

Low Level Genetic Diversity of Opalinid Morphotypes from the Digestive Tract of *Hoplobatrachus rugulosus* (Batrachia, Amphibia) in Thailand

Eleni GENTEKAKI^{1,2,&}, Chitchai CHANTANGSI^{1,&}

¹ Department of Biology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand; ² School of Science, Mae Fah Luang University, Chiang Rai, Thailand; & Both authors contributed equally to this work

Abstract. Opaline is an unusual group of protists, characterized by the presence of flagella covering their whole body. They reside in the intestinal tracts of various animals, most notably amphibians. While there is a wealth of data regarding the morphological features of opalines, molecular data are extremely sparse. Consequently, the extent of diversity of this understudied group remains unknown. Here, we examine opalines from the intestinal tract of the amphibian *Hoplobatrachus rugulosus* in Thailand. We provide micrographs obtained from light and scanning electron microscopy of various opalinid morphotypes. Furthermore, we enrich the database of opaline sequences by providing new molecular data of the small subunit ribosomal DNA gene of these species. In our phylogenetic analyses, the newly derived sequences form a cluster sister to *Protoopalina*.

Key words: endocommensal, molecular phylogeny, protist, Slopalinida, Thailand

INTRODUCTION

The Opaline taxon comprises multinucleated, osmotrophic eukaryotic microbes characterized by the lack of an oral apparatus and the presence of multiple flagella (Delvinquier and Patterson 1993). Opalines have a worldwide distribution inhabiting the digestive tract of anurans, reptiles and even fishes (Amiet and Affa'a 1985; Delvinquier and Patterson 1993; Delvinquier and Desser 1996; Mohammad *et al.* 2013).

In previous decades, extensive research around the globe led to the discovery and description of hundreds

of opaline species, all of which are classified in five genera: *Cepedea*, *Opalina*, *Protoopalina*, *Protozelleriella*, and *Zelleriella* (Delvinquier *et al.* 1991a, 1991b, 1992, 1995). The presence of only a few opalinid genera indicates that the level of morphological diversity at the genus level is low. At the same time, the existence of a plethora of opalinid species denotes a high morphological diversity at the species level. This raises the question as to whether the level of genetic diversity within opalinid morphotypes corresponds to the species level morphological diversity or not.

In the last decade, a lot of discussion has centered on morphological versus genetic diversity in microbial eukaryotes and how the two might not necessarily be complementary (Lahr *et al.* 2014). Specifically, certain groups of protists have a very high degree of genetic diversity even when using conserved gene markers, such

Address for correspondence: Chitchai Chantangsi, Department of Biology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand; Tel.: 66-2-218-5378; Fax: 66-2-218-5386; E-mail: Chitchai.C@chula.ac.th and chantangsi01@hotmail.com

as the small subunit ribosomal DNA (SSU rDNA). For instance, the diversity of insect trypanosomes is more aptly captured, when using the phylogenetic species concept rather than morphological features (Yurchenko *et al.* 2008; Maslov *et al.* 2013). Another example is that of *Blastocystis*, which belongs to the same clade as the opalines. Though the genus is morphologically static, it is extremely heterogeneous genetically with no less than 17 distinct genotypes, all of which are considered separate species (Alfellani *et al.* 2013).

Herein, we examine the genetic diversity of the SSUrDNA of a prevalent opalinid morphotype from the gut microbiota of the anuran *Hoplobatrachus rugulosus* in Thailand. We illustrate a low degree of intra-morphotype genetic diversity, which corresponds to that revealed by morphological features. Moreover, our molecular data and tree suggest that our sequences correspond to an opalinid that has not been previously sequenced. Furthermore, we present morphological information derived from light and scanning electron microscopy (SEM).

MATERIALS AND METHODS

Isolation of opaline cells and light microscopy

Opalines were isolated from 10 individuals of *Hoplobatrachus rugulosus* obtained from a local market in Bangkok, Thailand (Supplementary Fig. 1). The cloacal regions of the guts of the hosts were removed and placed in a Petri dish containing saline medium as previously described (Yang 1960) with the modification that there was no addition of liver concentrate or inactivated human serum. A longitudinal incision was made and the cloacae were then flipped so that the incision was touching the surface of the Petri dish. This ensured the slow, steady release of the opalines in the medium. Several opalines from as many observed morphotypes as possible were picked with drawn glass Pasteur pipettes and placed in a separate small petri dish filled with the same liquid medium. Subsequently, individual opaline cells were picked one at a time and transferred in depression slides where they were washed multiple times. Pipettes were changed between each wash to minimize transfer of host fecal matter and red blood cells. After the last wash, the cell was placed on a flat slide and observed under an inverted microscope to ensure that no other cells were present. This process was repeated for all isolates.

