Original

Analysis of an intron intervening the SSU rDNA of *Chlorella* sp. T-24-5, a photobiont of *Paramecium bursaria*

Ryo HOSHINA

Department of Biomedical Science, College of Life Sciences, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

SUMMARY

The intronic sequence intervening the small subunit ribosomal DNA (SSU rDNA) of *Chlorella* sp. T-24-5, an atypical photobiont of *Paramecium bursaria* was examined. The position of the insertion was found to be 10 nucleotides upstream from that of a major *P. bursaria* photobiont, *Micractinium reisseri* (S651), and was found to be a novel insertion site (S641; the numbering reflects the homologous position in the rRNA gene of *Escherichia coli*: S = SSU rRNA). A secondary structure diagram showed that the intron is classified as a group I intron (subgroup IC), characterized by an extended P5 helix. Phylogenetic analyses could not reveal its evolutionary relationships with other introns, but were suggestive of a monophyletic relationship with introns of some trebouxiophytes. These introns all share the insertion position S641, and their sequences are extremely conserved and are likely to have spread recently. The intron of *Chlorella* sp. T-24-5 had twelve-nucleotide sequence repeats lying at the head of the intron and after the insertion, which may play a role in intron invasion.

Key words: Group I intron, S641, Photobiont

INTRODUCTION

The green ciliate *Paramecium bursaria* (Ehrenberg) Focker is one of the best-studied protists because of its observable endosymbiosis. The symbiotic relationship can be restarted—algae-removed *P. bursaria* can absorb and fix algae as new photobionts (Kodama and Fujishima, 2009 and references therein). Despite the ability of *P. bursaria* to re-establish symbiosis, the diversity of its photobionts is limited. Nearly 50 strains of photobionts have been genetically identified; however, most of them belong to either *Chlorella variabilis* Shihira et Krauss or *Micractinium reisseri* Hoshina, Iwataki et Imamura (Chlorellaceae, Trebouxiophyceae) (Linz et al., 1999; Kvitko et al.,

Tel: +81-77-566-1111/Fax: +81-77-561-5203
E-mail: wwhoseena@hotmail.com
Present address: Department of Environmental Science and Engineering, Graduate School of Science and Engineering, Yamaguchi University, Yamaguchi 753-8512, Japan
Received: 1 February 2012; Accepted: 19 May 2012.

2001; Hoshina et al., 2004, 2005, 2010; Gaponova et al., 2007; Summerer et al., 2008; Hoshina and Imamura, 2008a; Vorobyev et al., 2009; Pröschold et al., 2011).

To determine which photobiont *P. bursaria* possesses, algal DNA amplifications directly from *P. bursaria* extracts have been introduced (Hoshina and Imamura, 2009a; Vorobyev et al., 2009). These methods are based on unique intron insertions at different sites in the small subunit ribosomal DNA (SSU rDNA) of these photobionts; polymorphisms in the lengths of PCR products containing (or not containing) these introns can be a useful tool for symbiont identification (intron insertion sites, see Hoshina et al., 2010).

Recently, Vorobyev et al. (2009) reported an atypical photobiont SSU rDNA sequence from *P. bursaria* collected from Tajikistan. This sequence (*Chlorella* sp. T-24-5, GenBank accession number EU281549) included an intron at nearly the same position as in *M. reisseri* (described as "Northern" ecotype), but the length and sequence were different (Vorobyev et al., 2009). They stated that the intron sequence did not show any significant similarity to other sequences using a BLAST search for nucleotide, whereas it was highly similar to the group I intron intervening the major capsid protein Vp54 gene (AB006978) of *C. variabilis* virus (CvV) under conditions of "somewhat similar sequences" in the viral nucleotide database.

Group I introns are a distinct RNA group that function as enzymes, splicing themselves out of precursor RNA transcripts and ligating exons. Another distinctive characteristic of group I introns is their mobility. Phylogenetic analyses have indicated that introns at homologous gene sites are related (position family), even among distantly related host organisms. This phenomenon is linked to intron spreading mechanisms (i.e. homing or reverse splicing). When an intron at a gene locus infects a different organism, the new intron will be inserted into the same locus where it was originally located (for the general characteristics of group I introns, see Cech, 2002; Haugen et al., 2005; Nielsen and Johansen, 2009 and the references therein). Group I introns are classified into subgroups IA through IE based on their structural diversity and phylogeny, and nuclear encoding introns belong to either subgroup IC or IE (Cannone et al., 2002).

