Slow evolution of 1506 group I intron in *Spirogyra* Link 1820 (Zygnematophyceae, Streptophyta), a fast evolving lineage in the Zygnemataceae

Charlotte Chen & Michael Schagerl*

*University of Vienna, Department Limnology, Althanstraße 14, A–1090 Vienna, Austria; * Corresponding author e–mail: michael.schagerl@univie.ac.at*

Abstract: Phylogenetic analyses of SSU rDNA sequences of 130 *Spirogyra* strains have revealed that these strains were subdivided into eight clades. Approximately 60% of the assessed strains (clades A–D) contain a 1506 group I intron, whereas strains without introns form individual clades (E–H). The *Spirogyra* intron shared the common insertion site of the Zygnematalean intron (position 1506 relative to the *Escherichia coli* small–subunit rRNA). Phylogenetic analyses of the *Spirogyra* group I intron showed the monophyletic origin within the Zygnematophyceae. Therefore, we assume the secondary loss of the intron in clades E–H is caused by the high evolutionary rate of *Spirogyra* and its long evolutionary history. The *Spirogyra* intron belongs to the IC group I introns and shares many common features with the intron of other Zygnematophyceae (the typical domain structure (P1–P9), the base composition, the highly conserved regions the U preceding the 5' splice site and the G to which it pairs, and the G preceding the 3' splice site are typical for IC group I intron). *Spirogyra* group I introns exhibit features of early desmids (optional P2 domain) as well as of later diverging desmids (variation from the typical L5b–GAAA tetraloop). The P2 domain shows an additional optional P2 sub–domain in clade B. Surprisingly, the mutation rate of the *Spirogyra* SSU rRNA exceeds the rate of the intron by far. Evolutionary rates differ significantly within the *Spirogyra* SSU rRNA accessions, but not within the respective group I intron sequences.

Key words: *Spirogyra*, group I intron, intron subgroup IC1, optional P2 domain, secondary structure, evolutionary rate

Abbreviations: (nt) nucleotide; (rDNA) ribosomal DNA; (SSU) small subunit; (A) adenine; (G) guanine; (C) cytidine; (U/T) uridine/ thymine; (MP) maximum parsimony; (ML) maximum likelihood; (NJ) neighbor joining; (PP) posterior probability; (Meso) Mesotaeniceae; (Zygn) Zygnemetaceae; (Desm) Desmidiaceae; (NHS) Non–Homoplasious Synapomorphy; [(H–) CBC] (Hemi–) Compensatory Base Change

Introduction

The genus *Spirogyra* Link (Zygnematophyceae, Streptophyta) is characterized by unbranched filaments, coiled chloroplasts and a special type of reproduction (conjugation) with complete absence of flagellated reproductive stages (Kolkwitz & Krieger 1941; Transeau 1951; Randhawa 1959; Kadlubowska 1984; McCourt et al. 1986). The filaments are covered by a layer of mucilage, which is responsible for the slimy appearance (Hoshaw & McCourt 1988). The genus is distributed worldwide in fresh to slightly brackish water bodies (Rieth 1983; Hoshaw & McCourt 1988; Simons & Van Beem 1990). *Spirogyra* covers a wide ecological range from dystrophic mountain lakes and bogs to highly eutrophicated systems (Hainz et al. 2009).

The term "intron" was introduced by Gilbert (1978) to describe intragenic regions that occur alternating with coding regions and that would subsequently be lost from the mature messenger RNA. Introns are of interest because they are used to gain insight into the evolution of protein synthesis and how primitive RNAs could have catalyzed their own replication (Saldanha et al. 1993). Beside group I introns, group II intron and spliceosomal introns exist, which are defined by their way of splicing (self–splicing vs. RNA operated splicing) and have a distinct secondary structure (Saldanha et al. 1993). Group I introns are an important class of RNA enzymes that also exhibit a typical secondary structure that is expressed in the sub–domains P1–P9 and the conserved core regions P, Q, R and S (BURKE et al. 1987; CECH 1988). Five major subgroups of group I introns (IA, IB, IC, ID and IE) and 10 minor subgroups can be distinguished (CANNONE et al. 2002). The first rDNA group I Intron was discovered in the LSU rRNA of the protist *Tetrahymena thermophila* (ZAUG et al. 1983). Since then, group I introns have been found in rRNAs, mitochondrial and chloroplast genomes in a great variety of organisms from fungi (MICHEL & DUJON 1983) to many algal groups (DAMBERGER & GUTELL 1994; BESENDAHL & BHATTACHARYA 1999; BHATTACHARYA et al. 2001; MULLER et al. 2001). So far, no function could be linked to group I introns nor do they affect the phenotype (DUJON 1989). Two theories about their manner of distribution have been formulated: vertical inheritance together with the gene they interrupt or, horizontal transfer by an unknown vector between individuals of different species, genera or families (BHATTACHARYA 1998). A common ancestry of all group I introns is likely but still remains obscure (OLIVEIRA & RAGAN 1994; BHATTACHARYA 1998). Group I introns are typically capable of self–splicing, but they have lost this ability within the Desmidiaceae (BESENDAHL & BHATTACHARYA 1999). *Spirogyra* is a very species–rich genus within the Zygnematales with an estimated species number of 386 (KADLUBOWSKA 1984). Older monographs had somewhat lower numbers between 275 (TRANSEAU 1951) and 289 (RANDHAWA 1959). Even though the current species number is questioned (HOSHAW et al. 1985), there are only few studies with adequate taxon sampling (DRUMMOND et al. 2005) that give more insight into genetic diversity in this genus. The group I intron of the genus *Spirogyra* has not yet been described in detail; Gontcharov who first discovered the *Spirogyra* group I intron, only made a brief characterization of it as a typical zygnematophycean intron, solely based on three intron sequences (GONTCHAROV et al. 2002).

Phylogenetic analyses of SSU rDNA sequences (GONTCHAROV et al. 2002; GONTCHAROV et al. 2004; GONTCHAROV & MELKONIAN 2004) have revealed that *Spirogyra* represents a so–called "long branch" taxon sensu Philippe (2000). We wanted to investigate if the long branch attraction phenomenon (LBA) is also present in a group I intron phylogeny and if evolutionary rates of the *Spirogyra* intron is comparable to the SSU rRNA region. In this study, we provide a detailed view of the characteristics of this RNA intron in the genus *Spirogyra* and on the variability of a group I intron within a single genus. Furthermore, the phylogenetic position of the *Spirogyra* intron is evaluated and its secondary structure is compared to known group I introns of other Zygnematalean taxa to test if the intron has taxonomical value as suggested by GONTCHAROV et al. (2002).

Material and Methods

Algal Material, DNA extraction, amplification and sequencing

Isolation of algal material, DNA extraction, amplification and sequencing were carried out as described in (CHEN et al. 2012). Four additional primers were used: 1380F (GCGTTGAWTACGTCCCTGCC), BR (TTGATCCTTCTGCAGGTTCACCTAC), ITS–GF (GGGATCCGTTTCCGTAGGTGAACCTGC) and 18SR3 (TGTTACGACTTCTCCTTCCTCTAAACG).

Sequence Alignment and Analysis

The group I intron sequences were manually aligned according to the secondary structures of the P, Q, R and S elements to identify homologous regions (MICHEL & WESTHOF 1990; BHATTACHARYA et al. 1994). The alignment was refined by comparison of the secondary structure of the sequences. Secondary structure was determined via the Rensselaer bioinformatics web server using mfold (ZUKER 2003). Secondary structure diagrams were generated with the software XRNA (WEISER & NOLLER 2009). Base composition was calculated from the alignment using BioEdit (HALL 1999). Only unambiguously aligned regions of the sequences were used for analyses, gap–rich regions were excluded. To compare the *Spirogyra* sequences with other Zygnematalean sequences, initial analyses of all *Spirogyra* sequences were performed to detect the variability among the *Spirogyra* strains (data not shown). Three different data sets were further analyzed: (1) For confirmation of the *Spirogyra* clade position among the other algae included, for checking the long branch status, and for demonstration of clade grouping in *Spirogyra*, an analysis of a SSU alignment comprising 12 Desmidiaceae, Mesotaeniaceae and Zygnemataceae sequences was conducted (Table 1, data set 6). Twelve *Spirogyra* sequences selected from the initial analysis were added; two sequences were taken from each clade containing the intron; one of the clades without intron. (2) For a phylogenetic understanding of the intron (Table 1, data set 7) in relation to other Zygnematophyceaen introns, an analysis of a data set comprising the same taxa as in the SSU alignment was performed. 8 *Spirogyra* sequences of the initial analysis from clades containing the intron were added. (3) The group I intron sequence of 51 *Spirogyra* strains (Table

1, data set 8). Identical sequences were excluded from analyses to reduce computing effort.

The alignment of the desmid SSU sequences consisted of 1717 unambiguously aligned bases, the respective intron alignment of 263 unambiguously aligned bases. To avoid LBA (Philippe 2000) and following Gontcharov's argument, that monophyly of the Zygnematophyceae is undoubted, but that its position in the streptophytes is unclear, and therefore no suitable outgroup can be chosen, unrooted phylogenies were used for calculations in this study (Gontcharov et al. 2004). The *Spirogyra* group I intron data set comprised 251 unambiguously aligned bases. To test for the best evolutionary model for the analyses, the log likelihood values of 56 models using Modeltest 3.7 were compared (Posada & Crandall 1998; Posada & Buckley 2004).

