Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/5/4/eaav1110/DC1

Supplementary Materials for

An aerobic eukaryotic parasite with functional mitochondria that likely lacks a mitochondrial genome

Uwe John*, Yameng Lu, Sylke Wohlrab, Marco Groth, Jan Janouškovec, Gurjeet S. Kohli, Felix C. Mark, Ulf Bickmeyer, Sarah Farhat, Marius Felder, Stephan Frickenhaus, Laure Guillou, Patrick J. Keeling, Ahmed Moustafa, Betina M. Porce, Klaus Valentin, Gernot Glöckner*

*Corresponding author. Email: uwe.john@awi.de (U.J.); gernot.gloeckner@uni-koeln.de (G.G.)

Published 24 April 2019, *Sci. Adv.* **5**, eaav1110 (2019)
DOI: 10.1126/sciadv.aav1110

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/5/4/eaav1110/DC1)

Data file S1 (.gz format). Contigs of the *A. ceratii* AT5.2 genome.
Data file S2 (.bz2 format). Scaffolds of the *A. ceratii* AT5.2 genome.
Data file S3 (.bz2 format). GFF of the *A. ceratii* AT5.2 genome.
Data file S4 (.gz format). Predicted protein sequences of the *A. ceratii* AT5.2 genome.
Data file S5 (.pdf format). Assembly statistics of the *A. ceratii* AT5.2 genome.
Data file S6 (Microsoft Excel format). Annotation table generated via Trinotate for the *A. ceratii* AT5.2 genome.

Figures

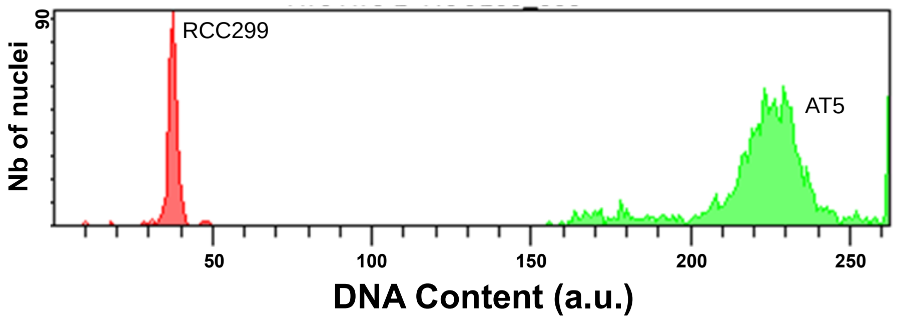


Fig. S1. Flow cytometry. Genome size estimation via flow cytometry. The genome size of *Amoebophrya ceratii* was calculated to be ~120Mb (green) compared to *Micromonas pusilla* RCC299 (1C = 20.9 fg). The predicted genome size resulting from the assembly (87.7. Mb) is smaller than the size indicated by flow cytometry (~120 Mb). This size difference between assembly and flow cytometry is partially due to repetitive elements, which collapse in the assembly to small contigs. Nuclei were extracted by mixing 50 µL of a culture of freshly produced dinospore with 450 µL of twice diluted NIB buffer, containing SYBR Green-I at final concentration of 1/5 000. Two µL of RCC299 in exponential growing phase were added as internal reference. The mixture was then incubated at least for 30 min in the dark, before being analysed using a FACS Canto II flow cytometer equipped with a 488 nm laser and the standard filter setup. Signal was triggered on the green fluorescence. The ratio between the mean distribution of the dinospores and the RCC299 is used for the evaluation of the DNA content.

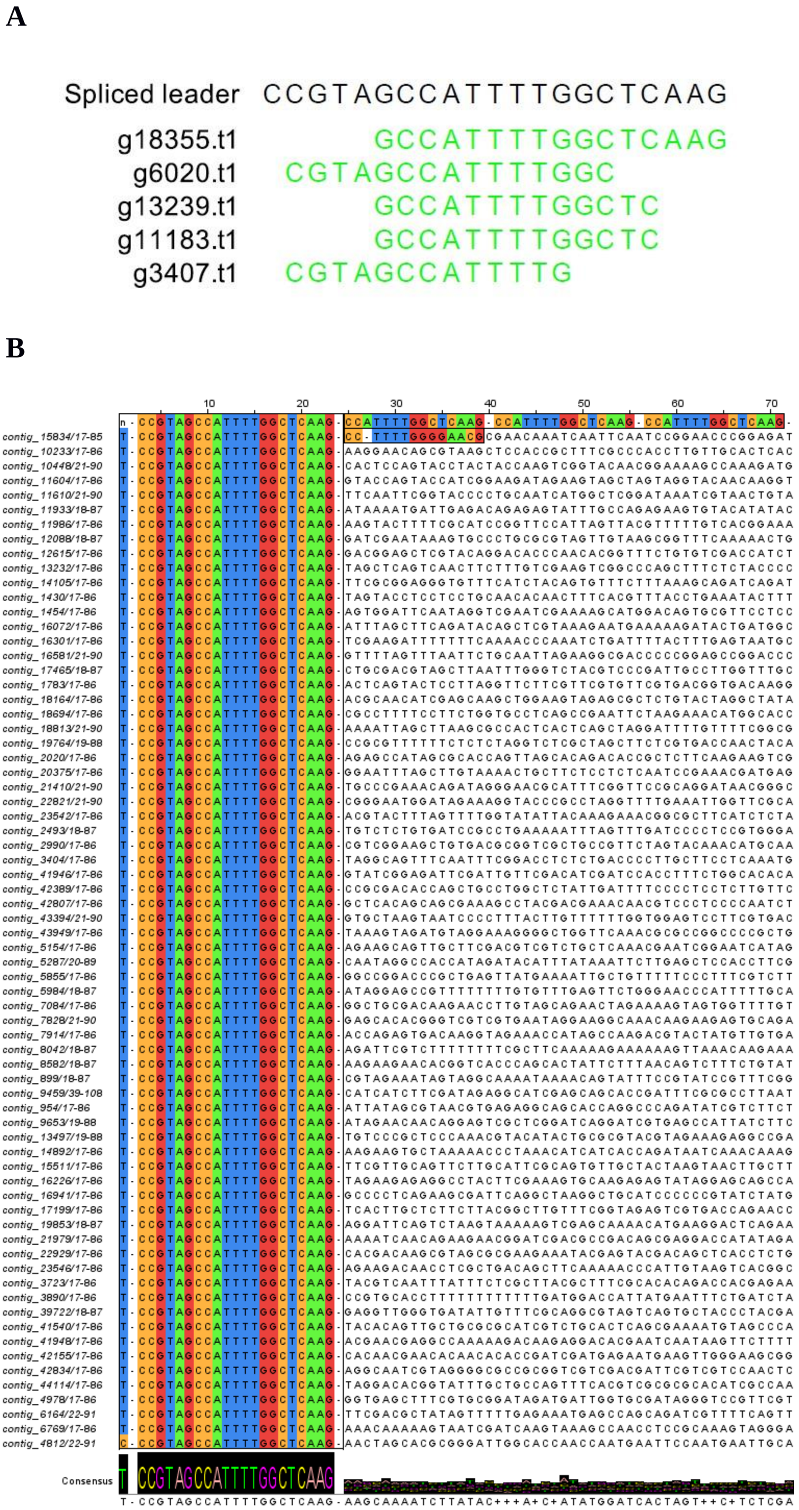


Fig. S2. Alignment of 5′ ends of transcripts of *A. ceratii* showing SL and relict SL repeats in (***A***) *A. ceratii* gene models and (***B***) the transcriptome dataset. The first sequence shows a reference SL consensus sequence. The calculated consensus sequence and the sequence logo are aligned under the transcripts.

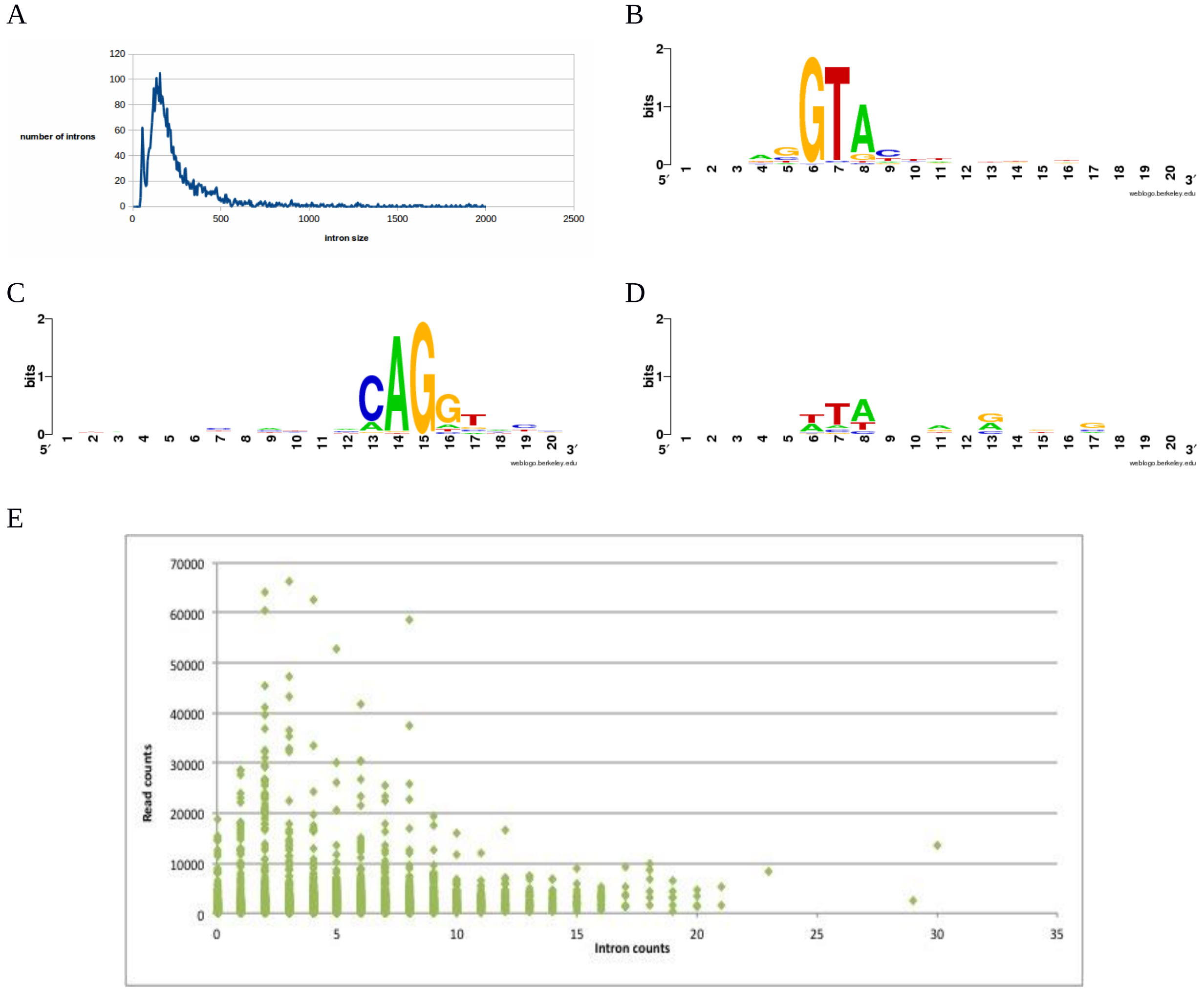


Fig. S3. The distribution of intron sizes in *A. ceratii* genes. (***A***) The distribution of intron sizes in *Amoebophrya ceratii* genes. Only introns confirmed by transcriptome data were taken into account to calculate donor-acceptor distances irrespective of predicted gene models. The mean length of introns is around 100 bases. A TopHat (Version 2.0.1) (https://ccb.jhu.edu/software/tophat/index.shtml) mapping of reads on the assembled scaffold data to resolve junction sites independently confirmed the gene-model based estimate of 51,066 introns in total. We defined consensus donor and acceptor sites of introns confirmed by transcripts: (***B***) donor site; (***C***) acceptor site; (***D***) splicing site AT type. Exon/intron boundaries were centered on the splice site (donors: position 5/6; acceptor: position 13/14). (***E***) Expression level dependency on intron number. RNAseq read counts were plotted against intron numbers per gene. *A. ceratii* also has some AT-AC splice boundaries, which are widely accepted to indicate U12 dependent splicing in eukaryotes.

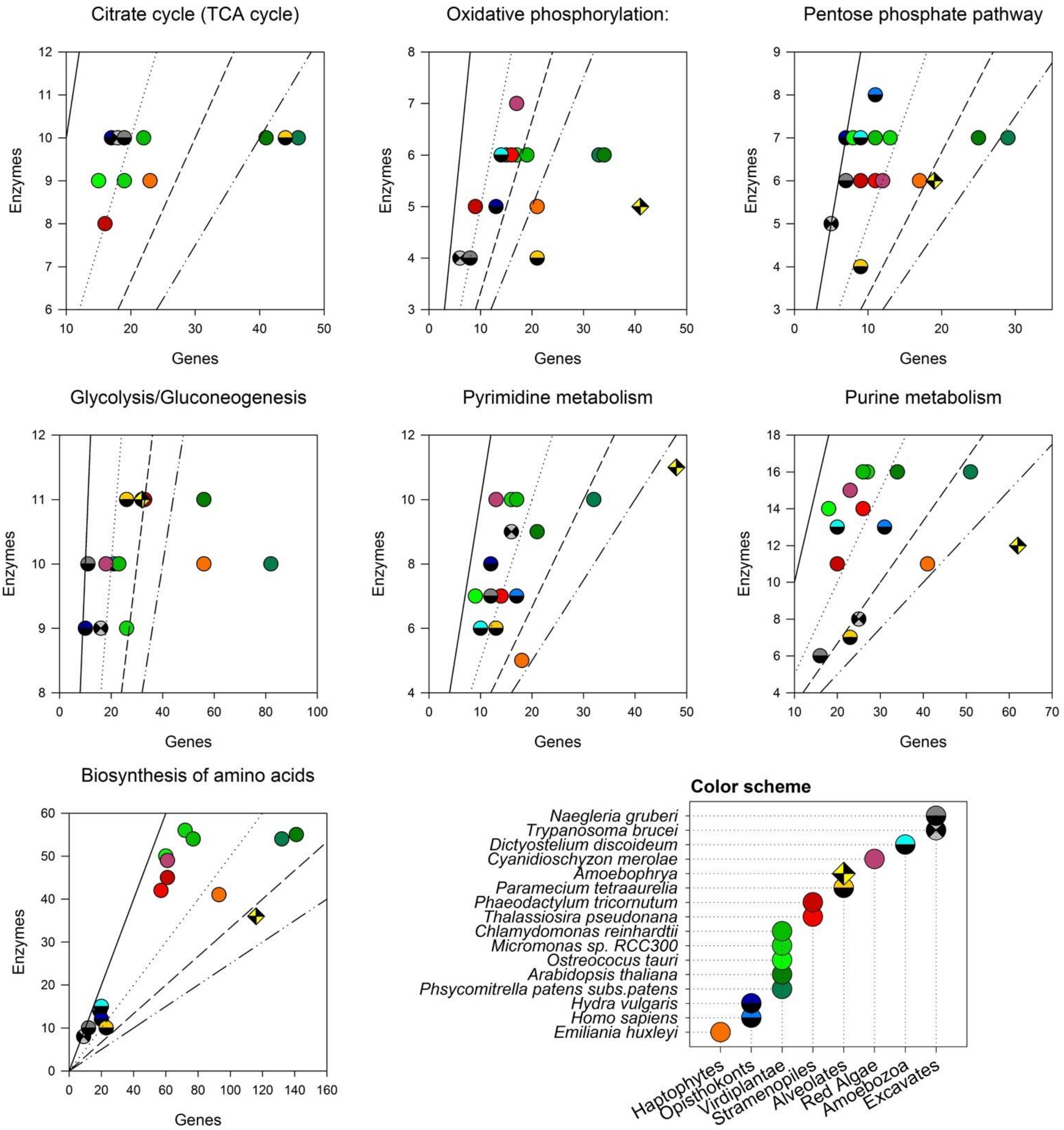


Fig. S4. Expansion of gene numbers per metabolic pathway in *A. ceratii*. Numbers were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG database) and compared with the EC annotations of *A. ceratii* (combined from the Trinotate annotation (http://trinotate.github.io/) and KAAS annotation pipeline Moriya (https://www.genome.jp/kegg/kaas/). For each species the numbers of genes per pathway were plotted against the number of essential enzymes in the respective pathway. Colours represent different taxonomic lineages, fill patterns different lifestyles (colour only: autotroph; half filled: heterotroph; quarter filled: parasitic).

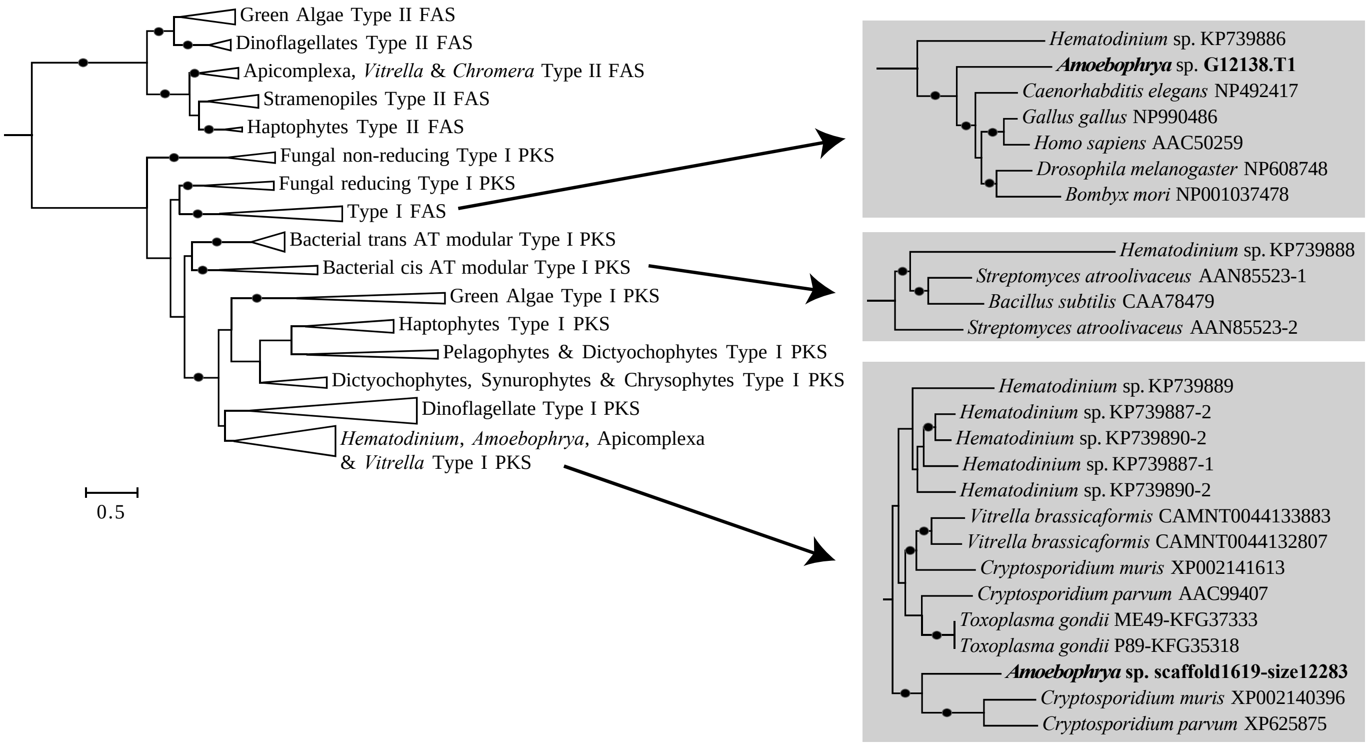


Fig. S5. Phylogenetic analysis of prokaryotic and eukaryotic PKS and FAS. Phylogenetic analysis of 26 type II 3-ketoacyl ACP synthase II and 74 type I ketosynthase domains from prokaryotic and eukaryotic polyketide synthase and fatty acid synthases, showing the position of each major group. To align sequences, MAFFT and ClustalW (within Geneious) were used and alignments were trimmed manually to ensure they spanned the same coding region of each enzyme. The final alignment had a length of 585. Maximum likelihood phylogenetic analysis was carried out using RAxML with 1000 bootstraps using the GAMMA and LG model of rate heterogeneity. Phylogenetic trees were visualised using Geneious (https://www.geneious.com/) and MEGA:Version 6 (https://www.megasoftware.net/). Solid circles indicate bootstrap values above 75.

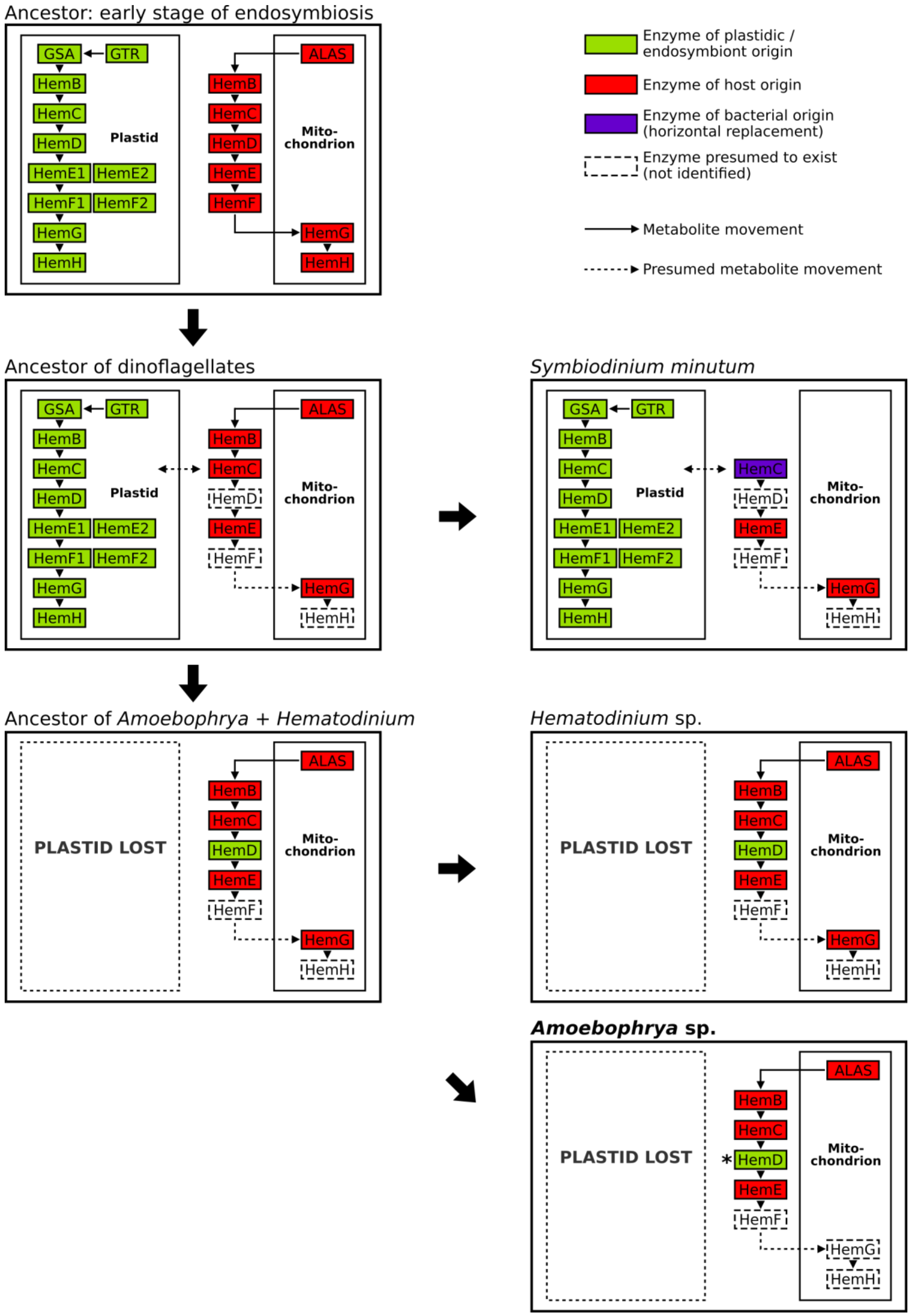


Fig. S6. Predicted evolution of tetrapyrrole biosynthesis in *A. ceratii* and other dinoflagellates. The *in silico* reconstruction combines the results of genome and transcriptome data searches, single-protein maximum likelihood phylogenies and N-terminal predictions of protein targeting (Materials and Methods). Enzymes are depicted by boxes color-coded as to their evolutionary origin. Lines and arrows show the movement and conversion of metabolites. All *A. ceratii* genes except HemE were expressed. ALAS = g4769, HemB = g9658, HemC = g11423, HemD = g15992, HemE = g5941 + g5942. A second HemD-like paralog was identified in *A. ceratii*, a putative pseudogene (g16908; shown by asterisk).

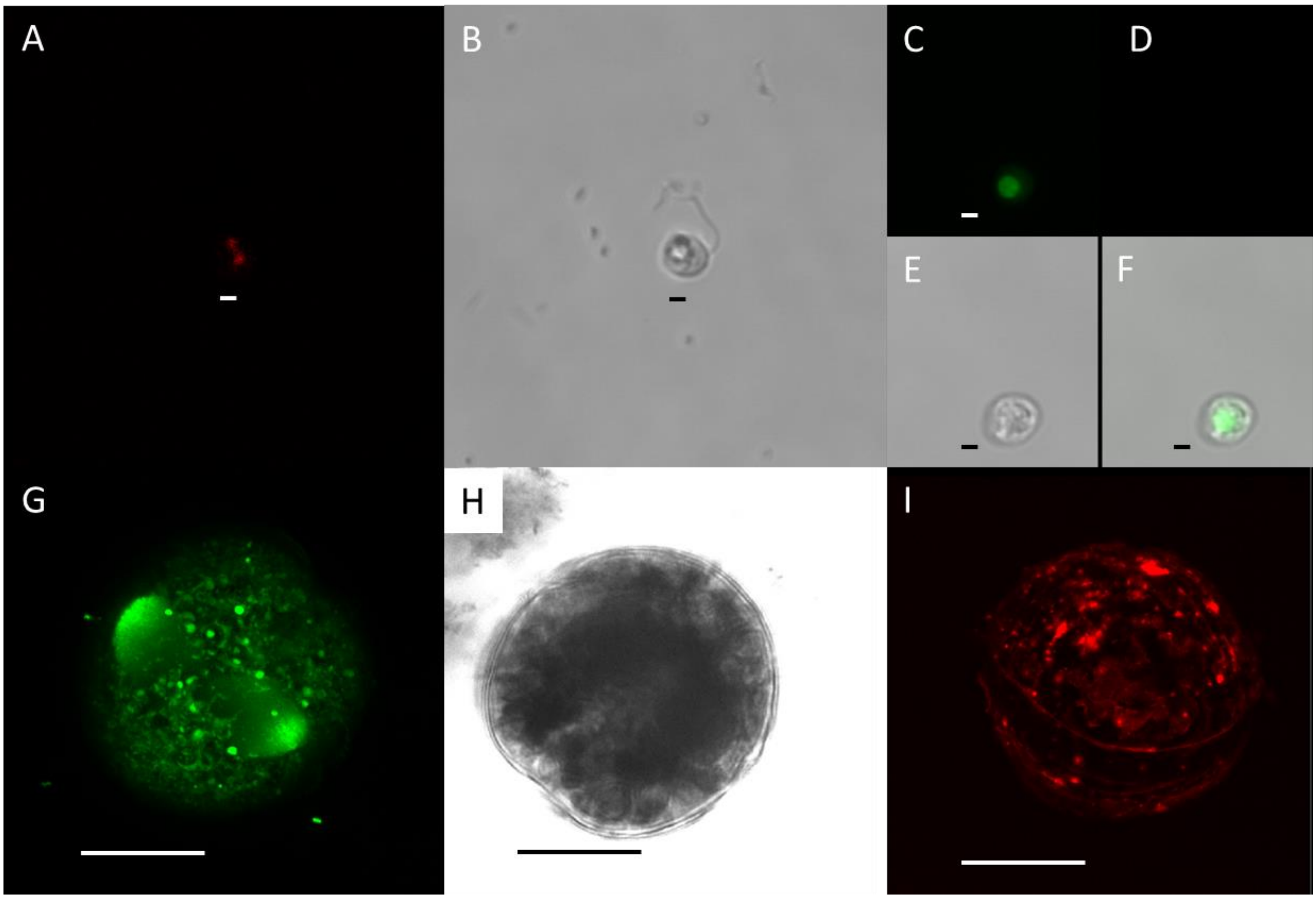


Fig. S7. Confocal microscopy images of *A. ceratii* and *A. catenella* as a control. The dyes Syto-13 and JC-10 possess overlapping fluorescence profiles, therefore cells are stained with one dye each. **A**) *A. ceratii* stained with JC-10, red fluorescence show two mitochondria. **B**) Image in transmission light of the same cell. **C**) *A. ceratii* stained with Syto 5µM showing green fluorescence nucleus. **D**) *A. ceratii* without stain (check for autofluorescence). **E**) *A. ceratii* image in transmission light. **F**) Overlay image of Syto fluorescence and transmission image. **G**) Syto stained DNA of *A. catenella*. **H**) Corresponding image in transmission light to (H). **I**) *A. catenella* stained with JC-10 red fluorescence show active organelles. More images in Figure 3.

