Selection of genes that have not been horizontally transferred for prokaryote phylogenetic inferences is regarded as a challenging task. The markers internal transcribed spacer of ribosomal genes (16S–23S ITS) and phycocyanin intergenic spacer (PC-IGS), based on the operons of ribosomal and phycocyanin genes respectively, are among the most used markers in cyanobacteria. The region of the ribosomal genes has been considered stable, whereas the phycocyanin operon may have undergone horizontal transfer. To investigate the occurrence of horizontal transfer of PC-IGS, phylogenetic trees of *Geitlerinema* and *Microcystis* strains were generated using PC-IGS and 16S–23S ITS and compared. Phylogenetic trees based on the two markers were mostly congruent for *Geitlerinema* and *Microcystis*, indicating a common evolutionary history among ribosomal and phycocyanin genes with no evidence for horizontal transfer of PC-IGS. Thus, PC-IGS is a suitable marker, along with 16S–23S ITS for phylogenetic studies of cyanobacteria.

**Key index words:** 16S–23S ITS; cyanobacteria; lateral gene transfer; molecular phylogeny; phycocyanin; ribosomal genes

**Abbreviations:** 16S–23S ITS, internal transcribed spacer of ribosomal genes marker; BCCUSP, Brazilian Cyanobacteria Collection of the University of Sao Paulo; HGT, horizontal gene transfer; PC-IGS, phycocyanin intergenic spacer marker; PP, posterior probability

The morphological characteristics of cyanobacteria do not always correspond to their taxonomic diversity (Wilson et al. 2000, Premanandh et al. 2006, Willame et al. 2006) and therefore the use of molecular markers for phylogenetic studies have become essential (Bolch et al. 1996, Thomazeau et al. 2010, Henson et al. 2011, Schirrmeister et al. 2011, Lopes et al. 2012).

One of the greatest challenges in the selection of markers for phylogenetic inference in cyanobacteria is targeting markers that have not undergone horizontal gene transfer (HGT). Phylogenies based on horizontally transferred markers might not correspond to the evolutionary history of the target host organism (Gribaldo and Brochier 2009).

HGT and orthologous gene substitutions are relatively common among prokaryotes and have been important processes in the evolution of this group (Zhaxybayeva and Gogarten 2002, Boucher et al. 2003, Gribaldo and Brochier 2009, Zhaxybayeva and Doolittle 2011). However, HGT events in cyanobacteria may still be underestimated, and genes with several functions could have been subjected to this process (Zhaxybayeva et al. 2006).

The markers phycocyanin intergenic spacer marker (PC-IGS), based on the phycocyanin loci arrangement (including the intergenic spacer between *cpcB* and *cpcA* genes and the flanking regions), and internal transcribed spacer of ribosomal genes marker (16S–23S ITS), based on the ribosomal gene spacer region, are among the most commonly used for investigating evolutionary relationships in cyanobacteria (Bolch et al. 1999, Barker...

There is no reported evidence that the operons of ribosomal genes have undergone horizontal transfer among cyanobacteria. However, the variability observed among the multiple copies of the ribosomal operon found within a single individual can hinder their use in phylogenetic studies (Iteman et al. 2000, 2002, Boyer et al. 2001, Engene et al. 2010).

Nontypical distribution patterns of nucleotides were observed in sequences of the PC-IGS marker (Janson and Granéli 2002, Manen and Falquet 2002). Manen and Falquet (2002) analyzed sequences from the genus Arthrospira Stützenberger ex Gomez and observed a mosaic clustering pattern that appeared to be the result of horizontal transfers of blocks of sequences, and a point of intragenic recombination close to the stop codon of the rpoB gene was proposed. Janson and Granéli (2002) also found unusual patterns when applying a likelihood method to test PC-IGS sequences for recombination events and observed sequences that, in contrast to what was expected, presented a higher variability in the coding sequences than in the spacer. Furthermore, consistent differences among phylogenetic proposals were observed when comparing PC-IGS to 16S–23S rDNA. These results led some authors to hypothesize that PC-IGS could have been horizontally transferred (Janson and Granéli 2002, Six et al. 2007, Haverkamp et al. 2008, 2009). Bittencourt-Oliveira et al. (2009) also demonstrated that in a group of morphologically similar strains of Geitlerinema (Anagnostidis et Komárek) Anagnostid, a fine oscillatory cyanobacteria, some had shorter PG-IGS sequences (84 base pairs) compared to others (294 base pairs), suggesting a HGT event in the gene.

