosomes. *Mol. Biochem. Parasitol.* **166**, 190–193
45. Bannerman, B. P., Kramer, S., Dorrell, R. G., and Carrington, M. (2018) Multispecies reconstructions uncover widespread conservation, and lineage-specific elaborations in eukaryotic mRNA metabolism. *PLoS One* **13**, e0192633
46. Serpeloni, M., Vidal, N. M., Goldenberg, S., Ávila, A. R., and Hoffmann, F. G. (2011) Comparative genomics of proteins involved in RNA nucleo-cytoplasmic export. *BMC Evol. Biol.* **11**, 7
47. Kramer, S. (2021) Nuclear mRNA maturation and mRNA export control: From trypanosomes to opisthokonts. *Parasitology* **148**, 1196–1218
48. Serpeloni, M., Moraes, C. B., Muniz, J. R. C., Motta, M. C. M., Ramos, A. S. P., Kessler, R. L., Inoue, A. H., DaRocha, W. D., Yamada-Ogatta, S. F., Fragoso, S. P., Goldenberg, S., Freitas-Junior, L. H., and Ávila, A. R. (2011) An essential nuclear protein in trypanosomes is a component of mRNA transcription/export pathway. *PLoS One* **6**, e20730
49. Schwede, A., Manful, T., Jha, B. A., Helbig, C., Bercovich, N., Stewart, M., and Clayton, C. (2009) The role of deadenylation in the degradation of unstable mRNAs in trypanosomes. *Nucleic Acids Res.* **37**, 5511–5528
50. Kramer, S., Kimblin, N. C., and Carrington, M. (2010) Genome-wide in silico screen for CCCH-type zinc finger proteins of Trypanosoma brucei, Trypanosoma cruzi and Leishmania major. *BMC Genomics* **11**, 283
51. Dostalova, A., Käser, S., Cristodero, M., and Schimanski, B. (2013) The nuclear mRNA export receptor Mex67-Mtr2 of Trypanosoma brucei contains a unique and essential zinc finger motif. *Mol. Microbiol.* **88**, 728–739
52. Neumann, N., Lundin, D., and Poole, A. M. (2010) Comparative genomic evidence for a complete nuclear pore complex in the last eukaryotic common ancestor. *PLoS One* **5**, e13241
53. Bühlmann, M., Walrad, P., Rico, E., Ivens, A., Capewell, P., Naguleswaran, A., Roditi, I., and Matthews, K. R. (2015) NMD3 regulates both mRNA and rRNA nuclear export in African trypanosomes *via* an XPOI-linked pathway. *Nucleic Acids Res.* **43**, 4491–4504
54. Rink, C., and Williams, N. (2019) Unique interactions of the nuclear export receptors TbMex67 and TbMtr2 with components of the 5S ribonuclear particle in Trypanosoma brucei. *mSphere* **4**, e00471-19
55. Hegedűsová, E., Kulkarni, S., Burgman, B., Alfonzo, J. D., and Paris, Z. (2019) The general mRNA exporters Mex67 and Mtr2 play distinct roles in nuclear export of tRNAs in Trypanosoma brucei. *Nucleic Acids Res.* **47**, 8620–8631
56. Obado, S. O., Brillantes, M., Uryu, K., Zhang, W., Ketaren, N. E., Chait, B. T., Field, M. C., and Rout, M. P. (2016) Interactome mapping reveals the evolutionary history of the nuclear pore complex. *PLoS Biol.* **14**, 1–30
57. DuBois, K. N., Alsford, S., Holden, J. M., Buisson, J., Swiderski, M., Bart, J. M., Ratushny, A. V., Wan, Y., Bastin, P., Barry, J. D., Navarro, M., Horn, D., Aitchison, J. D., Rout, M. P., and Field, M. C. (2012) NUP-1 is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. *PLoS Biol.* **10**, e1001287
58. Akiyoshi, B., and Gull, K. (2014) Discovery of unconventional kinetochores in kinetoplastids. *Cell* **156**, 1247–1258
59. D'Archivio, S., and Wickstead, B. (2017) Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. *J. Cell Biol.* **216**, 379–391
60. Inoue, A. H., Serpeloni, M., Hiraiwa, P. M., Yamada-Ogatta, S. F., Muniz, J. R. C., Motta, M. C. M., Vidal, N. M., Goldenberg, S., and Avila, A. R. (2014) Identification of a novel nucleocytoplasmic shuttling RNA helicase of trypanosomes. *PLoS One* **9**, e109521
61. Contreras, V. T., Salles, J. M., Thomas, N., Morel, C. M., and Goldenberg, S. (1985) *In vitro* differentiation of Trypanosoma cruzi under chemically defined conditions. *Mol. Biochem. Parasitol.* **16**, 315–327
62. Brun, R., and Schonenberger, M. (1979) Cultivation and *in vitro* cloning of procyclic culture forms of Trypanosoma brucei in a semi-defined medium. Short communication. *Acta Trop.* **36**, 289–292
63. Fridy, P. C., Li, Y., Keegan, S., Thompson, M. K., Nudelman, I., Scheid, J. F., Oeffinger, M., Nussenzweig, M. C., Fenyö, D., Chait, B. T., and Rout, M. P. (2014) A robust pipeline for rapid production of versatile nanobody repertoires. *Nat. Methods* **11**, 1253–1260
