学位論文

Molecular identification of rickettsial endosymbionts in the photosynthetic eukaryotes and horizontal transfers of the rickettsial genes to *Volvox carteri*

（植物細胞内共生リケッチアの分子同定とリケッチア遺伝子のボルボックスへの水平伝播）

平成 26 年 12 月博士（理学）申請

東京大学大学院理学系研究科

生物科学専攻

川舩　かおる

Contents

i

Abbreviations

BI: Bayesian inference

ccmF: cytochrome C type biogenesis protein CCMF

CDS: coding sequence

DAPI: 4′-6-deamidino-2- phenylindole

ddlB: D-alanine-D-alanine-ligase B gene

EGT: emdosymbiotic gene transfer

ftsQ: cell division septal protein FtsQ gene

HGT: horizontal gene transfer

HSP: high-scoring segment pair

ITS: internal transcribed spacer

ML: maximum likelihood

MP: maximum parsimony

murB: UDP-N-acetylenolpyruvylglucosamine reductase gene

murC: UDP-N-acetylmuramate-alanine ligase gene

PCR: polymerase chain reaction

phbB: acetoacetyl-CoA reductase gene

plsC: 1-acyl-sn-glycerol-3-phosphate acyltransferase gene

PP: posterior probability

rDNA: ribosomal DNA

rRNA: ribosomal RNA

TEM: transmission electron microscopy

tRNA: transfer RNA

Abstract

Endosymbiosis is a phenomenon that has often appeared in diverse species and has a great impact on evolution. During endosymbiosis, horizontal gene transfers from the genome of endosymbiont to that of host, which is called endosymbiotic gene transfer (EGT), may contribute to a closer symbiotic relationship and presumably may lead the endosymbionts to organelle. The mitochondria are regarded to originate from a bacterial endosymbiont related to rickettsiae (bacterial organisms in the order *Rickettsiales* [*Alphaproteobacteria*]). The bacterial family *Rickettsiaceae*, one of the families of *Rickettsiales*, comprises obligate intracellular bacteria and is principally associated with arthropods, containing human-pathogenic species studied with a great medical attention. By contrast, recently the bacterial endosymbionts of *Rickettsiaceae* have been found in cells of non-arthropodal organisms.

Two species of Volvocales (Chlorophyceae)*, Carteria cerasiformis* strain NIES-425 and *Pleodorina japonica* strain NIES-577 contain bacterial endosymbionts under the transmission electron microscopy. Previously in my Master's thesis, I determined the bacterial 16S ribosomal RNA (rRNA) gene sequences from *C. cerasiformis* NIES-425 and *P. japonica* NIES-577. These bacterial sequences phylogenetically belonged to the "hydra group", which is a clade comprising bacteria hosted by non-arthropodal organisms within *Rickettsiaceae*. For the molecular identification, however, it is required to be demonstrated that an intracellular localization of these rickettsial sequences is consistent with the bacterial endosymbionts, by cell examination of fluorescence *in situ* hybridization (FISH).

In the present thesis, I performed the molecular identification of the bacterial endosymbionts in the cells of *Volvox carteri* strain UTEX 2180, along with those in *C. cerasiformis* and *P. japonica* based on analyses of the 16S rRNA gene in Chapter 2. The bacterial 16S rRNA gene sequence obtained from *V. carteri* was located in the hydra group, where it formed a clade with the sequences of *C. cerasiformis* and *P. japonica*. Using a probe specific for the bacterial 16S rRNA sequences derived from these

endosymbiont-containing strains, FISH clearly indicated that bacterial endosymbionts belong to the hydra group within *Rickettsiaceae*. By 4'-6-diamidino-2-phenylidole (DAPI) staining and genomic PCR of 13 strains of *V. carteri*, the rickettsial endosymbionts were detected only in strain UTEX 2180. This result and similar observations of *Carteria* strains in my Master's thesis suggest that the rickettsia may have been transmitted to only a few algal strains very recently. Nevertheless, in preliminary work I detected a sequence similar to that of a rickettsial 16S rRNA gene in the nuclear genome of *V. carteri* strain EVE. It implies an involvement of rickettsial EGT, even though strain EVE was observed to lack a rickettsial endosymbiont.

In Chapter 3, I explored the origin of the rickettsial gene-like sequences in the endosymbiont-lacking *V. carteri* strain EVE, by performing comparative analyses on 13 strains of *V. carteri*. I constructed draft genomic sequences of rickettsial endosymbionts in *C. cerasiformis* strain NIES-425 cells. The draft genome contains 80 contigs (≥5 kbp) amounting to 1.27 Mbp. By reference to these rickettsial genomic sequences, I confirmed that an approximately 9-kbp DNA sequence encompassing a region similar to that of four rickettsial genes was present in the nuclear genome of *V. carteri* strain EVE. In addition, I identified rickettsial gene-like sequences in the nuclear genome of other eight strains of *V. carteri*. The compositions of them were classified in four types. Phylogenetic analyses, and comparisons of the synteny of rickettsial gene-like sequences indicated that the rickettsial gene-like sequences in the nuclear genome of seven strains of *V. carteri* including strain EVE were closely related to rickettsial gene sequences of *P. japonica,* rather than those of *V. carteri* strain UTEX 2180.

The rickettsial gene-like sequences found in the nuclear genomes of various *V. carteri* strains are considered to be acquired via EGT events from rickettsial endosymbionts, which might have been defunct currently. At least two different rickettsial organisms may have invaded the *V. carteri* lineage, one of which would be the direct ancestor of the endosymbiont of *V. carteri* strain UTEX 2180, whereas the other would be closely related to the endosymbiont of *P. japonica.* EGT from the latter rickettsial organism may have occurred in an ancestor of *V. carteri*. Thus, the rickettsiae

3

are thought to be widely associated with *V. carteri*, and likely have often been lost during host evolution. The present thesis implies the dynamic nature of the rickettsiae that frequently invade and disappear from the cells of Volvocales.

1. General introduction

Endosymbiosis, which represents a long-term association between one organism (host) and another organism(s) (endosymbiont) living intracellularly within the host, is the phenomenon often appeared in diverse species and has made a great impact on the biological evolution. The present definition of endosymbiosis covers a wide range of relationship levels: parasitic, commensal and mutualistic symbiosis. The endosymbiosis evolutionally merges two different lineages, often across the domains of life, and could be another driving force than that by mutations occurring on genetic information of each lineage (Wernegreen, 2012). Certain endosymbionts are known to confer the ability of primary energy production (photosynthesis and chemosynthesis), nutrient recycling and producing defensive chemicals to their hosts (Wernegreen, 2012; Clay, 2014), resulting in the successful adaptation of host organisms to the various environment.

During endosymbiotic process, the genes of the endosymbiont were considered to be finally transferred to the genome of its host (endosymbiotic gene transfer, EGT) (McCutcheon & Keeling, 2014). EGT has been observed in host genomes of diverse endosymbiotic relationships, but a majority of gene sequences acquired via EGT has no evidence of its translation or expression and would seem to collapse by mutation. In a few cases the translation and the expression of EGT genes on the genome of eukaryotic hosts were reported, where the EGT gene-derived products were relocated into the cells of endosymbionts (Nowack & Grossman, 2012; Nakabachi et al., 2014). Successful EGT and establishment of a retargeting system of expressed proteins would promote an integration of the host organism and the endosymbiont (McCutcheon & Keeling, 2014). Meanwhile, a genome of the endosymbiont tends to reduce by mutation and genetic drift that induced by small genetic population sizes and asexuality (McCutcheon & Moran, 2011). Eventually, only an almost minimum set of genes including those contributing to mutualistic effect for the host would be left in its genome (McCutcheon & Moran, 2011), as seen in the case of "*Candidatus* Nasuia deltocephalinicola" (*Gammaproteobacteria*), which is an endosymbiont of the leafhopper

5

and has a genome of only 112 kbp in length (Bennett & Moran, 2013). The reduction of endosymbiotic genome should also increase the dependence of endosymbionts on their hosts.

Mitochondria and plastids, the organelles with bilayer membranes and their own genome DNAs, are regarded to originate from bacterial endosymbionts (Figure 1). The base of this endosymbiosis theory was established by Dr. Lynn Margulis (Sagan, 1976; Margulis, 1970), and thereafter the physiological, biochemical and phylogenetical evidence have supported this thesis (reviewed in Zimorski et al., 2014). Though the phylogenetic details are still in controversy (Gray, 2014), it is broadly agreed that mitochondria might have evolved from an ancestor related to the bacteria of the order *Rickettsiales* (Andersson et al., 1998; Fitzpatrick et al., 2006), and that plastids from a cyanobacterial ancestor (Reyes-Prieto et al., 2007).

During the evolution of organelles from bacterial endosymbionts, EGT to the host ancestral eukaryotic genome would have contributed to establishing the host's control of the endosymbiont (McCutcheon & Keeling, 2014). The transferred genes began to be translated and expressed by the host's eukaryotic system, with the acquisition of eukaryotic promoters, introns and regulatory elements. The products of transferred gene would become available for the endosymbionts when targeting signals for transportation into the endosymbiotic compartments were added to them (Dyall et al., 2004). In parallel, the endosymbionts would have reduced their genomes in size during their transition to organelles. The genomes of the *Rickettsiales* and cyanobacteria contain ca. 900-1500 and ca. 6000 genes, respectively (Merhej & Raoult, 2011; Reyes-Prieto et al., 2007). By contrast, mitochondria and chloroplasts have now 5-91 genes and ca. 100-200 genes in their genome respectively (Nosek & Tomaska, 2003; Reyes-Prieto et al., 2007), and the ancestral bacteria of mitochondria and chloroplasts are predicted to devote ca. 630 and 700->4000 genes to nuclear genome (Kleine et al., 2009; Reyes-Prieto et al., 2007). Finally, the endosymbionts had almost completely depended on the host's expression systems in the result of EGT and reduction of the their genomes.

The order *Rickettsiales* (*Alphaproteobacteria*), generally termed the rickettsiae, contains Gram-negative obligate intracellular bacteria (Dumler & Walker, 2005). The order *Rickettsiales* contains three families, *Rickettsiaceae*, *Anaplasmataceae*, and *Holosporaceae* (Dumler and Walker, 2005; Euzéby, 2006) (Figure 2). In addition, recent study has proposed a new family, "*Candidatus* Midichloriaceae", which includes an intramitochondrial bacterium "*Candidatus* Midichloria mitochondrii" (Montagna et al., 2013). The *Rickettsiales* includes well-studied pathogens of mammals (Yu & Walker, 2012); and manipulators of host sexual reproduction (Engelstädter & Hurst, 2009).

The family *Rickettsiaceae*, one of the families of *Rickettsiales*, contains two genera, *Rickettsia* and *Orientia*, both of which are hosted principally by arthropod cells (Dumler & Walker, 2005). These genera include many agents causing tick-borne disease (Balraj et al., 2009). These bacteria of *Rickettsiaceae* invade the endothelial cells by inducing endocytosis, and then they became free from the phagosomal membrane in the host cytoplasm (Balraj et al, 2009).

In the recent decades, there were also found bacteria of the *Rickettsiaceae* inhabiting as endosymbionts of non-arthropod organisms including leeches (Kikuchi et al., 2002; Kikuchi & Fukatsu, 2005), a haplosporidian (Hine et al., 2002), an amoeba (Dykova et al., 2003), hydras (Fraune & Bosch, 2007), and ciliates (Vannini et al., 2005; Ferrantini et al., 2009; Sun et al., 2009; Schrallhammer et al., 2013; Vannini et al., 2014). Most of those endosymbionts hosted by non-arthropods form a monophyletic group (termed the "hydra group") within the family *Rickettsiaceae*, based on phylogenetic analyses of 16S ribosomal RNA (rRNA) genes (Weinert et al., 2009; Schrallhammer et al, 2013). In addition, a endosymbiont related to the *Rickettsiaceae* was discovered within the cells of the plastid-lacking heterotrophic euglenid flagellate *Petalomonas sphagnophila* (Kim et al., 2010). Little is known about the virulence or contribution of these endosymbionts to their non-arthropod hosts (Perlman et al., 2006). On the other hand, *Rickettsiaceae* endosymbionts have not been reported within the cells of non-phagotrophic, photosynthetic eukaryotes, such as primary photosynthetic eukaryotes, or Archaeplastida (green plants [land plants and green algae], red algae,

and glaucophytes).

The order Volvocales comprises flagellate green algae that are mainly found in freshwater environments (Nozaki, 2003), including unicellular *Chlamydomonas* and multicellular *Volvox*. Presence of endosymbiotic bacteria within the cytoplasm was first reported in *Volvox carteri* by transmission electron microscopy (TEM) (Kochert & Olson, 1970). These endosymbionts were rod-shaped and localized in the cytoplasm of the host cells without encompassing host-derived membranous structures (Kochert & Olson, 1970). Similar endosymbiotic bacteria were subsequently found in other volvocaleans, including two colonial species, *Pleodorina japonica* and *Eudorina illinoisensis*, and quadriflagellated unicellular *Carteria cerasiformis*, by TEM and/or 4′-6-diamidino-2-phenylindole (DAPI) staining (Nozaki et al., 1989; Nozaki & Kuroiwa, 1992; Nozaki et al., 1994), but the molecular identities and phylogenies of these bacterial endosymbionts had remained unresolved.

In my Master's Thesis (Kawafune, 2011), I performed molecular phylogenetic analyses of the intracellular bacteria in *C. cerasiformis* strain NIES-425 and *P. japonica* strain NIES-577 (Figure 3A,B). Bacterial 16S rRNA gene sequences that obtained from the cells of two green algal species are positioned within the bacterial family *Rickettsiaceae* and form a small clade with rickettsial endosymbionts hosted by freshwater hydra and freshwater and brackish ciliates within the hydra group (Kawafune, 2011). However, the intracellular localization of endosymbionts containing the bacterial sequences positioned within the family was not demonstrated by cell examination of fluorescence *in situ* hybridization (FISH).

In this thesis, I performed molecular identification of the endosymbionts in *C. cerasiformis*, *P. japonica* and *V. carteri* by FISH in Chapter 2, and demonstrated the first molecular evidence for rickettsial endosymbionts within non-phagotrophic photosynthetic cells. In Chapter 3, I explored the origin of the rickettsial gene-like sequences in the nuclear genome of the endosymbiont-lacking *V. carteri* strain EVE by the genome-wide similarity search using a newly constructed draft genome of *C. cerasiformis* rickettsial endosymbiont.

2. Molecular identification of bacterial endosymbiont in the cell of three species of Volvocales

Introduction

Previously, I demonstrated that the bacterial 16S rRNA sequences obtained from total DNA of *C. cerasiformis* and *P. japonica* strains that contain bacterial endosymbionts belong to the family *Rickettsiaceae* phylogenetically (Kawafune, 2011). Molecular identification of bacterial endosymbionts, which were often difficult to isolate and to cultivate in artificial medium, needs a clear evidence that the bacterial sequences are localized in the endosymbionts within their host cells, by the means for cell examination such as FISH.

The green alga *V. carteri* serves as a model organism for studies on multicellularity and the evolution of sexual reproduction (Herron et al., 2009; Ferris et al., 2010). In *V. carteri* strain UTEX 2180 (Figure 3C), Kochert and Olson (1970) observed intracellular bacteria using TEM. The endosymbionts were found in the cytoplasm of all cell types of the host *V. carteri*, including somatic cells, gonidia, androgonidia, and sperm (Kochert & Olson 1970; Lee & Kochert 1976). The endosymbiont formed rod-shaped cells, resided free in the host cytoplasm without being engulfed by host phagosomal membranes, and replicated by binary fission (Kochert & Olson 1970; Lee & Kochert 1976). In the intracellular location free from host phagosomal membranes, they are similar to the possible rickettsial endosymbionts in cells of *P. japonica* and *C. cerasiformis* (Nozaki et al., 1989; Nozaki et al., 1994), but the endosymbiont in *V. carteri* has not been identified using molecular methods.

In this chapter, I examined the bacterial endosymbiont of *V. carteri* in phylogenetic analyses of the 16S rRNA gene. The results clearly demonstrate that the bacterial sequence obtained from *V. carteri* cells harboring bacterial the endosymbionts belonged to the hydra group, as well as those from *C. cerasiformis* and *P. japonica*. I then performed FISH with the oligonucleotide probe specific to the hydra group, and

demonstrated that the cells of the hydra group bacteria were present in the cytoplasm of three species of Volvocales. In addition, a distribution of the rickettsial endosymbionts was examined in various strains of *V. carteri* by DAPI staining and genomic PCR.

Materials and Methods

Cultures.

All strains of Volvocales used in this thesis were supplied from Culture Collection of Algae at the University of Texas at Austin (UTEX, Austin, TX, USA; http://www.utex.org/) and the Microbial Culture Collection at the National Institute for Environmental Studies, Japan (NIES) (Kasai et al., 2009) (Table 1). Free-living bacterial contaminants in *V. carteri* strains were removed from the culture using the treatment described by Isaka et al. (2012). Purified cultures were inoculated into B-V medium (Ichimura & Watanabe, 1977) and incubated at 37°C in darkness for 1 week to verify sterility. No bacterial growth was detected outside the *V. carteri* cells in fixed and DAPI-stained materials using epifluorescence microscopy. The cultures of *V. carteri* were grown in screw-cap tubes (18x150 mm; Fujimoto Rika, Tokyo, Japan) containing 10 mL of VTAC medium (Nozaki et al., 1989) and then maintained at 25°C under irradiance with *ca*. 100-150 μmol photons/m^{2}s, with a 12-h:12-h light-dark (LD) photoperiod provided by cool-white fluorescent lamps. Cultures of strains of *P. japonica* and *C. cerasiformis* (Table 1) were grown as described above except that I used AF-6 medium (Kato, 1982) modified as described by Kasai et al. (2009). The cultures were maintained at 20°C, under irradiance with ca. 100–150 mmol photons/m^{2}s, with a 14:10 h LD photoperiod provided by cool-white fluorescent lamps.

PCR amplification and sequencing.

Total DNA from strains of Volvocales (Table 1) was extracted after cultured cells were boiled for 5 min in the extractin buffer containing 1M NaCl, 70mM Tris and 30mM Na^{2}EDTA (pH 8.0), according to the modified protocol (Nakada & Nozaki, 2007) of Fawley and Fawley (Fawley & Fawley, 2004) as follows. Boiled cells were mixed with equal volume of chlorofolm and x1/12 volume of 10% hexadecyltrimethylammonium bromide, and disrupted with ceramic beads using a Mixer Mill MM300 (Retsch, Haan, Germany). The GFX Genomic Blood Prification Kit (GE Healthcare, Buckinghamshire,

UK) was used to extract DNA from aqueous layer of the disrupted cells.

To sequence bacterial 16S rRNA genes, PCR (35 cycles of 94°C 20s, 55°C 30s and 72°C 2 min followed by 72°C for 7 minutes) was performed using total DNA of *V. carteri* UTEX 2180 with the universal primers 9F and 1492R (Weisburg et al., 1991) and *TaKaRa LA Taq* with GC I Buffer (Takara Bio Inc., Otsu, Japan). The product was purified using an illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) and directly sequenced on an ABI PRISM 3130xl Genetic Analyzer (Life Technologies, Carlsbad, CA) using a BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.1 (Life Technologies), with internal primers specific for the rickettsial endosymbiont of *C. cerasiformis* (Table 2; Kawafune, 2011) and two new internal primers specific for the endosymbiont of *V. carteri* (forward primers eveFC and FJ, and the reverse primer RH; Table 2).

Phylogenetic analysis.

The 16S rRNA gene-like sequences of *V. carteri* UTEX 2180 were aligned using ARB software (Ludwig et al., 2004) with those of 40 members of the hydra group and other members of the *Rickettsiaceae* family (Table 3). No chimeric trait was detected in the aligned sequences using UCHIME (Edger et al., 2011) and Mallard (Ashelford et al., 2006). The alignment, including 1403 bp corresponding to positions 38-1482 of *Escherichia coli* on ARB (J01695.2), was corrected manually referring to the secondary structure and is available from TreeBASE (http://www.treebase.org/treebase-web/; Submission ID: 16773, with additional two sequences for analysis in Chapter 3). Phylogenetic analyses were performed using the maximum likelihood (ML) method, the maximum parsimony (MP) method, and Bayesian interference (BI) with PhyML (Guindon et al., 2010), PAUP 4.0b10 (Swoffold, 2003), and MrBayes 3.2 (Ronquist et al. 2012), respectively. The GTR+gamma+I model was selected by jModelTest 2 (Guindon & Gascuel, 2003; Darriba et al., 2012) as the best evolutionary model that fits the data matrix. ML and MP analyses were performed with 1000 bootstrap replications. MrBayes 3.2 was performed using four different Markov Chain Monte Carlo runs with

three cold chains and one heated chain for 500,000 generations. Since the genus *Orientia* and *Candidatus* Cryptoprodotis polytropus are positioned most basally within the monophyletic family *Rickettsiaceae* (Ferrantini et al., 2009; Kawafune, 2011), they were designated as the outgroup.

In situ hybridization

Volv-835 (Table 4), an oligonucleotide probe targeting the 16S rRNA of bacterial endosymbionts of *C. cerasiformis* NIES-425, was designed using the probe-designing function of ARB. The sequence specificity of the probe was checked by probeCheck (Loy et al., 2008), and the optimal temperature and formamide concentration for specific hybridization was estimated using DINAMelt (Markham & Zuker, 2005). Volv-835 did not match the chloroplast 16S rRNA of host species (*C. cerasiformis* NIES-425: AB688625, *P. japonica* NIES-577: AB688626, *V. carteri* UTEX 1885: X53904) *in silico*. In the SILVA database (Pruesse et al., 2007), probe Volv-835 matched only five 16 S rRNA sequences, including that of an endosymbiont of *D. appendiculata* with 99% identity to the endosymbionts of the species of Volvocales; the other sequences contained at least two mismatches.