DNA isolation, amplification of SSU rDNA, cloning and sequencing

After photographing single cells, the cover slip was carefully lifted. The cell was then placed in 10 μ l of lysis buffer and heated at 60°C for 15 min followed by 5 min at 98°C. The universal primers EukA (5'-ACCTGGTTGATCCTGCCAGT-3') and EukB (5'-GATCCTTCTGCAGGTTACCTAC-3') (Medlin *et al.* 1988)

along with the OneTaq[®] Quick-Load[®] 2X Master Mix (New England Biolabs, Ipswich, MA, USA) were used for the initial PCR amplification reaction. Amplification cycle conditions were: 3 min at 95°C, followed by 5 cycles of 95°C for 30 s, 45°C for 60 s and 72°C for 120 s and 35 cycles of 95°C for 30 s, 55°C for 60 s and 72°C for 120 s with a final extension time of 10 min at 72°C. A second nested PCR reaction was carried out with the universal forward primer SSU-525F (5'-AAGTCTGGTGCCAGCAGCC-3'; McCallum and Maden 1985) and the reverse primers 18S-OP1113R (5'-CTACGAGGGTATCTAAGCCTCTTT-3') and 18S-OP1248R (5'-TTGCGACCGTACTCCCCCAT-3'), both of which were designed specifically for this study. Amplification cycle conditions were: 3 min at 95°C, followed by 40 cycles of 95°C for 30 s, 55°C for 60 s and 72°C for 120 s with a final extension time of 10 min at 72°C. The amplicons were purified and inserted into the StrataClone vector (Agilent Technologies, Inc., Santa Clara, CA, USA). Colonies were screened and sequenced bidirectionally with the forward T3 (5'-AATTAACCCTCACTAAAGGGAA-3') and reverse T7 (5'-TAATACGACTCACTATAGG-3') primer pair. Only amplicons for isolates OP66, OP90, BC15, BC16, BC17, BC18, BC21, and BC25 were cloned. The rest of the sequences were obtained by direct sequencing of the purified product using the PCR amplification primers.

In total, 15 isolates were sequenced. All new sequences have been deposited in GenBank under the accession numbers MG662426-MG662440.

Sequence and phylogenetic analysis

The Kimura two-parameter model (Kimura 1980) of base substitution was employed to evaluate gene sequence divergence. Positions with gaps and missing data were not considered. The analyses were conducted in MEGA5 (Tamura *et al.* 2011). Sixty-three sequences of the SSU rDNA spanning the breadth of the stramenopile lineages were aligned with MAFFT v. 6.903b (Katoh *et al.* 2002). Ambiguous regions were removed with trimAl v. 1.3 (Capella-Gutierrez *et al.* 2009) available online through Phylemon 2.0 (<http://www.phylemon.bioinfo.cipf.es>; Sanchez *et al.* 2011). Following masking, the alignment contained 1,915 sites. Initially, we included all available slopalinid sequences in our analysis. However, the extremely short size of the *Cepedea* (223–225 bp) and *Opalina* (229 bp) sequences along with the lack of barely any overlap with the rest of the opalines resulted in erroneous phylogenies. Thus, we removed these sequences and did not consider them further. The RAxML program v. 8 (Stamatakis 2006) was used to construct the Maximum Likelihood (ML) tree with the General Time Reversible (GTR) + Γ model of nucleotide substitution. The heuristic tree search was based on 20 starting trees and 500 bootstrap replicates were generated and used for the ML analysis. The MrBayes program v. 3.2.6 (Ronquist and Huelsenbeck 2003) was used to construct the Bayesian inference (BI) tree with the model of nucleotide evolution set to GTR + Γ . Two sets of four independent chains of Markov Chain Monte Carlo (MCMC) simulations were run for 1,500,000 generations with sampling occurring every 1,000 generations and 25% of them were discarded as burn-in. Convergence was confirmed from the standard deviation of split sequences (less than 0.01). All phylogenetic analyses were conducted on the Cipres Science Gateway (<http://www.phylo.org/portal2/home.action>; Miller *et al.* 2010).

Scanning electron microscopy

Cells from all observed morphotypes were placed in a petri dish and washed several times to remove fecal matter adhering to their flagella. After washing, cells were fixed in various glutaraldehyde concentrations ranging (in 0.5% increments) from 0.5 to 3% final concentration at room temperature for one hour. The optimal concentration for the majority of the opaline cells was 2.5%. The opalinids were subsequently washed twice with distilled water and transferred onto 10- μ m polycarbonate membrane filters and dehydrated in a graded series of ethanol. The dehydration process involved two washes of 15 minutes each in each of 10, 30, 50, 70, 80, 90, 95% ethanol and three washes of 20 minutes each in 100% ethanol. The filters containing the cells were dried with liquid CO₂ in a critical point dryer (Tousimis Samdri 780). Subsequently, the filters were mounted on brass stubs, coated with gold particles using a sputter coater and examined under a JSM-5410 LV scanning electron microscope.

RESULTS

The opaline cells for which we obtained photomicrographs differed in both shape and size (Figs 1, 2). The cells of all opalinids were elongated and cylindri-

cal and about 30–175 μ m in width and 150–370 μ m in length. The bodies were densely covered with flagella (Figs 1, 2) with one exception, where the body is flattened (Fig. 1C). The falx, the point from which the flagella radiate is clearly visible (Fig. 1A) and so are the nuclei (Figs 1B, 1C, 1E). The prominent ridges visible in some morphotypes correspond to the metachronal beat patterns of the flagella (Figs 1D, 1E, 2A).

