

A Complex Distribution of Elongation Family GTPases EF1A and EFL in Basal Alveolate Lineages

Kirill V. Mikhailov^{1,2}, Jan Janouškovec³, Denis V. Tikhonenkov^{3,4}, Gulnara S. Mirzaeva⁵, Andrei Yu. Diakin⁶, Timur G. Simdyanov⁷, Alexander P. Mylnikov⁴, Patrick J. Keeling^{3,*}, and Vladimir V. Aleoshin^{1,2,8,*}

¹Belozersky Institute for Physico-Chemical Biology, Lomonosov Moscow State University, Russian Federation

²Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russian Federation

³Botany Department, University of British Columbia, Vancouver, British Columbia, Canada

⁴Institute for Biology of Inland Waters, Russian Academy of Sciences, Borok, Yaroslavl Province, Russian Federation

⁵Institute of Gene Pool of Plants and Animals, Uzbek Academy of Sciences, Tashkent, Republic of Uzbekistan

⁶Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

⁷Faculty of Biology, Lomonosov Moscow State University, Russian Federation

⁸National Research Institute of Physiology, Biochemistry, and Nutrition of Farm Animals, Russian Academy of Agricultural Sciences, Borovsk, Kaluga Region, Russian Federation

*Corresponding author: E-mail: pkeeling@mail.ubc.ca; aleoshin@genebee.msu.su.

Accepted: August 25, 2014

Data deposition: Elongation factor-1 alpha and the related GTPase EF-like described in this article have been deposited at DDBJ/EMBL/GenBank under the accessions KF997847–KF997856.

Abstract

Translation elongation factor-1 alpha (EF1A) and the related GTPase EF-like (EFL) are two proteins with a complex mutually exclusive distribution across the tree of eukaryotes. Recent surveys revealed that the distribution of the two GTPases in even closely related taxa is frequently at odds with their phylogenetic relationships. Here, we investigate the distribution of EF1A and EFL in the alveolate supergroup. Alveolates comprise three major lineages: ciliates and apicomplexans encode EF1A, whereas dinoflagellates encode EFL. We searched transcriptome databases for seven early-diverging alveolate taxa that do not belong to any of these groups: colpodellids, chromerids, and colponemids. Current data suggest all seven are expected to encode EF1A, but we find three genera encode EFL: *Colpodella*, *Voromonas*, and the photosynthetic *Chromera*. Comparing this distribution with the phylogeny of alveolates suggests that EF1A and EFL evolution in alveolates cannot be explained by a simple horizontal gene transfer event or lineage sorting.

Key words: Elongation Factors, Alveolata, EF1A, EFL, Colpodellids, Chromerids, Colponemids.

Introduction

Translation elongation factor-1 a (EF1A) is a key component of protein synthesis in eukaryotes, recruiting charged tRNAs to the ribosome (and is homologous to archaeal EF1A and bacterial EF-Tu, which fulfill the same role). Despite this essential function, however, widespread genomic analysis of diverse eukaryotes has shown a significant number of lineages lack EF1A, and in all such cases encode a second related subfamily of GTPase called EF-like (EFL) (Gile et al. 2006; Ruiz-Trillo et al. 2006; Noble et al. 2007; Gile and Keeling 2008; Gile, Faktorová, et al. 2009; Sakaguchi et al. 2009; Kamikawa et al. 2011). EFL is assumed to fulfill the essential function

of EF1A because of their mutually exclusive distribution, and because EFL retains the main binding sites of functional significance (Keeling and Inagaki 2004).