EUB338MIX, which is a mixture of the probes EUB338 (Amann et al., 1990), EUB338-II (Daims et al., 1999), and -III (Daims et al., 1999), was used (Table 4) to detect a wide-range of eubacterial taxa. EUB338 matches the 16S rRNA sequences of the endosymbionts of *C. cerasiformis* NIES-425, *P. japonica* NIES-577 and *V. carteri* UTEX 2180, and their chloroplast 16S rRNA sequences *in silico*. These probes were labeled with 6-carboxy-fluorescein at their 5′ end. Two non-labeled oligonucleotides were also designed and used as helper probes (Fuchs et al., 2000) to Volv-835: help-volv1 and help-volv2 (Table 4). The sequences of the probes are also available from probeBase (http://www.microbial-ecology.net/probebase/).

FISH was performed according to the method of Noda et al. (2003) with some modifications. Actively growing, 5- to 7-day-old cultures were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) overnight at 4ºC and washed in

PBS twice. Fixed cells were incubated with enzyme solution (1% cellulase Onozuka RS, 1% macerozyme R-10 [Yakult Pharmaceutical industry, Tokyo, Japan] in PBS) for 5 min, washed in PBS, and suspended in 1% Tween 20 (Sigma-Aldrich, St Louis, MO) solution in PBS. After washing in PBS twice, the treated cells were spotted on an aminosilane-coated glass slide (Matsunami Glass, Osaka, Japan), air-dried, and treated with 0.25N HCl for 20 min at room temperature. Samples were then washed with distilled water and sequentially dehydrated in 50, 80, 90, and 100% ethanol. Oligonucleotide probes with 20% formamide in the hybridization buffer (0.1 M Tris-HCl, 0.9 M NaCl) were then applied, sealed in an incubation chamber (CoverWell; Grace Bio-Labs, Bend, OR), and incubated for 90 min at 51°C (*V. carteri* cells and the EUB338MIX probes were incubated in hybridization buffer without formamide). After washing in wash buffer (0.1 M Tris-HCl, 0.2 M NaCl) and high stringency wash buffer (20 mM Tris-HCl, 40 mM NaCl), slides were mounted with enclosing liquid (90% glycerol, 1% triethylenediamine in PBS) containing DAPI, and observed under an epifluorescence microscope, the Olympus BX-60 (Olympus, Tokyo, Japan).

Detection of rickettsial endosymbionts by DAPI staining and genomic PCR.
For DAPI staining, an actively growing, 5- to 8-day-old culture was fixed for 30 min with a final concentration of 2% glutaraldehyde at room temperature. The fixed materials were added to an equal volume of DAPI solution (1 mg/ml) in TAN buffer (Takahara et al., 1999) and squashed by pushing the cover glass onto the cells. These materials were observed under an epifluorescence microscope (BX-60; Olympus) equipped with Nomarski interference.

To detect the rickettsial endosymbionts, genomic PCR was carried out in 13 *V. carteri* strains (Table 1) using the hydra group 16S rRNA-specific primers eveFC and enRB and *TaKaRa LA Taq* with GC I Buffer, with 35 cycles of 94°C 20s, 55°C 30s and 72°C 2 min, followed by 72°C for 7 minutes. As positive control, the eukaryotic 18S rRNA gene was amplified in same conditions, with primers 18S-FA and 18S-RF (Table 2).

Results

<u>Phylogenetic analysis of the bacterial endosymbiont in *V. carteri*</u>

The nearly complete sequence of the bacterial 16S rRNA gene (1431 bp; AB861537) was obtained from *V. carteri* UTEX 2180 cells. A BLASTN search of the NCBI nucleotide database (http://www.ncbi.nlm.nih.gov/) revealed that the *V. carteri* endosymbiont had the highest blast score with the *Rickettsiaceae* endosymbiont of *C. cerasiformis* NIES-425 (AB688628) under a filter excluding uncultured/environmental sample sequences. The similarity in the 16S rRNA sequences of the endosymbionts between the hosts *V. carteri* and *C. cerasiformis* NIES-425 or *P. japonica* NIES-577 (AB688629.) was 99.6%.

A phylogenetic analysis of the endosymbiont from *V. carteri* UTEX 2180 and 40 rickettsial bacteria, including the two endosymbionts of the green algae *C. cerasiformis* NIES-425 and *P. japonica* NIES-577 (Table 3), showed that the family *Rickettsiaceae* was subdivided into two robust monophyletic groups (Figure 4). The first group included the genus *Rickettsia* and some endosymbionts of leeches. The second corresponded to the hydra group. The hydra group was subdivided into sister clades A and B. Clade A was supported with 69-91% bootstrap values by the ML and the MP methods and 1.00 posterior probability (PP) by BI, whereas clade B was robust with 91-97% bootstrap values and 1.00 PP. The endosymbiont of *V. carteri* was placed in clade B, with those of *P. japonica*, *C. cerasiformis*, the freshwater ciliates (*Euplotes octocarinatus*, *Spirostomum* sp., *Paramecium caudatum*), and the brackish ciliate *Diophrys oligothrix*, and with uncultured bacterial sequence clones originating from various aquatic environments. Clade A contained a bacterial sequence obtained from the coral *Montastraea faveolata*, and the endosymbionts of *Hydra oligactis*, and the freshwater ciliate *Ichthyophthirius multifiliis*. The endosymbiont from *V. carteri* shared 88.8-99.6% sequence similarity with bacteria of the *Rickettsiaceae* and 94.3-99.6% with those of the hydra group.

Cell identification of rickettsial endosymbiont in three species of Volvocales by FISH

To identify bacteria corresponding to the obtained rickettsial sequences, I designed a specific oligonucleotide probe, Volv-853, with helper probes help-volv1 and help-volv2, targeting 16S rRNA (see Materials and Methods). In *C. cerasiformis* NIES-425, endosymbiont-specific signals (Volv-835) were detected exclusively from the rod-shaped bodies within the cytoplasm (Figure 5A–C). These bodies corresponded almost exactly to the rod-shaped light blue fluorescence when visualized by DAPI-staining (Figure 5A, B). The EUB338MIX (targeting 16 S rRNA of eubacteria and chloroplasts, see Materials and Methods) signals showed net-like hybridization patterns within the chloroplasts of *C. cerasiformis* NIES-425 cells (Figure 6A–C). In contrast, within the cells of *C. cerasiformis* NIES-424, no fluorescent signal from probe Volv-835 was detected (Figure 5D–F), while EUB338MIX showed the same net-like pattern as in *C. cerasiformis* NIES-425 (Figure 6D–F). In *P. japonica* NIES-577, the endosymbiont-specific signal of Volv-835 showed rod-shaped bodies corresponding to the cytoplasmic DAPI signals, as seen in *C. cerasiformis* NIES-425 (Figure 5G-I). Similarly, endosymbiont-specific rod-shaped signals of volv-835 were observed in *V. carteri* UTEX 2180 somatic cells (Figure 5J-L). These signals were consistent with the high-intensity fluorescence of the DAPI-stained bacteria (Figure 5J, L). Fluorescent signals of EUB338MIX appeared as a combination of the two patterns in *V. carteri* somatic cells: a rod-shaped pattern in the cytoplasm (endosymbionts) and larger signals around the pyrenoid (chloroplast stroma) (Figure 6G-I). The intracellular localization of the rickettsial endosymbionts in *V. carteri* gonidia was also determined using FISH as described above (not shown).

Detecting the presence/absence of rickettsial endosymbionts in various strains of *V. carteri*

Under epifluorescence microscopy, the DAPI-stained DNA and chloroplasts exhibited light blue and red fluorescence (Figure 7) in all *V. carteri* strains examined (Table 1). Somatic cells of *V. carteri* UTEX 2180 contained one eukaryotic cell nucleus and nearly

circular bodies, identified as chloroplast nucleoids (Figure 7A). These somatic cells contained 5-12 bacterial endosymbionts as highly fluorescent rod-shaped bodies in the cytoplasm (Figure 7A). The rod-shaped bodies were distributed mainly at a peripheral region outside the chloroplast or around the cell nucleus. These epifluorescence microscopic features of the endosymbionts were essentially the same as those in *P. japonica* (Nozaki et al., 1989) and *C. cerasiformis* (Kawafune, 2011). By contrast, no rod-shaped highly fluorescent bodies were found in the cytoplasm of 12 other strains of *V. carteri* (Figure 7B-D,I-L). The chloroplast nucleoids in the somatic cells of these strains were circular or polygonal in shape (Figure 7B-D,I-L).

To detect rickettsial endosymbionts, genomic PCR using *Rickettsiaceae*-specific 16S rRNA primers was performed in 13 strains of *V. carteri* (Table 1). A single band indicating the presence of rickettsia was detected only in UTEX 2180 and the positive controls (*C. cerasiformis* NIES-425 and *P. japonica* NIES-577; Kawafune, 2011), while no amplification was observed in the remaining eight *V. carteri* strains and in the negative control *C. cerasiformis* NIES-424 (Kawafune, 2011) (Figure 8).

Discussion

My data clearly demonstrate that the cells of *C. cerasiformis* NIES-425, *P. japonica* NIES-577 and *V. carteri* UTEX 2180 harbor rickettsial bacteria. Phylogenetic analysis of 16S rRNA, and FISH with a specific probe indicated that the bacterial endosymbiont of three species of Volvocales belongs to the hydra group in the family *Rickettsiaceae* (Figures 4, 5). This result agrees with the morphological characteristics of *V. carteri* UTEX 2180 endosymbionts reported in previous TEM studies (Kochert & Olson, 1970; Lee & Kochert, 1976); the bacteria of the family *Rickettsiaceae* grow freely in the host cytoplasm surrounded by an electron-transparent area, and they are not enclosed in a host-derived phagosomal membrane (Dumler & Walker, 2005). These morphological features of the endosymbiont are essentially identical to those of the *C. cerasiformis* endosymbiont (Kawafune et al., 2013), which also belongs to clade B of the hydra group (Figure 4). All of the bacterial endosymbionts hosted by species of Volvocales that were identified by the present molecular data belong to clade B (Figure 4). Although information is limited, this fact suggests a level of host-environment specificity among rickettsial endosymbionts in the hydra group.

Present genomic PCR and DAPI staining showed that rickettsia are present only in the cells of UTEX 2180 among 13 strains of *V. carteri*. A similar situation exists in *Carteria*, in which endosymbionts were detected in only *C. cerasiformis* NIES-425 cells among 10 strains of the four related *Carteria* species, by TEM (Nozaki et al., 1994), DAPI-staining and PCR using *Rickettsiaceae*-specific primers (Kawafune, 2011). The results of these two lineages imply that rickettsiae might be horizontally transferred to the host strains independently, after the divergence of host strains.

3. Horizontal gene transfer of the rickettsial endosymbiont to *Volvox carteri* genome

Introduction

The nuclear genomic sequence of *V. carteri* f. *nagariensis* strain EVE determined recently (Prochnik et al., 2010). Although no rickettsial endosymbiont was evident in this strain (based on genomic PCR using the hydra group-specific primers, and DAPI-staining of algal cells; Chapter 2), I detected in preliminary work a sequence similar to that of rickettsial 16S rRNA genes in the nuclear genome of *V. carteri* f. *nagariensis* strain EVE. The presence of rickettsial gene-like sequences in the nuclear genome of this strain may indicate the occurrence of horizontal gene transfer (HGT) from a rickettsial bacterium to an ancestor of *V. carteri* f. *nagariensis* strain EVE.

In this chapter, I examined rickettsial gene-like sequences encoded in the *V. carteri* nuclear genomes using a draft genome of *C. cerasiformis* strain NIES-425 rickettsial endosymbiont newly constructed. I elucidated the origin of rickettsial gene-like sequences within the nuclear genome of *V. carteri* f. *nagariensis* strain EVE, based on comparison with 12 other strains of *V. carteri*. Phylogenetic analyses of rickettsial gene and gene-like sequences from the endosymbionts, and the nuclear genomes of species of Volvocales, indicated that the rickettsial sequence obtained from *V. carteri* f. *nagariensis* strain EVE nuclear genome was closely related to that of *P. japonica,* rather than that of *V. carteri* f. *weismannia* strain UTEX 2180.

Materials and Methods

DNA extraction of rickettsial endosymbiont for genome sequencing

C. cerasiformis strain NIES-425 was grown in a 500 mL Erlenmeyer flask containing about 400 mL AF-6 medium bubbled with air at 20°C, under irradiance with ca. 100–150 mmol photons/m^2s, with a 14:10 h LD photoperiod provided by cool-white fluorescent lamps. One litter of an actively growing, 1-2-week-old *C. cerasiformis* culture was mixed with 5 μL 10% tween20 and concentrated by centrifuge (1,000 g, 5 minutes). The concentrated cells were dissolved in 300 mL fracturing buffer (25mM HEPES pH 7.5, 20mM KCl, 20mM $MgCl_2$, 20.5 g Sucrose, 10 g Glycerol, x1 Protease inhibiter solution per 100 mL). Cells were then fractured in twice by BioNeb disruption system (Glas-Col, Terre Haute, IN) and examined by DAPI-staining to confirm the almost complete fracturing of cells. In order to remove the host cell nuclei, broken cells in fracturing buffer with 0.5% NP40 were centrifuged at 8,000 g for 10 minutes and the pellet was removed. The supernatant was centrifuged at 10,000 g for 20 minutes for collecting bacterial cells. After the supernatant including chloroplast nucleoids and mitochondria floated was discarded, the pellet was washed in fracturing buffer with 0.5% NP40 by centrifuge (15,000 g, 10 minutes). The pellet and supernatant were checked by DAPI-staining at each step. Genomic DNA of isolated bacterial cells was extracted from this pellet by the illustra bacteria genomicPrep Mini Spin Kit (GE Healthcare).

Genome sequencing, assembly and annotation

For single-end sequencing, 1 μg of bacterial genomic DNA was sheared by Covaris S2 (Covaris, Woburn, MA) at 700 bp setting. A library from 700 bp sheared DNAs was prepared using NEBNext DNA Reagent Set for Illumina (New England BioLabs, Ipswich, MA) according to the manufacturer's protocols, with a modification that KAPA HiFi HotStart PCR Kit (KAPABIOSYSTEMS, Boston, MA) was used instead of NEBNext High-Fidelity 2X PCR Master Mix supplied by NEBNext DNA Reagent Set. The library was then sequenced on an Illumina Genome Analyzer IIx platform (GAIIx;

Illumina, San Diego, CA) with single-end 70 bp reads.

Generated 29,855,149 reads (ca. 2,090M bases) were *de novo* assembled using Velvet (Zerbino & Birney, 2008) provided in the DDBJ Read Annotation Pipeline (DDBJ Pipeline; http://p.ddbj.nig.ac.jp/pipeline/) (Kaminuma et al., 2010). At the k-mer length of 41, Velvet assembles reads to 167,152 contigs (ca. 24 Mbp) and generates the largest maximum contig size (61,032 bp). The number of contigs that are longer than 5000 bp was 82. These 82 contigs were then automatically annotated using the Microbial Genome Annotation Pipeline (MiGap; www.migap.org/) (Sugawara et al., 2009). Two contigs, on which no genes coding prokaryotic protein were identified, have been excluded from subsequent BLAST searches.

BLAST-based screening of the *V. carteri* genome for transferred rickettsial genes
A BLASTN search (Altschul et al., 1997) was performed on the *V. carteri* f. *nagariensis* EVE genome data (version 2, 8x, not masked) (Prochnik et al., 2010) on Phytozome version 9.1 (http://www.phytozome.net) (Goodstein et al., 2012) using 80 contigs (>5 kb; Table 5) derived from my preliminary genome assembly database of the rickettsial endosymbiont hosted by *C. cerasiformis* strain NIES-425 (acquired as part of the Plant Global Education Project of the Nara Institute of Science and Technology) as queries. High-scoring segment pairs (HSPs) with E-values ≤1.0e-50 were annotated (Table 6) using BLASTN and BLASTX searches (Altschul et al., 1997) against nucleotide data and non-redundant protein sequences lodged in the National Center for Biotechnology Information database (NCBI, http://www.ncbi.nlm.nih.gov). Sequences associated with the three HSPs (similar to the rickettsial 16S rRNA gene, the gene encoding UDP-N-acetylenolpyruvylglucosamine reductase [*murB*], and the gene encoding D-alanine-D-alanine-ligase B [*ddlB*]) that afforded the highest scores and E-values, were concentrated around base no. 940,000 of scaffold 6 on the *V. carteri* EVE genome. Additional BLASTN searching (using bases 935,001-945,000 of scaffold 6 as the query) against the 80 contigs of the *C. cerasiformis* NIES-425 rickettsial endosymbiont genome found one further short HSP (similar to rickettsial the cell division septal protein FtsQ

gene *ftsQ*; Table 6); this HSP was also annotated as described above.

Cultures, DNA extraction, and sequencing of the ITS region

V. obversus strain UTEX 1865, 13 strains of *V. carteri*, *C. cerasiformis* strain NIES-425 and *P. japonica* strain NIES-577 (Table 1) were grown as described in Chapter 2. *Chlamydomonas reinhardtii* strain CC-503 (*cw92* mt+) was supplied from Chlamydomonas Resource Center (http://chlamycollection.org) and was grown as well as those of *V. carteri*, except that TAP medium (Harris, 1989) was used. Total DNA was extracted as described in Chapter 2.

The internal transcribed spacer (ITS) regions of nuclear ribosomal DNA (rDNA) sequences of *V. carteri* strains and *V. obversus* UTEX 1865 (Table 1) were identified as described in Setohigashi et al. (2011) as follows, except for those of *V. carteri* f. *weismannia* strain UTEX 1875 and UTEX 1876, which are publicably available (Hiraide et al., 2013). PCR with *TaKaRa LA Taq* with GC I Buffer and direct sequencing of its product were performed using total DNA of *Volvox* strains with primers for ITS regions (Table 2) as described in Chapter 2.

Sequencing of rickettsial gene/gene-like sequences

To determine genomic sequences (including those of *murB*, *ddlB*, and *ftsQ* of rickettsial endosymbionts), PCR primers ccmF-R02 and phbB-F01 (Table 2) were designed based on preliminary genomic information on the *C. cerasiformis* NIES-425 endosymbiont. Using these primers, ca. 10-kbp segments lying between the gene encoding cytochrome C type biogenesis protein CCMF (*ccmF*) and acetoacetyl-CoA reductase (*phbB*) were amplified from total DNAs of *C. cerasiformis* strain NIES-425, *P. japonica* strain NIES-577, and *V. carteri* f. *weismannia* strain UTEX 2180, via PCR (35 cycles of 98°C 10s, 60°C 15s and 68°C 5 min) with ca. 400 pg/μL DNA in 20 μL PCR solution, using Tks Gflex DNA Polymerase (Takara Bio Inc.).

In order to make sure the sequence of scaffold 6 of *V. carteri* f. *nagariensis* EVE genome, the sequence between bases 934,933-943,805 of scaffold 6 was amplified from

total DNA of that strain via PCR using a *TaKaRa LA Taq* with GC I buffer and specific primers (eveFZ, eveRJ, eveFX, eveRD, eveFV, eveRT, asmb81_F4, asmb82_R2; Table 2) with ca. 150 pg/μL DNA in 20 μL PCR solution, for 35 cycles of 94°C 20s, 53-55°C 30s and 72°C 2-5 min, followed by 72°C for 7 minutes. Each PCR product was purified and directly sequenced as described in Chapter 2, with internal sequencing primers (Table 2).

For sequencing the rickettsial gene-like sequence of other *V. carteri* strains, genomic PCR was performed using various combinations of the rickettsial gene and gene-like sequence-specific primers (Table 2). These primers were appropriately designed by reference to such sequences, as were primers amplifying 16S rRNA gene sequences of rickettsial endosymbionts (AB688628, AB688629, and AB861537) determined in Kawafune (2011) and Chapter 2 of this thesis. PCR was performed with ca. 6-400 pg/μL DNA in 20 μL PCR solution, with 35 cycles of 94°C 20s, 53-57°C 30s and 72°C 2-5 min, followed by 72°C for 7 minutes. PCR products were directly sequenced as described above. In order to determine sequences around the *murB*-like sequence, vectorette PCR was also performed as described in Ko et al. (2003) on the total DNA of *V. carteri* f. *weismannia* strain UTEX 2170, using *TaKaRa LA Taq* with GC I Buffer and the restriction enzymes *Eco*RI (TOYOBO Co., Ltd., Osaka, Japan) and *Apo*I (New England Biolabs). PCR products were directly sequenced as described in Chapter 2.

Detection of rickettsial gene and gene-like sequence by genomic PCR and semi-quantitative PCR

For confirming the detection of the rickettsial gene and gene-like sequence, genomic PCR was re-performed on total DNAs of *V. carteri* strains using the rickettsia-specific primer sets (16S rRNA 5′-region: eveFC and eveRD; 16S rRNA 3′-region: enFN and enRG; *murB*: murB-FP and murB-RK2; *ddlB*: ddlB-FQ and ddlB-RH2; 18S rRNA gene as positive control: 18S-FA and 18S-RD; see Table 2), with 35 cycles of 94°C 20s, 55°C 30s and 72°C 40 sec, followed by 72°C for 7 minutes. An initial DNA concentration of PCR solution was ca. 2-3 pg/μL in 20 μL PCR solution. Total DNA of *Chlamydomonas*

reinhardtii cc-503 was used as negative control; in the *Chlamydomonas reinhardtii* genome data (version 5.5, not masked) on Phytozome, no sequences similar to the rickettsial 16S rRNA gene, *murB*, *ddlB* and *ftsQ* were detected by a BLASTN search (E-values≤1.0e-50) using sequences of the rickettsial endosymbiont of *C. cerasiformis* strain NIES-425 as queries.