The resulting amplicon lengths were 600 bp for the SSU-525F + 18S-OP1113R primer combination and 700 bp for the SSU-525F + 18S-OP1248R combination. The sequences that we obtained correspond to the morphotypes in Figures 1A, 1D and 2A. We were unable to obtain sequences for the morphotypes depicted in Figs 2B–D. In total, we obtained 15 new sequences. The intraspecific divergence among the 15 isolates was less than 1% (Supplementary Table 1). All newly generated sequences have two insertions to the exclusion of all other taxa in the alignment (Supplementary Fig. 2). Our phylogenetic analysis of SSU rDNA sequences shows that the taxon Opalinata is monophyl-

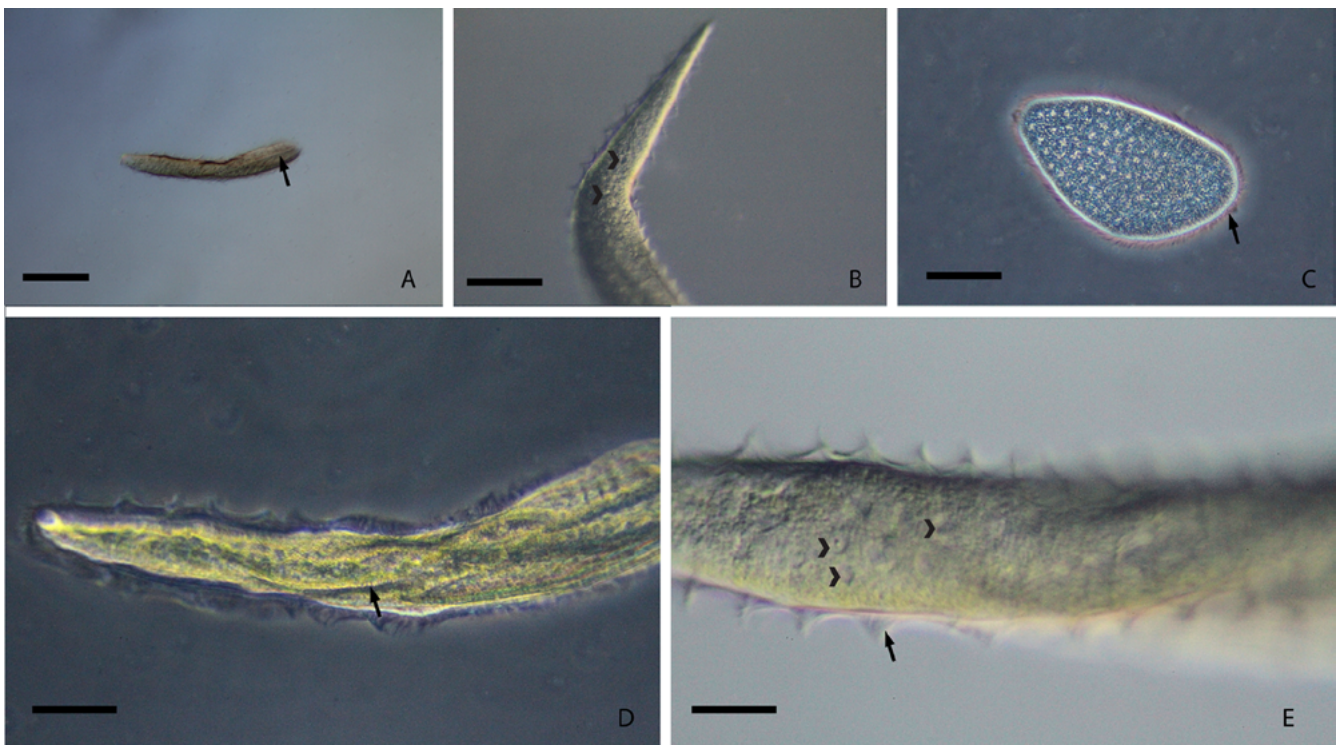


Fig. 1. Light micrographs of opaline cells from life. **(A)** The opaline falx (arrow). **(B)** Multiple nuclei throughout the cell (arrowheads). A bend at the mid-body level is noticeable. **(C)** Flagella covering the body of the opaline cell (arrow). Several scores of nuclei are clearly visible (small circular clear whitish areas). **(D)** Prominent ridges signifying the metachronal beating of the flagella cover the opaline body in a random arrangement (arrow). **(E)** Flagella beating in a metachronal fashion (arrow). Opalinid cell nuclei (arrowheads). Scale bar: A and C: 100 μ m; B, D and E: 50 μ m.

etic with maximum support (100ML/1.0BI) and groups as a sister clade to *Blastocystis* (100ML/1.0BI) (Fig. 3). The newly generated sequences form a cluster, which then groups as sister to the *Protoopalina* spp. (Fig. 3). The Opalinidae is also recovered as monophyletic (78ML/0.99BI). However, the Proteromonadidae is not monophyletic with *Proteromonas* grouping separately from *Karotomorpha*.

