A multi-gene approach in the systematics of the phototrophic purple sulfur bacteria genera *Marichromatium* and *Allochromatium* with description of the new species *Allochromatium humboldtianum* and the new biotype *Marichromatium gracile* biotype *thermosulfidiphilum*.

Dissertation
zur
Erlangung des Grades eines
Doktors der Naturwissenschaften
-Dr. rer. nat.-

dem Fachbereich Biologie/Chemie der
Universität Bremen vorgelegt von

Wilbert Serrano
aus
Lima-Peru

Bremen 2009

Die Untersuchungen zur vorlegelegten Doktorarbeit wurden in der Zeit von Juli 2005 bis Juni 2009 am UFT der Universität Bremen und am Max Planck Institut für Marine Mikrobiologie in Bremen durchgeführt.

1. Gutachter: Prof. Dr. Ulrich Fischer

2. Gutachter: Prof. Dr. Rudolf Amann

Tag des Promotionskolloquiums: 23. Juli 2009.

Für meine Eltern

Table of contents

Table of contents 3

List of abbreviations.. ..5

Summary..7

Zusammenfassung 9

Chapter 1

Introduction 11

1. A general overview of bacterial sytematics 11

2. A multi-gene taxonomy in prokaryotes 17

3. Photosynthetic bacteria 20

3.1. Antenna complex...24

3.2. Reaction center 25

4. The family *Chromatiaceae*..26

4.1. Environmental factors 27

4.2. Importance...28

4.3. Carbon fixation and sulfur oxidation by purple sulfur bacteria..........29

5. Material and methods..30

5.1. Source of microorganisms... 30

5.2. Media and cultivation conditions...32

5.3. Enrichment and isolation..34

5.4. Genomic DNA isolation and molecular analysis.........................35

5.4.1. Genomic DNA "fishing" extraction protocol...............................35

5.4.2. G+C mol% determination...37

5.4.3. DNA-DNA reassociation...38

5.4.4. Protein-coding gene analysis...39

5.4.5. Population genetic analysis..40

6. Aim of this study... 41

7. Results and discussion..*42*

7.1. Characterization of a new purple sulfur bacterium, *Marichromatium gracile* biotype *thermosulfidiphilum* **(Chapter 2)**. *42*

7.2. A multi-gene phylogeny of the genus *Marichromatium* **(Chapter 3)**.*44*

7.3. A multi-gene phylogeny of the genus *Allochromatium* **(Chapter 4)**....*47*

7.4. Description of a novel *Allochromatium* species; *A. humboldtianum* sp. nov. **(Chapter 5)**...*49*

8. This work (List of publications)..**50**

9. References *52*

Chapter 2

A new moderately thermophilic and high sulfide tolerant biotype of *Marichromatium gracile,* isolated from tidal sediments of the German Wadden Sea: *Marichromatium gracile* biotype *thermosulfidiphilum*.......................... *63*

Chapter 3

Evaluation of the use of Multilocus Sequence Analysis (MLSA) to resolve taxonomic conflicts within the genus *Marichromatium.* *79*

Chapter 4

The Genus *Allochromatium* (*Chromatiales*, *Chromatiaceae*) revisited: a study of its intragenic structure based on multilocus sequence analysis (MLSA) versus DNA-DNA Hybridization (DDH)..*117*

Chapter 5

Allochromatium humboldtianum sp. nov., a novel *Allochromatium* species isolated from soft marine sediments from Peruvian coast *137*

Appendices *149*

Acknowledgements *156*

List of abbreviations

AAI	Amino acid identity
Bchl	Bacteriochlorophyll
BOX A1R	Highly conserved repeated sequence, based on BOX repetitive elements present in the chromosome of *Streptococcus pneumoniae*
DDH	DNA-DNA hybridization
DIG DNA	Digoxigenin-dUTP labeling DNA probe
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
d*N*/d*S*	Ratio of the rate of non synonymous substitutions (d*N*) to the rate of synonymous substitutions (d*S*)
ERIC	Enterobacterial repetitive intergenic consensus sequences
H	gene diversity
HGT	Horizontal gene transfer
ITS	Internal transcriber spacer
LH	Photosynthetic **light harvesting** complex
MLEE	Multilocus enzyme electrophoresis
MLST	Multilocus sequence typing
MLSA	Multilocus sequence analysis
Mol% G+C	Percentage of guanine-cytosine content in the genome
M13-Core	Bacteriophage M13 core repeated sequence
π	nucleotide diversity
PhyML	Phylogenies by maximum likelihood
RAPD	Ramdom amplified polymorphic DNA
RC	Photosynthetic **reaction center** complex
SG	Segregation site
TD-PCR	Touch-down PCR
T'D	Tajima test of neutrality

Summary

In the present thesis two genera of the photosynthetic Purple Sulfur Bacteria (PSB) of the family *Chromatiaceae* are revised by molecular techniques which include extensive DNA-DNA hybridization (DDH) studies and a multi-gene analysis. Traditionally, species of PSB have been described based on some easily recognizable phenotypic properties. Separation of the PSB in two groups was possible by a diagnostical key property, e.g. internal or external storage of sulfur globules. This separation of PSB in the two families *Chromatiaceae* and *Ectothiorhosdopiraceae* was confirmed by comparing their 16S rRNA gene sequences (Imhoff *et al.,* 1998, Int. J. Syst. Bacteriol. 48, 1129-1143).

Within the *Chromatiaceae* the genus *Chromatium* was also separated in several genera being the most important the genera *Marichromatium* and *Allochromatium*. However, until now only a low number of species have been reported within each genus. Type species from both genera do not exhibit remarkable phenotypic differences and they seem to be closely related phylogenetically (16S rRNA-based) and genomically (Serrano *et al*., 2009, Syst. Appl. Microbiol. 32, 1-7).

Type species and strains of both genera examined in the present study were obtained from culture collections as well as isolated from geographically distant locations. A multi-gene analysis was performed based on PCR amplification and sequencing of a set of protein coding-genes such as *pufM*, *soxB*, *recA*, *fusA*, *dnaK*, and *gyrB* for *Marichromatium* and additionally *cbbL-1, mfd,* and *groES* genes for *Allochromatium*. In addition, extensive cross-DDH comparison among all the available strains within each genus were also made.

Strains within the genera *Marichromatium* and *Allochromatium* showed a high 16S rRNA gene sequence similarity ranging from 98 to 100% (mean 99.3%) and 97.0 to 100% (mean 98.0%) respectively. However, cross-DDH comparison, a standard tool for the definition of bacteria species, showed ambiguous results. On one hand they confirmed the 16S rRNA-based phylogenetic relationship, on the other hand they were contradicting, thus making species definition within both genera difficult. However, these phylo-genomic incongruences could be corrected by using a MLSA approach. The analysis of the concatenated protein-coding gene sequences for the genera *Marichromatium* (4331 bp) and *Allochromatium* (4427 bp) allowed the separation of all examined strains within each genus in several independent clusters of phylotypes.

Two new PSB strains, one of each genus, were collected from different geographical locations and completely characterized by using a polyphasic approach. The new strain *M. gracile* biotype *thermosulphidiphilum,* isolated from the northern German Wadden Sea, showed some remarkable phenotypic properties not reported for the corresponding genus. The new biotype shows a moderate high thermal (up to 44℃) and high sufide concentration (up to 16 mM) tolerances as well as a plasmid (see Chapter 2). However, despite of such unusual behaviour, this new strain could not be classified as a new species because of its high genetic similarity with respect to the type strain *M. gracile* DSM 203^T. This close relationship was also supported by multi-gene analysis. Thus, this strain was included within the Phylotype I of the genus *Marichromatium* as designed by MLSA (see Chapter 3).

On the other hand, the new strain AX1YPE, isolated from marine soft sediments from the Peruvian coast and belonging to the genus *Allochromatium* showed also some particular phenotypic features not described for this genus. The organism was able to use a wide range of nitrogen compounds as nitrogen sources and was able to grow at low light intensities as well (15 - 60 µE m^{-2} s^{-1} PAR) (see Chapter 5). Based on these phenotypic properties, the 16S rRNA phylogenetic analysis, and DDH relatedness, strain AX1YPE is reported as a new species of the genus *Allochromatium* for which the name *A. humboldtianum* is proposed. Moreover, the separation of strain AX1YPE as an independent lineage within the genus *Allochromatium* was also confirmed by MLSA (see chapter 4).

Thus, the MLSA approach could be used as a valuable tool for separating closely related strains. In the present work some examined strains within the genera *Marichromatium* and *Allochromatium* appeared to form a well identified and monophyletic unit, that in many cases reflect previous species circumscription based on phenotypic analysis.

Zusammenfassung

In der hier vorgelegten Arbeit wurden zwei Genera der photosynthetischen Roten Schwefelbakterien (RSB) aus der Familie der *Chromatiaceae* mittels molekularer Techniken überarbeitet, unter Einbeziehung extensiver Studien zur DNA-DNA Hybridisierung (DDH) und von Mutli-Gen-Analysen.
Herkömmlicherweise sind RSB-Spezies aufgrund einiger leicht zu erkennenden Eigenschaften beschrieben worden. Eine Auftrennung der RSB in zwei Gruppen erfolgte unter anderem aufgrund deren Fähigkeit, die gebildeten Schwefelkugeln entweder innerhalb oder außerhalb der Zelle zu lagern. Diese Aufteilung der RSB in die beiden Familien *Chromatiaceae* und *Ectothiorhodospiraceae* wurde durch einen Vergleich ihrer 16S rRNA Gensequenzen bestätigt (Imhoff *et al.,* 1998, Int. J. Syst. Bacteriol 48, 1129-1143).

Innerhalb der *Chromatiaceae* wurde auch das Genus *Chromatium* in mehrere Genera aufgeteilt, von denen die Genera *Marichromatium* und *Allochromatium* die wichtigsten sind. Bis heute jedoch ist erst eine kleine Anzahl an Spezies für jedes Genus benannt worden.
Typen-Spezies beider Genera weisen keine spektakulären phenotypischen Unterschiede auf, und sie scheinen mehr phylogenetisch (16S rRNA-basierend) und genomisch verwandt zu sein (Serrano *et a*l., 2009, Syst. Appl. Microbiol. 32, 1-7).

Die in der vorliegenden Arbeit untersuchten Typen-Spezies und Typenstämme beider Genera stammen entweder aus Stammsammlungen oder wurden von geographisch entfernten Standorten selbst isoliert. Mit den zu untersuchenden *Marichromatium*-Stämmen wurden sowohl Multi-Gen-Analysen mittels PCR-Amplifikationen durchgeführt, als auch Sequenzierungen Protein-codierender Gene wie *pufM, soxB, fusA, dnaK* und *gyrB*. Für die *Allochromatium*-Stämme wurden zusätzlich noch die *cbbL-1, mfd* und *groES* Gene untersucht.
Ferner wurde auch ein umfangreicher paarweiser Kreuzvergleich mittels DDH mit allen Stämmen innerhalb jedes Genus durchgeführt.

Die Stämme zu den Genera *Marichromatium* und *Allochromatium* zeigten eine 16S rRNA Gensequenz-Ähnlichkeit von 98 bis 100% (Durchschnitt 99,3%) bzw. 97,0 bis 100% (Durchschnitt 98%).
Der DDH-Kreuzvergleich, eingesetzt als Standardwerkzeug für die Definition einer Bakterienart, lieferte dagegen doppeldeutige Ergebnisse. Einerseits bestätigten die daraus erzielten Daten die 16S rRNA basierte phylogenetische Verwandtschaft, andererseits waren sie mit letzteren unvereinbar, was eine Spezies-Definition innerhalb beider Genera erschwerte.
Diese phylogenomische Nichtübereinstimmung konnte jedoch mittels MLSA korrigiert werden. Die Auswertungen der aneinanderhängenden Protein-codierenden Gensequenzen für *Marichromatium* (4331 bp) und *Allochromatium* (4427 bp) erlaubte jedoch eine Aufgliederung der Stämme innerhalb eines jeden Genus in mehrere unabhängige Gruppen oder Phylotypen.
Zwei neue RSB-Stämme, von denen jeweils einer zu einem der zuvor genannten Genera gehört, wurden von unterschiedlichen geographischen Standorten isoliert und mittels polyphasischem Ansatz vollständig charakterisiert. Der neue Stamm *M. gracile* Biotyp *thermosulfidiphilum*, der aus Sedimenten des deutschen nördlichen Wattenmeeres isoliert wurde, zeigte einige bemerkenswerte phenotypische Eigenschaften, die bisher noch nicht für das entsprechende Genus beschrieben

worden waren. Dazu gehörten u. a. Toleranzen gegenüber gemäßigt hohen Temperaturen (bis 44°C) und hohen Sulfidkonzentrationen (bis 16 mM) sowie das Vorhandensein eines Plasmids (Einzelheiten siehe Kapitel 2). Trotz solcher ungewöhnlichen Eigenschaften konnte der neue Stamm nicht als neue Spezies klassifiziert werden, da eine sehr hohe genetische Ähnlichkeit mit dem Typenstamm *M. gracile* DSM 203^T nachgewiesen werden konnte. Diese enge Verwandtschaft wurde auch mittels durchgeführter Multigen-Analysen unterstützt. Stamm SW26 konnte mittels MLSA dem Phylotyp I des Genus *Marichromatium* zugeordnet werden.

Der andere neue Stamm AX1YPE wurde aus marinen weichen Sedimenten vor der Peruanischen Küste isoliert und gehört zum Genus *Allochromatium*. Dieser neue Stamm zeigt ebenfalls besondere phenotypische Merkmale, die für dieses Genus noch nicht bekannt waren. So ist der Organismus in der Lage, viele Stickstoffverbindungen als Stickstoffquelle zu nutzen sowie bei schwachen Lichtintensitäten zu wachsen (15 - 60 µE m^{-2} s^{-1} PAR) (Einzelheiten siehe Kapitel 5). Aufgrund dieser phenotypischen Eigenschaften sowie der 16S rRNA phylogenetischen Analysen und der DDH Beziehung wird Stamm AX1YPE als neue Spezies im Genus *Allochromatium* beschrieben und als Name *A. humboldtianum* vorgeschlagen. Mittels MLSA konnte zusätzlich noch eine Separierung von Stamm AX1YPE als unabhängige Linie innerhalb des Genus *Allochromatium* bestätigt werden (siehe hierzu Kapitel 4).

Die im Rahmen der vorliegenden Arbeit durchgeführten Untersuchungen haben deutlich gezeigt, dass ein MLSA Ansatz als wertvolles Instrument genutzt werden kann, um nahe verwandte Stämme in den Genera *Marichromatium* und *Allochromatium* voneinander abzutrennen. Zusätzlich kann dieses molekulare Instrumentarium auch in den polyphasischen Ansatz aufgenommen werden, um eine Spezies-Abgrenzung in der Bakteriologie zu erreichen.

Chapter 1

Introduction

1. A general overview of bacterial systematics. Bacteria and Archaea constitute the main part of the existing biosphere on earth. Both are prokaryotes, which means that as a general rule they do not have a nucleus and there is no biological membrane separating the genetic material from the cytoplasm. However, within the domain bacteria the phyla *Planctomycetes* and *Poribacteria* are the exception, species of both phyla have a complex intercellular compartmentation (Fuerst, 2005). For example, the planctomycetes have their genetic material enclosed by an intracytoplasmatic membrane which is located in the internal region of the cell, forming the so-called nucleoid or pirellulosome in *Pirellula* species (Lindsay *et al.*, 2001; Fuerst, 2005). Although the term bacteria has been traditionally used to define all prokaryotes, nowadays is well accepted that prokaryotes evolved from a common ancestor forming two independent groups or domains (named Bacteria and Archaea). In the present review, the term bacteria will be applied indistinctly to define both prokaryotes, since the main part of this work deals with the study of two photosynthetic bacterial genera.

To date it is widely accepted that prokaryotes might represent far more than 50% of the Earth`s total biomass (Whitman *et al.*, 1998). These micro-organisms are found in many environments from temperate to extreme, they were also found at around 1600 meters below the seafloor (Roussel *et al.*, 2008) and also inhabiting on ocean crust (Santelli *et al.*, 2008). The ability to perform complex biochemical pathways allows these microorganisms to survive and proliferate in environments with a limited source of nutrients.

However, in spite of their importance as main regenerator in biogeochemical cycles, most of the existing bacterial species are still unknown. This could be a direct consequence of the particular simplicity of bacteria, they do not display a differentiated morphology and they are usually microscopic (exception to this rule are both giant sulfide oxidizers *Thioploca* and *Thiomargarita;* Schultz *et al*., 1996, 1999). Bacteria also are able to perform countless and still unknown biochemical reactions. Bacteria, as source of new active biocompounds are also intensively explored. On the other hand, genetic engineering has allowed scientists to produce medically relevant compounds and proteins from genetically modified bacteria. With this short introduction on bacterial capabilities, in the following chapters of the present work the issue of bacterial sytematics will be addresed in great detail.

The term systematics as main word in biology and related subjects was taken because of its meaning. According to Merrian-Webster's online English dictionary, the word sytematic derived from Latin systematicus and Greek systematikos – group of items forming a unified whole, an organized set of doctrines, armonious arrangement or pattern – has several meanings i) presented or formulated as a coherent body of ideas or principles ii) methodical in procedure and plan iii) related to, or concerned with classification (www.merriam-webster.com).

Systematics as a branch of the biological sciences, deals with the identification, classification, and comparative study of living organisms, their functions in nature and their evolutionary history. In biology the word "systematic" was originally used by Carl von Linné (1707-1778) in his "Systema naturae" which was the first natural system of classification created. Although, initially designed for classification of animals and plants, the Linnaean scheme is based on presumed homologies between compared organisms, summarizing different aspects such as morphology, physiology, and ecology. When more homologies are shared by two compared organisms they must be closer in terms of their evolutionary distance. Application of this philosophical scheme has resulted in the establishment of a hierarchical system of classification in which two main categories are considered: the top consisting in all living things (Domains, Kindoms) and the lower rank consisting in a single rank named species.

For prokaryotes this word is used as plural adjective e.g. systematics, and embrace all known ways to study microorganisms with the ultimate finality to circumvent the problem in biological classification of prokaryotes and arranging them in an ordered manner after Rosselló-Mora & Amann (1992).

Systematics consequently is an integrative approach in which several other disciplines e. g. taxonomy and ecology are also included. Since the organisms are adapting entities, the inclusion of ecology – the interaction of these organisms with their environment – was also necessary. However, the process of arranging organisms into categories or ranks according to their specific characteristics has necessarily to follow certain rules. These rules are given by taxonomy (theory of classification), a word derived from the Greek words, ***τάξις*** (taxis = order, arrangement) and ***νόμος*** (nomos = managing, law; nomia = method).

Differences between superior organisms (eukaryotes) and prokaryotes do not permit the application of biological species concept (BSC) to the latter. Comparing particular features, such as a multi-structured morphology, gametogenesis and meiosis, as well as embryogenesis, is not possible in prokaryotes. By this reason, traditionally, circunscription of bacterial species has been assesed using another type of information such as morphology, physiology, and biochemical data, which is strongly dependant on the ability of researchers to obtain bacteria in pure cultures for subsequent characterization.

Studies on genomic DNA reassociation started during the latest 1950s and continuing during the earliest 1960s. Intensive DNA-DNA renaturation experiments of genomic DNA of closely related *E. coli* strains (Doty *et al*., 1960) as well as of different strains of *Diplococcus* and *Streptococcus* (Marmur & Lane, 1960) were made. However, the well-known and standardized DDH method that is widely used nowadays was first reported by Britten & Kohne (1968). Thus, these studies left open the possibility to establish certain relationships between paired organisms based on their genome-wide comparison. Since that time, large collections of bacterial species, which were mainly defined by phenotypic criteria, were re-examined using the DDH approach. A general conclusion derived from those large comparisons was the fact that almost in all cases strains within the named species showed DDH homology

values at around 70% or more while strains from different species shared less than 70% (Johnson, 1973). Thus, DDH mired as indirect expression of whole genome identity has gradually replaced the insufficiently informative phenotypic and chemotaxonomic criteria to delineate bacterial species. Moreover, the level of 70% of homology soon had been commonly accepted as standard threshold for determining bacterial species (Wayne *et al*. 1987). Although DDH studies opened new insights in bacterial taxonomy, the method is time consuming, relies on pairwise comparison, and can only resolve taxonomic status at genus and species level (Staley, 2006). It has also been argued that results of DDH experiments represent an arbitrary number that does not contain any useful phylogenetic information. However, one should keep in mind that the purpose of a wide genomic comparison is to infer some relationships amog organisms. This point was critically evaluated and discussed in detail by Rosselló-Mora (2006).

Bacterial phylogeny was first investigated by comparing protein sequences such as ferredoxins and cytochromes (Fitch & Yasunobu, 1975; Schwartz & Dayhof, 1978). However, the analyses of the ribosomal rRNA molecules, which are essential in the process of protein biosynthesis within the cell, allowed a natural classification of all living organisms. The immediate consequence of using this new molecular tool is the fact that today's existing life can be divided in three lines or domains (*Archaea, Bacteria,* and *Eucarya*)*,* which are divergent from each other (Woese *et al*., 1987). This new molecular tool has also served as starting point for further development of fluorescently labelled oligonucleotides based-probes which allows the detection of target microorganisms directly under the microscope (DeLong *et al*., 1989; Amann *et al*., 1990 a, b). The use of molecular techniques allows the detection of microorganism in the absence of culturing and has demonstrated to be a valuable tool in studies of microbial diversity and ecology. Today, the 16S rRNA-based phylogenetic analysis is one important criterium widely used in modern bacterial systematics, enabling microbiologists to allocate new bacterial isolates to already existing taxon on the basis of their shared sequence similarity.