The distribution of EF1A and EFL has been difficult to explain by any simple model since it was first discovered. Neither protein is restricted to a group of closely related lineages, and instead both proteins are widely scattered among different subgroups in the tree of eukaryotes. This complex distribution led initial reports to question whether the current pattern was due to ancient paralogy and lineage sorting, more recent

horizontal gene transfer (HGT), or a combination of both. HGT has become the favored explanation (Keeling and Inagaki 2004; Kamikawa et al. 2008; Sakaguchi et al. 2009), although in at least one case the pattern is more consistent with lineage sorting (Gile, Faktorová, et al. 2009). Distinguishing between the two is difficult because the great diversity of taxa involved and ancient time scales of the events, which both contribute to insufficient resolution in the phylogenies to unequivocally document cases of HGT or lineage sorting (Cocquyt et al. 2009; Kamikawa et al. 2010). EFL has now been found in lineages from all major eukaryotic supergroups, and its overall distribution has become even more complex (Ruiz-Trillo et al. 2006; Cocquyt et al. 2009; Gile, Faktorová, et al. 2009; Sakaguchi et al. 2009; Cavalier-Smith and Chao 2012; Henk and Fisher 2012; Ishitani et al. 2012; Kamikawa et al. 2013). But more interestingly, deeper analyses into some EFL-containing lineages have shown that the distribution patterns between closely related lineages may also be complex; this is particularly well documented in green algae and euglenozoa, where very unusual distribution patterns conflict with known phylogenetic relationships (Gile, Faktorová, et al. 2009; Gile, Novis, et al. 2009).

In the alveolates, a major eukaryotic supergroup comprising the well-studied lineages ciliates, dinoflagellates, and apicomplexans, only the dinoflagellates and their close relative *Perkinsus* have EFL, whereas all other alveolates have EF1A (Gile et al. 2006). Previous work on EFL-containing taxa rejected the monophyletic origin of EFL gene in *Perkinsus* and dinoflagellates, suggesting independent transfers of EFL gene in closely related groups (Gile et al. 2006). Here, we show that sampling a number of early-diverging alveolate lineages requires the addition of lineage sorting events to this pattern.

Early-Diverging Alveolates Have Either EF1A or EFL

In addition to the three major lineages of alveolates, molecular and morphological data have both shown that photosynthetic chromerids (*Chromera* and *Vitrella*) and predatory colpodellids (*Colpodella*, *Voromonas*, and *Alphamonas*) are basal relatives of the apicomplexans (Kuvardina et al. 2002; Leander et al. 2003; Moore et al. 2008; Janouškovec et al. 2010; Obornik et al. 2012; Gile and Slamovits 2014), and the enigmatic colponemid predators also branch deeply within the alveolates (Janouškovec et al. 2013; Tikhonenkov et al. 2014). We searched transcriptome databases from representatives of all six genera (and in the case of *Acavomonas peruviana* targeted polymerase chain reaction [PCR]) for homologues of EF1A and EFL. Based on the organismal phylogeny and current distribution of the proteins, the expectation would be that all these taxa should encode EF1A and not EFL, but EF1A was only found in colponemids, *Vitrella*, and *Alphamonas*. Surprisingly, EFL was found in *Colpodella*, *Voromonas*, and *Chromera*. In no case were both genes found in the same

taxon suggesting that only one type is expressed or present. No additional EF1A or EFL paralogs were detected in the transcriptome assemblies or the previously prepared DNA library or a small-scale genome sequence survey of *A. peruviana* (Janouškovec et al. 2013). Incidentally, we also found two different elongation factors in bodonids that were used as prey for predatory colpodellids, with *Parabodo caudatus* encoding EF1A and *Procryptobia sorokini* encoding EFL, further supporting previous work on the distribution of the two proteins in kinetoplastids (Gile, Faktorová, et al. 2009).