DISCUSSION

In this study, we investigated the genetic diversity of a prevalent opalinid morphotype from the intestinal tract of several *H. rugulosus* anurans. We used light and scanning electron microscopy to examine morphological features of the opalines. Furthermore, we used molecu-

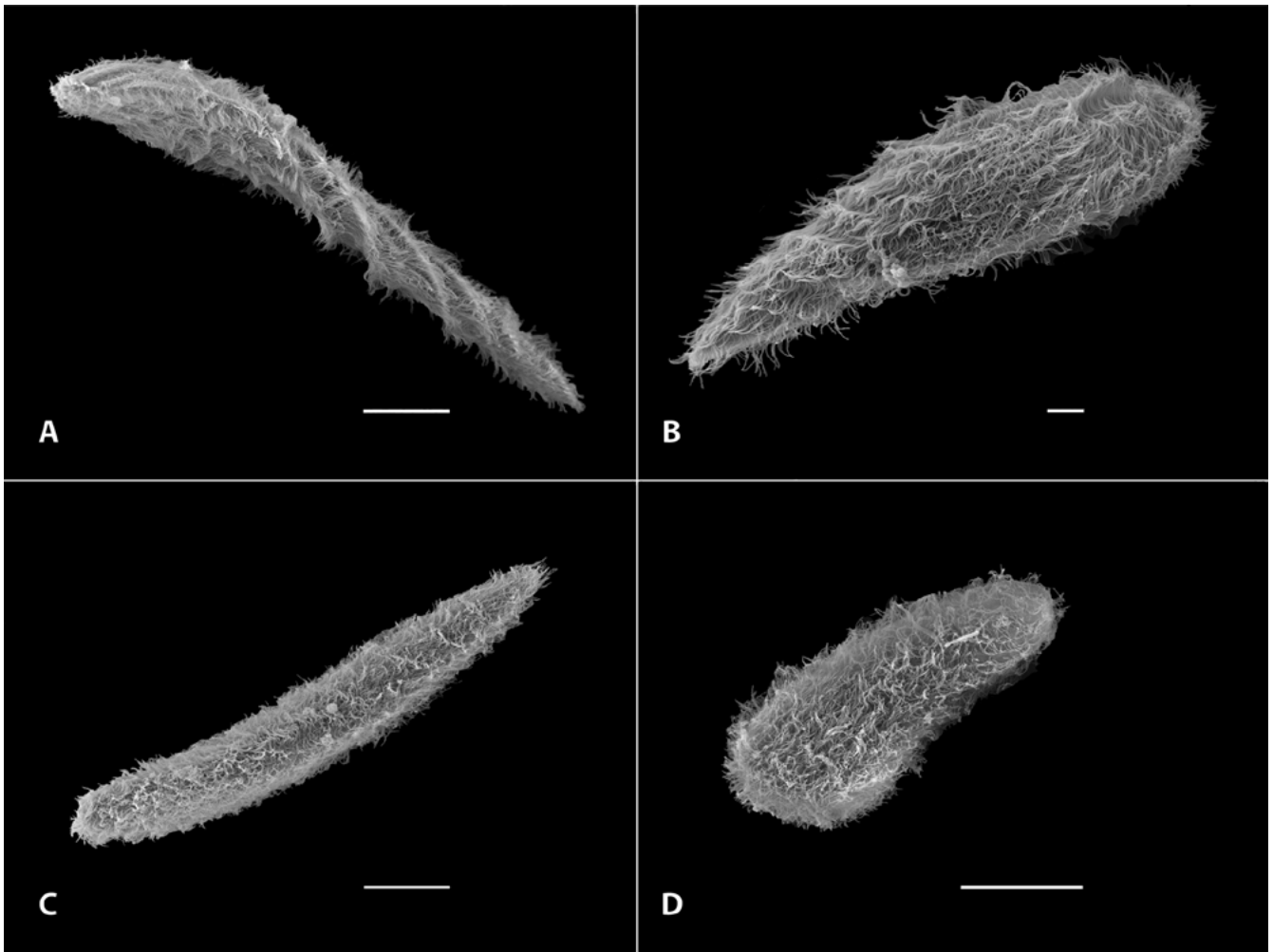
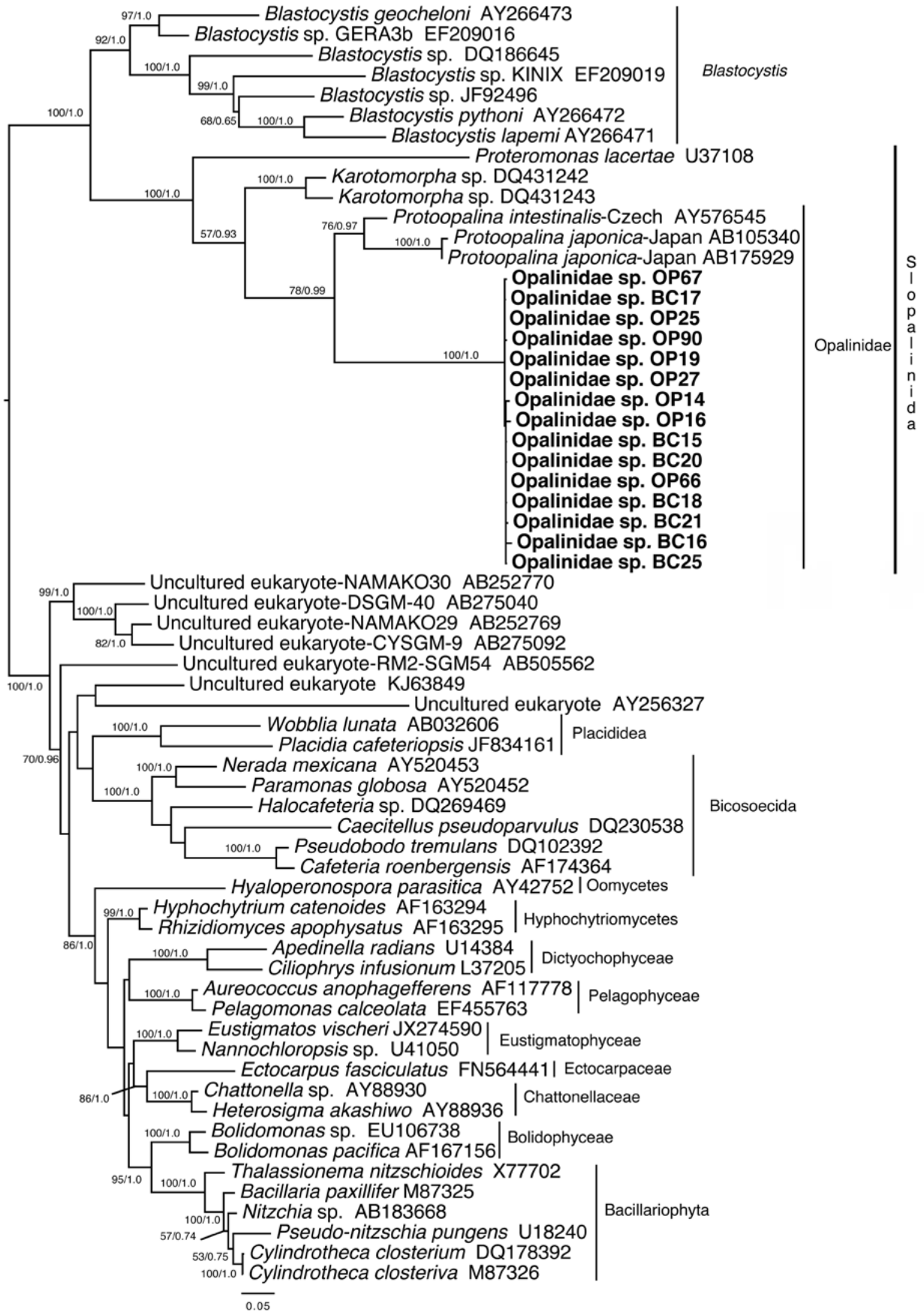


