Members of the genus *Nostoc* are the most commonly encountered cyanobacterial partners in terrestrial symbiotic systems. The objective of this study was to determine the taxonomic position of the various symbionts within the genus *Nostoc*, in addition to examining the evolutionary relationships between symbiont and free-living strains within the genus by analyzing the complete sequences of the nitrogen fixation (*nif*) genes. *NifD* was sequenced from thirty-two representative strains, and phylogenetically analyzed using the Maximum likelihood and Bayesian criteria. Such analyses indicate at least three well-supported clusters exist within the genus, with moderate bootstrap support for the differentiation between symbiont and free-living strains. Our analysis suggests 2 major patterns for the evolution of symbiosis within the genus *Nostoc*. The first resulting in the symbiosis with a broad range of plant groups, while the second exclusively leads to a symbiotic relationship with the aquatic water fern, *Azolla*. 
PHYLOGENETIC ANALYSIS OF THE SYMBIOTIC NOSTOC CYANOBACTERIA AS ASSESSED BY THE NITROGEN FIXATION (*NIFD*) GENE

A Thesis

Submitted to the
Faculty of Miami University
in partial fulfillment of
the requirements for the degree of
Master of Science
Department of Botany
by
Hassan S. Salem
Miami University
Oxford, Ohio
2010

Advisor________________________
(Susan Barnum)
Reader________________________
(Nancy Smith-Huerta)
Reader________________________
(Hank Stevens)
# Table of Contents

## CHAPTER 1

- **Cyanobacteria Life History and Distribution** .......................................................................................... 1
  - Taxonomy and Phylogeny .......................................................................................................................... 3
  - Genus *Nostoc* ....................................................................................................................................... 3
  - *Nostoc* Symbioses ................................................................................................................................. 7
  - Literature Cited...................................................................................................................................... 16

## CHAPTER 2

- **Introduction** ....................................................................................................................................... 26
  - Materials and Methods .......................................................................................................................... 30
  - Results.................................................................................................................................................. 32
  - Discussion............................................................................................................................................ 35
  - Literature Cited.................................................................................................................................. 39
Chapter 1
Introduction to the Nitrogen Fixing Cyanobacteria

Cyanobacteria Life History and Distribution

Cyanobacteria are a taxonomically distinct group of oxygenic photosynthetic, Gram-negative prokaryotes. They are defined by the presence of two photosystems (PSI and PSII) as well as their ability to utilize water as an electron donor during photosynthesis (Woes, 1987). All members within the phylum are capable of photoautotrophy; however, some species have also been found to exhibit facultative chemo-heterotrophy and photo-heterotrophy (Castenholz, 1992). Studies have suggested that cyanobacteria played a pivotal role in converting Earth’s ancient anaerobic atmosphere to an aerobic one, which to a large extent allowed for the diversity of life that exists today (Hayes, 1983; Schopf et al., 1983). Cyanobacteria are believed to have evolved in the Precambrian period, well before the Paleozoic boundary (Weiguo, 1987; Knoll, 1985). Such dating has been strongly supported by the existence of microfossils of the middle and late Proterozoic that are morphologically similar to a number of living cyanobacteria (Schopf, 1996).

Members of the phylum have been described to have one of the most ubiquitous distributions relative to any prokaryote. Given their versatile physiology and wide ecological tolerance, cyanobacteria have been reported to occupy diverse environments including extreme ones such as hot springs, polar zones, hyper-saline seas, and desiccant terrestrial systems (Ferris, 2003; Gordon, 2000; Dor, 1991; Budel, 1991). Cyanobacterial blooms have also been commonly encountered in freshwater environments, particularly when eutrophic conditions are dominant (Zohary, 1991). The domination of cyanobacteria over the world’s oligotrophic oceans, often characterized by clear, nutrient-deficient waters, has been attributed to their adaptation to low light, as well as their efficient nutrient uptake mechanisms (Waterbury, 1979; Olson, 1990; Castenholz, 1992). Their tolerance to desiccation has also been pivotal for their survival in terrestrial systems (Mollenhaur, 1970). Members of this phylum have also been characterized as one of the world’s leading natural suppliers of fixed nitrogen in both aquatic and terrestrial ecosystems (Sprent, 1990). Given the toxic effects that oxygen inflicts on the nitrogenase
enzyme complex, anaerobic conditions generally enhance the fixation of atmospheric dinitrogen (Wolk, 1994). Many cyanobacteria have developed the ability to differentiate into heterocysts, or nitrogen fixing cells, that lack a functioning photosystem II (Sprent, 1990; Wolk, 1994). Alternatively, many cyanobacteria have the ability to fix atmospheric nitrogen aerobically without heterocyst differentiation by forming microniches that are depleted of oxygen (Paerl, 1989; Currin, 1990; Mitsui, 1986). These microniches are formed when following an aggregate prokaryotic formation – similar to that of microbial mats - that is characterized by well-oxygenated peripheries and an anoxic environment in the center of the cyanobacterial clump (Currin, 1990; Mitsui, 1986). Given their photosynthetic capacities, cyanobacteria are also efficient primary producers, and in some instances represent the only carbon fixers within their environments (Chisholm, 1988).

Members of the cyanobacterial genus, _Nostoc_, possess a high degree of morphological complexity, with the ability to differentiate into a number of cell types including akinetes, hormogonia, and heterocysts (Rippka, 1979; Herdman, 1988; Tandeau de Marsac, 1994). Akinetes, or resting spores, are unique to a number of cyanobacteria within Subsection IV and are produced in response to nutrient deficiency and light limitation. The differentiation process that results in the formation of these thick-walled cells is usually coupled with the accumulation of cyanophycin, glycogen and lipids (Herdman, 1988). Depending on the cyanobacterial genus, akinete formation is tightly tied to the location of heterocysts, forming either adjacent to or at a distance from the nitrogen fixing cells (Herdman, 1988; Dok, 1996; 1997). For members of the genus _Nostoc_, akinetes have been shown to differentiate away from the heterocyst (Rippka, 1979; 1989; 2001).

Heterocyst differentiation is a unique developmental process in cyanobacteria from subsection IV and V, where this cell type is produced in trichomes at regular intervals, or the terminal ends of the filaments (Wolk, 1994). Differentiation from vegetative into heterocysts that lack a functional photosystem II is triggered by the decrease of combined nitrogen in the surrounding environment (Wolk, 1994; Chen, 1996; 1998). Throughout the differentiation process, the cell accumulates additional wall layers coupled with a modified thylakoid membrane, which ultimately promotes an anoxic
interior for the synthesis of the nitrogen fixing enzyme, nitrogenase (Wolk, 1994; Yoon, 1998). Once nitrogenase is produced, nitrogen fixation takes place typically in the presence of light where ATP is supplied by cyclic electron transport of photosystem I (Wolk, 1994; Yoon, 1998).

The migrating stage is usually defined by the formation of hormogonia. These short chains (5-15 cells) that move through gliding motility, possess smaller diameters than their vegetative counterparts, and usually form following a number of environmental quos (e.g. phosphors repletion) (Herdman and Rippka, 1988; Tandeau de Marsac, 1994). Studies have also shown that the process of cell division ceases until the hormogonia ends its dispersal period (Herdman, 1988).

**Taxonomy and Phylogeny**

Cyanobacteria classification and taxonomy has been debated and revised numerous times since it was first described in the early 19th century (Rippika, 1979; 1988; 1992; 2001, Castenholz, 1989; 1992; 2001). Early cyanobacterial classifications depended on morphological characters and the phylum was ultimately classified as algae and grouped with other eukaryotes (Golubic, 1976). Later, cyanobacteria were reclassified using bacterial principles, which depended primarily on developmental characteristics (Castenholz & Waterbury, 1989; Rippka, 1988). Cyanobacteria traditionally have been classified into five subsections, I-V. Subsection I species include unicellular strains and divide by binary fission (Rippka, 1979; 1988; 2001). Subsection II strains are unicellular as well, but differ in their ability to divide by multiple fission forming smaller internally formed cells, or baeocytes (Rippka, 1979; Castenholz & Waterbury, 1989). Filamentous cyanobacteria that regenerate by trichome fracture and are non-heterocysts are grouped in subsection III (Rippka, 1979; 1992; Hoffman, 2001). Filamentous cyanobacteria that can differentiate into akinetes and heterocysts are included in subsections IV and V, where they are classified according to their ability to divide in one or multiple planes, respectively. (Rippka, 1979; 2001; Hoffman, 2001).