Semi-quantitative genomic PCR was also performed on total DNAs of *V. carteri* strains and *Chlamydomonas reinhardtii* CC-503 (as negative control), using *TaKaRa LA Taq* with GC Buffer I, and the same specific primer sets as the genomic PCR described above, with 27 cycles of 94°C 20s, 55°C 30s and 72°C 40 sec, followed by 72°C for 7 minutes and with a modification that the actin gene was amplified as a control (primers ONact1 and ONact2 for *V. carteri* and ONact1_CR and CR_IDA5_R3 for *Chlamydomonas reinhardtii* were used; see Table 2) instead of 18S rRNA gene. An initial DNA concentration of PCR solution was ca. 2-3 pg/μL in 20 μL PCR solution.

Phylogenetic analysis

The 16S rRNA gene-like sequences of *V. carteri* were aligned using ARB software (Ludwig et al., 2004) with a data matrix that has been used in Chapter 2. Alignment was corrected manually by reference to secondary structure and is available from TreeBASE (Submission ID: 16773).

The *murB* and *ddlB* gene-like sequences of *V. carteri* were manually aligned with the *murB* and *ddlB* nucleotide sequences of bacteria belonging to the family *Rickettsiaceae*, including those of the preliminary genomic sequence of the *C. cerasiformis* strain NIES-425 endosymbiont, and sequences obtained from total DNAs of *P. japonica* strain NIES-577 and *V. carteri* f. *weismannia* strain UTEX 2180 (Table 7). These nucleotide sequences were next converted to the translated amino acid sequences (available from TreeBASE; ID: 16773).

Phylogenetic analyses of 16S rRNA gene/gene-like sequences, translated *murB*, translated *ddl,* and combined [translated *murB* and *ddlB*] genes/gene-like sequences, were performed using both ML and MP methods, employing PhyML 3.0 (Guindon et

al., 2010) and PAUP 4.0b10 (Swoffold, 2003), respectively, with 1,000 bootstrap replications. In addition, 16S rRNA genes/gene-like sequences were subjected to BI testing using MrBayes 3.2 (Ronquist et al., 2012). For both ML analyses and BI, the GTR+gamma+I model was selected by jModelTest 2 (Guindon & Gascuel, 2003; Darriba et al., 2012) to analyze the data matrix of 16S rRNA genes/gene-like sequences; and the CpREV+gamma, JTT+gamma+I+F, and JTT+gamma+I+F models were selected by ProtTest3 (Darriba et al., 2011; Darriba et al., 2012) to analyze the the data matrixes of translated *murB*, *ddlB,* and combined genes/gene-like sequences, respectively.

The nuclear rDNA ITS-2 sequences of *V. carteri* strains and *V. obversus* strain UTEX 1865 were aligned based on secondary structures, predicted using the RNA Folding Form on the mFold Web Server (http://mfold.rna.albany.edu/?q=mfold) (Zuker, 2003); revised based on the data of previous studies (Coleman & Mai, 1997; Caisová et al., 2013); and drawn using VARNA 3.9 (Darty et al., 2009) (Figure 9). Phylogenetic analyses of aligned sequences (available from TreeBASE; ID: 16773) were performed as described by Nozaki et al. (2014).

Results

<u>Draft genome of rickettsial endosymbionts</u>

The 29,855,149 single reads of 70 bp were generated from DNA library of the *C. cerasiformis* NIES-425 endosymbiont and assembled to 167,152 contigs (total size: 23.7 Mbp), most of which were short (N50 contig size: 139 bp). Eighty-one contigs that were ≥5 kbp in length cover 1.27 Mbp (Table 5). This genome size was similar to those of *Rickettsia* that range 1.11 - 1.52 Mbp (Merhej et al., 2011) but smaller than the sizes of *Orientia* being 2.01 - 2.13 Mbp in length (Nakayama et al., 2010). The total G+C content of 80 contigs was 33.9% (Table 5), slightly higher than those of *Rickettsia* (28 - 32%) and *Orientia* (30.5%) (Nakayama et al., 2010; Merhej et al., 2011). These 80 contigs were automatically annotated and predicted to encode 1246 coding DNA sequences (CDSs) with 34 tRNA genes and one set of rRNA genes (5S, 23S and 16S). The number of coding DNA was in the range of *Rickettsia* genomes (835-1476) (Merhej et al., 2011). Based on automatic annotation by MiGap (Sugawara et al., 2009), 60% (747) of the CDSs were annotated to be the closest to genes of *Rickettsiales*.

<u>Sequences and synteny of rickettsial gene/gene-like sequences in nuclear genomes, and endosymbionts harbored by cells of the Volvocales</u>

I performed a BLASTN search using long contig sequences (>5 kbp), from my ongoing genome data collection of the *C. cerasiformis* strain NIES-425 rickettsial endosymbiont, against the *V. carteri* f. *nagariensis* strain EVE nuclear genome. I found that the three rickettsial gene-like sequences with the highest E-values (Table 6) were localized in a synteny region within scaffold 6 of the *V. carteri* f. *nagariensis* strain EVE nuclear genome (Figure 10A). The three genes were a 16S rRNA gene, *murB*, and *ddlB*. In addition to these three rickettsial gene-like sequences, the synteny contained a short sequence similar to that of *ftsQ* (found upon additional BLASTN searching) and three *V. carteri*-specific sequences (see below) within a region approximately 9 kbp in length (Figure 10A). In order to make sure that the synteny region is not contaminated and

correctly assembled, this 9 kbp-region was confirmed via direct sequencing of genomic PCR products from total DNA of *V. carteri* f. *nagariensis* strain EVE (see Figure 10A and Materials and Methods). No nucleotide difference was found between the ca. 9 kbp-region sequenced and the corresponding region of scaffold 6 of strain EVE genome.

The rickettsial 16S rRNA gene-like sequence within the synteny region of the *V. carteri* f. *nagariensis* strain EVE genome was 412 bp in length and corresponded to the 5′-region of the *C. cerasiformis* strain NIES-425 endosymbiont 16S rRNA gene (including the 5′ end). The 3′-region of the 16S rRNA gene-like sequence was truncated, and lay adjacent to a sequence similar to that of the rickettsial *murB* gene. This *murB*-like sequence was 768 bp in length, contained the 3′-end of the coding region, and exhibited no frameshift or premature stop codon, although the sequence seemed to lack a 5′-coding-region of 117 bp (including the start codon) when compared with *murB* of the *C. cerasiformis* strain NIES-425 endosymbiont. The sequence similar to that of rickettsial *ddlB* was interrupted by a 1,021-bp insertion ("*V. carteri* multicopy A" in Figure 10A), but had both a start and a stop codon, and exhibited no frameshift or premature stop codon (the insertion was excluded from analysis). The insertion (*V. carteri* multicopy A) had no sequence similar to that of rickettsial genes, but had 103 DNA sequences similar to those distributed in the *V. carteri* f. *nagariensis* strain EVE nuclear genome (E-value=0; nucleotide identity, 84-100%; a BLASTN search yielded these data). Thus, it appeared specific to *V. carteri*. The *ftsQ*-like sequence was short (75 bp), but exhibited high-level similarity (nucleotide identity: 88%) to the 5′-region (including the start codon) of *ftsQ* of the endosymbiont of *C. cerasiformis* strain NIES-425. A sequence resembling the 5′-region of the *V. carteri*-specific transposon *Jordan* (Miller et al., 1993) was also found (E value=0; nucleotide identity 96%); the sequence was located near the *ftsQ*-like sequence, but lacked the 5′-inverted terminal repeat found in *Jordan* (Miller et al., 1993). A DNA sequence of 516 bp located in the 5′-upstream region of the rickettsial 16S rRNA-like sequence ("*V. carteri* multicopy B" in Figure 10A) may also be *V. carteri*-specific because it had 32 similar DNA sequences (E values=0.0, nucleotide identity 95-98%) within the *V. carteri* f. *nagariensis* strain EVE nuclear genome (found

using a BLASTN search). This region exhibited no similarity to the rickettsial genes or DNA inserted in the *ddlB*-like sequence. At both ends of the ca. 9 kbp-DNA region including rickettsial gene-like sequences, two regions (asmbl_81.volvox20_pasa2 and asmbl_82.volvox20_pasa2) that match with EST sequences are identified on Phytozome version 9.1 (http://www.phytozome.net) (Goodstein et al., 2012). These regions were not similar to any sequences in my preliminary genome assembly database of the rickettsial endosymbiont of *C. cerasiformis* strain NIES-425, and their functions are unknown.

In the rickettsial endosymbiont genome of *C. cerasiformis* strain NIES-425, *murB*, *ddlB,* and *ftsQ* formed a synteny, but the 16S rRNA gene was separated from the three genes, being composed of two separate DNA sequences (Figure 10B). One sequence was 1,422 bp long and encoded a 16S rRNA gene positioned between genes encoding nucleoside triphosphate pyrophosphohydrolase (*mutT*) and a M23 superfamily membrane-bound metallopeptidase (*nlpD2*). The other sequence was 6,230 bp long and encoded *murB*, *ddlB*, *ftsQ*, UDP-N-acetylmuramate-alanine ligase (*murC*), and three other proteins (Figure 10B, black double-headed arrows on the baseline). In the endosymbiont genome of *C. cerasiformis* strain NIES-425, a CDS for a hypothetical protein was inserted between *murB* and *ddlB*.

Partial genomic DNA sequences including the *murC*, *murB*, *ddlB,* and *ftsQ* genes were obtained using total DNAs of rickettsial endosymbiont-containing cells of *V. carteri* f. *weismannia* strain UTEX 2180 (Figure 10C) and *P. japonica* strain NIES-577 (Figure 10D). These sequences were similar to that of the endosymbiont of *C. cerasiformis* strain NIES-425 in terms of gene arrangement. However, the *P. japonica* strain NIES-577 sequence differed from the other two sequences in that an additional CDS was lacking between *murB* and *ddlB,* and a CDS encoding a DDE transposase was located between DNA encoding 1-acyl-sn-glycerol-3-phosphate acyltransferase (*plsC*) and *phbB* (Figure 10D).

Rickettsial gene-like sequences in various *V. carteri* strains

Performance of genomic PCR using 16S rRNA primers specific to the hydra group, and DAPI-staining of algal cells from 13 strains of three forms of *V. carteri,* showed that *V. carteri* f. *weismannia* strain UTEX 2180 harbored a rickettsial endosymbiont, whereas the 12 other strains (including *V. carteri* f. *nagariensis* strain EVE) did not (Chapter 2; Figures 7, 8). Sequences similar to those of rickettsial genes were detected in nine endosymbiont-lacking strains of *V. carteri* via genomic PCR using specific primers (Table 2) targeting endosymbiont genes and rickettsial gene-like sequences in the nuclear genome of *V. carter*i f. *nagariensis* strain EVE (Figure 10). We sequenced the PCR products from all nine strains (Figures 11E, 12). The sequences were classified into four types: (A)-(D). Type A: *V. carteri* f. *nagariensis* strains UTEX 1886, NIES-397, and NIES-398 had sequences of 6,284-6,290 bp that were completely or almost identical (99-100% nucleotide identity) to part of scaffold 6 of the published *V. carteri* f. *nagariensis* strain EVE nuclear genome (Figure 10A), encompassing sequences similar to those of rickettsial 16S rRNA, *murB*, *ddlB*, *ftsQ, V. carteri*-specific transposon *Jordan*, and two *V. carteri*-specific sequences (*V. carteri* multicopies A and B) (Figure 11A); Type B: *V. carteri* f. *weismannia* strains UTEX 1875 and 1876 contained DNA sequences including five regions that were very similar to a sequence of *C. cerasiformis* strain NIES-425 (containing the *murC*, *murB*, *ddlB*, *ftsQ,* and *plsC* genes; 88-92% nucleotide identity; Figure 11B). In these five regions, the sequences similar to *murB*, *ddlB,* and *ftsQ* exhibited intact open reading frames (ORFs), whereas the *murC*-like sequences had frameshift mutations in the 3′ regions. The gene arrangement of this DNA sequence was similar to that of *P. japonica* strain NIES-577, in that a CDS encoding a hypothetical protein, lying between the *murB* and *ddlB* genes/gene-like sequences, was lacking. Type C: A 678 bp sequence (only), similar to that of truncated rickettsial *murB*, was detected in *V. carteri* f. *weismannia* strain UTEX 2170 (Figure 11C). This *murB*-like sequence lacked both the 5′ and 3′ ends, although no frameshift or premature stop codon was evident. Type D: In *V. carteri* f. *kawasakiensis* strains NIES-732 and NIES-733, no sequence exhibiting similarity to those encoding rickettsial *murB*, *ddlB*, or *ftsQ* was

detected, but I found a rickettsial 16S rRNA gene-like sequence (Figure 11D). This sequence was 882 bp long and corresponded to the 3′-region of the 16S rRNA gene from the endosymbiont of *C. cerasiformis* strain NIES-425 (Figure 10B). However, sequence corresponding to the partial 16S rRNA gene-like sequence (the 5′-region) found in the nuclear genome of *V. carteri* f. *nagariensis* strains was lacking.

Semi-quantitative PCR of rickettsial gene/gene-like sequences from *Volvox carteri* strains

To determine whether the rickettsial gene/gene-like sequences were present in the nuclear genome or in endosymbionts of the cytoplasm, I performed semi-quantitative PCR on total DNA of 13 strains of *V. carteri* (Figure 11E). Of these *V. carteri* strains, the rickettsia-harboring *V. carteri* f. *weismannia* strain UTEX 2180 exhibited apparently higher amplification than did the other 12 strains lacking rickettsial endosymbionts, when primer sets targeting rickettsial gene and gene-like sequences were employed (Figure 11E). This indicated that the numbers of rickettsial gene or gene-like sequence molecules targeted by PCR in the rickettsia-harboring strain (*V. carteri* f. *weismannia* strain UTEX 2180) was greater than those of other cells lacking endosymbionts. Thus, high-level detection of *V. carteri* f. *weismannia* strain UTEX 2180 sequences may reflect the presence of many endosymbionts within host algal cells, and lower-level detection the presence of low copy-number sequences within the nuclear genome.

Phylogenetic analysis of rickettsial gene-like sequences from *V. carteri*

In a phylogenetic tree constructed using rickettsial 16S rRNA gene and gene-like sequences, the family *Rickettsiaceae* was divided into two robust monophyletic groups, as shown in Chapter 2 (Figure 4): the first group corresponded to "*Rickettsia*" that included certain endosymbionts of leeches (Kikuchi et al., 2002; Kikuchi & Fukatsu, 2005), and the second the hydra group (Figure 13). The second group was subdivided into two subclades A and B, as described in Chapter 2 (Figure 4) (supported by 76-93% bootstrap values upon ML and MP analysis, and a 1.00 PP by BI). Subclade A contained

bacterial sequences derived from coral, endosymbionts hosted by *Hydra oligactis*, marine green macroalgae (*Bryopsis* spp.), and the freshwater ciliate *Ichthyophthirius multifiliis*. Subclade B included 16S rRNA gene-like sequences from six endosymbiont-lacking strains of *V. carteri* f. *kawasakiensis* and f. *nagariensis*, and those of 16S rRNA genes from endosymbionts of *C. cerasiformis* strain NIES-425, *P. japonica* strain NIES-577, and *V. carteri* f. *weismannia* strain UTEX2180. Within subclade B, endosymbionts from *C. cerasiformis* strain NIES-425, and four ciliates, formed a robust monophyletic group (88-89% bootstrap values and a PP of 0.99). A weak bootstrap value (52%) upon ML analysis suggested that the endosymbiont of *P. japonica* strain NIES-577, and the 16S rRNA gene-like sequences from four strains of *V. carteri* f. *nagariensis,* were positioned into a small monophyletic group that did not include the endosymbiont of *V. carteri* f. *weismannia* strain UTEX 2180 or the gene-like sequence of *V. carteri* f. *kawasakiensis*.

Phylogenetic analyses of *murB* and *ddlB* genes and gene-like sequences yielded essentially the same results (Figures 14, 15). Rickettsial gene-like sequences from endosymbiont-lacking strains of *V. carteri* f. *nagariensis* and f. *weismannia* formed a robust monophyletic group, combined with sequences from three endosymbiont-containing strains (*C. cerasiformis* strain NIES-425, *P. japonica* strain NIES-577, and *V. carteri* f. *weismannia* strain UTEX 2180); the bootstrap values were 99-100% by both ML and MP analysis. In this monophyletic group, *V. carteri* f. *weismannia* strain UTEX 2180 and *C. cerasiformis* strain NIES-425 were (respectively) primarily and secondarily basal to all others members. Combined amino acid data from these two gene/gene-like sequences (Figure 16) showed that rickettsia-lacking strains of *V. carteri* f. *nagariensis* and f. *weismannia,* and the endosymbiont-containing *P. japonica* strain NIES-577, formed a robust small clade (with 97-99% bootstrap values upon ML and MP analyses), which did not include *V. carteri* f. *weismannia* strain UTEX 2180 or *C. cerasiformis* strain NIES-425.

<u>Phylogenetic relationships among various strains of three forms of *Volvox carteri* based on ITS-2 sequences of nuclear rDNA</u>

Coleman (Coleman, 1999) derived phylogenetic relationships among three forms of *V. carteri* based on ITS sequences of nuclear rDNA (ITS-1, the 5.8S rRNA gene, and ITS-2), using only five strains. Thus, I constructed a phylogenetic tree based on nuclear rDNA ITS-2 sequences from 13 strains of three forms of *V. carteri* (Table 1). As reported by Coleman (Coleman, 1999), strains of f. *nagariensis* and f. *kawasakiensis* formed a robust monophyletic group (91% bootstrap values) to which strains of f. *weismannia* were sister (Figure 17). Within f. *weismannia*, two sister clades were well-resolved (with 81-92% bootstrap values); one contained the rickettsia-lacking strains UTEX 1875, UTEX 1876, and UTEX 2170, in which (at least) rickettsial *murB* gene homologs were detected in the present thesis; and the other the rickettsia-lacking strains UTEX 1874 and UTEX 2904, and the rickettsia-containing strain UTEX 2180.

Discussion

The present phylogenetic analyses (Figures 13-16) indicated that the nuclear genomes of nine strains of *V. carteri* include the rickettsial sequences that apparently belong to the hydra group. On the other hand, PCR experiment of these nine strains using the hydra group-specific primers (Chapter 2, Figure 8) demonstrated the absence of the whole 16S rRNA gene that is essential for the living endosymbiotic bacteria. In addition, the DAPI-staining of the algal cells in these nine strains (Chapter 2, Figure 7) clearly shows the absence of bacterial cells within the host algal cells. Thus, the rickettsial gene-like sequences detected in the nine strains of *V. carteri* cannot be actual genes of endosymbionts. It should also be noted that, even if the 16S rRNA gene-specific primers and light microscopic examination by the DAPI-staining could have missed rickettsial endosymbiotic variants that do not belong to the hydra group, the sequences positioned within the possible hydra group (Figures 13-16) should not originate from such missing endosymbionts. In addition, semi-quantitative PCR results of the 12 endosymbiont-lacking strains of *V. carteri* (Figure 11E) indicated that the degree of amplification of their sequences is consistent with that of the low copy DNA such as that in the nuclear genome. Thus, the rickettsial gene-like sequences from the endosymbiont-lacking *V. carteri* strains are considered to be coded in the nuclear genome of the host cells. Although it could also be speculated that the rickettsial gene-like sequences in rickettsial endosymbiont-lacking strains of *V. carteri* might have been transferred to other endosymbionts still remaining in the host, this scenario seems very unlikely because the 6-9 kbp-DNA sequences encompassing both rickettsial and *V. carteri*-specific sequences were experimentally confirmed (Figures 10A, 11A).

The rickettsial 16S rRNA and/or *murB* gene homologs were found in various *V. carteri* strains lacking rickettsial endosymbionts in the cytoplasm (Figures 10, 11A-D). Upon phylogenetic analysis, the nuclear-encoded 16S rRNA gene homologs of *V. carteri* f. *nagariensis* (strains EVE, UTEX 1886, NIES-397, and NIES-398) and f. *kawasakiensis* (strains NIES-732 and NIES-733) belonged to subclade B within the hydra group, as did

the sequences of rickettsial endosymbionts of *V. carteri* f. *weismannia* strain UTEX 2180, *P. japonica* strain NIES-577, and *C. cerasiformis* strain NIES-425 (Figure 13). While rickettsial 16S rRNA gene-like sequences were not evident in three endosymbiont-lacking strains of *V. carteri* f. *weismannia* (UTEX 1875, UTEX 1876, and UTEX 2170), their *murB* and *ddlB* (except for UTEX 2170) homologs formed a monophyletic group in which endosymbiotic genes of *V. carteri* f. *weismannia* strain UTEX 2180 and *C. cerasiformis* strain NIES-425 were basally positioned (Figures 14-16). Therefore, rickettsial gene homologs in the nuclear genome of *V. carteri* may have originated from rickettsial bacteria of subclade B via HGT. As the *Rickettsiales* are obligate intracellular bacteria, and as their hosts (green algal cells) are not phagotrophic, the donor rickettsial organisms may have been harbored by ancestral cells of rickettsial endosymbiont-lacking strains of *V. carteri*. Thus, rickettsial gene-like sequences should have been transferred to the host nuclear genome from the genomic donor rickettsial endosymbionts; subsequently these rickettsial endosymbionts have apparently been lost from the cells of *V. carteri* strains.