Phylogenetic analysis places the EFL genes from the early-diverging alveolate lineages in a relatively well-supported clade (fig. 1), indicating a single common origin in these taxa. The relationship of these sequences to other alveolate EFLs from dinoflagellates and their close relative *Perkinsus* was not adequately supported, and they were found to branch very distantly in the tree: the coplodellid/chromerid branch fell much closer to green algae, cryptomonads and fungi, than they did to the dinoflagellates and *Perkinsus*. Using the EFL data, the monophyly of coplodellids/chromerids plus dinoflagellates/*Perkinsus* is rejected by the approximately unbiased (Shimodaira 2002), one sided Kishino–Hasegawa (Kishino and Hasegawa 1989), and Expected Likelihood Weight (Strimmer and Rambaut 2002) tests, but fails to be rejected by the two-sided Kishino–Hasegawa (Kishino and Hasegawa 1989) and Shimodaira–Hasegawa (Shimodaira and Hasegawa 1999) tests. Unlike previous analysis (Gile et al. 2006), the monophyly of *Perkinsus* and dinoflagellates was not rejected by the approximately unbiased and other tests with the taxon sampling used. In EF1A phylogenies, the ciliates have historically been shown to have a confounding covarion effect (Moreira et al. 1999), and so we analyzed this gene with and without ciliates included. In neither case was the alveolates recovered or were many of the relationships between the alveolate subgroups supported, but the *Vitrella* and *Alphamonas* sequences consistently branched at the base of the apicomplexans as one would expect (fig. 2 and supplementary fig. S1, Supplementary Material online). In both analyses, with and without ciliates, colponemid EF1A sequences formed two independent lineages, with *A. peruviana* having a loose association with the branch uniting stramenopiles and *Telonema*, whereas the other colponemid sequences branched elsewhere in the tree with no support (fig. 2 and supplementary fig. S1, Supplementary Material online). However, the monophyly of alveolates is not rejected with the EF1A data by the approximately unbiased and the more liberal tests. Resolution of EFs trees does not increase when the most variable positions are excluded from the alignments (supplementary figs. S2 and S3, Supplementary Material online). Overall, there is no evidence that any unusual evolutionary event such as HGT has affected alveolate EF1A, and the phylogeny of EFL can only be used to make a strong case that the colpodellid/chromerid EFLs arose in common, but whether they arose separately from dinoflagellate and *Perkinsus* EFLs cannot be concluded