**Genus Nostoc**

The genus of interest in this study is *Nostoc*. Members of this genus are one of the most widespread cyanobacterial groups and are common in both terrestrial and aquatic
habitats (Vincent 1988). Within terrestrial systems, the most commonly encountered are *Nostoc muscorum* and *Nostoc commune*, which have been shown to form colonies in a wide range of environments including deserts and polar zones (Potts, 1987; Wessels, 1991; Arif, 1992). That being said, wetlands have been established as the optimal environment for *Nostoc* given the high occurrence of this genus among rice paddies and other wetland habitats (Stewart, 1973; Peters, 1991; Vincent, 1988; Johansen, 1983). Additionally, filamentous *Nostoc* strains have also played pivotal roles in the formation of symbiotic relationships with a number of eukaryotic partners, which can be attributed to their photosynthetic and nitrogen-fixing abilities, as well as their capacity to differentiate into motile hormogonia, which is pivotal for the infection and colonization of the different host tissues and organs (Peters, 1986; Rai, 1990; Stewart, 1980; Svenning, 2005). The formation of hormogonia filaments was identified initially as a systematic criterion used to distinguish between *Nostoc* from *Anabaena*, its closest relative (Rippka, 1979). However, the use of this taxonomic feature has proved to be inconsistent in some laboratory cultures where *Nostoc* strains fail to differentiate into the migrating stage (Rippka, 2001; Papeafthimou, 2009).

According to the most recent classification, the genus *Nostoc* has been assigned to subsection IV in addition to genera *Anabaenopsis, Cyanospira, Aphanizomenon, Anabaena, Nodularia, Cylindrospermum, Cylindropermopsis, Scytonema, Calothrix, Galoeotrichia, Tolypothrix* and *Microchaete* (Wilmotte, 2001; Rippka, 1979; 1988; 2001). Cyanobacterial strains belonging to the genus *Nostoc* are the most commonly encountered cyanobacterial partners in exosymbiotic (symbiont inhabits the outer surface of its host) and endosymbiotic (symbiont inhabits the body or cells of its host) associations with ascomycetes, bryophytes, pteridophytes, gymnosperms, and angiosperms, in which they benefit the hosts by a supply of fixed nitrogen (Rasmussen, 1998; Nilsson, 2002; Guevara, 2002; Zheng, 2002; West, 1997; Costa, 2001; Rasmussen, 2002).

The ability to differentiate into heterocysts, akinetes and hormogonia has often been used to define the genus *Nostoc* (Lazarohh, 1966; Molelenhauer, 1970; Rippka, 1979). Terminal heterocysts are found to differentiate at both ends of the hormogonia (Komarek, 1989); giving rise to trichomes containing predominantly intercalary
heteocysts. Another defining feature of this genus is the manner by which akinetes are formed. Akinetes are always initiated distant from heterocysts, often occurring in long chains (Rippka, 1988; Castenholz & Waterbury, 1989). *Nostoc* is closely related to *Anabaena*, but is distinguished morphologically in two respects. First, *Nostoc* strains generally produce more clustered filaments, whereas *Anabaena* forms smaller clusters (Rippka, 1979). Additionally, a number of studies have indicated that hormogonia are more commonly produced by *Nostoc* than by *Anabaena* (Rippka, 1979; Castenholz, 1989), however, under laboratory conditions, some *Nostoc* strains fail to differentiate into their hormogonia stage, making it a less reliable differentiating character (Rippka, 2001; Bergman, 2002).

Cyanobacterial classifications that rely solely on morphological characteristics may not depict accurately the evolutionary relationships within the phylum because an increasing number of the morphological characters utilized to distinguish genera and species are phenotypically plastic and can vary with environmental conditions (Mollenhauer, 1988; Rippka, 1979; 1988; 2001; Rajaniemi, 2005). Examples of such placement errors can be illustrated by the fact that *Nostoc* PCC 7120 was formally classified within the *Anabaena* genus (Lachance, 1981), and *Nostoc* PCC 6720 was previously described as *Anabaenaopsis* (Rippka, 1979). This ultimately encouraged the utilization of molecular data to help resolve evolutionary relationships within the cyanobacterial lineage (Giovannoni, 1988; Wilmotte, 1994; Turner, 1996; Zehr, 1997; Henson, 2002; 2004; Lohtander, 2003; Rajaniemi, 2005; Svenning, 2005; Papeothilmou, 2009).

A number of studies have utilized the phylogenies of the nitrogen fixation (*nif*) genes, *nifH*, *nifD*, and *nifK*, to resolve the evolutionary relationships within the cyanobacterial lineage (Normand, 1989; Zehr, 1997; 1998; 2003; Henson, 2002; 2004). Recent studies have shown *nif* gene phylogenies supporting the monophyly of the heterocystous lineage (subsections IV and V) relative to other cyanobacterial groups (Zehr, 1997; Henson, 2004). This is congruent with 16S rRNA gene sequence data, which also have shown strong support for monophyly (Giovannoni, 1988; Wilmotte, 1994; Turner, 1997). The *nifHDK* operon that encodes the nitrogen fixing enzyme complex, nitrogenase (Figure 1), is composed of two major subunits, dinitrogenase
reductase (iron protein) and dinitrogenase (molybdenum-iron protein) (Ben-Portath, 1994; Tuner, 1997; Zehr, 1997; Henson, 2002; 2004). The smaller and much more conserved nifH gene (~ 0.8 kb) encodes the two identical subunits that comprise dinitrogenase reductase (Maverech, 1980; Postgate, 1982). This component is responsible for the maintenance of the structural integrity of the enzyme complex, and is further involved in mediating the ATP-dependent transfer of electrons to the dinitrogenase tetramer (Postgate, 1982). Dinitrogenase, on the other hand, is composed of two subunits encoded by nifD (~1.5 kb), and two subunits encoded by nifK (~1.58 kb) (Mazur and Chui, 1982; Lammers, 1984), ultimately forming a tetramer that is responsible for binding atmospheric nitrogen (N$_2$), and transferring electrons to it (Postgate, 1982). Given the oxidative effects of oxygen, the enzyme complex, nitrogenase, is irreversibly inhibited when exposed to the gas as a result of the oxidation of the enzyme’s iron-sulfur cofactor (Wolk, 1994; Modak, 2002). Such a condition is overcome in when nitrogenase can be synthesized and maintained in anoxic conditions (Wolk, 1994). Some cyanobacteria have adapted to such a condition by developing the nitrogen fixing cells, heterocysts, ultimately allowing for a spatial separation between oxygen and nitrogenase (Carr, 1983).

Cluster analysis of nif genes from a number of diazotrophic organisms has demonstrated that nifD possesses an intermediate divergence in comparison to the other genes within the nifHDK operon (Mathur, 1990). The study demonstrated a high degree of similarity among nifH genes (82%), when compared to nifD (77%) and nifK (62%). The authors concluded that the high conservation of nifH and nifD suggests important

\[
\begin{align*}
NifH & \quad NifD & \quad NifK \\
0.8 \text{ kb} & \quad 1.5 \text{ kb} & \quad 1.58 \text{ kb}
\end{align*}
\]

Figure 1. Schematic drawing of a contiguous nifHDK operon for nonheterocystous cyanobacteria with an arrow representing the direction of transcription.

Cluster analysis of nif genes from a number of diazotrophic organisms has demonstrated that nifD possesses an intermediate divergence in comparison to the other genes within the nifHDK operon (Mathur, 1990). The study demonstrated a high degree of similarity among nifH genes (82%), when compared to nifD (77%) and nifK (62%). The authors concluded that the high conservation of nifH and nifD suggests important
structural functions for the two genes, while the variability of nifK may have allowed for the adaptation to different physiological conditions.

The gene has been successfully utilized in previous studies and contributed to the determination of the monophyly of heterocystous cyanobacteria within the phylum, which was in congruence with 16S rRNA sequence data (Henson, 2004; Giovannoni, 1988; Wilmotte, 1994; Lohtander, 2003; Rajaniemi, 2005). Additionally, nifD sequence analyses have also been able to successfully differentiate between closely related cyanobacterial genera, Nostoc and Anabaena within subsection IV (Henson, 2002). For those reasons; nifD was selected as the molecular marker gene in this study to help elucidate the evolutionary relationships between free-living and symbiotic Nostoc strains (Chapter 2).

**Nostoc Symbioses**

A number of symbiotic associations with Nostoc have been thoroughly described. Associations include partnerships with algae, fungi, bryophytes, gymnosperms, and a single genus of angiosperms (Rasmussen, 1998; Nilsson, 2000; Guevara, 2002; Zheng, 2002; West, 1997; Costa, 2001; Rasamussen, 2002). Despite the representation of a wide segment of the plant kingdom, only a limited number of genera are involved in such partnerships with Nostoc.

The ability of Nostoc strains to develop heterocysts and hormogonia filaments has been often been cited as the two most important characters behind the successful establishment of plant symbioses (Campbell, 1989; Meeks, 1990; Johansson, 1994). Heterocysts provide the host with a constant supply of fixed nitrogen, while the hormogonia filaments allow the stationary Nostoc filaments, a means of dispersal (Meeks, 1990). Unlike other symbiotic associations, such as that of the legume-Rhizobium partnership, the specific structures occupied by the cyanobacteria in symbiotic partnerships are not a result of infection and the structures differentiate and develop in each host plant regardless of the presence of the symbiont (Peters, 1991; Rai, 1990).