Fig. 2. Scanning electron micrographs of opaline cells fixed in 2.5% glutaraldehyde. All opaline cells are covered with flagella throughout their body. **(A)** Opalinid cell with somatic ridges caused by the flagellar metachronal beating clearly defined in a spiral arrangement. **(B)** Fan-shaped opalinid cell with a broad anterior end and tapering posterior. **(C)** Uniformly elongated opalinid cell with a slightly tapering posterior end. **(D)** Fan-shaped opalinid cell of a similar morphotype to B but of a smaller size. Scale bar: A, C and D: 50 μ m; B: 10 μ m.



Fig. 3. Maximum likelihood (ML) phylogenetic tree of SSU rDNA spanning the breadth of the major lineages of stramenopiles inferred from 63 taxa and 1,915 sites. The General Time Reversible + Γ model of nucleotide substitution was used to construct the tree. New isolates derived from this study are depicted in bold. The ML bootstrap values and Bayesian Inference posterior probabilities are depicted on the nodes in that order. Only values higher than 50 are shown. The scale bar corresponds to 0.05



lar data to assess the biodiversity of the group and enriched the currently existing database of the SSU rDNA sequences of the opalines. In the course of our study, we also identified morphotypes for which we were unable to obtain any molecular data (Fig. 2B–D), despite using several combinations of specific and universal primer pairs and a variety of PCR conditions. This might imply a degree of intra-morphotype genetic heterogeneity that has yet to be described in the Opalinidae.

For our phylogenetic analyses, we had to regrettably omit the available sequences of *Cepedea* and *Opalina*, because of their extremely short length (less than 230 bp). Using these sequences resulted in erroneous and unreliable phylogenetic trees. Thus, the only opalinids used for the analyses are the newly sequenced ones and those belonging to *Protoopalina* spp. Our sequences group as sister to *Protoopalina* spp., though it is not possible to speculate, whether this relationship will hold once more opalinid sequences become available. The branch containing the new sequences is long relative to the *Protoopalina* spp., indicating a large genetic distance.