64. Obado, S. O., Field, M. C., Chait, B. T., and Rout, M. P. (2016) High-efficiency isolation of nuclear envelope protein complexes from trypanosomes. *Methods Mol. Biol.* **1411**, 67–80
65. Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **4**, 1265–1272
66. Olsen, J. V., de Godoy, L. M. F., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–2021
67. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372
68. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805
69. Luber, C. A., Cox, J., Lauterbach, H., Fancke, B., Selbach, M., Tschopp, J., Akira, S., Wiegand, M., Hochrein, H., O'Keeffe, M., and Mann, M. (2010) Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity* **32**, 279–289
70. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402
71. Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277
72. Price, M. N., Dehal, P. S., and Arkin, A. P. (2010) FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490
73. Chang, J., Tommaso, P. Di, Taly, J., and Notredame, C. (2012) Accurate multiple sequence alignment of transmembrane proteins with PSI-Coffee. *BMC Bioinformatics* **13**, 1–7
74. Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., and Huelsenbeck, J. P. (2012) Mrbayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542
75. Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) Mega X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549

76. Edgar, R. C. (2004) Muscle: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797
77. Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321
78. Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Gateway Computing Environments Workshop (GCE)*, IEEE, New Orleans, LA: 1–8
79. Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., Gardner, M. J., Gingle, A., Grant, G., Harb, O. S., Heiges, M., *et al*. (2010) TriTrypDB: A functional genomic resource for the trypanosomatidae. *Nucleic Acids Res.* **38**, D457–D462
80. Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., *et al*. (2014) InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240
81. Eddy, S. R. (1998) Profile hidden Markov models. *Bioinformatics* **14**, 755–763
82. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J., and Punta, M. (2014) Pfam: The protein families database. *Nucleic Acids Res.* **42**, D222–D230
83. Batista, M., Marchini, F. K., Celedon, P. A. F., Fragoso, S. P., Probst, C. M., Preti, H., Ozaki, L. S., Buck, G. A., Goldenberg, S., and Krieger, M. A. (2010) A high-throughput cloning system for reverse genetics in Trypanosoma cruzi. *BMC Microbiol.* **10**, 259
84. Kugeratski, F. G., Batista, M., Inoue, A. H., Ramos, B. D., Krieger, M. A., and Marchini, F. K. (2015) pTcGW plasmid vectors 1.1 version: A versatile tool for Trypanosoma cruzi gene characterisation. *Mem. Inst. Oswaldo Cruz* **110**, 687–690
85. Oberholzer, M., Morand, S., Kunz, S., and Seebeck, T. (2006) A vector series for rapid PCR-mediated C-terminal *in situ* tagging of Trypanosoma brucei genes. *Mol. Biochem. Parasitol.* **145**, 117–120
86. Redmond, S., Vadivelu, J., and Field, M. C. (2003) RNAit: An automated web-based tool for the selection of RNAi targets in Trypanosoma brucei. *Mol. Biochem. Parasitol.* **128**, 115–118
87. Wickstead, B., Ersfeld, K., and Gull, K. (2002) Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of Trypanosoma brucei. *Mol. Biochem. Parasitol.* **125**, 211–216
88. Pfaffl, M. W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45
89. Mani, J., Güttinger, A., Schimanski, B., Heller, M., Acosta-Serrano, A., Pescher, P., Späth, G., and Roditi, I. (2011) Alba-domain proteins of Trypanosoma brucei are cytoplasmic RNA-binding proteins that interact with the translation machinery. *PLoS One* **6**, e22463
90. Das, A., Bellofatto, V., Rosenfeld, J., Carrington, M., Romero-zaliz, R., and Estévez, A. M. (2015) High throughput sequencing analysis of Trypanosoma brucei DRBD3/PTB1-bound mRNAs. *Mol. Biochem. Parasitol.* **199**, 1–4
91. Wippel, H. H., Malgarin, J. S., Inoue, A. H., Leprevost, F. da V., Carvalho, P. C., Goldenberg, S., and Alves, L. R. (2019) Unveiling the partners of the DRBD2-mRNP complex, an RBP in Trypanosoma cruzi and ortholog to the yeast SR-protein Gbp2. *BMC Microbiol.* **19**, 128