The herbaceous angiosperm genus Gunnera represents the only existing flowering plant that engages in a symbiotic association with a cyanobacterium (Bonnell, 1990; Bergman, 1992; 2002). The geographic distribution of this angiosperm is limited to the
Southern Hemisphere with up to fifty species identified to date (Silverster, 1996; Osborne, 2002). Unlike most other cyanobacterial-plant symbioses, where the cyanobacteria are extracellularly associated with the host, the *Gunnera-Nostoc* association is intracellular with the symbiont surrounded by a plasma membrane-like layer (Silvester, 1976; Johansson, 1992). Regardless of the growth range and environment, all *Gunnera* species have been found to accommodate *Nostoc* endosymbionts exclusively (Reinke, 1872; Bergmann, 1992a; Wanntorp, 2001). This was further demonstrated in the laboratory through reconstitution experiments using *Gunnera manicata*, where only members of the genus *Nostoc* exhibited symbiotic competence relative to other closely related cyanobacterial groups, including the hormogonia producing cyanobacterial genera *Calothrix* and *Fischerella* (Bonnett, 1981; Johansson, 1994). However, more recent studies have established that more than one *Nostoc* strain can infect a single *Gunnera* plant demonstrating a lack of specificity for members of the genus (Nilsson, 2000; Bergman, 2007). This led to the conclusion that *Nostoc* infections are generally influenced by the presence of compatible symbionts as opposed to a selective determinate in the angiosperm (Zimmerman, 1990; Nilsson, 2000; Bergman, 2007). This was also further explained by the wider distribution of the genus relative to other hormogonia forming, heterocystous cyanobacteria, and has been cited as one of the factors behind the success of *Nostoc* for establishing a broad and diverse set of symbiotic relationships (Peters, 1991; Meeks, 1988; 1998).

The ability of vegetative cells to differentiate into hormogonia and heterocysts has been described as the two pivotal prerequisites for the establishment of the *Nostoc-Gunnera* symbiosis (Peters, 1991; 1998). Unlike with *Azolla* symbioses, *Nostoc* strains in the angiosperm have never been found to differentiate into their resting spore stage, the akinete (Soderback, 1990).

The nature of the *Nostoc-Gunnera* association had been perviously under much debate, but an experimental study demonstrated that given the transfer of $^{15}$N from cyanobacterium to the plant, the interaction was deemed a true symbiotic partnership (Silvester, 1969). Similar studies have pointed to the plant’s potential to exclude all other
potential symbionts except for a number of cyanobacteria, indicating that a strict specificity exists between the cyanobacteria and their host (Johansson, 1994; Bergman, 2002).

The cellular differentiation process of *Nostoc* endosymbionts during the establishment of symbiosis with their *Gunnera* host is dramatic, strongly influencing the cyanobiont’s behavioral responses and adaptations (Johansson, 1992; Bergman, 2002). Unlike other cyanobacteria-plant associations, the endosymbiont resides in the stem tissue (Bonnett, 1990; Johansson, 1992; 1994; Nilsson, 2000; Guevara, 2002; Chiu, 2005). Studies have shown that the origins of the *Nostoc*-harboring glands can be detected on the surface of the hypocotyl shortly following germination where usually a single gland appears at the base of each new leaf (Bergman, 2002). Typically, all glands are colonized by the *Nostoc* endosymbiont (Bonnett, 1990; Osborne, 1991). Morphological studies have indicated that the *Gunnera* glands may function as entry points for the *Nostoc* filaments into the angiosperm (Towata, 1985; Chiu, 2005). The formation of these glands was found to begin from a cluster of stem cells that continue mitotic activities and produce an external papilla that eventually ruptures the epidermal layer (Bonnett, 1990; Johansson, 1992). The glands are connected to the stem and the interior of the plant through channels that also aid in the entry of the *Nostoc* filaments (Bonnet; 1990; Uhda, 2001; Chiu, 2005). Cells lining the channels produce carbohydrate-rich mucilage that is thought to support the growth of the cyanobacteria and promote cell division following infection (Bergman, 1992; Rasmussen, 1994; Chiu, 2005). The continuation of cyanobacterial growth and development results in the formation of ‘symbiotic regions’ that provide the *Gunnera* plants with a steady release of fixed nitrogen (Rasmussen, 1994; Bergman, 1992; Uhda, 2001).

The symbiotic relationship between the water fern *Azolla*, and *Nostoc* represents the only known permanent symbiosis among cyanobacteria-plant associations (Peters, 1991; Lechno-Yossef, 2002; Papaefthimiou, 2008). The *Nostoc* symbiont occupies the highly specialized cavities in the dorsal leaf lobes, and is directly passed on from one generation to another (Hill, 1975; 1989; Peters, 1983; Lechno-Yossef, 2002). The plant further retains the cyanobacterial symbiont during sporulation (symbionts are packaged
into sporocaps) in addition to the reproduction of the water fern, which is unique relative to other cyanobacterial-plant symbiotic systems (Peters, 1991; Lechno-Yossef, 2002). The *Nostoc* symbionts occupy the edge of the leaf cavity of the water fern specifically enclosed in a viscous secretion between the internal and external envelopes (Nierzwicki-Bauer, 1989; Uheda, 1991). Once the extracellular cavity is fully formed, the cyanobacterial filaments are positioned into it through the quickly differentiating primary branched hair within the plant trichome (Calvert, 1983, Meeks, 1989; 2002). The trichome and its associated cyanobacteria are then completely engulfed within the cavity and the symbiont begins its differentiation into the heterocystous stage and nitrogen fixation within the host tissue is initiated (Kaplan, 1986; Meeks, 2002). Mature cavities benefit from maximal nitrogenase as the nitrogen fixing heterocysts represent 20-30% of the entire cell population, with a higher differentiation rate for heterocysts, coupled with decreased hormogonia formation (Hill, 1975, 1977; Kaplan, 1986; Kaplan, 1981; Peters, 1980). Despite the higher differentiation rates of heterocysts relative to other cell types, little is known regarding the percentage of the functioning nitrogen-fixing cells, since studies have pointed out the limited lifespan of heterocysts and the terminal nature of their differentiation (i.e. heterocysts, unlike akinetes, can not differentiate back to vegetative form) (Peters, 1991; Placios, 2005). This lead the authors to believe that despite the higher percentage of heterocysts, the nitrogen fixing capacities of symbiotic tissues might actually be uniform across the board (Meeks, 1998; 2002). Older host leaves have been found to harbor higher percentages of akinetes, or resting spores and significantly fewer heterocysts indicating patterns of parallel differentiation and development taking place during the symbiotic association (Braun-Howland, 1988; Braun-Howland, 1990; Peters, 1975; Meeks, 2002). Such a pattern has been attributed to the lower photosynthetic activity of the older leaves and the resulting carbohydrate provided to the cavities and the cyanobacterial partners (Braun-Howland, 1988; Meeks, 2002).

The taxonomy of the *Azolla* symbionts is still under debate. The symbiont was formerly described as *Anabaena azolla* (Strasburger, 1884), and while a number of studies still support such a placement (Wright, 2001), the majority of recent studies
involving morphological and phylogenetic analyses (16S rRNA, 23S rRNA) point towards a closer affiliation with the genus *Nostoc* (Meeks, 1988; Plazinski, 1990; Zheng, 1999; Pabby, 2002; Sood, 2008; Papaefthimiou, 2009). Zheng *et al.* have also pointed out that according to DNA fingerprinting of short tandemly repeated repetitive (STRR) sequences of representative *Azolla* symbionts and their hosts, that the association is highly specific and that phylogeny of the cyanobacteria taxa directly correlates to that of the host (Zheng, 1999). Such results highly support the hypothesis of co-evolution between *Azolla* and its cyanobacterial symbiont. Additionally, phylogenetic (16S rRNA data) and morphological cluster analysis in a recent study demonstrates a clear distinction between other *Azolla* symbionts and other symbionts within *Nostoc* (Papaefthimiou, 2009). Such results are a further indication that the symbiotic association between the aquatic fern *Azolla*, and its cyanobacterial symbionts is different from other *Nostoc*-plant symbioses.

*Nostoc* forms associations with tropical gymnosperms of the division *Cycadophyta*. This represents one of the most ancient seed plants with a total of three families and ten genera. This plant group often exhibits branched caralloid roots that supplement normal root patterning (Lindbald, 1990). The symbiotic *Nostoc* are predominantly found in the caralloid roots, usually within a separate cell layer that is recognizable in root cross sections (Milindasuta, 1975). The cyanobacterial filaments are often localized to intercellular spaces between the inner and outer cortical layers of the host plant (Lindbald, 1985). *Nostoc* symbionts have been found to infect the cycad, *Macrozamia communis*, directly following a break in the dermal layer of the root. This break propagates the bacteria to continue through a cortical channel of intercellular spaces leading to the symbiotic patches found within the caralloid roots (Milindasuta, 1975; Nathanielsz, 1975). Interestingly, *Nostoc* symbionts were found to exhibit a baspidetal mode of cellular differentiation where the apical region of the cycad root generally lacked heterocystous cells. Heterocyst frequencies were then found to dramatically increase below the apical portion and steadily increase gradually in the more mature root region (Grilli-Cailoa, 1980). Despite the predominantly extracellular nature of this association, a number of studies have also noted a limited form of intracellular
symbiosis involving *Nostoc* filaments and *Cycas revolute* and *Macrozamia cumunis* (Nathanielsz, 1975; Obukowicz, 1981).