The phylogenetic trees presented were inferred by maximum likelihood settings using PAUP* 4.0b10 (Swofford 2003). The best models were chosen according to the Akaike Information Criterion by Modeltest 3.7. For the desmid SSU alignment analyses the GTR+G model was chosen, for the desmid intron alignment analyses the GTR+I+G model and for the *Spirogyra* intron alignment the TrN+I+G model. Bootstrap support values were calculated by distance (neighbor joining (NJ; 1000 replicates), maximum parsimony (MP; 1000 replicates) and maximum likelihood methods (ML; 1000 replicates; using the appropriate evolutionary model for each data set). Bayesian inference was calculated using MrBayes 3.1.2 (using following settings: 1 – 3 million generations; MCMC chains = 4; all trees below the burnin value of 0,01 were excluded as burnin, the majority consensus tree was calculated using PAUP*; (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003)). Details of the evolutionary models and Bayesian analyses are given in the legend of the respective figures. Relative rate tests were carried out between all accessions used for phylogeny in GRate 0.4 (Müller et al. 2004).

Secondary structure analyses

To construct a putative secondary structure for *Spirogyra*, existing secondary structure models from other taxa were obtained from http://www.rna.ccbb.utexas.edu/ (Cannone et al. 2002). Homologous regions (P, Q, R and S (Michel & Westhof 1990; Bhattacharya et al. 1994)) were identified, aligned, and assigned to the respective P domains. For secondary structure calculation, the individual sequence fragments were uploaded to the Rensselaer bioinformatics web server using mfold (Zuker 2003).

Secondary structures were compared among the clades for *Spirogyra*. To find all non–homoplasious synapomorphies (NHS), the method described in Marin, Palm et al. (2003) was applied. With this method NHS, Compensatory Base Change (CBC) and Hemi–CBC (H–CBC; only one base changes, but the base pairing remains intact) can be detected. The two NHS criteria were applied: (1) absence of convergent evolution outside the clade, and (2) strict conservation within the clade.

Results

Phylogenetic analyses

Phylogenetic analyses of SSU rDNA sequences of 130 *Spirogyra* strains have revealed that these strains were subdivided into eight clades. Numerous strains of these clades (clades A – D) contain a 1506 group I intron, whereas strains without introns belong to an individual derived clade (clade E–H, Fig. 1). *Spirogyra* was analyzed together with other Zygnematalean taxa to confirm its position within the Zygnemataceae taxa, and to test whether the same position would be true for the *Spirogyra* group I intron. Seventy two out of 130 *Spirogyra* strains contained a 1506 group I intron. Additionally, base frequencies were compared among the groups and between the SSU and the intron. Accepted secondary structure suggestions of the taxa used for the desmid alignment were compared to putative secondary structures of *Spirogyra* group I introns.

For the desmid 18s SSU phylogeny shown in Fig. 1, a data set of the SSU with a total of 1715 unambiguously aligned positions was used. The data set was analyzed to depict the clade grouping in the genus *Spirogyra* and test the group I introns as molecular marker in comparison to the well established SSU. We were also interested, if the *Spirogyra* clades established in the 18s SSU phylogeny will be reflected in the group I intron phylogeny. The *Spirogyra* group I intron alignment was analyzed separately using 263 unambiguously aligned positions (Fig. 2B). In the tree derived from intron alignment analyses (Fig. 2A), two different lineages are visible in our analyses, the Desmidiaceaen branch and the Mesotaeniaceaen/ Zygnemataceaen lineage; *Closterium littorale* could not be placed in either lineages. The only phylogenetic relation supported by bootstrap and Bayesian values outside the *Spirogyra* clade is the *Gonatozygon/Genicularia* clade with very high support, but is placed in the Mesotaeniaceaen/ Zygnemataceaen clade. All other phylogenetic relationships in these groups remain questionable. Within *Spirogyra*, clade A, B and D receive high bootstrap support. There

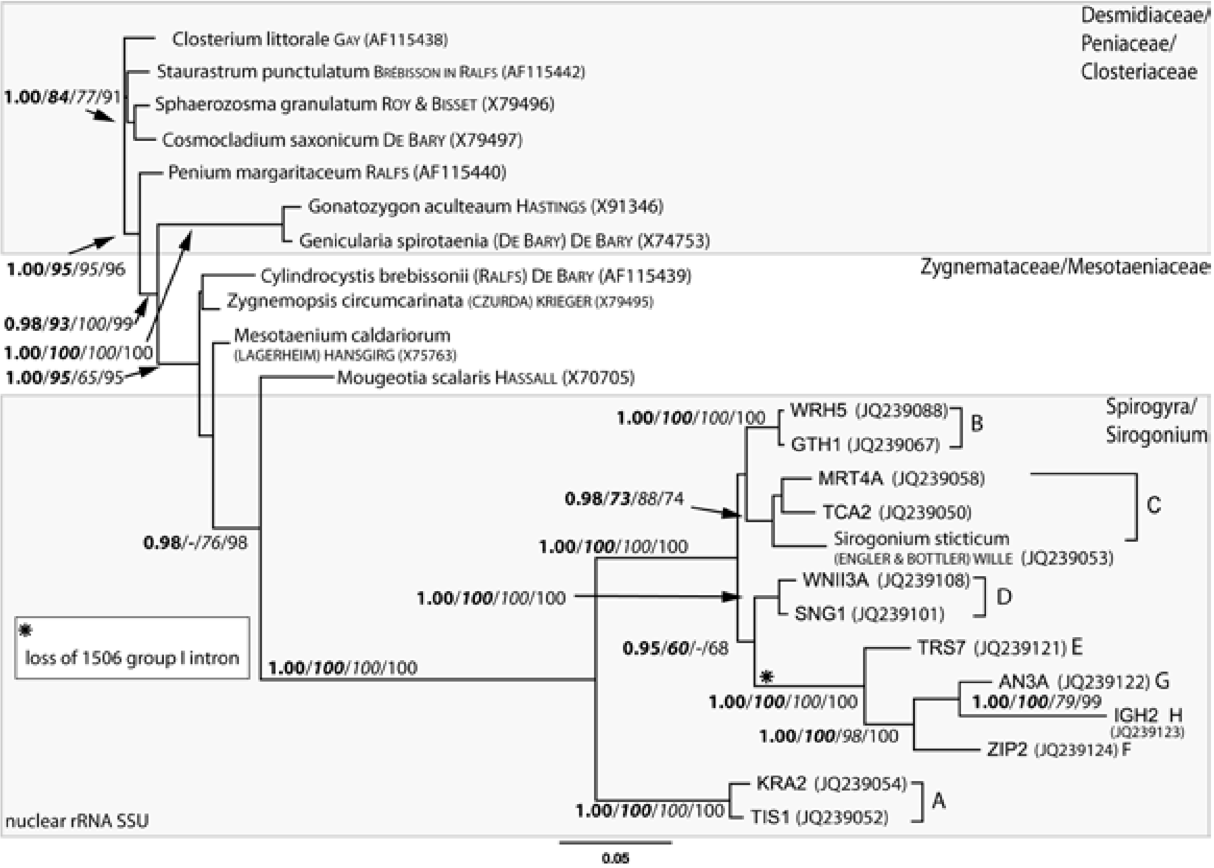

Fig. 1. Desm SSU: molecular phylogeny of Desmidiaceae, Mesotaeniaceae and Zygnemataceae based on a SSU alignment. The phylogenetic tree was inferred by maximum likelihood analyses of 1715 aligned positions of 24 taxa using PAUP* 4.0b10. As an evolutionary model, the GTR+I+G model was chosen (base frequencies: A 0.25, C 0.23, G 0.27, T 0.25; rate matrix: A–C 1. 4387, A–G 2. 3888, A–T 1. 4904, C–G 1. 2962, C–T 5.4133, G–T 1.0000) with the proportion of invariable sites (I=0. 5030) and gamma distribution parameter (G=0.7457). Bootstrap values (>50%) of the maximum likelihood (1000 replicates; bold italic), neighbor–joining (1000 replicates; italic), and maximum likelihood (1000 replicates; not bold) as well as Bayesian values (>95%;bold) were calculated by MrBayes 3.1.2 using the covariation settings (3 million generations, trees from 4100 – 30000) are given in the tree (bayesian/ML/NJ/MP).

is no clade C formed in this tree; the members of this clade are placed on individual branches basal to clade B and D**.** In tree derived from SSU alignment analyses (Fig. 1), more branches receive support by bootstrap and Bayesian values, still not all proposed phylogenetic relationships are reliable. Only two clades of closely related taxa are supported: the *Sphaerozosma*/*Cosmocladium* clade and the *Gonatozygon*/*Genicularia* clade. Unlike in the intron alignment, Desmidiaceae and Mesotaeniaceae/Zygnemataceae are not separated into two lineages. The Mesotaeniaceaen/Zygnemataceaen branch is placed within a larger clade with other taxa in basal positions. Bootstrap support within *Spirogyra* is very high–accessions that show no 1506 group I intron form a well supported clade.

Spirogyra group I introns were analyzed separately (Fig. 2B). 50 individual sequences obtained by the author and *Sirogonium sticticum* were included using 251 unambiguously aligned positions. Seven sequences had to be excluded from analyses because of sequencing problems. The tree was rooted with clade A. Clade B is not supported by all algorithms, but the taxa of clade B still form a branch together in the maximum likelihood tree. The individual branches of Clade C are well supported, but are not located in a collective branch. Clade D is well supported by Bayesian and bootstrap values.