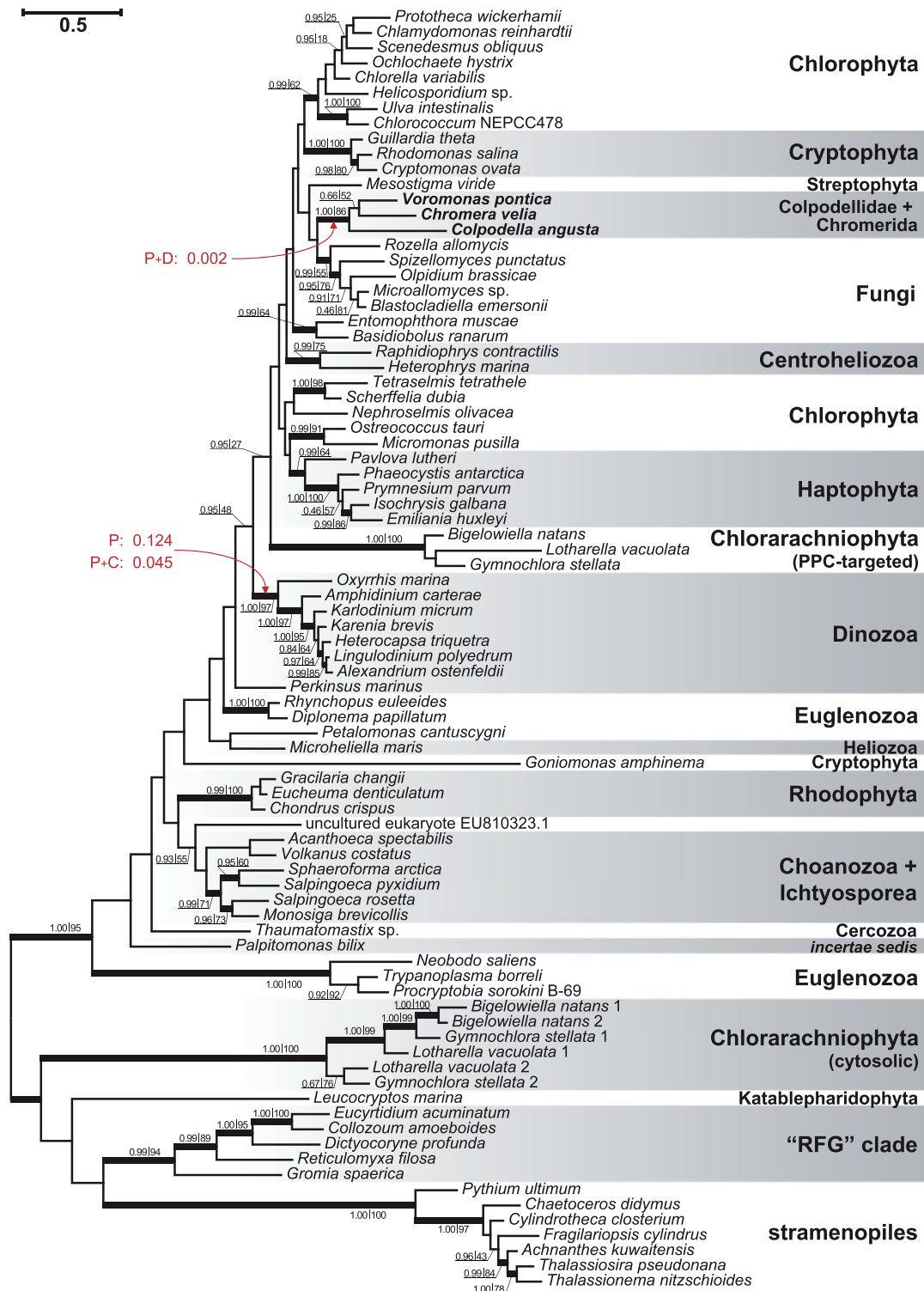


Fig. 1.—Phylogeny of EFL. The tree was reconstructed using Bayesian inference (PhyloBayes) under CAT profile mixture model with four discrete gamma categories and the exchange rates fixed by the LG model (maxdiff = 0.127; loglik effsize = 188). Node support values are given for two types of tree inference methods—Bayesian posterior probability (left) and maximum-likelihood (ML) bootstrap support value (right); bootstrap support was generated on the basis of 1,000 replicates using RAxML and LG+G+F model. Support values for nodes with Bayesian posterior probabilities <0.95 and ML bootstrap support <50% are omitted. Nodes with Bayesian posterior probabilities ≥ 0.95 and ML bootstrap support $\geq 50\%$ are given with thick lines. The “RFG” clade stands for Radiolaria, Foraminifera, and *Gromia*—a tentative group introduced in Ishitani et al. (2012). PPC, periplastid compartment.

Downloaded from <http://gbe.oxfordjournals.org/> at University of British Columbia on January 20, 2016