Cycad partnerships with cyanobacteria have been demonstrated to be specific to the genus *Nostoc*, with only two other associations described for an *Encephalartos* species with the filamentous *Calothrix* (Grobbelaar, 1984; Obukowicz, 1981). On the other hand, a number of studies have indicated a lack of specificity regarding the *Nostoc* symbiont, by demonstrating that several *Nostoc* strains can potentially infect a single cycad species (Lindbald, 1989; Lotti, 1996). This was further demonstrated by a recent study where a number of genetically distinct *Nostoc* strains were found in a single coralloid root (Zheng, 2002). This in obviously in congruence with previous reconstitution studies by Meeks et al (Meeks, 1988; 1992) and raises questions regarding the diversity among the various symbionts within the genus.

The symbiotic relationship between the *Geosiphon pyriformis* and *Nostoc puntiforme* is the only form of an endocytobiotic association linking a fungus with a cyanobacterium (lichens are ectocytobiotic) (Mollenhauer, 1992). *Geosiphon pyriformis* represents a coenocytic fungus with mycelia distributed in the surface layers of damp, oligotrophic soils (Wolf, 1997). *Geosiphon pyriformis* and its *Nostoc* endosymbiont live together in the same habitat where the fungal hypha incorporates the cyanobacterium by a mode of endocytosis (Mollenhauer, 1996; Bergman, 2002). The hyphal tip subsequently swells forming a multinucleate compartment where the incorporated *Nostoc* cells reside, multiply and continue to differentiate into heterocysts to provide the host with a constant supply of fixed carbon and nitrogen (Mollenhauer, 1996; Kluge, 2002). Through scanning microscopy and green light illumination of the multinucleate compartment, the enclosed *Nostoc* strains were found to exhibit a multiplication rate that is twice as high as the one maintained by free-living filaments (Wolf, 1997). The incorporated cells have also been found to increase in volume by up to tenfold relative to the primordial *Nostoc* cells outside the compartment. The incorporation of the *Nostoc* provides the fungus with an external supply of photosynthate, while the symbiont has been found to benefit from the stable water supply in its immediate surroundings (Kluge, 1992; 2002). It has also been reported that the *Nostoc* cells enclosed within the *Geosiphon* compartments retain
their full genomic integrity, suggesting that this symbiotic relationship represents a fairly early state among endocyanoses (Mollenhauer, 1992). *Geosiphon* has been labeled as an obligate symbiotic organism largely due to the inability to cultivate the fungus without its endosymbiont (i.e. the symbiont has been found in every life stage of the fungus) (Schubler, 2000). As opposed to *Nostoc* symbiosis with plants, where heterocyst frequency is evidently amplified as a result of the symbiont’s nitrogen fixing contribution, endosymbiont heterocysts differentiation frequencies are essentially unaffected in *Geosiphon* partnerships (Kluge, 1992; Mollenhauer, 1996). Such patterns seem to suggest that despite the nitrogen fixing capacities of the endosymbiont, the photosynthetic abilities of the cyanobacterium with the fungus plays a significant role (Kluge, 1992).

*Nostoc* has also been reported to be the most common cyanobiont in lichens, specifically with *Lecanorales* (Paulsrud, 1985). A number of molecular markers have been utilized to resolve the phylogenetic relationships of the *Nostoc* cyanobionts from a number of cyanolichen associations (Paulsrud, 1998; 2000; 2001; Rikkinen, 2002). In those studies, tRNA (UAA) intron and 16S rDNA sequence analyses have demonstrated a significantly high level of diversity among the lichen-forming *Nostoc* strains, and grouped them together with free-living strains within the genus. Interestingly, the studies have also shown that despite the high degree of variation among the cyanolichens, the group seems to form a monophyletic cluster among nostocalean cyanobacteria. This was further supported by a recent study where only a single *Nostoc* strain was found per cyanolichen association with *Peltigera*, which demonstrates a very high degree of symbiont specificity (Paulsrud, 2002). Recent studies have also shown that the cluster of cyanolichen symbionts is further divided into two major sub-clusters; one harboring epiphytic cyanolichens, and another that includes terricolous (deep soil) cyanolichens (Lohtander, 2001; Rikkinen, 2002). In an attempt to explain possible morphological differences between cyanolichens and *Nostoc* symbionts from other systems, Paulsrud et al. (2001) have focused on hormogonia formations. In their study, the authors noted that a high number of lichen-forming strains do not produce hormogonia in their associations with the fungi as well as on culture plates. This is obviously an aberration from other
Nostoc cyanobacteria since hormogonia formation has been described by Rippka et al. (1979) to be a one of the diagnostic features of the genus.

Bryophyte associations with cyanobacteria, predominantly Nostoc, are rare and are generally limited to a small group of mosses, liverworts and hornworts (Peters, 1991; Rai, 2002). Of all the known liverworts genera, only four have been found to form partnerships with cyanobacteria (West, 1997; Meeks, 1988). The nature of such relationships is either epiphytic or endophytic (Rai, 2002 Adams, 1999). The endophytic associations are limited to the genera Blasia and Cavicularia, while the Marchantia and Porella genera dominate the epiphytic partnerships (Meeks, 1990; West, 1997). On the other hand, Anthoceros, Phaeoceros, Notothylas, and Dendroceros have been shown to form endophytic associations (Meeks, 1990). The endophytic relationship that Anthoceros maintains with Nostoc has been perceived as an ancient form of plant-cyanobacterial association because of the ability to accommodate the majority of cyanobacterial symbionts through reconstitution experiments (Sprent, 1985). The Nostoc symbiont usually resides in hollowed spaces on the ventral plane of the gametophyte thalli (Enderline, 1983). The hormogonia filaments of the symbiont are thought to gain entry into the mucilage-filled cavities through slime pores.

A 2001 study by Costa et al examined the diversity of Nostoc symbionts in bryophytes, hornworts and liverworts using a number of molecular markers. Their findings demonstrated that a broad range of Nostoc symbionts are present under natural conditions (Costa, 2001; Rikkinen, 2008). The study would further support the hypothesis that given the lower degree of genetic heterogeneity among bryophyte symbionts, there must be a level of sequential continuity among the different populations of symbionts throughout the association with this plant group.

The aim of this present study is to utilize molecular markers, specifically nifD genes, to determine the phylogenetic positions of the various symbionts relative to each other, and to free-living taxa within the genus Nostoc. Such an analysis will help us address a number of questions relating to the evolutionary history of symbiotic behavior within Nostoc, and the evolutionary relationships of free-living and symbiotic strains. The methodologies and results of this study will be described in the following chapter,
which is formatted according to the guidelines of the International Journal of Systematic and Evolutionary Microbiology (IJSEM), where this work will be submitted.
Literature Cited


Henson, B. J., L. E. Watson, et al. (2002). "Molecular differentiation of the heterocystous
cyanobacteria, Nostoc and Anabaena, based on complete NifD sequences." Current Microbiology 45(3): 161-164.


Chapter 2

Phylogenetic Analysis of the Symbiotic *Nostoc* Cyanobacteria as Assessed by the Nitrogen Fixation Gene, *nifD*

Introduction

Cyanobacteria represent a taxonomically distinct phylum of photosynthetic prokaryotes that occupy a diverse range of environments (Castenholz, 1989). Evidence in the fossil record suggests that this group evolved 3.5 billion years ago (Castenholz, 1992). Studies have suggested that cyanobacteria played a pivotal role in converting Earth’s ancient anerobic atmosphere to an aerobic one, which to a large extent allowed for the diversity of life that exists today (Hayes, 1983; Schopf et al., 1983). Members of the phylum have been described to have one of the most ubiquitous distributions of any prokaryote and are known as major suppliers of fixed nitrogen in numerous ecosystems (Schopf, 2000). Cyanobacteria have often been found to form a number of stable symbiotic associations with a range of partners in terrestrial and aquatic environments (Rai, 2000; Carpenter, 2002; Janson, 2002).

Recent cyanobacteria classifications divided the phylum into five subsections, I-V. *Nostoc*, the genus of this study, has been assigned to subsection IV alongside the genera *Anabaenopsis, Cyanospira, Aphanizomenon, Anabaena, Nodularia, Cylindrospermum, Cylindrospermopsis, Scytonema, Calothrix, Galoeotrichia, Tolypothrix* and *Microchaete* (Wilmotte, 2001; Rippka, 1979; 1988; 2001). Members of this subsection are filamentous and divide by binary fission on a single plane (Rippka, 1979; 1988; 2001). The ability to differentiate into heterocysts, akinetes and hormogonia has often been used to define the genus *Nostoc* (Lazarohh, 1966; Molelenhauer, 1970; Rippka, 1979). Terminal heterocysts are found to differentiate at both ends of the hormogonia (Komarek, 1989); giving rise to trichomes containing predominantly intercalary heterocysts. Another defining feature of this genus is the manner by which akinetes are formed. Akinetes are always initiated distant from heterocysts, often occurring in long chains (Rippka, 1988; Castenholz & Waterbury, 1989).