Our goal, using strains of Geitlerinema and Microcystis Küting ex Lemmermann as models, was to evaluate the possible occurrence of HGT in PC-IGS by comparing phylogenetic trees built with PC-IGS and 16S–23S ITS.

**MATERIAL AND METHODS**

**Strains and cultivation conditions.** The clonal and nonaxenic sequenced strains of Geitlerinema and Microcystis belonged to the Brazilian Cyanobacteria Collection of the University of São Paulo (BCCUSP). Strains were maintained in climate chambers with controlled conditions of light (30 μmol photons·m⁻²·s⁻¹, photometer LI-COR [Lincoln, NE, USA], mod. LI-250), photoperiod (14:10 h light:dark), and temperature (29 ± 0.5°C) in BG-11 cultivation medium. (Rippka et al. 1979), of pH 7.4, modified by Bittencourt-Oliveira (2000). The morphospecies nomination (Bittencourt-Oliveira et al. 2001, 2009), culture collection numbers, original site of strains, and GenBank sequence accession numbers are presented in Table 1.

**DNA extraction, 16S–23S ITS PCR amplification, and sequencing.** Total genomic DNA was extracted from fresh cells harvested in the exponential phase as described by Bittencourt-Oliveira et al. (2012). Amplification of the 16S–23S ITS fragment was performed in a GeneAmp PCR System 9700 thermocycler (Applied Biosystems, Foster City, CA, USA).

**Table 1. Morphospecies, code, and place of origin of the cyanobacteria strains sequenced in this work and GenBank accession number of the obtained sequences.**

