Experimental Examination of EFL and MATX Eukaryotic Horizontal Gene Transfers: Coexistence of Mutually Exclusive Transcripts Predates Functional Rescue

Jana Szabová,§,[1] Petr Růžička,†§,[2] Zdeněk Verner,‡,[2] Vladimír Hampl,§,[1] and Julius Lukeš*§,[2]

[1]Charles University in Prague, Faculty of Science, Department of Parasitology, Prague, Czech Republic

[2]Biology Centre, Institute of Parasitology, Czech Academy of Sciences, and Faculty of Sciences, University of South Bohemia, České Budějovice (Budweis), Czech Republic

†Present address: Institute of Experimental Botany, Czech Academy of Sciences, Prague, Czech Republic.

‡Present address: Sanatorium Helios s.r.o., Brno, Czech Republic.

§These authors contributed equally to this work.

***Corresponding author:** E-mail: jula@paru.cas.cz.

Associate editor: Hervé Philippe

Research article

Abstract

Many eukaryotic genes do not follow simple vertical inheritance. Elongation factor 1α (EF-1α) and methionine adenosyl transferase (MAT) are enzymes with complicated evolutionary histories and, interestingly, the two cases have several features in common. These essential enzymes occur as two relatively divergent paralogs (EF-1α/EFL, MAT/MATX) that have patchy distributions in eukaryotic lineages that are nearly mutually exclusive. To explain such distributions, we must invoke either multiple eukaryote-to-eukaryote horizontal gene transfers (HGTs) followed by functional replacement or presence of both paralogs in the common ancestor followed by long-term coexistence and differential losses in various eukaryotic lineages. To understand the evolution of these paralogs, we have performed in vivo experiments in *Trypanosoma brucei* addressing the consequences of long-term coexpression and functional replacement. In the first experiment of its kind, we have demonstrated that EF-1α and MAT can be simultaneously expressed with EFL and MATX, respectively, without affecting the growth of the flagellates. After the endogenous MAT or EF-1α was downregulated by RNA interference, MATX immediately substituted for its paralog, whereas EFL was not able to substitute for EF-1α, leading to mortality. We conclude that MATX is naturally capable of evolving patchy paralog distribution via HGTs and/or long-term coexpression and differential losses. The capability of EFL to spread by HGT is lower and so the patchy distribution of EF-1α/EFL paralogs was probably shaped mainly by deep paralogy followed by long-term coexistence and differential losses.

Key words: EFL, MATX, horizontal gene transfer, functional rescue, RNAi, *Trypanosoma*.

Introduction

The transfer of genetic information among distantly related organisms called horizontal (= lateral) gene transfer (HGT) represents one of the major driving forces of evolution (Keeling and Palmer 2008). The pervasive occurrence of HGT among prokaryotic organisms is apparent in their genomes and can be easily experimentally demonstrated, thus disputing the actual existence of stable bacterial species (Welch et al. 2002; Doolittle and Papke 2006; Doolittle and Bapteste 2007; Papke 2009). Recently, the role of HGT is becoming recognized also as an important force in the evolution of eukaryotes and an increasing number of examples are being reported (Andersson 2005, 2009; Richards et al. 2006; Watkins and Gray 2006; Andersson et al. 2007; Keeling 2009; Whitaker et al. 2009; Stairs et al. 2011). Even though several mechanisms of HGT have been proposed (Gogarten 2003), two are believed to be prevalent: endosymbiotic gene transfer (Martin and Schnarrenberger 1997) and "you-are-what-you-eat" (Doolittle 1998). The former occurred upon endosymbiosis of the bacterial ancestors of mitochondria and plastids and represents HGT on the largest scale because more than 90% of the endosymbiosed genomes were subsequently lost or transferred to the host cell nucleus and thus form a substantial part of the coding capacity of the nuclei in extant eukaryotes (Esser et al. 2004; Bock and Timmis 2008). Most of these organelle-derived proteins remain functionally associated with the organelle of their evolutionary origin (Kurland and Andersson 2000). However, some of these proteins have eventually found their way to other cellular compartments, being responsible for the mosaic pattern of most metabolic pathways in a typical eukaryotic cell (Gabaldón and Huynen 2004; Oborník and Green 2005).

Independently of these massive endosymbiosis-driven HGT events, most eukaryotes were subject to intermittent acquisitions of genomic material from prokaryotes or other eukaryotes. According to the "you-are-what-you-eat" concept, the digested prey is the pervasive source of these transferred genes (Doolittle 1998). Various eukaryotic groups differ in the extent of HGT from prokaryotes, and their life style and environment seem to be an important factor—HGTs occurred frequently in the evolutionary

history of rumen ciliates (Ricard et al. 2006), anaerobic protists (Andersson 2006; Andersson et al. 2007), or diatoms (Bowler et al. 2008), whereas very few such events have so far been documented in yeasts (Dujon et al. 2004) or animals (Kondrashov et al. 2006).

Although the eukaryote-to-eukaryote HGTs are likely to be underestimated (Keeling and Palmer 2008), this process can hardly be considered a frequent one. Particularly interesting are highly conserved essential genes that have a surprisingly complex evolutionary history—elongation factor 1α (EF-1α) and methionine adenosyltransferase (MAT) (Keeling and Inagaki 2004; Gile et al. 2006; Ruiz-Trillo et al. 2006; Noble et al. 2007; Kamikawa et al. 2008, 2009, 2010, 2011; Sanchez-Perez et al. 2008; Cocquyt et al. 2009; Gile, Faktorová, et al. 2009; Gile, Novis, et al. 2009; Sakaguchi et al. 2009), being the best known examples. The main function of EF-1α is to bring an aminoacyl-transfer RNA into the A site of the ribosome (Andersen et al. 2003). This extremely abundant protein has also been implicated in ubiquitin-dependent protein degradation (Chuang et al. 2005) and localization of selected transcripts via simultaneous binding of EF-1α to actin (Liu et al. 2002). MAT is the only enzyme synthesizing S-adenosyl-L-methionine, which is one of the key metabolites, as it donates the methyl group for most methylation reactions in prokaryotic and eukaryotic cells (Chiang et al. 1996).