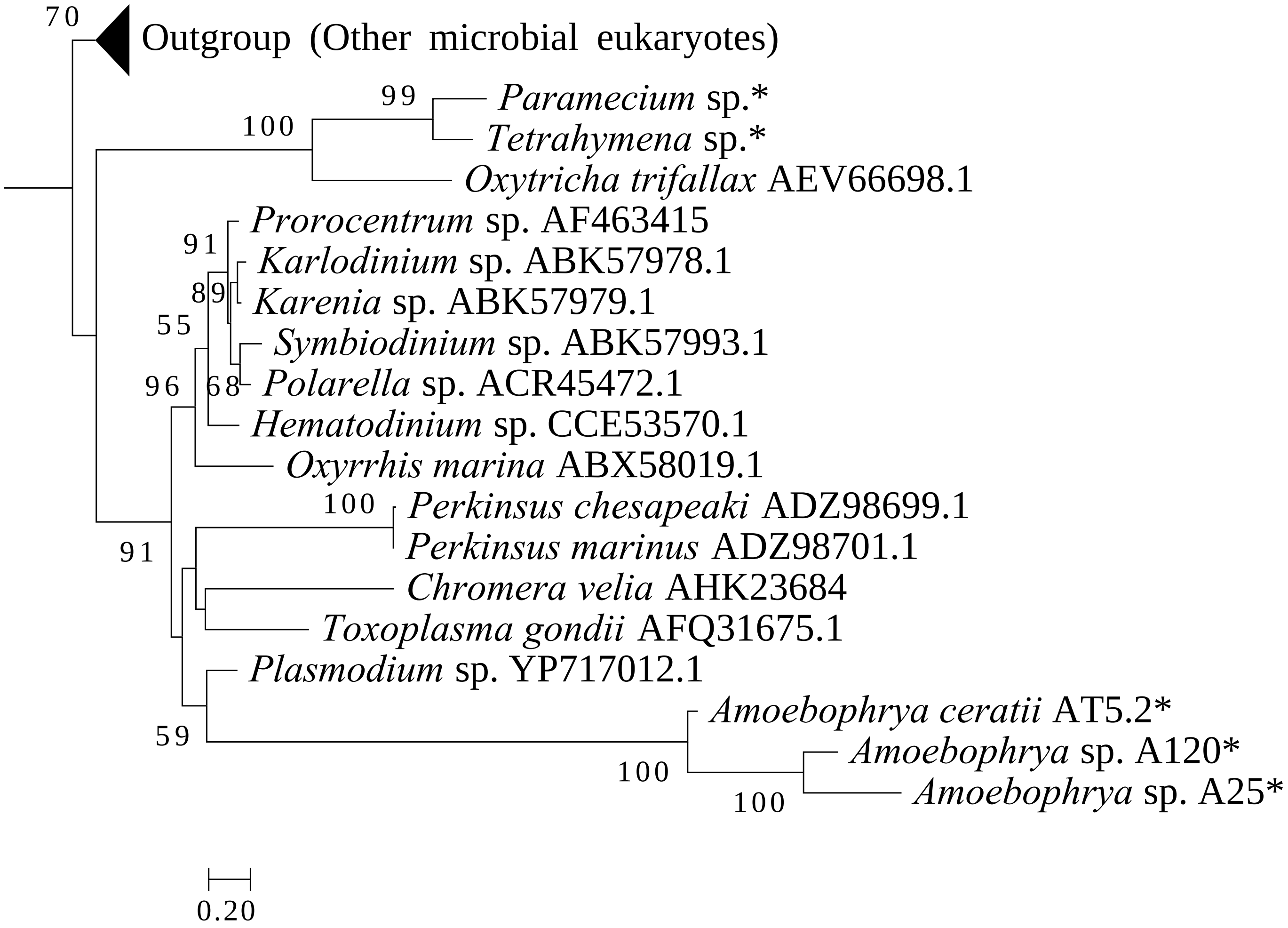
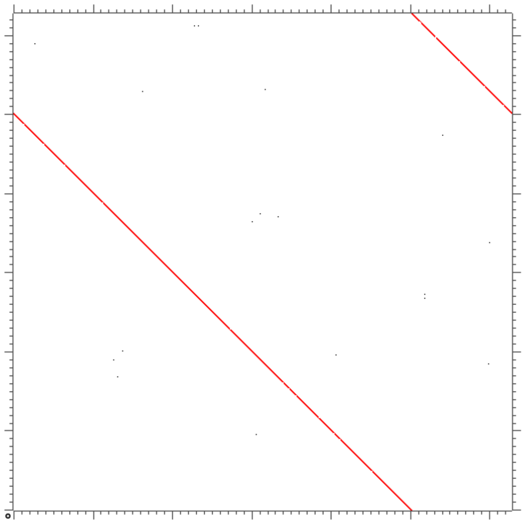


Fig. S8. *CoxI* phylogeny. Concatenated *coxI* gene fragments (g833 and g15932) used for mitochondrial phylogeny of 31 taxa were calculated from an alignment of *coxI* genes. Spuriously aligned sites of the original alignment were removed using Gblocks (http://molevol.cmima.csic.es/castresana/Gblocks_server.html) to use only conserved sequences for the calculation of the phylogeny (200 AA after removal of spurious sites). The Maximum likelihood (ML) analysis (PROTGAMMAILGF, 1000 bootstrap replicates) was conducted with RAxML. The tree was rooted with algal mitochondria (13 taxa) outside the Alveolata. *coxI* fragments for the all three *Amoebophrya* strains were sourced from their reference genomes. *coxI* genes for all other dinoflagellates were recovered from various EST datasets. (*) Sequences were obtained from their respective reference genomes.

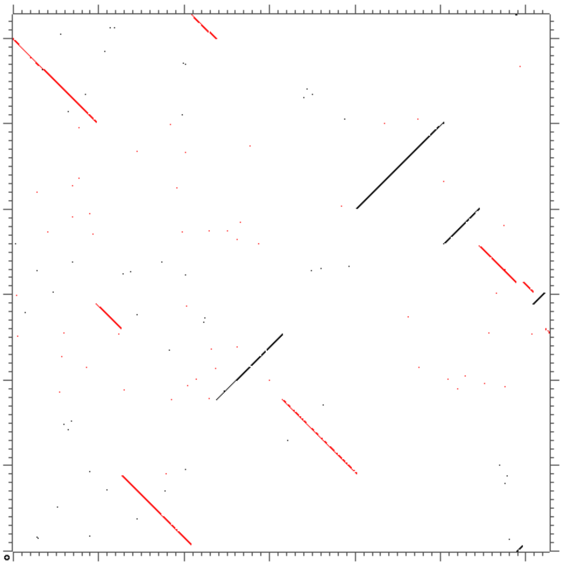


```
SEQUENCES, IMAGE DIMENSIONS
 seq 1, x axis : contig_6535
 seq 2, y axis : NC_002508.1 Physarum polycephalum mitochondrion, complete genome
 len seq 1 : 62862 bp
 len seq 2 : 62862 bp
 picture size : 500x500, 250000 pixels
 scaling x : 0.007954, 125.7 bases per pixel
 scaling y : 0.007954, 125.7 bases per pixel
 bp^2 per pixel : 15806

COMPARISON PRELUDE
 tuple size : 9 (automatic)
 tuple space : 262144
 expected tuple inst. E(N_tuple_i|seq1,seq2) : 0.479599
 expected hits E(N_tuple_i|seq1) * E(N_tuple_i|seq2) : 0.057504
 tuple_overrep_cutoff (option -i) : 1.000000
 tuple_throwoff : N_tuple_i > 2.00000
 tuples ignored due to overrepresentation : 14114 (67620 fwd instances)

COMPARISON, IMAGE CONSTRUCTION
 + : min_value=0.0 max_value=5.0
 - : min_value=0.0 max_value=70.0
 cutoff method : adaptive, diagonal portion (option -A) 1.000000
 cutoff method : adaptive, max. portion positive 0.004000
 at most pixel to set : 1000
 cutoff pixel score : >3.00000
 tic step x (minor) : 1000 bp, 7.953931 pixels
 tic step y (minor) : 1000 bp, 7.953931 pixels
 pixels set : 999, portion 0.003996
```

Fig. S9. Plot obtained by assembling the PE dataset including spike-in mitochondrial reads of *P. polycephalum*. Tuple_plot of contig_6535 obtained by assembling the paired-end dataset including spike-in of simulated mitochondrial reads of *Physarum polycephalum* with a coverage of 226-fold. In order to check the theoretical ability to find mitochondrial sequences, paired-end reads of *Physarum polycephalum* (Genbank accession: NC_002508) mitochondria were simulated using *maq* (http://maq.sourceforge.net/maq-man.shtml). Doing so the simulation was trained using *maq simutrain* by introducing the real paired-end data. Afterwards the simulated reads were created using *maq simulate* (parameter: -d 375 -s 50 -N NO_OF_READS -1 92 -2 86). Insert length as well as lengths of reads_1 and reads_2 were chosen to be similar to those of the real paired-end reads. Reads were simulated representing a coverage of 5-fold to around 500-fold (5, 11, 22, 33, 45, 56, 159, 226, 339, 556-fold), . Since the mitochondrion is circular the simulation was performed twice per dataset using different linearized mitochondrial reference sequences, e.g. for simulate 226-fold coverage dataset 80,000 read pairs are necessary. So the simulation was done twice creating 40,000 read pairs each. After the simulation the reads were concatenated with the real paired-end dataset and introduced in the *clc_assembly_cell* as described. Using BLAT the mitochondrion representing contigs were determined. Using tuple_plot (https://genome.leibniz-fli.de/software/) the position and direction of contigs with respect to the mitochondrion was analysed. In summary assemblies of datasets of all coverages (except the 5-fold coverage dataset) resulted in a single contig representing the complete mitochondrial sequence (fig. S9, 226-fold coverage dataset). The assembly approach of the 5-fold coverage dataset resulted in 13 contigs which represent almost the complete mitochondrial sequence (fig. S10).



```
SEQUENCES, IMAGE DIMENSIONS
 seq 1, x axis : assembly_contig_numbers_13
 seq 2, y axis : NC_002508.1 Physarum polycephalum mitochondrion, complete genome
 len seq 1 : 62834 bp
 len seq 2 : 62862 bp
 picture size : 499x500, 249500 pixels
 scaling x : 0.007942, 125.9 bases per pixel
 scaling y : 0.007954, 125.7 bases per pixel
 bp^2 per pixel : 15831

COMPARISON PRELUDE
 tuple size : 9 (automatic)
 tuple space : 262144
 expected tuple inst. E(N_tuple_i|seq1,seq2) : 0.479492
 expected hits E(N_tuple_i|seq1) * E(N_tuple_i|seq2) : 0.057478
 tuple_overrep_cutoff (option -i) : 1.000000
 tuple_throwoff : N_tuple_i > 2.00000
 tuples ignored due to overrepresentation : 13778 (68737 fwd instances)

COMPARISON, IMAGE CONSTRUCTION
 + : min_value=0.0 max_value=65.5
 - : min_value=0.0 max_value=91.0
 cutoff method : adaptive, diagonal portion (option -A) 1.000000
 cutoff method : adaptive, max. portion positive 0.004003
 at most pixel to set : 998
 cutoff pixel score : >2.50000
 tic step x (minor) : 1000 bp, 7.941560 pixels
 tic step y (minor) : 1000 bp, 7.953931 pixels
 pixels set : 992, portion 0.003976
```

Fig. S10. Plot obtained by assembling the PE dataset including spike-in mitochondrial reads of *P. polycephalum.* Tuple_plot of 13 contigs obtained by assembling the paired-end dataset including spike-in of simulated mitochondrial reads of *Physarum polycephalum* with a coverage of 5-fold. In order to check the theoretical ability to find mitochondrial sequences, paired-end reads of *Physarum polycephalum* (Genbank accession: NC_002508) mitochondria were simulated using *maq* (http://maq.sourceforge.net/maq-man.shtml). Doing so the simulation was trained using *maq simutrain* by introducing the real paired-end data. Afterwards the simulated reads were created using *maq simulate* (parameter: -d 375 -s 50 -N NO_OF_READS -1 92 -2 86). Insert length as well as lengths of reads_1 and reads_2 were chosen to be similar to those of the real paired-end reads. Reads were simulated representing a coverage of 5-fold to around 500-fold (5, 11, 22, 33, 45, 56, 159, 226, 339, 556-fold), . Since the mitochondrion is circular the simulation was performed twice per dataset using different linearized mitochondrial reference sequences, e.g. for simulate 226-fold coverage dataset 80,000 read pairs are necessary. So the simulation was done twice creating 40,000 read pairs each. After the simulation the reads were concatenated with the real paired-end dataset and introduced in the *clc_assembly_cell* as described. Using BLAT the mitochondrion representing contigs were determined. Using tuple_plot (https://genome.leibniz-fli.de/software/) the position and direction of contigs with respect to the mitochondrion was analysed. In summary assemblies of datasets of all coverages (except the 5-fold coverage dataset) resulted in a single contig representing the complete mitochondrial sequence (fig. S9, 226-fold coverage dataset). The assembly approach of the 5-fold coverage dataset resulted in 13 contigs which represent almost the complete mitochondrial sequence (fig. S10).

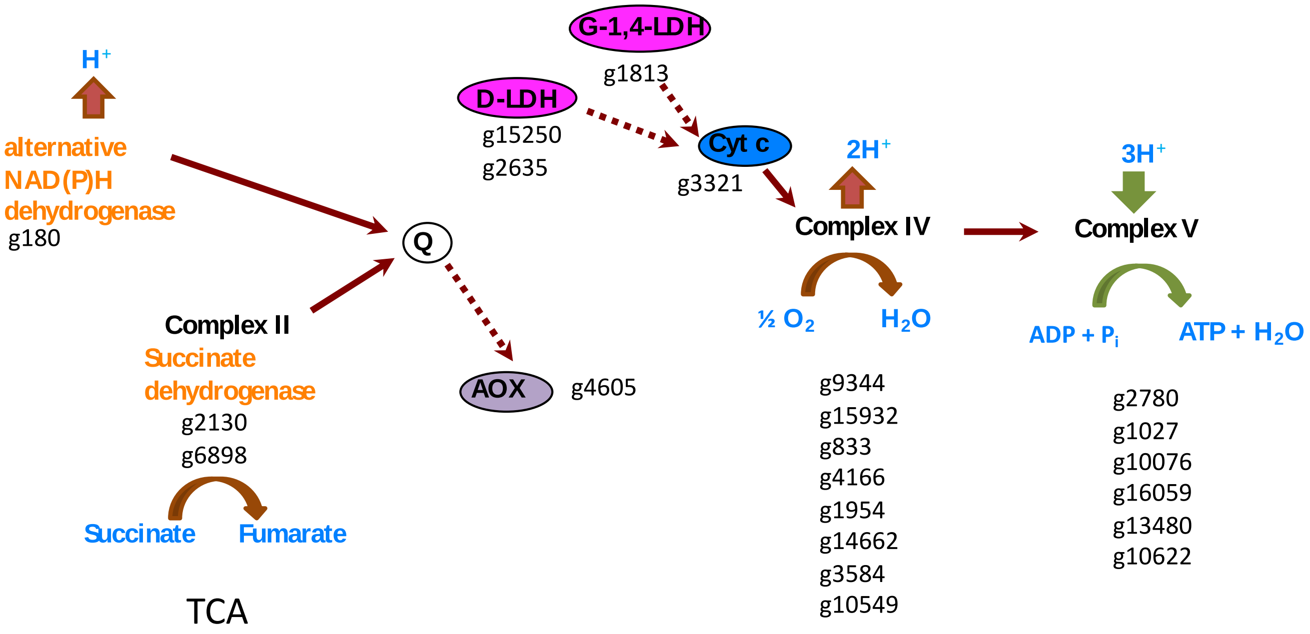


Fig. S11. Model of the mitochondrial respiratory chain in *A. ceratii* in accordance to Figure 5. Gene identifiers according to the *A. ceratii* database encoding components of complexes are shown below the complexes.

Tables

Table S1. Enriched gene families. We clustered all predicted proteins using OrthoMCL (http://www.orthomcl.org/orthomcl/). The families (clusters) were then examined for functions according to their domain structure. Furthermore, we combined the gene families according to the functions of their members if the majority of members of a specific family contained the same indicative domain. Families without any definable function were combined to one group, which comprises the highest number of families and proteins. Among the most prominent functions found in families with at least 10 members in *Amoebophrya ceratii* are many transposon derived sequences followed by proteins involved in carbohydrate or protein binding. The prominent occurrence of carbohydrate degradation in this list indicates an important role of these functions in the parasites life cycle.

Function	Number of proteins	Number of families
Calcium binding	10	1
Cation channel	11	1
Nucleotide-diphospho-sugar transferase	12	1
Galactosyltransferase	13	1
ABC transporter	16	1
Peptidase	16	1
Cell surface	21	1
Phosphoesterase	24	1
Cell cycle or growth phase-related regulation	32	1
Kinase	32	3
Protease or protease inhibitor	32	3
Cytoskeletton	36	1
Detoxification and degradation	44	2
Carbohydrate degradation	50	3
Protein-protein or protein-carbohydrate interactions	107	6
Transposon	499	12
No definition	1,424	76

Table S2. Number of genes encoding transcription factors in genomes/transcriptomes of 48 genera. (**A**). In order to find protein domains that might function as transcription factors in *Amoebophrya ceratii*, the predicted proteins were searched using Pfam profiles (https://pfam.xfam.org/) and in total 57 unique protein domains that are described as being involved in gene regulation were detected. In order to compare these results, genomes/transcriptomes for 47 other microbial eukaryotic genera belonging to alveolates, Stramenopiles, Rhizaria, Haptophyta, Viridiplantae, Rhodophyta, Euglenids and Choanoflagellatea were also screened for these 57 transcription factors. All other transcription factors except one were found in almost all other genera. One protein/transcription factor "PF06167- Glucose-regulated metallo-peptidase M90" was only found in the *A. ceratii* genome. B shows species used in this analysis.

A

Organism (No. of peptides in library)	**Transcription Factors**				
	TF TFIIB repeat PF00382	**SNF2 family N-terminal domain PF00176**	**Ubiquitin family PF00240**	**DEAD/DEAH box helicase PF00270**	**TATA-binding protein TBP PF00352**
Alveolates (basal[1] and core[2] dinoflagellates, Ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	1	38	19	88	2
Noctiluca scintillans[1] (40911)	2	44	72	92	1
Oxyrrhis marina[1] (34417)	0	34	36	71	3
*Perkinsus marinus**[1] (23654)	6	86	28	124	5
Alexandrium catenella[2] (1108345)	3	124	226	179	4
Amphidinium carterae[2] (36236)	1	56	85	107	1
Crypthecodinium cohnii[2] (48368)	2	70	66	97	2
Gambierdiscus australes[2] (83402)	2	112	147	174	2
Gymnodinium catenatum[2] (83508)	2	136	193	260	2
Heterocapsa triquetra[2] (57789)	1	52	68	99	1
Karenia brevis[2] (93029)	2	104	258	193	2
Karlodinium micrum[2] (58585)	3	94	155	151	1
Kryptoperidinium foliaceum[2] (162643)	3	156	208	249	9
Lingulodinium polyerda[2] (99361)	1	118	201	155	3
Polarella glacialis[2] (59754)	1	118	97	231	0
Scrippsiella hangoei[2] (75339)	1	108	163	188	4
Symbiodinium sp.[2] (44733)	1	50	95	76	2
Togula jolla[2] (42755)	3	66	100	151	4
Favella ehrenbergii[3] (14034)	2	1	13	21	0
Strombidinopsis acuminatum[3] (33155)	1	9	67	65	1
*Eimeria tenella**[4] (8597)	1	21	15	41	2
*Neospora caninum**[4] (7122)	2	22	14	56	2

*Toxoplasma gondii**[4] (8460)	2	22	14	53	2
Vitrella brassicaformis[5] (23743)	2	38	40	40	2
Chromera velia[6] (19244)	1	18	29	43	1
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	1	23	15	42	1
Chromulina nebulosa[8] (3839)	1	3	6	18	0
*Aureococcus anophagefferens**[9] (11501)	1	40	24	65	1
Schizochytrium aggregatum[10] (13373)	1	9	16	33	1
Chattonella subsalsa[11] (18256)	3	19	36	51	1
Heterosigma akashiwo[11] (32905)	2	12	43	42	2
Dinobryon sp. [12] (18284)	2	15	19	65	1
Ochromonas sp.[12] (16703)	2	14	23	48	2
Pinguiococcus pyrenoidosus[13] (10811)	2	26	14	45	0
Vaucheria litorea[14] (8258)	1	6	35	39	2
Rhizaria					
Bigelowiella longifila (15959)	1	1	37	15	2
Lotharella oceanica (17354)	2	5	51	24	3
Haptophytes					
*Guillardia theta** (24840)	4	46	43	78	11
*Emiliania huxleyi** (39125)	2	95	53	142	1
Imantonia sp. (14800)	2	5	32	44	0
Isochrysis galbana (39744)	2	60	44	158	2
Viridiplantae					
Chlamydomonas euryale (19516)	2	25	19	39	1
Dunaliella tertiolecta (13636)	2	30	14	62	1
Micromonas sp. (8956)	2	27	19	82	1
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	1	14	32	66	1
Acanthoeca-like sp. [16] (21655)	3	25	20	50	1
Rhodosorus marinus[17] (9995)	4	26	27	74	1
Porphyridium aerugineum[17] (8890)	3	22	13	35	1

	Leucine zipper PF15924	**Staphylococcal nuclease homologue PF00565**	**F-box domain PF00646**	**BTB/POZ domain PF00651**	**Integrase core domain PF00665**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	1	2	3	7	20
Noctiluca scintillans[1]	1	0	2	52	0

(40911)					
Oxyrrhis marina[1] (34417)	2	1	0	26	0
*Perkinsus marinus**[1] (23654)	0	0	0	12	175
Alexandrium catenella[2] (1108345)	2	2	1	117	0
Amphidinium carterae[2] (36236)	1	1	0	50	1
Crypthecodinium cohnii[2] (48368)	1	0	0	38	0
Gambierdiscus australes[2] (83402)	2	3	0	84	1
Gymnodinium catenatum[2] (83508)	1	1	1	144	0
Heterocapsa triquetra[2] (57789)	1	1	0	40	0
Karenia brevis[2] (93029)	2	0	1	134	0
Karlodinium micrum[2] (58585)	1	0	1	83	0
Kryptoperidinium foliaceum[2] (162643)	1	3	4	96	6
Lingulodinium polyerda[2] (99361)	1	1	0	112	0
Polarella glacialis[2] (59754)	2	2	2	54	11
Scrippsiella hangoei[2] (75339)	1	2	1	146	0
Symbiodinium sp.[2] (44733)	1	0	0	103	0
Togula jolla[2] (42755)	2	3	0	63	0
Favella ehrenbergii[3] (14034)	0	1	0	1	0
Strombidinopsis acuminatum[3] (33155)	1	4	7	12	2
*Eimeria tenella**[4] (8597)	0	3	0	6	0
*Neospora caninum**[4] (7122)	0	3	1	6	0
*Toxoplasma gondii**[4] (8460)	0	3	1	8	0
Vitrella brassicaformis[5] (23743)	1	2	4	30	3
Chromera velia[6] (19244)	0	2	0	20	7
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	0	3	4	22	0
Chromulina nebulosa[8] (3839)	0	0	0	3	4
*Aureococcus anophagefferens**[9] (11501)	0	2	5	5	2
Schizochytrium aggregatum[10] (13373)	1	1	1	10	0
Chattonella subsalsa[11] (18256)	1	4	4	37	9
Heterosigma akashiwo[11] (32905)	1	4	1	22	19
Dinobryon sp. [12] (18284)	1	0	0	12	11
Ochromonas sp.[12] (16703)	1	0	0	9	4
Pinguiococcus	0	1	2	11	2

pyrenoidosus[13] (10811)					
Vaucheria litorea[14] (8258)	0	5	2	6	1
Rhizaria					
Bigelowiella longifila (15959)	1	3	2	5	0
Lotharella oceanica (17354)	1	4	4	6	0
Haptophytes					
*Guillardia theta** (24840)	0	2	5	14	2
*Emiliania huxleyi** (39125)	0	6	25	183	0
Imantonia sp. (14800)	1	3	0	9	2
Isochrysis galbana (39744)	2	5	20	23	6
Viridiplantae					
Chlamydomonas euryale (19516)	1	2	6	8	8
Dunaliella tertiolecta (13636)	1	2	7	8	5
Micromonas sp. (8956)	0	1	0	5	0
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	2	2	4	13	2
Acanthoeca-like sp. [16] (21655)	0	1	3	4	4
Rhodosorus marinus[17] (9995)	0	3	5	7	10
Porphyridium aerugineum[17] (8890)	0	4	4	2	4

	Nucleosome assembly protein PF00956	**TF S-II (TFIIS) PF01096**	**D-1 ubiquitin protease PF01398**	**IBR domain PF01485**	**GAF domain PF01590**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	2	2	8	2	10
Noctiluca scintillans[1] (40911)	5	3	9	18	11
Oxyrrhis marina[1] (34417)	3	0	5	11	23
*Perkinsus marinus**[1] (23654)	6	9	10	5	1
Alexandrium catenella[2] (1108345)	3	7	9	39	17
Amphidinium carterae[2] (36236)	4	1	8	17	12
Crypthecodinium cohnii[2] (48368)	4	2	8	15	8
Gambierdiscus australes[2] (83402)	5	6	12	45	9
Gymnodinium catenatum[2] (83508)	10	4	14	30	25
Heterocapsa triquetra[2] (57789)	5	2	7	13	7
Karenia brevis[2] (93029)	9	6	8	34	31
Karlodinium micrum[2] (58585)	5	3	8	34	18
Kryptoperidinium foliaceum[2] (162643)	15	5	21	23	19
Lingulodinium polyerda[2]	5	5	9	21	19

(99361)					
Polarella glacialis[2] (59754)	8	4	20	12	14
Scrippsiella hangoei[2] (75339)	8	4	9	31	8
Symbiodinium sp.[2] (44733)	2	2	9	8	6
Togula jolla[2] (42755)	4	2	9	13	11
Favella ehrenbergii[3] (14034)	2	0	5	6	2
Strombidinopsis acuminatum[3] (33155)	9	3	20	5	0
*Eimeria tenella**[4] (8597)	2	3	3	1	0
*Neospora caninum**[4] (7122)	3	4	6	3	3
*Toxoplasma gondii**[4] (8460)	3	4	6	3	3
Vitrella brassicaformis[5] (23743)	2	1	8	10	7
Chromera velia[6] (19244)	1	4	8	6	0
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	1	2	5	6	3
Chromulina nebulosa[8] (3839)	1	1	2	1	2
*Aureococcus anophagefferens**[9] (11501)	1	4	10	7	1
Schizochytrium aggregatum[10] (13373)	1	2	4	1	4
Chattonella subsalsa[11] (18256)	2	1	9	8	8
Heterosigma akashiwo[11] (32905)	2	2	9	8	13
Dinobryon sp. [12] (18284)	2	0	7	12	4
Ochromonas sp.[12] (16703)	1	2	5	7	5
Pinguiococcus pyrenoidosus[13] (10811)	0	2	7	5	6
Vaucheria litorea[14] (8258)	1	1	5	1	0
Rhizaria					
Bigelowiella longifila (15959)	4	6	6	10	4
Lotharella oceanica (17354)	2	3	9	8	2
Haptophytes					
*Guillardia theta** (24840)	2	4	9	17	11
*Emiliania huxleyi** (39125)	4	6	10	19	3
Imantonia sp. (14800)	2	3	11	5	3
Isochrysis galbana (39744)	7	9	16	9	3
Viridiplantae					
Chlamydomonas euryale (19516)	4	2	10	3	3
Dunaliella tertiolecta (13636)	3	3	11	4	5
Micromonas sp. (8956)	1	2	8	2	2
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					

Eutreptiella gymnastica[15] (21956)	4	1	9	4	11
Acanthoeca-like sp. [16] (21655)	1	3	8	9	1
Rhodosorus marinus[17] (9995)	1	1	7	1	0
Porphyridium aerugineum[17] (8890)	1	4	6	5	0

	BolA-like protein PF01722	**IPT/TIG domain PF01833**	**DJ-1/PfpI family PF01965**	**Sir2 family PF02146**	**Transcriptional Coactivator p15 (PC4) PF02229**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], Chromera[6])					
*Amoebophrya ceratii**[1] (20969)	2	3	4	5	3
Noctiluca scintillans[1] (40911)	7	1	7	15	1
Oxyrrhis marina[1] (34417)	3	1	3	14	2
*Perkinsus marinus**[1] (23654)	5	5	2	9	2
Alexandrium catenella[2] (1108345)	6	7	6	46	1
Amphidinium carterae[2] (36236)	4	7	5	12	1
Crypthecodinium cohnii[2] (48368)	7	2	1	16	1
Gambierdiscus australes[2] (83402)	5	5	6	43	1
Gymnodinium catenatum[2] (83508)	7	2	5	35	1
Heterocapsa triquetra[2] (57789)	1	2	4	12	0
Karenia brevis[2] (93029)	7	9	5	50	2
Karlodinium micrum[2] (58585)	3	2	7	30	2
Kryptoperidinium foliaceum[2] (162643)	9	11	27	25	3
Lingulodinium polyerda[2] (99361)	2	1	9	21	1
Polarella glacialis[2] (59754)	5	6	8	23	2
Scrippsiella hangoei[2] (75339)	5	0	5	25	1
Symbiodinium sp.[2] (44733)	5	6	4	16	1
Togula jolla[2] (42755)	3	1	7	15	2
Favella ehrenbergii[3] (14034)	3	10	0	7	0
Strombidinopsis acuminatum[3] (33155)	3	13	5	6	0
*Eimeria tenella**[4] (8597)	1	1	1	1	0
*Neospora caninum**[4] (7122)	2	0	1	2	0
*Toxoplasma gondii**[4] (8460)	2	0	1	2	0
Vitrella brassicaformis[5] (23743)	3	9	1	4	1
Chromera velia[6] (19244)	3	0	1	5	1
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					

Thalassiosira pseudonana[7] (17050)	3	1	3	7	1
Chromulina nebulosa[8] (3839)	0	1	0	0	0
*Aureococcus anophagefferens**[9] (11501)	2	3	3	3	0
Schizochytrium aggregatum[10] (13373)	2	1	2	5	1
Chattonella subsalsa[11] (18256)	3	8	5	8	1
Heterosigma akashiwo[11] (32905)	5	0	7	5	2
Dinobryon sp. [12] (18284)	1	12	1	3	2
Ochromonas sp.[12] (16703)	2	1	1	2	0
Pinguiococcus pyrenoidosus[13] (10811)	1	1	0	3	1
Vaucheria litorea[14] (8258)	2	1	3	3	1
Rhizaria					
Bigelowiella longifila (15959)	3	0	2	7	4
Lotharella oceanica (17354)	4	0	6	7	4
Haptophytes					
*Guillardia theta** (24840)	5	68	4	6	2
*Emiliania huxleyi** (39125)	8	5	12	24	0
Imantonia sp. (14800)	3	0	3	11	1
Isochrysis galbana (39744)	6	14	8	17	0
Viridiplantae					
Chlamydomonas euryale (19516)	6	4	0	2	1
Dunaliella tertiolecta (13636)	2	6	2	2	1
Micromonas sp. (8956)	1	12	2	3	1
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	4	6	3	12	0
Acanthoeca-like sp. [16] (21655)	4	4	0	7	0
Rhodosorus marinus[17] (9995)	1	1	2	7	0
Porphyridium aerugineum[17] (8890)	1	2	2	2	1

	Pirin PF02678	**Phosphotyrosyl phosphate activator protein PF03095**	**CAS/CSE protein, C-terminus PF03378**	**LEM3 (ligand-effect modulator 3) CDC50 family PF03381**	**Early transcription elongation factor of RNA pol II, NGN PF03439**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	3	1	2	4	1
Noctiluca scintillans[1] (40911)	3	1	1	3	1
Oxyrrhis marina[1] (34417)	6	0	1	4	1
*Perkinsus marinus**[1] (23654)	0	2	7	7	1

Alexandrium catenella[2] (1108345)	5	2	3	8	1
Amphidinium carterae[2] (36236)	11	1	1	10	1
Crypthecodinium cohnii[2] (48368)	3	1	2	6	1
Gambierdiscus australes[2] (83402)	10	1	2	6	1
Gymnodinium catenatum[2] (83508)	4	1	2	6	1
Heterocapsa triquetra[2] (57789)	2	1	1	9	1
Karenia brevis[2] (93029)	6	2	2	12	1
Karlodinium micrum[2] (58585)	4	1	1	7	1
Kryptoperidinium foliaceum[2] (162643)	3	3	5	16	2
Lingulodinium polyerda[2] (99361)	4	1	3	7	1
Polarella glacialis[2] (59754)	6	0	1	9	1
Scrippsiella hangoei[2] (75339)	5	1	3	7	1
Symbiodinium sp.[2] (44733)	7	1	1	7	1
Togula jolla[2] (42755)	4	1	2	9	2
Favella ehrenbergii[3] (14034)	1	1	0	1	0
Strombidinopsis acuminatum[3] (33155)	3	2	1	2	0
*Eimeria tenella**[4] (8597)	0	1	1	3	1
*Neospora caninum**[4] (7122)	0	1	1	4	1
*Toxoplasma gondii**[4] (8460)	0	1	1	4	1
Vitrella brassicaformis[5] (23743)	3	1	1	2	1
Chromera velia[6] (19244)	2	1	1	3	1
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	1	1	1	1	1
Chromulina nebulosa[8] (3839)	2	0	1	0	1
*Aureococcus anophagefferens**[9] (11501)	4	1	1	2	1
Schizochytrium aggregatum[10] (13373)	2	1	1	1	1
Chattonella subsalsa[11] (18256)	3	2	1	2	1
Heterosigma akashiwo[11] (32905)	1	1	2	2	0
Dinobryon sp. [12] (18284)	2	1	1	3	1
Ochromonas sp.[12] (16703)	5	0	1	4	1
Pinguiococcus pyrenoidosus[13] (10811)	4	1	1	1	1
Vaucheria litorea[14] (8258)	2	1	1	3	1
Rhizaria					
Bigelowiella longifila	4	1	2	4	0

