Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 1009

The tRNA^{Leu} (UAA) Intron of Cyanobacteria

Towards Understanding a Genetic Marker

BY

JOSÉ LUIS COSTA

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2004

Dissertation presented at Uppsala University to be publicly examined in Lecture Hall, room 3041, Department of Physiological Botany, Uppsala, Friday, October 15, 2004 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Costa, J L. 2004. The tRNA^{Leu} (UAA) Intron of Cyanobacteria. Towards Understanding a Genetic Marker. Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1009. 49 pp. Uppsala. ISBN 91-554-6033-X

The tRNA^{Leu} (UAA) intron has been recorded in the plastid genome of many algae and land plants and was the first intron to be discovered in cyanobacteria. In all known cases it interrupts the tRNA^{Leu} anticodon loop at a conserved position (U-intron-AA). Cyanobacteria are a diverse group of photosynthetic prokaryotes, some involved in symbiotic associations with a wide range of organisms. The most studied associations are those with plants, where strains of *Nostoc* are the common cyanobacterial partner. In this thesis two aspects of the biology of the cyanobacterial tRNA^{Leu} (UAA) intron are focused: first, the use of the intron as a genetic marker for studying the diversity and specificity of two cyanobacterial symbiosis (bryophytes and cycads) and second, the evolutionary patterns of the intron by using the unique data set generated from the diversity analysis.

From the studies, many different *Nostoc* strains are involved in the two symbiotic associations, although no variation was observed within a single bryophyte cavity or cycad coralloid root. Furthermore, a certain level of temporal stability in the cyanobiont composition of the bryophyte population was found and, in the cycad association different coralloid roots from a single specimen may harbor different cyanobacteria. That a minor cyanobiont could have avoided detection is still possible but unlikely. The sequence alignment of the *Nostoc* tRNA^{Leu} (UAA) introns reveals great sequence similarity with size variation only found in the structural element P6b. This element was found to consist of heptanucleotide repeats and of other non-repetitive genetic elements (NIS elements). The sporadic occurrence of the NIS elements indicates recent origins and a mechanism for its dispersal is proposed.

In this thesis new insights are given concerning cyanobacterial symbioses and also on the mechanisms involved in the evolution of an old genetic element: the tRNA^{Leu} (UAA) intron in cyanobacteria.

José Luis Costa, Department of Evolution, Genomics and Systematics, Department of Physiological Botany, Villav. 6, Uppsala University, SE-752 36 Uppsala, Sweden

ISSN 1104-232X
ISBN 91-554-6033-X
urn:nbn:se:uu:diva-4537 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-4537)

This thesis is based on the following papers, which will be referred to in the text by their respective Roman numerals:

I. José Luis Costa, Per Paulsrud and Peter Lindblad. (1999) Cyanobiont diversity within coralloid roots of selected cycad species. *FEMS Microbiology Ecology* 28: 85-91.

II. José Luis Costa, Esperanza Martínez Romero and Peter Lindblad. (2004) Sequence based data supports a single *Nostoc* strain in coralloid roots of cycads. *FEMS Microbiology Ecology* 49: 481-487.

III. José Luis Costa, Per Paulsrud, Jouko Rikkinen and Peter Lindblad. (2001) Genetic diversity of *Nostoc* endophytically associated with two bryophyte species. *Applied and Environmental Microbiology* 67: 4393-4396.

IV. José Luis Costa, Per Paulsrud and Peter Lindblad. (2002) Cyanobacterial $tRNA^{Leu}$ (UAA) intron: evolutionary patterns in a genetic marker. *Molecular Biology and Evolution* 19: 850-857.

V. José Luis Costa, Jeff Elhai, Sarah Cousins and Peter Lindblad. (2004) Dispersal of iterated sequences in the genome of the cyanobacterium *Nostoc punctiforme* ATCC 29133 (manuscript).

TABLE OF CONTENTS

PREFACE

"Almost all aspects of life are engineered at the molecular level and without understanding molecules we can only have a very sketchy understanding of life itself"

FRANCIS CRICK

The evolution of life on Earth has been one of the most intriguing questions for the scientific community. It took billions of years for life to evolve from simple molecules to today's complex life forms. During this development several key processes have taken place, two of which can be associated with cyanobacteria. First was the release of molecular oxygen to earth's atmosphere, making a radical change for how organisms could, or had to evolve. A second key process was the development of the eukaryotic cell. This process, according to the endosymbiotic theory, resulted from three consecutive symbioses, in which the proto-eukaryote became host: first to aerobic bacteria which evolved into mitochondria; second to spirochetes which evolved into cilia and flagella; and third to cyanobacteria which became the plastids of algae and plants (109).

Indeed, symbiosis is one of the most interesting and important ways of living. Today, several symbiotic associations have great importance for us humans and many others are of profound importance for ecosystems. Cyanobacteria are unique in their capacity to participate in symbiotic associations with a remarkable range of eukaryotic hosts including fungi, protists, sponges and plants. I have in this study focused on the symbiotic associations of cyanobacteria with plants (bryophytes and cycads).

The concept of symbiosis, for me, can be understood in a broader sense. For example, the way some genetic elements (e.g. transposons) interact with their host DNA can be interpreted as two genetic entities, associated together and benefiting from each other. The double stranded helix described by Watson and Crick (127) consists, as we now know, of a combination of several sequence/structural elements: coding sequences, non-coding sequences, exons, introns, transposons, retrotransposons, repetitive sequences, etc. The way all these elements interact and evolve can be viewed as symbiotic or as parasitic associations.

This thesis is just a small step towards understanding the use and the evolutionary patterns of one such genetic element – the cyanobacterial $tRNA^{Leu}$ (UAA) intron. This intron was used as a valuable instrument for the study of symbiotic associations, both at the organism level and at the genomic level.

INTRODUCTION

INTRONS

A genome is the complete DNA sequence of an organism. The size of genomes varies greatly (compare e.g. 1.83 Mb of *Haemophilus influenzae* Rd with the 3.15 Gb of the *Homo sapiens* genome) (33, 74, 124). Until the first complete genome sequence became available (33) it was mostly unknown how the large amount of information was organized. Since then, more and more genomes have been sequenced, and what was though to be only information storage in coding and non-coding DNA sequences, has been revealed to be much more complex. It is now known that genomic DNA is a very dynamic system with many different genetic elements interacting and evolving together with the host DNA. Introns are such a genetic element.

Introns are stretches of non-coding DNA that interrupt the coding sequences of many genes. Introns can be defined as part of a primary transcript (or the DNA encoding it) that is removed by splicing during RNA processing and is not included in the mature, functional mRNA, rRNA, or tRNA. Sometimes they are also referred to as intervening sequences (56). Introns are a heterogeneous group of genetic elements found in all major groups of organisms. They have different functional capacities and structural features and, depending on the requirements for their splicing and on the nature of the splicing mechanism, they can be classified into four major groups (8, 78): (i) spliceosomal introns, (ii) group I introns, (iii) group II introns, and (iv) archaeal introns. (i) Spliceosomal introns are the conventional introns of eukaryotic cells. They are transcribed into RNA and are excised at the RNA level during RNA processing by spliceosomes (protein-RNA complex) (78). This type of intron is thought to have arisen rather recently in evolution (61), and might have evolved from group II introns (16). (ii) Group I introns can catalyze their own excision and do thus not require the presence of a spliceosome (19). They are widely distributed, interrupting a variety of mitochondrial and plastid genes, nuclear rRNA genes, bacterial tRNA genes, and protein coding genes of eukaryotic organisms and viruses (110). Some may also act as mobile elements (53). An example of a group I intron is the $tRNA^{Leu}$ (UAA) intron, which has been studied in this thesis. (iii) Group II introns, like group I introns, can catalyze their own excision without the need for a spliceosome. They differ from group I introns in the nature of the splicing reaction which is more similar to that of spliceosomal introns (70). They are also less abundant than group I introns. Group II introns have been found in protein coding, tRNA and

rRNA genes of organelles, but not in genes encoded outside of organelles (70, 110). Some introns from this group are also mobile (53). (iv) Archael introns are only found in tRNA and rRNA genes of archaea and, unlike group I and group II introns, are not self-splicing (46). Archael introns splice via an endoribonuclease that cuts a "bulge-helix-bulge" motif which is formed at the exon-intron junction (64).

Since the discovery of introns in the 1970s, their origin has been intensely debated. A number of proposals have been put forward, but the main debate has been between two opposing hypotheses. The "intron early hypothesis" proposes that introns are ancient genetic elements already present in the universal ancestor (34, 114). In such a scenario, the observed distribution of introns could be explained by multiple losses in different lineages during evolution (24) and by their mobility, which is assumed to be a derived feature (10). The other theory, the "intron late hypothesis" proposes that introns evolved relatively recently, splitting originally uninterrupted genes and that they are gradually accumulating in eukaryotic genomes (84). In this scenario, horizontal transfer and transposition of introns are frequent events, accounting for the scattered distribution of introns (61).

Group I introns

Group I introns have the capacity to splice themselves out of the genes where they are located (19). Indeed, they were the first example of an RNA molecule with catalytic capabilities discovered, making them the first RNA enzyme or ribozyme to be found. Group I introns are found in both bacteria and eukaryotes (50, 68, 100, 134). The first group I intron to be discovered in bacteria was the $tRNA^{Leu}$ (UAA) intron (50, 134). Typically, in bacterial chromosomes, they are found interrupting structural RNA genes like tRNA or rRNA and not in protein coding sequences (14). Four different tRNAs are known to be interrupted by group I introns: $tRNA^{Leu}$ (UAA), $tRNA^{Arg}$ (CCU), $tRNA^{fMet}$, and $tRNA^{Ile}$ (CAU). In contrast, group I introns present on bacterial phages are localized in phage protein coding genes, but not on structural RNA genes (28). The only known exception to this so far is the intron interrupting the *recA* gene of *Bacillus anthracis* (48).

The splicing mechanism common to group I introns was first described for the *Tetrahymena thermophila* large rRNA intron (17, 137). The self-splicing reaction is initiated by a nucleophilic attack of an exogenous guanosine at the 5' splice site, and proceeds by a second trans-esterification reaction between the generated free 3' end of the 5' exon and the 3' splice site (17, 19, 137). Apart from the cofactor guanosine, only magnesium ions are required for the splicing to take place (21, 37, 95). After

excision from the precursor RNA and ligation of the flanking exon sequences, some group I introns circularize via an additional trans-esterification (36, 137). The ability of group I introns to catalyze their own splicing is dependent on their highly conserved secondary and tertiary structure (19, 118). The folding of the intron results in the formation of a core structure bringing together key residues that are separated in primary sequence. This core structure corresponds to circa 120 nucleotides and is composed of two domains, the P4-P6 (P5-P4-P6) and P3-P9 (P8-P3-P7-P9), formed by coaxially stacked helices (Fig. 1). The P3-P9 domain wraps around the P4-P6 domain, forming a cleft into which the substrate helices P1 or P10 dock (18). Long range tertiary interactions connecting these domains are required to achieve the overall globular shape (19, 71, 110, 121, 126). The boundaries of group I introns are marked by a U residue at the 3' end of the 5' exon and a G residue at the 3' end of the intron (18).

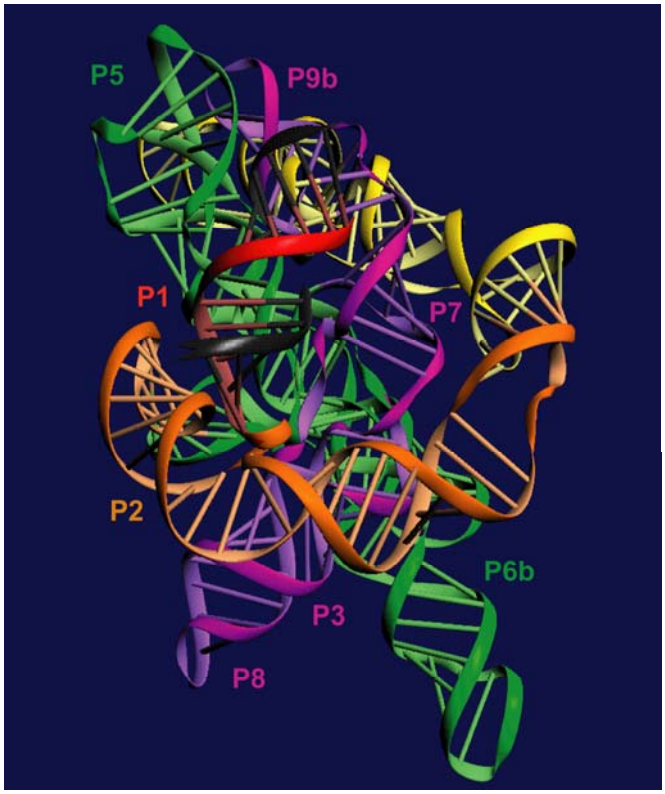


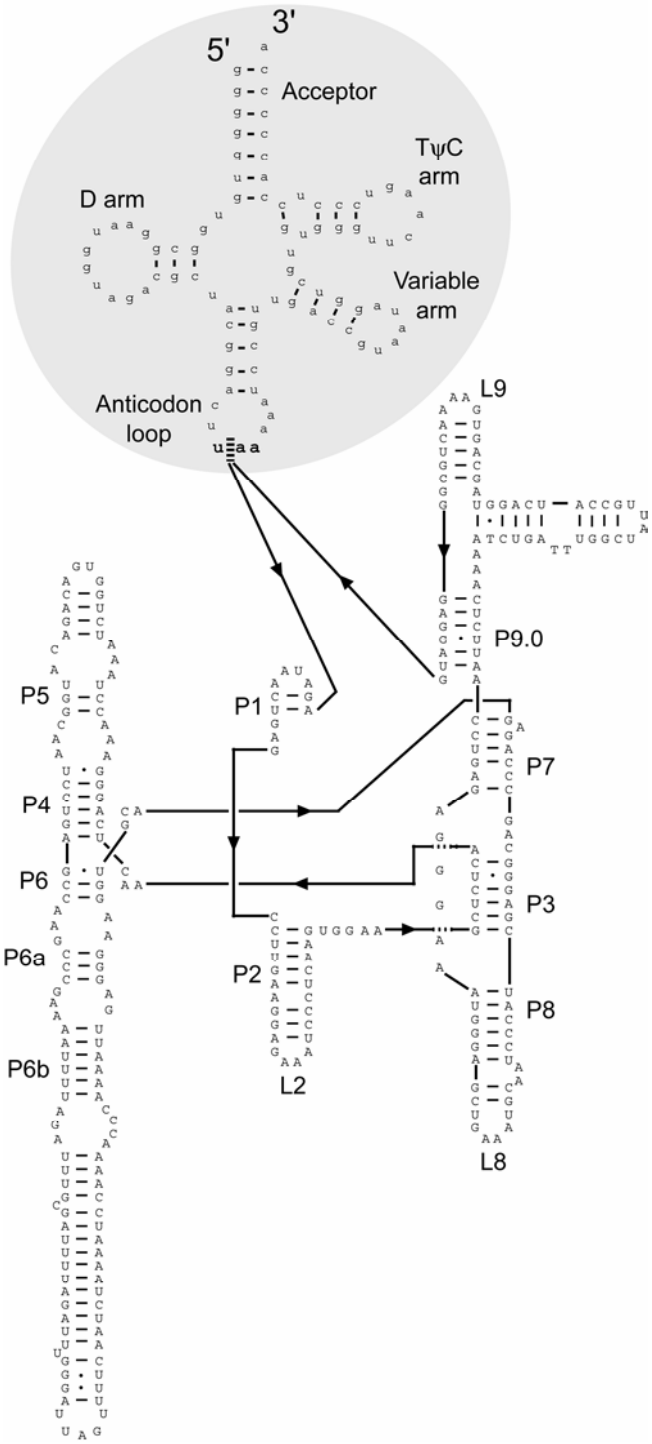
Fig. 1. Three-dimensional architecture of the catalytic core of the group I intron of the *Tetrahymena thermophila* large ribosomal RNA in the classical orientation with the different domains color coded. The P4-P6 domain is shown in green, the P3-P8 domain in purple, domain 2 in orange, domain 9 in yellow, and the two exons in black base pairing with the internal guide sequence (P1) at the 5' end of the intron (red). Modified with permission from Lehnert et al (55).