Although elongation factor like (EFL) and MATX were initially considered to have evolved by vertical descent, the respective phylogenetic trees were inconsistent with such a simple scenario. In both cases, the analyses revealed that subsets of unrelated organisms possess a divergent version of the gene—EFL and MATX (Keeling and Inagaki 2004; Sanchez-Perez et al. 2008). In general, the patchy distribution of two paralogs can be explained by two outermost scenarios: 1) deep paralogy—presence of both paralogs in the common ancestor—followed by differential loss of one variant in individual lineages or 2) more recent origin of one paralog in one lineage of eukaryotes followed by its spread by eukaryote-to-eukaryote HGT. It is difficult to distinguish between these alternatives purely on the basis of phylogenetic analyses of protein sequences and their distributions. Theoretically, because the scenarios differ in the types of events they invoke, we might use the principle of parsimony and prefer the one that requires fewer improbable events. The first scenario minimizes the events of HGT replacements and expects long-term coexistence of both paralogs. The assumption of this scenario therefore is that the two paralogs are (or were in the past) capable of long-term coexpression in a single cell compartment without negative effect on the fitness of the organism. The second scenario expects that one paralog (probably the less frequent and less diversified one, namely MATX and EFL) spreads among eukaryotes via HGT. The assumption of this scenario is that this paralog is in a very short time able to substitute the function of its counterpart. Unfortunately, in the cases of EFL/EF-1α and MATX/MAT, we have no information how they fulfill one assumption or the other.

The coexistence of these paralogs under natural conditions is rare if not totally lacking. The distribution of EFL/EF-1α paralogs and MATX/MAT paralogs is almost strictly mutually exclusive, that is, organisms have either EF-1α or EFL but not both, the same applying to MAT and MATX. Strangely enough, the exceptional group in both cases is the diatoms. The genome of *Thalassiosira pseudonana* harbors both variants of these genes (EFL, EF-1α, MAT, and MATX) (Armbrust et al. 2004), transcripts of both EFL and EF-1α have been detected in *Th. pseudonana* and in five other diatom species (Kamikawa et al. 2008), and finally transcripts of both MAT and MATX were revealed in another four diatom species (Kamikawa et al. 2009). This almost strict mutually exclusive distribution led to a proposal that the long-term coexistence of both paralogs in one compartment is detrimental for the cell, probably due to problems with regulation, competition for substrate, or in the case of MAT/MATX, formation of less functional heteromers (Sanchez-Perez et al. 2008).

The process of HGT replacement of these essential proteins by their paralogs is, however, potentially problematic as well. It is difficult to envisage a smooth switch, during which these essential proteins are replaced with their horizontally acquired paralogs that take instantly over their functions. Moreover, the replacement is inevitably preceded by the potentially hazardous period of coexpression of both variants.

We have decided to experimentally test on the model of *Trypanosoma brucei* how the two paralog couples (EF-1α/EFL and MAT/MATX) satisfy assumptions of the two evolutionary scenarios. For the first time, we have simulated step-by-step the process of HGT under laboratory conditions. We have shown that EFL and MATX can coexist with EF-1α and MAT, respectively. Moreover, the MATX gene from *Euglena gracilis* was able to rescue the RNA interference (RNAi) knockdown for MAT in *Tr. brucei*, but in the same organism, the EFL gene from *Diplonema papillatum* failed to rescue the knockdown of EF-1α. Although MAT/MATX fulfills assumptions of both scenarios, EF-1α/EFL apparently fulfills just one of them.

Materials and Methods

EF-1α and EFL Constructs

Oligonucleotides for generation of gene fragments suitable for generation of RNAi knockdown cell lines were designed using the RNAit online tool available on the TrypanoFAN web site (http://trypanofan.path.cam.ac.uk/software/RNAit.html). The 453 bp-long 5′ region of the *Tr. brucei* EF-1α gene was amplified using oligonucleotides EF-1-F (5′-GGATCCTGGAGGCACTAGACATGCTG-3′) and EF-1-R (5′-AAGCTTCGATCTTCGACTCGATCTCC-3′) (added BamH*I* and Hind*III* restriction sites are underlined) and cloned via these restriction sites into the p2T7-177 RNAi vector carrying phleomycin resistance.

For constitutive expression in *Tr. brucei* of the full-size exogenous EFL, genes from *D. papillatum* or *Isochrysis galbana* were used. EFL from *D. papillatum* was expressed using

either a pABPURO vector containing an HA_3-tag and puromycin resistance (Long et al. 2008) or a pHD1344tub vector with TAP-tag and puromycin resistance (Carnes et al. 2008). For the expression of EFL from *I. galbana*, the pABPURO vectors with or without HA_3-tag were used. The entire open reading frame of the EFL gene (accession number ACO50119) was polymerase chain reaction (PCR) amplified from the cDNA of the diplonemid *D. papillatum* using oligonucleotides Dp-HA-F (5′-TCACATCGATATGGCTAACGCTACCGA-3′) and Dp-HA-R (5′-AGTGGCTAGCCTTCTTCTTGCCCTTGC-3′) (added ClaI and NheI restriction sites are underlined) for pABPURO, and oligonucleotides Dp-TAP-F (5′-TCACAAGCTTATGGCTAACGCTACCGA-3′) and Dp-TAP-R (5′-AGTGGGATCCCTTCTTCTTGCCCTTGC-3′) (added BamHI and HindIII restriction sites are underlined) for pHD1344tub.

In the case of the *I. galbana* EFL gene (accession number AAV34146), its entire open reading frame was PCR amplified from the total DNA using oligonucleotides Ig1-F (5′-AAGCTTATGGCCTCCGAGAAA-3′) and Ig1-R (5′-GGATCCCTACTTCTTCTTCTT-3′) (added HindIII and BamHI restriction sites are underlined), the amplicon was cloned into pCRII TOPO (Invitrogen) and subsequently recloned into the puromycin resistance–carrying pABPURO vectors with or without the HA_3-tag (Long et al. 2008). Proper integration of each construct was confirmed by sequencing. Comparison of the *I. galbana* EFL sequence with other EFL sequences has not revealed any introns that could potentially preclude the proper expression in trypanosomes.