The photobionts of *P. bursaria*, *C. variabilis* has four subgroup IC and four IE introns and *M. reisseri* has two IE introns in their nuclear rDNA. These introns are interesting for two reasons. First, some viruses possess group I introns, of which the only eukaryotic virus is CvV (Zhou et al., 2008). This virus infects only *C. variabilis*, which is a photobiont of *P. bursaria*. Since their introns are classified as IC type (like the introns intervening eukaryotic nuclear rDNA), some researchers have suggested that viruses play a role in intron transfer (Yamada et al., 1994; Bhattacharya et al., 1996; Nishida et al., 1998; Friedl et al., 2000). Second, subgroup IE introns commonly inserted into *C. variabilis* and *M. reisseri* are monophyletic and independent from other IE intron lineages. However, the insertion positions are straggled in their rDNA, and this cannot be explained by existing intron transmission mechanisms. In other terms, unknown intron transmissions could have occurred in *P. bursaria* (Hoshina and Imamura, 2009b). This study examined the sequence, structure, and phylogenetic relationship of the *Chlorella* sp. T-24 -5 intron.

MATERIALS AND METHODS

Structure prediction

A secondary structure diagram of the *Chlorella* sp. T-24-5 intron (EU281549) was prepared according to previously reported models (e.g. Lehnert et al., 1996; Haugen et al., 2002). The structure of the central catalytic intron core (e.g. from P3 to P7) was solved previously; the remaining helices, predicted using the Mfold (Mathews et al., 1999; Zuker, 2003) web server (http://mfold.rna.albany.edu/?q=mfold), were subsequent-

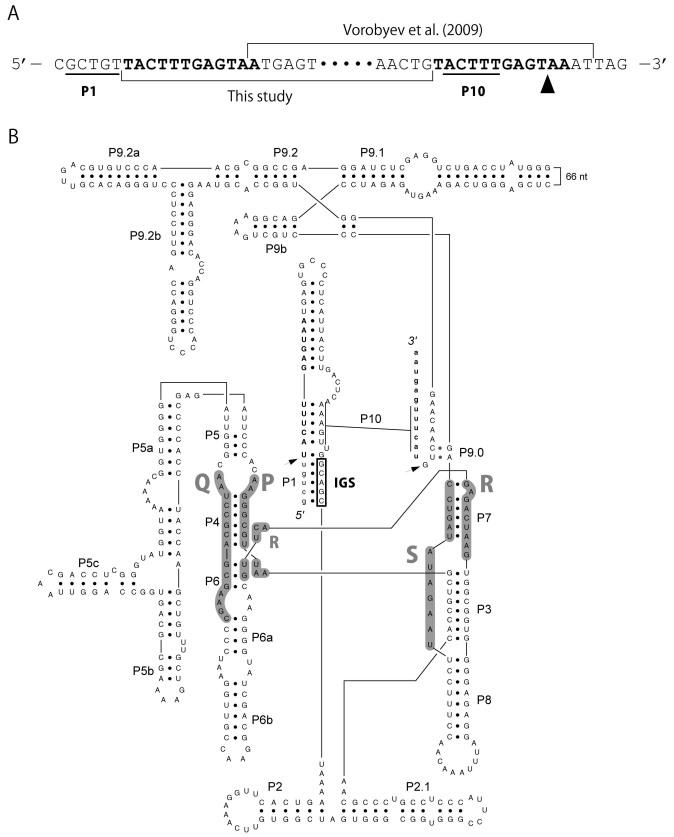

Fig. 1. Flanking sequences and secondary structure prediction of the *Chlorella* sp. T-24-5 SSU rDNA group I intron. A. Differences in intron annotations between Vorobyev et al. (1999) and this study. The replicated sequence —TACTTTGAGTAA— is in bold. P1 and P10 are pairing exon segments constructing helices P1 and P10 in the secondary structure. The arrowhead indicates the S651 insertion position. B. Structure diagram of the intron. Capital letters represent the sequence of the intron; lower case letters indicate flanking exon sequences. Conserved sequence elements P, Q, R and S are shaded. Arrows point to the 5' and 3' splice sites.

ly appended to the core.

Phylogenetic analyses

Intron sequences were aligned through juxtaposition, taking into account the secondary structures. Although some unalignable regions remained between ingroup and outgroup (fungal L2449 introns) taxa, this study emphasized the alignment accuracy between ingroup taxa. A total of 154 aligned positions (138 parsimony-informative) were selected for IC intron analysis. These selected positions were limited conserved base paring elements P3 to P7 and the comparatively conserved base paring elements (determined in this study) of P.2.1, P5a, P5b, P8, P9, and P9.1. The sequence alignments are available from the author upon request.

Two phylogenetic trees were constructed using the p-distance neighbor-joining (NJ) method in MEGA 5 (Tamura et al. 2011) and the maximum likelihood (ML) method in PAUP 4.0b10 (Sinauer Associates, MA). Based on the Akaike's Information Criterion, the best-fit evolutionary model for ML analysis was determined via Modeltest 3.7 (Posada and Crandall, 1998), which selected the TrN + I + G evolutionary model with the following parameters: substitution-rate matrix of AC = 1, AG = 2.5291, AT = 1, CG = 1, CT = 4.0889, and GT = 1; proportion of sites assumed to be invariable = 0.0522; rates for variable sites assumed to follow a gamma distribution with shape parameter = 0.7832; and number of rate categories = 4. With these settings, a heuristic search was performed using the neighbor-joining tree as the starting tree and a nearest-neighbor interchange swapping algorithm. Bootstrap probabilities were

Fig. 2. Comparisons of *Chlorella* sp. T-24-5 sequence with *C. variabilis* and *M. reisseri* under the intron annotation presented here. The nucleotide number corresponds to the sequence coordinate of *Chlorella* sp. T-24-5 (EU 281549). Nucleotides between 201 and 710 are omitted. Intron sequence of *M. reisseri* is excluded. Differences from the uppermost variant are shown by nucleotide, identical nucleotides by a dot, and a dash donates a missing nucleotides. The arrows indicate start and end points of the intron (SSU rDNA: <1..172,726..>857, Intron: 173..725).

computed for 1000 (NJ) and 100 (ML) replicates with these settings.