The conserved secondary structure of group I intron was first deduced from sequence comparisons, with many of the base-pairing regions of the model confirmed by site-specific mutagenesis and structure mapping (18, 71). This secondary structure representation provided a model for considering questions of RNA structure, folding pathways, phylogenetic relationships and evolution (20). Based on structural and folding studies, group I introns are divided into 5 major groups, with 12 subclasses, designated IA1-3, IB1-4, IC1-3, ID and IE (71). For example, subclass IA introns contain extra pairings between P3 and P7, whereas subclass IB and IC (e.g. $tRNA^{Leu}$ (UAA) intron) have an extended P5 domain. After Michel and Westhof (71) modeled the catalytic centre of the group I introns, several studies established new tertiary interactions involving peripheral structures of the intron (40, 41, 55, 63, 69, 75). These new interactions are important for the self splicing capacity of group I introns and were the basis for the currently accepted two-dimensional diagram that more accurately represents the domain organization and helices orientation within the intron (20). Figure 2 illustrates such a diagram for the *Nostoc* sp. strain PCC 73102 $tRNA^{Leu}$ (UAA) intron. Recently, the X-ray crystal structure of a complete group I intron in complex with both the 5' and the 3' exons was obtained (3).

Some group I introns have been shown to be mobile genetic elements that have developed mechanisms to insert themselves into intron-less genes in a process called homing (27). This site-specific insertion is achieved using an endonuclease encoded within the intron. Indeed, approximately 30% of group I introns contain open reading frames (ORFs) (54). Some of these ORFs encode site-specific DNA endonucleases that function in intron mobility and in some cases have adapted to function in RNA splicing (9, 53). The location of these ORFs varies within the conserved RNA secondary structure, but generally they are found in positions that do not interfere with the folding of the catalytic core (54). The finding of mobile, non-intron-encoded endonucleases strongly supported the proposal that ORFs encoding such endonucleases are independent genetic elements that were inserted into previously existing group I introns (53, 62). A typical group I intron is the $tRNA^{Leu}$ (UAA) intron in cyanobacteria. This genetic element is the center of this thesis and will be further described.

→

Fig. 2. Secondary structure of the cyanobacterium *Nostoc* sp. PCC 73102 pre-$tRNA^{Leu}$ (UAA). Upper and lower case letters indicate the intron and the exon (tRNA) sequences, respectively. The anticodon triplet is bold and a dashed line indicates the exon-intron boundaries. The tRNA sequence is folded in the traditional cloverleaf shape and the conserved base pairings are indicated. The intron structure is based on the secondary structure prediction for the *Anabaena* sp. PCC 7120 $tRNA^{Leu}$ (UAA) intron. Regions of base pairing (P1 to P9) and three loops (L2, L8-9) are indicated.


3'
5'
Acceptor
TψC
arm
D arm
Variable
arm
Anticodon
loop
L9
P9.0
P5
P1
P7
P4
P6
P3
P6a
P2
P8
P6b
L2
L8

tRNA^{Leu} (UAA) intron

The tRNA^{Leu} (UAA) intron is an important group I intron. The intron has been recorded in *Proteobacteria* (*Pseudomonas*), in various cyanobacteria and in the plastid genomes of many algae and land plants (50, 86, 115, 125). In all known cases, the intron interrupts the tRNA gene at a conserved position in the anticodon triplet, between the second and third base (U-intron-AA). This intron contains two characteristic peripheral elements, P2 and P9, with GAAA tetra-loops at their termini, feature of the IC3 subclass of group I introns (20). The catalytic core is conserved among the different organisms. However, the plastid introns of land plants were shown not to be able to self-splice *in vitro*, presumably dependent on a host factor needed to facilitate *in vivo* excision (115). The tRNA^{Leu} (UAA) intron was shown not to contain any endonuclease encoding ORFs (50, 86, 115, 125), thus reducing the possibility that it was acquired independently by horizontal transfer in the various lineages that possess it.

The discovery of this intron in cyanobacteria came as a new argument in the debate of intron origin (50, 134). The comparison of the intron sequences from cyanobacteria and chloroplasts revealed great homology and indicated that they are more closely related to each other than to any other known sequences. This observation was used both by the supporters of the intron early and late hypotheses. While the late origin supporters used this as an evidence for recent propagation of a mobile element, the early origin supporters suggested that this intron must have been present before the endosymbiotic incorporation of chloroplasts, thus making the tRNA^{Leu} (UAA) intron the most ancient intron found so far (86). Still, the evolutionary history of the tRNA^{Leu} (UAA) intron is under debate. Although its present day distribution is consistent with a history of strictly vertical transmission (12, 86), it has been suggested that the intron is of more recent origin and has been introduced through lateral transfer (105-107).

The best studied tRNA^{Leu} (UAA) intron in cyanobacteria is the one present in *Anabaena/Nostoc* sp. strain PCC7120[*] (20, 32, 63, 71, 86, 121, 134, 136, 138). This intron is shorter compared to other group I introns. This is caused by the absence of large peripheral domains known to stabilize larger group I introns. These peripheral domains include the P5abc domain and the P2.1 and P9.1 domains that nucleate folding and stabilize the catalytic core of the *Tetrahymena* intron (55). A number of studies have been

[*]The classification of this strain, either as *Anabaena* or *Nostoc*, has been controversial (52, 103, 120). Throughout this thesis it will be referred to as *Anabaena* sp. PCC 7120, the name used in the genome project (44).

dedicated to the folding and splicing mechanism of this particular intron (32, 63, 121, 136, 138). Using phosphorothioate interference assays it was shown that even though the positions affecting splicing were located primarily in the conserved core of the intron, distinct sites outside of the core were also found to interfere with the folding and stabilization of the structure of the catalytic core itself (63). Furthermore, it has been shown that the formation of the anticodon stem in the pre-tRNA is required for optimal splicing of the intron (138). This interaction between the tRNA and the intron is believed to compensate for the short exon-intron pairings in the substrate helix (P1) and apparent absence of helix P10 (138). This was suggested to reflect a possible adaptation of the intron to its location (86, 138).

CYANOBACTERIA

Cyanobacteria are a group of microorganisms that have existed for approximately 3500 million years (112). Their capacity for oxygenic photosynthesis is responsible for one of the major changes on Earth - the introduction of oxygen into the atmosphere. Formerly known as blue-green algae, cyanobacteria are a diverse group of Gram-negative bacteria containing two photosystems (PSII and PSI) and chlorophyll *a* (15). All cyanobacteria can live photoautotrophically, deriving ATP and NADPH required for carbon dioxide fixation from photosynthesis. Many cyanobacteria can also grow photoheterotrophically, obtaining energy from light and the carbon and reducing power from organic compounds. Some strains can also grow anaerobically in the dark by fermentation (15). A small number of cyanobacteria can even switch from oxygenic to anoxygenic photosynthesis using hydrogen sulfide as an electron source (15). In addition to photosynthesis, many cyanobacteria also have the capacity to reduce atmospheric nitrogen to ammonium, a capacity restricted to a small number of prokaryotic organisms (135). The enzymes involved in nitrogen fixation are very oxygen sensitive, making this process difficult to combine with oxygenic photosynthesis (39). Cyanobacteria have evolved different strategies to cope with this problem. Some strains separate the two processes in time, photosynthesizing during the day and fixing nitrogen during night time (e.g. *Gloeothece*) (73). Another strategy is the separation of the two processes in space. One example of this is *Trichodesmium*, which restricts nitrogen fixation to a subset of cells of the filaments (57). Some other cyanobacteria (e.g. *Nostoc*) have developed a specialized type of cell, the heterocyst, where nitrogen fixation takes place (2). The heterocyst is both structurally and physiologically different from the adjacent vegetative cells where oxygen-evolving photosynthesis takes place. Characteristics of the

heterocyst include lack of an active PSII, a high rate of respiration and extra cell wall layers limiting oxygen diffusion (31).

The physiological flexibility of cyanobacteria has made them a successful group in nature. Indeed, cyanobacteria are present in almost all environments on Earth (132). They are important components of many soils ranging from the tropics to the polar regions. They are also abundant in areas which are partially wet or submerged during certain periods of the year (e.g. rice fields) (131). In the vast oceanic environments, with nitrogen depleted waters, planktonic nitrogen fixing cyanobacteria like *Trichodesmium* and *Richelia* can successfully compete and form massive blooms. In subsurface waters, cyanobacteria like *Synechococcus*, *Synechocystis* and *Prochlorococcus* are common, and often dominate, accounting for more than 50% of the biomass (83). Freshwater cyanobacterial blooms are also frequent and are formed by filamentous strains (e.g. *Anabaena*, *Gloeotrichia*, *Spirulina*), but also non-filamentous strains are present (e.g. *Microcystis*) (81). Cyanobacteria are also found in some of the most extreme habitats on earth, such as thermal springs, hypersaline environments, and deserts (132). They are also involved in symbiotic associations with different organisms (1, 94). In these associations, the cyanobacteria provide the different hosts with fixed carbon (e.g. dinoflagellates, lichens and sponges), with the product of nitrogen fixation (e.g. plants and lichens) or with both products of carbon and nitrogen fixation (e.g. diatoms and lichens) (1). The presence of chloroplasts in plants is also the consequence of an old cyanobacterial symbiotic event. In fact, chloroplasts are still similar to cyanobacteria, even though most of their genetic material has been transferred to the host nucleus (108).

Classification of cyanobacteria

Cyanobacteria show great morphological variation. They may be unicellular (e.g. *Synechocystis* and *Cyanocystis*) or filamentous (e.g. *Nostoc* and *Stigonema*), and filaments may be branched (e.g. *Stigonema* and *Fischerella*) or not (e.g. *Spirulina* and *Nostoc*). Many filamentous cyanobacteria also form hormogonia. Hormogonia are modified short filaments, motile, and associated with dispersal, also playing a crucial role in the formation of symbiotic associations. Some cyanobacteria also differentiate akinetes and heterocysts. Akinetes are formed under adverse conditions and function as spores. Heterocysts are a specialized type of cell involved in nitrogen fixation.

A first attempt for a cyanobacterial taxonomy based on the Code of Nomenclature of Bacteria was developed in the 1970s (103). In the second edition of Bergey's manual of systematic bacteriology, the cyanobacterial phylum is divided into five subsections (15):

- Unicellular strains
  - Subsection I: reproduction by binary fission or by budding
  - Subsection II: reproduction by internal multiple fission or multiple followed by binary fission
- Filamentous strains
  - Subsection III: reproduction by binary fission in one plane; cells do not differentiate into heterocysts and akinetes
  - Subsection IV: reproduction by binary fission in one plane; cells may differentiate into heterocysts and akinetes
  - Subsection V: reproduction by binary fission in more than one plane; cells may differentiate into heterocysts and akinetes; branched filaments

This more recent division of cyanobacteria does not correspond to a phylogenetic grouping of genera because the phylogeny of cyanobacteria is still unresolved (15).