Evolutionary rates

Evolutionary rates were inferred by pair wise comparison of unambiguously aligned positions of the Desmidiaceae/ Mesotaeniaceae/ Zygnemataceae intron (Table 1, data set 7) and the Desmidiaceae/ Mesotaeniaceae/ Zygnemataceae SSU alignment (Table 1, data set 6) in GRate

Table 1. Comparison of modeltest results among the individual data sets; highest and lowest value in rate matrix marked in light grey [(1) all taxa of the SSU rRNA alignment aside *Spirogyra*; (2) *Spirogyra* accessions of the SSU rRNA alignment that contained a group I intron; (3) all *Spirogyra* accessions of the SSU rRNA alignment; (4) all taxa of the group I intron alignment aside *Spirogyra*; (5) all *Spirogyra* accessions of the intron alignment; (6) complete 18s SSU rRNA alignment; (7) complete 1506 group I intron alignment; and (8) *Spirogyra* 1506 group I intron alignment].

	incl.	incl.	pars–inf.	model			base frequencies				Substitution model/rate matrix				
	taxa	char.	char.	selected	I	G	A	C	G	U	A–C	A–G	A–U	C–G	C–U
1	11	1717	147	GTR+I+G	0.6752	0.9038	0.2627	0.2009	0.2619	0.2745	1.2538	2.0361	1.5494	0.6242	6.4562
2	9	1717	195	GTR+I+G	0.6392	0.5762	0.2569	0.2193	0.2676	0.2562	6.0200	8.3107	5.5401	3.2841	20.8136
3	13	1717	266	GTR+I+G	0.5234	0.6118	0.2438	0.2369	0.2789	0.2404	1.8712	3.1395	1.6342	1.9982	6.3806
4	11	263	124	TrN+I+G	0.2727	1.7394	0.2286	0.2655	0.2955	0.2103	1.0000	2.6770	1.0000	1.0000	6.5834
5	9	263	91	GTR+G	0.0000	0.2809	0.2392	0.2983	0.3125	0.1500	0.7473	1.4149	1.8670	0.4333	3.2024
6	24	1717	425	GTR+I+G	0.5030	0.7457	0.2499	0.2262	0.2707	0.2531	1.4387	2.3888	1.4904	1.2962	5.4133
7	20	263	168	GTR+G	0.0000	0.6668	0.2163	0.2944	0.3105	0.1788	0.8651	2.2251	1.9833	0.7013	4.6364
8	51	251	108	TrN+I+G	0.1481	0.8150	0.2381	0.2934	0.2627	0.2059	1.0000	1.4160	1.0000	1.0000	1.8766

(Table 2). In both data sets, a significant difference between the evolutionary rate of *Spirogyra* and all other taxa analyzed could be detected. Within the *Spirogyra* intron sequences, many accessions exhibit a rate that is not significantly different (Table 2; highlighted in grey). Whereas in the *Spirogyra* SSU sequences, more significantly different evolutionary rates could be found. When including more taxa to the intron alignment and using not only the unambiguously aligned bases, there is no significant difference between *Spirogyra* and other genera (data not shown). However, the trend in the SSU alignment is the same – different evolutionary rate to the remaining taxa, diversity of evolutionary rates within *Spirogyra*.

Additionally, a modeltest analysis was conducted to compare substitution models for 5 data sets (Table 1), which are (1) all taxa of the SSU rRNA alignment except for *Spirogyra*, (2) *Spirogyra* accessions of the SSU rRNA alignment that contained a group I intron, (3) all *Spirogyra* accessions of the SSU rRNA alignment, (4) all taxa of the group I intron alignment except for *Spirogyra* and (5) all *Spirogyra* accessions of the intron alignment. The Desmidiaceae/ Mesotaeniaceae/ Zygnemataceae SSU alignment (data set 1) contains the largest proportion of invariant sites, the Desmidiaceae/ Mesotaeniaceae/ Zygnemataceae intron data set (data set 4) the smallest. The largest base frequency differences were found between the Desmidiaceae/ Mesotaeniaceae/ Zygnemataceae SSU (data set 1) and the *Spirogyra* intron data set (data set 5). The *Spirogyra* SSU +I data set (data set 2) has the highest base substitution rates; the *Spirogyra* intron data set (data set 5) the lowest (except for the A–U substitution rate). The *Spirogyra* intron data set exhibits the highest average G–C content (0.6108), the Desmidiaceae/ Mesotaeniaceae/ Zygnemataceae SSU the lowest (0.4628).

Base composition

Relative frequencies of the four bases were compared for the complete SSU and the intron (Fig. 3). The base composition differed considerably between the SSU rRNA and the 1506 group I intron. The SSU is relatively U–rich and C–poor; in the intron the opposite was observed. The G–content is higher than the A – content in both SSU and intron. The content of individual bases varies more in the intron than in the SSU eg., G–content 25–34% in introns, in the SSU from 26–29%; the same holds true for the other bases. These were

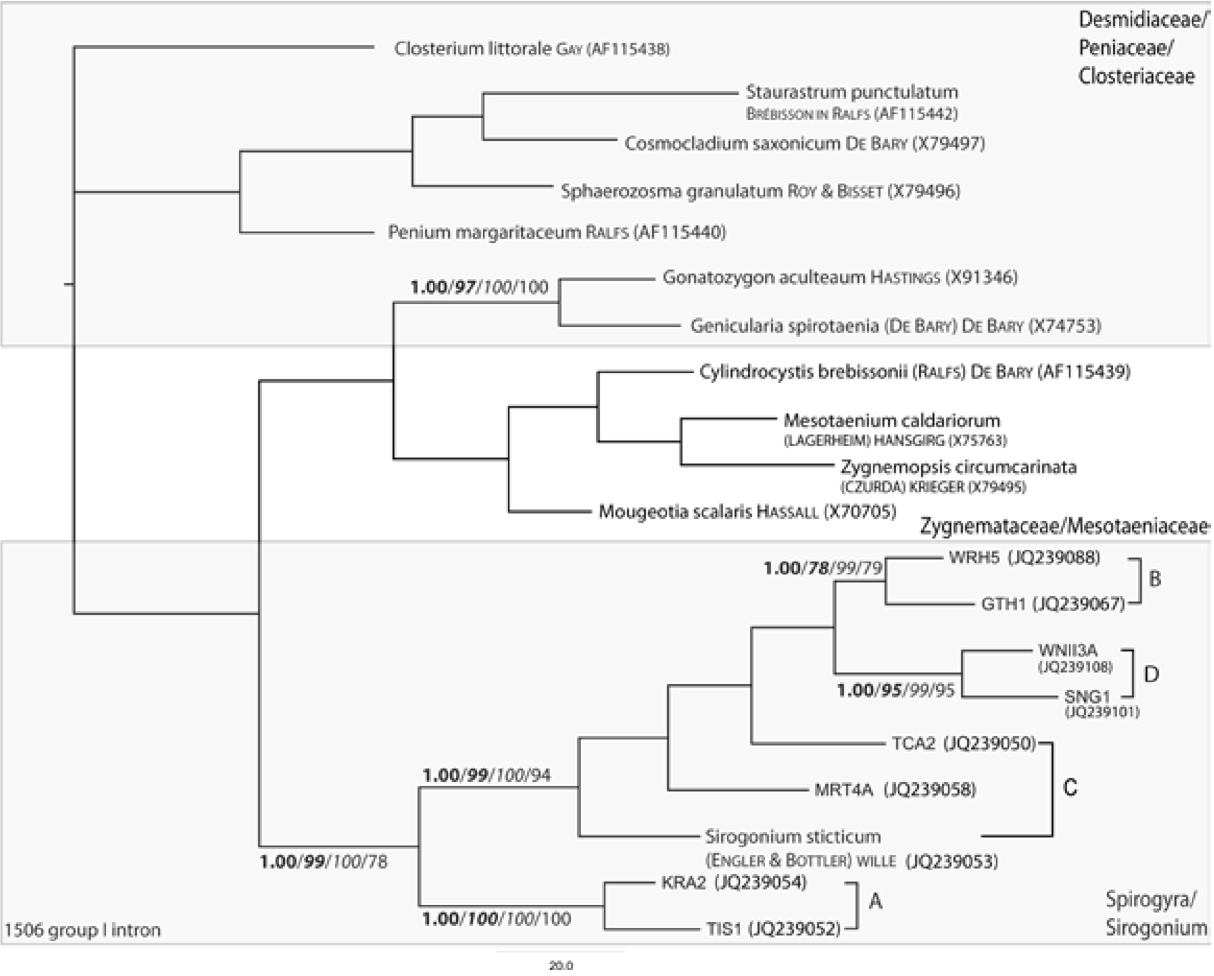

Fig. 2A. Desm Intron: molecular phylogeny of Desmidiaceae, Mesotaeniaceae and Zygnemataceae based on a group I intron alignment. The phylogenetic tree was inferred by maximum likelihood analyses of 263 aligned positions of 20 taxa using PAUP* 4.0b10. As an evolutionary model, the GTR+G model was chosen (base frequencies: A 0.22, C 0.29, G 0.31, T 0.18; rate matrix: A–C 0. 8651, A–G 2.2251, A–T 1. 9833, C–G 0. 7013, C–T 4.6364, G–T 1.0000) with the proportion of invariable sites (I=0) and gamma distribution parameter (G=0. 6668). Bootstrap values (>50%) of the maximum likelihood (1000 replicates; bold italic), neighbor–joining (1000 replicates; italic), and maximum likelihood (1000 replicates; not bold) as well as Bayesian values (>95%;bold) were calculated by MrBayes 3.1.2 using the covariation settings (1 million generations, trees from 52100 – 67300, 77000 – 100000) are given in the tree (bayesian/ML/NJ/MP).

also reflected in the Modeltest settings (see Fig. 1, 2A + B, Table 1). The relative content of U and A are very similar in the 18s SSU, whereas the content of G and C show a relatively big difference. In the intron, G and C show similar relative contents, and the percentage of A and U differ from each other.

Sequence identity of *Spirogyra*

Seventy two sequenced strains exhibited the 1506 group I intron. The total length of both the 18s SSU and the 1506 group I intron were compared individually. Out of these 72 SSU sequences 19 sequences were unique, 13 were found more than once. The group I intron exhibited 50 unique variants and only 8 types were found more than once. All *Spirogyra* strains that had identical intron sequences also had identical SSU sequences – 22 individual accessions forming 6 different groups. 58% of the strains that had identical SSU sequences, differed in intron sequences. In Fig. 2B, sequences that were obtained from isolates from the same sampling site are marked with open grey boxes, sequences from different isolation sites with closed boxes. Not all isolates from the same site had the same intron, but only small differences (up to 5 nt) could be observed.