Although the statistical support was weak (52% upon ML analysis), the phylogenetic analysis of 16S rRNA gene/gene-like sequences suggested that gene-like sequences in rickettsial endosymbiont-lacking strains of *V. carteri* f. *nagariensis* (EVE, UTEX 1886, NIES-397, and NIES-398) were closely related to that of the endosymbiont *P. japonica* strain NIES-577, rather than those of the endosymbionts of *V. carteri* f. *weismannia* strain UTEX 2180 or *C. cerasiformis* strain NIES-425 (Figure 13). The close relationship between rickettsial genes/gene-like sequences of *V. carteri* f. *nagariensis* and the endosymbiont of *P. japonica* strain NIES-577 was robustly supported by phylogenetic analyses (Figures 14-16) of *murB* and *ddlB* genes/gene-like sequences. In addition, the synteny of *murB* and *ddlB* in *P. japonica* strain NIES-577 was similar to that of *V. carteri* f. *nagariensis* gene-like sequences, lacking a CDS for a hypothetical protein encoded between *murB* and *ddlB* in endosymbionts of *V. carteri* f. *weismannia* strain UTEX 2180 and *C. cerasiformis* strain NIES-425 (Figures 10, 11A). Thus, rickettsial gene-like sequences in the nuclear genome of *V. carteri* f. *nagariensis* may have been

transmitted from an endosymbiont closely related to that of *P. japonica* strain NIES-577. It appears that such sequences were not derived directly from an endosymbiont closely related to that of *V. carteri* f. *weismannia* strain UTEX 2180. A similar EGT event featuring an endosymbiont closely related to that of the *P. japonica* strain NIES-577 may explain the origin of rickettsial gene-like sequences in three endosymbiont-lacking strains of *V. carteri* f. *weismannia* (Figure 11B, C). *MurB* and *ddlB* gene/gene-like sequences from *V. carteri* f. *weismannia* strains, *V. carteri* f. *nagariensis* strains, and *P. japonica* strain NIES-577 formed a small clade distinct from those of *V. carteri* f. *weismannia* strain UTEX 2180 (Figures 14-16). However, the phylogenetic positions of rickettsial 16S rRNA gene-like sequences from *V. carteri* f. *kawasakiensis* strains NIES-732 and NIES-733 remain ambiguous because the sequences are short and no other rickettsial gene-like sequences were detected in these strains (Figure 13).

MurB and *ddlB* gene/gene-like sequences from endosymbiont-lacking strains of *V. carteri* f. *nagariensis* and f. *weismannia* were closely related to those of the endosymbiont (or rickettsial gene-like sequences) of *P. japonica* NIES-577 (Figures 14-16). However, these two forms of *V. carteri* were robustly separated upon phylogenetic analysis (Figure 17). Thus, rickettsial gene-like sequences may have been independently transmitted to ancestors of the two forms of *V. carteri,* from endosymbiont(s) containing a sequence closely related to that of *P. japonica* strain NIES-577. Alternatively, such transmission may have occurred once, in a common ancestor of the three forms of *V. carteri*. A similar EGT event occurring during historical endosymbiosis has been reported in an aphid genome; genes and pseudogenes were transferred from not only the primary endosymbiont *Buchnera aphidicola* (*Gammaproteobacteria*), but also from organisms related to *Wolbachia* spp. (*Anaplasmataceae*, *Rickettsiales*) and *Orientia tsutsugamushi*, neither of which exist in aphid cells (Nikoh & Nakabachi, 2009; Nikoh et al., 2010). Nonetheless, the EGT event that occurred in *V. carteri* remains unique, in that two different but closely related endosymbionts may have invaded closely related hosts. Host specificity is assumed to be not high among endosymbionts within subclade B (Figure 13) of the *Rickettsiaceae.*

The genome of *V. carteri* f. *weismannia* strain UTEX 2180 and of other strains of the Volvocales harboring rickettsial endosymbionts are conceivable to contain rickettsial gene-like EGT sequences aquired from their endosymbionts. Both rickettsial genes and gene-like EGT sequences are potentially amplified by PCR performed on total DNA of such strains. In the present thesis, rickettsial 16S rRNA sequences amplified from *V. carteri* f. *weismannia* strain UTEX 2180 and *P. japonica* strain NIES-577 (Figure 10C, D) came from the rickettsial endosymbionts, as confirmed by FISH (Chapter 2, Figure 5). The other rickettsial gene homologs (e.g., *murB*, *ddlB*) of *V. carteri* f. *weismannia* strain UTEX 2180 (Figure 10C) are also thought to come from endosymbionts, as revealed by semi-quantitative PCR (Figure 11E). However, I could not detect possible nuclear-encoded rickettsial gene-like sequences by direct sequencing of this strain, possibly because I studied only a limited number of genes employing selective primer sets. I speculate that the presence of high numbers of rickettsial endosymbiont cells may inhibit amplification of low copy-number EGT sequences in the nuclear genomes of host cells. It is impossible to distinguish between sequences derived from existing rickettsial endosymbionts and nuclear genomic sequences when the sizes of amplified fragments are near-identical.

Of *V. carteri* strains lacking rickettsial endosymbionts, no amplification of rickettsial gene-like sequences was observed in *V. carteri* f. *nagariensis* strain UTEX 2903 and f. *weismannia* strains UTEX 1874 and UTEX 2904 (Figure 12). Nevertheless, it remains possible that the genomes of these strains harbor other EGT sequences derived from past rickettsial endosymbionts. Whole-genome analyses of these *V. carteri* strains will reveal other EGT sequences.

Because rickettsial gene-like sequences are present in various endosymbiont-lacking strains of *V. carteri* (Chapter 3), and because endosymbiont distribution is sporadic in the phylogenetic trees of *V. carteri* (Chapter 2; see Figure 17) and *Carteria* (Kawafune, 2011), endosymbionts appear to have been frequently lost. Of the lineages of *V. carteri* f. *weismannia*, only UTEX 2180 contains the rickettsial endosymbiont (Figure 17). Nuclear rDNA ITS-2 sequences from UTEX 2180, UTEX 1874,

and UTEX 2904 of *V. carteri* f. *weismannia* were almost identical (only two nucleotide deletions were noted). Thus, transmission of the rickettsial endosymbiont to UTEX 2180 may have been very recent, occurring after divergence of the three strains. Alternatively, transmission might have taken place prior to divergence of the three strains, with subsequent loss of endosymbionts in UTEX 1874 and UTEX 2904. Support for the latter scenario is afforded by the frequent loss of rickettsial endosymbionts in *V. carteri*.

4. General discussion

A role of rickettsial endosymbiont associated with species of Volvocales for ecology and evolution of rickettsiae

Recent studies have expanded the diversity of rickettsiae. Weinert et al. (2009) resolved rickettsial endosymbionts in various non-arthropod hosts, particularly in protists as the hydra group, and later some endosymbionts of another ciliate species were added to this group (Sun et al., 2009; Schrallhammer et al., 2013). In the present thesis, I identified the bacterial endosymbionts of three species of Volvocales, based on analyses of the 16S rRNA gene (Chapter 2). It represents the first report of *Rickettsiaceae*/*Rickettsiales* endosymbionts harbored within photosynthetic eukaryotic cells. Bacterial 16S rRNA gene sequence which newly determined from endosymbiont-containing *V. carteri* strain UTEX 2180 belongs to the subclade B of the hydra group in the *Rickettsiaceae*, as those of *C. cerasiformis* and *P. japonica* (Figure 4). This result was confirmed by FISH that demonstrated the identity of these rickettsial sequences and the bacteria inside the algal host cells (Figure 5). Focusing on subclade B of hydra group (Figure 4), the major hosts of the bacterial endosymbionts are freshwater ciliates including *Euplotes octocarinatus*, *Spirostomum* sp., and *Paramecium caudatum*, which can pray the green microalgae, while there is an exception where the ciliate *Diophrys oligothrix* collected from a brackish environment feeds marine algae (Vannini et al., 2005; Schrallhammer et al., 2013). It is possible that the species of Volvocales mediate horizontal transmission of rickettsial endosymbionts to ciliates via food web interactions in freshwater environment. Moreover, it can imply that other phagotrophic protists and organisms in higher trophic levels that harbor rickettsial endosymbionts would be distributed more widely. Also the rickettsial endosymbionts harbored in the ciliate *Ichthyophthrius multifilliis* and hydras in subclade A (Figure 4) are expected to be associated with species of Volvocales where rickettsial endosymbionts are currently undiscovered.

Possible role of rickettsial genes in *V. carteri* genomes

I also demonstrated that the possible EGT-derived rickettsial gene-like sequences are present on the genomes of nine *V. carteri* strains (Chapter 3), while these strains were observed to lack rickettsial endosymbionts (Chapter 2). Among the rickettsial genes, *murC*, *murB*, *ddlB,* and *ftsQ* are component of the dcw (division and cell wall) cluster, which includes genes involved in the peptidoglycan synthesis and cell division (Mingorance et al., 2004). Some genes of dcw cluster have previously been shown to transfer from endosymbiotic bacteria to host eukaryote genomes; e.g., that of an adzuki bean beetle (a gene encoding cell division protein FtsZ [*ftsZ*]) (Kondo et al., 2002); that of a rotifer (*ddl*) (Gladyshev et al., 2008); and that of *Trichoplax adherens* (UDP-N-acetylglucosamine enoylpyruvyl transferase =*murA*) (Driscoll et al., 2013). In the *T. adherens* model, it was suggested that the transferred *murA* gene is expressed, and limits the growth of endosymbionts of the family *Midichloriaceae* (*Rickettsiales*) (Driscoll et al., 2013). In addition, the genome of the mealybug *Planococcs citri* contains several dcw genes acquired via HGT from bacteria including *Rickettsia*. This mealybug was considered to be capable of controlling a cell wall of its bacterial endosymbiont *Candidatus* Moranella endobia (*Enterobacteriales*, *Gammaproteobacteria*) by expressing and regulating horizontally transferred dcw genes (Husnik et al., 2013). Thus, host-encoded dcw genes seem to have a potential for controlling growth of intracellular bacteria.

In this thesis, however, rickettsial dcw gene-like sequences encoded on the nuclear genome are appeared to be non-functional at least in cells of *V. carteri* f. *nagariensis* strains, because most of these sequences were incomplete when compared with the CDS of rickettsial genes of *C. cerasiformis* strain NIES-425 and *V. carteri* f. *weismannia* strain UTEX 2180 endosymbionts (Figures 10, 11A). Nuclear genomes of two *V. carteri* f. *weismannia* strains UTEX 1875 and UTEX 1876 contain intact CDSs of the dcw genes (*murB*, *ddlB*, and *ftsQ*; Figure 11B), but yet it is difficult to determined whether these CDSs obtain capabilities to be expressed functionally through eukaryotic system because the presence of eukaryotic-specific promoters and regulatory elements is unclear.

In addition, rickettsial 16S rRNA gene-like sequences in the nuclear genomes of *V. carteri* f. *nagariensis* strains lack sequences corresponding to the 3′ ends (Figures 10A, 11A), and should be non-functional. Similarly, the genome of *Trichonympha agilis*, an eukaryotic symbiont in the termite gut, contains 16S rRNA pseudogenes that were likely transferred from bacterial endosymbionts, but have large deletions. It remains unknown whether such pseudogenes have a certain function as non-cording DNA or not (Sato et al., 2014). Examining expression of rickettsial gene-like sequences in the nuclear genome of *V. carteri* strains will reveal their function.

Conclusion

This thesis demonstrates the presence of rickettsial endosymbionts inside cells of three species of Volvocales, especially multiple trajectories of the rickettsial endosymbiosis occurring in *V. carteri*, where one of endosymbionts has continued its association with *V. carteri* f. *weismannia* strain UTEX 2180 and the others have been lost from the cells of *V. carteri*. It implies the dynamic nature of the rickettsiae that frequently invade and disappear from the cells of various species in Volvocales. Moreover, it is conceivable that the rickettsial endosymbiosis could have occurred more widely in the species in Volvocales than previously believed. In this thesis, most of the rickettsial gene-like sequences obtained by the nuclear genome of *V. carteri* via the past rickettsial endosymbiosis are considered to be non-functional, but it still remains possible that such a transferred sequence obtained from the "past" rickettsial endosymbiosis could have contributed to the host evolution, as shown in the relationship between plant cells and the *Chlamydiales* (Ball et al., 2013). To resolve the past endosymbiosis, it requires comparing genomes of host and endosymbiont candidates, whereas available genome sequences in Volvocales are limited to those of *V. carteri* and *Chlamydomonas reinhardtii*. In further studies, increasing available genome data of Volvocales will reveal more elaborate, complex evolutionary relationships between them and prokaryotes including those of the *Rickettsiaceae*/*Rickettsiales*.

Addendum

After the publication of Kawafune et al. (2012, PLOS ONE 7: e31749) reporting a part of results in Chapter 2 (the molecular identification of the rickettsial endosymbionts of *C. cerasiformis* and *P. japonica*), Hollants et al. (2013, Proc R Soc Lond B Biol Sci 280: 20122659) reported intracellular rickettsiae from the marine green macroalgae *Bryopsis* spp. They resolved that the endosymbionts of *Bryopsis* spp., *Hydra oligactis*, and the freshwater ciliate *Ichthyophthirius multifiliis* belong to subclade A of the hydra group.

Acknowledgments

First of all, I would like to express my sincere appreciation to Assoc. Prof. Hisayoshi Nozaki (University of Tokyo) for his valuable suggestions, fruitful discussions and great encouragement throughout this thesis and for the critical proofreading of this manuscript.

I am also extremely grateful to Dr. Yuichi Hongo (Tokyo Institute of Technology) for his technical assistance with molecular identification of bacteria, insightful discussions, and encouragement.

I profoundly thank Dr. Takashi Hamaji (Donald Danforth Plant Science Center) for his supports and helpful comments especially about genome evolution; Drs. Tetsuya Kurata and Tomoaki Sakamoto (Nara Institute of Science and Technology) for their technical assistance and advice in sequencing; Drs. Shin-ya Miyagishima and Shunsuke Hirooka (National Institute of Genetics) for their technical assistance in cell organelle fraction, and for their encouragement and helpful discussion.

I also wish to thank Dr. Akira Yokota (University of Indonesia) for helpful comments for bacterial identification, and Dr. Chikako Shingyoji (University of Tokyo) for her encouragement and helpful comments. I am also grateful to all the current and former members of the Laboratory of ORIGIN OF EUKARYOTE BIODIVERSITY (Department of Biological Sciences, Graduate School of Science, University of Tokyo) for their hearty support and helpful advice.

A part of Chapter 2 was reproduced from Kawafune et al. (2014, Phycologia 53: 95-99.)

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.

Amann RI, Krumholz L, Srahl DA (1990) Fluorescent oligonucleotide probing of whole cells for determinative, phylogenetic and environmental studies in microbiology. J Bacteriol 172: 762–770.

Andersson S, Zomorodipour A (1998) The genome sequence of Rickettsia prowazekii and the origin of mitochondria. Nature 296: 133–143.

Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, Weightman AJ (2006) New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. Appl Environ Microbiol 72: 5734-5741.

Ball SG, Subtil A, Bhattacharya D, Moustafa A, Weber AP, Gehre L, Colleoni C, Arias MC, Cenci U, Dauvillee D (2013) Metabolic effectors secreted by bacterial pathogens: essential facilitators of plastid endosymbiosis? Plant Cell 25: 7–21.

Balraj P, Renesto P, Raoult D (2009) Advances in *Rickettsia* pathogenicity. Ann NY Acad Sci 1166: 94–105.

Becker B (2007) Function and evolution of the vacuolar compartment in green algae and land plants (Viridiplantae). Int Rev Cytol 264: 1–24.

Bennett GM, Moran N (2013) Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a Phloem-feeding insect. Genome Biol Evol 5: 1675–1688.

Caisová L, Marin B, Melkonian M (2013) A consensus secondary structure of ITS2 in the Chlorophyta identified by phylogenetic reconstruction. Protist 164: 482-496.

Clay K (2014) Defensive symbiosis: a microbial perspective. Functional Ecol 28: 293–298.

Coleman AW (1994) Molecular delination of species and syngens in Volvocacean green algae. J Phycol 30: 80–90.

Coleman AW (1999) Phylogenetic analysis of "Volvocacae" for comparative genetic

studies. Proc Natl Acad Sci USA 96: 13892-13897.

Coleman AW, Mai JC (1997) Ribosomal DNA and ITS-2 sequence comparisons as a tool for predicting genetic relatedness. J Mol Evol 45: 168-177.

Daims H, Bruhl A Amann R, Schleifer KH, Wagner M (1999) The domain- specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. Syst Appl Microbiol 22: 434–444.

Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27: 1164-1165.

Darriba D, Taboada GL, Doallo R, Posada D. (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9: 772-772.

Darty K, Denise A, Ponty Y (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics 25: 1974-1975.

Driscoll T, Gillespie JJ, Nordberg EK, Azad AF, Sobral BW (2013) Bacterial DNA sifted from the *Trichoplax adhaerens* (Animalia: Placozoa) genome project reveals a putative rickettsial endosymbiont. Genome Biol Evol 5: 621-645.

Dumler JS, Walker DH (2005) Order II. Rickettsiales Gieszczykiewicz 1939, 25AL emend. Dumler, Barbet, Bekker, Dasch, Palmer, Ray, Rikihisa and Rurangirwa 2001, 2156. In: Garrity GM, Boone DR, Castenholz RW, editors. Bergey's manual of systematic bacteriology. Volume 2. 2nd edition. East Lansing, MI. Springer.

Dyall SD, Brown MT, Johnson PJ (2004) Ancient invasions: from endosymbionts to organelles. Science 304: 253–257.

Dykova I, Ververkova M, Fiala I, Machackova B, Peckova H (2003) *Nuclearia pattersoni* sp. n. (Filosea), a new species of amphizoic amoeba isolated from gills of roach (Rutilus rutilus), and its rickettsial endosymbiont. Folia Parasitologica 50: 161–170.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27: 2194-2200.

Engelstädter J, Hurst GDD (2009) The ecology and evolution of microbes that manipulate host reproduction. Annu Rev Ecol Evol Syst 40: 127-149.

Euzéby J (2006) List of new names and new combinations previously effectively, but not

validly, published. Int J Syst Evol Microbiol 56: 2507–2508.

Fawley MW, Fawley KP (2004) A simple and rapid technique for the isolation of DNA from microalgae. J Phycol 40: 223–225.

Ferla MP, Thrash JC, Giovannoni SJ, Patrick WM (2013) New rRNA gene-based phylogenies of the *Alphaproteobacteria* provide perspective on major groups, mitochondrial ancestry and phylogenetic instability. PLOS ONE 8: e83383.

Ferrantini F, Fokin SI, Modeo L, Andreoli I, Dini F, et al. (2009) "*Candidatus* Cryptoprodotis polytropus," a novel *Rickettsia*-like organism in the Ciliated protist *Pseudomicrothorax dubius* (Chiliophora, Nassophorea). J Eukaryot Microbiol 56: 119–129.

Ferris P, Olson BJ, De Hoff PL, Douglass S, Casero D, et al. (2010) Evolution of an expanded sex-determining locus in *Volvox*. Science 328: 351-354.

Fitzpatrick DA, Creevey CJ, Mcinerney JO (2006) Genome phylogenies indicate a meaningful alpha-proteobacterial phylogeny and support a grouping of the mitochondria with the Rickettsiales. Mol Biol Evol 23: 74–85.

Fraune S, Bosch TCG (2007) Long-term maintenance of species-specific bacterial microbiota in the basal metazoan *Hydra*. Proc Natl Acad Sci USA 104: 13146–13151.

Fuchs BM, Glochner FO, Wulf J, Amann R (2000) Unlabeled helper oligonucleotides increase the in situ accessibility to 16 S rRNA of fluorescently labeled oligonucleotide probes. Appl Environ Microbiol 66: 3603–3607.

Gladyshev EA, Meselson M, Arkhipova IR (2008) Massive horizontal gene transfer in *Bdelloid rotifers*. Science 320: 1210–1213.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, et al. (2012) Phytozome: a comparative platform for green plant genomics. Ncl Acids Res 40: D1178-D1186.

Gray MW (2014) The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria. Cold Spring Harb Perspect Biol 6: a016097–a016097.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New

algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307–321.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704.

Harris EH (1989) The Chlamydomonas Source Book. Oxford, Burlington, San Diego.

Herron MD, Hackett JD, Aylward FO, Michod RE (2009) Triassic origin and early radiation of multicellular volvocine algae. Proc Natl Acad Sci USA 106: 3254-3258.

Hine PM, Wakefield S, Diggles BK, Webb VL, Maas EW (2002) Ultrastructure of a haplosporidian containing Rickettsiae, associated with mortalities among cultured paua *Haliotis iris*. Dis Aquat Org 49: 207–219.

Hiraide R, Kawai-Toyooka H, Hamaji T, Matsuzaki R, Kawafune K et al. (2013) The evolution of male–female sexual dimorphism predates the gender-based divergence of the mating locus gene *MAT3/RB*. Mol Biol Evol 30: 1038-1040.

Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, et al. (2013) Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. Cell 153: 1567–1578.

Ichimura T, Watanabe MM (1977) An axenic clone of *Microcystis aeruginosa* Kütz. Emend. Elenkin from Lake Kasumigaura. Bull Jpn Soc Phycol 25: 177–181.

Isaka N, Kawai-Toyooka H, Matsuzaki R, Nakada T, Nozaki H (2012) Description of two new monoecious species of *Volvox* sect. *Volvox* (Volvocaceae, Chlorophyceae), based on comparative morphology and molecular phylogeny of cultured material. J Phycol 48: 759-767.