Modern bacterial systematics is an integrative approach in which different sorts of information (phenotypic, genotypic, and phylogentic) are merged together. This integrative taxonomy is also known as "polyphasic taxonomy". The term polyphasic

was mentioned the first time during the 1970s and was used for defining a consensus numerical approach in the taxonomy of Vibrios (Colwell, 1970). Table 1 summarizes all main characters used in polyphasic taxonomy according to Busse *et al*. (1996). Some of the listed characters are used in the present work for the identification of new bacterial isolates. In addition, some other parameters which are mainly focused on the taxonomy of anoxygenic phototrophic bacteria according to Imhoff & Caumette (2004) are also included.

Table 1: Categories and characters applied in bacterial systematics (from Busse *et al*., 1996). Characters sudied in the present work are shown in italics.

Categories	Examples
Cultural	*Colony morphology*
	Color of colonies
	Fruiting bodies
	Mycelia
Morphological	*Cell morphology*
	Cell size
	Motility
	Flagellation type
	Reserve material
	Gram stain
	Acid fast stain
Physiological	*Temperature range*
	pH range
	Salinity tolerance
Biochemical	*Carbon source utilization*
	Oxidation and fermentation of carbohydrate
	Enzyme profile
Inhibitory tests	Selective media
	Antibiotics
	Dyes
Serological	Agglutination
	Immunodiffusion
Chemotaxonomy	Fatty acid
	Polar lipid
	Micolic acid
	Lipopolysacharide composition
	Cell wall diaminoacid
	Cell wall aminoacid composition
	Whole cell sugar
	Cellular pigment
	Quinone system
	Polyamine content
	Whole cell protein (PAGE)
Genotypic	*DNA base ratio* (G + C content)
	Random amplified polymorphic DNA (RAPD)
	Restriction fragment length polimorphism (RFPL)
	Pulsed field gel electrophoresis (PFGE)
Phylogenetic	DNA probes
	DNA:DNA hybridization
	DNA:rRNA hybridization
	16S rRNA sequence
	23 rRNA sequence
	Sequence of the b-subunit of ATP-synthase
	GroEL(chaperonin) sequence

2. A multi-gene taxonomy in Prokaryotes

The prevalent division of life summarized in the dichotomy Prokaryote-Eukaryote (Stanier & Van Niel, 1962) was widely accepted by biologists but was strongly challenged recently with the introduction and analysis of ribosomal molecules. These analyses allowed the division of former prokaryotes in two new domains Archaea and Bacteria (Woese, 1987). In addition, the study of phylogeny of microorganisms was also possible by comparing the 16S rRNA gene sequences. However, over the time it also became clear that the use of the 16S rRNA gene as sole phylogenetic marker does not help to resolve the taxonomic status of bacteria below the genus level (Fox *et al.*, 1992; Rosello-Mora & Amann, 1992) conflicting in many cases with the phenotypic and genomic (DDH) criteria for delineation of species (Martinez-Murcia *et al.*, 1992; Jaspers & Overman, 2004; Badger *et al.*, 2005).

It was in the field of medical research that phenotypic differences within a named species have been taken as main criteria for practical diagnostic reason. For accurate differentiation of benign strains from pathogens, methods such as serotyping, monoclonal antibody typing, whole protein electrophoresis, and fingerprinting have been commonly tested. However, the introduction of multilocus enzyme electrophoresis (MLEE) - a molecular tool first developed to study genetic variation in natural population in eukaryotes - to routine diagnostic bacteriology, allowed to study genetic variation of pathogens at intraspecies level. Although, gel electrophoresis methods were also used early to separate bacterial strains (Norris & Burger, 1963), the technique was not applied extensively in bacterial taxonomy because of its poor resolution in comparison to the genomic DDH. MLEE is also an electrophoretic-based technique in which water-soluble enzymes (proteins) migrate differently during the electrophoresis. Such mobility is ultimately governed by the net electrostatic charge of amino acid sequences, thus variants of an enzyme (allozymes) corresponding to structural locus can easily be detected (Selander *et al.*, 1986). The applicability of multilocus enzyme electrophoresis to bacterial population genetics and systematics as an alternative to DDH was also first discussed by Selander *et al.*, (1986), however, this technique though more accurate and discriminative than DDH had also as main drawback that it is also based on electrophoretic banding pattern.

The idea of studying polymorphic variation by analysing nucleotide sequences was a reality in the 1990s when a first multilocus sequence typing (MLST) for pathogenic

bacteria was introduced by Maiden *et al*. (1998). Soon it became clear that the use of nucleotide sequences was superior than any other banding pattern-based technique. Perhaps the main advantage that nucleotide sequences could offer is the possibility that sequences can be stored in public data bases being easily accessible to the scientific community and suitable for comparison and further analysis.

Differences detected in the sequences at a single locus in closely related microorganisms (usually within species level) and their assignment as independent sequence typing (ST) of a gene is the fundamental idea behind the MLST. Each ST representing a unique allele is identified with an arbitrary number. In this way, different strains are defined on the basis of their allelic profile at all the examined loci (Feil, 2004). In other words, MLST represents a better version of MLEE in which banding patterns are replaced by nucleotide sequences. MLST has been used for discriminating populations of pathogens in which sequence variations may represent phenotypic variants within a known strain. Such variants might eventually be the causative agents of epidemic episodes. MLST was successfully applied for the first time concerning the study of both known human pathogens *Neisseria* and *Staphylococcus* (Maiden *et al.,* 1998; Enright *et al.,* 2000; Clarke *et al.,* 2001a).

However, this multi-gene approach also was found to be useful for interspecies comparison within a genus or even family (see Table 2). Such an extensive use of multi-gene analysis above species level has been denominated as Multilocus Sequence Analysis MLSA (Gevers *et al*., 2005) and is nowadays extensively applied in bacterial systematics as complementary tool for species delimitation. A basic premise of the MLSA is that a set of core genes (universal core genes, frozen genes, house-keeping genes) must be shared by all compared strains or species under scrutiny. Although it is a discussible issue whether the value or usefulness of the same set of genes are still informative above the species level (because the possibility of horizontal gene transfer (HGT) increases proportionally to the taxonomic level), the searching of a set of universal core genes which are applicable to all taxa is still an ultimate goal in biology. Thus, identification of a set of genes that are vertically inherited is essential in classification at least within a group or taxa, but even so homologous recombination in closely related strains, and potential HGT in far related species or genera can be seen as a form of noise that could eventually break and obscure any phylogenetic signal in the reconstruction of evolution of a group of organisms (Bapteste *et al*., 2003).

Table 2. Techniques used presently for the classification of Bacteria and Archaea (their utility for description of a level in the taxonomic hierarchy is noted by +, - or ±). Taken from Staley, 2006.

Present taxon level	Phenotypic properties	16S rRNA sequences	DNA-DNA hybridization	Genomics	MLSA
Domain	+	+	-	+	+
Kindom	+	+	-	+	+
Phylum	+	+	-	+	+
Class	+	+	-	+	+
Order	+	+	-	+	+
Family	+	+	-	+	+
Genus	+	±	±	+	+
Species	+	-	±	+	+
Subspecies	±	-	-	+	+

This fact was clearly exemplified in the MLSA of the vibrios. Historically, the number of housekeeping genes that have been used in phylogenetic studies and identification of *Vibrio* species totalize 14 loci (Thompson *et al*., 2004). These genes were used in different combinations and variable numbers by many research groups around the world. Among these analysed genes, the *recA*, *rpoA,* and *pyrH* genes were commonly reported as reliable alternative phylogenetic markers for almost all cultured vibrios, however, further studies also revealed that even these genes were not suitable for differentiating some *Vibrio* species, e. g. *Vibrio neptunis* from *V. coralliilyticus* and *V. anguilarum* from *V. ordallii* (Thompson, 2005). In light of these results, it is clear that no single gene among species within a genus is completely resistant to HGT, paralogy (homologous gene within a single species that diverged by gene duplication), or differences in their evolutionary clock as was pointed out by Thompson *et al*. (2004).

The amount of protein-coding genes that are necessary in MLSA should also be taken into consideration. The most recommended number of gene loci to be used in MLSA is seven as found in literature. Certainly, this number is directly related to the number of loci used in MLEE in *Neisseria meningitidis* for which the analysis of seven loci has provided a good level of resolution (Maiden *et al*. 1998; Feil, 2004). Searching for a minimal number of genes capable to produce an accurate phylogenetic tree has also been investigated; phylogeny based on selection of three “best genes” was found reliable in *Escherichia coli* groups and also reproducible in *Salmonella*, *Shewanella,* and *Burkholderia* (Konstantinidis *et al*., 2006). However, recently it was also reported that 12 protein-coding genes are necessary to reach a stable concatenated phylogenetic tree (>90% bootstrap support) in the extremely halophilic bacterium *Salinibacter rubber* (Soria-Carrasco *et al*., 2007). It seems that a

minimal number of gene loci is directly linked with the taxonomic level and population size of the target bacterial species. However, as a matter of fact, a common agreement among bacteriologists is that the concatenation of several protein-coding genes, ideally seven, could help to minimize the fuzziness observed in phylogenetic tree reconstruction based on a single gene.

3. Photosynthetic bacteria

Since the main part of the present work deals with the description of new species of anoxygenic phototrophic purple sulfur bacteria, some important features of this group of microorganisms are described.

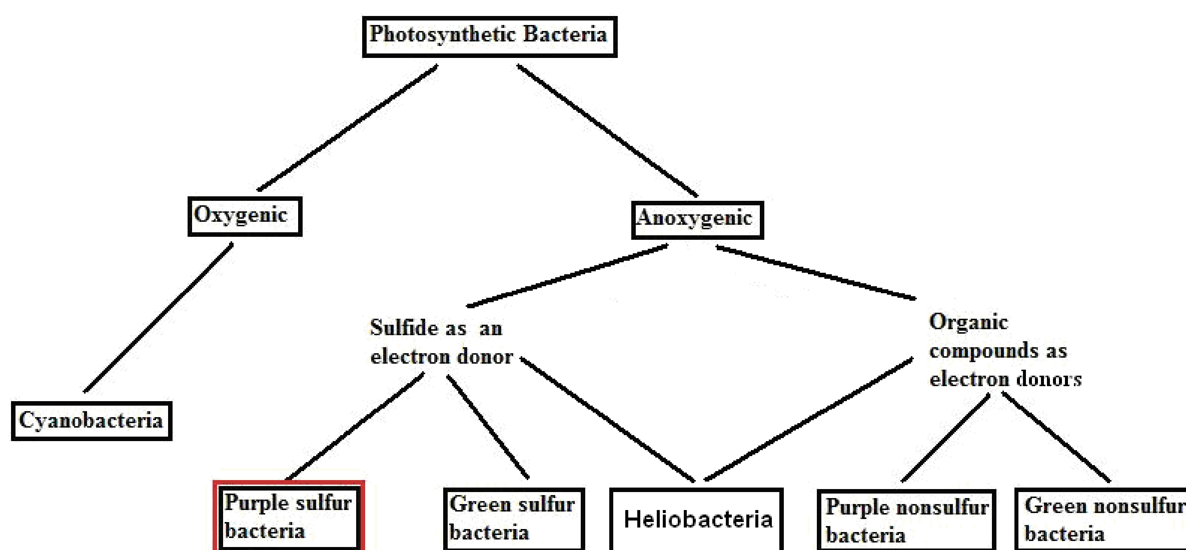


Figure 1. Schematic representation of the main groups of photosynthetic bacteria present today. Only the anaerobic anoxygenic phototrophs are represented in this figure. Red box represents the group studied in the present work. (Purple nonsulfur and green nonsulfur bacteria are able to tolerate and utilize low concentrations of sulfide as an electron donor in photoautotrophic growth).

The ability to produce organic molecules by photosynthesis (process in which electromagnetic radiation is converted to chemical energy) is widely extended in nature. Basically, the photosynthesis involves two main steps; i) transformation of light energy (photons) in a gradient of energy (proton motive force) and ii) fixation of CO_2 through different mechanisms such as Calvin cycle, reverse citric acid cycle, and 3-hydroxypropionate cycle.

Because both key enzymes of the Calvin cycle e.g. ribulose bisphosphate, carboxylase and phosphoribulokinase are widely present in nature, this type of CO_2 fixation is present in almost all autotrophs from green plants, algae, cyanobacteria,

purple bacteria, as well as in chemolitothophic bacteria and archaea (Madigan & Martinko, 2006). However, photosynthetic bacteria are able to use two alternative pathways to fix carbon dioxide. The reverse tricarboxylic acid (rTCA) pathway can be performed by *Chlorobium* species from the green sulfur bacteria. The whole rTCA uses four ATP molecules for generating one molecule of oxalacetate from four molecules of CO_2 (Fuchs, 1989). Since the rTCA is an energetically unfavorable reaction, this mechanism requires the participacion of key enzymes such as the ATP dependent citrate lyase, which catalyzes the cleavage of citrate into acetyl-CoA and oxalacetate using CoA and ATP, as well as the carbon dioxide-fixing enzymes 2-oxoglutarate synthase (ferredoxin dependent), NAD(P)H dependent isocitrate dehydrogenase, and the pyruvate ferredoxin oxidoreductase (Antranikian *et al*, 1982; Fuchs, 1989; Fuchs, 2007). Although originally the rTCA cycle was described in the photosynthetic *Chlorobium thiosulfatophilum* (Evans *et al.*, 1966), this mechanism was also reported in some delta proteobacteria (Schauder *et al*, 1987) as well as in thermophilic bacteria and archaea (Beh *et al.*, 1993). Relatively recently the rTCA pathway was also reported for the epsilon proteobacteria isolated from hydrothermal vents (Campbell & Cary, 2004), reinforcing the theory that the rTCA was the first autotrophic CO_2 fixation mechanism evolved on Earth (Wächterhauser, 1990).

The second alternative CO_2 fixation is the so-called hydroxypropionate pathway because one special intermediate product of the CO_2 reduction to glyoxylate is the 3-hydroxypropionate (Madigan & Martinko, 2006). Altough this pathway was originally described and confirmed in the photosynthetic green nonsulfur bacterium *Chloroflexus aurantiacus* (Holo, 1989), an unknown variant of the 3-hydroxypropionate cycle was also reported for the thermoacidophiliautotrophic archaea *Metallosphaera sedula*, and other species of *Sulfulobus* and *Archaeoglobus* genera (Berg *at al.*, 2007).

During the photo-assimilation, when oxygen is released, the process is known as oxygenic photosynthesis and usually is carried out by eukaryotic organisms (green plants and algae) as well as some prokaryotes, such as cyanobacteria (Fig. 1). When no oxygen is formed during photo-assimilation, the process is called anoxygenic photosynthesis and is usually performed by purple and green photosynthetic bacteria (Fig. 1). The basic differences between the two processes involve the use of a

different electron-transport machinery. The flux of electrons in the anoxygenic photosynthesis is cyclic. Briefly, this process is initialized by energy light photons exiting bacteriochlorophyll molecules in the reaction center which in its turn emit electrons. These electrons are taken up by the primary electron acceptor (bacteriophaeophytin) and are passed to other electron carriers through the membrane creating a proton motive force necessary for the ATP formation (Fig. 2). This process is also known as cyclic photophosphorylation.

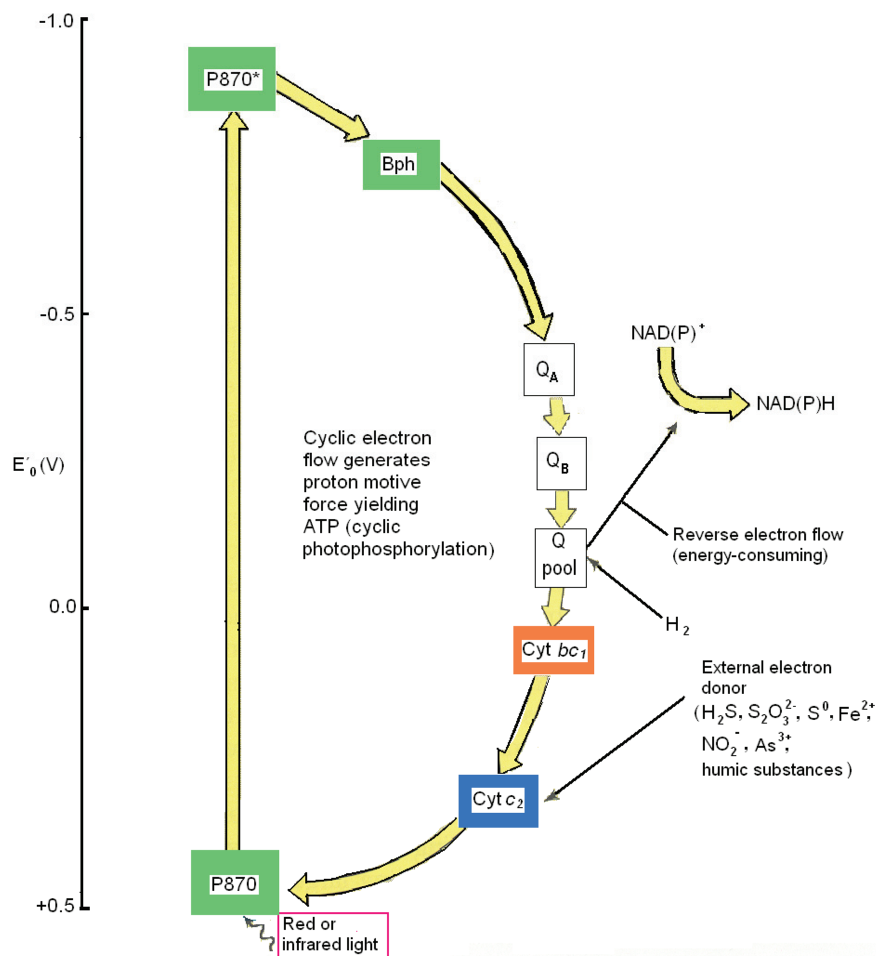


Figure 2. General scheme of electron flow in anoxygenic photosynthesis in a purple bacterium. Adopted and modified from Madigan & Martinko (2006).

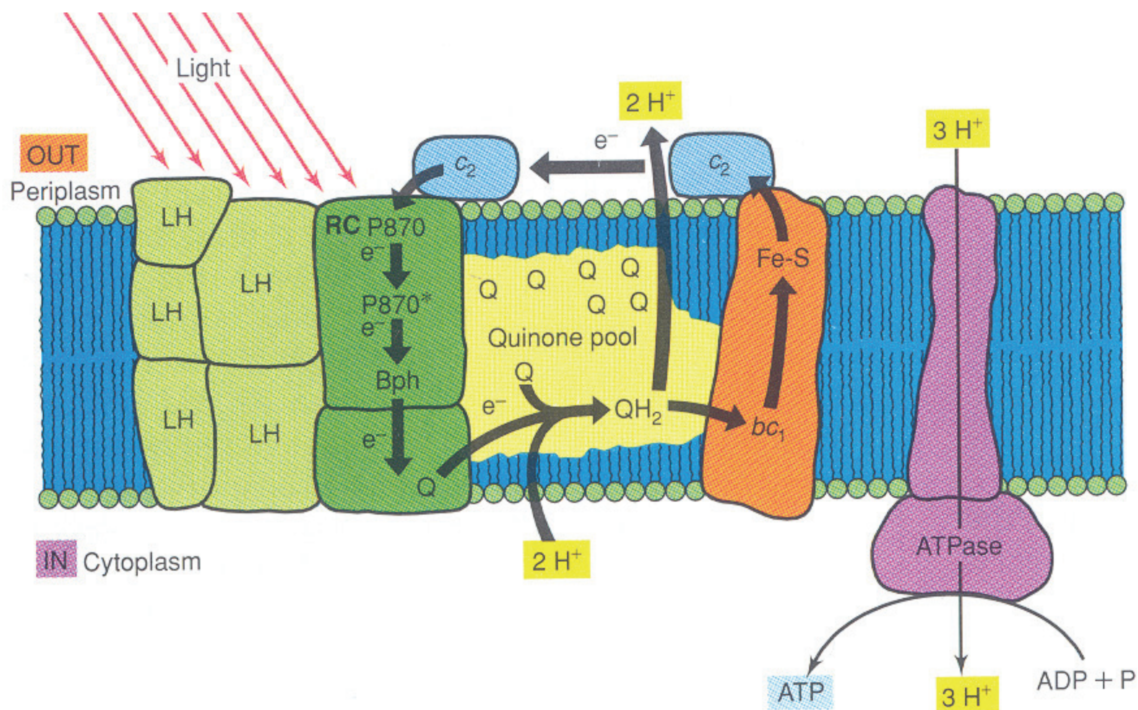


Figure 3. Arrangement of protein complexes in the photosynthetic membrane of a purple phototrophic bacterium. The light, generated electron and proton flux, are indicated by arrows. Adopted from Madigan & Martinko (2006).

The final step of this process is the return of these electrons to the bacteriochlorophyll molecule again enabling them to start a new cycle. Since during the whole process electrons only travel in a closed route, there is not a net input and loss of electrons (Fig. 2+3). It has been proposed that photosyntesis had evolved very early in Earth's history. Molecular analysis indicated that photosytems in the anoxygenic purple photosynthetic bacteria are the most ancient known (Blankenship, 2001).

Anoxygenic photosynthetic bacteria typically use sulfide or other reduced sulfur compounds as primary electron donors for the reduction of CO_2 to organic matter; however, these organisms also are able to utilize iron II (Widdel *et al.*,1993), nitrite (Griffin *et al.*, 2007), and arsenite (arsenic III) (Budinoff & Hollibaugh, 2008; Kulp *et al.*, 2008) as electron donors. More recently it has been reported also that one anoxygenic phototroph of the genus *Thiocystis* is able to use humic substances as electron donor (Tran, 2008).