Secondary structures

Spirogyra group I introns showed the typical secondary structure (Fig. 4). The secondary structure of the *Spirogyra* group I intron exhibits

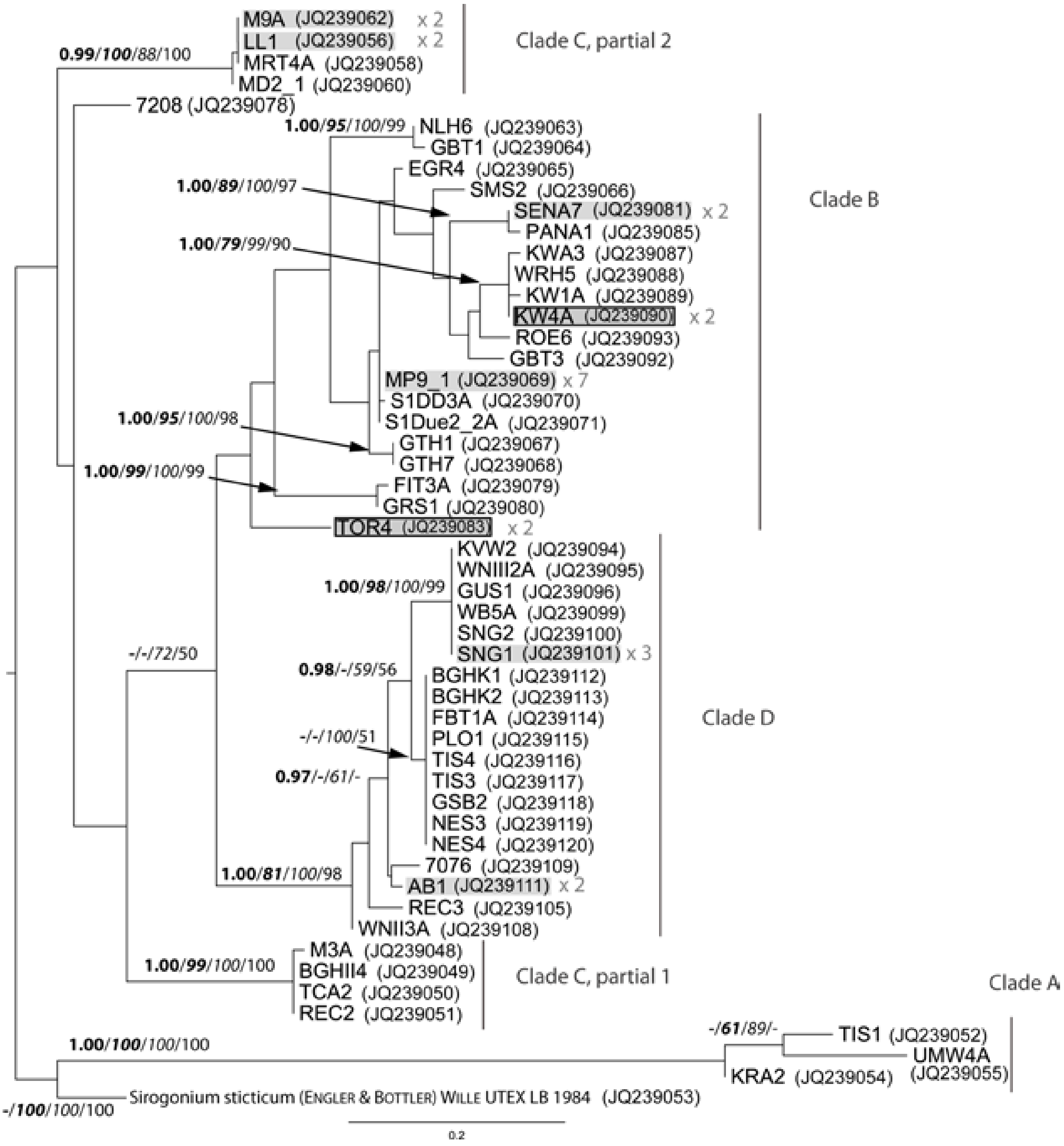

Fig. 2B. *Spirogyra* Intron: molecular phylogeny of a *Spiroyra* group I intron alignment. The phylogenetic tree was inferred by maximum likelihood analyses of 251 aligned positions of 52 taxa using PAUP* 4.0b10. As an evolutionary model, the TrN+I+G model was chosen (base frequencies: A 0.24, C 0.29, G 0.26, T 0.21; rate matrix: A–C 1.0000, A–G 1.4160, A–T 1.0000, C–G 1.0000, C–T 1.8766, G–T 1.0000) with the proportion of invariable sites (I=0.1481) and gamma distribution parameter (G=0.8150). Bootstrap values (>50%) of the maximum likelihood (1000 replicates; bold italic), neighbor–joining (1000 replicates; italic), and maximum likelihood (1000 replicates; not bold) as well as Bayesian values (>95%; bold) were calculated by MrBayes 3.1.2 using the covariation settings (2.5 million generations, trees from 8320 – 250000) are given in the tree (bayesian/ML/NJ/MP). Sequences marked with grey boxes were found more than once (number of clones with identical sequence given in grey after the respective sequence), for details see Results.

the typical P5 extension of the IC introns (Cech 1988; Michel & Westhof 1990), which together with the P2.1 domain places them in the subgroup IC1. The overall sequence length of the intron varies considerably from 387 to 771 nt. Four taxa (all members of Clade C_1, Fig. 2B) contained an insertion in P2 (229–316 nt). Short open reading frames were found, but do not correspond to known genes.

Tetraloops are the most frequent forms of terminal loops, yet *Spirogyra* tetraloops do not fit the proposed consensus sequences GNRA or

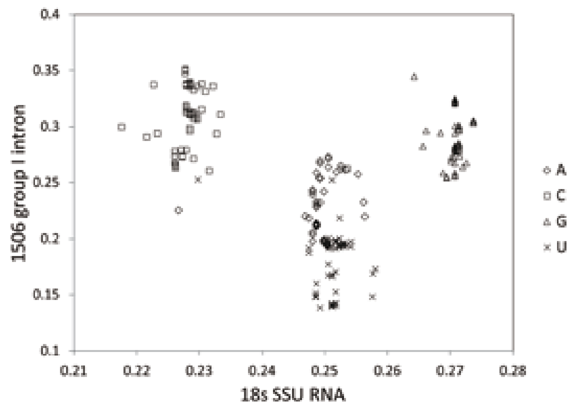

Fig. 3. Comparison of base composition in *Spirogyra* SSU rRNAs (x–axis) and 1506 group I introns (y–axis); relative base frequencies are plotted on the axes; single data point may overlap and create pseudo–filled symbols.

UNCG (MURPHY & CECH 1994) (where N = A, G, C or U and R =A or G).

The intron shows the highly conserved U preceding the 5' splice site and the G to which it pairs (Fig. 4, 1), and the G preceding the 3' splice site (CECH 1988). The most conserved regions in the *Spirogyra* 1506 group I intron are the intron core regions: P3, P7 and J7/8 and P4, P6 and J4/5.

Clade A differs in the sequence of P3 and P7 from the typical *Spirogyra* group I intron. The typical P3 sequence is CGCGACC (see Fig. 4), while the intron sequences of clade A have a mutation in the 6th nt from C to U. The typical P7 sequence is GAGGC–GCCUC. Clade A strains show a different sequence (GAGAC– GUCCU/G). The first four base pairs in P1 and P8 are highly conserved (Fig. 4/2,3). They are identical to a putative consensus structure of zygnematalean taxa (BHATTACHARYA et al. 1996). P2 has a highly conserved sub–helix that is present in all taxa (Figs 4/4) and shows an additional optional sub–helix that is only present in strains of clade B (Fig. 4/5). The expected GAAA tetraloop in the P5 extension and the A–rich bulge are not present. Instead, L5b shows highly divergent sequences (Fig. 4/6). The remaining intron core region is highly conserved as well (P4, P6, J4/5); the peripheral extensions are more divergent (MURPHY & CECH 1994). The J4/5 contains AAA in 65 cases (Fig. 4/7); in the other sequences the first A is changed to a G. The most frequent terminal loop in L5c (39/72) is an AACG–tetraloop (Fig. 4/8). Domains P8 and P9 are longer in clade D (58 to 82 nt, and 80 to 105 nt, respectively). Also the P4–6 domain is longer in clade D (197 to 207 nt) than in other clades.

The search for NHS revealed 58 NHS, 4 H–CBCs and 5 CBCs (Table 3). Clade A showed the most diverging pattern from the other clades; it was characterized by 26 NHS, 3 H–CBCs (P3, P5, P7) and 2 CBCs (P3, P7). Clade B had 2 NHS and 1 CBC (P6), Clade C 18 NHS and 2 CBCs (P2.1, P6), Clade D 8 NHS and 1 H–CBC (P6) and *Sirogonium* showed 4 NHS.

Discussion

The *Spirogyra* intron is a typical Zygnematalean intron (GONTCHAROV et al. 2002) and belongs to the CI subgroup of the group I introns based on the characterization of the primary and secondary structure features by Michel & Westhof (1990) and Damberger & Gutell (1994). The introns show most features of the characteristic catalytic cores for enzymatic activity consisting of P3, P4, J4/5, P6, P7, P8 and P9 which are highly conserved (SALDANHA et al. 1993). They also exhibit typical linker regions J3/4: AAC, J6/7: UCA, J8/7: AAGAUA (GANESAN & KESAVAN 2009). As expected, the more peripheral extensions are more divergent (MURPHY & CECH 1994). Large extensions of the P5 domain are typical for such introns. However, terminal loop of P5 (L5b) does not have the expected GAAA tetraloop and also the A–rich bulge is missing, which is a characteristic of later–diverging Desmidiaceae according to Bhattacharya, Damberger et al. (1996). All sequenced *Spirogyra* introns showed the optional P2 helix that is characteristic for saccoderm and filamentous desmids as well as for the Zygnematales (BESENDAHL & BHATTACHARYA 1999) which is highly conserved in sequence. In the strains of Clade B, an additional sub–domain in P2 could be found, but it varies in sequence and size. Four strains of Clade C displayed an insertion in P2 (229–316 nt), which is shorter compared to insertions of other groups such as in the Bangiales (MULLER et al. 2001), but much longer than the zygnematalean 1506 group I intron insertion found in *Mougeotia* (147 nt in P6) (BHATTACHARYA et al. 1996).