Subsection IV - Nostoc

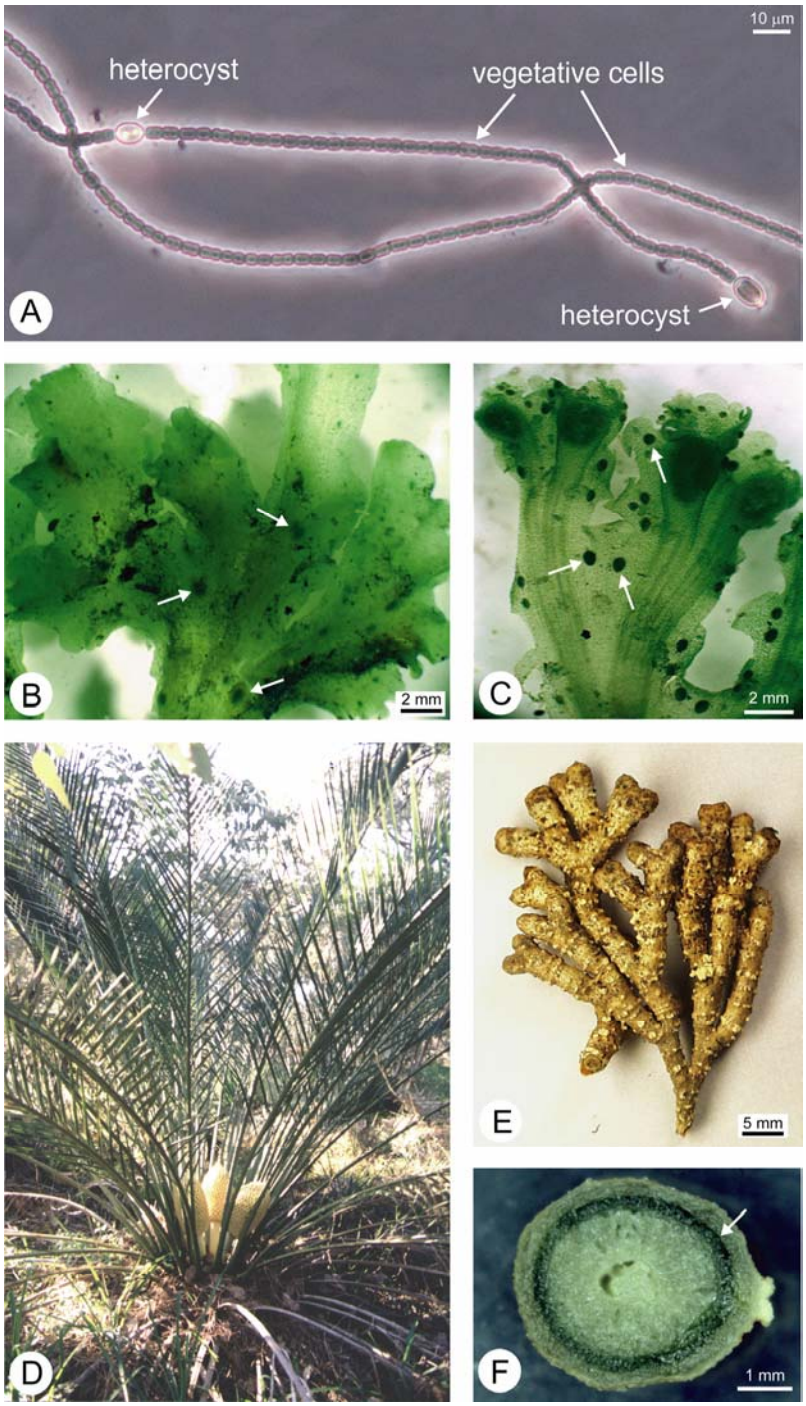
Species of *Nostoc* are common in a large number of ecosystems, including fresh waters, tropical, temperate and polar terrestrial biotopes, but they are rarely found in marine habitats (93). The filaments often occur in colonies ranging in size from microscopic to macroscopic dimensions. *Nostoc* strains are characterized by a complex life cycle and by growth as unbranched filaments and the production of up to four kinds of differentiated cells (vegetative cells, heterocysts, hormogonia, and akinetes) (15, 103). Heterocysts develop in response to the lack of combined nitrogen at regular intervals along the filaments (15). Motile hormogonia are formed in response to certain environmental stimuli (e.g. phosphorus depletion, light conditions and plant signals) (93). This type of short filaments consists of smaller cells, a consequence of their formation by cell division from vegetative cells without cell growth or DNA replication. In addition, *Nostoc* species may differentiate spore-like structures termed akinetes in response to nutrient limitation. The complex life cycle of *Nostoc* represents physiological and genetic capacities that are unknown in other genera (93). Furthermore, many *Nostoc* species are involved in symbiotic associations with a variety of organisms from different taxonomic groups such as fungi and plants (1). In

the initiation of these associations, hormogonia play an important role as the mobile infective unit.

The type example of a *Nostoc* is the strain *Nostoc* sp. PCC 73120 (or ATCC 29133, the coidentical strain in the American Type Culture Collection, also known as *Nostoc punctiforme*) (Fig. 3a). The vegetative cells of this strain are 5 to 6 µm in diameter, the cells of hormogonia between 1.5 to 2 µm, heterocysts 6 to 10 µm and the akinetes range from 10 to 20 µm. This strain was originally isolated from a symbiotic association with the gymnosperm cycad *Macrozamia* sp. (103), and it was the strain of choice when JGI (Joint Genome Institute - www.jgi.doe.gov) decided to sequence a *Nostoc* genome (the genome project refers to this strain as *Nostoc punctiforme*). The *N. punctiforme* genome is approximately 9.78 Mb with a 45.2 mol % GC (66). It is significantly larger than any other cyanobacterial genome that has been sequenced and is also the largest microbial genome ever sequenced. Sequence analyses of the *N. punctiforme* genome reveal an unusual amount of repeated DNA sequences. About 1.5% of the genome (approximately 7.5% of intergenic sequences) is taken up by tandem repeats at least 20-bp in length, with repeating units of 7-bp being by far the most common (66).. The complete understanding of this genome is likely to reveal information about the regulation of the multiple developmental pathways that this genus is capable of, about the regulation of carbon and nitrogen fixation, and also about the regulation of the symbiotic associations that *Nostoc* can be involved in.

→

Fig. 3. Biological material studied in this thesis. (A) Phase contrast micrograph of the filamentous cyanobacterium *Nostoc* sp. strain PCC 73102 showing vegetative cells and heterocysts. (B) The gametophyte thallus of *Anthoceros fusiformes* with the *Nostoc* cyanobiont in slime cavities (arrows). (C) The gametophyte thallus of *Blasia pusilla* with the *Nostoc* cyanobiont in hemispherical auricles (arrows). (D) Male specimen of *Macrozamia riedlei* in its natural habitat in Western Australia. (E) *M. riedlei* coralloid root with established symbiotic cyanobacteria. (F) Cross section of a *M. riedlei* coralloid root. The *Nostoc* cyanobiont is located in the cyanobacterial zone, seen as a green circle (arrow).


10 µm
heterocyst
vegetative cells
heterocyst
A
B
2 mm
C
2 mm
D
E
5 mm
F
1 mm

Symbioses

Cyanobacteria form symbiotic associations with a variety of organisms such as fungi (in the formation of lichens), sponges, protists and plants (1). The best studied cyanobacterial associations are those with plants (Table 1) (67, 94). Another ecologically important symbiosis is the association with fungi in the formation of lichens (87). *Nostoc* species are the most common cyanobacterial genus present in symbiotic associations. With plants, *Nostoc* strains can be found in associations with representatives from bryophytes (mosses, hornworts and liverworts), pteridophytes (the water fern *Azolla*), gymnosperms (cycads) and angiosperms (*Gunnera*) (1, 67, 94). In all these cases it is the capability of *Nostoc* to fix atmospheric nitrogen which is of interest for the photosynthesizing plant host. In plants, the structure where the cyanobacterial partner is housed is formed in the absence of the symbiont (42, 92, 94, 128). The plant may enhance their chances of invasion by producing and releasing chemical signals that induce and attract the hormogonia into the plant tissue (11, 13, 35, 47, 96). Once inside the plant structures, the cyanobacterial filaments undergo considerable morphological and physiological changes. These changes result in a reduced growth rate, reduced carbon dioxide fixation, enhanced frequency of heterocyst differentiation and an increase in nitrogen fixation. In general, it is believed that the cyanobacterial cells obtain their carbon from the host in exchange for the products of nitrogen fixation (1, 94).

Different lines of study have tried to address the identity of the *Nostoc* symbiont. Reconstitution experiments with axenically grown plants have shown that several, but not all strains of *Nostoc* can form symbiotic associations with different hosts. These experiments, performed in both liverworts, hornworts, *Gunnera* and cycads, revealed that free-living and symbiotic strains isolated from lichens, bryophytes, cycads and *Gunnera* could efficiently infect the plant partner (30, 43, 82). Several molecular/genetic methods have also been employed to establish the identity of the cyanobacterial partner. However, since different methodologies have been used, some based on banding patterns (RFLP, fingerprinting, and DGGE) and others on DNA sequences, the results are difficult to compare (38, 45, 60, 79, 89, 98, 99, 102, 129, 139-141).

Part of this thesis is focused on the study of the diversity/specificity of two cyanobacterial associations, the bryophytes (hornworts and liverworts) and the cycads. Due to this, a more detailed description of these two systems follows.

Table 1. Characteristic features of some *Nostoc* associations.

Host		Structure infected	Cyanobiont location	Heterocyst frequency[a]	Form of combined nitrogen transferred
Fungi	Bipartite	Thallus	Intercellular	4-8%	NH_4^+
	Tripartite	Cephalodia	Intercellular	10-55%	NH_4^+
Bryophytes	Mosses	-	Epiphytic	-	-
	Liverwort	Auricles	Intercellular	30-50%	NH_4^+
	Hornworts	Cavities	Intercellular	30-50%	NH_4^+
Pteridophytes	*Azolla*	Leaf cavity	Intercellular	25-40%	NH_4^+
Gymnosperms	Cycads	Coralloid roots	Intercellular	40-45%	Citrulline/glutamine
Angiosperms	*Gunnera*	Stem glands	Intracellular	60-80%	NH_4^+

[a]Heterocysts frequencies expressed as percentage of total cells. Typical values for free-living cyanobacteria are 3-10%. There is often an increasing gradient in heterocysts frequencies from newly colonized structures to older tissues.

Bryophytes

Bryophytes are generally considered to be primitive plants. Nevertheless, they form a successful group which is able to survive under environmental extremes where few other plants survive. This success has lead to a world wide distribution, from arctic and alpine environments to the humid tropics, from semiarid sites to submerged in water (111).

In general, the cyanobacteria form epiphytic associations with mosses and are endophytic with liverworts (e.g. *Blasia pusilla* L.) and hornworts (e.g. *Anthoceros punctatus* L.). In all cases, the associations appear to be confined to the bryophyte gametophyte (111). The symbiotic colonies are visible as dark spots within the plant tissue (Fig. 3b-c). In *Anthoceros*, the cyanobacterial colonies occur in the so called "slime cavities" on the ventral side of the thallus. Slime pores make it possible for the cyanobacterium to reach the cavities (25, 104). The *Nostoc* colonies associated with *Blasia* inhabit specialized hemispherical auricles on the ventral side of the thallus (25, 116). Multicellular bryophytic filaments extensively penetrate the *Nostoc* colonies in both *Blasia* and *Anthoceros*. While, in *Anthoceros*, these are comprised of highly vacuolated thin walled cells, in *Blasia* the cells of the filaments are thick-walled and often develop labyrinthine wall ingrowths (25). These specialized types of bryophytic cells are thought to be responsible for the transfer of metabolites between the symbionts (116). In general, the gametophyte of Bryophytes produces

sporophytes and sexual spores. In *Blasia*, the gametophyte can also produce two types of asexual propagules. Large stellate gemmae produced on the dorsal epidermis and smaller ovoid gemmae produced in long-necked flask-shaped receptacles (7, 26, 101). It is of particular interest that the apically derived stellate gemmae reach a complex level of development before they detach and germinate. In fact, they are essentially miniature *Blasia* gametophytes, equipped with two lobes and two auricles, regularly containing symbiotic *Nostoc* colonies (26, 101). With the help of these symbiotic diaspores the whole symbiotic consortium consisting of the bryophyte and cyanobiont can disperse together, maintaining the symbiotic consortia over time and space. As the symbiosis is restricted to the gametophyte, each bryophyte generation needs to be infected by new cyanobacteria. Gametophytes tend to show much dieback during winter. In the spring there is a rapid recolonization via spores and, later via the two types of asexual gemmae (26).

Gymnosperms - Cycads

Cycads are an ancient group of seed plants that first appeared in the late Carboniferous and have thus existed for approximately 300 million years. Today they are found on every continent except Europe and Antarctica, but are restricted to small populations in the tropics and subtropics of both hemispheres. Cycads have an aerial stem, covered with persistent leaf bases, producing either pollen-bearing or ovule-bearing modified leaves called sporophylls (Fig. 3d). The ovules, like those of true gymnosperms, are naked (130).

Cycads produce three types of roots: (i) a tap-root that is equivalent to the primary root system found in most types of plants, (ii) lateral roots, and (iii) a highly specialized type of lateral root usually termed “coralloid roots”. It is in the coralloid roots that the symbiotic cyanobacteria are found (Fig. 3e). Instead of a downward growth pattern, these roots show a marked negative geotropism and grow laterally and upward toward the surface of the soil (4, 77). When infected, the cyanobacteria are found in a specific cortical layer inside the root, the cyanobacterial zone (Fig. 3f). The presence of cyanobacteria inside the root induces irreversible modifications of the growth and development of the root (4, 77). The growth pattern is quite different from normal roots of the same age. Growth in length is much reduced while growth in diameter increases noticeably in these roots (72). Furthermore, the cycad cells in the cyanobacterial zone undergo marked differentiation, elongating radially to interconnect the two adjacent cortical layers (59). It has been suggested that these elongated cells are specialized cells responsible for the transfer of metabolites between the partners (59).

The process of infection is still unclear. Invasion of filamentous cyanobacteria may occur at any stage of development of the root, but the precise time and location of the invasion is unpredictable. Several suggestions for how infection takes place have been made: (i) through injured parts of the root, (ii) through lenticels, (iii) through the papillose, and (iv) through breaks in the dermal layer (76). More detailed studies are needed to establish if there is a single process of infection in cycads or if different cycad genera/families have developed different strategies.

Recently the "Cycad Genomics Project" was started by The Plant Genomics Consortium (http://genomics.nybg.org/). This project in combination with the *N. punctiforme* genome may reveal the unique characteristics behind this association.

AIM OF THIS PROJECT

The questions addressed in this thesis concern two aspects of the cyanobacterial $tRNA^{Leu}$ (UAA) intron.

The **first aspect** is the use of the $tRNA^{Leu}$ (UAA) intron to study the distribution and specificity of cyanobacteria associated with different hosts (the bryophytes *Anthoceros fusiformes* Austin and *Blasia pusilla* L., and various cycad species). This can be analyzed at different levels:

- one symbiotic structure
- one individual host
- one host population
- different host species
- different symbiotic systems

Also, the study of the diversity can be done in different geographic regions and at different time points.

The **second aspect** concerns the study of the evolutionary patterns of the $tRNA^{Leu}$ (UAA) intron itself. The diversity studies have provided a large number of sequences from natural populations (see below) of closely related *Nostoc* strains. Using this unique data set, as well as other $tRNA^{Leu}$ (UAA) intron sequences from databases, the study of the evolutionary patterns in this genetic marker can be performed and thus interesting insights into its evolution can be obtained.

BIOLOGICAL MATERIAL

Field collected material was used in this thesis (Fig. 3b-d). The *Anthoceros* specimens were collected in 1998 from three different sites in Oregon, USA (on a sandy soil along a creek bank; 44°16'N, 124°06'W, elev. 15-20 m). The *Blasia* specimens were collected in 1996 (one site) and in 2000 (three sites including the original site from 1996) in Saarijärvi, central Finland (on mineral soil along road banks; 62°47'N, 25°03'E, elev. 140-170 m).

The study of the diversity of the cycad-*Nostoc* association was based on samples collected from cycads growing in both botanical gardens and natural populations. The use of plants from a botanical garden has the advantage that, different cycads naturally occurring in different geographical regions can be investigated at the same location. However, some disadvantages may also be considered. It is not known when the coralloid roots were infected. Some may have been infected in their natural habitat whereas others might have been infected in the botanical garden. In addition, perhaps not the same events take place in natural populations of cycads as in cycads growing in a botanical garden. To overcome these problems, a more extensive study was performed using a natural population of the cycad *Macrozamia riedlei* (Fischer ex Gaudichaud-Beaupré) C.A. Gardn. growing in Perth, Western Australia (Fig. 3d).