The associations of *Nostoc* strains with ascomycetes (to form lichens),
bryophytes, *Pteridophytes*, gymnosperms, and an angiosperm (*Gunnera*) proclaim them the most widely encountered cyanobacterial partners in terrestrial symbiotic systems with documented distributions across all major continents (Rasmussen, 1998; Nilsson, 2000; Guevara, 2002; Zheng, 2002; West, 1997; Costa, 2001; Rasamussen, 2002). The ability to form heterocysts and hormogonia allows for the successful establishment of plant symbiosis (Campbell, 1989; Meeks, 1990; Johansson, 1994). Heterocysts provide the host with a constant supply of fixed nitrogen, while the hormogonia filaments allow a means of dispersal (Meeks, 1990). Unlike the symbiotic association of the legume-*Rhizobium* partnership, the specific structures occupied by the cyanobacteria in symbiotic partnerships are not a result of infection and such structures differentiate and develop in each host plant regardless of the presence of the symbiont (Peters, 1991, Chiu, 2005). It has also been proposed that the preference of *Nostoc* as a symbiotic partner in terrestrial systems can be explained by the widespread geographic distribution of the genus (Vincent, 1988; Meeks, 1988; 2002).

The ability of *Nostoc* strains to form symbioses with representatives from all divisions within the plant kingdom has prompted a number of studies to investigate the nature of those associations and the diversity of the cyanobacterial strains involved (West, 1997; Costa, 2001; Rasamussen, 2002; Svenning, 2005; Papaefthimou, 2008). *Nostoc* symbionts have been mostly described as secondary symbionts, except for the *Azolla* partnership, where an obligate symbiotic relationship has been recognized (Peters, 1991, Meeks, 1998, 2002). Primary bacterial symbionts are characterized by the continuous symbiotic partnership maintained throughout the various life cycles of the host, and contribute directly to the host fitness (Oliver, 2004). Secondary symbionts also present the host with nutritional benefits following infection, but the symbiosis is considered facultative, as the presence of the prokaryotic partner is not essential to the host’s survival and in many cases the symbiont is not present in the all of the host’s life stages (Baumann, 2005).

Recent studies have demonstrated a high level of strain diversity among the various host species, except for *Azolla*, where there seems to be a high level of genetic homogeneity among the symbionts of this plant group (Rasmussen, 1998; Nilsson, 2000; Guevara, 2002; Zheng, 2002; West, 1997; Costa, 2001; Rasamussen, 2002; Zheng, 1999).
The nature of the symbiotic partnership involving the water fern, *Azolla*, and *Nostoc* has been described as the most obligate of any of the cyanobacterial-plant symbioses (Peters and Meeks, 1986; Papaefthimou, 2008), where the cyanobacterial symbionts have never been found to naturally exist in a free-living state (Peters, 1991; Meeks, 1988; Gebhardt, 1991). Nonetheless, a number of studies have been able to isolate and culture cyanobacteria from the association (Arad, 1985; Gates, 1980; Newton, 1979; Subaramanian, 1988; Tel-Or, 1983). In those studies, the researchers noted that despite the close resemblance of the cultured *Azolla* symbionts to other *Nostoc* strains, there are morphological differences that set them apart from the freshly isolated, and otherwise unculturable, symbionts of the water fern (Newton, 1979, Tel-Or, 1983). Such differences sparked the debate regarding whether the cultured symbionts of *Azolla* are in fact the true primary and obligate symbionts that occupy the leaf cavities of the water fern. Subsequently, a range of immunological studies have demonstrated that the cultured strains from the water fern express a set of different surface proteins than the freshly isolated strains (Arad, 1985; McCown, 1987). Additionally, a couple of studies have used restriction fragment length polymorphism (RFLP) to contrast the two groups and reported that the cultured isolates shared only a few hybridization bands with fresh isolates form the leaf cavity (Franche, 1985; 1987). Moreover, reconstitution experiments with *Anthoceros* have also been conducted on a number of cyanobacteria form the *Azolla* leaf cavities in an attempt to selectively isolate symbiotically competent cyanobacteria (Meeks, 1988; 1990). Such studies have yielded a single cyanobacterial strain (*Nostoc* sp. PCC 9304), which was included in our phylogenetic analysis in addition to the two culturable *Azolla* symbionts (*Nostoc* sp. PCC 9303 and PCC 8307). Previous classifications of the *Azolla* symbionts have identified them as *Anabaena* (Moore, 1969; Lumpkin, 1980), although the debate regarding their true taxonomic affiliation is still ongoing (Plazinski, 1990; Gebhardt, 1991; Svenning, 2005).

Symbiotic associations that *Nostoc* strains form with other plant and fungal groups are facultative partnerships, unlike the *Azolla* association. Here, the cyanobacterial partner undergoes most of the morphological and physiological alterations to accommodate the host’s nitrogen and carbon fixing needs, with minor changes affecting the host’s growth and developmental progressions (Meeks, 2002). Examples of
the various metabolic and physiological modifications that the cyanobacterial partner undergoes in response to the symbiotic association are numerous, and can range from growth inhibition, and an increased rate of the nitrogen fixation (Enderline, 1983; Steinberg, 1991), to a decrease in the rate of photosynthetic activity (Meeks, 1998; Lindbald, 1987; Soderback, 1992; Steinberg, 1989). This raises the interesting question regarding the phylogenetic affiliation of the *Nostoc* symbionts relative to the free-living strains within the genus. Recent studies that have addressed this question have yielded seemingly dissimilar results ranging from a clear distinction between the two groups (Sood, 2008), to an intermixed distribution across the genus (Papaefthimiou, 2008). Given the importance of nitrogenase in the life history of heterocystous cyanobacteria, and the prior success of *nif* genes as conserved and informative molecular markers, much can be gained from this study regarding the nature of the relationship of the two groups, as well as the evolution of symbiotic behavior within the genus *Nostoc*.

Molecular marker genes have been employed in a number of taxonomic studies to resolve relationships within the cyanobacterial lineage (Ranjaniemi, 2005; Svenning, 2005; Lohtander, 2003; Henson, 2002; 2004; Zehr, 1997). In addition to some of the more commonly utilized marker genes (small subunit rRNA, rpoB, rbcLX), a number of recent studies have used the nitrogen fixation (*nif*) genes, with particular attention to the *nif*HDK operon (Ranjaniemi, 2005; Svenning, 2005; Lohtander, 2003; Henson, 2002; 2004; Zehr, 1997). *Nif*D (∼1.5 kb) - which encodes the alpha subunit of dinitrogenase (Lammers, 1984) - has been chosen as the molecular marker for this study. Cluster analysis of *nif* genes from a number of diazotrophic organisms has demonstrated that *nif*D possesses an intermediate divergence in comparison to the other genes within the *nif*HDK operon (Mathur, 1990). The study demonstrated a high degree of similarity among *nif*H genes (82%), when compared to *nif*D (77%) and *nif*K (62%). The authors concluded that the high conservation of *nif*H and *nif*D suggests important structural functions for the two genes, while the variability of *nif*K may have allowed for the adaptation to different physiological conditions. The *nif*D gene has been successfully utilized in previous studies and contributed to the monophyly of heterocystous cyanobacteria within the phylum, which was in congruence with 16S rRNA sequence
data (Henson, 2004; Giovannoni, 1988; Wilmotte, 1994). Additionally, \textit{nifD} sequence analysis has successfully differentiated between the closely related cyanobacterial genera, \textit{Nostoc} and \textit{Anabaena}, within subsection IV (Henson, 2002). Thus, \textit{nifD} may be able to provide sufficient phylogenetic signal to resolve the evolutionary relationship among the different groups within the genus \textit{Nostoc}.

The objective of this study was to employ \textit{nifD} as a molecular marker for a comparative phylogenetic analysis of symbiotic and free-living strains within the genus \textit{Nostoc} to address four main questions: 1) Can \textit{nif} genes phylogenetically differentiate between the free-living and symbiont strains within the genus \textit{Nostoc}? 2) What are the phylogenetic affiliations of the symbiont strains? 3) Is there evidence for convolution between the cyanobacteria and its host? 4) Given the physiological importance of nitrogenase for establishing a functional symbiotic association, is there evidence for positive selection acting on the \textit{nif} genes of the various symbionts relative to free-living \textit{Nostoc}?

\textbf{Materials and Methods}

\textbf{Cultures and growth conditions.} Thirty \textit{Nostoc} strains were obtained from the Pasteur Culture Collection (PCC), Paris, France, including 12 cyanobionts (Table 1). The cultures were grown in lighted shaking incubators at 29°C in two different deviations of BG-11 media (Rippka, 1979; Henson, 2002).

\textbf{DNA extraction.} Genomic DNA was extracted by using a Purgene DNA Isolation kit from Gentra Systems (Minneapolis, MN, U.S.A) with minor modification as illustrated by Henson \textit{et al} (2002).