Nowadays the existence of six phylogenetic groups of anoxygenic phototrophs is known. The Purple Sulfur Bacteria, Purple Nonsulfur Bacteria, Green Sulfur Bacteria (Imhoff, 1995), the Green Nonsulfur Bacteria (Pierson & Castenholz, 1995), and the Heliobacteria (Madigan & Ormerod, 1995) can perform anoxygenic photosynthesis at

strictly anaerobic conditions. The sixth group of photosynthetic prokaryotes is able to carry out anoxygenic photosynthesis only under aerobic conditions, the so-called aerobic anoxygenic phototrophs (AAPs) Yurkov & Beatty (1998). Phototrophic bacteria can utilize different strategies for harvesting light energy, which include the presence of the photosynthetic pigment bacteriochlorophyll (Bchl) in the antenna complex and in a reaction center (RC) in which the presence of carotenoid pigments play important functions preventing cellular damage caused by photo-oxidation (Pfennig & Trüper, 1992). Phototrophic purple bacteria, so far represent the most diverse anoxygenic phototrophic prokaryotes with representatives belonging to the α, β, and γ subclasses of the proteobacteria. As main feature, the purple bacteria have Bchls *a* or *b* as main pigments*;* while the green photosynthetic bacteria show a wide variety of Bchl types with Bchl *a* located mainly in the reaction center and *c, d*, or *e* as antennae located in chlorosomes. The Heliobacteria which are the unique representative of Gram-positive phototrophic bacteria show as main photosynthetic pigment Bchl *g* (Madigan & Ormerod, 1995; Madigan, 2000). On the other hand, the AAPs show as main pigment Bchl *a* which is synthethized in the presence of oxygen (Shiba *et al*., 1979; Yurkov & Beatty, 1988). Although all first reported AAPs belong to the α proteobacteria such as the cultivated *Erythrobacter* spp. as main representatives, it has been reported also that AAPs could be more diverse with representative strains belonging to the β and γ subclasses of the proteobacteria (Beja *et al*, 2002).

The process of photosynthesis can be subdivided in two reactions: the **light reaction** which requires light energy to operate and the **dark reaction** in which the chemical energy accumulated during the light reaction is used to reduce CO_2 to organic compounds. Bacteriochlorophylls are the primary light harvesting pigments in photosynthetic bacteria, however, in order to enhance the ability of capturing light energy more efficiently, bacteriochlorophylls are organized in association with proteins and carotenoids forming the antenna complex and reacion center (RC).

3.1. **Antenna complex**. The principal element for harvesting light energy is a membrane-associated pigment-protein complex, which is known as antenna or light harvesting (LH) complex. This complex was first characterized in purple sulfur bacteria and is formed by polypeptides of α and β chain which are part of an integral membrane protein associated with photopigments Bchl *a* or *b* and carotenoids

(Blankenship *et al*., 1995). The most important function of the antennae complex is to serve as an intermediate in the transfer of the harvested photon energy to the reaction center. As a consequence, absorption of the light spectrum is highly dependant on the amount and type of the photopigments present in this complex.
When a closely association between the antenna complex and reaction center occurs, it is known as "core complex" and its absorption spectrum usually has a wavelegth range between 870 and 1020 nm (Ketelaars *et al*., 2002; Ma *et al*., 2008). Conversely, when a lack association occurs and an LH has a variable distribution, this feature is known as "peripheral antenna" complex, and their absorption spectrum wavelength range falls between 800 and 860 nm (Hess *et al*., 1993, Fujii *et al*., 2007).
Endospore-forming Gram-positive Heliobacteria are anoxygenic phototrophs, which contain Bchl *g* as sole pigment and they possess a small size antenna, directly associated to the RC in the cytoplasmatic membrane (Madigan & Ormerod, 1995). These organisms do neither contain chlorosomes nor any other intracytoplasmatic membrane system (Madigan & Ormerod, 1995). The absorbance spectrum of Bchl *g* is unique and allows Heliobacteria to harvest near-ultraviolet and near-infrared light, growing best at high light intensites in its terrestrial environment (Gest & Favinger, 1983). Althoug Heliobacteria were originally isolated from garden soil on the campus of Indiana University, many species were subsequently isolated from rice field soils and even hot springs (Madigan & Ormerod, 1995; Heinnickel & Golbeck, 2007). Heliobacteria are able to fix N_2, but conditions of diazotrophy can vary among species. They are able to use organic and inorganic (hydrogen and sulfide) electron donors. No heliobacteria has been shown to grow autotrophically by the use of sulfide as reductant in the assimilation of organc substrates (Madigan & Ormerod, 1995; Heinnickel & Golbeck, 2007).
In nature, the most effective light-harvesting structure is the so-called chlorosome, which also contains Bchl. However, differently to the other light-harvesting complex above mentioned, no protein is present in in the chlorosome assemblages (Madigan & Martinko, 2006). So far, these structures have been reported only in green sulfur bacteria and green filamentous bacteria, allowing these organisms to live at very low light conditions (Beatty *et a*l., 2005; Manske *et al*., 2005).

3.2. **Reaction center**. Reaction center is defined as a complex structure which contains several membrane proteins designated as: H (heavy), M (medium), and L

(light) sub-units according to their molecular size. A substantial part of the RC complex is constituted by pigment molecules (Bchl, bacteriophaeophytins, and carotenoids) and cofactors such as quinones. It is well established that subunits M and L bind to the pigment molecules and cofactors, while subunit H helps in the stabilization of the reaction center complex (Deisenhofer *et al.*, 1985, Deisenhofer & Michel, 1989).

The reaction centers are integrated into the photosynthetic membrane and are classified as RC I and RC II according to the identity of the final electron acceptor (Allen & Williams, 1998). RC I has as final electron acceptor an iron sulfur cluster (Fe-S), whereas the pheophytin-quinone is the terminal electron acceptor for RC II (Heathcoat et al., 2002). As a general feature, the “purple” photosynthetic proteobacteria have a particular electron transport chain which is composed of a reaction center (RCII), pool quinones, a cytochrome bc_1 complex, and a soluble or membrane-bound electron carrier (cytochromes c_2) (Madigan & Martinko, 2006, Fig 2). Although they also contain a non-heme iron sulfur cluster, the involvement of this cluster as a final electron acceptor in the electron transfer (ET) chain is not clear (Remy & Gervert, 2003).

4. **The Family *Chromatiaceae*.**

The family *Chromatiaceae* comprises a wide range of genera and species closely related. This bacterial group is characterized by its ability to carry out anoxygenic photosynthesis and the use of sulfide (H_2S) as primary electron donor (Pfennig and Trüper, 1992). A recent reclassification of the purple sulfur bacteria include the two families ***Chromatiaceae*** and ***Ectothiorhodospiraceae*** as separate phylogenetic lines of the γ-proteobacteria (Imhoff *et al.*, 1998)

The purple sulfur bacteria grow anaerobically in the presence of light, use H_2S as electron donor, and they are able to fix carbon by photosynthesis. A common feature of the *Chromatiaceae* is the presence of *bacteriochlorophyll* (Bchl) *a or b* (pigment antenna and RC) which are membrane protein-associated. Bchls harvest light energy and transfer electrons to the ETS (electron transfer system) by generating a proton gradient. They use light energy for creating a reducing power (ATP and NADPH) through a cyclic photophosphorylation. Purple sulfur bacteria have also characteristic carotenoid pigments such as okenones, spirilloxanthins, lycopene, and

rhodopin which play an important role preventing antenna pigment photo-oxidation (Pfennig & Trüper, 1992).

Purple sulfur bacteria are found in nearly all aquatic environments both marine and fresh-water, where light reaches anoxic H_2S rich zones. They are predominant in stratified fresh water meromictic lakes, coastal lagoons, and marine sediments, as well as in sewage poultry and cattle lagoon plant treatments (Wenke & Vogt, 1981).

4.1.Environmental factors

The environmental factors are discussed briefly: light intensity, anaerobic conditions, sulfide concentration, temperature, and pH have been reported to have an enormous influence on growth behaviour of purple sulfur bacteria (Pfennig & Trüper, 1992).

In the environment, light is first scattered by water. Light can penetrate up to 200 m into clear water but in most lakes, where dissolved and particulate organic matter are abundant, light penetrates only at the upper 10 m. In wastewater treatment lagoons where the amounts of dissolved and suspended solids are usually high, light penetrates even less. It has been observed in anoxic animal waste treatment lagoons that purple sulfur bacteria can be isolated at a maximum depth of 2.5 m (Wenke & Vogt, 1981). Meromictic lakes (permanently stratified) seem to be the most favorable environment for purple sulfur bacteria development. If sulfate is present in sediments to support sulfate reduction, the resulting sulfide production can eventually support growth of purple sulfur bacteria. In such lakes, *Chromatiaceae* were isolated in the chemoclines at different depths (Overmann *et al.*, 1991; Lüthy *et al., 2*000; Peduzzi *et al.*, 2003). This feature is directly correlated with the fact that light intensity reaches the anoxic H_2S rich zone, which means that at depths where light can penetrate, the purple sulfur bacteria can grow anaerobically and are able to use H_2S as an electron donor for photosynthesis. Then, sulfide is oxidized to sulfur or sulfate, being in this particular case both the light intensity and sulfide concentrations the most important factors (Overmann, 2001).

Purple sulfur bacteria are able to use part of the infrared region, 800-900 nm wavelength through Bchls. Bchl *a* absorb light below 880 nm while Bchl *b* shows a

maximum of 1030 nm. It has been reported also that carotenoids in PSB have also an important light harvesting function which is transferred to the Bchls (Overmann, 2001). Considering that the infrared region of visible light is strongly absorbed by water in the top few centimeters (0 - 50 cm) while in sediments this happens in the first milimeters, PSB are only significant in shallow ponds and estuarine environments. The green sulfur bacteria (GSB) as a possible competitor have also a great affinity for H_2S but their light absorption maximum lies between 705-750 nm (Pfennig & Trüper,1992).

A pH between 7.2 to 7.4 is convenient for the enrichment of *Chromatiaceae.* Some species of the purple sulfur bacteria such as *Thiocapsa roseopersicina* and *Allochromatium vinosum* are able to grow in a lithotrophic fashion (Friedrich *et al*., 2001). On the other hand, other species of the PSB such as *Marichromatium gracile* are able to survive under micro-aereophilic conditions, moving to the micro-oxic regions during the night (Thar & Kühl, 2001).

4.2. Importance

The most studied group of purple sulfur bacteria belongs to the genera *Chromatium, Thiocapsa, Thiospirillum,* and *Thiopedia* within the family *Chromatiaceae* and usually they are isolated from waste treatment lagoons. The most known species are *Allochromatium vinosum* (formerly *Chromatium vinosum*)*, Thiocapsa roseopersicina, Thiocapsa rosea* (formerly *Thiopedia rosea*), and the marine species *Marichromatium purpuratum* (formerly *Chromatium purpuratum*) (Imhoff *et al*., 1998).

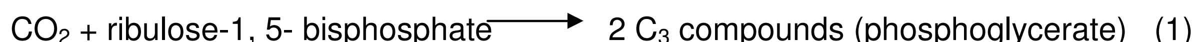
The ability of purple sulfur bacteria to oxidize H_2S can be utilized potentially for controlling odor in waste poultry lagoons. Bacterial biomass production can be a potential source of proteins. Cell protein of the purple sulfur bacteria could be used as by-product that can be included in the diet for poultry and fish (Sasikala & Ramana, 1995; Ponsano *et al*., 2002).

On the other hand, it has been reported that *Marichromatium purpuratum* is able to produce antimicrobial compounds with a broad spectrum activity against filamentous fungi, yeast as well as Gram-negative and Gram-positive bacteria (Burgess *et al*., 1991). *Marichromatium purpuratum* is also a subject of intensive studies concerning characterization of its photosynthetic system and soluble electron transfer proteins

(Kerfeld *et al.*, 1996). Isolation of the metalloenzyme hydrogenase, which catalyze the uptake and production of molecular hydrogen, is also important for biotechnological applications. A stable hydrogenase activity has been reported for *Thiocapsa roseopersicina* (Gotob, 1986; Rakhely *et al.*, 1999).

4.3. Carbon fixation and sulfur oxidation by purple sulfur bacteria

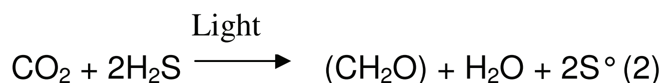
The biological carbon fixation is responsible for the conversion of CO_2 into living biomass. CO_2 fixation via the Calvin cycle resulted in the formation of the C_3 compound phosphoglycerate (equation 1).



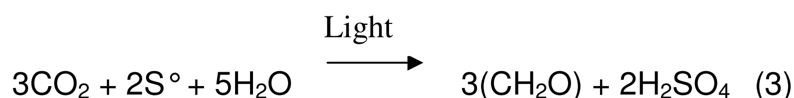
$$CO_2 + \text{ribulose-1, 5- bisphosphate} \longrightarrow 2\ C_3\ \text{compounds (phosphoglycerate)} \quad (1)$$

It has been established for purple sulfur bacteria that the light–dependent conversion of 1 mol of sulfide to 1 mol of sulfate is coupled to the fixation of 2 moles of carbon dioxide (Wolfe, 1999). Photosynthetic carbon dioxide fixation and sulfide or thiosulfate oxidation are stoichiometrically related as can be seen from the equations below (2-7).

For sulfide:

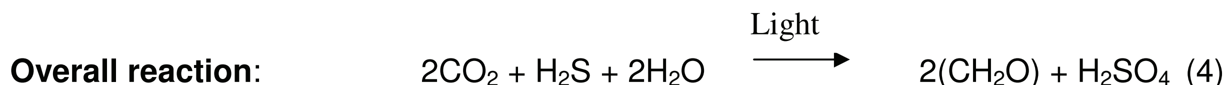


$$CO_2 + 2H_2S \xrightarrow{\text{Light}} (CH_2O) + H_2O + 2S° \quad (2)$$



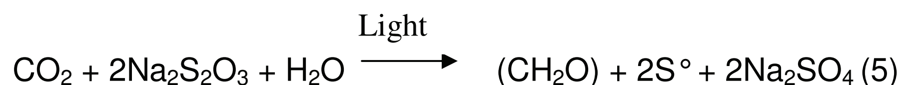
$$3CO_2 + 2S° + 5H_2O \xrightarrow{\text{Light}} 3(CH_2O) + 2H_2SO_4 \quad (3)$$

Overall reaction:

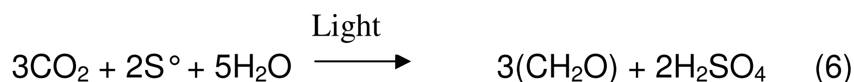


$$2CO_2 + H_2S + 2H_2O \xrightarrow{\text{Light}} 2(CH_2O) + H_2SO_4 \quad (4)$$

For thiosulfate:

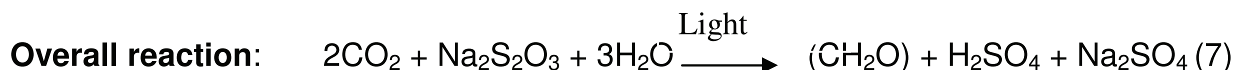


$$CO_2 + 2Na_2S_2O_3 + H_2O \xrightarrow{\text{Light}} (CH_2O) + 2S° + 2Na_2SO_4 \quad (5)$$



$$3CO_2 + 2S° + 5H_2O \xrightarrow{\text{Light}} 3(CH_2O) + 2H_2SO_4 \quad (6)$$

Overall reaction:



$$2CO_2 + Na_2S_2O_3 + 3H_2O \xrightarrow{\text{Light}} (CH_2O) + H_2SO_4 + Na_2SO_4 \quad (7)$$

5. Material and methods

In this chapter, only the methodological details which are not included within the manuscripts are enumerated.

5.1. Source of microorganisms. Bacterial strains used in this study are listed in Table 1. Type strains of both genera *Allochromatium* and *Marichromatium* were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ). Type strains *Marichomatium indicum* JA100^T and *M. bheemlicum* JA124^T were kindly donated by Dr. Ch. Sasikala from Jawaharlan Nehru Technological University, India. In addition, new isolates from sediment samples from various regions around the world were also included (see Table 3 and Fig. 4).

Table 3. Locations and geographical coordinates of isolated *Allochromatium* and *Marichromatium* strains used in this study.

Locations	Strain/species name	Latitude	Longitude
NR	*Allochromatium vinosum* DSM 180^T *	NR	NR
NR	*Allochromatium minutissimum* DSM 1376^T *	NR	NR
America [a]	*Allochromatium sp.* strain AX1YPE	12°02'S	72°13'W
America [b]	*Allochromatium sp.* strain BH-2*	23°58'N	77°42'W
Asia [c]	*Allochromatium renukae* DSM 18713^T	16°54'N	82°14'E
NR	*Allochromatium vinosum* DSM 183*	NR	NR
NR	*Allochromatium vinosum* DSM 1686*	NR	NR
Europe [d]	*Marichromatium gracile* DSM 203^T*	NR	NR
Europe [e]	*Marichromatium purpuratum* DSM 1591^T*	NR	NR
Europe [f]	*Marichromatium purpuratum* DSM 1711*	NR	NR
Europe [g]	*M. gracile* biotype *thermosulfidiphilum* DSM 21765	55°02'N	08°25'E
Europe [h]	*Marichromatium sp.* strain CLA37	45°41'N	13°07'E
NR [i]	*Marichromatium sp.* WS-2008*	NR	NR
Asia [j]	*Marichromatium indicum* JA100^T	NR	NR
Asia [k]	*Marichromatium bheemlicum* JA124^T	17°51'N	83°25'E
America [l]	*Marichromatium sp.* strain MP1	03°29'S	80°23'W

[a] Originally isolated from soft marine sediments, Callao, Peru. [b] Originally isolated from South Andros Black Hole, Bahamas. [c] Originally isolated from a brackish water pond, Kakinada India.
[d] Originally isolated from streams around Cracow, Poland (Strzeszewski, 1913). [e] Originally isolated near Split harbour, Adriatic sea, Croatia (Imhoff & Trüper, 1980). [f] Unknown origin. [g] Strain isolated from sediment samples from the North Wadden Sea, Germany (Serrano *et al.*, 2009). [h] Strain isolated from sediment samples from Marano Lake, northern Italy, this strain was kindly donated by Ms. Clelia Dona, Department of Marine Microbiology, University of Bremen. [i] Unknown origin, DSMZ deposited as *Allochromatium vinosum* DSM 185 = NCBI data base 16S rRNA sequence deposited as *Allochromatium sp.* strain WS-2008. [j] Originally isolated from mangrove sediments, nearby Goa, India. [k] Aquaculture pond, India. [l] Original isolate from mangrove sediments northern coast of Peru.
NR= not reported
* Type strains obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ).

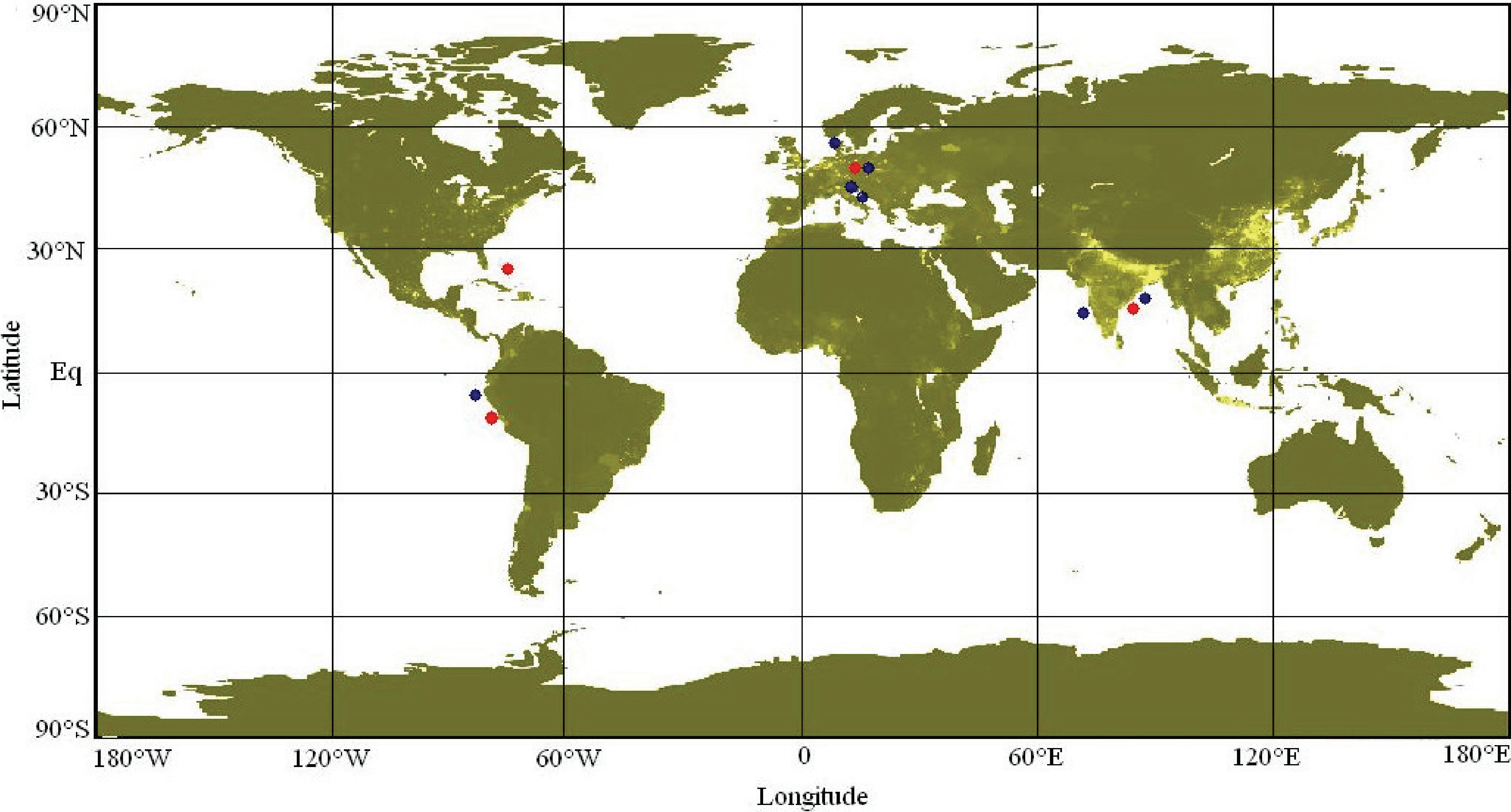


Figure 4. Map of the geographical distribution of photosynthetic purple sulfur bacteria examined in this study for which geographical coordinates or reference locations were available. Blue spots represent *Marichromatium* strains and red spots represent *Allochromatium* strains.

