Comparison of gene expression of *Paramecium bursaria* with and without *Chlorella variabilis* symbionts

Yuuki Kodama¹, Haruo Suzuki²†, Hideo Dohra³, Manabu Sugii⁴, Tatsuya Kitazume⁵, Katsushi Yamaguchi⁵, Shuji Shigenobu⁵,⁶ and Masahiro Fujishima²*

**Abstract**

**Background:** The ciliate *Paramecium bursaria* harbors several hundred cells of the green-alga *Chlorella* sp. in their cytoplasm. Irrespective of the mutual relation between *P. bursaria* and the symbiotic algae, both cells retain the ability to grow without the partner. They can easily reestablish endosymbiosis when put in contact with each other. Consequently, *P. bursaria* is an excellent model for studying cell–cell interaction and the evolution of eukaryotic cells through secondary endosymbiosis between different protists. Despite the importance of this organism, no genomic resources have been identified for *P. bursaria* to date. This investigation compared gene expressions through RNA-Seq analysis and de novo transcriptome assembly of symbiont-free and symbiont-bearing host cells.

**Results:** To expedite the process of gene discovery related to the endosymbiosis, we have undertaken Illumina deep sequencing of mRNAs prepared from symbiont-bearing and symbiont-free *P. bursaria* cells. We assembled the reads de novo to build the transcriptome. Sequencing using Illumina HiSeq2000 platform yielded 232.3 million paired-end sequence reads. Clean reads filtered from the raw reads were assembled into 68,175 contig sequences. Of these, 10,557 representative sequences were retained after removing *Chlorella* sequences and lowly expressed sequences. Nearly 90% of these transcript sequences were annotated by similarity search against protein databases. We identified differentially expressed genes in the symbiont-bearing *P. bursaria* cells relative to the symbiont-free cells, including heat shock 70 kDa protein and glutathione S-transferase.

**Conclusions:** This is the first reported comprehensive sequence resource of *Paramecium*–*Chlorella* endosymbiosis. Results provide some keys for the elucidation of secondary endosymbiosis in *P. bursaria*. We identified *P. bursaria* genes that are differentially expressed in symbiont-bearing and symbiont-free conditions.

**Keywords:** *Paramecium bursaria*, *Chlorella variabilis*, Secondary symbiosis, Transcriptome analysis

**Background**

As demonstrated by the evolution of mitochondria and chloroplasts, endosymbiosis is a major driving force behind eukaryotic cell evolution leading to acquisition of new intracellular components and cell diversity [1,2]. Although endosymbiosis is an important and widespread phenomenon, the mechanisms controlling the establishment of endosymbiosis between different eukaryotic cells are not well understood. In fact, *P. bursaria* cells harbor about 700 symbiotic algae in their cytoplasm [3]. Each alga is enclosed in a perialgal vacuole (PV) membrane derived from the host digestive vacuole (DV) membrane, which protects the alga from the host’s lysosomal fusion [4-6]. Irrespective of the mutual relations between *P. bursaria* and symbiotic algae [7-11], the symbiont-free cells and the symbiotic algae retain the ability to grow without a partner. Symbiont-free cells can be prepared by various means: cultivation under constant dark conditions [12-14], treatment with cycloheximide [3,15,16], and treatment with the photosynthesis inhibitor dichlorophenyl dimethylurea (DCMU) [17]. However, symbiotic algae can be
isolated by homogenization or by sonication or by the treatment of symbiotic cells with detergent. They can grow outside host cells [18]. Symbiont-free cells are easily reinfected with symbiotic algae by mixing the two together. Therefore, *P. bursaria* has been considered an excellent model for studying cell–cell interaction and the evolution of eukaryotic cells through secondary endosymbiosis between different protists [19]. However, neither genomic nor transcriptomic information has been available to elucidate the establishment of endosymbiosis in *P. bursaria* to date. To expedite the process of gene discovery related to the endosymbiosis, we have undertaken Illumina deep sequencing of mRNAs prepared from symbiont-bearing and symbiont-free *P. bursaria* cells in this study. Our data provide a comprehensive sequence resource for the advancement of *P. bursaria* study.