Kaminuma E, Mashima J, Kodama Y, Gojobori T, Ogasawara O, et al. (2010) DDBJ launches a new archive database with analytical tools for next-generation sequence data. Nucleic Acids Res 38: D33-D38

Kasai F, Kawachi M, Erata M, Mori F, Yumoto K, et al. (2009) NIES-Collection. List of Strains. 8th Edition. Jpn J Phycol (Sôrui) 57: suppl.1–350, pls1–7.

Kato S (1982) Laboratory culture and morphology of *Colacium vesiculosum* Ehrb. (Euglenophyceae). Jpn J Phycol (Sôrui) 30: 63–67 (in Japanese with English

abstract).

Kawafune K (2011) An evolutional study of bacterial endosymbionts in volvocalean green algae (Master's thesis). The University of Tokyo, Tokyo, Japan.

Kawafune K, Sato M, Toyooka K, Nozaki H (2013) Ultrastructure of the rickettsial endosymbiont "*MIDORIKO*" in the green alga *Carteria cerasiformis* as revealed by high-pressure freezing and freeze-substitution fixation. Protoplasma 250: 949-953.

Kikuchi Y, Fukatsu T (2005) *Rickettsia* infection in natural leech populations. Microb Ecol 49: 265–271.

Kikuchi Y, Sameshima S, Kitade O, Kojima J, Fukatsu T (2002) Novel clade of *Rickettsia* sp. from leeches. Appl Environm Microbiol 68: 999–1004.

Kim E, Park JS, Simpson AG, Matsunaga S, Watanabe M, et al. (2010) Complex array of endobionts in *Petalomonas sphagnophila*, a large heterotrophic euglenid protist from Sphagnum-dominated peatlands. The ISME Journal 4: 1108–1120.

Kleine T, Maier UG, Leister D (2009) DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. Ann Rev Plant Biol 60: 115–138.

Ko WY, David RM, Akashi H (2003) Molecular phylogeny of the *Drosophila melanogaster* species subgroup. J Mol Evol 57: 562-573.

Kobl I, Kirk DL, Schmitt R. (1998) Quantitative PCR data falsify the chromosomal endoreduplication hypothesis for *Volvox carteri* (Volvocales, Chlorophyta). J Phycol 34: 981-988.

Kochert G, Olson LW (1970) Endosymbiotic bacteria in Volvox carteri. Trans Am Microsc Soc 89: 475–478.

Kondo N, Nikoh N, Ijichi N, Shimada M, Fukatsu T (2002) Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect. Proc Natl Acad Sci USA 99: 14280-14285.

Lee W-SB, Kochert G (1976) Bacterial endosymbionts in *Volvox carteri* (Chlorophyceae). J Phycol 12: 194-197.

Loy A, Arnold R, Tischler P, Rattei T, Wagner M, et al. (2008) probeCheck - a central resource for evaluating oligonucleotide probe coverage and specificity. Environ

Microbiol 10: 2894–2898.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, et al. (2004) ARB: a software enviroment for sequence data. Ncl Acids Res 32: 1363–1371.

Margulis L (1970) Origin of eukaryotic cells: evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the Precambrian earth. New Haven: Yale University Press.

Markham NR, Zuker M (2005) DINAMelt web server for nucleic acid melting prediction. Ncl Acids Res 33: W577–W581.

McCutcheon JP, Keeling PJ (2014) Endosymbiosis: Protein Targeting Further Erodes the Organelle/Symbiont Distinction. Curr Biol 24: R654–R655.

McCutcheon J, Moran N (2011) Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol 10: 13–26.

Merhej V, Raoult D (2011) Rickettsial evolution in the light of comparative genomics. Biol Rev 86: 379–405.

Miller SM, Schmitt R, Kirk DL (1993) *Jordan*, an active Volvox transposable element similar to higher plant transposons. Plant Cell 5: 1125-1138.

Mingorance J, Tamames J, Vicente M (2004) Genomic channeling in bacterial cell division. J Mol Recognit 17: 481-487.

Montagna M, Sassera D, Epis S, Bazzocchi C, Vannini C, et al. (2013) "*Candidatus* Midichloriaceae" fam. nov. (*Rickettsiales*), an ecologically widespread clade of intracellular alphaproteobacteria. Appl Environ Microbiol 79: 3241–3248.

Nakabachi A, Ishida K, Hongoh Y, Ohkuma M, Miyagishima S (2014) Aphid gene of bacterial origin encodes a protein transported to an obligate endosymbiont. Curr Biol 24: R640-641.

Nakada T, Nozaki H (2007) Re-evaluation of three *Chlorogonium* (Volvocales, Chlorophyceae) species based on 18 S ribosomal RNA gene phylogeny. Eur J Phycol 42: 177–182.

Nakayama K, Kurokawa K, Fukuhara M, Urakami H, Yamamoto S, et al. (2010) Genome comparison and phylogenetic analysis of *Orientia tsutsugamushi* strains.

DNA Res 17: 281–291.

Nakazawa A, Nozaki H (2004) Phylogenetic analysis of the Tetrasporalean genus
Asterococcus (Chlorophyceae) based on 18S ribosomal RNA gene sequences. J Jpn
Bot 79: 255–261.

Nikoh N, McCutcheon JP, Kudo T, Miyagishima SY, Moran NA, Nakabachi A (2010)
Bacterial genes in the aphid genome: absence of functional gene transfer from
Buchnera to its host. PLOS Genet 6: e1000827.

Nikoh N, Nakabachi A (2009) Aphids acquired symbiotic genes via lateral gene transfer.
BMC Biol 7: 12.

Noda S, Ohkuma M, Yamada A, Hongoh Y, Kudo T (2003) Phylogenetic position and in
situ identification of ectosymbiotic spirochetes on protists in the termite gut. Appl
Environ Microbiol 69: 625–633.

Nozaki H (2003) Flagellated green algae. In Freshwater Algae of North America. In:
Wehr JD, Shearh RG, eds. Academic Press: San Diego, California. pp 225–252.

Nozaki H, Aizawa K, Watanabe MM (1994) A taxonomic study of four species of
Carteria (Volvocales, Chlorophyta) with cruciate anterior papillae, based
on cultured material. Phycologia 33: 239–247.

Nozaki H, Kuroiwa H, Mita T, Kuroiwa T (1989) *Pleodorina japonica* sp. Nov. (Volvocales,
Chlorophyta) with bacteria-like endosymbionts. Phycologia 28: 252–267.

Nozaki H, Kuroiwa T (1992) Ultrastructure of the extracellular matrix and taxonomy of
Eudorina, *Pleodorina* and *Yamagishiella* gen. nov. (Volvocaceae, Chlorophyta).
Phycologia 31: 529–541.

Nozaki H, Yamada TK, Takahashi F, Matsuzaki R, Nakada T (2014) New "missing link"
genus of the colonial volvocine green algae gives insights into the evolution of
oogamy. BMC Evol Biol 14: 37.

Nosek J, Tomáska L (2003) Mitochondrial genome diversity: evolution of the molecular
architecture and replication strategy. Curr Genet 44: 73–84.

Nowack ECM, Grossman AR (2012) Trafficking of protein into the recently established
photosynthetic organelles of *Paulinella chromatophora*. Proc Nat Acad Sci USA 109:

5340–5345.

Perlman SJ, Hunter MS, Zchori-Fein E (2006) The emerging diversity of *Rickettsia*. Proc R Soc B 273: 2097–2106.

Prochnik SE, Umen J, Nedelcu AM, Hallmann A, Miller SM, et al. (2010) Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. Science 329: 223–226.

Pruesse EC, Quast K, Knittel B, Fuchs W, Ludwig J, et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nuc Acid Res 35: 7188–7196.

Reyes-Prieto A, Weber APM, Bhattacharya D (2007) The origin and establishment of the plastid in algae and plants. Ann Rev Genet 41: 147–168.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61: 539-542.

Sagan L (1967). On the origin of mitosing cells. J Theoretic Biol 14: 255–74.

Sato T, Kuwahara H, Fujita K, Noda S, Kihara K, Yamada A, et al. (2014) Intranuclear verrucomicrobial symbionts and evidence of lateral gene transfer to the host protist in the termite gut. ISME J 8: 1008-1019.

Schrallhammer M, Ferrantini F, Vannini C, Galati S, Schweikert M, et al. (2013) '*Candidatus* Megaira polyxenophila'gen. nov., sp. nov.: considerations on evolutionary history, host range and shift of early divergent rickettsiae. PLOS ONE 8: e72581.

Setohigashi Y, Hamaji T, Hayama M, Matsuzaki R, Nozaki H (2011) Uniparental inheritance of chloroplast DNA is strict in the isogamous Volvocalean *Gonium*. PLOS ONE 6: e19545.

Sugawara H, Ohyama A, Mori H, Kurokawa K (2009) Microbial genome annotation pipeline (MiGAP) for diverse users. In 20th International Conference on Genome Informatics (pp. 1-2).

Sun HY, Noe J, Barber Coyne RS, Cassidy-Hanley D, Clark TG, et al. (2009)

Endosymbiotic bacteria in the parasitic ciliate *Ichthyophthirius multifiliis*. Appl Environ Microbiol 75: 7445–7452.

Swofford DL (2003) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. Sunderland, U.S.: Sinauer Associates.

Takahara M, Takahashi H, Matsunaga S, Sakai A, Kawano S, Kuroiwa T (1999) Two types of *ftsZ* genes isolated from the unicellular primitive red alga *Galdieria sulphuraria*. Plant Cell Physiol 40: 784-791.

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5: 123–135.

Vannini C, Boscaro V, Ferrantini F, Benken KA, Mironov TI, et al. (2014) Flagellar movement in two bacteria of the family *Rickettsiaceae*: a re-evaluation of motility in an evolutionary perspective. PLOS ONE 9: e87718.

Vannini C, Petroni G, Verni F, Rosati G (2005) A bacterium belonging to the *Rickettsiaceae* Family inhabits the cytoplasm of the marine ciliate *Diophrys appendiculata* (Ciliophora, Hypotrichia). Microbial Ecol 49: 434–442.

Weinert LA, Werren JH, Aebi A, Stone GN, Jiggins FM (2009) Evolution and diversity of *Rickettsia* bacteria. BMC Biol 7: 6–20.

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetics study. J Bacteriol 173: 697–703.

Wernegreen JJ (2012) Endosymbiosis. Curr Biol 22: R555–61.

Yu XJ, Walker DH (2012) Rickettsia and Rickettsial diseases. In: Morse SA, editor. Bioterrorism. InTech. pp. 179-192.

Zerbino DR, Ewan B (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18: 821-829.

Zimorski V, Ku C, Martin WF, Gould SB (2014) Endosymbiotic theory for organelle origins. Curr Opin Microbiol 22: 38–48.

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucl Acids Res 31: 3406-3415.

Tables

Table 1. List of *Volvox carteri* and other algal strains used in this thesis.

DDBJ/EMBL/GenBank accession numbers of nuclear ribosomal DNA internal transcribed spacer regions and rickettsial gene-like sequences determined in this thesis are also shown.

Taxon	Strain designation	Accession No. ITS sequence	Rickettsial gene/gene-like sequences From rickettsiae-lacking strains	From rickettsiae-containing strains
Volvox carteri f. *nagariensis*	EVE	LC004701[a]	LC004713[a]	-
	UTEX[b] 1886	LC004702[a]	LC004714[a]	-
	NIES[c]-397	LC004703[a]	LC004715[a]	-
	NIES-398	LC004704[a]	LC004716[a]	-
	UTEX 2903	LC004705[a]	-	-
Volvox carteri f. *kawasakiensis*	NIES-732	LC004706[a]	LC004717[a]	-
	NIES-733	LC004707[a]	LC004718[a]	-
Volvox carteri f. *weismannia*	UTEX 1874	LC004708[a]	-	-
	NIES-866 (=UTEX 1875)	AB771953	LC004719[a]	-
	UTEX 1876	AB771954	LC004720[a]	-
	UTEX 2170	LC004709[a]	LC004721[a]	-
	UTEX 2904	LC004710[a]	-	-
	UTEX 2180	LC004711[a]	-	AB861537[a] (16S rRNA), LC004722[a] (other genes)
Volvox obversus	UTEX 1865	LC004712[a]	-	-
Pleodorina japonica	NIES-577	-	-	AB688629 (16S rRNA), LC004723[a] (other genes)
Carteria cerasiformis	NIES-424	-	-	-

	NIES-425	-	-	LC004724[a] (including 16S rRNA); LC004725[a] (including *murB*, *ddlB*, etc.)
Chlamydomonas reinhardtii	CC[d]-503 (*cw92* mt+)	-	-	-

[a]Sequence determined in this thesis.

[b]Microbial Culture Collection at the National Institute for Environmental Studies, Japan (Kasai et al., 2009).

[c]Culture Collection of Algae at the University of Texas at Austin (Austin, TX, USA; http://www.utex.org/).

[d]Chlamydomonas Resource Center (St. Paul, MN, USA; http://chlamycollection.org/).

Table 2. The list of primers used in genomic PCR, semi-quantitative genomic PCR and sequencing of this thesis.

Primer designation	Primer sequence (5′ to 3′)	Target	Forward/ Reverse
9F[a]	GAGTTTGATCCTGGCTCAG	Universal 16S rRNA gene	F
1492R[a]	GCTTACCTTGTTACGACTT	Universal 16S rRNA gene	R
enRB[b]	TAGCTCACCAGCTTCGGGTAAA	Rickettsial 16S rRNA gene	R
enFE[b]	CGCGTAGGCGGATTAGTAAGTTGG	Rickettsial 16S rRNA gene	F
enRG[b]	CGTCTTGCTTCCCTCTGTAAGT	Rickettsial 16S rRNA gene	R
16S-mkFJ	TGACATGGTGATCGTAGGTTACAG	Rickettsial 16S rRNA gene	F
16S-enFK	ACTTACAGAGGGAAGCAAGACG	Rickettsial 16S rRNA gene	F
16S-enRL	CACTAAAATTGGAGCAAGCCCC	Rickettsial 16S rRNA gene	R
16S-enRM	AGCGATAAATCTTTCCTCCTTGG	Rickettsial 16S rRNA gene	R
16S-enFN	GCTTAACCTCGGAACTGCTTTC	Rickettsial 16S rRNA gene	F
18S-FA[c]	AACCTGGTTGATCCTGCCAGT	Chlorophytes 18S rRNA gene	F
18S-RD[c]	GCTGGCACCAGACTTGCCCTC	Chlorophytes 18S rRNA gene	R
ONact1[d]	GTGTGCGACAACGGCTCGGGC	*Volvox carteri* actin gene	F
ONact2[d]	GCCTCCGTGAGAAGGACTGGGT	*V. carteri* actin gene	R
ONact1_CR	GTGTGCGACAATGGTTCGGGC	*Chlamydomonas reinhardtii* actin gene *IDA5*	F
CR_IDA5_R3	GTGCTCAATGGGGTAGCGCAGA	*C. reinhardtii* actin gene *IDA5*	R
eveFA	ATCGACAATCGTTGAGGTCCTT	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
eveFB	GTCCTTGCAGTCAAGAATAAGTCC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
eveFC	GGGCTTGCTCCAATTTTAGTGA	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsial 16S rRNA gene	F
eveRD	ATTCACGCGGTATTGCTGGA	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsial 16S rRNA gene	R
eveFE	AGTAGGAGGAGGTGTAGCGATG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *murB* gene	F
eveRG	TGGTTGCGAAGCACTCCTAATA	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *murB* gene	R
eveRH	CGCCTAAATTTTCCATATCAGC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *murB* gene	R

eveRH2	ATTTTCCATATCAGCAGCAGTAGC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *murB* gene	R
eveRJ	CACCTAACTTGCTGAACGAATC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *ddlB* gene	R
eveFZ	CAACGGACAGGGAGACGATTAC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
eveRY	CGCTTCTGCAAGCTATTTATATGG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
eveFX	TTTCGCGGATAGGAGGAATCTG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
eveRL	AGGCGGCTAGTCTTTTGTTG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
eveFK	CAGGCCCAGAAATGATGATAGC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
eveFM	TGGGAGCCGATATTGCTGTACT	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *ddlB* gene	F
eveRN	TGCTCGGCAAAACCTTCAGTA	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *ddlB* gene	R
eveRP	GGTAATGCTTGTGGAGCGAAG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
eveFQ	CCAGAATCGCCAAGTCTTTTCA	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
eveRS	TGGCAAAGCAGTTACGAAGG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
eveRT	CGAGGTCTTTTGACCGGTTG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
eveFU	GGAGGTGTAGCGATGAATGC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *murB* gene	F
eveFV	ATGTCTGCAGAACGCGAAG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *ddlB* gene	F
eveRW	GCAATTTCTGGGCAAATAGATAA	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome/ rickettsia *ddlB* gene	R
eve-Jd-F1	GCCCATTGTCACAAGGTTCATGCC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
asmb81_F4	AAGAGGCGTGTACGGACGAGAATA	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
asmb81_F5	TACGGCGGAAAGTCAAAAACAGAT	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
asmb82_R1	TTCGATTTACCGAGCTCCAATGTT	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
asmb82_R2	TAATTTCACCCCGGACCTTAAACC	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
asmb81_R6	ATAGTGTTGTTGGCTGACGGATTG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
asmb82_F3	CATCGTACCAACGGTACATCCAAG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
asmb82_F4	GCAGGGTGAGTTGCACGAAGA	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F

asmb82_F5	CAGGAACTACTCTCGCTCGGTCAG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	F
asmb82_R6	GTATATTGGGTGCAGTCGGAGGAG	Scaffold 6 of *V. carteri* f. *nagariensis* strain EVE genome	R
ITS-Fa[e]	GGGATCCGTTTCCGTAGGTGAACCTGC	Internal transcribed spacer (ITS) region of nuclear rDNA	F
ITS-Rb[e]	GGGATCCATATGCTTAAGTTCAGCGGGT	ITS region of nuclear rDNA	R
ITS-Fc[e]	GCATCGATGAAGAACGCAGC	ITS region of nuclear rDNA	F
ITS-Rd[e]	GCTGCGTTCTTCAGCGAT	ITS region of nuclear rDNA	R
RcBl-murB-FA	ATTGGCATTGGCGGCGTTGG	Rickettsial *murB*	F
U2180-murB-FC	TCTGAATGGGCTACCAAAGAGG	Rickettsial *murB*	F
MkCc-murB-FE	TTTTCAGGGTATCCGAGGGGTA	Rickettsial *murB*	F
MkCc-murB-FG	TACCGTTATTGGTGCTGGCTCA	Rickettsial *murB*	F
MkCc-murB-FK	TCAGAAATGCACTGTAACTTCA	Rickettsial *murB*	F
RcMk-murB-FP	TATTCCGGGRACAGTAGGAGGAGG	Rickettsial *murB*	F
RcMk-murB-RK2	AGTGCATTTCTGACATACATGC	Rickettsial *murB*	R
RcMk-murB-FQ	AATCTGAARCACCTAACYTGGCTTAAAG	Rickettsial *murB*	F
RcMk-murB-FK3	CAGAAATGCACTGTAACTTCATGATTAAT	Rickettsial *murB*	F
MkPj-murB-RG	TGAACCAGCACCGATAACGGTA	Rickettsial *murB*	R
RcBl-ddlB-RB	GGGAAAATTCTTTTTGCTGTTTTA	Rickettsial *ddlB*	R
U2180-ddlB-RD	ATACCGCCGCACCTAAAGCA	Rickettsial *ddlB*	R
MkCc-ddlB-RH	CATGCCGGTAACACAATCAAGG	Rickettsial *ddlB*	R
MkCc-ddlB-RJ	TGGTGTCATACCAGGGTGAGTGTT	Rickettsial *ddlB*	R
RcMk-ddlB-R01	GAAACTTCCCGTTCAGCAGACA	Rickettsial *ddlB*	R
RcMk-ddlB-R02	GCAATATCTGCTCCCATATC	Rickettsial *ddlB*	R
MkCc-ddlB-RL	GACTTCCTTTTGTTTAAAGCAA	Rickettsial *ddlB*	R
MkCc-ddlB-RM	TCTTTGGTATAACCATCTGATTT	Rickettsial *ddlB*	R
RcMk-ddlB-FQ	YGGAGAAGACGGTTGCCT	Rickettsial *ddlB*	F
RcMk-ddlB-RH2	CATRCCRGTAACACAATCAAGG	Rickettsial *ddlB*	R
RcMk-ddlB-RN	CCGAAGCYAAAACCCCAGGACC	Rickettsial *ddlB*	R
MkPj-ddlB-RP	TTTTTCYTTATATACAGTTTTAAAATTTCCGTTAG	Rickettsial *ddlB*	R
RcMk-ddlB-FH3	CCTTGATTGTGTTACYGGYATGATAAG	Rickettsial *ddlB*	F
MkPj-ddlB-FJ2	AGYTAATACTCATCCCGGTATGACAC	Rickettsial *ddlB*	F

56

RcMk-murC-FA	GGTTATTGCTATTTTCCAGCCKCA	Rickettsial *murC*	F
RcMk-murC-FB	TTATGATGGGTGCTGGTARTATTT	Rickettsial *murC*	F
RcMk-murC-FC	GATYTGATAGTTATGATGGGWGCWGG	Rickettsial *murC*	F
RcMk-murC-FD	CTTAATTCTCTTGCGGCCATTG	Rickettsial *murC*	F
RcMk-murC-R05	TTTATCRGGTAATTGCCARGCCCA	Rickettsial *murC*	R
RcMk-murC-R6	TGGYCTTGCACCTAATTTTAYCATTTC	Rickettsial *murC*	R
RcMk-murC-R7	TACYGCAATAGTAGADGGRATTTTGA	Rickettsial *murC*	R
RcMk-murC-F08	GTTAAGAGRCGTTTTACCAAAGTTTGT	Rickettsial *murC*	F
RcMd-murC-F09	GTTAMAGAACGCCGGATAATTACTTACG	Rickettsial *murC*	F
RcMk-murC-F10	GGYCTTGCYCCTACTGTTATAAACG	Rickettsial *murC*	F
RcMk-murC-R10	CGTTTATAACAGTAGGRGCAAGRCC	Rickettsial *murC*	R
MkPj-murC-F11	TTGGATATTTATGGGGCAGGGGAAA	Rickettsial *murC*	F
RcMk-ftsQ-R01	ATTGGCGTTCGCTCAGTAATCG	Rickettsial *ftsQ*	R
RcMk-ftsQ-R02	GCATCCTGACCCACAACGTGTA	Rickettsial *ftsQ*	R
RcMk-ftsQ-F03	TGAAGAAGGTAATCGAATCAGCTMA	Rickettsial *ftsQ*	F
MkCc-ccmF-R01	CCTCCCCAGCCAAGTTCCCTAT	Rickettsial *ccmF*	F
MkCc-ccmF-R02	CAATAGCCTCCCCAGCCAAGTTCC	Rickettsial *ccmF*	F
MkCc-phbB-F01	CAAGCCGGACAAATAGGACAGACT	Rickettsial *phbB*	R
MkCc-phbB-F02	AAGACTTGGCCGCCCTGAAGAA	Rickettsial *phbB*	R
MkCc-phbB-F03	GCCGCCCTGAAGAAGTAGCAAG	Rickettsial *phbB*	R
PjMk-phbB-TF1	AATGCTGCCTATTGCTTCAGGAAA	Rickettsial *phbB*	F
RcMk-phbB-F4	CCAAAATGACTGATGATTAGCRAGAAC	Rickettsial *plsC*	R
RcMk-phbB-R5	ATGTTCCCAGAAGGYGGYACRG	Rickettsial *plsC*	F

[a]Weisburg et al., 1991.