<table>
<thead>
<tr>
<th>Morphospecies</th>
<th>Code</th>
<th>Place</th>
<th>GenBank accession number—16S–23S ITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. amphibium</td>
<td>BCCUSP 31</td>
<td>Albufera lagoon, Valencia, Spain</td>
<td>KJ754563</td>
</tr>
<tr>
<td>G. amphibium</td>
<td>BCCUSP 79</td>
<td>Jucazinho reservoar, Caruaru, PE</td>
<td>KJ754572</td>
</tr>
<tr>
<td>G. amphibium</td>
<td>BCCUSP 80</td>
<td>Tapacura reservoar, São Lorenco da Mara, PE</td>
<td>KJ754553</td>
</tr>
<tr>
<td>G. amphibium</td>
<td>BCCUSP 84</td>
<td>Jucazinho reservoar, Surubim, PE</td>
<td>KJ754544</td>
</tr>
<tr>
<td>G. amphibium</td>
<td>BCCUSP 85</td>
<td>Jucazinho reservoar, Surubim, PE</td>
<td>KJ754555</td>
</tr>
<tr>
<td>G. amphibiun</td>
<td>BCCUSP 86</td>
<td>Jucazinho reservoar, Surubim, PE</td>
<td>KJ754556</td>
</tr>
<tr>
<td>G. amphibium</td>
<td>BCCUSP 87</td>
<td>Jucazinho reservoar, Surubim, PE</td>
<td>KJ754557</td>
</tr>
<tr>
<td>G. amphibium</td>
<td>BCCUSP 91</td>
<td>Ornamental lagoon, Piracica, SP</td>
<td>KJ754558</td>
</tr>
<tr>
<td>G. unigranulatum</td>
<td>BCCUSP 94</td>
<td>Usina Sta. Rita lagoon, Maracaí, SP</td>
<td>KJ754559</td>
</tr>
<tr>
<td>G. amphibiun</td>
<td>BCCUSP 338</td>
<td>Usina Sta. Rita lagoon, Maracaí, SP</td>
<td>KJ754560</td>
</tr>
<tr>
<td>G. unigranulatum</td>
<td>BCCUSP 350</td>
<td>Usina Sta. Rita lagoon, Maracaí, SP</td>
<td>KJ754561</td>
</tr>
<tr>
<td>M. aeruginosa</td>
<td>BCCUSP 352</td>
<td>Usina Sta. Rita lagoon, Maracaí, SP</td>
<td>KJ754562</td>
</tr>
<tr>
<td>M. aeruginosa</td>
<td>BCCUSP 232</td>
<td>Garças lagoon, São Paulo, SP</td>
<td>KJ754570</td>
</tr>
<tr>
<td>M. aeruginosa</td>
<td>BCCUSP 009</td>
<td>Garças lagoon, São Paulo, SP</td>
<td>KJ754566</td>
</tr>
<tr>
<td>M. aeruginosa</td>
<td>BCCUSP 30</td>
<td>Garças lagoon, São Paulo, SP</td>
<td>KJ754571</td>
</tr>
<tr>
<td>M. paniforme</td>
<td>BCCUSP 299</td>
<td>Garças lagoon, São Paulo, SP</td>
<td>KJ754546</td>
</tr>
<tr>
<td>M. aeruginosa</td>
<td>BCCUSP 299</td>
<td>Garças lagoon, São Paulo, SP</td>
<td>KJ754567</td>
</tr>
<tr>
<td>M. aeruginosa</td>
<td>BCCUSP 011</td>
<td>Usina Sta. Rita lagoon, Maracaí, SP</td>
<td>KJ754565</td>
</tr>
</tbody>
</table>

*Morphological description present in Bittencourt-Oliveira et al. (2009).


*Sequences obtained in this work.

BCCUSP, Brazilian Cyanobacteria Collection of the University of São Paulo; 16S–23S ITS, internal transcribed spacer of ribosomal genes marker.
using 2.5–10 ng of DNA, 20 μM of each oligonucleotide primer in a total volume of 25 μL with pure Taq Ready-To-Go PCR Beads™ kit (GE Healthcare, Fairfield, CT, USA) and primers 322F and 340R (Itoyan et al. 2000). Annealing temperature was increased to 58°C when amplifying DNA from *Microcystis* strains.

Negative control reactions were performed using the same reaction conditions and primers without DNA; no PCR products were detected by agarose electrophoresis. PCR products were purified using the Purelink Kit (Invitrogen, Carlsbad, CA, USA). Bands were extracted and purified from gels using the QIAquick kit (Qiagen, Hilden, Germany). When DNA concentrations from a single PCR reaction were too low for sequencing, four individual amplification reactions were pooled during the purification step. Amplified fragments were directly sequenced on both strands using the forward and reverse primers with a 3100 ABI sequencer (Applied Biosystems). Selected chromatograms are presented in Figure S1 in the Supporting Information. Automated base calls for both strands were checked by manual inspection, and ambiguous sequences for each strain. Alignments (Table S1, in the Supporting Information) were made using ClustalW in the BioEdit program (Hall 1999) to establish consensus sequences. In the Supporting Information, we have included all sequences used for the phylogenetic analyses. Base calls and conflicts were resolved by alignment and comparison using BioEdit (Hall 1999) to establish consensus sequences for each strain. Alignments (Table S1, in the Supporting Information) were made using ClustalW in the BioEdit program (Hall 1999) and manually inspected. Evolutionary distances between sequences were calculated by the Evolutionary Distance program (Hall 1999). 