(15959)					
Lotharella oceanica (17354)	4	1	1	5	0
Haptophytes					
*Guillardia theta** (24840)	3	1	1	2	1
*Emiliania huxleyi** (39125)	2	2	1	2	2
Imantonia sp. (14800)	2	1	0	2	1
Isochrysis galbana (39744)	6	2	2	2	3
Viridiplantae					
Chlamydomonas euryale (19516)	2	1	1	2	1
Dunaliella tertiolecta (13636)	1	0	1	1	1
Micromonas sp. (8956)	2	0	2	1	1
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	2	0	1	0	1
Acanthoeca-like sp. [16] (21655)	3	1	1	2	1
Rhodosorus marinus[17] (9995)	2	1	2	2	1
Porphyridium aerugineum[17] (8890)	1	1	1	1	1

	Structure-specific recognition protein (SSRP1) PF03531	**TF Tfb2 PF03849**	**TF Tfb4 PF03850**	**BSD domain PF03909**	**Leo1-like protein PF04004**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	1	1	1	1	1
Noctiluca scintillans[1] (40911)	1	0	1	3	1
Oxyrrhis marina[1] (34417)	1	0	1	2	1
*Perkinsus marinus**[1] (23654)	2	1	2	2	2
Alexandrium catenella[2] (1108345)	1	1	1	2	1
Amphidinium carterae[2] (36236)	1	1	1	3	1
Crypthecodinium cohnii[2] (48368)	1	1	1	3	1
Gambierdiscus australes[2] (83402)	2	1	1	5	2
Gymnodinium catenatum[2] (83508)	1	1	2	4	1
Heterocapsa triquetra[2] (57789)	1	0	0	1	1
Karenia brevis[2] (93029)	2	2	1	3	1
Karlodinium micrum[2] (58585)	1	0	1	2	1
Kryptoperidinium foliaceum[2] (162643)	3	3	1	1	3
Lingulodinium polyerda[2] (99361)	1	1	3	2	1
Polarella glacialis[2] (59754)	3	0	0	1	1

Scrippsiella hangoei[2] (75339)	1	1	1	2	1
Symbiodinium sp.[2] (44733)	1	1	1	1	1
Togula jolla[2] (42755)	1	1	0	1	1
Favella ehrenbergii[3] (14034)	1	0	0	3	0
Strombidinopsis acuminatum[3] (33155)	1	0	0	4	1
*Eimeria tenella**[4] (8597)	1	0	1	0	1
*Neospora caninum**[4] (7122)	1	1	1	3	1
*Toxoplasma gondii**[4] (8460)	1	1	1	2	1
Vitrella brassicaformis[5] (23743)	1	2	1	4	1
Chromera velia[6] (19244)	1	1	0	3	1
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	2	0	1	1	1
Chromulina nebulosa[8] (3839)	1	0	0	0	0
*Aureococcus anophagefferens**[9] (11501)	1	1	0	0	0
Schizochytrium aggregatum[10] (13373)	2	1	1	3	1
Chattonella subsalsa[11] (18256)	1	1	0	3	1
Heterosigma akashiwo[11] (32905)	0	1	1	2	1
Dinobryon sp. [12] (18284)	1	1	0	2	1
Ochromonas sp.[12] (16703)	2	2	1	2	1
Pinguiococcus pyrenoidosus[13] (10811)	2	1	0	0	1
Vaucheria litorea[14] (8258)	1	0	1	3	1
Rhizaria					
Bigelowiella longifila (15959)	1	1	1	3	0
Lotharella oceanica (17354)	1	0	1	4	0
Haptophytes					
*Guillardia theta** (24840)	1	1	1	4	1
*Emiliania huxleyi** (39125)	4	1	1	9	2
Imantonia sp. (14800)	2	1	0	7	1
Isochrysis galbana (39744)	3	3	2	11	2
Viridiplantae					
Chlamydomonas euryale (19516)	1	1	0	4	1
Dunaliella tertiolecta (13636)	2	1	0	3	1
Micromonas sp. (8956)	1	0	0	3	1
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	1	1	0	3	1
Acanthoeca-like sp. [16]	1	1	1	2	1

(21655)					
Rhodosorus marinus[17] (9995)	2	2	1	3	2
Porphyridium aerugineum[17] (8890)	1	1	1	3	1

	CCR4-Not complex component, Not1 PF04054	**Cell differentiation family, Rcd1-like PF04078**	**NOT2 / NOT3 / NOT5 family PF04153**	**EAP30/Vps36 family PF04157**	**Transcriptional repressor TCF25 PF04910**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	1	1	2	1	1
Noctiluca scintillans[1] (40911)	1	1	2	2	2
Oxyrrhis marina[1] (34417)	1	1	2	2	1
*Perkinsus marinus**[1] (23654)	2	3	5	4	2
Alexandrium catenella[2] (1108345)	1	1	2	2	2
Amphidinium carterae[2] (36236)	1	3	2	2	1
Crypthecodinium cohnii[2] (48368)	1	1	2	1	2
Gambierdiscus australes[2] (83402)	3	2	3	2	1
Gymnodinium catenatum[2] (83508)	5	1	3	2	3
Heterocapsa triquetra[2] (57789)	0	1	2	0	0
Karenia brevis[2] (93029)	2	1	2	2	1
Karlodinium micrum[2] (58585)	1	1	2	1	1
Kryptoperidinium foliaceum[2] (162643)	4	4	5	2	4
Lingulodinium polyerda[2] (99361)	1	0	2	4	1
Polarella glacialis[2] (59754)	0	1	2	2	1
Scrippsiella hangoei[2] (75339)	1	2	3	0	2
Symbiodinium sp.[2] (44733)	1	2	2	2	0
Togula jolla[2] (42755)	1	1	3	0	1
Favella ehrenbergii[3] (14034)	1	1	3	2	0
Strombidinopsis acuminatum[3] (33155)	1	1	3	0	0
*Eimeria tenella**[4] (8597)	1	1	2	0	1
*Neospora caninum**[4] (7122)	1	1	2	0	1
*Toxoplasma gondii**[4] (8460)	1	1	2	0	1
Vitrella_brassicaformis[5] (23743)	2	0	3	1	1
Chromera velia[6] (19244)	1	1	2	1	2
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	1	1	2	1	1

Chromulina nebulosa[8] (3839)	0	0	2	0	0
*Aureococcus anophagefferens**[9] (11501)	0	1	2	2	1
Schizochytrium aggregatum[10] (13373)	1	1	2	2	1
Chattonella subsalsa[11] (18256)	1	1	2	2	1
Heterosigma akashiwo[11] (32905)	2	1	1	2	0
Dinobryon sp. [12] (18284)	1	1	2	2	1
Ochromonas sp.[12] (16703)	1	1	2	2	2
Pinguiococcus pyrenoidosus[13] (10811)	1	1	3	1	1
Vaucheria litorea[14] (8258)	1	0	2	2	1
Rhizaria					
Bigelowiella longifila (15959)	1	1	3	2	1
Lotharella oceanica (17354)	1	1	2	2	0
Haptophytes					
*Guillardia theta** (24840)	1	1	1	1	0
*Emiliania huxleyi** (39125)	2	2	2	0	2
Imantonia sp. (14800)	1	0	4	0	1
Isochrysis galbana (39744)	2	2	4	0	2
Viridiplantae					
Chlamydomonas euryale (19516)	2	1	2	1	1
Dunaliella tertiolecta (13636)	1	1	2	2	1
Micromonas sp. (8956)	1	1	2	2	1
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	1	3	2	0	0
Acanthoeca-like sp. [16] (21655)	1	1	2	2	1
Rhodosorus marinus[17] (9995)	1	0	5	2	2
Porphyridium aerugineum[17] (8890)	1	1	1	1	1

	unknown function (DUF667) PF05018	**TF elongation Elf1 like PF05129**	**RNA polI specific TIF RRN3 PF05327**	**NmrA-like family PF05368**	**SOH1 PF05669**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	3	1	1	1	1
Noctiluca scintillans[1] (40911)	2	1	1	2	1
Oxyrrhis marina[1] (34417)	2	2	1	10	0
*Perkinsus marinus**[1] (23654)	4	1	1	1	2
Alexandrium catenella[2] (1108345)	3	1	1	6	1
Amphidinium carterae[2] (36236)	2	1	1	2	0

Crypthecodinium cohnii[2] (48368)	4	1	1	0	0
Gambierdiscus australes[2] (83402)	4	1	2	6	1
Gymnodinium catenatum[2] (83508)	7	1	1	8	1
Heterocapsa triquetra[2] (57789)	5	0	1	2	0
Karenia brevis[2] (93029)	5	0	0	3	1
Karlodinium micrum[2] (58585)	5	0	0	0	0
Kryptoperidinium foliaceum[2] (162643)	5	2	3	11	1
Lingulodinium polyerda[2] (99361)	2	1	1	2	0
Polarella glacialis[2] (59754)	11	1	1	3	0
Scrippsiella hangoei[2] (75339)	6	1	1	1	1
Symbiodinium sp.[2] (44733)	3	0	1	4	0
Togula jolla[2] (42755)	2	1	2	2	1
Favella ehrenbergii[3] (14034)	1	1	0	0	0
Strombidinopsis acuminatum[3] (33155)	5	2	0	3	1
*Eimeria tenella**[4] (8597)	2	0	1	0	1
*Neospora caninum**[4] (7122)	2	1	1	0	1
*Toxoplasma gondii**[4] (8460)	2	1	1	0	1
Vitrella_brassicaformis[5] (23743)	0	1	1	6	0
Chromera velia[6] (19244)	1	0	0	2	1
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	1	1	1	2	1
Chromulina nebulosa[8] (3839)	0	0	0	1	0
*Aureococcus anophagefferens**[9] (11501)	3	1	1	1	0
Schizochytrium aggregatum[10] (13373)	3	0	1	0	0
Chattonella subsalsa[11] (18256)	1	0	1	1	0
Heterosigma akashiwo[11] (32905)	2	0	4	2	0
Dinobryon sp. [12] (18284)	1	1	1	1	0
Ochromonas sp.[12] (16703)	3	1	1	6	0
Pinguiococcus pyrenoidosus[13] (10811)	1	1	1	1	0
Vaucheria litorea[14] (8258)	1	1	1	2	0
Rhizaria					
Bigelowiella longifila (15959)	1	1	0	2	1
Lotharella oceanica (17354)	2	1	0	4	0
Haptophytes					

*Guillardia theta** (24840)	3	1	1	7	1
*Emiliania huxleyi** (39125)	2	2	2	8	0
Imantonia sp. (14800)	1	1	0	1	1
Isochrysis galbana (39744)	4	2	3	4	2
Viridiplantae					
Chlamydomonas euryale (19516)	4	0	3	7	0
Dunaliella tertiolecta (13636)	3	1	2	2	0
Micromonas sp. (8956)	2	0	1	2	0
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	1	0	0	1	0
Acanthoeca-like sp. [16] (21655)	2	1	1	3	0
Rhodosorus marinus[17] (9995)	0	0	1	4	1
Porphyridium aerugineum[17] (8890)	0	1	1	1	0

	Tex-like protein N-terminal domain PF09371	**RNA polIII TF IIIC subunit PF09734**	**TF e(y)2 PF10163**	**TF/nuclear export subunit protein 2 PF11262**	**Splicing factor SF3a60 binding domain PF12108**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	1	1	1	1	1
Noctiluca scintillans[1] (40911)	0	0	1	1	1
Oxyrrhis marina[1] (34417)	0	0	1	2	0
*Perkinsus marinus**[1] (23654)	1	2	1	1	0
Alexandrium fundayense[2] (1108345)	0	3	1	1	1
Amphidinium carterae[2] (36236)	0	2	1	1	1
Crypthecodinium cohnii[2] (48368)	0	2	1	1	1
Gambierdiscus australes[2] (83402)	1	2	2	1	1
Gymnodinium catenatum[2] (83508)	1	1	1	1	2
Heterocapsa triquestra[2] (57789)	0	1	1	1	1
Karenia brevis[2] (93029)	0	2	1	1	1
Karlodinium micrum[2] (58585)	0	1	1	1	3
Kruptoperidinium foliaceum[2] (162643)	2	3	3	3	2
Lingulodinium polyerda[2] (99361)	1	2	0	1	1
Polarella glacialis[2] (59754)	1	5	0	2	2
Scrippsiella hangoei[2] (75339)	1	3	1	1	1
Symbiodinium sp.[2] (44733)	1	1	1	1	1
Togula jolla[2] (42755)	0	2	2	1	1

Favella ehrenbergii[3] (14034)	0	0	0	0	0
Strombidinopsis acuminatum[3] (33155)	0	1	4	0	1
*Eimeria tenella**[4] (8597)	0	1	1	1	0
*Neospora caninum**[4] (7122)	0	1	1	1	1
*Toxoplasma gondii**[4] (8460)	0	1	0	1	1
Vitrella brassicaformis[5] (23743)	0	0	1	1	0
Chromera velia[6] (19244)	0	1	0	1	1
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	0	0	1	1	0
Chromulina nebulosa[8] (3839)	0	0	0	1	0
*Aureococcus anophagefferens**[9] (11501)	1	0	0	2	0
Schizochytrium aggregatum[10] (13373)	0	0	1	1	0
Chattonella subsalsa[11] (18256)	0	1	0	1	1
Heterosigma akashiwo[11] (32905)	1	0	1	0	1
Dinobryon sp. [12] (18284)	0	0	1	1	0
Ochromonas sp.[12] (16703)	0	2	0	3	0
Pinguiococcus pyrenoidosus[13] (10811)	0	1	0	1	0
Vaucheria litorea[14] (8258)	0	1	0	1	0
Rhizaria					
Bigelowiella longifila (15959)	0	0	1	0	1
Lotharella oceanica (17354)	0	1	1	0	1
Haptophytes					
*Guillardia theta** (24840)	0	2	2	2	1
*Emiliania huxleyi** (39125)	2	1	2	2	1
Imantonia sp. (14800)	0	0	1	0	0
Isochrysis galbana (39744)	1	3	2	2	1
Viridiplantae					
Chlamydomonas euryale (19516)	0	1	1	1	1
Dunaliella tertiolecta (13636)	0	1	0	2	1
Micromonas sp. (8956)	0	1	0	1	0
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	0	0	1	0	1
Acanthoeca-like sp. [16] (21655)	0	1	1	1	0
Rhodosorus marinus[17] (9995)	0	1	0	1	0
Porphyridium aerugineum[17] (8890)	0	0	0	0	0

	GC-rich DNA-binding factor PF07842	Apoptosis-antagonizing TF, C-terminal PF08164	TATA-binding protein interacting PF08623	FACT complex subunit (SPT16/CDC68) PF08644	TFIIS helical bundle-like domain PF08711
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	1	1	1	2	5
Noctiluca scintillans[1] (40911)	0	1	1	1	12
Oxyrrhis marina[1] (34417)	0	1	1	1	6
*Perkinsus marinus**[1] (23654)	0	1	0	2	8
Alexandrium catenella[2] (1108345)	2	1	1	1	22
Amphidinium carterae[2] (36236)	1	1	1	1	8
Crypthecodinium cohnii[2] (48368)	1	1	1	2	12
Gambierdiscus australes[2] (83402)	1	1	1	3	11
Gymnodinium catenatum[2] (83508)	2	1	4	2	20
Heterocapsa triquetra[2] (57789)	1	1	1	1	3
Karenia brevis[2] (93029)	1	1	2	1	10
Karlodinium micrum[2] (58585)	1	1	3	1	8
Kryptoperidinium foliaceum[2] (162643)	2	1	5	3	16
Lingulodinium polyerda[2] (99361)	1	1	1	1	15
Polarella glacialis[2] (59754)	2	1	1	2	10
Scrippsiella hangoei[2] (75339)	1	1	1	1	17
Symbiodinium sp.[2] (44733)	1	1	1	1	8
Togula jolla[2] (42755)	1	0	2	2	8
Favella ehrenbergii[3] (14034)	0	0	2	2	0
Strombidinopsis acuminatum[3] (33155)	0	0	0	0	5
*Eimeria tenella**[4] (8597)	2	1	0	1	1
*Neospora caninum**[4] (7122)	1	1	0	1	2
*Toxoplasma gondii**[4] (8460)	1	1	1	1	3
Vitrella brassicaformis[5] (23743)	2	1	1	1	1
Chromera velia[6] (19244)	1	1	1	1	2
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	0	2	1	1	5
Chromulina nebulosa[8] (3839)	0	0	0	0	0
*Aureococcus anophagefferens**[9] (11501)	0	0	0	2	4

Schizochytrium aggregatum[10] (13373)	1	1	1	1	3
Chattonella subsalsa[11] (18256)	1	1	1	1	7
Heterosigma akashiwo[11] (32905)	0	1	1	1	2
Dinobryon sp. [12] (18284)	5	1	1	4	3
Ochromonas sp.[12] (16703)	0	1	1	1	4
Pinguiococcus pyrenoidosus[13] (10811)	2	0	1	1	3
Vaucheria litorea[14] (8258)	2	1	1	1	2
Rhizaria					
Bigelowiella longifila (15959)	2	1	2	0	2
Lotharella oceanica (17354)	2	1	2	1	2
Haptophytes					
*Guillardia theta** (24840)	0	1	1	1	2
*Emiliania huxleyi** (39125)	1	2	1	0	5
Imantonia sp. (14800)	0	0	1	1	3
Isochrysis galbana (39744)	3	1	2	2	11
Viridiplantae					
Chlamydomonas euryale (19516)	2	1	1	1	2
Dunaliella tertiolecta (13636)	2	1	1	1	3
Micromonas sp. (8956)	2	0	1	1	5
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	1	0	1	1	3
Acanthoeca-like sp. [16] (21655)	2	1	1	1	2
Rhodosorus marinus[17] (9995)	2	1	1	2	2
Porphyridium aerugineum[17] (8890)	1	0	1	1	0

	Spt4/RpoE2 zinc finger PF06093	**Glucose-regulated metallo-peptidase M90 PF06167**	**Sin3 associated polypeptide p18 (SAP18) PF06487**	**DEAD_2 PF06733**	**HCNGP-like protein PF07818**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])					
*Amoebophrya ceratii**[1] (20969)	1	1	1	10	1
Noctiluca scintillans[1] (40911)	1	0	1	4	1
Oxyrrhis marina[1] (34417)	0	0	2	2	1
*Perkinsus marinus**[1] (23654)	1	0	4	3	1
Alexandrium catenella[2] (1108345)	1	0	1	9	1
Amphidinium carterae[2] (36236)	1	0	1	4	0
Crypthecodinium cohnii[2] (48368)	1	0	0	4	1
Gambierdiscus australes[2]	2	0	2	5	1

(83402)					
Gymnodinium catenatum[2] (83508)	0	0	3	8	1
Heterocapsa triquetra[2] (57789)	1	0	1	3	1
Karenia brevis[2] (93029)	1	0	1	6	1
Karlodinium micrum[2] (58585)	1	0	3	4	1
Kryptoperidinium foliaceum[2] (162643)	2	0	2	6	1
Lingulodinium polyerda[2] (99361)	0	0	1	4	0
Polarella glacialis[2] (59754)	2	0	1	6	0
Scrippsiella hangoei[2] (75339)	1	0	2	5	1
Symbiodinium sp.[2] (44733)	1	0	2	5	1
Togula jolla[2] (42755)	1	0	1	6	1
Favella ehrenbergii[3] (14034)	0	0	0	0	0
Strombidinopsis acuminatum[3] (33155)	1	0	1	0	0
*Eimeria tenella**[4] (8597)	1	0	1	3	1
*Neospora caninum**[4] (7122)	1	0	1	4	1
*Toxoplasma gondii**[4] (8460)	1	0	1	4	1
Vitrella brassicaformis[5] (23743)	0	0	1	0	0
Chromera velia[6] (19244)	0	0	1	2	1
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])					
Thalassiosira pseudonana[7] (17050)	1	0	0	2	0
Chromulina nebulosa[8] (3839)	0	0	0	0	0
*Aureococcus anophagefferens**[9] (11501)	1	0	1	3	0
Schizochytrium aggregatum[10] (13373)	1	0	0	2	0
Chattonella subsalsa[11] (18256)	1	0	1	2	1
Heterosigma akashiwo[11] (32905)	1	0	1	0	1
Dinobryon sp. [12] (18284)	1	0	0	1	0
Ochromonas sp.[12] (16703)	1	0	1	1	1
Pinguiococcus pyrenoidosus[13] (10811)	0	0	0	1	0
Vaucheria litorea[14] (8258)	1	0	0	0	1
Rhizaria					
Bigelowiella longifila (15959)	1	0	1	0	1
Lotharella oceanica (17354)	1	0	1	0	1
Haptophytes					
*Guillardia theta** (24840)	1	0	1	4	1
*Emiliania huxleyi** (39125)	2	0	1	4	0

Imantonia sp. (14800)	1	0	1	1	2
Isochrysis galbana (39744)	2	0	2	6	2
Viridiplantae					
Chlamydomonas euryale (19516)	1	0	1	1	1
Dunaliella tertiolecta (13636)	1	0	1	2	1
Micromonas sp. (8956)	0	0	0	5	0
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]					
Eutreptiella gymnastica[15] (21956)	0	0	1	0	0
Acanthoeca-like sp. [16] (21655)	1	0	2	2	1
Rhodosorus marinus[17] (9995)	1	0	0	5	3
Porphyridium aerugineum[17] (8890)	0	0	0	3	0

	(DUF3819) in CCR4 Not1. PF12842	**Hinge domain PF14327**
Alveolates (basal[1] and core[2] dinoflagellates, ciliates[3], Apicomplexa[4], *Vitrella*[5], *Chromera*[6])		
*Amoebophrya ceratii**[1] (20969)	1	1
Noctiluca scintillans[1] (40911)	1	3
Oxyrrhis marina[1] (34417)	0	2
*Perkinsus marinus**[1] (23654)	3	4
Alexandrium catenella[2] (1108345)	1	2
Amphidinium carterae[2] (36236)	1	3
Crypthecodinium cohnii[2] (48368)	1	3
Gambierdiscus australes[2] (83402)	1	3
Gymnodinium catenatum[2] (83508)	4	3
Heterocapsa triquetra[2] (57789)	0	1
Karenia brevis[2] (93029)	0	3
Karlodinium micrum[2] (58585)	1	3
Kryptoperidinium foliaceum[2] (162643)	3	5
Lingulodinium polyerda[2] (99361)	0	4
Polarella glacialis[2] (59754)	0	1
Scrippsiella hangoei[2] (75339)	1	3
Symbiodinium sp.[2] (44733)	1	2
Togula jolla[2] (42755)	0	2
Favella ehrenbergii[3] (14034)	0	0
Strombidinopsis acuminatum[3] (33155)	6	0

*Eimeria tenella**[4] (8597)	0	0
*Neospora caninum**[4] (7122)	1	1
*Toxoplasma gondii**[4] (8460)	1	1
Vitrella brassicaformis[5] (23743)	2	1
Chromera velia[6] (19244)	0	1
Stramenopiles (diatoms[7], Chrysophyceae[8], Pelagophyceae [9], Labyrinthulids[10], Raphidophyceae[11], Synurophyceae[12], Dictyochophyceae[13], Xanthophyceae[14])		
Thalassiosira pseudonana[7] (17050)	1	0
Chromulina nebulosa[8] (3839)	0	1
*Aureococcus anophagefferens**[9] (11501)	1	0
Schizochytrium aggregatum[10] (13373)	0	1
Chattonella subsalsa[11] (18256)	1	2
Heterosigma akashiwo[11] (32905)	0	1
Dinobryon sp. [12] (18284)	1	0
Ochromonas sp.[12] (16703)	1	0
Pinguiococcus pyrenoidosus[13] (10811)	1	0
Vaucheria litorea[14] (8258)	1	1
Rhizaria		
Bigelowiella longifila (15959)	0	1
Lotharella oceanica (17354)	0	1
Haptophytes		
*Guillardia theta** (24840)	1	1
*Emiliania huxleyi** (39125)	1	0
Imantonia sp. (14800)	0	1
Isochrysis galbana (39744)	2	2
Viridiplantae		
Chlamydomonas euryale (19516)	1	2
Dunaliella tertiolecta (13636)	1	1
Micromonas sp. (8956)	1	0
Euglenids[15], Choanoflagellatea[16], Rhodophyta[17]		
Eutreptiella gymnastica[15] (21956)	1	0
Acanthoeca-like sp. [16] (21655)	1	1
Rhodosorus marinus[17] (9995)	1	2
Porphyridium aerugineum[17] (8890)	1	0

B

Organism	Strain number	Library number	No. of peptides
Noctiluca scintillans	NC	MMETSP0253	40911
Oxyrrhis marina	LB1974	combined MMETSP1424-25-26	34417
*Perkinsus marinus**	ATCC50983	Accession numbers EEQ97081-EER20735	23654
Alexandrium catenella	38-3	Accession numbers GAIV01000001-01108345	1108345
Amphidinium carterae	CCMP1314	combined MMETSP0258-259-398	36236
Crypthecodinium cohnii	Seligo	combined MMETSP0323-326	48368
Gambierdiscus australes	CAWD149	Kohli et al 2015	83402
Gymnodinium catenatum	GC744	MMETSP0784	83508
Heterocapsa triquetra	CCMP448	MMETSP0448	57789
Karenia brevis	Wilson	combined MMETSP0201-02-648-49	93029
Karlodinium micrum	CCMP2283	combined MMETSP1015-16-17	58585
Kryptoperidinium foliaceum	CCMP1326	combined MMETSP0120-121	162643
Lingulodinium polyerda	CCMP1738	combined MMETSP1032-33-34-35	99361
Polarella glacialis	CCMP1383	MMETSP0227	59754
Scrippsiella hangoei	SHTV-5	combined MMETSP0359-360-361	75339
Symbiodinium sp.	Ccmp2430	combined MMETSP1115-16-17	44733
Togula jolla	CCCM725	MMETSP0224	42755
Favella ehrenbergii	fehre1	MMETSP123	14034
Strombidinopsis acuminatum	MMETSP0123	MMETSP0126	33155
*Eimeria tenella**	ETH	toxoDB.com	8597
*Neospora caninum**	NCLIV	toxoDB.com	7122
*Toxoplasma gondii**	GT1	toxoDB.com	8460
Vitrella brassicaformis	CCMP3346	MMETSP1451	23743
Chromera velia	CCMP2878	MMETSP0290	19244
Thalassiosira pseudonana	GMp14c1	combined MMETSP0492-93-94	17050
Chromulina nebulosa	UTEXLB2642	MMETSP1095	3839
*Aureococcus anophagefferens**	CCMP1850	Gobler et al., 2011	11501
Schizochytrium aggregatum	ATCC28209	MMETSP0962-63-64-65	13373
Chattonella subsalsa	CCMP2191	combined MMETSP0947-48-49-50	18256
Heterosigma akashiwo	NB	combined MMETSP0414-15-16	32905
Dinobryon sp.	UTEXLB2267	combined MMETSP0019, MMETSP0020, MMETSP0812	18284
Ochromonas sp.	CCMP1393	combined MMETSP0004-5	16703
Pinguiococcus pyrenoidosus	CCMP2078	MMETSP1160	10811
Vaucheria litorea	CCMP2940	combined MMETSP0945-46	8258
Bigelowiella longifila	CCMP242	MMETSP1359	15959
Lotharella oceanica	CCMP622	MMETSP0040	17354
*Guillardia theta**	CCMP2712	Curtis et al., 2012	24840
*Emiliania huxleyi**	CCMP1516	Read et al., 2013	39125
Imantonia sp.	RCC918	MMETSP1474	14800
Isochrysis galbana	CCMP1323	combined MMETSP0943-44-45	39744
Chlamydomonas euryale	CCMP219	MMETSP0063	19516
Dunaliella tertiolecta	CCMP1320	combined MMETSP1126-28	13636
Micromonas sp.	CCMP2099	combined MMETSP1082, MMETSP1386	8956
Eutreptiella gymnastica	CCMP1594	MMETSP0809-10-11	21956
Acanthoeca-like sp.	10tr	MMETSP0105-06	21655
Rhodosorus marinus	CCMP769	MMETSP0011	9995
Porphyridium aerugineum	SAG1380	MMETSP0313	8890

Table S3. SL detected in the genome of *A ceratii*. (***A***) Predicted genes with 5’ adjacent SL sequences. (***B***) Distribution of full length SL on scaffolds of *A. ceratii*. There are three scaffolds (*Italic, Underline*) which contain more than one SL-sequence. (***C***) Distribution of truncated SL (length of match > 15 nt, accuracy > 70%) in scaffolds of *A. ceratii* genome. (**D**) Position of truncated SL (length of match > 15 nt, accuracy > 70%) in scaffolds of *A. ceratii* genome.