[b]Kawafune, 2011.

[c]Nakazawa & Nozaki, 2004.

[d]Kobl et al., 1998.

[e]Coleman, 1994.

Table 3. DDBJ/EMBL/GenBank accession numbers of bacterial 16S rRNA gene sequences used in this thesis.

Taxon	Accession no.
Bacteria from *Montastraea faveolata* diseased tissue, clone SHFG464	FJ203077
Candidatus Cryptoprodotis polytropus	FM201293
Endosymbiont of *Bryopsis* sp., clone WB4.44	HE648947
Endosymbiont of *Bryopsis* sp., clone WE1.5	HE648945
Endosymbiont of *Bryopsis* sp., clone WE2.2	HE648946
Endosymbiont of *Carteria cerasiformis*	AB688628
Endosymbiont of *Diophrys oligothrix* BOD9	AJ630204
Endosymbiont of *Diophrys oligothrix* DS12/4	FR823001
Endosymbiont of *Empoasca papayae*	U76910
Endosymbiont of *Euplotes octocarinatus*	FR823004
Endosymbiont of *Hemicrepsis marginata*	AB066352
Endosymbiont of *Hydra oligactis*, clone Ho(lakePloen)_1	EF667899
Endosymbiont of *Hydra oligatics*, clone Ho(lakePohlsee)_4	EF667896
Endosymbiont of *Ichthyophthirius multifiliis*	GQ870455
Endosymbiont of *Nephotettix cinctiseps*	AB702995
Endosymbiont of *Paramecium caudatum*	FR822997
Endosymbiont of *Pleodorina japonica*	AB688629
Endosymbiont of *Spirostomum* sp.	FR822998
Endosymbiont of *Torix tagoi*	AB066351
Endosymbiont of *Torix tsukubana*	AB113214
Endosymbiont of *Volvox carteri* UTEX 2180	AB861537[a]
Orientia tsutsugmushi strain *Boryong*	AM494475
Orientia tsutsugmushi strain *Ikeda*	AP008981
Rickettsia akari	CP000847
Rickettsia australis	L36101
Rickettsia bellii	CP000087
Rickettsia canadensis	L36104
Rickettsia conorii	L36105
Rickettsia felis	CP000053
Rickettsia helvetica	L36212
Rickettsia japonica	L36213
"*Rickettsia limoniae*"	AF322442
Rickettsia prowazekii	AJ235272
Rickettsia rickettsii	CP000766.2
Rickettsia typhi	AE017197
Uncultured clone (Chinese freshwater lake, clone AS94)	JN869203
Uncultured clone (Chinese waste water, clone 214)	JF828749
Uncultured clone (Kalahari Shield South Africa subsurface water, clone EV221H2111601SAH71)	DQ223223
Uncultured clone (ornamental fish aquariar water, clone T60-Ps-25C-1)	JX105713
Uncultured clone (ornamental fish aquariar water, clone T60-Ps-25C-27)	JX105706
Uncultured clone (U.S. freshwater spring, clone Wat77)	KC189769

[a]Sequenced in this thesis.

Table 4. The list of oligonucleotide probes used in fluorescence *in situ* hybridization of this thesis.

Probe designation	Probe sequence (5′ to 3′)	Target
volv-835	CCGAAAGAAAAACTCCCG	Rickettsial 16S rRNA gene
help-volv1	ATATCTAGCACTCATCGT	Rickettsial 16S rRNA gene
help-volv2	AACGTGTTAACTACGAAA	Rickettsial 16S rRNA gene
EUB338[a]	GCTGCCTCCCGTAGGA	Universal 16S rRNA gene
EUB338-II[b]	GCAGCCACCCGTAGGTGT	Universal 16S rRNA gene
EUB338-III[b]	GCTGCCACCCGTAGGTGT	Universal 16S rRNA gene

[a]Amann et al., 1990.

[b]Daims et al., 1999.

Table 5. Draft genome contigs (≥5 kbp) of rickettsial endosymbiont of *Carteria cerasiformis* NIES-425, used for the BLASTN search.

#	Length [bp]	GC%	CDS	rRNA	tRNA
			Annotated features		
1	52518	33.4	47	0	1
2	39775	33.8	39	0	1
3	37827	33.1	37	0	1
4	35490	33.0	40	0	1
5	35440	34.4	35	0	2
6	35382	33.7	38	0	0
7	35023	33.5	26	0	3
8	34943	38.8	38	0	0
9	34873	32.6	32	0	0
10	34637	33.8	31	0	1
11	31687	34.1	33	0	2
12	30818	35.1	26	0	1
13	29507	34.4	27	0	0
14	26744	33.9	23	0	0
15[a]	26476	33.5	23	0	1
16	25154	34.7	30	0	1
17	24229	35.6	37	0	2
18	23959	34.1	24	0	1
19	22860	34.3	19	0	0
20	22334	32.7	20	0	1
21	21779	34.4	20	0	1
22	17607	31.9	21	0	0
23	17333	34.5	18	0	1
24	17253	33.1	17	0	0
25	17125	34.1	16	0	0
26	17012	34.8	15	0	0
27	16310	32.7	17	0	0
28	15101	34.8	13	0	1
29	14858	35.2	16	0	0
30	14759	33.9	11	0	2
31	14743	33.9	13	0	0
32	14568	32.6	15	0	0
33	14449	33.0	11	0	0
34	14409	33.7	14	0	2
35	13973	32.3	14	0	0
36	13214	34.1	9	0	0
37	13146	32.6	12	0	0
38	13130	34.2	10	0	0
39	12880	35.2	13	0	1
40	12813	32.9	12	0	0
41	12090	35.2	17	0	0

42	11834	33.6	8	0	1
43	11716	34.1	15	0	0
44	11713	32.4	19	0	0
45	11673	33.7	6	0	0
46	11438	35.1	6	0	0
47	11426	35.2	13	0	0
48	10460	32.9	9	0	1
49	9782	31.9	11	0	0
50	9589	36.0	7	0	0
51	9094	32.3	10	0	1
52	8843	33.4	10	0	1
53	8662	32.5	9	0	0
54	8651	32.2	8	0	0
55	8640	33.2	9	0	0
56	8538	34.9	8	0	1
57	8177	35.2	7	0	0
58	8170	32.0	9	0	1
59	8091	32.2	8	0	0
60	7864	34.2	8	0	0
61	7747	33.9	7	0	0
62	7590	33.0	12	0	0
63	7454	32.1	9	0	0
64	7382	38.2	7	0	0
65	7277	30.4	13	0	0
66	7251	39.4	6	2 (23S, 5S)	0
67	7231	32.9	9	0	0
68	7182	34.8	7	0	0
69	6827	33.9	7	0	0
70	6593	31.1	9	0	0
71	6588	37.7	5	0	0
72	6208	33.2	7	0	0
73	6073	31.5	7	0	0
74	5951	31.0	5	0	0
75	5786	41.8	2	0	0
76	5731	31.6	5	0	0
77	5549	32.3	3	0	0
78[b]	5466	36.3	5	1 (16S)	0
79	5367	30.4	6	0	0
80	5077	35.0	6	0	1
Total	1272919	33.9	1246	3	34

[a]including *murC, murB, ddlB, ftsQ, plsC* and *phbB* (Figure 9B).

[b]including 16S rRNA gene (Figure 9B).

Table 6. The high scoring hits of *Volvox carteri* f. *nagariensis* strain EVE genome sequence by BLASTN searches using *Carteria cerasiformis* rickettsial endosymbiont draft genome.

BLASTN searches were performed on Phytozome v9.1, using 80 contigs (>5 kb) from preliminary genome assembly database of *C. cerasiformis* NIES-425 rickettsial endosymbiont. Hits with the E-value ≤1.0e-50 of BLASTN searches and the information of *ftsQ*-like sequence found in following BLASTN search in scaffold 6 of EVE genome are shown.

Scaffold	Position		Score	E-value	Identity (%)	(bp/bp)	Annotation
6	940432	939419	1373.7	0	90.0	913/1014	D-alanyl-alanine ligase [*Rickettsia japonica*]
6	938261	937504	1014.8	0	89.6	682/761	UDP-N-acetylenolpyruvoylglucosamine reductase [*Rickettsia helvetica*]
6	937505	937077	747.9	0	98.6	423/429	*Rickettsiaceae* endosymbiont of *Pleodorina japonica* gene for 16S ribosomal RNA
36	56105	56609	446.7	5.90E-121	78.9	426/540	Hypothetical protein [*Rickettsia rhipicephali*]
30	1311260	1311562	380	7.20E-101	87.8	266/303	Putative bifunctional glutamate synthase subunit beta/2-polyprenylphenol hydroxylase [*Rickettsia canadensis*]
1	13232743	13232265	352.9	1.00E-92	76.2	374/491	Surface antigen (D15) [*Desulfohalobium retbaense*]
60	166248	166583	336.7	7.70E-88	82.4	277/336	DNA repair protein RadC [*Rickettsia massiliae*]
1	8394848	8394414	333.1	9.30E-87	76.8	350/456	Heat shock protein GrpE [*Amorphus coralli*]
114	9149	9491	324.1	4.80E-84	81.0	278/343	Hypothetical protein [*Rickettsiaceae* bacterium Os18]
7	3741662	3741429	309.7	1.10E-79	89.3	209/234	MFS transporter [*Rickettsia bellii*]
11	2289762	2289993	300.6	5.50E-77	84.1	238/283	RND efflux transporter [*Pseudomonas aeruginosa* PA99]
24	1251646	1251958	298.8	1.90E-76	82.5	259/314	Small multidrug resistance (SMR) family efflux pump [*Rhizobium tropici*]
6	2460215	2459921	255.6	2.10E-63	79.2	236/298	Hypothetical protein GUITHDRAFT_133732 [*Guillardia theta* CCMP2712]
13	350503	350293	241.1	4.50E-59	85.8	182/212	No significant hits
17	2041301	2041462	232.1	2.40E-56	92.0	150/163	UDP-3-O-(3-hydroxymyristoyl) glucosamine N-acyltransferase (LpxD) [*Rickettsia typhi*]

6	2459685	2459520	219.5	1.50E-52	89.2	148/166	Recombination protein RecJ [*Rickettsia bellii*]
11	57355	57566	217.7	5.20E-52	83.0	176/212	Phenylalanyl-tRNA synthetase subunit beta [*Rickettsia bellii*]
48	507707	507884	217.7	5.20E-52	87.1	155/178	Elongation factor Tu family protein [*Wolbachia* endosymbiont of *Drosophila ananassae*]
6[a]	940359	940432	93.3	8.40E-18	87.8	65/74	Cell division protein *ftsQ* [*Rickettsia canadensis*]

[a]Hit found in additional search (see Materials and Methods).

Table 7. DDBJ/EMBL/GenBank accession numbers of rickettsial *murB* and *ddlB* genes/gene-like sequences used in this thesis.

Taxon	Accession No.	
	murB	*ddlB*
Rickettsia massiliae str. MTU5	NC_009900	NC_009900
Rickettsia massiliae str. AZT80	-	NC_016931
Rickettsia rickettsii str. Iowa	NC_010263.2	NC_010263.2
Rickettsia conorii str. Malish 7	AE006914	AE006914
Rickettsia japonica str. YH	NC_016050	NC_016050
Rickettsia slovaca str. D-CWPP	CP003375	CP003375
Rickettsia slovaca str. 13-B	-	CP002428
Rickettsia parkeri str. Portsmouth	CP003341	CP003341
Rickettsia rhipicephali str. 3-7-female6-CWPP	CP003342	CP003342
Rickettsia heilongjiangensis str. 54	CP002912	CP002912
Candidatus Rickettsia amblyommii str. GAT-30V	CP003334	CP003334
Rickettsia philipii str. 364D	CP003308	CP003308
Rickettsia africae str. ESF-5	CP001612	CP001612
Rickettsia montanensis str. OSU 85-930	CP003340	CP003340
Rickettsia peacockii str. Rustic	CP001228	CP001228
Rickettsia typhi str. Wilmington	NC_006142	NC_006142
Rickettsia prowazekii str. MadridE 2/4	AJ235271	AJ235271
Rickettsia australis str. Cutlack	CP003338	CP003338
Rickettsia akari str. Hartford chromosome	NC_009881	NC_009881
Rickettsia felis str. URRWXCal2	NC_007109	NC_007109
Rickettsia canadensis str. McKiel	NC_009879	NC_009879
Rickettsia canadensis str. CA410	CP003304	CP003304
Rickettsia bellii str. OSU 85-389	NC_009883	NC_009883
Rickettsia bellii str. RML369-C	NC_007940	NC_007940
Orientia tsutsugamushi str. Boryong	NC_009488	NC_009488
Orientia tsutsugamushi str. Ikeda	NC_010793	NC_010793
Endosymbiont of *Carteria cerasformis* NIES-425	LC004725[a]	LC004725
Endosymbiont of *Pleodorina japonica* NIES-577	LC004723[a]	LC004723
Endosymbiont of *Volvox carteri* f. *weismannia* UTEX 2180	LC004722[a]	LC004722
Volvox carteri f. *nagariensis* EVE[b]	LC004713[a]	LC004713
Volvox carteri f. *nagariensis* UTEX 1886[b]	LC004714[a]	LC004714
Volvox carteri f. *nagariensis* NIES-397[b]	LC004715[a]	LC004715
Volvox carteri f. *nagariensis* NIES-398[b]	LC004716[a]	LC004716
Volvox carteri f. *weismannia* NIES-866[b] (=UTEX 1875)	LC004719[a]	LC004719
Volvox carteri f. *weismannia* UTEX 1876[b]	LC004720[a]	LC004720
Volvox carteri f. *weismannia* UTEX 2170[b]	LC004721[a]	-

[a]Sequence determined in this thesis.

[b]Strain lacking rickettsial endosymbionts.
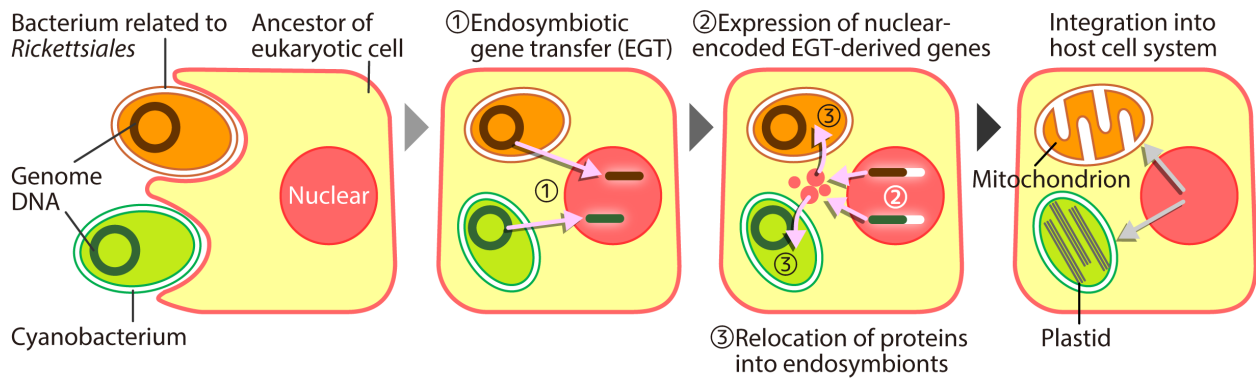
Figures

Figure 1. A model of the evolution of organelles.

Bacterial endosymbionts inside an ancestor of eukaryotic cell undergo an integrative change toward organelles, through the three steps: (1) endosymbiotic gene transfer (EGT) to the host nuclear genome, (2) an establishment of the expression system to express and regulate nuclear-encoded EGT-derived genes, and (3) an acquisition of the retargeting system that relocates expressed proteins into compartments of the endosymbionts. (Based on: Timmis et al., 2004; Kleine et al., 2009)

Figure 2. Schematic phylogeny of the hydra group within the order *Rickettsiales* (*Alphaproteobacteria*), based on Weinert et al. (2009) and Montagna et al. (2013).

Figure 3. Nomarski differential interference images of bacterial endosymbiont-containing strains of Volvocales used in this thesis.

A. *Carteria cerasiformis* NIES-425. B. *Pleodorina japonica* NIES-577. C. *Volvox carteri* f. *weismannia* UTEX 2180.

Figure 4. Phylogenetic tree of the endosymbiont of *Volvox carteri* UTEX 2180 within the family *Rickettsiaceae*.

The tree was inferred using the maximum-likelihood (ML) method (with the GTR+GAMMA+I model) based on 41 sequences and 1403 sites in the 16S rRNA gene from endosymbionts (En) of eukaryotic hosts and other environmental samples in the family *Rickettsiaceae*. Bootstrap values (≥50%) for the ML and maximum parsimony (MP) analyses and posterior probability (≥0.90) for Bayesian interferrence (BI) are indicated at the respective nodes. The scale bar shows 0.01 nucleotide substitutions per position. The accession numbers of the sequences are shown in Table 3. The hydra group refers to Weinert et al. (2009).

Figure 5. FISH identification of *Rickettsiaceae* endosymbionts in cells of *Carteria cerasiformis*, *Pleodorina japonica*, and *Volvox carteri*.

A–C. *C. cerasiformis* NIES-425. D–F. *C. cerasiformis* NIES-424. G-I. *P. japonica* NIES-577. J-L. *V. carteri* UTEX 2180. Horizontal panels show the same cells at the same magnification, composed of epifluorescence images with the volv-835 probe specific for the 16S rRNA sequence of *C. cerasiformis* NIES-425 endosymbiont (A, D, G, J), epifluorescence images with DAPI staining (B, E, H, K) and Nomarski differential interference images (C, F, I, L). White arrowheads indicate the signals from the

69

endosymbionts. The green signals (A, G, J) represent endosymbiont-specific probes and the yellow background (A, D, G, J) is autofluorescence. 'n' and 'p' indicate host cell nuclei and pyrenoids, respectively. Scale bar = 10 μm.

Figure 6. FISH images with bacterial universal probes EUB338MIX in cells of *Carteria cerasiformis* and *Volvox carteri*.

A–C. *C. cerasiformis* NIES-425. D–F. *C. cerasiformis* NIES-424. G-I. *V. carteri* UTEX 2180. Horizontal panels show the same cells at the same magnification, composed of epifluorescence images with 16S rRNA probes EUB338 MIX (A, D, G), epifluorescence images with DAPI staining (B, E, H; n, the host cell nuclei) and Nomarski differential interference images (C, F, I; p, the pyrenoids). The green fluorescence and yellow background (A, D, G) are the probe signal and autofluorescence respectively. bar = 10 μm.

Figure 7. Somatic cells of *Volvox carteri* UTEX 2180 and seven other *V. carteri* strains stained with DAPI.

Each pair of two vertical panels shows the same cells shown at the same magnification, composed of epifluorescence images (A-D, I-L) and Nomarski differential interference images (E-H, M-P). The arrowhead, arrow, 'n' and 'p' indicate the bacterial endosymbiont, chloroplast nucleoid, host cell nuclei and pyrenoids, respectively. Scale bar = 10 μm. Bacteria-like rod-shaped bodies were observed only in the cells of strain UTEX 2180 (A, E), but not in the cells of strain NIES-398 (B, F), strain NIES-733 (C, G), strain UTEX 2170 (D, H), strain UTEX 1874 (I, M), strain UTEX 1886 (J, N), strain UTEX 2903 (K, O) and strain UTEX 2904 (L, P).

Figure 8. Detection of 16S rRNA gene of the *Rickettsiaceae* in *Volvox carteri* strains by genomic PCR.