When comparing base frequencies of *Spirogyra* SSU to the values for desmids given in Bhattacharya, Damberger et al. (1996), a higher C–content (Spiro: 22% to 25%, desmid: ~20% to 21%), a lower U–content (Spiro: 23% to 26%,

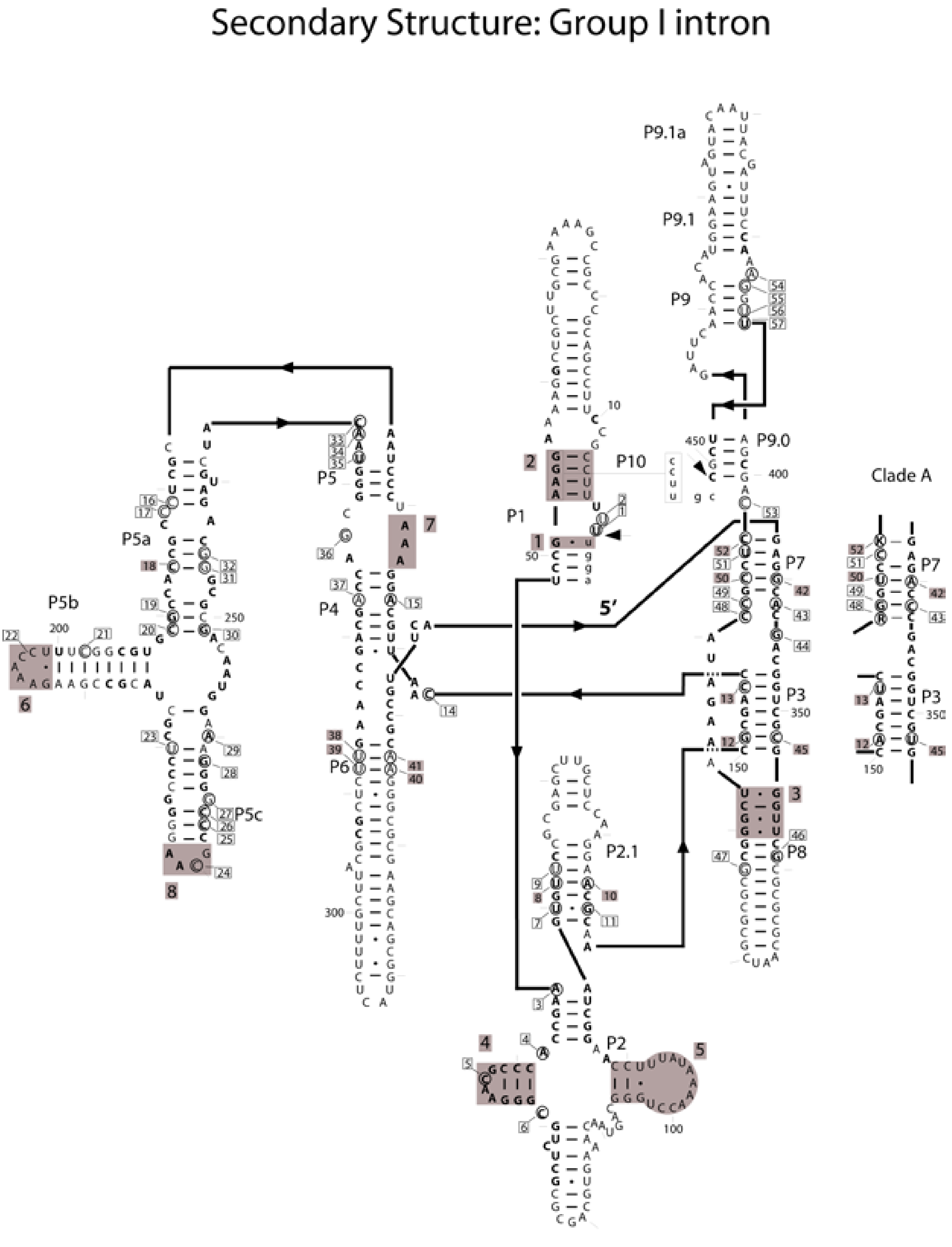

Fig. 4. Putative secondary structure of *Spirogyra*, shown by way of example of a representative sequence WRH5 (Acc.No. JQ239088; Clade B). Exon sequence in lower case, intron in upper case letters; 80% consensus sequence in bold face letters. Subdomains are numbered according to Burke, Belfort et al. 1987 and Cech 1988, nucleotide marks are set every 10 nts. NHS are numbered following Tab.1. H–CBCs and CBCs are marked in grey. Parts discussed in the results are numbered (1 – 8, see text for details) and highlighted in grey.

Table 2. Relative Rate Test (Müller et al., 2004); using only unambiguously aligned positions; *Gonatozygon aculteaum* as outgroup; genus *Spirogyra* marked in light grey [Intron – above diagonal, SSU – below diagonal; not significant: N.S. (p > 0.05; relative rates not significantly different at 5% level). Asterisks: p = 0.05 > * > 0.01 > ** > 0.005 > *** (relative rates significantly different).; (Genic|X74753) *Genicularia spirotaenia*; (Meso|X75763) *Mesotaenium caldariorum*; (Zygn|X79495) *Zygnemopsis circumcarinata*; (Moug|X70705) *Mougeotia scalaris;* (Siro|AJ428076) *Sirogonium sticticum*].

	taxa	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Genic|X74753	–	**	***	**	***	***	***	***	***	***	***	***	***
2	Meso|X75763	*	–	N.S.	N.S.	***	***	***	***	***	***	**	***	**
3	Zygn|X79495	N.S.	N.S.	–	N.S.	**	**	***	***	***	**	*	**	*
4	Moug|X70705	*	N.S.	*	–	**	**	***	***	***	**	**	**	**
5	WRH5	N.S.	N.S.	N.S.	N.S.	–	*	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
6	WNII3A	***	***	***	***	***	–	N.S.	N.S.	N.S.	*	*	*	N.S.
7	MRT4A	***	***	***	***	***	N.S.	–	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
8	SNG1	***	***	***	***	***	N.S.	*	–	N.S.	***	N.S.	N.S.	N.S.
9	GTH1	***	***	***	***	***	N.S.	N.S.	**	–	N.S.	N.S.	N.S.	N.S.
10	TCA2	***	***	***	***	***	N.S.	N.S.	N.S.	N.S.	–	N.S.	N.S.	N.S.
11	KRA2	***	***	***	***	***	N.S.	**	N.S.	*	*	–	N.S.	N.S.
12	TIS1	***	***	***	***	***	***	***	***	***	***	***	–	N.S.
13	Siro|AJ428076	***	***	***	***	***	***	***	***	***	***	***	N.S.	–

desmid: ~27%), a higher G–content (Spiro: 26% to 29%, desmid: ~26%) and a slightly lower A–content (Spiro: 23% to 26%, desmid: ~26%) was observed. A shift in relative base frequencies from U to C and from A to G is characteristic for the *Spirogyra* SSU. The percentual G+C content of the group I intron ranges from 52% to 68%. Compared to the typical desmid intron (Bhattacharya et al. 1996), the *Spirogyra* intron exhibits a much higher G–C value, but for the SSU, the G+C percentage of the unambiguously aligned positions ranged from 56% to 64% and therefore agrees well with the observed range of G+C percent contents of group I intron sequences given in the ML tree in Bhattacharya, Surek et al. (1994). This C – U bias (Fig. 3) is reflected also in the modeltest results (Fig. 2A + B), which explains why *Spirogyra* is a so–called long branch taxon.

The *Spirogyra* intron is different from the typical Desm/Meso/Zygn intron, but its role among this group of algae may not be exceptional. *Closterium littorale* also exhibits a very long branch in the intron phylogeny in Fig. 2A. The Desmidiaceae do not form a clade in Fig. 1+2A, unlike to the findings using *rbcL* phylogenetic analyses (McCourt et al. 2000). Only the *Gonatozygon/ Genicularia* clade and the *Sphaerozosma/ Cosmocladium* clade agrees with their *rbcL* tree and with phylogenetic trees presented by Besendahl & Bhattacharya (1999). These dissimilarities can be attributed to different taxon sampling as we did not include the same taxa to our study. However, similar to those results, *Mougeotia* and *Mesotaenium* could not be placed in a defined relationship. Because there is no bootstrap or Bayesian support for any of the relationships among the Mesotaeniaceae/ Zygnemataceae, except for the *Spirogyra/ Sirogonium* branch, the phylogenetic relationship remains unclear.

The genus *Spirogyra* shares some similarities with the genus *Mougeotia*. Besendahl did not find the 1506 group I intron in one analyzed species of *Mougeotia,* whereas other species do contain the intron (Besendahl & Bhattacharya 1999). This genus also shows both, species with and without intron. Maybe these species differences are a characteristic feature for certain Zygnematalean genera, but this hypothesis has to be solved with enhanced within–genus taxon sampling.

Although there are suggestions to use group I introns as a taxonomic marker to provide more divergent sequences for resolving evolutionary relationships (Bhattacharya 1998), we assume that analyses from the intron alone do not provide a reliable source for phylogenetic interpretation. As shown in our present study, introns are not necessarily present in all members of a genus, which makes it difficult to judge genetic relatedness between taxa with and without intron. Furthermore, they are more difficult to align, decisions on including or excluding characters are more crucial for analyses, and resulting longer branches and low bootstrap/Bayesian support pave the way for misinterpretation. Even within a genus, results are not sufficient for clear statements on difficult taxa. However, when intron sequences are used as a supporting marker to the conserved SSU, it may enhance bootstrap/Bayesian values. The problem with phylogenetic analyses of SSU and group I introns is that both represent rather fast evolving parts of the genome for *Spirogyra* (Gontcharov et al. 2004). To level all long branch effects, it would be better to combine markers with different evolutionary rates (Gontcharov et al. 2004). The use of the intron as a taxonomic marker within the genus *Spirogyra* is problematic as faster evolving taxa that experienced the loss of the intron would be excluded from analyses and interpretation.

