Supporting Information

Janouškovec et al.: Plastid dependency and the early evolution of apicomplexans and their free-living relatives.

SI Materials and Methods:

1. Environmental sequencing and data analysis:

Environmental samples were collected in sterile Eppendorfs or Falcon tubes at several locations around Vancouver, BC, Canada (Table S1). DNA was extracted as soon as possible on the day of collection using using 0.5 - 1 g of sample material. Corals were healty specimen sampled at the Birch Aquarium, Scripps Institution of Oceanography, San Diego (records were not available to length of time in captivity or place of origin; photos of each are available from A. T. Howe). Coral mucus was sampled by a 10 ml syringe and concentrated by centrifugation at 10,000 g for 15 min in 1.5 ml Eppendorf tubes. Coral tissue DNA was extracted from a small 0.5 g crushed coral fragment (tissue, skeleton and mucus). DNA from all samples was extracted by the standard protocol of the PowerSoil DNA Isolation Kit (MoBio Laboratories). The V4 and its flanking regions of 18S rDNA was PCR-amplified by using different combinations of the following custom primers sets: Apicomlexa-wide primers: 346F-mix (CADCGACGGGTAACGGGGAATTA; CAGYGACGGGTAACGGGGAATTA; CAGYGACGGGTAACGGGGAATTA; CAGYGACGGGTAACGGGGAATTA) and 785R-mix (IIITATTCCATGCTGIAGTATTCA; IIITATTCCATGCTAAASTATTCA). *Colpodella angusta* group-specific primers: Ctet151 (GAATTCTAGGGCTAATACATCGC); *Alphamonas edax* group-specific primers: Cedax122 (GATGGTYRCTCACTACWTGGATAA); *Voromonas pontica* group-specific primers: Cpont123 (TGGTCACTTACTACATGGATAACCGT); Genotype N group-specific primers: GN123 (TGGCTATATTACTACATGGATAACCG); *Chromera velia*-specific primers: Chrom219 (CCAATCCATCTCTGATGGTWT); *Chromera* and *Vitrella*-specific primers: ChrRMF (TCAMGRCCYGACTTCATGGAAG). PCR products were purified using the Polyethylene glycol (PEG) protocol and cloned using StrataClone PCR Cloning Kit (Agilent) following the instructions of the manufacturer. Clones were PCR-screened using the standard M13F and M13R primers and those containing the corresponding inserts were purified using the PEG protocol and Sanger-sequenced. The total of 78 apicomplexan-related partial 18S rDNA sequences were generated following the removal of duplicates (3), chimeric or low quality sequences (3), and sequences belonging to eukaryotes other than apicomplexans (7 dinoflagellate, 7 ciliate, and 2 rhizarian-like). Among the 78 sequences, 52 branched outside established apicomplexan clades (Figure 1) and 26 branched within them (Figure S2). All sequences were deposited in GenBank under accessions KP213181-KP213258 (Table S1).

Environmental 18S rDNAs sequences were obtained from GenBank by iterative BLASTn searches using both organismal and environmental queries. Hits shorter than 350 nucleotides were removed. Alignment-based check identified several chimaeric entries which were either removed completely (GIs: 309753092, 356459426, 529105829, 529105876, 408773533, 401668321, 213058553) or where the apicomplexan-related part of the sequence was removed (GIs: 89572015, 213058464, 171199596, 355346671, 213058384, 393405443, 60476916, 118420061, 374532687). Several entries that did not well overlap well with the rest of the alignment were removed including 9 sequences from the *Vitrella brassicaformis* clade (Gis: 300088848, 300088897, 300088909, 300088918, 300088921, 300088938, 300088949, 300088950, 300088963). This resulted in 174 environmental sequences from GenBank (generated in 41 studies, 36 published and 9 unpublished; Table S1) that were subsequently combined with the 78 sequences generated by us. The divergent ends of some sequences were trimmed upon alignment. The apicomplexan affiliation of all 252 seqtutences was verified in a Maximum Likelihood phylogeny (ML) using a broad representative sampling of other eukaryotes (those distantly related to alveolates were not used futher). Another ML analysis was used to verify that none of the 226 was specifically related to alveolates with very fast-evolving sequences (gregarines from crustacean hosts, *Trichotokara*, *Lecudina* - *Veloxidinium* clade, *Paralecudina polymorpha*, *Filipodium*, *Selenidium vivax*/ *serpulae*/ cf. e*chinatum*, ellobiopsids; none were used further).

2. Culturing and transcriptome sequencing:

Chromera velia CCMP2878 and *Vitrella brassicaformis* CCMP3155 were cultured in the marine f/2 medium at 22°C. *Voromonas pontica* isolate G-3, *Alphamonas edax* isolate BE-2, and *Colpodella angusta* isolates Spi-2 and BE-6 were cultured on enrichments of the marine bodonid *Procryptobia sorokini* strain B-69, the freshwater heterotrophic chrysophyte *Spumella* sp. isolate OF-40, respectively, and the freshwater bodonid *Parabodo caudatus* strain BAS-1, respectively. Prey organisms were raised on the suspension of Pseudomonas fluorescens (strain ICISC19) in the freshwater Pratt medium or marine Schmalz-Pratt medium at room temperature (see (1) for details). Total RNA was extracted using the RNAqueous kit (Ambion). Chromerid cells were first pelleted, the medium was removed, and the cells were broken by grinding in liquid nitrogen before being resuspend in the Lysis buffer. Colpodellid cells were separated from the medium by a gentle centrifugation (1000 rpm or less) on a 0.7 micrometer filter column and subsequently resuspended in the Lysis buffer. Chromerid transcriptomes were sequenced and assembled as described in Text S1 of Keeling et al. 2014 (2) at http://www.plosbiology.org/article/info%3Adoi%2F10.1371%2Fjournal.pbio.1001889 and the data is available at: http://camera.crbs.ucsd.edu/mmetsp/list.php. Total RNA from colpodellid cultures was reverse-transcribed and amplified using the SMARTer Pico PCR cDNA Synthesis Kit (Clontech). Transcriptome of *Spumella* sp. prey was prepared using the same protocol as those of colpodellids. Transcriptomes were generated by Illumina HiSeq sequencing (100bp paired-end reads; multiplexing 0.25 lane per each of the four samples) producing the following number of raw reads (forward+reverse) 177,414,112 raw reads (*A. edax* BE-2); 50,933,148 (*V. pontica* G-3), 42,068,699 raw reads (*C. angusta* Spi-2), and 53,155,624 raw reads (*C. angusta* BE-6). Cutadapt 1.2.1 was used to trim SMARTer adaptors and 3' Illumina sequencing adaptor (in short fragments). Trimmed reads were assembled in Inchworm (Trinity r2012-06-08) using the default settings (K25 L25) and the minimum contig length of 150 nucleotides. Contamination by *Spumella* prey sequence in *A. edax* transcriptome and cross-contamination between colpodellid transcriptomes (a likely consequence of sequencing all samples in a single lane) was removed by using a custom script as follows: for any two contigs that were ≥ 95% similar between two samples, contig kmer coverages (produced by Inchworm) were compared; if ≥ 50 times difference in coverage was found then the low-coverage contig was discarded as a probable contaminant, otherwise both contigs were discarded. Contigs highly similar (95% or more) to known bacterial sequences were also removed by using a custom script, however, many bacterial contaminants were retained due to the lack of reference sequences. Purified contigs from the four colpodellid transcriptomes were deposited in GenBank Sequence Read Archive. Contigs of the two *C. angusta* strains (Spi-2 and BE-6) were pooled in further analysis after being found to be nearly identical at the nucleotide level (differences if present were typically limited to non-coding terminal regions). *Chromera velia* and *Voromonas pontica* contigs were pooled with those generated in independent studies (3, 4). Open reading frames (ORFs) were six-frame translated into amino acids using GetORF (EMBOSS). DNA surveys from *A. edax* and

C. angusta cells were prepared by Epicentre DNA extraction kit, followed by sequencing and assembly as above.

3. Analysis of plastid dependency:

In order to study distribution of diverse characteristics in apicomplexans and relatives, a custom BLAST database was built. This included protein sequences predicted from the 5 chrompodellid transcriptomes, 11 apicomplexan genomes (*Babesia bovis* T2Bo, *Babesia microti* RI, *Cryptosporidium muris* RN66, *Cryptosporidium parvum* Iowa II, *Eimeria maxima*, *Eimeria tenella* Houghton, *Neospora caninum* LIV, *Plasmodium berghei* ANKA, *Plasmodium falciparum* 3D7, *Plasmodium vivax* SaI1, *Theileria annulata* Ankara, *Theileria parva* Muguga, *Toxoplasma gondii* ME49), the small-scale genome sequence survey of *Ascogregarina taiwanensis*, the unpublished genome draft from *Gregarina niphandrodes*, 11 dinoflagellate transcriptomes (*Alexandrium catenella*, *A. minutum*, *A. ostenfeldii*, *A. tamarense*, *Amphidinium carterae*, *Heterocapsa triquetra*, *Karenia brevis, Karlodinium micrum*, *Lingulodinium polyedrum*, *Oxyrrhis marina*, *Symbiodinium* sp. kb8), the unpublished draft genome of *Perkinsus marinus* ATCC50983 and the draft genome of the dinoflagellate *Symbiodinium minutum*, and 53 genomes from representative eukaryotes (*Acanthamoeba castellanii* Neff, *Albugo laibachii* Nc14, *Allomyces macrogynus* ATCC 38327, *Aplanochytrium kerguelense* PBS07, *Arabidopsis thaliana*, *Aspergillus fumigatus* A1163, *Aureococcus anophagefferens* CCMP1984, *Bathycoccus prasinos* RCC1105, *Bigelowiella natans* CCMP2755, *Blastocystis hominis*, *Calliarthron tuberculosum*, *Capsaspora owczarzaki* ATCC30864, *Chlamydomonas reinhardtii*, *Chlorella variabilis* NC64A, *Chlorokybus atmophyticus* UTEX2591, *Chondrus crispus* PCG, *Coccomyxa subellipsoidea* C169, *Cyanidioschyzon merolae* 10D, *Cyanophora paradoxa*, *Dictyostelium discoideum* AX4, *Ectocarpus siliculosus* Ec32, *Emiliania huxleyi* CCMP1516, *Entamoeba dispar*, *Euglena gracilis*, *Galdieria sulphuraria*, *Giardia intestinalis*, *Glaucocystis nostochinearum*, *Guillardia theta* CCMP2712, *Ichthyophthirius multifilis*, *Leishmania major* Friedlin, *Mesostigma viride*, *Micromonas* sp. RCC299, *Monosiga brevicollis* MX1, *Naegleria gruberi* NEG-M, *Nannochloropsis gaditana* CCMP526, *Ostreococcus* sp. RCC809, *Paramecium tetraurelia*, *Phaeodactylum tricornutum* CCAP1055-1, *Physcomitrella patens patens*, *Porphyridium purpureum*, *Pyropia yezoensis*, *Pythium ultimum* BR144, *Rhizopus oryzae*, *Saprolegnia parasitica* ATCC90214, *Spizelomyces punctatus* DAOM BR117, *Sterkiella histriomuscorum*, *Tetrahymena thermophila*, *Thalassiosira pseudonana* CCMP1335, *Thecamonas trahens* ATCC50062, *Trichomonas vaginalis*, *Trichoplax adhaerens* Grell-BS-1999, *Trypanosoma brucei gambiense*, *Ustilago maydis*). Homologs to eukaryotic queries were retrieved based on BLASTp searches with a variable cutoff. Prokaryotic homologs were retrieved from a custom database (5-10 representatives per phylum) and GenBank (additional sequences most closely related to apicomplexan or chrompodellid queries). Plastid targeting signals were analysed in selected plastid proteins with complete N-termini. Proteins that screened positively for signal peptides in SignalP 4.1 were further tested for the presence of transit peptides in ChloroP 1.1 and listed in Table S3.

4. Analysis of the distribution of apicomplexan-specific proteins:

a) 21 bacterial genes predicted to have been acquired in the apicomplexan ancestor (Fig. 4A) were identified from published literature and during preliminary analyses in this study (Table S5). Out of the 23 bacterial transfers previously identified in *Cryptosporidium* ((5); malate and lactate dehydrogenase represent the same HGT event), 12 were examined here (Table S5). The remaining 11 were either specific to *Cryptosporidium* (EAK89684.1,

AAU15128.1, EAK88388.1, EAK88222.1, AAS47699.1, EAK87293.1, AAL83208.1, EAK87294.1, EAK89739.1, EAK89015.1) or excluded because of the divergent character of the sequences (acetyltransferase, EAK87438.1; a homolog was identified in *Alphamonas*, however, suggesting that neither this gene represents an apicomplexan-specific HGT). Retrieval of eukaryotic and bacterial sequences was done as described above. Alignments, removal of variable sites, RAxML, and PhyML phylogenies were done as described in Materials and Methods (in final alignments shorter than 150 amino acids only aLRT supports were calculated; Fig. S5). Results were summarized in Table S5 and Fig. S5.

b-d) 185 OrthoMCL clusters (OG5 edition; (6)) were downloaded from http://www.orthomcl.org. 527 Apiloc proteins (representing 382 OrthoMCL clusters) were downloaded from http://apiloc.biochem.unimelb.edu.au/apiloc. 173 Wasmuth et al. clusters were downloaded from the Supplement of the paper (7). A query representing each cluster or Apiloc protein (most frequently the sequence from *Toxoplasma gondii*) was then used to search homologs in our eukaryotic database using BLASTp searches, as described in Materials and Methods. All hits were evaluated by eye to prevent misassignment owing to the presence of repeats and amino acid bias both of which are common in secreted proteins. Proteins that contained any regions or domains with significant hits (<1e-5 evalue similarity) to eukaryotes other than apicomplexans or other myzozoans were removed from further searches as 'eukaryotic' (all control proteins localized to Golgi belonged here). Proteins that were limited to certain apicomplexan subgroups only were likewise removed (many Apiloc proteins; Table S5). These strict homology assignments identified 102 protein cluster candidates that appeared to have originated during early apicomplexan or myzozoan evolution. Further in-depth similarity searches by using these proteins as queries and even stricter criteria (all hits 1e-3 to 1e-5 were thoroughly evaluated) led to removal of some of the candidates and produced the final count of 74 protein clusters that appeared specific to either apicomplexans (n=19), apicomplexans and chrompodelids (n=29), or all myzozoans (n=26). Alveolins and oocyst wall proteins (OWPs) clusters identified through this process did not share significant similarity to other eukaryotes, however, they belonged to larger families, and some members of these families shared sequence similary with ciliates (8), and fungi/insects/bacteria (9), respectively. In the case of alveolins, the pattern is consistent with the origin in early alveolates followed by diversification prior to myzozoan divergence (Fig. 4B). In the case of OWPs, the similarity may result from convergent evolution on the same function (9). The distribution of additional proteins (e.g., MyoA, subtilisin proteases, ribonucleotide reductase, proteins lost in apicomplexans) was determined in an analogous fashion and their phylogenies were done as described in Materials and Methods.

SI References:

1. 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(4):e95467.

2. 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(6):e1001889.

3. 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(5):e96258.

4. Woehle C, Dagan T, Martin WF, Gould SB (2011) Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related Chromera velia. *Genome Biol Evol* 3(Botanik III):1220–30.

5. Huang J, et al. (2004) Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in Cryptosporidium parvum. *Genome Biol* 5(11):R88–R88.

6. Chen F (2006) OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res* 34(90001):D363–D368.

7. Wasmuth J, Daub J, Peregrin-Alvarez JM, Finney CAM, Parkinson J (2009) The origins of apicomplexan sequence innovation. *Genome Res* 19:1202–1213.

8. Gould SB, Tham W-H, Cowman AF, McFadden GI, Waller RF (2008) Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. *Mol Biol Evol* 25(6):1219–30.

9. Templeton TJ, et al. (2004) The Cryptosporidium Oocyst Wall Protein Is a Member of a Multigene Family and Has a Homolog in Toxoplasma. *Infect Immun* 72(2):980–987.

10. Seow F, et al. (2005) The plastidic DNA replication enzyme complex of Plasmodium falciparum. *Mol Biochem Parasitol* 141(2):145–153.

11. Sato S, Wilson RJM (2003) Proteobacteria-like ferrochelatase in the malaria parasite. *Curr Genet* 42(5):292–300.

12. Madern D, Cai X, Abrahamsen MS, Zhu G (2003) Evolution of Cryptosporidium parvum Lactate Dehydrogenase from Malate Dehydrogenase by a Very Recent Event of Gene Duplication. *Mol Biol Evol* 21(3):489–497.

13. Kicska GA, et al. (2002) Transition State Analogue Inhibitors of Purine Nucleoside Phosphorylase from Plasmodium falciparum. *J Biol Chem* 277(5):3219–3225.

14. Huang J, Mullapudi N, Sicheritz-Ponten T, Kissinger JC (2004) A first glimpse into the pattern and scale of gene transfer in Apicomplexa. *Int J Parasitol* 34(3):265–74.

15. Hu G, et al. (2010) Transcriptional profiling of growth perturbations of the human malaria parasite Plasmodium falciparum. *Nat Biotechnol* 28(1):91–98.

16. Proellocks NI, et al. (2009) Characterisation of PfRON6, a Plasmodium falciparum rhoptry neck protein with a novel cysteine-rich domain. *Int J Parasitol* 39(6):683–692.

17. Bradley PJ, et al. (2005) Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in Toxoplasma gondii. *J Biol Chem* 280(40):34245–58.

18. Tufet-Bayona M, et al. (2009) Localisation and timing of expression of putative Plasmodium berghei rhoptry proteins in merozoites and sporozoites. *Mol Biochem Parasitol* 166(1):22–31.

19. Van Ooij C, et al. (2008) The Malaria Secretome: From Algorithms to Essential Function in Blood Stage Infection. *PLoS Pathog* 4(6):e1000084.

20. Lamarque MH, et al. (2012) Identification of a New Rhoptry Neck Complex RON9/RON10 in the Apicomplexa Parasite Toxoplasma gondii. *PLoS ONE* 7(3):e32457.

21. Bullen HE, et al. (2009) A Novel Family of Apicomplexan Glideosome-associated Proteins with an Inner Membrane-anchoring Role. *J Biol Chem* 284(37):25353–25363.

22. Possenti A, et al. (2010) Molecular characterisation of a novel family of cysteine-rich proteins of Toxoplasma gondii and ultrastructural evidence of oocyst wall localisation. *Int J Parasitol* 40(14):1639–1649.

23. Schneider I, et al. (2007) Identification, molecular characterization and subcellular localization of a Theileria annulata parasite protein secreted into the host cell cytoplasm. *Parasitol Res* 101(6):1471–1482.

24. Schneider I, Haller D, Seitzer U, Beyer D, Ahmed JS (2004) Molecular genetic characterization and subcellular localization of a putative Theileria annulata membrane protein. *Parasitol Res* 94(6):405–415.

25. Beck JR, et al. (2010) A Novel Family of Toxoplasma IMC Proteins Displays a Hierarchical Organization and Functions in Coordinating Parasite Division. *PLoS Pathog* 6(9):e1001094.

26. Tremp AZ, Khater EI, Dessens JT (2008) IMC1b Is a Putative Membrane Skeleton Protein Involved in Cell Shape, Mechanical Strength, Motility, and Infectivity of Malaria Ookinetes. *J Biol Chem* 283(41):27604–27611.

SI Figure Legends:

Figure S1: Phylogeny of the reference 18S rDNA alignment. Maximum likelihood phylogeny (RAxML; GTR+GAMMA+4 model) of 109 eukaryotic 18S rDNA sequences with sequence gi accessions after the @ symbol. Branch support values correspond to 300 non-parametric bootstraps. Supports ≥60 are shown. Full black circles correspond to complete support (100%).

Figure S2: Phylogeny of environmental sequences generated in this study that cluster with known apicomplexan clades. Maximum likelihood phylogeny (RAxML; GTR+GAMMA+4 model) of eukaryotic and environmental 18S rDNA sequences with non-parametric bootstrap supports (300 replicates) and environment types listed according to the key. Supports ≥60 are shown. Full black circles correspond to complete support (100%). Sequence accessions (gi) are specified following the @ sign.

Figure S3: SufB phylogeny. Maximum likelihood phylogeny (RAxML; LG+GAMMA+4 model) with non-parametric bootstraps (300 replicates) / PhyML aLRT supports at relevant nodes. Upper values correspond to the dataset shown, lower values correspond to dataset with fast-evolving sequences excluded (*Vitrella* and *Colpodella*, indicated by asterisks). The position of apicomplexan plastid-encoded SufB was established in a separate analysis and is shown by a dashed line.

Figure S4: ClpC phylogeny. Maximum likelihood phylogeny of ClpC and the closest ClpB paralogs (RAxML; LG+GAMMA+4 model) with PhyML aLRT supports at relevant nodes. First values correspond to the dataset shown, second values correspond to dataset with fast-evolving sequences excluded (*Chromera* and apicomplexans, indicated by asterisks).

Figure S5: Phylogenies of 13 bacterial proteins in apicomplexans and relatives. Maximum likelihood phylogeny (RAxML; LG+GAMMA+4 model) with non-parametric bootstraps (300 replicates) / PhyML aLRT supports at relevant nodes (dasets shorter than 150 amino acid positions were not bootstrapped). Predicted horizontal gene transfer (HGT) events are highlighted.

Figure S6: Phylogeny of ribonucleotide reductase. Maximum likelihood phylogeny (RAxML; LG+GAMMA+4 model) reveals that the R2e2 type (previously thought to be unique to apicomplexans and not closely related to the R2e1 type) is also found in chrompodellids and originated by lineage-specific duplication from the R2e1 type.

Figure S7: Phylogeny of subtilisin proteases. Maximum likelihood phylogeny (RAxML; LG+GAMMA+4 model) shows that the majority of apicomplexan subtilisin proteases (e.g. SUB1-12 in *Toxoplasma*) share an ortholog in chrompodellids (or other myzozoans).

Table S1: Environmental sequences related to apicomplexans that were used in this study.

Table S2: Abbreviations and full names of proteins and pathways used in Fig. 3.

Table S3: Chrompodellid plastidic proteins with complete N-terminal sequence.

Table S4: Classificiation of 527 proteins from Apiloc (correspoding to 382 OrthoMCL clusters) based on their putative origin.

Table S5: Classification of 21 bacterial genes in apicomplexans and relatives acquired by horizontal gene transfer (HGT).

Table S6: Classification of 74 unique proteins specific to apicomplexans, apicomplexans and chrompodellids, and myzozoans.

Table S7: Domain superfamilies in chrompodellid and dinoflagellate PKSI/FASI proteins identified by NCBI Conserved Domains searches.

Figure S1: Phylogeny of the reference 18S rDNA alignment.

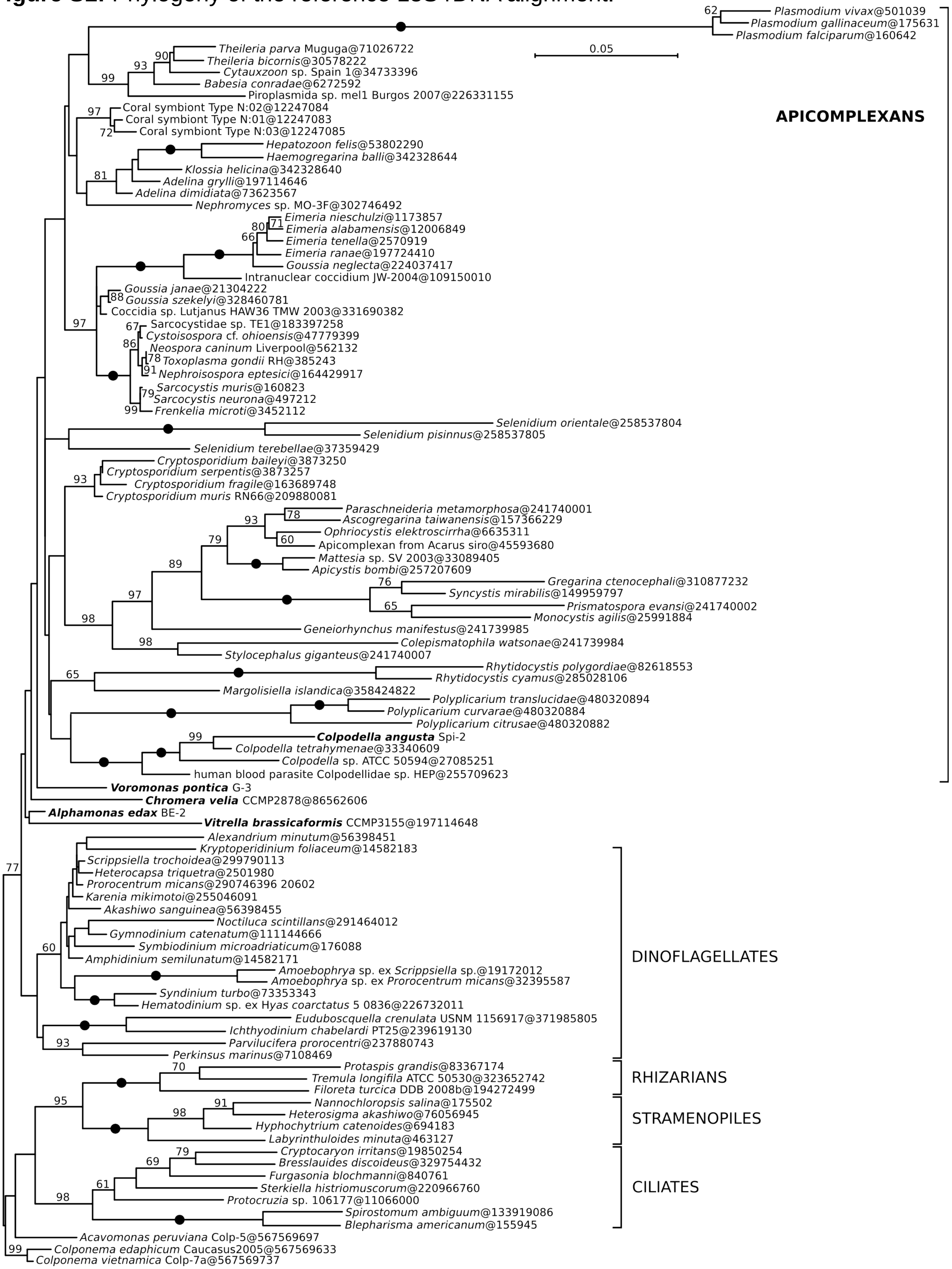


Figure S2: Phylogeny of environmental sequences generated in this study that cluster with known apicomplexan clades.

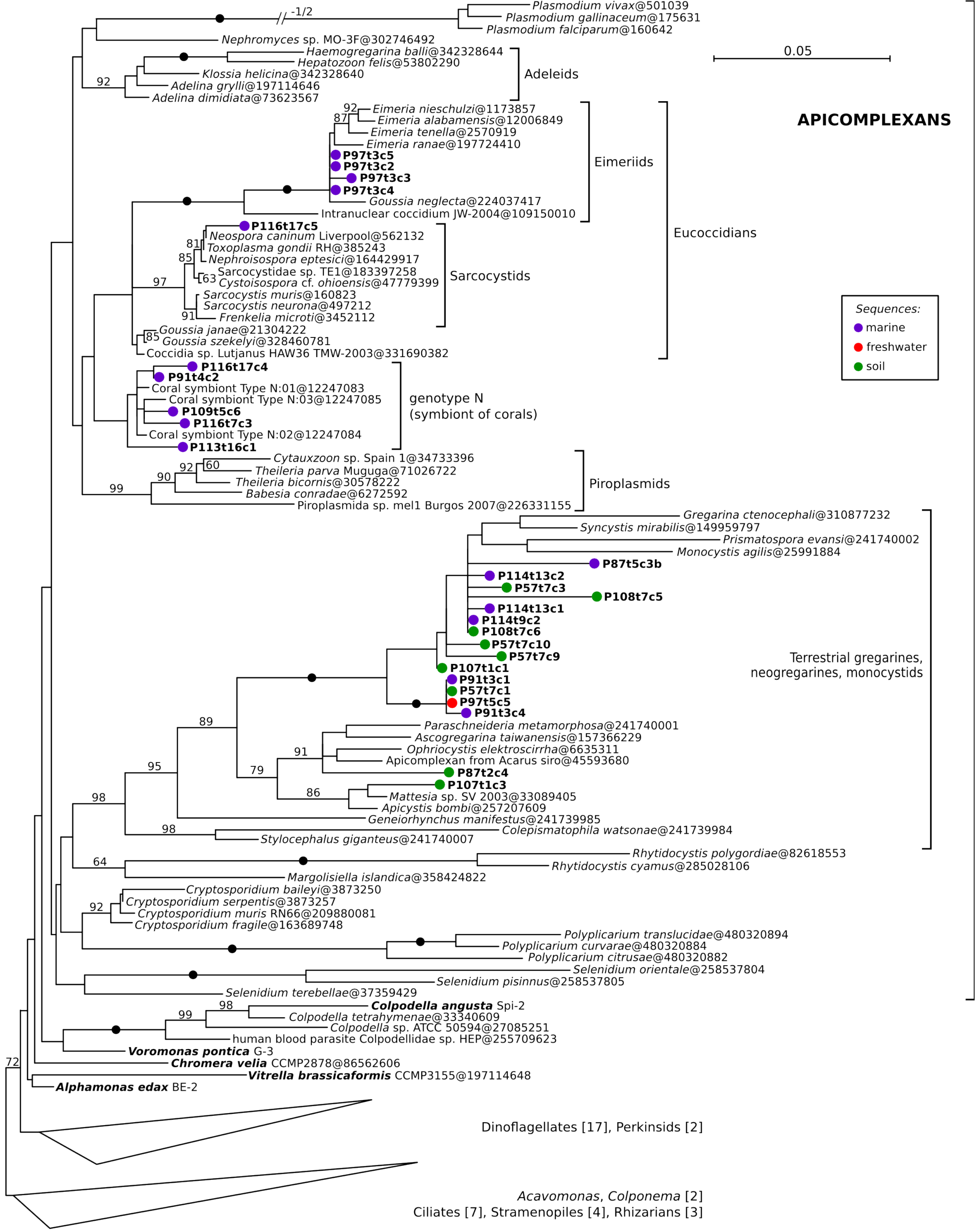


Figure S3: SufB phylogeny.

SufB (445/446 amino acid positions)

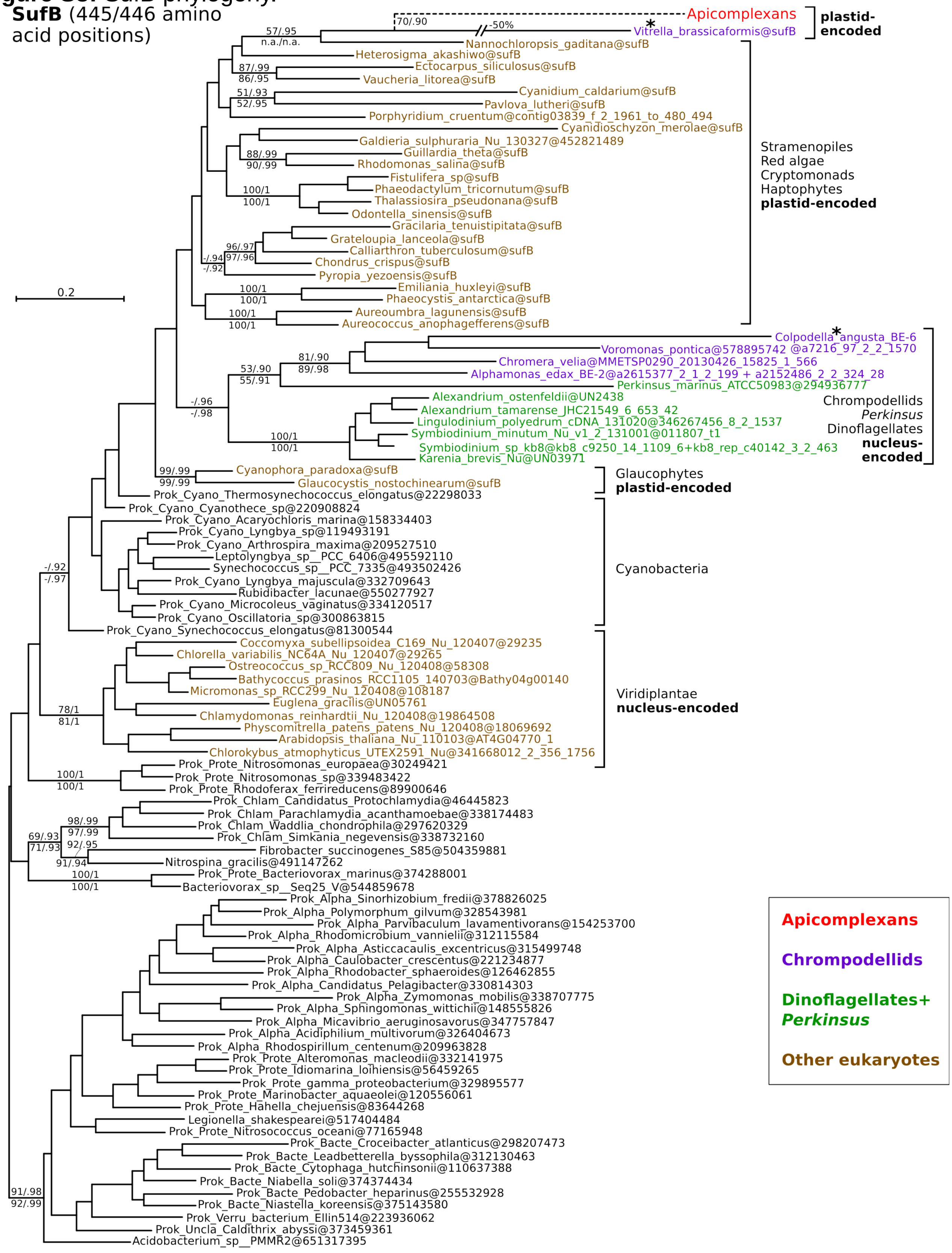


Figure S4: ClpC phylogeny.