Results and discussion

Deep-sequencing and assembly

We constructed three RNA-seq libraries from mRNA of *P. bursaria* harboring symbiotic alga, *Chlorella variabilis*, and three libraries from symbiont-free *P. bursaria*. Sequencing using Illumina HiSeq2000 platform yielded 232.3 million 101-by-101 bp paired-end sequence reads. After trimming the low-quality parts and removing reads of less than 50 bp, 436.9 million reads (42.9 Gb) remained. To obtain a comprehensive sequence set of the *P. bursaria* transcriptome, all the clean reads of symbiont-bearing and symbiont-free libraries were assembled together using the Trinity program [20]. The *de novo* assembly produced 68,175 contigs, clustering into 40,805 subcomponents (i.e. unigenes). We selected the longest transcript as the representative for each cluster. The unigene sizes were 200 bp up to 22,858 bp, with mean length of 904 bp, N50 of 1,503 (Additional file 1). The corresponding values for lengths of ORFs for *P. bursaria* were, respectively, 1,161, 1,155, and 1,503 (Additional file 2).

We excluded unigenes derived from the symbiotic *Chlorella* and other contaminants. Of the 68,175 contig sequences, 11,256 were matched to the *C. variabilis* sequences, and were therefore removed. Unigenes lowly expressed with log-counts-per-million (logCPM) < 0 were also discarded because they are likely to be contaminant sequences or poor assembly models. Based on the database search, the small amount of the contaminant sequences appears to be derived from some bacteria such as *Methylobacterium* and *Burkholderiales*, which are likely to be included in the culture media in which we grew *P. bursaria*. These procedures produced *P. bursaria* transcript reference sequences composed of 10,557 unigenes.

Annotation of the assembled contigs

We performed similarity searches of the 10,557 *P. bursaria* unigenes against the Swiss-Prot and UniRef90 protein sequence databases [21] using BLASTX [22] with the E-value cutoff of 1e-5 and assigned the functional annotations of the most similar protein sequences. Of the 10,557 unigenes, 7,051 (67%) had matches with 4,102 unique records in the Swiss-Prot database; 9,536 (90%) had matches with 8,189 unique records in the UniRef90 database. The species distribution of the BLASTX best hits in the UniRef90 database showed that 8,710 (91.7%) of the 9,502 hits had top matches with sequences from *P. tetraurelia*, followed by *Tetrahymena thermophila* with 153 (1.6%) best BLASTX hits.

We predicted open reading frames (ORFs) from the 10,557 *P. bursaria* unigene sequences using OrfPredictor [23]. Of the 10,557 ORFs, 10,535 were longer than 50 amino acids, 10,134 were longer than 100 amino acids, and 3,425 were longer than 500 amino acids. Although whole genome sequences have been clarified in *P. tetraurelia* [24] and *T. thermophila* [25], endosymbiotic algae including *Chlorella* species have not yet been detected in these ciliates. Therefore, we tried to compare their ORFs length, GC%, and shared gene clusters among these two ciliates and *P. bursaria* to elucidate the genomic features of *P. bursaria* as a potential host cell for the symbiotic alga.

We compared ORFs (>100 amino acids) of *P. bursaria* with those of its close relatives *P. tetraurelia* and *T. thermophila*. The maximum values for lengths of ORFs for *P. bursaria*, *P. tetraurelia*, and *T. thermophila* were, respectively, 19,640, 21,570, and 34,740. The median values for lengths of ORFs for *P. bursaria*, *P. tetraurelia*, and *T. thermophila* were, respectively, 1,161, 1,155, and 1,503 (Additional file 1). The corresponding values for G+C contents (%) of ORFs were 30.1, 29.7, and 27.8 (Additional file 2).

We identified 29,835 orthologous gene clusters containing 70,114 ORFs from the three organisms: i.e. 10,134, 37,382, and 22,598 ORFs longer than 100 amino acids long, respectively, from *P. bursaria*, *P. tetraurelia*, and *T. thermophila*. Figure 1 shows the number of orthologous gene clusters shared among the three organisms. Of the 29,835 orthologous gene clusters, 3,421 were common to all three organisms, 2,854 were unique to *Paramecium* species, and 2,330 were unique to *P. bursaria*.

Differential gene expression between symbiont-bearing and symbiont-free conditions