5.2. Media and cultivation conditions. The basic medium for the initial enrichment of marine strains was prepared according to Widdel & Bak (1992) and contained (per litre): 26.37 g NaCl, 5.67 g $MgCl_2 \times 6H_2O$, 6.8 g $MgSO_4 \times 7H_2O$, 1.47 g $CaCl_2 \times 2H_2O$, and 0.72 g KCl. After autoclaving and cooling, a Widdel flask was used to create oxygen free N_2-atmosphere and the basic medium was supplemented with the following sterile stock solutions (per litre): 10 ml KH_2PO_4 (50 g/l), 10 ml NH_4Cl (50 g/l), 30 ml $NaHCO_3$ (1M, CO_2 equilibrated), trace elements (2 ml), and 1 ml each of selenite-tungstate (1 ml). Seven vitamins solution was prepared according to Widdel & Pfennig (1981) and contains per 100 ml: NaH_2PO_4 3 mM, Na_2HPO_4 7 mM, 4-aminobenzoic acid 4 mg, D-(+)-biotin 2 mg, Nicotinic acid (niacin) 10 mg, Ca-D-(+)-pantothenate 5 mg, pyridoxine hydrochloride 15 mg, folic acid 4mg, lipoic acid 2 mg, pH ~ 7.7, vitamin B_{12} (5 mg/100 ml), riboflavin (5 mg/100 ml), and thiamine (5 mg/100 ml). After adding 1M Na_2S (2 ml), the pH of the medium was checked and adjusted to 7.3 using sterile 2M NaOH. All the time, the medium was kept under an N_2 and CO_2 (90:10) atmosphere. In order to control anoxic conditions inside the Widdel flask, 1 ml of a sterile resazurin solution (1 g/l) was also added. The medium thus prepared was distributed into 50 ml screw-cap bottles avoiding presence of air bubbles inside. To cultivate an inoculum, 10 ml of the medium was transferred into 15 ml Hungate culture tubes and saturated with a mixture of N_2/CO_2 (90:10).

During the course of this study, methodical modifications were made. Basically the anoxic medium was prepared according to DSMZ medium 28 (Pfennig's medium I, modified, 1988, for purple sulfur bacteria, see Appendix 1).

Modifications were made regarding preparation and use of solution F of medium 28:

- $Na_2S \times 9H_2O$ crystals were washed in distilled oxygen free water (saturated with N_2). After drying the crystals on towel paper, they were carefully weighted. A stock solution (1M sulfide) was prepared according to the method of Cline *et al.* (1965).
- This non-neutralized sulfide stock solution was basic and was used in experiments of sulfide tolerance (see chapter 2) and for the preparation of anoxic Pfennig's medium I.

Sulfide tolerance experiments

- Experiments of sulfide tolerance were carried out using a set of 15 ml Hungate tubes and were conducted under nitrogene atmosphere. Different sulfide concentrations were calculated based on the standardized and titrated Na_2S x $9H_2O$ stock solution (1M). For this purpose, Pfennig's medium, containing all supplements as indicated in Medium 28 (except sulfide) and saturated with N_2 and CO_2 in a Widdel flask, was dispensed into 250 ml bottles, sealed with rubber and screw caps to avoid air bubbles inside.
- From these 250 ml bottles, 15 ml of the medium were then dispensed into Hungate tubes. Different concentrations of sulfide (0.5 to 24mM) were prepared using the titrated sulfide stock solution (1M) as described above.
- Since the preparation of the medium is based on a non-neutralized sulfide solution, it had to be neutralized by adding various amounts of concentrated sulfuric acid. It is important to consider that the process of adding the titrated sulfide stock solution followed by neutralization should be performed as fast as possible to avoid loss of sulfide as gas in a neutralized solution. The amounts of sulfuric acid necessary to neutralize the alkaline sulfidic solution should be calculated before running the experiment. The final pH for each desired sulfide concentration must be checked carefully by using a standard stock of reagents.

Pfennig's medium I preparation (modified)

- Once all ingredients of Pfennig's I DSMZ medium 28 are incorporated into the Widdel flask in a saturated N_2 atmosphere, some amounts of titrated non-neutralized sulfide stock solution can be added (final sulfide concentration will depend on the strains requirement and the final volume of medium prepared). Usually a final sulfide concentration of 4 mM is described in the literature for PSB.
- Immediately after adding the non-neutralized sulfide, the medium will become reddish and turbid indicating alkalinity (Fig. 5, left panel). Now, change the gas phase to CO_2 and and let it remain in flux until the medium turns from reddish to clear and transparent (Fig. 5, right panel). The medium will now become acidic. After switching once again to an N_2 gas phase, let the solution flux for ~15 minutes. The medium will continue to be transparent but the pH will

change to ~7.0 Using this method, anoxic Pfennig's medium can be prepared faster and it is not necessary to add any other reagent e.g. resazurin as oxygen indicator, NaOH or HCl for adjusting the final pH. Additionally, this procedure is ideal when a gas mixture of N_2/CO_2 (90/10) is not available.

Figure 5. Photographs of Pfennig's medium immediately after addition of the sulfide stock solution into a Widdel flask with nitrogen gas flux (left panel). Pfennig's medium after changing the gas phase to CO_2 (right panel).

5.3. Enrichment and Isolation. For the enrichment of *Chromatiaceae,* the basic medium was supplemented with thiosulfate (5 mM) and acetate (2 mM) as recommended by Pfennig & Trüper (1992). First enrichments of new isolates were performed at room temperature (25ºC) and illuminated by a 40 W tungsten lamp placed at a distance of 30 cm or day light. From a well developed red turbid culture, several liquid dilution series from one to ten were set up in order to obtain pure cultures. This procedure was repeated several times (4), every time starting a new dilution series from the previous lower dilution series, where growth was detected. Purity of the culture was checked continuously under a phase-contrast microscopy. Additionally, deep agar dilution series (1 to 10) were also prepared according to Pfennig & Trüper (1992).

Further physiological experiments were performed in an incubator at a constant temperature of 30°C under continuous illumination of 50 to 100 µE m^{-2} s^{-1} photon flux. Experiments of thermal and sulfide tolerance were performed in a water bath at

40 to 45°C. The temperature was controlled carefully with digital and conventional thermometers. During these experiments, the photon flux was 200 μE m^{-2} s^{-1} in order to compensate light quenching by water.

5.4. Genomic DNA isolation and molecular analysis. The conventional "classical method" reported by Marmur (1961) for extracting genomic DNA has been modified several times by different research groups. However, the main basic framework of the method remains almost unaltered.

In the present study, Marmur's method was modified once more by adjusting it to further experimental requirements, since DDH and G+C mol% determinations studies require a high concentration of native DNA. In addition, the quality of the extracted DNA – an important factor to consider - may be seriously compromised when organic solvents such as phenol:chloroform are used. The modified method presented here avoids the use of a phenol:chloroform mixture and, if some other organic solvents are used during nucleic acid extraction, the amounts are carefully adjusted to each step in order to minimize their presence in the final dissolved DNA.

5.4.1. Genomic DNA "fishing" extraction protocol

Based on the method of Marmur (1961), with modifications.

Approximately 10 ml of pure cultures of the log growth phase were harvested by centrifugation at 10.000 rpm. This method was optimized (reagent concentrations and volume) to work at a final volume of 2 ml in an Eppendorf reaction tube for the G+C mol% determination.

However, if necessary e.g for DDH experiments, the method can be applied for the extraction of genomic DNA of large culture volumes, always maintaining the proportion (reagents used are given in Appendix 2).

Samples required: 10 ml or more of an active culture.

Yield: depending on the amount of culture and the organisms

Size of recovered DNA: native and of high molecular weight.

Quality: Good, checked by spectrophotometry, nanodrop, and gel electrophoresis

1. Wash the pellets by centrifugation at 13.000 rpm with 500 μl of PBS 1X buffer, discard the supernatant, repeat three times.

2. Add 300 μl of suspention buffer #1 (see appendix 2) to the pellet and homogenize by pippeting up and down several times. Add 20 μl lysozyme (solution ready to use 10 mg/ml) and vortex shortly.
3. Incubate at 37℃ for 10-60 min.
4. Take an aliquote of the solution (step 2, this protocol) and mix it with an equal volume of the lysis buffer #2 (see appendix 2). If the mixed turbid solution changes to translucient and viscous consistency, continue adding 400 μl of solution 2 (see appendix 2). Complete this step by adding 200 μl of 5M NaCl approximately (1/4 of the total volume). At this step, the final volume should be 1 ml.
5. Incubate at 65℃ for 10 to 60 min. It is recommendable to increase the incubation time e.g. overnight. Sometimes, if the organism has an unusually high content of polysaccharides, a solution of 2X CTAB buffer (50-100 μl) can be added. This solution will help to precipitate polysaccharides maintaining nucleic acids in the supernatant.
6. Once the lysis solution becomes completely clear, start precipitation of organic debris by adding an equal volume (1 ml) of chloroform: isoamyl alcohol mix (24:1, v/v). Homogenize gently by hand and centrifuge at 10,000 rpm for 10 min.
7. Carefully collect the supernatant (must be transparent) in a new tube with a truncated tip avoiding the contact with the organic phase contaminant present in the middle layer.
8. Precipitate the nucleic acids by adding 0.6 vol of cold isopropanol. Mix gently the tube up to down, in this step a precipitate (loose clot, see Fig. 6) will be observed.
9. Take the precipitate by “fishing” it with a fine tip (don’t suck it) and dehydrate it in 200 μl absolute ethanol in a separate Eppendorf tube. Take out the absolute ethanol by a pipette after 2 min and replace it by 70% ethanol and let stand for 2 min more.
10. After the washing step, take out all 70% ethanol and let the pellet dry in the air or at 37℃ in an oven.
11. Once dry, the DNA is resuspended in sterile 200 μl of 0.1x SSC buffer #3 (see appendix 2) and let dissolve it at 4℃ overnight.

12. A final RNase treatment step will be necessary. Add 10 µl of RNase (30 µg/ml; Molzym GmbH, Bremen) to each 100 µl of dissolved DNA and incubate at 37°C for 1 h. A final step to inactivate the enzyme will be necessary by incubating the dissolved nucleic acid at 65°C for 10 min.
13. Genomic DNA, as collected in step 11, is suitable for PCR based techniques and nuclease digestion.
14. For genotypic (DNA base ratio G+C content) and phylogenetic (DNA-DNA hybridization) analysis, it is recommend to dialyze the pellet (step 10) directly on a dialysis membrane (0.025 µm pore size) in a Petri dish against 0.1x SSC for at least 12 h.

Figure 6. Photograph of the precipitated nucleic acid in an Eppendorf tube after addition of 0.6 vol of cold isopropanol (step 8).

5.4.2. G+C mol% determination

There are different methods available to determine the mol% G+C content, however, the method developed by Marmur & Doty (1962), based on thermal denaturation, has been used in many articles in which descriptions of new species are reported. Although the spectrophotometric-based technique is being replaced by the HPLC-based method (Mesbah *et al.*, 1989), the thermal denaturation method still is applied in extensive strain comparisons.

For more detailed information about this method see appendix 3.

5.4.3. DNA-DNA hybridization

One of the most commonly used methods by bacterial taxonomists is the genomic DNA-DNA hybridization (DDH). Studies on DDH are still considered an important criterion for determining phenetic relationship and species delineation between bacteria (Stackebrandt & Goebel, 1994, Rosselló-Mora, 2006). A basic assumption in DNA similarity studies is the fact that deoxyribonucleic acid can be denatured and then renatured into the native molecule. When a competitor DNA is introduced after the denaturation of the DNA molecule, it is expected that the competitor DNA will renature or hybridize with the original molecule to some extent. The amount that hybrid DNA renatures, correlates with the amount of similarity between the sequences of the two different molecules.
Several methods have been developed for measuring the extent of DNA similarity. One of them is based on the use of labelled DNA for the detection of heteroduplex formation. Two different methods for labelling are still in use: The radioactive labelling DNA is performed by using radioactive iodine (iodination). Iodine can be bound to cytosine residues in the presence of tallium chloride at pH 4.8, and detection of the hybrid DNA is performed by radioactive counter scientillation. In the second method, the radioactive iodine is replaced by the analogous non-radioactive digoxygenin DIG-11-dUTP and Biotin-16-dUTP (Ziemke *et al.*, 1998) and the hybrid DNA detection is made by using a colorimetric method. However, in both cases separation of double strand and degradation of single stranded DNA (treated with S1 nuclease) is performed by the hydroxyapatite method (Britten & Kohne, 1968). A review of the literature where description of new bacterial species are reported indicated that the labeling methods so far represent 67% of the total hybridization experiments carried out during the year 2004 (Rosselló-Mora, 2006).

The non-labelling methods or so-called optical method utilize individual and a mixture of DNA which are melted. Thus, the degree of DNA-DNA binding is determined quantitatively by spectrophotometry (DeLey *et al.*, 1970). The kinetic of renaturation rates should be rigorously measured in stringent condition e.g. renaturation temperature and size of sheared DNA. The advantage of this method is the use of a non-radioactive tracer, consequently, no contamination is produced. However, one of its major disadvantages is that this technique requires high concentrated genomic DNA being time consuming. In addition, it also requires many experimental repetitions in order to assure a reliable result.

Genomic DNA to be used for DNA-DNA reassociation experiments was diluted to a concentration of 0.4 mg/ml in 0.1x SSC and 1 ml aliquots were sheared by sonication. Fragment sizes determined by electrophoresis on 1% agarose gel should be in the range of 200 to 800 bp as maximum (see Fig. 7). Samples that had fragments sizes larger than 800 bp were sonicated again. Denaturation was performed by filling 1.5 ml glass vials with 500 µl of single and equally mixed DNA samples in a boiling water bath for at least 15 min. DNA-DNA hybridization was performed following the renaturation kinetic method described by De Ley *et al.* (1970) using a Beckman DU 640 model UV spectrophotometer equipped with a thermostated multicell changer and a temperature controller.

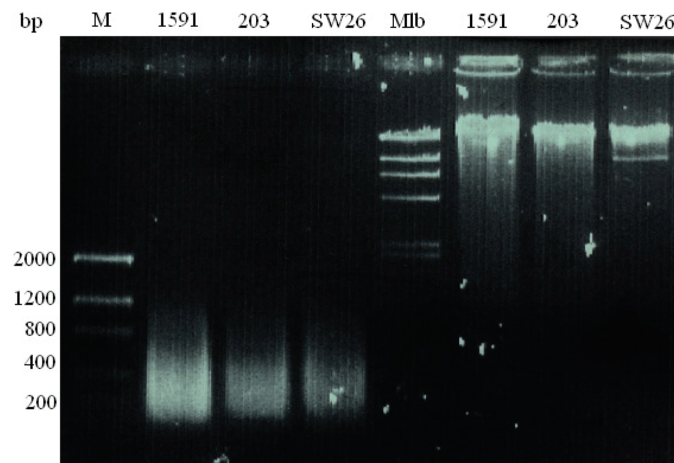


Figure 7. Fragment size of sonicated genomic DNA (left side) and native genomic DNA (right side) from different *Marichromatium* type strains as revealed by agarose gel electrophoresis (1%). M, low molecular marker; M1b, lambda marker. 1591, 203 are the type strains *M. purpuratum* and *M. gracile* respectively, SW26 is the new biotype *M. gracile* biotype *thermosulfidiphilum.*

5.4.4. Protein-coding gene analysis

Details about PCR amplification and sequencing are given in the corresponding chapters 3 and 4. Prior to the translation into amino acid sequences, nucleotide sequences were aligned and trimmed at both ends in order to create blocks of aligned sequences. Translation was performed by using the BioEdit package (Hall, 1992) and GeneDoc (Nicholas & Nicholas, 1997). Triplets or codons of nucleotide sequences were then translated to amino acid sequences and were compared to those proteins deposited in GenBank using BlastX. Quality of alignments in relation to the corresponding amino acid sequence and detection of hypervariable regions within each block of sequences was performed by using Gblocks program

(Castresana, 2000). After the cleaning procedure of all aligned sequences, further population genetics analysis and phylogenetic tree reconstruction were carried out.

5.4.5. Population genetics analysis

Polymorphic sites, haplotype diversity (H), nucleotide diversity (π), and the test of neutrality based on Tajima's D method (with 0 = neutral equilibrium, Tajima, 1989), were calculated with the DnaSP package version 4.0 (Rozas *et al.*, 2003). Multilocus linkage disequilibrium and the corresponding index of association (I_A) and (r_D) with I_A values ranging from **0 (= panmixia) to 1 (= total linkage disequilibrium**) were assessed using the program MULTILOCUS version 1.3 (Agapow & Burst, 2001). The allelic profile for all strains used as input for calculating linkage disequilibrium was generated with the program DnaSP 4.0. Reticulate network trees for each individual gene sequences as well as on the concatenated sequences were constructed by split decomposition analysis as implemented in the software Splits Tree version 4.0 (Huson & Bryant, 2006) using Hamming distances. The probability of recombination occurring in each gene was analysed by phi statistics as implemented also in the Splits Tree 4.0 program. The presence of parallelograms (splits) in the reticulate tree may indicate in some instance the occurrence of recombination. Pairwise estimations of nonsynonymous (d*N*) and synonymous (d*S*) substitutions per site were calculated by using the DnaSP 4.0 software package.

6. Aim of this study

The main subjects explored in the present thesis were:

i) to characterize new isolates of anoxygenic phototrophic purple sulfur bacteria. Therefore, samples were obtained from a tidal sediment in the northern German Wadden Sea, from shallow water marine sediments in the central Peruvian coast as well as samples from mangrove sediments from northern Peru.

ii) to compare these new isolates with type strains deposited in culture collections using a polyphasic approach.

iii) to fill a gap in the description of species within the two genera to which these new isolates affiliated (*Marichromatium* and *Allochromatium*). So far, species from both genera were described based solely on phenotypic and 16S rRNA-based phylogenetic criteria.

iv) to study the intra-and interspecies relationship of several strains within each genus (*Marichromatium* and *Allochromatium*) using a multi-gene approach (MLSA). With this aim, all available strains belonging to both genera were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) and from private collections.

v) to study population genetic among closely related strains belonging to the genera *Marichromatium* and *Allochromatium* isolated from distant locations.

7. Results and discussion

In this item, only the results that are not available in the manuscripts will be presented and discussed.

7.1. Characterization of a new purple sulfur bacterium, *Marichromatium gracile* biotype *thermosulfidiphilum* (chapter 2).

During the course of the MarMic program field training work in October 2003, samples from inter-tidal sediments from Sylt-Rømø Basin, northern Wadden Sea, Germany, were collected. Some samples were incubated at room temperature under continuous illumination of a 40W tungsten lamp (more details see material and methods of this chapter). Further enrichments and purifications were performed in order to isolate a pure culture of a rod shaped and fast motile bacterium, which shows sulfur globules inclusion inside the cell, and which was designated as strain SW26. The 16S rRNA comparison produced a sequence similarity of >99% with some uncultured *Marichromatium* strains and below 97% with the type species *Marichromatium purpuratum* DSM 1591 and *M. gracile* DSM 203 (Imhoff *et al*., 1998). At that time, *M. indicum* was also reported as a new species within the genus *Marichromatium* (Arunasri *et al*., 2005). The 16S rRNA gene sequence of strain SW26 showed a high similarity >99% to this new isolate, contradicting some preliminary results based on phenotypic and genotypic studies. These incongruencies could have been resolved by re-sequencing both type strains *M. purpuratum* and *M. gracile*, and the resulting polyphasic approach indicated that strain SW26 was a biotype of the described species *M. gracile* DSM 203^T. The complete description of the new biotype is given in chapter 2.

However, two remarkable phenotypic properties were observed in this new biotype e. g. its moderate thermal tolerance (growth up to 44°C) and its high tolerance to sulfide concentrations up to 16 mM which have not been reported for members of this genus before. Results of the thermal tolerance experiments are also reported in detail in chapter 2. The strain also contains a small plasmid, not found in other *Marichromatium* species.

Mechanism of regulation to thermal shock in the new biotype was also investigated Serrano *et al*. (2006). A short experiment was designed in order to analyse gene

expression of some molecular chaperons in active cultures of *M. gracile* biotype *thermosulfidiphilum* under thermal stress. Therefore, northern blot experiments were performed. As can be seen in Fig. 8, the amounts of molecular chaperon *dnaK* increase by heat shock (incubation at 55°C) and was proportional to the time of incubation. Thus, it could clearly be demonstrated that the molecular mechanism for protecting cells against thermal shock are highly active in this new biotype, however, the compared *dnaK* gene sequences (1100 bp) between both closely related strains *Marichromatium gracile* DSM 203 and *M. gracile* biotype *thermosulfidiphilum* showed only two nucleotide differences.