RESULTS AND DISCUSSION

A large insertion in a nuclear rRNA gene is generally categorized as a group I intron. The intron sequence annotated by Vorobyev et al. (2009) deviates from the fundamental rules for group I introns; that is, the last exon nucleotide (insertion point) is T (U) and the last intron nucleotide is G (Burke et al., 1987; Lambowitz and Belfort, 1993) (Fig. 1A). Upon closer inspection of the insertion, the same twelve-nucleotide sequences, —TACTTTGAGTAA—, was found flanking the insertion. This sequence repeat did not follow the correct intron annotation by Vorobyev et al. (2009). An annotation pattern was highlighted compared to some *Chlorella* rDNA sequences. It is possible that the intron started with —TACTTTGAGTAA— and ended before the second —TACTTTGAGTAA— (intron sequence length of 553 nucleotides) since this annotation follows the fundamental rules for group I introns, as discussed above (Fig. 1A). Under this annotation, there was no gap when compared to other chlorellacean exon sequences (Fig. 2). The exon sequence (having removed the first four nucleotides [nucleotide numbers 1–4] and the last nucleotide [857]), which is 299 nucleotides in length, differed from some *Chlorella* (including *C. variabilis*) and *Micractinium* species by one transition, and differed from *M. reisseri* by one transition and one transversion.

Figure 1B shows the secondary structure diagram of the intron. It was constructed by tracing the well-known core region (as a ribozyme) and adding physically folded peripheral loops. The structure contained base paring helices P1–P10. The extended P5 helix is typical for IC introns. The P1 helix was constructed with a 5' flanking exon sequence to compensate for the internal guide sequence (IGS). Similarly, the P10 helix was constructed with a 3' flanking exon sequence. These helices play an important role in excising the intron from rRNA (Cech, 1990; Suh et al., 1999). These results demonstrated that the above annotation is more reasonable. Therefore, the presumable insertion position is ten nucleotides upstream from that of *M. reisseri* (S651) (Fig. 1A). This insertion position has never been reported as an intron insertion site, and this study tentatively proposes the position to be S641, compared to *Escherichia coli* rRNA (the numbering reflects their homologous

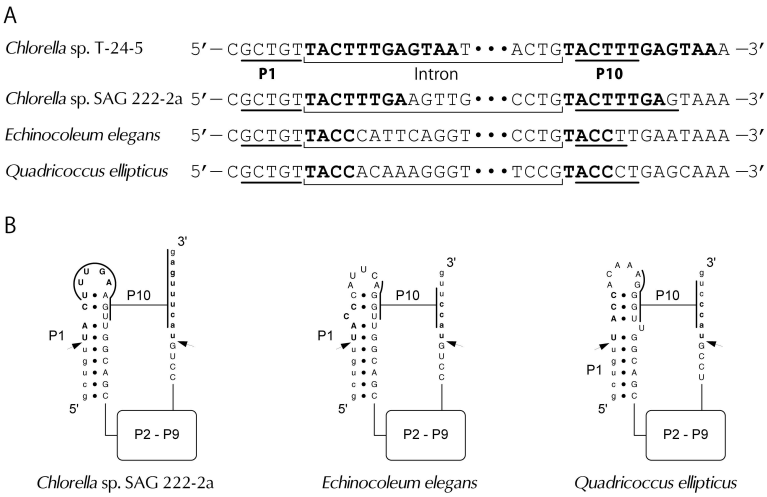

Fig. 3. Annotations for introns similar to the *Chlorella* sp. T-24-5 S641 intron. A. Annotation results considering the fundamental rules for group I introns (see text) and secondary structures. Sequence repeats lying at the head of the intron and after the insertion are in bold. B. Structure predictions for the introns; P1 and P10 pairing segments are only shown. Capital letters represent the sequence of the intron; lower case letters indicate flanking exon sequences. Arrows point to the 5' and 3' splice sites.

positions in the *E. coli* rRNA gene; S = SSU rRNA). The intron has characteristics of an extended, degenerative P9.1 helix and a long P9.2b (P9.2b is the original name in this study). The area similar to the viral intron (Vorobyev et al., 2009) corresponds to helices P4 and P5, although it does not have a structurally distinctive feature.