We compared gene expressions of symbiont-bearing and symbiont-free *P. bursaria* to elucidate the genetic control for establishment of secondary symbiosis. Of the 10,557 transcripts, 6,698 were significantly differentially expressed between symbiont-bearing and symbiont-free cells with false discovery rates (FDR) < 0.05 (Additional file 3). The positive and negative values of log2 Fold Change (logFC) show that the sequences were up-regulated and down-regulated in symbiont-bearing cells compared to
symbiont-free cells. The parametric analysis of gene set enrichment (PAGE) [27] detected enrichment in the 17 gene ontology (GO) terms, i.e. 8 biological processes, 3 cellular components, and 6 molecular functions based on the logFC between symbiont-bearing and symbiont-free cells with false discovery rates (FDR) < 0.05 (Table 1). Figure 2 portrays a parent–child relation between GO molecular function terms including the six highest ranking GO terms based on PAGE P-values: oxidoreductase activity (Z = −6.879), structural constituent of ribosome (Z = −4.153), pyridoxal phosphate binding (Z = −4.015), phosphorelay response regulator activity (Z = 3.310), chromatin binding (Z = 3.107), and phosphorelay sensor kinase activity (Z = 2.901). A closer examination of individual proteins assigned to these GO terms revealed that algal proteins with functions equivalent to those of the host Paramecium cells are transferred to the host through the PV membrane. Consequently transcriptional activity of the host was reduced. However, no report to date has described a demonstration showing that the proteins synthesized by the algae and transferred to the host cell through photosynthetic products, mainly maltose, are transferred from the algae [7,9,11].

Up-regulation of signal transduction histidine kinase in symbiont-bearing P. bursaria cells suggests that the histidine kinases play an important role in signal perception in this secondary symbiosis, as shown in the primary symbiosis by bacteria [28] for the adaptation and survival of various organisms to harsh environmental conditions [29]. Sensor histidine kinases are known to play important roles in several eukaryotic species including yeasts, fungi, slime molds, and higher plants [30-32]. Symbiont-bearing P. bursaria cells acquire various stress resistance through endosymbiosis with Chlorella spp [33-36].

In addition to the genes of enriched GO terms discussed above, heat shock 70 kDa protein (Hsp70) and glutathione S-transferase (GST) genes were up-regulated and down-regulated as shown by the positive and negative values of logFC, respectively, in symbiont-bearing cells compared to symbiont-free cells (Table 2). Of the 10,557 unigenes, 8 were annotated as Hsp70 with logFC of −1.3 to 5.6, with a median of 0.92. Symbiont-bearing P. bursaria cells are known to show a higher survival ratio against nickel chloride, high temperature, and hydrogen peroxide than the symbiont-free cells show [35,36]. Furthermore, P. caudatum cells reportedly acquire heat-shock resistance by infection of endonuclear symbiotic bacteria Holospora [37-39], and osmotic-shock resistance [40]. Hori and Fujishima [39] reported that H. obtusa-bearing paramecia expressed high levels of Hsp70 mRNA even at 25°C. The up-regulation of the transcripts encoding Hsp70 might be related to the host’s tolerance to environmental fluctuations. Of the 10,557 transcripts, 7 were annotated as GST and tended to be down-regulated with logFC of −5.7 to −0.12, with a median of −0.85 (Table 2). This enzyme is related to protection of cells from oxidative stress, as shown by McCord and Fridovich [41] and by Veal et al. [42]. Although it was conceivable that photo-oxidative stress in symbiont-bearing P. bursaria cells is greater than that in symbiont-free ones, our data showed opposite results from the prediction. A similar result was obtained by Hörttnagl and Sommaruga [33]. They suggested that the presence of algal symbionts minimizes photo-oxidative stress [33]. Consequently, different expression levels in these genes between symbiont-free and symbiont-bearing P. bursaria agree well with differences in cytological phenomena observed in these paramecia, suggesting that these proteins appear to be involved in the symbiosis. Immunological detections of the gene
products and comparisons of the amount of the antigens or qualitative PCR between the symbiont-free and the symbiont-bearing paramecia are necessary for future experiments. To ascertain the molecular mechanisms controlling the secondary symbiosis between *Paramecium* and *Chlorella* cells, transcriptome analyses and proteome analyses of the symbiotic *Chlorella* alone cultivated in algal culture medium and the same *Chlorella* inside the host *Paramecium* are necessary for future studies.

### Conclusion

This study is the first whole transcriptome analysis conducted between symbiont-free and symbiont-bearing *P. bursaria*. Results showed *P. bursaria* genes that differentially expressed in symbiotic and symbiont-free conditions. We showed that genes for glutathione S-transferase, trans-2-enoyl-CoA, aminotransferases, and ribosomal proteins are down-regulated, and that genes for Hsp70, transcriptional activator Myb-related proteins, and signal transduction histidine kinase are up-regulated in the symbiont-bearing *P. bursaria*. Our results enable us to understand the molecular mechanism for establishment of the secondary symbiosis and for the host evolutionary adaptation to global climate change.

