Chapter 1

General introduction
**Microbial diversity**

Life on Earth is subdivided in three major domains, the *Archaea*, the *Bacteria* and the *Eukarya* (Woese and Fox, 1977). The *Eukarya* comprises both macro- and microorganisms while the other two domains are exclusively microbial. With the invention of the microscope in the 17th century by Antonie van Leeuwenhoek, microbial life became visible for the human eye and allowed the exploration of this fascinating world. For a very long time, the microscope was the only instrument that allowed researchers to study microorganisms in the natural environment without the need for isolation and cultivation. However, the morphological characteristics of microorganisms are few and their isolation into laboratory cultures yielded only low numbers of species. For many years the known species diversity in the microbial world was therefore low, although microbiologists realized that the vast majority remained undiscovered. Only with the application of culture independent molecular biological techniques it became clear that there is a staggering diversity in the microbial world, which will probably prove to be beyond our comprehension.

Microorganisms are the dominant component of any ecosystem on Earth and many ecosystems are exclusively microbial. Microbes are found in a wide range of environmental conditions. For instance, microorganisms grow in environments with temperatures ranging from below 0 to above 100°C (Blank *et al.*, 2002; Sattley and Madigan, 2007; van der Meer *et al.*, 2007). Other environmental extremes under which microorganisms may proliferate include hypersaline, extremely acidic or alkaline and arid environments as well as systems where hyperbaric pressure or electromagnetic radiation prevent any other lifeform (van der Wielen *et al.*, 2005; Appukuttan *et al.*, 2006; Dong *et al.*, 2007; Nogi *et al.*, 2007; Islam *et al.*, 2008). Microorganisms evolved specific adaptations that enable them to live in these harsh environments (Makarova *et al.*, 2001; Swire, 2007).

The high species diversity among microbes is not only related to their survival in a wide range of different ecosystems, but also within a particular ecosystem the microbial diversity is usually very large. This is especially intriguing for the largely unstructured aquatic ecosystems such as oceans, seas and lakes. In aquatic ecosystems the plankton community consists of many coexisting species competing for the same resources. Hutchinson (1961) called this “The paradox of the plankton”. This is an interesting paradox because of the nutrient deficiencies that occur in summer in most aquatic ecosystems. One would expect that nutrient deficiency leads to fierce competition among species depending on these nutrients. Nevertheless, even under these resource limiting conditions the species diversity in aquatic environments usually remains high. In order to solve the paradox formulated by Hutchinson it is necessary to understand the mechanisms responsible for the existence of the high species diversity within the plankton specifically and for microbial communities in general. These mechanisms can be studied using clearly distinguishable microorganisms (morphological, phenotypical or genotypic differences) that can be compared for their adaptation to specific environmental conditions. Since many microorganisms have only limited morphological differences it was difficult to study the mechanisms underlying their diversity until the introduction of molecular techniques.
Chapter 1

The introduction of culture independent molecular biological techniques in microbial ecology in the 1980’s revealed far greater microbial species diversity in many ecosystems than had been expected. For instance, the polymerase chain reaction (PCR) enabled the amplification of ribosomal genes in DNA samples extracted directly from the environment without the need to cultivate microorganisms (Olsen et al., 1986; Pace et al., 1986). After the initial work of Norman Pace and collaborators many researchers have used PCR and PCR-dependent techniques in order to describe the microbial diversity for the three domains of life in a wide range of ecosystems. This revealed that hundreds of species might coexist in one milliliter of ocean water, while in one gram of soil more than 10,000 different species have been detected (Torsvik et al., 1996; Curtis and Sloan, 2004; Schloss and Handelsman, 2005). More recently, it was shown that the number of species could be in the range of millions for marine and soil ecosystems (Gans et al., 2005; Sogin et al., 2006). These very large numbers of species present in the environment contrast strongly with the low number of microorganisms that have been cultivated. Culture-independent approaches indicated that many of the newly identified taxa and species have not been cultivated and are only known by their 16S rRNA sequences (Ward et al., 1990; Rappe and Giovannoni, 2003; Sogin et al., 2006). We do not know who they are, what they look like and what they are doing.