92. Pérez-díaz, L., Caroline, T., and Teixeira, S. M. (2017) Involvement of an RNA binding protein containing Alba domain in the stage-specific regulation of beta-amastin expression in Trypanosoma cruzi. *Mol. Biochem. Parasitol.* **211**, 1–8
93. Subota, I., Rotureau, B., Blisnick, T., Ngwabyt, S., Durand-Dubief, M., Engstler, M., and Bastin, P. (2011) ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation. *Mol. Biol. Cell* **22**, 4205–4219
94. Gupta, S. K., Chikne, V., Eliaz, D., Tkacz, I. D., Naboishchikov, I., Carmi, S., Ben-Asher, H. W., and Michaeli, S. (2014) Two splicing factors carrying serine-arginine Two splicing factors carrying serine-arginine motifs, TSR1 and TSR1IP, regulate splicing, mRNA stability, and rRNA processing in Trypanosoma brucei. *RNA Biol.* **11**, 715–731
95. Holden, J. M., Koreny, L., Obado, S., Ratushny, A. V., Chen, W., Bart, J., Navarro, M., Chait, B. T., Aitchison, J. D., Rout, M. P., and Field, M. C. (2018) Involvement in surface antigen expression by a moonlighting FG-repeat nucleoporin in trypanosomes. *Mol. Biol. Cell* **29**, 1100–1110
96. Elvira-Matelot, E., Bardou, F., Ariel, F., Jauvion, V., Bouteiller, N., Masson, L., Cao, J., Crespi, M. D., and Vaucheret, H. (2016) The nuclear ribonucleoprotein SmD1 interplays with splicing, RNA quality control, and posttranscriptional gene silencing in Arabidopsis. *Plant Cell* **28**, 426–438
97. Jiang, J., Liu, X., Liu, C., Liu, G., Li, S., and Wang, L. (2017) Integrating omics and alternative splicing reveals insights into grape response to high temperature 1 [OPEN]. *Plant Physiol.* **173**, 1502–1518
98. Buscemi, G., Saracino, F., Masnada, D., and Carbone, M. L. A. (2000) The Saccharomyces cerevisiae SDA1 gene is required for actin cytoskeleton organization and cell cycle progression. *J. Cell Sci.* **113**, 1199–1211
99. Michaeli, S. (2012). In: Bindereif, A., ed. *RNA Metabolism in Trypanosomes*, Springer, Berlin, Heidelberg: 123–148
100. Cristodero, M., Schimanski, B., Heller, M., and Roditi, I. (2014) Functional characterization of the trypanosome translational repressor SCD6. *Biochem. J.* **457**, 57–67
101. Holetz, F. B., Correa, A., Avila, A. R., Nakamura, C. V., Krieger, M. A., and Goldenberg, S. (2007) Evidence of P-body-like structures in Trypanosoma cruzi. *Biochem. Biophys. Res. Commun.* **356**, 1062–1067
102. Ebenezer, T. E., Carrington, M., Lebert, M., Kelly, S., and Field, M. C. (2017) Euglena gracilis genome and transcriptome: Organelles, nuclear genome assembly strategies and initial features. *Adv. Exp. Med. Biol.* **979**, 125–140
103. Kressler, D., de la Cruz, J., Rojo, M., and Linder, P. (1997) Fal1p is an essential DEAD-box protein involved in 40S-ribosomal-subunit biogenesis in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **17**, 7283–7294
104. Chang, C.-T., Hautbergue, G. M., Walsh, M. J., Viphakone, N., van Dijk, T. B., Philipsen, S., and Wilson, S. A. (2013) Chtop is a component of the dynamic TREX mRNA export complex. *EMBO J.* **32**, 473–486
105. van Dijk, T. B., Gillemans, N., Stein, C., Fanis, P., Demmers, J., van de Corput, M., Essers, J., Grosveld, F., Bauer, U.-M., and Philipsen, S. (2010) Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol. Cell. Biol.* **30**, 260–272
106. Azizi, H., Dumas, C., and Papadopoulou, B. (2017) The Pumilio-domain protein PUF6 contributes to SIDER2 retroposon-mediated mRNA decay in Leishmania. *RNA* **23**, 1874–1885
107. Dallagiovanna, B., Correa, A., Probst, C. M., Holetz, F., Smircich, P., De Aguiar, A. M., Mansur, F., Vieira, C., Mortara, R. A., Garat, B., Buck, G. A., Goldenberg, S., and Krieger, M. A. (2008) Functional genomic characterization of mRNAs associated with TcPUF6, a Pumilio-like protein from Trypanosoma cruzi. *J. Biol. Chem.* **283**, 8266–8273
108. Simon, M. J., and Bindu, N. (2020) Functions, mechanisms and regulation of Pumilio/Puf family RNA binding proteins: A comprehensive review. *Mol. Biol. Rep.* **47**, 785–807