The published *Protoopalina* sequences belong to two species that come from two separate localities: *P. japonica* comes from Japan and *P. intestinalis* from Czech Republic (Kostka *et al.* 2004; Nishi *et al.* 2005). The genetic distance between these two species is 10%. The distance for the same fragment between our isolates and *P. intestinalis* ranges from 11 to 13%, while that with *P. japonica* ranges from 16 to 18%. If the new sequences are indeed *Protoopalina*, then this would put the interspecific genetic distance to 10–18%. While this is considered high for metazoans, it is not unusual for protists. For instance, *Blastocystis*, which is sister to the slopalinids, comprises several subtypes (species) that show a remarkable degree of inter-subtype genetic heterogeneity of up to 12.9% (Stensvold *et al.* 2012). Nonetheless, there are some notable differences that set apart our sequences from the rest of the *Protoopalina* spp. The aligned fragment of our isolates is longer by about 200 bases, due to insertions in the hypervariable V4 regions. These insertions are unique to our isolates and no other species in the alignment have them. Hypervariable regions are commonly of similar length within the same species (e.g. *Karotomorpha* and *Blastocystis*), however this is not the case between our isolates and the available sequences of *Protoopalina* spp. Collectively, these findings strongly hint that our sequences belong to either a previously non sequenced, highly divergent *Protoopalina* or another genus altogether. The short axial falx running almost parallel to the anteroposterior

axis of the cell, which is characteristic of *Protoopalina*, further support this statement. Nonetheless, the possibility of our isolates being *Cepedea* cannot be excluded either. Examination of additional morphological features and availability of more individuals would aid in definitively identifying our isolates. Regrettably, lack of Protargol impregnation setup at the time and sensitivity of opalinids to oxygen exposure, greatly hampered our efforts.

While the new isolates are phylogenetically distinctive from all others in the tree, they are almost identical to each other as indicated by the short branches and genetic distances, all of which were less than 1%. A possible explanation is that this particular opalinid morphotype is not genetically divergent. Though genetic heterogeneity of very similar morphotypes (i.e. cryptic speciation) is not unusual in protists, there are many cases where morphological and molecular diversity are in agreement. A second explanation could be that the SSU rDNA is too conserved and therefore not an appropriate marker to reveal intra-morphotype differences. Typically, SSU rDNA is not used for lower than species level analyses. Rather, other less conserved markers are preferred, including the internal transcribed spacers (ITS), large subunit rDNA (LSUrDNA) and cytochrome *c* oxidase subunit 1 (*cox1*) (Gentekaki and Lynn 2010; Katz *et al.* 2011; Saez *et al.* 2003). Nonetheless, SSU rDNA has sometimes revealed substantial differences among seemingly identical isolates (Finlay *et al.* 2006; Evans *et al.* 2008). Regrettably, based on our data it is not possible to tell which of the two scenarios holds. Sequencing of more markers and more opalinid morphotypes is needed to shed light to this question.

The data presented in this study come from an opalinid morphotype from an anuran host in Thailand. Though it was not possible to reach a definitive morphological identification, it was obvious from the molecular data that our isolates have not been sequenced previously. Nonetheless molecular data from this group of protists remain scant. Obtaining longer fragments of *Opalina* and *Cepedea*, as well as, sequences from *Protozelleriella* and *Zelleriella* would help greatly in not only discovering more taxonomically relevant characters, but also in disentangling the evolutionary history of this unusual group of protists.

Acknowledgements. E. Gentekaki was supported through C. Chantangsi by the Postdoctoral Fellowship, Rachadapisaek Sompote Fund, Chulalongkorn University (2014). The research was also supported by a grant to C. Chantangsi from Rachadapisaek Sompote Fund for New Faculty, Chulalongkorn University.