The discrepancy between the number of microbial species in the environment and the number of cultivated microorganisms was known long before molecular biology techniques were applied in microbiology. It has been noted for a long time that there is a large discrepancy between the number of colony-forming bacteria and the number of bacteria that were counted microscopically. The cultivation efficiency of bacteria in marine water samples using standard marine media ranges between 0.1 and 0.01% (Kogure et al., 1979). This difference between cell counts using microscopy or flow-cytometry and the number of cultivated bacteria in the same sample is known as “The great plate count anomaly” (Staley and Konopka, 1985). Moreover, we now know that many of the uncultivated bacteria appear to be dominant in the natural environment while most cultivated bacteria are rare (for reviews, see e.g. Rappe and Giovannoni, 2003; Pedros-Alio, 2006). This led to the development of new and improved methods to enhance the number of bacteria that can be isolated and cultivated in the laboratory, specifically those species that are numerically important in natural communities (Button et al., 1993; Bruns et al., 2002; Connon and Giovannoni, 2002; Zengler et al., 2002; Ferrari et al., 2005). With the introduction of these novel cultivation techniques several new microorganisms have been isolated and can now be grown in the laboratory, but they still represent only a tiny fraction of the enormous diversity found in the microbial world.
The microbial species definition

Historically, bacterial species are defined based on morphology, physiology and metabolic capacities of pure cultures (Gevers et al., 2005). However, taxonomy of microorganisms based on these characteristics do not reflect their phylogenetic relationships. For instance, Gram-negative bacteria belonging to the group of pseudomonads were characterized by the usage of only one carbon source, acetate. Later on it was discovered through 16S rRNA gene sequencing that these organisms belonged to three different evolutionary groups, the Alpha-, Beta-, and Gammaproteobacteria (Staley, 2006). Nowadays, phylogenetic inference of prokaryotes utilizes often the ribosomal genes (particularly the 16S and 18S rRNA genes in Bacteria/Archaea and Eukarya, respectively) since they are present in all forms of life and are highly conserved. The usage of the 16S rRNA gene sequences led to the discovery of many unknown taxa at various taxonomic levels from species to genera and even to the discovery of the third kingdom of life: the Archaea (Woese and Fox, 1977; Pace, 1997; Rappe and Giovannoni, 2003). With the sequencing of the gene coding for the 16S small ribosomal subunit and the rapidly expanding number of such sequences deposited in databases, it became possible to produce a reasonable sound phylogenetic taxonomy of Bacteria and Archaea (Woese et al., 1990). The problem with this 16S rRNA taxonomy was that it did neither reflect morphology, nor the physiology and metabolic capacities of the organisms in a consistent way. In general, a prokaryotic species would include a collection of strains with approximately 70% or greater DNA-DNA hybridization (DDH) values (Wayne et al., 1987; Stackebrandt et al., 2002). This agrees in many cases with the proposed limit of 97% similarity between the 16S rRNA gene sequences to separate species. Strains showing less than 97% 16S rRNA similarity are considered different species, while strains with more than 97% 16S rRNA similarity could be the same species (Stackebrandt and Goebel, 1994). Likewise 95% similarity of the 16S rRNA has been considered to mark the level of the genus (Konstantinidis and Tiedje, 2007). Recently, genomic analysis of closely related species indicated that the average nucleotide identity (ANI) of the shared genes between genomes could be around 94%. This measure corresponded with the 70% DDH standard of the above species definition (Konstantinidis and Tiedje, 2005).

The discussion above shows that the bacterial species definition changed from a phenotypically based approach to one that is pure phylogenetic. The most recent species definitions for prokaryotes use distinct DNA sequences of one or more genes, or genomes, to describe monophyletic clusters (Gevers et al., 2005; Staley, 2006; Cohan and Perry, 2007). The reasoning behind this is that through time each cluster independently evolved of other clusters while acquiring distinct adaptations that are only shared by the entire cluster (Cohan and Perry, 2007).
Ecotypes