MAT and MATX Constructs

To generate the RNAi knockdown cells, a 438 bp-long 5′ fragment of the *Tr. brucei* MAT gene was amplified using oligonucleotides IF-F (5′-TCACTCTAGAACGACGGTGTGTCAAATGAA-3′) and IF-R (5′-AGTGAAGCTTGCAGTCGGAAGTTTTTCTGC-3′) (added XbaI and HindIII restriction sites are underlined) and cloned into the p2T7-177 RNAi vector. Furthermore, the full-size MATX gene from the euglenid *E. gracilis* (accession number GU989640) was amplified from a cDNA clone using oligonucleotides RE-F (5′-TCACATCGATATGGCTGAATCTGCTTC-3′) and RE-R (5′-AGTGGCTAGCGTCCA CCCACTTCTGCA-3′) (added NheI and ClaI restriction sites are underlined). The amplicon was cloned into the pABPURO vector containing HA_3-tag and puromycin resistance as described above.

Transfection, Cloning, and RNAi Induction

The HA_3-tagged *E. gracilis* MATX in pABPURO was digested with MluI. Digestion with NotI was used to linearize all the other constructs. After digestion, 10 μg of each linearized vector was individually transfected into exponentially growing (at 27 °C in SDM-79 medium) procyclic *Tr. brucei* 29-13 strain or cell lines derived from thereof, using 2-mm cuvettes and a BTX electroporator with the settings of 1600 V, 25 μfarads, and 500 ohms. The clones were obtained after about 2-week cultivation by limiting dilution in 24-well plates at 27 °C in the presence of 5% CO_2, with 1 μg/ml puromycin or 1 μg/ml phleomycin as the selectable agent depending on the type of construct.

The following clonal cell lines derived from the 29-13 strain were prepared for the EF-1α/EFL experiments: 1) RNAi knockdowns containing *Tr. brucei* EF-1α in p2T7-177; 2) cells constitutively overexpressing HA_3-tagged *D. papillatum* EFL in pABPURO; 3) cells constitutively overexpressing HA_3-tagged *D. papillatum* EFL cotransfected with p2T7-177 containing EF-1α; 5) cells constitutively overexpressing TAP-tagged *D. papillatum* EFL in pHD1344tub; 5) cells constitutively overexpressing TAP-tagged *D. papillatum* EFL cotransfected with p2T7-177 containing EF-1α; 6) cells constitutively overexpressing HA_3-tagged *I. galbana* EFL in pABPURO; 7) cells constitutively overexpressing HA_3-tagged *I. galbana* EFL cotransfected with p2T7-177 containing EF-1α; 8) cells constitutively overexpressing nontagged *I. galbana* EFL in pABPURO; and 9) cells constitutively overexpressing nontagged *I. galbana* EFL cotransfected with p2T7-177 containing EF-1α.

The following clonal cell lines derived from the 29-13 strain were made for the MAT/MATX experiments: 1) RNAi knockdowns containing *Tr. brucei* MAT in p2T7-177; 2) cells constitutively overexpressing HA_3-tagged *E. gracilis* MATX in pABPURO; and 3) cells constitutively overexpressing HA_3-tagged *E. gracilis* MATX cotransfected with p2T7-177 containing MAT construct.

From each cell line containing the RNAi p2T7-177 vector, always a single clone was selected for further experiments based on the tightness of tetracycline-inducible expression of target double-stranded (ds) RNA and the corresponding robust elimination of target mRNA, as determined by Northern blot analysis using the *Tr. brucei* EF-1α or MAT gene as a probe. Synthesis of dsRNA was induced by the addition of 1 μg/ml tetracycline to the medium.

Northern Blot Analysis

Approximately 5 μg of total RNA/lane was loaded on a 1% formaldehyde agarose gel, blotted, and cross-linked following standard protocols. After prehybridization in NaPi solution (0.25 M Na_2HPO_4 and 0.25 M NaH_2PO_4, pH 7.2, 1 mM ethylenediaminetetraacetic acid, and 7% sodium dodecyl sulfate [SDS]) for 30 min at 60 °C, hybridization was performed overnight in the same solution at the same temperature. A wash in 2× saline sodium citrate (SSC) + 0.1% SDS at room temperature for 20 min was followed by three washes in 0.2× SSC + 0.1% SDS for 20 min each at 55 °C.

Western Blot Analysis

Cell lysates corresponding to 2.5×10^6 cells/lane were separated on a 15% sodium dodecyl sulfate–polyacrylamide gel electrophoresis, blotted, and for HA_3-tagged constructs, the membranes were treated with an anti-HA_3-tag mouse monoclonal antibody, followed by chicken anti-mouse antibody coupled to horseradish peroxidase. For TAP-tagged construct, the membranes were treated with anti-TAP-tag mouse monoclonal antibody, followed by rabbit anti-mouse antibody. Signal in Western blots was quantified with the Bio-Rad quantity one software.

Growth Analysis

Growth curves of selected *Tr. brucei* clones representing the (non)-induced RNAi knockdowns and the other genetically manipulated cells, obtained over a period of 12 days after RNAi induction were established using the Beckman Z2 Cell Counter.

Results

EFL and EF-1α mRNAs Are Fully Compatible

We have first tested whether EFL from *D. papillatum*, a diplonemid flagellate closely related to kinetoplastids, will be compatible with the *Tr. brucei* proteome, which contains the EF-1α. For that purpose, the full-size *D. papillatum* EFL gene was cloned into a vector that allows its constitutive expression under the procyclin promoter. The NotI-linearized vector was transfected into the 29-13 *Tr. brucei* strain where it was integrated into the tubulin locus. Total RNA was isolated from a puromycin-resistant clonal cell line and analyzed by Northern blotting using the full-size *D. papillatum* EFL gene as a probe. The analysis showed that the introduced gene was strongly transcribed in the transfectants, whereas no signal was detected in the wild type 29-13 cells (fig. 1*A*). At stringent hybridization conditions (60 °C), no cross-hybridization with the EF-1α mRNA, transcribed from three endogenous copies of the EF-1α gene, was observed. Western blot analysis using specific anti-HA_3 or anti-TAP tag monoclonal antibodies detected a tagged protein translated from the heterologous gene (fig. 1*B*). The morphology of *Tr. brucei* cells containing both the endogenous EF-1α and exogenous EFL appeared normal by light microscopy (data not shown), and their growth was similar in comparison with the wild-type cells (fig. 2*B* and *C*) indicating that EF-1α and EFL are fully compatible.