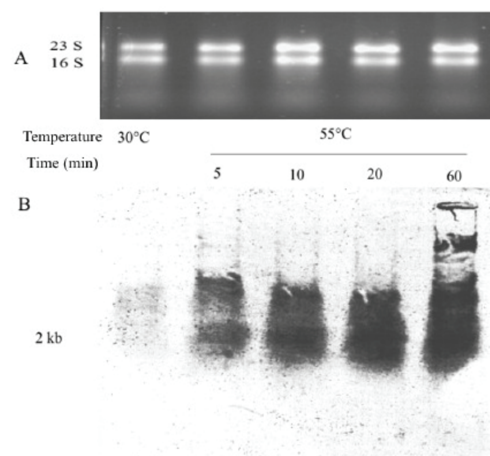


Figure 8. Northern hybridization of the *dnaK*: (A) Total RNA of strain *SW26* before (line 1, control 30°C) and after heat treatment for 5 to 60 min at 55°C (lines 2 to 5), (B) size and intensity of the bands reflecting mRNA for *dnaK* (2 kb).

Although both strains are almost similar by their 16S rRNA and *dnaK* gene sequences, expression of the *dnaK* molecular chaperon in *M. gracile* biotype *thermosulfidiphilum* could be due to the fact that some gene promoters were activated upon heat shock, which was not the case in its closest relative *M. gracile* DSM 203. These results may give hints to the molecular adaptation of microorganisms in the environment and could give new insights in gene expression studies, microdiversity, and microevolution. It was not possible to repeat the experiment with *M. gracile* DSM 203, because this organism does'n grow very well at temperatures above 40°C.

7.2 A multi-gene phylogeny of the genus *Marichromatium* (chapter 3).

In this sub-item only population genetic analyses within the genus *Marichromatium* are shown. Individual and concatenated phylogenetic trees are shown and discussed in detail in chapter 3. This analysis was based on the study of six protein-coding genes from nine *Marichromatium* strains.

Table 4 summarizes all genetic diversity parameters. As can be seen, the analysed protein-coding genes comprise eight fragments with average sizes ranging from 225 bp (*pufM*) to ~1400 bp (16S rRNA). The strength and direction of selection operating on each examined genes were assessed by calculating their respective (d*N*/d*S*) ratios. The resulting (d*N*/d*S*) ratios were between 0.004 (*recA*) and 0.194 (*dnaK*) with an average value of 0.082. In all cases, the observed values were significantly <1 (Table 4, Fig. 9), revealing a possible purifiying (stabilizing) selection operating on each protein-coding gene. However, it can not be discarded that these low d*N*/d*S* ratios could also be due to a result of neutral replacements accumulated over a long period of time (Rocha *et al.* 2006). These low d*N*/d*S* ratios are in agreement with the fact that essential protein-coding genes must be under purifying selection, consequently the observed accumulation of mutation (fixed over time) mainly synonymous substitutions should not affect the essential activity of the encoded proteins. That's the reason why prior to MLSA, the d*N*/d*S* ratio of candidate genes should be checked first.

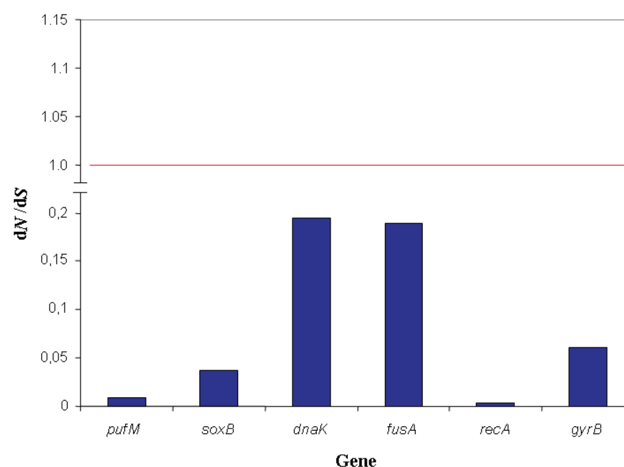


Figure 9. Non synonymous (*dN*) and synonymous (*dS*) substitution ratios of the six examined protein-coding genes in nine *Marichromatium* strains.

The (G + C) % content for each gene, which represents the average value of all nine strains analysed, has varied between 61.33% (*pufM*) to 67.11% (*soxB*). The overall average of the six housekeeping genes was 64.79%.

The observed gene diversity indices (H) were = 1 for both genes *dnaK* and *gyrB,* indicating that all nine strains analysed have different allelic profiles at both loci. The estimates of the nucleotide diversity (π) ranged from 0.013 for *dnaK* to 0.055 for *fusA* with an average value of 0.034. The probability that the mutations or nucleotide replacements found on all analysed housekeeping genes are neutral was also evaluated by using the Tajima's D test of neutrality (Tajima, 1989). As can be seen in Table 4, these values do not differ significantly from 0 (P> 0.10). The obtained allelic profiles from all analysed strains were used as input for calculating the linkage disequilibrium. The resulting values of I_A =1.271 and R_D = 0.439 indicate the clonal nature for all compared strains.

Table 4. Genetic diversity indices

Locus	Size bp	SG	N	H	π	Tajima'D*	Recombination (Phi) [a]	d*N*/d*S*	G+C(mol%)
pufM	225	18	7	0.944	0.033	0.318*	0.81	0.008	61.33
soxB	411	29	8	0.972	0.028	- 0.926*	0.55	0.037	67.11
fusA	330	55	8	0.972	0.055	- 0.632*	0.20	0.190	64.00
recA	366	32	7	0.944	0.027	- 0.888*	0.66	0.004	65.77
dnaK	741	32	9	1.000	0.013	- 0.923*	0.52	0.194	64.55
gyrB [b]	573	83	8	1.000	0.050	- 0.865*	0.06	0.061	66.00
average				0.972	0.034			0.082	64.79

SG=number of segregating sites; N=number of alleles; H=gene diversity; π =Nucleotide diversity; D*=Tajima test values not significantly different from zero (P>0.10). [a] = Phi recombination statistics calculated according to Husson & Bryant, (2006). [b]The *gyrB* analysis was based only on the study of eight strains.

The possibility of recombination among the analysed strains was also investigated. The phi test for recombination was applied to each independent protein-coding gene. The result indicates that statistically there is no evidence of recombination among the nine compaired strains for all analysed loci (Table 4). Nevertheless, as can be seen in Fig 10, the concatenated split network tree shows clearly the presence of parallelograms between both *M. purpuratum* strains and between *M. indicum* and *M. bheemlicum* respectively, suggesting that in some instance recombination occurres between each paired strains. Since the detection of recombination is in essence a statistical work, our results in this subject can not be conclusive because our study was conducted only on a low number of strains.

However, the multi-gene analysis has proved to be a good approach for separating the nine *Marichromatium* strains in four phylogroups or lineages based on the concatenated sequences (4331 bp) (see Fig. 10). As was pointed out by Gevers *et al.* (2005), a multi-gene analysis provides a buffer against the distorting effects of recombination at a single locus. Thus, the MLSA method can be used for discriminating closely related species and strains within the named species.

<u>Conclusions (Chapter 3)</u>

- Based on the MLSA, the nine analysed strains appeared to form clearly four separate lineages or phylogroups.
- The new *M. gracile* biotype *thermosulfidiphilum* is confirmed as a new strain of *M. gracile* DSM 203 and forms together with the other three closely related strains a new and independent phylogroup (lineage I) (Fig. 10).
- The separation of both strains of *M. purpuratum* in an independent phylogroup (lineage II) is supported by MLSA, though they are closely related to *M. gracile* phylogroup I by their >99% of the 16S rRNA similarity and their DDH relatedness above 70%.
- Strains isolated from India appear far related to other phylogroups, possibly forming two independent lineages.
- No major taxonomic changes are recommended, because MLSA is in agreement with the species assignment of the genus *Marichromatium* as was previously determined using classical methods.

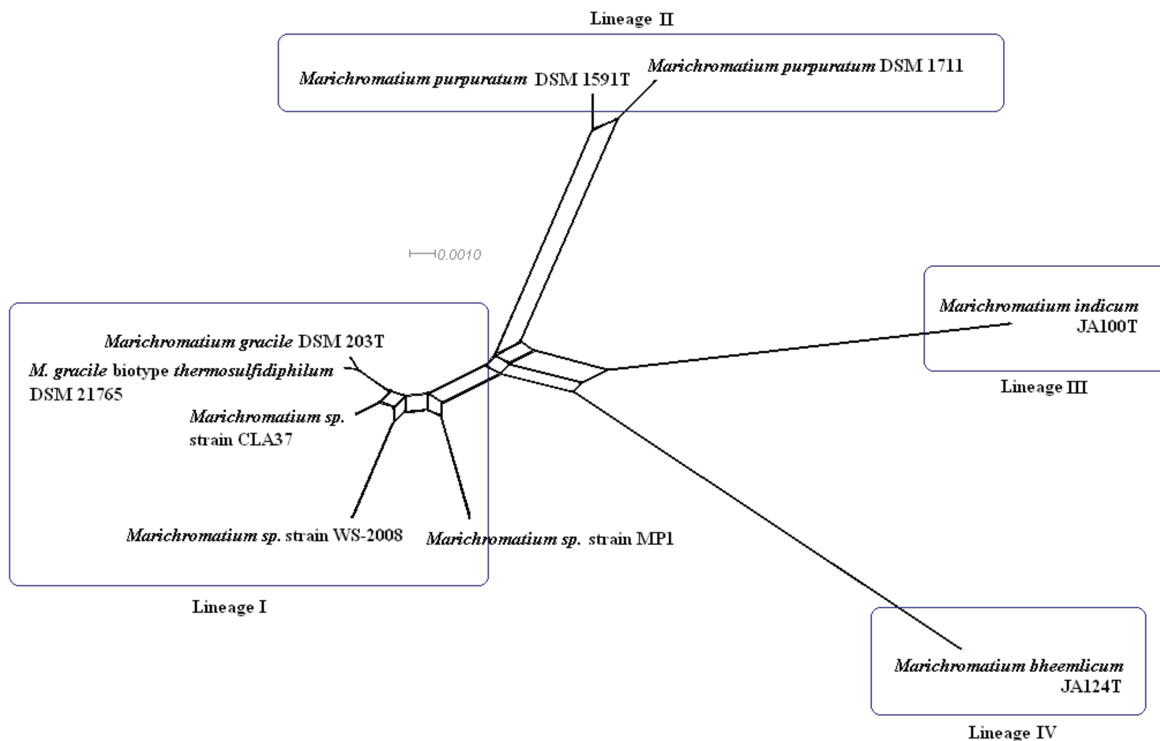


Figure 10. Concatenated split network tree constructed on the basis of the 16S rRNA, ITS, *pufM, soxB, fusA, recA,* and *dnaK* gene sequences (4331 bp). The nework-like tree based on the concatenated sequences separate the nine *Marichromatium* strains in four lineages. The presence of parallelograms may indicate that recombination occurs between some strains. Sequence similarities were corrected using the Jukes & Cantor method. The scale bar indicates 0.1% sequence difference.

7.3 Multi-gene phylogeny of the genus *Allochromatium* (chapter 4).

This sub-item is fully developed in Chapter 4 and includes Tables with population genetic analyses of seven *Allochromatium* strains. Individual and concatenated phylogenetic trees are also shown and discussed in detail in chapter 4.

MLSA was based on the study of eight protein-coding genes (Table S1, chapter 4). These protein-coding genes were also selected following the criteria above mentioned. As can be seen in Fig. 11, all chosen protein-coding genes showed d*N*/d*S* ratios <1, thus, assuming that these genes are under stabilizing selection, they were selected as good candidates for further MLSA.

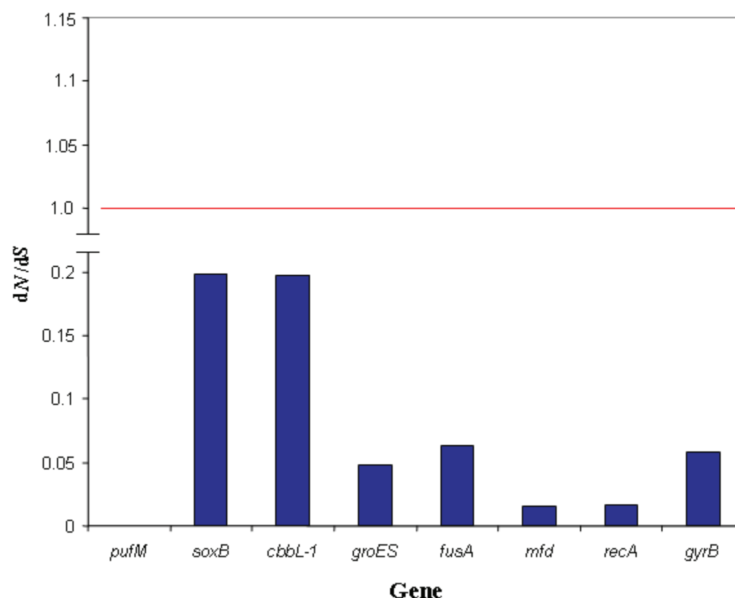


Figure 11. Non synonymous (*dN*) and synonymous (*dS*) substitution ratios of the eight examined protein-coding genes in seven *Allochromatium* strains.

Basically, the methodology employed was the same as described above for the genus *Marichromatium.* The concatenated gene sequences upon which a multigene tree was constructed totalize 4427 bp.

Conclusions (chapter 4)

- MLSA supports the formation of an independent phylogroup in which some closely related *A. vinosum* strains were included (see Fig. S1, chapter 4).
- The taxonomic status of *A. minutissimum* needs to be revised. This type strain is identical to *A. vinosum* strains based on the 16S rRNA similarity >99% and DDH values above 80% relatedness (Table 2, chapter 4). This finding was in agreement with previous studies based on genotyping methods in which a comparison of both type strains showed identical banding patterns (Mas-Castella *et al*., 1996; Pavon & Gaju, 1997).
- The multi-gene phylogeny confirms the closely relationship between *A. minutissimum* and *A. vinosum* strains. However, although *A. minutissimum* cluster within the *A. vinosum* phylotype, it branches separately to the other *A. vinosum* strains (Fig. 3, chapter 4), suggesting that the morphological difference of *A. minutissimum* could be a result of cryptic speciation or another speciation process.
- Strain BH-2 was first described to be one of the species *Allochromatium vinosum,* based on its high DDH relatedness above 80%. However, nucleotide divergence of

the 16S rRNA gene and concatenated multi-gene sequence analysis between strain BH-2 and its closest *Allochromatium* species examined in this study showed the lowest values (see Chapter 4). Regarding these results and the compared phenotypic properties, strain BH-2 might deserve the status of a new species within this genus.
- Based on MLSA, the new isolate strain AX1YPE is confirmed as a new species of the genus *Allochromatium.* A detailed description of strain AX1YPE is given in chapter 5.

7.4 Description of a novel *Allochromatium* species, *A. humboldtianum* sp. nov.

A new anoxygenic phototroph that contains Bchl *a* as pigment was isolated from soft marine sediments at 48 m depth of shallow waters of the Peruvian coast. The organism exhibits some new features not yet described for this genus such as growth at low light intensities and use of a wide range of nitrogen compounds as nitrogen sources (see Chapter 5). The capability of strain AX1YPE to grow at low light conditions could reflect the characteristic of the environment where the organism was retrieved from, e.g. high primary productivity and suspended organic particles in the upper part of the water column as well as a constant chemocline at 20 m depth. Under these conditions, light is usually scattered in the surface. On the other hand, the ability to use different nitrogen sources could also reflect the adaptation of the organism to a high productive environment where different nitrogen species are abundantly available as a result of high organic matter degradation. Strain AX1YPE showed a 16S rRNA gene sequence similarity of 98.9% with *A. vinosum* and *A. minutissimum,* 96% and 95% with *A. renukae* and *A. warmingii* respectively, but only 59 and 64% DDH-relatedness to the closest relatives *A. vinosum* and *A. minutissimum*.
Although strain AX1YPE was typically retrieved from marine sediments, additional enrichment cultures were also performed using as inoculum the overlying water which was concentrated on polycarbonate membranes; but no PSB were obtained. Because of its particular phenotype and the polyphasic approach applied in the characterization of strain AX1YPE, we describe it as a new species of the *Allochromatium* genus for which the name *Allochromatium humboldtianum* is proposed (see Chapter 5).

8. This work

This thesis is presented in form of one published article, a submitted one, and two manuscripts in preparation.

List of publications

Contribution to the manuscripts presented in this thesis

Chapter 2

Serrano W., Amann, R. & Fischer, U. (2009). A new moderately thermophilic and high sulfide tolerant biotype of *Marichromatium gracile*, isolated from tidal sediments of the German Wadden Sea: *Marichromatium gracile* biotype *thermosulfidiphilum*. Syst. Appl. Microbiol. 32: 1-7.

Enrichment cultures were obtained by W. S. during the MarMic training course in general microbiology, complete characterization of the new isolate was done by W. S. The manuscript was written by W. S. with editorial assistance of the co-authors.

Chapter 3

Serrano W., Amann, R., Rossello-Mora, R. & Fischer, U. (2009). Evaluation of the use of multilocus sequence analysis (MLSA) to resolve taxonomic conflicts within the genus *Marichromatium*. Syst. Appl. Microbiol. Accepted.

Concept by W. S., practical work and writing by W. S., editorial assistance and comments ***on*** *the final text by the co-authors.*

Chapter 4

Serrano W., Amann, R. & Fischer, U. (2009). The genus *Allochromatium* (*Chromatiales Chromatiaceae*) revisited: a study on its intragenic structure based on Multilocus Sequence Analysis (MLSA) and DNA-DNA hybridization (DDH). Manuscript in revision.

Concept by W. S., practical work and writing by W. S., editorial assistance and comments by R. A., and U. F.

Chapter 5

Serrano, W., Schrübbers, J., Amann, R. & Fischer, U. (2009). *Allochromatium humboldtianum* sp. nov., a novel *Allochromatium* species isolated from soft marine sediments from Peruvian coast. Manuscript in revision.

Concept by W. S., practical work and writing by W. S., electron microscopy by J. S., editorial help by U. F. and R. A.

Part of this thesis work was presented as poster communications

Serrano, W., Amann, R. & Fischer, U. (2006). Thermal tolerance and heat-shock response of a new purple sulfur bacterium, strain SW26 isolated from Wadden Sea. Book of programm and abstracts. 12th International Symposium of Phototrophic Prokaryotes (Tandeau de Marsac, N., Stal, L. J., Caumette, P. eds.), Pau, France, p 161.

Serrano, W., Amann, R. & Fischer, U. (2008). The use of Multilocus Sequence Analysis in bacterial sytematics: A multilocus phylogeny of the genus *Allochromatium.* Poster presented at the International Gordon Research Conference of Marine Microbes, July 21-25, Il Ciocco, Italy.

9. References.

Agapow, P.-M. & Burst, A. (2001). Indices of multilocus linkage disequilibrium. Mol. Ecol. Notes 1, 101-102.

Allen, J. P. & Williams, J. C. (1998). Photosynthetic reaction centers. FEBS Lett. 438, 5-9.

Amann, R. I., Binder, B. J., Olson, R. J., Chrisholm, S. W., Devereaux R. & Stahl, D. A. (1990a). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl. Environ. Microbiol. 56: 1919-1925.

Amann, R., Krumholz, L. & Stahl, D. A. (1990b). Fluorescent oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. J. Bacteriol. 172: 762-770.

Antranikian, G., Herzberg, C. & Gottschalk, G. (1982). Characterization of ATP citrate lyase from *Chlorobium limicola*. J. Bacteriol. 152, 1284-1287.

Arunasri, K., Sasikala, Ch., Ramana, V. C., Süling, J. & Imhoff, J. F. (2005). *Marichromatium indicum* sp. nov., a new purple sulfur *Gammaproteobacterium* from mangrove soil of Goa, India. Int. J. Syst. Evol. Microbiol. 55, 673-679.

Badger, J. H., Jonathan, A., Eisen, A. & Ward, N. L. (2005). Genomic analysis of *Hyphomonas neptunium* contradicts 16S rRNA gene-based phylogenetic analysis: implications for the taxonomy of the order '*Rhodobacterales*' and *Caulobacterales*. Int. J. Syst. Evol. Microbiol. 55, 1021-1026.

Bapteste, E., Boucher, Y., Leigh, J. & Dooloitle, W. F. (2004). Phylogenetic reconstruction and lateral gene transfer. Trends Microbiol. 12, 406-411.

Beatty, J. T., Overmann, J., Lince, M. T., Manske, A. K., Lang, A. S., Blankenship, R. E., Van Dover, C. L., Martinson, T. A. & Plumley, F. G. (2005). An obligate photosynthetic bacteria anaerobe from deep-sea hydrothermal vent. Proc. Natl. Acad. Sci. USA. 102, 9306-9310.

Beh, M., Strauss, G., Huber, R., Stetter, K. O. & Fuchs, G. (1993). Enzymes of the reductive citric cycle in the autotrophic eubacterium *Aquifex pyrophilus* and in the achaebacterium *Thermoproteus neutrophilus*. Arch. Microbiol. 160, 306-311.

Beja, O., Suzuki, M. T., Heidelberg, J. F., Nelso, W. C., Preston, C. M., Hamada, T., Eisen, J. A., Fraser, C. M. & DeLong, E. F. (2002). Unsuspected diversity among marine aerobic anoxygenic phototrophs. Nature 450, 630-633.