The ease of use of the 16S rRNA gene sequences to describe novel microbial species has led to many discoveries and has revolutionized our knowledge on microbial diversity. But the conserved nature of the 16S rRNA gene has also a downside, since it is possible that microorganisms with identical 16S rRNA genes should be assigned to different species when DNA-DNA hybridization reveals less than 70% similarity (Fox et al., 1992). Furthermore, with the increase of the number of ribosomal sequences in the databases it became apparent that closely related 16S or 18S rRNA sequences (similarity > 96%) often belong to clusters of species with different physiologies that inhabit distinct niches in the environment (Rappe and Giovannoni, 2003). The 16S rRNA microdiversity clusters are often accompanied by large differences between the genomes of closely related species. Processes like Horizontal Gene Transfer (HGT), which causes genes to be exchanged between bacteria, can be responsible for the differences found between closely related genomes. In this way, HGT could complicate the analysis of phylogenetic relationships between microorganisms because species boundaries become blurred (Zhaxybayeva et al., 2006; Choi and Kim, 2007). This suggests that the microdiversity clusters of 16S rRNA, and other genes, are of evolutionary and of ecological importance (Cohan, 2001; Rappe and Giovannoni, 2003; Acinas et al., 2004).

An example of the ecological distinctiveness of strains with closely related 16S rRNA gene sequences is the comparison of two groups of picocyanobacteria both belonging to the genus Prochlorococcus. Prochlorococcus differs from all other cyanobacteria because they use divinyl chlorophyll $a_2$ and $b$ for light harvesting instead of the phycobiliproteins. The strains MED4 and MIT9313 differ with regard to the optimum light intensity for growth (Moore et al., 1998). Strain MED4 tolerates high light intensities, has a low Chl $b/a_2$ ratio and belongs phylogenetically to the low B/A clade. Strain MIT9313 is sensitive to high light, has a high Chl $b/a_2$ ratio and belongs to the high B/A clade (Moore and Chisholm, 1999). Despite these differences the sequences of the 16S rRNA genes of both strains diverge less than 3% (Rocap et al., 2003). However, the genomes are much more divergent. The genome of strain MED4 has 1716 genes, while that of strain MIT9313 contains 2275 genes. Both genomes share 1352 genes. The remaining genes are strain specific. However, the 923 genes specific of strain MIT9313 are shared with Synechococcus WH8102 (Rocap et al., 2003; Hess, 2004), another marine picocyanobacterium. The differences between the genomes of both Prochlorococcus strains were attributed to the specific ecological niches where they thrive. Prochlorococcus MED4 occurs in the surface layers where it is exposed to high light while MIT9313 thrives at greater depth with lower light intensities (Moore et al., 1998; West and Scanlan, 1999). Based on the different ecological niches of these Prochlorococcus strains, they were assigned as “ecotypes” of the same species (Rocap et al., 2003).

Cohan and Perry (2007) defined “ecotype” in the following way: “A group of bacteria that are ecologically similar to one another, so similar that genetic diversity within the ecotype is limited by a cohesive force, either periodic selection or genetic drift, or both.” The above example on Prochlorococcus ecotypes, and other examples not discussed here, indicate that a microbial species in the microbial world consists of assemblages of closely related 16S rRNA
genomes, but that these assemblages may proliferate as ecologically distinct populations or
ecotypes (Rocap et al., 2003; Lopez-Lopez et al., 2005; Staley, 2006; Cohan and Perry, 2007;
Ward et al., 2008).

The color of light

The ecotype differentiation in Prochlorococcus is mainly based on its ability to grow optimal
at high or low light levels. However, the light gradient in the water column alone does not explain
the huge diversity of the phytoplankton. The phytoplankton community consists of many different
species, each with their own nutrient requirements and nutrient uptake characteristics, their strategies
to escape predation, differences in light harvesting properties and pigmentation, and many other
factors (Irigoien et al., 2004; Wilson et al., 2007). Stomp et al. (2004) used two strains of Synechococcus
for competition experiments in chemostats. One strain (BS4) contained only the blue-green pigment phycocyanin (PC), while the other (BS5) contained mainly the red phycocerythrin (PE). These authors demonstrated that red and green
cyanobacteria can coexist under white light.