### Methods

#### Strains and cultures

Symbiont-free *P. bursaria* strain Yad1w was produced from *Chlorella* sp.-bearing *P. bursaria* strain Yad1g cells (syngen 3, mating type I) through repeated cloning and cultivation under dark conditions. The Yad1g cell strain was collected in Yamaguchi, Japan in 2004. Symbiont-bearing strain of Yad1g1N cells was used for symbiont-bearing cells. The Yad1g1N strain was produced by infection of cloned symbiotic *Chlorella variabilis* (formerly *C. vulgaris*) strain 1 N cells to the Yad1w cell [15]. *Paramecium* strains used for this study were provided by Symbiosis Laboratory, Yamaguchi University with support in part by the National Bio-Resource Project of the Ministry of Education, Culture, Sports, Science and Technology, Japan. The culture medium used was 1.25% (w/v) fresh lettuce juice in modified Dryls solution (MDS) [43] (KH2PO4 was used instead of NaH2PO4 · 2H2O), inoculated with a non-pathogenic strain of *Klebsiella pneumoniae* one day before use [44]. In ordinary cultures, several hundred cells were inoculated into 2 ml culture medium. Then 2 ml of fresh culture medium was added on each of the next 12 days. One day after the final feeding, the cultures were in the early stationary phase of growth. All cells used in the present experiments were at this phase. Cultivation and all experiments were performed at 25 ± 1°C. In the case of the symbiont-bearing strain of Yad1g1N, the cells were cultivated under constant light conditions: 20–30 μmol photon/m²/s.

#### Transcriptome sequencing

Total RNA was isolated from 400,000 cells of symbiont-bearing and symbiont-free *P. bursaria* cells using Trizol...
reagent (Invitrogen Corp.) according to the manufacturer's protocol. To construct three RNA-seq libraries from mRNAs of *P. bursaria*, the total RNA was isolated from three independent cultures of symbiont-free and symbiont-bearing *P. bursaria*. After suspension in Trizol reagent, the symbiont-bearing and symbiont-free cells were stirred in the presence of 600 μl of 0.5 mm zirconia/silica beads (BioSpec Products Inc.) to break cell walls of the symbiotic algae. The RNA was treated with RNase-free DNase and was cleaned up using RNeasy Mini (Qiagen Inc.) according to the manufacturer's protocol. RNA integrity was confirmed (2100 Bioanalyzer; Agilent Technologies Inc.) with a minimum RNA integrated number value of 7.6. The samples for transcriptome analyses were prepared using Illumina's kit following the manufacturer's recommendations. First, mRNA was purified from 0.5 μg of total RNA of symbiont-bearing or symbiont-free cells using oligo (dT) magnetic beads. Following purification, the mRNA was fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments were used for first strand cDNA synthesis using SuperScript II Reverse Transcriptase (Invitrogen Corp.) and random primers. Second strand cDNA synthesis was conducted next. These cDNA fragments then went through an end repair process and ligation of adapters. These products were purified and enriched with PCR to create the final cDNA library. Multiplex sequencing of paired-end reads was performed on an Illumina Hiseq2000 instrument at NIBB Core Research Facilities, followed by raw data processing, base-calling, and quality-control by manufacturer's standard pipeline using RTA, OLB, and CASAVA. The output sequence quality was inspected using the FastQC program (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/).
De novo assembly
The reads were cleaned up with cutadapt program by trimming low-quality ends (<QV30) and adapter sequences, and by discarding reads shorter than 50 bp. De novo assembly of the clean reads was conducted using Trinity [20] in the paired-end mode with an option ‘–min_kmer_cov = 2’.

Differential expression analysis
Data of two biological replicates were used in this analysis for each condition. Using scripts included the Trinity package suite [20], cleaned reads were aligned to the Trinity-assembled transcripts using Bowtie [45]. Then transcript abundance was estimated using RSEM [46]. We used the edgeR package [47] of Bioconductor to identify genes that are differentially expressed between conditions. To adjust for library sizes and skewed expression of transcripts, the estimated abundance values were normalized using the Trimmed Mean of M-values (TMM) normalization method [48] included in the edgeR package. Based on a negative binomial model implemented in edgeR, genes that were differentially expressed between symbiont-bearing and symbiont-free P. bursaria samples were identified.

Functional annotation of assembled contigs
Nucleotide sequences of Chlorella variabilis NC64A were obtained from the DOE Joint Genome Institute (JGI) site (http://genome.jgi.doe.gov/ChlNC64A_1/ChlNC64A_1.home.html). The assembled contigs that matched the Chlorella sequences indicated by MEGABLAST search (E-value cutoff of 1e-40) were excluded from analyses. We performed a BLASTX search (with the ciliate nuclear genetic code and E-value cutoff of 1e-5) of the P. bursaria transcripts against the UniProtKB Swiss-Prot and Uniref90 protein sequence databases [21] and assigned the functional annotations of the most similar protein sequences.