Silencing of EF-1α Inhibits Growth

The addition of tetracycline into the medium triggers synthesis of dsRNA in trypanosomes transfected with the EF-1α-containing inducible RNAi construct. The extent of EF-1α mRNA silencing and the tightness of its inducible ablation was determined by Northern blotting, which revealed lack of leaky dsRNA transcription (fig. 1*C*, lanes 3 and 7) and virtually complete ablation of EF-1α mRNA upon the induction of RNAi (fig. 1*C*, lanes 4 and 8). Next, we have followed the growth of cells constitutively expressing exogenous *D. papillatum* EFL in which RNAi against EF-1α was induced. The elimination of EF-1α mRNA triggers an almost instant cessation of growth, eventually causing death regardless of the absence (fig. 2*A*) or presence (fig. 2*B* and *C*) of exogenous EFL that is efficiently translated (fig. 1*B*), yet still fails to rescue the growth phenotype.

To confirm these findings, similar experiments were performed with cells with constitutive expression of the EFL gene from a different organism, the haptophyte *I. galbana*. In this series of experiments, we also wanted to establish whether the attachment of a tag to the C-terminus of expressed exogenous EFL protein does (not) interfere with its enzymatic function. Therefore, two parallel experiments were performed; one using construct with HA_3 tagged EFL and the other using nontagged EFL. Because results of both experiments were very similar, only the experiments without tag are presented below. Successful constitutive expression of exogenous EFL and inducible downregulation of endogenous EF-1α was confirmed using Northern blotting (supplementary fig. S1, Supplementary Material online). As in the case of EFL from *D. papillatum*, trypanosomes in which both EF-1α and *I. galbana* EFL were expressed grew at the same rate as the wild-type cells, with the exogenous EFL failing to rescue the growth of the cells with RNAi-ablated EF-1α (fig. 2*D*).

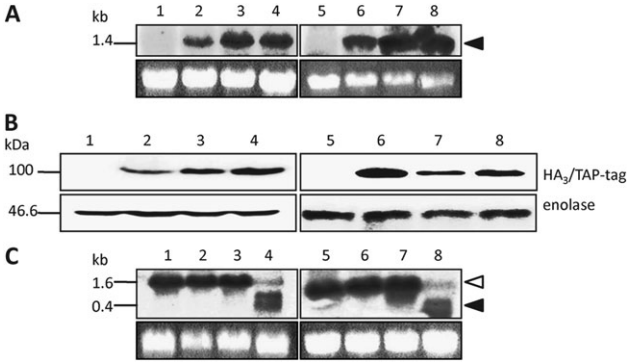

Fig. 1. Expression of exogenous EFL from *Diplonema papillatum* and (parallel) RNAi silencing of EF-1α in *Trypanosoma brucei*. (*A*) The EFL mRNA is expressed in *Trypanosoma brucei* cells. Upper panels: Levels of EFL mRNA with HA3-tag (lanes 1–4) and EFL mRNA with TAP-tag (lanes 5–8) were analyzed by blotting 10 μg of total RNA/lane extracted from 29-13 wild-type cells (lanes 1 and 5), cells constitutively expressing exogenous EFL (lanes 2, 6), noninduced cells expressing EFL and also containing RNAi vector against EF-1α (lanes 3 and 7), and same cells as in lanes 3 and 7 in which RNAi was induced (lanes 4 and 8). The full-length EFL gene was used as a probe, and hybridization was performed at 60 °C, at which no cross-hybridization with EF-1α mRNA occurs. Lower panels: As a loading control, the gel was stained with ethidium bromide to visualize ribosomal RNA (rRNA) bands. (*B*) The EFL protein is expressed in *Tr. brucei* cells. Upper panels: The levels of the HA3-tagged (lanes 1–4) and TAP-tagged (lanes 5–8) exogenous EFL protein were followed using specific mouse monoclonal antibodies. The levels were analyzed in total lysates (from 5×10^6cells) from 29-13 wild-type cells (lanes 1 and 5), cells constitutively expressing EFL (lanes 2 and 6), noninduced cells constitutively expressing exogenous EFL and containing RNAi vector against EF-1α (lanes 3 and 7), and the same cells as in lanes 3 and 7 in which RNAi was induced (lanes 4 and 8). Lower panels: Enolase visualized by specific rabbit polyclonal antibodies was used as loading control. (*C*) Down regulation of EF-1α. Upper panels: Level of EF-1α mRNA in the cells with HA3-tagged (lanes 1–4) and TAP-tagged exogenous EFL (lanes 5–8) were analyzed in total RNA extracted from 29-13 wild-type cells (lanes 1 and 5), cells constitutively expressing EFL (lanes 2 and 6), noninduced cells containing RNAi vector against EF-1α and constitutively expressing EFL (lanes 3 and 7); and same cells as in lanes 3 and 7 in which RNAi was induced (lanes 4 and 8). The full-length EF-1α gene was used as a probe, and hybridization was performed at 60 °C, at which no cross-hybridization with EFL mRNA occurs. The positions of the EF-1α mRNAs and respective dsRNA are indicated with white and gray arrowheads, respectively. Lower panels: As a loading control, the gel was stained with ethidium bromide to visualize rRNA bands.