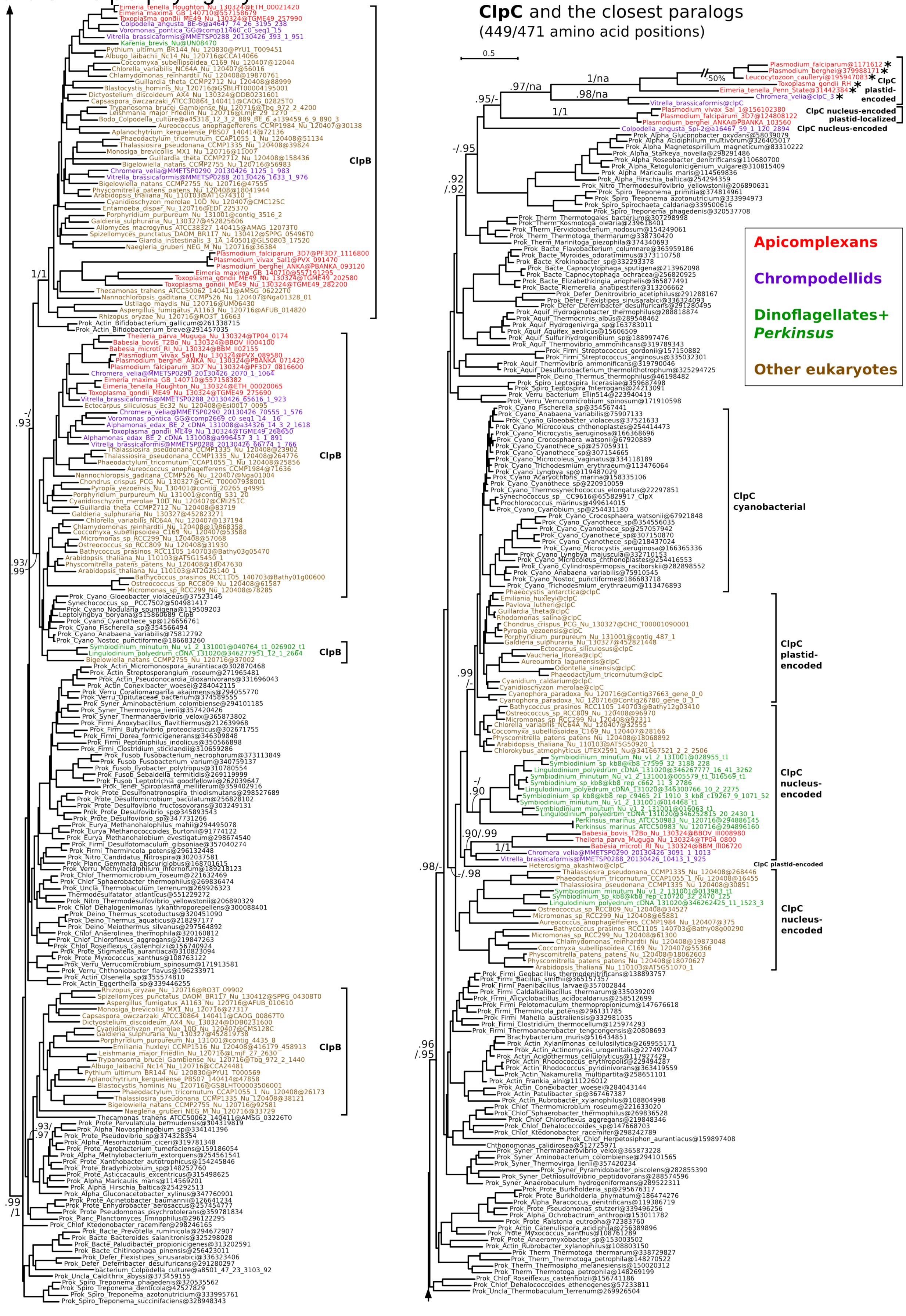
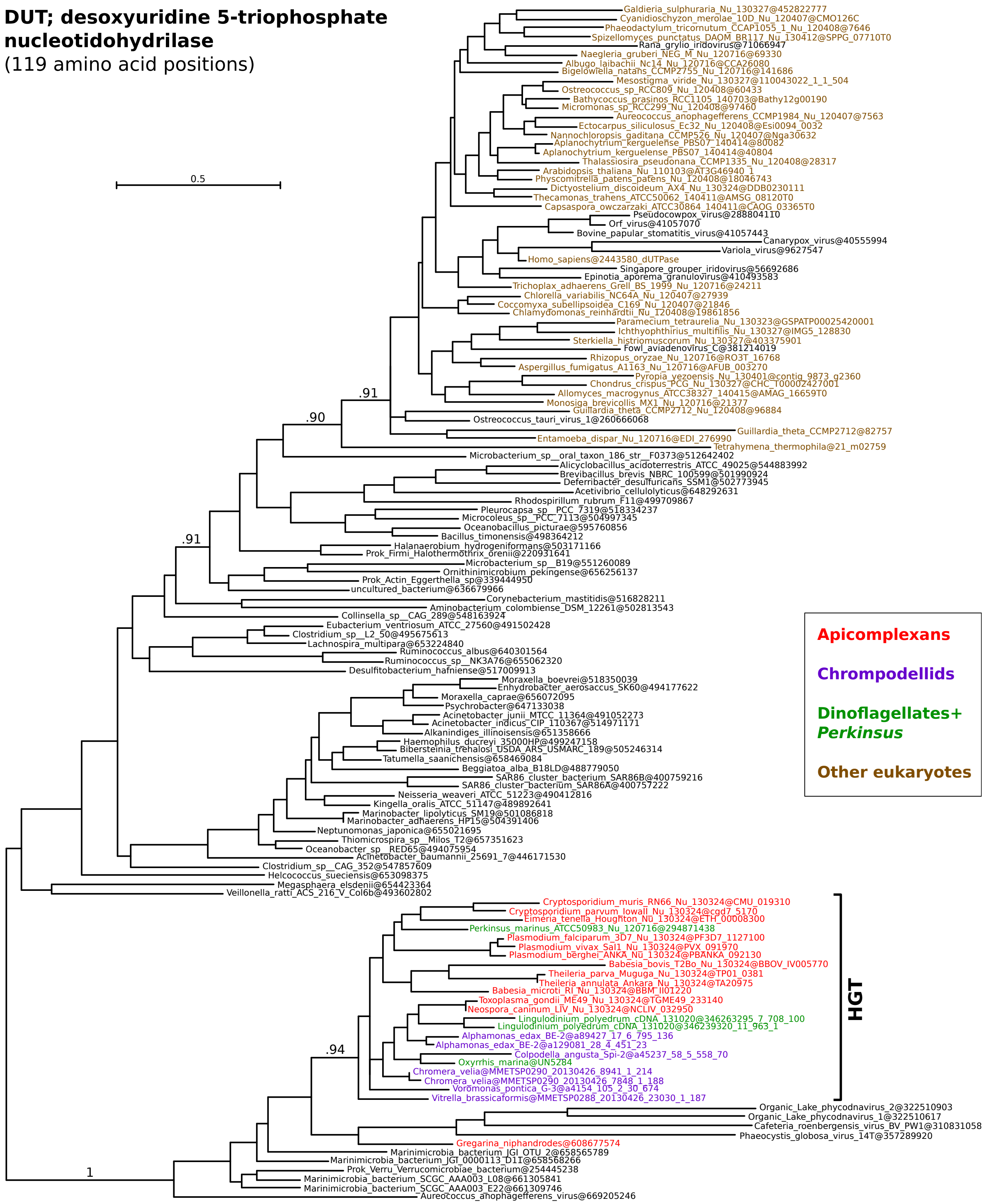
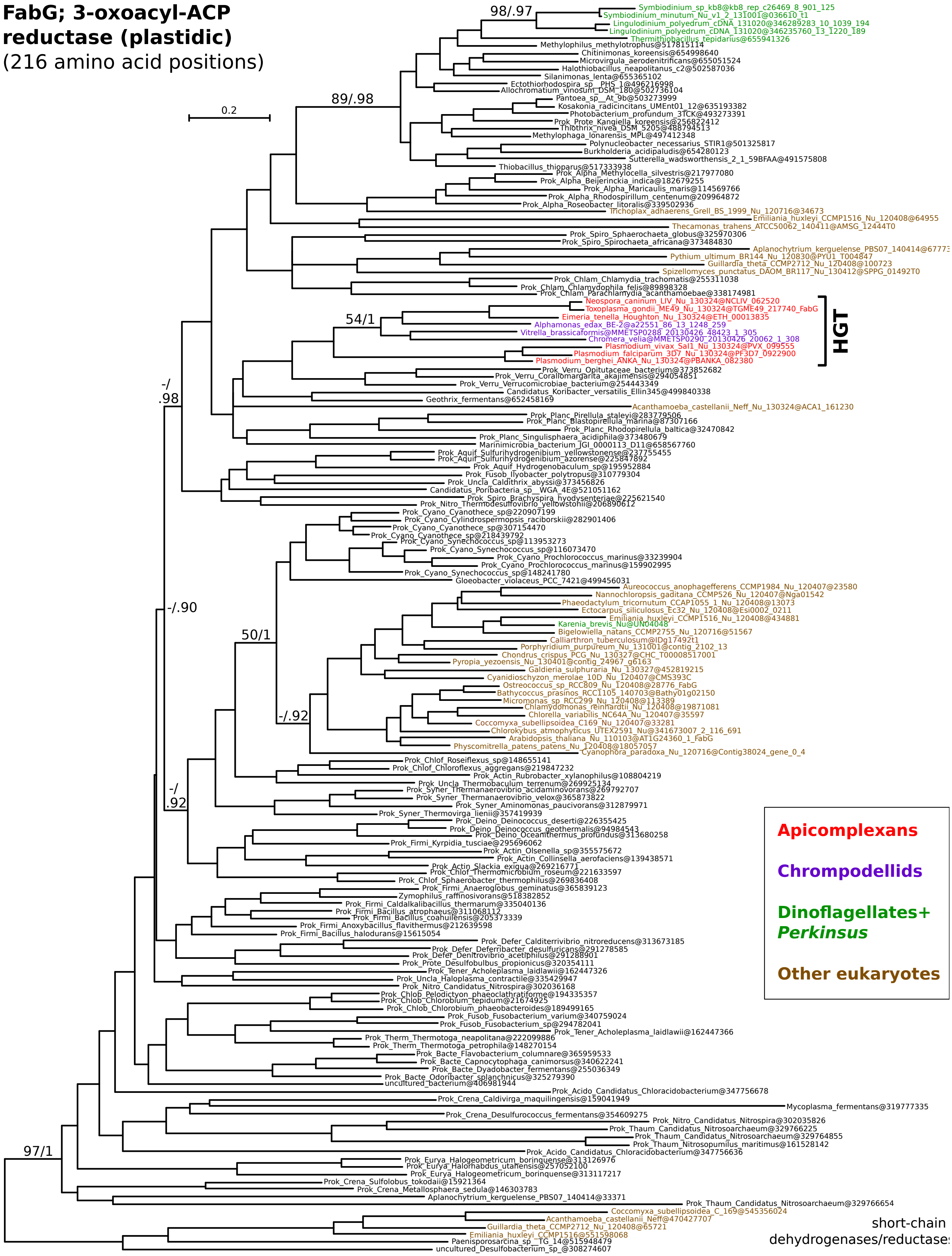
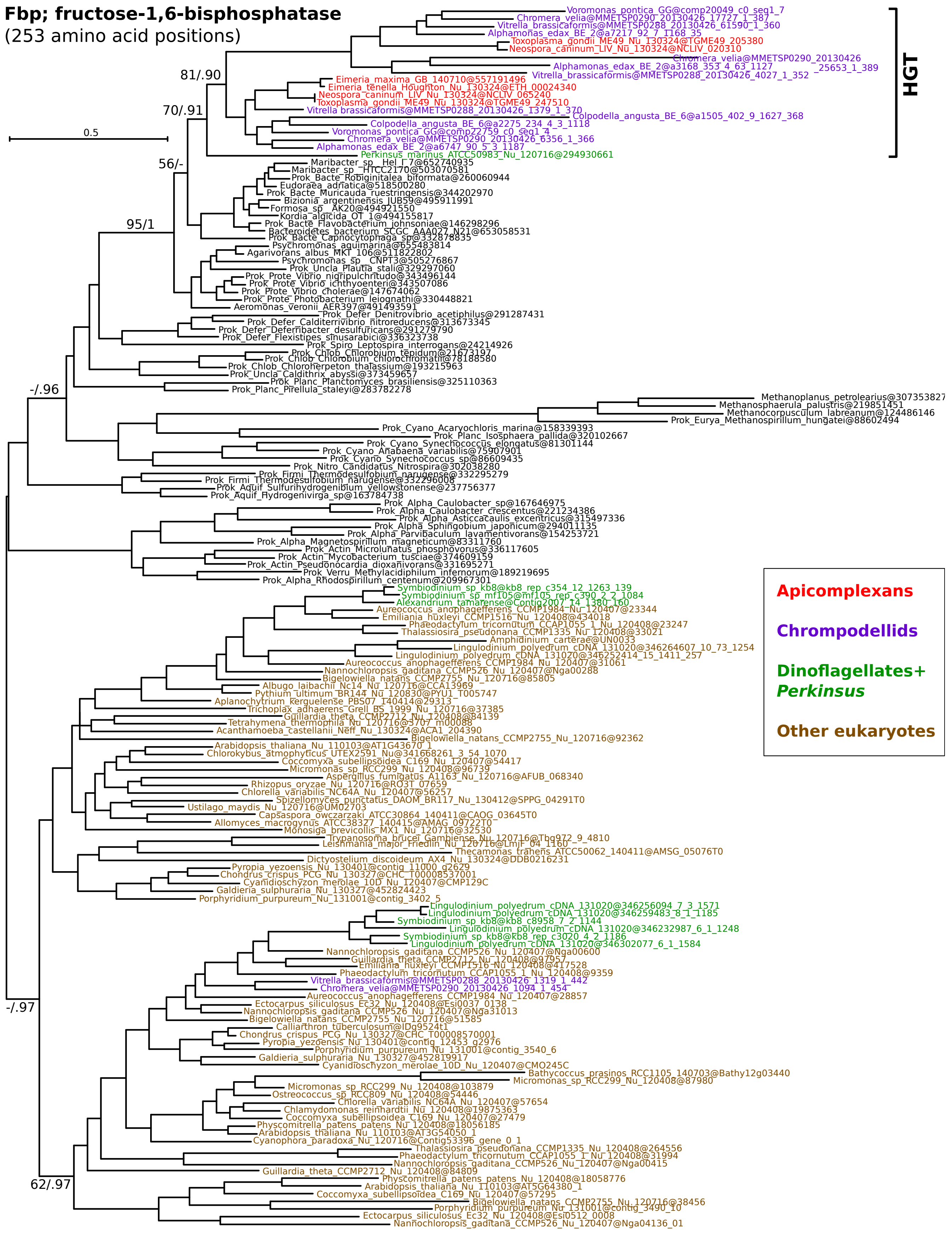


Figure S5: Phylogenies of 13 bacterial proteins in apicomplexans and relatives.

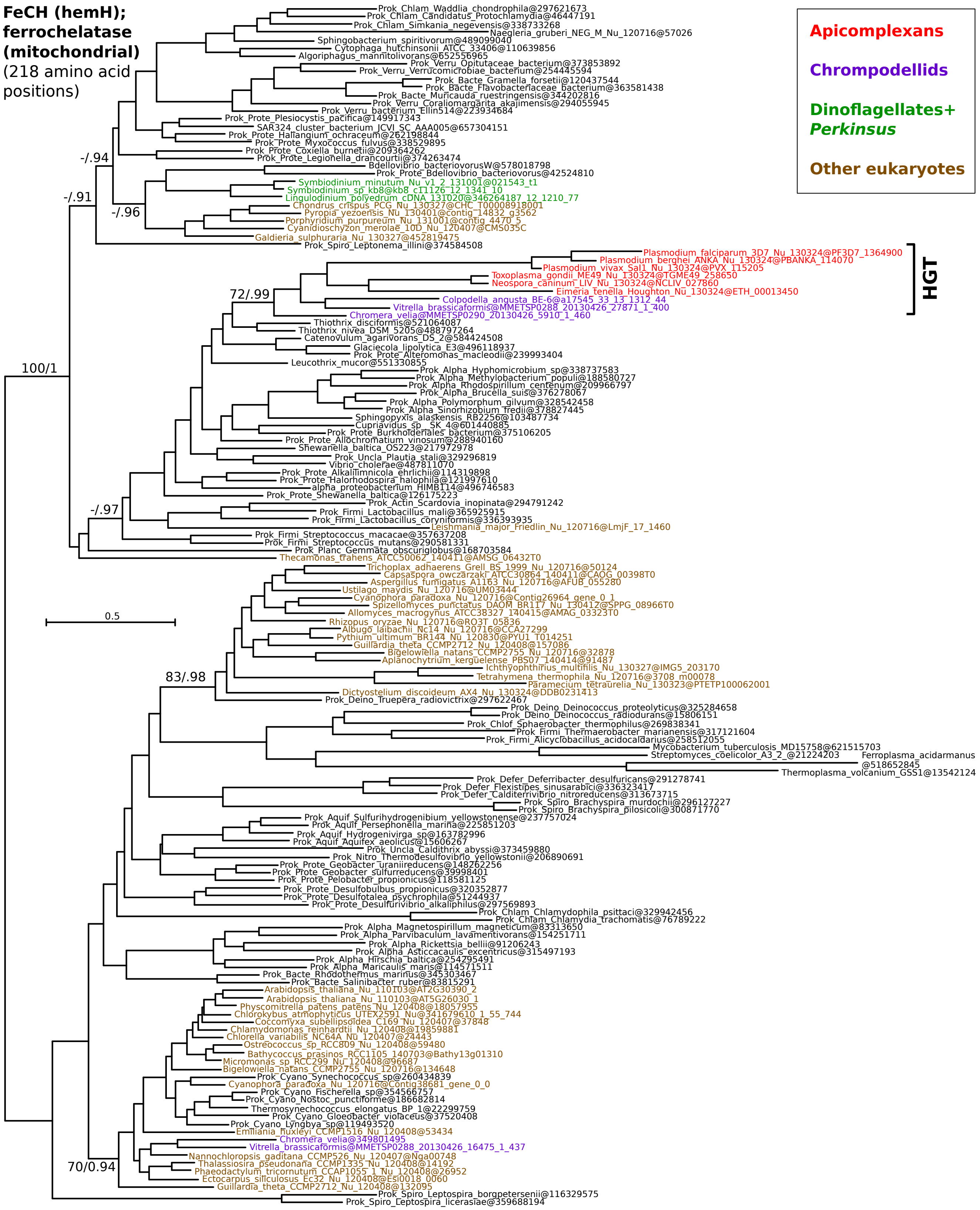
DUT; desoxyuridine 5-triophosphate nucleotidohydrilase
(119 amino acid positions)

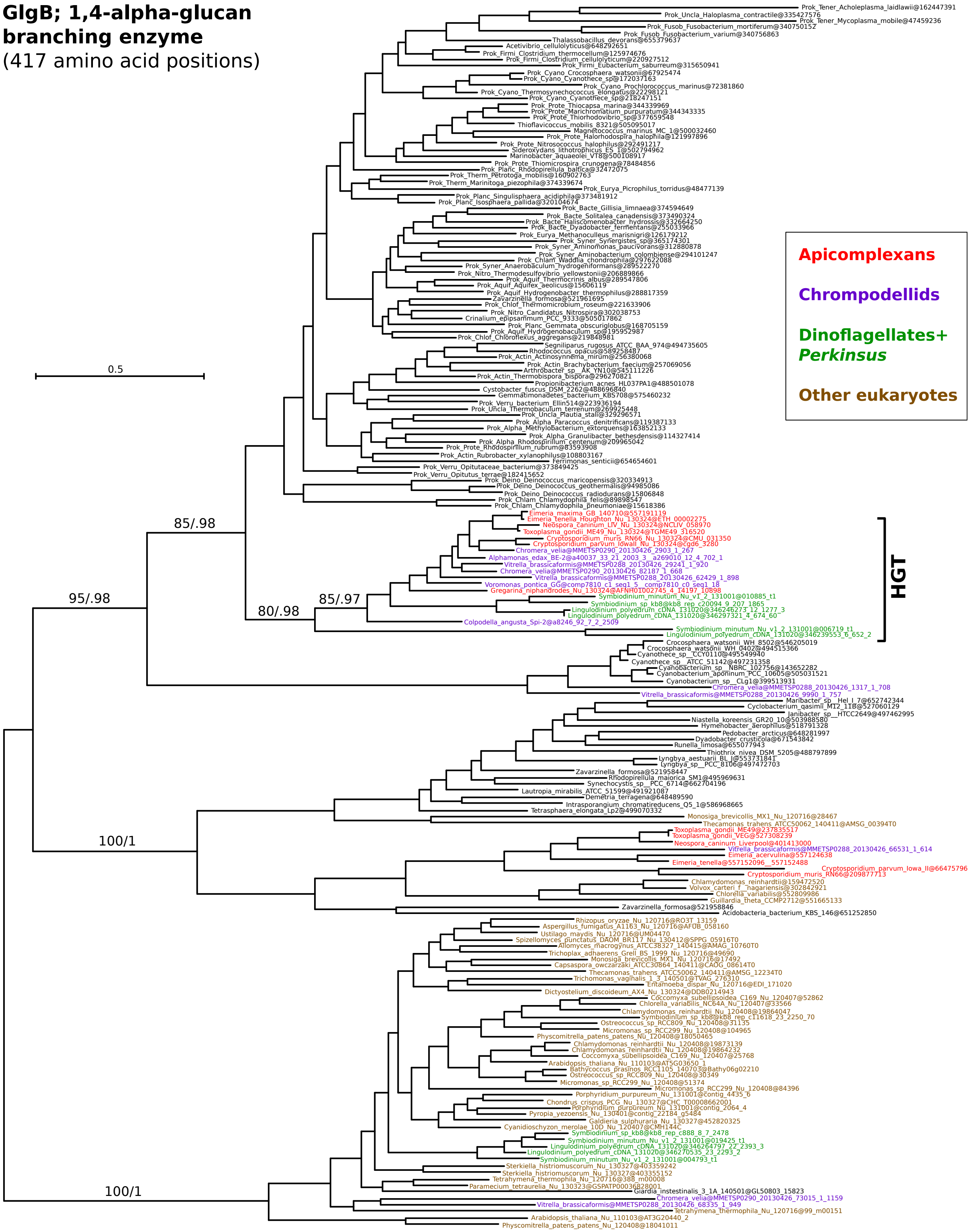


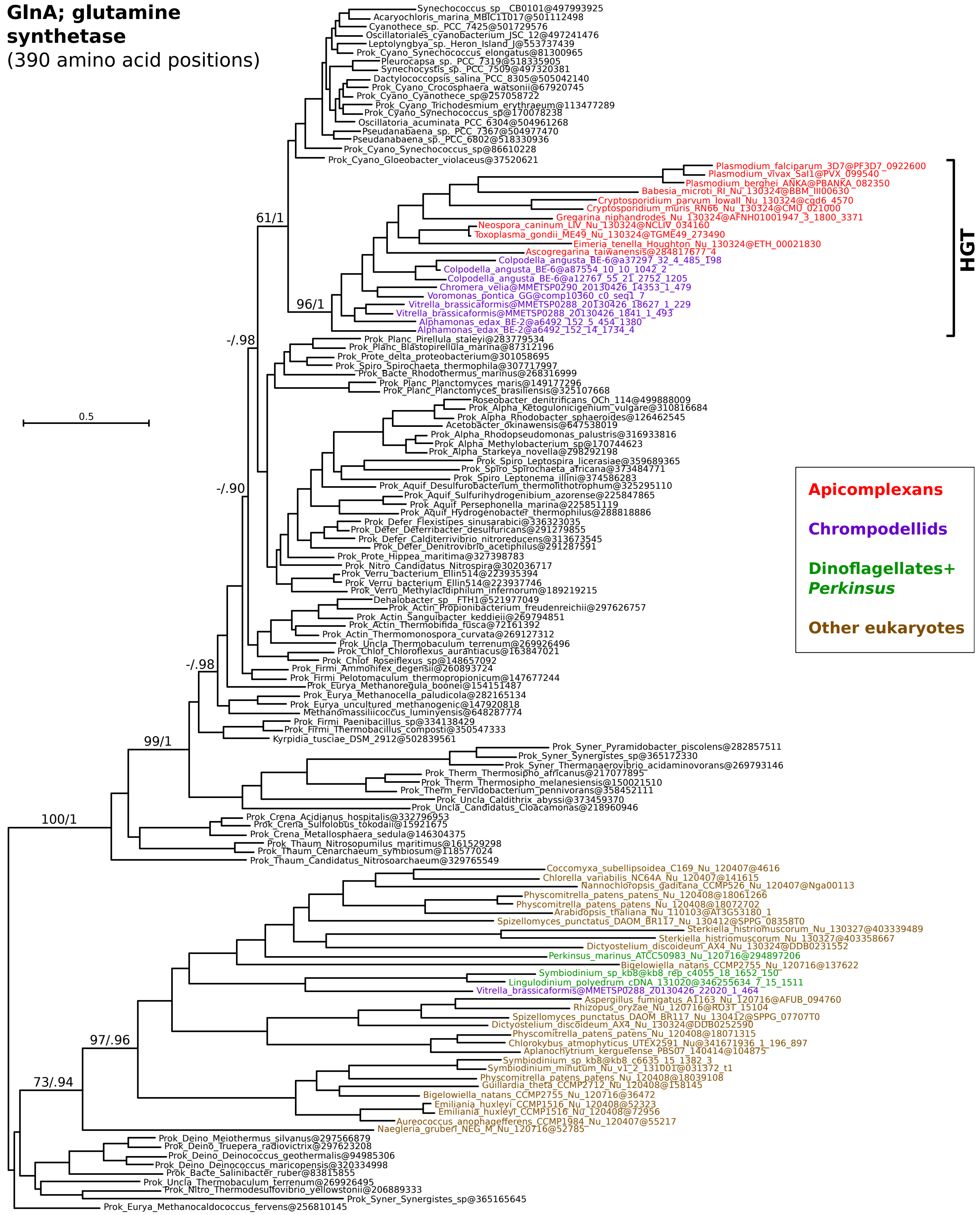

FabG; 3-oxoacyl-ACP reductase (plastidic)
(216 amino acid positions)
0.2
98/.97
89/.98
54/1
-/ .98
-/.90
50/1
-/.92
-/ .92
97/1
HGT
Symbiodinium_sp_kb8@kb8_rep_c26469_8_901_125
Symbiodinium_minutum_Nu_v1_2_131001@036610_t1
Lingulodinium_polyedrum_cDNA_131020@346289283_10_1039_194
Lingulodinium_polyedrum_cDNA_131020@346235760_13_1220_189
Thermithiobacillus_tepidarius@655941326
Methylophilus_methylotrophus@517815114
Chitinimonas_koreensis@654998640
Microvirgula_aerodenitrificans@655051524
Halothiobacillus_neapolitanus_c2@502587036
Silanimonas_lenta@655365102
Ectothiorhodospira_sp__PHS_1@496216998
Allochromatium_vinosum_DSM_180@502736104
Pantoea_sp__At_9b@503273999
Kosakonia_radicincitans_UMEnt01_12@635193382
Photobacterium_profundum_3TCK@493273391
Prok_Prote_Kangiella_koreensis@256822412
Thiothrix_nivea_DSM_5205@488794513
Methylophaga_lonarensis_MPL@497412348
Polynucleobacter_necessarius_STIR1@501325817
Burkholderia_acidipaludis@654280123
Sutterella_wadsworthensis_2_1_59BFAA@491575808
Thiobacillus_thioparus@517333938
Prok_Alpha_Methylocella_silvestris@217977080
Prok_Alpha_Beijerinckia_indica@182679255
Prok_Alpha_Maricaulis_maris@114569766
Prok_Alpha_Rhodospirillum_centenum@209964872
Prok_Alpha_Roseobacter_litoralis@339502936
Trichoplax_adhaerens_Grell_BS_1999_Nu_120716@34673
Emiliania_huxleyi_CCMP1516_Nu_120408@64955
Thecamonas_trahens_ATCC50062_140411@AMSG_12444T0
Prok_Spiro_Sphaerochaeta_globus@325970306
Prok_Spiro_Spirochaeta_africana@373484830
Aplanochytrium_kerguelense_PBS07_140414@67773
Pythium_ultimum_BR144_Nu_120830@PYU1_T004847
Guillardia_theta_CCMP2712_Nu_120408@100723
Spizellomyces_punctatus_DAOM_BR117_Nu_130412@SPPG_01492T0
Prok_Chlam_Chlamydia_trachomatis@255311038
Prok_Chlam_Chlamydophila_felis@89898328
Prok_Chlam_Parachlamydia_acanthamoebae@338174981
Neospora_caninum_LIV_Nu_130324@NCLIV_062520
Toxoplasma_gondii_ME49_Nu_130324@TGME49_217740_FabG
Eimeria_tenella_Houghton_Nu_130324@ETH_00013835
Alphamonas_edax_BE-2@a22551_86_13_1248_259
Vitrella_brassicaformis@MMETSP0288_20130426_48423_1_305
Chromera_velia@MMETSP0290_20130426_20062_1_308
Plasmodium_vivax_Sal1_Nu_130324@PVX_099555
Plasmodium_falciparum_3D7_Nu_130324@PF3D7_0922900
Plasmodium_berghei_ANKA_Nu_130324@PBANKA_082380
Prok_Verru_Opitutaceae_bacterium@373852682
Prok_Verru_Coraliomargarita_akajimensis@294054851
Prok_Verru_Verrucomicrobiae_bacterium@254443349
Candidatus_Koribacter_versatilis_Ellin345@499840338
Geothrix_fermentans@652458169
Acanthamoeba_castellanii_Neff_Nu_130324@ACA1_161230
Prok_Planc_Pirellula_staleyi@283779506
Prok_Planc_Blastopirellula_marina@87307166
Prok_Planc_Rhodopirellula_baltica@32470842
Prok_Planc_Singulisphaera_acidiphila@373480679
Marinimicrobia_bacterium_JGI_0000113_D11@658567760
Prok_Aquif_Sulfurihydrogenibium_yellowstonense@237755455
Prok_Aquif_Sulfurihydrogenibium_azorense@225847892
Prok_Aquif_Hydrogenobaculum_sp@195952884
Prok_Fusob_Ilyobacter_polytropus@310779304
Prok_Uncla_Caldithrix_abyssi@373456826
Candidatus_Poribacteria_sp__WGA_4E@521051162
Prok_Spiro_Brachyspira_hyodysenteriae@225621540
Prok_Nitro_Thermodesulfovibrio_yellowstonii@206890612
Prok_Cyano_Cyanothece_sp@220907199
Prok_Cyano_Cylindrospermopsis_raciborskii@282901406
Prok_Cyano_Cyanothece_sp@307154470
Prok_Cyano_Cyanothece_sp@218439792
Prok_Cyano_Synechococcus_sp@113953273
Prok_Cyano_Synechococcus_sp@116073470
Prok_Cyano_Prochlorococcus_marinus@33239904
Prok_Cyano_Prochlorococcus_marinus@159902995
Prok_Cyano_Synechococcus_sp@148241780
Gloeobacter_violaceus_PCC_7421@499456031
Aureococcus_anophagefferens_CCMP1984_Nu_120407@23580
Nannochloropsis_gaditana_CCMP526_Nu_120407@Nga01542
Phaeodactylum_tricornutum_CCAP1055_1_Nu_120408@13073
Ectocarpus_siliculosus_Ec32_Nu_120408@Esi0002_0211
Emiliania_huxleyi_CCMP1516_Nu_120408@434881
Karenia_brevis_Nu@UN04048
Bigelowiella_natans_CCMP2755_Nu_120716@51567
Calliarthron_tuberculosum@IDg17492t1
Porphyridium_purpureum_Nu_131001@contig_2102_13
Chondrus_crispus_PCG_Nu_130327@CHC_T00008517001
Pyropia_yezoensis_Nu_130401@contig_24967_g6163
Galdieria_sulphuraria_Nu_130327@452819215
Cyanidioschyzon_merolae_10D_Nu_120407@CMS393C
Ostreococcus_sp_RCC809_Nu_120408@28776_FabG
Bathycoccus_prasinos_RCC1105_140703@Bathy01g02150
Micromonas_sp_RCC299_Nu_120408@113389
Chlamydomonas_reinhardtii_Nu_120408@19871081
Chlorella_variabilis_NC64A_Nu_120407@35597
Coccomyxa_subellipsoidea_C169_Nu_120407@33281
Chlorokybus_atmophyticus_UTEX2591_Nu@341673007_2_116_691
Arabidopsis_thaliana_Nu_110103@AT1G24360_1_FabG
Physcomitrella_patens_patens_Nu_120408@18057057
Cyanophora_paradoxa_Nu_120716@Contig38024_gene_0_4
Prok_Chlof_Roseiflexus_sp@148655141
Prok_Chlof_Chloroflexus_aggregans@219847232
Prok_Actin_Rubrobacter_xylanophilus@108804219
Prok_Uncla_Thermobaculum_terrenum@269925134
Prok_Syner_Thermanaerovibrio_acidaminovorans@269792707
Prok_Syner_Thermanaerovibrio_velox@365873822
Prok_Syner_Aminomonas_paucivorans@312879971
Prok_Syner_Thermovirga_lienii@357419939
Prok_Deino_Deinococcus_deserti@226355425
Prok_Deino_Deinococcus_geothermalis@94984543
Prok_Deino_Oceanithermus_profundus@313680258
Prok_Firmi_Kyrpidia_tusciae@295696062
Prok_Actin_Olsenella_sp@355575672
Prok_Actin_Collinsella_aerofaciens@139438571
Prok_Actin_Slackia_exigua@269216771
Prok_Chlof_Thermomicrobium_roseum@221633597
Prok_Chlof_Sphaerobacter_thermophilus@269836408
Prok_Firmi_Anaeroglobus_geminatus@365839123
Zymophilus_raffinosivorans@518382852
Prok_Firmi_Caldalkalibacillus_thermarum@335040136
Prok_Firmi_Bacillus_atrophaeus@311068112
Prok_Firmi_Bacillus_coahuilensis@205373339
Prok_Firmi_Anoxybacillus_flavithermus@212639598
Prok_Firmi_Bacillus_halodurans@15615054
Prok_Defer_Calditerrivibrio_nitroreducens@313673185
Prok_Defer_Deferribacter_desulfuricans@291278585
Prok_Defer_Denitrovibrio_acetiphilus@291288901
Prok_Prote_Desulfobulbus_propionicus@320354111
Prok_Tener_Acholeplasma_laidlawii@162447326
Prok_Uncla_Haloplasma_contractile@335429947
Prok_Nitro_Candidatus_Nitrospira@302036168
Prok_Chlob_Pelodictyon_phaeoclathratiforme@194335357
Prok_Chlob_Chlorobium_tepidum@21674925
Prok_Chlob_Chlorobium_phaeobacteroides@189499165
Prok_Fusob_Fusobacterium_varium@340759024
Prok_Fusob_Fusobacterium_sp@294782041
Prok_Tener_Acholeplasma_laidlawii@162447366
Prok_Therm_Thermotoga_neapolitana@222099886
Prok_Therm_Thermotoga_petrophila@148270154
Prok_Bacte_Flavobacterium_columnare@365959533
Prok_Bacte_Capnocytophaga_canimorsus@340622241
Prok_Bacte_Dyadobacter_fermentans@255036349
Prok_Bacte_Odoribacter_splanchnicus@325279390
uncultured_bacterium@406981944
Prok_Acido_Candidatus_Chloracidobacterium@347756678
Prok_Crena_Caldivirga_maquilingensis@159041949
Mycoplasma_fermentans@319777335
Prok_Crena_Desulfurococcus_fermentans@354609275
Prok_Nitro_Candidatus_Nitrospira@302035826
Prok_Thaum_Candidatus_Nitrosoarchaeum@329766225
Prok_Thaum_Candidatus_Nitrosoarchaeum@329764855
Prok_Thaum_Nitrosopumilus_maritimus@161528142
Prok_Acido_Candidatus_Chloracidobacterium@347756636
Prok_Eurya_Halogeometricum_borinquense@313126976
Prok_Eurya_Halorhabdus_utahensis@257052100
Prok_Eurya_Halogeometricum_borinquense@313117217
Prok_Crena_Sulfolobus_tokodaii@15921364
Prok_Crena_Metallosphaera_sedula@146303783
Aplanochytrium_kerguelense_PBS07_140414@33371
Prok_Thaum_Candidatus_Nitrosoarchaeum@329766654
Coccomyxa_subellipsoidea_C_169@545356024
Acanthamoeba_castellanii_Neff@470427707
Guillardia_theta_CCMP2712_Nu_120408@65721
Emiliania_huxleyi_CCMP1516@551598068
Paenisporosarcina_sp__TG_14@515948479
uncultured_Desulfobacterium_sp_@308274607
short-chain
dehydrogenases/reductases
Apicomplexans
Chrompodellids
Dinoflagellates+
Perkinsus
Other eukaryotes


Fbp; fructose-1,6-bisphosphatase

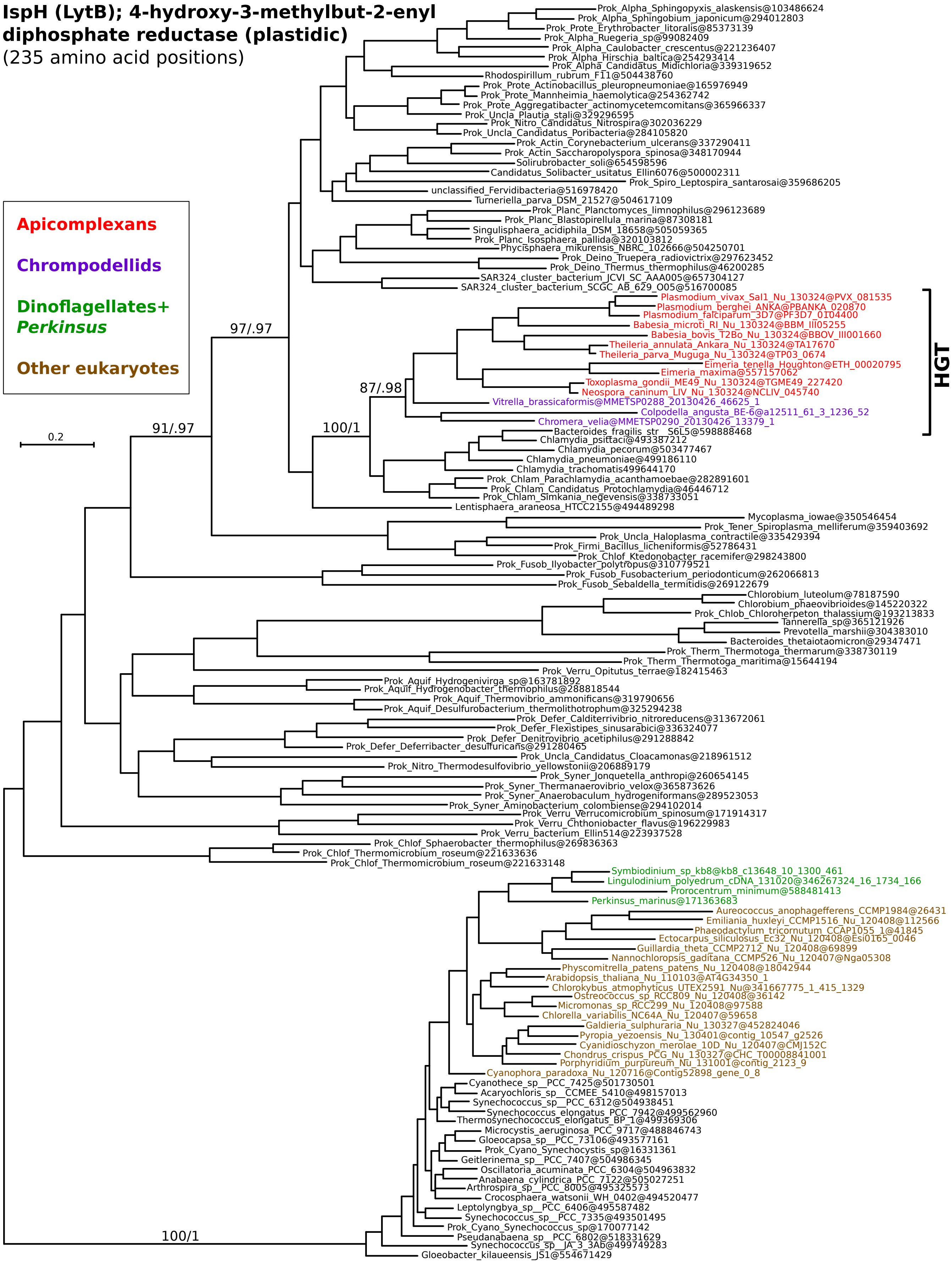
(253 amino acid positions)