A

Sequence ID	Hit length	Hit start	Hit end	Hit strand	Function
g18355.t1	16.00	4738	4753	Plus	Unknown
g6020.t1	15.00	25	11	Minus	Uncharacterized protein containing DHHC-type Zn finger
g13239.t1	13.00	292	304	Plus	Unknown
g11183.t1	13.00	348	336	Minus	Pentatricopeptide repeat-containing protein At1g62930
g3407.t1	13.00	3292	3304	Plus	Unknown

B

Sequence ID	Hit length	Hit start	Hit end	Hit strand
scaffold1189	21	32704	32724	Minus
scaffold1205	21	43000	43020	Minus
scaffold1270	21	12201	12221	Plus
scaffold1289	21	18168	18188	Plus
scaffold1308	21	40822	40842	Plus
scaffold132	21	41109	41129	Plus
scaffold1320	21	19617	19637	Minus
scaffold1358	21	1354	1374	Plus
scaffold1359	21	11485	11505	Plus
scaffold1385	21	14549	14569	Plus
scaffold1398	21	23563	23583	Plus
scaffold1458	21	309	329	Minus
scaffold1489	21	25315	25335	Plus
scaffold17	21	6576	6596	Minus
scaffold1737	21	4962	4982	Plus
scaffold180	21	268396	268416	Plus
scaffold2066	21	3630	3650	Plus
scaffold2171	21	7973	7993	Plus
scaffold2275	21	4514	4534	Minus
scaffold2289	21	8172	8192	Minus
scaffold233	21	81864	81884	Plus
scaffold2671	21	36	56	Minus
scaffold283	21	211262	211282	Plus
scaffold286	21	34005	34025	Minus
scaffold286	21	25201	25221	Minus
scaffold297	21	106066	106086	Plus
scaffold3000	21	7663	7683	Plus
scaffold307	21	85828	85848	Plus

scaffold309	21	71658	71678	Plus
scaffold309	21	74700	74720	Plus
scaffold329	21	225792	225812	Minus
scaffold3442	21	973	993	Minus
scaffold37	21	271411	271431	Minus
scaffold381	21	4775	4795	Minus
scaffold385	21	8435	8455	Minus
scaffold3942	21	282	302	Minus
scaffold4061	21	1083	1103	Plus
scaffold48	21	17146	17166	Plus
scaffold48	21	215298	215318	Plus
scaffold544	21	63960	63980	Plus
scaffold549	21	2921	2941	Minus
scaffold574	21	85014	85034	Plus
scaffold600	21	7053	7073	Plus
scaffold61	21	114241	114261	Plus
scaffold656	21	1686	1706	Plus
scaffold68	21	131099	131119	Minus
scaffold811	21	61758	61778	Minus
scaffold820	21	12792	12812	Minus
scaffold912	21	56882	56902	Plus
scaffold94	21	33762	33782	Minus
scaffold956	21	41864	41884	Minus
scaffold958	21	1772	1792	Minus

C

Sequence ID	Sequence length	Number of truncated SL*
scaffold61	406243	56
scaffold544	230577	53
scaffold180	286137	47
scaffold329	318215	42
scaffold48	276268	41
scaffold68	248039	40
scaffold37	324791	37
scaffold17	333242	35
scaffold286	196863	35
scaffold283	211282	28
scaffold1358	40547	26
scaffold309	124822	24
scaffold233	133269	23
scaffold94	168647	22
scaffold811	107907	20
scaffold297	150960	18
scaffold381	76577	18
scaffold132	124252	16
scaffold1189	72863	15
scaffold307	88209	13
scaffold1385	64388	12
scaffold574	85034	12

scaffold820	86658	11
scaffold956	84797	11
scaffold1489	38893	10
scaffold958	45976	9
scaffold1205	59972	8
scaffold385	40098	8
scaffold1289	45660	7
scaffold1308	47888	7
scaffold600	54455	7
scaffold1737	13558	6
scaffold549	37577	6
scaffold1398	42967	5
scaffold912	59273	5
scaffold1458	27774	4
scaffold656	38666	4
scaffold1359	45323	3
scaffold2066	8694	3
scaffold2171	8014	3

* positions were provided in D

D

Sequence	Length	Start	End	Match	Length match	Accuracy (%)	Found at strand
scaffold544	230577	211334	211348	CCATTTTGGCTCAAG	15	100	Negative
scaffold385	40098	8420	8434	CCATTTTGGCTCAAG	15	100	Negative
scaffold1489	38893	21759	21773	CCATTTTGGCTCAAG	15	100	Negative
scaffold180	286137	179354	179368	CCATTTGGGCTCAAG	15	93	Positive
scaffold286	196863	31255	31269	CCATTTTGGCTCAAT	15	93	Negative
scaffold286	196863	185923	185937	CCATCTTGGCTCAAG	15	93	Negative
scaffold1308	47888	11001	11015	GCATTTTGGCTCAAG	15	93	Negative
scaffold1489	38893	25395	25409	CCATTTTGGCTGAAG	15	93	Positive
scaffold2066	8694	4637	4651	CCATTTTGGCTCCAG	15	93	Positive
scaffold2066	8694	4609	4623	CCATTTTGGCTCATG	15	93	Positive
scaffold17	333242	292748	292762	CCATTGTGGCTCAAT	15	86	Negative
scaffold329	318215	24682	24696	CCATTTTCGTTCAAG	15	86	Negative
scaffold48	276268	242180	242194	CCGTTTTGCCTCAAG	15	86	Negative
scaffold48	276268	9364	9378	CCATTCTGGCTCGAG	15	86	Positive
scaffold286	196863	31283	31297	ACATTTTGGCTCGAG	15	86	Negative
scaffold286	196863	171104	171118	CCGTTTTAGCTCAAG	15	86	Negative
scaffold286	196863	28762	28776	ACATTTTGGCGCAAG	15	86	Negative
scaffold297	150960	150412	150426	CCAATTTGGCGCAAG	15	86	Negative
scaffold297	150960	126026	126040	CCATTCTGACTCAAG	15	86	Negative
scaffold309	124822	124727	124741	CCTTTTTGGCTCCAG	15	86	Negative
scaffold381	76577	61475	61489	CCATTTTGGATGAAG	15	86	Positive
scaffold1289	45660	35483	35497	CCATTTTTACTCAAG	15	86	Negative
scaffold1358	40547	27947	27961	CCGCTTTGGCTCAAG	15	86	Positive

scaffold1358	40547	33550	33564	CCGCTTTGGCTCAAG	15	86	Positive
scaffold61	406243	206523	206537	CCAGGTAGGCTCAAG	15	80	Positive
scaffold61	406243	301054	301068	CCTGTTTGGCTGAAG	15	80	Negative
scaffold61	406243	118867	118881	CCTTCTTGGGTCAAG	15	80	Positive
scaffold61	406243	149440	149454	GCTTTTTGGCACAAG	15	80	Negative
scaffold61	406243	79206	79220	CCATTTTGGCACCCG	15	80	Negative
scaffold61	406243	370658	370672	CCATGTTGGCCCAAA	15	80	Negative
scaffold61	406243	102710	102724	CCATTTTGGCACCAC	15	80	Positive
scaffold61	406243	383212	383226	CCATGTTGGCGCACG	15	80	Negative
scaffold61	406243	392255	392269	CCATGTTGGCGCACG	15	80	Negative
scaffold17	333242	59946	59960	CCATTTGCGCTGAAG	15	80	Positive
scaffold37	324791	30131	30145	GCATTTCGGCTCAAT	15	80	Positive
scaffold37	324791	145129	145143	CCTTTTTGGCGCAAC	15	80	Negative
scaffold180	286137	7591	7605	CCATGTTAGCACAAG	15	80	Positive
scaffold180	286137	91640	91654	CCATTGTGGTTTAAG	15	80	Negative
scaffold48	276268	217604	217618	ACATTTTGGCTTCAG	15	80	Negative
scaffold48	276268	83004	83018	CCATTTTAGCACCAG	15	80	Negative
scaffold48	276268	211413	211427	CCATCTTCGCTCACG	15	80	Negative
scaffold68	248039	32078	32092	CCAACTTGCCTCAAG	15	80	Negative
scaffold68	248039	92228	92242	CCGTTCTGGCTCCAG	15	80	Negative
scaffold68	248039	34517	34531	CCAACTTGCCTCAAG	15	80	Negative
scaffold68	248039	160746	160760	CCTTTTTGGCTAATG	15	80	Positive
scaffold68	248039	2872	2886	CCAATTAGGCTCCAG	15	80	Negative
scaffold544	230577	138324	138338	CCATTTCTGCGCAAG	15	80	Negative
scaffold544	230577	123516	123530	CCGTGTTGGCTCAAT	15	80	Negative
scaffold544	230577	164342	164356	CATTTTTGTCTCAAG	15	80	Negative
scaffold544	230577	9762	9776	CCAAGTTGCCTCAAG	15	80	Positive
scaffold544	230577	71672	71686	GCATGTTGGCACAAG	15	80	Negative
scaffold283	211282	137409	137423	CCATCTTGGCTTGAG	15	80	Negative
scaffold283	211282	54892	54906	ACGTGTTGGCTCAAG	15	80	Positive
scaffold283	211282	202306	202320	CGATTTCGGCTCGAG	15	80	Positive
scaffold286	196863	76855	76869	CCATTTTGCGCCAAG	15	80	Positive
scaffold286	196863	58975	58989	CTTTTTTGGCTCAGG	15	80	Negative
scaffold94	168647	100137	100151	CCATTTTGGCTGCGG	15	80	Negative
scaffold94	168647	54972	54986	CCATTCTGGGTCAAA	15	80	Positive
scaffold297	150960	48649	48663	CCATTTTGGCCTATG	15	80	Negative
scaffold297	150960	134185	134199	CGATTATGGCTCAAC	15	80	Positive
scaffold233	133269	11202	11216	CGATTTTGGTTGAAG	15	80	Negative
scaffold233	133269	23344	23358	CCATTCTTGCTCAAC	15	80	Negative
scaffold309	124822	110851	110865	CCCTTTTGGTTCACG	15	80	Negative
scaffold132	124252	97975	97989	CTTCTTTGGCTCAAG	15	80	Negative
scaffold132	124252	78188	78202	CCACTTTGGCTCGGG	15	80	Negative
scaffold811	107907	92463	92477	CCACTTTGGCGCGAG	15	80	Negative
scaffold307	88209	84052	84066	CCATTTTCGCGCAGG	15	80	Positive
scaffold820	86658	63136	63150	CCAGTTGGGGTCAAG	15	80	Negative
scaffold820	86658	56178	56192	CCATTTTTGCGTAAG	15	80	Negative
scaffold574	85034	2120	2134	CCATTTTGGGTAGAG	15	80	Negative
scaffold574	85034	6166	6180	CCCTTTTTGCTCCAG	15	80	Negative
scaffold381	76577	52339	52353	CCATTTTAGCTTCAG	15	80	Negative
scaffold1189	72863	23558	23572	CCATTTCTGGTCAAG	15	80	Positive

scaffold1189	72863	24308	24322	CCCTTTTGGCGAAAG	15	80	Negative
scaffold1385	64388	31917	31931	GCATTGTGGCTGAAG	15	80	Negative
scaffold1205	59972	11331	11345	CCATTTTGGCCCACT	15	80	Positive
scaffold1205	59972	11493	11507	CCATTTTGGCCCACT	15	80	Positive
scaffold912	59273	57031	57045	CGATTTTGGCGCAGG	15	80	Negative
scaffold600	54455	19672	19686	GCATTTCGGCTCAAC	15	80	Positive
scaffold1308	47888	14302	14316	CTATCTCGGCTCAAG	15	80	Negative
scaffold1308	47888	17972	17986	CTATTTTGGCAGAAG	15	80	Positive
scaffold1359	45323	35718	35732	GCCTTTTGCCTCAAG	15	80	Positive
scaffold1358	40547	3647	3661	CCCTCGTGGCTCAAG	15	80	Positive
scaffold1358	40547	21296	21310	CGATCTTGACTCAAG	15	80	Negative
scaffold1358	40547	8167	8181	ACATTTTGACTCCAG	15	80	Positive
scaffold385	40098	19780	19794	CCTTCTTGGCTCCAG	15	80	Positive
scaffold1489	38893	25236	25250	ATAATTTGGCTCAAG	15	80	Positive
scaffold1270	21076	17105	17119	CCCGTTTCGCTCAAG	15	80	Negative
scaffold3442	4912	133	147	CGATTTTGGCGCACG	15	80	Positive
scaffold3442	4912	15	29	CGATTTTGGCGCACG	15	80	Positive
scaffold61	406243	361928	361942	ACCTTGTGGCTAAAG	15	73	Positive
scaffold61	406243	123227	123241	CAATTTTGGCCCGAT	15	73	Positive
scaffold61	406243	237477	237491	TGATTCTGGTTCAAG	15	73	Negative
scaffold61	406243	289708	289722	GCATTTTGGCGGAAC	15	73	Negative
scaffold61	406243	301437	301451	ACATTTTGGCTTGGG	15	73	Negative
scaffold61	406243	373566	373580	CAATTCCGGCTCAAC	15	73	Negative
scaffold61	406243	183349	183363	TCATGTTGAGTCAAG	15	73	Negative
scaffold61	406243	399904	399918	GCGTTTCGGTTCAAG	15	73	Positive
scaffold61	406243	226387	226401	TCATCTTCGCTTAAG	15	73	Negative
scaffold61	406243	391051	391065	AGAATTTGGCTGAAG	15	73	Negative
scaffold61	406243	223285	223299	TCATCTTCGCTGAAG	15	73	Negative
scaffold61	406243	3157	3171	GCGATCTGGCTCAAG	15	73	Positive
scaffold61	406243	110202	110216	CTATTTTGGTTCATT	15	73	Positive
scaffold61	406243	257704	257718	TCATTTTGGTTTAAT	15	73	Negative
scaffold61	406243	366538	366552	CCATTTTATCTCAGT	15	73	Positive
scaffold61	406243	354968	354982	CGATTTTGGATCAGT	15	73	Negative
scaffold61	406243	97166	97180	CCCGTTCTGCTCAAG	15	73	Positive
scaffold61	406243	174871	174885	CCATTTTTGTCCAGG	15	73	Negative
scaffold61	406243	183828	183842	CCACTTCTGCTCAAT	15	73	Positive
scaffold61	406243	111303	111317	CCATTTTGCTTCTTG	15	73	Negative
scaffold61	406243	268850	268864	CCACTTTGGTTCTCG	15	73	Positive
scaffold61	406243	40165	40179	CCACCTTGGTTCAAT	15	73	Negative
scaffold61	406243	125121	125135	CCATTTTGACACCGG	15	73	Negative
scaffold61	406243	225000	225014	CCCTTCTGGCGGAAG	15	73	Negative
scaffold61	406243	72488	72502	CCATTTGGACGAAAG	15	73	Negative
scaffold61	406243	218068	218082	CCAGTTCGGCACCAG	15	73	Positive
scaffold61	406243	222746	222760	CCATTCTAGCTCTTG	15	73	Negative
scaffold61	406243	292158	292172	CCATTCGGACTCAAA	15	73	Positive
scaffold61	406243	337936	337950	CCATTATTGAACAAG	15	73	Negative
scaffold61	406243	182664	182678	CCATTATAGCCCGAG	15	73	Negative
scaffold61	406243	141812	141826	CCATCTTGTCTTCAG	15	73	Positive
scaffold61	406243	212412	212426	ACAAATTGGCTCGAG	15	73	Positive
scaffold61	406243	165429	165443	CCAGAATGGCGCAAG	15	73	Negative

scaffold61	406243	35771	35785	CCATCATTGTTCAAG	15	73	Positive
scaffold61	406243	383586	383600	CGATTTTGCCACCAN	15	73	Positive
scaffold61	406243	392638	392652	CGATTTTGCCACCAG	15	73	Positive
scaffold61	406243	156027	156041	CGATTTGGGGTCTAG	15	73	Negative
scaffold61	406243	4244	4258	CCTTTTTGCTCCAAG	15	73	Negative
scaffold61	406243	33781	33795	CCGTATTGGAGCAAG	15	73	Positive
scaffold61	406243	398010	398024	CCCATTTGACTCCAG	15	73	Positive
scaffold61	406243	361972	361986	CCTTTCAGGCACAAG	15	73	Negative
scaffold61	406243	207247	207261	CCGTTTTTGCGCATG	15	73	Negative
scaffold61	406243	286794	286808	CCGGTCGGGCTCAAG	15	73	Negative
scaffold61	406243	304518	304532	CCGTTTTGTCTCATC	15	73	Positive
scaffold61	406243	5964	5978	CCGTTTTGCCTCCTG	15	73	Negative
scaffold61	406243	222462	222476	CCACCATGGCTCCAG	15	73	Negative
scaffold61	406243	338334	338348	CCACCATGCCTCAAG	15	73	Positive
scaffold17	333242	248994	249008	CCTTTTTCGTTCCAG	15	73	Positive
scaffold17	333242	258466	258480	CAGTTTTAGCTCCAG	15	73	Positive
scaffold17	333242	325525	325539	CTGTTTTTGCTCACG	15	73	Positive
scaffold17	333242	220969	220983	CAGTTTTGGATCAGG	15	73	Negative
scaffold17	333242	6531	6545	CCTTTTTTGCGAAAG	15	73	Negative
scaffold17	333242	90238	90252	CCATTTCTGATGAAG	15	73	Negative
scaffold17	333242	138535	138549	ACATATTAGCGCAAG	15	73	Positive
scaffold17	333242	258268	258282	CCATCTTGGCTCGTC	15	73	Positive
scaffold17	333242	52561	52575	CCATTCCGGCTCCAC	15	73	Negative
scaffold17	333242	317780	317794	CCAGTTCGGATCACG	15	73	Positive
scaffold17	333242	78321	78335	CCATTCTGTCTCGAT	15	73	Negative
scaffold17	333242	137030	137044	CGGTTTTGGATCATG	15	73	Positive
scaffold17	333242	313481	313495	CCAGTGTGGCGCAGG	15	73	Negative
scaffold17	333242	85817	85831	CCACCTTGGGTCACG	15	73	Positive
scaffold17	333242	116643	116657	CCAGTCTGGCCCAAC	15	73	Negative
scaffold17	333242	327152	327166	CCAGTACGGCACAAG	15	73	Positive
scaffold17	333242	51344	51358	CCGTTTTGGTTGCAG	15	73	Negative
scaffold17	333242	205644	205658	CCGTTTTAGAGCAAG	15	73	Negative
scaffold17	333242	245338	245352	CCATTTTCGCAGCAG	15	73	Negative
scaffold17	333242	64010	64024	CGATTTGAGCCCAAG	15	73	Positive
scaffold17	333242	249013	249027	CCATTCTGTTTCGAG	15	73	Positive
scaffold17	333242	221248	221262	GCAGATTGCCTCAAG	15	73	Negative
scaffold17	333242	250820	250834	ACCTCTTGCCTCAAG	15	73	Negative
scaffold17	333242	114047	114061	TGGGTTTGGCTCAAG	15	73	Positive
scaffold17	333242	239375	239389	GCATCTTGCCTCAAT	15	73	Negative
scaffold17	333242	150499	150513	TCAGTTTGCCGCAAG	15	73	Negative
scaffold17	333242	300832	300846	TCATTTAGGCTGCAG	15	73	Negative
scaffold17	333242	173552	173566	GCGTTTCGGCTGAAG	15	73	Negative
scaffold17	333242	93877	93891	ACTTTTCGGCTCTAG	15	73	Negative
scaffold17	333242	116235	116249	TAATTTTGGCCCAAC	15	73	Negative
scaffold17	333242	79604	79618	GCTTTTGGGCTCACG	15	73	Negative
scaffold17	333242	49231	49245	TCGTTTTGGTTGAAG	15	73	Positive
scaffold17	333242	141643	141657	TCATAGTGGCTCACG	15	73	Positive
scaffold37	324791	127637	127651	ACTTTTTGTCGCAAG	15	73	Negative
scaffold37	324791	192865	192879	CCAGTTTGCATCTAG	15	73	Negative
scaffold37	324791	27408	27422	CAATTTTGCGGCAAG	15	73	Negative

scaffold37	324791	130639	130653	CCACATTTGCTCCAG	15	73	Positive
scaffold37	324791	129852	129866	CCATTCTGGACCACG	15	73	Negative
scaffold37	324791	298822	298836	GCCTTTTGGCTGCAG	15	73	Negative
scaffold37	324791	196273	196287	CGATTTTGCCGGAAG	15	73	Negative
scaffold37	324791	120543	120557	CGATTTGGGCATAAG	15	73	Negative
scaffold37	324791	205130	205144	CGTTTTTGGCTGAGG	15	73	Negative
scaffold37	324791	268616	268630	CCATGGTGGCTGAGG	15	73	Negative
scaffold37	324791	256812	256826	CCATTTCGGCGCACT	15	73	Positive
scaffold37	324791	172851	172865	CCATTCTGTCGCAGG	15	73	Positive
scaffold37	324791	266204	266218	CGATTTCGGATGAAG	15	73	Negative
scaffold37	324791	233478	233492	CGATTTCGCCTCTAG	15	73	Negative
scaffold37	324791	46645	46659	CCCTTTTTGCTCACC	15	73	Positive
scaffold37	324791	299216	299230	CCCTTTCGTCTCAAT	15	73	Negative
scaffold37	324791	15604	15618	CCGTTTTTGCACGAG	15	73	Positive
scaffold37	324791	84594	84608	CCTGTTTGATTCAAG	15	73	Positive
scaffold37	324791	33343	33357	CCAATTTGGCGGAAA	15	73	Positive
scaffold37	324791	274051	274065	CGACTTTCGCACAAG	15	73	Negative
scaffold37	324791	288066	288080	CCGTTTTGGCTCGCT	15	73	Positive
scaffold37	324791	59661	59675	CCCTTTTTCCTCAAC	15	73	Negative
scaffold37	324791	292807	292821	CCTTTTTTGCTCCAA	15	73	Negative
scaffold37	324791	280175	280189	CCCTTTGGTCACAAG	15	73	Positive
scaffold37	324791	143121	143135	AAATTTTGGCTGATG	15	73	Positive
scaffold37	324791	324501	324515	TCATCTTGACTCACG	15	73	Positive
scaffold37	324791	92009	92023	GCATTTTGGCGCTTG	15	73	Negative
scaffold37	324791	59562	59576	CTTTTTTGGCGCAAC	15	73	Positive
scaffold37	324791	248905	248919	CCACTTTGGCACATC	15	73	Positive
scaffold37	324791	212974	212988	CCATTTTGGAGCCAT	15	73	Positive
scaffold37	324791	115816	115830	CTATTTCGTCTCTAG	15	73	Negative
scaffold37	324791	84046	84060	ACGATTCGGCTCAAG	15	73	Negative
scaffold37	324791	80186	80200	CCGCTTTGGCTCGAC	15	73	Negative
scaffold37	324791	299739	299753	GCTTTGTGGCTGAAG	15	73	Positive
scaffold37	324791	297390	297404	AGATTTTGGCGCCAG	15	73	Positive
scaffold329	318215	100826	100840	CCATTCTGATTAAAG	15	73	Positive
scaffold329	318215	286856	286870	CCATTTTGCCTCCTC	15	73	Positive
scaffold329	318215	286149	286163	CCATTTTGCCACCAT	15	73	Negative
scaffold329	318215	284929	284943	CGTTTTCAGCTCAAG	15	73	Negative
scaffold329	318215	15548	15562	CCATTTTTGGACGAG	15	73	Negative
scaffold329	318215	20536	20550	CCATTTTTGGCCCAG	15	73	Positive
scaffold329	318215	173571	173585	CTATTTTGGCGCTTG	15	73	Negative
scaffold329	318215	317375	317389	TGAATTTCGCTCAAG	15	73	Negative
scaffold329	318215	119049	119063	CCAGTTGGGCTCGAC	15	73	Negative
scaffold329	318215	309210	309224	CCAGGTTGGCTGAGG	15	73	Negative
scaffold329	318215	69402	69416	CCACTTAGGCACGAG	15	73	Negative
scaffold329	318215	175678	175692	CCATTTTTGCCAATG	15	73	Negative
scaffold329	318215	142454	142468	CCACTTCCGCTGAAG	15	73	Positive
scaffold329	318215	196069	196083	CGATTCTGGTTCAAC	15	73	Positive
scaffold329	318215	181552	181566	CCAGTTTGATGCAAG	15	73	Negative
scaffold329	318215	285145	285159	CGATCTTGGTTCAAT	15	73	Negative
scaffold329	318215	278659	278673	CCCGATAGGCTCAAG	15	73	Negative
scaffold329	318215	313695	313709	TCGTTTTCGCCCAAG	15	73	Positive

scaffold329	318215	55368	55382	CCAATCTGGCGCACG	15	73	Positive
scaffold329	318215	74258	74272	CCTTTCAGGCTCCAG	15	73	Positive
scaffold329	318215	218401	218415	GCATTTTGATTCCAG	15	73	Positive
scaffold329	318215	33327	33341	GCATGTTTGCTCTAG	15	73	Positive
scaffold329	318215	32133	32147	CCCTTTTGCTTCACG	15	73	Positive
scaffold329	318215	280476	280490	CCCTTTTGTCCCAAC	15	73	Positive
scaffold329	318215	141906	141920	CCGTTTTGGAGCAGG	15	73	Negative
scaffold329	318215	199056	199070	CGTTTTTGGCGCAAC	15	73	Positive
scaffold329	318215	307437	307451	CCCTTTTGGCACATT	15	73	Positive
scaffold329	318215	174328	174342	TCATTATGGTGCAAG	15	73	Positive
scaffold329	318215	209500	209514	TCACTTTTGCTCGAG	15	73	Positive
scaffold329	318215	137838	137852	CCAATTTGGCTGCTG	15	73	Negative
scaffold329	318215	172450	172464	CTCTTTTAGCTCCAG	15	73	Positive
scaffold329	318215	164358	164372	CCGGTATGACTCAAG	15	73	Negative
scaffold329	318215	63883	63897	CAATTTTCGCTGAAT	15	73	Negative
scaffold329	318215	92251	92265	CCAATACGGTTCAAG	15	73	Negative
scaffold329	318215	203995	204009	GCAATTCGGCTGAAG	15	73	Positive
scaffold329	318215	42001	42015	CCTTCTTGGCCCCAG	15	73	Negative
scaffold329	318215	21670	21684	CCACGGTGGCGCAAG	15	73	Negative
scaffold329	318215	200318	200332	GCAATTCGGCTGAAG	15	73	Positive
scaffold329	318215	184848	184862	CGATTTCGACCCAAG	15	73	Negative
scaffold329	318215	219120	219134	CCATTTGGAGGCAAG	15	73	Negative
scaffold329	318215	63605	63619	GCATGGTGGCTCGAG	15	73	Positive
scaffold180	286137	143134	143148	CCAGTTCGGCTTTAG	15	73	Positive
scaffold180	286137	144171	144185	CCCTTTTGGCGCACC	15	73	Positive
scaffold180	286137	21103	21117	CCATTCTAGCTCGTG	15	73	Positive
scaffold180	286137	30750	30764	AGATTTTGGCACAAA	15	73	Positive
scaffold180	286137	34264	34278	TCATTATCGCTCAAC	15	73	Negative
scaffold180	286137	79209	79223	CCTCTTCTGCTCAAG	15	73	Positive
scaffold180	286137	191225	191239	CCACTTCTGCTCGAG	15	73	Positive
scaffold180	286137	51048	51062	CCGTTTTGGCGGAGG	15	73	Positive
scaffold180	286137	118387	118401	CAATTTTGTCACACG	15	73	Positive
scaffold180	286137	70794	70808	GCATTTTCGCGCACG	15	73	Positive
scaffold180	286137	159386	159400	CCGTTTTGTCAGAAG	15	73	Positive
scaffold180	286137	202911	202925	CCATTTTCTTTGAAG	15	73	Positive
scaffold180	286137	176345	176359	CCAAATTCGATCAAG	15	73	Positive
scaffold180	286137	171421	171435	GCGTCTTTGCTCAAG	15	73	Negative
scaffold180	286137	159421	159435	CCGTCTTGGCTCGCG	15	73	Negative
scaffold180	286137	128438	128452	CCGTTTTTGCGGAAG	15	73	Positive
scaffold180	286137	274476	274490	CCCTTTTGGGTCGCG	15	73	Positive
scaffold180	286137	13572	13586	CCAGGCTGGCGCAAG	15	73	Positive
scaffold180	286137	273783	273797	CCTGTTTTGCTCCAG	15	73	Negative
scaffold180	286137	74157	74171	CGATTTCAGATCAAG	15	73	Positive
scaffold180	286137	202011	202025	CCATTTAGTCTCCTG	15	73	Positive
scaffold180	286137	84411	84425	GCATTTTTGTTCGAG	15	73	Negative
scaffold180	286137	102869	102883	CCGTTTTCGGTCGAG	15	73	Positive
scaffold180	286137	106845	106859	GCATTTTTGCTCACC	15	73	Positive
scaffold180	286137	46788	46802	CCATTTTGGCACTTT	15	73	Negative
scaffold180	286137	262274	262288	ACATTTTGACTCTTG	15	73	Positive
scaffold180	286137	216538	216552	TGAGTTCGGCTCAAN	15	73	Positive

scaffold180	286137	27808	27822	ACATGTTGGCTAATG	15	73	Negative
scaffold180	286137	170749	170763	CCCGCGTGGCTCAAG	15	73	Negative
scaffold180	286137	180256	180270	CCACGTAGGCTCACG	15	73	Positive
scaffold180	286137	120639	120653	CCATCTTGACTGACG	15	73	Positive
scaffold180	286137	105302	105316	CCATAGTTGATCAAG	15	73	Negative
scaffold180	286137	225580	225594	ACCTTTTGGCACCAG	15	73	Negative
scaffold180	286137	268126	268140	CTATTTTGGCACACA	15	73	Positive
scaffold180	286137	2609	2623	CCAGTTTGGGAGAAG	15	73	Positive
scaffold180	286137	80155	80169	CCATCTTCGTTCTAG	15	73	Negative
scaffold180	286137	76565	76579	TCATTTTCGTTCGAG	15	73	Negative
scaffold180	286137	159208	159222	CCGATTTTGCGCAAG	15	73	Positive
scaffold180	286137	252001	252015	CTACTCTGGCTAAAG	15	73	Negative
scaffold180	286137	170694	170708	GTATTATCGCTCAAG	15	73	Positive
scaffold180	286137	264084	264098	CCACTTCGGGTCCAG	15	73	Negative
scaffold180	286137	78421	78435	CCATGTTGAAGCAAG	15	73	Positive
scaffold180	286137	267111	267125	CGTTTTTGCCTCAAA	15	73	Negative
scaffold180	286137	6540	6554	CCTTTTTGGTTTAGG	15	73	Positive
scaffold48	276268	63356	63370	CTATTGTGGCTCTAC	15	73	Negative
scaffold48	276268	243918	243932	CCAGTTTGCGTCCAG	15	73	Positive
scaffold48	276268	148415	148429	CCATGTTGCGCCAAG	15	73	Negative
scaffold48	276268	143111	143125	CTAACTTCGCTCAAG	15	73	Negative
scaffold48	276268	185508	185522	GCATTCTTGCTGAAG	15	73	Positive
scaffold48	276268	243539	243553	CCCTTTCGGTTGAAG	15	73	Negative
scaffold48	276268	232930	232944	GCATTTTCGATCAAA	15	73	Negative
scaffold48	276268	242763	242777	CCACTTGGGCTCTAC	15	73	Negative
scaffold48	276268	163751	163765	CGATGTTGGCAGAAG	15	73	Positive
scaffold48	276268	66153	66167	CCACTTCTGCTCCAG	15	73	Negative
scaffold48	276268	47266	47280	CCTTTTTGACTCTGG	15	73	Negative
scaffold48	276268	57187	57201	CCTTTTAGGGTGAAG	15	73	Positive
scaffold48	276268	135937	135951	GCTTTTGGGCTCCAG	15	73	Positive
scaffold48	276268	249142	249156	CCCGTTTTGCTCACG	15	73	Negative
scaffold48	276268	23173	23187	CCTTTTTTGCTCGTG	15	73	Positive
scaffold48	276268	163592	163606	TCATTCACGCTCAAG	15	73	Negative
scaffold48	276268	210068	210082	CCCTTTTGACTGGAG	15	73	Negative
scaffold48	276268	94190	94204	CCCTTTTGGCGCCCG	15	73	Positive
scaffold48	276268	162864	162878	CCGATTTGGCTGCAG	15	73	Positive
scaffold48	276268	146322	146336	CCGTTTGAGCTGAAG	15	73	Positive
scaffold48	276268	223125	223139	ACATGTTGCATCAAG	15	73	Negative
scaffold48	276268	169579	169593	TCCTTTTGGCTCTAC	15	73	Positive
scaffold48	276268	71792	71806	TCGTTTTCGTTCAAG	15	73	Negative
scaffold48	276268	179678	179692	CCAGTTGGCCGCAAG	15	73	Negative
scaffold48	276268	185584	185598	TCGTTTTCGCTCACG	15	73	Negative
scaffold48	276268	159600	159614	CCATTCTCGGTCATG	15	73	Positive
scaffold48	276268	167042	167056	CGATTTGTGCACAAG	15	73	Positive
scaffold48	276268	38040	38054	TCATTTTGCATCAAC	15	73	Negative
scaffold48	276268	143670	143684	CACTTTTGGCGAAAG	15	73	Negative
scaffold48	276268	178852	178866	CCATTTTGGATGGAC	15	73	Negative
scaffold48	276268	165539	165553	CCATTTTGGAGGAGG	15	73	Positive
scaffold48	276268	268574	268588	CGATTGCCGCTCAAG	15	73	Negative
scaffold48	276268	125222	125236	CGATTCTGCGTCAAG	15	73	Negative