However, it is likely that a common ancestor of the *Spirogyra* contained the group I intron that was subsequently lost in more derived clades like clade E (Fig. 1). The single origin of the Zygnematalean group I intron has already been stated and discussed by Bhattacharya (Bhattacharya et al. 1994; Bhattacharya 1996; Bhattacharya et al. 1996; Besendahl & Bhattacharya 1999) and Gontcharov (2008). Haugen, Simon et al. (2005) state that once an intron invades a population and becomes fixed, the homing endonuclease gene responsible for splicing activity accumulates mutations and is consequently inactivated or lost. This loss of self–splicing capacity of the 1506 intron may therefore have originated within the common ancestor of the Zygnematales (Besendahl & Bhattacharya 1999). *Spirogyra* exhibits both – presence and absence of a 1506 intron. This disposition for a secondary loss is potentially caused by accumulation of mutations because of a long evolutionary history and/or a high evolutionary rate. The loss of an optional subdomain in P2 defines the derived desmids (Besendahl & Bhattacharya 1999). The presence of this optional P2 domain – therefore a characteristic of early Desmidiaceae – supports the theory of a long evolutionary history. The absence of the GAAA tetraloop in L5b and the high divergence of the intron and the SSU rRNA among the *Spirogyra* strains and from other Desmidiaceae/ Zygnemataceae support the theory of a high evolution rate, as Gontcharov, Marin et al. (2002) already mentioned for the Zygnematales. The phylogenies generated from the *Spirogyra* sequence data of the SSU rRNA and the 1506 group I intron are largely congruent (Fig. 1+2A), and support the hypothesis of a long term immobility and vertical inheritance of the intron for the Zygnematales (Bhattacharya 1998).

The variability of the intron between the individual *Spirogyra* clades is high. The intron of Clade A is characterized by 2 CBCs and 3 H–CBCs, all located in highly conserved regions of the intron core (Fig. 4), which explains the long branch formed by this clade. The remaining clades do not show as many H–CBCs or CBCs nor NHS. This reflects not only the generally high evolutionary rate in *Spirogyra* but also shows that in different clades have different evolutionary rates.

When comparing evolutionary rates among the different groups of taxa and different sequence regions used in this study, significant differences in evolutionary rate can be found among different data sets and taxa, respectively. A higher rate of base substitution can be found in the Desm/Meso/Zygn intron data set than in the Desm/Meso/Zygn SSU data set. The Desm/Meso/Zygn SSU data set also has the highest portion of invariant sites, marking it as the least variable data set. Mutations in the group I intron of the Desm/Meso/Zygn group are fixed at a higher rate than in the respective SSU rRNA sequence. However, in *Spirogyra* SSU rRNA, much higher substitution rates were found than in the group I intron sites – up to 8 times higher. *Spirogyra* not only differs significantly in evolutionary rate from the other taxa used; it also shows more individual, significantly different evolutionary rates within the SSU rRNA than within the intron. The 1506 group I introns of *Spirogyra* strains used here are more similar in evolutionary rate than the respective SSU rRNA sequences of the same *Spirogyra* strain. Selective constraints seem to be stronger in the SSU rRNA, where mutations influence the functionality of the resulting ribosomal subunit. Whereas in the group I intron with no known function (Nielsen & Johansen 2009), mutations are more likely to be silent. *Spirogyra* is subjected to the same selective constraints as other algae; still a different mutation pattern is established.

Spirogyra is an early divergent genus of the Zygnematophyceae that possesses a very high evolutionary rate concerning the intron sequences and even more in the SSU rRNA sequences.

Table 3. NHS, H–CBCs and CBC identified in Spirogyra 1506 group I intron; numbers highlighted in grey refer to marked bases in Fig. 4.

Number														
Clade A	1		3	4	5	6	7	8		9			11	12
Clade B														
Clade C					5				8		9	10		
Clade D		2												
Sirogo-nium														
Pos. in alignment	2	3	178	194	200	210	592	595	595	596	596	729	731	740
Domain	P1	P1	J1/2	P2	P2	P2	P2.1	P2.1	P2.1	P2.1	P2.1	P2.1	P2.1	P3
									CBC			CBC		CBC
NHS	A	G	U	C	A	G/G/A	A	A	C	G/G/A	C	G	A	A
original base	U	A/U	A	A	C	C	U	U	U	U	U	A	G	G

Number														
Clade A	13	14	15		17		18	19		20		21		23
Clade B														
Clade C		14				17		19	19		20			
Clade D				16	17								22	
Sirogo-nium														
Pos. in alignment	746	750	761	800	807	807	813	824	824	825	825	835	845	951
Domain	P3	J3/4	P4	P5	P5	P5	P5	P5	P5	P5	P5	P5b	L5b	P5c
	H–CBC						H–CBC							
NHS	U	A	G	G	G	U	U	A	U	A	G	A	G	G
original base	C	C	A	C	C	C	C	G	G	C	C	C/U	A/C/U/–	A/U

Number														
Clade A							30	31						
Clade B														
Clade C		25	26	27	28	29			31					
Clade D										31	32			
Sirogo-nium	24											33	34	35
Pos. in alignment	982	997	999	1001	1009	1011	1024	1031	1031	1031	1034	1075	1076	1078
Domain	L5c	P5c	P5c	P5c	P5c	P5c	P5a	P5a	P5a	P5a	P5a	P5	P5	P5
NHS	U	G/U	A	U	U	U/G	A	C	A	U	C	U	C	C
original base	A/C/G	C	C	G/K	G/–	A	G	G	G	G	A/G/–	C	A	U

Table 3 Cont.

Number											
Clade A		37					42	43		45	46
Clade B	36			39	40						
Clade C		37	38			41			44		
Clade D						41					
Sirogonium											
Pos. in alignment	1089	1096	1113	1114	1196	1197	1220	1222	1226	1239	1248
Domain	J4/5	P4	P6	P6	P6	P6	P7	P7	J7/3	P3	P8
			CBC	CBC	CBC	H–CBC/ CBC	CBC			CBC	
NHS	G (C)	C	C	U	A (G, U)	G	A	C	C	U	A
original base	A/–/U	A	U	C	G (A)	A	G	A	G	C	G

Number												
Clade A		48	49	50	51	52				56		57
Clade B								54				
Clade C									55		56	57
Clade D	47						53					
Sirogonium												
Pos. in alignment	1331	1350	1351	1354	1357	1361	1362	1486	1487	1489	1489	1490
Domain	P8	J7/8	J7/8	P7	P7	P7	J7/9	P9	P9	P9	P9	P9
				CBC		H – CBC						
NHS	C/U	G/A	G	U	C	U/G	G	A	A/U	C/–	A	A
original base	A/G	C	C	C	U	C	A/C/U	C/G/–	G	U/G	U/G	U

Similar to *Mougeotia*, for which a unique insertion in the P6 region was reported (Bhattacharya et al. 1996), some *Spirogyra* strains have large insertions. This variability in the intron and the number of mutations in otherwise conserved regions (P7) lead to the theory that the loss of the intron is a logical and likely step in evolutionary history, and that genera that have not yet developed taxa without group I introns are evolutionary younger or slower than *Spirogyra*. Thus the genus *Spirogyra* may be a helpful tool for tracking and evaluating evolutionary processes.

Acknowledgements

We are very grateful to Prof. Annette Coleman, Brown University, for her long–lasting support and cooperativeness. This study was supported by the Austrian Science Fund, Project number P18465.

References

Besendahl, A. & Bhattacharya, D. (1999): Evolutionary analyses of small–subunit rDNA coding regions and the 1506 group I introns of the Zygnematales (Charophyceae, Streptophyta). – J. Phycol. 35: 560–569.

Bhattacharya, D. (1996): Analysis of the distribution of bootstrap tree lengths using the maximum parsimony method. – Mol. Phylogenet. Evol. 6: 339–350.

Bhattacharya, D. (1998): The origin and evolution of protist group I introns. – Protist 149: 113–122.

Bhattacharya, D.; Cannone, J. & Gutell, R. (2001): Group I intron lateral transfer between red and brown algal ribosomal RNA. – Curr. Genet. 40: 82–90.

Bhattacharya, D.; Damberger, S.; Surek, B. & Melkonian, M. (1996a): Primary and second-

Table 4. Origin of isolates [(Abbr.) Abbreviation; (Coll. date) Collection date; (repr. by) represented in SSU phylogeny by; (AUT) Austria; (GER) Germany; (HR) Croatia; (C.C.) culture collection; (n.a.) not available].