PCR amplification using the forward primer eveFC and reverse primer enRB (specific 16S rRNA primers specific to the bacteria belonging to the hydra group, see Table 2) corresponds the presence or absence of rickettsial endosymbionts. *Carteria cerasiformis* NIES-425 and *Pleodorina japonica* NIES-577 (harboring rickettsial endosymbionts), and *C. cerasiformis* NIES-424 (lacking rickettsial endosymbionts) (Kawafune, 2011) are shown as positive and negative controls respectively. As a PCR control, the eukaryotic 18S rRNA gene was amplified as described in Chapter 2.

A *Volvox obversus* UTEX 1865

B *Volvox carteri* f. *nagariensis*
EVE, UTEX 1886,
NIES-397, NIES-398, UTEX 2903
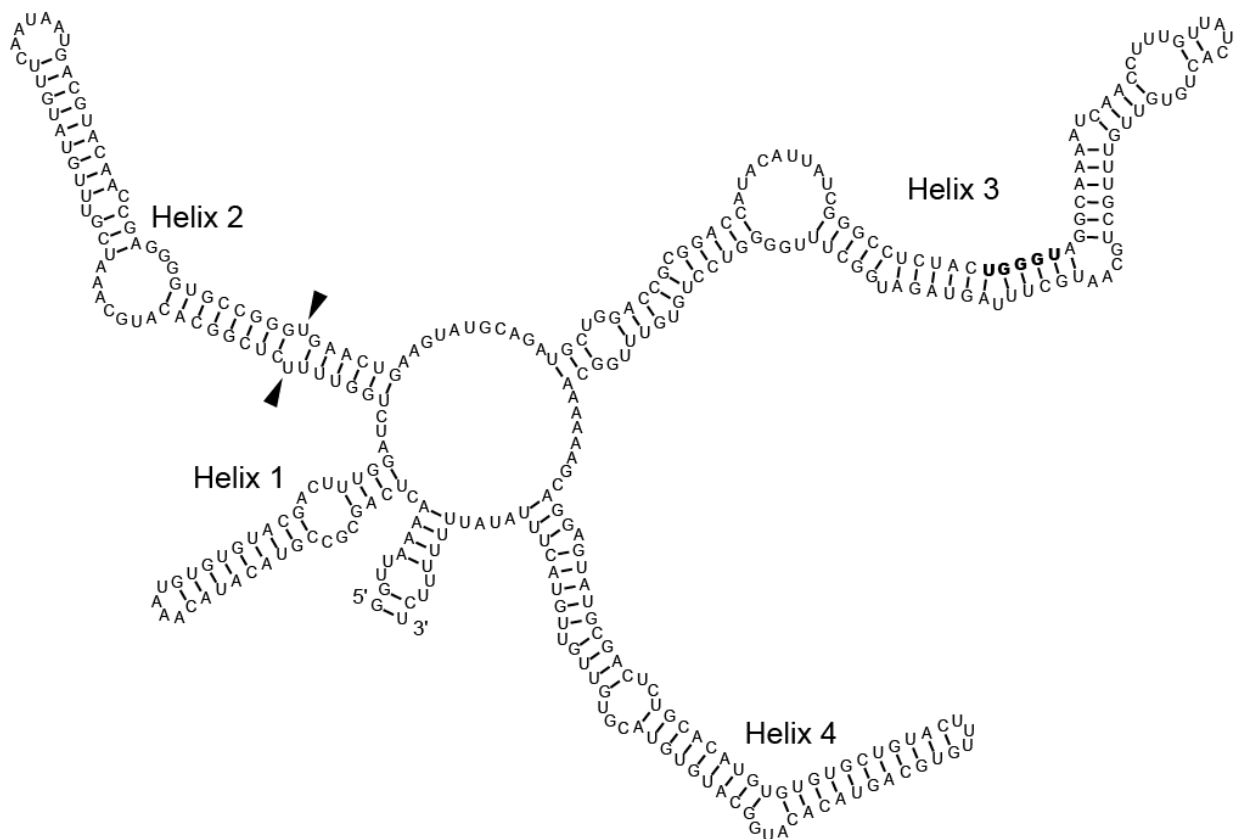
C *V. carteri* f. *kawasakiensis* NIES-732, NIES-733

74

(continued)
E *Volvox carteri* f. *weismannia*
UTEX 1875, UTEX 1876

F *Volvox carteri* f. *weismannia*
UTEX 2170

(continued)

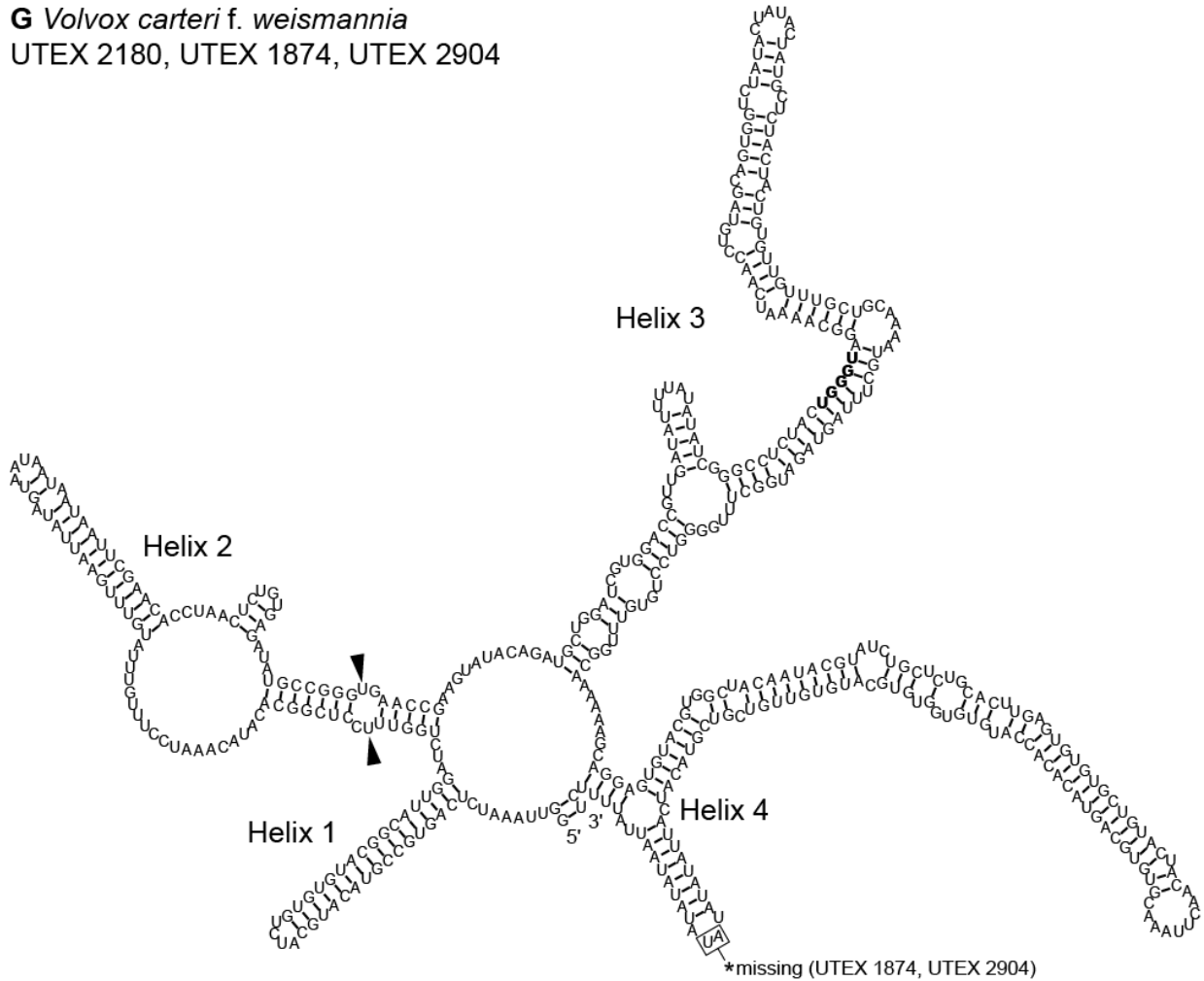**G** *Volvox carteri* f. *weismannia*
UTEX 2180, UTEX 1874, UTEX 2904

Figure 9. Whole secondary structure of nuclear ribosomal DNA internal transcribed spacer 2 region of the strains of *Volvox obversus* and *V. carteri*.

The structure was predicted and drawn as described in Chapter 3. The U-U mismatch in helix 2 (arrowheads) and the UGGU motif on the 5′ side near the apex of helix 3 (boldface) are the universally conserved features (Coleman & Mai, 1997). Among the ITS-2 sequences of *V. carteri* f. *nagariensis* strains, one nucleotide of UTEX 1886 and NIES-398, and three of UTEX 2903 are different from those of strains EVE and NIES-397 (shown around the structures). ITS-2 sequences of *V. carteri* f. *weismannia* strains UTEX 1874 and UTEX 2904 differ from that of *V. carteri* f. *weismannia* strain UTEX 2180 in missing of one pair of "AU" from AU repeats in helix 4 (asterisk).
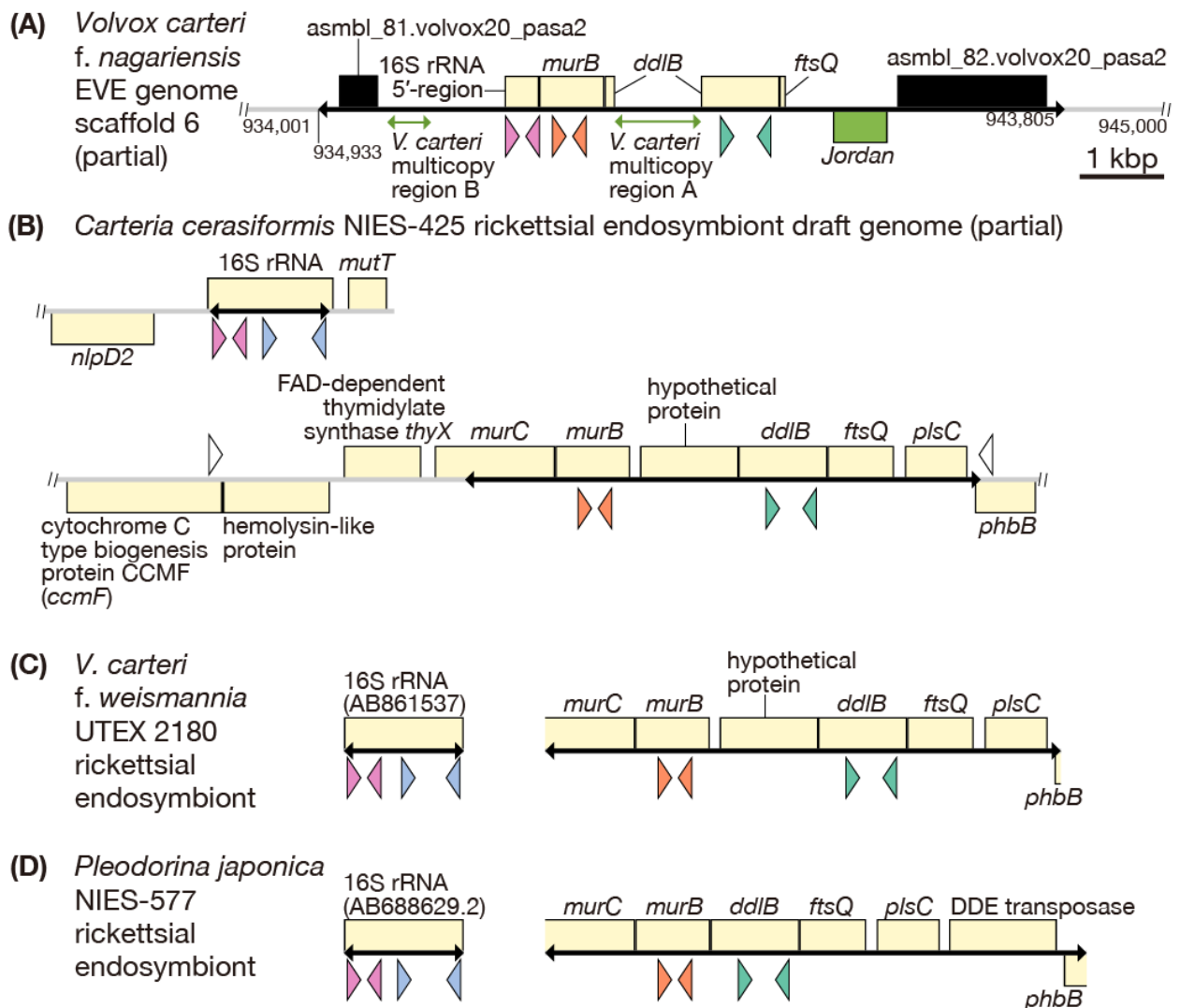
Figure 10. Rickettsial genes and gene-like sequences in the genomes of *Volvox carteri* and rickettsial possible endosymbionts.

Schematic representations of arrangements/synteny of several rickettsial genes and gene-like sequences present in DNA of the nuclear genome of *V. carteri* (A) and in the genomes of rickettsial possible endosymbionts harbored by three volvocalean species (B-D). Coding DNA sequences (CDSs) and CDS-like regions are shown as boxes. Rickettsial CDSs/CDS-like regions are shown in pale yellow, the *V. carteri* transposon *Jordan*-like region in green and others in black. Placement of boxes above/below the line indicates gene direction (from left-to-right or right-to-left, respectively). Black double-headed arrows on the baseline indicate the regions sequenced in this thesis. Colored triangles under boxes indicate the locations of primers used for semi-quantitative genomic PCR (16S rRNA gene 5'-region: magenta, 16S rRNA gene

3′-region: light blue, *murB*: orange, *ddlB*: green; Figure 11E). For accession numbers of sequences used in this figure, see Table 1. (A) Part of scaffold 6 of the *V. carteri* f. *nagariensis* strain EVE nuclear genome. (B) Part of the *Carteria cerasiformis* NIES-425 draft endosymbiont genome, including 16S rRNA (first line) and *murB-ftsQ* (second line). White triangles indicate primers used to amplify the sequencing templates (ccmF-R02 and phbB-F01; see Chapter 3). (C) Part of the genome of a possible endosymbiont of *V. carteri* f. *weismannia* strain UTEX 2180, including *murB* and *ddlB* (right). The 16S rRNA gene of the endosymbiont is also shown (left). (D) Part of the genome of a possible endosymbiont of *Pleodorina japonica* strain NIES-577, including *murB* and *ddlB* (right). The 16S rRNA gene is also shown (left).
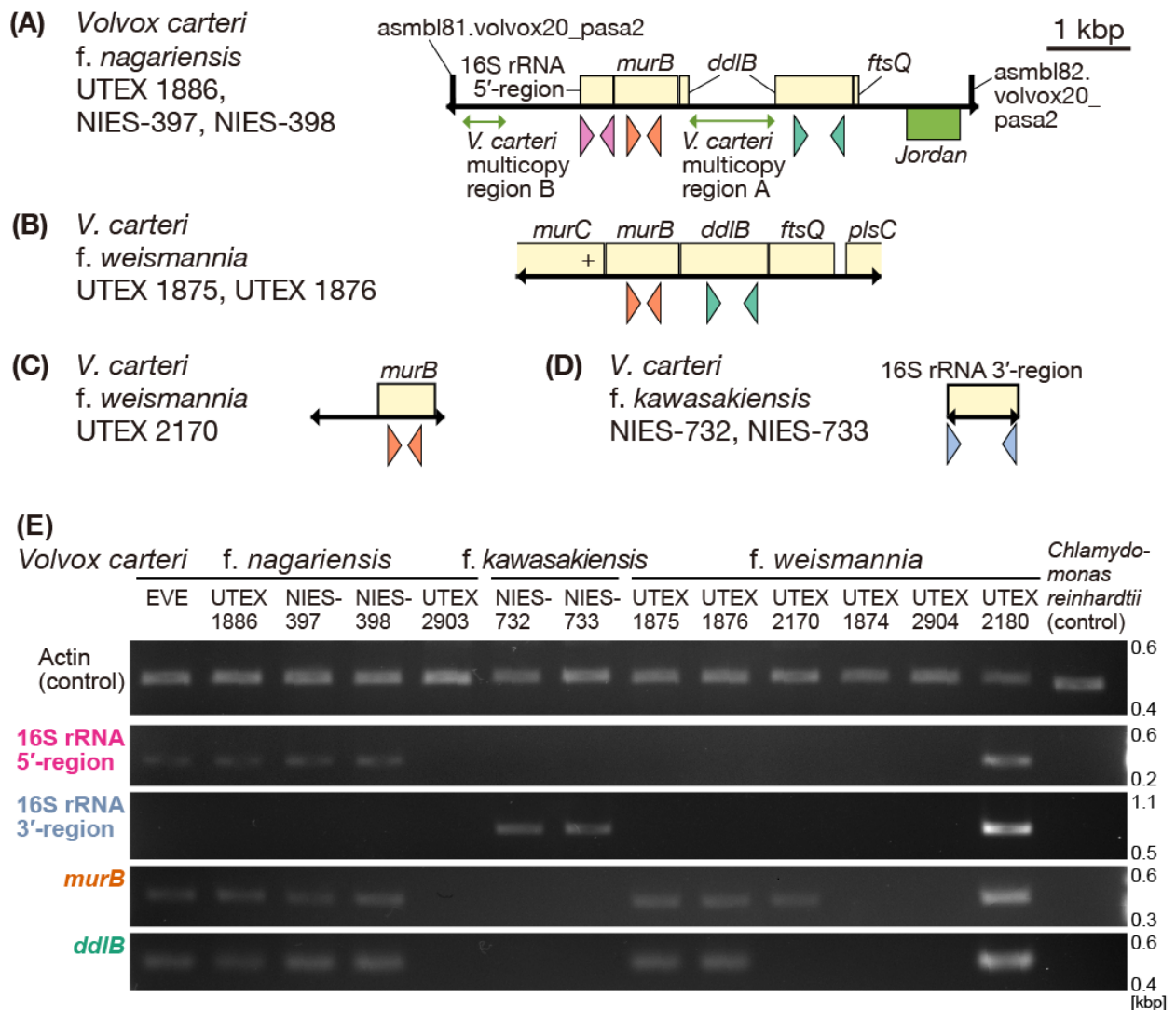
Figure 11. Rickettsial gene-like sequences possibly located in the nuclear genomes of *Volvox carteri* strains.

Schematic representations of arrangements/synteny (A-D) in, and semi-quantitative genomic PCR data (E) from, several rickettsial gene-like sequences possibly located in the nuclear genomes of *V. carteri* strains. Coding DNA sequence (CDS)-like regions are shown as boxes. Rickettsial CDS-like regions are shown in pale yellow, the *V. carteri* transposon *Jordan*-like region in green and others in black. Placement of boxes above/below the line indicates the gene direction (from left-to-right or right-to-left, respectively). Black double-headed arrows on the baseline indicate the regions sequenced in this thesis. Colored triangles under boxes indicate the primers used for semi-quantitative genomic PCR (16S rRNA gene 5'-region: magenta, 16S rRNA gene 3'-region: light blue, *murB*: orange, *ddlB*: green; Figure 11E). For accession numbers of