The protein-coding region of RNA sequences was predicted using OrfPredictor with the ciliate nuclear genetic code [23]. For sequences having BLASTX hits, the frames used by BLASTX were used for identifying the coding regions of the sequences. For sequences without BLASTX hits, the coding regions were estimated using RSEM [46]. We used the pfam2go conversion table (http://www.geneontology.org/external2go/pfam2go) [50]. We performed a BLASTX search (with the ciliate nuclear genetic code [23]) of the protein sequences against the Pfam database release 26.0 [49]. Gene ontology (GO) terms were assigned to each transcript using the pfam2go conversion table (http://www.geneontology.org/external2go/pfam2go) [50]. We performed parametric analysis of gene set enrichment or PAGE [27] to test over-

| Table 2 Transcripts encoding glutathione S-transferase and heat shock 70 kDa protein in P. bursaria |
|-----------------------------------------------|
| **Trinity transcript name** | **Annotation from the SwissProt database** | **logFC** |
| comp34044_c0 | sp[P09446]HSP7A_CAEL| Heat shock 70 kDa protein A OS = Caenorhabditis elegans | 5.601 |
| comp36402_c4 | sp[P09446]HSP7A_CAEL| Heat shock 70 kDa protein A OS = Caenorhabditis elegans | 4.183 |
| comp36402_c6 | sp[P14834]HSP70_LEIMA| Heat shock 70 kDa protein (Fragment) OS = Leishmania major | 1.975 |
| comp36402_c1 | sp[Q95SN1]HSP7E_ARATH| Heat shock 70 kDa protein 5 OS = Arabidopsis thaliana | 1.555 |
| comp37280_c1 | sp[P37899]HSP70_PYRSA| Heat shock 70 kDa protein OS = Pyrenomonas salina | 0.287 |
| comp3771_c0 | sp[P09446]HSP7A_CAEL| Heat shock 70 kDa protein A OS = Caenorhabditis elegans | –0.594 |
| comp1901_c0 | sp[Q957C0]HSP70_ARATH| Heat shock 70 kDa protein 14 OS = Arabidopsis thaliana | –1.076 |
| comp1912_c0 | sp[F4JM11]HSP7R_ARATH| Heat shock 70 kDa protein 17 OS = Arabidopsis thaliana | –1.337 |
| **Glutathione S-transferase** | | |
| comp37410_c0 | sp[P78417]GSTO1_HUMAN| Glutathione S-transferase omega-1 OS = Homo sapiens | –0.119 |
| comp32377_c0 | sp[Q9ZRT5]GSTT1_ARATH| Glutathione S-transferase T1 OS = Arabidopsis thaliana | –0.288 |
| comp36943_c0 | sp[Q9ZVQ3]GSTZ1_ARATH| Glutathione S-transferase Z1 OS = Arabidopsis thaliana | –0.748 |
| comp37841_c0 | sp[Q9ZRT5]GSTT1_ARATH| Glutathione S-transferase T1 OS = Arabidopsis thaliana | –0.851 |
| comp36483_c0 | sp[Q9ZRT5]GSTT1_ARATH| Glutathione S-transferase T1 OS = Arabidopsis thaliana | –1.557 |
| comp35816_c1 | sp[P78417]GSTO1_HUMAN| Glutathione S-transferase omega-1 OS = Homo sapiens | –1.564 |
| comp36242_c0 | sp[P16413]GSTMU_CAVPO| Glutathione S-transferase B OS = Cavia porcellus | –5.749 |
representation and under-representation of GO terms based on the logFC between symbiont-bearing and symbiont-free cells.

Availability of supporting data
The data sets supporting the results of this article are available in the DDBJ Sequence Read Archive (DRA) (accession number DRA000907, http://trace.ddbj.nig.ac.jp/DRASearch/submission?acc=DRA000907).

Additional files

Additional file 1: Histogram showing the distribution of lengths for protein-coding sequences in *P. bursaria* (A), *P. tetraurelia* (B), and *T. thermophila* (C). The x-axis shows the length (bp) of the sequences. The y-axis shows their frequencies.

Additional file 2: Histogram showing the distribution of G + C contents for protein-coding sequences in *P. bursaria* (A), *P. tetraurelia* (B), and *T. thermophila* (C). The x-axis shows the G + C content (%) of the sequences. The y-axis shows their frequencies. G + C content is defined as 100 x (G + C)/(A + T + G + C).