\textbf{Amplification of \textit{nifD} gene.} Amplification of \textit{nifD} was accomplished using GoTaq® DNA Polymerase (Promega, Madison, U.S.A.). The PCR primers used to amplify \textit{nifD} are shown in table 2. PCR was conducted under the following conditions: 94°C for 2 min, 10 cycles of denaturation at 94°C for 45s, annealing at 53.5°C for 45s, and extension at 72°C for 10 min with each cycle increasing 20s in duration before a final extension of 72 °C for 5min. PCR products were visualized by 0.8% agarose gel using standard methods. PCR products were purified with the QIAquick Gel Extraction Kit (QIAGEN, Valencia, CA).
DNA sequencing and analysis of *nifD*. DNA sequencing was conducted using the Big Dye Terminator Cycle Sequencing Ready Reaction kit and the ABI 310 automated sequencer (Applied Biosystems, Foster City, CA). Complete sequencing of the gene was performed in the forward and reverse directions according to the guidelines and specifications listed in the manufacturer’s manual (Applied Biosystems, Foster City, CA) to ensure the accuracy of the sequence. *NifD* sequences were aligned using Clustal W 1.83 (Thompson, 1997) and verified manually in SeaView (Thompson, 2007).

The best evolutionary model for our Maximum likelihood analysis selected using Modeltest 3.70 (Posada, 2004) was the general time-reversible model with gamma distribution of rates and a proportion of invariable sites (GTR+I+G) with a –\(\ln L\) = 5894.3756. The best tree was determined using the analyses “heuristic” search algorithm with PAUP*4.0 (Swofford, 2007). A molecular clock was not imposed, and the initial branch lengths were computed according to the Rogers-Swofford (Swofford, 2007) approximation method and equal measuring of all characters. Additionally, starting trees were obtained via stepwise addition and a lone tree was recovered at each step. Topological constrains were not enforced. Bootstrap analysis with heuristic search was conducted, and 1000 replicates were used to calculate the statistical confidence for the 50% majority-rule consensus tree (Figure 2).

Bayesian analysis were performed using MrBayes v3.1.1 1 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003) and the default value of 4 Markov chains with preset model parameters (Figure 3). Temperature parameters were set to a value of 0.2 and the Markov chain Monte Carlo length was 1,000,000 generations with resampling taking place every 50 generations. The proportion of trees sampled after burnin that contained each of the observed bipartitions was used to determine the posterior probabilities. Markov chain Monte Carlo (mcmc) chains ran until a convergence stop value of 0.01 was reached.

Protein sequence analysis were performed through the neighbor-joining method, and were statistically tested using bootstrap analysis from 1000 replicates (Swofford, 2007) (Figure 4).

*NifD* structure and sequence divergence. A SeaView (Thompson, 1997) multi-alignment of all the *nifD* amino acid sequences obtained in this study was used to identify
the conserved regions and detect the degrees of sequence similarity among the various symbiotic and free-living *Nostoc* strains. An estimation of DNA sequence divergence for the *nif*D gene was conducted for the thirty *Nostoc* strains utilized in this study and a DNA matrix was generated using pairwise comparisons with each strain (Swofford, 2007). However, for the purpose of space, the values of only eight representative strains from each of the three major phylogenetic clusters are listed in Table 6.

**Analysis of selective pressures on nifD.** To address the types of selective pressures that act on nitrogenase, the rates of non-synonymous substitutions (dN) and synonymous substitutions (dS) were estimated across the completed sequences of *nif*D. For each cluster, a simple star tree was generated with a bifurcation at the deepest node to assign the root. The input file was then executed in codeml (Yang, 1997), where the four models were tested (M1), (M2), (M7), and (M8). (M1) details an independent substitution ratio for each branch, and is identified as the free-ratios model. (M2) is similar to (M1), but it allows for the analysis to feature several ratios, and requires branches in the tree file to be designated using branch labels. (M7) and (M8) compare the null hypothesis from two distributions to the maximum likelihood tree and vary only in the degree by which the branch lengths are specified (Yang, 1997). Based on likelihood ratio tests, the positive selection model (M2) generated the highest likelihood value and was selected as the evolutionary model for our analysis (Figure 4).

**Results**

Comparison of the alpha subunit sequences using SeaView’s multi-alignment analysis indicated five conserved cysteine residues at positions 59, 85, 151, and 245 that were consistent throughout the various *Nostoc* strains of our study, except for the minor symbionts, sp. PCC 9303, and sp. PCC 8307. Both stains diverged slightly at positions 59 and 85 and possessed a slightly longer *nif*D sequence (1510 bp as opposed to 1495). Such variance was further supported in our phylogenetic analysis as well as DNA pairwise comparisons. The results of the DNA pairwise comparisons (Table 5) indicate that the divergence in the *nif*D gene between the free-living strains *Nostoc* PCC 7118 and PCC 7906 to be about 0.1% which signified a high level of homogeneity among the *nif*D sequences of free-living strains. This is a significantly lower than the divergence values
between either of the free-living strains and their symbiotic counterparts *Nostoc* sp. PCC 7422, PCC 9304, PCC 9303, and PCC 8307, in which case, the percentage for sequence divergence ranged between 18.6% and 21.9% for *Nostoc* sp. PCC 7118 and PCC 7906. Interestingly, both the secondary *Azolla* symbionts, *Nostoc* sp. PCC 9303 and *Nostoc* sp. PCC 8307 uniformly demonstrated divergence values of about 20% when the pairwise comparison was conducted with free-living or symbiotic strains. Additionally, the DNA sequence divergence between the two *Azolla* symbionts and the reconstituted symbiont from *Azolla caroliniana* (*Nostoc* sp. PCC 9304) was also around 21%.

The extent of similarity in tree topology obtained from Maximum-likelihood and Bayesian analyses was very high, and the bootstrap values correlated directly to the posterior probabilities, indicating that the phylogenetic relationships among the studied strains inferred from our analyses were largely independent of method-related biases. This was also the case when comparing both nucleotide-based methods to the amino acid analysis using Neighbor-joining. However, because of the conserved nature of *nifD*, a slightly lower resolution of detected. Therefore, in order to avoid statistical redundancy, and lower phylogenetic signals, only the Maximum-likelihood analysis will be thoroughly described in this section. Maximum likelihood analysis based on *nifD* generated three major clusters (Figure 1). Sixteen of the 20 free-living strains were placed in cluster 1. Support for the monophyly of this cluster was moderate, with a bootstrap value of 81% (Figure 2). However, apart from *Nostoc* sp. strain PCC 6302, 100% bootstrap value was assigned to the remaining seventeen strains within cluster 1 (Figure 2). The latter, strongly supported subcluster provided further evidence for the monophyly of the majority of free-living strains utilized in this study relative to their symbiotic counterparts.

Two smaller branches were present within cluster 1; subclusters 1a and 1b. Subcluster 1a was composed entirely of free-living strains, while subcluster 1b contained a *Gunnera* symbiont, in addition to four free-living representatives. In cluster 1a, free-living *Nostoc* sp. strains PCC 7121, PCC 7906, PCC 6310, PCC 6720, PCC 7120, PCC 6411, PCC 7413, PCC 6719, PCC 7118, PCC 7119 and PCC 6705 occurred as an unresolved polytomy. In cluster 1b, free-living *Nostoc* sp. strains PCC 7107 and PCC
7416 were presented as sister strains and were closely grouped with the *Gunnera* symbiont, *Nostoc* sp. PCC 9405. In a separate branch within cluster 1b, free-living strains PCC 7423 and PCC 7524 were grouped together (Figure 1). The separation of clusters 1a and 1b is moderately supported with a bootstrap value of 81-93% (Figure 2).

The second group, referred to as cluster 2, was supported by bootstrap values of 85-100% and contained thirteen *Nostoc* sp. strains, nine of which originated from symbiotic associations, and four from free-living conditions. Of the symbionts, two were from cycads, *Gunnera*, and *Peltigera*, in addition to a single representative from *Geosiphon*, *Anthoceros*, and *Azollae* (Figure 1). Within this cluster, two major branches were supported by a bootstrap value of 87% and are designated as subclusters 2a and 2b. Subcluster 2a demonstrated that the symbionts from cycad PCC 73102, *Azollae* PCC 9304, *Gunnera* PCC 9229, *Anthoceros* PCC 9305, and *Geosiphon* PCC 9503 were closely related. Additionally, very high bootstrap values (95%) supported the monophyly of the lichen symbionts; PCC 9709 and PCC 9316 (Figure 2) separate from the other symbionts within the subcluster. Subcluster 2b contained symbiont strains PCC 7422 isolated from *Cycas*, and PCC 9230 isolated from *Gunnera*, together with free-living sister strains PCC 6314 and PCC 6314/1 (Figure 1).

Cluster 3, was composed entirely of *Azolla* secondary symbionts, including both *Nostoc* sp. strains PCC 9303 and PCC 8307 which form associations with *Azolla caroliniana* and *Azolla filiculoides*, respectively. High bootstrap values (98%) supported the separation of this cluster from the remaining *Nostoc* strains utilized in this study (Figure 2).

To investigate whether the enzyme complex, nitrogenase, was under positive selection in symbiotic strains, the selective pressures acting on the complete sequences of *nifD* across the three major phylogenetic clusters within the genus was examined (Figure 4). Typically, a dN/dS ratio higher than 1 implies positive selection, and given the increased DNA sequence divergence values observed for the symbionts of clusters 2 and 3 (15%-20%), which is usually a signature of positive selection (Baldo, 2010), we expected to dN/dS ratio higher than 1. However, the highest dN/dS ratio that was
obtained was 0.35, which signifies selection acted on the enzyme, but not enough to support positive selection.