GENETIC MARKER

During the 1960's and 1970's, many studies addressed the question of the identity of the *Nostoc* cyanobiont(s) of plants. These studies involved the identification of cyanobacteria by their morphological characteristics and required pure cultures, omitting all strains that could not be cultured (5). Since then, the development and use of different molecular methods has created new possibilities for studying the diversity and specificity of symbiotic cyanobacteria. The methods used on symbiotic cyanobacteria vary from the comparison of specific sequences (e.g. 16S rDNA gene or $tRNA^{Leu}$ (UAA) intron) to RFLP and fingerprinting methods (38, 45, 60, 79, 89, 97-99, 102, 129, 139-141). The use of different methods may render different results when analyzing bacterial diversity making comparisons difficult. The genetic marker most extensively used for symbiotic systems is the $tRNA^{Leu}$ (UAA) intron (80, 88-91, 102, 119, 133). This marker fulfills the criteria required to answer questions about bacterial diversity: (i) it can be specifically amplified from complex biological samples, (ii) it has sufficient sequence variability to distinguish biological variation from variation caused by possible PCR errors, (iii) it is present in a single copy and it is not mobile;

and (iv) its size is most advantageous for easy sequencing. Furthermore, this intron has been used in the lichen association, where it revealed both interesting patterns of diversity and spatial distribution of the cyanobacterial symbionts (80, 88-91, 102, 119). For these reasons the $tRNA^{Leu}$ (UAA) intron was selected as the genetic marker to be used in this study.

RESULTS AND DISCUSSION

DIVERSITY AND SPECIFICITY (PAPERS I TO III)

Methodology

The study of the genetic diversity of cyanobacterial symbioses has been based on different methodologies, some based on banding patterns (RFLP, fingerprinting, and DGGE) and others on sequencing analyzes (for a review see (97). This fact has proven to be problematic because the different techniques render different results making comparisons difficult.

Concerning the bryophyte-*Nostoc* association, the studies revealed the presence of a single cyanobacterial strain within each bryophyte cavity (paper III). Furthermore, in the *Blasia* sites of central Finland, some *Nostoc* strains were only found from single sites, while others were found from all three localities (paper III). The finding that some symbiotic *Nostoc* strains were shared by bryophytes growing 2000 meters apart differed from the results of West and Adams (129), who detected a more localized distribution of *Nostoc* strains in *Phaeoceros laevis*. However, one cannot directly compare the results of these studies, since different molecular methods with different levels of resolution were used.

On the cycad-*Nostoc* association, it was demonstrated that a single *Nostoc* strain is present in an individual coralloid root (papers I and II). This is in contrast with some earlier studies. The use of different methodologies over the years has resulted in conflicting results, some more easily explained than others. The use of southern hybridizations using probes against conserved genes and genomic cyanobacterial DNA from natural populations of different cycad species revealed the presence of several cyanobionts in a single cycad species (60). However, to get sufficient amounts of genomic DNA, samples from many coralloid roots were pooled for each analyzed sample. Thus, no conclusion can be drawn concerning whether a single plant or coralloid root, hosts multiple cyanobionts. Using a method based on banding patterns generated using PCR (fingerprinting) and a single primer (based on a repeated DNA element - HIP1) Zheng et al concluded that several different cyanobacterial strains are present inside individual coralloid roots (139). However, the use of fingerprinting methods has limitations. To be able to compare banding patterns, a high level of standardization is required (29). Artifactual variation represents a potential problem in surveys of genetic variation in natural populations and must be distinguished from true polymorphism. The cyanobacterial filaments inside the coralloid roots are embedded in a chemically complex mucilage that needs to be removed in

order to consistently obtain PCR products (paper I). In addition, it is known that the cyanobacterial cells are in different physiological states along the coralloid root (22). In an attempt to evaluate these results, we used a single cyanobacterial strain, *Nostoc* sp. PCC 73102, and the same primer (STRR*mod*) as Zheng et al (139). By varying the quality of the *Nostoc* template used (purified DNA, whole cells, cells grown under different conditions or cells in different physical states) different banding patterns were obtained, even though the same cyanobacterial strain was always used (paper II). This demonstrates that fingerprinting methods require great caution to avoid artefactual results arising from chemical or physiological differences in the samples. Furthermore, the primer must be tested on host DNA because when the cyanobacterium is collected it is unavoidable to also get host cells. So, at least the host's DNA will be present in the sample and may be a template generating DNA fragments. Indeed this was shown to be the case when using the STRR*mod* primer and genomic cycad DNA (paper II). Thus, since fingerprinting is based only on comparing obtained DNA banding patterns, one can not e.g. distinguish the source of template DNA. In contrast, sequence information (e.g. $tRNA^{Leu}$ (UAA) intron) can confirm cyanobacterial origin of variation, something that is of great importance when symbiotic systems and field samples are used.

Patterns of diversity

Many different *Nostoc* strains are involved in the associations with hornworts, liverworts and cycads, although no variation has ever been observed within a single bryophyte cavity or cycad coralloid root () (Fig. 4). The fact that a single strain was consistently found in each cavity/coralloid root indicates that each endophytic *Nostoc* within a bryophyte thallus or coralloid root results from a single infection event. In other words, new *Nostoc* strains do not seem to enter individual cavities or coralloid roots after these have already acquired a specific strain. The symbiotic cyanobacteria can potentially come from several different sources, including a population of free-living *Nostoc* strains in the soil, from neighboring hosts, or from other symbiotic associations, such as cyanolichens. The relative importance of these different sources may vary and should be further investigated.

The level of *Nostoc* diversity within both bryophyte species was not consistent; some gametophytes contained only one *Nostoc* strain, while others housed several different strains (Fig. 4). In the cases where only one *Nostoc* strain was present within a gametophyte, all the symbionts may well have originated from the same source. Possibly, there had been only one symbiotically competent *Nostoc* strain in the immediate surroundings of the growing gametophyte, or at least one strain had dominated within the

microhabitat. It is also possible that during the growth of the bryophyte, mature cyanobacterial colonies had released hormogonia, which then had infected new symbiotic cavities at the thallus apex. This phenomenon could have promoted low cyanobacterial diversity within individual bryophyte thalli, provided that these had initially been infected by single *Nostoc* strains. Conversely, in cases where several *Nostoc* strains were detected in single thalli, a mixture of symbiotically competent cyanobacteria must have existed during thallus growth. The assemblage of competent *Nostoc* strains must also have coexisted for some time, as the distribution of specific strains within individual *Blasia* thalli did not correlate with the developmental sequence of the symbiotic structures. In these cases, random infection by hormogonia from old cyanobacterial colonies may have acted to maintain cyanobacterial diversity within growing thalli, provided that these had initially been infected by more than one *Nostoc* strain.

In the cycad association, the level of *Nostoc* diversity within coralloid roots was consistent, with all individual coralloid roots containing a single cyanobiont (Fig. 4). That a minor cyanobiont could have avoided detection is still possible but unlikely. However, different coralloid roots from a single *Encephalartos villosus* Lem. specimen harbored different cyanobacteria (paper I). Additionally, cyanobionts in coralloid roots of different cycad species possessed the same intron sequence indicating that the same cyanobiont is present in different cycad species. Furthermore, when analyzing the natural populations of *M. riedlei*, plants that were growing close together (corresponding to females and their descendants) shared the same *Nostoc* strain (paper II). It is tempting to state that the selectivity of the plant towards the *Nostoc* strain is genetically transmitted. However this observation could merely correspond to the availability of *Nostoc* strains in sandy soil.

Temporal stability

One *Blasia* site in central Finland was first sampled in August 1996 and again in August 2000. One of the two *Nostoc* strains found in 1996 was also detected in 2000 revealing a level of temporal stability in the cyanobiont composition of the bryophyte population. The sampling dates were separated by three full gametophyte generations. The fact that mature *Blasia* thalli did not overwinter at the site was confirmed in May 2000, when only dying shoot apices with immature sporophytes, and minute thallus primordia, apparently developing from gemmae, were seen. Semidecayed gametophytes were taken to the laboratory for close examination, but no endophytic *Nostoc* colonies were detected. Later during the summer, there was a rapid recolonization and abundant production of new gametophytes at the site.

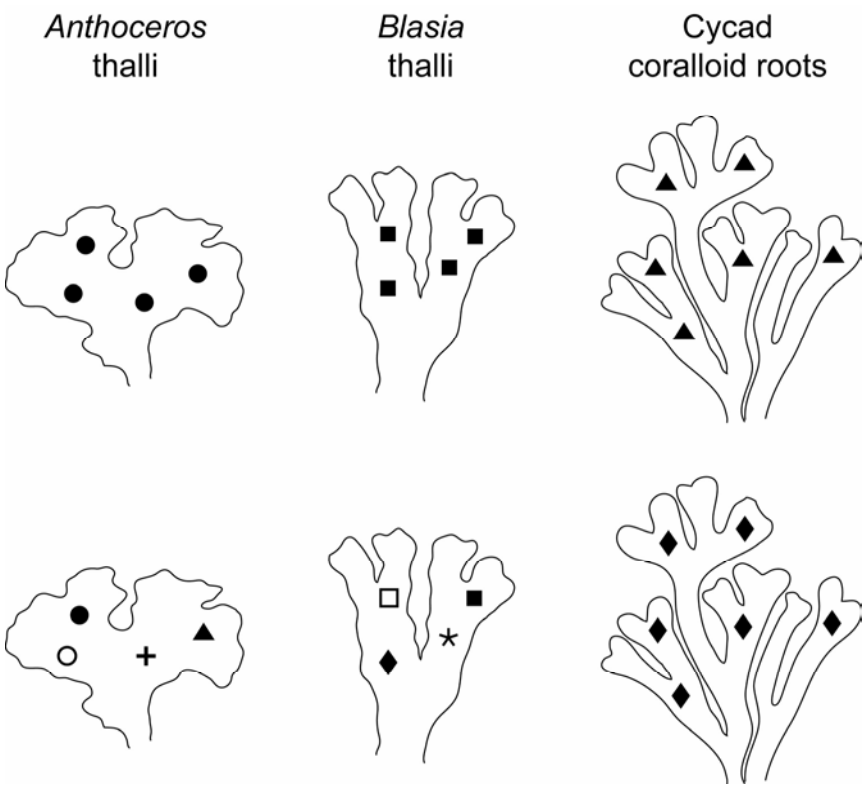


Fig. 4. Simplified representation of results obtained using field material showing the diversity of cyanobacterial symbionts based on the $tRNA^{Leu}$ (UAA) intron sequences (papers I, II and III). Individual cyanobacterial samples are indicated with a symbol and each symbol represents a specific nucleotide sequence. The level of cyanobacterial diversity varies:

- Each sample contained a single sequence implying a single *Nostoc* strain per cavity/coralloid root.
- Different bryophyte thalli possessed either the same *Nostoc* or different *Nostoc* strains.
- Different coralloid roots from the same cycad individual can contain different *Nostoc* strains.
- Identical sequences are found in both bryophytes and cycads implicating that one *Nostoc* can be present in both types of associations.

Although further studies are needed, the results indicate that there is a certain level of spatial and temporal continuity in the *Blasia*-*Nostoc* symbiosis. This finding is understandable in the light of what is known about the seasonality and reproductive biology of *B. pusilla*. Most symbiotic bryophytes are pioneer species, exploiting temporary niches (113). The ability of *Blasia* to effectively colonize temporary habitats hinges partly from the fact that the gametophytes can produce, not only sporophytes and

sexual spores, but also the two types of asexual propagules described earlier. With the help of these symbiotic diaspores, the whole symbiotic consortium, consisting of the bryophyte and up to two different *Nostoc* strains, can disperse together. Thereby, unlike the other bryophyte symbioses, the *Blasia*-*Nostoc* association has the potential for maintaining specific symbiotic consortia over time and space. The potential significance of this special mode of propagation has not been recognized in the literature on cyanobacterial symbioses.

When comparing the sequences from the specimens of *Anthoceros*, *Blasia* and the different cycad species which were collected from different parts of the world, identical sequences could be found (Fig. 4). This points to the fact that some *Nostoc* strains might be more cosmopolitan than others. This may be quite significant as the thalloid liverworts, hornworts and cycads together represent a fair proportion of all extant plant lineages that can be traced back to pre-Permian times (49, 122).

$tRNA^{LEU}$ (UAA) INTRON (PAPER IV)

The studies on the distribution and specificity of symbiotic cyanobacteria provided a unique data set of $tRNA^{Leu}$ (UAA) introns sequences from natural populations of *Nostoc* strains. Using these sequences together with other available sequences from the databases, the evolutionary patterns of this genetic element was studied.

The sequence alignment of the cyanobacterial $tRNA^{Leu}$ (UAA) introns reveals great sequence similarity. Sequence variation is mostly confined to certain regions that, when the alignment is compared to the secondary structure predictions, are localized to some of the loops or hairpin structures. Fig. 5 shows the two-dimensional structure of the consensus sequences of the $tRNA^{Leu}$ (UAA) intron from 88 *Nostoc* strains. All size variation within the genus *Nostoc* can be found in the structural element P6b. The size of this hairpin ranges from 50 to 127 nucleotides and the nature of its size variation will be further discussed below. When comparing intron sequences from cyanobacteria belonging to the different subsections, several others regions with size variation, especially element P2, were found (paper IV).

In the cases where variation in single nucleotides occurs, these often retain the secondary structure of the intron (Fig. 5). This can be seen both in cases where a change in one position is accompanied by a change on the base pairing strand so that different sequences either have a G:C or A:U base pair in this position, as well as where base pairing of G·U type allows changes to occur on one strand without disturbing the base pairing structure.