Additional file 3: Data for *Paramecium bursaria* transcripts. The first five columns show the Trinity sequence name, length (bp), and functional annotations from different databases: Swiss-Prot, UniRef100, and gene ontology (GO). The remaining four columns show data of differential expression analysis using edgeR log2 Fold Changes (logFC), log-counts-per-million (logCPM), P-values, and False Discovery Rate (FDR) adjusted P-values.

Competing interests
The authors declare that there are no competing interests.

Authors’ contributions
YK and MF cultured paramecia and isolated RNA from paramecia. YK and MF cultured paramecia and isolated RNA from paramecia. TK, YK, and SS conducted the bioinformatics analysis using edgeR: log2 Fold Changes (logFC), log-counts-per-million (logCPM), P-values, and False Discovery Rate (FDR) adjusted P-values.

Acknowledgements
We thank Dr. Motohide Aoki, School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, for his kind technical assistance in isolation of Chlorella RNAs. This work was supported by a Grant-in-Aid for Research Activity Start-up (No. 22870023) and by a Grant-in-Aid for Young Scientists (B) from the Japan Society for the Promotion of Science (JSPS), by a grant from the Inoue Foundation for Science (Inoue Research Award for Young Activity Start-up (No. 22870023) and by a Grant-in-Aid for Young Scientists (B) from the Japan Society for the Promotion of Science (JSPS), by a grant from the Inoue Foundation for Science (Inoue Research Award for Young Scientists), by a Narishige Zoological Science Award, and by a Grant for Basic Science Research Projects from The Sumitomo Foundation to Y. Kodama. This work was also supported by a Grant-in-Aid for Scientific Research (B) (No. 2337008) and a Grant-in-Aid for Challenging Exploratory Research (No. 23657157) from JSPS to M. Fujishima.

Author details
1Department of Biological Science, Faculty of Life and Environmental Science, Shizuoka University, 1166 Higashi-Shizuoka, Shizuoka 422-8529, Japan. 2Department of Environmental Science and Engineering, Graduate School of Science and Engineering, Yamaguchi University, 1677-1 Yoshida, Yamaguchi 753-8512, Japan. 3Instrumental Research Support Office, Research Institute of Green Science and Technology, Shizuoka University, 836 Oya, Suruga-ku, Shizuoka 422-8529, Japan. 4Media and Information Technology Center, Yamaguchi University, 2-16-1 Tokiwadai, Ube 755-8611, Japan. 5National Institute for Basic Biology, Nishigonaka 38, Myodaiji, Okazaki 444-8858, Japan. 6Department of Basic Biology, School of Life Science, Graduate University for Advanced Studies, Okazaki 444-8858, Japan.