Abbr.	Origin	Source	Coll. date	GPS coordinates	Accession No.	Clade	Intron	repr. by
7076	n.a.	C.C.	n.a.	n.a.	JQ239109	D	+	7076
7208	n.a.	C.C.	n.a.	n.a.	JQ239078	C	+	7208
AB1	between Ahrensdorf & Behrensdorf	GER	2007-04-17	14° 03‘ 32“/ 52° 10‘ 52“	JQ239111	D	+	AB4
AB4	between Ahrensdorf & Behrensdorf	GER	2007-04-17	14° 03‘ 32“/ 52° 10‘ 52“	JQ239110	D	+	AB4
AN3A	Alte Naufahrt	AUT	2006-05-10	16° 28‘ 15“/ 48° 11‘ 46“	JQ239122	G	-	AN3A
BEA1	Benda pond	AUT	2006-05-15	16° 20‘ 58“/ 48° 09‘ 11“	JQ239098	D	+	KVW2
BGHII4	Botanical garden Hamburg near entrance	GER	2007-04-26	09° 51‘ 37“/ 53° 33‘ 43“	JQ239049	C	+	REC2
BGHK1	Botanical garden Hamburg limestone alpine area	GER	2007-04-26	09° 51‘ 36“/ 53° 33‘ 50“	JQ239112	D	+	BGHK1
BGHK2	Botanical garden Hamburg limestone alpine area	GER	2007-04-26	09° 51‘ 36“/ 53° 33‘ 50“	JQ239113	D	+	NES3
BTSA1	Pond Suessenbrunn	AUT	2006-05-25	16° 29‘ 09“/ 48° 16‘ 36“	JQ239106	D	+	VNE2
DRS3	Dranser lake	GER	2007-04-19	12° 37‘ 53“/ 53° 10‘ 48“	JQ239059	C	+	MRT4A
EGR4	drainage trench Reddelich	GER	2007-04-23	11° 49‘ 52“/ 54° 04‘ 38“	JQ239065	B	+	EGR4
FBT1A	Fuchsbodenteich	AUT	2006-06-16	15° 51‘ 45“/ 48° 21‘ 54“	JQ239114	D	+	TIS3
FIT3A	Filmteich	AUT	2006-05-29	16° 24‘ 17“/ 48° 09‘ 01“	JQ239079	B	+	FIT3A
GBT1	gr Brunnerteich	AUT	2006-10-23	16° 18‘ 30“/ 48° 07‘ 00“	JQ239064	B	+	GBT1
GBT3	gr Brunnerteich	AUT	2006-10-23	16° 18‘ 30“/ 48° 07‘ 00“	JQ239092	B	+	GBT3
GRS1	Hamburg Großensee (pier)	GER	2007-04-25	10° 20‘ 51“/ 53° 36‘ 46“	JQ239080	B	-	GRS1
GSB2	gr. Segeberger Lake	GER	2007-04-25	10° 20‘ 23“/ 53° 57‘ 27“	JQ239118	D	+	GSB2
GTH1	garden pond Hamburg	GER	2007-04-25	10° 08‘ 44“/ 53° 37‘ 21“	JQ239067	B	+	GTH1
GTH7	garden pond Hamburg	GER	2007-04-25	10° 08‘ 44“/ 53° 37‘ 21“	JQ239068	B	+	GTH1
GUS1	Grubensee	GER	2007-04-17	13° 59‘ 40“/ 52° 09‘ 30“	JQ239096	D	+	KVW2
IGH2	Industrial area Höltingbaum	GER	2007-04-25	10° 11‘ 01“/ 53° 36‘ 44“	JQ239123	H	-	IGH2
KRA2	watercourse between Krauswitz & Schlagnitz	GER	2007-04-17	13° 52‘ 50“/ 52° 01‘ 48“	JQ239054	A	+	KRA2
KVW2	Pond near traffic circle A19 Wittstock	GER	2007-04-19	12° 27‘ 20“/ 53° 09‘ 26“	JQ239094	D	+	KVW2
KVW4	Pond near traffic circle A19 Wittstock	GER	2007-04-19	12° 27‘ 20“/ 53° 09‘ 26“	JQ239097	D	+	KVW2
KW1A	Kaiserwasser	AUT	2006-04-25	16° 25‘ 29“/ 48° 13‘ 54“	JQ239089	B	+	KWA3
KW4A	Kaiserwasser	AUT	2006-04-25	16° 25‘ 29“/ 48° 13‘ 54“	JQ239090	B	+	KWA3
KWA3	Kaiserwasser	AUT	2006-04-25	16° 25‘ 29“/ 48° 13‘ 54“	JQ239087	B	+	KWA3

Table 4 Cont.

LL1	local fire service pond Limsdorf	GER	2007-04-17	14° 00‘ 51“/ 52° 09‘ 34“	JQ239056	C	+	MRT4A
LL2	local fire service pond Limsdorf	GER	2007-04-17	14° 00‘ 51“/ 52° 09‘ 34“	JQ239057	C	+	MRT4A
LSB4A	Langenschoenbichl	AUT	2006-06-16	15° 59‘ 45“/ 48° 19‘ 52“	JQ239091	B	+	KWA3
M3A	St. Pölten	AUT	2006-03-20	n.a.	JQ239048	C	+	M3A
M5A	St. Pölten	AUT	2006-03-20	n.a.	JQ239075	B	+	M6A
M6A	St. Pölten	AUT	2006-03-20	n.a.	JQ239076	B	+	M6A
M9A	St. Pölten	AUT	2006-03-20	n.a.	JQ239062	C	+	MRT4A
M9B	St. Pölten	AUT	2006-03-20	n.a.	JQ239074	B	+	M6A
MD2-1	St. Pölten	AUT	2006-03-20	n.a.	JQ239060	C	+	MRT4A
MDA1	St Pölten	AUT	2006-03-20	n.a.	JQ239061	C	+	MRT4A
MP9-1	St Pölten	AUT	2006-04-11	n.a.	JQ239069	B	+	M6A
MP923	St. Pölten	AUT	2006-04-11	n.a.	JQ239073	B	+	M6A
MP92-6	St. Pölten	AUT	2006-04-11	n.a.	JQ239072	B	+	M6A
MRT4A	Mausrodlteich	AUT	2006-08-24	15° 02‘ 52“/ 47° 52‘ 25“	JQ239058	C	+	MRT4A
NES3	Nebelsee	GER	2007-04-19	12° 39‘ 58“/ 53° 14‘ 54“	JQ239119	D	+	NES3
NES4	Nebelsee	GER	2007-04-19	12° 39‘ 58“/ 53° 14‘ 54“	JQ239120	D	+	NES3
NLH5	watercourse between Neu-Lübbenau & Hohenbrück	GER	2007-04-19	13° 53‘ 38“/ 52° 04‘ 57“	JQ239084	B	+	NLH5
NLH6	watercourse between Neu-Lübbenau & Hohenbrück	GER	2007-04-17	13° 53‘ 38“/ 52° 04‘ 57“	JQ239063	B	+	NLH6
OSOR-IV-A1	Osor	HR	2006-05-24	14° 23‘ 87“/ 44° 41,732‘	JQ239104	D	+	SNG1
PANA1	Panozzalacke	AUT	2006-05-10	16° 29‘ 16“/ 48° 10‘ 51“	JQ239085	B	+	PANA1
PLO1	Plötinsee Rostock	GER	2007-04-20	12° 58‘ 22“/ 53° 13‘ 14“	JQ239115	D	+	TIS3
REC2	Meadow pond Rechlin	GER	2007-04-20	12° 45‘ 45“/ 53° 18‘ 43“	JQ239051	C	+	REC2
REC3	Meadow pond Rechlin	GER	2007-04-20	12° 45‘ 45“/ 53° 18‘ 43“	JQ239105	D	+	VNE2
ROE6	Pond near Rödel	GER	2007-04-19	12° 35‘ 51“/ 53° 23‘ 06“	JQ239093	B	+	ROE6
S1DD3A	St. Pölten	AUT	2006-03-20	n.a.	JQ239070	B	+	M6A
S1DUe2-2A	St. Pölten	AUT	2006-03-20	n.a.	JQ239071	B	+	M6A
S1Due2A	St Pölten	AUT	2006-03-20	n.a.	JQ239077	B	+	M6A
SENA4	Sengsee	GER	2006-07-04	11° 18‘ 58“/ 47° 46‘ 37“	JQ239082	B	+	SENA7
SENA7	Sengsee	GER	2006-07-04	11° 18‘ 58“/ 47° 46‘ 37“	JQ239081	B	+	SENA7
SMS2	Scharmützelsee	GER	2007-04-17	14° 00‘ 50“/ 52° 12‘ 49“	JQ239066	B	+	SMS2

Table 4 Cont.

SNG1	Marsh north Großensee	GER	2007-04-25	10° 21‘ 27“/ 53° 37‘ 51“	JQ239101	D	+	SNG1
SNG2	Marsh north Großensee	GER	2007-04-25	10° 21‘ 27“/ 53° 37‘ 51“	JQ239100	D	+	SNG1
SNG3	Marsh north Großensee	GER	2007-04-25	10° 21‘ 27“/ 53° 37‘ 51“	JQ239103	D	+	SNG1
SNG5	Marsh north Großensee	GER	2007-04-25	10° 21‘ 27“/ 53° 37‘ 51“	JQ239102	D	+	SNG1
TCA2	Canow fish farming	GER	2007-04-20	12° 53‘ 23“/ 53° 11‘ 32“	JQ239050	C	+	REC2
TIS1	Tiefensee	GER	2007-04-17	13° 59‘ 22“/ 52° 09‘ 23“	JQ239052	A	+	TIS1
TIS3	Tiefensee	GER	2007-04-17	13° 59‘ 22“/ 52° 09‘ 23“	JQ239117	D	+	TIS3
TIS4	Tiefensee	GER	2007-04-17	13° 59‘ 22“/ 52° 09‘ 23“	JQ239116	D	+	TIS3
TOR4	Torsee	GER	2007-04-18	13° 39‘ 56“/ 52° 14‘ 24“	JQ239083	B	+	TOR4
TRS7	Tristacher Lake	AUT	2006-08-10	12° 47‘ 40“/ 46° 48‘ 29“	JQ239121	E	-	TRS7
UMW4A	Lower Muehlwasser	AUT	2006-05-09	16° 27‘ 24“/ 48° 12‘ 47“	JQ239055	A	+	UMW4A
UMW4A	Lower Muehlwasser	AUT	2006-05-09	16° 27‘ 24“/ 48° 12‘ 47“	JQ239055	A	+	UMW4A
UTEX LB 1984	n.a.	C.C.	n.a.	n.a.	JQ239053	C	+	UTEX LB 1984
VNE2	accretion zone Neversdorfer lake	GER	2007-04-25	10° 15‘ 31“/ 53° 51‘ 56“	JQ239107	D	+	VNE2
WB5A	Wienerberg pond2	AUT	2006-05-15	16° 20‘ 49“/ 48° 09‘ 43“	JQ239099	D	+	KVW2
WNII3A	local fire service pond Gerasdorf	AUT	2006-05-25	16° 27‘ 58“/ 48° 18‘ 00“	JQ239108	D	+	WNII3A
WNII4A	local fire service pond Gerasdorf	AUT	2006-04-25	16° 27‘ 58“/ 48° 18‘ 00“	JQ239086	B	+	PANA1
WNIII2A	Pond recreation area Seeschlacht (Vienna North III)	AUT	2006-05-25	16° 21‘ 49“/ 48° 17‘ 56“	JQ239095	D	+	KVW2
WRH5	Wandse detention reservoir	GER	2007-04-25	10° 11‘ 07“/ 53° 37‘ 17“	JQ239088	B	+	KWA3
ZIP2	Zipke	GER	2007-04-22	12° 47‘02“/ 54° 20‘ 30“	JQ239124	F	-	ZIP2

ary structure analyses of the rDNA group–I introns of the Zygnematales (Charophyta). – Curr. Genet. 29: 282–286.