109. Skaar, J. R., Ferris, A. L., Wu, X., Saraf, A., Khanna, K. K., Florens, L., Washburn, M. P., Hughes, S. H., and Pagano, M. (2015) The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res.* **25**, 288–305
110. Rienzo, M., and Casamassimi, A. (2016) Integrator complex and transcription regulation: Recent findings and pathophysiology. *Biochim. Biophys. Acta* **1859**, 1269–1280
111. Liu, Y., Li, S., Chen, Y., Kimberlin, A. N., Cahoon, E. B., and Yu, B. (2016) snRNA 3′ end processing by a CPSF73-containing complex essential for development in Arabidopsis. *PLoS Biol.* **14**, e1002571
112. Wu, Y., Albrecht, T. R., Baillat, D., Wagner, E. J., and Tong, L. (2017) Molecular basis for the interaction between Integrator subunits IntS9 and IntS11 and its functional importance. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4394–4399
113. Brogna, S., and Wen, J. (2009) Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat. Struct. Mol. Biol.* **16**, 107–113
114. Kataoka, N., Yong, J., Kim, V. N., Velazquez, F., Perkinson, R. A., Wang, F., and Dreyfuss, G. (2000) Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Mol. Cell* **6**, 673–682
115. Luo, M., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M., and Reed, R. (2001) Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* **413**, 644–647

116. Palacios, I. M., Gatfield, D., Johnston, D. S., and Izaurralde, E. (2004) An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature* **427**, 753–757
117. Nott, A., Hir, H. Le, and Moore, M. J. (2004) Splicing enhances translation in mammalian cells: An additional function of the exon junction complex. *Genes Dev.* **18**, 210–222
118. Delhi, P., Queiroz, R., Inchaustegui, D., Carrington, M., and Clayton, C. (2011) Is there a classical nonsense-mediated decay pathway in trypanosomes? *PLoS One* **6**, e25112
119. Strässer, K., and Hurt, E. (2001) Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature* **413**, 648–652
120. Naguleswaran, A., Gunasekera, K., Schimanski, B., Heller, M., Hemphill, A., Ochsenreiter, T., and Roditi, I. (2015) Trypanosoma brucei RRM1 is a nuclear RNA-binding protein and modulator of chromatin structure. *mBio* **6**, e00114
121. Wippel, H. H., Malgarin, J. S., De Toledo, S., Vidal, N. M., Marcon, B. H., Miot, H. T., Marchini, F. K., Goldenberg, S., and Alves, L. R. (2019) The nuclear RNA-binding protein RBSR1 interactome in Trypanosoma cruzi. *J. Eukaryot. Microbiol.* **66**, 244–253
122. Mishra, A., Kaur, J. N., McSkimming, D. I., Hegedűsová, E., Dubey, A. P., Ciganda, M., Paris, Z., and Read, L. K. (2021) Selective nuclear export of mRNAs is promoted by DRBD18 in Trypanosoma brucei. *Mol. Microbiol.* **116**, 827–840
123. Rout, M. P., Obado, S. O., Schenkman, S., and Field, M. C. (2017) Specialising the parasite nucleus: Pores, lamins, chromatin, and diversity. *PLoS Pathog.* **13**, 1–16
124. Bayliss, R., Leung, S. W., Baker, R. P., Quimby, B. B., Corbett, A. H., and Stewart, M. (2002) Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO J.* **21**, 2843–2853
125. Fribourg, S., Braun, I. C., Izaurralde, E., and Conti, E. (2001) Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor. *Mol. Cell* **8**, 645–656
126. Leandro de Jesus, T. C., Calderano, S. G., Vitorino, F. N. de L., Llanos, R. P., Lopes, M. de C., Araújo, C. B. de, Thiemann, O. H., Reis, M. da S., Elias, M. C., and da Cunha, J. P. C. (2017) Quantitative proteomic analysis of replicative and nonreplicative forms reveals important insights into chromatin biology of Trypanosoma cruzi. *Mol. Cell. Proteomics* **16**, 23–38
127. Coller, J. M., Tucker, M., Sheth, U., Valencia-sanchez, M. A., and Parker, R. (2001) The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* **7**, 1717–1727
128. Kramer, S., Queiroz, R., Ellis, L., Hoheisel, J. D., Clayton, C., and Carrington, M. (2010) The RNA helicase DHH1 is central to the correct expression of many developmentally regulated mRNAs in trypanosomes. *J. Cell Sci.* **123**, 699–711
129. Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010) Article decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. *Mol. Cell* **39**, 773–783
130. Köhler, A., and Hurt, E. (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.* **8**, 761–773