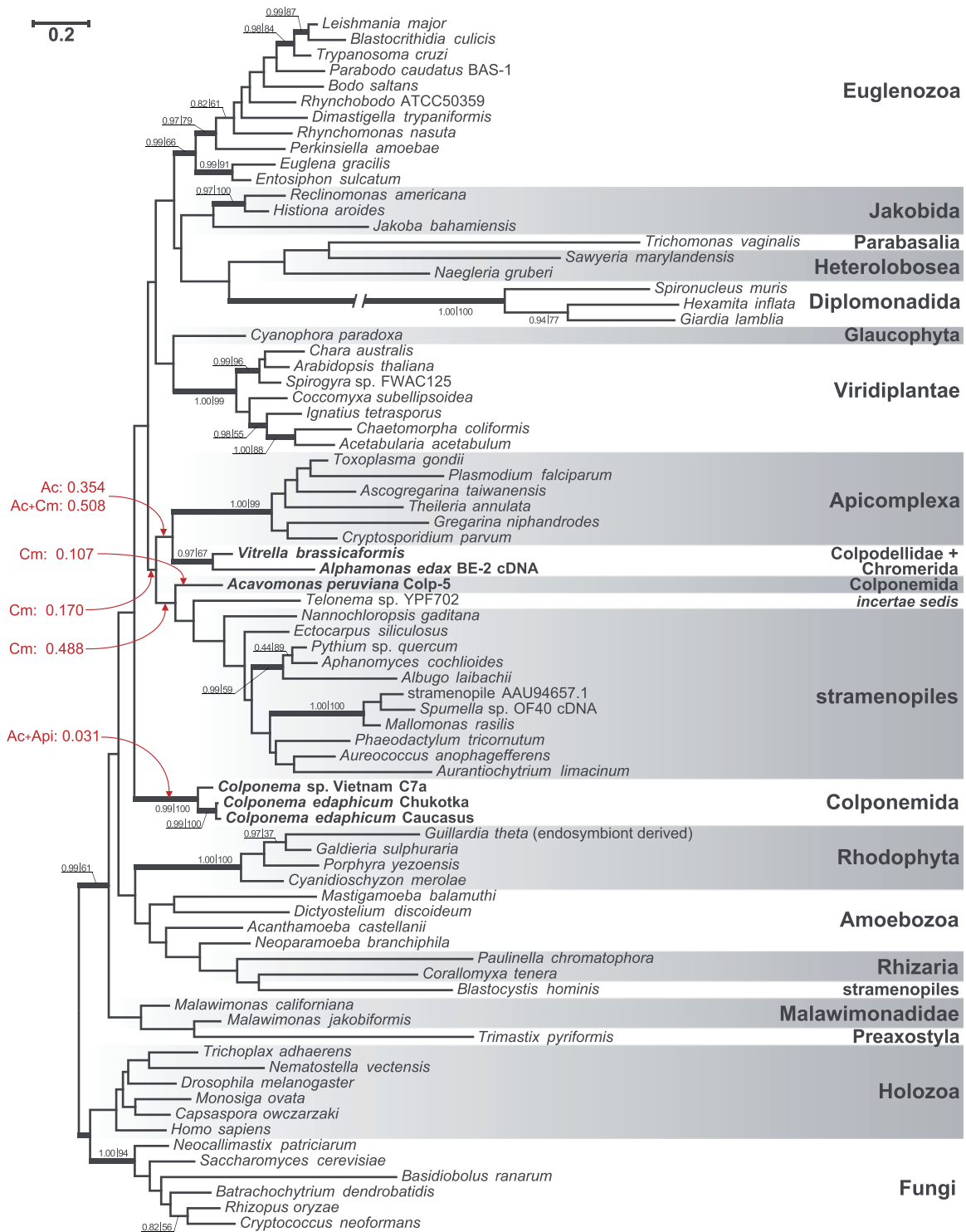


FIG. 2.—Phylogeny of EF1A. The tree was reconstructed using Bayesian inference (PhyloBayes) under CAT profile mixture model with four discrete gamma categories and the exchange rates fixed by the LG model (maxdiff = 0.244; loglik effsize = 113). Node support values are given for two types of tree inference methods—Bayesian posterior probability (left) and maximum-likelihood (ML) bootstrap support value (right); bootstrap support was generated on the basis of 1,000 replicates using RAxML and LG+G+I model. Support values for nodes with Bayesian posterior probabilities <0.95 and ML bootstrap support <50% are omitted. Nodes with Bayesian posterior probabilities ≥ 0.95 and ML bootstrap support $\geq 50\%$ are given with thick lines. The branch leading to diplomonads, marked with a hatch, is artificially shortened.

Downloaded from <http://gbe.oxfordjournals.org/> at University of British Columbia on January 20, 2016

with strong support (at face value, the trees suggest separate origins).

Alveolates EF1A and EFL Have a Complex Evolutionary History

Plotting the presence and absences of EF1A and EFL on alveolate phylogeny reveals no easy explanation for the current distribution of the translation elongation GTPases (fig. 3). If all alveolate EFLs originated only by HGT, there must have been two (colpodellids and all dinoflagellates), or perhaps even three (colpodellids, *Perkinsus*, and core dinoflagellates) individual events. If the distribution is only due to ancient paralogy and lineage sorting, then both genes must have coexisted for a prolonged period in the lineage leading to apicomplexans, leading to three independent sorting events at a minimum. The reality may lay between these two extremes, with a mix of HGT events followed by a period of redundancy and lineage sorting. This makes sense from a functional standpoint as well, as it seems unrealistic to expect an incoming transferred gene to simply replace its existing analog in an instant, so some period of more or less gradual change of function is not unreasonable, and during this period the functional assignment could likely go either way. The deepest alveolate lineages (colponemids, ciliates) seem to contain only EF1A gene but this circumstance does not lead to a definite conclusion, as both EF1A and EFL genes are present in the outgroup (at stramenopiles and Rhizaria). So the ultimate origin of EFL in alveolates may be HGT, but lineage sorting could still play a role in the current distribution if, for example, the ancestor of apicomplexan and dinoflagellate lineages acquired a copy of EFL by HGT, and the descendent lineages retained one or the other resulting in the pattern seen here. This complexity