**FeCH (hemH); ferrochelatase (mitochondrial)**
(218 amino acid positions)
Apicomplexans
Chrompodellids
Dinoflagellates+ *Perkinsus*
Other eukaryotes
HGT
100/1
-/.91
-/.94
-/.96
72/.99
-/.97
0.5
83/.98
70/0.94
Prok_Chlam_Waddlia_chondrophila@297621673
Prok_Chlam_Candidatus_Protochlamydia@46447191
Prok_Chlam_Simkania_negevensis@338733268
Naegleria_gruberi_NEG_M_Nu_120716@57026
Sphingobacterium_spiritivorum@489099040
Cytophaga_hutchinsonii_ATCC_33406@110639856
Algoriphagus_mannitolivorans@652556965
Prok_Verru_Opitutaceae_bacterium@373853892
Prok_Verru_Verrucomicrobiae_bacterium@254445594
Prok_Bacte_Gramella_forsetii@120437544
Prok_Bacte_Flavobacteriaceae_bacterium@363581438
Prok_Bacte_Muricauda_ruestringensis@344202816
Prok_Verru_Coraliomargarita_akajimensis@294055945
Prok_Verru_bacterium_Ellin514@223934684
Prok_Prote_Plesiocystis_pacifica@149917343
SAR324_cluster_bacterium_JCVI_SC_AAA005@657304151
Prok_Prote_Haliangium_ochraceum@262198844
Prok_Prote_Myxococcus_fulvus@338529895
Prok_Prote_Coxiella_burnetii@209364262
Prok_Prote_Legionella_drancourtii@374263474
Bdellovibrio_bacteriovorusW@578018798
Prok_Prote_Bdellovibrio_bacteriovorus@42524810
Symbiodinium_minutum_Nu_v1_2_131001@021543_t1
Symbiodinium_sp_kb8@kb8_c11126_12_1341_10
Lingulodinium_polyedrum_cDNA_131020@346264187_12_1210_77
Chondrus_crispus_PCG_Nu_130327@CHC_T00008918001
Pyropia_yezoensis_Nu_130401@contig_14832_g3562
Porphyridium_purpureum_Nu_131001@contig_4470_5
Cyanidioschyzon_merolae_10D_Nu_120407@CMS035C
Galdieria_sulphuraria_Nu_130327@452819475
Prok_Spiro_Leptonema_illini@374584508
Plasmodium_falciparum_3D7_Nu_130324@PF3D7_1364900
Plasmodium_berghei_ANKA_Nu_130324@PBANKA_114070
Plasmodium_vivax_Sal1_Nu_130324@PVX_115205
Toxoplasma_gondii_ME49_Nu_130324@TGME49_258650
Neospora_caninum_LIV_Nu_130324@NCLIV_027860
Eimeria_tenella_Houghton_Nu_130324@ETH_00013450
Colpodella_angusta_BE-6@a17545_33_13_1312_44
Vitrella_brassicaformis@MMETSP0288_20130426_27871_1_400
Chromera_velia@MMETSP0290_20130426_5910_1_460
Thiothrix_disciformis@521064087
Thiothrix_nivea_DSM_5205@488797264
Catenovulum_agarivorans_DS_2@584424508
Glaciecola_lipolytica_E3@496118937
Prok_Prote_Alteromonas_macleodii@239993404
Leucothrix_mucor@551330855
Prok_Alpha_Hyphomicrobium_sp@338737583
Prok_Alpha_Methylobacterium_populi@188580727
Prok_Alpha_Rhodospirillum_centenum@209966797
Prok_Alpha_Brucella_suis@376278067
Prok_Alpha_Polymorphum_gilvum@328542458
Prok_Alpha_Sinorhizobium_fredii@378827445
Sphingopyxis_alaskensis_RB2256@103487734
Cupriavidus_sp__SK_4@601440885
Prok_Prote_Burkholderiales_bacterium@375106205
Prok_Prote_Allochromatium_vinosum@288940160
Shewanella_baltica_OS223@217972978
Prok_Uncla_Plautia_stali@329296819
Vibrio_cholerae@487811070
Prok_Prote_Alkalilimnicola_ehrlichii@114319898
Prok_Prote_Halorhodospira_halophila@121997610
alpha_proteobacterium_HIMB114@496746583
Prok_Prote_Shewanella_baltica@126175223
Prok_Actin_Scardovia_inopinata@294791242
Prok_Firmi_Lactobacillus_mali@365925915
Prok_Firmi_Lactobacillus_coryniformis@336393935
Leishmania_major_Friedlin_Nu_120716@LmjF_17_1460
Prok_Firmi_Streptococcus_macacae@357637208
Prok_Firmi_Streptococcus_mutans@290581331
Prok_Planc_Gemmata_obscuriglobus@168703584
Thecamonas_trahens_ATCC50062_140411@AMSG_06432T0
Trichoplax_adhaerens_Grell_BS_1999_Nu_120716@50124
Capsaspora_owczarzaki_ATCC30864_140411@CAOG_00398T0
Aspergillus_fumigatus_A1163_Nu_120716@AFUB_055280
Ustilago_maydis_Nu_120716@UM03444
Cyanophora_paradoxa_Nu_120716@Contig26964_gene_0_1
Spizellomyces_punctatus_DAOM_BR117_Nu_130412@SPPG_08966T0
Allomyces_macrogynus_ATCC38327_140415@AMAG_03323T0
Rhizopus_oryzae_Nu_120716@RO3T_05836
Albugo_laibachii_Nc14_Nu_120716@CCA27299
Pythium_ultimum_BR144_Nu_120830@PYU1_T014251
Guillardia_theta_CCMP2712_Nu_120408@157086
Bigelowiella_natans_CCMP2755_Nu_120716@32878
Aplanochytrium_kerguelense_PBS07_140414@91487
Ichthyophthirius_multifilis_Nu_130327@IMG5_203170
Tetrahymena_thermophila_Nu_120716@3708_m00078
Paramecium_tetraurelia_Nu_130323@PTETP100062001
Dictyostelium_discoideum_AX4_Nu_130324@DDB0231413
Prok_Deino_Truepera_radiovictrix@297622467
Prok_Deino_Deinococcus_proteolyticus@325284658
Prok_Deino_Deinococcus_radiodurans@15806151
Prok_Chlof_Sphaerobacter_thermophilus@269838341
Prok_Firmi_Thermaerobacter_marianensis@317121604
Prok_Firmi_Alicyclobacillus_acidocaldarius@258512055
Mycobacterium_tuberculosis_MD15758@621515703
Streptomyces_coelicolor_A3_2_@21224203
Ferroplasma_acidarmanus @518652845
Thermoplasma_volcanium_GSS1@13542124
Prok_Defer_Deferribacter_desulfuricans@291278741
Prok_Defer_Flexistipes_sinusarabici@336323417
Prok_Defer_Calditerrivibrio_nitroreducens@313673715
Prok_Spiro_Brachyspira_murdochii@296127227
Prok_Spiro_Brachyspira_pilosicoli@300871770
Prok_Aquif_Sulfurihydrogenibium_yellowstonense@237757024
Prok_Aquif_Persephonella_marina@225851203
Prok_Aquif_Hydrogenivirga_sp@163782996
Prok_Aquif_Aquifex_aeolicus@15606267
Prok_Uncla_Caldithrix_abyssi@373459880
Prok_Nitro_Thermodesulfovibrio_yellowstonii@206890691
Prok_Prote_Geobacter_uraniireducens@148262256
Prok_Prote_Geobacter_sulfurreducens@39998401
Prok_Prote_Pelobacter_propionicus@118581125
Prok_Prote_Desulfobulbus_propionicus@320352877
Prok_Prote_Desulfotalea_psychrophila@51244937
Prok_Prote_Desulfurivibrio_alkaliphilus@297569893
Prok_Chlam_Chlamydophila_psittaci@329942456
Prok_Chlam_Chlamydia_trachomatis@76789222
Prok_Alpha_Magnetospirillum_magneticum@83313650
Prok_Alpha_Parvibaculum_lavamentivorans@154251711
Prok_Alpha_Rickettsia_bellii@91206243
Prok_Alpha_Asticcacaulis_excentricus@315497193
Prok_Alpha_Hirschia_baltica@254295491
Prok_Alpha_Maricaulis_maris@114571511
Prok_Bacte_Rhodothermus_marinus@345303467
Prok_Bacte_Salinibacter_ruber@83815291
Arabidopsis_thaliana_Nu_110103@AT2G30390_2
Arabidopsis_thaliana_Nu_110103@AT5G26030_1
Physcomitrella_patens_patens_Nu_120408@18057955
Chlorokybus_atmophyticus_UTEX2591_Nu@341679610_1_55_744
Coccomyxa_subellipsoidea_C169_Nu_120407@37848
Chlamydomonas_reinhardtii_Nu_120408@19859881
Chlorella_variabilis_NC64A_Nu_120407@24443
Ostreococcus_sp_RCC809_Nu_120408@59480
Bathycoccus_prasinos_RCC1105_140703@Bathy13g01310
Micromonas_sp_RCC299_Nu_120408@96687
Bigelowiella_natans_CCMP2755_Nu_120716@134648
Prok_Cyano_Synechococcus_sp@260434839
Cyanophora_paradoxa_Nu_120716@Contig38681_gene_0_0
Prok_Cyano_Fischerella_sp@354566757
Prok_Cyano_Nostoc_punctiforme@186682814
Thermosynechococcus_elongatus_BP_1@22299759
Prok_Cyano_Gloeobacter_violaceus@37520408
Prok_Cyano_Lyngbya_sp@119493520
Emiliania_huxleyi_CCMP1516_Nu_120408@53434
Chromera_velia@349801495
Vitrella_brassicaformis@MMETSP0288_20130426_16475_1_437
Nannochloropsis_gaditana_CCMP526_Nu_120407@Nga00748
Thalassiosira_pseudonana_CCMP1335_Nu_120408@14192
Phaeodactylum_tricornutum_CCAP1055_1_Nu_120408@26952
Ectocarpus_siliculosus_Ec32_Nu_120408@Esi0018_0060
Guillardia_theta_CCMP2712_Nu_120408@132095
Prok_Spiro_Leptospira_borgpetersenii@116329575
Prok_Spiro_Leptospira_licerasiae@359688194

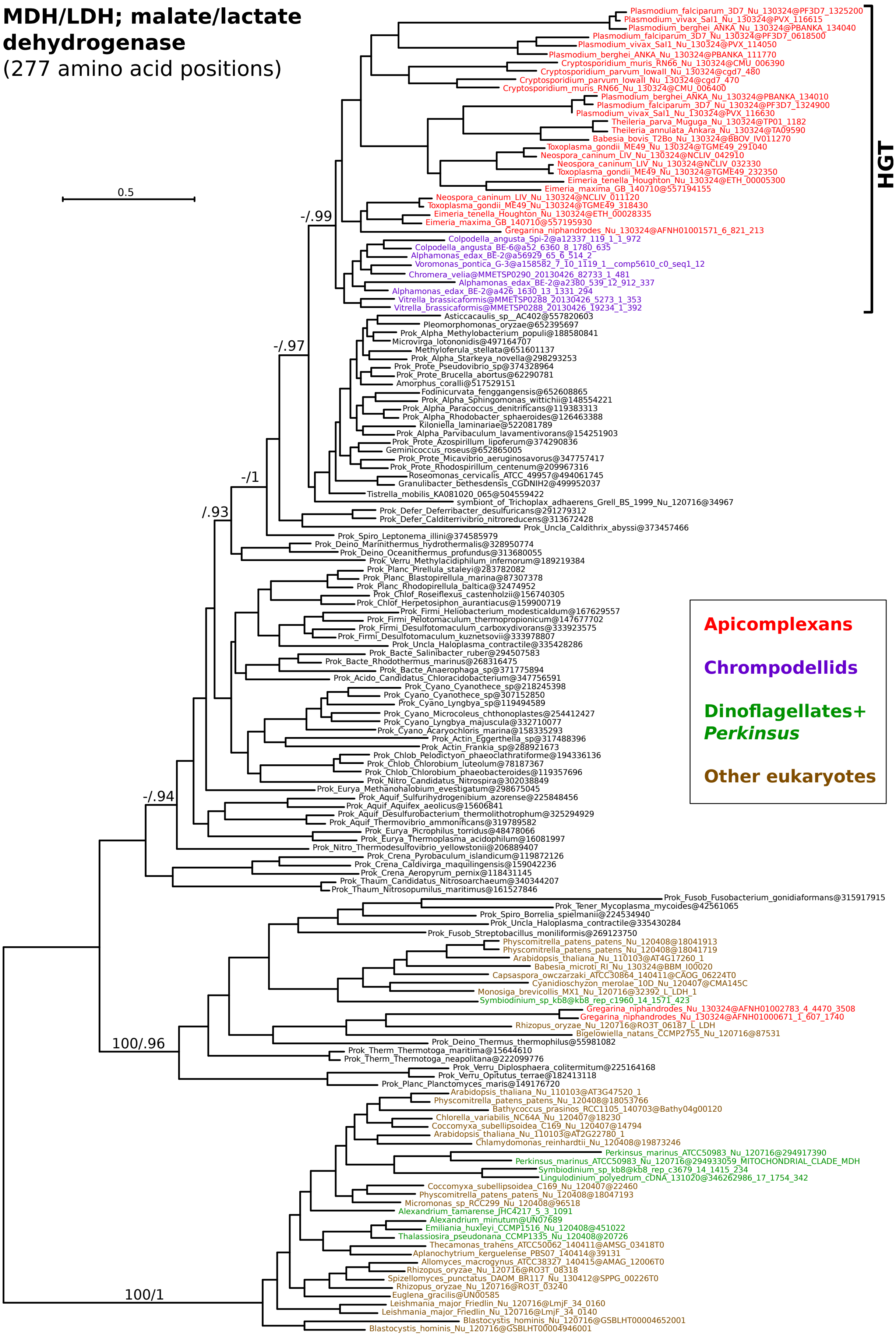

GlgB; 1,4-alpha-glucan branching enzyme
(417 amino acid positions)

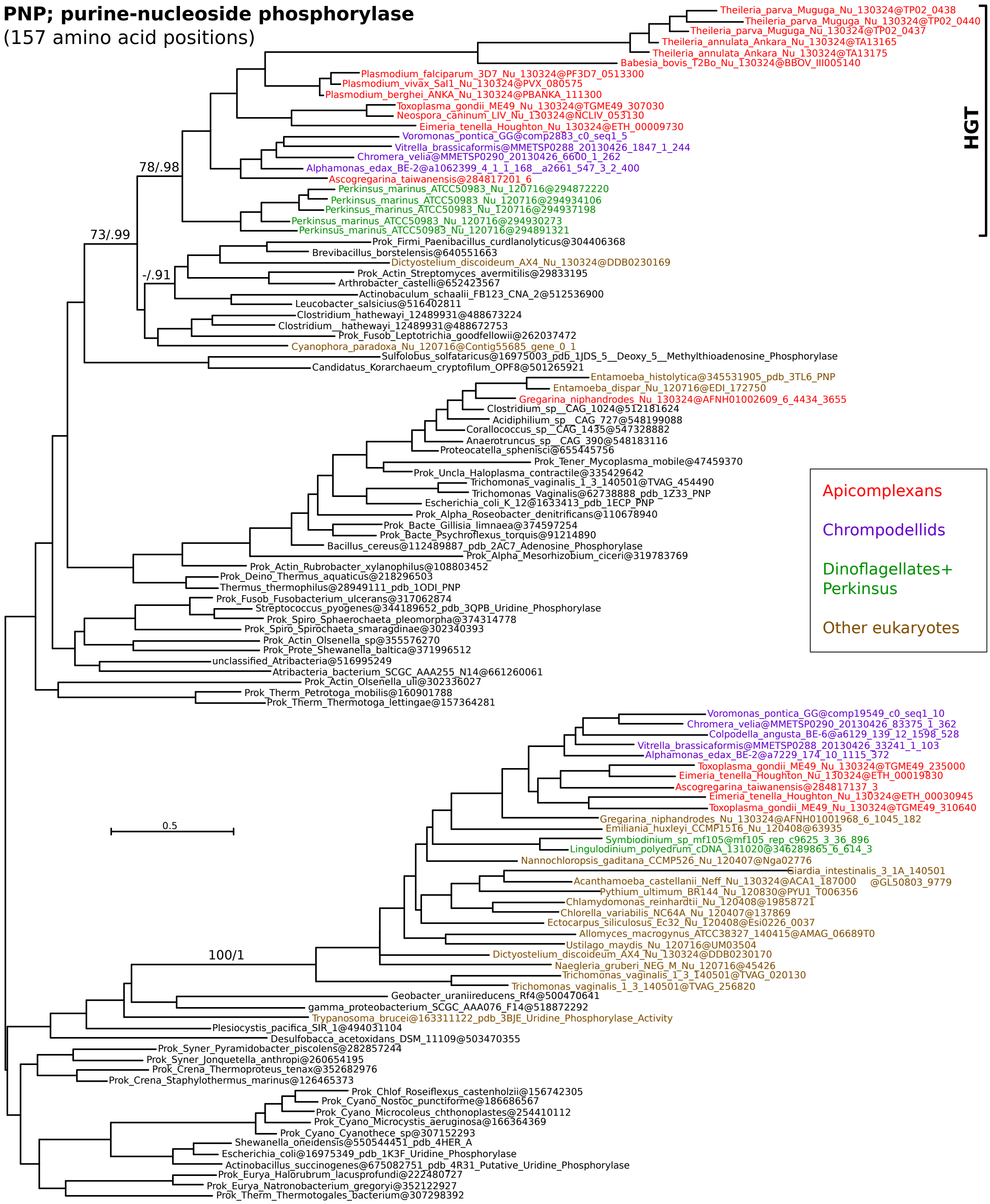

GlnA; glutamine synthetase
(390 amino acid positions)
Synechococcus_sp__CB0101@497993925
Acaryochloris_marina_MBIC11017@501112498
Cyanothece_sp._PCC_7425@501729576
Oscillatoriales_cyanobacterium_JSC_12@497241476
Leptolyngbya_sp._Heron_Island_J@553737439
Prok_Cyano_Synechococcus_elongatus@81300965
Pleurocapsa_sp._PCC_7319@518335905
Synechocystis_sp._PCC_7509@497320381
Dactylococcopsis_salina_PCC_8305@505042140
Prok_Cyano_Crocosphaera_watsonii@67920745
Prok_Cyano_Cyanothece_sp@257058722
Prok_Cyano_Trichodesmium_erythraeum@113477289
Prok_Cyano_Synechococcus_sp@170078238
Oscillatoria_acuminata_PCC_6304@504961268
Pseudanabaena_sp._PCC_7367@504977470
Pseudanabaena_sp._PCC_6802@518330936
Prok_Cyano_Synechococcus_sp@86610228
Prok_Cyano_Gloeobacter_violaceus@37520621
61/1
Plasmodium_falciparum_3D7@PF3D7_0922600
Plasmodium_vivax_Sal1@PVX_099540
Plasmodium_berghei_ANKA@PBANKA_082350
Babesia_microti_RI_Nu_130324@BBM_III00630
Cryptosporidium_parvum_Iowall_Nu_130324@cgd6_4570
Cryptosporidium_muris_RN66_Nu_130324@CMU_021000
Gregarina_niphandrodes_Nu_130324@AFNH01001947_3_1800_3371
Neospora_caninum_LIV_Nu_130324@NCLIV_034160
Toxoplasma_gondii_ME49_Nu_130324@TGME49_273490
Eimeria_tenella_Houghton_Nu_130324@ETH_00021830
Ascogregarina_taiwanensis@284817677_4
Colpodella_angusta_BE-6@a37297_32_4_485_198
Colpodella_angusta_BE-6@a87554_10_10_1042_2
Colpodella_angusta_BE-6@a12767_55_21_2752_1205
Chromera_velia@MMETSP0290_20130426_14353_1_479
Voromonas_pontica_GG@comp10360_c0_seq1_7
Vitrella_brassicaformis@MMETSP0288_20130426_18627_1_229
Vitrella_brassicaformis@MMETSP0288_20130426_1841_1_493
Alphamonas_edax_BE-2@a6492_152_5_454_1380
Alphamonas_edax_BE-2@a6492_152_14_1734_4
96/1
HGT
-/.98
Prok_Planc_Pirellula_staleyi@283779534
Prok_Planc_Blastopirellula_marina@87312196
Prok_Prote_delta_proteobacterium@301058695
Prok_Spiro_Spirochaeta_thermophila@307717997
Prok_Bacte_Rhodothermus_marinus@268316999
Prok_Planc_Planctomyces_maris@149177296
Prok_Planc_Planctomyces_brasiliensis@325107668
Roseobacter_denitrificans_OCh_114@499888009
Prok_Alpha_Ketogulonicigenium_vulgare@310816684
Prok_Alpha_Rhodobacter_sphaeroides@126462545
Acetobacter_okinawensis@647538019
Prok_Alpha_Rhodopseudomonas_palustris@316933816
Prok_Alpha_Methylobacterium_sp@170744623
Prok_Alpha_Starkeya_novella@298292198
Prok_Spiro_Leptospira_licerasiae@359689365
Prok_Spiro_Spirochaeta_africana@373484771
Prok_Spiro_Leptonema_illini@374586283
Prok_Aquif_Desulfurobacterium_thermolithotrophum@325295110
Prok_Aquif_Sulfurihydrogenibium_azorense@225847865
Prok_Aquif_Persephonella_marina@225851119
Prok_Aquif_Hydrogenobacter_thermophilus@288818886
Prok_Defer_Flexistipes_sinusarabici@336323035
Prok_Defer_Deferribacter_desulfuricans@291279855
Prok_Defer_Calditerrivibrio_nitroreducens@313673545
Prok_Defer_Denitrovibrio_acetiphilus@291287591
Prok_Prote_Hippea_maritima@327398783
Prok_Nitro_Candidatus_Nitrospira@302036717
Prok_Verru_bacterium_Ellin514@223935394
Prok_Verru_bacterium_Ellin514@223937746
Prok_Verru_Methylacidiphilum_infernorum@189219215
-/.90
Dehalobacter_sp__FTH1@521977049
Prok_Actin_Propionibacterium_freudenreichii@297626757
Prok_Actin_Sanguibacter_keddieii@269794851
Prok_Actin_Thermobifida_fusca@72161392
Prok_Actin_Thermomonospora_curvata@269127312
Prok_Uncla_Thermobaculum_terrenum@269926496
Prok_Chlof_Chloroflexus_aurantiacus@163847021
Prok_Chlof_Roseiflexus_sp@148657092
Prok_Firmi_Ammonifex_degensii@260893724
Prok_Firmi_Pelotomaculum_thermopropionicum@147677244
Prok_Eurya_Methanoregula_boonei@154151487
Prok_Eurya_Methanocella_paludicola@282165134
Prok_Eurya_uncultured_methanogenic@147920818
Methanomassiliicoccus_luminyensis@648287774
Prok_Firmi_Paenibacillus_sp@334138429
Prok_Firmi_Thermobacillus_composti@350547333
Kyrpidia_tusciae_DSM_2912@502839561
-/.98
99/1
Prok_Syner_Pyramidobacter_piscolens@282857511
Prok_Syner_Synergistes_sp@365172330
Prok_Syner_Thermanaerovibrio_acidaminovorans@269793146
Prok_Therm_Thermosipho_africanus@217077895
Prok_Therm_Thermosipho_melanesiensis@150021510
Prok_Therm_Fervidobacterium_pennivorans@358452111
Prok_Uncla_Caldithrix_abyssi@373459370
Prok_Uncla_Candidatus_Cloacamonas@218960946
100/1
Prok_Crena_Acidianus_hospitalis@332796953
Prok_Crena_Sulfolobus_tokodaii@15921675
Prok_Crena_Metallosphaera_sedula@146304375
Prok_Thaum_Nitrosopumilus_maritimus@161529298
Prok_Thaum_Cenarchaeum_symbiosum@118577024
Prok_Thaum_Candidatus_Nitrosoarchaeum@329765549
Coccomyxa_subellipsoidea_C169_Nu_120407@4616
Chlorella_variabilis_NC64A_Nu_120407@141615
Nannochloropsis_gaditana_CCMP526_Nu_120407@Nga00113
Physcomitrella_patens_patens_Nu_120408@18061266
Physcomitrella_patens_patens_Nu_120408@18072702
Arabidopsis_thaliana_Nu_110103@AT3G53180_1
Spizellomyces_punctatus_DAOM_BR117_Nu_130412@SPPG_08358T0
Sterkiella_histriomuscorum_Nu_130327@403339489
Sterkiella_histriomuscorum_Nu_130327@403358667
Dictyostelium_discoideum_AX4_Nu_130324@DDB0231552
Perkinsus_marinus_ATCC50983_Nu_120716@294897206
Bigelowiella_natans_CCMP2755_Nu_120716@137622
Symbiodinium_sp_kb8@kb8_rep_c4055_18_1652_150
Lingulodinium_polyedrum_cDNA_131020@346255634_7_15_1511
Vitrella_brassicaformis@MMETSP0288_20130426_22020_1_464
Aspergillus_fumigatus_A1163_Nu_120716@AFUB_094760
Rhizopus_oryzae_Nu_120716@RO3T_15104
Spizellomyces_punctatus_DAOM_BR117_Nu_130412@SPPG_07707T0
Dictyostelium_discoideum_AX4_Nu_130324@DDB0252590
Physcomitrella_patens_patens_Nu_120408@18071315
Chlorokybus_atmophyticus_UTEX2591_Nu@341671936_1_196_897
Aplanochytrium_kerguelense_PBS07_140414@104875
97/.96
Symbiodinium_sp_kb8@kb8_c6635_15_1382_3
Symbiodinium_minutum_Nu_v1_2_131001@031372_t1
Physcomitrella_patens_patens_Nu_120408@18039108
Guillardia_theta_CCMP2712_Nu_120408@158145
Bigelowiella_natans_CCMP2755_Nu_120716@36472
Emiliania_huxleyi_CCMP1516_Nu_120408@52323
Emiliania_huxleyi_CCMP1516_Nu_120408@72956
Aureococcus_anophagefferens_CCMP1984_Nu_120407@55217
73/.94
Naegleria_gruberi_NEG_M_Nu_120716@52785
Prok_Deino_Meiothermus_silvanus@297566879
Prok_Deino_Truepera_radiovictrix@297623208
Prok_Deino_Deinococcus_geothermalis@94985306
Prok_Deino_Deinococcus_maricopensis@320334998
Prok_Bacte_Salinibacter_ruber@83815855
Prok_Uncla_Thermobaculum_terrenum@269926495
Prok_Nitro_Thermodesulfovibrio_yellowstonii@206889333
Prok_Syner_Synergistes_sp@365165645
Prok_Eurya_Methanocaldococcus_fervens@256810145
0.5
Apicomplexans
Chrompodellids
Dinoflagellates+
Perkinsus
Other eukaryotes

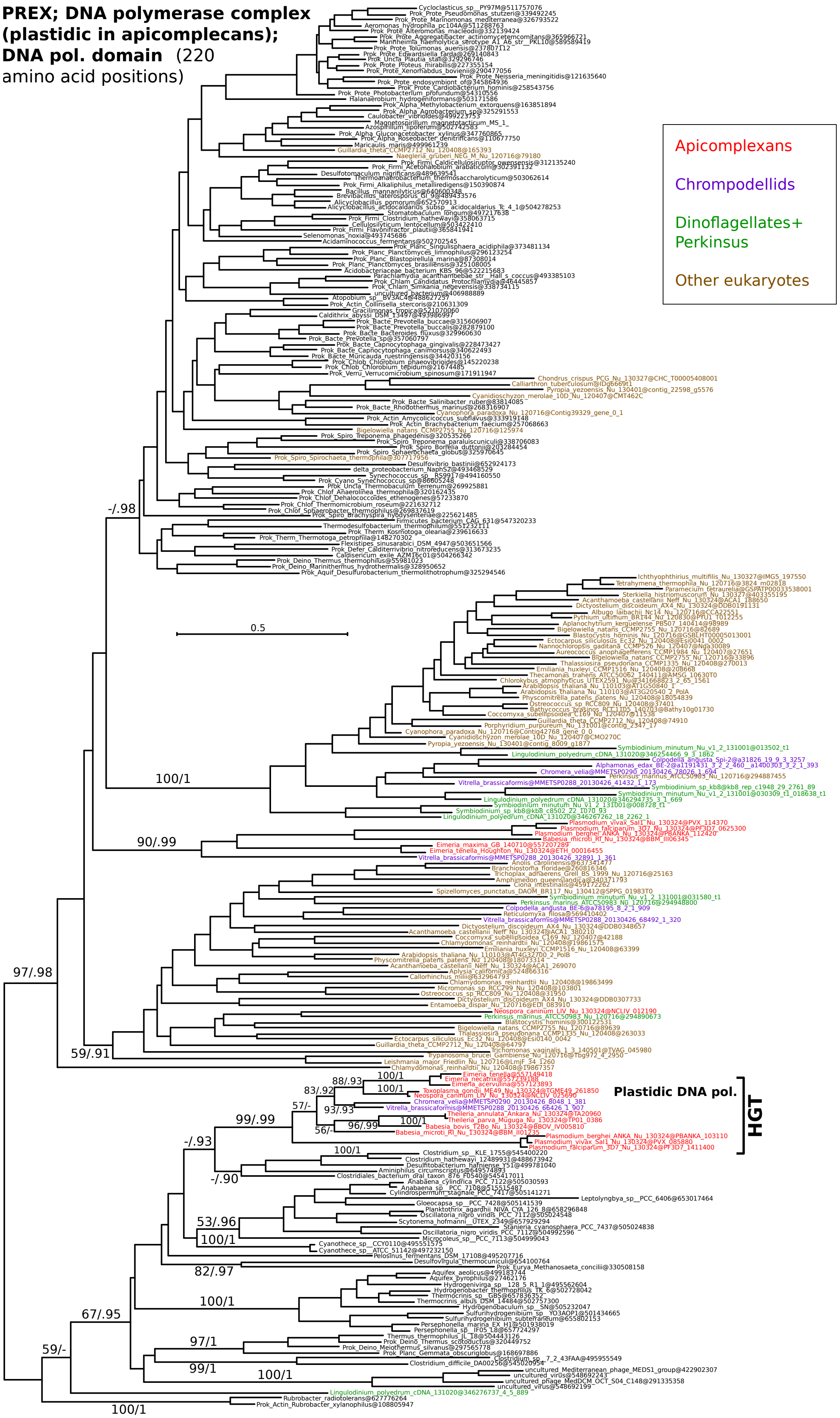
IspH (LytB); 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (plastidic)
(235 amino acid positions)



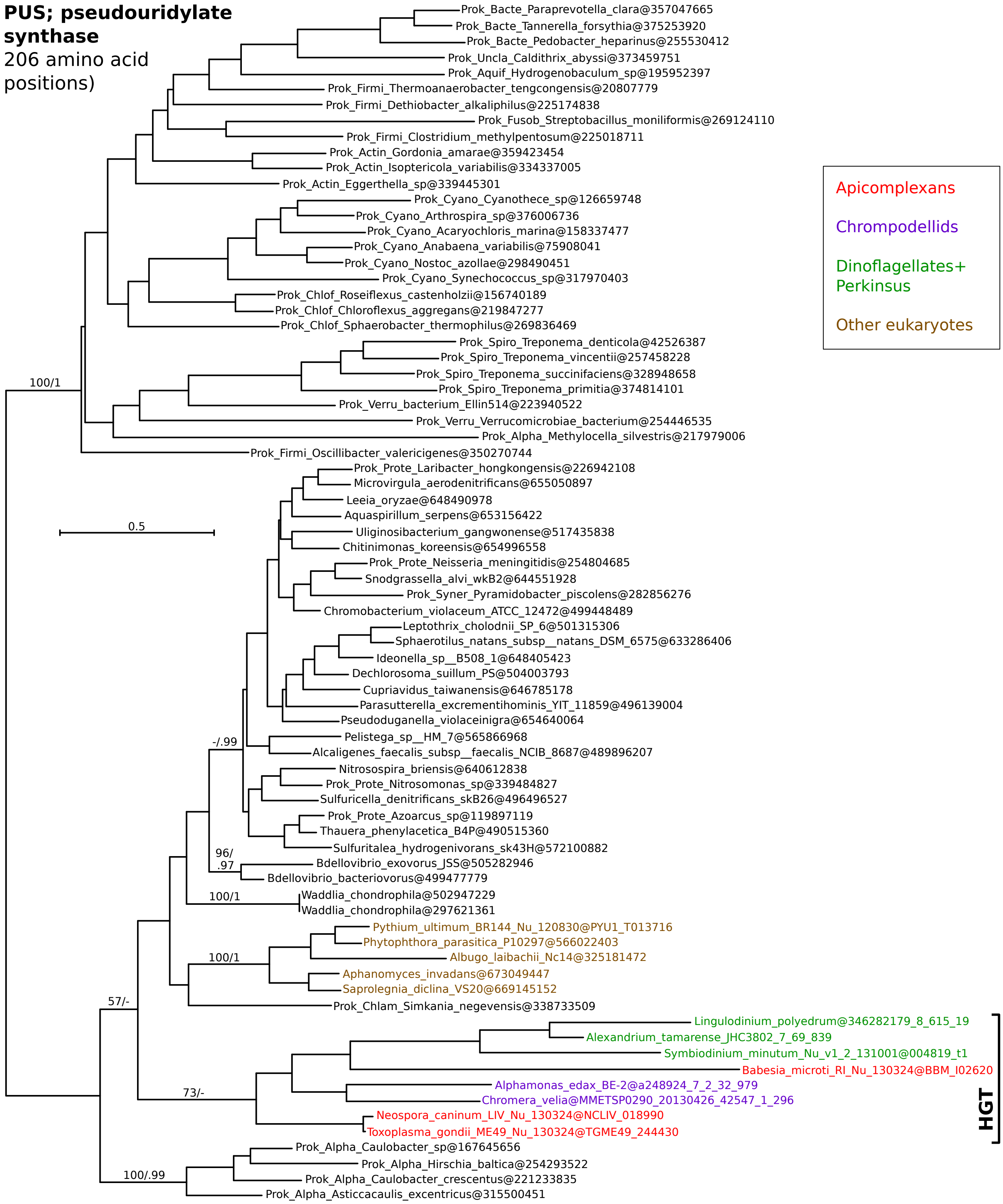
MDH/LDH; malate/lactate dehydrogenase
(277 amino acid positions)




PNP; purine-nucleoside phosphorylase
(157 amino acid positions)
Theileria_parva_Muguga_Nu_130324@TP02_0438
Theileria_parva_Muguga_Nu_130324@TP02_0440
Theileria_parva_Muguga_Nu_130324@TP02_0437
Theileria_annulata_Ankara_Nu_130324@TA13165
Theileria_annulata_Ankara_Nu_130324@TA13175
Babesia_bovis_T2Bo_Nu_130324@BBOV_III005140
Plasmodium_falciparum_3D7_Nu_130324@PF3D7_0513300
Plasmodium_vivax_Sal1_Nu_130324@PVX_080575
Plasmodium_berghei_ANKA_Nu_130324@PBANKA_111300
Toxoplasma_gondii_ME49_Nu_130324@TGME49_307030
Neospora_caninum_LIV_Nu_130324@NCLIV_053130
Eimeria_tenella_Houghton_Nu_130324@ETH_00009730
Voromonas_pontica_GG@comp2883_c0_seq1_5
Vitrella_brassicaformis@MMETSP0288_20130426_1847_1_244
Chromera_velia@MMETSP0290_20130426_6600_1_262
Alphamonas_edax_BE-2@a1062399_4_1_1_168__a2661_547_3_2_400
Ascogregarina_taiwanensis@284817201_6
Perkinsus_marinus_ATCC50983_Nu_120716@294872220
Perkinsus_marinus_ATCC50983_Nu_120716@294934106
Perkinsus_marinus_ATCC50983_Nu_120716@294937198
Perkinsus_marinus_ATCC50983_Nu_120716@294930273
Perkinsus_marinus_ATCC50983_Nu_120716@294891321
HGT
78/.98
73/.99
-/.91
Prok_Firmi_Paenibacillus_curdlanolyticus@304406368
Brevibacillus_borstelensis@640551663
Dictyostelium_discoideum_AX4_Nu_130324@DDB0230169
Prok_Actin_Streptomyces_avermitilis@29833195
Arthrobacter_castelli@652423567
Actinobaculum_schaalii_FB123_CNA_2@512536900
Leucobacter_salsicius@516402811
Clostridium_hathewayi_12489931@488673224
Clostridium__hathewayi_12489931@488672753
Prok_Fusob_Leptotrichia_goodfellowii@262037472
Cyanophora_paradoxa_Nu_120716@Contig55685_gene_0_1
Sulfolobus_solfataricus@16975003_pdb_1JDS_5__Deoxy_5__Methylthioadenosine_Phosphorylase
Candidatus_Korarchaeum_cryptofilum_OPF8@501265921
Entamoeba_histolytica@345531905_pdb_3TL6_PNP
Entamoeba_dispar_Nu_120716@EDI_172750
Gregarina_niphandrodes_Nu_130324@AFNH01002609_6_4434_3655
Clostridium_sp__CAG_1024@512181624
Acidiphilium_sp__CAG_727@548199088
Corallococcus_sp__CAG_1435@547328882
Anaerotruncus_sp__CAG_390@548183116
Proteocatella_sphenisci@655445756
Prok_Tener_Mycoplasma_mobile@47459370
Prok_Uncla_Haloplasma_contractile@335429642
Trichomonas_vaginalis_1_3_140501@TVAG_454490
Trichomonas_Vaginalis@62738888_pdb_1Z33_PNP
Escherichia_coli_K_12@1633413_pdb_1ECP_PNP
Prok_Alpha_Roseobacter_denitrificans@110678940
Prok_Bacte_Gillisia_limnaea@374597254
Prok_Bacte_Psychroflexus_torquis@91214890
Bacillus_cereus@112489887_pdb_2AC7_Adenosine_Phosphorylase
Prok_Alpha_Mesorhizobium_ciceri@319783769
Prok_Actin_Rubrobacter_xylanophilus@108803452
Prok_Deino_Thermus_aquaticus@218296503
Thermus_thermophilus@28949111_pdb_1ODI_PNP
Prok_Fusob_Fusobacterium_ulcerans@317062874
Streptococcus_pyogenes@344189652_pdb_3QPB_Uridine_Phosphorylase
Prok_Spiro_Sphaerochaeta_pleomorpha@374314778
Prok_Spiro_Spirochaeta_smaragdinae@302340393
Prok_Actin_Olsenella_sp@355576270
Prok_Prote_Shewanella_baltica@371996512
unclassified_Atribacteria@516995249
Atribacteria_bacterium_SCGC_AAA255_N14@661260061
Prok_Actin_Olsenella_uli@302336027
Prok_Therm_Petrotoga_mobilis@160901788
Prok_Therm_Thermotoga_lettingae@157364281
Voromonas_pontica_GG@comp19549_c0_seq1_10
Chromera_velia@MMETSP0290_20130426_83375_1_362
Colpodella_angusta_BE-6@a6129_139_12_1598_528
Vitrella_brassicaformis@MMETSP0288_20130426_33241_1_103
Alphamonas_edax_BE-2@a7229_174_10_1115_372
Toxoplasma_gondii_ME49_Nu_130324@TGME49_235000
Eimeria_tenella_Houghton_Nu_130324@ETH_00019830
Ascogregarina_taiwanensis@284817137_3
Eimeria_tenella_Houghton_Nu_130324@ETH_00030945
Toxoplasma_gondii_ME49_Nu_130324@TGME49_310640
Gregarina_niphandrodes_Nu_130324@AFNH01001968_6_1045_182
Emiliania_huxleyi_CCMP1516_Nu_120408@63935
Symbiodinium_sp_mf105@mf105_rep_c9625_3_36_896
Lingulodinium_polyedrum_cDNA_131020@346289865_6_614_3
Nannochloropsis_gaditana_CCMP526_Nu_120407@Nga02776
Giardia_intestinalis_3_1A_140501
@GL50803_9779
Acanthamoeba_castellanii_Neff_Nu_130324@ACA1_187000
Pythium_ultimum_BR144_Nu_120830@PYU1_T006356
Chlamydomonas_reinhardtii_Nu_120408@19858721
Chlorella_variabilis_NC64A_Nu_120407@137869
Ectocarpus_siliculosus_Ec32_Nu_120408@Esi0226_0037
Allomyces_macrogynus_ATCC38327_140415@AMAG_06689T0
Ustilago_maydis_Nu_120716@UM03504
Dictyostelium_discoideum_AX4_Nu_130324@DDB0230170
Naegleria_gruberi_NEG_M_Nu_120716@45426
Trichomonas_vaginalis_1_3_140501@TVAG_020130
Trichomonas_vaginalis_1_3_140501@TVAG_256820
100/1
0.5
Geobacter_uraniireducens_Rf4@500470641
gamma_proteobacterium_SCGC_AAA076_F14@518872292
Trypanosoma_brucei@163311122_pdb_3BJE_Uridine_Phosphorylase_Activity
Plesiocystis_pacifica_SIR_1@494031104
Desulfobacca_acetoxidans_DSM_11109@503470355
Prok_Syner_Pyramidobacter_piscolens@282857244
Prok_Syner_Jonquetella_anthropi@260654195
Prok_Crena_Thermoproteus_tenax@352682976
Prok_Crena_Staphylothermus_marinus@126465373
Prok_Chlof_Roseiflexus_castenholzii@156742305
Prok_Cyano_Nostoc_punctiforme@186686567
Prok_Cyano_Microcoleus_chthonoplastes@254410112
Prok_Cyano_Microcystis_aeruginosa@166364369
Prok_Cyano_Cyanothece_sp@307152293
Shewanella_oneidensis@550544451_pdb_4HER_A
Escherichia_coli@16975349_pdb_1K3F_Uridine_Phosphorylase
Actinobacillus_succinogenes@675082751_pdb_4R31_Putative_Uridine_Phosphorylase
Prok_Eurya_Halorubrum_lacusprofundi@222480727
Prok_Eurya_Natronobacterium_gregoryi@352122927
Prok_Therm_Thermotogales_bacterium@307298392
Apicomplexans
Chrompodellids
Dinoflagellates+
Perkinsus
Other eukaryotes