**Phylogenetic analyses.** Phylogenetic analyses were performed using MrBayes v3.1.2 (Ronquist and Huelsenbeck 2003). Optimal nucleotide substitution models were selected using jModeltest (Posada 2008) under the Akaike Information Criterion. For the *Geitlerinema* analysis, K80+G (Hasegawa et al. 1985 with invariant sites and rate variations among sites) was selected for the PC-IGS matrix, and TPM (3-parameter model, Kimura 1981) for the 16S–23S ITS matrix. For *Microcystis*, HKY + I + G (Hasegawa et al. 1985 with invariant sites and rate variations among sites) was selected for the PC-IGS matrix, and F81+ I + G (Felsenstein 1981 with invariant sites and rate variations among sites) for the 16S–23S ITS matrix. For the Bayesian analysis, two Metropolis-coupled Markov Chain Monte Carlo (MC) searches with four chains (three heated and one cold) were run. The analysis started with a random tree and was run for 5,000,000 generations. Trees and parameters were sampled every 100th generation. The first 2,500 generations were excluded as the burn-in. For all analyses, posterior probability (PP; ×100) values were considered low up to 70, moderate from 70 to 90, and high above 90. Only PP (×100) values higher than 70 were shown.

Phylogenetic trees of *Geitlerinema* and *Microcystis* strains, based on the PC-IGS and 16S–23S ITS markers, were generated and topologically compared. All sequences used for the PC-IGS trees were selected from the GenBank database. Sequences obtained for this study in the 16S–23S ITS trees were used in combination with other sequences from GenBank. *Spiroplana subsalsa* PD2002/gca (accession number AY575949 for PC-IGS and AY575953 for 16S–23S ITS) and *Synechocystis* sp. PCC6803 (accession number NC_020286 for PC-IGS and 16S–23S ITS) were used as outgroups for *Geitlerinema* and *Microcystis* analyses respectively. RNA transfer tRNA(Ile) was observed in the 16S–23S ITS of all *Microcystis* and *Geitlerinema* strains sequenced and used in this study.

**RESULTS**

The *Geitlerinema* phylogenetic trees based on the markers PC-IGS (Fig. 1) and 16S–23S ITS (Fig. 2) showed similar topologies. In both trees, two major monophyletic groups were observed (Figs. 1 and 2, clades A and B) and the distance calculated by *P distance* between the strains of clade A and those of clade B was approximately twice that observed among strains from clade B. Both trees presented conserved external branch topology, except for the position of BCCUSP 80.

For *Microcystis*, the phylogenetic trees obtained using PC-IGS (Fig. 1) and 16S–23S ITS (Fig. 4) were similar, but differences in the terminal branch positions of some strains were observed. The terminal branches labeled A, B, and C were congruent. However, strains of clades D and E of the PC-IGS tree were clustered into clade D in 16S–23S ITS, and the strains BCCUSP 011, UAM 241, UAM 259, NIES 843, and NIES 102 occurred in divergent positions in the two trees.

**DISCUSSION**

HGT is relatively common among prokaryotes, but it does not affect all genes in the same way. For some genomes, gene clusters have a lower probability of being transferred (Shi and Falkowski 2008, Gribaldo and Brochier 2009). In this study, phylogenies proposed using the phycocyanin genes and intergenic spacer (PC-IGS) and the ribosomal gene spacer (16S–23S ITS) as molecular markers were compared to investigate the occurrence of horizontal transfer of PC-IGS.

The phylogenies for *Geitlerinema* and *Microcystis* based on these markers were mostly congruent, indicating a common evolutionary pathway for the phycocyanin and ribosomal genes. Thus, no clear HGT signal was found for PC-IGS. This result is consistent with those of Sanchis et al. (2005) and Dadheech et al. (2010), who found that PC-IGS and 16S–23S ITS regions of *Microcystis* and *Arthrospira* strains also showed a high similarity between marker topologies.

The main point of discordance between the *Geitlerinema* phylograms was the location of *Geitlerinema amphibia* (C.Agardh ex Gomont) Anagnostidis BCCUSP 80 strain. This sequence was sister to clade A in the PC-IGS tree (Fig. 1), but to clade B in the 16S–23S ITS tree (Fig. 2). However, it is noteworthy that in the 16S–23S ITS proposed phylogeny, there is a low Bayesian posterior probability to support its location (Fig. 2).