The light harvesting complexes of cyanobacteria, the phycobilisomes, consist of
phycobiliproteins with different chromophores and, hence, different light harvesting capacities
(Fig. 1.1). These protein complexes differ in their light absorbance maxima by the chromophores
they bind. PE binds phycerythrobilin (PEB) ($\lambda_{\text{max}}$: 540 to 560 nm) which results in a maximal
absorption peak at ~565nm. PC binds phycocyanobilin (PCB) which results in an absorption
maximum at ~620nm (Ong and Glazer, 1991). The difference in light absorption capacities of
the pigments enables the Synechococcus strains BS4 and BS5 to coexist under white light in the
laboratory (Stomp et al., 2004) because they each use a different part of the light spectrum.
These differently pigmented Synechococcus species can be regarded as ecotypes.

Another pigment ecotype exists among Synechococcus. This ecotype binds an additional
pigment, phycourobilin (PUB) ($\lambda_{\text{max}}$ ~ 490 nm) to its phycoerythrin protein complex (Wood
et al., 1985; Ong and Glazer, 1991). Synechococcus species that possess PUB as well as PEB
chromophores exhibit a large range of variation in the PUB/PEB ratio (Fuller et al., 2003),
resulting in optically different phenotypes (red to orange). Some marine Synechococcus strains
are capable of a special form of chromatic adaptation by which they change the PUB/PEB ratio
in response to changes in the light spectrum (Palenik, 2001). Cyanobacteria possessing PUB
are exclusively found in the oceans, suggesting that the evolution of PUB is an adaptation to
blue light that prevails in the marine environment.

In contrast to the Prochlorococcus ecotypes, Synechococcus pigment ecotypes are not clearly
distinguishable from their 16S rRNA and ITS-1 phylogenies. Many clusters contain isolates
that possess different pigmentation phenotypes (Crosbie et al., 2003; Ernst et al., 2003; Fuller
et al., 2003). For example, the two Synechococcus strains used in the experiments of Stomp
et al. (2004) were identical for the 16S rRNA gene and nearly identical (similarity > 99%)
for the ribosomal internal transcribed spacer region (Crosbie et al., 2003; Ernst et al., 2003).
Figure 1.1 A schematic representation of the phycobilisomes in picocyanobacteria. The phycobilisome has a core consisting of allophycocyanin (APC) protein complexes. The allophycocyanin core is attached to photosystem I or II present in the thylakoids membrane and it transfers the light energy absorbed by the rods to the photosynthetic reaction centers. Attached to the APC core are the rods that harvest the light energy. Close to the core, discs consisting of hexameric complexes of phycocyanin proteins with PCB as the pigment (A). The next discs are composed of hexameric complexes with phycoerythrin (PE) I (B) and the most distal discs contain PE II proteins (C and D). In general, PE I binds only PEB (B), while PUB binds to PE II (C and D). Picture reproduced from Six et al. (2007) under the Creative Commons Attribution License version 2.0.
Furthermore, *Synechococcus* strains may not only vary with respect to their pigment composition, but they may also use different nitrogen sources and may differ with respect to motility. None of these properties are reflected in the 16S rRNA phylogeny (Fuller *et al.*, 2003; Scanlan, 2003; Ernst *et al.*, 2005).

Hence, it is difficult to explain the evolution of *Synechococcus* ecotypes from the level of the ribosomal operon as was done in the case of *Prochlorococcus*. The *Synechococcus* strains BS4 and BS5 were isolated from the Bornholm Sea together with a thin brown filamentous cyanobacterium that is capable of complementary chromatic adaptation (CCA) (Stal *et al.*, 2003). This cyanobacterium was identified as *Pseudanabaena* sp. Cyanobacteria capable of CCA are able to change their relative amounts of phycoerythrin and phycocyanin, in response to the color of the light they are exposed to. Under white light they absorb light using both PE and PC, rendering the organism a brown to black appearance. These cyanobacteria are able to compensate for changes in the light spectrum by adjustment of the PE and PC composition in their phycobilisomes (Kehoe and Gutu, 2006). Hence, in red light the organism turns (blue) green and in green light it turns red.