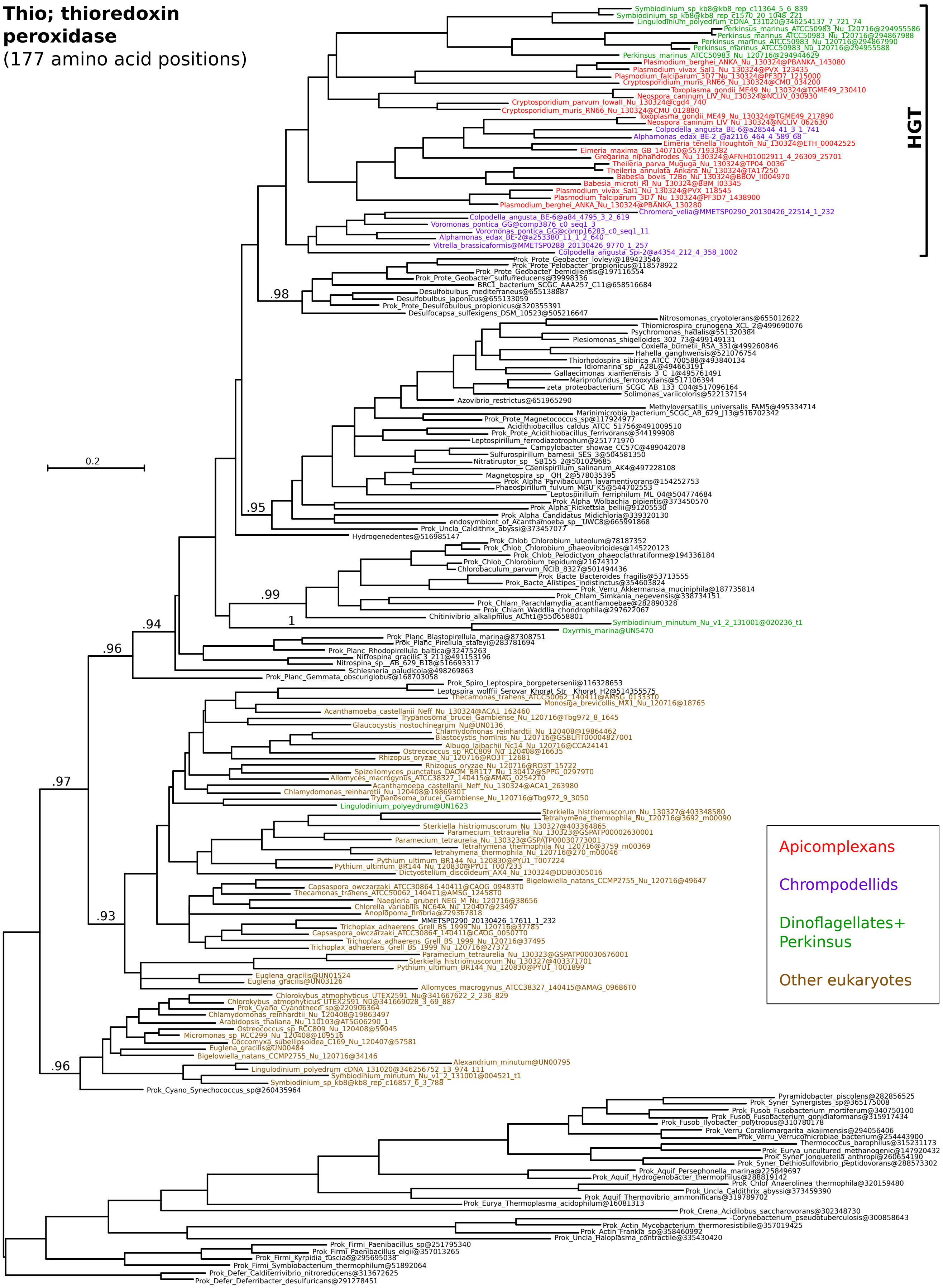


PREX; DNA polymerase complex (plastidic in apicomplecans); DNA pol. domain (220 amino acid positions)


PUS; pseudouridylate synthase
206 amino acid positions)
Apicomplexans
Chrompodellids
Dinoflagellates+ Perkinsus
Other eukaryotes
0.5
100/1
Prok_Bacte_Paraprevotella_clara@357047665
Prok_Bacte_Tannerella_forsythia@375253920
Prok_Bacte_Pedobacter_heparinus@255530412
Prok_Uncla_Caldithrix_abyssi@373459751
Prok_Aquif_Hydrogenobaculum_sp@195952397
Prok_Firmi_Thermoanaerobacter_tengcongensis@20807779
Prok_Firmi_Dethiobacter_alkaliphilus@225174838
Prok_Fusob_Streptobacillus_moniliformis@269124110
Prok_Firmi_Clostridium_methylpentosum@225018711
Prok_Actin_Gordonia_amarae@359423454
Prok_Actin_Isoptericola_variabilis@334337005
Prok_Actin_Eggerthella_sp@339445301
Prok_Cyano_Cyanothece_sp@126659748
Prok_Cyano_Arthrospira_sp@376006736
Prok_Cyano_Acaryochloris_marina@158337477
Prok_Cyano_Anabaena_variabilis@75908041
Prok_Cyano_Nostoc_azollae@298490451
Prok_Cyano_Synechococcus_sp@317970403
Prok_Chlof_Roseiflexus_castenholzii@156740189
Prok_Chlof_Chloroflexus_aggregans@219847277
Prok_Chlof_Sphaerobacter_thermophilus@269836469
Prok_Spiro_Treponema_denticola@42526387
Prok_Spiro_Treponema_vincentii@257458228
Prok_Spiro_Treponema_succinifaciens@328948658
Prok_Spiro_Treponema_primitia@374814101
Prok_Verru_bacterium_Ellin514@223940522
Prok_Verru_Verrucomicrobiae_bacterium@254446535
Prok_Alpha_Methylocella_silvestris@217979006
Prok_Firmi_Oscillibacter_valericigenes@350270744
Prok_Prote_Laribacter_hongkongensis@226942108
Microvirgula_aerodenitrificans@655050897
Leeia_oryzae@648490978
Aquaspirillum_serpens@653156422
Uliginosibacterium_gangwonense@517435838
Chitinimonas_koreensis@654996558
Prok_Prote_Neisseria_meningitidis@254804685
Snodgrassella_alvi_wkB2@644551928
Prok_Syner_Pyramidobacter_piscolens@282856276
Chromobacterium_violaceum_ATCC_12472@499448489
Leptothrix_cholodnii_SP_6@501315306
Sphaerotilus_natans_subsp__natans_DSM_6575@633286406
Ideonella_sp__B508_1@648405423
Dechlorosoma_suillum_PS@504003793
Cupriavidus_taiwanensis@646785178
Parasutterella_excrementihominis_YIT_11859@496139004
Pseudoduganella_violaceinigra@654640064
-/.99
Pelistega_sp__HM_7@565866968
Alcaligenes_faecalis_subsp__faecalis_NCIB_8687@489896207
Nitrosospira_briensis@640612838
Prok_Prote_Nitrosomonas_sp@339484827
Sulfuricella_denitrificans_skB26@496496527
Prok_Prote_Azoarcus_sp@119897119
Thauera_phenylacetica_B4P@490515360
Sulfuritalea_hydrogenivorans_sk43H@572100882
96/ .97
Bdellovibrio_exovorus_JSS@505282946
Bdellovibrio_bacteriovorus@499477779
100/1
Waddlia_chondrophila@502947229
Waddlia_chondrophila@297621361
Pythium_ultimum_BR144_Nu_120830@PYU1_T013716
Phytophthora_parasitica_P10297@566022403
100/1
Albugo_laibachii_Nc14@325181472
Aphanomyces_invadans@673049447
Saprolegnia_diclina_VS20@669145152
57/-
Prok_Chlam_Simkania_negevensis@338733509
Lingulodinium_polyedrum@346282179_8_615_19
Alexandrium_tamarense_JHC3802_7_69_839
Symbiodinium_minutum_Nu_v1_2_131001@004819_t1
Babesia_microti_RI_Nu_130324@BBM_I02620
73/-
Alphamonas_edax_BE-2@a248924_7_2_32_979
Chromera_velia@MMETSP0290_20130426_42547_1_296
Neospora_caninum_LIV_Nu_130324@NCLIV_018990
Toxoplasma_gondii_ME49_Nu_130324@TGME49_244430
HGT
Prok_Alpha_Caulobacter_sp@167645656
Prok_Alpha_Hirschia_baltica@254293522
100/.99
Prok_Alpha_Caulobacter_crescentus@221233835
Prok_Alpha_Asticcacaulis_excentricus@315500451

Thio; thioredoxin peroxidase

(177 amino acid positions)



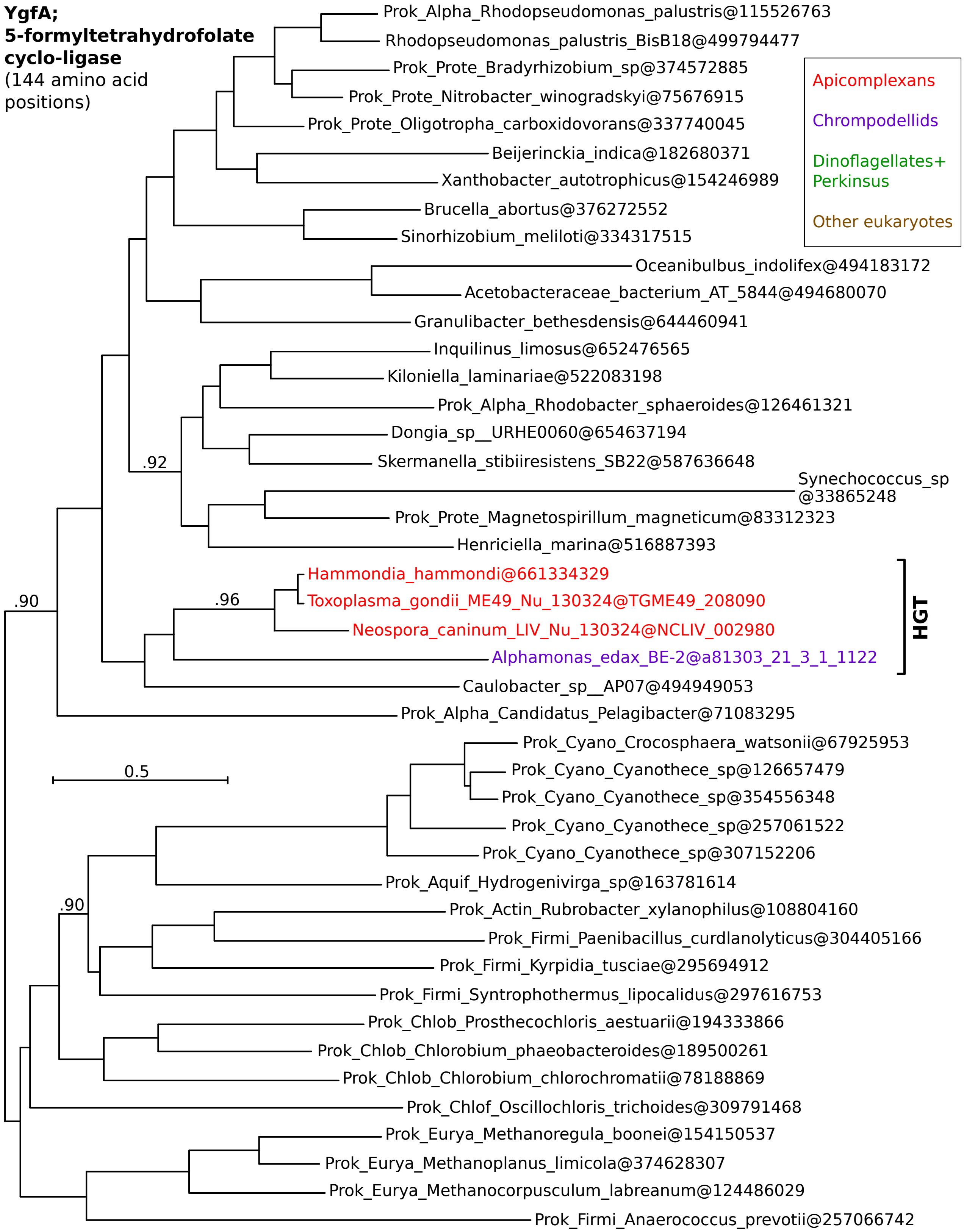

YgfA;
5-formyltetrahydrofolate
cyclo-ligase
(144 amino acid
positions)
Apicomplexans
Chrompodellids
Dinoflagellates+
Perkinsus
Other eukaryotes
Prok_Alpha_Rhodopseudomonas_palustris@115526763
Rhodopseudomonas_palustris_BisB18@499794477
Prok_Prote_Bradyrhizobium_sp@374572885
Prok_Prote_Nitrobacter_winogradskyi@75676915
Prok_Prote_Oligotropha_carboxidovorans@337740045
Beijerinckia_indica@182680371
Xanthobacter_autotrophicus@154246989
Brucella_abortus@376272552
Sinorhizobium_meliloti@334317515
Oceanibulbus_indolifex@494183172
Acetobacteraceae_bacterium_AT_5844@494680070
Granulibacter_bethesdensis@644460941
Inquilinus_limosus@652476565
Kiloniella_laminariae@522083198
Prok_Alpha_Rhodobacter_sphaeroides@126461321
Dongia_sp__URHE0060@654637194
.92
Skermanella_stibiiresistens_SB22@587636648
Synechococcus_sp
@33865248
Prok_Prote_Magnetospirillum_magneticum@83312323
Henriciella_marina@516887393
Hammondia_hammondi@661334329
.96
Toxoplasma_gondii_ME49_Nu_130324@TGME49_208090
.90
Neospora_caninum_LIV_Nu_130324@NCLIV_002980
Alphamonas_edax_BE-2@a81303_21_3_1_1122
HGT
Caulobacter_sp__AP07@494949053
Prok_Alpha_Candidatus_Pelagibacter@71083295
Prok_Cyano_Crocosphaera_watsonii@67925953
0.5
Prok_Cyano_Cyanothece_sp@126657479
Prok_Cyano_Cyanothece_sp@354556348
Prok_Cyano_Cyanothece_sp@257061522
Prok_Cyano_Cyanothece_sp@307152206
Prok_Aquif_Hydrogenivirga_sp@163781614
.90
Prok_Actin_Rubrobacter_xylanophilus@108804160
Prok_Firmi_Paenibacillus_curdlanolyticus@304405166
Prok_Firmi_Kyrpidia_tusciae@295694912
Prok_Firmi_Syntrophothermus_lipocalidus@297616753
Prok_Chlob_Prosthecochloris_aestuarii@194333866
Prok_Chlob_Chlorobium_phaeobacteroides@189500261
Prok_Chlob_Chlorobium_chlorochromatii@78188869
Prok_Chlof_Oscillochloris_trichoides@309791468
Prok_Eurya_Methanoregula_boonei@154150537
Prok_Eurya_Methanoplanus_limicola@374628307
Prok_Eurya_Methanocorpusculum_labreanum@124486029
Prok_Firmi_Anaerococcus_prevotii@257066742

Figure S6: Phylogeny of ribonucleotide reductase.

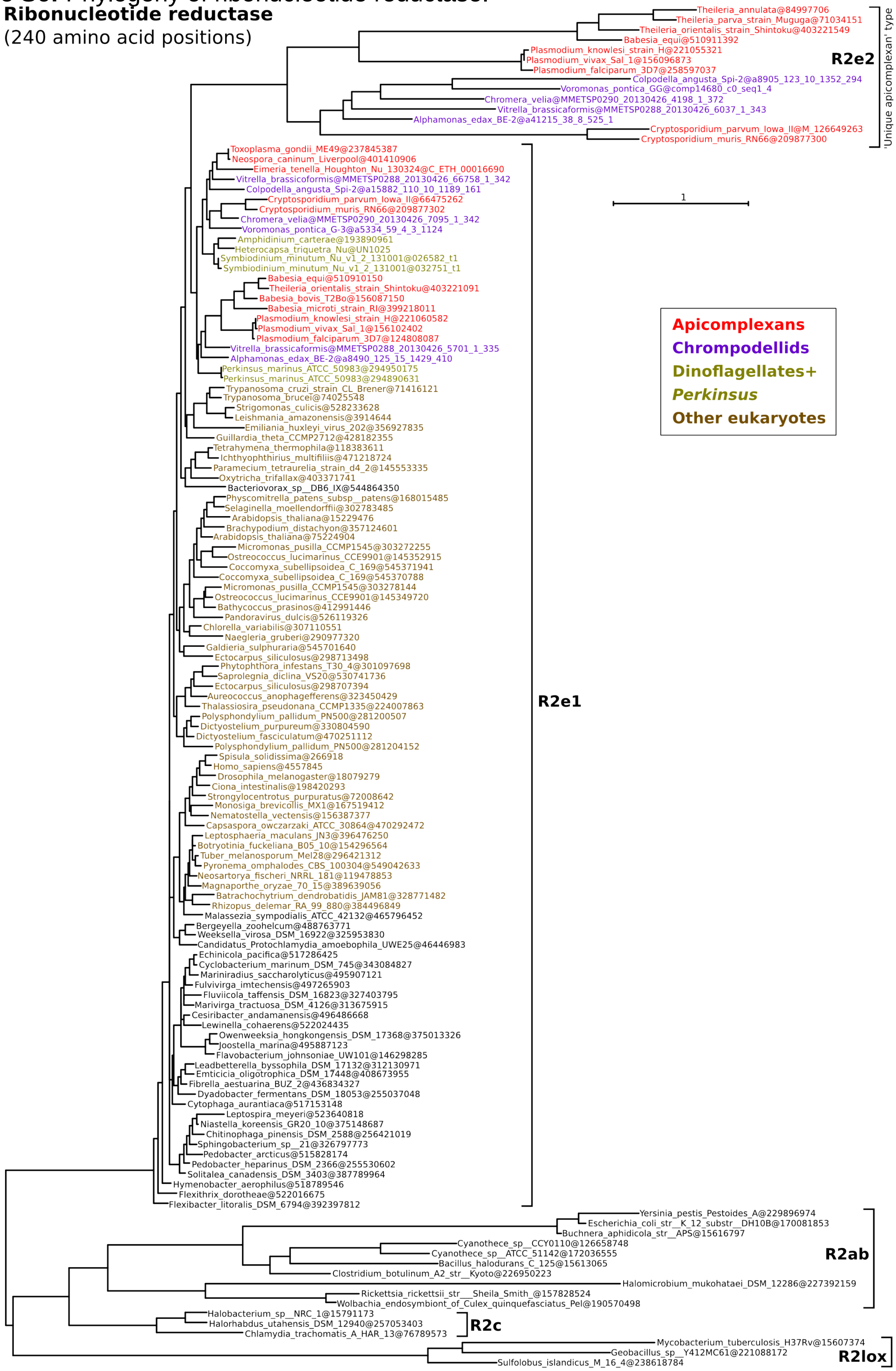


Figure S7: Phylogeny of subtilisin proteases.

Table S1: Environmental sequences related to apicomplexans that were used in this study.

Figure	Clone name	NCBI GI or Accession	Enviroment type[1]	Low O2[2]	Isolation source; Location	Study[3]	Clade affiliation[4]
Fig. 1	P107t1c2	KP213182	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Env. clade IX
Fig. 1	P107t1c6	KP213184	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Env. clade IX
Fig. 1	P107t6c1	KP213185	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Colpodella angusta
Fig. 1	P107t6c6	KP213186	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Colpodella angusta
Fig. 1	P108t7c1	KP213187	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Colpodella angusta
Fig. 1	P108t7c2	KP213188	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Env. clade IX
Fig. 1	P109t4c2	KP213191	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Colpodella angusta
Fig. 1	P109t4c6	KP213192	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Colpodella angusta
Fig. 1	P113t2c1	KP213195	marine		woodchip on the beach; Canada: Boundary Bay, BC	THIS STUDY	Colpodella angusta
Fig. 1	P113t2c6	KP213196	marine		woodchip on the beach; Canada: Boundary Bay, BC	THIS STUDY	Colpodella angusta
Fig. 1	P114t10c3	KP213197	marine		mucus from Acropora formosa; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Colpodella angusta
Fig. 1	P114t13c5	KP213200	marine		coral Acropora yongei; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Env. clade IX
Fig. 1	P114t9c5	KP213202	marine		mucus from Acropora yongei; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Env. clade IX
Fig. 1	P115t1c2	KP213203	marine		woodchip on the beach; Canada: Boundary Bay, BC	THIS STUDY	Chromera velia
Fig. 1	P115t1c3	KP213204	marine		woodchip on the beach; Canada: Boundary Bay, BC	THIS STUDY	Colpodella angusta
Fig. 1	P116t19c4	KP213207	marine		JAR_04_2010_ARMS_A_Filter corners	THIS STUDY	Env. clade VIII
Fig. 1	P116t19c5	KP213208	marine		JAR_04_2010_ARMS_A_Filter corners	THIS STUDY	Env. clade VIII
Fig. 1	P116t19c7	KP213209	marine		JAR_04_2010_ARMS_A_Filter corners	THIS STUDY	Unresolved: Api.
Fig. 1	P122t3c2	KP213211	marine		coral Echinophyllia sp.; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Env. clade V
Fig. 1	P122t3c3	KP213212	marine		coral Echinophyllia sp.; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Env. clade V
Fig. 1	P122t3c4	KP213213	marine		coral Echinophyllia sp.; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Env. clade V
Fig. 1	P23t21c1	KP213214	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Unresolved:Api. relatives
Fig. 1	P23t21c2	KP213215	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Unresolved:Api. relatives
Fig. 1	P23t21c5	KP213216	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Unresolved:Api. relatives
Fig. 1	P23t5c3	KP213217	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Unresolved:Api. relatives
Fig. 1	P23t5c7	KP213218	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Unresolved:Api. relatives
Fig. 1	P23t5c8	KP213219	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Unresolved:Api. relatives
Fig. 1	P43t2c8	KP213220	marine		rocky intertidal spit; Canada: Spanish Banks, Vancouver, BC	THIS STUDY	Voromonas pontica
Fig. 1	P46t3c16	KP213221	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Unresolved: Api.
Fig. 1	P56t2c1	KP213222	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Env. clade VI
Fig. 1	P57t7c8	KP213226	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Env. clade VI
Fig. 1	P71t4c5	KP213228	marine		rocky intertidal spit; Canada: Spanish Banks, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P71t4c7	KP213229	marine		rocky intertidal spit; Canada: Spanish Banks, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P79t3c2	KP213230	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Unresolved:Api. relatives
Fig. 1	P79t3c8	KP213231	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Colpodella angusta

Fig. 1	P83t1c3	KP213232	marine		rocky intertidal spit; Canada: Spanish Banks, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P85t4c2	KP213233	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Colpodella angusta
Fig. 1	P85t4c5	KP213234	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Colpodella angusta
Fig. 1	P87t1c4	KP213235	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Chromera velia
Fig. 1	P87t2c1	KP213236	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Env. clade VI
Fig. 1	P87t2c5	KP213238	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Chromera velia
Fig. 1	P87t3c4	KP213239	marine		rocky intertidal spit; Canada: Spanish Banks, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P87t4c1	KP213240	marine		rocky intertidal spit; Canada: Spanish Banks, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P87t4c2	KP213241	marine		rocky intertidal spit; Canada: Spanish Banks, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P87t5c2	KP213242	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P87t5c3a	KP213243	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P87t5c4	KP213245	marine		damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Chromera velia
Fig. 1	P87t8c3	KP213246	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Env. clade VI
Fig. 1	P87t8c4	KP213247	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Env. clade VI
Fig. 1	P97t1c1	KP213251	marine		Dry high-beach sand; Canada: Boundary Bay, BC	THIS STUDY	Env. clade IX
Fig. 1	P97t5c2	KP213256	freshwater		freshwater duck pond near the beach; Canada: Boundary Bay, BC	THIS STUDY	Chromera velia
Fig. 1	P97t6c3	KP213258	marine		woodchip on the beach; Canada: Boundary Bay, BC	THIS STUDY	Voromonas pontica
Fig. 1	6	472401464	animal feces/manure	YES	Megalapteryx didinus coprolite, sample 10198a; New Zealand: Dart River Valley	PLoS ONE 8(2): E57315 (2013)	Colpodella angusta
Fig. 1	7	472401465	animal feces/manure	YES	Megalapteryx didinus coprolite, sample 10198a; New Zealand: Dart River Valley	PLoS ONE 8(2): E57315 (2013)	Colpodella angusta
Fig. 1	56059	6606066	marine		European oyster, Ostrea edulis	Unpublished	Unresolved: Api.
Fig. 1	1_133	171199596	marine		marine sediment; East Sea	J. Microbiol. 46 (3), 244-249 (2008)	Env. clade I
Fig. 1	1_41	156466089	marine		marine sediment; East Sea	J. Microbiol. 46 (3), 244-249 (2008)	Env. clade I
Fig. 1	1_52	156466099	marine		marine sediment; East Sea	J. Microbiol. 46 (3), 244-249 (2008)	Env. clade I
Fig. 1	1_73	156466110	marine		marine sediment; East Sea	J. Microbiol. 46 (3), 244-249 (2008)	Env. clade I
Fig. 1	9_12	156466122	marine		marine sediment; East Sea	J. Microbiol. 46 (3), 244-249 (2008)	Env. clade I
Fig. 1	9_39	156466143	marine		marine sediment; East Sea	J. Microbiol. 46 (3), 244-249 (2008)	Env. clade I
Fig. 1	9_61	156466155	marine		marine sediment; East Sea	J. Microbiol. 46 (3), 244-249 (2008)	Env. clade VIII
Fig. 1	A_1_87	304361468	freshwater		paddy field soil; Japan: Aomori	Microbes Environ. 25 (4), 281-287 (2010)	Unresolved:Api. relatives
Fig. 1	A2_E039	18025924	marine	YES	hydrothermal sediment; Guaymas Basin	PNAS 99 (11), 7658-62 (2002)	Unresolved: Api.
Fig. 1	AA5F13RJ1A10	291497875	marine		oxygenated water column; Cariaco Basin, Caribbean Sea	ISME J 5 (8), 1344-1356 (2011)	Unresolved:Api. relatives
Fig. 1	AE-2012_FL1	411025944	freshwater		laboratory dishes with mosquito larvae; USA: Rutgers university	Aquat.Ecol.:10.1007/s10452-014-9474-3(2014)	Colpodella angusta
Fig. 1	ANEK03F02	345648463	soil		soil in front of the Brazilian Antarctic Station; Antarctica	Unpublished	Colpodella angusta
Fig. 1	BAQA40	20377996	marine	YES	anoxic marine sediment; USA: Berkeley Aquatic Park, Berkeley, CA	PNAS 99 (12), 8324-29 (2002)	Env. clade VII
Fig. 1	BCB5F13RM1D02	291498055	marine	YES	anoxic water column sample; Caribbean Sea: Cariaco Basin	ISME J 5 (8), 1344-1356 (2011)	Api.: genotype N
Fig. 1	BCB3F14RJ3H04	291495817	marine	YES	anoxic water column sample; Caribbean Sea: Cariaco Basin	ISME J 5 (8), 1344-1356 (2011)	Env. clade X
Fig. 1	BCB5F14RJ1F07	291495355	marine	YES	anoxic water column sample; Caribbean Sea: Cariaco Basin	ISME J 5 (8), 1344-1356 (2011)	Env. clade X
Fig. 1	BIO10_B6	270381369	marine	YES	deep water, depth 500m; Pacific Ocean: South Pacific	Environ. Microbiol. 12 (11), 2946-64 (2010)	Voromonas pontica
Fig. 1	BIO3_B2	270381403	marine	YES	deep water, depth 500m; Pacific Ocean: South Pacific	Environ. Microbiol. 12 (11), 2946-64 (2010)	Voromonas pontica

Fig. 1	BOLA176	20378006	marine	YES	anoxic marine sediment; USA: Bolinas Tidal Flat, Bolinas, CA	PNAS 99 (12), 8324-29 (2002)	Colpodella angusta
Fig. 1	BOLA553	20378005	marine	YES	anoxic marine sediment; USA: Bolinas Tidal Flat, Bolinas, CA	PNAS 99 (12), 8324-29 (2002)	Env. clade V
Fig. 1	BOLA566	20378000	marine	YES	anoxic marine sediment; USA: Bolinas Tidal Flat, Bolinas, CA	PNAS 99 (12), 8324-29 (2002)	Env. clade VI
Fig. 1	BOLA914	20377992	marine	YES	anoxic marine sediment; USA: Bolinas Tidal Flat, Bolinas, CA	PNAS 99 (12), 8324-29 (2002)	Env. clade IV
Fig. 1	Buttons_D03	300088832	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	C1_E039	18025812	marine	YES	hydrothermal sediment; Guaymas Basin	PNAS 99 (11), 7658-62 (2002)	Api.:Polyplicarum -related
Fig. 1	C1_E047	18025817	marine	YES	hydrothermal sediment; Guaymas Basin	PNAS 99 (11), 7658-62 (2002)	Api.:Polyplicarum -related
Fig. 1	C17_250P117_0239Tplate02_P2	389549566	marine (*)		marine sediment from shallow water hydrothermal system; Papua New Guinea	Pacific Science 64 (3), 391–404 (2010)	Vitrella brassicaformis
Fig. 1	CCI31	27802552	marine	YES	oxygen-depleted marine env.; USA: Great Sippewisset salt marsh, Cape Cod, MA	Appl. Environ. Microbiol. 69(5),2657-63(2003)	Api.:Polyplicarum -related
Fig. 1	CPSGM-1	133778656	marine	YES	deep-sea methane cold seep sediment; Japan: Sagami Bay	Extremophiles 11 (4), 563-576 (2007)	Unresolved: Api.
Fig. 1	'Cryptosporidium struthionis'	58371885	animal feces/manure	YES	oocyst in Struthio camelus feces; Spain	Unpublished	Colpodella angusta
Fig. 1	'Cryptosporidium sp.' NEV10	342918948	animal feces/manure	YES	feces of calves with diarrhea; Turkey: Nevsehir	Unpublished	Colpodella angusta
Fig. 1	D1P01G01	118420045	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Api.:Polyplicarum -related
Fig. 1	D1P01H03	118420061	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Voromonas pontica
Fig. 1	D1P02A01	118420063	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Api.:Polyplicarum -related
Fig. 1	D1P02B08	118420240	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Colpodella angusta
Fig. 1	D1P02C06	118420241	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Api.:Polyplicarum -related
Fig. 1	D1P02D12	118420066	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Colpodella angusta
Fig. 1	D1P02E03	118420067	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade VII
Fig. 1	D1P02F05	118420046	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Api.:Polyplicarum -related
Fig. 1	D2P03B10	118420105	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade VII
Fig. 1	D2P03D02	118420082	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Colpodella angusta
Fig. 1	D2P03D06	118420106	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade I
Fig. 1	D2P03E11	118420245	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Api.:Polyplicarum -related
Fig. 1	D3P05C07	118420125	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade XI
Fig. 1	D3P05F04	118420152	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade VII
Fig. 1	D3P06B02	118420161	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade XI
Fig. 1	D4P07B05	118420253	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade VII
Fig. 1	D4P07C10	118420192	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade XI
Fig. 1	D4P07G08	118420173	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade I
Fig. 1	D4P07H09	118420191	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade XI
Fig. 1	D4P08D06	118420260	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade VII
Fig. 1	D4P08E10	118420178	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Api.:Polyplicarum -related
Fig. 1	D4P08E11	118420179	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade I
Fig. 1	D4P08G08	118420181	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade XI

Fig. 1	D5P09C09	118420206	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Colpodella angusta
Fig. 1	D5P10A10	118420214	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Api.:Polyplicarum-related
Fig. 1	D5P10D11	118420226	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Env. clade VII
Fig. 1	D5P10G04	118420223	marine	YES	oxygen-depleted intertidal marine sediment, upper 2 cm; Greenland: Arctic	PLoS ONE 2(8): e728 (2007)	Colpodella angusta
Fig. 1	DGGE_gel_band_3mFR1L9B9	544587044	animal feces/manure	YES	composting dairy manure	Unpublished	Colpodella angusta
Fig. 1	DGGE_gel_band_3mFR11011	544587051	animal feces/manure	YES	composting dairy manure	Unpublished	Colpodella angusta
Fig. 1	He000427_15535	89572115	marine		marine sample; Germany: Helgoland	Microb. Ecol. 52 (1), 53-71 (2006)	Colpodella angusta
Fig. 1	He000427_20	89572015	marine		marine sample; Germany: Helgoland	Microb. Ecol. 52 (1), 53-71 (2006)	Colpodella angusta
Fig. 1	dpeuk39	60592951	freshwater	YES	freshwater anaerobic pond; USA: Zodletone Spring, Oklahoma	Appl. Environ.Microbiol.71(10),6175-84(2005)	Unresolved:Api. relatives
Fig. 1	DSGM-12	133778560	marine	YES	deep-sea methane cold seep sediment; Japan: Sagami Bay	Extremophiles 11 (4), 563-576 (2007)	Unresolved: Api.
Fig. 1	DSGM-13	133778561	marine	YES	deep-sea methane cold seep sediment; Japan: Sagami Bay	Extremophiles 11 (4), 563-576 (2007)	Api.:Polyplicarum-related
Fig. 1	DSGM-8	133778556	marine	YES	deep-sea methane cold seep sediment; Japan: Sagami Bay	Extremophiles 11 (4), 563-576 (2007)	Unresolved: Api.
Fig. 1	E08_250P220_0239Tplate03_P3	389549864	marine (*)		marine sediment from shallow water hydrothermal system; Papua New Guinea	Pacific Science 64 (3), 391–404 (2010)	Vitrella brassicaformis
Fig. 1	E12_180P126_0239Uplate02_P2	389549895	marine (*)		marine sediment from shallow water hydrothermal system; Papua New Guinea	Pacific Science 64 (3), 391–404 (2010)	Vitrella brassicaformis
Fig. 1	E1aD71	168203738	marine		marine plankton; South China Sea, northern part	J. Plankton Res. 33 (3), 445-456 (2011)	Voromonas pontica
Fig. 1	E6	30144456	marine	YES	permanently anoxic deep-sea basin; Caribbean Sea: Cariaco Basin	Appl. Environ. Microbiol. 69(9),5656-63(2003)	Api.:Polyplicarum-related
Fig. 1	EBA140.9	213058639	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Env. clade II
Fig. 1	EBA43.25	213058601	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Env. clade IV
Fig. 1	EBA47.26	213058602	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Env. clade IV
Fig. 1	EBA52.36	213058604	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Env. clade IV
Fig. 1	EBA9.8	213058584	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Unresolved:Api. relatives
Fig. 1	EBF1.128	213058470	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Env. clade IV
Fig. 1	EBF1.90	213058469	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Unresolved:Api. relatives
Fig. 1	EBM3.51	213058518	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Unresolved:Api. relatives
Fig. 1	EC12	55831924	animal feces/manure	YES	pig manure storage pit	FEMS Microbiol. Ecol. 52(2), 229-42(2005)	Colpodella angusta
Fig. 1	EC54	148743803	animal feces/manure	YES	pig manure storage pit	FEMS Microbiol. Ecol. 52(2), 229-42(2005)	Colpodella angusta
Fig. 1	EC67	148743812	animal feces/manure	YES	pig manure storage pit	FEMS Microbiol. Ecol. 52(2), 229-42(2005)	Colpodella angusta
Fig. 1	elb6-t0-otu7	294829215	freshwater		ice-covered lake; Antarctica	ISME J 5, 1559-1564 (2011)	Unresolved:Api. relatives
Fig. 1	Elev_18S_5171	166084485	soil		trembling aspen rhizosphere, elevated CO2 conditions; USA: UMBS, Michigan	Environ. Microbiol. 10 (4), 926-941 (2008)	Colpodella angusta
Fig. 1	ESS270706.033	299789590	freshwater		lake water filtered through 3 um from lake Esch sur Sure, depth 0m; Luxembourg	Unpublished	Unresolved:Api. relatives
Fig. 1	H11C09cA	261290035	marine (*)		stromatolite (marine microbial mat); Bahamas: Highborne Cay	Environ. Microbiol. 11 (10), 2710-19 (2009)	Vitrella brassicaformis
Fig. 1	He000427_1	89572118	marine		marine sample; Germany:Helgoland	Microb. Ecol. 52 (1), 53-71 (2006)	Colpodella angusta
Fig. 1	I10_180P53_0239Rplate06_P6	389550613	marine (*)		marine sediment from shallow water hydrothermal system; Papua New Guinea	Pacific Science 64 (3), 391–404 (2010)	Vitrella brassicaformis
Fig. 1	I9S18K_ML_148	223034114	freshwater		Industrial Canal, water, greater than 3 micron; USA: New Orleans, LA	Environ. Sci. Technol. 42(24), 9072-8(2008)	Env. clade IV
Fig. 1	K1OCT2011	430726581	freshwater		freshwater lake; Japan: Gunma, lake Kusaki	Unpublished	Env. clade IV