Surprisingly, a BLAST search identified trebouxiophytes with similar introns, although Vorobyev et al. (2009) did not identify any significantly similar introns. These sequences have been published recently (Luo et al., 2010; Pažoutová et al., 2010; Krienitz and Bock, 2011). This study examined these sequences; the annotation results are shown in Figs. 3A and B. The introns were inserted in the same position (S641), and the repeat sequences were commonly found at the intron heads and 3' flanking exon sequences, although the lengths of the repeat sequences varied.

The phylogenetic relationships of IC introns were determined via NJ and ML analyses. When multiple parameters are expected to estimate highly associated variances for short alignment data (i.e. short and highly divergent sequences like this intron alignment), the parameter-rich ML analyses may give incorrect topologies (Nei et al., 1998). The "simple" models such as the single parameter NJ method are potentially useful for these cases (Bruno and Halpern, 1999; Piontkivska, 2004). Most of previous works used NJ method to construct the phylogenetic tree of introns (Bhattacharya et al., 2005; Haugen et al., 2004b; Nikoh and Fukatsu, 2001). By contrast, some literatures indicated the superiority of ML method with appropriate parameters (Holder and Lewis, 2003; Som, 2009). The present study, therefore, shows both simple NJ and parameter-rich ML trees (Figs. 4A and B). The trees were rooted with fungal L2449 introns that have been indicated the introns with properties intermediate between IC and IE (Hoshina and Imamura, 2009b). Topologies of NJ and ML trees were somewhat different from each other. In NJ, IC introns were diverged into S516 introns of pelagophytes and rhodophytes and the others (Clade A). ML tree indicated fugal and green algal S1506 introns diverged one by one, and the others (Clade B) involved the S516 introns. These S1506 introns were scattered in NJ tree. Bootstrap supports for the branches in Clade A (NJ) or Clade B (ML) were basically low. Clade C was indicated by both analyses, but bootstrap values of ML was below 50%.

Subgroup IC intron analyses have been advanced with focus on fungal rDNA introns. In fun-