Stomp *et al.* (2004) used as a model organism the filamentous cyanobacterium *Tolypothrix* that is capable of CCA. These authors showed that *Tolypothrix* coexisted with either BS4 or BS5. In competition against the green *Synechococcus* strains BS4, *Tolypothrix* turned green. Conversely, in competition against the red *Synechococcus* strain BS5, *Tolypothrix* turned green. Thus, the species capable of CCA took advantage of the photons that were not used by its competitors. The rate at which species can change their pigmentation also plays a key role in phytoplankton competition, as has been shown in competition experiments with *Pseudanabaena* strains capable of CCA (Stomp *et al.*, in press). To assess the pigmentation of *Synechococcus* and *Pseudanabaena* in their natural habitat, many strains of *Synechococcus* and *Pseudanabaena* were isolated from the Baltic Sea. The investigation of these isolates raised questions with regard to their diversity, abundance and distribution in the natural environment. This thesis aimed at answering some of these questions.
Aim of this thesis

The aim of this thesis was to investigate the distribution of differently pigmented coexisting picocyanobacteria in the natural environment and to study the mechanisms of ecotype differentiation in *Synechococcus* and *Pseudanabaena* populations at the genetic level using culture independent methods as well as cultivated isolates.

Outline

This thesis consists of four scientific papers and a synthesis.

Chapter 2: Colourful coexistence of red and green picocyanobacteria in lakes and seas.

In this chapter 70 different aquatic ecosystems (ranging from the oligotrophic ocean to turbid brown peat lakes) were analyzed with respect to the underwater light spectrum and the abundance of red and green picocyanobacteria (Stomp et al. 2007). The results were compared with a parameterized competition model that predicted opportunities for coexistence of red and green phytoplankton. The field data were consistent with laboratory experiments showing the coexistence of red and green picocyanobacteria (Stomp et al., 2004), and proved coexistence of red and green phytoplankton species in lakes and seas by niche differentiation along the light spectrum.

![Diagram](image.png)

**Figure 1.2** The operon structure encoding the alpha and beta subunit genes of the phycoerythrin (cpe) and phycocyanin (cpc) proteins. The transcription of the genes is from left to right. Indicated by the arrows are the forward (black) and reverse (white) primers used in this thesis. Primer names accompany the arrows. The information of the mRNA is translated in amino-acid chains encoding the alpha and beta subunits. Both amino-acid chains are subsequently assembled into one protein.
Chapter 3: Diversity and phylogeny of Baltic Sea picocyanobacteria inferred from their ITS and phycobiliprotein operons.

In this chapter the distribution of red and green picocyanobacteria in the Baltic Sea is reported (Haverkamp et al. 2008). Using flow-cytometry and clone libraries derived from DNA extracted from environmental samples it is demonstrated that phycoerythrin (PE) (red) and phycocyanin (PC) (green) rich Synechococcus coexist. The Baltic Sea picocyanobacteria are related to freshwater Synechococcus. Furthermore, analysis of the genes encoding for the PE and PC proteins (Fig 1.2) showed that the Synechococcus group separated phylogenetically into three different ecotypes, each characterized by a different pigmentation.

Chapter 4: Rapid diversification of red and green Synechococcus strains in the Baltic Sea.

Almost 50 closely related Synechococcus isolates were analyzed in order to assess the importance of microdiversity within this genus. By using a multi-locus sequence approach it was demonstrated that closely related genotypes at the 16S rRNA level can have different pigmentation phenotypes. Due to this microdiversity it is difficult to interpret environmental studies using clone libraries of Synechococcus 16S rRNA genes. Closely related 16S rRNA genotypes may exhibit totally different phenotypes. This can only be resolved through cultures that provide the whole genome. Using an extensive culture collection of Baltic Sea Synechococcus strains it was discovered that horizontal gene transfer was probably involved in the generation of microdiversity among this group of picocyanobacteria.

Chapter 5: Phenotypic and genetic diversification of Pseudanabaena.

This chapter focuses on the extension of the present genetic knowledge of the small filamentous cyanobacteria Pseudanabaena. Isolates of Pseudanabaena from the Baltic Sea and from the Albufera de Valencia (Spain) were used to identify various lineages that showed high levels of microdiversity. The results hinted to the presence of endemic and cosmopolitan species present in both ecosystems. Furthermore, it was found that purifying selection at the locus of the phycocyanin operon promoted evolutionary diversification in populations of Pseudanabaena.

Chapter 6: General Discussion.

This final chapter integrates the research presented in this thesis and discusses the results in their connection to one another, reaching some overall conclusions.