Fig. 1	K2MAY2011	430726574	freshwater		freshwater lake; Japan: Gunma, lake Kusaki	Unpublished	Env. clade IV
Fig. 1	K5AUG2009	327247873	freshwater		freshwater lake; Japan: Gunma, lake Kusaki	Unpublished	Env. clade IV
Fig. 1	KRL01E25	353255973	freshwater		freshwater lake; Greece: Lake Karla	ScientificWorldJournal 2012, 504135 (2012)	Unresolved:Api. relatives
Fig. 1	KRL03E03	449890141	freshwater		freshwater lake; Greece: Lake Karla	ScientificWorldJournal 2012, 504135 (2012)	Env. clade IV
Fig. 1	KRL03E57	449890156	freshwater		freshwater lake; Greece: Lake Karla	Hydrobiologia 718, 73-83 (2013)	Env. clade III
Fig. 1	KRL09E06	449890170	freshwater		freshwater lake; Greece: Lake Karla	Hydrobiologia 718, 73-83 (2013)	Env. clade III
Fig. 1	KRL09E21	449890174	freshwater		freshwater lake; Greece: Lake Karla	Hydrobiologia 718, 73-83 (2013)	Env. clade IV
Fig. 1	L07_250P64_0239DSNP001F_P3	389551156	marine (*)		marine sediment from shallow water hydrothermal system; Papua New Guinea	Pacific Science 64 (3), 391–404 (2010)	Vitrella brassicaformis
Fig. 1	LG02-01	60476886	freshwater		oligotrophic freshwater lake; USA: Lake George, Adirondack Park	Environ. Microbiol. 7 (9), 1413-1425 (2005)	Env. clade II
Fig. 1	LG05-01	60476898	freshwater		oligotrophic freshwater lake; USA: Lake George, Adirondack Park	Environ. Microbiol. 7 (9), 1413-1425 (2005)	Env. clade II
Fig. 1	LG09-02	60476916	freshwater		oligotrophic freshwater lake; USA: Lake George, Adirondack Park	Environ. Microbiol. 7 (9), 1413-1425 (2005)	Env. clade II
Fig. 1	LG25-01	60476973	freshwater		oligotrophic freshwater lake; USA: Lake George, Adirondack Park	Environ. Microbiol. 7 (9), 1413-1425 (2005)	Env. clade II
Fig. 1	LG32-03	60476996	freshwater		oligotrophic freshwater lake; USA: Lake George, Adirondack Park	Environ. Microbiol. 7 (9), 1413-1425 (2005)	Env. clade II
Fig. 1	LT19_H1	529105019	hypersaline		hypersaline lake; Australia: Lake Tyrrell, Victoria	Front Microbiol 4, 115 (2013)	Unresolved:Api. relatives
Fig. 1	LT35_C14	529105426	hypersaline		hypersaline lake; Australia: Lake Tyrrell, Victoria	Front Microbiol 4, 115 (2013)	Unresolved:Api. relatives
Fig. 1	LT37_G4	529105789	hypersaline		hypersaline lake; Australia: Lake Tyrrell, Victoria	Front Microbiol 4, 115 (2013)	Chromera velia
Fig. 1	LT41_e05_2	529106034	hypersaline		hypersaline lake; Australia: Lake Tyrrell, Victoria	Front Microbiol 4, 115 (2013)	Unresolved:Api. relatives
Fig. 1	LT50_P1	529106341	hypersaline		hypersaline lake; Australia: Lake Tyrrell, Victoria	Front Microbiol 4, 115 (2013)	Unresolved:Api. relatives
Fig. 1	LT62_D14	529106413	hypersaline		hypersaline lake; Australia: Lake Tyrrell, Victoria	Front Microbiol 4, 115 (2013)	Unresolved:Api. relatives
Fig. 1	Mat131	356459394	marine (*)	YES	anoxic sink; Gulf of Mexico: Jewfish Sink	Hydrobiologia 677 (1), 65-87 (2011)	Vitrella brassicaformis
Fig. 1	Mat140	356459402	marine (*)	YES	anoxic sink; Gulf of Mexico: Jewfish Sink	Hydrobiologia 677 (1), 65-87 (2011)	Vitrella brassicaformis
Fig. 1	Mat179	356459439	marine (*)	YES	anoxic sink; Gulf of Mexico: Jewfish Sink	Hydrobiologia 677 (1), 65-87 (2011)	Vitrella brassicaformis
Fig. 1	ML1_dino18S-9	309753096	freshwater		5 micron fraction of surface freshwater lake; USA: Connecticut, Storrs, Mirror Lake	PNAS 107 (46), 20033-38 (2010)	Env. clade IV
Fig. 1	MLBA54.107	213058446	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Unresolved:Api. relatives
Fig. 1	MLBA74.128	213058464	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Unresolved:Api. relatives
Fig. 1	MLBM116.37	213058384	freshwater		freshwater lake; China	Acta Protozool. 48, 245–264 (2009)	Unresolved:Api. relatives
Fig. 1	MPE1-2	374532643	freshwater		aquatic moss pillars, lake; Antarctica: Skarvsnes, Hotoke-Ike lake	Polar Biol. 35, 1495-1504 (2012)	Unresolved:Api. relatives
Fig. 1	MPE1-46	374532687	freshwater		aquatic moss pillars, lake; Antarctica: Skarvsnes, Hotoke-Ike lake	Polar Biol. 35, 1495-1504 (2012)	Unresolved:Api. relatives
Fig. 1	MPE2-2	374532695	freshwater		aquatic moss pillars, lake; Antarctica: Skarvsnes, Hotoke-Ike lake	Polar Biol. 35, 1495-1504 (2012)	Unresolved:Api. relatives
Fig. 1	Out-556	356459667	marine		anoxic sink; Gulf of Mexico: Jewfish Sink	Hydrobiologia 677 (1), 65-87 (2011)	Vitrella brassicaformis
Fig. 1	Out-623	356459729	marine		anoxic sink; Gulf of Mexico: Jewfish Sink	Hydrobiologia 677 (1), 65-87 (2011)	Vitrella brassicaformis
Fig. 1	Out-678	356459781	marine		anoxic sink; Gulf of Mexico: Jewfish Sink	Hydrobiologia 677 (1), 65-87 (2011)	Vitrella brassicaformis
Fig. 1	PAA22AU2004	82470184	freshwater		deep freshwater meromictic lake; France: Lake Pavin, Massif Central	Environ. Microbiol. 9 (1), 61-71 (2007)	Env. clade IV
Fig. 1	PAF8AU2004	82470188	freshwater		deep freshwater meromictic lake; France: Lake Pavin, Massif Central	Environ. Microbiol. 9 (1), 61-71 (2007)	Unresolved:Api. relatives
Fig. 1	PC04.G3_Pla_1	393405643	freshwater		high mountain lake; Spain: Pyrenees, Lake Plan	Environ. Microbiol. 14 (9), 2445-56 (2012)	Env. clade II
Fig. 1	PC05.B12_Ger_12	393405440	freshwater		high mountain lake; Spain: Pyrenees, Lake Plan	Environ. Microbiol. 14 (9), 2445-56 (2012)	Env. clade II
Fig. 1	PC05.F10_Ger_1	393405443	freshwater		high mountain lake; Spain: Pyrenees, Lake Plan	Environ. Microbiol. 14 (9), 2445-56 (2012)	Env. clade II

Fig. 1	PC06.B8_Ger_1	393405451	freshwater		high mountain lake; Spain: Pyrenees, Lake Plan	Environ. Microbiol. 14 (9), 2445-56 (2012)	Env. clade II
Fig. 1	PC06.D4_Ger_9	393405454	freshwater		high mountain lake; Spain: Pyrenees, Lake Plan	Environ. Microbiol. 14 (9), 2445-56 (2012)	Env. clade II
Fig. 1	PC06.H4_Ger_6	393405472	freshwater		high mountain lake; Spain: Pyrenees, Lake Plan	Environ. Microbiol. 14 (9), 2445-56 (2012)	Env. clade II
Fig. 1	Pink_A01	300088886	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_C09	300088917	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_D06	300088926	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_D07	300088927	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_D12	300088932	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_E09	300088940	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_E10	300088941	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_F04	300088947	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_F05	300088948	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_F11	300088953	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_G07	300088960	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	Pink_G11	300088964	marine (*)		intertidal thrombolites; Bahamas: Highborne Cay	Geobiology 8 (4), 337-354 (2010)	Vitrella brassicaformis
Fig. 1	PR2_3E_116	316926548	freshwater		water from Praz Rodet peat bog; Switzerland: Jura Mountains	Protist 162, 14–32 (2011)	Unresolved: Api.
Fig. 1	PR2_3E_74	299778455	freshwater		water from Praz Rodet peat bog; Switzerland: Jura Mountains	Protist 162, 14–32 (2011)	Unresolved:Api. relatives
Fig. 1	PR3_4E_33	299778456	freshwater		water from Praz Rodet peat bog; Switzerland: Jura Mountains	Protist 162, 14–32 (2011)	Env. clade VI
Fig. 1	PS13A10	262357995	marine (*)		Lakshwadeep coral reef sand; Arabian Sea	FEMS Microbiol. Ecol. 71(3),399–412(2010)	Vitrella brassicaformis
Fig. 1	PS13C9	262358012	marine (*)		Lakshwadeep coral reef sand; Arabian Sea	FEMS Microbiol. Ecol. 71(3),399–412(2010)	Vitrella brassicaformis
Fig. 1	PS13D9	262358015	marine (*)		Lakshwadeep coral reef sand; Arabian Sea	FEMS Microbiol. Ecol. 71(3),399–412(2010)	Vitrella brassicaformis
Fig. 1	SB16_2010	387912501	freshwater		water purification plant; Japan:Gunma, Kiryu	Unpublished	Voromonas pontica
Fig. 1	SB19_2010	387912504	freshwater		water purification plant; Japan:Gunma, Kiryu	Unpublished	Chromera velia
Fig. 1	SB21_2010	387912506	freshwater		water purification plant; Japan:Gunma, Kiryu	Unpublished	Unresolved:Api. relatives
Fig. 1	SB7_2010	387912492	freshwater		water purification plant; Japan:Gunma, Kiryu	Unpublished	Unresolved:Api. relatives
Fig. 1	SC8051102_25	225548510	freshwater		freshwater reservoir; South Korea: Sangcheon reservoir	J. Microbiol. Methods 81 (1), 61-68 (2010)	Unresolved:Api. relatives
Fig. 1	SEO81101_35	225548493	freshwater		freshwater reservoir; South Korea: Seoho reservoir	J. Microbiol. Methods 81 (1), 61-68 (2010)	Env. clade IV
Fig. 1	SEO81107_36	225548503	freshwater		freshwater reservoir; South Korea: Seoho reservoir	J. Microbiol. Methods 81 (1), 61-68 (2010)	Env. clade IV
Fig. 1	Sey061	49356800	freshwater		river sediment; Switzerland: Seymaz River, Geneva	BMC Biol. 2, 13 (2004)	Env. clade VII
Fig. 1	TB4-1	259906333	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	TB4-15	259906347	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	TB4-24	259906354	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	TB4-27	259906357	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	TB4-39	259906369	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	TB4-43	259906371	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives

Fig. 1	TB4-44	259906372	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	TB4-45	259906373	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	TB4-48	259906376	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	TB4-51	259906379	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	V10-5	355346671	freshwater		freshwater reservoir; France, Villerest reservoir	Unpublished	Env. clade IV
Fig. 1	VP1	100289290	freshwater		lacustrine community in hyper-eutrophic conditions; France, Villerest reservoir	FEMS Microbiol. Ecol. 61(3), 483-95(2007)	Unresolved:Api. relatives
Fig. 1	WD0-1	259906092	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Env. clade II
Fig. 1	WD0-50	259906128	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Env. clade IV
Fig. 1	WD4-1	259906134	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	WD4-110	259906236	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	WD4-128	259906254	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	WD4-138	259906264	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	WD4-66	259906196	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. 1	WD4-80	259906208	freshwater		freshwater lake; China: Lake Taihu	FEMS Microbiol. Ecol. 74 (1), 19-31 (2010)	Unresolved:Api. relatives
Fig. S2	P107t1c1	KP213181	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P107t1c3	KP213183	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P108t7c5	KP213189	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P108t7c6	KP213190	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P109t5c6	KP213193	marine		coral Porites compressa; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.: genotype N
Fig. S2	P113t16c1	KP213194	marine		Coral 3, identified as Seriatopora hystrix; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.: genotype N
Fig. S2	P114t13c1	KP213198	marine		coral Acropora yongei; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P114t13c2	KP213199	marine		coral Acropora yongei; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P114t9c2	KP213201	marine		mucus from Acropora yongei; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P116t17c4	KP213205	marine		Coral 3, identified as Seriatopora hystrix; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.: genotype N
Fig. S2	P116t17c5	KP213206	marine		Coral 3, identified as Seriatopora hystrix; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.: Sarcocystids
Fig. S2	P116t7c3	KP213210	marine		coral Seriatopora hystrix; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.: genotype N
Fig. S2	P57t7c1	KP213223	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P57t7c10	KP213224	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P57t7c3	KP213225	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P57t7c9	KP213225	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P87t2c4	KP213237	soil		soil from UBC endowment lands; Canada: Vancouver	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P87t5c3b	KP213244	marine		Damp woodchip and surface sand; Canada: Locarno Beach, Vancouver, BC	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P91t3c1	KP213248	marine		coral Porites astreoides; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P91t3c4	KP213249	marine		coral Porites astreoides; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P91t4c2	KP213250	marine		coral Porites compressa; USA: Birch Aquarium, San Diego, CA	THIS STUDY	Api.: genotype N

Fig. S2	P97t5c5	KP213257	freshwater		Freshwater duck pond near the beach; Canada: Boundary Bay, BC	THIS STUDY	Api.:Terrestrial gregarines
Fig. S2	P97t3c2	KP213252	marine		Damp mid-beach sand under broken shell and debris; Canada: Boundary Bay, BC	THIS STUDY	Api.: Eimeriids
Fig. S2	P97t3c3	KP213253	marine		Damp mid-beach sand under broken shell and debris; Canada: Boundary Bay, BC	THIS STUDY	Api.: Eimeriids
Fig. S2	P97t3c4	KP213254	marine		Damp mid-beach sand under broken shell and debris; Canada: Boundary Bay, BC	THIS STUDY	Api.: Eimeriids
Fig. S2	P97t3c5	KP213255	marine		Damp mid-beach sand under broken shell and debris; Canada: Boundary Bay, BC	THIS STUDY	Api.: Eimeriids

[1] Marine, CaCO3-dominated environments specifically associated with sequences from the *Vitrella brassicaformis* clade are indicated by the (*) symbol. These include coral reef sediment (Arabian Sea), thrombolites and stromatolites at Higborne Cay (Bahamas), CaCO3-precipitating sink off the coast of Florida, and coastal carbonate sediment in Papua New Guinea. See the respective publications for further details.

[2] Low oxygen environement (anoxic or microaerophilic)

[3] THIS STUDY = sequences generated in this study by apicomplexan-specific primers

[4] Based on Fig. 1; Env. = Environmental; Api. = Apicomplexa

Table S2: Abbreviations and full names of proteins and pathways used in Fig. 3.

Pathway	Localization	Short names	E.C.	Protein name
Isoprenoid precursor biosynthesis (Isopentenyl diphosphate= IPP/ Dimethylallyl diphosphate= DMAP)	Cytosolic (mevalonate pathway = MEV)	HMGCS	2.3.3.10	hydroxymethylglutaryl-CoA synthase
		HMGCR	1.1.1.34	hydroxymethylglutaryl-CoA reductase
		MVK	2.7.1.36	mevalonate kinase
		PMVK	2.7.4.2	phosphomevalonate kinase
		MVD	4.1.1.33	diphosphomevalonate decarboxylase
	Plastidic (non-mevalonate pathway = MEP/DOXP)	DXS	2.2.1.7	1-deoxy-D-xylulose-5-phosphate synthase
		IspC	1.1.1.267	1-deoxy-D-xylulose-5-phosphate reductoisomerase
		IspD	2.7.7.60	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
		IspE	2.7.1.148	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
		IspF	4.6.1.12	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
		IspG	1.17.7.1	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase
		IspH (LytB)	1.17.1.2	4-hydroxy-3-methylbut-2-enyl diphosphate reductase
Tetrapyrrole biosynthesis (heme, chlorophyll, etc.)	Mitochondrial (C4 pathway)	ALAS	2.3.1.37	5-aminolevulinate synthase
		ALAD (HemB)	4.2.1.24	5-aminolevulinate dehydratase/ porphobilinogen synthase
	Cytosolic/ Mitochondrial (C4 pathway) or Plastidic (C5 pathway)	PBGD (HemC)	2.5.1.61	porphobilinogen deaminase hydroxymethylbilane synthase
		UROS (HemD)	4.2.1.75	uroporphyrinogen-III synthase
		UROD (HemE)	4.1.1.37	uroporphyrinogen decarboxylase
		CPOX (HemF)	1.3.3.3	coproporphyrinogen III oxidase
		PPOX (HemY)	1.3.3.4	protoporphyrinogen oxidase
		FECH (HemH)	4.99.1.1	ferrochelatase
	Plastidic (C5 pathway)	GTS (HemA)	1.2.1.70	glutamyl-tRNA reductase
		GTR (HemL)	5.4.3.8	glutamate-1-semialdehyde 2,1-aminomutase
Fatty acid biosynthesis and elongation	Cytosolic (polyketide synthase/ fatty acid synthase type I pathway = PKSI/FASI)	FAAL	2.3.1.86	fatty acid synthase type I, fatty acyl-AMP ligase domain
		KS	2.3.1.86	fatty acid synthase type I, ketoacyl synthase domain
		AT	2.3.1.86	fatty acid synthase type I, acyl transferase domain
		DH	2.3.1.86	fatty acid synthase type I, dehydrase domain
		ER	2.3.1.86	fatty acid synthase type I, enoyl reductase domain
		KR	2.3.1.86	fatty acid synthase type I, ketoacyl reductase domain
		ACP	2.3.1.86	fatty acid synthase type I, acyl carrier protein domain
		SDR	2.3.1.86	fatty acid synthase type I, terminal reductase domain
	Endoplasmic reticulum (ER fatty acid elongation pathway)	ELO	2.3.1.199	beta-ketoacyl-CoA synthase
		KCR	1.1.1.330	beta-ketoacyl-CoA reductase
		PHS	4.2.1.134	beta-hydroxyacyl-CoA dehydratase
		TECR	1.3.1.93	trans-2-enoyl-CoA reductase
	Plastidic (fatty acid synthase type II pathway = FASII)	FabD	2.3.1.39	malonyl-CoA-acyl carrier protein transacylase
		FabG	1.1.1.100	beta-ketoacyl-acyl carrier protein reductase
		FabH	2.3.1.180	beta-ketoacyl-acyl carrier protein synthase III
		FabZ	4.2.1.59	D-3-hydroxyoctanoyl-acyl carrier protein dehydratas
		FabI	1.3.1.9	enoyl acyl carrier protein reductase
		FabB/F	2.3.1.41	beta-ketoacyl-acyl carrier protein synthetase
		ACP		acyl-carrier protein
Other functions	Plastidic (plastid-encoded)	SufB		cysteine desulfurase activator complex subunit SufB
		ClpC		ATP-dependent Clp protease ATP-binding subunit ClpC
		Ycf93		plastid membrane protein of unknown function

Table S3: Chrompodellid plastidic proteins with complete N-terminal sequence.

Chrompodellid species	**Short name**	**Contig/Protein reference**	**SP[1]**	**SP score[1]**	**SP length[1]**	**cTP[2]**	**cTP score[2]**	**cTP length[2]**
Alphamonas edax BE-2	FabF	a7716_61_4_2_1591	Y	0.81	23		0.46	16
Alphamonas edax BE-2	FabG	a22551_86_13_1248_259	Y	0.9	24		0.49	38

Alphamonas edax BE-2	FabZ	a263933_12_2_2_448	Y	0.88	20		0.47	35
Alphamonas edax BE-2	SufE	a110568_24_3_80_688	Y	0.94	20	Y	0.52	34
Chromera velia	ALAD	MMETSP0290_20130426_17951_1_402	Y	0.55	15	Y	0.55	47
Chromera velia	ClpC	MMETSP0290_20130426_3091_1_1013	Y	0.76	18	Y	0.56	92
Chromera velia	CPOX	MMETSP0290_20130426_4127_1_409	Y	0.8	17	Y	0.53	36
Chromera velia	CPOX	MMETSP0290_20130426_4154_1_411	Y	0.7	23		0.49	51
Chromera velia	FabD	MMETSP0290_20130426_21839_1_349	Y	0.85	19	Y	0.51	35
Chromera velia	FabF	MMETSP0290_20130426_1976_1_463	Y	0.59	19	Y	0.55	51
Chromera velia	FabG	MMETSP0290_20130426_20062_1_308	Y	0.88	18		0.49	54
Chromera velia	FabH	MMETSP0290_20130426_17765_1_393	Y	0.77	21	Y	0.55	57
Chromera velia	FabI	MMETSP0290_20130426_14303_1_355	Y	0.82	19	Y	0.54	38
Chromera velia	FabZ	MMETSP0290_20130426_4631_1_213	Y	0.89	17		0.49	54
Chromera velia	Fdx	MMETSP0290_20130426_1982_1_149	Y	0.75	15	Y	0.54	50
Chromera velia	FECH	MMETSP0290_20130426_5910_1_460	Y	0.79	33	Y	0.51	63
Chromera velia	IspD	MMETSP0290_20130426_22696_1	Y	0.82	19	Y	0.51	64
Chromera velia	IspE	MMETSP0290_20130426_14275_1	Y	0.8	18	Y	0.57	53
Chromera velia	PBGD	MMETSP0290_20130426_12195_1_411	Y	0.85	18	Y	0.57	56
Chromera velia	PPOX	MMETSP0290_20130426_18890_1_615	Y	0.58	19	Y	0.52	15
Chromera velia	SufB	MMETSP029020130426_15825_1_566	Y	0.7	18		0.48	25
Chromera velia	SufS	contig12900_JO801918_blast_full	Y	0.72	19	Y	0.5	34
Chromera velia	UROS	MMETSP0290_20130426_22563_1_399	Y	0.48	17	Y	0.52	24
Colpodella angusta BE-6	ALAD	a1810_243_9_1431_172	Y	0.65	19		0.45	14
Colpodella angusta BE-6	CPOX	a5011_127_9_1611_262	Y	0.76	19		0.44	13
Colpodella angusta BE-6	Fdx	a63516_19_2_1_489	Y	0.7	15		0.44	14
Colpodella angusta BE-6	FECH	a17545_33_13_1312_44	Y	0.73	17		0.5	16
Colpodella angusta BE-6	PPOX	a34610_19_5_1_1764	Y	0.7	25		0.45	20
Colpodella angusta BE-6	SufB	a7216_97_2_21570	Y	0.74	19		0.43	15
Colpodella angusta BE-6	SufC	a30602_28_6_898_62	Y	0.61	15		0.43	26
Colpodella angusta BE-6	UROD	a13302_39_9_1650_67	Y	0.54	22		0.47	40
Colpodella angusta Spi-2	FNR	a13692_81_9_976_65	Y	0.58	22		0.45	19
Colpodella angusta Spi-2	IspC	a44398_21_7_1598_168	Y	0.8	20		0.47	57
Colpodella angusta Spi-2	IspD	a18350_62_2_44_820	Y	0.62	18		0.44	25
Colpodella angusta Spi-2	IspE	a31756_25_7_1128_46	Y	0.72	16		0.44	26
Colpodella angusta Spi-2	PBGD	a53819_34_11_1198_62	Y	0.81	21		0.45	6
Vitrella brassicaformis	ClpC	MMETSP0290_20130426_10413_1_925	Y	0.58	19	Y	0.51	74
Vitrella brassicaformis	CPOX	MMETSP0290_20130426_65767_1_398	Y	0.83	19		0.48	41
Vitrella brassicaformis	CPOX	MMETSP0290_20130426_28340_1_400	Y	0.82	23	Y	0.51	71
Vitrella brassicaformis	DXS	MMETSP0290_20130426_5902_1	Y	0.87	19	Y	0.55	73
Vitrella brassicaformis	FabG	MMETSP0290_20130426_48423_1_305	Y	0.82	24	Y	0.54	47
Vitrella brassicaformis	FabZ	MMETSP0290_20130426_8033_1_242	Y	0.7	22	Y	0.55	67
Vitrella brassicaformis	Fdx	MMETSP0290_20130426_11262_1_158	Y	0.82	17		0.48	59
Vitrella brassicaformis	Fdx	MMETSP0290_20130426_12802_1_160	Y	0.62	17		0.49	14
Vitrella brassicaformis	Fdx	MMETSP0290_20130426_5440_1_197	Y	0.9	24		0.46	50

Vitrella brassicaformis	FECH	MMETSP0290_20130426_27871_1_400	Y	0.47	19		0.47	12
Vitrella brassicaformis	FNR	MMETSP0290_20130426_14229_1_357	Y	0.91	18		0.48	37
Vitrella brassicaformis	IspD	MMETSP0290_20130426_1323_1	Y	0.71	16	Y	0.53	50
Voromonas pontica G-3	Fdx	a49492_12_5_308_784_M1	Y	0.86	17	Y	0.53	56
Voromonas pontica GG	ALAD	comp4582_c0_seq1_7	Y	0.7	17		0.43	82
Voromonas pontica GG	IspE	comp26767_c0_seq1_4	Y	0.89	20	Y	0.53	41
Voromonas pontica GG	SufB	578895742 (GI)	Y	0.62	20	Y	0.5	51

[1] Signal peptide (SP) presence, score and length as predicted by SignalP 4.1
[2] Transit peptide (cTP) presence, score and length as predicted by ChloroP 1.1

Table S4: Classification of 527 proteins from Apiloc (correspoding to 382 OrthoMCL clusters) based on their putative origin.

Putative origin 1	OrthoMCL group	Protein EuPathDb name or GenBank GI accession	Organism	Protein annotation	Length	Apical 2	Exported 2	Membrane 2	IMC 2	Golgi 2	Oocyst wall 2	CHrompod.3	Dino + Perk.3	Eukaryot.3	Note4
Aconoidasids	OG5_153539	PBANKA_100620	Plasmodium berghei ANKA	sporozoite invasion-associated protein 1 (SIAP1)	999	y		y				-	-	-	
Aconoidasids	OG5_156194	PF3D7_1116000	Plasmodium falciparum 3D7	rhoptry neck protein 4 (RON4)	1201	y			y			-	-	-	
Aconoidasids	OG5_162079	PF3D7_0929400	Plasmodium falciparum 3D7	high molecular weight rhoptry protein 2 (RhopH2)	1378	y		y				-	-	-	
Apicomplexans	OG5_141738	PF3D7_1463900	Plasmodium falciparum 3D7	conserved Plasmodium membrane protein, unknown function	1071		y					-	-	-	X
Apicomplexans	OG5_141751	PF3D7_0210600	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	446	y						-	-	-	X
Apicomplexans	OG5_143206	PF3D7_1030200	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	450	y						-	-	-	X,R
Apicomplexans	OG5_145102	PF3D7_1318000	Plasmodium falciparum 3D7	conserved Plasmodium membrane protein, unknown function	192	y						-	-	-	X
Apicomplexans	OG5_150575	PF3D7_0214900	Plasmodium falciparum 3D7	rhoptry neck protein 6 (RON6)	950	y						-	-	-	X,R
Apicomplexans	OG5_166877	PBANKA_141830	Plasmodium berghei ANKA	rhoptry protein 2, putative (PRP2)	1272	y						-	-	-	X,R
Apicomplexans	OG5_222773	TGME49_310010	Toxoplasma gondii ME49	rhoptry neck protein RON1 (RON1)	1158	y						-	-	-	X,R
Babesia	NO_GROUP	BBOV_I004210	Babesia bovis T2Bo	spherical body protein 3	1089	y	y					-	-	-	
Babesia	NO_GROUP	BBOV_IV005390	Babesia bovis T2Bo	spherical body protein 4 SBP4	372	y	y					-	-	-	
Babesia	NO_GROUP	gi|14646759	Babesia gibsoni	P50	466			y				-	-	-	
Babesia	NO_GROUP	gi|223036620	Babesia gibsoni	merozoite 22-kDa protein, partial	207			y				-	-	-	
Babesia	OG5_140337	BBOV_II000740	Babesia bovis T2Bo	Spherical Body Protein 2 (SBP2)	2248	y	y					-	-	-	
Babesia	OG5_164986	BBOV_II002880	Babesia bovis T2Bo	85 kDa protein	574	y	y					-	-	-	R
Babesia	OG5_171201	BBOV_I002990	Babesia bovis T2Bo	merozoite surface antigen-2b (MSA-2b)	287			y				-	-	-	
Babesia	OG5_171201	BBOV_I003000	Babesia bovis T2Bo	merozoite surface antigen-2a2 (MSA-2a2)	316			y				-	-	-	
Babesia	OG5_171201	BBOV_I003010	Babesia bovis T2Bo	merozoite surface antigen-2a1 (MSA-2a1)	316			y				-	-	-	
Babesia	OG5_171201	BBOV_I003020	Babesia bovis T2Bo	merozoite surface antigen-2c (MSA-2c)	265			y				-	-	-	
Babesia	OG5_171201	BBOV_I003060	Babesia bovis T2Bo	merozoite surface antigen-1 (MSA-1)	319			y				-	-	-	
Apicomplexans + Chrompodellids	OG5_140542	TGME49_202500	Toxoplasma gondii ME49	GAPM1a	305				y			y	-	-	X
Apicomplexans + Chrompodellids	OG5_140542	PF3D7_1323700	Plasmodium falciparum 3D7	glideosome associated protein with multiple membrane spans 1 (GAPM1)	303				y			y	-	-	X
Apicomplexans + Chrompodellids	OG5_141755	PF3D7_0423500	Plasmodium falciparum 3D7	glideosome associated protein with multiple membrane spans 2 (GAPM2)	372				y			y	-	-	X
Apicomplexans + Chrompodellids	OG5_141755	TGME49_206690	Toxoplasma gondii ME49	glideosome-associated protein with multiple-membrane spans GAPM2B (GAPM2B)	426				y			y	-	-	X
Apicomplexans + Chrompodellids	OG5_143227	PCHAS_103620	Plasmodium chabaudi chabaudi	glideosome associated protein with multiple membrane spans 3, putative (GAPM3)	283	y		y				y	-	-	X
Apicomplexans + Chrompodellids	OG5_143227	PF3D7_1406800	Plasmodium falciparum 3D7	glideosome associated protein with multiple membrane spans 3 (GAPM3)	284				y			y	-	-	X
Apicomplexans + Chrompodellids	OG5_143227	TGME49_271970	Toxoplasma gondii ME49	glideosome-associated protein with multiple-membrane spans GAPM3	310				y			y	-	-	X
Apicomplexans + Chrompodellids	OG5_145116	PF3D7_0414600	Plasmodium falciparum 3D7	conserved protein, unknown function	235	y						-	-	-	X
Apicomplexans + Chrompodellids	OG5_225269	cgd6_200	Cryptosporidium parvum Iowa II	oocyst wall protein 8	457						y				X
Apicomplexans + Chrompodellids	OG5_239307	TGME49_209610	Toxoplasma gondii ME49	oocyst wall protein OWP2 (OWP2)	462						y				X
Coccidians + Aconoidasids	OG5_147452	BBOV_IV011230	Babesia bovis T2Bo	apical membrane antigen 1	605	y						-	-	-	
Coccidians + Aconoidasids	OG5_126585	PF3D7_1149000	Plasmodium falciparum 3D7	antigen 332, DBL-like protein (Pf332)	6093		y					-	-	-	R
Coccidians + Aconoidasids	OG5_126854	PBANKA_090130	Plasmodium berghei ANKA	merozoite adhesive erythrocytic binding protein (MAEBL)	1799	y						-	-	-	R
Coccidians + Aconoidasids	OG5_126854	PF3D7_1147800	Plasmodium falciparum 3D7	merozoite adhesive erythrocytic binding protein (MAEBL)	2055	y		y				-	-	-	R
Coccidians + Aconoidasids	OG5_126854	PY05977	Plasmodium yoelii yoelii 17XNL	erythrocyte binding protein	1701	y						-	-	-	R
Coccidians + Aconoidasids	OG5_138272	PF3D7_0220800	Plasmodium falciparum 3D7	cytoadherence linked asexual protein 2 (CLAG2)	1440	y						-	-	-	
Coccidians + Aconoidasids	OG5_138272	PF3D7_0302200	Plasmodium falciparum 3D7	cytoadherence linked asexual protein 3.2 (CLAG3.2)	1416	y						-	-	-	