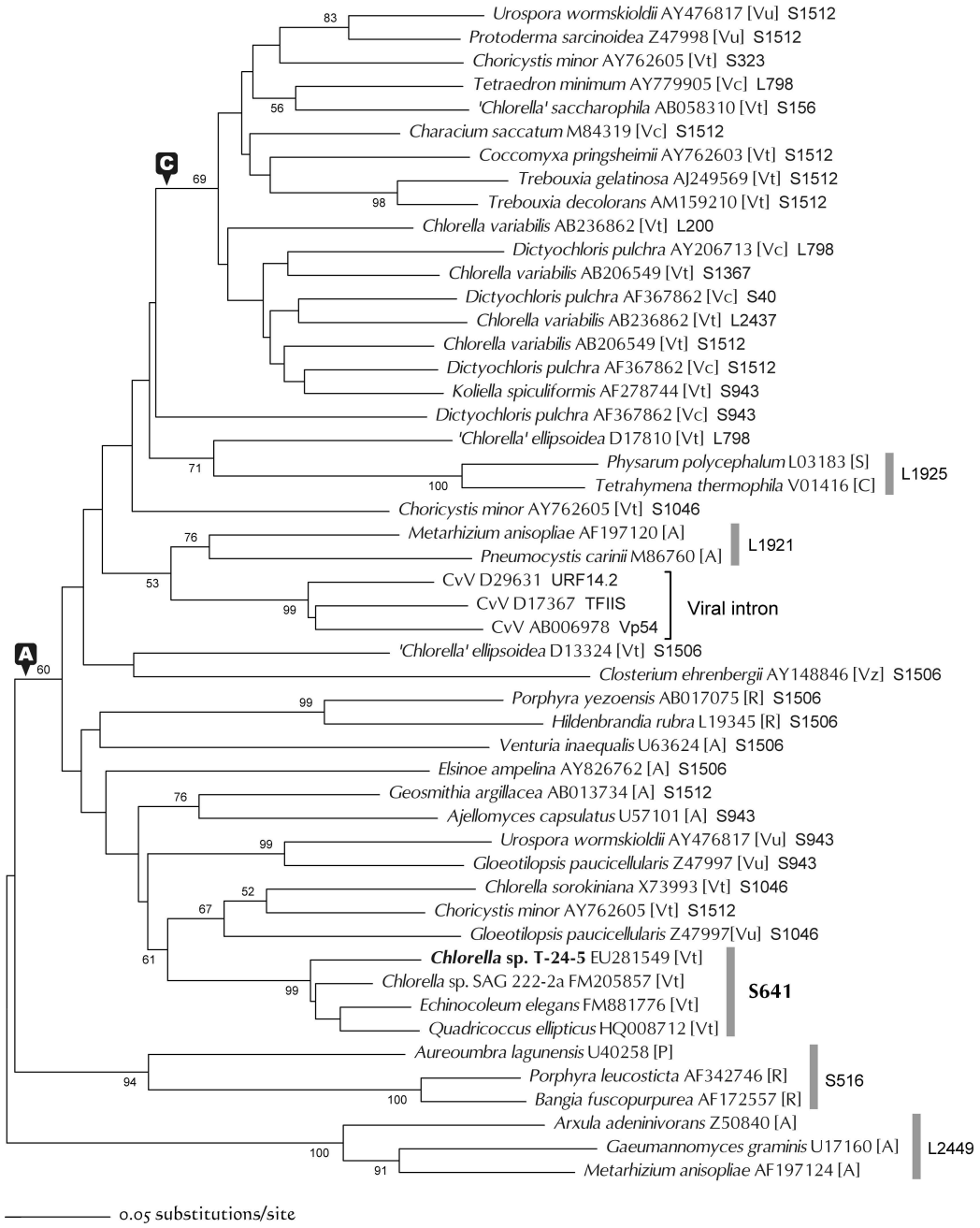

Fig. 4. Phylogenetic trees of selected IC introns. The numerals at each node are bootstrap probabilities of NJ/ME (above the node) and ML/MP (below the node) analyses; only values with >50% support are shown. Three clades are named A, B and C (see text). The letters in brackets indicate taxonomic affiliations: Vc, Chlorophyceae; Vt, Trebouxiophyceae; Vu, Ulvophyceae; Vz, Zygnemophyceae; A, ascomycetes; C, ciliates; P, pelagophytes; R, rhodophytes; S, plasmodial slime molds. Inserted positions are given on the right side of thick vertical bars or after brackets. A. P-distance neighbor-joining (NJ) tree. B. Maximum likelihood (ML) tree using the TrN + I + G evolutionary model.

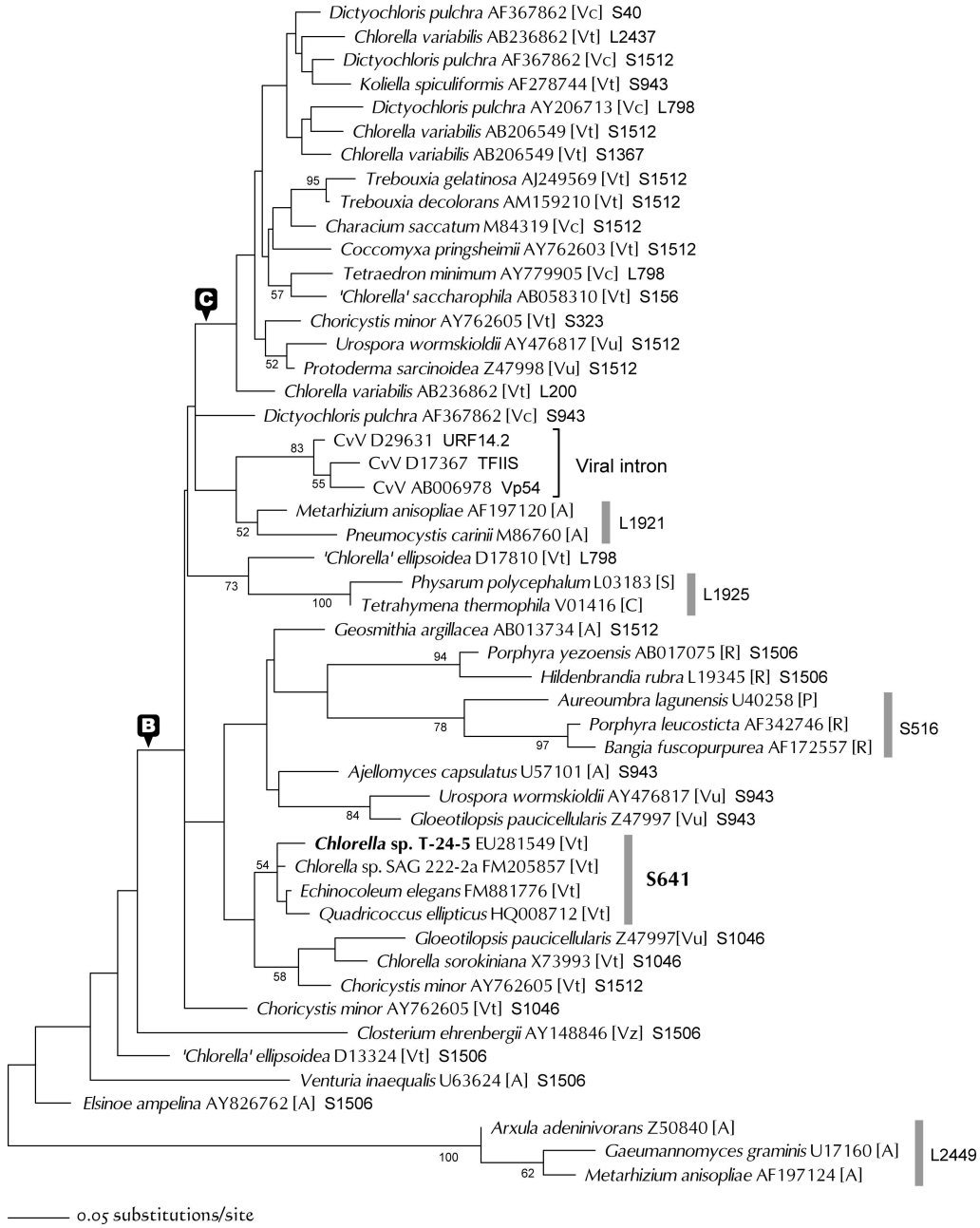

Fig. 4—*continued*.