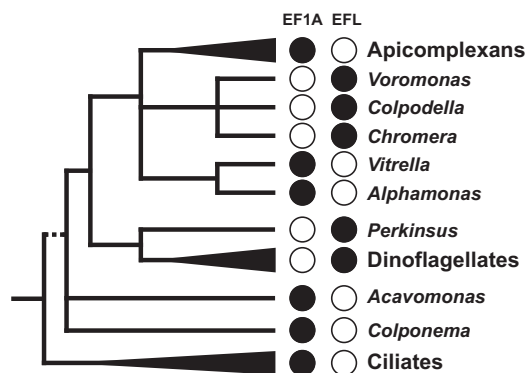


Fig. 3.—Schematic diagram of prospective relationships between the three main alveolate lineages and the early-diverging colponemids, perkinsids, colpodellids, and chromerids. The relationships are based on the rDNA phylogeny (supplementary fig. S4, Supplementary Material online and Gile and Slamovits 2014): Polytomies are unknown, and dotted lines less certain. The presence (filled circle) or absence (open circle) of EF1A and EFL is indicated for each branch.

suggests ancestors of organisms with EF1A (e.g., apicomplexans) may have once also encoded EFL, and underscores the importance of sampling a broad taxonomic diversity when reconstructing such events. The early-diverging alveolate lineages are still poorly studied compared with model organisms from the three main lineages, but without more data from these organisms our understanding of the evolution of the main lineages will remain incomplete.

Materials and Methods

Predatory flagellates *Colpodella angusta* isolate Spi-2 and *Colponema vietnamica* isolate Colp-7a were cultured with free-living bodonid prey *Parabodo caudatus* strain BAS-1; *Voromonas pontica* isolate G-3 and *Alphamonas edax* isolate BE-2 were cultured with bodonid *P. sorokini* strain B-69 and heterotrophic chrysophyte *Spumella* sp. isolate OF-40, respectively. EFs from the prey and predator organisms were identified in the isolated prey and mixed transcriptomes generated using SMARTer Pico PCR cDNA Synthesis Kit (Clontech) and Illumina HiSeq sequencing, and assembled in Inchworm (Trinity v. r2012-06-08) using default parameters, according to the pipeline (Keeling et al. 2014). Between 14 and 48 million 100-bp paired-end raw reads were obtained per sample. The number of contigs assembled for each species ranged from 4×10^4 to 1.7×10^5 (of which a minimum was 40,668 contigs in *V. pontica* assembly). Up to 15% of contigs were discarded during filtering for prey contamination and even more are bacterial contaminants that could not be filtered out due to the lack of reference genomes. Predatory flagellates *A. peruviana* isolate Colp-5 and two isolates of *Colponema edaphicum* (Chukotka and Caucasus) were cultured with *P. sorokini* strain B-69 and *Spumella* sp. isolate OF-40, respectively. Alveolate EF1A was generated by PCR using Encyclo PCR kit (Evrogen) and a pair of degenerate primers (5'-GTTAARTAYGCNTGGGTNYTNGA-3', 5'-ATRTGVGMIGTRTGR CARTC-3'), and sequenced directly on Applied Biosystems 3730 DNA Analyzer. No EFL sequences of *A. peruviana* or *C. edaphicum* were detected by PCR or found in their transcriptomes. Sequences obtained in this study were deposited in GenBank with accession numbers KF997847–KF997856. EFs genes from *Chromera velia* and *Vitrella brassicaformis* were identified in transcriptomes generated through the Marine Microbial Eukaryote Transcriptome Sequencing Project (Gordon and Betty Moore Foundation).