scaffold48	276268	100386	100400	CCATTTCTGCTTCAG	15	73	Negative
scaffold48	276268	144330	144344	TCCTTTTGGTTCTAG	15	73	Positive
scaffold48	276268	79658	79672	CCATTTTCGTGCACG	15	73	Positive
scaffold68	248039	164141	164155	ACATTTTCGCTCTCG	15	73	Positive
scaffold68	248039	115948	115962	ACATTGTGGTTCAAA	15	73	Positive
scaffold68	248039	210325	210339	CCATTTTAGCGCTTG	15	73	Negative
scaffold68	248039	86336	86350	CCCTTCAGGCTCATG	15	73	Positive
scaffold68	248039	112484	112498	CCATGTGGGGTCAAC	15	73	Positive
scaffold68	248039	225580	225594	CCATTTTTGCTCTTC	15	73	Positive
scaffold68	248039	829	843	CCATTATGGCTTGAC	15	73	Negative
scaffold68	248039	94302	94316	CGATTTTGGCTAAGC	15	73	Negative
scaffold68	248039	204684	204698	CCTTTCAGGCTCCAG	15	73	Negative
scaffold68	248039	109179	109193	CCTACTTGGCCCAAG	15	73	Positive
scaffold68	248039	29973	29987	CGTTTTTGGTGCAAG	15	73	Positive
scaffold68	248039	141031	141045	CCGTTTGAGCGCAAG	15	73	Positive
scaffold68	248039	7526	7540	CCATGTCGGCTATAG	15	73	Positive
scaffold68	248039	120393	120407	CCGGTTTGGCGCACG	15	73	Negative
scaffold68	248039	80846	80860	CCATTTTCGCTCCTC	15	73	Positive
scaffold68	248039	134569	134583	CCATTTTGCAGCATG	15	73	Positive
scaffold68	248039	142317	142331	CCATTTTGCATCGTG	15	73	Negative
scaffold68	248039	183331	183345	CCCCTTTGGCGCCAG	15	73	Negative
scaffold68	248039	168918	168932	CCATTTTGTCTGCCG	15	73	Positive
scaffold68	248039	49821	49835	CCATTTTTACACAAC	15	73	Positive
scaffold68	248039	212252	212266	CCATTTGGGCACGGG	15	73	Negative
scaffold68	248039	48776	48790	TCATTGTGGCCCAGG	15	73	Negative
scaffold68	248039	226318	226332	ACTTTTCTGCTCAAG	15	73	Negative
scaffold68	248039	193175	193189	TTACTTTGGCTCGAG	15	73	Negative
scaffold68	248039	16702	16716	TGTTTTTAGCTCAAG	15	73	Negative
scaffold68	248039	216836	216850	TGATTTCGCCTCAAG	15	73	Negative
scaffold68	248039	98075	98089	CAATTCTTGCTCCAG	15	73	Negative
scaffold68	248039	220008	220022	TCTTTTTGGCTCGCG	15	73	Positive
scaffold68	248039	169589	169603	TCATTTTGGCGAACG	15	73	Positive
scaffold68	248039	61602	61616	TCATTTTGGATGATG	15	73	Positive
scaffold68	248039	155532	155546	CCATCTTTGTTGAAG	15	73	Positive
scaffold68	248039	190853	190867	CTATTTTTGCTCCAT	15	73	Positive
scaffold68	248039	13875	13889	CTTTTTTGGGTGAAG	15	73	Negative
scaffold68	248039	219588	219602	CCATATTTGCCCAAC	15	73	Positive
scaffold68	248039	148266	148280	CCATATTGATTGAAG	15	73	Positive
scaffold544	230577	213296	213310	CAATTTTTGCGCATG	15	73	Negative
scaffold544	230577	80382	80396	CAATATTCGCACAAG	15	73	Positive
scaffold544	230577	46858	46872	CCACTTTTCCTCAAT	15	73	Positive
scaffold544	230577	222873	222887	CCAATTTTGCTTCAG	15	73	Positive
scaffold544	230577	172686	172700	ACTTTTTGGCGCCAG	15	73	Positive
scaffold544	230577	63689	63703	CCAGTGTGGCACAAT	15	73	Positive
scaffold544	230577	211557	211571	CCAGTTTGGCTCGTC	15	73	Negative
scaffold544	230577	202826	202840	CCACTCCGGCTCGAG	15	73	Positive
scaffold544	230577	94626	94640	CCACTCCGGCTCAAT	15	73	Negative
scaffold544	230577	6952	6966	CAATTATGGCGGAAG	15	73	Negative
scaffold544	230577	158496	158510	CCACCTCGGCTTAAG	15	73	Negative
scaffold544	230577	120265	120279	CCGTCTTGGCTCGAC	15	73	Positive

scaffold544	230577	157388	157402	CCATTTTGAATCGCG	15	73	Negative
scaffold544	230577	99813	99827	TTCTTTTGCCTCAAG	15	73	Negative
scaffold544	230577	199273	199287	CCGTTTTTGCGCAGG	15	73	Negative
scaffold544	230577	184940	184954	GCATTTTTGCTCGAA	15	73	Positive
scaffold544	230577	69586	69600	GCATTTTGTCTCACC	15	73	Positive
scaffold544	230577	211825	211839	CCGTTTTGGGCAAAG	15	73	Negative
scaffold544	230577	185895	185909	CCGTTTTGGACCACG	15	73	Positive
scaffold544	230577	23690	23704	CCGTTTTGCCCGAAG	15	73	Negative
scaffold544	230577	156791	156805	CCATTTTCATACAAG	15	73	Positive
scaffold544	230577	13271	13285	CCATTTTTGATCAGC	15	73	Negative
scaffold544	230577	68912	68926	TCAGGTTGGATCAAG	15	73	Negative
scaffold544	230577	176939	176953	CGATTTTGGATCAGA	15	73	Positive
scaffold544	230577	2034	2048	GCATCTTGGCTAAAT	15	73	Negative
scaffold544	230577	177235	177249	GCATCCGGGCTCAAG	15	73	Positive
scaffold544	230577	139644	139658	GCAGTTTGACTCGAG	15	73	Positive
scaffold544	230577	203970	203984	CCGATCTGGCTCACG	15	73	Negative
scaffold544	230577	85250	85264	GCAGTTGGGCTCAAA	15	73	Positive
scaffold544	230577	183579	183593	GATTTTTGTCTCAAG	15	73	Negative
scaffold544	230577	26186	26200	GAGTTTTGGCTCATG	15	73	Negative
scaffold544	230577	106919	106933	CTTTTTTGGCTAAGG	15	73	Positive
scaffold544	230577	200084	200098	CCTTTTTTTGCTCCCG	15	73	Positive
scaffold544	230577	211821	211835	CCATCTTTGCCCAAA	15	73	Positive
scaffold544	230577	83008	83022	CCATGATGGCTCATC	15	73	Negative
scaffold544	230577	165983	165997	CCATGCTGGCGCACG	15	73	Negative
scaffold544	230577	60239	60253	CCATGTTGATACAAG	15	73	Negative
scaffold544	230577	90612	90626	CGATTATCGCTGAAG	15	73	Positive
scaffold544	230577	69010	69024	TCAGGTTGGCTCAGG	15	73	Negative
scaffold544	230577	91819	91833	CCTTTTTGGAAGAAG	15	73	Positive
scaffold544	230577	215333	215347	CCTTTTGGGTTCAAA	15	73	Positive
scaffold544	230577	95593	95607	CGCTTTTTGCTCAGG	15	73	Negative
scaffold544	230577	23017	23031	CCATTCTTGCGGAAG	15	73	Positive
scaffold544	230577	54225	54239	CTATTTTGCGTCAAT	15	73	Positive
scaffold544	230577	20385	20399	CGCTTTCGGCTCCAG	15	73	Negative
scaffold544	230577	85743	85757	TCATATTGGCACAAT	15	73	Positive
scaffold544	230577	89697	89711	CGCTTTTTGCTCAGG	15	73	Negative
scaffold283	211282	205076	205090	CCATTTTGCTGCAGG	15	73	Positive
scaffold283	211282	116503	116517	CCATTTTGTTTCGAA	15	73	Negative
scaffold283	211282	201697	201711	CCATTTTGGAGCGAC	15	73	Negative
scaffold283	211282	12645	12659	ACATTTTCGTTCAAC	15	73	Negative
scaffold283	211282	14438	14452	CCATTTTGGGGCGGG	15	73	Negative
scaffold283	211282	165017	165031	TCATGTTGGAACAAG	15	73	Positive
scaffold283	211282	147264	147278	GAATTATGGCTCAAC	15	73	Negative
scaffold283	211282	6638	6652	CTTCTTTAGCTCAAG	15	73	Negative
scaffold283	211282	111915	111929	ACATTTTGAGTCATG	15	73	Negative
scaffold283	211282	95251	95265	CAATGTTGGCGCGAG	15	73	Positive
scaffold283	211282	48214	48228	CCATCGTCGCTCCAG	15	73	Negative
scaffold283	211282	140596	140610	ACAATTGGGTTCAAG	15	73	Positive
scaffold283	211282	48908	48922	GCATTTCGGCGCGAG	15	73	Positive
scaffold283	211282	26581	26595	CCATTCTTGCTGACG	15	73	Negative
scaffold283	211282	199806	199820	CAATTTCGGCTCGTG	15	73	Positive

scaffold283	211282	13755	13769	CGAGTTTCGCTGAAG	15	73	Positive
scaffold283	211282	183884	183898	CCGTTTTGGTTCTCG	15	73	Negative
scaffold283	211282	31689	31703	CGATGTTCGCTCATG	15	73	Negative
scaffold283	211282	167539	167553	CCATGTTGGTACATG	15	73	Negative
scaffold283	211282	78817	78831	GCCTTCTGGCTCCAG	15	73	Negative
scaffold283	211282	172244	172258	TCATGTTGGATCGAG	15	73	Positive
scaffold283	211282	63603	63617	CCATTGTCGCTCGGG	15	73	Positive
scaffold283	211282	164909	164923	GCATTTTGGCTCCGT	15	73	Positive
scaffold283	211282	122592	122606	AGACTGTGGCTCAAG	15	73	Positive
scaffold283	211282	20018	20032	CCCTTTTGTCACCAG	15	73	Negative
scaffold286	196863	192008	192022	CCAATTCGACTCACG	15	73	Negative
scaffold286	196863	190543	190557	CCGTTTTGTCGCATG	15	73	Negative
scaffold286	196863	175097	175111	CCCTTTTGCATCCAG	15	73	Positive
scaffold286	196863	75043	75057	TCATTCTGGATGAAG	15	73	Negative
scaffold286	196863	112717	112731	CCACGTTGGCACTAG	15	73	Negative
scaffold286	196863	16176	16190	CCGTTTCGTCTCCAG	15	73	Positive
scaffold286	196863	106026	106040	CCATTCTGACCCGAG	15	73	Positive
scaffold286	196863	48375	48389	TCCTTCTGGCTCACG	15	73	Positive
scaffold286	196863	127942	127956	CCAGTCTGGCCCAAC	15	73	Positive
scaffold286	196863	133109	133123	CCATTTCGAATCAAA	15	73	Positive
scaffold286	196863	24674	24688	CCATTGTGGGCCAAC	15	73	Negative
scaffold286	196863	183682	183696	CGATTCTCGCGCAAG	15	73	Positive
scaffold286	196863	28081	28095	TCGTGTTGGCTGAAG	15	73	Negative
scaffold286	196863	42455	42469	CGATTTTCGCACAAT	15	73	Negative
scaffold286	196863	29606	29620	CGCTTTTGCCTCAAT	15	73	Positive
scaffold286	196863	181858	181872	CCATGTTGCTTCCAG	15	73	Positive
scaffold286	196863	69712	69726	AAATACTGGCTCAAG	15	73	Negative
scaffold286	196863	19130	19144	TTATTTTGCCTCATG	15	73	Negative
scaffold286	196863	124695	124709	CGATTTTTGATCCAG	15	73	Positive
scaffold286	196863	45967	45981	CCATTTTGAATCAGA	15	73	Positive
scaffold286	196863	56251	56265	CCATTTTTGCGCTAT	15	73	Positive
scaffold286	196863	83983	83997	CCTTTTCTGCTCCAG	15	73	Negative
scaffold286	196863	101875	101889	CGATTCGGGCTAAAG	15	73	Positive
scaffold286	196863	147927	147941	CCATTTTCGATCTAT	15	73	Positive
scaffold286	196863	100195	100209	GTCGTTTGGCTCAAG	15	73	Negative
scaffold286	196863	169426	169440	CCATTTTGGATTTAC	15	73	Positive
scaffold286	196863	158673	158687	CGAGTTTGACTCGAG	15	73	Positive
scaffold286	196863	9314	9328	CCAGCTTGGTACAAG	15	73	Positive
scaffold94	168647	29533	29547	GCGTTCTGGATCAAG	15	73	Negative
scaffold94	168647	5205	5219	GCATTCTGGCTCATC	15	73	Positive
scaffold94	168647	88124	88138	CTGTTTTGGTTTAAG	15	73	Negative
scaffold94	168647	112624	112638	CCATGTTGGGTTCAG	15	73	Positive
scaffold94	168647	81051	81065	CTATTTTTGCTGATG	15	73	Positive
scaffold94	168647	85717	85731	CTACTTTGGTCCAAG	15	73	Negative
scaffold94	168647	116182	116196	CCATTTTTGCCGTAG	15	73	Negative
scaffold94	168647	66122	66136	CGATTTTCGCTCCAC	15	73	Negative
scaffold94	168647	92727	92741	CCATTTTCGATCGAA	15	73	Positive
scaffold94	168647	50429	50443	CCATTTTTGCTGCAT	15	73	Positive
scaffold94	168647	8798	8812	CCCCTTTGGCTCTGG	15	73	Positive
scaffold94	168647	102734	102748	CCCTCTTGGCTACAG	15	73	Negative

scaffold94	168647	95391	95405	ACGTTTTTGCTCCAG	15	73	Positive
scaffold94	168647	62296	62310	CCATTTCGGCTCTCT	15	73	Positive
scaffold94	168647	116705	116719	CCGCCTTGGCTCAGG	15	73	Negative
scaffold94	168647	113576	113590	CATTTTTTTCTCAAG	15	73	Positive
scaffold94	168647	91711	91725	CCATTGTAGCTGACG	15	73	Positive
scaffold94	168647	95472	95486	TCGTTTTGGCTCAGC	15	73	Positive
scaffold94	168647	116527	116541	CCATCTGGTCCCAAG	15	73	Positive
scaffold94	168647	35272	35286	CCATTTTGGAAGATG	15	73	Positive
scaffold297	150960	67697	67711	CCAGTTTGGCTGCTG	15	73	Negative
scaffold297	150960	12464	12478	CGGTTTTTGCTCCAG	15	73	Negative
scaffold297	150960	26105	26119	GCATTCTGCATCAAG	15	73	Negative
scaffold297	150960	13161	13175	CCAATTTTGCCCCAG	15	73	Positive
scaffold297	150960	66575	66589	CCTTTTTCGATCACG	15	73	Negative
scaffold297	150960	114894	114908	CTCATTTCGCTCAAG	15	73	Negative
scaffold297	150960	146010	146024	CCATAAGGGCTCACG	15	73	Positive
scaffold297	150960	26741	26755	CCTTTTCGTCTCATG	15	73	Negative
scaffold297	150960	120344	120358	CCATGATGACTCCAG	15	73	Positive
scaffold297	150960	142819	142833	ACATTGTAGCTCGAG	15	73	Negative
scaffold297	150960	143374	143388	CGATTTTTGCGCACG	15	73	Positive
scaffold297	150960	32374	32388	CCCTTTGGGCCAAAG	15	73	Positive
scaffold297	150960	35147	35161	CCATTCTGTTTCTAG	15	73	Negative
scaffold297	150960	93064	93078	CCACGATGGGTCAAG	15	73	Positive
scaffold233	133269	103909	103923	GCGCTTTGGCTCCAG	15	73	Negative
scaffold233	133269	26177	26191	CCGTTTTTGCTCGCG	15	73	Negative
scaffold233	133269	55627	55641	CTGTTCTGGCACAAG	15	73	Negative
scaffold233	133269	27802	27816	CCGGATTGGCTGAAG	15	73	Positive
scaffold233	133269	75406	75420	CCATATTGGCTTTCG	15	73	Negative
scaffold233	133269	100668	100682	CAATCTGGGATCAAG	15	73	Negative
scaffold233	133269	25821	25835	CCTTTTTCGCTCTGG	15	73	Negative
scaffold233	133269	81567	81581	ACGCTTTGGCTCCAG	15	73	Positive
scaffold233	133269	90956	90970	CGGTTCTGGCACAAG	15	73	Positive
scaffold233	133269	123275	123289	GCATTTTCGCTGGAG	15	73	Positive
scaffold233	133269	78727	78741	CCATGATGGCACAGG	15	73	Positive
scaffold233	133269	49118	49132	GTATTCTGGCTCCAG	15	73	Negative
scaffold233	133269	78996	79010	TCGTCTTGACTCAAG	15	73	Negative
scaffold233	133269	69246	69260	CCCAATTGGGTCAAG	15	73	Negative
scaffold233	133269	38944	38958	CCATTGTCCCTCCAG	15	73	Positive
scaffold233	133269	111919	111933	TCATTCTCGCTCCAG	15	73	Positive
scaffold233	133269	47270	47284	TCGTATGGGCTCAAG	15	73	Positive
scaffold233	133269	66345	66359	ACATTGTGGATCAAC	15	73	Positive
scaffold233	133269	73764	73778	GCATGTTGGCTCCAC	15	73	Negative
scaffold233	133269	5257	5271	CCATCTTCGCAGAAG	15	73	Positive
scaffold233	133269	63769	63783	CGATTTTGCCTCCAC	15	73	Positive
scaffold309	124822	72675	72689	ACATTTTGGATGACG	15	73	Positive
scaffold309	124822	85651	85665	CGATTTTTGCACAGG	15	73	Positive
scaffold309	124822	92771	92785	CCATTTGGGCCCCCG	15	73	Negative
scaffold309	124822	121907	121921	TCTTTTTGTTTCAAG	15	73	Negative
scaffold309	124822	106382	106396	CCATCTTCGATAAAG	15	73	Negative
scaffold309	124822	65833	65847	AAATTTCGGATCAAG	15	73	Positive
scaffold309	124822	9430	9444	CCTCTTTGGCTTATG	15	73	Positive

scaffold309	124822	37125	37139	TCATTTTTGCTCCTG	15	73	Negative
scaffold309	124822	61927	61941	TCATTTTCACTTAAG	15	73	Positive
scaffold309	124822	87000	87014	TCTTTTTGGTTCCAG	15	73	Negative
scaffold309	124822	114586	114600	GTTTCTTGGCTCAAG	15	73	Negative
scaffold309	124822	90608	90622	CCATTTTGTGTCGTG	15	73	Positive
scaffold309	124822	1092	1106	CCTCGTCGGCTCAAG	15	73	Negative
scaffold309	124822	68746	68760	TCATTTTGGCGAACG	15	73	Positive
scaffold309	124822	45713	45727	CAATTTGGTCTCCAG	15	73	Negative
scaffold309	124822	78261	78275	CCGTTTTGGCGCTTG	15	73	Negative
scaffold309	124822	78532	78546	CCGTTGGGGCTCGAG	15	73	Positive
scaffold309	124822	15216	15230	CAGCTGTGGCTCAAG	15	73	Negative
scaffold309	124822	59752	59766	GCATTGTAGCTCATG	15	73	Negative
scaffold309	124822	34512	34526	CCGTTTGGCTTCAAG	15	73	Negative
scaffold309	124822	120847	120861	ACATTGTGACTCAAA	15	73	Negative
scaffold309	124822	64030	64044	CCATGTTGGGTGACG	15	73	Negative
scaffold132	124252	102880	102894	ACATTTTAGTACAAG	15	73	Negative
scaffold132	124252	88680	88694	TCCTTTTTGCTCAGG	15	73	Positive
scaffold132	124252	23553	23567	TGATTTTCGGTCAAG	15	73	Negative
scaffold132	124252	99852	99866	CCACCTTGGTTCAGG	15	73	Positive
scaffold132	124252	32903	32917	GCATTTTGGTGTAAG	15	73	Positive
scaffold132	124252	29222	29236	CAACCTTGGCACAAG	15	73	Negative
scaffold132	124252	116831	116845	CAATGTTGACACAAG	15	73	Negative
scaffold132	124252	64766	64780	CAATCTTGGCTAAAT	15	73	Negative
scaffold132	124252	8767	8781	CAATATTCTCTCAAG	15	73	Positive
scaffold132	124252	2202	2216	CCATTCTCGGTCAAT	15	73	Negative
scaffold132	124252	95311	95325	CCATGCTGGCTCCCG	15	73	Negative
scaffold132	124252	46009	46023	GCATGTTGGTCCAAG	15	73	Negative
scaffold132	124252	82572	82586	CTGTTTTTGCTGAAG	15	73	Positive
scaffold132	124252	65839	65853	CCACTTTCGTTCGAG	15	73	Negative
scaffold811	107907	70365	70379	GCTTTTTGGATCAGG	15	73	Positive
scaffold811	107907	48058	48072	CGTTTTTGGCGCAGG	15	73	Positive
scaffold811	107907	29907	29921	CCATTTTTGGCCGAG	15	73	Positive
scaffold811	107907	25261	25275	CCCTTTTTCCGCAAG	15	73	Positive
scaffold811	107907	44356	44370	CCATTTTTGCTCTGC	15	73	Positive
scaffold811	107907	91928	91942	CCGTGGTGGCACAAG	15	73	Negative
scaffold811	107907	99108	99122	GCTTTTCTGCTCAAG	15	73	Negative
scaffold811	107907	59009	59023	TCCTTGTGCCTCAAG	15	73	Negative
scaffold811	107907	35833	35847	CGATTTTTCCTCGAG	15	73	Positive
scaffold811	107907	68476	68490	CGAAATTGTCTCAAG	15	73	Negative
scaffold811	107907	105186	105200	TCAGTGTGGCTCCAG	15	73	Negative
scaffold811	107907	62690	62704	GCAATCTGGATCAAG	15	73	Negative
scaffold811	107907	934	948	CCATTTGGTTTCAAA	15	73	Negative
scaffold811	107907	15444	15458	CCATTTTGTCGCTCG	15	73	Positive
scaffold811	107907	104416	104430	CCATTTTGGTGCTTG	15	73	Positive
scaffold811	107907	62165	62179	ACATTTTACCTCGAG	15	73	Negative
scaffold811	107907	1286	1300	CCATTTTGGCGCCCT	15	73	Positive
scaffold811	107907	43187	43201	CCATCTTTGATGAAG	15	73	Positive
scaffold811	107907	101287	101301	ACTTTTTGCCTCACG	15	73	Negative
scaffold307	88209	16668	16682	TCATTTTGGATCAGC	15	73	Negative
scaffold307	88209	55004	55018	CCATTTTCGTTCCGG	15	73	Positive

scaffold307	88209	71924	71938	CCATGATGGCTCATC	15	73	Negative
scaffold307	88209	26219	26233	CCCTTGTGCCTCATG	15	73	Positive
scaffold307	88209	7950	7964	CAACATTCGCTCAAG	15	73	Negative
scaffold307	88209	46548	46562	CAGTTTTGGGTTAAG	15	73	Positive
scaffold307	88209	38289	38303	CGACTTTGGCTGGAG	15	73	Positive
scaffold307	88209	36815	36829	CAATTTTTGCTCCAC	15	73	Positive
scaffold307	88209	66696	66710	CTATTTTTGCTCTAT	15	73	Positive
scaffold307	88209	30521	30535	CCATTTTTGCGGATG	15	73	Positive
scaffold307	88209	35566	35580	CGATTATGGCTAATG	15	73	Negative
scaffold307	88209	41453	41467	TCAATTTCGATCAAG	15	73	Negative
scaffold820	86658	51078	51092	CCACATTGAGTCAAG	15	73	Positive
scaffold820	86658	78301	78315	CCATTTATGCTCAGT	15	73	Positive
scaffold820	86658	86404	86418	CCGATCTGGTTCAAG	15	73	Positive
scaffold820	86658	12469	12483	CCATTGAGGCTCCCG	15	73	Negative
scaffold820	86658	57480	57494	CCGTTTTCGCTCGAC	15	73	Positive
scaffold820	86658	49911	49925	CCGTTTCTGCTCACG	15	73	Negative
scaffold820	86658	72713	72727	ACTTTTTGGCTCAGA	15	73	Positive
scaffold820	86658	51822	51836	GGATTTAGGCTGAAG	15	73	Positive
scaffold820	86658	76323	76337	CCGTTGTGGCTGATG	15	73	Positive
scaffold574	85034	36975	36989	CTTTTTTATCTCAAG	15	73	Negative
scaffold574	85034	30447	30461	CTGTTTTGGGTCCAG	15	73	Negative
scaffold574	85034	41211	41225	CGATTTTTTCTAAAG	15	73	Positive
scaffold574	85034	35006	35020	CCACATTGGTTAAAG	15	73	Negative
scaffold574	85034	54643	54657	GCATTTTGGCCGATG	15	73	Negative
scaffold574	85034	9719	9733	CCATTTTGGCGGCCG	15	73	Positive
scaffold574	85034	47590	47604	CCATTCTGCAGCAAG	15	73	Positive
scaffold574	85034	11819	11833	CCATTTGGGCGTGAG	15	73	Positive
scaffold574	85034	49251	49265	CCACTTTTGCTCCAT	15	73	Negative
scaffold574	85034	66947	66961	CAATCTTGGCAGAAG	15	73	Positive
scaffold956	84797	47764	47778	TCACTTGGCCTCAAG	15	73	Positive
scaffold956	84797	17208	17222	CCATCTCGGCTCTGG	15	73	Negative
scaffold956	84797	41254	41268	CAATTTCGTCTCGAG	15	73	Positive
scaffold956	84797	34752	34766	CCATCTTGGCCAATG	15	73	Positive
scaffold956	84797	28189	28203	CCGATTCGGCTGAAG	15	73	Positive
scaffold956	84797	11415	11429	CGATTCAGGATCAAG	15	73	Negative
scaffold956	84797	60270	60284	CCATTGTTGCGCAGG	15	73	Positive
scaffold956	84797	61901	61915	CCGTTTTGTAGCAAG	15	73	Positive
scaffold956	84797	62744	62758	CCGTTTTGTAGCAAG	15	73	Positive
scaffold956	84797	51036	51050	CCGTTTTTCCTCCAG	15	73	Positive
scaffold956	84797	1885	1899	CCGTTTTTGTTCAAC	15	73	Positive
scaffold381	76577	4342	4356	CCATCCCGGCACAAG	15	73	Negative
scaffold381	76577	17232	17246	CGATTTTGCCGCAGG	15	73	Positive
scaffold381	76577	27947	27961	TCAATTTGGTTCACG	15	73	Positive
scaffold381	76577	26131	26145	CCATCTTCGCACAAC	15	73	Positive
scaffold381	76577	37124	37138	ATATTTTGGCTGGAG	15	73	Positive
scaffold381	76577	15580	15594	CCCTTATGGCTAGAG	15	73	Negative
scaffold381	76577	62039	62053	CCATTTTGTTCCAAA	15	73	Positive
scaffold381	76577	75340	75354	ACATTCTGGAGCAAG	15	73	Negative
scaffold381	76577	43273	43287	CCATATCGACTCACG	15	73	Negative
scaffold381	76577	48579	48593	CCATTTTGGCGGTAT	15	73	Positive

scaffold381	76577	48751	48765	CCATTCCGGCCCGAG	15	73	Positive
scaffold381	76577	23587	23601	CAATTATGGCTCCAC	15	73	Positive
scaffold381	76577	32166	32180	CCATTTTGGATGACT	15	73	Negative
scaffold381	76577	8926	8940	GCATTTTGCCTCCAA	15	73	Negative
scaffold381	76577	29627	29641	CGATTACGGCTCGAG	15	73	Negative
scaffold381	76577	6512	6526	CCTTCTTTGCTCATG	15	73	Positive
scaffold1189	72863	45089	45103	CCATTTTGGTGCCGG	15	73	Negative
scaffold1189	72863	40463	40477	CGACTGTGGCTCCAG	15	73	Positive
scaffold1189	72863	56498	56512	TCAATTTGGATCAAC	15	73	Positive
scaffold1189	72863	28184	28198	CCGTTTAGGGTCAGG	15	73	Positive
scaffold1189	72863	27742	27756	CCGTTTAGGGTCAGG	15	73	Positive
scaffold1189	72863	53305	53319	CCGTTCTCGCTCCAG	15	73	Positive
scaffold1189	72863	19816	19830	CCCTTTTGTTTCAGG	15	73	Negative
scaffold1189	72863	24035	24049	GCGTTTTGCCTCATG	15	73	Negative
scaffold1189	72863	60867	60881	CCATTTTTGCAGAGG	15	73	Negative
scaffold1189	72863	25239	25253	GCATTTTGCCTGTAG	15	73	Positive
scaffold1189	72863	22479	22493	ACATTTTTGCTCATT	15	73	Positive
scaffold1189	72863	22981	22995	CCATTTCTGTACAAG	15	73	Negative
scaffold1189	72863	54113	54127	CTCTTGTCGCTCAAG	15	73	Positive
scaffold1385	64388	18389	18403	CAATTCTGTCGCAAG	15	73	Positive
scaffold1385	64388	37767	37781	CGATTCCGGCTCCAG	15	73	Negative
scaffold1385	64388	2852	2866	CAATTTTTACTCAAC	15	73	Negative
scaffold1385	64388	53420	53434	ATCTTCTGGCTCAAG	15	73	Positive
scaffold1385	64388	4762	4776	TGGTTTTGGCTCCAG	15	73	Negative
scaffold1385	64388	50488	50502	CCATGCTCGCGCAAG	15	73	Negative
scaffold1385	64388	43584	43598	CCATGCTCGCGCAAG	15	73	Negative
scaffold1385	64388	30870	30884	ACATTTTGGCGCGCG	15	73	Negative
scaffold1385	64388	24257	24271	CCGTTTTGGCGGAGG	15	73	Negative
scaffold1385	64388	52390	52404	TCATTTTGGCCCCGG	15	73	Negative
scaffold1385	64388	38426	38440	TCATTTTGCCCGAAG	15	73	Positive
scaffold1205	59972	53275	53289	CAATTTTGGATGAGG	15	73	Negative
scaffold1205	59972	46679	46693	CCATCTTGAAACAAG	15	73	Negative
scaffold1205	59972	32401	32415	CTATTGTAGATCAAG	15	73	Positive
scaffold1205	59972	33177	33191	CAATTCTGCATCAAG	15	73	Negative
scaffold1205	59972	4825	4839	CTATTTTGGCTGTTG	15	73	Positive
scaffold1205	59972	21852	21866	CCGTTGTGGCCCAGG	15	73	Positive
scaffold912	59273	52725	52739	CAATTTCGACTCAAC	15	73	Positive
scaffold912	59273	13948	13962	CATTTTTGGAACAAG	15	73	Negative
scaffold912	59273	16819	16833	TCCTTTTCGCTCAGG	15	73	Negative
scaffold912	59273	51996	52010	AAATTTTGGCTGTAG	15	73	Negative
scaffold600	54455	24233	24247	CCATTTTATCGCATG	15	73	Positive
scaffold600	54455	5358	5372	CGATTTTTGCACAAC	15	73	Negative
scaffold600	54455	38363	38377	GCATTTTGCCACAAC	15	73	Negative
scaffold600	54455	53769	53783	CGGTTTTGGCGCCAG	15	73	Negative
scaffold600	54455	17571	17585	CTTTTGTGGCTCACG	15	73	Positive
scaffold600	54455	40855	40869	AGATTCTGGCTAAAG	15	73	Positive
scaffold1308	47888	34143	34157	CAGCTTTGGCTGAAG	15	73	Negative
scaffold1308	47888	38334	38348	CCGTGTCGTCTCAAG	15	73	Negative
scaffold1308	47888	4250	4264	TCACTTTGGCCCACG	15	73	Negative
scaffold1308	47888	18903	18917	TCATTTTGGATAAGG	15	73	Negative