Bittencourt-Oliveira et al. (2009) noted that all *Geitlerinema* strains that they studied were morphologically similar to each other, but that two of them (BCCUSP 352 and BCCUSP 94) showed shorter PC-IGS sequences (84 base pairs against 294 presented by the other strains) and formed an isolated clade in the PC-IGS phylogram. In this study, both strains also formed an isolated clade in 16S–23S ITS, indicating that despite the morphological similarity, they represent genetically divergent strains. Thus, the hypothesis that the divergence of the strains observed in the PC-IGS tree could have been due to HGT was rejected.
Despite the overall correspondence between the PC-IGS and the 16S–23S ITS trees for *Microcystis*, some differences were observed. Among the strains that presented divergences in their positioning, BCCUSP 011, UAM 241, and UAM 259 showed no significant statistical support for their position.
FIG. 3. Bayesian inference phylogram for PC-IGS sequences from *Microcystis* strains. Posterior probability (×100) is shown on each branch. Outgroup (data not shown) was *Synechocystis* sp. PCC 6803 (NC_020286). Bar represents 0.3 substitutions.
in at least one of the trees. On the other hand, NIES 843 and NIES 102 were included in clade B of the PC-IGS tree (Fig. 3), with a high value of PP and clustered with a moderate statistic support into clade C in the 16S–23S ITS tree (Fig. 4).

One possible explanation for the differences in topology and terminal branches is that some of these strains, especially those with well-supported divergent positions, might be subjected to a speciation process and undergo incomplete lineage sorting. The selection of different evolutionary models...
for the alignment matrices containing the same strains indicated that both markers might not show the same evolutionary pattern. Different regions of the genome may have experienced different evolutionary pressure, promoting an incomplete lineage sorting, leading to a divergence in the fixation rates among the locus and the taxon itself (Maddison and Knowles 2006, Gogarten and Zhaixybayeva 2008). Maddison and Knowles (2006) pointed out that the smaller the divergence time between taxa, the greater the chances of the genetic strains to be subjected to incomplete lineage sorting. Therefore, studies involving taxa at an infrageneric level, such as this study, are more susceptible to this problem.

Another issue related to the use of the marker 16S–23S ITS is the intragenomic variability of its multiple copies present in the genome of a single individual (Boyer et al. 2001, Engene et al. 2010). Boyer et al. (2001) argued that copies with certain sequences may be favored in the annealing of primers and might be preferably amplified at the expense of others, which would pose problems for phylogenies based on these sequences. The use of direct sequencing of the PCR product may have masked this diversity in the studied strains and may also be related to differences between phylogenies obtained in comparison with the PC-IGS marker.

Our analyses do not corroborate the presence of HGT, but this event cannot be neglected as a hypothesis for explaining divergences in phylogenies. A study on the genome of Synechococcus spp., indicated that genes encoding phycocyanin may have evolved independently from genes of the core genome such as the allo-PC gene or the ribosomal regions (Six et al. 2007).

The search for more stable markers, not biased by HGT, has become essential for understanding the phylogeny and taxonomy of prokaryotes (Makarova et al. 1999, Gribalbo and Brochier 2009). The results presented herein strongly support PC-IGS as a marker of choice for cyanobacterial phylogenetic studies and emphasize the importance of using multiple molecular markers to prevent erroneous conclusions based on HGT.

This research was supported by grants from the Brazilian funding agencies FAPESP (State of São Paulo Foundation for the promotion of Research process 2007/57338-2 and 2011/00417-3) and CNPq (National Council for Research and Development 301739/2011-0).

Janson, S. & Granéli, E. 2002. Phylogenetic analyses of nitrogen-fixing cyanobacteria from the Baltic Sea reveal sequence