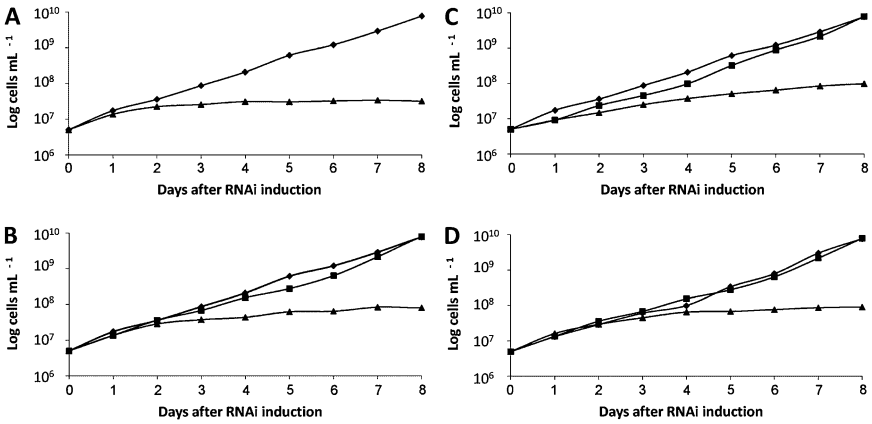

FIG. 2. Lethality RNAi-ablation of EF-1α is not rescued by exogenous EFL in *Trypanosoma brucei*. Cell numbers were measured using a Coulter Counter Z2. The *y* axis is labeled by a logarithmic scale and represents the product of cell densities measured and total dilution. Growth curves are one representative set from three independent experiments. (*A*) The growth of cells with ablated EF-1α mRNA is inhibited (triangles), as compared with the 29-13 wild-type cells (diamonds). (*B*) The growth of cells with inducibly ablated EF-1α mRNA that also constitutively express *Diplonema papilatum* EFL with HA3-tag is inhibited (triangles) in comparison with the noninduced cells constitutively expressing the same EFL (squares) and 29-13 wild-type cells (diamonds), which grow at the same rate. (*C*) The growth of cells with inducibly ablated EF-1α mRNA that also constitutively express *D. papilatum* EFL with TAP-tag is inhibited (triangles) in comparison with the noninduced cells constitutively expressing the same EFL (squares) and 29-13 wild-type cells (diamonds), which grow at the same rate. (*D*) The growth of cells with inducibly ablated EF-1α mRNA that also constitutively express *Isochrysis galbana* EFL is inhibited (triangles) in comparison with the noninduced cells constitutively expressing the same EFL (squares) and 29-13 wild-type cells (diamonds), which grow at the same rate.

MAT and MATX mRNAs Are Compatible

The same strategy as the one described above for the EFL and EF-1α genes was also used for the MAT/MATX system. First, using Northern blot analysis and the MATX gene as a probe, we have shown that in transfected *Tr. brucei*, the MATX gene from *E. gracilis* is indeed expressed (fig. 3*A*). This approach also excluded the unlikely yet possible presence of another MATX gene in the 29-13 wild-type cells, which contain nine copies of the MAT gene in their genome. As shown by Western blot analysis using specific anti-HA_3 tag monoclonal antibody, MATX is not only transcribed but also efficiently translated in *Tr. brucei* transfected with the respective construct (fig. 3*B*). Growth curve analysis of wild-type cells and those overexpressing MATX clearly demonstrated that trypanosomes fully tolerate expression of this exogenous gene (fig. 4*B*).

MATX Rescues MAT Deficiency

Next, we have downregulated endogenous MAT mRNA using RNAi. In selected clones, the ablation was very efficient, because after 48 h of RNAi induction, virtually no target transcript was detectable by Northern blot analysis (fig. 3*C*). We have then followed the consequences of such depletion on cell growth. As shown in figure 4*A*, MAT is clearly an essential protein for trypanosomes, as they were unable to propagate in its absence. The situation was, however, strikingly different with cells depleted for MAT but overexpressing *E. gracilis* MATX. The constitutive expression of MATX in the inducible *Tr. brucei* RNAi MAT knockdown fully rescued the growth, which differed neither from the 29-13 wild-type cells nor from the noninduced knockdowns (fig. 4*B*). This experiment shows that when the euglenid MATX is concurrently expressed with the endogenous MAT, following the depletion of the latter protein, MATX quickly takes over its function(s) and rescues the cells, which would otherwise die.

Discussion

Using in vivo experiments, we have mimicked the process of acquisition of an exogenous paralog of an enzyme by HGT followed by a period of simultaneous expression and eventually leading to functional replacement of the endogenous paralog. Our experiments have several important implications for the evolution of the studied paralogs.

For two gene couples (MAT/MATX and EF-1α/EFL) whose members are virtually never found simultaneously in one cell, we have demonstrated for the first time that their products can cohabitate in *Tr. brucei* in the same compartment at least for several weeks, which is a substantially long period for an organism with 8 h-long generation time. Although our results do not attach any selective advantage to this highly risky and cumbersome process, they show that it can indeed happen under experimentally controlled conditions. Growth phenotype of trypanosomes coexpressing MAT and MATX or EF-1α and EFL is similar and comparable with the wild-type cells. The results of our experiments are in agreement with the observations that MAT and MATX mRNA as well as EF-1α and EFL mRNA are present in several diatom species at the same time (Kamikawa et al. 2008, 2009). In the light of these findings, the long-term cohabitation followed by slow and more or less random differential losses of one or the other paralog is a plausible scenario for both paralog couples.

MATX and EFL differ in their ability to substitute the endogenous paralog in trypanosomes. After downregulation of MAT by RNAi, the expressed exogenous MATX was able