scaffold958	45976	27838	27852	CAATTTTTGCTGAAT	15	73	Negative
scaffold958	45976	42304	42318	CCGTATTGGGTCAAT	15	73	Positive
scaffold958	45976	41786	41800	CGATTTCGGTTCAGG	15	73	Positive
scaffold958	45976	1731	1745	TCATTTTTCCTCCAG	15	73	Negative
scaffold958	45976	2787	2801	GCGTTTTGGCAGAAG	15	73	Negative
scaffold958	45976	26652	26666	CCTTTCTCGCTGAAG	15	73	Positive
scaffold958	45976	9464	9478	CCATGCTTGTTCAAG	15	73	Positive
scaffold958	45976	28985	28999	CCATGCTAGCACAAG	15	73	Positive
scaffold958	45976	37348	37362	CCTTTCGAGCTCAAG	15	73	Positive
scaffold1289	45660	16206	16220	CCATCTTCGCACAAC	15	73	Positive
scaffold1289	45660	38740	38754	CGCTTTCGACTCAAG	15	73	Negative
scaffold1289	45660	5357	5371	CCCTTTTCGTTCAAC	15	73	Negative
scaffold1289	45660	24186	24200	CGATTTGGGGTCAAC	15	73	Positive
scaffold1289	45660	16225	16239	CCATTACGGCTTACG	15	73	Positive
scaffold1289	45660	30961	30975	CTAGTTTAGCCCAAG	15	73	Negative
scaffold1359	45323	30379	30393	CCGCTTTGGGTGAAG	15	73	Positive
scaffold1359	45323	25021	25035	CCATCGCGTCTCAAG	15	73	Positive
scaffold1398	42967	12438	12452	TCATTTTTCCTCTAG	15	73	Negative
scaffold1398	42967	37121	37135	CGGTTGTGGCTCAGG	15	73	Negative
scaffold1398	42967	36744	36758	GCGTTTTGGCGCGAG	15	73	Positive
scaffold1398	42967	4020	4034	CGATTTTGGCCGACG	15	73	Positive
scaffold1398	42967	11434	11448	CCCTTTTCGCTGATG	15	73	Negative
scaffold1358	40547	21473	21487	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	21566	21580	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	21206	21220	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	21029	21043	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	20755	20769	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	20848	20862	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	21743	21757	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	19297	19311	GCATTTAGCCTCTAG	15	73	Negative
scaffold1358	40547	20488	20502	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	21157	21171	CGGCTTTGGGTCAAG	15	73	Positive
scaffold1358	40547	20311	20325	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	20218	20232	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	4757	4771	CCGTTTTTGCTCGGG	15	73	Negative
scaffold1358	40547	20980	20994	CGGCTTTGGGTCAAG	15	73	Positive
scaffold1358	40547	21656	21670	CGACCTTGACTCAAG	15	73	Negative
scaffold1358	40547	17386	17400	CCATTTTCTCTGGAG	15	73	Negative
scaffold1358	40547	21903	21917	CTATCTTGACGCAAG	15	73	Negative
scaffold1358	40547	20169	20183	CGGCTTTGGGTCAAG	15	73	Positive
scaffold1358	40547	18231	18245	CCAGTTTCGATCACG	15	73	Negative
scaffold1358	40547	20439	20453	CAGCTTTGGGTCAAG	15	73	Positive
scaffold1358	40547	13498	13512	CAGCTTTCGCTCAAG	15	73	Negative
scaffold385	40098	14012	14026	GCTTCTTGGCGCAAG	15	73	Positive
scaffold385	40098	27179	27193	ACATTTTTGATCGAG	15	73	Positive
scaffold385	40098	39249	39263	TCATTTTGGTTGAAC	15	73	Negative
scaffold385	40098	27777	27791	CCATCTTGCTTCAGG	15	73	Negative
scaffold385	40098	24503	24517	GCATTGTGGATCACG	15	73	Positive
scaffold385	40098	30135	30149	CCACTTTGACCGAAG	15	73	Negative
scaffold1489	38893	307	321	CTGTTTTGACTCCAG	15	73	Negative

scaffold1489	38893	7848	7862	CCAGTTTGGCTTCCG	15	73	Negative
scaffold1489	38893	10094	10108	CCATCACGGCTCAAA	15	73	Positive
scaffold1489	38893	37148	37162	CCAATCGGGCTCATG	15	73	Negative
scaffold1489	38893	23309	23323	ACATTTTGAATCAAC	15	73	Positive
scaffold1489	38893	30158	30172	CAATTTCTGCTCCAG	15	73	Negative
scaffold1489	38893	5706	5720	CAATTTTTGCCCGAG	15	73	Negative
scaffold656	38666	7444	7458	CGATTTTGTCGGAAG	15	73	Positive
scaffold656	38666	30715	30729	CCAGTTTCGATCTAG	15	73	Positive
scaffold656	38666	32534	32548	GCATTTTAGCTCCCG	15	73	Positive
scaffold656	38666	21068	21082	CCATTTTAGCTGGTG	15	73	Positive
scaffold549	37577	336	350	CCATTTTGGTCTAAA	15	73	Negative
scaffold549	37577	31620	31634	CTATTTTGGATTTAG	15	73	Positive
scaffold549	37577	13097	13111	CCTTCTTGGCTTCAG	15	73	Negative
scaffold549	37577	20102	20116	TCATTTTGTCTGTAG	15	73	Negative
scaffold549	37577	35287	35301	CAATCTTGGCTAAAT	15	73	Positive
scaffold549	37577	28732	28746	TCATTTTTGCGGAAG	15	73	Positive
scaffold1458	27774	7593	7607	CGCTTTTGGATGAAG	15	73	Positive
scaffold1458	27774	24629	24643	TTATTTTGGCTGACG	15	73	Positive
scaffold1458	27774	14029	14043	GCGCTTTGGATCAAG	15	73	Positive
scaffold1458	27774	3041	3055	GCATCTTGGCTCGTG	15	73	Positive
scaffold1320	25307	6334	6348	CCAGAATGGCTCATG	15	73	Positive
scaffold1320	25307	12349	12363	CCAATGTGGCTCCAT	15	73	Negative
scaffold1270	21076	20550	20564	GCATTCTGCCTCCAG	15	73	Negative
scaffold1737	13558	5807	5821	TCTTTTTGCCTCTAG	15	73	Negative
scaffold1737	13558	10208	10222	TCATCTTGCCTCAAC	15	73	Negative
scaffold1737	13558	10457	10471	TCGTGGTGGCTCAAG	15	73	Positive
scaffold1737	13558	12701	12715	CCAATTTGGCTTTCG	15	73	Positive
scaffold1737	13558	8679	8693	CTGTTTTTGCTCGAG	15	73	Negative
scaffold1737	13558	8508	8522	GCCTTGTGGCTCAAC	15	73	Negative
scaffold2289	9361	4726	4740	CCAGTCTGGGACAAG	15	73	Negative
scaffold2066	8694	6169	6183	CCCTTCTGGCTGTAG	15	73	Negative
scaffold2171	8014	4144	4158	CCCTTCGGGATCAAG	15	73	Negative
scaffold2171	8014	7533	7547	CGAATTATGCTCAAG	15	73	Negative
scaffold2171	8014	7463	7477	ACATCTTGACTCTAG	15	73	Negative
scaffold3000	7973	7023	7037	CCCTTTCGGCTCTAC	15	73	Positive
scaffold3000	7973	6583	6597	AAATGTTGGCTCAAC	15	73	Positive
scaffold3000	7973	3907	3921	CCATTTTCGCGCGCG	15	73	Negative
scaffold2275	7391	5035	5049	CGGTTTTGGTTCAAT	15	73	Negative
scaffold3442	4912	357	371	AGATTTTGGCGCACG	15	73	Positive
scaffold4061	3138	1123	1137	GCTCGTTGGCTCAAG	15	73	Positive

Table S4. Illumina RNAseq reads of *A. ceratii* at 6 and 96 hours during the infection. (***A***) Illumina RNAseq reads of *A. ceratii* sp. ex *A. catenella* at 6 and 96 hours during the infection. RNA was extracted following the manufacturer's protocol (Sigma-Aldrich, Steinheim, Germany). RNA quality and quantity check were performed using a NanoDrop ND-100 spectrometer (PeqLab, Erlangen, Germany) and a RNA Nano Chip Assay by the 2100 Bioanalyzer (Agilent Technologies, Böblingen, Germany). Library preparation was done using Illumina's TruSeq RNA sample prep kit (Illumina, San Diego, CA, USA). Libraries were sequenced with a HiSeq2000/2500 (high-throughput mode) in single-read/50 cycle mode. After quality checking the reads were aligned to the *A. ceratii* predicted genes (19,925 genes) by using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) to obtain the read counts. (***B***) KOG enrichments of *A. ceratii* predicted genes with intronless (0 intron), low-intron (1-3 introns) and high-intron (>3 introns) (***Bold Italic***: p-value < 0.01). Significant enrichments of the transcripts were tested by calculating the *P* value from a hypergeometric distribution at the background level of the annotation of eukaryotic orthologous groups (KOGs) (https://genome.jgi.doe.gov/Tutorial/tutorial/kog.html). KOGs were considered significantly enriched for a given intron number when test statistics gave a *P* value of < 0.01. All RNASeq raw sequencing data have been submitted to NCBI under Bioproject PRJNA274490.

A

Time point	Name	Sequencing approach	No. of reads	Read length
6 hours	Sample_6h_A	single	34,132,086	50
	Sample_6h_B	single	27,920,490	50
	Sample_6h_C	single	29,078,452	50
96 hours	Sample_96h_A	single	33,502,911	50
	Sample_96h_B	single	36,356,519	50
	Sample_96h_C	single	37,047,919	50

B

intron counts	0			0-3			>3		
	Count	*p*-value	%	Count	*p*-value	%	Count	*p*-value	%
Cell wall/membrane/envelope biogenesis	16.33	0.079633653	3.12	29.00	0.093435737	2.66	19.33	0.055213818	2.08
Cell motility	1.50	0.389494754	0.29	2.00	0.315209359	0.18	0.00	0.255190159	0.00
Posttranslational modification, protein turnover, chaperones	53.00	0.064681456	10.13	115.50	0.041555137	10.59	88.00	0.041136207	9.46
Signal transduction mechanisms	***40.83***	***1.93014E-07***	***7.81***	150.00	0.041466774	13.75	***168.33***	***4.60454E-06***	***18.10***
Intracellular trafficking, secretion, and vesicular transport	2.50	0.293493683	0.48	***9.00***	***0.008696886***	***0.82***	0.00	0.00661944	0.00
Defense mechanisms	***28.50***	***4.25763E-06***	***5.45***	33.00	0.021217713	3.02	***0.00***	***5.77826E-13***	***0.00***
Cytoskeleton	1.00	0.362025869	0.19	4.50	0.220721174	0.41	1.50	0.166697724	0.16
RNA processing and modification	0.00	0.31608484	0.00	1.00	0.228131368	0.09	4.00	0.056509636	0.43
Chromatin structure and dynamics	1.50	0.328614056	0.29	2.50	0.178715282	0.23	4.50	0.202707915	0.48
Translation, ribosomal structure and biogenesis	68.50	0.05350139	13.10	***157.00***	***0.001891763***	***14.39***	***91.50***	***0.000367389***	***9.84***
Transcription	16.00	0.109138109	3.06	31.33	0.058576815	2.87	35.50	0.046144398	3.82
Replication, recombination and repair	44.00	0.040346179	8.41	***89.33***	***0.005604752***	***8.19***	***111.00***	***0.000583883***	***11.94***
Energy production and conversion	31.00	0.0655722	5.93	51.00	0.033679218	4.67	53.33	0.057782765	5.73
Cell cycle control, cell division, chromosome partitioning	***0.83***	***0.009829731***	***0.16***	8.00	0.174292284	0.73	11.83	0.0432854	1.27
Amino acid transport and metabolism	***42.00***	***0.001087038***	***8.03***	61.00	0.063035812	5.59	***33.57***	***0.000568139***	***3.61***
Nucleotide transport and metabolism	18.00	0.03486392	3.44	21.33	0.031284004	1.96	23.83	0.10524688	2.56
Carbohydrate transport and metabolism	31.50	0.020896894	6.02	44.33	0.045920928	4.06	39.95	0.06673197	4.30
intron counts	0			0-3			>3		
	Count	*p*-value	%	Count	*p*-value	%	Count	*p*-value	%
Lipid transport and metabolism	14.17	0.047791046	2.71	48.20	0.016068817	4.42	30.00	0.06497314	3.23

Inorganic ion transport and metabolism	20.50	0.062218657	3.92	***37.00***	***0.003067602***	***3.39***	***59.75***	***0.000561138***	***6.42***
Secondary metabolites biosynthesis, transport and catabolism	13.17	0.042586633	2.52	21.83	0.089222102	2.00	***8.50***	***0.00536165***	***0.91***
General function prediction only	60.17	0.060949467	11.50	124.15	0.047574857	11.38	110.75	0.049142591	11.91
Function unknown	4.00	0.073973436	0.76	18.00	0.079998993	1.65	13.00	0.13907704	1.40

Table S5. A list of enzymes involved in various metabolic pathways in the *A. ceratii* genome.

Enzymes	Sequence Names	Number of genes
Glycolysis		
1.2.1.12: Glyceraldehyde 3-phosphate dehydrogenase	g5105.t1, g5318.t1, g1268.t1, g15719.t1	4
2.7.1.11: 6-phosphofructokinase	g4532.t1, g11905.t1, g13580.t1, g3565.t1, g16188.t1	5
2.7.1.2: Glucokinase	g6467.t1, g9076.t1, g17422.t1, g1820.t1, g6898.t1	3
2.7.1.40: Pyruvate kinase	g4450.t1, g7810.t1, g7811.t1, g13951.t1, g13651.t1, g1953.t1, g3035.t1	7
2.7.2.3: Phosphoglycerate kinase	g16261.t1, g6745.t1	2
4.1.2.13: Fructose-bisphosphate aldolase, class I	g1265.t1, g3323.t1	2
4.2.1.11: Enolase	g4531.t1, g10562.t1	2
5.3.1.1: Triosephosphate isomerase (TIM)	g2077.t1	1
5.3.1.9: Glucose-6-phosphate isomerase	g15821.t1, g3214.t1	2
5.4.2.1: Phosphoglycerate mutase	g11930.t1, g6352.t1	2
Other enzymes that might be involved in Glycolysis		
1.1.1.1/1.1.1.2: alcohol dehydrogenase; aldehyde reductase;	g17476.t1, g17477.t1, g17478.t1, g8973.t1, g1813.t1, g14245.t1, g3976.t1, g13812.t1, g10793.t1, g4523.t1, g5180.t1, g8245.t1, g4882.t1, g12688.t1, g11613.t1, g2755.t1, g2425.t1, g9223.t1	18
1.2.1.3: aldehyde dehydrogenase (NAD+);	g10929.t1, g15115.t1, g6908.t1, g9892.t1, g405.t1, g7274.t1, g5125.t1, g8354.t1, g4882.t1	9
1.2.4.1: pyruvate dehydrogenase (acetyl-transferring);	g4674.t1	1
1.8.1.4: dihydrolipoyl dehydrogenase;	g2390.t1, g11161.t1, g7447.t1	3
2.3.1.12: dihydrolipoyllysine-residue acetyltransferase;	g14980.t1, g467.t1, g872.t1	3
2.7.1.1: hexokinase	g1577.t1	1
3.1.3.11: fructose-bisphosphatase	g14685.t1, g14101.t1, g8034.t1, g17521.t1, g17520.t1	5
3.1.3.13: bisphosphoglycerate phosphatase	g18320.t1	1
4.1.1.49: phosphoenolpyruvate carboxykinase	g9192.t1, g834.t1, g12237.t1	3
5.1.3.15: glucose-6-phosphate 1-epimerase	g7980.t1, g10797.t1	2
5.1.3.3: aldose 1-epimerase; mutarotase;	g14385.t1, g5710.t1	2
5.4.2.2: phosphoglucomutase (alpha-D-glucose-1,6-bisphosphate-dependent);	g3214.t1, g13858.t1	2
5.4.2.4: bisphosphoglycerate mutase	g18320.t1	1
6.2.1.1: acetate---CoA ligase	g11834.t1, g4449.t1, g6814.t1, g7882.t1	4

Enzymes	Sequence Names	Number of genes
TCA cycle		
6.2.1.4/6.2.1.5: Succinyl-CoA synthetase alpha subunit	g7046.t1, g14826.t1, g4515.t1, g7046.t1, g14826.t1, g5924.t1, g4150.t1, g5450.t1, g5451.t1, g5449.t1	10
1.1.1.37: Malate dehydrogenase	g17707.t1, g3242.t1, g10767.t1, g17251.t1, g2848.t1, g4769.t1, g2054.t1, g14670.t1, g19916.t1, g4492.t1	10
1.1.1.42: Isocitrate dehydrogenase	g19124.t1, g629.t1, g7997.t1, g9270.t1	4
1.2.4.1/2.3.1.12: Pyruvate dehydrogenase E1 component/ Pyruvate dehydrogenase E2 component (dihydrolipoamide acetyltransferase)	g4674.t1, g14980.t1, g467.t1, g872.t1	4
2.3.1.61: 2-oxoglutarate dehydrogenase E2 component (dihydrolipoamide succinyltransferase)	g12217.t1, g13311.t1	2
1.3.5.1: Succinate dehydrogenase (ubiquinone) flavoprotein subunit	g6898.t1, g2130.t1, g7864.t1, g5319.t1	4
2.3.3.1: Citrate synthase	g16719.t1, g19126.t1, g474.t1, g477.t1, g8175.t1, g12081.t1, g6553.t1, g6197.t1, g5893.t1, g18242.t1	10
4.2.1.2: Fumarate hydratase, class I	g14808.t1, g13023.t1, g17892.t1, g9027.t1	4
4.2.1.3: Aconitate hydratase 1	g12259.t1, g15984.t1, g15985.t1, g12236.t1, g12237.t1, g12238.t1, g18949.t1, g4475.t1, g9355.t1, g9467.t1, g14059.t1, g10767.t1, g2848.t1, g17251.t1, g18949.t1, g4475.t1, g9355.t1, g9467.t1, g14059.t1, g10767.t1, g2848.t1, g17251.t1, g9467.t1, g14059.t1	24
Other enzymes that might be involved in TCA cycle		
1.1.5.4: malate dehydrogenase (quinone)	g3976.t1, g9723.t1, g4347.t1, g13513.t1, g4347.t1, g13513.t1	6
1.8.1.4: dihydrolipoyl dehydrogenase	g2390.t1, g11161.t1, g7447.t1	3
2.3.3.1: citrate (Si)-synthase	g13345.t1	1
2.3.3.8: ATP citrate synthase	g5236.t1, g5237.t1, g4150.t1, g5450.t1, g5451.t1, g5449.t1	6
4.1.1.49: phosphoenolpyruvate carboxykinase (ATP)	g9192.t1, g834.t1, g12237.t1	3
Oxidative phosphorylation		
1.6.99.3: NADH dehydrogenase/ NADH dehydrogenase (complex I)	g180.t1	1
1.3.5.1: Succinate dehydrogenase (ubiquinone)	g6898.t1, g2130.t1, g7864.t1,	4

Enzymes	**Sequence Names**	**Number of genes**
flavoprotein subunit (complex II)	g5319.t1	
1.10.2.2 : Ubiquinol-cytochrome c reductase iron-sulfur subunit (complex III)	--	--
1.9.3.1: Cb-type cytochrome c oxidase subunit I (complex IV)	g15932.t1, g17513.t1, g16451.t1, g3512.t1	4
3.6.3.6/3.6.3.14: H+ transporting ATPase/F-type H+-transporting ATPase subunit a (complex V)	g1027.t1, g10076.t1, g16058.t1, g16059.t1, g13480.t1, g10622.t1, g2780.t1, g4415.t1, g15533.t1, g15706.t1, g14951.t1, g7935.t1, g5142.t1, g5143.t1, g19157.t1, g7051.t1, g8137.t1, g16352.t1, g17037.t1, g13176.t1, g2225.t1, g2226.t1, g2227.t1, g2228.t1, g3147.t1, g4262.t1, g4504.t1, g10609.t1, g19394.t1, g10527.t1, g1265.t1, g799.t1	32
Other enzymes that might be involved in oxidative phosphorylation		
2.7.4.1: polyphosphate kinase; polyphosphoric acid kinase	g5951.t1	1
3.6.1.1: inorganic diphosphatase	g10466.t1, g16180.t1, g1721.t1, g4304.t1, g9124.t1, g12655.t1, g8814.t1, g1974.t1, g5266.t1, g2036.t1, g7849.t1, g5341.t1, g8766.t1, g14558.t1, g6231.t1, g11387.t1	16
$_D$-Lactate: cytochrome c oxidoreductase ($_D$-LDH)	g15250.t1, g2635.t1	2
Galacto-1,4-lactone: cytochrome c oxidoreductase (G-1,4-LDH)	g1813.t1	1
Pentose phosphate pathway		
1.1.1.49 glucose-6-phosphate dehydrogenase	--	--
3.1.1.31 6-phosphogluconolactonase	g19069.t1, g14685.t1, g14101.t1, g8034.t1, g17521.t1, g17520.t1	6
1.1.1.44 phosphogluconate dehydrogenase (NADP+-dependent, decarboxylating)	g18894.t1, g9293.t1, g10565.t1, g16978.t1, g15112.t1, g14530.t1, g8992.t1	7
5.3.1.6 ribose-5-phosphate isomerase	g12957.t1, g565.t1, g11785.t1, g10558.t1	4
5.1.3.1 ribulose-phosphate 3-epimerase	g11989.t1, g16002.t1	2
2.2.1.1 transketolase	g7317.t1, g5062.t1, g7241.t1	3
2.2.1.2 transaldolase	g12919.t1, g15227.t1, g265.t1, g13038.t1	4
Other enzymes that might be involved in pentose phosphate pathway		
1.1.1.44: phosphogluconate dehydrogenase	g9530.t1, g10943.t1, g14830.t1, g17477.t1, g17476.t1	5
2.7.1.11: 6-phosphofructokinase	g4532.t1, g11905.t1, g13580.t1, g3565.t1, g16188.t1	5

Enzymes	Sequence Names	Number of genes
2.7.1.15: ribokinase; deoxyribokinase;	g20772.t1, g18924.t1	2
2.7.6.1: ribose-phosphate diphosphokinase	g14804.t1, g18249.t1, g19930.t1, g9663.t1, g5424.t1, g13783.t1, g16201.t1	7
4.1.2.13: fructose-bisphosphate aldolase;	g1265.t1, g3323.t1	2
5.3.1.9: glucose-6-phosphate isomerase;	g15821.t1, g3214.t1	2
5.4.2.2: phosphoglucomutase (alpha-D-glucose-1,6-bisphosphate-dependent);	g3214.t1, g13858.t1	2
Purine nucleotide synthesis		
Inosine monophosphate synthesis (precursor)		
2.1.2.2 : GAR transformylase	--	--
2.1.2.3: AICAR transformylase	g8877.t1, g3736.t1	1
2.4.2.14: Amidophosphoribosyltransferase	--	--
2.7.6.1: Ribose-phosphate pyrophosphokinase	g14804.t1, g18249.t1, g19930.t1, g9663.t1, g5424.t1, g13783.t1, g16201.t1	7
3.5.4.10: IMP cyclohydrolase	--	--
4.1.1.21: AIR carboxylase	--	--
4.3.2.2 : Adenylosuccinate lyase	g2152.t1, g17162.t1, g17162.t1	3
6.3.2.6 : SAICAR synthetase	g1905.t1, g7103.t1	2
6.3.3.1: AIR synthetase	--	--
6.3.4.13 : Phosphoribosylamine--glycine ligase	--	--
6.3.5.3 : FGAM synthetase	g136.t1, g16665.t1	2
Synthesis of AMP, ADP and ATP		
6.3.4.4: Adenylosuccinate synthase	g2976.t1, g10842.t1	2
4.3.2.2 : Adenylosuccinate lyase	g2152.t1, g17162.t1, g17162.t1	3
2.7.4.3: Adenylate kinase	g10891.t1, g10926.t1, g11284.t1, g14085.t1, g15913.t1, g15914.t1, g18135.t1, g19077.t1, g3270.t1, g6743.t1, g8645.t1, g18723.t1, g11828.t1, g6398.t1, g7817.t1, g6265.t1, g16279.t1, g15238.t1, g19172.t1	19
2.7.4.6: Nucleoside-diphosphate kinase	g14086.t1, g7640.t1, g6274.t1, g108.t1, g14746.t1	5
Synthesis of dADP and dATP		
1.17.4.1: Ribonucleotide reductase, class II	g1133.t1, g9496.t1, g10928.t1, g8955.t1, g3413.t1, g3231.t1, g12136.t1, g7603.t1	8
2.7.4.6: Nucleoside-diphosphate kinase	g14086.t1, g7640.t1, g6274.t1, g108.t1, g14746.t1	5
Synthesis of GMP, GDP and GTP		
1.1.1.205: IMP dehydrogenase	g19874.t1, g3209.t1, g5837.t1, g8008.t1, g3593.t1, g18773.t1	6

Enzymes	Sequence Names	Number of genes
6.3.5.2: GMP synthase (glutamine-hydrolysing)	g1284.t1, g7103.t1, g8052.t1, g5566.t1, g12338.t1	5
2.7.4.8: Guanylate kinase	g11449.t1, g1263.t1	2
2.7.4.6: Nucleoside-diphosphate kinase	g14086.t1, g7640.t1, g6274.t1, g108.t1, g14746.t1	5
Synthesis of dGDP and dGTP		
1.17.4.1: Ribonucleotide reductase, class II	g1133.t1, g9496.t1, g10928.t1, g8955.t1, g3413.t1, g3231.t1, g12136.t1, g7603.t1	8
2.7.4.6: Nucleoside-diphosphate kinase	g14086.t1, g7640.t1, g6274.t1, g108.t1, g14746.t1	5
Other enzymes that might be involved in purine nucleotide synthesis		
1.17.1.4: xanthine dehydrogenase;	g11271.t1, g8501.t1	2
1.17.3.2: xanthine oxidase;	g11271.t1, g8501.t1	2
1.7.1.7: GMP reductase; guanosine 5'-monophosphate reductase	g10.t1, g7456.t1	2
1.7.3.3: factor-independent urate hydroxylase	g11270.t1	1
2.4.2.1: purine-nucleoside phosphorylase	g8354.t1, g6545.t1	2
2.4.2.7: adenine phosphoribosyltransferase	g2876.t1, g20801.t1, g17488.t1	3
2.4.2.8: hypoxanthine phosphoribosyltransferase	g15002.t1, g17096.t1, g2007.t1, g372.t1	4
2.7.1.20: adenosine kinase	g12038.t1, g3551.t1, g6389.t1	3
2.7.1.25: adenylyl-sulfate kinase	g8333.t1, g2374.t1, g17468.t1	3
2.7.1.40: pyruvate kinase	g4450.t1, g7810.t1, g7811.t1, g13951.t1, g13651.t1, g1953.t1, g3035.t1	7
2.7.7.4: sulfate adenylyltransferase	g2374.t1, g13133.t1	2
2.7.7.6: DNA-directed RNA polymerase; RNA polymerase; RNA nucleotidyltransferase (DNA-directed); RNA polymerase I; RNA polymerase II; RNA polymerase III; C RNA formation factors; deoxyribonucleic acid-dependent ribonucleic acid polymerase; DNA-dependent ribonucleate nucleotidyltransferase; DNA-dependent RNA nucleotidyltransferase; DNA-dependent RNA polymerase;	g14482.t1, g1531.t1, g1532.t1, g4726.t1, g16589.t1, g1891.t1, g9211.t1, g9212.t1, g16532.t1, g10236.t1, g18457.t1, g15754.t1, g19351.t1, g19545.t1, g7252.t1, g1283.t1, g8730.t1, g18584.t1, g2323.t1, g10447.t1, g5153.t1, g3168.t1, g20444.t1, g5828.t1, g15829.t1, g12096.t1, g18256.t1, g9055.t1, g14088.t1, g7800.t1	30
2.7.7.7: DNA-directed DNA polymerase; DNA polymerase I; DNA polymerase II; DNA polymerase III; DNA polymerase alpha; DNA polymerase beta; DNA polymerase gamma; DNA nucleotidyltransferase (DNA-directed); DNA nucleotidyltransferase (DNA-directed); deoxyribonucleate nucleotidyltransferase; deoxynucleate polymerase; deoxyribonucleic	g12267.t1, g9894.t1, g15977.t1, g854.t1, g3400.t1, g17205.t1, g19366.t1, g5761.t1, g7237.t1, g14931.t1, g5760.t1, g9686.t1, g17850.t1, g11189.t1, g18621.t1, g3464.t1, g4640.t1, g2546.t1, g8974.t1, g9037.t1, g9038.t1, g9056.t1, g2713.t1, g5671.t1,	34