REFERENCES

- Alfellani M. A., Taner-Mulla D., Jacob A. S., Imeede C. A., Yoshikawa H., Stensvold C. R., Clark C. G. (2013) Genetic diversity of *Blastocystis* in livestock and zoo animals. *Protist* **164**: 497–509
- Amiet J.-L., Affa'a F.-M. (1985) A propos des strategies d'infestation chez les protozoaires parasites ou endocommensaux des amphibiens anoures du Cameroun. *Rev. Ecol. (Terre Vie)*. **40**: 389–398
- Capella-Gutierrez S., Silla-Martinez J. M., Gabaldon T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973
- Delvinquier B. L. J., Desser S. S. (1996) Opalinidae (Sarcostigophora) in North American Amphibia. Genus *Opalina* Purkinje and Valentin, 1835. *Syst. Parasitol.* **33**: 33–51
- Delvinquier B. L. J., Patterson D. J. (1993) The Opalines. In: Kreier, J. P., Baker, J. R. (Eds), Parasitic Protozoa. Academic Press, San Diego, pp: 247–325
- Delvinquier B. L. J., Markus M. B., Passmore N. I. (1991a) Opalinidae in African Anura I. Genus *Opalina*. *Syst. Parasitol.* **19**: 119–146
- Delvinquier B. L. J., Markus M. B., Passmore N. I. (1991b) Opalinidae in African Anura II. Genera *Protozelleriella* n. g. and *Zelleriella*. *Syst. Parasitol.* **19**: 159–185
- Delvinquier B. L. J., Markus M. B., Passmore N. I. (1992) Opalinidae in African Anura III. Genus *Cepedea*. *Syst. Parasitol.* **24**: 53–80
- Delvinquier B. L. J., Markus M. B., Passmore N. I. (1995) Opalinidae in African Anura IV. Genus *Protoopalina*. *Syst. Parasitol.* **30**: 81–120
- Evans K. M., Wortley A. H., Simpson G. E., Chepurnov V. A., Man D. G. (2008) A molecular systematic approach to explore diversity within the *Sellaphora pupula* species complex (Bacillariophyta). *J. Phycol.* **44**: 215–231
- Finlay B. J., Esteban G. F., Brown S., Fenchel T., Hoef-Emden K. (2006) Multiple cosmopolitan ecotypes within a microbial eukaryote morphospecies. *Protist*. **157**: 377–390
- Gentekaki E., Lynn D. H. (2010) Evidence for cryptic speciation in *Carchesium polypinum* Linnaeus, 1758 (Ciliophora: Peritrichia) inferred from mitochondrial, nuclear and morphological markers. *J. Eukaryot. Microbiol.* **57**: 508–519
- Katoh K., Misawa K., Kuma K. I., Miyata T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**: 3059–3066
- Katz L. A., de Berardinis J., Hall M. S., Kovner A. M., Dunthorn M., Muse S. V. (2011) Heterogeneous rates of molecular evolution among cryptic species of the ciliate morphospecies *Chilodonella uncinata*. *J. Mol. Evol.* **73**: 266–272
- Kimura M. (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**: 111–120
- Kostka M., Hampl V., Cepicka I., Flegr J. (2004) Phylogenetic position of *Protoopalina intestinalis* based on SSU rRNA gene sequence. *Mol. Phylogenet. Evol.* **33**: 220–224
- Lahr D. J. G., Laughinghouse H. D., Oliverio A. M., Gao F., Katz L. A. (2014) How discordant morphological and molecular evolution among microorganisms can revise our notions of biodiversity on Earth. *Bioessays* **36**: 950–959
- Maslov D. A., Votypka J., Yurchenko V., Lukes J. (2013) Diversity and phylogeny of insect trypanosomatids: all that is hidden shall be revealed. *Trends Parasitol.* **29**: 43–52
- McCallum F., Maden B. E. H. (1985) Human 18S ribosomal RNA sequence inferred from DNA sequence. *Biochem. J.* **232**: 725–733
- Medlin L., Elwood H. J., Stickel S., Sogin M. L. (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* **71**: 491–499
- Miller M. A., Pfeiffer W., Schwartz T. (2010) “Creating the CIPRES science gateway for inference of large phylogenetic trees” in proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, LA pp: 1–8
- Mohammad K. N., Badrul M. M., Mohamad N., Zainal-Abidin A. H. (2013) Protozoan parasites of four species of wild anurans from a local zoo in Malaysia. *Trop. Biomed.* **30**: 615–620
- Nishi A., Ishida K., Endoh H. (2005) Reevaluation of the evolutionary position of opalinids based on 18S rDNA and α - and β -tubulin gene phylogenies. *J. Mol. Evol.* **60**: 695–705
- Ronquist F., Huelsenbeck J. P. (2003) MrBayes3, Bayesian inference under mixed models. *Bioinformatics* **19**: 1572–1574
- Saez A. G., Probert I., Geisen M., Quinn P., Young J. R., Medlin L. K. (2003). Pseudo-cryptic speciation in coccolithophores. *Proc. Natl. Acad. Sci. U.S.A.* **100**: 7163–7168
- Sanchez R., Serra F., Tarraga J., Medina I, Carbonell J., Pulido L., de Maria A., Capella-Gutierrez S., Huerta-Cepas J., Gabaldon T., Dopazo J., Dopazo H. (2011) Phylemon 2.0: a suite of web-tools for molecular evolution, phylogenetics, phylogenomics and hypotheses testing. *Nucleic Acids Res.* **39**: W470–W474
- Stamatakis A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690
- Stensvold C. R., Alfellani M., Clark G. C. (2012) Levels of genetic diversity vary dramatically between *Blastocystis* subtypes. *Infect. Genet. Evol.* **12**: 263–273
- Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* **28**: 2731–2739
- Yang W. C. T. (1960) On the continuous culture of opalinids. *J. Parasitol.* **46**: 32
- Yurchenko V., Lukes J., Tesarova M., Jirku M., Maslov D. A. (2008) Morphological discordance of the new trypanosomatid species phylogenetically associated with the genus *Crithidia*. *Protist* **159**: 99–114

Received on 25th May, 2017; revised on 12th December, 2017; accepted on 14th December, 2017