Berg, I. A., Kockelkorn, D., Buckel, W. & Fuchs, G. (2007). A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxyde assimilation pathway in archaea. Science. 318, 1782-1786.

Blankenship, R. E., Madigan, M. T. & Bauer, C. E. (eds.) (1995). Anoxygenic photosynthetic bacteria, Kluver Academic Pub., Dordrecht, The Netherlands.

Blankenship, R. E. (2001). Molecular evidence for the evolution of photosynthesis. Trends Plant Sci. 6, 4-6.

Britten, R. J. & Kohne, D. E. (1968). Repeated sequences in DNA. Science. 161, 529-540.

Budinoff, C. R. & Hollibaugh, J. T. (2008). Arsenite-dependant photoautotrophy by an *Ectothiorhodospira*-dominated consortium. ISME J. 2, 340-343.

Burgess, J. G., Miyashita, H., Sudo, H. & Matsunaga, T. (1991). Antibiotic production by the marine photosynthetic bacterium *Chromatium purpuratum* NKPB 031704 localization of activity to the chromatophores. FEMS Microbiol. Lett. 84: 301-306.

Busse, H. J., Denner, E. B. M. & Lubitz, W. (1996). Classification and identification of bacteria: current approaches to an old problem. Overview of methods used in bacterial systematics. J. Biotech. 47, 3-38.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540-552.

Campbell, B. J. & Cary, S. C. (2004). Abundance of reverse tricarboxylic acid cycle genes in free-living microorganisms at deep-sea hydrothermal vents. Appl. Environ. Microbiol. 70, 6282-6289.

Clarke, S. C., Diggle, M. A. & Edwards, G. F. (2001). Semiautomation of Multilocus Sequence Typing for the characterization of clinical isolates of *Neisseria meningitidis*. J. Clin. Microbiol. 39, 3066-3071.

Cline, J. D. (1969). Spectrophotometric determination of hydrogen sulfide in natural waters. Limnol. Oceanogr. 14, 454-458.

Colwell, R. R. (1970). Polyphasic taxonomy of the genus *Vibrio*: numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species. J. Bacteriol. 104, 410-433.

De Ley, J., Catoir, H. & Reynaerts, A. (1970). The quantitative measurement of DNA hybridization from renaturation rates. Eur. J. Biochem. 12, 133-142.

DeLong, E. F., Wickham, G. S. & Pace, N. R. (1989). Phylogenetic stains: ribosomal RNA-based probes for identification of single cells. Nature 243, 1360-1363.

Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. (1985). Structure of the protein subunits in the photosyntheic reaction centre of *Rhodopseudomonas viridis* at 3Å resolution. Nature 318, 618-624.

Deisenhofer, J. & Michel, H. (1989). Nobel lecture. The photosynthetic reaction center from the purple bacterium *Rhodopseudomonas viridis*. EMBO J. 8: 2149-2170.

Doty, P., Marmur, J., Eigner, J., & Schildkraut, C. (1960). Strand separation and specific recombination in deoxyribonucleic acids: physical chemical studies. Proc. Natl. Acad. Sci. USA. 46, 461-476.

Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J. & Spratt, B. G. (2000). Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J. Clin. Microbiol. 38, 1008-1015.

Evans, M. C. W., Buchanan, B. B. & Arnon, D. I. (1966). A new ferredoxin-dependent carbon reduction cycle in a photosynthetic bacterium. Proc. Natl. Acad. Sci. USA. 55, 928-934.

Feil, E. J. (2004). Small change: Keeping pace with microevolution. Nat. Rev. 2, 483-495.

Fitch, W. M. & Yasunobu, K. T. (1975). Phylogenies from amino acid sequences aligned with gaps: the problems of gaps weighting. J. Mol. Evol. 5, 1-24.

Fox, G. E., Wisotzkey, J. D. & Jurtshuk, P. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int. J. Syst. Bacteriol. 42, 166-170.

Friedrich, C. G., Rother, D., Bardischewsky, F., Quentmeier, A. & Fischer, J. (2001). Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism. Appl. Environ. Microbiol. 67, 2873-2882.

Fuchs, G. (1989). Alternate pathways of autotrophic CO_2 fixation. In: Schlegel, H. G. & Bowien, B. (eds.). Autotrophic bacteria. Springer-Verlag, New york, pp. 365-382.

Fuchs, G. (2007). Allgemeine Mikrobiologie. Thieme, G., Stuttgart.

Fuerst, J. A. (2005). Intracellular compartmentation in Planctomycetes. Annu. Rev. Microbiol. 59, 299-238.

Fujii, R., Shimonaka, S., Uchida, N., Gardiner, A. T., Cogdel, R. J., Sugisaki, M. & Hashimoto, H. (2008). Construction of hybrid photosynthetic units using peripheral and core antennae from two different species of photosynthetic bacteria: detection of the energy transfer from bacteriochlorophyll *a* in LH2 to bacteriochlorophyll *b* in LH1. Photosynth. Res. 95, 327-337.

Gest, H. & Favinger, J. L. (1983). *Heliobacterium chlorum*, an anoxygenic brownish-green photosynthetic bacterium containing a “new” form of bacteriochrlorophyll. Arch. Microbiol. 136, 11-16.

Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F. L. & Swing, J. (2005). Re-evaluating prokaryotic species. Nature Rev. 3, 733-739.

Gotov, I. N. (1986). Hydrogenases of phototrophic microorganisms. Biochem. 68:181-187.

Griffin, B. M., Schott, J. & Schink, B. (2007). Nitrite, an electron donor for anoxygenic photosynthesis. Science. 316, 1870.

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

Heathcoat, P., Fyke, P. K. & Jones, M. R. (2002). Reaction centres: the structural and evolution of biological solar power. Trends Biochem. Sci. 27, 79-87.

Heinnickel, M. & Golbeck, J. H. (2007). Heliobacterial photosynthesis. Photosynth. Res. 92, 35-53.

Hess, S., Visscher, K., Ulander, J., Pullerits, T., Jones, M. R., Hunter, C. N. & Sundtröm, V. (1993). Direct energy transfer from peripheral LH2 antenna to the reaction center in a mutant of *Rhodobacter sphaeroides* that lacks the core LH1 antenna. Biochemistry. 32, 10314-10322.

Holo, H. (2004). *Chloroflexus aurantiacus* secretes 3-hydroxypropionate, a possible intermediate in the assimilation of CO_2 and acetate. Arch. Microbiol. 151, 252-256.

Huson, D. H. & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254-267.

Imhoff, J. F. (1995). Taxonomy and physiology of phototrophic purple bacteria and green sulfur bacteria. *In*: Blankenship, R. E., Madigan, M. T., Bauer, C. E. (eds.)

Anoxygenic photosynthetic bacteria, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 1–15.

Imhoff, J. F., Süling, J. & Petri, R. (1998). Phylogenetic relationships among *Chromatiaceae*, their taxonomic reclassification and description of the new genera *Allochromatium, Halochromatium, Isochromatium, Marichromatium, Thiococcus, Thiohalocapsa and Thermochromatium*. Int. J. Syst. Bacteriol. 48, 1129-1143.

Imhoff, J. F. & Caumette, P. (2004). Recommended standards for the description of new species of anoxygenic phototrophic bacteria. Int. J. Syst. Evol. Microbiol. 54, 1415-1421.

Jaspers, E. & Overman, J. (2004). Ecological significance of microdiversity: Identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysiologies. App. Environ. Microbiol. 70: 4831-4839.

Johnson, J. L. (1973). Use of nucleic-acid homologies in the taxonomy of anaerobic bacteria. Int. J. Syst. Bacteriol. 23, 308-315.

Ketelaars, M., Hofman, C., Köhler, Howard, T. D., Cogdell, R. J., Schmidt, J. & Aartsma, T. (2002). Spectroscopy on individual light-harvesting 1 complexes of *Rhodopseudomonas acidophila*. Biophys. J. 83, 1701-1715.

Konstantinidis, K. T., Ramette, A. & Tiedje, J. (2006). Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. Appl. Environ. Microbiol. 72, 7912-7915.

Kulp, T. R., Hoeft, S. E., Asao, M., Madigan, M. T., Hollibaugh, J. T., Fisher, J. C., Stolz, J. F., Culberston, C. W., Miller, L. G. & Oremland, R. S. (2008). Arsenic (III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. Science. 231, 967-970.

Lindsay, M. R., Webb, R. I., Strous, M., Jetten, M. S. M., Butler, M. K., Forde, R. J. & Fuerst, J. A. (2001). Cell compartmentalisation in planctomycetes : novel types of structural organisation for the bacterial cell. Arch. Microbiol. 175, 413-429.

Lüthy, L., Fritz, M. & Bachofen, R. (2000). *In situ* determination of sulfide turnover rates in a meromictic alpine lake. Appl. Environ. Microbiol. 66, 712-717.

Ma, F., Kimura, Y., Zhao, X-H., Wu, Y-S., Wang, P., Fu, L-M., Wang, Z-Y. & Zhang J-P. (2008). Excitation dynamics of two spectral forms of the core complexes from photosynthetic bacterium *Thermochromatium tepidum*. Biophys. J. 95, 3349-3357.

Madigan, M. T. & Ormerod, J. G. (1995). Taxonomy, physiology and ecology pf Heliobacteria. *In*: Blankenship, R. E., Madigan, M. T., Bauer, C. E. (eds.) Anoxygenic photosynthetic bacteria, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 17–30.

Madigan, M. T. & Martinko, J. M. (2006). Brock biology of microorganisms. 11th edn. New Jersey, Pearson Prentice-Hall, Inc.

Madigan, M. T. (2000). The family *Heliobacteriaceae*. In: Garrity, G. (ed.), Bergey´s Manual of Systematic Bacteriology, vol. 1. Springer, New York, pp. 615-620.

Maiden, M. C. J., Bygraves, J. A., Feil, J. A., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K. & Caugant, D. A. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. USA. 95, 3140-3145.

Manske, A. K., Glaeser, J., Kuypers, M. M. M. & Overmann, J. (2005). Physiology and phylogeny of green sulfur bacteria forming a monospecific phototrophic assemblage at a depth of 100 meters in the Black Sea. Appl. Environ. Microbiol. 71, 8049-8060.

Marmur, J. & Lane, D. (1960). Strand separation and specific recombination in deoxyribonucleic acids (DNA): biological studies. Proc. Natl. Acad. Sci. 46, 453-461.

Marmur, J. (1961). A procedure for the isolation of desoxyribonucleic acid from micro-organisms. J. Mol. Biol. 3, 208-218.

Marmur, J. & Doty, P. (1962). Determination of the base composition of desoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. 5, 109-118.

Martinez-Murcia, A. J., Benlloch, S. & Collins, M. D. (1992). Phylogenetic interrelationships of members of the genera *Aeromonas* and *Plesiomonas* as determined by 16S ribosomal DNA sequencing: Lack of congruence with results of the DNA-DNA hybridizations. Int. J. Syst. Bacteriol. 42, 412-421.

Mas-Castella, J., Guerrero, R. & De Jonckheere, J. F. (1996). High degree of similarity between *Chromatium vinosum* and *Chromatium minutissimum* as revealed by riboprinting. Int. J. Syst. Bacteriol. 46, 922-925.

Mesbah, M., Premachandran, U. & Whitman, W. B. (1989). Precise measurement of the G+C content of desoxyribonucleic acid by high-performance liquid chromatography. Int. J. Syst. Bacteriol. 39: 159-167.

Nicholas, K. B. & Nicholas, H. B. Jr. (1997). GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the authors.

Norris, J. R. & Burger, H. D. (1963). Esterases of crystalliferous bacteria pathogenic for insect: epizootiological application. J. Insect. Pathol. 5, 460-472.

Overmann, J., Beatty, J. T., Hall, K. J., Pfennig, N. & Northcote, T. G. (1991). Characterization of a dense, purple sulfur bacterial layer in a meromictic salt lake. Limnol. Oceanogr. 36, 846-859.

Overmann, J. (2001). Diversity and ecology of phototrophic sulfur bacteria. Microbiol. Today 28, 116-118.

Pavon, V. & Gaju, N. (1997). Genomic relationship of five species of the genus *Chromatium* by analysis of large restriction fragments (macrorestriction analysis) using Pulsed-Field Gel Electrophoresis. Int. J. Syst. Bacteriol. 47, 586-589.

Peduzzi, S., Tonolla, M. & Hahn, D. (2003). Isolation and characterization of aggregate-forming sulfate-reducing and purple sulfur bacteria from the chemocline of meromictic Lake Cadagno, Switzerland. FEMS Microbiol. Ecol. 45, 29-37.

Pierson, B. K. & Castenholz, R. W. (1995). Taxonomy and physiology of filamentous anoxygenic phototrophs. *In*: Blankenship, R. E., Madigan, M. T., Bauer, C. E. (eds.) Anoxygenic photosynthetic bacteria, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 31–47.

Pfennig, N. & Trüper, H. G. (1992). The family *Chromatiaceae*. In: Ballows, A., Trüper, H. G., Dworkin, M., Harder, W., Schleifer, K. H. (eds.), The Prokaryotes, Springer, Berlin, Heidelberg, New York, pp. 3200-3221.

Ponsano, E. H. G., Lacava, P. M. & Pinto, M. F. (2002). Isolation of *Rhodocyclus gelatinosus* biomass produced in poultry slaughterhouse wastewater. Brazilian Arch. Biol. Technol. 46, 1-11.

Rakhely, G., Colbeau, A., Garin, J., Vignais, P. M. & Kovacs, K. L. (1998). Unusual organization of the genes coding for HydSL, the stable (NiFe) hydrogenase in the photosynthetic bacterium *Thiocapsa roseopersicina*, BBS. J. Bacteriol. 180: 1460-1465.

Remy, A. & Gerwert, K. (2003). Coupling of light-induced electron transfer to proton uptake in photosyntheis. Nature Struct. Biol. 10, 637-644.

Rocha, E. P. C., Maynard Smith, J., Hurst, L. D., Holden, T. G., Cooper, J. E., Smith, N. H. & Feil, E. J. (2006). Comparison of d*N*/d*S* are time dependent for closely related bacterial genomes. J. Theor. Biol. 239, 226-235.

Rozas, J., Sánchez-DelBarrio, J.C, Messeguer, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Rosselló-Mora, R. & Amann, R. (2001). The species concept for prokaryotes. FEMS Microbiol. Rev. 25, 39-67.

Rosselló-Mora R. (2006). DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In: Stackebrandt, E. (ed.), Molecular Identification, Systematics and Population Structure of Prokaryotes. Springer-Verlag, Berlin, pp. 23-50.

Roussel, E. G., Bonavita, M-A. C., Querellou, J., Cragg, B. A., Webster, G., Prieur & Parkes, R. J. (2008). Extending the sub-sea-floor biosphere. Science 320, 1046.

Santelli, C. M., Orcutt, B. N., Banning, E., Bach, W., Moyer, C. L., Sogin, M. L., Staudigel, H. & Edwars, K. J. (2008). Abundance and diversity of microbial life in ocean crust. Nature 453, 653-656.

Sasikala, G. H. & Ramana, C. H. V. (1995). Biotechnological potential of anoxygenic phototrophic bacteria. I. production of single-cell protein, vitamins, ubiquinones, hormones, and enzymes and use in waste treatment. Adv. Appl. Microbiol. 41, 173-226.

Selander, R. K., Caugant, D. A., Ochman, H., Musser, J. M., Gimour, M. N. & Whittam, T. S. (1986). Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl. Environ. Microbiol. 51, 873-884.

Serrano, W., Amann, R. & Fischer, U. (2006). Thermal tolerance and heat-shock response of a new purple sulfur bacterium, strain SW26 isolated from Wadden Sea. *In*: Tandeau de Marsac, N., Stal, L. J. & Caumette, P. (eds), Abstract book, 12th ISPP, Pau, France, p 161.

Serrano, W., Amann, R. & Fischer, U. (2009). A new moderately thermophilic and high sulfide tolerant biotype of *Marichromatium gracile*, isolated from tidal sediments of the German Wadden Sea: *Marichromatium gracile* biotype *thermosulfidiphilum*. Syst. Appl. Microbiol. 32, 1-7.

Schauder, R., Widdel, F. & Fuchs, G. (1987). Carbon assimilation pathways in the sulfate-reducing bacteria. II. Enzymes of a reductive citric acid cycle in the autotrophic *Desulfobacter hydrogenophylus*. Arch. Microbiol. 148, 218-225.

Schulz, H. N. Jørgensen, B.B., Fossing, H. A. & Ramsing, N. B. (1996). Community structure of filamentous, sheath-building sulfur bacteria, *Thioploca spp.*, off the coast of Chile. Appl. Environ. Microbiol. 62, 1855-1862.

Schulz, H. N., Brinkhoff, T., Ferdelman, T. G., Hernandez-Marine, M., Teske, A. & Jørgensen, B. B. (1999). Dense populations of a giant sulfur bacterium in Namibian shelf sediments. Science 284, 493-495.

Schwartz, R. M. & Dayhoff, M. O. (1978). Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. Science 199, 395-403.

Shiba, T., Simidu, U. & Taga, N. (1979). Distribution of aerobic bacteria which contain bacteriochlorophyll a. Appl. Environ. Microbiol. 38, 43-48.

Soria-Carrasco, V., Valens-Vadell, M., Peña, A., Anton, J., Amann, R., Castresana, J. & Rosello-Mora, R. (2007). Phylogenetic position of *Salinibacter ruber* based on concatenated protein alignments. Syst. Appl. Microbiol. 30, 171-179.

Staley, J. T. (2006). The bacterial species dilemma and the genomic-phylogenetic species concept. Phil. Trans. R. Soc. B. 361, 1899-1909.

Stackebrandt, E. & Goebel, B. M. (1994). Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int. J. Syst. Bacteriol. 44, 846-849.

Stanier, R. Y. & Van Niel, C. B. (1962). The concept of bacterium. Arch. Microbiol. 42, 17-35.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585-595.

Thar, R. & Kühl, M. (2001). Motility of *Marichromatium gracile* in Response to Light, Oxygen, and Sulfide. Appl. Environ. Microbiol. 67, 5410-5419.

Thompson, F. L., Iida, T. & Swings, J. (2004). Biodiversity of Vibrios. Microbiol. Mol. Biol. Rev. 68, 403-431.

Thompson, F. L., Gevers, D., Thompson, C. C., Dawyindt, P., Naser, S., Hoste B. & Munn, C. B. (2005). Phylogeny and molecular identification of Vibrios on the basis of Multilocus sequence analysis. Appl. Environ. Microbiol. 71, 5107-5115.

Tran, N. H. (2008). A photoautotrophic bacterium utilizing antraquinol or humic substances as electron donor. In: New organic substrates for anoxygenic phototrophic bacteria. PhD thesis, Bremen, University of Bremen, Germany, pp. 78-94.

Trüper, H. G. & Pfennig, N. (1981). Characterization and identification of the anoxygenic phototrophic bacteria. *In:* Starr, M. P., Stolp, H., Trüper, H. G., Balows, A. & Schlegel, H. G. (eds.), The Prokaryotes. Springer-Verlag, New York, pp. 299-312

Wächtershauser, G. (1990). Evolution of the first metabolic cycles. Proc. Natl. Acad. Sci. USA. 87, 200-204.

Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. & Trüper, H. G. (1987). Report of the *Ad Hoc* committee on reconciliation of approaches to bacterial systematics. Int. J. Syst. Bacteriol. 37, 463-464.

Wenke, T. L. & Vogt, J. C. (1981). Temporal changes in a pink feedlot lagoon. Appl. Environ. Microbiol. 41, 381-385.

Widdel, F. & Pfennig, N. (1981). Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids. I. Isolation of a new sulfate-reducer enriched with acetate from saline environment. Description of *Desulfobacter postgatei* gen. nov. sp. nov. Arch. Microbiol. 129, 395-400.

Widdel, F. & Bak, F. (1992). Gram-negative mesophilic sulfate-reducing bacteria. *In*: Balows, A., Trüper, H. G., Dworkin, M., Harder, W. & Schleifer, K. H. (eds.), The Prokaryotes. Springer, New York, pp. 3353-3378.

Widdel, F., Schnell, S., Heising, S., Ehrenreich, A., Assmus, B. & Schink, B. (1993). Ferrous iron oxidation by anoxygenic phototrophic bacteria. Nature. 362, 834-835.

Whitman, W. B, Coleman, D. C. & Wiebe, W. J. (1998). Prokaryotes: The unseen majority. Proc. Natl. Acad. Sci. USA 95, 6778-6583.

Woese, C. R. (1987). Bacterial evolution. Microbiol. Rev. 51, 221-271.

Yurkov, V. V. & Beatty, J. T. (1998). Aerobic anoxygenic phototrophic bacteria. Microbiol. Mol. Biol. Rev. 62, 695-724.

Ziemke, F., Höfle, M. G., Lalucat, J. & Rossello-Mora, R. (1998). Reclassification of *Shewanella putrefaciens* Owen´s genomic group as *Shewanella baltica* sp. nov. Int. J. Syst. Bacteriol. 48, 179-186.