Coccidians + Aconoidasids	OG5_142870	PBANKA_131570	Plasmodium berghei ANKA	rhoptry neck protein 2 (RON2)	2081	y						-	-	-	R
Coccidians + Aconoidasids	OG5_142870	PF3D7_1452000	Plasmodium falciparum 3D7	rhoptry neck protein 2 (RON2)	2189	y						-	-	-	R
Coccidians + Aconoidasids	OG5_142870	PVX_117880	Plasmodium vivax Sal-1	hypothetical protein, conserved	2203	y						-	-	-	R
Coccidians + Aconoidasids	OG5_142870	TGME49_300100	Toxoplasma gondii ME49	rhoptry neck protein RON2 (RON2)	1479	y	y		y			-	-	-	
Coccidians + Aconoidasids	OG5_144162	PBANKA_134980	Plasmodium berghei ANKA	sporozoite surface protein 2,thrombospondin-related anonymous protein (TRAP)	606	y						-	-	-	R
Coccidians + Aconoidasids	OG5_144162	PY03052	Plasmodium yoelii yoelii 17XNL	sporozoite surface protein 2 precursor	827	y		y				-	-	-	R
Coccidians + Aconoidasids	OG5_147428	PF3D7_1222700	Plasmodium falciparum 3D7	glideosome-associated protein 45 (GAP45)	204				y			-	-	-	
Coccidians + Aconoidasids	OG5_147428	TGME49_223940	Toxoplasma gondii ME49	GAP45	245				y			-	-	-	
Coccidians + Aconoidasids	OG5_147452	PF3D7_1133400	Plasmodium falciparum 3D7	apical membrane antigen 1 (AMA1)	622	y		y	y			-	-	-	
Coccidians + Aconoidasids	OG5_147452	PVX_092275	Plasmodium vivax Sal-1	apical merozoite antigen 1	562	y						-	-	-	
Coccidians + Aconoidasids	OG5_147452	PY01581	Plasmodium yoelii yoelii 17XNL	apical membrane antigen-1	557	y						-	-	-	
Coccidians + Aconoidasids	OG5_147452	TGME49_300130	Toxoplasma gondii ME49	apical membrane antigen 1 domain-containing protein	553	y						-	-	-	
Coccidians + Aconoidasids	OG5_150068	PF3D7_0718100	Plasmodium falciparum 3D7	exported serine/threonine protein kinase (EST)	2695		y					-	-	-	
Coccidians + Aconoidasids	OG5_153587	PF3D7_0817700	Plasmodium falciparum 3D7	rhoptry neck protein 5 (RON5)	1156	y						-	-	-	
Coccidians + Aconoidasids	OG5_153621	PF3D7_1460600	Plasmodium falciparum 3D7	inner membrane complex protein, putative	148				y			-	-	-	
Coccidians + Aconoidasids	OG5_158050	PF3D7_0212600	Plasmodium falciparum 3D7	secreted protein altered thrombospondin repeat protein (SPATR)	250	y		y				-	-	-	
Coccidians + Aconoidasids	OG5_158050	PKH_040760	Plasmodium knowlesi strain H	secreted protein with altered thrombospondin repeat domain, putative (SPATR)	262			y				-	-	-	
Coccidians + Aconoidasids	OG5_158050	TGME49_293900	Toxoplasma gondii ME49	sporozoite protein with an altered thrombospondin repeat SPATR	534	y		y				-	-	-	
Coccidians + Aconoidasids	OG5_158060	PF3D7_1459900	Plasmodium falciparum 3D7	conserved Plasmodium membrane protein, unknown function	193	y						-	-	-	
Coccidians + Aconoidasids	OG5_158061	TGME49_243930	Toxoplasma gondii ME49	hypothetical protein	943	y						-	-	-	
Coccidians + Aconoidasids	OG5_160940	PF3D7_0405200	Plasmodium falciparum 3D7	ag-1 blood stage membrane protein homologue (PfMAg-1)	579		y	y				-	-	-	R
Coccidians + Aconoidasids	OG5_166882	PF3D7_1252100	Plasmodium falciparum 3D7	rhoptry neck protein 3 (RON3)	2215	y						-	-	-	
Coccidians + Aconoidasids	OG5_167220	PF3D7_0405900	Plasmodium falciparum 3D7	apical sushi protein (ASP)	731	y						-	-	-	
Coccidians + Aconoidasids	OG5_204863	TGME49_309590	Toxoplasma gondii ME49	rhoptry protein ROP1 (ROP1)	446	y						-	-	-	W
Coccidians + Aconoidasids	OG5_240105	NCLIV_028680	Neospora caninum Liverpool	putative apical membrane antigen 1	569	y						-	-	-	
Coccidians + Aconoidasids	OG5_240105	TGME49_255260	Toxoplasma gondii ME49	apical membrane antigen AMA1	569	y			y			-	-	-	
Coccidians + Aconoidasids	OG5_240150	TGME49_229010	Toxoplasma gondii ME49	rhoptry neck protein RON4 (RON4)	984	y	y		y			-	-	-	
Coccidians + Aconoidasids	OG5_240764	TGME49_223920	Toxoplasma gondii ME49	rhoptry neck protein RON3 (RON3)	1958	y						-	-	-	
Coccidians + Aconoidasids	OG5_240985	TGME49_311470	Toxoplasma gondii ME49	rhoptry neck protein RON5 (RON5)	1702	y	y		y			-	-	-	
Coccidians	NO_GROUP	TGME49_227810	Toxoplasma gondii ME49	rhoptry kinase family protein ROP11 (incomplete catalytic triad) (ROP11)	525	y						-	-	-	
Coccidians	OG5_149164	TGME49_215775	Toxoplasma gondii ME49	rhoptry protein ROP8 (ROP8)	575	y						-	-	-	W
Coccidians	OG5_149164	TGME49_295110	Toxoplasma gondii ME49	rhoptry protein ROP7 (ROP7)	575	y						-	-	-	W
Coccidians	OG5_149164	NCLIV_001970	Neospora caninum Liverpool	unspecified product	594	y						-	-	-	
Coccidians	OG5_166898	TGME49_277240	Toxoplasma gondii ME49	NTPase I	203	y		y			y	-	-	-	
Coccidians	OG5_204884	NCLIV_070010	Neospora caninum Liverpool	hypothetical protein, conserved	2995	y						-	-	-	R
Coccidians	OG5_204884	TGME49_306060	Toxoplasma gondii ME49	rhoptry neck protein RON8 (RON8)	2980	y	y		y			-	-	-	
Coccidians	OG5_239269	gi|310347	Sarcocystis muris	major microneme antigen	241	y						-	-	-	
Coccidians	OG5_239527	TGME49_319560	Toxoplasma gondii ME49	microneme protein MIC3 (MIC3)	383	y		y				-	-	-	
Coccidians	OG5_240045	gi|40549148	Eimeria tenella	microneme protein MIC3	448	y						-	-	-	
Coccidians	OG5_240045	TGME49_260190	Toxoplasma gondii ME49	microneme protein MIC13 (MIC13)	468	y						-	-	-	
Coccidians	OG5_240150	NCLIV_030050	Neospora caninum Liverpool	unspecified product	890	y						-	-	-	
Coccidians	OG5_240528	TGME49_288650	Toxoplasma gondii ME49	dense granule protein GRA12 (GRA12)	436	y						-	-	-	
Coccidians	OG5_240553	TGME49_289630	Toxoplasma gondii ME49	microneme protein MIC16 (MIC16)	668	y						-	-	-	
Coccidians	OG5_240594	NCLIV_043270	Neospora caninum Liverpool	putative microneme protein MIC1	460	y						-	-	-	
Coccidians	OG5_240594	TGME49_291890	Toxoplasma gondii ME49	microneme protein MIC1 (MIC1)	456	y						-	-	-	
Coccidians	OG5_240802	TGME49_234380	Toxoplasma gondii ME49	hypothetical protein	347	y						-	-	-	
Coccidians	OG5_241199	TGME49_218520	Toxoplasma gondii ME49	microneme protein MIC6 (MIC6)	349	y						-	-	-	
Cryptosporidium	OG5_169027	cgd7_4020	Cryptosporidium parvum Iowa II	cryptopsoridial mucin, large thr stretch, signal peptide sequence	1937	y		y				-	-	-	
Cryptosporidium	OG5_195453	cgd6_2330	Cryptosporidium parvum Iowa II	hypothetical protein with signal peptide	302	y						-	-	-	

Cryptosporidium	OG5_222599	cgd4_2330	Cryptosporidium parvum Iowa II	hypothetical protein	104	y		y				-	-	-	
Cryptosporidium	OG5_222676	cgd6_1080	Cryptosporidium parvum Iowa II	conserved hypothetical protein	324	y		y				-	-	-	
Cryptosporidium	OG5_222709	cgd6_5410	Cryptosporidium parvum Iowa II	signal peptide plus thr stretch, charged repeats, likely mucin	711	y		y			y	-	-	-	
Eimeria	NO_GROUP	gi|1279427	Eimeria tenella	acidic microneme protein precursor	342	y	y	y				-	-	-	
Eimeria	NO_GROUP	gi|37699772	Eimeria tenella	surface antigen 4	253			y				-	-	-	
Eimeria	NO_GROUP	gi|37699776	Eimeria tenella	surface antigen 8	265			y				-	-	-	
Eimeria	NO_GROUP	gi|37699788	Eimeria tenella	surface antigen 13	263			y				-	-	-	
Eimeria	NO_GROUP	gi|38602661	Eimeria tenella	surface antigen 1	253			y				-	-	-	
Eukaryotic	OG5_139232	TA17315	Theileria annulata strain Ankara	surface protein precursor (TaSP), putative	314			y				y	y	y	X
Eukaryotic	NO_GROUP	PF3D7_0113700	Plasmodium falciparum 3D7	heat shock protein 40, type II (HSP40)	402		y					y	y	y	
Eukaryotic	NO_GROUP	PF3D7_0902200	Plasmodium falciparum 3D7	serine/threonine protein kinase, FIKK family (FIKK9.3)	633		y					y	-	y	
Eukaryotic	NO_GROUP	PF3D7_1200800	Plasmodium falciparum 3D7	serine/threonine protein kinase, FIKK family (FIKK12)	567		y					y	-	y	
Eukaryotic	NO_GROUP	TGME49_205250	Toxoplasma gondii ME49	rhoptry protein ROP18 (ROP18)	554	y						y	y	y	
Eukaryotic	OG5_126561	cgd1_700	Cryptosporidium parvum Iowa II	ATP-binding cassette protein	1431	y						y	y	y	
Eukaryotic	OG5_126585	TGME49_244470	Toxoplasma gondii ME49	hypothetical protein	2595				y			y	y	y	
Eukaryotic	OG5_126585	PY04858	Plasmodium yoelii yoelii 17XNL	ookinete protein-related	2401	y						y	y	y	R
Eukaryotic	OG5_126595	PBANKA_145930	Plasmodium berghei ANKA	actin I	376	y						y	y	y	
Eukaryotic	OG5_126596	TGME49_260310	Toxoplasma gondii ME49	ATP-binding cassette transporter ABC.B1 (ABCB1)	1345			y				y	y	y	
Eukaryotic	OG5_126600	PF3D7_0717500	Plasmodium falciparum 3D7	calcium-dependent protein kinase 4 (CDPK4)	528			y				y	y	y	
Eukaryotic	OG5_126600	TGME49_305860	Toxoplasma gondii ME49	calcium-dependent protein kinase CDPK3 (CDPK3)	537	y						y	y	y	
Eukaryotic	OG5_126605	TGME49_316400	Toxoplasma gondii ME49	alpha tubulin TUBA1 (TUBA1)	453	y			y			y	y	y	
Eukaryotic	OG5_126607	PF3D7_1115400	Plasmodium falciparum 3D7	cysteine proteinase falcipain 3 (FP3)	492		y					y	y	y	
Eukaryotic	OG5_126607	PF3D7_1115700	Plasmodium falciparum 3D7	cysteine proteinase falcipain 2a (FP2A)	484					y		y	y	y	
Eukaryotic	OG5_126607	TGME49_321530	Toxoplasma gondii ME49	cathepsin CPL (CPL)	422	y						y	y	y	
Eukaryotic	OG5_126611	TGME49_266960	Toxoplasma gondii ME49	beta tubulin	449	y			y			y	y	y	
Eukaryotic	OG5_126615	PBANKA_091560	Plasmodium berghei ANKA	aquaglyceroporin (AQP)	258			y				y	y	y	
Eukaryotic	OG5_126619	gi|187340651	Eimeria tenella	microneme protein 4	2340	y		y				y	-	y	
Eukaryotic	OG5_126619	gi|31322550	Eimeria maxima	TFP250	2360	y						y	-	y	
Eukaryotic	OG5_126619	TGME49_267680	Toxoplasma gondii ME49	microneme protein MIC12 (MIC12)	2182	y						y	y	y	
Eukaryotic	OG5_126623	gi|2792527	Eimeria tenella	heat shock protein 90	713	y						y	y	y	
Eukaryotic	OG5_126626	NCLIV_028990	Neospora caninum Liverpool	hypothetical protein	1240	y						y	y	y	
Eukaryotic	OG5_126628	PF3D7_1462800	Plasmodium falciparum 3D7	glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	337	y						y	y	y	
Eukaryotic	OG5_126628	TGME49_289690	Toxoplasma gondii ME49	glyceraldehyde-3-phosphate dehydrogenase GAPDH1 (GAPDH1)	489				y			y	y	y	
Eukaryotic	OG5_126635	PF3D7_1246900	Plasmodium falciparum 3D7	RAC-beta serine/threonine protein kinase (PKB)	735	y		y				y	y	y	
Eukaryotic	OG5_126645	TGME49_312100	Toxoplasma gondii ME49	plasma membrane-type Ca(2+)-ATPase A1 PMCAA1 (PMCAA1)	1822			y				y	y	y	
Eukaryotic	OG5_126660	PF3D7_1145400	Plasmodium falciparum 3D7	dynamin-like protein (DYN1)	837		y	y				y	y	y	
Eukaryotic	OG5_126660	TGME49_321620	Toxoplasma gondii ME49	dynamin-related protein DRPB (DRPB)	860					y		y	y	y	
Eukaryotic	OG5_126672	PBANKA_135570	Plasmodium berghei ANKA	myosin A (MyoA)	817	y						y	y	y	
Eukaryotic	OG5_126672	PF3D7_1342600	Plasmodium falciparum 3D7	myosin A (MyoA)	818	y			y			y	y	y	
Eukaryotic	OG5_126672	PY01232	Plasmodium yoelii yoelii 17XNL	myosin A	817				y			y	y	y	
Eukaryotic	OG5_126672	TGME49_235470	Toxoplasma gondii ME49	myosin A	831				y			y	y	y	
Eukaryotic	OG5_126683	TGME49_272500	Toxoplasma gondii ME49	sugar transporter ST2 (ST2)	689			y				y	y	y	
Eukaryotic	OG5_126685	gi|2854033	Eimeria tenella	pyruvate kinase	531	y						y	y	y	
Eukaryotic	OG5_126685	TGME49_256760	Toxoplasma gondii ME49	pyruvate kinase PyK1 (PYKI)	531				y			y	y	y	
Eukaryotic	OG5_126696	TGME49_276140	Toxoplasma gondii ME49	ADP ribosylation factor ARF1 (ARF1)	554					y		y	y	y	
Eukaryotic	OG5_126698	gi|13774515	Eimeria tenella	enolase	445	y						y	y	y	
Eukaryotic	OG5_126698	PF3D7_1015900	Plasmodium falciparum 3D7	enolase (ENO)	446			y				y	y	y	
Eukaryotic	OG5_126698	PY06644	Plasmodium yoelii yoelii 17XNL	enolase	455			y				y	y	y	
Eukaryotic	OG5_126706	TGME49_263090	Toxoplasma gondii ME49	14-3-3 protein	323			y				y	y	y	
Eukaryotic	OG5_126724	TGME49_267810	Toxoplasma gondii ME49	Rab 5	241					y		y	y	y	
Eukaryotic	OG5_126728	PF3D7_1340900	Plasmodium falciparum 3D7	sodium-dependent phosphate transporter (PiT)	687			y				y	y	y	
Eukaryotic	OG5_126729	TGME49_299060	Toxoplasma gondii ME49	sodium/hydrogen exchanger NHE2	823	y						y	y	y	
Eukaryotic	OG5_126744	TA14920	Theileria annulata strain Ankara	heat shock protein HSP70 homologue, putative	681		y					y	y	y	
Eukaryotic	OG5_126773	PF3D7_0501100	Plasmodium falciparum 3D7	heat shock protein 40, type II (HSP40)	402		y					y	y	y	

Eukaryotic	OG5_126843	PBANKA_070280	Plasmodium berghei ANKA	protein disulfide isomerase	482	y						y	y	y	
Eukaryotic	OG5_126855	PF3D7_0904900	Plasmodium falciparum 3D7	Cu2 -transporting ATPase, putative (CUP)	2568		y	y				y	y	y	
Eukaryotic	OG5_126868	PF3D7_0210700	Plasmodium falciparum 3D7	syntaxin, Qa-SNARE family (SYN17)	314			y				y	y	y	
Eukaryotic	OG5_126885	TGME49_201840	Toxoplasma gondii ME49	aspartyl protease ASP1 (ASP1)	619	y			y			y	y	y	
Eukaryotic	OG5_126888	PF3D7_0312400	Plasmodium falciparum 3D7	glycogen synthase kinase 3 (GSK3)	440		y					y	y	y	
Eukaryotic	OG5_126911	TGME49_232350	Toxoplasma gondii ME49	lactate dehydrogenase LDH1 (LDH1)	329				y			y	y	y	
Eukaryotic	OG5_126934	PF3D7_1121700	Plasmodium falciparum 3D7	protein GCN20 (GCN20)	815		y					y	y	y	
Eukaryotic	OG5_126935	TGME49_289600	Toxoplasma gondii ME49	heat shock protein HSP29 (HSP29)	272			y				y	-	y	
Eukaryotic	OG5_126971	PBANKA_141890	Plasmodium berghei ANKA	Rab GTPase 11a (Rab11a)	216	y						y	y	y	
Eukaryotic	OG5_126971	PF3D7_1320600	Plasmodium falciparum 3D7	Rab GTPase 11a (RAB11a)	216	y		y				y	y	y	
Eukaryotic	OG5_126971	TGME49_289680	Toxoplasma gondii ME49	Ras-related protein Rab11	219	y						y	y	y	
Eukaryotic	OG5_127051	PF3D7_1130200	Plasmodium falciparum 3D7	60S ribosomal protein P0 (PfP0)	316			y				y	y	y	
Eukaryotic	OG5_127053	PF3D7_0505800	Plasmodium falciparum 3D7	small ubiquitin-related modifier, putative (SUMO)	100		y					y	y	y	
Eukaryotic	OG5_127061	PF3D7_0406100	Plasmodium falciparum 3D7	vacuolar ATP synthase subunit b	494		y	y				y	y	y	
Eukaryotic	OG5_127126	PF3D7_0516600	Plasmodium falciparum 3D7	translation initiation factor IF-2,sporozoite surface antigen MB2 (MB2)	1610			y				y	y	y	
Eukaryotic	OG5_127135	PBANKA_030250	Plasmodium berghei ANKA	hexose transporter (HT)	505			y				y	y	y	
Eukaryotic	OG5_127135	PF3D7_0204700	Plasmodium falciparum 3D7	hexose transporter (HT)	504			y				y	y	y	
Eukaryotic	OG5_127135	TGME49_214320	Toxoplasma gondii ME49	facilitative glucose transporter GT1 (GT1)	568			y				y	y	y	
Eukaryotic	OG5_127143	TGME49_236040	Toxoplasma gondii ME49	fructose-1,6-bisphosphate aldolase	433	y			y			y	y	y	
Eukaryotic	OG5_127203	PF3D7_0822600	Plasmodium falciparum 3D7	protein transport protein Sec23 (SEC23)	759		y					y	y	y	
Eukaryotic	OG5_127226	TGME49_285980	Toxoplasma gondii ME49	glucosephosphate-mutase GPM1 (GPM1)	637	y						y	y	y	
Eukaryotic	OG5_127237	PF3D7_1311900	Plasmodium falciparum 3D7	vacuolar ATP synthase subunit a (vapA)	611			y				y	y	y	
Eukaryotic	OG5_127253	TGME49_252640	Toxoplasma gondii ME49	P-type ATPase PMA1	1151			y				y	y	y	
Eukaryotic	OG5_127256	PF3D7_1144900	Plasmodium falciparum 3D7	Rab GTPase 6 (RAB6)	207					y		y	y	y	
Eukaryotic	OG5_127256	TGME49_310460	Toxoplasma gondii ME49	Rab6	209	y				y		y	y	y	
Eukaryotic	OG5_127279	PF3D7_1353600	Plasmodium falciparum 3D7	ER lumen protein retaining receptor (ERD2)	221					y		y	y	y	
Eukaryotic	OG5_127279	TGME49_256000	Toxoplasma gondii ME49	endoplasmic reticulum retention receptor	223	y						y	y	y	
Eukaryotic	OG5_127331	PF3D7_0416800	Plasmodium falciparum 3D7	small GTP-binding protein sar1 (SAR1)	192		y					y	y	y	
Eukaryotic	OG5_127374	PF3D7_1012400	Plasmodium falciparum 3D7	hypoxanthine-guanine phosphoribosyltransferase (HGPRT)	231		y					y	y	y	
Eukaryotic	OG5_127374	TGME49_200320	Toxoplasma gondii ME49	hypoxanthine-xanthine-guanine phosphoribosyl transferase HXGPRT (HXGPRT)	279				y			y	y	y	
Eukaryotic	OG5_127474	TGME49_289770	Toxoplasma gondii ME49	mu1 adaptin	430	y				y		y	y	y	
Eukaryotic	OG5_127498	TGME49_223000	Toxoplasma gondii ME49	dynein light chain DLC	138	y			y			y	y	y	
Eukaryotic	OG5_127506	PF3D7_0214100	Plasmodium falciparum 3D7	protein transport protein sec31 (SEC31)	1471		y					y	y	y	
Eukaryotic	OG5_127599	TGME49_311360	Toxoplasma gondii ME49	protein kinase G AGC kinase family member PKG (PKG)	994			y				y	y	y	
Eukaryotic	OG5_127617	PF3D7_0511000	Plasmodium falciparum 3D7	translationally-controlled tumor protein homolog (TCTP)	171			y				y	y	y	
Eukaryotic	OG5_127665	PF3D7_0303000	Plasmodium falciparum 3D7	N-ethylmaleimide sensitive fusion protein (NSF)	783		y					y	y	y	
Eukaryotic	OG5_127790	PF3D7_0106700	Plasmodium falciparum 3D7	small ribosomal subunit assembling AARP2 protein (AARP2)	1434		y					y	y	y	
Eukaryotic	OG5_127800	TGME49_249670	Toxoplasma gondii ME49	cathepsin B	572	y						y	y	y	
Eukaryotic	OG5_127887	PF3D7_1027300	Plasmodium falciparum 3D7	peroxiredoxin (nPrx)	393				y			y	y	y	
Eukaryotic	OG5_127926	PF3D7_1332000	Plasmodium falciparum 3D7	syntaxin, Qa-SNARE family (SYN5)	281					y		y	y	y	
Eukaryotic	OG5_128071	PF3D7_1113100	Plasmodium falciparum 3D7	protein tyrosine phosphatase (PRL)	218	y						y	y	y	
Eukaryotic	OG5_128083	PF3D7_1017300	Plasmodium falciparum 3D7	golgi re-assembly stacking protein 2 (GRASP)	573					y		y	y	y	
Eukaryotic	OG5_128083	TGME49_261980	Toxoplasma gondii ME49	gorasp2-prov protein	475					y		y	y	y	
Eukaryotic	OG5_128181	PF3D7_0515300	Plasmodium falciparum 3D7	phosphatidylinositol 3-kinase (PI3K)	2133		y					y	y	y	
Eukaryotic	OG5_128249	cgd6_4840	Cryptosporidium parvum Iowa II	serine protease, subtilase family, signal peptide	1197	y						y	y	y	
Eukaryotic	OG5_128374	PF3D7_1457000	Plasmodium falciparum 3D7	signal peptide peptidase (SPP)	412	y						y	y	y	
Eukaryotic	OG5_128390	PF3D7_0320100	Plasmodium falciparum 3D7	SNARE protein (SEC22)	221		y			y		y	y	y	
Eukaryotic	OG5_128720	PF3D7_1412000	Plasmodium falciparum 3D7	p1/s1 nuclease, putative	320	y						y	y	y	

Eukaryotic	OG5_128905	PF3D7_1235200	Plasmodium falciparum 3D7	V-type K -independent H -translocating inorganic pyrophosphatase (VP2)	1057			y				y	y	y	
Eukaryotic	OG5_128905	PF3D7_1456800	Plasmodium falciparum 3D7	V-type H( )-translocating pyrophosphatase, putative (VP1)	717			y				y	y	y	
Eukaryotic	OG5_128905	TGME49_248670	Toxoplasma gondii ME49	V-type H(+)-translocating pyrophosphatase VP1	816	y		y				y	y	y	
Eukaryotic	OG5_128917	TGME49_259530	Toxoplasma gondii ME49	GalNac	751					y		y	y	y	
Eukaryotic	OG5_128972	PF3D7_1031200	Plasmodium falciparum 3D7	MORN repeat-containing protein 1 (MORN1)	364	y						y	y	y	
Eukaryotic	OG5_128972	TGME49_310440	Toxoplasma gondii ME49	membrane occupation and recognition nexus protein MORN1 (MORN1)	363	y			y			y	y	y	
Eukaryotic	OG5_128979	cgd4_2720	Cryptosporidium parvum Iowa II	P-type ATpase involved in cation transport	1528	y						y	y	y	
Eukaryotic	OG5_129142	TGME49_234570	Toxoplasma gondii ME49	sterol carrier protein-2 HAD-2SCP-2 (HAD2SCP2)	625	y						y	y	y	
Eukaryotic	OG5_129186	PBANKA_020930	Plasmodium berghei ANKA	actin-related protein (ARP1)	384	y						y	y	y	
Eukaryotic	OG5_129627	PF3D7_0918000	Plasmodium falciparum 3D7	secreted acid phosphatase (GAP50)	396	y			y			y	y	y	
Eukaryotic	OG5_129627	TGME49_219320	Toxoplasma gondii ME49	acid phosphatase GAP50 (GAP50)	431				y			y	y	y	
Eukaryotic	OG5_129645	BBOV_IV007730	Babesia bovis T2Bo	cysteine protease 2	445		y					y	y	y	
Eukaryotic	OG5_129767	PF3D7_1229400	Plasmodium falciparum 3D7	macrophage migration inhibitory factor (MIF)	116		y					y	y	y	
Eukaryotic	OG5_130026	PF3D7_1211400	Plasmodium falciparum 3D7	heat shock protein DNAJ homologue Pfj4 (PfJ4)	244		y					y	y	y	
Eukaryotic	OG5_130494	TGME49_289620	Toxoplasma gondii ME49	cathepsin CPC1 (CPC1)	733	y						y	y	y	
Eukaryotic	OG5_130542	PBANKA_141030	Plasmodium berghei ANKA	M1-family aminopeptidase, putative	1064	y						y	y	y	
Eukaryotic	OG5_131119	TGME49_248830	Toxoplasma gondii ME49	phosphoinositide phospholipase PIPLC (PIPLC)	1097			y				y	y	y	
Eukaryotic	OG5_131558	PF3D7_0718300	Plasmodium falciparum 3D7	cysteine repeat modular protein 2 (CRMP2)	2543		y	y				y	y	y	
Eukaryotic	OG5_131558	PF3D7_0911300	Plasmodium falciparum 3D7	cysteine repeat modular protein 1 (CRMP1)	3303		y					y	y	y	
Eukaryotic	OG5_132039	TGME49_242720	Toxoplasma gondii ME49	aspartyl protease ASP5 (ASP5)	1012					y		y	y	y	
Eukaryotic	OG5_132295	PF3D7_1343000	Plasmodium falciparum 3D7	phosphoethanolamine N-methyltransferase (PMT)	266					y		y	y	y	
Eukaryotic	OG5_132467	PBANKA_030490	Plasmodium berghei ANKA	serine repeat antigen 3 (SERA3)	1096		y					y	y	y	
Eukaryotic	OG5_132467	PBANKA_030500	Plasmodium berghei ANKA	serine repeat antigen 2 (SERA2)	1122		y					y	y	y	
Eukaryotic	OG5_132803	TGME49_315560	Toxoplasma gondii ME49	ATP-binding cassette G family transporter ABCG77 (ABCG77)	1163			y		y		y	y	y	
Eukaryotic	OG5_133090	TGME49_315220	Toxoplasma gondii ME49	rhoptry protein ROP14 (ROP14)	1061	y						y	y	y	
Eukaryotic	OG5_133546	PF3D7_0808200	Plasmodium falciparum 3D7	plasmepsin X	573	y						y	y	y	
Eukaryotic	OG5_133546	TGME49_246550	Toxoplasma gondii ME49	aspartyl protease ASP3 (ASP3)	643					y		y	y	y	
Eukaryotic	OG5_134280	PBANKA_082420	Plasmodium berghei ANKA	perforin like protein 3 (PPLP3)	815	y						-	y	y	
Eukaryotic	OG5_134280	PBANKA_100630	Plasmodium berghei ANKA	perforin like protein 1,sporozoite micronemal protein essential for cell traversal (SPECT2)	810	y						-	y	y	
Eukaryotic	OG5_134280	PY00454	Plasmodium yoelii yoelii 17XNL	MAC/Perforin domain, putative	727	y						-	y	y	
Eukaryotic	OG5_134280	TGME49_204130	Toxoplasma gondii ME49	perforin-like protein PLP1 (PLP1)	1161	y						-	-	y	
Eukaryotic	OG5_134639	TGME49_259200	Toxoplasma gondii ME49	Na+/H+ exchanger NHE1	2288			y				y	y	y	
Eukaryotic	OG5_134918	PF3D7_0522600	Plasmodium falciparum 3D7	inner membrane complex protein	483				y			y	-	y	
Eukaryotic	OG5_134971	gi|5708122	Eimeria tenella	microneme protein 5	932	y		y				y	-	-	X
Eukaryotic	OG5_135032	PBANKA_103520	Plasmodium berghei ANKA	LCCL domain-containing protein (CCp3)	1304	y		y				y	y	y	X
Eukaryotic	OG5_135032	PF3D7_1407000	Plasmodium falciparum 3D7	LCCL domain-containing protein (CCp3)	1272		y	y				y	y	y	X
Eukaryotic	OG5_135662	PF3D7_0109100	Plasmodium falciparum 3D7	LCCL domain-containing protein (CCp5)	1029			y				y	y	y	
Eukaryotic	OG5_135662	PF3D7_1451600	Plasmodium falciparum 3D7	LCCL-like protein (FNPA)	865			y				y	y	y	
Eukaryotic	OG5_135731	PF3D7_0506900	Plasmodium falciparum 3D7	rhomboid protease ROM4 (ROM4)	759			y				y	y	y	
Eukaryotic	OG5_135731	TGME49_268590	Toxoplasma gondii ME49	rhomboid protease ROM4 (ROM4)	665			y				y	y	y	
Eukaryotic	OG5_136613	PF3D7_1479000	Plasmodium falciparum 3D7	acyl-CoA synthetase (ACS1a)	820		y					y	y	y	
Eukaryotic	OG5_136898	PF3D7_1227200	Plasmodium falciparum 3D7	potassium channel protein (K1)	1966		y					y	y	y	
Eukaryotic	OG5_137241	PBANKA_093350	Plasmodium berghei ANKA	rhomboid protease ROM1 (ROM1)	278	y		y				y	y	y	
Eukaryotic	OG5_137241	TGME49_200290	Toxoplasma gondii ME49	rhomboid protease ROM1 (ROM1)	293	y				y		y	y	y	
Eukaryotic	OG5_137241	TGME49_263290	Toxoplasma gondii ME49	rhomboid protease ROM2 (ROM2)	283			y		y		y	y	y	
Eukaryotic	OG5_138369	NCLIV_021050	Neospora caninum Liverpool	unspecified product	865	y						y	y	y	
Eukaryotic	OG5_138369	PF3D7_0507500	Plasmodium falciparum 3D7	subtilisin-like protease 1 (SUB1)	688	y	y					y	y	y	
Eukaryotic	OG5_138369	TGME49_204050	Toxoplasma gondii ME49	subtilisin SUB1 (SUB1)	761	y						y	y	y	
Eukaryotic	OG5_138427	gi|110681634	Plasmodium gallinaceum	circumsporozoite protein/thrombospondin-related anonymous protein-related protein	1875	y						y	y	y	
Eukaryotic	OG5_138427	gi|156257452	Eimeria tenella	microneme 1 precursor	675	y						y	y	y	

Eukaryotic	OG5_138427	NCLIV_022970	Neospora caninum Liverpool	unspecified product	765	y						y	y	y	
Eukaryotic	OG5_138427	NCLIV_033690	Neospora caninum Liverpool	hypothetical protein	756	y						y	y	y	X
Eukaryotic	OG5_138427	PBANKA_041290	Plasmodium berghei ANKA	circumsporozoite- and TRAP-related protein (CTRP)	1905	y						y	y	y	
Eukaryotic	OG5_138427	PVX_095475	Plasmodium vivax Sal-1	CTRP adhesive protein (invasive stage), putative	2106	y						y	y	y	X
Eukaryotic	OG5_138427	TGME49_201780	Toxoplasma gondii ME49	microneme protein MIC2 (MIC2)	769	y		y				y	y	y	X
Eukaryotic	OG5_138807	TGME49_250340	Toxoplasma gondii ME49	centrin 2	170	y						y	y	y	
Eukaryotic	OG5_139187	PF3D7_0530900	Plasmodium falciparum 3D7	formin 1	2675	y			y			y	-	y	
Eukaryotic	OG5_139187	TGME49_206430	Toxoplasma gondii ME49	formin FRM1 (FRM1)	5009	y		y	y			y	y	y	
Eukaryotic	OG5_139194	NCLIV_056880	Neospora caninum Liverpool	conserved hypothetical protein	1525	y						y	y	y	
Eukaryotic	OG5_139606	cgd7_1730	Cryptosporidium parvum Iowa II	CpCCp1/Cpa135	1556	y						y	y	-	X
Eukaryotic	OG5_139606	PF3D7_1475500	Plasmodium falciparum 3D7	LCCL domain-containing protein (CCp1)	1620		y	y				y	y	-	X
Eukaryotic	OG5_140063	TGME49_320480	Toxoplasma gondii ME49	Rab11b	226				y	y		y	y	y	
Eukaryotic	OG5_140247	PF3D7_0628200	Plasmodium falciparum 3D7	protein kinase PK4 (PK4)	3072	y						y	y	y	
Eukaryotic	OG5_141680	PF3D7_0303400	Plasmodium falciparum 3D7	palmitoyl transferase (AnkDHHC)	616					y		y	y	y	
Eukaryotic	OG5_141740	TGME49_249850	Toxoplasma gondii ME49	GAP40 protein (GAP40)	385				y			y	y	y	
Eukaryotic	OG5_141742	PF3D7_1455800	Plasmodium falciparum 3D7	LCCL domain-containing protein (CCp2)	1617		y	y				y	y	y	X
Eukaryotic	OG5_141753	TGME49_219280	Toxoplasma gondii ME49	actin like protein ALP1 (ALP1)	414				y			y	y	y	
Eukaryotic	OG5_141757	PF3D7_1246400	Plasmodium falciparum 3D7	myosin light chain 1,myosin A tail domain interacting protein (MTIP)	204				y			y	y	y	
Eukaryotic	OG5_141757	PY00409	Plasmodium yoelii yoelii 17XNL	myosin A tail domain interacting protein MTIP	204				y			y	y	y	
Eukaryotic	OG5_141757	TGME49_257680	Toxoplasma gondii ME49	myosin light chain MLC1	213			y	y			y	y	y	
Eukaryotic	OG5_143200	cgd2_70	Cryptosporidium parvum Iowa II	ABC transporter, with 12 x transmembrane domains and 2x AAA domains	1587	y						y	y	y	
Eukaryotic	OG5_143207	TGME49_243730	Toxoplasma gondii ME49	rhoptry protein ROP9 (ROP9)	353	y						y	y	y	
Eukaryotic	OG5_144162	PF3D7_1335900	Plasmodium falciparum 3D7	sporozoite surface protein 2 (TRAP)	574	y		y				y	y	y	X
Eukaryotic	OG5_144162	BBOV_II002650	Babesia bovis T2Bo	thrombospondin-related anonymous protein	660	y						-	-	y	
Eukaryotic	OG5_144357	TGME49_299070	Toxoplasma gondii ME49	pyruvate kinase PyKII (PYKII)	987	y						y	y	y	
Eukaryotic	OG5_144441	gi|164710612	Babesia divergens	HSP-20	177		y					y	-	y	
Eukaryotic	OG5_144441	TGME49_232940	Toxoplasma gondii ME49	heat shock protein HSP20 (HSP20)	185	y		y				y	-	y	
Eukaryotic	OG5_144709	PF3D7_0112200	Plasmodium falciparum 3D7	multidrug resistance-associated protein 1 (MRP1)	1822			y				y	y	y	
Eukaryotic	OG5_144709	PF3D7_1229100	Plasmodium falciparum 3D7	ABC transporter, (CT family) (MRP2)	2108			y				y	-	y	
Eukaryotic	OG5_145156	TGME49_228170	Toxoplasma gondii ME49	inner membrane complex protein IMC2A (IMC2A)	1491				y			y	y	y	
Eukaryotic	OG5_145286	PF3D7_0404700	Plasmodium falciparum 3D7	dipeptidyl peptidase 3 (DPAP3)	939	y						-	-	y	X
Eukaryotic	OG5_145286	TGME49_276130	Toxoplasma gondii ME49	cathepsin CPC2 (CPC2)	753	y						y	y	y	
Eukaryotic	OG5_146338	cgd6_2090	Cryptosporidium parvum Iowa II	CpCOWP1, oocyst wall protein with type I and type II cysteine-rich repeats	1629						y	y	-	y	
Eukaryotic	OG5_146451	TGME49_200230	Toxoplasma gondii ME49	microneme protein MIC17C (MIC17C)	603	y						y	-	-	X
Eukaryotic	OG5_146451	TGME49_200250	Toxoplasma gondii ME49	microneme protein MIC17A (MIC17A)	351	y						y	y	-	X
Eukaryotic	OG5_146451	NCLIV_038110	Neospora caninum Liverpool	hypothetical protein	351	y						y	y	-	X
Eukaryotic	OG5_146451	TGME49_200240	Toxoplasma gondii ME49	microneme protein MIC17B (MIC17B)	351	y						y	y	-	X
Eukaryotic	OG5_146451	TGME49_200270	Toxoplasma gondii ME49	PAN/Apple domain-containing protein	351	y						y	y	-	X
Eukaryotic	OG5_146472	PF3D7_0102200	Plasmodium falciparum 3D7	ring-infected erythrocyte surface antigen (RESA)	1085	y	y					y	y	y	
Eukaryotic	OG5_146508	PF3D7_1347200	Plasmodium falciparum 3D7	nucleoside transporter 1 (NT1)	422			y				-	y	y	
Eukaryotic	OG5_146654	PF3D7_1465500	Plasmodium falciparum 3D7	potassium channel protein (K2)	1461		y					-	-	y	
Eukaryotic	OG5_149511	PF3D7_1136900	Plasmodium falciparum 3D7	subtilisin-like protease 2 (SUB2)	1341	y						y	y	y	
Eukaryotic	OG5_151343	PF3D7_1315100	Plasmodium falciparum 3D7	serine/threonine protein kinase (PK9)	367	y						y	y	y	
Eukaryotic	OG5_153039	PF3D7_0310400	Plasmodium falciparum 3D7	parasite-infected erythrocyte surface protein (PIESP1)	1294		y					-	-	y	X
Eukaryotic	OG5_156053	PVX_111175	Plasmodium vivax Sal-1	ookinete surface protein Pvs25	219			y				y	-	y	
Eukaryotic	OG5_156776	PF3D7_0910600	Plasmodium falciparum 3D7	SNARE protein (YKT6.1)	199					y		y	y	y	
Eukaryotic	OG5_158979	TGME49_314500	Toxoplasma gondii ME49	subtilisin SUB2 (SUB2)	1300	y						y	y	y	
Eukaryotic	OG5_163911	PF3D7_0503600	Plasmodium falciparum 3D7	myosin B (MyoB)	801	y						y	y	y	
Eukaryotic	OG5_164672	PF3D7_1206000	Plasmodium falciparum 3D7	shewanella-like protein phosphatase 2, putative (SHLP2)	304	y						y	y	y	
Eukaryotic	OG5_165002	PF3D7_1339900	Plasmodium falciparum 3D7	ABC transporter (MDR family), putative (MDR5)	925			y				y	-	y	