gal intron phylogeny, introns are separated into intron group that share an insertion site (position family), although relationships of those groups are unclear (e.g. Bhattacharya et al. 2005). There may be only one phylogenetic study covering comprehensive green algal IC introns, where defined groups were hardly indicated (Hoshina and Imamura, 2008b). The present study also could not resolve this problem. Perhaps, irregular evolutionary events have occurred among the IC introns

such as transpositional intron copy within a genome seen in subgroup IE introns (Hoshina and Imamura, 2009b). Indeed, some mildly related introns of single species (*C. variabilis* and *Dictyochloris pulchra* Deason and Herndon) can be seen in Clade C.

Although phylogenetic relationships of the introns in Clade A (Fig. 4A) or Clade B (Fig. 4B) are unclear, the monophyly of the introns sharing S641 insertion site are supported with higher or moderate bootstrap values. Whereas, S641 introns have a tenuous connection with viral introns, although the alignment data set used in the phylogenetic analyses includes regions P4, P5, and 5' side sequence of P5a and P5b. Also, S641 introns and the introns of *C. variabilis* were not phylogenetically related.

With the genome analyses by Blanc et al. (2010), *Chlorella variabilis* (46.2 Mb) became the first genome sequenced species in the Trebouxiophyceae. This study revealed surprising observations with respect to genes involved in cell wall metabolism. The *Chlorella* cell wall contains chitin and chitosan. These genes were derived from CvV and are not related to those of higher plants (Blanc et al., 2010). Thus, it is not surprising if the viral intron was transferred into *Chlorella* genome.

The P4 and P5a, b and c helices of the *Chlorella* sp. T-24-5 intron are similar to those of CvV. Approximately 70% of the sequences are identical to their P5a, b and c helices (comparatively variable region). However, the monophyletic relationship between them could not be determined from this study (Figs. 4A and B). To determine whether CvV truly mediates intron transfer requires further investigation.

This study identified more reasonable insertion position (novel insertion site S641) of *Chlorella* sp. T-24-5 intron (Fig. 1A), and the intron constructed a monophyletic clade with those same insertion positions (Figs. 3A, 3B, 4A and 4B). Such a close relationship between introns inserted at the same site is common (known as a position family), and is the result of intron spreading mechanisms. Introns may spread via homing or reverse splicing. In homing, the intron encodes a homing endonuclease (HE) gene at the terminus of a peripheral helix. HE recognises specific sequences of double-stranded DNA, 14–40 bp in length, and then cleaves the target site and inserts the intron. Reverse splicing is the inverse process of intron splicing at the RNA level. This pathway also requires between four and six specific nucleotides upstream of the insertion site, which the intron recognizes as the optimal site for self-insertion. Therefore, group I intron movement typically occurs at homologous sites due to these sequence recognition mechanisms. Consequently, the number of insertion sites is fairly limited (it is exceptional that the insertion sites in green algal IC introns are diverse), and introns at homologous sites are often more closely related than introns at different sites (for a review, see Haugen et al., 2005). Namely, the S641 position family can be explained by either homing or reverse splicing. HE genes were not found in these S641 introns. The introns that include the HE gene are a tiny minority, which were once fixed in the population, the HE genes no longer have active functions and were ultimately lost (Goddard and Burt, 1999; Haugen et al., 2004a). Instead, long sequence insertions (>50 nucleotides) at peripheral helices, as seen in P9.1 of the *Chlorella* sp. T-24-5 intron, can be regarded as HE gene remnants (Haugen et al., 2004a).

Meanwhile, the origin of the S641 introns remains an important question. Members of the S641 position family exhibit extremely strong sequence conservation. In the core regions of P, Q, R and S, and pairing region P3, the S641 introns had 67 invariants and two 75% conserved positions out of 69 juxtaposed positions (data are not shown). This indicates that these introns have undergone a recent, rapid spread (see Wikmark et al., 2007). The first invasion to position S641 may relate to the sequence repeat, as shown in Figs. 1A and B.

Specific twelve-nucleotide-long sequences emerge stochastically once per 16 Mb (4^{12}). This is unlikely to be coincidental, and may play an as yet unidentified role in intron invasion. This study could not determine whether *P. bursaria* underwent intron spread or invasion to S641.

ACKNOWLEDGEMENTS

This study was supported by an Environmental Research Grant (103180) from The Sumitomo Foundation.