sequences used in this figure, see Table 1. (A) Sequences including rickettsial CDS-like regions of *V. carteri* f. *nagariensis* strains UTEX 1886, NIES-397 and NIES-398. (B) Sequences including rickettsial gene homologs of *V. carteri* f. *weismannia* strains UTEX 1875 and UTEX 1876. Plus (+) indicates a frameshift deletion. (C) Sequence including rickettsial *murB*-like sequence of *V. carteri* f. *weismannia* strain UTEX 2170. (D) Sequences including rickettsial 16S rRNA gene-like sequences of *V. carteri* f. *kawasakiensis* NIES-732 and NIES-733. (E) Semi-quantitative genomic PCR of rickettsial genes and gene-like sequences. Each rickettsial gene-like sequence was amplified via genomic PCR using rickettsia-specific primer sets (see Chapter 3). The positions of primer sets with reference to target positions are shown in both Figure 10 and this Figure 11. As a control, the actin gene was amplified. *Chlamydomonas reinhardtii* strain CC-503 was used as negative control.
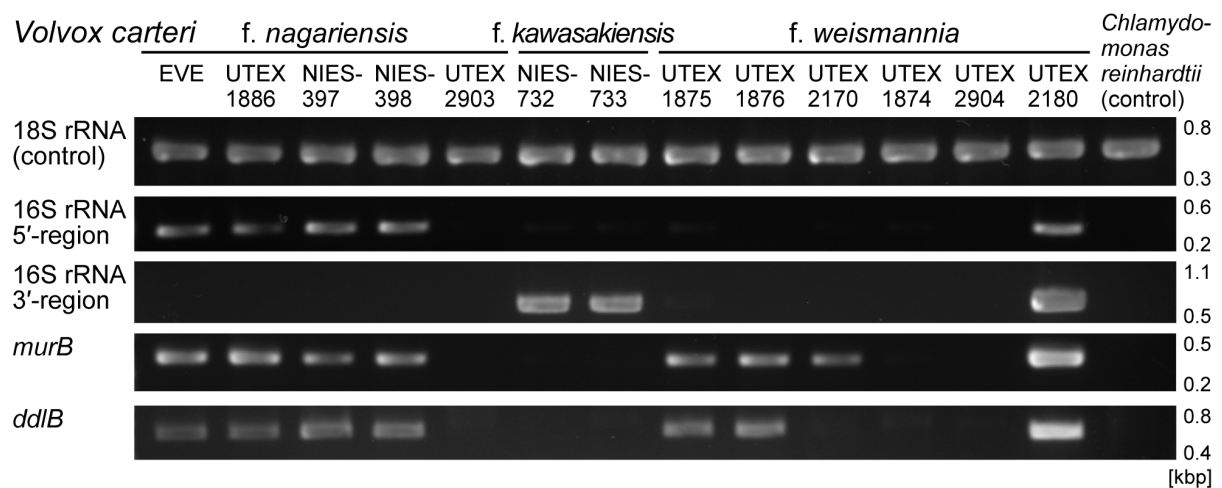
Figure 12. Detection of rickettsial gene-like sequences in 13 strains of *Volvox carteri*. Rickettsial gene-like sequences were amplified via genomic PCR using rickettsia-specific primer sets (see in Chapter 3). For PCR amplification, 12 endosymbiont-lacking strains of *V. carteri*, endosymbiont-containing *V. carteri* f. *weismannia* strain UTEX 2180 (positive control) and *Chlamydomonas reihnardtii* strain CC-503 (negative control) were used. As a control, the eukaryotic 18S rRNA gene was amplified.
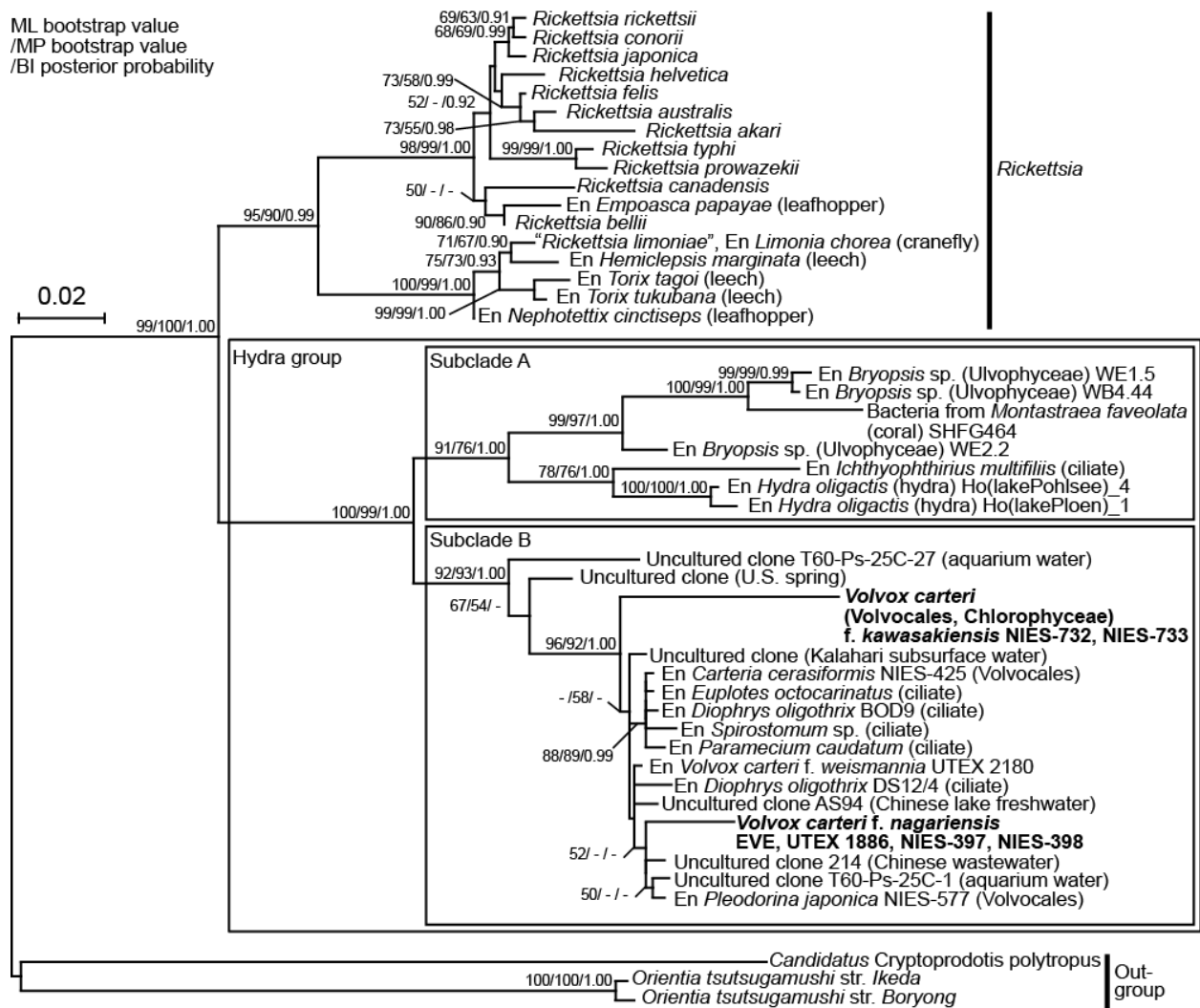
Figure 13. Phylogenetic positions of rickettsial 16S rRNA gene-like sequences in
Volvox carteri **nuclear genomes.**

The tree was inferred using the maximum-likelihood (ML) method based on 43
sequences and 1,403 nucleotides of the 16S rRNA genes from bacteria, endosymbionts
(En) of eukaryotic hosts, and other environmental samples, of the family *Rickettsiaceae*,
including rickettsial 16S rRNA gene-like sequences obtained from
endosymbiont-lacking strains of *V. carteri* (bold). Bootstrap values (≥50%) for the ML
and maximum parsimony analyses, and posterior probabilities (≥0.90) for Bayesian
interference, are indicated at the respective nodes. The scale bar corresponds to 0.02
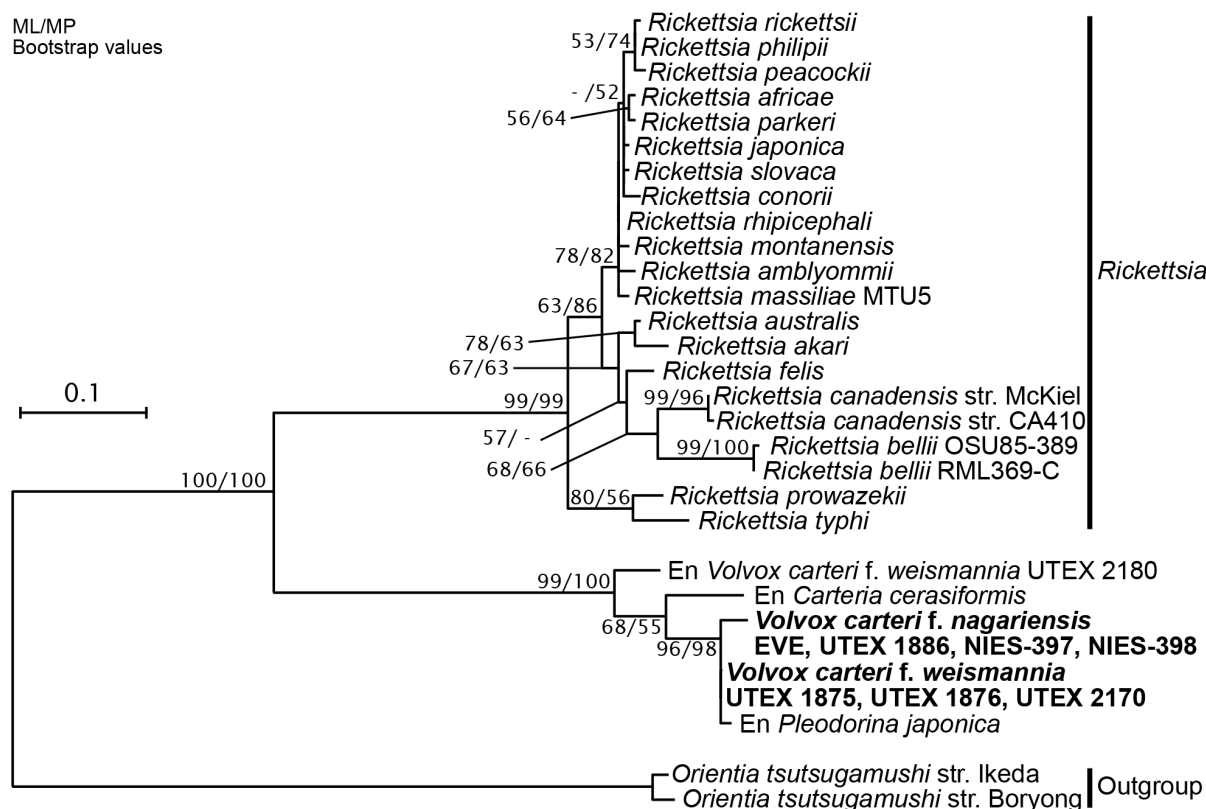nucleotide substitutions per position.

Figure 14. Phylogenetic positions of rickettsial *murB* gene-like sequences from endosymbiont-lacking strains of *Volvox carteri*.

The tree was inferred based on translated rickettsial *murB* genes and gene-like sequences (227 amino acid sites) from endosymbiont-lacking strains of *V. carteri* (boldface), with 26 translated *murB* sequences from bacteria and possible endosymbionts (En) of algal hosts in the family *Rickettsiaceae,* using the maximum-likelihood (ML) method. Bootstrap values (≥50%) for the ML and maximum parsimony analyses are indicated at the respective nodes. The scale bar corresponds to 0.1 amino acid substitutions per position.

Figure 15. Phylogenetic positions of rickettsial *ddlB* gene-like sequences from endosymbiont-lacking strains of *Volvox carteri*.

The tree was inferred based on translated rickettsial *ddlB* genes and gene-like sequences (361 amino acid sites) from endosymbiont-lacking strains of *V. carter* (boldface) with 28 translated *ddlB* sequences from bacteria and possible endosymbionts (En) of algal hosts in the family *Rickettsiaceae*, using the maximum-likelihood (ML) method. Bootstrap values (≥50%) for the ML and maximum parsimony analyses are indicated at the respective nodes. The scale bar corresponds to 0.1 amino acid substitutions per position.
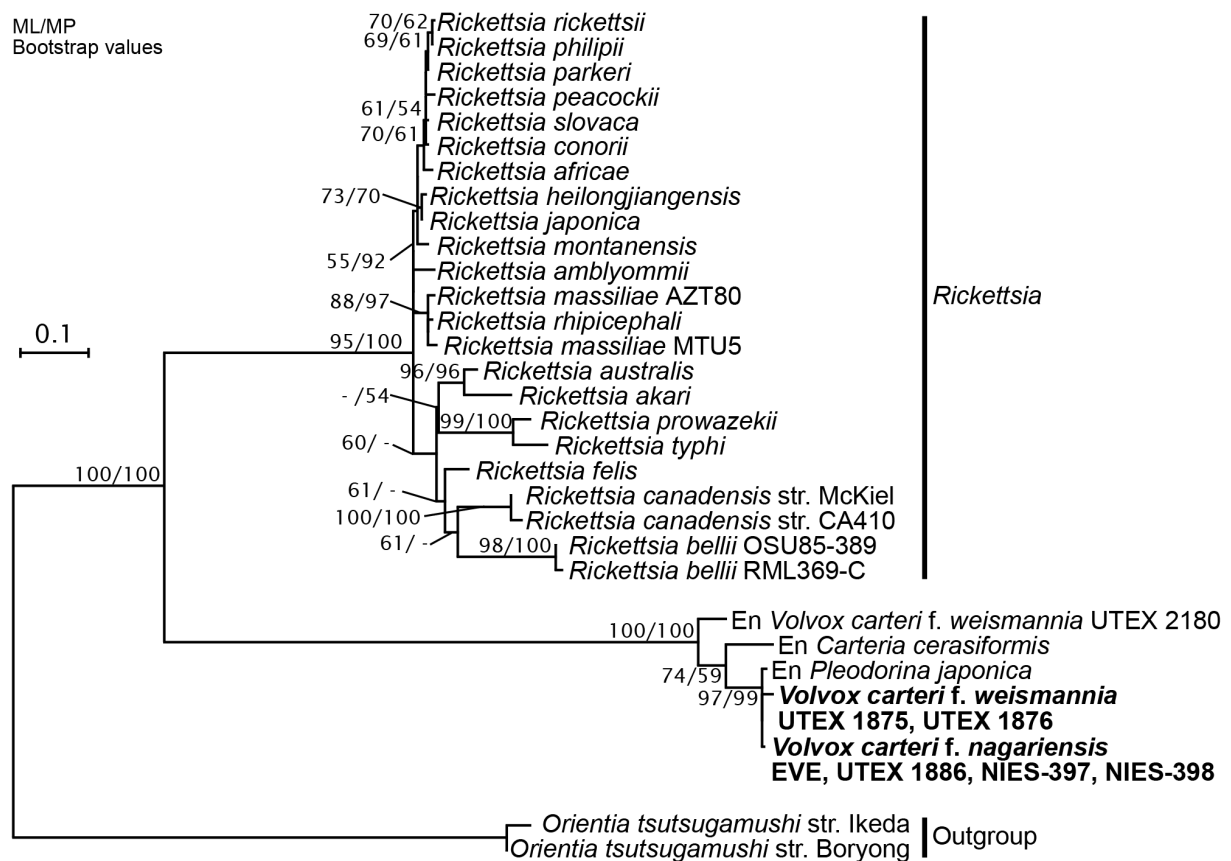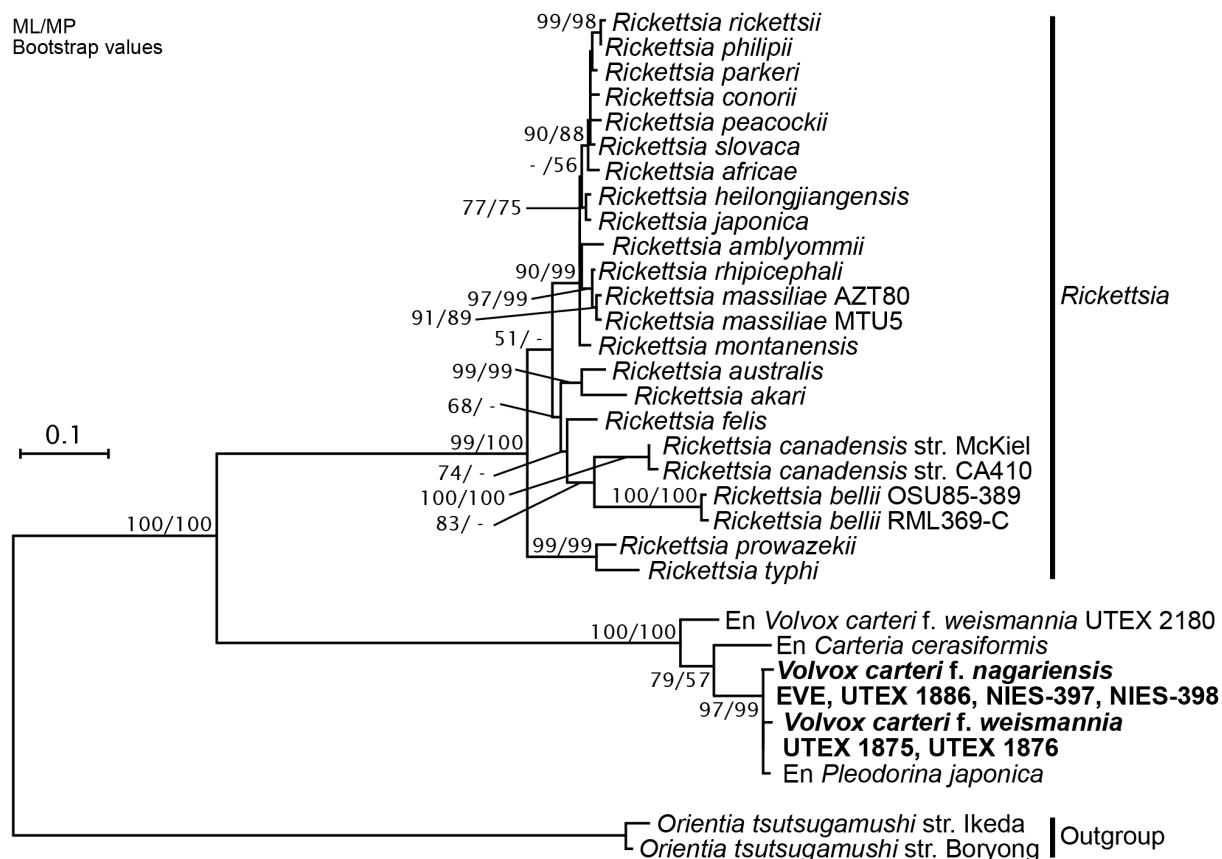
Figure 16. Phylogenetic positions of rickettsial *murB* and *ddlB* gene-like sequences from endosymbiont-lacking strains of *Volvox carteri*.

The tree was inferred using the maximum-likelihood (ML) method based on 637 amino acid sites in translated and combined *murB* and *ddlB* gene / gene-like sequences from 30 operational taxonomic units of bacteria in the *Rickettsiaceae,* including possible endosymbionts (En) of algal hosts, and possible nuclear-encoded sequences from endosymbiont-lacking strains of *V. carteri* (bold). Bootstrap values (≥50%) for ML and maximum parsimony analyses are indicated at the respective nodes. The scale bar corresponds to 0.1 amino acid substitutions per position.
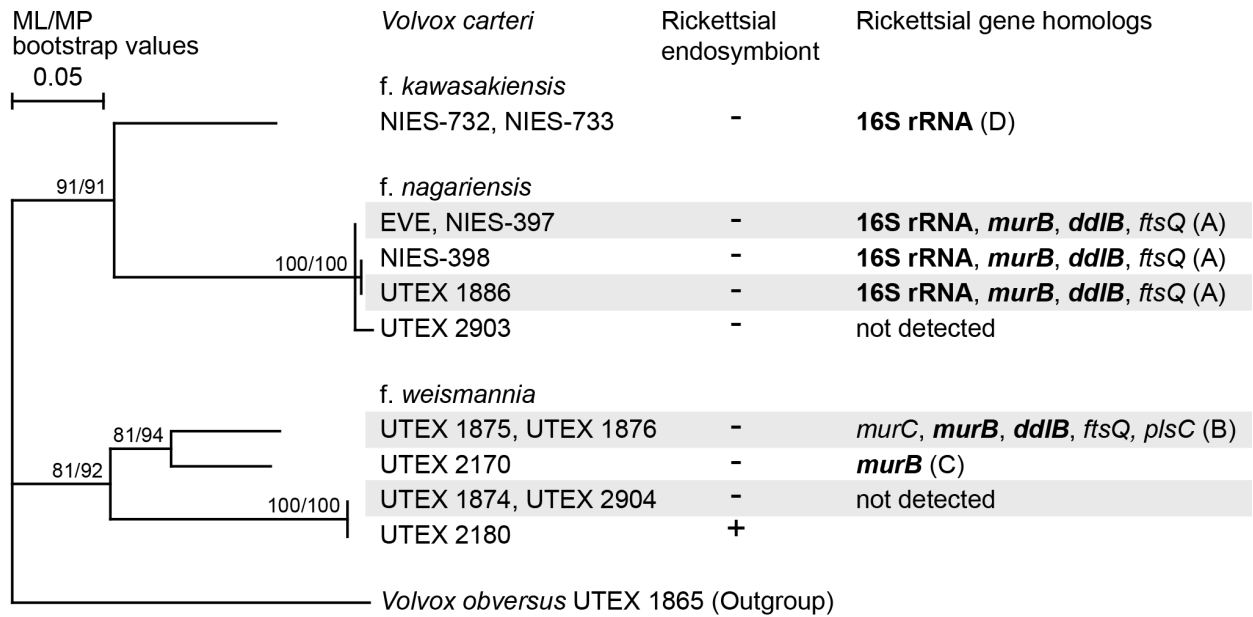
ML/MP bootstrap values	*Volvox carteri*	Rickettsial endosymbiont	Rickettsial gene homologs
0.05	f. *kawasakiensis* NIES-732, NIES-733	-	**16S rRNA** (D)
91/91	f. *nagariensis*		
	EVE, NIES-397	-	**16S rRNA**, **murB**, **ddlB**, *ftsQ* (A)
100/100	NIES-398	-	**16S rRNA**, **murB**, **ddlB**, *ftsQ* (A)
	UTEX 1886	-	**16S rRNA**, **murB**, **ddlB**, *ftsQ* (A)
	UTEX 2903	-	not detected
	f. *weismannia*		
81/94	UTEX 1875, UTEX 1876	-	*murC*, **murB**, **ddlB**, *ftsQ*, *plsC* (B)
81/92	UTEX 2170	-	**murB** (C)
100/100	UTEX 1874, UTEX 2904	-	not detected
	UTEX 2180	+	
	Volvox obversus UTEX 1865 (Outgroup)		

Figure 17. Phylogenetic relationships among 13 strains of three forms of *Volvox carteri*.

The tree was inferred using the maximum-likelihood (ML) method based on alignment of 471 nucleotide sites in the internal transcribed spacer 2 sequences of 10 operational taxonomic units of *V. carteri* strains and *V. obversus* strain UTEX 1865 (the outgroup). Bootstrap values (50% or more) for the ML and maximum parsimony analyses are indicated at the respective nodes. The scale bar shows 0.05 nucleotide substitutions per position. The presence (+) or absence (-) of rickettsial endosymbionts based on the data of Chapter 2 (Figures 7, 8) is shown in the central column. Possible nuclear-encoded, rickettsial gene homologs detected in this thesis (Figures 11E, 12) are shown in the column on the right. The gene names shown in bold were used in phylogenetic analyses (Figures 13-16). The letters A-D following gene names correspond to the forms of genetic composition shown in Figure 11. No rickettsial gene-like sequences were detected (in this thesis) in strains UTEX 2903, UTEX 1874 or UTEX 2904.