Eukaryotic	OG5_165382	PF3D7_0104200	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	466		y					y	-	y	
Eukaryotic	OG5_165464	TGME49_204420	Toxoplasma gondii ME49	oocyst wall protein OWP1 (OWP1)	499						y	y	-	y	
Eukaryotic	OG5_166914	PF3D7_0410000	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	833		y					y	-	y	
Eukaryotic	OG5_168383	cgd1_3500	Cryptosporidium parvum Iowa II	thrombospondin related adhesive protein	687	y						y	y	y	
Eukaryotic	OG5_171965	TGME49_206510	Toxoplasma gondii ME49	toxolysin TLN4 (TLN4)	2340	y						y	y	y	
Eukaryotic	OG5_171990	TGME49_268310	Toxoplasma gondii ME49	oocyst wall protein OWP3 (OWP3)	640						y	y	-	y	
Eukaryotic	OG5_175386	TGME49_206580	Toxoplasma gondii ME49	formin FRM2 (FRM2)	4709				y			y	y	y	
Eukaryotic	OG5_193087	TGME49_259020	Toxoplasma gondii ME49	bradyzoite antigen BAG1 (BAG1)	229			y				y	-	y	
Eukaryotic	OG5_195418	cgd6_780	Cryptosporidium parvum Iowa II	CpTSP8,predicted extracellular protein with 3 TSP1 repeats, an EGF domain and a C-terminal transmembrane domain. adjacent TSP domain containing gene.	625	y		y				y	y	y	
Eukaryotic	OG5_204795	TGME49_293770	Toxoplasma gondii ME49	chitinase-like protein CLP1 (CLP1)	714	y						-	-	y	
Eukaryotic	OG5_205532	PF3D7_0902500	Plasmodium falciparum 3D7	serine/threonine protein kinase, FIKK family (FIKK9.6)	591		y					y	-	y	
Eukaryotic	OG5_209056	TGME49_247540	Toxoplasma gondii ME49	ATP-binding cassette G family transporter ABCG107 (ABCG107)	981			y				y	y	y	
Eukaryotic	OG5_225241	TGME49_246930	Toxoplasma gondii ME49	calmodulin CAM1 (CAM1)	179	y			y			y	y	y	
Eukaryotic	OG5_231889	TGME49_245490	Toxoplasma gondii ME49	microneme protein MIC8 (MIC8)	684	y						y	y	y	
Eukaryotic	OG5_239228	TGME49_294690	Toxoplasma gondii ME49	rhomboid protease ROM5 (ROM5)	841			y				y	y	y	
Eukaryotic	OG5_239269	TGME49_208030	Toxoplasma gondii ME49	microneme protein MIC4 (MIC4)	580	y						y	-	-	X
Eukaryotic	OG5_239269	gi|7108680	Sarcocystis muris	microneme protein SML2	138	y						y	-	-	X
Eukaryotic	OG5_239636	TGME49_286150	Toxoplasma gondii ME49	PAN/Apple domain-containing protein	407	y						y	y	-	X
Eukaryotic	OG5_239946	TGME49_282055	Toxoplasma gondii ME49	protein phosphatase PP2C-hn (PP2CHN)	479	y	y					y	y	y	
Eukaryotic	OG5_239994	TGME49_262730	Toxoplasma gondii ME49	rhoptry protein ROP16 (ROP16)	707	y						y	y	y	
Eukaryotic	OG5_240010	TGME49_262010	Toxoplasma gondii ME49	calmodulin CAM2 (CAM2)	142	y			y			y	y	y	
Eukaryotic	OG5_240014	TGME49_261780	Toxoplasma gondii ME49	microneme protein MIC7 (MIC7)	340	y						y	y	y	
Eukaryotic	OG5_241405	gi|32816824	Sarcocystis neurona	nucleoside triphosphate hydrolase	714	y						-	-	y	
Eukaryotic	OG5_243453	PF3D7_0424500	Plasmodium falciparum 3D7	serine/threonine protein kinase, FIKK family (FIKK4.1)	622		y					y	-	y	
Eukaryotic	OG5_143115	PF3D7_0414900	Plasmodium falciparum 3D7	armadillo-domain containing rhoptry protein (ARO)	275	y						y	y	y	X,C
Eukaryotic	OG5_136403	TGME49_231630	Toxoplasma gondii ME49	alveolin domain containing intermediate filament IMC4 (ALV4)	452	y			y			y	y	y	X,C
Eukaryotic	OG5_139161	TGME49_275670	Toxoplasma gondii ME49	alveolin domain containing intermediate filament IMC15 (ALV5)	692				y			y	y	y	X
Eukaryotic	OG5_141746	NCLIV_000610	Neospora caninum Liverpool	putative profilin family protein	163	y						y	y	y	X
Eukaryotic	OG5_144120	TGME49_216000	Toxoplasma gondii ME49	alveolin domain containing intermediate filament IMC3 (ALV3)	538				y			y	y	y	X,C
Eukaryotic	OG5_146323	PF3D7_0525800	Plasmodium falciparum 3D7	membrane skeletal protein IMC1-related	300				y			y	y	y	X,C
Eukaryotic	OG5_156798	PF3D7_0304000	Plasmodium falciparum 3D7	inner membrane complex protein 1a, putative (IMC1a)	861				y			y	y	y	X,C
Eukaryotic	OG5_156798	TGME49_231640	Toxoplasma gondii ME49	alveolin domain containing intermediate filament IMC1 (ALV1)	609				y			y	y	y	X,C
Eukaryotic	OG5_240413	TGME49_268900	Toxoplasma gondii ME49	dense granular protein GRA10 (GRA10)	894	y						y	y	y	X
Myzozoans	OG5_153786	PBANKA_090710	Plasmodium berghei ANKA	inner membrane complex protein 1b (IMC1b)	535	y						-	-	-	X
Myzozoans	OG5_165465	PF3D7_1321900	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	292	y						y	y	-	X,W
Myzozoans	OG5_147397	TA20205	Theileria annulata strain Ankara	schizont protein e	441		y					y	y	-	X,W
Myzozoans	OG5_144710	PF3D7_1003600	Plasmodium falciparum 3D7	membrane skeletal protein IMC1-related (ALV5)	281				y			y	y	-	X
Myzozoans	OG5_191678	TA19865	Theileria annulata strain Ankara	surface protein d precursor	178		y	y				y	y	-	X
Piroplasmids	OG5_139232	TP04_0051	Theileria parva strain Muguga	polymorphic immunodominant molecule	480			y				-	-	-	R
Piroplasmids	OG5_153115	PF3D7_1233600	Plasmodium falciparum 3D7	asparagine and aspartate rich protein 1 (AARP1)	5439		y					-	-	-	R
Piroplasmids	OG5_156194	PY02159	Plasmodium yoelii yoelii 17XNL	Drosophila melanogaster CG15040 gene product	908	y						-	-	-	R
Piroplasmids	OG5_191821	BBOV_IV009860	Babesia bovis T2Bo	rhoptry-associated protein 1 (RAP-1)	565		y					-	-	-	W
Piroplasmids	OG5_191821	gi|255708246	Babesia gibsoni	rhoptry-associated protein-1b	484	y						-	-	-	
Piroplasmids	OG5_191821	gi|255708248	Babesia gibsoni	rhoptry-associated protein-1c	437	y						-	-	-	
Piroplasmids	OG5_191821	gi|88770794	Babesia gibsoni	rhoptry-associated protein	474	y	y	y				-	-	-	W
Plasmodium	NO_GROUP	PF3D7_0202200	Plasmodium falciparum 3D7	Plasmodium exported protein, unknown function	290		y					-	-	-	
Plasmodium	NO_GROUP	PF3D7_0424100	Plasmodium falciparum 3D7	reticulocyte binding protein homologue 5 (RH5)	526	y			y			-	-	-	
Plasmodium	NO_GROUP	PF3D7_0424200	Plasmodium falciparum 3D7	reticulocyte binding protein homologue 4 (RH4)	1716	y						-	-	-	
Plasmodium	NO_GROUP	PF3D7_0501200	Plasmodium falciparum 3D7	parasite-infected erythrocyte surface protein (PIESP2)	408		y					-	-	-	R
Plasmodium	NO_GROUP	PF3D7_0935900	Plasmodium falciparum 3D7	ring-exported protein 1 (REX1)	713		y					-	-	-	
Plasmodium	NO_GROUP	PF3D7_0936800	Plasmodium falciparum 3D7	Plasmodium exported protein (PHISTc), unknown function	383		y					-	-	-	R

Plasmodium	NO_GROUP	PF3D7_1016900	Plasmodium falciparum 3D7	early transcribed membrane protein 10.3 (ETRAMP10.3)	108		y					-	-	-	
Plasmodium	NO_GROUP	PF3D7_1035300	Plasmodium falciparum 3D7	glutamate-rich protein (GLURP)	1233	y	y					-	-	-	R
Plasmodium	NO_GROUP	PF3D7_1115800	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	304	y						-	-	-	
Plasmodium	NO_GROUP	PF3D7_1243400	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	443	y						-	-	-	
Plasmodium	NO_GROUP	PF3D7_1353200	Plasmodium falciparum 3D7	membrane associated histidine-rich protein (MAHRP2)	137		y					-	-	-	
Plasmodium	OG5_126585	PF3D7_0500800	Plasmodium falciparum 3D7	mature parasite-infected erythrocyte surface antigen,erythrocyte membrane protein 2 (MESA)	1434		y					-	-	-	R
Plasmodium	OG5_126585	PF3D7_1038400	Plasmodium falciparum 3D7	gametocyte-specific protein (Pf11-1)	9563		y					-	-	-	R
Plasmodium	OG5_126825	PF3D7_0901000	Plasmodium falciparum 3D7	rifin (RIF)	309		y					-	-	-	
Plasmodium	OG5_127056	PBANKA_124630	Plasmodium berghei ANKA	BIR protein	294		y					-	-	-	
Plasmodium	OG5_128193	PF3D7_1035700	Plasmodium falciparum 3D7	duffy binding-like merozoite surface protein (DBLMSP)	697	y		y				-	-	-	
Plasmodium	OG5_128193	PF3D7_1240600	Plasmodium falciparum 3D7	erythrocyte membrane protein 1, PfEMP1 (VAR)	2359		y					-	-	-	
Plasmodium	OG5_129533	PF3D7_1471200	Plasmodium falciparum 3D7	inorganic anion exchanger, inorganic anion antiporter (SulP)	664			y				-	-	-	
Plasmodium	OG5_130526	PF3D7_1335300	Plasmodium falciparum 3D7	reticulocyte binding protein 2 homologue b (RH2b)	3254	y						-	-	-	R
Plasmodium	OG5_130526	PF3D7_1335400	Plasmodium falciparum 3D7	reticulocyte binding protein 2 homologue a (RH2a)	3130	y						-	-	-	R
Plasmodium	OG5_130526	PY01365	Plasmodium yoelii yoelii 17XNL	rhoptry protein	2719	y						-	-	-	
Plasmodium	OG5_131915	PF3D7_1477300	Plasmodium falciparum 3D7	Plasmodium exported protein, unknown function (Pfg14-744)	280		y					-	-	-	
Plasmodium	OG5_131915	PF3D7_1477700	Plasmodium falciparum 3D7	Plasmodium exported protein (PHISTa), unknown function (Pfg14-748)	291		y					-	-	-	
Plasmodium	OG5_132200	PF3D7_1040200	Plasmodium falciparum 3D7	stevor	305		y					-	-	-	
Plasmodium	OG5_132376	PF3D7_0201900	Plasmodium falciparum 3D7	erythrocyte membrane protein 3 (EMP3)	2441		y					-	-	-	R
Plasmodium	OG5_132606	PF3D7_0702300	Plasmodium falciparum 3D7	sporozoite threonine and asparagine-rich protein (STARP)	594			y				-	-	-	R
Plasmodium	OG5_136473	PKH_146980	Plasmodium knowlesi strain H	reticulocyte binding protein, putative (fragment)	683	y						-	-	-	
Plasmodium	OG5_136836	PF3D7_0202000	Plasmodium falciparum 3D7	knob-associated histidine-rich protein (KAHRP)	654		y	y				-	-	-	
Plasmodium	OG5_137681	PF3D7_0424400	Plasmodium falciparum 3D7	surface-associated interspersed protein 4.2 (SURFIN 4.2) (SURF4.2)	2380	y	y	y				-	-	-	
Plasmodium	OG5_138272	PF3D7_0935800	Plasmodium falciparum 3D7	cytoadherence linked asexual protein 9 (CLAG9)	1340	y						-	-	-	
Plasmodium	OG5_138272	PVX_086930	Plasmodium vivax Sal-1	cytoadherence linked asexual protein (CLAG), putative	1343	y						-	-	-	
Plasmodium	OG5_138427	PY04764	Plasmodium yoelii yoelii 17XNL	duffy receptor, beta form precursor	786	y						-	-	-	
Plasmodium	OG5_138835	PCHAS_100150	Plasmodium chabaudi chabaudi	erythrocyte membrane antigen 1	433		y					-	-	-	W
Plasmodium	OG5_141486	gi|7530424	Plasmodium gallinaceum	endochitinase precursor	587	y						-	-	-	W
Plasmodium	OG5_143975	PF3D7_0114100	Plasmodium falciparum 3D7	Pfmc-2TM Maurer's cleft two transmembrane protein (MC-2TM)	229		y					-	-	-	
Plasmodium	OG5_144320	PY06203	Plasmodium yoelii yoelii 17XNL	blood-stage membrane protein Ag-1	549		y					-	-	-	R
Plasmodium	OG5_145193	PF3D7_1035800	Plasmodium falciparum 3D7	probable protein, unknown function (M712)	712	y						-	-	-	R
Plasmodium	OG5_145259	PF3D7_1035200	Plasmodium falciparum 3D7	S-antigen	585	y	y					-	-	-	R
Plasmodium	OG5_148188	PF3D7_0102500	Plasmodium falciparum 3D7	erythrocyte binding antigen-181 (EBA181)	1567	y						-	-	-	
Plasmodium	OG5_148188	PF3D7_0731500	Plasmodium falciparum 3D7	erythrocyte binding antigen-175 (EBA175)	1502	y		y				-	-	-	
Plasmodium	OG5_148188	PF3D7_1301600	Plasmodium falciparum 3D7	erythrocyte binding antigen-140 (EBA140)	1210	y						-	-	-	
Plasmodium	OG5_148188	PKH_062300	Plasmodium knowlesi strain H	erythrocyte binding protein (alpha)	1073	y						-	-	-	
Plasmodium	OG5_148188	PVX_110810	Plasmodium vivax Sal-1	Duffy receptor precursor (DBP)	1070	y						-	-	-	
Plasmodium	OG5_148188	PY03625	Plasmodium yoelii yoelii 17XNL	secreted blood-stage antigen pAg-3	354		y					-	-	-	
Plasmodium	OG5_151950	PF3D7_1335100	Plasmodium falciparum 3D7	merozoite surface protein 7 (MSP7)	351			y				-	-	-	
Plasmodium	OG5_153607	PF3D7_1366400	Plasmodium falciparum 3D7	rhoptry protein (Rhop148)	1262	y						-	-	-	
Plasmodium	OG5_153639	PF3D7_0707300	Plasmodium falciparum 3D7	rhoptry-associated membrane antigen (RAMA)	861	y				y		-	-	-	R
Plasmodium	OG5_154795	PF3D7_0423700	Plasmodium falciparum 3D7	early transcribed membrane protein 4 (ETRAMP4)	136		y					-	-	-	
Plasmodium	OG5_155060	PF3D7_0108300	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	2221		y	y				-	-	-	
Plasmodium	OG5_156053	PBANKA_051500	Plasmodium berghei ANKA	25 kDa ookinete surface antigen precursor (P25)	217			y				-	-	-	
Plasmodium	OG5_156053	PF3D7_1031000	Plasmodium falciparum 3D7	25 kDa ookinete surface antigen precursor (Pfs25)	217			y				-	-	-	
Plasmodium	OG5_156053	PY00522	Plasmodium yoelii yoelii 17XNL	28 kDa ookinete surface protein	217			y				-	-	-	
Plasmodium	OG5_156112	PBANKA_051490	Plasmodium berghei ANKA	28 kDa ookinete surface protein (P28)	213			y				-	-	-	
Plasmodium	OG5_158090	PBANKA_010020	Plasmodium berghei ANKA	Plasmodium exported protein, unknown function	341		y					-	-	-	

Plasmodium	OG5_158440	PF3D7_0920700	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	593	y						-	-	-	
Plasmodium	OG5_158725	PF3D7_1033200	Plasmodium falciparum 3D7	early transcribed membrane protein 10.2 (ETRAMP10.2)	355		y					-	-	-	
Plasmodium	OG5_159311	PF3D7_0501300	Plasmodium falciparum 3D7	skeleton-binding protein 1 (SBP1)	337		y					-	-	-	R
Plasmodium	OG5_159573	PF3D7_0930300	Plasmodium falciparum 3D7	merozoite surface protein 1 (MSP1)	1720	y		y				-	-	-	R
Plasmodium	OG5_159573	PY04986	Plasmodium yoelii yoelii 17XNL	250-270 copies of a 13 AA repeat, NSSTPITSSSIL	2689	y						-	-	-	R
Plasmodium	OG5_159573	PY05748	Plasmodium yoelii yoelii 17XNL	merozoite surface protein 1 precursor	1772			y				-	-	-	
Plasmodium	OG5_159725	PF3D7_0304600	Plasmodium falciparum 3D7	circumsporozoite (CS) protein (CSP)	397			y				-	-	-	R
Plasmodium	OG5_159725	gi|160221	Plasmodium malariae	circumsporozoite protein	429			y				-	-	-	R
Plasmodium	OG5_159725	PBANKA_040320	Plasmodium berghei ANKA	circumsporozoite (CS) protein (CSP)	340	y	y	y				-	-	-	R
Plasmodium	OG5_159781	PF3D7_0402300	Plasmodium falciparum 3D7	reticulocyte binding protein homologue 1,normocyte binding protein 1 (RH1)	2971	y						-	-	-	
Plasmodium	OG5_160444	PF3D7_0206800	Plasmodium falciparum 3D7	merozoite surface protein 2 (MSP2)	272	y						-	-	-	
Plasmodium	OG5_160582	PF3D7_1017100	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	310	y						-	-	-	
Plasmodium	OG5_160600	PF3D7_1212100	Plasmodium falciparum 3D7	peripheral plastid protein 1, putative (PPP1)	558		y					-	-	-	R
Plasmodium	OG5_160733	gi|68521876	Plasmodium gallinaceum	SES	550			y				-	-	-	
Plasmodium	OG5_160733	PBANKA_081300	Plasmodium berghei ANKA	inhibitor of cysteine proteases (ICP)	354	y	y					-	-	-	
Plasmodium	OG5_160791	PF3D7_0722200	Plasmodium falciparum 3D7	rhoptry-associated leucine zipper-like protein 1 (RALP1)	749	y						-	-	-	
Plasmodium	OG5_161018	PKH_146970	Plasmodium knowlesi strain H	reticulocyte binding protein, putative (fragment)	1612	y						-	-	-	W
Plasmodium	OG5_161429	PF3D7_0905400	Plasmodium falciparum 3D7	high molecular weight rhoptry protein 3 (RhopH3)	897	y	y					-	-	-	
Plasmodium	OG5_161676	PBANKA_121810	Plasmodium berghei ANKA	oocyst capsule protein (Cap380)	3204			y				-	-	-	
Plasmodium	OG5_162067	PF3D7_0209000	Plasmodium falciparum 3D7	6-cysteine protein (P230)	3135			y				-	-	-	R
Plasmodium	OG5_162072	PF3D7_0903800	Plasmodium falciparum 3D7	LCCL domain-containing protein (CCp4)	1523			y				-	-	-	
Plasmodium	OG5_162087	PF3D7_1410400	Plasmodium falciparum 3D7	rhoptry-associated protein 1 (RAP1)	782	y						-	-	-	
Plasmodium	OG5_162087	PVX_085930	Plasmodium vivax Sal-1	rhoptry-associated protein 1, putative	640	y						-	-	-	
Plasmodium	OG5_162090	PBANKA_110140	Plasmodium berghei ANKA	rhoptry-associated protein 2/3 (RAP2/3)	388	y						-	-	-	
Plasmodium	OG5_162090	PF3D7_0501600	Plasmodium falciparum 3D7	rhoptry-associated protein 2 (RAP2)	398	y						-	-	-	
Plasmodium	OG5_162112	PF3D7_1346800	Plasmodium falciparum 3D7	6-cysteine protein (P47)	439			y				-	-	-	R
Plasmodium	OG5_162815	PF3D7_1012200	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	267	y						-	-	-	
Plasmodium	OG5_164102	PF3D7_1028700	Plasmodium falciparum 3D7	merozoite TRAP-like protein (MTRAP)	498	y						-	-	-	
Plasmodium	OG5_166171	PF3D7_0620400	Plasmodium falciparum 3D7	merozoite surface protein 10 (MSP10)	525	y		y				-	-	-	
Plasmodium	OG5_166471	PY07092	Plasmodium yoelii yoelii 17XNL	hypothetical protein	175	y						-	-	-	
Plasmodium	OG5_166583	PF3D7_1228600	Plasmodium falciparum 3D7	merozoite surface protein 9 (MSP9)	743			y				-	-	-	
Plasmodium	OG5_166908	PF3D7_0502400	Plasmodium falciparum 3D7	ring-stage membrane protein 1,merozoite surface protein 8 (MSP8)	597			y				-	-	-	
Plasmodium	OG5_166908	PY06415	Plasmodium yoelii yoelii 17XNL	merozoite surface protein precursor	371	y		y				-	-	-	
Plasmodium	OG5_166985	PF3D7_0612700	Plasmodium falciparum 3D7	6-cysteine protein (P12)	347			y				-	-	-	
Plasmodium	OG5_167007	PF3D7_0207000	Plasmodium falciparum 3D7	merozoite surface protein 4 (MSP4)	272			y				-	-	-	
Plasmodium	OG5_167007	PKH_041300	Plasmodium knowlesi strain H	merozoite surface protein, putative	201			y				-	-	-	
Plasmodium	OG5_167007	PVX_003775	Plasmodium vivax Sal-1	merozoite surface protein 4, putative	250			y				-	-	-	R
Plasmodium	OG5_167041	PF3D7_1023000	Plasmodium falciparum 3D7	conserved Plasmodium membrane protein, unknown function	413	y						-	-	-	
Plasmodium	OG5_167067	PF3D7_0423400	Plasmodium falciparum 3D7	asparagine-rich protein (AARP)	217	y						-	-	-	
Plasmodium	OG5_167191	PF3D7_1129100	Plasmodium falciparum 3D7	parasitophorous vacuolar protein 1 (PV1)	452		y					-	-	-	
Plasmodium	OG5_167197	PBANKA_092670	Plasmodium berghei ANKA	circumsporozoite-related antigen (EXP1)	166			y				-	-	-	
Plasmodium	OG5_167197	PF3D7_1121600	Plasmodium falciparum 3D7	circumsporozoite-related antigen,exported protein 1 (EXP1)	162		y					-	-	-	
Plasmodium	OG5_167218	PF3D7_0404900	Plasmodium falciparum 3D7	6-cysteine protein (P41)	378	y		y				-	-	-	
Plasmodium	OG5_167218	PVX_000995	Plasmodium vivax Sal-1	transmission-blocking target antigen Pfs230, putative	384			y				-	-	-	
Plasmodium	OG5_167243	PF3D7_1420700	Plasmodium falciparum 3D7	surface protein, Pf113 (Pf113)	969	y		y				-	-	-	
Plasmodium	OG5_167268	PF3D7_0508000	Plasmodium falciparum 3D7	6-cysteine protein (P38)	349	y						-	-	-	
Plasmodium	OG5_167333	PF3D7_1016000	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	453	y						-	-	-	
Plasmodium	OG5_167430	PF3D7_1345600	Plasmodium falciparum 3D7	inner membrane complex protein	370				y			-	-	-	
Plasmodium	OG5_167433	PF3D7_1346700	Plasmodium falciparum 3D7	6-cysteine protein (P48/45)	448			y				-	-	-	

Plasmodium	OG5_167441	PY03011	Plasmodium yoelii yoelii 17XNL	early transcribed membrane protein family	241	y						-	-	-	
Plasmodium	OG5_167472	PBANKA_143220	Plasmodium berghei ANKA	male development gene 1 (MDV1)	214	y						-	-	-	
Plasmodium	OG5_167472	PF3D7_1216500	Plasmodium falciparum 3D7	male development gene 1 (MDV1)	221		y	y				-	-	-	
Plasmodium	OG5_167477	PBANKA_143360	Plasmodium berghei ANKA	thrombospondin-related apical membrane protein (TRAMP)	348			y				-	-	-	
Plasmodium	OG5_167477	PF3D7_1218000	Plasmodium falciparum 3D7	thrombospondin-related apical membrane protein (TRAMP)	352	y		y				-	-	-	
Plasmodium	OG5_167477	PVX_123575	Plasmodium vivax Sal-1	thrombospondin-related protein 3 precursor, putative	340	y						-	-	-	
Plasmodium	OG5_171217	PBANKA_143230	Plasmodium berghei ANKA	cell traversal protein for ookinetes and sporozoites (CelTOS)	185	y						-	-	-	
Plasmodium	OG5_171475	PY03168	Plasmodium yoelii yoelii 17XNL	circumsporozoite protein	137	y	y	y				-	-	-	
Plasmodium	OG5_171544	PF3D7_1001900	Plasmodium falciparum 3D7	Plasmodium exported protein (hyp16), unknown function (PfJ23)	276		y					-	-	-	
Plasmodium	OG5_172406	PVX_090265	Plasmodium vivax Sal-1	tryptophan-rich antigen (Pv-fam-a)	326		y					-	-	-	
Plasmodium	OG5_173153	PF3D7_1035400	Plasmodium falciparum 3D7	merozoite surface protein 3 (MSP3)	354	y		y				-	-	-	R
Plasmodium	OG5_173153	PF3D7_1035500	Plasmodium falciparum 3D7	merozoite surface protein 6 (MSP6)	371	y		y				-	-	-	
Plasmodium	OG5_173153	PF3D7_1035600	Plasmodium falciparum 3D7	merozoite surface protein (H101)	424	y						-	-	-	
Plasmodium	OG5_173153	PF3D7_1036000	Plasmodium falciparum 3D7	merozoite surface protein (MSP11)	405	y						-	-	-	
Plasmodium	OG5_173153	PF3D7_1036300	Plasmodium falciparum 3D7	merozoite surface protein (DBLMSP2)	762	y						-	-	-	
Plasmodium	OG5_173160	PF3D7_0419700	Plasmodium falciparum 3D7	apical merozoite protein (Pf34)	325	y						-	-	-	
Plasmodium	OG5_173160	PVX_090075	Plasmodium vivax Sal-1	hypothetical protein, conserved	362	y						-	-	-	R
Plasmodium	OG5_173858	PBANKA_103780	Plasmodium berghei ANKA	secreted ookinete adhesive protein (SOAP)	166	y						-	-	-	
Plasmodium	OG5_173941	PF3D7_0830700	Plasmodium falciparum 3D7	Plasmodium exported protein (hyp9), unknown function	218			y				-	-	-	
Plasmodium	OG5_176615	PKH_041310	Plasmodium knowlesi strain H	merozoite surface protein 5	374	y						-	-	-	R
Plasmodium	OG5_176615	PVX_003770	Plasmodium vivax Sal-1	merozoite surface protein 5	387	y						-	-	-	R
Plasmodium	OG5_179285	PF3D7_1016300	Plasmodium falciparum 3D7	glycophorin binding protein (GBP)	824		y	y				-	-	-	W
Plasmodium	OG5_180086	PBANKA_052480	Plasmodium berghei ANKA	early transcribed membrane protein (SEP1)	115			y				-	-	-	
Plasmodium	OG5_184624	PF3D7_1102800	Plasmodium falciparum 3D7	early transcribed membrane protein 11.2 (ETRAMP11.2)	94		y					-	-	-	
Plasmodium	OG5_194730	PCHAS_000190	Plasmodium chabaudi chabaudi	early transcribed membrane protein (ETRAMP)	172	y						-	-	-	
Plasmodium	OG5_205356	PY00204	Plasmodium yoelii yoelii 17XNL	hypothetical protein	224	y						-	-	-	
Plasmodium	OG5_205424	PF3D7_0308300	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	337	y						-	-	-	
Plasmodium	OG5_205493	PF3D7_1370300	Plasmodium falciparum 3D7	membrane associated histidine-rich protein (MAHRP1)	249		y					-	-	-	
Plasmodium	OG5_205510	PF3D7_1364100	Plasmodium falciparum 3D7	cysteine-rich surface protein (Pf92)	796	y		y				-	-	-	
Plasmodium	OG5_205520	PF3D7_1478600	Plasmodium falciparum 3D7	Plasmodium exported protein (hyp17), unknown function	1263		y					-	-	-	
Plasmodium	OG5_205525	PF3D7_0406200	Plasmodium falciparum 3D7	sexual stage-specific protein precursor (Pfs16)	157		y					-	-	-	
Plasmodium	OG5_205526	PF3D7_0423800	Plasmodium falciparum 3D7	cysteine-rich protective antigen (CyRPA)	362	y						-	-	-	
Plasmodium	OG5_205541	PF3D7_0936300	Plasmodium falciparum 3D7	ring-exported protein 3 (REX3)	326		y					-	-	-	R
Plasmodium	OG5_210391	PF3D7_0831800	Plasmodium falciparum 3D7	histidine-rich protein II (HRPII)	305		y					-	-	-	R
Plasmodium	OG5_243443	PF3D7_0113900	Plasmodium falciparum 3D7	Plasmodium exported protein (hyp8), unknown function (GEXP10)	243		y					-	-	-	
Plasmodium	OG5_243444	PF3D7_0731100	Plasmodium falciparum 3D7	Plasmodium exported protein (PHISTc), unknown function (GEXP11)	890		y					-	-	-	
Plasmodium	OG5_243452	PF3D7_1401400	Plasmodium falciparum 3D7	early transcribed membrane protein 14.1 (ETRAMP14)	107		y					-	-	-	
Plasmodium	OG5_243471	PF3D7_0936000	Plasmodium falciparum 3D7	(REX2) ring-exported protein 2	94		y					-	-	-	
Plasmodium	OG5_243472	PF3D7_0936400	Plasmodium falciparum 3D7	ring-exported protein 4 (REX4)	214		y					-	-	-	
Theileria	OG5_147143	TP04_0437	Theileria parva strain Muguga	104 kDa antigen	924	y						-	-	-	
Theileria	OG5_149405	gi|63079801	Theileria lestoquardi	unknown protein long splice variant	884			y				-	-	-	R
Theileria	OG5_185087	TA20082	Theileria annulata strain Ankara	Tashat3 protein	994		y					-	-	-	R
Theileria	OG5_185087	TA20085	Theileria annulata strain Ankara	Tashat1 protein	465		y					-	-	-	R
Theileria	OG5_185087	TA20095	Theileria annulata strain Ankara	Tashat2 protein	1163		y					-	-	-	R
Theileria	OG5_247131	TA06510	Theileria annulata strain Ankara	hypothetical protein	311			y				-	-	-	
Theileria	OG5_247131	TP01_0939	Theileria parva strain Muguga	hypothetical protein	307			y				-	-	-	
Theileria	OG5_247231	gi|3142237	Babesia equi	merozoite antigen 2	274		y					-	-	-	
Toxoplasma + Neospora	NO_GROUP	TGME49_233450	Toxoplasma gondii ME49	SAG-related sequence SRS29A (SRS29A)	420		y	y				-	-	-	
Toxoplasma + Neospora	NO_GROUP	TGME49_286450	Toxoplasma gondii ME49	dense granule protein GRA5 (GRA5)	120	y					y	-	-	-	
Toxoplasma + Neospora	OG5_126585	TGME49_252880	Toxoplasma gondii ME49	hypothetical protein	3168	y						-	-	-	R

Toxoplasma + Neospora	OG5_137815	TGME49_233030	Toxoplasma gondii ME49	gliding-associated protein GAP70 (GAP70)	353				y			-	-	-	R
Toxoplasma + Neospora	OG5_173666	TGME49_207130	Toxoplasma gondii ME49	SAG-related sequence SRS49A (SRS49A)	367			y				-	-	-	
Toxoplasma + Neospora	OG5_173666	TGME49_207160	Toxoplasma gondii ME49	SAG-related sequence SRS49D (SRS49D)	364		y					-	-	-	
Toxoplasma + Neospora	OG5_234994	TGME49_254720	Toxoplasma gondii ME49	dense granule protein GRA8 (GRA8)	269	y						-	-	-	
Toxoplasma + Neospora	OG5_239565	TGME49_211290	Toxoplasma gondii ME49	rhoptry protein ROP15 (ROP15)	333	y						-	-	-	
Toxoplasma + Neospora	OG5_239707	TGME49_239740	Toxoplasma gondii ME49	dense granule protein GRA14 (GRA14)	408	y						-	-	-	
Toxoplasma + Neospora	OG5_239773	TGME49_244030	Toxoplasma gondii ME49	inner membrane complex protein IMC3 (IMC3)	696				y			-	-	-	
Toxoplasma + Neospora	OG5_239843	TGME49_204530	Toxoplasma gondii ME49	microneme protein MIC11 (MIC11)	204	y						-	-	-	
Toxoplasma + Neospora	OG5_239857	TGME49_203990	Toxoplasma gondii ME49	rhoptry protein ROP12 (ROP12)	236	y						-	-	-	
Toxoplasma + Neospora	OG5_239876	NCLIV_021640	Neospora caninum Liverpool	unspecified product	217	y		y				-	-	-	
Toxoplasma + Neospora	OG5_239876	TGME49_203310	Toxoplasma gondii ME49	dense granule protein GRA7 (GRA7)	236	y					y	-	-	-	
Toxoplasma + Neospora	OG5_240082	gi|160822	Sarcocystis muris	31-kDa major surface antigen	280	y						-	-	-	
Toxoplasma + Neospora	OG5_240258	TGME49_233460	Toxoplasma gondii ME49	SAG-related sequence SRS29B (SRS29B)	336			y				-	-	-	
Toxoplasma + Neospora	OG5_240260	TGME49_233480	Toxoplasma gondii ME49	SAG-related sequence SRS29C (SRS29C)	372		y	y				-	-	-	
Toxoplasma + Neospora	OG5_240341	TGME49_271050	Toxoplasma gondii ME49	SAG-related sequence SRS34A (SRS34A)	187			y				-	-	-	
Toxoplasma + Neospora	OG5_240366	NCLIV_036400	Neospora caninum Liverpool	unspecified product	190	y		y				-	-	-	
Toxoplasma + Neospora	OG5_240366	TGME49_270250	Toxoplasma gondii ME49	dense granule protein GRA1 (GRA1)	190	y					y	-	-	-	
Toxoplasma + Neospora	OG5_240367	NCLIV_036410	Neospora caninum Liverpool	putative cyst matrix protein	456	y		y				-	-	-	
Toxoplasma + Neospora	OG5_240367	TGME49_270240	Toxoplasma gondii ME49	unspecified product (MAG1)	452						y				
Toxoplasma + Neospora	OG5_240666	NCLIV_045650	Neospora caninum Liverpool	unspecified product	211	y		y				-	-	-	
Toxoplasma + Neospora	OG5_240666	TGME49_227620	Toxoplasma gondii ME49	dense granule protein GRA2 (GRA2)	185	y					y	-	-	-	
Toxoplasma + Neospora	OG5_240860	TGME49_214080	Toxoplasma gondii ME49	toxofilin	243	y	y	y				-	-	-	
Toxoplasma + Neospora	OG5_240880	TGME49_214940	Toxoplasma gondii ME49	MIC2-associated protein M2AP	330	y		y	y	y		-	-	-	
Toxoplasma + Neospora	OG5_240905	NCLIV_052880	Neospora caninum Liverpool	unspecified product	193	y						-	-	-	
Toxoplasma + Neospora	OG5_240905	TGME49_275440	Toxoplasma gondii ME49	dense granule protein GRA6 (GRA6)	224	y		y			y	-	-	-	
Toxoplasma + Neospora	OG5_240919	TGME49_308840	Toxoplasma gondii ME49	SAG-related sequence SRS51 (SRS51)	472		y	y				-	-	-	
Toxoplasma + Neospora	OG5_240965	TGME49_310780	Toxoplasma gondii ME49	dense granule protein GRA4 (GRA4)	345	y					y	-	-	-	
Toxoplasma + Neospora	OG5_240990	TGME49_311670	Toxoplasma gondii ME49	transmembrane surface antigen TSA1 (TSA1)	440			y				-	-	-	
Toxoplasma + Neospora	OG5_241009	TGME49_312270	Toxoplasma gondii ME49	rhoptry protein ROP13 (ROP13)	400	y						-	-	-	
Toxoplasma + Neospora	OG5_241095	TGME49_315490	Toxoplasma gondii ME49	rhoptry protein ROP10 (ROP10)	586	y						-	-	-	
Toxoplasma + Neospora	OG5_241151	TGME49_215980	Toxoplasma gondii ME49	hypothetical protein	210	y						-	-	-	
Toxoplasma + Neospora	OG5_241166	TGME49_308020	Toxoplasma gondii ME49	SAG-related sequence SRS57 (SRS57)	385		y					-	-	-	
Toxoplasma + Neospora	OG5_241346	gi|22347779	Sarcocystis neurona	microneme antigen 10	255	y						-	-	-	
Toxoplasma + Neospora	OG5_241346	TGME49_250710	Toxoplasma gondii ME49	microneme protein MIC10 (MIC10)	198	y						-	-	-	
Toxoplasma + Neospora	OG5_241357	TGME49_251540	Toxoplasma gondii ME49	dense granule protein GRA9 (GRA9)	318	y						-	-	-	
Toxoplasma + Neospora	OG5_241411	TGME49_277080	Toxoplasma gondii ME49	microneme protein MIC5 (MIC5)	181	y						-	-	-	
Toxoplasma	NO_GROUP	TGME49_204340	Toxoplasma gondii ME49	hypothetical protein	877	y						-	-	-	
Toxoplasma	NO_GROUP	TGME49_314250	Toxoplasma gondii ME49	bradyzoite rhoptry protein BRP1 (BRP1)	162	y						-	-	-	
Toxoplasma	OG5_127711	TGME49_207690	Toxoplasma gondii ME49	programmed cell death 5 protein (PDCD5)	249	y		y				-	-	-	
Toxoplasma	OG5_240675	TGME49_227280	Toxoplasma gondii ME49	dense granule protein GRA3 (GRA3)	220	y					y	-	-	-	

[1] 'Eukaryotic' origin denotes that one or more regions in the protein has a significant hit to eukaryotes outside myzozoans

[2] Localization based on Apiloc categories: Apical (anterior part of the cells, conoid, rhoptries, micronemes, etc), Exported, Mebrane (of the parasite), IMC (inner membrane complex), Golgi (control; protein of eukaryotic origin)

[3] Similarity at evalue 1e-5 to chrompodellids (Chrompod.), dinoflagellates or Perkinsus marinus (Dino. + Perk.), other eukaryotes (Eukaryot.) in our database (Materials and Methods)

[4] Note: R = contains repeat that gives a non-specific hit to eukaryotes other than indicated; W = contains a single weak hit to eukaryotes other than indicated; X = classified based on a second round of more in-depth searches (Materials and Methods).