Chapter 2

A new moderately thermophilic and high sulfide tolerant biotype of *Marichromatium gracile,* isolated from tidal sediments of the German Wadden Sea: *Marichromatium gracile* biotype *thermosulfidiphilum*

Wilbert Serrano[1, 2, *], Rudolf Amann[2] and Ulrich Fischer[1]

Systematic and Applied Microbiology (2009) 32:1-7

ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

Systematic and Applied Microbiology 32 (2009) 1–7

SYSTEMATIC AND APPLIED MICROBIOLOGY

www.elsevier.de/syapm

A new moderately thermophilic and high sulfide tolerant biotype of *Marichromatium gracile,* isolated from tidal sediments of the German Wadden Sea: *Marichromatium gracile* biotype *thermosulfidiphilum*☆

Wilbert Serrano[a,b,*], Rudolf Amann[b], Ulrich Fischer[a]

[a]*Zentrum für Umweltforschung und nachhaltige Technologien, Fachbereich Biologie/Chemie, Abteilung Marine Mikrobiologie, Universität Bremen, D-28359 Bremen, Germany*
[b]*Max Planck-Institute for Marine Microbiology, Celsiusstraße 1, D-28359 Bremen, Germany*

Received 7 July 2008

Dedicated to the memory of the famous microbiologist Norbert Pfennig who passed away in February 2008

Abstract

A purple sulfur bacterium, strain SW26, was isolated in pure culture from intertidal sediments from the Sylt-Rømø Basin, German Wadden Sea, sharing many properties with validated *Marichromatium* species, but differing significantly by possessing a plasmid, by tolerating up to 16 mM sulfide, and up to 44 °C for growth. Strain SW26 has a DNA base composition of 68.3 mol% G+C, a 16S rRNA gene sequence similarity of >99% to those of *Marichromatium* species, and shows the highest level of genomic relationship with *Marichromatium gracile*, despite its remarkably different phenotypic characters. Based upon high genomic similarity but different physiological properties of strain SW26 with respect to the type strain of *M. gracile,* a novel biotype, designated as *M. gracile* biotype *thermosulfidiphilum* is described.
© 2008 Elsevier GmbH. All rights reserved.

Keywords: *Chromatiaceae*; Sulfide tolerance; *Marichromatium gracile* biotype

Introduction

The true so far described four marine species of the family *Chromatiaceae* were placed into the genus *Marichromatium* [5]. While *Marichromatium gracile*, *Marichromatium indicum*, and *Marichromatium bheemlicum* were isolated from sediments and contain carotenoids of the spirilloxanthin series (see [9]), *M. purpuratum* (formerly: *Chromatium purpuratum*) originated from the marine sponge *Ircinia* sp. and possesses okenone [7].

The initial descriptions of *M. gracile* (formerly: *Chromatium gracile*) and *M. purpuratum* were mainly based on phenotypic criteria [7,14], while those of *M. indicum* and *M. bheemlicum* also include the analysis of the 16S rRNA gene as a phylogenetic marker [1,9]. The rearrangement of the family *Chromatiaceae* in

☆The nearly complete 16S rRNA sequence has been deposited at the GenBank databases under accession number EU377479.

*Corresponding author at: Zentrum für Umweltforschung und nachhaltige Technologien, Fachbereich Biologie/Chemie, Abteilung Marine Mikrobiologie, Universität Bremen, D-28359 Bremen, Germany. Fax: +49 421 2187222.
E-mail address: wserrano@mpi-bremen.de (W. Serrano).

0723-2020/$ - see front matter © 2008 Elsevier GmbH. All rights reserved.
doi:10.1016/j.syapm.2008.11.002

several genera and the designation of the new genus *Marichromatium* were based on comparative sequence analysis of the 16S rRNA gene [5]. However, the use of the 16S rRNA gene as phylogenetic marker has low resolution in defining closely related species [3], for that reason a DNA–DNA hybridization (DDH) experiment is still a standard method for clarifying the taxonomy of new isolates.

Here we describe a novel *Marichromatium* strain by using a polyphasic approach for which the name *M. gracile* biotype *thermosulfidiphilum* is proposed.

Materials and methods

Intertidal sediments were collected from the Sylt-Rømø Basin (55°02′26.08″N, 8°25′03.29″E, German Wadden Sea) in sterile polypropylene bottles. Pfennig's medium [15], supplemented with 4 mM sulfide, 2 mM thiosulfate, 2 mM acetate (final concentration each), and NaCl (3%, w/v), was used for phototrophic growth (220 μmol $m^{-2} s^{-1}$ photon flux) at 30 °C. Repetitive agar-shake dilution series according to [15] were applied to purify strain SW26. All kinds of phototrophic growth of strain SW26 were performed in a completely filled 15 ml Hungate tubes and estimated turbidometrically at 650 nm. *In vivo* absorption spectra were recorded with a Beckman DU 640 spectrophotometer in sucrose solution [15]. Morphological properties (cell shape, size, sulfur globules), flagella, and intracytoplasmic membrane system were studied by phase-contrast- or electron-microscopy. The use of alternative nitrogen sources were tested by replacing amonium chloride with urea, glutamate or glutamine (see Table 1). Diazotrophy of the culture was determined by growth under N_2 headspace and confirmed by repeated subculturing (three times). Utilization of different carbon compounds (4 mM or 0.1% w/v) and electron donors (2–4 mM) was examined in the medium described above. For sulfide tolerance experiments, complete anoxic mineral medium with acetate (2 mM) and different sulfide concentrations as electron donor at a final pH of 7.0–7.4 were prepared.

Genomic DNA for mol% G + C analysis and DDH experiments was extracted and purified according to [11]. The mol% G + C content was determined by the thermal denaturation method as described by Marmur and Doty [12]. Genomic DNA from *Escherichia coli* K12 (G + C content: 51 mol%) and lambda DNA (G + C content: 49.9 mol%) served as internal reference standards. Renaturation rate experiments were performed according to [2], using a Beckman DU 640 spectrophotometer equipped with a thermostated multicell changer and a temperature controller.

Almost complete 16S rRNA gene sequences (1400 bp) were amplified by PCR using general primers (GM3, GM4) and the analyses of nucleotide sequences were made according to [16]. The 16S rRNA gene sequence of strain SW26 was first compared to the most similar sequences available in the GenBank databases from NCBI. Only sequences longer than 1300 bp with <0.5% ambiguous nucleotides [17] from deposited members of *Chromatiaceae* were taken into account for the subsequent reconstruction of a phylogenetic tree by using MEGA 3.1 software package [8]. Multiple sequence alignment was carried out by using the CLUSTAL X program [18].

Results and discussion

The complete morphological and physiological characteristics of the strictly anaerobic phototrophic strain SW26 can be taken from Table 1, supplementary Figs. S1–S3, and the protologue. Strain SW26 shows all characteristics typical for members of the genus *Marichromatium*, additionally also supported by the value of 68.3 mol% G + C of its DNA composition.

However, when compared with reference strains, isolate SW26 clearly differs in some features such as (i) the presence of a small 4 kb plasmid (supplementary Fig. S5), which was not found in any other *Marichromatium* type strains examined; (ii) photoassimilation of glycolate, sucrose, and fructose, which are not utilized by the *M. gracile* type strain, while *M. purpuratum* and *M. indicum* only assimilate glucose or fructose, respectively; (iii) ability to grow at relatively high temperatures and sulfide concentrations (supplementary Figs. S1, S2).

So far known, thermotolerant marine anoxygenic phototrophs occur only among the *Ectothiorhodospiraceae* [6], while *Thermochromatium tepidum* (formerly: *Chromatium tepidum*), a member of the *Chromatiaceae*, with an upper growth limit of 55 °C, originated from a Yellowstone National Park freshwater hot spring [10]. Therefore, strain SW26 can be considered as the first moderate thermotolerant marine *Chromatiaceae* representative, since its growth temperature ranges from 25 to 44 °C with an optimum at 40 °C. The ability of strain SW26 to grow at sulfide concentrations up to 16 mM seems to be unique within the *Chromatiaceae* (supplementary Fig. S2). Cells occurred singly with two to three small intracellular sulfur globules when grown with 8 mM sulfide, but tend to form chains with larger sulfur globules in the presence of 16 mM sulfide (supplementary Fig. S1B, C). High sulfide tolerances have so far only been reported for green sulfur bacteria [13] and for the purple sulfur bacteria *M. indicum* (8 mM) [1] and *Thiorhococcus drewsii* (11 mM) [19].

Strain SW26 grew photolithoautotrophically with sulfide and CO_2, chemolithoheterotrophycally under microaerophilic conditions in the dark with sulfide (4 mM), thiosulfate (4 mM), and acetate (2 mM), but not chemo-organotrophically. Oxygen sensibility was

W. Serrano et al. / Systematic and Applied Microbiology 32 (2009) 1–7

Table 1. Comparison of morphological, physiological, and molecular characteristics of strain SW26 with its closest phylogenetic relatives of the genus *Marichromatium*.

Character	Strains				
	M. gracile DSM 203[T a]	*M. purpuratum* DSM 1591[T a]	*M. indicum* JA100[T a]	*M. bheemlicum* JA124[T a]	Strain SW26[b]
Cell shape	Rod	Rod	Rod	Rod	Rod
Cell size (width × length; μm)	1–1.3 × 2–6	1.2–1.7 × 3–4	1.2–1.7 × 3–4	0.8–1 × 2–4	1–1.2 × 2.5–4
Sulfur granules per cell	Two	Two	Single	Two	Three or more
Flagellation	Single polar	Single polar	Single polar	Single polar	Single polar
Rosette formation	Absent	Absent	Present	Absent	Absent
Color of cell suspension	Reddish-brown	Purple red	Reddish-brown	Reddish-brown	Reddish-pink
Carotenoids	Spirilloxanthin series	Okenone series	Spirilloxanthin series	Spirilloxanthin series	Spirilloxanthin series
NaCl range (%)	NR	NR	0.05–13	1.5–11	0.5–10
NaCl optimum (%)	1	5	1–4	1.5–8.5	3
Vitamin B_{12} requirement	–	–	–	–	–
Optimum growth temperature (°C)	25–30 (35)[d]	30–35 (35)	30–35 (35)	30–35 (40)	25–44
Optimum pH	6.5–7.6	7.2–7.6	6–7.5	7.0–7.6	6.5–7.8
Growth modes[c]	POH, PLA, COH, CLA, CLH	POH, PLA, COH, CLA, CLH	PLH, PLA, CLH, POH	PLH, PLA, POH (with cysteine)	PLH, PLA, POH
Nitrogen source utilized	Ammonium salts, N_2	Ammonium salts	NH_4Cl, urea, N_2	NH_4Cl, glutamine, glutamate	NH_4Cl, urea, N_2
Sulfide concentrations (mM)[e]	4	4	4	8	16
Organic substrates utilized					
Acetate	+	+	+	+	+
Propionate	+	+	+	–	+
Butyrate	+	+	+	–	–
Citrate	–	–	–	–	–
Succinate	+	+	–	+	+
Fumarate	+	+	–	+	+
Malate	+	+	+	+	+
Pyruvate	+	+	+	NR	+
Glycolate	–	–	–	NR	+

Table 1. (*continued*)

Character	Strains				
	M. gracile DSM 203[T a]	*M. purpuratum* DSM 1591[T a]	*M. indicum* JA100[T a]	*M. bheemlicum* JA124[T a]	Strain SW26[b]
Glucose	–	+	–	+	+
Lactate	+	+	+	–	+
Glycerol	–	–	–	+	–
Methanol	–	–	+	+	–
Ethanol	–	–	–	+	–
Molecular features					
DDH (%) relatedness[f]	81 ± 5.55	71 ± 3.07	63 ± 4.82	65 ± 0.91	100
16S rRNA sequence similarity (%)	99.86	99.58	99.30	98.13	100
Number of nucleotide differences	2	5	10	27	0
DNA G+C content (mol%)	70.4 (Tm)	68.9 (Tm)	67.1 (Tm)	67 (HPLC)	68.3 (Tm)
Plasmid presence[g]	–	–	–	–	+

NR = not reported.

[a]Data taken from [1,9,15].

[b]All organic substrates were tested during photolithoheterotrophic growth with $Na_2S \times 9H_2O$ (4 mM). Organic substrates were utilized at a final concentration of 4 mM, except for fatty acids (propionate, butyrate), ethanol, and glycerol which were used at a concentration of 0.1% v/v.

[c]CLA, chemolithoautotrophy; CLH, chemolithoheterotrophy; COH, chemo-organoheterotrophy; PLA, photololithoautotrophy; PLH, photolithoheterotrophy; POH, photo-organoheterotrophy.

[d]The maximum growth temperature for each type of strain determined in this study is shown in parentheses.

[e]Sulfide tolerance for each type of strain determined in this study.

[f]The reassociation values are the mean ± 1 SD of two independent determinations.

[g]Detected by gel electrophoresis (Fig. S5, supplementary online material).

demonstrated by photolithoheterotrophic growth under microaerophilic conditions (data not shown).

Intact cells showed absorption maxima at 371, 490, 590, 798, and 854 nm referring to bacteriochlorophyll *a*. Absorption maxima at 472 and 502 nm of methanol extracted pigments indicate the presence of carotenoids of the spirilloxanthin series.

The 16S rRNA gene sequence of strain SW26 had a similarity of 99.3% to that of *Marichromatium indicum* JA100^T, but showed unexpected low values of 96% and 94% similarities to *M. gracile* DSM 203^T and *M. purpuratum* DSM 1591^T. Based on re-sequencing of the 16S rRNA genes of the last mentioned organisms, we could correct these values to 99.86% and 99.58%, respectively, and have deposited both gene sequences determined in the GenBank databases under accession numbers EU850806 and EU850807. The reconstructed phylogenetic tree, based on the neighbour-joining method, shows that strain SW26 is arranged within the genus *Marichromatium* of the *Chromatiaceae* (Fig. 1).

Since the 16S rRNA gene sequence of the novel isolate SW26 has similarity values higher than 99% to that of *Marichromatium* type strains, a DDH study was made as recommended [4]. This yielded reassociation values of 81%, 71%, and 63% with the reference strains of *M. gracile, M. purpuratum,* and *M. indicum*, respectively (Table 1). The calculated DDH values also corresponded very well with the highest 16S rRNA gene sequence similarities of above 99% (Table 1). Since DDH relatedness values can only be reproduced in a range of about 10%, values around 70% must be considered as critical borderline to define species. Thus, the relative high degree of DDH relatedness still observed when comparing the species pair

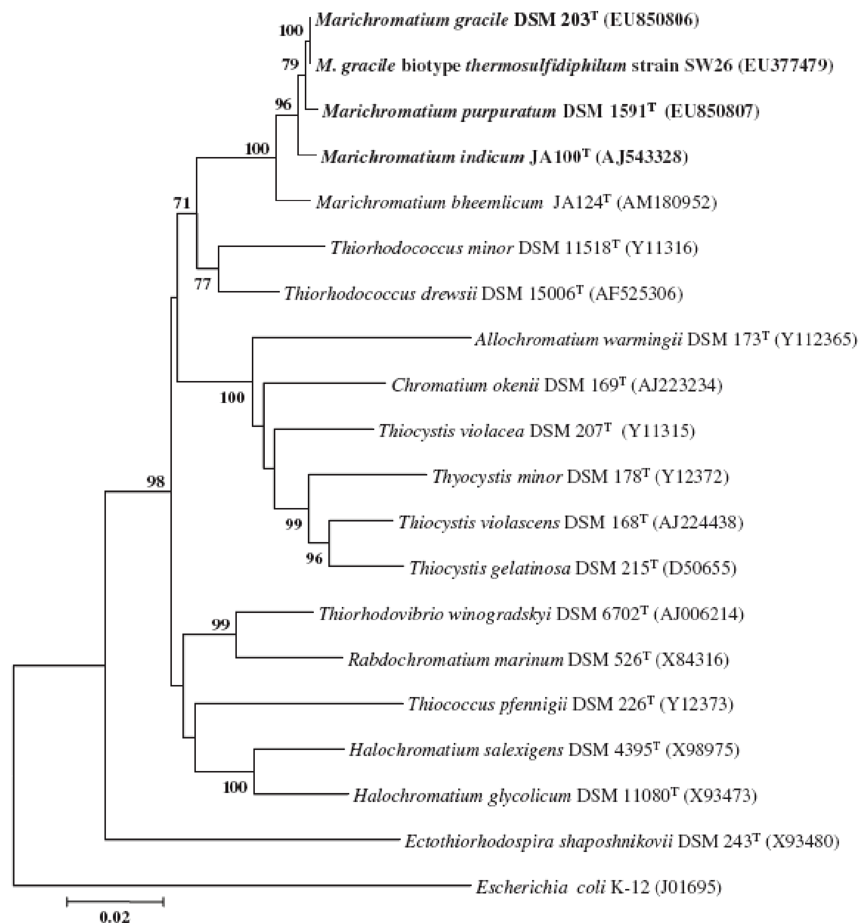


Fig. 1. Phylogenetic tree based on partial 16S rRNA (1400 bp) gene sequences. The constructed tree is the result of a distance matrix analysis. Similar affiliation of strain SW26 using the 16S rRNA gene sequences was observed when the tree was calculated using the maximum likelihood approach. The bar indicates 2% estimated sequence divergence. Accession numbers of the reference sequences are given in parentheses. Strains that share more than 99% of the 16S rRNA similarities are marked in bold. *Escherichia coli* K-12 was used as outgroup.

M. purpuratum vs. strain SW26 (71%) and their respective 16S rRNA sequence similarities above 99% could suggest that these species pairs have genomes that are closely related. We therefore argue that the high temperature and sulfide tolerance as well as the utilization of some carbon compounds and the presence of an extrachromosomal element clearly distinguish strain SW26 as a new biotype from the type strain *M. gracile*. The description of this new biotype additionally is supported by the fact that strain SW26 is readily differentiated from the type strain *M. gracile* by various PCR-based genomic fingerprinting (supplementary Table S1 and Fig. S4).

Description of *Marichromatium gracile* biotype *thermosulfidiphilum*

Marichromatium gracile biotype *thermosufidiphilum* (ther.mo.sul.fi.di'phi.lum Gr. n. *thermos* heat, N.L. neut. n. *sulfidum* sulfide, Gr. adj. *philos* friendly to, N. L. neut. adj. *thermosulfidiphilum* loving heat and sulfide).

Cells are Gram-negative, rod-shaped, 1–1.2 μm wide, and 2.5–4 μm long, and are motile by a polar flagellum. Chains are formed, when cells grow with high sulfide concentrations. The culture suspension is reddish-pink. Pigments include bacteriochlorophyll *a* and carotenoids of the spirilloxanthin series, and the photosynthetic membrane system is of the vesicular type. Growth occurs under strictly anaerobic conditions (photolithoautotrophy, photolithoheterotrophy, and chemolithoheterotrophy). No vitamins are required for growth. Sulfide and thiosulfate are electron donors, and elemental sulfur is formed as intermediate oxidation product and stored inside the cells. Organic substrates, photoassimilated in the presence of sulfide without carbonate supplement, are acetate, propionate, lactate, malate, succinate, fumarate, pyruvate, and glycolate. Additional substrates photoassimilated in the presence of sulfide plus carbonate are fructose and glucose. Strain SW26 is not able to utilize butyrate. Diazotrophic growth is possible. The strain is moderate thermophilic (temperature range 25–44 °C, optimum 40 °C), with a pH optimum at 6.8–7.6, requires NaCl (0.5–10%, optimum 3%) for growth, tolerates up to 16 mM sulfide, and has a 4 kb plasmid. The G+C content of the genomic DNA is 68.3 mol% (T_m). The new biotype strain is SW26 (= DSM 21765).

Habitat: Strain SW26 was isolated from an intertidal sandflat in the Sylt-Rømø Basin, North Sea, Germany.

Acknowledgements

We are grateful to Hans G. Trüper, Universität Bonn, Germany, for suggesting the specific epithet and Dr. Ch. Sasikala, Jawaharlal Nehru Technological University, India, for providing the original type strains of *M. indicum* JA 100^T and *M. bheemlicum* JA124^T. Support of the Bremen University and the Max Planck Society is gratefully acknowledged. W.S. was a member of the International Max Planck Research School of Marine Microbiology (MarMic).

Appendix A. Supplementary Materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.syapm.2008.11.002.

References

[1] K. Arunasri, Ch. Sasikala, V.C. Ramana, J. Süling, J.F. Imhoff, *Marichromatium indicum* sp. nov., a new purple sulfur *Gammaproteobacterium* from mangrove soil of Goa, India, Int. J. Syst. Evol. Microbiol. 55 (2005) 673–679.

[2] J. De Ley, H. Catoir, A. Reynaerts, The quantitative measurement of DNA hybridization from renaturation rates, Eur. J. Biochem. 12 (1970) 133–142.

[3] G.E. Fox, J.D. Wisotzkey, P. Jurtshuk, How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity, Int. J. Syst. Bacteriol. 42 (1992) 166–170.

[4] J.F. Imhoff, P. Caumette, Recommended standards for the description of new species of anoxygenic phototrophic bacteria, Int. J. Syst. Evol. Microbiol. 54 (2004) 1415–1421.

[5] J.F. Imhoff, J. Süling, R. Petri, Phylogenetic relationships among *Chromatiaceae*, their taxonomic reclassification and description of the new genera *Allochromatium*, *Halochromatium*, *Isochromatium*, *Marichromatium*, *Thiococcus*, *Thiohalocapsa* and *Thermochromatium*, Int. J. Syst. Bacteriol. 48 (1998) 1129–1143.

[6] J.F. Imhoff, H.G. Trüper, *Ectothiorhodospira halochloris* sp. nov., a new extremely halophilic phototrophic bacterium containing bacteriochlorophyll *b*, Arch. Microbiol. 114 (1977) 115–121.

[7] J.F. Imhoff, H.G. Trüper, *Chromatium purpuratum* sp. nov., a new species of the *Chromatiaceae*, Zentbl. Bakteriol. I Abt. Orig., Reihe C1 (1980) 61–69.

[8] S. Kumar, K. Tamura, M. Nei, MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment, Brief. Bioinform. 5 (2004) 150–163.