Received: 7 December 2013 Accepted: 3 March 2014
Published: 10 March 2014

References
1. Fujishima M: Infection and maintenance of *Holospora* species in *Paramecium caudatum*. In Endosymbionts in Paramecium vol. 12. Edited by Fujishima M, Heidelberg: Springer Berlin, 2009:201–225.
2. Nowack ECM, Melkonian M: Endosymbiotic associations within protists. Philos Trans R Soc Lond B Biol Sci 2010, 365(1516):999–1012.
3. Kodama Y, Fujishima M: Cyclorchimide induces synchronous swelling of perialgal vacuoles enclosing symbiotic *Chlorrella vulgaris* and digestion of the algae in the ciliate *Paramecium bursaria*. Protist 2008, 159(3):483–494.
4. Gu F, Chen L, Ni B, Zhang X: A comparative study on the electron microscopic enzyme-cytchemistry of *Paramecium bursaria* from light and dark cultures. Eur J Protistol 2002, 38(2):267–278.
5. Karakashian SJ, Rudzinska MA: Inhibition of lysosomal fusion with symbiont-containing vacuoles in *Paramecium bursaria*. Exp Cell Res 1981, 131(2):387–393.
6. Kodama Y, Fujishima M: Localization of perialgal vacuoles beneath the host cell surface is not a prerequisite phenomenon for protection from the host’s lysosomal fusion in the ciliate *Paramecium bursaria*. Protist 2006, 160(2):319–329.
7. Brown JA, Nielsen PI: Transfer of photosynthetically produced carbohydrate from endosymbiotic *Chlorella* to *Paramecium bursaria*. J Protozool 1974, 21(4):569–570.
8. Reisser W: The metabolic interchanges between *Paramecium bursaria* Ehrbg. and *Chlorella* spec. in the *Paramecium bursaria*-symbiosis. I. Nitrogen and the carbon metabolism (author’s transl). Arch Microbiol 1976, 97(5):357–360.
9. Reisser W: Endosymbiotic Associations of Freshwater Protozoa and Algae, Volume 1: Bristol: Biopress Ltd; 1986.
10. Albers D, Wiessner W: Nitrogen nutrition of endosymbiotic *Chlorella* spec. Endocytobiol Cell Res 1985, 25:65–64.
11. Reisser W: The metabolic interchanges between *Paramecium bursaria* Ehrbg. and *Chlorella* spec. in the *Paramecium bursaria*-symbiosis. II. Symbiosis-specific properties of the physiology and the cytology of the symbiotic unit and their regulation (author’s transl). Arch Microbiol 1976, 111(1–2):161–170.
12. Karakashian SJ: Growth of *Paramecium bursaria* as influenced by the presence of algal symbionts. Physiol Zool 1963, 36:52–68.
13. Padó R: Mutual relation of protozoans and symbiotic algae in *Paramecium bursaria*. I. The influence of light on the growth of symbionts. Folia Biol 1965, 13:217–182.
14. Weis DS: Regulation of host and symbiont population size in *Paramecium bursaria*. Experientia 1966, 25(6):664–666.
15. Kodama Y, Nakahara M, Fujishima M: Symbiotic alga *Chlorella vulgaris* of the ciliate *Paramecium bursaria* shows temporary resistance to host lysosomal enzymes during the early infection process. Protistoplasma 2007, 201:261–267.
16. Weis DS: The effect of accumulation time of separate cultivation on the frequency of infection of aposymbiotic ciliates by symbiotic algae in *Paramecium bursaria*. Protozool 1984, 31:14A.
17. Reisser W: Die stoffwechselsphysiologischen Beziehungen zwischen *Paramecium bursaria* Ehrbg. und *Chlorella* spec. in der *Paramecium bursaria*-Symbiose. I. Der Stickstoff- und der Kohlenstoff-Stoffwechsel. Arch Microbiol 1976, 107(3):357–360.
18. Siegel R, Karakashian SJ: Dissociation and restoration of endocellular symbiosis in *Paramecium bursaria*. Anat Rec 1959, 134:639.
19. Kodama Y, Fujishima M: Secondary symbiosis between *Paramecium* and *Chlorera* cells. In International Review of Cell and Molecular Biology vol. 279. Edited by Leon KW: Elsevier Inc, 2010:33–77.
20. Grabher MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mausel E, Hacohen N, Gnirke A, Rhind N, di Palma F, Biren BW, Nussbaum C, Lindblad-Toh K, Friedman N, Regev A: Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 2011, 29(4):444–452.
21. UniProt Consortium: Update on activities at the universal protein resource (UniProt) in 2013. Nucleic Acids Res 2013, 41(Database issue):D43–D47.
22. Autschl SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997, 25(17):3389–3402.
23. Min XJ, Butler G, Storms R, Tsang A: OrthPred: predicting protein-coding regions in EST-derived sequences. Nucleic Acids Res 2005, 33:677–680. Web Server issue.
24. Aury JM, Jallion O, Duret L, Noel B, Jubin C, Porcel BM, Segurens B, Daubin V, Anthouard V, Alach N, Arañoa Q, Billaut A, Beisson J, Blanc I, Bouhoucha K, Cámara F, Duharcourt S, Guigo R, Gogendeau D, Katinka M, Keller AM, Kissmehl R, Klotz C, Koll F, Le Mouell A, Lepère G, Malinsky S, Nowacki M, Nowak JK, Plattner H. Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. *Nature* 2008, 444(7116):171–178.

25. Eisen JA, Coyne RS, Jones RB, Badger JH, Wortman JR, Aury JM, DeSalle R, Echave J, Eisenhaber F, Emsley K, Feng Z, Feltus FA, Flanigan KM, Frings I, Friedrich M, Fujita S, Gish WR, Gray NS, Greenblatt JF, Grogan P, Gunsalus IC, Haas BJ, Haft DH, Han C, Han B, Hartman M, Hebl C, Hedrick L, Hernandez-Zavaleta M, Hu Z, Hulcr J, Illum L, Iyer LV, Jiang H, Jones S, Joo MY, Karp PD, Keeling PJ, Kerlavage AR, Klein DF, Kyrpides NC, Langmead B, Lipman DJ, Lutkin S, Ma J, Ma B, Marks D, Mastral A, May AK, McVey TR, Meigs JB, Mevius D, Mekhedov S, Methe BA, Michel DT, Miller W, Miller W, Myburg AA, Myronov A, Nelson D, Nusbaum C, Oakley G, Olde P, Orsavo I, Papamichail M, Petrosino JF, Pevzner PA, Phillips RM, Pollock JS, Poskin K, Proctor BD, Quackenbush J, Reiners AS, Riesch RS, Roberts JA, Rokhsar DS, Roth FR, Ruvkun G, Schadt E, Schmutz JD, Schug J, Schuldiner M, Shen ASY, Shokat KM, Sirok RT, Spellman PT, Stoeckert CJ Jr, Storz G, Talwell LA, Tatusova TA, Tett A, Theisen H, Tettelin H, Toth PA, Trask BJ, Treangen TJ, Turcotte JA, Turcotte J, Venter CS, Venter JC, Vinit P, Volek M, Wang Q, Wang Z, Ward D, Waterhouse AM, Weinstock GM, Weinstock GM, Weideman LA, White O, Wilcox D, Williams JF, Wilson RK, Wolkenstein P, Wortman JR, Yang G, Yao X, Young L, Zhou H, Zhu M, Zichichi A. Alignment of short DNA sequences to the human genome. *Genome Res.* 2000, 10(3):R25–R36.