Table S5: Classification of 21 bacterial genes in apicomplexans and relatives acquired by horizontal gene transfer (HGT).

Origin[1]	EuPathDb accession	Short name	Protein name	EC	Pathway	Original reference for HGT
API+CHP	TGME49_261850	PREX	DNA polymerase complex (plastid)	2.7.7.7	DNA replication	Seow et al., 2005 (10)
API+CHP	TGME49_227420	IspH, (LytB)	4-hydroxy-3-methylbut-2-enyl diphosphate reductase (plastid)	1.17.1.2	Isoprenoid biosynthesis	Identified in this study
API+CHP	TGME49_258650	FeCH (HemH)	ferrochelatase (mitochondrial)	4.99.1.1	Tetrapyrrole biosynthesis	Sato & Wilson, 2003 (11)
API+CHP	TGME49_217740	FabG	3-oxoacyl-ACP reductase (plastid)	1.1.1.100	Fatty acid biosynthesis	Identified in this study
API+CHP	cgd7_480	MDH/LDH	malate/lactate dehydrogenase	1.1.1.37	Carbohydrate metabolism	Madern et al., 2003 (12)
API+CHP	cgd6_4570	GlnA	glutamine synthetase	6.3.1.2	Amino acid and nitrogen metabolism	Huang et al., 2004a (5)
API+CHP [2]	TGME49_208090	YgfA	5-formyltetrahydrofolate cyclo-ligase	6.3.3.2	Folate metabolism	Identified in this study
MYZ	TGME49_307030	PNP	purine-nucleoside phosphorylase	2.4.2.1	Purine metabolism	Kicska et al., 2003 (13)
MYZ	TGME49_244430	PUS	pseudouridylate synthase	4.2.1.70	Pyrimidine metabolism	Identified in this study
MYZ	TGME49_233140	DUT	desoxyuridine 5-triophosphate nucleotidohydrilase	3.6.1.23	Pyrimidine metabolism	Identified in this study
MYZ	cgd4_740	Thio	thioredoxin peroxidase	1.11.1.15	Redox reactions	Huang et al., 2004a (5)
MYZ	TGME49_247510	Fbp	fructose-1,6-bisphosphatase	3.1.3.11	Carbohydrate metabolism	Huang et al., 2004b (14)
MYZ	cgd6_3280	GlgB	1,4-alpha-glucan branching enzyme	2.4.1.18	Carbohydrate metabolism	Huang et al., 2004a (5)
EUK	cgd2_570	TrmH	tRNA/rRNA methyltransferase	2.1.1.34	RNA modification	Huang et al., 2004a (5)
EUK	cgd6_3750	AMY	alpha-amylase	3.2.1.1	Carbohydrate metabolism	Huang et al., 2004a (5)
EUK	cgd1_730	QPCT	glutamine cyclotransferase	2.3.2.5	Peptide modification	Huang et al., 2004a (5)
EUK	cgd8_3430	ANPEP	aminopeptidase N	3.4.11.2	Peptide modification	Huang et al., 2004a (5)
EUK	cgd2_2020		hypothetical protein		Unknown	Huang et al., 2004a (5)
EUK	cgd5_3230	SOD	superoxide dismutase	1.15.1.1	Superoxide conversion	Huang et al., 2004a (5)
EUK	cgd8_350	NARS	asparaginyl-tRNA synthetase	6.3.5.6	Aminoacyl-tRNA synthesis	Huang et al., 2004a (5)
EUK [3]	cgd6_2130	RsmE	16S RNA methyltransferase	2.1.1.193	RNA modification	Huang et al., 2004a (11)

[1] Acquired in the ancestor of apicomplexans and chrompodellids (API+CHP), myzozoans (MYZ) or shares a common origin with one or more other eukaryotes (EUK).

[2] YgfA is absent in dinoflagellates and most other eukaryotes.

[3] The relationship of apicomplexan RsmE to that of bacteria and other eukaryotes is unresolved in phylogenies (however a related homolog is found in chrompodellids suggesting that neither this gene represents an apicomplexan-specific HGT).

Table S6: Classification of 74 unique proteins specific to apicomplexans, apicomplexans and chrompodellids, and myzozoans.

Origin	Localization	Protein EuPathDb accession	Organism	Protein annotation	Length	OrthoMCL Group	Note	Putative function (reference)
Apicomplexan-specific	Apical	PF3D7_1318000	Plasmodium falciparum 3D7	conserved Plasmodium membrane protein, unknown function	192	OG5_145102		Apical protein linked to merozoite invasion (15)
Apicomplexan-specific	Apical	PF3D7_1030200	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	450	OG5_143206	R	Apical protein linked to merozoite invasion (15)
Apicomplexan-specific	Apical	PF3D7_0210600	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	446	OG5_141751		Apical protein linked to merozoite invasion (15)
Apicomplexan-specific	Apical	PF3D7_0214900	Plasmodium falciparum 3D7	rhoptry neck protein 6 (RON6)	950	OG5_150575	R	Secreted from rhoptries during merozoite invasion (16)
Apicomplexan-specific	Apical	TGME49_310010	Toxoplasma gondii ME49	rhoptry neck protein RON1 (RON1)	1158	OG5_222773	R	GPI anchor-containing rhoptry membrane protein (17)
Apicomplexan-specific	Apical	PBANKA_141830	Plasmodium berghei ANKA	rhoptry protein 2, putative (PRP2)	1272	OG5_166877	R	Essential rhoptry protein possibly involved in rhoptry assembly (18)
Apicomplexan-specific	Exported	PF3D7_1463900	Plasmodium falciparum 3D7	conserved Plasmodium membrane protein, unknown function	1071	OG5_141738		Exported to erythrocyte from merozoite (19)
Apicomplexan-specific		TGME49_261750	Toxoplasma gondii ME49	rhoptry neck protein RON10 (RON10)	835	OG5_171370		Possibly involved in sorting proteins to rhoptries (20)
Apicomplexan-specific		cgd7_4030	Cryptosporidium parvum Iowa II	hypothetical low complexity protein	586	OG5_160601		Unknown
Apicomplexan-specific		TGME49_200330	Toxoplasma gondii ME49	hypothetical protein	589	OG5_142982		Unknown

Apicomplexan-specific		TGME49_202120	Toxoplasma gondii ME49	hypothetical protein	1322	OG5_142642		Unknown
Apicomplexan-specific		TGME49_222350	Toxoplasma gondii ME49	hypothetical protein	410	OG5_143198		Unknown
Apicomplexan-specific		TGME49_222940	Toxoplasma gondii ME49	hypothetical protein	489	OG5_171982		Unknown
Apicomplexan-specific		TGME49_235670	Toxoplasma gondii ME49	hypothetical protein	558	OG5_156810		Unknown
Apicomplexan-specific		TGME49_244080	Toxoplasma gondii ME49	hypothetical protein	376	OG5_171957		Unknown
Apicomplexan-specific		TGME49_274160	Toxoplasma gondii ME49	hypothetical protein	525	OG5_145098		Unknown
Apicomplexan-specific		TGME49_282020	Toxoplasma gondii ME49	hypothetical protein	394	OG5_180667		Unknown
Apicomplexan-specific		TGME49_318390	Toxoplasma gondii ME49	hypothetical protein	1278	OG5_171978		Unknown
Apicomplexan-specific		TGME49_202290	Toxoplasma gondii ME49	membrane protein, putative	349	OG5_141728		Unknown
Apicomplexan-specific	IMC	TGME49_202500	Toxoplasma gondii ME49	GAPM1a	305	OG5_140542		Six-pass transmembrane IMC protein with IMC-anchoring role in gliding (21)
Apicomplexan-specific	IMC	TGME49_206690	Toxoplasma gondii ME49	glideosome-associated protein with multiple-membrane spans GAPM2B (GAPM2B)	426	OG5_141755		Six-pass transmembrane IMC protein with IMC-anchoring role in gliding (21)
Apicomplexan-specific	IMC	TGME49_271970	Toxoplasma gondii ME49	glideosome-associated protein with multiple-membrane spans GAPM3	310	OG5_143227		Six-pass transmembrane IMC protein with IMC-anchoring role in gliding (21)
Apicomplexan-specific	Oocyst wall	cgd6_200	Cryptosporidium parvum Iowa II	oocyst wall protein 8	457	OG5_225269		Oocyst wall-specific protein (9)
Apicomplexan-specific	Oocyst wall	TGME49_209610	Toxoplasma gondii ME49	oocyst wall protein OWP2 (OWP2)	462	OG5_239307		Oocyst wall-specific protein (22)
Apicomplexan-specific	Apical	PF3D7_0414600	Plasmodium falciparum 3D7	conserved protein, unknown function	235	OG5_145116		Apical protein linked to merozoite invasion (15)
Api+Chromp-specific		TGME49_200460	Toxoplasma gondii ME49	hypothetical protein	436	OG5_171961		Unknown
Api+Chromp-specific		TGME49_202040	Toxoplasma gondii ME49	hypothetical protein	1047	OG5_143225		Unknown
Api+Chromp-specific		TGME49_205680	Toxoplasma gondii ME49	hypothetical protein	805	OG5_156805	R	Unknown
Api+Chromp-specific		TGME49_209490	Toxoplasma gondii ME49	hypothetical protein	505	OG5_171967		Unknown
Api+Chromp-specific		TGME49_209600	Toxoplasma gondii ME49	hypothetical protein	451	OG5_171975		Unknown
Api+Chromp-specific		TGME49_217480	Toxoplasma gondii ME49	hypothetical protein	1589	OG5_169915		Unknown
Api+Chromp-specific		TGME49_219160	Toxoplasma gondii ME49	hypothetical protein	349	OG5_145100		Unknown
Api+Chromp-specific		TGME49_240470	Toxoplasma gondii ME49	hypothetical protein	409	OG5_171976		Unknown
Api+Chromp-specific		TGME49_245520	Toxoplasma gondii ME49	hypothetical protein	235	OG5_169642		Unknown
Api+Chromp-specific		TGME49_255410	Toxoplasma gondii ME49	hypothetical protein	165	OG5_143229		Unknown
Api+Chromp-specific		TGME49_258920	Toxoplasma gondii ME49	hypothetical protein	481	OG5_150770		Unknown
Api+Chromp-specific		TGME49_262890	Toxoplasma gondii ME49	hypothetical protein	555	OG5_144475		Unknown
Api+Chromp-specific		TGME49_263440	Toxoplasma gondii ME49	hypothetical protein	600	OG5_163433		Unknown
Api+Chromp-specific		TGME49_264020	Toxoplasma gondii ME49	hypothetical protein	647	OG5_171222		Unknown
Api+Chromp-specific		TGME49_266630	Toxoplasma gondii ME49	hypothetical protein	582	OG5_141750		Unknown
Api+Chromp-specific		TGME49_268690	Toxoplasma gondii ME49	hypothetical protein	577	OG5_160615		Unknown
Api+Chromp-specific		TGME49_273040	Toxoplasma gondii ME49	hypothetical protein	311	OG5_143217		Unknown
Api+Chromp-specific		TGME49_273705	Toxoplasma gondii ME49	hypothetical protein	424	OG5_145099		Unknown
Api+Chromp-specific		TGME49_311880	Toxoplasma gondii ME49	hypothetical protein	728	OG5_160577		Unknown
Api+Chromp-specific		TGME49_313380	Toxoplasma gondii ME49	hypothetical protein	274	OG5_141745		Unknown
Api+Chromp-specific		TGME49_309850	Toxoplasma gondii ME49	insulin-degrading enzyme	456	OG5_147636		Unknown
Api+Chromp-specific		cgd8_5160	Cryptosporidium parvum Iowa II	uncharacterized apicomplexan-specific protein	170	OG5_143199		Unknown
Api+Chromp-specific		cgd7_1550	Cryptosporidium parvum Iowa II	uncharacterized protein	840	OG5_195587		Unknown
Myzozoan specific	Exported, Membrane	TA20205	Theileria annulata strain Ankara	schizont protein e	441	OG5_147397	W	Exported protein associated with host cell microtubules (23)
Myzozoan specific	Exported, Membrane	TA19865	Theileria annulata strain Ankara	surface protein d precursor	178	OG5_191678		Exported membrabe protein (24)
Myzozoan specific	Apical	PF3D7_1321900	Plasmodium falciparum 3D7	conserved Plasmodium protein, unknown function	292	OG5_165465	W	Apical protein linked to merozoite invasion (15)
Myzozoan specific	IMC	TGME49_237820	Toxoplasma gondii ME49	IMC sub-compartment protein ISP2 (ISP2)	160	OG5_141732		IMC protein involved in cell division (25)
Myzozoan specific	Apical	PBANKA_090710	Plasmodium berghei ANKA	inner membrane complex protein 1b, putative (IMC1b)	504	OG5_153786		Membrane skeleton protein involved in cell shape maintenance, gliding motility, and infection (26)
Myzozoan specific	IMC	PF3D7_1003600	Plasmodium falciparum 3D7	IMC-related (ALV5)	281	OG5_144710		Alveolin domain-containing protein linked to merozoite invasion (15)
Myzozoan specific		cgd8_5030	Cryptosporidium parvum Iowa II	articulin family protein, Pfs77 protein-related	324	OG5_147639		Alveolin domain-containing protein
Myzozoan specific		TGME49_260500	Toxoplasma gondii ME49	COPI associated protein, putative	270	OG5_143228		Unknown
Myzozoan specific		cgd8_2990	Cryptosporidium parvum Iowa II	hypothetical protein	180	OG5_151232		Unknown
Myzozoan specific		TGME49_203358	Toxoplasma gondii ME49	hypothetical protein	395	OG5_143204		Unknown
Myzozoan specific		TGME49_205320	Toxoplasma gondii ME49	hypothetical protein	655	OG5_141752		Unknown
Myzozoan specific		TGME49_207240	Toxoplasma gondii ME49	hypothetical protein	108	OG5_171986		Unknown
Myzozoan specific		TGME49_209890	Toxoplasma gondii ME49	hypothetical protein	193	OG5_138808		Unknown
Myzozoan specific		TGME49_215430	Toxoplasma gondii ME49	hypothetical protein	244	OG5_143223		Unknown
Myzozoan specific		TGME49_216410	Toxoplasma gondii ME49	hypothetical protein	188	OG5_143208		Unknown
Myzozoan specific		TGME49_225840	Toxoplasma gondii ME49	hypothetical protein	686	OG5_156809		Unknown
Myzozoan specific		TGME49_244360	Toxoplasma gondii ME49	hypothetical protein	383	OG5_145121		Unknown

Myzozoan specific	TGME49_258820	Toxoplasma gondii ME49	hypothetical protein	129	OG5_141758	Unknown
Myzozoan specific	TGME49_265510	Toxoplasma gondii ME49	hypothetical protein	272	OG5_146989	Unknown
Myzozoan specific	TGME49_267510	Toxoplasma gondii ME49	hypothetical protein	346	OG5_150760	Unknown
Myzozoan specific	TGME49_269310	Toxoplasma gondii ME49	hypothetical protein	968	OG5_141231	Unknown
Myzozoan specific	TGME49_293720	Toxoplasma gondii ME49	hypothetical protein	336	OG5_147633	Unknown
Myzozoan specific	TGME49_297870	Toxoplasma gondii ME49	hypothetical protein	872	OG5_171974	Unknown
Myzozoan specific	TGME49_306980	Toxoplasma gondii ME49	hypothetical protein	226	OG5_147631	Unknown
Myzozoan specific	TGME49_310210	Toxoplasma gondii ME49	hypothetical protein	788	OG5_150590	Unknown
Myzozoan specific	TGME49_310270	Toxoplasma gondii ME49	hypothetical protein	1357	OG5_141163	Unknown

Table S7: Domain superfamilies in chrompodellid and dinoflagellate PKSI/FASI proteins identified by NCBI Conserved Domains searches.

Organism	Contig translation or Protein name	Start	End	E-value	Bit Score	Super-family	Superfamily description	Domain[1]	Note[2]
Colpodella angusta Spi-2	a2699_83_57_7353-79	1	44	1.97E-015	74.98	cl09936	PP-binding	ACP	N
Colpodella angusta Spi-2	a2699_83_57_7353-79	77	474	2.07E-160	502.86	cl09938	cond_enzymes	KS	
Colpodella angusta Spi-2	a2699_83_57_7353-79	527	826	3.96E-107	346.7	cl08282	Acyl_transf_1	AT	
Colpodella angusta Spi-2	a2699_83_57_7353-79	1235	1456	8.06E-048	178.25	cl09931	NADB_Rossmann	KR	N
Colpodella angusta Spi-2	a2699_83_57_7353-79	1504	1564	1.51E-012	66.5	cl09936	PP-binding	ACP	N
Colpodella angusta Spi-2	a2699_83_57_7353-79	1629	1926	6.30E-069	236.78	cl09931	NADB_Rossmann	SDR	
Colpodella angusta Spi-2	a6216_94_7_1-3726	120	209	1.99E-005	46.58	cl00509	hot_dog	DH	N
Colpodella angusta Spi-2	a6216_94_7_1-3726	610	903	3.61E-083	274.06	cl16912	MDR	ER	
Colpodella angusta Spi-2	a6216_94_7_1-3726	950	1121	6.87E-035	133.03	cl09931	NADB_Rossmann	KR	
Colpodella angusta Spi-2	a6216_94_7_1-3726	1184	1239	3.77E-007	49.11	cl09936	PP-binding	ACP	
Colpodella angusta BE-6	a1114_130_47_6309-79	1	126	9.03E-044	166.97	cl09938	cond_enzymes	KS	N
Colpodella angusta BE-6	a1114_130_47_6309-79	179	478	7.33E-108	348.24	cl08282	Acyl_transf_1	AT	
Colpodella angusta BE-6	a1114_130_47_6309-79	887	1108	3.48E-048	179.02	cl09931	NADB_Rossmann	KR	N
Colpodella angusta BE-6	a1114_130_47_6309-79	1156	1216	1.16E-012	66.5	cl09936	PP-binding	ACP	N
Colpodella angusta BE-6	a1114_130_47_6309-79	1281	1578	3.73E-069	237.16	cl09931	NADB_Rossmann	SDR	
Colpodella angusta BE-6	a1273513_2_3_336-1	1	44	2.66E-008	46.86	cl09936	PP-binding	ACP	N
Colpodella angusta BE-6	a1273513_2_3_336-1	64	111	3.31E-014	66.81	cl09938	cond_enzymes	KS	C
Colpodella angusta BE-6	a172836_7_22_2511-1	1	209	1.18E-052	185.08	cl16912	MDR	ER	N
Colpodella angusta BE-6	a172836_7_22_2511-1	222	453	5.26E-067	228.71	cl09931	NADB_Rossmann	KR	N
Colpodella angusta BE-6	a172836_7_22_2511-1	484	562	3.24E-022	92.7	cl09936	PP-binding	ACP	
Colpodella angusta BE-6	a172836_7_22_2511-1	595	833	6.71E-095	306.02	cl09938	cond_enzymes	KS	C
Colpodella angusta BE-6	a490947_4_4_1-738	1	124	8.98E-044	153.87	cl09938	cond_enzymes	KS	N
Colpodella angusta BE-6	a490947_4_4_1-738	153	224	2.27E-014	69.3	cl00509	hot_dog	DH	C
Colpodella angusta BE-6	a511308_3_4_2-496	1	79	1.15E-032	120.36	cl09938	cond_enzymes	KS	N
Colpodella angusta BE-6	a511308_3_4_2-496	112	165	5.45E-012	60.88	cl08282	Acyl_transf_1	AT	C
Colpodella angusta BE-6	a950_137_16_23-7138	53	497	1.80E-154	486.3	cl09938	cond_enzymes	KS	
Colpodella angusta BE-6	a950_137_16_23-7138	701	1006	5.75E-045	167.58	cl08282	Acyl_transf_1	AT	
Colpodella angusta BE-6	a950_137_16_23-7138	1098	1383	4.51E-012	67.76	cl00509	hot_dog	DH	
Colpodella angusta BE-6	a950_137_16_23-7138	1740	2033	2.01E-082	275.6	cl16912	MDR	ER	
Colpodella angusta BE-6	a950_137_16_23-7138	2080	2251	2.05E-034	133.03	cl09931	NADB_Rossmann	KR	
Colpodella angusta BE-6	a950_137_16_23-7138	2314	2369	3.82E-007	49.88	cl09936	PP-binding	ACP	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_11567_1_305	1	64	2.50E-009	56.04	cl16912	MDR	ER	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_11567_1_305	82	305	7.82E-078	244.12	cl09931	NADB_Rossmann	KR	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_3923_1_1196	2	101	3.17E-028	118.16	cl09931	NADB_Rossmann	KR	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_3923_1_1196	134	212	9.58E-019	83.45	cl09936	PP-binding	ACP	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_3923_1_1196	370	669	2.32E-074	249.11	cl09931	NADB_Rossmann	SDR	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_6061_1_544	43	442	0.00E+000	534.45	cl09938	cond_enzymes	KS	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_6061_1_544	514	544	5.07E-011	62.42	cl08282	Acyl_transf_1	AT	C
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_65964_1_2899	25	442	2.42E-171	536.37	cl09938	cond_enzymes	KS	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_65964_1_2899	502	797	1.96E-109	354.02	cl08282	Acyl_transf_1	AT	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_65964_1_2899	847	1156	5.94E-050	182.17	cl00509	hot_dog	DH	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_65964_1_2899	1402	1512	3.90E-011	66.54	cl09931	NADB_Rossmann	KR	C
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_65964_1_2899	1635	1949	6.75E-081	271.75	cl16912	MDR	ER	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_65964_1_2899	1977	2211	1.28E-061	219.08	cl09931	NADB_Rossmann	KR	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_65964_1_2899	2252	2335	1.47E-019	87.31	cl09936	PP-binding	ACP	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_65964_1_2899	2496	2796	2.12E-079	267.6	cl09931	NADB_Rossmann	SDR	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66433_1_1647	52	1010	7.86E-129	415.06	cl17068	AFD_class_I	FAAL	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66433_1_1647	1105	1176	2.48E-009	56.49	cl09936	PP-binding	ACP	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66433_1_1647	1248	1547	5.09E-047	172.45	cl09931	NADB_Rossmann	KR	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66484_1_586	1	115	2.49E-043	158.99	cl09931	NADB_Rossmann	KR	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66484_1_586	218	273	1.07E-015	73.05	cl09936	PP-binding	ACP	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66484_1_586	473	585	5.66E-035	133.16	cl09931	NADB_Rossmann	SDR	C
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66893_1_2173	2	53	5.26E-006	46.86	cl09936	PP-binding	ACP	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66893_1_2173	82	604	1.95E-150	473.58	cl09938	cond_enzymes	KS	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66893_1_2173	636	961	1.69E-105	341.69	cl08282	Acyl_transf_1	AT	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66893_1_2173	1242	1740	2.81E-069	240.65	cl09931	NADB_Rossmann	KR	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66893_1_2173	1791	1874	3.16E-013	68.43	cl09936	PP-binding	ACP	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_66893_1_2173	1910	2169	3.56E-025	107.91	cl19140	Hydrolase_4		
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_67019_1_1077	3	291	8.22E-030	120.92	cl00509	hot_dog	DH	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_67019_1_1077	402	563	4.11E-008	55.37	cl09931	NADB_Rossmann	KR	C
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_67019_1_1077	599	896	2.89E-102	325.29	cl16912	MDR	ER	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_67019_1_1077	930	1076	1.93E-036	137.27	cl09931	NADB_Rossmann	KR	C
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_67680_1_938	605	653	4.63E-008	54.87	cl09938	cond_enzymes	KS	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_67680_1_938	681	820	3.27E-040	151.02	cl08282	Acyl_transf_1	AT	C

Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_68652_1_1566	122	182	6.80E-008	54.5	cl16912	MDR	ER	C
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_68652_1_1566	342	591	9.27E-063	220.62	cl09931	NADB_Rossmann	KR	N
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_68652_1_1566	696	778	7.57E-019	84.22	cl09936	PP-binding	ACP	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_68652_1_1566	837	1161	2.14E-078	262.59	cl09931	NADB_Rossmann	SDR	
Vitrella brassicaformis CCMP3155	MMETSP0288-20130426_68652_1_1566	1310	1493	3.87E-008	54.75	cl19140	Hydrolase_4		
Chromera velia CCMP2878	MMETSP0290-20130426_10064_1_839	2	195	3.77E-041	155.23	cl09931	NADB_Rossmann	KR	N
Chromera velia CCMP2878	MMETSP0290-20130426_10064_1_839	261	345	1.75E-020	87.69	cl09936	PP-binding	ACP	
Voromonas pontica GG	comp4108_c1_seq1_4	3	157	4.83E-026	101.79	cl09931	NADB_Rossmann	KR	N
Voromonas pontica GG	comp4108_c1_seq1_4	234	295	2.58E-009	52.57	cl09936	PP-binding	ACP	
Voromonas pontica GG	comp7328_c0_seq1_8	1	51	2.34E-010	56.91	cl09931	NADB_Rossmann	KR	N
Voromonas pontica GG	comp7328_c0_seq1_8	93	166	1.95E-022	86.92	cl09936	PP-binding	ACP	
Voromonas pontica GG	comp7328_c2_seq1_7	1	120	1.98E-026	103.03	cl16912	MDR	ER	N
Voromonas pontica GG	comp7328_c2_seq1_7	135	209	6.99E-010	54.45	cl09931	NADB_Rossmann	KR	C
Lingulodinium polyedrum	346235801_9_1-2160	35	175	1.58E-016	80.79	cl09931	NADB_Rossmann	KR	C
Lingulodinium polyedrum	346235801_9_1-2160	207	501	5.60E-099	309.11	cl16912	MDR	ER	-
Lingulodinium polyedrum	346235801_9_1-2160	510	720	2.56E-053	189.03	cl09931	NADB_Rossmann	KR	N
Lingulodinium polyedrum	346254191_19_1415-3	1	254	1.83E-096	299.47	cl09938	cond_enzymes	KS	N
Lingulodinium polyedrum	346254191_19_1415-3	273	471	1.67E-084	264.65	cl08282	Acyl_transf_1	AT	C
Lingulodinium polyedrum	346254191_19_1415-3	349	785	7.38E-111	356.1	cl09938	cond_enzymes	KS	-
Lingulodinium polyedrum	346305629_13_1-2883	65	496	1.36E-123	384.6	cl09938	cond_enzymes	KS	-
Lingulodinium polyedrum	346305629_13_1-2883	593	628	1.66E-006	47.21	cl09101	E_set		C
Symbiodinium minutum	003781.t1	285	334	2.41E-012	68.35	cl00615	Membrane-FADS-like		NC
Symbiodinium minutum	003781.t1	554	906	2.24E-108	356.87	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	003781.t1	927	1234	1.61E-082	277.36	cl08282	Acyl_transf_1	AT	-
Symbiodinium minutum	003781.t1	1272	1558	1.19E-033	135.17	cl00509	hot_dog	DH	-
Symbiodinium minutum	003781.t1	1640	1757	1.36E-014	77.71	cl09931	NADB_Rossmann	KR	C
Symbiodinium minutum	003781.t1	1785	2060	1.10E-074	254.8	cl16912	MDR	ER	-
Symbiodinium minutum	003781.t1	2068	2305	7.17E-067	235.26	cl09931	NADB_Rossmann	KR	N
Symbiodinium minutum	003781.t1	2344	2390	2.37E-005	45.7	cl09936	PP-binding	ACP	C
Symbiodinium minutum	003781.t1	3477	3792	3.35E-017	84.79	cl19095	Cyt_C5_DNA_methylase		-
Symbiodinium minutum	003781.t1	4109	4158	9.76E-007	50.46	cl19088	ANK		NC
Symbiodinium minutum	008781.t1	1	114	1.93E-027	116.89	cl09938	cond_enzymes	KS	N
Symbiodinium minutum	008781.t1	142	440	4.58E-026	111.29	cl00509	hot_dog	DH	-
Symbiodinium minutum	008781.t1	511	649	4.79E-009	59.22	cl09931	NADB_Rossmann	KR	C
Symbiodinium minutum	008781.t1	677	1000	6.24E-082	274.06	cl16912	MDR		-
Symbiodinium minutum	008781.t1	1011	1240	1.32E-058	209.45	cl09931	NADB_Rossmann	KR	N
Symbiodinium minutum	008781.t1	1279	1360	2.42E-016	77.29	cl09936	PP-binding	ACP	-
Symbiodinium minutum	008781.t1	1367	1441	1.09E-015	75.36	cl09936	PP-binding	ACP	-
Symbiodinium minutum	008781.t1	1477	1880	2.15E-160	501.7	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	008781.t1	1924	2080	8.14E-020	92.03	cl00509	hot_dog	DH	C
Symbiodinium minutum	008782.t1	9	336	2.44E-080	272.24	cl09931	NADB_Rossmann	KR	-
Symbiodinium minutum	008782.t1	375	449	2.97E-016	76.91	cl09936	PP-binding	ACP	-
Symbiodinium minutum	008782.t1	476	888	8.22E-149	467.42	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	008782.t1	999	1374	3.45E-050	184.8	cl09931	NADB_Rossmann	KR	-
Symbiodinium minutum	008782.t1	1407	1484	9.88E-015	72.28	cl09936	PP-binding	ACP	-
Symbiodinium minutum	008782.t1	1521	1840	1.65E-118	381.91	cl09938	cond_enzymes	KS	C
Symbiodinium minutum	008783.t1	15	283	8.37E-032	125.16	cl00509	hot_dog	DH	-
Symbiodinium minutum	008783.t1	376	492	8.68E-013	69.24	cl09931	NADB_Rossmann	KR	C
Symbiodinium minutum	008783.t1	521	557	1.37E-009	58.35	cl16912	MDR	ER	C
Symbiodinium minutum	008784.t1	2	87	1.94E-011	63.74	cl16912	MDR	ER	N
Symbiodinium minutum	008784.t1	89	331	1.29E-059	203.29	cl09931	NADB_Rossmann	KR	N
Symbiodinium minutum	008784.t1	383	424	1.74E-005	43.01	cl09936	PP-binding	ACP	C
Symbiodinium minutum	008784.t1	452	514	2.85E-007	48.4	cl09936	PP-binding	ACP	C
Symbiodinium minutum	012436.t1	88	303	3.77E-041	158.72	cl08282	Acyl_transf_1	AT	N
Symbiodinium minutum	012436.t1	441	723	7.46E-014	75.42	cl19241	Condensation	KS	-
Symbiodinium minutum	012436.t1	884	1295	1.79E-130	424.7	cl17068	AFD_class_I	FAAL	-
Symbiodinium minutum	012436.t1	1795	2124	3.78E-073	256.33	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	012436.t1	2149	2425	6.60E-023	103.97	cl00509	hot_dog	DH	-
Symbiodinium minutum	012436.t1	2518	2846	1.14E-054	200.97	cl09931	NADB_Rossmann	KR	-
Symbiodinium minutum	012436.t1	2907	2958	1.14E-005	47.56	cl09936	PP-binding	ACP	N
Symbiodinium minutum	012436.t1	3054	3457	3.32E-124	404.25	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	012436.t1	3749	3976	4.90E-016	82.4	cl00509	hot_dog	DH	-
Symbiodinium minutum	012436.t1	4156	4254	4.97E-013	70.1	cl17173	AdoMet_MTases		-
Symbiodinium minutum	012436.t1	4267	4576	5.15E-036	145.6	cl09931	NADB_Rossmann	KR	N
Symbiodinium minutum	012436.t1	4706	5110	2.51E-095	320.66	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	012436.t1	5220	5265	1.90E-003	43.55	cl08282	Acyl_transf_1	AT	C
Symbiodinium minutum	012436.t1	5536	5844	1.60E-043	165.56	cl19241	Condensation	KS	-
Symbiodinium minutum	012436.t1	5870	5929	6.09E-013	68.8	cl07150	NRPS		-
Symbiodinium minutum	012436.t1	6040	6470	1.42E-143	461.14	cl17068	AFD_class_I	FAAL	-
Symbiodinium minutum	012436.t1	6829	6868	2.81E-007	56.4	cl17068	AFD_class_I	FAAL	N
Symbiodinium minutum	012436.t1	6979	7249	1.77E-042	162.48	cl19241	Condensation	KS	-
Symbiodinium minutum	012436.t1	7273	7375	1.25E-007	54.03	cl16371	HxxPF_rpt		-
Symbiodinium minutum	012436.t1	7467	7723	1.45E-069	246.59	cl17068	AFD_class_I	FAAL	N
Symbiodinium minutum	012436.t1	7852	8257	6.08E-100	334.14	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	012436.t1	8331	8571	8.52E-017	84.71	cl00509	hot_dog	DH	-
Symbiodinium minutum	012436.t1	8667	9037	2.53E-067	237.95	cl09931	NADB_Rossmann	KR	-
Symbiodinium minutum	012436.t1	9082	9142	2.62E-003	40.63	cl09936	PP-binding	ACP	-
Symbiodinium minutum	012436.t1	9172	9582	1.62E-128	416.96	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	012436.t1	9730	10068	5.56E-051	189.9	cl09931	NADB_Rossmann	KR	-
Symbiodinium minutum	012436.t1	10213	10427	4.85E-019	91.73	cl19140	Hydrolase_4		-
Symbiodinium minutum	015786.t1	1	207	3.45E-061	209.73	cl16912	MDR	ER	N
Symbiodinium minutum	015786.t1	233	452	3.01E-047	173.24	cl09931	NADB_Rossmann	KR	N

Symbiodinium minutum	015786.t1	498	581	5.97E-021	89.23	cl09936	PP-binding	ACP	-
Symbiodinium minutum	015786.t1	625	893	6.91E-020	90.19	cl19140	Hydrolase_4		-
Symbiodinium minutum	015788.t1	30	195	4.03E-048	169.2	cl09931	NADB_Rossmann	KR	-
Symbiodinium minutum	015788.t1	288	363	3.06E-015	72.67	cl09936	PP-binding	ACP	-
Symbiodinium minutum	015788.t1	393	805	1.60E-148	447.39	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	015789.t1	104	503	1.75E-070	241.42	cl09931	NADB_Rossmann	KR	-
Symbiodinium minutum	015789.t1	541	615	1.13E-015	74.59	cl09936	PP-binding	ACP	-
Symbiodinium minutum	015789.t1	639	1067	7.74E-144	443.92	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	026959.t1	1	150	1.78E-011	65.07	cl00509	hot_dog	DH	N
Symbiodinium minutum	026959.t1	280	373	4.65E-004	42.27	cl09931	NADB_Rossmann	KR	C
Symbiodinium minutum	026959.t1	404	694	1.52E-096	310.66	cl16912	MDR	ER	-
Symbiodinium minutum	026959.t1	701	930	7.00E-058	205.31	cl09931	NADB_Rossmann	KR	N
Symbiodinium minutum	026959.t1	973	1074	1.02E-014	71.51	cl09936	PP-binding	ACP	-
Symbiodinium minutum	029496.t1	1252	1316	3.11E-007	51.09	cl15299	VCBS		-
Symbiodinium minutum	029496.t1	1305	1371	3.71E-006	48.01	cl15299	VCBS		-
Symbiodinium minutum	029496.t1	1633	1733	2.87E-006	48.96	cl19082	TNFR		-
Symbiodinium minutum	029496.t1	1716	1812	9.09E-009	56.67	cl19082	TNFR		-
Symbiodinium minutum	029496.t1	2220	2618	5.65E-123	399.63	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	029496.t1	2651	2928	1.46E-034	137.87	cl00509	hot_dog	DH	-
Symbiodinium minutum	029496.t1	3030	3389	3.77E-070	244.89	cl09931	NADB_Rossmann	KR	-
Symbiodinium minutum	029496.t1	3422	3495	1.19E-015	76.52	cl09936	PP-binding	ACP	-
Symbiodinium minutum	029496.t1	3523	3931	1.81E-134	432.75	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	029496.t1	4078	4378	7.99E-037	146.66	cl09931	NADB_Rossmann	KR	N
Symbiodinium minutum	029496.t1	4409	4489	1.25E-016	79.22	cl09936	PP-binding	ACP	-
Symbiodinium minutum	029496.t1	4533	4940	4.89E-140	448.55	cl09938	cond_enzymes	KS	-
Symbiodinium minutum	039996.t1	1	253	2.61E-085	279.83	cl09938	cond_enzymes	KS	C
Symbiodinium minutum	039996.t1	289	591	3.51E-111	343.62	cl08282	Acyl_transf_1	AT	-
Symbiodinium minutum	039996.t1	627	780	1.07E-020	92.8	cl00509	hot_dog	DH	C
Symbiodinium minutum	040601.t1	53	347	1.41E-071	230.15	cl16912	MDR	ER	-
Symbiodinium minutum	040601.t1	359	448	8.97E-017	77.52	cl09931	NADB_Rossmann	KR	C

[1] See Fig. 3A and Table S2 for domain organization and names, respectively. Note that FAAL domains in *Symbiodinium* are highly divergent and phylogenetically distant to those in *Cryptosporidium* and chrompodellids.

[2] N=incomplete at N-terminus; C=incomplete at C-terminus