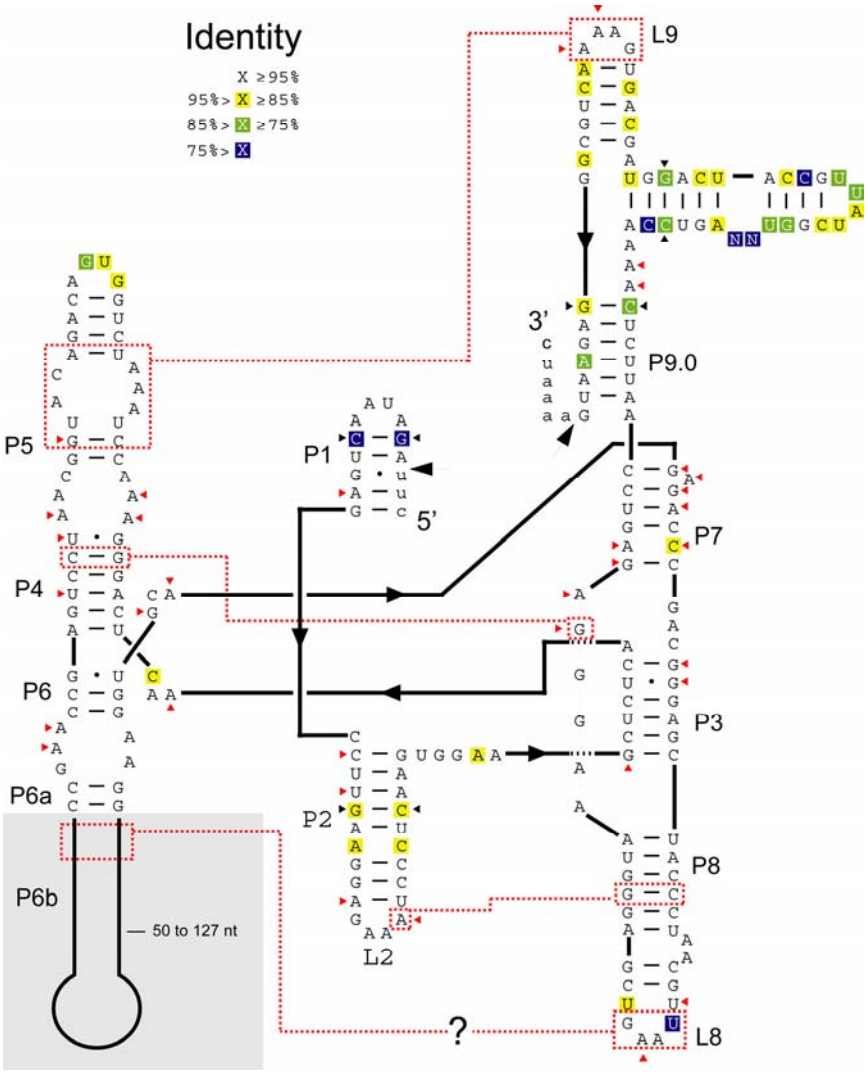


Fig. 5. Variability of different positions in *Nostoc* $tRNA^{Leu}$ (UAA) introns visualized in the secondary structure prediction of a consensus sequence based on 88 intron sequences (papers II and IV). Intron sequence is shown in upper case letters and sequence of the tRNA in lower case letters. Arrows mark the intron/exon boundaries. For each specific nucleotide, different colors are used to reflect how large a fraction of the different sequences supports the consensus sequence. Black arrowheads indicate positions where variation between differing sequences still retain the base pairing structure seen in the figure. Red arrowheads indicate positions known to interfere with folding and/or splicing of the *Anabaena* sp. PCC 7120 $tRNA^{Leu}$ (UAA) intron (63). Tertiary interactions are indicated by red dotted lines. The structural element P6b is shown as a schematic loop due to its variation in size.

In Fig. 5 these degenerate bases, where different sequences among those used to construct the consensus sequence show either different base pairs (that is either A:U or G:C) or differences between sequences involving G·U base pairing in combination with either A:U or G:C base pairing are labeled. The latter type of base pairing thus provides a bridge from a G:C to an A:U base pairing, or vice versa, without loss of base pairing in the intermediate structure. Using phosphorothioate substitution interference with RNA activity and structure, several positions were identified that interfere with folding and splicing of the $tRNA^{Leu}$ (UAA) intron in *Anabaena* sp. PCC 7120 (63) (Fig. 5). These positions (red arrowheads) were mostly located within the catalytic core region (P3, P4, P5, P6 and P7), but some were also found in peripheral parts of the intron (L2, L8 and L9). These peripheral parts are proposed to be involved in tertiary interactions stabilizing group I introns (40) and correspond to contacts between the L2 tetraloop and the minor groove of P8, the L9 tetraloop and the tetraloop receptor of the P5 domain and possibly the L8 loop and the lower part of P6b (63) (Fig. 5). When comparing all these positions with the consensus intron sequence from the *Nostoc* strains, it is clear that all positions, with the exception of a position in structural element P7, are highly conserved. Furthermore, the positions present in the proposed regions of tertiary interactions are also conserved. The only exception might be the tertiary relation between the L8 loop and the lower part of P6b due to the high variability in size of this element in the *Nostoc* strains (Fig. 5). The comparison and superimposition of results indicates that the positions where sequence variability is found, there might be a requirement for base pairing to retain the structure needed for autocatalysis of the intron, which would give a selection pressure for compensatory mutations. Conversely, the positions that were demonstrated to interfere with folding and splicing of the intron are under a stronger selection pressure, which is seen by the fact that these positions are highly conserved among the different *Nostoc* strains. The results from these two different types of studies, the phosphorothioate substitution assays (63) and the analyses of sequences from natural populations of *Nostoc* strains, thus work as a complement to each other.

INTRON VARIABLE REGION P6B (PAPERS IV AND V)

When the *Nostoc* $tRNA^{Leu}$ (UAA) intron sequences are compared, most differences are found in the structural element P6b (paper IV). This element was found to consist of degenerate heptanucleotide repeats (Fig. 6a) that folds into a hairpin structure (Fig. 6b) allowing the repeats to base pair. Figure 6a shows the different heptanucleotide repeats in this element for a selected group of *Nostoc* $tRNA^{Leu}$ (UAA) intron sequences. Size variation in

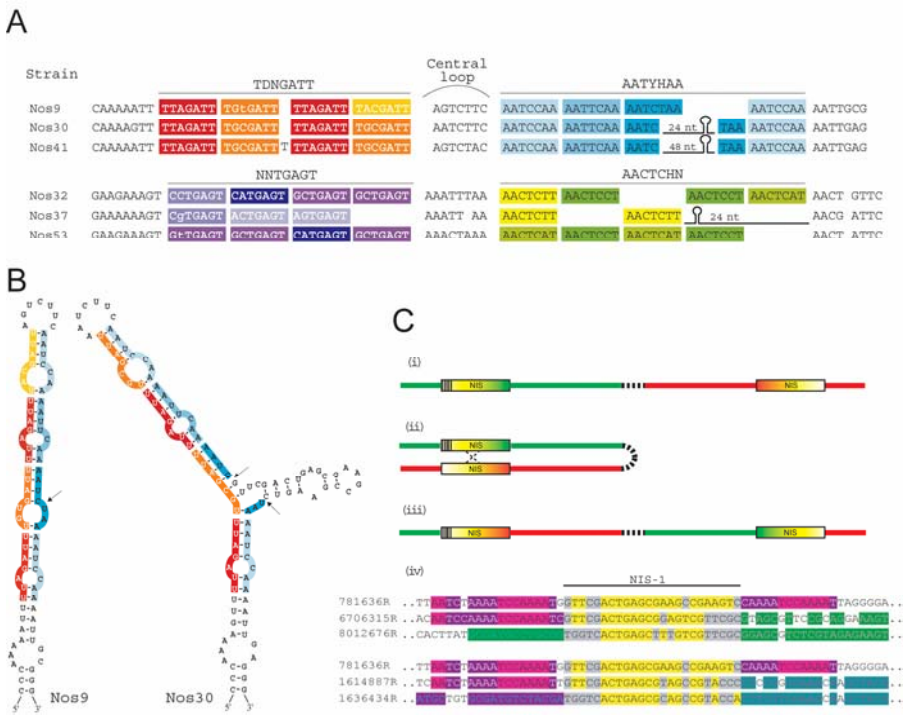


Fig. 6. Patterns of distribution of heptanucleotide repeats and NIS elements in *Nostoc* (paper IV).

(A) Alignment of the P6b region of the $tRNA^{Leu}$ (UAA) intron from a selected group of *Nostoc* strains. The region is built from degenerate heptanucleotide repeats, which are represented as colored blocks. Sequences from the two classes of repeated sequences, TDNGATT/AATYHAA and NNTGAGT/AACTCHN, are represented. NIS elements are represented as schematic hairpins with their sizes indicated.

(B) Representation of the hairpin formed by the intron structural element P6b from *Nostoc* symbionts of the lichen *Nephroma articum* collected in Sweden (Nos9) and *N. resupinatum* collected in Finland (Nos30). The different repeats are represented as in (A). Arrows indicate the positions where the extra hairpin formed by the NIS element is inserted.

(C) Schematic representation of the putative dispersal mechanism of NIS elements and two examples retrieved from the *N. punctiforme* genome. (i) Two unspecified regions of the *Nostoc* genome (green and red) each with a NIS element. (ii) Due to sequence similarity, the two NIS elements are brought together and recombination can occur between the NIS elements (dashed cross). (iii) The result from the recombination event is two hybrid NIS elements, containing parts of both original NIS, and an inversion of the region between the NIS elements. Notice that the regions flanking the recombinant NIS are the same as for the original NIS elements. (iv) Two groups of three sequences retrieved from the *N. punctiforme* genome illustrating the mechanism described above. Each sequence is identified and colored as in table 2. The middle sequences (6706315R and 161488R) may correspond to recombination events between the upper (781636R) and lower (8012676R and 1636434R) sequences, respectively.

this region is caused by different numbers of repeats and, in some cases, by insertion of other genetic elements not having the heptanucleotide repeat motif. These elements are represented as schematic hairpins in Fig. 6 and will be referred to in the text as NIS (*Nostoc* Iterated Sequences). Figure 6b shows the P6b region of the intron from two symbiotic *Nostoc* strains. These two sequences are very similar with the exception of a NIS element that forms an extra hairpin in this region.

Based on the observation that different groups of degenerate heptanucleotide repeats could be found, the sequences were divided into two classes: one class with the repeated consensus sequence TDNGATT and its pairing repeat AATYHAA and, another class with the repeated unit NNTGAGT and its base pairing unit AACTCHN (paper IV). Between the two sets of base pairing repeats, a central loop with different sequences in the two classes is found. Among the different heptanucleotide repeats, several groups could be identified based on the degree of sequence similarity (Fig. 6a). Increasing numbers of short sequence DNA repeats are being identified in prokaryotes as more genomes become available and algorithms that allow a simple way of detecting them are developed. Also in cyanobacteria, several short sequence repeats have been found, but the significance of their existence is still unknown (6, 58, 65). A possible mechanism involved in their formation is slipped strand mispairing occurring in combination with inadequate DNA mismatch repair pathways (117). Also, the peculiar secondary structure of repetitive DNA allows mismatching of neighboring repeats and, depending on the strand orientation, repeats can be inserted or deleted during DNA duplication mediated by DNA polymerase (123). These mechanisms can be responsible for the variation observed in the number of repeats in the P6b structural element of the *Nostoc* $tRNA^{Leu}$ (UAA) intron.

In some instances, the repeats in the structural element P6b are interrupted by NIS elements. The NIS elements can be found interrupting a single repeat or between two consecutive repeats (Fig. 6a-b). They are of different lengths (24, 42, 45 and 48 nucleotides) and their sporadic occurrence indicates recent origins and raises the question of how such interpolations might arise. In order to answer this question, the recently completed genomic sequence of *Nostoc punctiforme* was searched for sequences similar to the three 24-bp NIS elements found in the $tRNA^{Leu}$ (UAA) introns of *Nostoc* strains from natural populations. The $tRNA^{Leu}$ (UAA) intron of *N. punctiforme* does not have any interpolation within its heptameric repeats, but a Blast search of the *N. punctiforme* genome found multiple copies of sequences closely related to each of the three 24-bp segments. Most were situated in non-coding regions, but some were found interrupting ORFs. The search was repeated using a function within BioLingua (http://nostoc.stanford.edu/Docs/) that returns sequences

according to the number of mismatches with the query. The three sets of related sequences were extended by searching for sequences with four or fewer mismatches relative to an element of the set. The threshold of four mismatches was chosen because one would expect to find a match of this stringency about 0.1% of the times when searching through random sequences the same size and base composition as the genome of *N. punctiforme*. The sequences were grouped according to the original 24-bp segment that led to their discoveries and have been named NIS-1, NIS-2, and NIS-3 (Table 2).

Sequences termed NIS-1 generally have pairs of 4-nt inverted repeats at characteristic positions. The sequence of the repeats varies, but usually in such a way to preserve the potential for base-pairing. Sequences within NIS-2 fall into two closely related subgroups, NIS-2a and NIS-2b, sharing the same 18-nt palindromic core with the exception of an insertion/deletion. Sequences classified as NIS-3 are the least numerous. NIS-2a is about as common in *Anabaena* sp. PCC 7120 as in *N. punctiforme*. However, NIS-1 and NIS-2b are not present in the former cyanobacterium, and NIS-3 is rare (Table 2). The sequences are predominantly on the chromosome with only 2.4% residing on one of the five plasmids that constitute 9.1% of the total DNA. The 48-bp segment noted originally in the cyanobiont (Nos-41) of the lichen *Peltigera venosa* is clearly related to one of the 24-bp long NIS-1, differing by the interpolation of an additional 24 nucleotides in the central region of the segment, which itself is an element of NIS-1 (paper V). Several other instances of such compound segments were found in the *N. punctiforme* genome all formed from nested NIS-1 sequences. No compound segments were found composed of other NIS sequences (paper V).

The flanking sequences surrounding the most common elements of NIS-1, NIS-2a, and NIS 3 are generally heptameric repeats (NIS-2b constitutes a special case as its is instead flanked by a relatively conserved sequence resulting in a unit typically of 101 nucleotides but as many as 157 nucleotides). This suggests that the intervening sequence were dispersed in the genome through a mechanism that depended on flanking sequences. Such a mechanism is described on Fig. 6c. This mechanism has similarities with conservative site-specific recombination (23, 51), but with a longer region of homology (24 bp compared to the 6-8 bp for site-specific recombination). A consequence of the action of this mechanism is the rearrangement of the DNA region between the two recombining NIS elements. It is difficult to obtain direct evidence for the existence of such a mechanism. However, one fact points in this direction. The gene sequences of *Anabaena* sp. PCC 7120 are very similar to the homologous genes in *N. punctiforme*. Indeed, the genetic information of the two organisms is similar, the major difference between the two is on how the genes are organized and

positioned in the two genomes. If we consider this, together with the fact that NIS-1 elements are only present in *N. punctiforme*, it is tempting to infer that NIS-1 could indeed be involved in the genome reorganization *N. punctiforme*.