Enzymes	Sequence Names	Number of genes
acid duplicase; deoxyribonucleic acid polymerase; deoxyribonucleic duplicase; deoxyribonucleic polymerase; deoxyribonucleic polymerase I	g4661.t1, g7831.t1, g11609.t1, g6682.t1, g18957.t1, g7922.t1, g9536.t1, g1812.t1, g11075.t1, g16276.t1	
2.7.7.8: polyribonucleotide nucleotidyltransferase; polynucleotide phosphorylase;	g16455.t1, g7714.t1	2
3.1.4.17: 3',5'-cyclic-nucleotide phosphodiesterase; cyclic 3',5'-mononucleotide phosphodiesterase; PDE; cyclic 3',5'-nucleotide phosphodiesterase; cyclic 3',5'-phosphodiesterase; 3',5'-nucleotide phosphodiesterase; 3':5'-cyclic nucleotide 5'-nucleotidohydrolase; 3',5'-cyclonucleotide phosphodiesterase;	g10974.t1, g15075.t1, g18884.t1, g2808.t1, g3208.t1, g9101.t1, g19192.t1, g2756.t1, g9244.t1, g13561.t1, g7175.t1, g844.t1, g19192.t1, g2756.t1, g9244.t1, g874.t1, g844.t1, g810.t1, g9484.t1, g9483.t1, g17640.t1, g10502.t1	22
3.1.4.53: 3',5'-cyclic-AMP phosphodiesterase; cAMP-specific phosphodiesterase;	g5868.t1, g2047.t1	2
3.2.2.2: inosine nucleosidase; inosinase; inosine-guanosine nucleosidase	g3292.t1	1
3.5.4.10: IMP cyclohydrolase; inosinicase; inosinate cyclohydrolase	g8877.t1, g3736.t1	2
3.5.4.3: guanine deaminase; guanase; guanine aminase;	g3919.t1, g4901.t1	2
3.5.4.4: adenosine deaminase; deoxyadenosine deaminase	g13973.t1, g17512.t1, g15531.t1	3
3.5.4.6: AMP deaminase; adenylic acid deaminase; AMP aminase; adenylic deaminase;	g13078.t1, g19884.t1, g6019.t1, g5425.t1, g2504.t1, g3274.t1	6
3.6.1.11: exopolyphosphatase;	g10650.t1	1
3.6.1.13: ADP-ribose diphosphatase;	g19921.t1, g5405.t1, g13021.t1, g15536.t1	4
3.6.1.15: nucleoside-triphosphate phosphatase; nucleoside-triphosphatase; nucleoside triphosphate phosphohydrolase; nucleoside-5-triphosphate phosphohydrolase; nucleoside 5-triphosphatase; unspecific diphosphate phosphohydrolase	g9812.t1, g1327.t1, g3113.t1, g8960.t1, g15605.t1, g7850.t1, g6705.t1, g471.t1, g14654.t1, g3112.t1, g16627.t1, g13530.t1, g3776.t1, g2002.t1, g2948.t1, g18609.t1, g12851.t1, g2000.t1, g470.t1, g12985.t1, g9528.t1, g13636.t1, g2349.t1, g14780.t1, g15395.t1, g6913.t1, g2593.t1, g15603.t1, g6704.t1, g12848.t1, g4208.t1, g1975.t1, g8144.t1, g9255.t1, g15398.t1, g6856.t1, g15604.t1, g7354.t1, g6078.t1, g3603.t1, g2476.t1, g14451.t1	42
3.6.1.19: nucleoside-triphosphate diphosphatase;	g602.t1, g2594.t1, g593.t1	3
3.6.1.29: bis(5'-adenosyl)-triphosphatase;	g2586.t1, g4157.t1	2

Enzymes	Sequence Names	Number of genes
dinucleosidetriphosphatase;		
3.6.1.3: adenosinetriphosphatase; adenylpyrophosphatase; ATP monophosphatase; triphosphatase; ATPase (ambiguous); SV40 T-antigen; adenosine 5'-triphosphatase; ATP hydrolase; complex V (mitochondrial electron transport); (Ca2+ + Mg2+)-ATPase; HCO3--ATPase; adenosine triphosphatase	g12016.t1, g19880.t1, g390.t1, g15112.t1, g8145.t1, g1611.t1, g16364.t1, g6390.t1, g9293.t1, g185.t1, g6108.t1, g3614.t1, g13199.t1, g19951.t1, g11723.t1, g5342.t1, g4867.t1, g5707.t1, g522.t1, g2365.t1, g18585.t1, g18611.t1, g4158.t1, g6180.t1, g10651.t1, g13615.t1, g12494.t1, g13528.t1, g3405.t1, g9507.t1, g6232.t1, g14297.t1, g12513.t1, g18627.t1, g4438.t1, g13534.t1, g11070.t1, g2001.t1, g15393.t1, g18610.t1, g16212.t1, g8815.t1, g12176.t1, g15518.t1, g14876.t1, g7064.t1, g2937.t1, g15399.t1, g13531.t1, g15396.t1, g14420.t1, g10652.t1, g2096.t1, g16211.t1, g6233.t1, g4869.t1, g19233.t1, g8195.t1, g468.t1, g18264.t1, g9763.t1, g18103.t1, g6347.t1, g15394.t1, g1977.t1, g15397.t1, g6703.t1, g20698.t1, g469.t1, g4215.t1, g12664.t1, g12178.t1, g13529.t1, g6183.t1, g12850.t1, g15667.t1, g12035.t1, g13405.t1, g3042.t1, g4935.t1, g14522.t1, g18354.t1, g9462.t1, g2707.t1, g9835.t1, g3913.t1, g3492.t1, g639.t1, g1287.t1, g9994.t1, g7419.t1, g12255.t1, g1785.t1, g9071.t1, g18257.t1, g1175.t1, g7935.t1, g5143.t1, g5142.t1, g10141.t1, g2092.t1, g14995.t1, g10340.t1, g15553.t1, g10676.t1, g2743.t1, g12538.t1, g2811.t1, g983.t1, g16137.t1, g7398.t1, g9845.t1, g19466.t1, g9712.t1, g5688.t1, g20053.t1, g5689.t1, g5687.t1, g10677.t1, g2228.t1, g16059.t1, g8137.t1, g16058.t1, g4058.t1, g1174.t1, g4059.t1, g9481.t1, g9482.t1, g5647.t1, g8786.t1, g7952.t1, g11137.t1, g12053.t1	133

Enzymes	**Sequence Names**	**Number of genes**
3.6.1.41: bis(5'-nucleosyl)-tetraphosphatase (symmetrical);	g1078.t1	1
3.6.1.5: apyrase; ATP-diphosphatase;	g9835.t1	1
3.6.1.6: nucleoside diphosphate phosphatase;	g8837.t1, g9293.t1	2
4.6.1.1: adenylate cyclase; adenylylcyclase; adenyl cyclase; 3',5'-cyclic AMP synthetase; ATP diphosphate-lyase (cyclizing)	g10142.t1, g15906.t1, g10333.t1, g14367.t1, g7060.t1, g5488.t1, g14994.t1, g13232.t1, g5558.t1	9
4.6.1.2: guanylate cyclase; guanylyl cyclase; guanyl cyclase; GTP diphosphate-lyase (cyclizing)	g10334.t1, g9027.t1, g17610.t1	3
5.4.2.2: phosphoglucomutase (alpha-D-glucose-1,6-bisphosphate-dependent);	g3214.t1, g13858.t1	2
3.1.3.5 5:'-nucleotidase; uridine 5'-nucleotidase; 5'-adenylic phosphatase;	g18957.t1, g7714.t1, g17652.t1, g8859.t1, g4447.t1	5
Pyrimidine nucleotide synthesis		
Uridine monophosphate synthesis (precursor)		
6.3.5.5: carbamoyl-phosphate synthase	g647.t1, g645.t1, g11082.t1, g14530.t1, g9630.t1, g7317.t1	6
2.1.3.2: aspartate carbamoyltransferase	g11723.t1, g645.t1, g11082.t1, g14530.t1, g9630.t1, g7317.t1, g8083.t1	7
3.5.2.3: dihydroorotase	g14586.t1, g14587.t1, g645.t1, g14302.t1, g11082.t1, g14530.t1, g9630.t1	7
1.3.5.2: dihydroorotate dehydrogenase	g6389.t1	1
2.4.2.10: orotate phosphoribosyltransferase	g19813.t1, g19814.t1, g12085.t1, g9000.t1	4
4.1.1.23: orotidine-5'-monophosphate decarboxylase	g12085.t1, g11791.t1	2
UDP, UTP, CTP and CDP synthesis		
2.7.4.4: nucleoside phosphate kinase	--	--
2.7.4.6: nucleoside-diphosphate kinase	g14086.t1, g7640.t1, g6274.t1, g108.t1, g14746.t1	5
6.3.4.2: CTP synthase	g3467.t1, g136.t1	2
dCDP, dCTP, dUDP and dUTP synthesis		
1.17.4.1: ribonucleotide reductase, class II	g1133.t1, g9496.t1, g10928.t1, g8955.t1, g3413.t1, g3231.t1, g12136.t1, g7603.t1	8
2.7.4.6: nucleoside-diphosphate kinase	g14086.t1, g7640.t1, g6274.t1, g108.t1, g14746.t1	5
dTMP, dTDP and dTTP synthesis		
2.7.4.4 nucleoside phosphate kinase	--	--
2.1.1.45: thymidylate synthase	g3876.t1, g10928.t1	2
2.7.4.9: dTMP kinase	g5946.t1, g986.t1, g352.t1, g11606.t1	4

Enzymes	Sequence Names	Number of genes
2.7.4.6: nucleoside-diphosphate kinase	g14086.t1, g7640.t1, g6274.t1, g108.t1, g14746.t1	5
Other enzymes that might be involved in pyrimidine synthesis and metabolism		
1.8.1.9: thioredoxin-disulfide reductase; NADP-thioredoxin reductase; NADPH-thioredoxin reductase; thioredoxin reductase (NADPH);	g1067.t1, g19251.t1, g11401.t1, g7370.t1, g16727.t1, g20723.t1, g16238.t1	7
2.4.2.1: purine-nucleoside phosphorylase;	g8354.t1, g6545.t1	2
2.4.2.3: uridine phosphorylase;	g19470.t1, g3885.t1	2
2.4.2.9: uracil phosphoribosyltransferase;	g5319.t1, g539.t1, g2635.t1	3
2.7.1.48: uridine kinase	g18950.t1, g9372.t1, g9373.t1, g2130.t1, g2635.t1	5
2.7.4.14: UMP/CMP kinase; cytidylate kinase; deoxycytidylate kinase;	g6807.t1, g7627.t1, g18723.t1, g11828.t1	4
2.7.7.6: DNA-directed RNA polymerase; RNA polymerase; RNA nucleotidyltransferase (DNA-directed); RNA polymerase I; RNA polymerase II; RNA polymerase III; C RNA formation factors; deoxyribonucleic acid-dependent ribonucleic acid polymerase; DNA-dependent ribonucleate nucleotidyltransferase; DNA-dependent RNA nucleotidyltransferase; DNA-dependent RNA polymerase; ribonucleate nucleotidyltransferase;	g14482.t1, g1531.t1, g1532.t1, g4726.t1, g16589.t1, g1891.t1, g9211.t1, g9212.t1, g16532.t1, g10236.t1, g18457.t1, g15754.t1, g19351.t1, g19545.t1, g7252.t1, g1283.t1, g8730.t1, g18584.t1, g2323.t1, g10447.t1, g5153.t1, g3168.t1, g20444.t1, g5828.t1, g15829.t1, g12096.t1, g18256.t1, g9055.t1, g14088.t1, g7800.t1	30
2.7.7.7: DNA-directed DNA polymerase; DNA polymerase I; DNA polymerase II; DNA polymerase III; DNA polymerase alpha; DNA polymerase beta; DNA polymerase gamma; DNA nucleotidyltransferase (DNA-directed); DNA nucleotidyltransferase (DNA-directed); deoxyribonucleate nucleotidyltransferase; deoxynucleate polymerase; deoxyribonucleic acid duplicase; deoxyribonucleic acid polymerase; deoxyribonucleic duplicase; deoxyribonucleic polymerase; deoxyribonucleic polymerase I; DNA duplicase;	g12267.t1, g9894.t1, g15977.t1, g854.t1, g3400.t1, g17205.t1, g19366.t1, g5761.t1, g7237.t1, g14931.t1, g5760.t1, g9686.t1, g17850.t1, g11189.t1, g18621.t1, g3464.t1, g4640.t1, g2546.t1, g8974.t1, g9037.t1, g9038.t1, g9056.t1, g2713.t1, g5671.t1, g4661.t1, g7831.t1, g11609.t1, g6682.t1, g18957.t1, g7922.t1, g9536.t1, g1812.t1, g11075.t1, g16276.t1	34
2.7.7.8: polyribonucleotide nucleotidyltransferase;	g16455.t1, g7714.t1	2
3.1.3.5 5:'-nucleotidase; uridine 5'-nucleotidase;	g18957.t1, g7714.t1, g17652.t1, g8859.t1, g4447.t1	5
3.2.2.3: uridine nucleosidase; uridine hydrolase	g14928.t1, g3292.t1	2
3.5.4.12: dCMP deaminase; deoxycytidylate deaminase; deoxy-CMP-deaminase;	g2753.t1, g2586.t1, g10378.t1, g593.t1	4
3.6.1.19: nucleoside-triphosphate diphosphatase;	g602.t1, g2594.t1, g593.t1	3
3.6.1.23: dUTP diphosphatase;	g4127.t1, g5812.t1	2
3.6.1.5: apyrase; ATP-diphosphatase;	g9835.t1	1
3.6.1.6: nucleoside diphosphate phosphatase;	g8837.t1, g9293.t1	2

Enzymes	Sequence Names	Number of genes
Amino Acid Biosynthesis		
Tyrosine (Y), Phenylalanine (F), Tryptopham (W) synthesis		
Chorismate synthesis (precursor)		
2.5.1.54: 3-deoxy-7-phosphoheptulonate synthase	g3103.t1, g8371.t1	2
4.2.3.4: 3-dehydroquinate synthase	g4426.t1	1
4.2.1.10: 3-dehydroquinate dehydratase I	g4426.t1	1
1.1.1.25: Shikimate dehydrogenase	g4426.t1	1
2.7.1.71: Shikimate kinase	g4426.t1	1
2.5.1.19: 3-phosphoshikimate 1-carboxyvinyltransferase	g4426.t1	1
4.2.3.5: Chorismate synthase	g6770.t1	1
Y synthesis		
5.4.99.5: Chorismate mutase	--	--
1.3.1.13: Prephenate dehydrogenase (NADP+)	--	--
2.6.1.1: Aspartate aminotransferase, cytoplasmic	g13273.t1, g13812.t1, g15236.t1, g18784.t1, g18785.t1, g20207.t1, g4650.t1, g4792.t1, g5612.t1, g8199.t1, g8358.t1	11
F synthesis		
5.4.99.5: Chorismate mutase	--	--
4.2.1.51: Prephenate dehydratase	g8371.t1	1
2.6.1.1: Aspartate aminotransferase, cytoplasmic	g13273.t1, g13812.t1, g15236.t1, g18784.t1, g18785.t1, g20207.t1, g4650.t1, g4792.t1, g5612.t1, g8199.t1, g8358.t1	11
W synthesis		
4.1.3.27: Anthranilate synthase	g6970.t1	1
2.4.2.18: Anthranilate phosphoribosyltransferase	g13589.t1 (trpD)	1
5.3.1.24: Phosphoribosylanthranilate isomerase		
4.1.1.48: Indole-3-glycerol phosphate synthase	g13589.t1	1
4.2.1.20: Tryptophan synthase alpha chain	g13589.t1, g15906.t1	2
Serine (S), Glycine (G), Threonien (T) synthesis		
S synthesis from 3-phosphoglycerate		
1.1.1.95: D-3-phosphoglycerate dehydrogenase	g7538.t1, g8967.t1	2
2.6.1.52: Phosphoserine aminotransferase	g11274.t1, g11695.t1, g16539.t1, g16727.t1, g17487.t1, g1993.t1, g3685.t1, g6496.t1	8
3.1.3.3: Phosphoserine phosphatase	g17330.t1, g8974.t1, g9804.t1	3
G synthesis from serine		
2.1.2.1: Glycine hydroxymethyltransferase	g12557.t1, g12987.t1, g13120.t1, g16051.t1, g16632.t1, g4150.t1, g5450.t1, g5451.t1	8
T synthesis from L-aspartate		

Enzymes	**Sequence Names**	**Number of genes**
2.7.2.4: Aspartate kinase	g14757.t1, g2286.t1, g2287.t1, g330.t1, g5534.t1, g8709.t1, g9518.t1, g9723.t1	8
1.2.1.11: Aspartate-semialdehyde dehydrogenase	g10051.t1, g11148.t1, g1591.t1, g3078.t1, g538.t1	5
1.1.1.3: Homoserine dehydrogenase	g11082.t1, g13797.t1, g14757.t1, g16165.t1, g2286.t1, g2287.t1, g8123.t1, g9723.t1	8
2.7.1.39: Homoserine kinase	g2239.t1, g8768.t1	2
4.2.3.1: Threonine synthase	g14064.t1, g14474.t1, g15894.t1, g9263.t1	4
Arginine (R), Proline (P) synthesis		
R synthesis from glutamate		
2.3.1.1: Amino-acid N-acetyltransferase	--	--
2.7.2.8: Acetylglutamate kinase	--	--
1.2.1.38: N-acetyl-gamma-glutamyl-phosphate reductase	g1591.t1, g538.t1	2
2.6.1.11: Acetylornithine aminotransferase	g9949.t1	1
3.5.1.16: Acetylornithine deacetylase	g11470.t1, g15979.t1, g5682.t1, g7835.t1	4
2.1.3.3: Ornithine carbamoyltransferase (urea cycle)	g11723.t1	1
6.3.4.5: Argininosuccinate synthase (urea cycle)	--	--
4.3.2.1: Argininosuccinate lyase (urea cycle)	g7980.t1	1
P synthesis from glutamate		
2.7.2.11: Glutamate 5-kinase	g12819.t1, g3326.t1	2
1.2.1.41: Glutamate-5-semialdehyde dehydrogenase	g4536.t1	1
1.5.1.2: Pyrroline-5-carboxylate reductase	g15260.t1, g6808.t1	2
Alanine (A), Aspartic acid (D), Asparginine (N), Glutamic acid (E), Glutamine (Q) synthesis		
A synthesis from pyruvate		
2.6.1.44: Alanine-glyoxylate transaminase	g11401.t1, g12516.t1, g16262.t1, g3045.t1, g7370.t1, g9503.t1	6
D synthesis from oxaloacetate		
2.6.1.1: Aspartate aminotransferase, cytoplasmic	g5612.t1, g16262.t1, g11161.t1, g19987.t1, g4792.t1, g18784.t1, g7370.t1, g18785.t1	8
N synthesis from aspartate		
6.3.5.4: Asparagine synthase (glutamine-hydrolysing)	g5566.t1, g6052.t1, g413.t1, g14518.t1, g10842.t1, g3142.t1, g8052.t1	7
E synthesis from 2-oxoglutarate		
1.4.1.13 /1.4.1.14: Glutamate synthase NADPH/ Glutamate synthase NADH	g5164.t1, g3030.t1, g5164.t1	3
Q synthesis from glutamate		

Enzymes	Sequence Names	Number of genes
6.3.1.2: Glutamine synthetase	g13643.t1, g17262.t1, g3340.t1, g3575.t1, g5100.t1, g6410.t1, g8513.t1	7
Cysteine (C) , Methionine (M) synthesis		
C synthesis from serine		
2.3.1.30: Serine O-acetyltransferase	--	--
2.5.1.47: Cysteine synthase A	g3030.t1	1
M synthesis from aspartate		
2.7.2.4: Aspartate kinase	g14757.t1, g2286.t1, g2287.t1, g330.t1, g5534.t1, g8709.t1, g9518.t1, g9723.t1	8
1.2.1.11: Aspartate-semialdehyde dehydrogenase	g10051.t1, g11148.t1, g1591.t1, g3078.t1, g538.t1	5
1.1.1.3: Homoserine dehydrogenase	g11082.t1, g13797.t1, g14757.t1, g16165.t1, g2286.t1, g2287.t1, g8123.t1, g9723.t1	8
2.3.1.31: Homoserine O-acetyltransferase	--	--
2.5.1.48: Cystathionine gamma-synthase	g16732.t1	1
4.4.1.8: Cystathionine beta-lyase	g5174.t1, g9342.t1	2
2.1.1.5/2.1.1.10/2.1.1.13/2.1.1.14: Betaine-homocysteine S-methyltransferase /Homocysteine S-methyltransferase/5-methyltetrahydrofolate--homocysteine methyltransferase /5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	g1133.t1, g4961.t1, g8955.t1	3
Valine (V), Leucine (L), Isoleucine (I) synthesis		
V synthesis from pyruvate		
2.2.1.6: Acetolactate synthase I/II/III large subunit	g14751.t1	1
1.1.1.86: Ketol-acid reductoisomerase	--	--
4.2.1.9: Dihydroxy-acid dehydratase	--	--
2.6.1.42: Branched-chain amino acid aminotransferase	g1847.t1, g8320.t1	2
L synthesis from pyruvate		
2.2.1.6: Acetolactate synthase I/II/III large subunit	g14751.t1	1
1.1.1.86: Ketol-acid reductoisomerase	--	--
4.2.1.9: Dihydroxy-acid dehydratase	--	--
2.3.3.13: 2-isopropylmalate synthase	--	--
4.2.1.33: 3-isopropylmalate dehydratase	--	--
1.1.1.85: 3-isopropylmalate dehydrogenase	--	--
2.6.1.42: Branched-chain amino acid aminotransferase	g1847.t1, g8320.t1	2
I synthesis from 2-oxobutanoate		

Enzymes	Sequence Names	Number of genes
2.2.1.6: Acetolactate synthase I/II/III large subunit	g14751.t1	1
1.1.1.86: Ketol-acid reductoisomerase		
4.2.1.9: Dihydroxy-acid dehydratase		
2.6.1.42: Branched-chain amino acid aminotransferase	g1847.t1, g8320.t1	2
Lysine (K) synthesis		
K synthesis from aspartate		
2.7.2.4: Aspartate kinase	g14757.t1, g2286.t1, g2287.t1, g330.t1, g5534.t1, g8709.t1, g9518.t1, g9723.t1	8
1.2.1.11: Aspartate-semialdehyde dehydrogenase	g10051.t1, g11148.t1, g1591.t1, g3078.t1, g538.t1	5
4.2.1.52: Dihydrodipicolinate synthase		
1.3.1.26: Dihydrodipicolinate reductase		
2.3.1.117: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase		
2.6.1.17: N-succinyldiaminopimelate aminotransferase	g10836.t1	1
3.5.1.18: Succinyl-diaminopimelate desuccinylase		
5.1.1.7: Diaminopimelate epimerase		
4.1.1.20: Diaminopimelate decarboxylase	g12946.t1	1
Histidine (H) synthesis		
H synthesis from PRPP		
2.4.2.17: ATP phosphoribosyltransferase		
3.6.1.31: Phosphoribosyl-ATP pyrophosphohydrolase	g21.t1	1
3.5.4.19: Phosphoribosyl-AMP cyclohydrolase		
5.3.1.16: Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase	--	--
2.6.1.16: Glutamine amidotransferase	g6052.t1, g5164.t1	2
4.2.1.19: Imidazoleglycerol-phosphate dehydratase		
2.6.1.9: Histidinol-phosphate aminotransferase	g13273.t1, g13812.t1, g15236.t1, g20207.t1, g4650.t1, g8199.t1, g8358.t1	7
3.1.3.15: Histidinol-phosphatase	g9452.t1	1
1.1.1.23: histidinol dehydrogenase		
Other enzymes that might be involved in amino acid biosynthesis		
1.1.1.87: aksF; methanogen homoisocitrate dehydrogenase	g10497.t1, g7376.t1	2
1.1.1.42: IDH1, IDH2, icd; isocitrate dehydrogenase	g19124.t1, g629.t1, g7997.t1, g9270.t1	4
1.14.16.1: phhA, PAH; phenylalanine-4-hydroxylase	g10912.t1, g9123.t1	2

Enzymes	Sequence Names	Number of genes
1.2.1.-: lysY; N-acetyl-gamma-glutamyl-phosphate/LysW-gamma-L-alpha-aminoadipyl-6-phosphate reductase	g17342.t1, g17343.t1, g2816.t1, g7597.t1	4
1.2.1.12: GAPDH, gapA; glyceraldehyde 3-phosphate dehydrogenase	g1268.t1, g15719.t1, g5105.t1, g5318.t1	4
1.2.1.3: ALDH7A1; aldehyde dehydrogenase family 7 member A1	g10929.t1, g15115.t1, g405.t1, g4882.t1, g5125.t1, g6908.t1, g7274.t1, g8354.t1, g9892.t1	9
1.2.1.31: ALDH7A1; aldehyde dehydrogenase family 7 member A1	g20352.t1, g5125.t1, g15116.t1, g2093.t1, g2093.t1, g5125.t1	6
2.1.1.-: yrrT; putative AdoMet-dependent methyltransferase	g10472.t1, g12038.t1, g13735.t1, g1577.t1, g19021.t1, g20152.t1, g3051.t1, g3395.t1, g4450.t1, g4942.t1, g5375.t1, g5638.t1, g6685.t1, g7116.t1, g7694.t1, g8036.t1, g8150.t1, g8333.t1, g8768.t1, g9076.t1	20
2.2.1.1: tktA, tktB; transketolase	g5062.t1, g7241.t1, g7317.t1	3
2.2.1.2: tal-pgi; transaldolase / glucose-6-phosphate isomerase	g12919.t1, g13038.t1, g15227.t1, g265.t1	4
2.3.3.1: citrate (Si)-synthase	g12081.t1, g13345.t1, g16719.t1, g18242.t1, g19126.t1, g474.t1, g477.t1, g5893.t1, g6197.t1, g6553.t1, g8175.t1	11
2.5.1.6: metK; S-adenosylmethionine synthetase	g12039.t1, g13225.t1, g15738.t1, g17234.t1, g6035.t1, g6739.t1	6
LYSN; 2-aminoadipate transaminase patA; aminotransferase lysJ; acetylornithine/LysW-gamma-L-lysine aminotransferase GGAT;	g11016.t1, g3104.t1, g4982.t1, g4983.t1, g6033.t1, g6229.t1	6
2.6.1.2; GPT, ALT; alanine transaminase alaA; alanine-synthesizing transaminase GGAT;	g13226.t1, g2414.t1	2
2.6.1.2: TAT; tyrosine aminotransferase ARO8; aromatic amino acid aminotransferase I	g12493.t1, g13273.t1, g13812.t1, g15236.t1, g20207.t1, g4650.t1, g8199.t1, g8358.t1	8
2.6.1.57: ARO8; aromatic amino acid aminotransferase I tyrB; aromatic-amino-acid transaminase	g13273.t1, g13812.t1, g15236.t1, g20207.t1, g4650.t1, g8199.t1, g8358.t1	7
2.6.1.7: AADAT, KAT2; kynurenine/2-aminoadipate aminotransferase	g10836.t1	1
2.6.1.83: LL-diaminopimelate aminotransferase	g15640.t1	1
2.7.1.11: pfkA, PFK; 6-phosphofructokinase 1 pfkB; 6-phosphofructokinase 2	g11905.t1, g13580.t1, g16188.t1, g3565.t1, g4532.t1	5
2.7.1.40: PK, pyk; pyruvate kinase	g13651.t1, g13951.t1, g1953.t1, g3035.t1, g4450.t1, g7810.t1, g7811.t1	7
2.7.2.3: PGK, pgk; phosphoglycerate kinase	g16261.t1, g6745.t1	2

Enzymes	Sequence Names	Number of genes
2.7.6.1: PRPS, prsA; ribose-phosphate pyrophosphokinase	g13783.t1, g14804.t1, g16201.t1, g18249.t1, g19930.t1, g5424.t1, g9663.t1	7
3.1.3.11: K01622; fructose 1,6-bisphosphate aldolase/phosphatase	g14101.t1, g14685.t1, g17520.t1, g17521.t1, g8034.t1	5
3.2.2.9: K18284; adenosylhomocysteine /aminodeoxyfutalosine nucleosidase	g3893.t1, g8334.t1	2
3.5.1.14: ACY1; aminoacylase	g1045.t1, g16485.t1, g18485.t1, g4698.t1	4
3.5.3.1: rocF, arg; arginase	g17079.t1, g4127.t1	2
4.1.2.13: K16306; fructose-bisphosphate aldolase / 2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase	g1265.t1, g3323.t1	2
4.1.2.5: ltaE; threonine aldolase	g799.t1	1
4.3.2.10: hisF; cyclase HIS7; glutamine amidotransferase / cyclase	g11320.t1	1
4.2.1.11: ENO, eno; enolase	g10562.t1, g4531.t1	2
2.5.1.47: cysO; cysteine synthase / O-phosphoserine sulfhydrylase / cystathionine beta-synthase	g13209.t1, g9095.t1	2
4.2.1.3: ACO, acnA; aconitate hydratase	g10767.t1, g12236.t1, g12237.t1, g12238.t1, g12259.t1, g14059.t1, g15984.t1, g15985.t1, g17251.t1, g18949.t1, g2848.t1, g4475.t1, g9355.t1, g9467.t1	14
4.2.1.99: acnB; aconitate hydratase 2 / 2-methylisocitrate dehydratase	g10767.t1, g12236.t1, g12237.t1, g12238.t1, g14059.t1, g17251.t1, g2848.t1, g9467.t1	8
4.3.1.12: ocd; ornithine cyclodeaminase	g18055.t1, g16049.t1, g20353.t1, g7203.t1	4
4.3.1.19: ilvA, tdcB; threonine dehydratase	g16049.t1	1
4.4.1.1: mccB; cystathionine gamma-lyase / homocysteine desulfhydrase	g5175.t1	1
4.4.1.8: metC; cystathionine beta-lyase	g14385.t1	1
5.1.3.1: rpe, RPE; ribulose-phosphate 3-epimerase	g11989.t1, g16002.t1	2
5.3.1.1: TPI, tpiA; triosephosphate isomerase (TIM)	g2077.t1	1
5.3.1.6: rpiB; ribose 5-phosphate isomerase B	g10558.t1, g11785.t1, g12957.t1, g565.t1	4
2.2.1.2: tal-pgi; transaldolase / glucose-6-phosphate isomerase	g15821.t1, g3214.t1	2
5.4.2.11: PGAM, gpmA; 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	g11930.t1	1
5.4.2.12: gpmI; 2,3-bisphosphoglycerate-	g6352.t1	1

Enzymes	Sequence Names	Number of genes
independent phosphoglycerate mutase		
Fatty acid biosynthesis		
1.1.1.100: 3-oxoacyl-[acyl-carrier-protein] reductase;	g10497.t1, g14059.t1, g3962.t1, g5691.t1	4
1.3.1.10: enoyl-[acyl-carrier-protein] reductase (NADPH, Si-specific);	g5691.t1	1
2.3.1.39: [acyl-carrier-protein] S-malonyltransferase;	g5691.t1	1
2.3.1.41: beta-ketoacyl-[acyl-carrier-protein] synthase I	g5691.t1	1
2.3.1.85: fatty-acid synthase; FASN (gene name)	g12138.t1	1
3.1.2.14: oleoyl-[acyl-carrier-protein] hydrolase	g5691.t1, g5691.t1, g5691.t1	3
4.2.1.59: 3-hydroxyacyl-[acyl-carrier-protein] dehydratase;	g5691.t1	1
6.2.1.3: long-chain-fatty-acid---CoA ligase;	g7296.t1, g7880.t1, g7882.t1, g10624.t1, g6814.t1, g8060.t1, g7294.t1, g14826.t1, g8513.t1, g6410.t1, g17262.t1, g13748.t1	12
6.4.1.2: acetyl-CoA carboxylase; acetyl coenzyme A carboxylase	g16665.t1, g645.t1	2
Fatty acid elongation		
1.1.1.330: very-long-chain 3-oxoacyl-CoA reductase;	g7376.t1	1
1.1.1.35: 3-hydroxyacyl-CoA dehydrogenase;	g10565.t1, g6421.t1	2
2.3.1.16: acetyl-CoA C-acyltransferase;	g20084.t1, g12162.t1, g3411.t1, g9558.t1	4
2.3.1.199: very-long-chain 3-oxoacyl-CoA synthase;	g1887.t1, g8424.t1	2
3.1.2.22: palmitoyl[protein] hydrolase;	g3997.t1, g11079.t1	2
4.2.1.17: enoyl-CoA hydratase	g5758.t1, g10565.t1, g5174.t1	3
Fatty acid degradation		
1.1.1.1: alcohol dehydrogenase; aldehyde reductase; ADH;	g17476.t1, g17477.t1, g17478.t1, g8973.t1, g1813.t1, g14245.t1, g3976.t1	7
1.1.1.35: 3-hydroxyacyl-CoA dehydrogenase;	g10565.t1, g6421.t1	2
1.14.14.1: unspecific monooxygenase;	g15444.t1	1
1.2.1.3: aldehyde dehydrogenase (NAD+);	g10929.t1, g15115.t1, g6908.t1, g9892.t1, g405.t1, g7274.t1, g5125.t1, g8354.t1, g4882.t1	9
1.2.1.48: long-chain-aldehyde dehydrogenase;	g7274.t1	1
1.3.3.6: acyl-CoA oxidase;	g15580.t1, g8950.t1, g8951.t1, g8952.t1, g1484.t1, g4622.t1, g17278.t1, g2229.t1	8
1.3.8.1: short-chain acyl-CoA dehydrogenase;	g323.t1, g18766.t1, g4114.t1, g7262.t1	4