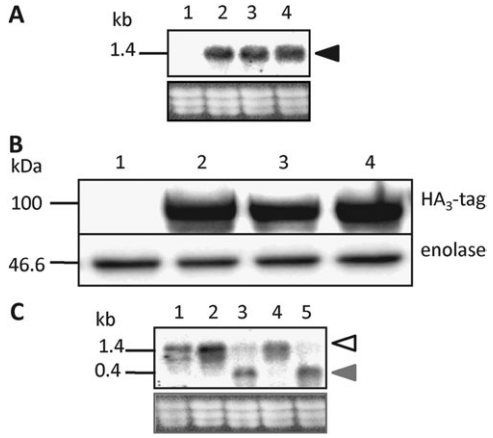

FIG. 3. Expression of exogenous MATX from *Euglena gracilis* and (parallel) RNAi silencing of MAT in *Trypanosoma brucei*. (*A*) The MATX mRNA is expressed in *T. brucei* cells. Upper panel: Level of MATX mRNA was analyzed by blotting 10 μg of total RNA extracted from 29-13 cells (lane 1), cells constitutively expressing MATX from *E. gracilis* (lane 2), noninduced cells constitutively expressing exogenous MATX and containing RNAi vector against endogenous MAT (lane 3), and the same cells as in lane 3 in which RNAi was induced (lane 4). The full-size MATX gene was used as a probe, and hybridization was performed at 60 °C. The position of the MATX mRNA is indicated with a black arrowhead. Lower panel: As a loading control, the gel was stained with ethidium bromide to visualize ribosomal RNA (rRNA) bands. (*B*) The MATX protein is expressed in *T. brucei*. Upper panel: The levels of the HA3-tagged exogenous MATX protein were followed using specific mouse monoclonal antibodies. The levels were analyzed in total lysates from 29-13 wild-type cells (lane 1), cells constitutively expressing MATX (lane 2), noninduced cells constitutively expressing exogenous MATX and containing RNAi vector against MAT (lane 3), and the same cells as in lane 3 in which RNAi was induced (lane 4). Lower panel: Enolase visualized by specific rabbit polyclonal antibodies was used as loading control. (*C*) Down regulation of MAT. Upper panel: Levels of MAT mRNA and respective dsRNA were analyzed in total RNA extracted from the following cell lines: 29-13 wild-type cells (lane 1); noninduced cells containing RNAi vector against endogenous MAT and constitutively expressing exogenous MATX (lane 2); same cells as in lane 2 in which RNAi was induced (lane 3); noninduced cells containing RNAi vector against MAT (lane 4); same cells as in lane 4 in which RNAi was induced (lane 5). The 5′ region of the *T. brucei* MAT gene was used as a probe, and hybridization was performed at 60 °C. The positions of the targeted MAT mRNA and the dsRNA are indicated with white and gray arrowheads, respectively. Lower panel: As a loading control, the gel was stained with ethidium bromide to visualize rRNA bands.

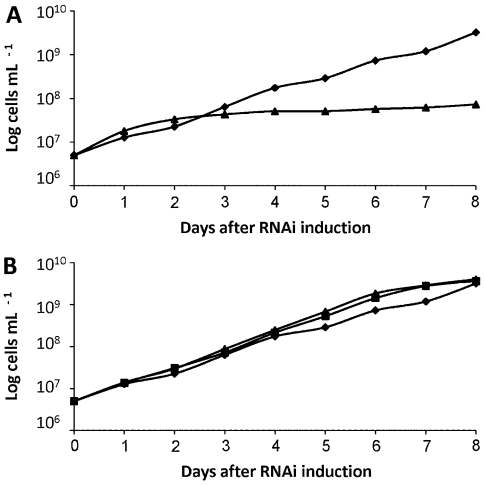

FIG. 4. Lethality of RNAi-ablated MAT is rescued by exogenous MATX in *Trypanosoma brucei*. The growth curves were performed as described in figure 2. (*A*) The growth of cells with ablated MAT mRNA is inhibited (triangles), as compared with 29-13 wild-type cells (diamonds). (*B*) The growth of cells with inducibly ablated MAT mRNA is rescued by the expression of exogenous MATX (triangles), and the cells grow at about the same rate as the noninduced cells constitutively expressing the same MATX (squares) and 29-13 wild-type cells (diamonds).

to take over the function of its counterpart. The functional replacement of MAT by MATX happened immediately with no effect on the growth phenotype. This result is in agreement with the work of Ho et al. (2007), who showed that MATX from a dinoflagellate *Crypthecodinium cohnii* rescued the MAT knockout of yeast. On the other hand, cells expressing EFL but depleted for EF-1α died in our experiments irrespective of the presence or absence of tags on the EFL protein and irrespective of the *Diplonema* or *Isochrysis* origin of the EFL gene.

Because horizontally transferred genes may be disadvantaged in codon usage, we have compared this parameter for both gene couples. Codon usage of MATX and both EFLs departs from the kinetoplastid consensus to about the same extent as the endogenous MAT and EF-1α (supplementary figs. S2–S8, Supplementary Material online). We therefore conclude that codon usage bias does not play significant role in establishment of horizontally acquired MATX and EFL.

The analysis of functional divergence in case of EFL (Keeling and Inagaki 2004) and the conservation of all functional residues in case of MATX (Sanchez-Perez et al. 2008) suggest that these distant paralogs can perform the function of their counterparts. Nevertheless, the overall amino acid identity between EF-1α and EFL (40–45%) is generally lower than the amino acid identity between MAT and MATX (55–64%, excluding insertions). The amino acid identities between the particular gene pairs that have been used in our experiments fell into the aforementioned ranges, *Euglena* MATX with *Trypanosoma* MAT, *Diplonema* EFL with *Trypanosoma* EF-1α, and *Isochrysis* EFL with *Trypanosoma* EF-1α share 55%, 44%, and 42% of amino acids, respectively. The lower identity between EFL and EF-1α might contribute to the incapability of EFL to functionally substitute EF-1α. The fact that EFL was not able to substitute its paralog while MATX was could also be explained by the complexity hypothesis (Cohen et al. 2010). This hypothesis posits that genes, whose products are involved in many interactions with other proteins or molecules (like elongation factors), are less prone to transfers than genes with less interactions (enzymes like MATX).