Table 2. Common instances of NIS-1, NIS-2, and NIS-3

NIS group[a]	Sequence[b]	Occurrences[c] A	B	C
Group NIS-1				
Nostoc punctiforme 1418927D	TGGGGCAT GGG ACTTGTACTGAGCGAAGTCGAAGTATTGGGCAT GAAAA	10	10	3
4605950D	GAGCAGGGAGCAGGCTTGCCGTGAGCGTAGCCGAACGGGAGCAGGGAGCAGG	10	10	10
2 80235R	TGGGGAATGGGGAACTTGTACTGAGCGCAGTCGAAGTATGGGGAATGGGGTA	8	7	6
5891985D	TGGGGAATGGGGAACTTGTACTGAGCGCAGTCGTAAAGCCTGCGGCATAGCT	7	7	5
7692787R	ATGGGCAT GGGCACTTGTACTGAGCGGAGTCGAAGTATGGGGCATGAGGCA	5	5	2
7099863R	CAAAATCCAAAATCGTTCGACTGAGCGAAGCCGAAGTCCAAAAGCCAAAATC	3	3	3
Nostoc Nos30/Nos54 tRNA^{Leu} intron	CAAAATYCAAAAYCGTTCGACTGAGCGAAGCCGAAGTCYAAAAYCAAAAA			
Group NIS-2a				
Nostoc punctiforme 3582136D	AACTTCTAACTCCTGTACAGACGCGATTAATCGCGTCTCTCCTAACTCCTAA	22	22	19
Nostoc Nos37/Nos38 tRNA^{Leu} intron	AACTCTTAACTCTTGTACAGACGTGATTAATCGCGTCTAACGATTCGGAAGG			
Anabaena sp. PCC 7120 4514797R	GACAACTGACAACTGTACAGACGCGATTAATCGCGTCTCTGACAACTGACAA	22	22	21
Group NIS-2b				
Nostoc punctiforme 1305306D	GACCAAAACACTTGTAGAGACGGCGATTTATCGCGTCTCAAAACCCACGATT	50	49	48
Group NIS-3				
Nostoc punctiforme 7450815D	GTGCTGAGTGCTGTTAGCGGTAGCGGGGCGTTTAGCCCGTAAGGAGTGTTGA	9	8	9
Nostoc Nos51 tRNA^{Leu} intron	GTCATGAGTGCTGTTAGCGGTAGCAGGGCGTTTAGCCCGTGCTGAGTCAATT			
Anabaena sp. PCC 7120 3457249D	TTTGAAGGCGCTGTTAGCGATAGCGGGGCGTTCAGCCCGTGCTGAGTACTGA	2	2	2

[a]The different groups of NIS sequences found in the genome of *N. punctiforme*, from the tRNA^{Leu} (UAA) intron of different *Nostoc* strains, and the genome of *Anabaena* sp. PCC 7120 (when present). The low-order chromosomal coordinate of a typical sequence is given, followed by the orientation of the sequence (D, direct; R, reverse).

[b]Selected examples of NIS from the different groups (in yellow, with gray highlighting indicating deviance from a reference sequence). Flanking sequences are highlighted to accentuate repeat units. The most represented NIS sequence variants are shown: all with greater than three occurrences and others of special interest. Many other less represented instances exist in the genome. Horizontal lines indicate regions potentially involved in self-base pairing.

[c]Total number of occurrences of the indicated NIS sequence (A) in the genome of *N. punctiforme* or *Anabaena* sp. PCC 7120, (B) in the chromosome, and (C) flanked at least on one side by the given sequences.

INTRON EVOLUTION

The question of the origin of introns has been controversial (84, 114). One view is that introns arose late in the evolutionary history and invaded originally uninterrupted genes. Another view is that introns are ancient components of genomes, which have been lost in many cases. The apparent lack of introns in the eubacterial lineage was used as an argument for the intron late hypothesis. When, in 1990, the first eubacterial intron was discovered in the $tRNA^{Leu}$ (UAA) gene of cyanobacteria (50, 134), the intron early hypothesis was given an important argument. Since then, several eubacterial introns (of both group I and group II) have been found in different eubacterial lineages (85, 100). The question of the origin of these rare eubacterial introns has remained controversial. The $tRNA^{Leu}$ (UAA) intron has been important in this debate. In contrast to other eubacterial introns interrupting tRNA genes, the $tRNA^{Leu}$ (UAA) intron has features expected for an evolutionarily old genetic element. The $tRNA^{Leu}$ (UAA) introns from cyanobacteria and chloroplasts are more closely related to each other than to any other introns. Their position between the second and third base of the anticodon is conserved, which also implies evolutionary age. The lack of intron in some cyanobacterial lineages and algal chloroplasts would then be explained by a later loss of the intron (12, 85). However, data has been presented that suggests a polyphyletic origin of $tRNA^{Leu}$ (UAA) introns within the cyanobacterial radiation (105-107). The evidence presented for a complex origin of cyanobacterial $tRNA^{Leu}$ (UAA) introns includes the presence of two different types of $tRNA^{Leu}$ introns in some strains and the lack of intron in other strains. The interpretation that this implies a complex pattern of intron evolution is however not unproblematic. The two different types of $tRNA^{Leu}$ introns were found to be situated in different types of tRNA genes. The new intron type was thus inserted into a new type of tRNA gene; and this would thereby not be a case of lateral transfer of introns, but rather a case of a different tRNA gene containing a different intron. When the "traditional" $tRNA^{Leu}$ (UAA) gene contained introns, they were always of the traditional type. It is also worth noting that all those sequences were obtained as PCR products from cultures from the NIVA collection (106, 107), which consists of unialgal, but not axenic, cultures. The cyanobacterial origin of this second type of tRNA gene and intron thus remains to be shown by cloning of a greater genomic fragment containing both the new type of intron and clearly cyanobacterial genes. This is especially important since $tRNA^{Leu}$ introns have been detected in other eubacteria such as *Pseudomonas* (GenBank accession number AY029760), which is a potential contaminant of cyanobacterial cultures.

The intron studied in this thesis interrupts the tRNA gene with the anticodon UAA for leucine. In the *N. punctiforme* genome six different

codons can be used for leucine. When analyzing the frequencies of these six codons in protein encoding genes, it is striking that the codon UUA is by far the most used. In other words, during translation, the most common tRNA used for leucine is the $tRNA^{Leu}$ (UAA) which contains an intron interrupting its gene. It is striking that this intron has been maintained and favored through evolution interrupting the gene of a tRNA that is extensively used. The reason for this is puzzling, but some evolutionary advantage must exist. This could be a case of an old "genome symbioses" in which the intron is maintained during evolution in exchange for being a place where different genetic activities may occur without interfering with vital genes.

CONCLUSIONS AND FUTURE DIRECTIONS

The usefulness of the $tRNA^{Leu}$ (UAA) intron as a genetic marker to study the diversity and specificity of two *Nostoc* associations was confirmed. These studies gave several new insights into the complex associations between the plant partner (thalloid bryophytes or cycads) and symbiotic cyanobacteria.

- Only one *Nostoc* strain is present in each symbiotic cavity or coralloid root.
- The level of diversity of *Nostoc* strains may vary when analyzing different bryophyte thalli; some only contain one *Nostoc* strain while others contain several strains.
- The level of diversity of *Nostoc* strains may vary in cycads. Individual coralloid roots only contain one *Nostoc* strain, but different coralloid roots from a given plant may contain different *Nostoc* strains.
- There is a level of temporal stability in the bryophyte association.
- Some *Nostoc* strains seem more cosmopolitan than others.

These results form a platform from which our knowledge about how these symbiotic systems work can be expanded. Possible next steps include the following studies: (i) how the host influences the cyanobacterium, (ii) how the metabolism is changed in the symbiotic cyanobacterium compared to its free-living state, and (iii) how the two different symbiotic partners communicate. In a preliminary attempt I have tried to address some of these questions. Using two-dimensional gel electrophoresis, I have compared the protein profiles of freshly collected symbiotic *Nostoc* and free-living cultures. Protein profiles with clear differences were obtained. Two proteins which showed different expression patterns under symbiotic and non-symbiotic conditions were further analyzed by sequencing. The two sequences showed no similarities to known proteins. Recently, the genes encoding these proteins were included in a microarray experiment where RNA collected from cells grown under different conditions was compared. The results are still not available but might reveal possible functions of these proteins. Although this data is not included in this thesis, it is a possible approach for increasing our understanding of these cyanobacterial-plant associations.

In addition to the understanding of symbiotic systems, the unique data set which was generated, allowed interesting insights into the patterns of evolution of the cyanobacterial $tRNA^{Leu}$ (UAA) intron.

- The intron sequence is highly conserved in the genus *Nostoc*.

- Nucleotides that were shown to be important for the correct folding and splicing of the *Anabaena* sp. PCC 7120 $tRNA^{Leu}$ (UAA) intron are conserved in the *Nostoc* strains studied.
- In *Nostoc* strains, intron size variation only occurs on the structural element P6b.
- The element P6b in *Nostoc* is built up of heptanucleotide repeats and this is the cause for size variation in this element.
- Some *Nostoc* P6b sequences, in addition to the heptanucleotide repeats, contain other sequences – NIS elements.
- NIS elements are abundant in the *N. punctiforme* genome.

The study of this intron is interesting because of the intron's old origin, and its value in revealing how evolution occurs in such genomic elements. An aspect I would give priority in future studies is the influence that the size variation of the element P6b has on the folding and splicing of the intron. The crystal structure of introns from *Nostoc* strains containing different patterns in this element (e.g. with and without NIS elements) would reveal not only if the putative tertiary interaction between the L8 loop and P6b indeed exists, but also show the influence that P6b has on the correct folding and splicing of the intron. Another topic of interest that arose from these studies was the discovery of the NIS elements and their possible relation to other repetitive sequences. Although the function of repetitive sequences is in general unknown, I believe that such elements are sequences that are being shuffled around in the genome until they are recruited to perform a determinate function. Some functions for these elements could involve protein binding or transcription termination. Otherwise, if these sequence elements have no functions, "…then it seems like a terrible waste of space."*

The work described in this thesis is a result of my learning process during the past years. When I started my Ph.D. education I thought, maybe naively, that I would start a project and after four years it would be finished. I was wrong. In science you are never finished, you always end up with new and even more interesting questions to answer. Indeed, this thesis is precisely that. Although several questions were answered, the number of new questions is even greater. This is why science is such a fascination for me.

*Quote by Carl Sagan concerning the possibility of life elsewhere in the Universe.

ACKNOWLEDGMENTS

I would like to thank:
My supervisor Peter Lindblad for sharing his knowledge and experience, and also for teaching me that problems do not exist only challenges!! Thank you also for the help with the Swedish summary of this thesis.

Per Paulsrud, my other supervisor, for the science, the important discussions, the help, the support, the enthusiasm and ideas, but above all for being such a friend!

All my co-authors and co-workers.

Peter, Per, Paulo, Rikard and Lígia for reading and commenting on this thesis.

The two Cyanogroups – the old one with Per, Rikard, Röbbe, Kjell, Pia, Olga, Mehtap and Caisa; and the new with Paulo, Fernando, Ellenor, Karin, Marie, Taras and Åsa. You all made life easier around the lab. A special thanks to Rikard and Röbbe for your friendship. We still have "The Cyanobacteriologist" to do! Also a special thanks to the Portuguese embassy at the department for bringing Portugal to the lab!

Professor Peter Engström for accepting me as a PhD student at the department.

All "plant people" past and present…….

Birgitta Norman-Ebendal, Anita Wallin and Björn Andersson for all practical help at the department.

To all of you "ett stort tack" for making Sweden my home for 4 years.

Por último um grande obrigado à minha familia e aos meus amigos.
Em especial aos meus pais pelo apoio nesta fase da minha vida e, desculpem não ter estado com vocês em alturas que aí devia ter estado. Ao meu irmão um obrigado especial pelas dicas informáticas (ex: como subir para o telhado do pipeline). Aos pais da Lígia pelas palavras de incentivo dadas e por terem aturado a Lígia nos momentos menos bons. A ti Lígia um muito obrigado por todas as opiniões, discussões e ideias, mas acima de tudo por todo o amor, sem ele isto não teria sido possível.

Financial support was provided by the Elliassons fund and The Royal Swedish Academy of Sciences.

SUMMARY IN SWEDISH

En organisms arvsanlag, DNA, består av många olika genetiska element. Ett sådant genetiskt element är introner, delar av icke-kodande DNA som ofta ligger inom den kodande sekvensen av olika gener. Introner kan defineras som en del av ett primärt transkript, eller den del av arvsanlaget som kodar för det genetiska element som avlägsnas när RNA processas och som inte är en del av det funktionella mRNA, rRNA eller tRNA. Beroende på vad som behövs för intronets avlägsnande, och inblandade mekanismer, klassificeras introner i fyra olika huvudgrupper. Grupp I introner kan avlägsna sig själva från den del av genomet som de sitter i och de var det första kända exemplet på en RNA molekyl med katalytisk förmåga, dvs det första RNA enzymet eller ribozymet. Grupp I introner finns hos både bakterier och eukaryoter. Det först beskrivna bakteriella grupp I intronet finns inom $tRNA^{Leu}$ (UAA), ett intron som är lokaliserat till en konserverad position i antikodon-tripletten, mellan den andra och den tredje basen (U-intron-AA). Det mest kända och bäst studerade $tRNA^{Leu}$ (UAA) intronet hos cyanobakterier finns hos den filamentösa stammen *Anabaena* sp. PCC 7120. Cyanobakterier är en evolutionärt äldre grupp fotoautotrofa mikroorganismer som idag uppvisar en stor morfologisk variation, från encelliga till filamentösa former. Den fysiologiska flexibiliteten har gjort dem mycket framgångsrika i naturen. Förutom terrestiska, limnisiska och marina stammar så kan cyanobakterier även bilda symbioser med olika eukaryota organismer. De bäst kända cyanobakteriesymbioserna är med växter, filamentösa stammar av *Nostoc* som lever tillsammans med mossor (t.ex. *Anthoceros* och *Blasia*), ormbunkar (*Azolla*), gymnospermer (kottepalmer) och angiospermer (*Gunnera*). Hos alla dessa symbioser är det förmågan hos *Nostoc* att omvandla och leverera bundet kväve från atmosfären som är av intresse hos den fotosyntetiska värden.

Den här avhandlingen handlar om mångfalden och specificiteten hos *Nostoc* i symbioser med mossor och med kottepalmer. Målsättningen kan delas upp i två; (i) $tRNA^{Leu}$ (UAA) som en genetisk markör för att studera utbredning och specificitet hos cyanobakterier i naturliga symbioser med olika värdar, samt (ii) evolutionära mönster inom $tRNA^{Leu}$ (UAA) intronet baserat på intronsekvenser från olika naturliga symbiotiska *Nostoc* stammar.