Enzymes	Sequence Names	Number of genes
1.3.8.6: glutaryl-CoA dehydrogenase (ETF);	g15267.t1, g16656.t1	2
2.3.1.16: acetyl-CoA C-acyltransferase; beta-ketothiolase; 3-ketoacyl-CoA thiolase;	g20084.t1, g12162.t1, g3411.t1, g9558.t1	4
2.3.1.9: acetyl-CoA C-acetyltransferase;	g15576.t1, g4193.t1	2
4.2.1.17: enoyl-CoA hydratase;	g5758.t1, g10565.t1, g5174.t1	3
5.3.3.8: dodecenoyl-CoA isomerase;	g19464.t1, g10565.t1, g17157.t1	3
6.2.1.3: long-chain-fatty-acid---CoA ligase; acyl-CoA synthetase; fatty acid thiokinase (long chain); acyl-activating enzyme; palmitoyl-CoA synthase; lignoceroyl-CoA synthase;	g7296.t1, g7880.t1, g7882.t1, g10624.t1, g6814.t1, g8060.t1, g7294.t1, g14826.t1, g8513.t1, g6410.t1, g17262.t1, g13748.t1	12
Pyruvate metabolism		
1.1.1.37: malate dehydrogenase; malic dehydrogenase; L-malate dehydrogenase; NAD-L-malate dehydrogenase; malic acid dehydrogenase; NAD-dependent malic dehydrogenase;	g17707.t1, g3242.t1, g10767.t1, g17251.t1, g2848.t1, g4769.t1, g2054.t1, g14670.t1, g19916.t1, g4492.t1	10
1.1.1.38: malate dehydrogenase (oxaloacetate-decarboxylating); 'malic' enzyme (ambiguous);	g16222.t1, g16220.t1, g12521.t1, g1030.t1, g8330.t1	5
1.1.1.40: malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+); 'malic' enzyme (ambiguous);	g18320.t1, g20352.t1, g20353.t1, g20354.t1, g9873.t1, g16222.t1, g16220.t1, g12521.t1, g1030.t1, g8330.t1	10
1.1.1.79: glyoxylate reductase (NADP+);	g16491.t1	1
1.1.2.3: L-lactate dehydrogenase (cytochrome);	g2637.t1, g613.t1, g15576.t1, g7475.t1, g6734.t1, g7285.t1, g4136.t1	7
1.1.5.4: malate dehydrogenase (quinone); FAD-dependent malate-vitamin K reductase;	g3976.t1, g9723.t1, g4347.t1, g13513.t1, g4347.t1, g13513.t1	6
1.2.1.3: aldehyde dehydrogenase (NAD+); CoA-independent aldehyde dehydrogenase; m-methylbenzaldehyde dehydrogenase;	g10929.t1, g15115.t1, g6908.t1, g9892.t1, g405.t1, g7274.t1, g5125.t1, g8354.t1, g4882.t1	9
1.2.4.1: pyruvate dehydrogenase (acetyl-transferring);	g4674.t1	1
1.8.1.4: dihydrolipoyl dehydrogenase	g2390.t1, g11161.t1, g7447.t1	3
2.3.1.12: dihydrolipoyllysine-residue acetyltransferase	g14980.t1, g467.t1, g872.t1	3
2.3.1.9: acetyl-CoA C-acetyltransferase	g15576.t1, g4193.t1	2
2.3.3.9: malate synthase; L-malate glyoxylate-lyase (CoA-acetylating);	g7986.t1, g7987.t1, g3061.t1, g3889.t1	4
2.7.1.40: pyruvate kinase; phosphoenolpyruvate kinase; phosphoenol transphosphorylase	g4450.t1, g7810.t1, g7811.t1, g13951.t1, g13651.t1, g1953.t1, g3035.t1	7
2.7.9.1: pyruvate, phosphate dikinase; pyruvate, orthophosphate dikinase;	g16026.t1, g2241.t1	2
3.1.2.6: hydroxyacylglutathione hydrolase;	g19845.t1, g14384.t1, g17478.t1	3
3.6.1.7: acylphosphatase; acetylphosphatase;	g8323.t1	1

Enzymes	Sequence Names	Number of genes
4.1.1.49: phosphoenolpyruvate carboxykinase (ATP)	g9192.t1, g834.t1, g12237.t1	3
4.2.1.2: fumarate hydratase; fumarase; L-malate hydro-lyase; (S)-malate hydro-lyase	g14808.t1, g13023.t1, g17892.t1, g9027.t1	4
4.4.1.5: lactoylglutathione lyase;	g13761.t1, g15637.t1	2
6.2.1.1: acetate---CoA ligase; acetyl-CoA synthetase;	g11834.t1, g4449.t1, g6814.t1, g7882.t1	4
6.4.1.2: acetyl-CoA carboxylase	g16665.t1, g645.t1	2
Endocytosis		
2.7.1.68: 1-phosphatidylinositol-4-phosphate 5-kinase; diphosphoinositide kinase; PIP kinase; phosphatidylinos	g16179.t1, g16854.t1, g2116.t1, g4883.t1, g5536.t1, g9051.t1, g11452.t1, g15684.t1, g8518.t1, g5196.t1, g232.t1, g10965.t1, g13746.t1, g13745.t1, g16770.t1, g18891.t1	16
2.7.10.2: non-specific protein-tyrosine kinase;	g5164.t1, g4826.t1	2
2.7.11.16: G-protein-coupled receptor kinase;	g13016.t1, g1025.t1, g3553.t1	3
3.1.2.15: ubiquitin thiolesterase; ubiquitin carboxy-terminal esterase	g1263.t1, g10673.t1, g6274.t1, g9008.t1, g13783.t1, g352.t1, g5424.t1, g9663.t1, g10674.t1, g17597.t1, g108.t1, g14746.t1, g15238.t1, g20.t1, g7322.t1, g12318.t1, g11078.t1, g14168.t1, g15144.t1, g19068.t1, g19447.t1, g10503.t1, g16249.t1, g12656.t1, g8042.t1, g17067.t1, g16252.t1, g16250.t1, g6727.t1, g7708.t1, g4168.t1, g8284.t1, g13912.t1, g4523.t1	34
3.6.5.5: dynamin GTPase	g10661.t1, g10662.t1, g12791.t1, g4857.t1, g9636.t1	5
6.3.2.19: ubiquitin---protein ligase; ubiquitin-activating enzyme	g13350.t1, g324.t1, g9376.t1, g12917.t1, g8263.t1, g15902.t1, g5100.t1, g17183.t1, g3340.t1, g3575.t1, g11395.t1, g20628.t1, g4994.t1, g6060.t1, g9284.t1, g149.t1, g5149.t1, g14073.t1, g5071.t1, g3267.t1, g9208.t1, g1990.t1, g8759.t1, g4515.t1, g9401.t1, g7149.t1, g1054.t1, g11794.t1, g10838.t1, g11063.t1, g13643.t1, g2579.t1, g6216.t1, g13004.t1, g18583.t1, g116.t1, g10400.t1, g8850.t1, g12048.t1, g4607.t1, g5302.t1, g9820.t1, g3882.t1, g10181.t1, g12066.t1,	110

Enzymes	Sequence Names	Number of genes
	g8023.t1, g7983.t1, g8895.t1, g147.t1, g9219.t1, g12171.t1, g619.t1, g1055.t1, g7512.t1, g3647.t1, g8702.t1, g16504.t1, g11097.t1, g3065.t1, g6782.t1, g1528.t1, g7830.t1, g14590.t1, g148.t1, g12116.t1, g18155.t1, g2297.t1, g9108.t1, g396.t1, g4071.t1, g1943.t1, g7580.t1, g4339.t1, g15379.t1, g16395.t1, g579.t1, g6105.t1, g13351.t1, g1969.t1, g460.t1, g9218.t1, g981.t1, g15246.t1, g9105.t1, g5749.t1, g8471.t1, g8215.t1, g8006.t1, g19210.t1, g11462.t1, g18992.t1, g10056.t1, g3654.t1, g18763.t1, g4388.t1, g19807.t1, g5841.t1, g5840.t1, g2539.t1, g11376.t1, g1708.t1, g5859.t1, g12901.t1, g185.t1, g6108.t1, g3614.t1, g13199.t1, g19951.t1, g11723.t1, g647.t1	
Enzymes related to Phagosome processes		
2.7.1.137: phosphatidylinositol 3-kinase;	g2465.t1, g17818.t1	2
2.7.1.150: 1-phosphatidylinositol-3-phosphate 5-kinase	g7094.t1	1
3.6.3.14: H+-transporting two-sector ATPase	g1027.t1, g10076.t1, g16058.t1, g16059.t1, g13480.t1, g10622.t1, g2780.t1, g4415.t1, g15533.t1, g15706.t1, g14951.t1, g7935.t1, g5142.t1, g5143.t1, g19157.t1, g7051.t1, g8137.t1, g16352.t1, g17037.t1, g13176.t1, g2225.t1, g2226.t1, g2227.t1, g2228.t1, g3147.t1, g4262.t1, g4504.t1, g10609.t1	28
Enzymes related to	**Lysosome processes**	
2.7.8.17: UDP-N-acetylglucosamine---lysosomal-enzyme N-acetylglucosaminephosphotransferase;	g13117.t1, g6737.t1, g10248.t1	3
3.1.1.13: sterol esterase; cholesterol esterase;	g15814.t1, g3553.t1	2
3.1.1.5: lysophospholipase; lecithinase B; lysolecithinase; phospholipase B; lysophosphatidase; lecitholipase	g4946.t1, g9609.t1, g13411.t1, g7620.t1, g9921.t1, g9922.t1, g13807.t1, g10549.t1, g5180.t1	9
3.1.2.22: palmitoyl[protein] hydrolase; palmitoyl-protein thioesterase; palmitoyl-	g3997.t1, g11079.t1	2

Enzymes	Sequence Names	Number of genes
(protein) hydrolase		
3.1.4.12: sphingomyelin phosphodiesterase; neutral sphingomyelinase	g19104.t1	1
3.2.1.18: exo-alpha-sialidase; neuraminidase; sialidase; alpha-neuraminidase; acetylneuraminidase	g5968.t1	1
3.2.1.20: alpha-glucosidase; maltase; glucoinvertase; glucosidosucrase; maltase-glucoamylase; alpha-glucopyran	g2087.t1, g6007.t1, g6007.t1, g5639.t1, g11886.t1	5
3.2.1.22: alpha-galactosidase; melibiase; alpha-D-galactosidase; alpha-galactosidase A;	g17703.t1, g18976.t1, g7708.t1, g13397.t1, g734.t1, g4888.t1, g13397.t1, g734.t1, g4888.t1, g10638.t1	10
3.2.1.23: beta-galactosidase; lactase (ambiguous); beta-lactosidase; maxilact; hydrolact; beta-D-lactosidase;	g13920.t1, g14384.t1, g19713.t1, g2617.t1, g18481.t1, g5023.t1, g5876.t1, g925.t1, g217.t1, g3387.t1, g18612.t1, g10627.t1, g10940.t1, g4271.t1, g12380.t1	15
3.2.1.24: alpha-mannosidase; alpha-D-mannosidase; p-nitrophenyl-alpha-mannosidase; alpha-D-mannopyranosidase;	g16103.t1, g13332.t1	2
3.2.1.31: beta-glucuronidase; beta-glucuronide glucuronohydrolase glucuronidase;	g242.t1	1
3.2.1.45: glucosylceramidase; psychosine hydrolase; glucosphingosine glucosylhydrolase;	g17141.t1	1
3.2.1.49: alpha-N-acetylgalactosaminidase; alpha-acetylgalactosaminidase;	g11079.t1, g4168.t1, g10638.t1	3
3.2.1.52: beta-N-acetylhexosaminidase; hexosaminidase;	g7925.t1, g8511.t1	2
3.4.14.1: dipeptidyl-peptidase I; cathepsin C; dipeptidyl aminopeptidase I; dipeptidyl transferase; cathepsin	g11345.t1, g16487.t1, g1667.t1, g227.t1, g3896.t1, g4039.t1, g7229.t1, g885.t1, g9958.t1	9
3.4.16.5: carboxypeptidase C; carboxypeptidase Y; serine carboxypeptidase I; cathepsin A;	g10021.t1, g10282.t1	2
3.4.18.1: cathepsin X; cathepsin B2; cysteine-type carboxypeptidase; cathepsin IV; cathepsin Z;	g10940.t1, g11428.t1, g16871.t1	3
3.4.22.-: Cysteine endopeptidase	g4523.t1, g13912.t1, g17440.t1, g12181.t1, g9921.t1, g18485.t1, g16485.t1, g7835.t1, g5682.t1, g15805.t1, g1693.t1, g99.t1, g9451.t1, g19487.t1, g6951.t1, g18039.t1	16
3.4.22.1: cathepsin B; cathepsin B1 (obsolete); cathepsin II cathepsin L;	g12788.t1	1
Aspartic endopeptidases	g7922.t1, g9536.t1, g1812.t1,	42

Enzymes	Sequence Names	Number of genes
	g11075.t1, g16276.t1, g10.t1, g11270.t1, g18475.t1, g2390.t1, g18959.t1, g11581.t1, g9770.t1, g5644.t1, g19251.t1, g1067.t1, g17256.t1, g15070.t1, g2824.t1, g3268.t1, g2624.t1, g17398.t1, g3598.t1, g14587.t1, g14586.t1, g558.t1, g12166.t1, g17079.t1, g8877.t1, g2753.t1, g13728.t1, g3919.t1, g17512.t1, g13078.t1, g6019.t1, g19884.t1, g10152.t1, g11540.t1, g17955.t1, g14137.t1, g4304.t1, g16180.t1, g6812.t1	
3.4.23.5: cathepsin D	g11717.t1, g11718.t1, g1446.t1, g14621.t1, g17420.t1, g18498.t1, g2894.t1, g4451.t1, g4481.t1, g5120.t1, g6224.t1, g7260.t1, g797.t1, g9042.t1	14
3.5.1.23: ceramidase; acylsphingosine deacylase;	g14304.t1	1
3.6.1.6: nucleoside diphosphate phosphatase;	g8837.t1, g9293.t1	2
Enzymes related to peroxisome processes		
1.1.1.35: 3-hydroxyacyl-CoA dehydrogenase; beta-hydroxyacyl dehydrogenase;	g10565.t1, g6421.t1	2
1.1.1.42: isocitrate dehydrogenase (NADP+);	g19124.t1, g629.t1, g7997.t1, g9270.t1	4
1.1.3.15: (S)-2-hydroxy-acid oxidase;	g3321.t1	1
1.11.1.15: peroxiredoxin; thioredoxin peroxidase; tryparedoxin peroxidase	g16286.t1, g4347.t1, g9074.t1, g711.t1, g10793.t1, g4348.t1	6
1.15.1.1: superoxide dismutase; superoxidase dismutase; copper-zinc superoxide dismutase;	g1299.t1, g1327.t1, g18085.t1, g15343.t1, g12437.t1, g14067.t1	6
1.17.1.4: xanthine dehydrogenase	g11271.t1, g8501.t1	2
1.17.3.2: xanthine oxidase;	g11271.t1, g8501.t1	2
1.3.3.6: acyl-CoA oxidase; fatty acyl-CoA oxidase; acyl coenzyme A oxidase; fatty acyl-coenzyme A oxidase	g15580.t1, g8950.t1, g8951.t1, g8952.t1, g1484.t1, g4622.t1, g17278.t1, g2229.t1	8
2.3.1.16: acetyl-CoA C-acyltransferase; beta-ketothiolase; 3-ketoacyl-CoA thiolase;	g20084.t1, g12162.t1, g3411.t1, g9558.t1	4
2.3.1.176: propanoyl-CoA C-acyltransferase	g11613.t1, g6813.t1	2
2.3.1.42: glycerone-phosphate O-acyltransferase	g7214.t1, g7215.t1, g5105.t1, g5318.t1	4
2.5.1.18: glutathione transferase; glutathione S-transferase; glutathione S-alkyltransferase;	g11239.t1, g15471.t1, g16768.t1, g55.t1, g8864.t1, g18240.t1, g2249.t1, g622.t1, g11856.t1, g15155.t1, g15115.t1, g7321.t1	12
2.6.1.44: alanine---glyoxylate transaminase; AGT; alanine-glyoxylate aminotransferase	g11401.t1, g16262.t1, g7370.t1, g3045.t1, g9503.t1, g12516.t1	6

Enzymes	Sequence Names	Number of genes
2.6.1.45: serine---glyoxylate transaminase	g11401.t1, g16262.t1, g7370.t1	3
2.6.1.51: serine---pyruvate transaminase; SPT; hydroxypyruvate:L-alanine transaminase	g11401.t1, g16262.t1, g7370.t1, g3045.t1, g9503.t1, g12516.t1	6
3.1.2.27: choloyl-CoA hydrolase; PTE-2 (ambiguous)	g3273.t1	1
3.3.2.10: soluble epoxide hydrolase;	g19945.t1, g13358.t1, g13357.t1	3
Regulation of autophagy		
2.7.1.137: phosphatidylinositol 3-kinase; 1-phosphatidylinositol 3-kinase;	g2465.t1, g17818.t1	2
2.7.11.1: non-specific serine/threonine protein kinase; A-kinase; AP50 kinase;	g2362.t1, g10597.t1, g10995.t1, g13038.t1, g20486.t1, g4043.t1, g5893.t1, g6167.t1, g322.t1, g4534.t1, g1784.t1, g6500.t1, g5866.t1, g3166.t1, g4270.t1, g4985.t1, g696.t1, g17104.t1, g17187.t1, g371.t1, g372.t1, g373.t1, g1159.t1, g12987.t1, g1274.t1, g18711.t1, g4025.t1, g506.t1, g6417.t1, g3165.t1, g15943.t1, g11099.t1, g11236.t1, g11695.t1, g13189.t1, g13460.t1, g1390.t1, g14361.t1, g1550.t1, g15640.t1, g15947.t1, g16108.t1, g16163.t1, g16537.t1, g16539.t1, g16960.t1, g17182.t1, g17443.t1, g17606.t1, g18614.t1, g209.t1, g2963.t1, g3021.t1, g3612.t1, g4193.t1, g4418.t1, g7241.t1, g7325.t1, g8971.t1, g9469.t1, g5788.t1, g7957.t1, g15051.t1, g2345.t1, g8885.t1, g3132.t1, g345.t1, g13236.t1, g9558.t1, g4052.t1, g3061.t1, g8175.t1, g3729.t1, g16718.t1, g6813.t1, g9188.t1, g508.t1, g6778.t1, g946.t1, g19076.t1, g407.t1, g1820.t1, g13459.t1, g12310.t1, g5604.t1, g3884.t1, g3885.t1, g4231.t1, g977.t1, g13687.t1, g10360.t1, g12557.t1, g12643.t1, g12715.t1, g14336.t1, g16064.t1, g16093.t1, g265.t1, g7329.t1, g1776.t1, g4243.t1, g11652.t1, g14581.t1, g14695.t1, g1598.t1, g16013.t1, g16051.t1, g17025.t1, g17759.t1, g1796.t1, g18025.t1,	129

Enzymes	Sequence Names	Number of genes
	g18242.t1, g2519.t1, g4652.t1, g7049.t1, g7561.t1, g9302.t1, g2757.t1, g4834.t1, g7442.t1, g7486.t1, g14978.t1, g16770.t1, g16870.t1, g5051.t1, g7837.t1, g1454.t1, g1922.t1, g5080.t1	
2.7.11.11: cAMP-dependent protein kinase; PKA; PKA C; protein kinase A;	g11856.t1, g13225.t1, g15155.t1, g16731.t1, g16732.t1, g17196.t1, g17486.t1, g17487.t1, g17488.t1, g18254.t1, g3103.t1, g3686.t1, g539.t1, g6035.t1, g622.t1, g6739.t1, g12081.t1, g12523.t1, g19575.t1, g3411.t1, g3993.t1, g6553.t1, g7242.t1, g3400.t1, g5857.t1, g16455.t1, g3613.t1, g12834.t1, g12808.t1, g2395.t1, g17122.t1, g3862.t1	32
6.3.2.19: ubiquitin---protein ligase; ubiquitin-activating enzyme	g13350.t1, g324.t1, g9376.t1, g12917.t1, g8263.t1, g15902.t1, g5100.t1, g17183.t1, g3340.t1, g3575.t1, g11395.t1, g20628.t1, g4994.t1, g6060.t1, g9284.t1, g149.t1, g5149.t1, g14073.t1, g5071.t1, g3267.t1, g9208.t1, g1990.t1, g8759.t1, g4515.t1, g9401.t1, g7149.t1, g1054.t1, g11794.t1, g10838.t1, g11063.t1, g13643.t1, g2579.t1, g6216.t1, g13004.t1, g18583.t1, g116.t1, g10400.t1, g8850.t1, g12048.t1, g4607.t1, g5302.t1, g9820.t1, g3882.t1, g10181.t1, g12066.t1, g8023.t1, g7983.t1, g8895.t1, g147.t1, g9219.t1, g12171.t1, g619.t1, g1055.t1, g7512.t1, g3647.t1, g8702.t1, g16504.t1, g11097.t1, g3065.t1, g6782.t1, g1528.t1, g7830.t1, g14590.t1, g148.t1, g12116.t1, g18155.t1, g2297.t1, g9108.t1, g396.t1, g4071.t1, g1943.t1, g7580.t1, g4339.t1, g15379.t1, g16395.t1, g579.t1, g6105.t1, g13351.t1, g1969.t1, g460.t1, g9218.t1, g981.t1, g15246.t1, g9105.t1, g5749.t1, g8471.t1, g8215.t1, g8006.t1, g19210.t1, g11462.t1,	110

Enzymes	**Sequence Names**	**Number of genes**
	g18992.t1, g10056.t1, g3654.t1, g18763.t1, g4388.t1, g19807.t1, g5841.t1, g5840.t1, g2539.t1, g11376.t1, g1708.t1, g5859.t1, g12901.t1, g185.t1, g6108.t1, g3614.t1, g13199.t1, g19951.t1, g11723.t1, g647.t1	

Table S6. Mitochondrion respiration assay. Respiration analysis of live dinospores in concentrations from 2-3*10^4 cells/ml. Respiration analysis was carried out in pure filtered medium, without any additional substrates. In addition, we provide cell respiration data from other systems for comparison.

System	**Temperature [°C]**	**MO_2 [nmol O_2*10^4 $cells^{-1}$*min^{-1}]**				**Study**
		Run 1	Run 2	Run 3	Mean ± SEM	
Dinospores	15	0.0140	0.0128	0.0164	0.0144 ± 0.0018	This Study
Gymnodinium nelsoni (clone GSBL)	21				0.21-0.65	P. W. Hochachka, J. M. Teal, The Biological Bulletin 126, 274-281 (1964).
Human endothelial cell culture	37				0.012	E. Gnaiger, R. Steinlechner-Maran, G. Méndez, T. Eberl, R. Margreiter, Journal of bioenergetics and biomembranes 27, 583-596 (1995).
Crassostrea gigas, Isolated Gill cells	15				0.004	G. Lannig, S. Eilers, H. O. Pörtner, I. M. Sokolova, C. Bock, Marine drugs 8, 2318-2339 (2010).
Antarctic fish hepatocytes	12				0.004	F. C. Mark, T. Hirse, H.-O. Pörtner, Polar biology 28, 805-814 (2005).

Table S7. The search for components of the electron transport system of *A. ceratii*. The search was sequential from top to bottom of the table. Sequences identified not to be *A. ceratii* derived were removed from further searches. Mitochondrial sequences from *Plasmodium falciparum*, *Alexandrium tamarense*, *Oxyrrhis marina*, *Tetrahymena thermophila*, *Hematodinium*, and *Perkinsus marinus* were used for Blast searches.

Search criteria	**Method**	**target**	**Found**
Known Dinoflagellate mitochondrial genomes	blastn	genome assembly	Host mitogenome fragments (*Alexandrium*) of *cox*1 and *cyt*b
Mitochondrial genomes of Apicomplexa and Dinoflagellata	blastx	predicted proteins from genome assembly	48 kb mitochondrial genome of a chrysophyte; 2 matches to nuclear genome parts of *cox*1…
Mitochondrial genomes of Apicomplexa and Dinoflagellata	blastx	predicted proteins from transcriptome assembly	*cox*1 fragments
Functional annotation	KEGG/KAAS	genome assembly	identification of oxidative phosphorylation components (see Fig. 5 and S7)
Potential mitochondrial sequences from other *A. ceratii* genome strains	BWA*	raw reads	No matches

*Burrows-Wheeler Aligner

Table S8. Mitochondrial functions encoded in the genome of *A. ceratii*. We searched the whole genome for presence of components of the respiratory chain using BLAST with known components from *Tetrahymena thermophila*. Furthermore, pfam domain structures were detected with InterProScan (http://www.ebi.ac.uk/interpro/interproscan.html), and finally, the whole predicted gene set was submitted to the KAAS annotation server. The table lists all identified components of the respiratory chain in *A. ceratii*. All components of Complex I and III are missing.

Scaffold ID	Position	KEGG ID	Name	Protein ID
Complex components				
Succinate dehydrogenase (complex 2)				
scaffold294_size101967	179..3514	K00234	SDHA	g6898.t1
scaffold87_size276801	53643..55411	K00235	SDHB	g2130.t1
cytochrome c oxidase (complex 4)				
scaffold432_size187470	185280..186478	K02257	COX10	g9344.t1
scaffold1091_size62708	1..700	K02256	COX1	g15932.t1
scaffold46_size328227	41768..42440	K02256	COX1	g833.t1
scaffold161_size93542	33982..34842	K02261	COX2	g4166.t1
scaffold82_size396561	164857..171534	K02265	COX5B	g1954.t1
scaffold919_size47424	31843..33387	K02267	COX6B	g14662.t1
scaffold137_size61657	11329..12202	K02258	COX11	g3584.t1
scaffold516_size65844	39254..41893	K02259	COX15	g10549.t1
ATPase (complex 5)				
scaffold106_size145008	102075..103026	K02137	ATPeF0O	g2780.t1
scaffold52_size130106	21982..22614	K02128	ATPeF0C	g1027.t1
scaffold481_size177949	92371..95061	K02132	ATPeF1A	g10076.t1
scaffold1109_size67627	27142.. 28910	K02133	ATPeF1B	g16059.t1
scaffold792_size49622	46482..48426	K02134	ATPeF1D	g13480.t1
scaffold528_size76787	14694..15944	K02136	ATPeF1G	g10622
Additional identified components as in (*31*)				
Alternative NAD(P)H dehydrogenase				
scaffold17_size333242	300947..303618			g180.t1
Cytochrome b2				
scaffold126_size117765	68257..69291			g3321.t1
Glycerol-3-phosphate dehydrogenase				
scaffold809_size125823	113948..115819			g13621.t1
Dihydroorotate dehydrogenase				
scaffold268_size171261	54808..55797			g6389.t1

Table S9. Primers used in this study to amplify cox from *A. ceratii* cDNA. Primer sequences marked with (*) were obtained from *(33)* and the others were designed during this study.

Name	Primer sequence (5‘-3‘)	Target/position
Dinocox1F*a**	AAAAATTGTAATCATAAACGCTTAGG	cox1 forward
Dinocox1F2*b**	GAATTATATCTCCTGAAAACCAGAACTTC	cox1 forward
Dinocox1F3*b**	TTATGATCTTCTTYTTWRTWATGCC	cox1 forward
Dinocox1R*b**	TGTTGAGCCACCTATAGTAAACATTA	cox1 reverse
Dinocox1R2*b**	AGTTATTCCTGATCCAATAGATGACAG	cox1 reverse
Dinocox1R3*b**	CTGATCCAATAGATGACAGAAAATTCC	cox1 reverse
Dinocob1F*a**	ATGAAATCTCATTTACAWWCATATCCTTGTCC	cob forward
Dinocob2F*b**	ATWAATTYTTTYTGGAATHTBGGTTT	cob forward
Dinocob1R*a**	TCTCTTGAGGKAATTGWKMACCTATCCA	cob reverse
Dinocob2R*b**	CGAGCATAAGATAKAAACWTCTCTTGAGG	cob reverse
	Cox1 gene fragment 1 (g15932)	
G15932_Set1_F	ACGCAGCAGATACGGTCTTT	37
G15932_Set1_R	CAGTGATGAACAAGGCTCCA	471
G15932_Set2_F	TGACTTCGCTTCCAGAGGTT	200
G15932_Set2_R	TGCCGAACACAATTACCTCA	839
G15932_Set3_F	GAGTCCGTCTCCGTGAAGAG	732
G15932_Set3_R	TGGTACTGCACGAGGAACTG	1555
	Cox1 gene fragment 2 (g833)	
G833_set1_F	CGGCGGCAGTATTAGATT	1035
G833_set1_R	TCCTCGTCTTGCTCTTGT	147
G833_set2_F	CAACAGCTGCGACTACTCC	572
G833_set2_R	TTCGAAGCTCCACCAACA	1197

Table S10. Primers used to amplify various shikimate pathway transcripts from *A. ceratii* cDNA.

Name	Primer sequence (5‘-3‘)	Target
	g6770	
aroC_gen1_F	CTTTGACCGAGCAGGATA	chorismate synthase (aroC)
aroC_gen1_R	ATGTTTTCGCCCGCTGTA	chorismate synthase (aroC)
aroB_gen2_F	CAGTACAGCGGGCGAAAA	chorismate synthase (aroB)
aroB_gen2_R	ATCTTCGTCTTGAACACACT	chorismate synthase (aroB)
aroA_gen3_F	ACACTATGCTTTCCCACAA	EPSP synthase (aroA)
aroA_gen3_R	GTGCCCGAAGTTCAAAAT	EPSP synthase (aroA)
aroK_gen4_F	ATTTTGAACTTCGGGCAC	shikimate kinase (aroK)
aroK_gen4_R	ACCAGCCCTCCTTCTCTA	shikimate kinase (aroK)
aroD_gen5_F	TAGAGAAGGAGGGCTGGT	shikimate dehydrogenase (aroD)
aroD_gen5_R	ATTATGTCGAAGATCCCGT	shikimate dehydrogenase (aroD)
aroE_gen6_F	AGATGAACGCCGGAAAAA	shikimate dehydrogenase (aroE)
aroE_gen6_R	AAAACCTGCTCCTTCAAC	shikimate dehydrogenase (aroE)
	g13589	
aroG_gen7_F	CGCAAAGAGTGGAAAGAG	chorismate binding DAHP (aroG)
aroG_gen7_R	AACCGAATCCTTCCACAC	chorismate binding DAHP (aroG)