**Discussion**

This study addressed the relationship between free-living and symbiotic cyanobacteria of the genus *Nostoc*, with emphasis on differentiating between the two physiologically different groups, in addition to elucidating the phylogenetic affiliation of the different *Nostoc* symbionts in an attempt to investigate whether symbiont distributions correlates to host evolutionary patterns (i.e. whether coevolution is evident in our phylogenetic analysis).

Phylogenetic analyses based on complete *nifD* sequences resulted in three well-supported clusters. Cluster 1, which was found to contain the majority of the free-living strains, is moderately supported with a bootstrap value of 81%. This was contrary to recent phylogenetic studies that utilized the 16S rRNA gene sequences, where no evidence for discrimination between symbiotic and free-living strains (Papaefthimiou, 2008). In this recent study, free-living strains were distributed across the three major clades generated with no clear monophyletic clustering (Papaefthimiou, 2008). The grouping of free-living *Nostoc* strains in cluster 1 may reflect some of the physiological differences known to separate free-living from symbiotic strains within the genus. Those differences include decreased rates of growth, as well as CO₂ and ammonium assimilation for symbiotic cyanobacteria relative to their free-living counterparts (Lindbald, 1987; Soderback, 1992; Meeks, 2002).

A number of studies have shown that the fixation of atmospheric carbon is significantly reduced in cyanobacterial partnerships with photosynthetic eukaryotes (Meeks, 1998, Lindbald, 1987, Soderback, 1992; Steinberg, 1989). Carbon fixation rates in freshly isolated symbionts from Cycads, *Gunnera*, and *Anthoceros* exhibited 0%, 1%, and 12%, respectively, of the rate of a typical free-living *Nostoc* or *Anabaena* (Lindbald, 1987; Soderback, 1992; Steinberg, 1989). It is unclear if this is the case with older symbiont cultures, nonetheless it is clear that cyanobacterial symbionts significantly suspend their photosynthetic activity throughout their association with their host and transform into a heterotrophic state where nitrogen fixation is optimized. Additionally,
when examining growth rates, Nostoc symbionts in Anthoceros were found to exhibit a decreased growth rate coupled with an increase in nitrogenase activity (Steinberg, 1991).

Further studies have shown that the nitrogen metabolism of symbiotic Nostoc strains to be quite distinct from those cultured in free-living conditions. This was initially demonstrated by the significant increase (up to fivefold) in nitrogenase activity in Nostoc symbionts of Anthoceros, Cycads and Gunnera (Steinberg, 1991; Lindbald, 1985; Soderback, 1990). Additionally, free-living cyanobacteria were found to release lower amounts of fixed nitrogen compared to symbionts associated with their eukaryotic hosts (Meeks, 1985; Meeks, 1987, Silvester, 1996).

Interestingly, the inclusion of free-living Nostoc strains PCC 6314, PCC 6314/1, PCC 7807 and PCC 7706 within cluster 2 - which includes 9 of the 13 symbionts utilized in this study - seems to suggest that they share a similar nitrogenase structure and perhaps similar nitrogen fixing capacities. This raises the question regarding the symbiotic competence of these traditionally free-living strains, which was also addressed in West et al. 2007. In their study, a small number of free-living Nostoc strains demonstrated symbiotic competence through a range of reconstitution experiments with axenic cultures of the hornwort, Phaeoceros, and the liverwort, Blasia (West, 1997). The authors noted, however, that when the plants were later infected with their native symbiont, the symbiotically competent free-living strain was often lost within a few generations, indicating that despite their symbiotic competence, the free-living strains were not likely to enter into symbiosis under natural conditions.

Within subcluster 2a, lichen forming symbionts Nostoc sp. PCC 9709 and Nostoc sp. PCC 9316 were found to form a monophyletic group within the subcluster with bootstrap values of 95%. Such a pattern is in congruence with previous findings by Rikkinen et al. (2002) and Svenning et al (2005), where symbiotic cyanobacteria from epiphytic lichen form highly specific groups relative to other symbionts. However, when considering the remaining strains that constitute cluster 2a, we note that no additional patterns for host specificity can be detected given the grouping of symbionts from cycad (PCC 73102), Azollae (Nostoc sp. PCC 9304), Gunnera (Nostoc sp. PCC 9229), Anthoceros (Nostoc sp. PCC 9305), and Geosiphon (Nostoc sp. PCC 9503). This
distribution is in congruence with previous studies addressing the diversity of symbiotic *Nostoc* strains, in which a high level of genetic diversity was demonstrated coupled with low host specificity (Costa, 1999; 2001; Guevara, 2002; Nilsson, 2000; Paulsrud, 1998; Rasmussen, 1998; West, 1997).

According to *nifD* analyses, cluster 3 was established as an entirely separate clade relative to other *Nostoc* strains (Figure 1). Members of this cluster included the *Azolla* symbionts *Nostoc* sp. PCC 9303 and *Nostoc* sp. PCC 8307. Based on Southern hybridizations of *nif* genes from multiple *Azolla* symbionts, Franche et al. (1985) noted a different restriction pattern for *Nostoc* sp. PCC 8307 which set it apart from the identical DNA restriction patterns observed for the otherwise unculturable and obligate *Azolla* primary symbionts. This led the authors to conclude that *Nostoc* sp. PCC 8307 was not the primary symbiont of its respective host, *Azolla filiculoides*, since this strain was 1) culturable in a free-living state, and 2) demonstrated different DNA hybridization patterns from the freshly isolated obligate symbionts of the water fern. The same DNA hybridization patterns were observed for the *Nostoc* sp. PCC 9303, where multiple DNA probes firmly established that the major symbiont of *Azolla caroliniana* was different from the cultured symbiont (Nierzwicky-Bauer, 1986). This led Meeks *et al.* (1988) to establish that *Nostoc* sp. PCC 9303 and *Nostoc* sp. PCC 8307 were in fact minor *Azolla* symbionts. This was further confirmed when *Nostoc* sp. PCC 9303 failed to demonstrate symbiotic competence through reconstitution with *Anthoceros* (Enderline and Meeks, 1983).

Meeks *et al.* (1988) would further differentiate the minor symbionts by indicating that *Nostoc* sp. PCC 9303 possesses an altered *nifHDK* organization, which can explain their strong divergence in our analyses. In their study, they provide evidence through Southern hybridization that the *xisA* gene – a gene required for the excision of an 11 kb DNA segment separating *nifD* and *nifK* – is absent in the minor symbiont. Additionally, by mapping their restriction fragments, Meeks *et al.* (1988) were able to establish that the minor symbiont possess a contiguous *nifHDK* organization, which is contrary to the organization pattern observed in their study for free-living *Nostoc* sp. PCC 7801 and the cultured isolate obtained from *Azolla caroliniana* via association with *Anthoceros*,
*Nostoc* sp PCC 9304 (Meeks, 1988). Such findings are strongly reflected in our *nif*D analyses in terms of the placement of the minor *Azolla* symbionts, *Nostoc* sp. PCC 9303 and *Nostoc* sp. PCC 8307 relative to other *Nostoc* strains (Figure 1). This was further demonstrated in the percentage of sequence divergence between the minor symbionts and the other *Nostoc* strains of our study where the percentage of sequence divergence ranged between 20-21%, in addition to the conserved residues of *The branching pattern of our nifD trees supports the phylogenetic differentiation between free-living and symbiotic *Nostoc* strains, which is congruent with a recent cluster analyses preformed by Sood et al. (2008). Such findings contradict results by Papaefthimou et al (2008) where symbionts and free-living strains were clustered together across three major clades. Previous studies addressing the diversity and specificity of *Nostoc* symbionts have demonstrated a high level of diversity among symbionts from bryophytes, lichen, *Azolla*, *Gunnera* and cycads (West, 1997; Paulsrud, 1998a; 1998b; Costa, 1999; Janson, 1999; Zheng, 1999; 2002; Nilsson, 2000; Papaefthimmiou, 2008; 2009), which, by and large, correlates with the branching patterns of our *nif*D analysis where a high level of genetic heterogeneity was confirmed. That being said, the distribution of the various symbionts across our phylogenetic trees does not correlate with the evolutionary relationships of their respective hosts, suggesting that coevolution was not a factor in the evolution of symbiotic behavior within the genus *Nostoc*. Interestingly, despite the high sequence divergence values between symbionts and free-living strains within the genus, we were unable to obtain a significant dN/dS ratio indicating positive selection. This demonstrates that despite different living habits and some physiological differences of symbiotic and free-living cyanobacterial strains, the structural integrity of the *nif*D gene is much conserved across the genus.
Literature Cited


Meeks, J. C. (2003). "Symbiotic interactions between Nostoc punctiforme, a multicellular


Table 1. Cyanobacterial strains included in this study and references to literature and to GenBanck sequence accession numbers. Sequences designated by (*) were obtained in this study.