REFERENCES

Bhattacharya, D., Friedl, T. and Damberger, S. (1996) Nuclear-encoded rDNA group I introns: origin and phylogenetic relationships of insertion site lineages in the green algae. Mol. Biol. Evol., 13, 978–989.

Bhattacharya, D., Reeb, V., Simon, D. M. and Lutzoni, F. (2005) Phylogenetic analyses suggest reverse splicing spread of group I introns in fungal ribosomal DNA. BMC Evol. Biol., 5, 68.

Blanc, G., Duncan, G., Agarkova, I., Borodovsky, M., Gurnon, J., Kuo, A., Lindquist, E., Lucas, S., Pangilinan, J., Polle, J., Salamov, A., Terry, A., Yamada, T., Dunigan, D. D., Grigoriev, I. V., Claverie, J. M. and Van Etten, J. L. (2010) The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. Plant Cell, 22, 2943–2955.

Bruno, W. J. and Halpern, A. L. (1999) Topological bias and inconsistency of maximum likelihood using wrong models. Mol. Biol. Evol., 16, 564–546.

Burke, J. M., Belfort, M., Cech, T. R., Davies, R. W., Schweyen, R. J., Shub, D. A., Szostak, J. W. and Tabak, H. F. (1987) Structural conventions for group I introns. Nucleic Acids Res., 15, 7217–7221.

Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Müller, K. M., Pande, N., Shang, Z., Yu, N. and Gutell, R. R. (2002) The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics, 3, 2.

Cech, T. R. (1990) Self-splicing of group I introns. Annu. Rev. Biochem., 59, 543–568.

Cech, T. R. (2002) Ribozymes, the first 20 years (Ribozyme mechanisms and folding). Biochem. Soc. Trans., 30, 1162–1166.

Friedl, T., Besendahl, A., Pfeiffer, P. and Bhattacharya, D. (2000) The distribution of group I introns in lichen algae suggests that lichenization facilitates intron lateral transfer. Mol. Phylogenet. Evol., 14, 342–352.

Gaponova, I. N., Andronov, E. E., Migunova, A. V., Vorobyev, K. P., Chizhevskaja, E. P. and Kvitko, K. V. (2007) Genomic dactyloscopy of *Chlorella* sp., symbionts of *Paramecium bursaria*. Protistology, 4, 311–317.

Goddard, M. R. and Burt, A. (1999) Recurrent invasion and extinction of a selfish gene. Proc. Natl. Acad. Sci. USA, 96, 13880–13885.

Haugen, P., De Jonckheere, J. F. and Johansen, S. (2002) Characterization of the self-splicing products of two complex *Naegleria* LSU rDNA group I introns containing homing endonuclease genes. Eur. J. Biochem., 269, 1641–1649.

Haugen, P., Reeb, V., Lutzoni, F. and Bhattacharya, D. (2004a) The evolution of homing endonuclease genes and group I introns in nuclear rDNA. Mol. Biol. Evol., 21, 129–140.

Haugen, P., Runge, H. and Bhattacharya, D. (2004b). Long-term evolution of the S788 fungal nuclear small subunit rRNA group I introns. RNA, 10, 1084–1096.

Haugen, P., Simon, D. M. and Bhattacharya, D.

(2005) The natural history of group I introns. Trends Genet., 21, 111–119.

Holder, M. and Lewis, P. O. (2003). Phylogeny estimation: traditional and Bayesian approaches. Nat. Rev. Genet., 4, 275–284.

Hoshina, R. and Imamura, N. (2008a) Multiple origins of the symbioses in *Paramecium bursaria*. Protist, 159, 53–63.

Hoshina, R. and Imamura, N. (2008b) Eu-*Chlorella* large subunit rDNA sequences and group I introns in ribosomal DNA of the paramecian symbiotic alga NC64A. Phycol. Res., 56, 21–32.

Hoshina, R. and Imamura, N. (2009a) Origins of algal symbionts of *Paramecium bursaria*. *In:* Endosymbionts in *Paramecium*. Fujishima, M. (ed.). Springer-Verlag, Berlin Heidelberg, pp. 1–29.

Hoshina, R. and Imamura, N. (2009b) Phylogenetically close group I introns with different positions among *Paramecium bursaria* photobionts imply a primitive stage of intron diversification. Mol. Biol. Evol., 26, 1309–1319.

Hoshina, R., Iwataki, M. and Imamura, N. (2010) *Chlorella variabilis* and *Micractinium reisseri* sp. nov. (Chlorellaceae, Trebouxiophyceae): redescription of the endosymbiotic green algae of *Paramecium bursaria* (Peniculia, Oligohymenophorea) in the 120th year. Phycol. Res., 58, 188–201.

Hoshina, R., Kamako, S. and Imamura, N. (2004) Phylogenetic position of endosymbiotic green algae in *Paramecium bursaria* Ehrenberg from Japan. Plant Biol., 6, 447–453.

Hoshina, R., Kato, Y., Kamako, S. and Imamura, N. (2005) Genetic evidence of "American" and "European" type symbiotic algae of *Paramecium bursaria* Ehrenberg. Plant Biol., 7, 526–532.