New sequences were translated and aligned with a taxonomically broad sample of EF1A and EFL sequences collected from GenBank (nr, est, wgs), Joint Genome Institute, Broad Institute, and TBestDB databases. Elongation factor sequences of *Nannochloropsis gaditana*, *Pythium ultimum*, and *Galdieria sulphuraria* were extracted in their respective genome project databases (supplementary table S1, Supplementary Material online). The alignments of EF1A and EFL amino acid sequences were prepared separately using MUSCLE alignment program

(Edgar 2004) and manually refined using BioEdit (Hall 1999). After the exclusion of ambiguously aligned positions, the EF1A data set contained 81 sequences and 419 positions, and the EFL data set contained 84 sequences and 452 positions. Tree search for both data sets was performed using the Bayesian method implemented by PhyloBayes 3.3 (Lartillot et al. 2009). Tree reconstruction for both data sets used the CAT profile mixture model with four discrete gamma categories and the exchange rates fixed by the LG model. For each data set, four independent chains were run for 50,000 cycles sampling trees every 100 cycles after discarding the first 10,000 cycles as burn-in. The maximum discrepancy (maxdiff parameter) had values less than 0.3, and the effective sizes (for loglik parameter) ranged from 58 to 188. For the specific parameter values related to individual trees, see the figure captions. The sampled trees were used to generate the majority rule consensus tree with Bayesian posterior probabilities. Bootstrap support values for the consensus tree reconstructed by PhyloBayes were generated using RAxML 7.2.6 (Stamatakis 2006) on the basis of 1,000 replicates under the LG+G+I model for the EF1A data set and LG+G+F model for the EFL data set. The models for each data set were chosen as best-fit by ModelGenerator 0.85 (Keane et al. 2006). The alternative topologies were tested using the CONSEL program (Shimodaira and Hasegawa 2001). The topologies were visualized using TreeView (Page 1996), and site-wise log-likelihood values were computed with TREE-PUZZLE program under the LG+G model (Schmidt 2009).

Supplementary Material

Supplementary tables S1 and figures S1–S4 are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

Acknowledgments

This study used the Supercomputer Center of the Moscow State University (<http://parallel.ru/cluster>) and the Bioportal of University of Oslo (www.bioportal.uio.no) to make phylogenetic inferences. This work was supported by grants from the Russian Foundation for Basic Research to K.V.M., A.P.M., and V.V.A., the Russian Science Foundation to D.V.T., a grant from Ministry of Education and Science of Russian Federation to V.V.A., a grant from GAČR No. GBP505/12/G112 to A.Yu.D., a grant from the Canadian Institutes for Health Research (MOP-42517) to P.J.K., and by the Gordon and Betty Moore Foundation through Grant #2637 to the National Center for Genome Resources. Samples MMETSP0288 and MMETSP0290 were sequenced, assembled, and annotated at the National Center for Genome Resource. J.J. is a Global Scholar and P.J.K. is a Senior Fellow of the Canadian Institute for Advanced Research.