<table>
<thead>
<tr>
<th>Classification/Strain</th>
<th>Accession Numbers</th>
<th>Host</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cyanobionts</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nostoc PCC 73102</td>
<td>E47298871*</td>
<td>Macrozamia sp.</td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7422</td>
<td>E47298871*</td>
<td>Cycas sp.</td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 8307</td>
<td>E47298871*</td>
<td>Azolla filiculoides</td>
<td>Tel-Or et al. (1984); Franche &amp; Cohen-Bazire (1985)</td>
</tr>
<tr>
<td>Nostoc PCC 9230</td>
<td>E47298871*</td>
<td>Gunnera chilensis</td>
<td>E. Söderbäck</td>
</tr>
<tr>
<td>Nostoc PCC 9231</td>
<td>E47298871*</td>
<td>Gunnera dentata</td>
<td>Rasmussen &amp; Svenning (1998)</td>
</tr>
<tr>
<td>Nostoc PCC 9232</td>
<td>E47298871*</td>
<td>Gunnera monoika</td>
<td>E. Söderbäck</td>
</tr>
<tr>
<td>Nostoc PCC 9303</td>
<td>E47298871*</td>
<td>Azolla caroliniana</td>
<td>Newton &amp; Herman (1979)</td>
</tr>
<tr>
<td>Nostoc PCC 9304</td>
<td>E47298871*</td>
<td>Azolla caroliniana via Anthoceros</td>
<td>Meeks et al. (1988)</td>
</tr>
<tr>
<td>Nostoc PCC 9305</td>
<td>E47298871*</td>
<td>Anthoceros punctatus</td>
<td>Enderlin &amp; Meeks (1983)</td>
</tr>
<tr>
<td>Nostoc PCC 9316</td>
<td>E47298871*</td>
<td>Peltigaria</td>
<td>Wagner &amp; Rippka</td>
</tr>
<tr>
<td>Nostoc PCC 9405</td>
<td>E47298871*</td>
<td>Gunnera chilensis</td>
<td>E. Söderbäck</td>
</tr>
<tr>
<td>Nostoc PCC 9709</td>
<td>E47298871*</td>
<td>Peltigaria membranacea</td>
<td>Miao et al. (1997)</td>
</tr>
<tr>
<td><strong>Free-living</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nostoc PCC 6302</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 6314</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 6314/1</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 6310</td>
<td>E47298871</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 6411</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 6705</td>
<td>E47298871*</td>
<td></td>
<td>Rippka (1979)</td>
</tr>
<tr>
<td>Nostoc PCC 6719</td>
<td>E47298871*</td>
<td></td>
<td>Rippka (1979)</td>
</tr>
<tr>
<td>Nostoc PCC 6720</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7107</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7118</td>
<td>E47298871</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7119</td>
<td>E47298871</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7120</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7121</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7413</td>
<td>E47298871</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7416</td>
<td>E47298871*</td>
<td></td>
<td>Rippka &amp; Herdman (1992)</td>
</tr>
<tr>
<td>Nostoc PCC 7423</td>
<td>E47298871</td>
<td></td>
<td>Rippka (1979)</td>
</tr>
<tr>
<td>Nostoc PCC 7524</td>
<td>E47298871</td>
<td></td>
<td>Rippka (1979)</td>
</tr>
<tr>
<td>Nostoc PCC 7706</td>
<td>E47298871*</td>
<td></td>
<td>Rippka (1979)</td>
</tr>
<tr>
<td>Nostoc PCC 7906</td>
<td>E47298871</td>
<td></td>
<td>Rippka (1979)</td>
</tr>
</tbody>
</table>
Table 2. Primers used for PCR amplification of *nifD*.

<table>
<thead>
<tr>
<th>Forward Primer</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>H495</td>
<td>*Nostoc PCC 7807, Nostoc PCC 7906, Nostoc PCC 6310, Nostoc PCC 7107, Nostoc PCC 6411, Nostoc PCC 6705, Nostoc PCC 7118, Nostoc PCC 7119, Nostoc PCC 7413, Nostoc PCC 6314, Nostoc PCC 6314/1, Nostoc PCC 7416, Nostoc PCC 7524, Nostoc PCC 7706, *</td>
</tr>
<tr>
<td>NifDK</td>
<td><em>Nostoc PCC 7803</em></td>
</tr>
<tr>
<td>H750</td>
<td><em>Nostoc PCC 6719, Nostoc PCC 7422</em></td>
</tr>
<tr>
<td>H621</td>
<td><em>Nostoc PCC 7121</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reverse Primer</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>K300B</td>
<td><em>Nostoc PCC 7807, Nostoc PCC 7906, Nostoc PCC 6302, Nostoc PCC 6310, Nostoc PCC 7107, Nostoc PCC 6411, Nostoc PCC 7118, Nostoc PCC 6314, Nostoc PCC 7416, Nostoc PCC 6314/1</em></td>
</tr>
<tr>
<td>K169</td>
<td><em>Nostoc PCC 6719, Nostoc PCC 7119, Nostoc PCC 7413, Nostoc PCC 7121</em></td>
</tr>
<tr>
<td>11KBRI</td>
<td><em>Nostoc PCC 7803, Nostoc PCC 6705, Nostoc PCC 7422, Nostoc PCC 7524, Nostoc PCC 7706</em></td>
</tr>
</tbody>
</table>

Table 3. Detailed primer sequences.

<table>
<thead>
<tr>
<th>Primer (F/R)</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K300B (R)</td>
<td>GTACGAAGTAAGCAACGCAACCTTG</td>
</tr>
<tr>
<td>K169 (R)</td>
<td>AGTTCTTTTCACGGTAGTCCC</td>
</tr>
<tr>
<td>11KBRI (R)</td>
<td>TAACCGTGATAAGGGCCGGAGTAATCC</td>
</tr>
<tr>
<td>H495 (F)</td>
<td>GATGCGAAGTAGTATGCTGTAACCC</td>
</tr>
<tr>
<td>H750 (F)</td>
<td>CACTCGTACCTCGTGACACATCG</td>
</tr>
<tr>
<td>H621 (F)</td>
<td>CAAATGATTCATTCGTACC</td>
</tr>
<tr>
<td>NifDK (F)</td>
<td>TGGTCTGGTCGTCGTAATCTA</td>
</tr>
</tbody>
</table>
Table 5. Matrix of percent DNA sequence divergence in the nifD region of eight representative symbiotic and free-living Nostoc strains. Symbiotic strains are designated by an asterisk.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Nostoc PCC 6302</th>
<th>Nostoc PCC 7118</th>
<th>Nostoc PCC 7906</th>
<th>Nostoc PCC 7422</th>
<th>Nostoc PCC 9230</th>
<th>Nostoc PCC 9304</th>
<th>Nostoc PCC 8307</th>
<th>Nostoc PCC 9303</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nostoc PCC 6302 (I)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nostoc PCC 7118 (I)</td>
<td>12.71</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nostoc PCC 7906 (I)</td>
<td>12.31</td>
<td>0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nostoc PCC 7422* (II)</td>
<td>10.64</td>
<td>18.6</td>
<td>18.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nostoc PCC 9230* (II)</td>
<td>11.1</td>
<td>18.47</td>
<td>18.47</td>
<td>3.07</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nostoc PCC 9304* (II)</td>
<td>13.37</td>
<td>19.21</td>
<td>19.21</td>
<td>10.49</td>
<td>9.96</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nostoc PCC 8307* (III)</td>
<td>21.16</td>
<td>21.72</td>
<td>21.97</td>
<td>20.96</td>
<td>20.69</td>
<td>21.39</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nostoc PCC 9303* (III)</td>
<td>20.98</td>
<td>21.9</td>
<td>21.9</td>
<td>21.09</td>
<td>20.56</td>
<td>21.26</td>
<td>0.37</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 1. Maximum-likelihood tree of nifD nucleic acid sequences with a $-\ln L$ score of 5894.3756. Branches are drawn proportional to the inferred amount of change. The tree was obtained using the general time reversible model with a gamma distribution of rates and a proportion of invariable sites (GTR+G+I). Numbers on the right refer to the various clusters and subclusters. Symbionts are designated by *.
Figure 2. Majority rule consensus (50%) of the trees sampled in the Maximum-likelihood analysis. Bootstrap values are designated above the branches.
**Figure 3.** Bayesian tree of *nif/D* nucleic acid sequences. Free-living strains are illustrated in black font, symbionts in red, and minor *Azolla* symbionts in blue. Numbers above the branches are the posterior probabilities generated.
Figure 4. Rooted Majority rule consensus (50%) of the trees generated by the neighbor-joining algorithm inferred from the amino acid sequence of the nifD gene.
Figure 5. Phylogenetic tree generated using the Maximum-likelihood criterion and utilized to display the dN/dS ratios for all three major clusters.
Acknowledgments

We are grateful to Dr. Rosemarie Rippka of the Pasteur Culture Collection for kindly providing rare *Nostoc* strains from her private collection. We thank Dr. John Hawes and his staff at the Center of Bioinformatics and Functional Genomics at Miami University for all their diligent help and insightful comments regarding our sequencing data. We gratefully acknowledge funding by the Academic Challenge research grant provided for by the Department of Botany at Miami University.