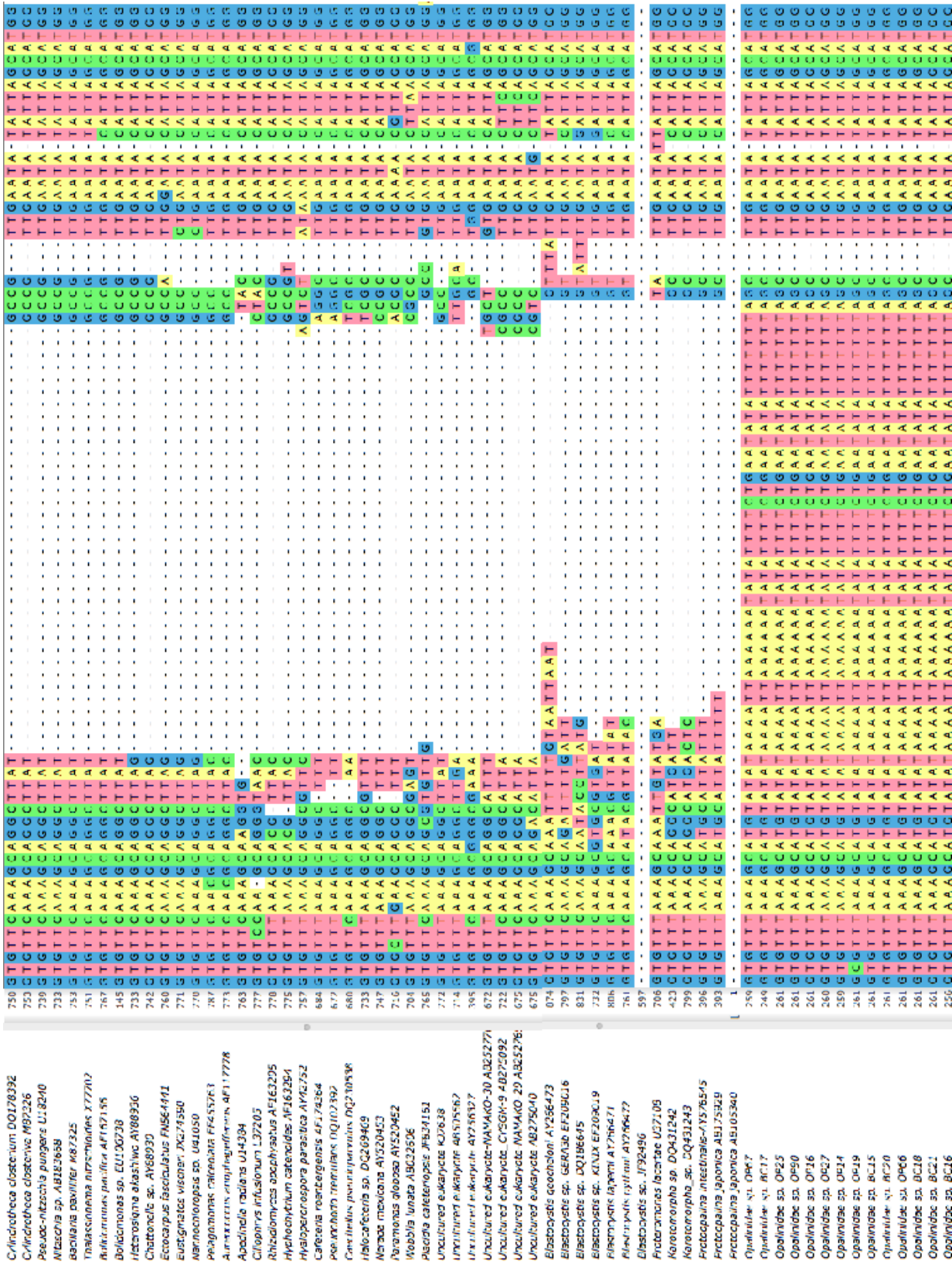
Supplementary



Supplementary Fig. 1. Photograph of the anuran *Hoplobatrachus rugulosus* from which the opalines were sampled.

Supplementary Table 1. Estimates of evolutionary divergence between sequences. Analyses were conducted using the Kimura two-parameter model. Positions containing gaps and missing data were not considered

DQ431243_Karotomorpha_																				
DQ431242_Karotomorpha_	0.04																			
Protoopalina_intestinalis	0.14	0.12																		
Protoopalina_japonica	0.21	0.19	0.10																	
Proteromonas_lacertae	0.23	0.22	0.22	0.30																
OP27	0.19	0.17	0.11	0.16	0.22															
OP16	0.19	0.17	0.11	0.16	0.22	0.00														
BC16	0.21	0.18	0.13	0.18	0.23	0.01	0.01													
OP19	0.20	0.18	0.12	0.17	0.23	0.01	0.01	0.02												
BC17	0.19	0.17	0.12	0.17	0.22	0.00	0.00	0.01	0.01											
OP66	0.19	0.17	0.11	0.16	0.22	0.00	0.00	0.01	0.01	0.00										
BC25	0.19	0.17	0.11	0.16	0.22	0.00	0.00	0.01	0.01	0.00	0.00									
BC15	0.19	0.17	0.12	0.17	0.22	0.00	0.00	0.01	0.01	0.01	0.00	0.00								
BC21	0.20	0.18	0.12	0.17	0.22	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01							
OP67	0.19	0.17	0.12	0.17	0.22	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.01	0.01						
OP90	0.19	0.17	0.12	0.17	0.22	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01					
OP25	0.19	0.17	0.11	0.16	0.22	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.01				
BC18	0.19	0.17	0.12	0.17	0.21	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.00			
BC20	0.19	0.17	0.12	0.17	0.22	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.01	
OP14	0.19	0.17	0.11	0.16	0.22	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00



Supplementary Fig. 2. Alignment of the V4 region of the small subunit rDNA showing the taxon-specific insertion.