Judging by their phylogenies and distribution among eukaryotes, the cases of EF-1α/EFL and MAT/MATX look very similar, yet our laboratory experiments indicate that this similarity may only be superficial. In the case of MAT/MATX, both long-term coexistence and horizontal transfer

followed by instantaneous functional replacement are plausible. MAT/MATX pair therefore satisfies assumptions of both outermost scenarios how the patchy distribution might evolve, that is, ancestral presence of both paralogs followed by differential losses or origin in one lineage followed by spread via HGTs. Results of our experiments thus do not help judging which of these scenarios are more plausible in this particular case. In fact, it is reasonable to expect that both phenomena—differential loss and HGT—contributed to the evolution of MAT/MATX patchy distribution. In the case of EF-1α/EFL, we have demonstrated that the coexistence of both variants is possible. This result is in agreement with the hypotheses that this dual state could be maintained for millions of years in euglenids (Gile, Faktorová, et al. 2009), diatoms (Kamikawa et al. 2008), green algae (Noble et al. 2007; Cocquyt et al. 2009), or even for much longer time since the common ancestor of all extant eukaryotes (Kamikawa et al. 2010). On the contrary, the instantaneous functional replacement of endogenous EF-1α by exogenous EFL was not successful, so at least in our experimental setting, we were not able to show that the EF-1α/EFL pair fulfills the assumption of the multiple-HGT scenario, and this scenario therefore seems in this particular case less probable. This does not mean that after a sufficiently long period of coexpression, when the cells become adapted to the exogenous paralog, EFL could not be able to substitute EF-1α. This situation, depending on the length of the adaptation period, however, approaches to the long-term coexpression followed by differential losses scenario. It is theoretically possible that the EFL gene was horizontally transferred at some points of its evolutionary history yet there is better evidence that coexpression followed by losses played a major role in the shaping the distribution of EFL and EF1-α.

In summary, the process of HGT and functional replacement of paralogs was simulated in a step-by-step fashion, which allowed to directly demonstrate that two relatively divergent variants of essential proteins can be coexpressed in vivo. A trouble-free simultaneous expression represents a necessary assumption of scenarios invoking deep paralogy and differential losses to explain the complex distribution of paralogs in the eukaryotic tree. Our experiments thus increase plausibility of this scenario for both EF-1α/EFL and MAT/MATX paralog pairs. Unlike EFL, MATX exhibits also a natural capability to spread among eukaryotes by horizontal transfer and instantaneous functional replacement indicating that also this mechanism might play a role in the evolutionary history of this particular paralog pair.

Supplementary Materials

Supplementary figures S1–S8 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We would like to acknowledge Gabino Sanchez-Perez and Dion G. Durnford for kindly providing the *E. gracilis* MATX clone and Bryony Williams and Patrick Keeling for sharing the *I. galbana* EFL clone and Zuzana Vavrova (Biology Centre) for help with some experiments. This work was supported by the Czech Science Foundation 204/09/1667 (to J.L.), the Czech Science Foundation P506/11/1320 (to V.H.), the Grant Agency of the Charles University 63409 (to J.S.), the Ministry of Education of the Czech Republic (2B06129, 6007665801 and 0021620828)(to J.L. and V.H.), and the Praemium Academiae award (to J.L).

References

Andersen GR, Nissen P, Nyborg J. 2003. Elongation factors in protein biosynthesis. *Trends Biochem Sci*. 28:434–441.

Andersson JO. 2005. Lateral gene transfer in eukaryotes. *Cell Mol Life Sci*. 62:1182–1197.

Andersson JO. 2006. Genome evolution of anaerobic protists: metabolic adaptation via gene acquisition. In: Katz LA, Bhattacharya D, editors. Genomics and evolution of microbial eukaryotes. Oxford: Oxford University press. p. 109–122.

Andersson JO. 2009. Horizontal gene transfer between microbial eukaryotes. In: Gogarten MB, Gogarten JP, Olendzenski L, editors. Horizontal gene transfer. New York: Humana Press. p. 473–487.

Andersson JO, Sjogren AM, Horner DS, Murphy CA, Dyal PL, Svard SG, Logsdon JM Jr., Ragan MA, Hirt RP, Roger AJ. 2007. A genomic survey of the fish parasite *Spironucleus salmonicida* indicates genomic plasticity among diplomonads and significant lateral gene transfer in eukaryote genome evolution. *BMC Genomics*. 8:51.

Armbrust EV, Berges JA, Bowler C, et al. (45 co-authors). 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306:79–86.

Bock R, Timmis JN. 2008. Reconstructing evolution: gene transfer from plastids to the nucleus. *Bioessays* 30:556–566.

Bowler C, Allen AE, Badger JH, et al. (77 co-authors). 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456:239–244.

Carnes J, Trotter JR, Peltan A, Fleck M, Stuart K. 2008. RNA editing in *Trypanosoma brucei* requires three different editosomes. *Mol Cell Biol*. 28:122–130.

Chiang PK, Gordon RK, Tal J, Zeng GC, Doctor BP, Pardhasaradhi K, McCann PP. 1996. S-adenosylmethionine and methylation. *FASEB J*. 10:471–480.

Chuang SM, Chen L, Lambertson D, Nand M, Kinzy TG, Madura K. 2005. Proteasome-mediated degradation of cotranslationally damaged proteins involves translation elongation factor 1A. *Mol Cell Biol*. 25:403–413.

Cocquyt E, Verbruggen H, Leliaert F, Zechman FW, Sabbe K, De Clerck O. 2009. Gain and loss of elongation factor genes in green algae. *BMC Evol Biol*. 9:39.

Cohen O, Gophna U, Pupko T. 2010. The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer. *Mol Biol Evol*. 21:1643–1660.

Doolittle WF. 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet*. 14:307–311.

Doolittle WF, Bapteste E. 2007. Pattern pluralism and the Tree of Life hypothesis. *Proc Natl Acad Sci U S A*. 104:2043–2049.

Doolittle WF, Papke RT. 2006. Genomics and the bacterial species problem. *Genome Biol*. 7:116.

Dujon B, Sherman D, Fischer G, et al. (67 co-authors). 2004. Genome evolution in yeasts. *Nature* 430:35–44.

Esser C, Ahmadinejad N, Wiegand C, et al. (15 co-authors). 2004. A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol Biol Evol*. 21:1643–1660.

Gabaldón T, Huynen MA. 2004. Shaping the mitochondrial proteome. *Biochim Biophys Acta*. 1659:212–220.

Gile GH, Faktorová D, Castlejohn CA, Burger G, Lang BF, Farmer MA, Lukeš J, Keeling PJ. 2009. Distribution and phylogeny of EFL and EF-1α in Euglenozoa suggest ancestral co-occurrence followed by differential loss. *PLoS One*. 4:e5162.