Kodama, Y. and Fujishima, M. (2009) Infection of *Paramecium bursaria* by symbiotic *Chlorella* species. *In:* Endosymbionts in *Paramecium*. Fujishima, M. (ed.). Springer-Verlag, Berlin Heidelberg, pp. 31–55.

Krienitz, L. and Bock, C. (2011) *Elongatocystis ecballocystiformis* gen. et comb. nov., and some reflections on systematics of Oocystaceae (Trebouxiophyceae, Chlorophyta). Fottea, 11, 271–278.

Kvitko, K., Migunova, A. and Karelov, D. (2001) Molecular taxonomy of virus sensitive *Chlorella* sp. – symbionts of *Paramecium bursaria*. Protistology, 2, 96–104.

Lambowitz, A.M. and Belfort, M. (1993) Introns as mobile genetic elements. Annu. Rev. Biochem., 62, 587–622.

Lehnert, V., Jaeger, L., Michel, F. and Westhof, E. (1996) New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. Chem. Biol., 3, 993–1009.

Linz, B., Linz, A., Migunova, A. and Kvitko, K. (1999) Correlation between virus-sensitivity and isoenzyme spectrum in symbiotic *Chlorella*-like algae. Protistology, 1, 76–81.

Luo, W., Pröschold, T., Bock, C. and Krienitz, L. (2010) Generic concept in *Chlorella*-related coccoid green algae (Chlorophyta, Trebouxiophyceae). Plant Biol., 12, 545–553.

Mathews, D. H., Sabina, J., Zuker, M. and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol., 288, 911–940.

Nei, M., Kumar, S. and Takahashi, K. (1998) The optimization principle in phylogenetic analysis tends to give incorrect topologies when the number of nucleotides or amino acids used is small. Proc. Natl. Acad. Sci. USA, 95, 12390–12397.

Nielsen, H. and Johansen, S. D. (2009) Group I introns: Moving in new directions. RNA Biol., 6, 375–383.

Nikoh, N. and Fukatsu, T. (2001) Evolutionary dynamics of multiple group I introns in nuclear ribosomal RNA genes of endoparasitic fungi of the genus *Cordyceps*. Mol. Biol. Evol., 18, 1631–1642.

Nishida, K., Suzuki, S., Kimura, Y., Nomura, N., Fujie, M. and Yamada, T. (1998) Group I introns found in *Chlorella* viruses: biological implications. Virology, 242, 319–326.

Pažoutová, M., Škaloud, S. and Nemjová, K. (2010) Phylogenetic position of *Ooplanctella planoconvexa*, gen. et comb. nova and *Echinocoleum elegans* (Oocystaceae, Trebouxiophyceae, Chlorophyta). Fottea, 10, 75–82.

Posada, D. and Crandall, K. A. (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics, 14, 817–818.

Piontkivska, H. (2004) Efficiencies of maximum likelihood methods of phylogenetic inferences when different substitution models are used. Mol. Phylogenet. Evol., 31, 865–873.

Pröschold, T., Darienko, T., Silva, P. C., Reisser, W. and Krienitz, L. (2011) The systematics of *Zoochlorella* revisited employing an integrative approach. Environ. Microbiol., 13, 350–364.

Som, A. (2009). ML or NJ-MCL? A comparison between two robust phylogenetic methods. Comp. Biol. Chem., 33, 373–378.

Suh, S. O., Jones, K. G. and Blackwell, M. (1999) A Group I intron in the nuclear small subunit rRNA gene of *Cryptendoxyla hypophloia*, an ascomycetous fungus: evidence for a new major class of group I introns. J. Mol. Evol., 48, 493–500.

Summerer, M., Sonntag, B. and Sommaruga, R. (2008) Ciliate-symbiont specificity of freshwater endosymbiotic *Chlorella* (Trebouxiophyceae, Chlorophyta). J. Phycol., 44, 77–84.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol., 28, 2731–2739.

Vorobyev, K., Andronov, E., Rautian, M., Skoblo, I., Migunova, A. and Kvitko, K. (2009) An atypical *Chlorella* symbiont from *Paramecium bursaria*. Protistology, 6, 39–44.

Wikmark, O.-G., Haugen, P., Lundblad, E. W., Kaugli, K. and Johansen, S. D. (2007) The molecular evolution and structural organization of group I introns at position S1389 in nuclear small subunit rDNA of myxomycetes. J. Eukaryot. Microbiol., 54, 49–56.

Yamada, T., Tamura, K., Aimi, T. and Songsri, P. (1994) Self-splicing group I introns in eukaryotic viruses. Nucleic Acids Res., 22, 2532–2537.

Zhou, Y., Lu, C., Wu, Q.-J., Wang, Y., Sun, Z.-T., Deng, J.-C. and Zhang, Y. (2008) GISSD: Group I Intron Sequence and Structure Database. Nucleic Acid Res., 36, D31–D37.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res., 31, 3406–3415.