Literature Cited

- Cavalier-Smith T, Chao EE. 2012. *Oxnerella micra* sp. n. (Oxnerellidae fam. n.), a tiny naked centrohelid, and the diversity and evolution of heliozoa. *Protist* 163:574–601.
- Cocquyt E, et al. 2009. Gain and loss of elongation factor genes in green algae. *BMC Evol Biol.* 9:39.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.
- Gile GH, Faktorová D, et al. 2009. Distribution and phylogeny of EFL and EF-1alpha in Euglenozoa suggest ancestral co-occurrence followed by differential loss. *PLoS One* 4:e5162.
- Gile GH, Keeling PJ. 2008. Nucleus-encoded periplastid-targeted EFL in chlorarachniophytes. *Mol Biol Evol.* 25:1967–1977.
- 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.
- Gile GH, Slamovits CH. 2014. Transcriptomic analysis reveals evidence for a cryptic plastid in the colpodellid *Voromonas pontica*, a close relative of chromerids and apicomplexan parasites. *PLoS One* 9:e96258.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser.* 41:95–98.
- Henk DA, Fisher MC. 2012. The gut fungus *Basidiobolus ranarum* has a large genome and different copy numbers of putatively functionally redundant elongation factor genes. *PLoS One* 7:e31268.
- Ishitani Y, et al. 2012. Evolution of elongation factor-like (EFL) protein in Rhizaria is revised by radiolarian EFL gene sequences. *J Eukaryot Microbiol.* 59:367–373.
- Janouškovec J, et al. 2013. Colponemids represent multiple ancient alveolate lineages. *Curr Biol.* 23:2546–2552.
- Janouškovec J, Horák A, Obornik M, Lukeš J, Keeling PJ. 2010. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc Natl Acad Sci U S A.* 107:10949–10954.
- Kamikawa R, et al. 2011. Cercozoa comprises both EF-1alpha-containing and EFL-containing members. *Eur J Protistol.* 47:24–28.
- Kamikawa R, et al. 2013. Parallel re-modeling of EF-1alpha function: divergent EF-1alpha genes co-occur with EFL genes in diverse distantly related eukaryotes. *BMC Evol Biol.* 13:131.
- 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.
- Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McLnerney JO. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol Biol.* 6:29.
- Keeling PJ, et al. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* 12:e1001889.
- Keeling PJ, Inagaki Y. 2004. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1alpha. *Proc Natl Acad Sci U S A.* 101: 15380–15385.
- Kishino H, Hasegawa M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J Mol Evol.* 29:170–179.

- Kuvarina ON, et al. 2002. The phylogeny of colpodellids (Alveolata) using small subunit rRNA gene sequences suggests they are the free-living sister group to apicomplexans. *J Eukaryot Microbiol.* 49: 498–504.
- Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25:2286–2288.
- Leander BS, Kuvarina ON, Aleshin VV, Mylnikov AP, Keeling PJ. 2003. Molecular phylogeny and surface morphology of *Colpodella edax* (Alveolata): insights into the phagotrophic ancestry of apicomplexans. *J Eukaryot Microbiol.* 50:334–340.
- Moore RB, et al. 2008. A photosynthetic alveolate closely related to apicomplexan parasites. *Nature* 451:959–963.
- Moreira D, Le Guyader H, Philippe H. 1999. Unusually high evolutionary rate of the elongation factor 1 alpha genes from the Ciliophora and its impact on the phylogeny of eukaryotes. *Mol Biol Evol.* 16:234–245.
- 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.
- Obornik M, et al. 2012. Morphology, ultrastructure and life cycle of *Vitrella brassicaformis* n. sp., n. gen., a novel chromerid from the Great Barrier Reef. *Protist* 163:306–323.
- Page RDM. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci.* 12:357–358.
- Ruiz-Trillo I, Lane CE, Archibald JM, Roger AJ. 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.
- Schmidt HA. 2009. Testing tree topologies. In: Lemey P, Salemi M, Vandamme AM, editors. *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*, 2nd ed. Cambridge: Cambridge University Press. p. 381–404.
- Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst Biol.* 51:492–508.
- Shimodaira H, Hasegawa M. 1999. Multiple comparisons of loglikelihoods with applications to phylogenetic inference. *Mol Biol Evol.* 16: 1114–1116.
- Shimodaira H, Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246–1247.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
- Strimmer K, Rambaut A. 2002. Inferring confidence sets of possibly misspecified gene trees. *Proc R Soc Lond B Biol Sci.* 269: 137–142.
- Tikhonenkov DV, et al. 2014. Description of *Colponema vietnamica* sp. n. and *Acavomonas peruviana* n. gen. n. sp., two new alveolate phyla (Colponemidia nom. nov. and Acavomonidia nom. nov.) and their contributions to reconstructing the ancestral state of alveolates and eukaryotes. *PLoS One* 9:e95467.

Associate editor: John Archibald