26. Li L, Stoeckert CJ Jr, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 2003, 13(9):2178–2189.

27. Kim SY, Volsky DJ. PAGE: parametric analysis of gene set enrichment. *BMC Bioinformatics* 2005, 6:144.

28. Parkinson JS, Kafodi EC. Communication modules in bacterial signaling proteins. *Annu Rev Genet.* 1992, 26:11–112.

29. Itoh H, Shoda A, Okumura H, Imai T, Oishi K, Akiyama A, Yamagishi K, Yonezawa K, Shimada Y, Nishimura T. A novel family of histidine kinase genes in *A. thaliana*. *Plant Cell Physiol.* 2001, 42(2):231–235.

30. Chang C, Kwok SF, Bleecker AB, Meyerowitz EM. Communication modules in bacterial signaling revealed by the ciliate Paramecium tetraurelia. *Nature* 2008, 444(7116):171–178.

31. Eisentraut C, Kofoid EC. The two-component system. Regulation of diverse signal transduction pathways in prokaryotes and eukaryotes. *Plant Physiol.* 1996, 112(3):273–279.

32. Wurgler-Murphy SM, Saito H. Two-component signal transducers and MAPK cascades. *Trends Biochem Sci.* 1997, 22(5):172–176.

33. Hökfelt T, Feletti M, Tatemoto K, Pasternak GW, Yaksh TL. Central and peripheral actions of brain peptides and enkephalins. *Nature* 1982, 298(5871):56–60.

34. Horii H, Fuji K, Fujishima M. Micronucleus-specific bacterium *Holospora obtusa* elegans irreversibly enhances stress gene expression of the host *Paramecium caudatum*. *J. Eukaryot. Microbiol.* 2008, 55(6):515–521.

35. Hori M, Fuji K, Fujishima M. The endosymbiotic bacterium *Holospora obtusa* enhances heat-shock gene expression of the host *Paramecium caudatum*. *J. Eukaryot. Microbiol.* 2003, 50(4):293–298.

36. Smurov AO, Fokin SI. Resistance of *Paramecium caudatum* infected with endosymbiotic bacteria *Holospora* against salinity impact. *Proc. Soil Biol. Inst. RAN* 1998, 276:175–178.

37. McCord JM, Fridovich I. Peroxidase dismutase. An enzymatic function for erythrocytopenic (hemoglobin). *J. Biol. Chem.* 1969, 244(22):6049–6055.

38. Veal EA, Toone WM, Jones N, Morgan BA. Distinct roles for glutathione S-transferases in the oxidative stress response in *Schizosaccharomyces pombe*. *J. Biol. Chem.* 2002, 277(35):35523–35531.

39. Dyli S. Antigenic transformation in *Paramecium aurelia* after homologous antisemir treatment during autogamy and conjugation. *J. Protozool.* 1959, 6:Suppl2:25.

40. Ishii M, Nagahara K, Kojima Y. Changes in morphology, buoyancy, density and protein composition in differentiation from the reproductive short form to the infectious long form of *Holospora obtusa*, a micronucleus-specific symbiont of the ciliate *Paramecium caudatum*. *Zool. Sci.* 1990, 7:839–860.

41. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009, 10(3):R25.

42. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011, 12:323.

43. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26(1):139–140.

44. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010, 11(3):R25.

45. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Sardy SR, Bateman A, Finn RD. The Pfam protein families database. *Nucleic Acids Res.* 2012, 40(Database issue):D290–D296.

46. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010, 11(3):R25.

47. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet.* 2000, 25(1):25–29.

Cite this article as: Kodama et al. Comparison of gene expression of *Paramecium bursaria* with and without Chlorella variabilis symbionts. *BMC Genomics* 2014:15–183.