Genom att använda $tRNA^{Leu}$ (UAA) intronet som en genetisk markör har jag visat att många olika *Nostoc* stammar kan etablera symbioser med både mossor, exemplifierat genom *Anthoceros* och *Blasia*, och med kottepalmer. Däremot visade studierna att det inte finns någon variation

inom en individuel symbiotisk koloni, varken i en mossa eller en kottepalms koralloida rot. Nivån på diversiteten hos *Nostoc* varierar inom de både mossorna; några gametofyter innhåller enbart en *Nostoc* stam medan andra kan innhålla flera olika stammar. Inom symbiosen med kottepalmer innehåller varje enskild koralloid rot en enda cyanobakterie. Däremot kan olika koralloida rötter från samma kottepalmsindivid innehålla olika cyanobakterier. Dessutom finns det exempel på identiska intronsekvenser när man jämför *Nostoc* i symbios med både *Anthoceros* och *Blasia* samt de olika kottepalmsarterna från olika delar av världen. Det visar att några *Nostoc* stammar verkar vara med kosmopolitiska än andra.

Jämförelser av intronsekvenser visar på stora likheter, även identiska konserverade regioner. Sekvensvariationer finns framförallt i vissa regioner som när man jämför med den sekundära strukturen består av specifika loopar eller så kallade hairpins. Alla storleksskillnader, eller skillnader i antal baspar, är lokaliserade till strukturelementet P6b, som i sin tur består av upprepningar av degenererade heptanukleotider med en förmåga att baspara och därmed bilda hairpins. Storleksvariationer inom intronet består av olika antal heptanukleotider, eller i några fall av unika icke kodande, ej repetetivt DNA, så kallade NIS (*Nostoc* Iterated Sequence). NIS element kan bilda extra hairpins och finns i ett repetetivt element eller mellan två konsekutiva element. Den sporadiska förekomsten av NIS element indikerar ett evolutionärt yngre ursprung och aktualiserar frågan hur dessa element har utvecklats. I *Nostoc punctiforme* genomet kan man identifiera många närbesläktade NIS element. De flesta är lokaliserade till icke kodande regioner medan några finns i gener. NIS elementen omges av repetetiva sekvenser vilket indikerar en spridnings mekanism som innefattar sekvenserna på båda sidorna om NIS elementet.

REFERENCES

1. ADAMS D. 2000. Symbiotic interactions. In *The ecology of cyanobacteria. Their diversity in time and space*, ed. B Whitton, M Potts, pp. 523-561. Dordrecht: Kluwer Academic Publishers. ISBN0792347552
2. ADAMS DG, DUGGAN PS. 1999. Heterocyst and akinete differentiation in cyanobacteria. *New Phytol* 144: 3-33
3. ADAMS PL, STAHLEY MR, KOSEK AB, WANG J, STROBEL SA. 2004. Crystal structure of a self-splicing group I intron with both exons. *Nature* 430: 45-50
4. AHERN CP, STAFF IA. 1994. Symbiosis in cycads: The origin and development of coralloid roots in *Macrozamia communis* (Cycadaceae). *Amer J Bot* 81: 1559-1570
5. AKKERMANS ADL, MIRZA MS, HARMSEN HJM, BLOK HJ, HERRON PR, et al. 1994. Molecular ecology of microbes: A review of promises, pitfalls and true progress. *FEMS Microbiol Rev* 15: 185-194
6. ASAYAMA M, KABASAWA M, TAKAHASHI I, AIDA T, SHIRAI M. 1996. Highly repetitive sequences and characteristics of genomic DNA in unicellular cyanobacterial strains. *FEMS Microbiol Lett* 137: 175-181
7. BARTHOLOMEW SE. 1986. The sporeling development of *Blasia pusilla* L. *J Hattori Bot Lab* 60: 255-261
8. BELFORT M, REABAN ME, COETZEE T, DALGAARD JZ. 1995. Prokaryotic introns and inteins: a panoply of form and function. *J Bacteriol* 177: 3897-3903
9. BELFORT M, ROBERTS RJ. 1997. Homing endonucleases: keeping the house in order. *Nucleic Acids Res* 25: 3379-3388
10. BELL-PEDERSEN D, QUIRK S, CLYMAN J, BELFORT M. 1990. Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. *Nucleic Acids Res* 18: 3763-3770
11. BERGMAN B, MATVEYEV A, RASMUSSEN U. 1996. Chemical signalling in cyanobacterial-plant symbioses. *Trends Plant Sci* 1: 191-197
12. BESENDAHL A, QIU YL, LEE J, PALMER JD, BHATTACHARYA D. 2000. The cyanobacterial origin and vertical transmission of the plastid $tRNA^{Leu}$ group-I intron. *Curr Genet* 37: 12-23
13. CAMPBELL EL, MEEKS JC. 1989. Characteristics of hormogonia formation by symbiotic *Nostoc* spp. in responce to the presence of *Anthoceros punctatus* or its extracellular products. *Appl Environ Microbiol* 55: 125-131
14. CANNONE JJ, SUBRAMANIAN S, SCHNARE MN, COLLETT JR, D'SOUZA LM, et al. 2002. The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* 3
15. CASTENHOLZ RW. 2001. Phylum BX. Cyanobacteria - Oxygenic photosynthetic bacteria. In *Bergey's manual of systematic bacteriology*, ed. GM Garrity, pp. 474-599. New York: Springer-Verlag. ISBN0387987711
16. CAVALIER-SMITH T. 1991. Intron phylogeny: a new hypothesis. *Trends Genet* 7: 145-148

17. CECH TR. 1987. The chemistry of self-splicing RNA and RNA enzymes. *Science* 236: 1532-1539
18. CECH TR. 1988. Conserved sequences and structures of group I introns: building an active site for RNA catalysis--a review. *Gene* 73: 259-271
19. CECH TR. 1990. Self-splicing of group I introns. *Annu Rev Biochem* 59: 543-568
20. CECH TR, DAMBERGER SH, GUTELL RR. 1994. Representation of the secondary and tertiary structure of group I introns. *Nat Struct Biol* 1: 273-280
21. CHRISTIAN EL, YARUS M. 1993. Metal coordination sites that contribute to structure and catalysis in the group I intron from *Tetrahymena*. *Biochemistry* 32: 4475-4480
22. COSTA JL, LINDBLAD P. 2002. Cyanobacteria in symbiosis with cycads. In *Cyanobacteria in symbiosis*, ed. AN Rai, B Bergman, U Rasmussen, pp. 195-205. Dordrecht: Kluwer Academic Publishers. ISBN1402007779
23. CRAIG N. 2002. Mobile DNA: an introduction. In *Mobile DNA II*, ed. N Craig, R Craigie, M Gellert, AM Lambowitz, pp. 3-11. Washington: ASM Press. ISBN1555812090
24. DARNELL JE, DOOLITTLE WF. 1986. Speculations on the early course of evolution. *Proc Natl Acad Sci U S A* 83: 1271-1275
25. DUCKETT JG, PRASAD AK, DAVIES DA, WALKER S. 1977. A cytological analysis of the *Nostoc*-Bryophyte relationship. *New Phytol* 79: 349-362
26. DUCKETT JG, RENZAGLIA KS. 1993. The reproductive biology of the liverwort *Blasia pusilla* L. *J Bryol* 17: 541-552
27. DUJON B, BELFORT M, BUTOW RA, JACQ C, LEMIEUX C, et al. 1989. Mobile introns: definition of terms and recommended nomenclature. *Gene* 82: 115-118
28. EDGELL DR, BELFORT M, SHUB DA. 2000. Barriers to intron promiscuity in bacteria. *J Bacteriol* 182: 5281-5289
29. ELLSWORTH DL, RITTENHOUSE KD, HONEYCUTT RL. 1993. Artifactual variation in randomly amplified polymorphic DNA banding patterns. *Biotechniques* 14: 214-217
30. ENDERLIN CS, MEEKS JC. 1983. Pure culture and reconstitution of the *Anthoceros* and *Nostoc* symbiotic association. *Planta* 158: 157-165
31. FAY P. 1992. Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev* 56: 340-373
32. FILIP ZA, ROHLMAN CE. 2001. Thermodynamic characterization of the *Anabaena* group I intron structure. *Abst Am Chem Soc* 221: U198-U199
33. FLEISCHMANN RD, ADAMS MD, WHITE O, CLAYTON RA, KIRKNESS EF, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269: 496-512
34. GILBERT W, MARCHIONNI M, MCKNIGHT G. 1986. On the antiquity of introns. *Cell* 46: 151-153
35. GORELOVA OA, BAULINA OI, KORZHENEVSKAYA TG, GUSEV MV. 1997. Formation of hormogonia and their taxis during the interaction of cyanobacteria and plants. *Mikrobiologiya* 66: 800-806
36. GRABOWSKI PJ, ZAUG AJ, CECH TR. 1981. The intervening sequence of the ribosomal RNA precursor is converted to a circular RNA in isolated nuclei of *Tetrahymena*. *Cell* 23: 467-476

37. GROSSHANS CA, CECH TR. 1989. Metal ion requirements for sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme. *Biochemistry* 28: 6888-6894
38. GUEVARA R, ARMESTO JJ, CARU M. 2002. Genetic diversity of *Nostoc* microsymbionts from *Gunnera tinctoria* revealed by PCR-STRR fingerprinting. *Microb Ecol* 44: 127-136
39. HASELKORN R, BUIKEMA WJ. 1992. Nitrogen fixation in cyanobacteria. In *Biological nitrogen fixation*, ed. G Stacey, RH Burris, HJ Evans, pp. 166-190. London: Chapman and Hall. ISBN0412024217
40. JAEGER L, MICHEL F, WESTHOF E. 1994. Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J Mol Biol* 236: 1271-1276
41. JAEGER L, WESTHOF E, MICHEL F. 1991. Function of P11, a tertiary base pairing in self-splicing introns of subgroup IA. *J Mol Biol* 221: 1153-1164
42. JOHANSSON C, BERGMAN B. 1992. Early events during the establishment of the *Gunnera-Nostoc* symbiosis. *Planta* 188: 403-413
43. JOHANSSON C, BERGMAN B. 1994. Reconstitution of the symbiosis of *Gunnera manicata* Linden: cyanobacterial specificity. *New Phytol* 126: 643-652
44. KANEKO T, NAKAMURA Y, WOLK CP, KURITZ T, SASAMOTO S, et al. 2001. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res* 8: 205-213
45. KIM J, KRAWCZYK K, LORENTZ W, ZIMMERMAN W. 1997. Fingerprinting cyanobionts and hosts of the *Azolla* symbioses by DNA amplification. *World J Microb Biot* 13: 97-100
46. KJEMS J, GARRETT RA. 1991. Ribosomal RNA introns in archaea and evidence for RNA conformational changes associated with splicing. *Proc Natl Acad Sci U S A* 88: 439-443
47. KNIGHT CD, ADAMS DG. 1996. A method for studying chemotaxis in nitrogen fixing cyanobacterium-plant symbioses. *Physiol Mol Plant P* 49: 73-77
48. KO M, CHOI H, PARK C. 2002. Group I self-splicing intron in the *recA* gene of *Bacillus anthracis*. *J Bacteriol* 184: 3917-3922
49. KRASSILOV V, SCHUSTER R. 1984. Paleozoic and Mesozoic fossils. In *New manual of bryology*, ed. R Schuster, pp. 1172-1193. Nichinan: The Hattori Botanical Laboratory
50. KUHSEL MG, STRICKLAND R, PALMER JD. 1990. An ancient group I intron shared by eubacteria and chloroplasts. *Science* 250: 1570-1573
51. KUNKEL B, LOSICK R, STRAGIER P. 1990. The *Bacillus subtilis* gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev* 4: 525-535
52. LACHANCE M-A. 1981. Genetic relatedness of heterocystous cyanobacteria by deoxyribonucleic acid-deoxyribonucleic acid reassociation. *Int J Syst Bacteriol* 31: 139-147
53. LAMBOWITZ AM, BELFORT M. 1993. Introns as mobile genetic elements. *Annu Rev Biochem* 62: 587-622
54. LAMBOWITZ AM, CAPRARA MG, ZIMMERLY S, PERLMAN PS. 1999. Group I and group II ribozymes as RNPs: clues to the past and guides to the future. In

The RNA world, ed. RF Gesteland, TR Cech, JF Atkins, pp. 451-485. New York: Cold Spring Harbor Laboratory Press. ISBN0879695897

55. LEHNERT V, JAEGER L, MICHELE F, WESTHOF E. 1996. New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem Biol* 3: 993-1009
56. LEWIN B. 2000. *Genes VII*. New York: Oxford University Press. 990 pp. ISBN0198792778
57. LIN S, HENZE S, LUNDGREN P, BERGMAN B, CARPENTER EJ. 1998. Whole-cell immunolocalization of nitrogenase in marine diazotrophic cyanobacteria *Trichodesmium* spp. *Appl Environ Microbiol* 64: 3052-3058
58. LINDBERG P, HANSEL A, LINDBLAD P. 2000. *hupS* and *hupL* constitute a transcription unit in the cyanobacterium *Nostoc* sp. PCC 73102. *Arch Microbiol* 174: 129-133.
59. LINDBLAD P, BERGMAN B, HOFSTEN AV, HALLBOM L, NYLUND JE. 1985. The cyanobacterium and *Zamia skinneri* symbiosis: An ultrastructural study. *New Phytol* 101: 707-716
60. LINDBLAD P, HASELKORN R, BERGMAN B, NIERZWICKI-BAUER SA. 1989. Comparison of DNA restriction fragment length polymorphisms of *Nostoc* strains in and from cycads. *Arch Microbiol* 152: 20-24
61. LOGSDON JM, JR. 1998. The recent origins of spliceosomal introns revisited. *Curr Opin Genet Dev* 8: 637-648
62. LOIZOS N, TILLIER ER, BELFORT M. 1994. Evolution of mobile group I introns: recognition of intron sequences by an intron-encoded endonuclease. *Proc Natl Acad Sci U S A* 91: 11983-11987
63. LUPTÁK A, DOUDNA JA. 2004. Distinct sites of phosphorothioate substitution interfere with folding and splicing of the *Anabaena* group I intron. *Nucleic Acids Res* 32: 2272-2280
64. LYKKE-ANDERSEN J, AAGAARD C, SEMIONENKOV M, GARRETT RA. 1997. Archaeal introns: splicing, intercellular mobility and evolution. *Trends Biochem Sci* 22: 326-331
65. MAZEL D, HOUMARD J, CASTETS AM, TANDEAU DE MARSAC N. 1990. Highly repetitive DNA sequences in cyanobacterial genomes. *J Bacteriol* 172: 2755-2761.
66. MEEKS J, ELHAI J, THIEL T, POTTS M, LARIMER F, et al. 2001. An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosynth Res* 70: 85-106
67. MEEKS JC. 1998. Symbiosis between nitrogen-fixing cyanobacteria and plants. *Bioscience* 48: 266-276
68. MICHEL F, DUJON B. 1983. Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear-encoded members. *Embo J* 2: 33-38
69. MICHEL F, JAEGER L, WESTHOF E, KURAS R, TIHY F, et al. 1992. Activation of the catalytic core of a group I intron by a remote 3' splice junction. *Genes Dev* 6: 1373-1385
70. MICHEL F, UMESONO K, OZEKI H. 1989. Comparative and functional anatomy of group II catalytic introns - a review. *Gene* 82: 5-30