Bhattacharya, D.; Friedl, T. & Damberger, S. (1996b): 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.; Surek, B.; Rüsing, M.; Damberger, S. & Melkonian, M. (1994): Group I introns are inherited through common ancestry in the nuclear–encoded rRNA of Zygnematales (Charophyceae). – Proceedings of the National Academy of Sciences of the United States of America 91: 9916–9920.

Burke, J.M.; Belfort, M.; Cech, T.R.; Davies, R.W.; Schweyen, R.J.; Shub, D.A.; Szostak, J.W. & Tabak, H.F. (1987): Structural conventions for group I introns. – Nucl. 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.; Mueller, K.M.; Pande, N.; Shang, Z.; Yu, N. & 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. – BioMed Central Bioinformatics 3: 15.

Cech, T.R. (1988): Conserved sequences and structures of group I introns: building an active site for RNA catalysis – a review. – Gene 73: 259–271.

Chen, C.; Barfuss, M.H.J.; Coleman, A.W.; Pröschold,

T. & Schagerl, M. (2012): Hidden genetic diversity of the green alga *Spirogyra* (Conjugatophyceae, Streptophyta). – BMC Evolutionary Biology, in press.

Damberger, S.H. & Gutell, R.R. (1994): A comparative database of group I intron structures. – Nucl. Acids Res. 22: 3508–3510.

Drummond, C.S.; Hall, J.; Karol, K.G.; Delwiche, C.F. & McCourt, R.M. (2005): Phylogeny of *Spirogyra* and *Sirogonium* (Zygnematophyceae) based on rbcL sequence data. – J. Phycol. 41: 1055–1064.

Dujon, B. (1989): Group I introns as mobile genetic elements: facts and mechanistic speculations: a review. – Gene 82: 91–114.

Ganesan, N. & Kesavan, C. (2009): Phylogeny of conserved adenines in linkers of group I introns. – Available from Nature Preceedings <http://hdl.handle.net/10101/npre.2009.2909.1>.

Gilbert, W. (1978): Why genes in pieces? – Nature 271: 501.

Gontcharov, A.A. (2008): Phylogeny and classification of Zygnematophyceae (Streptophyta): Current state of affairs. – Fottea 8: 87–104.

Gontcharov, A.A.; Marin, B. & Melkonian, M. (2002): Molecular phylogeny of conjugating green algae (Zygnemophyceae, Streptophyta) inferred from SSU rDNA sequence comparisons. – J. Mol. Evol. 56: 89–104.

Gontcharov, A.A.; Marin, B. & Melkonian, M. (2004): Are combined analyses better than single gene phylogenies? A case study using SSU rDNA and rbcL sequence comparisons in the Zygnematophyceae (Streptophyta). – Mol. Biol. Evol. 21: 612–624.

Gontcharov, A.A. & Melkonian, M. (2004): Unusual position of the genus *Spirotaenia* (Zygnematophyceae) among streptophytes revealed by SSU rDNA and rbcL sequence comparisons. – Phycologia 43: 105–113.

Hainz, R.; Wöber, C. & Schagerl, M. (2009): The relationship between *Spirogyra* (Zygnematophyceae, Streptophyta) filament type groups and environmental conditions in Central Europe. – Aquat. Bot. 91: 173–180.

Hall, T.A. (1999): BioEdit: a user–friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. – Nucleic Acids Symposium Series 41: 95–98.

Haugen, P.; Simon, D.M. & Bhattacharya, D. (2005): The natural history of group I introns. – Trends Genet. 21: 111–119.

Hoshaw, R.W. & McCourt, R.M. (1988): The Zygnemataceae (Chlorophyta): A twenty–year update of research. – Phycologia 27: 511–548.

Hoshaw, R.W.; Wang, J.–C.; McCourt, R.M. & Hull, H.M. (1985): Ploidal changes in clonal cultures of *Spirogyra communis* and implications for species definition. – Am. J. Bot. 72: 1005–1011.

Huelsenbeck, J.P. & Ronquist, F. (2001): MRBAYES: Bayesian inference of phylogeny. – Bioinformatics 17: 754–755.

Kadlubowska, J.Z. (1984): Conjugatophyceae I – Zygnemales. – In: Ettl, H., Gerloff, H., Heynig, H. & Mollenhauer, D. (eds): Süßwasserflora von Mitteleuropa, Chlorophyta VIII. – 532 pp., Gustav Fischer Verlag.Stuttgart, New York.

Kolkwitz, R. & Krieger, H. (1941): Zygnemales. – In: Kolkwitz, R. (ed.): Dr. L Rabenhorst's Kryptogamen–Flora von Deutschland und der Schweiz XIII, 2 Abteilung. – 499 pp., Akademische Verlagsgesellschaft Becker & Erler, Leipzig.

Marin, B.; Palm, A.; Klingberg, M.a.x. & Melkonian, M. (2003): Phylogeny and Taxonomic Revision of Plastid–Containing Euglenophytes based on SSU rDNA Sequence Comparisons and Synapomorphic Signatures in the SSU rRNA Secondary Structure. – Protist 154: 99–145.

McCourt, R.M.; Hoshaw, R.W. & Wang, J.–C. (1986): Distribution, morphological diversity and evidence for polyploidy in North American Zygnemataceae (Chlorophyta). – J. Phycol. 22: 307–315.

McCourt, R.M.; Karol, K.G.; Bell, J.; Helm–Bychowski, K.M.; Grajewska, A.; Wojciechowski, M.F. & Hoshaw, R.W. (2000): Phylogeny of the conjugating green algae (Zygnemophyceae) based on rbcL sequences. – J. Phycol. 36: 747–758.

Michel, F. & Dujon, B. (1983): Conservation of RNA secondary structures in two intron families including mitochondrial–, chloroplast– and nuclear–encoded members. – The EMBO Journal 2: 33–38.

Michel, F. & Westhof, E. (1990): Modeling of the three–dimensional architecture of group I catalytic introns based on comparative sequence analysis. – J. Mol. Biol. 216: 585–610.

Müller, K.; Borsch, T.; Legendre, L.; Porembski, S.; Theisen, I. & Barthlott, W. (2004): Evolution of Carnivory in Lentibulariaceae and the Lamiales. – Plant Biology 6: 477–490.

Muller, K.M.; Cannone, J.J.; Gutell, R.R. & Sheath, R.G. (2001): A structural and phylogenetic analysis of the group IC1 introns in the order Bangiales (Rhodophyta). – Mol. Biol. Evol. 18: 1654–1667.

Murphy, F.L. & Cech, T.R. (1994): GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. – J. Mol. Biol. 236: 49–63.

Nielsen, H. & Johansen, S.D. (2009): Group I introns:

Moving in new directions. – RNA Biology 6: 375–383.

OLIVEIRA, M.C. & RAGAN, M.A. (1994): Variant forms of a group I intron in nuclear small–subunit rRNA genes of the marine red alga *Porphyra spiralis var. amplifolia*. – Mol. Biol. Evol. 11: 195–207.

PHILIPPE, H. (2000): Opinion: Long branch attraction and protist phylogeny. – Protist 151: 307–316.

POSADA, D. & BUCKLEY, T. (2004): Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. – Syst. Biol. 53: 793–808.

POSADA, D. & CRANDALL, K. (1998): Modeltest: testing the model of DNA substitution. – Bioinformatics 14: 817–818.

RANDHAWA, M.S. (1959): Zygnemaceae. – 478 pp., Indian Council of Agricultural Research, New Delhi.

RIETH, A. (1983): Eine *Spirogyra* von der Ostsee bei Zingst. – Genet. Resour. Crop. Ev. 31: 317–326.

RONQUIST, F. & HUELSENBECK, J.P. (2003): MRBAYES 3: Bayesian phylogenetic inference under mixed models. – Bioinformatics 19: 1572–1574.

SALDANHA, R.; MOHR, G.; BELFORT, M. & LAMBOWITZ, A.M. (1993): Group I and group II introns. – FASEB J. 7: 15–24.

SIMONS, J. & VAN BEEM, A. (1990): *Spirogyra* species and accompanying algae from pools and ditches in the Netherlands. – Aquat. Bot. 37: 247–269.

SWOFFORD, D.L. (2003): PAUP*. Phylogenetic Analysis Using Parsimony (* and other methods). Sunderland, Massachusetts, Sinauer Associates.

TRANSEAU, E.N. (1951): The Zygnemataceae (Fresh–water conjugate algae). – 327 pp., The Ohio State University Press.

WEISER, B. & NOLLER, H. (2009): XRNA, University of Santa Cruz.

ZAUG, A.J.; GRABOWSKI, P.J. & CECH, T.R. (1983): Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage–ligation reaction. – Nature 301: 578–583.

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

© Czech Phycological Society (2012)

Received September 3, 2011
Accepted October 8, 2011