Gile GH, Novis PM, Cragg DS, Zuccarello GC, Keeling PJ. 2009. The distribution of Elongation factor-1 alpha (EF-1alpha), Elongation factor-like (EFL), and a non-canonical genetic code in the ulvophyceae: discrete genetic characters support a consistent phylogenetic framework. *J Eukaryot Microbiol*. 56:367–372.

Gile GH, Patron NJ, Keeling PJ. 2006. EFL GTPase in cryptomonads and the distribution of EFL and EF-1alpha in chromalveolates. *Protist* 157:435–444.

Gogarten JP. 2003. Gene transfer: gene swapping craze reaches eukaryotes. *Curr Biol*. 13:R53–R54.

Ho P, Kong KF, Tang JSH, Wong JTY. 2007. An unusual S-adenosylmethionine synthetase gene from dinoflagellate is methylated. *BMC Mol Biol*. 8:87.

Kamikawa R, Inagaki Y, Sako Y. 2008. Direct phylogenetic evidence for lateral transfer of elongation factor-like gene. *Proc Natl Acad Sci U S A*. 105:6965–6969.

Kamikawa R, Sakaguchi M, Matsumoto T, Hashimoto T, Inagaki Y. 2010. Rooting for the root of elongation factor-like protein phylogeny. *Mol Phylogenet Evol*. 56:1082–1088.

Kamikawa R, Sanchez-Perez GF, Sako Y, Roger AJ, Inagaki Y. 2009. Expanded phylogenies of canonical and non-canonical types of methionine adenosyltransferase reveal a complex history of these gene families in eukaryotes. *Mol Phylogenet Evol*. 53: 565–570.

Kamikawa R, Yabuki A, Nakayama T, Ishida K, Hashimoto T, Inagaki Y. 2011. Cercozoa comprises both EF-1α-containing and EFL-containing members. *Eur J Protistol*. 47:24–28.

Keeling PJ. 2009. Functional and ecological impacts of horizontal gene transfer in eukaryotes. *Curr Opin Genet Dev*. 19: 613–619.

Keeling PJ, Inagaki Y. 2004. A class of eukaryotic GTPase with a punctuate distribution suggesting multiple functional replacements of translation elongation factor 1alpha. *Proc Natl Acad Sci U S A*. 101:15380–15385.

Keeling PJ, Palmer JD. 2008. Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet*. 9:605–618.

Kondrashov FA, Koonin EV, Morgunov IG, Finogenova TV, Kondrashova MN. 2006. Evolution of glyoxylate cycle enzymes in metazoa: evidence of multiple horizontal transfer events and pseudogene formation. *Biol Direct*. 1:31.

Kurland CG, Andersson SG. 2000. Origin and evolution of the mitochondrial proteome. *Microbiol Mol Biol Rev*. 64:786–820.

Liu G, Grant WM, Persky D, Latham VM Jr, Singer RH, Condeelis J. 2002. Interactions of elongation factor 1alpha with F-actin and beta-actin mRNA: implications for anchoring mRNA in cell protrusions. *Mol Biol Cell*. 13:579–592.

Long S, Jirků M, Mach J, Ginger ML, Sutak R, Richardson D, Tachezy J, Lukeš J. 2008. Ancestral roles of eukaryotic frataxin: mitochondrial frataxin function and heterologous expression of hydrogenosomal *Trichomonas* homologs in trypanosomes. *Mol Microbiol*. 69:94–109.

Martin W, Schnarrenberger C. 1997. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. *Curr Genet*. 32:1–18.

Noble GP, Rogers MB, Keeling PJ. 2007. Complex distribution of EFL and EF-1alpha proteins in the green algal lineage. *BMC Evol Biol*. 7:82.

Oborník M, Green BR. 2005. Mosaic origin of the heme biosynthesis pathway in photosynthetic eukaryotes. *Mol Biol Evol*. 22: 2343–2353.

Papke RT. 2009. A critique of prokaryotic species concepts. In: Gogarten MB, Gogarten JP, Olendzenski L, editors. Horizontal gene transfer. New York: Humana Press. p. 379–395.

Ricard G, McEwan NR, Dutilh BE, et al. (17 co-authors). 2006. Horizontal gene transfer from bacteria to rumen ciliates indicates adaptation to their anaerobic, carbohydrate-rich environment. *BMC Genomics*. 7:22.

Richards TA, Dacks JB, Campbell SA, Blanchard JL, Foster PG, McLeod R, Roberts CW. 2006. Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements. *Eukaryot Cell*. 5:1517–1531.

Ruiz-Trillo I, Lane CE, Archibald JM, Roger AJ (17 co-authors). 2006. Insights into the evolutionary origin and genome architecture of the unicellular opisthokonts *Capsaspora owczarzaki* and *Sphaeroforma arctica*. *J Eukaryot Microbiol*. 53:379–384.

Sakaguchi M, Takishita K, Matsumoto T, Hashimoto T, Inagaki Y. 2009. Tracing back EFL gene evolution in the cryptomonads–haptophytes assemblage: separate origins of EFL genes in haptophytes, photosynthetic cryptomonads, and goniomonads. *Gene* 441:126–131.

Sanchez-Perez GF, Hampl V, Simpson AGB, Roger AJ. 2008. A new divergent type of eukaryotic methionine adenosyltransferase is present in multiple distantly related secondary algal lineages. *J Eukaryot Microbiol*. 55:374–381.

Stairs CW, Roger AJ, Hampl V. 2011. Eukaryotic pyruvate formate lyase and its activating enzyme were acquired laterally from a firmicute. *Mol Biol Evol*. 28:2087–2099.

Watkins RF, Gray MW. 2006. The frequency of eubacterium-to-eukaryote lateral gene transfers shows significant cross-taxa variation within amoebozoa. *J Mol Evol*. 63:801–814.

Welch RA, Burland V, Plunkett G 3rd, et al. (19 co-authors). 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A*. 99:17020–17024.

Whitaker JW, McConkey GA, Westhead DR. 2009. Prediction of horizontal gene transfers in eukaryotes: approaches and challenges. *Biochem Soc Trans*. 37:792–795.