71. MICHEL F, WESTHOF E. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* 216: 585-610
72. MILINDASUTA B-E. 1975. Developmental anatomy of coralloid roots in cycads. *Amer J Bot* 65: 468-472
73. MILLINEAUX PM, GALLON JR, CHAPLIN AE. 1981. Acetylene reduction (nitrogen fixation) by cyanobacteria grown under alternating light-dark cycles. *FEMS Microbiol Lett* 10: 245-247
74. MORGAN MJ. 2001. Initial sequencing and analysis of the human genome. *Nature* 409: 860-921
75. MURPHY FL, CECH TR. 1994. GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J Mol Biol* 236: 49-63
76. NATHANIELSZ CP, STAFF IA. 1975. Mode of entry of blue-green algae into the apogeotropic roots of *Macrozamia communis*. *Amer J Bot* 62: 232-235
77. NATHANIELSZ CP, STAFF IA. 1975. On the occurence of intracellular blue-green algae in cortical cells of the apogeotropic roots of *Macrozama communis*. *Ann Bot* 39: 363-368
78. NILSEN TW. 2003. The spliceosome: the most complex macromolecular machine in the cell? *Bioessays* 25: 1147-1149
79. NILSSON M, BERGMAN B, RASMUSSEN U. 2000. Cyanobacterial diversity in geographically related and distant host plants of the genus *Gunnera*. *Arch Microbiol* 173: 97-102
80. OKSANEN I, LOHTANDER K, PAULSRUD P, RIKKINEN J. 2002. A molecular approach to cyanobacterial diversity in a rock-pool community involving gelatinous lichens and free-living *Nostoc* colonies. *Ann Bot Fen* 39: 93-99
81. OLIVER RL, GANF GG. 2000. Freshwater blooms. In *The ecology of cyanobacteria. Their diversity in time and space*, ed. B Whitton, M Potts, pp. 149-194. Dordrecht: Kluwer Academic Publishers. ISBN0792347552
82. OW MC, GANTAR M, ELHAI J. 1999. Reconstitution of the cycad-cyanobacterial association. *Symbiosis* 27: 125-134
83. PAERL HW. 2000. Marine plankton. In *The ecology of cyanobacteria. Their diversity in time and space*, ed. B Whitton, M Potts, pp. 121-148. Dordrecht: Kluwer Academic Publishers. ISBN0792347552
84. PALMER JD, LOGSDON JM, JR. 1991. The recent origins of introns. *Curr Opin Genet Dev* 1: 470-477
85. PAQUIN B, HEINFLING A, SHUB DA. 1999. Sporadic distribution of $tRNA^{Arg}$ CCU introns among alpha-purple bacteria: evidence for horizontal transmission and transposition of a group I intron. *J Bacteriol* 181: 1049-1053
86. PAQUIN B, KATHE SD, NIERZWICKI-BAUER SA, SHUB DA. 1997. Origin and evolution of group I introns in cyanobacterial tRNA genes. *J Bacteriol* 179: 6798-6806
87. PAULSRUD P. 2001. *The Nostoc symbiont of lichens. Diversity, specificity, and cellular modifications*. Comprehensive summaries of Uppsala dissertations from the Faculty of Science and Technology 662. Uppsala University, Uppsala. 55 pp. ISBN9155451365
88. PAULSRUD P, LINDBLAD P. 1998. Sequence variation of the $tRNA^{Leu}$ intron as a marker for genetic diversity and specificity of symbiotic cyanobacteria in some lichens. *Appl Environ Microbiol* 64: 310-315

89. PAULSRUD P, RIKKINEN J, LINDBLAD P. 1998. Cyanobiont specificity in some *Nostoc*-containing lichens and in a *Peltigera aphthosa* photosymbiodeme. *New Phytol* 139: 517-524
90. PAULSRUD P, RIKKINEN J, LINDBLAD P. 2000. Spatial patterns of photobiont diversity in some *Nostoc*-containing lichens. *New Phytol* 146: 291-299
91. PAULSRUD P, RIKKINEN J, LINDBLAD P. 2001. Field investigations on cyanobacterial specificity in *Peltigera aphthosa*. *New Phytol* 152: 117-123
92. PETERS GA, MEEKS JC. 1989. The *Azolla-Anabaena* symbiosis: basic biology. *Annu Rev Plant Physiol Plant Mol Biol* 40: 193-210
93. POTTS M. 2000. *Nostoc*. In *The ecology of cyanobacteria. Their diversity in time and space*, ed. B Whitton, M Potts, pp. 465-504. Dordrecht: Kluwer Academic Publishers. ISBN0792347552
94. RAI A, SÖDERBÄCK E, BERGMAN B. 2000. Cyanobacterium-plant symbioses. *New Phytol* 147: 449-481
95. RANGAN P, WOODSON SA. 2003. Structural requirement for Mg^{2+} binding in the group I intron core. *J Mol Biol* 329: 229-238
96. RASMUSSEN U, JOHANSSON C, BERGMAN B. 1994. Early communication in the *Gunnera-Nostoc* symbiosis: Plant-induced cell differentiation and protein synthesisin the cyanobacterium. *Mol Plant Microbe Interact* 7: 696-702
97. RASMUSSEN U, NILSSON M. 2002. Cyanobacterial diversity and specificity in plant symbioses. In *Cyanobacteria in symbiosis*, ed. A Rai, B Bergman, U Rasmussen, pp. 313-328. Dordrecht: Kluwer Academic Publishers. ISBN1402007779
98. RASMUSSEN U, SVENNING MM. 1998. Fingerprinting of cyanobacteria based on PCR with primers derived from short and long tandemly repeated repetitive sequences. *Appl Environ Microbiol* 64: 265-272
99. RASMUSSEN U, SVENNING MM. 2001. Characterization by genotypic methods of symbiotic *Nostoc* strains isolated from five species of *Gunnera*. *Arch Microbiol* 176: 204-210
100. REINHOLD-HUREK B, SHUB DA. 1992. Self-splicing introns in tRNA genes of widely divergent bacteria. *Nature* 357: 173-176
101. RENZAGLIA JG. 1982. A comparative developmental investigation of the gametophyte generation in the Metzgeriales (Hepatophyta). *Bryophyt Bibl* 24: 1-253
102. RIKKINEN J, OKSANEN I, LOHTANDER K. 2002. Lichen guilds share related cyanobacterial symbionts. *Science* 297: 357
103. RIPPKA R, DERUELLES J, WATERBURY JB, HERDMAN M, STANIER RY. 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol* 111: 1-61
104. RODGERS GA, STEWART WD. 1977. The cyanophyte-hepatic symbiosis I. Morphology and physiology. *New Phytol* 78: 441-458
105. RUDI K, FOSSHEIM T, JAKOBSEN KS. 2002. Nested evolution of a $tRNA^{Leu}$ (UAA) group I intron by both horizontal intron transfer and recombination of the entire tRNA locus. *J Bacteriol* 184: 666-671
106. RUDI K, JAKOBSEN KS. 1997. Cyanobacterial $tRNA^{Leu}$ (UAA) group I introns have polyphyletic origin. *FEMS Microbiol Lett* 156: 293-298
107. RUDI K, JAKOBSEN KS. 1999. Complex evolutionary patterns of $tRNA^{Leu}$ (UAA) group I introns in the cyanobacterial radiation. *J Bacteriol* 181: 3445-3451

108. RUJAN T, MARTIN W. 2001. How many genes in Arabidopsis come from cyanobacteria? An estimate from 386 protein phylogenies. *Trends Genet* 17: 113-120
109. SAGAN L. 1967. On the origin of mitosing cells. *J Theor Biol* 14: 255-274
110. SALDANHA R, MOHR G, BELFORT M, LAMBOWITZ AM. 1993. Group I and group II introns. *Faseb J* 7: 15-24
111. SCHOFIELD WB. 1985. *Introduction to bryology*. New York: Collier Macmillan. 431 pp. ASIN0029496608
112. SCHOPF JW. 1993. Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life. *Science* 260: 640-646
113. SCHUSTER R. 1969. *Hepaticae and Anthocerotae of North America*: Columbia University Press. 1062 pp. ISBN0231089821
114. SHUB DA. 1991. The antiquity of group I introns. *Curr Opin Genet Dev* 1: 478-484
115. SIMON D, FEWER D, FRIEDL T, BHATTACHARYA D. 2003. Phylogeny and self-splicing ability of the plastid $tRNA^{Leu}$ group I Intron. *J Mol Evol* 57: 710-720
116. STEWART WD, RODGERS GA. 1977. The cyanophyte-hepatic symbiosis II. Nitrogen fixation and the interchange of nitrogen and carbon. *New Phytol* 78: 459-471
117. STRAND M, PROLLA TA, LISKAY RM, PETES TD. 1993. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* 365: 274-276
118. STRAUSS-SOUKUP JK, STROBEL SA. 2000. A chemical phylogeny of group I introns based upon interference mapping of a bacterial ribozyme. *J Mol Biol* 302: 339-358
119. SUMMERFIELD TC, GALLOWAY DJ, EATON-RYE JJ. 2002. Species of cyanolichens from Pseudocyphellaria with indistinguishable ITS sequences have different photobionts. *New Phytol* 155: 121-129
120. TAMAS I, SVIRCEV Z, ANDERSSON SG. 2000. Determinative value of a portion of the *nifH* sequence for the genera *Nostoc* and *Anabaena* (cyanobacteria). *Curr Microbiol* 41: 197-200
121. TANNER MA, CECH TR. 1997. Joining the two domains of a group I ribozyme to form the catalytic core. *Science* 275: 847-849
122. TAYLOR T, TAYLOR E. 1993. *The biology and evolution of fossil plants*. New Jersey: Prentice Hall. 982 pp. ASIN0136515894
123. VAN BELKUM A, SCHERER S, VAN ALPHEN L, VERBRUGH H. 1998. Short-sequence DNA repeats in prokaryotic genomes. *Microbiol Mol Biol Rev* 62: 275-293.
124. VENTER JC, ADAMS MD, MYERS EW, LI PW, MURAL RJ, et al. 2001. The sequence of the human genome. *Science* 291: 1304-1351
125. VEPRITSKIY AA, VITOL IA, NIERZWICKI-BAUER SA. 2002. Novel group I intron in the $tRNA^{Leu}$ (UAA) gene of a gamma-proteobacterium isolated from a deep subsurface environment. *J Bacteriol* 184: 1481-1487
126. WALDSICH C, MASQUIDA B, WESTHOF E, SCHROEDER R. 2002. Monitoring intermediate folding states of the td group I intron in vivo. *Embo J* 21: 5281-5291
127. WATSON JD, CRICK FH. 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171: 737-738

128. WEBB DT, NEVAREZ M, DE JESUS S. 1984. Further in vitro studies of light-induced root nodulation in the cycadales. *Environ Exp Bot* 24: 37-41
129. WEST N, ADAMS D. 1997. Phenotypic and genotypic comparison of symbiotic and free-living cyanobacteria from a single field site. *Appl Environ Microbiol* 63: 4479-4484
130. WHITELOCK LM. 2002. *The cycads*. Portland: Timber Press. 374 pp. ISBN0881925225
131. WHITTON B, POTTS M. 2000. Introduction to the cyanobacteria. In *The ecology of cyanobacteria. Their diversity in time and space*, ed. B Whitton, M Potts, pp. 1-11. Dordrecht: Kluwer Academic Publishers. ISBN0792347552
132. WHITTON BA, POTTS M. 2000. *The ecology of cyanobacteria. Their diversity in time and space*. Dordrecht: Kluwer Academic Publishers. 669 pp. ISBN0792347552
133. WIRTZ N, LUMBSCH HT, GREEN TGA, TURK R, PINTADO A, et al. 2003. Lichen fungi have low cyanobiont selectivity in maritime Antarctica. *New Phytol* 160: 177-183
134. XU MQ, KATHE SD, GOODRICH-BLAIR H, NIERZWICKI-BAUER SA, SHUB DA. 1990. Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria. *Science* 250: 1566-1570
135. YOUNG JPW. 1992. Phylogenetic classification of nitrogen fixing organisms. In *Biological nitrogen fixation*, ed. G Stacey, RH Burris, HJ Evans, pp. 43-86. London: Chapman and Hall. ISBN0412024217
136. ZAUG AJ, DAVILA-APONTE JA, CECH TR. 1994. Catalysis of RNA cleavage by a ribozyme derived from the group I intron of *Anabaena* pre-tRNA^{Leu}. *Biochemistry* 33: 14935-14947
137. ZAUG AJ, GRABOWSKI PJ, CECH TR. 1983. Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage-ligation reaction. *Nature* 301: 578-583
138. ZAUG AJ, MCEVOY MM, CECH TR. 1993. Self-splicing of the group-I intron from *Anabaena* pre-transfer-RNA - Requirement for base-pairing of the exons in the anticodon stem. *Biochemistry* 32: 7946-7953
139. ZHENG W, SONG T, BAO X, BERGMAN B, RASMUSSEN U. 2002. High cyanobacterial diversity in coralloid roots of cycads revealed by PCR fingerprinting. *FEMS Microbiol Ecol* 40: 215-222
140. ZHENG WW, NILSSON M, BERGMAN B, RASMUSSEN U. 1999. Genetic diversity and classification of cyanobacteria in different *Azolla* species by the use of PCR fingerprinting. *Theor Appl Genet* 99: 1187-1193
141. ZIMMERMAN W, BERGMAN B. 1990. The *Gunnera* symbiosis: DNA restriction fragment length polymorphism and protein comparisions of *Nostoc* symbionts. *Microb Ecol* 19: 291-302

Acta Universitatis Upsaliensis

Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology*. (Prior to October, 1993, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science".)

Distribution:
Uppsala University Library
Box 510, SE-751 20 Uppsala, Sweden
www.uu.se, acta@ub.uu.se

ISSN 1104-232X
ISBN 91-554-6033-X