[9] P.A. Kumar, T.S.S. Jyothsna, T.N.R. Srinivas, Ch. Sasikala, Ch.V. Ramana, J.F. Imhoff, *Marichromatium bheemlicum* sp. nov., a non-diazotrophic, photosynthetic gammaproteobacterium from a marine aquaculture pond, Int. J. Syst. Evol. Microbiol. 57 (2007) 1261–1265.

[10] M.T. Madigan, *Chromatium tepidum* sp. nov., a thermophilic photosynthetic bacterium of the family *Chromatiaceae*, Int. J. Syst. Bacteriol. 36 (1986) 222–227.

[11] J. Marmur, A procedure for the isolation of desoxyribonucleic acid from micro-organisms, J. Mol. Biol. 3 (1961) 208–218.

[12] J. Marmur, P. Doty, Determination of the base composition of desoxyribonucleic acid from its thermal denaturation temperature, J. Mol. Biol. 5 (1962) 109–118.

[13] N. Pfennig, The phototrophic bacteria and their role in the sulfur cycle, Plant Soil 43 (1975) 1–16.

[14] N. Pfennig, H.G. Trüper, The family *Chromatiaceae*, in: J.T. Staley, M.P. Bryant, N. Pfennig, J.G. Holt (Eds.), Bergey's Manual of Determinative Bacteriology, vol. 3, Williams & Wilkins, Baltimore, 1989, pp. 1637–1653.

[15] N. Pfennig, H.G. Trüper, The family *Chromatiaceae*, in: A. Balows, H.G. Trüper, M. Dworkin, W. Harder, K.H. Schleifer (Eds.), The Prokaryotes, Springer, Berlin, Heidelberg, New York, 1992, pp. 3200–3221.

[16] F.S. Sanger, S. Niklen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors, Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

[17] E. Stackebrandt, W. Frederiksen, G.M. Garrity, P.A.D. Grimont, P. Kämpfer, M.C.J. Maiden, X. Nesme, R. Rosello-Mora, J. Swings, H.G. Trüper, L. Vauterin, A.C. Ward, W.B. Whitman, Report of the *ad hoc* committee for the re-evaluation of the species definition in bacteriology, Int. J. Syst. Evol. Microbiol. 52 (2002) 1043–1047.

[18] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876–4882.

[19] A. Zaar, G. Fuchs, J.R. Golecki, J. Overmann, A new purple sulfur bacterium isolated from a littoral microbial mat, *Thiorhodococcus drewsii*, sp. nov., Arch. Microbiol. 179 (2003) 174–183.

Supplementary tables and figures

Table S1: Primer sets used for PCR based fingerprinting profiling. PCR program was as follows: 5 min at 95°C for initial denaturation, 35 cycles (94°C, 1 min; annealing temperatures as indicated in the table, 1 min each, 72°C, 2 min elongation), except for primer BOX A1R in which the elongation time was 5 min at 65 °C. Banding profiles were determined by electrophoresis in a 1 % agarose gel.

Primer	Annealing temperature (°C)	Sequence 5'
BOX A1R	50	CTACGCCAAGGCGACGCTGACG
RAPD (SMO3)	36	CGGGTCGATC
M13-CORE	38	GAGGGTGGCGTTCT
ERIC 1	52	ATGTAAGCTCCTGGGGATTCAC
ERIC2	52	AAGTAAGTGACTGGGGTGAGCG

Table 2: Chemotrophic growth of strain SW26 and *Marichromatium* type strains as deep agar shake cultures.

Strains	Chemoautotrophic (a)		Chemoorganotrophic (a)		Chemolithoheterotrohic (a)		Photolithoheterotrophic (b)	
	cm (c)	growth (d)	cm (c)	growth (d)	cm (c)	growth (d)	cm (c)	growth (d)
SW26	-	-	-	-	1- 1.5	+ + +	7	+ + +
M. indicum	-	-	-	-	0.7- 1.2	+ + +	1.3	+ + +
M. gracile	3 - 4	+	0.3	+ (e)	1.2- 1.6	+ + +	1.5	+ + +
M. purpuratum	-	-	-	-	0.5- 1	+ + +	1.5	+ + +

(a) Chemoautotrophy: CO_2 plus sulfide (4mM, final concentration); chemoorganotrophy: mineral medium with only acetate (2mM, final concentration); chemolithoheterotrophy: sulfide (4mM), thiosulfate (4mM), and acetate (2mM) final concentrations, respectively, all experiments were conducted under microaerobic conditions in dark at 30°C.

(b) Same preparation as for chemolithoheterotrophic growth but incubated microaerobically in the light and at 30° C.

(c) Distance of the growth zone from surface in cm.

(d) Intensity of growth in the growth zone: - no growth, + weak, + + + intense growth, evidenced by coloured colony formation.

(e) Data taken from Kämpf and Pfennig (1980).

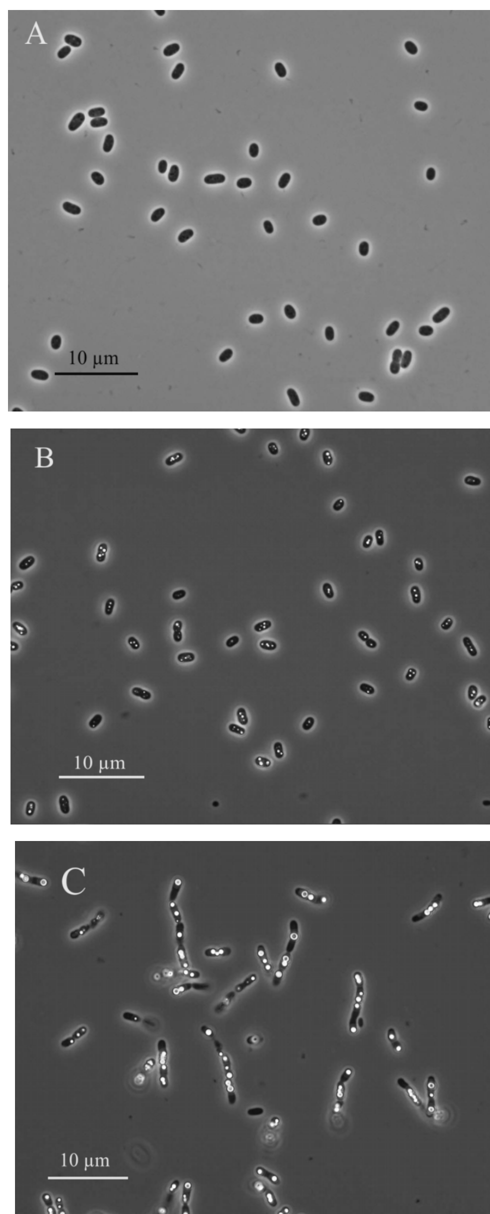


Figure S1: Phase contrast photomicrographs of strain SW26. (A) Sulfur free cells from stationary phase, (B) cells with sulfur globules from the exponential growth phase, grown with 8mM sulfide, (C) chain-shaped cells with large sulfur globules from exponential growth phase, grown with 16 mM sulfide.

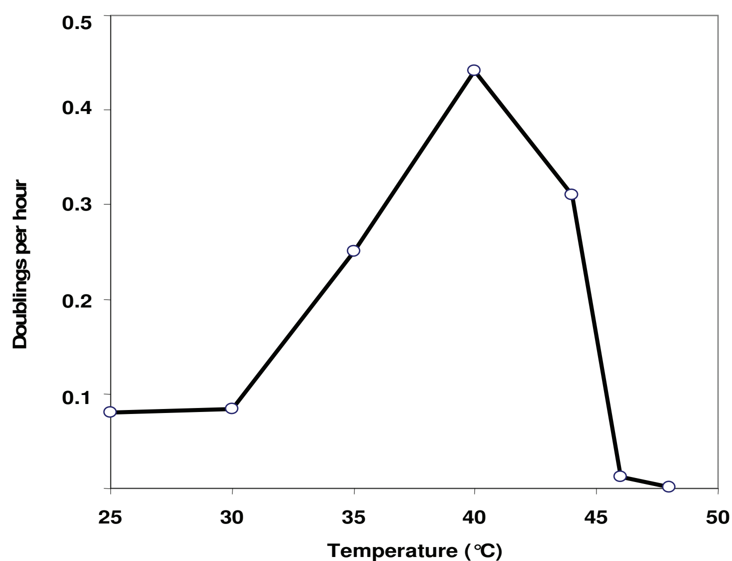


Figure S2: Temperature dependent growth behaviour (doubling per hour) of strain SW26.

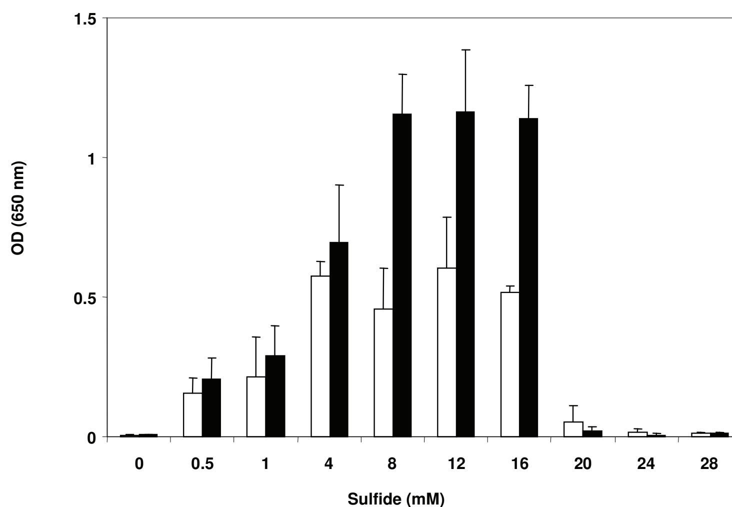


Figure S3: Sulfide and time dependent growth of strain SW26 at 40° C and 220 µE $m^{-2} s^{-1}$ photon flux. Incubation times: 20 h (white bars); 30 h (black bars). The error bars indicate one standard deviation

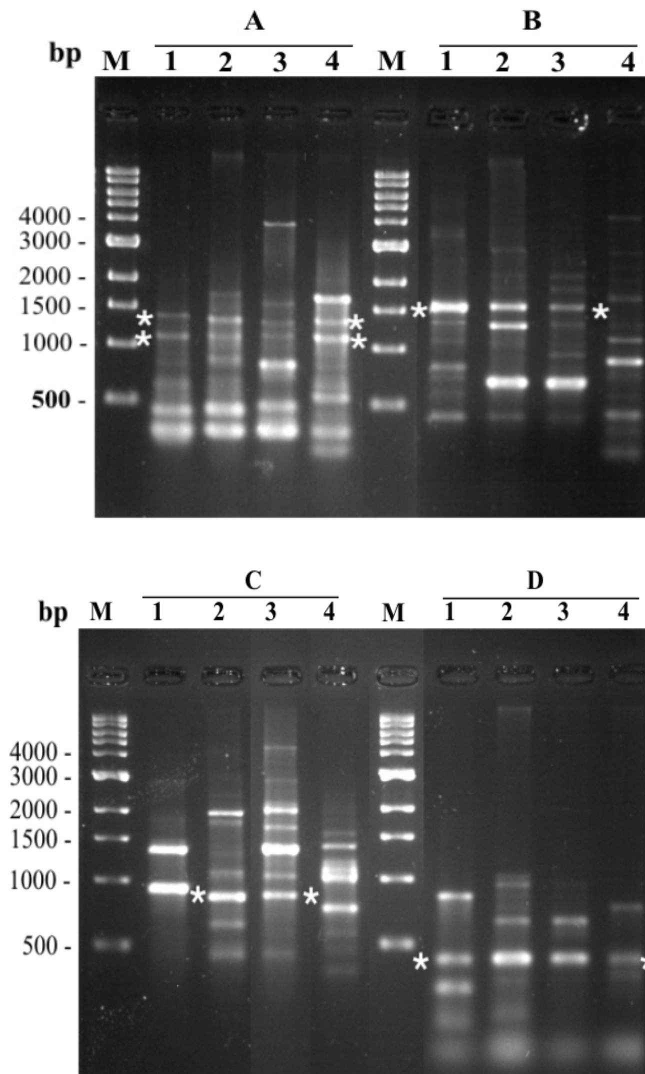


Figure S4: Variation in the fingerprinting pattern among four *Marichromatium* spp. generated by using different primer sets: BOX-A1R (**A**), SMO3 (**B**), M13 CORE (**C**), and ERIC 1, 2 (**D**), Lanes 1, *Marichromatium purpuratum* DSM 1591^T; lanes 2, M. gracile DSM 203^T; lanes 3, strain SW26; lanes 4, M. indicum $JA100^T$. Lines M, 1kb ladder. Asterisks show bands shared by the different strains.

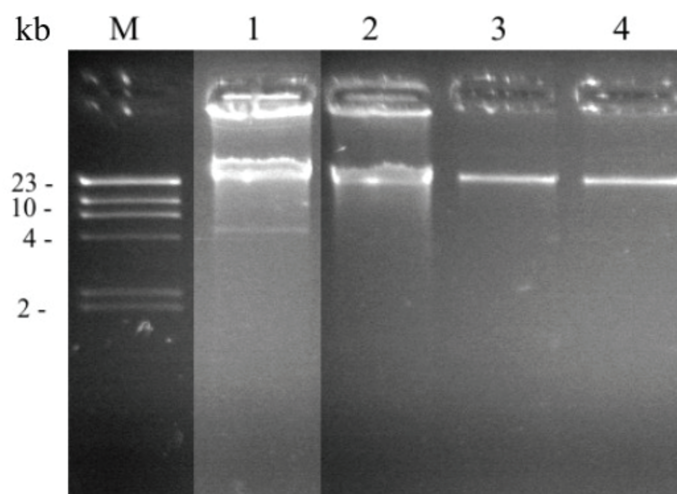


Figure S5: Agarose gel electrophoresis (1.5%) of genomic DNA from different *Marichromatium* type strains. Lane 1, strain SW26 containing a small plasmid (approximate size of 4 kb); lane 2, *M. purpuratum* DSM 1591^T; lane 3, *M. gracile* DSM 203^T; lane 4, *M. indicum* JA100T. Lane M, phage λ HindIII molecular weight marker.

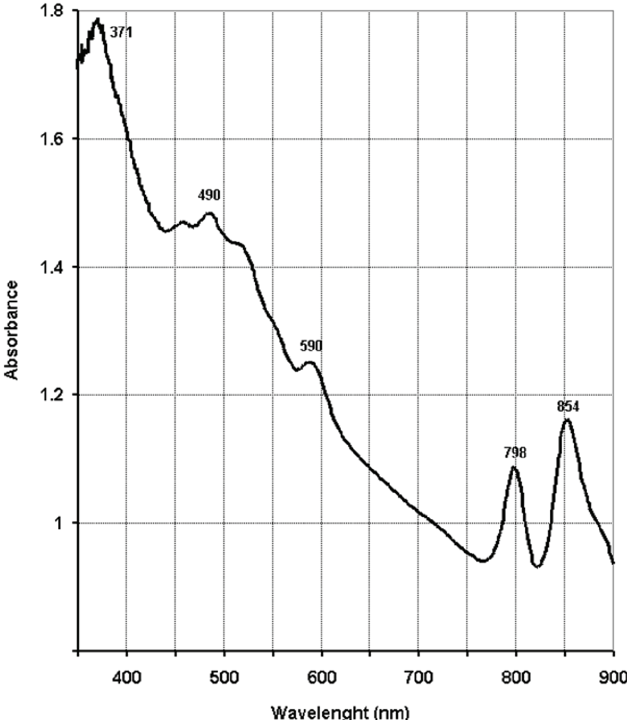


Figure S6 Whole cell absorption spectrum of a living cell suspension of strain SW26.

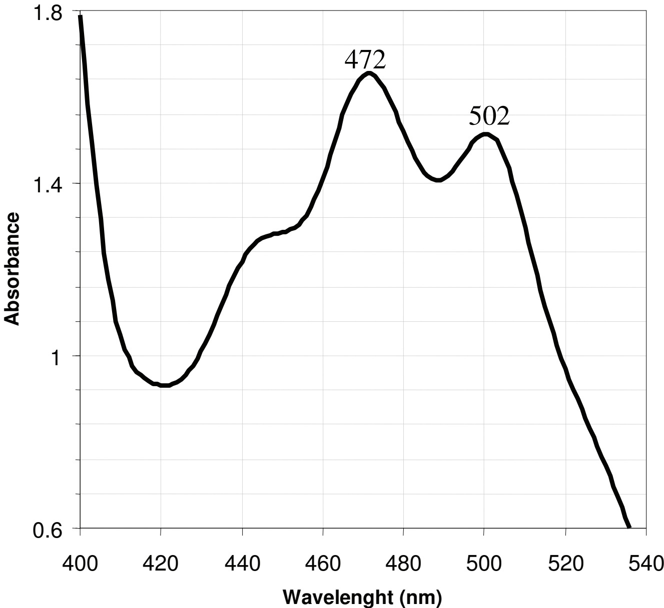


Figure S7. Acetone:Methanol spectra of extracted pigments of strain SW26.

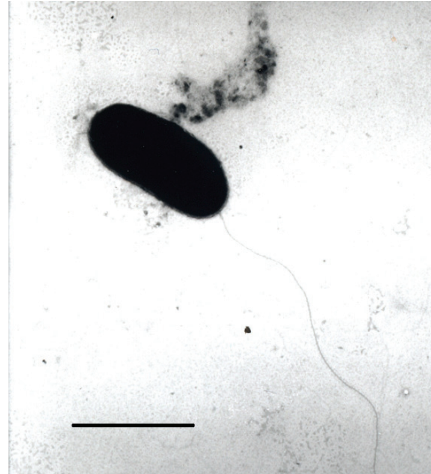


Figure S8: Electron microphotograph of a negative stained cell of strain SW26. Bar = 2 µm.

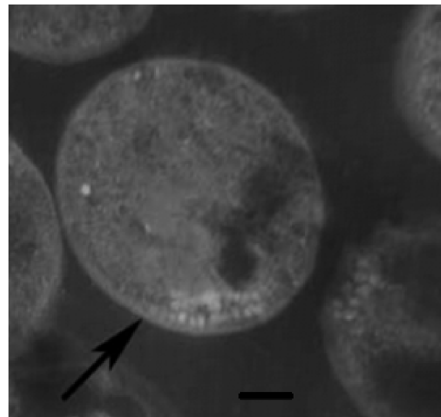


Figure S9: Electron micrograph of ultrathin-section of strain SW26. The arrow indicates the intracellular membrane of the vesicular type. Bar = 0.2 µm

Literature

Kämpf, C. and N. Pfennig. (1980). Capacity of *Chromatiaceae* for Chemotrophic Growth. Specific Respiration Rates of *Thiocystis violacea* and *Chromatium vinosum*. Arch. Microbiol. 127: 125-135.

Chapter 3

Evaluation of the use of multilocus sequence analysis (MLSA) to resolve taxonomic conflicts within the genus *Marichromatium*

Wilbert Serrano[1, 2, *], Rudolf Amann[2] Ramon Rossello-Mora[3] and Ulrich Fischer[1]

Accepted for publication in Systematics and Applied Microbiology

Abbreviations: TD-PCR, touchdown-PCR; DDH, DNA-DNA hybridization; MLSA, multilocus sequence analysis; (π), nucleotide diversity; (H), genetic diversity.

The GenBank/EMBL/DDBJ accession numbers for the *Marichromatium* strains sequences determined in this study are listed in Table S3.

[1] Zentrum für Umweltforschung und nachhaltige Technologien, Fachbereich Biologie/Chemie, Abteilung Marine Mikrobiologie, Universität Bremen, D-28359 Bremen, Germany
[2] Max Planck-Institute for marine Microbiology, Celsiusstraße 1, D-28359 Bremen, Germany
3 Institut Mediterrrani d'Estudis Avançats (CSIC-UIB), E-07190 Esporles, Mallorca, Spain.
* Corresponding author, email: wserrano@mpi-bremen.de

Abstract – Four species of marine purple sulfur bacteria of the genus *Marichromatium* have been validly described. A recent re-analysis of the 16S rRNA-based similarity and genomic DNA-DNA hybridizations (DDH) of the type strains (Serrano *et al.*, 2009 System. Appl. Microbiol. 32:1-7) suggested that some of them are so closely related that they can be considered heterotypic synonyms. Here, we report the evaluation of the multilocus sequence analysis approach (MLSA) of nine *Marichromatium* strains to resolve the intrageneric genealogical relationships. MLSA was based on six protein-coding genes (*gyrB, recA*, *fusA, dnaK, pufM,* and *soxB*), and the results were comparable to DDH. The phylogenetic tree constructed with the concatenated sequences, which also include the 16S rRNA gene and the internal transcriber spacer ITS region (4331 bp), separated the nine strains in four lineages reflecting the four *Marichromatium* species. The reconstructed phylogenetic tree based on concatenation of six protein-coding genes was also highly congruent with the tree topology based on the 16S rRNA gene.

Keywords

Anoxygenic photosynthetic bacteria, *Chromatiaceae, Marichromatium,* MLSA, DDH.

Introduction

The genus *Marichromatium* of anoxygenic photosynthetic bacteria was first established by Imhoff el al. [17] as an independent branch of the purple sulfur bacteria (PSB) of the family *Chromatiaceae*. To date, only four *Marichromatium* species have been validly described: *M. purpuratum*, *M. gracile*, *M. indicum*, and *M. bheemlicum*. The latter two were isolated from mangrove sediments and aquaculture ponds in India respectively [2, 23]. Species belonging to this genus appear to be phylogenetically and genomically closely related [33]. In addition, these species do not exhibit clear-cut phenotypic differences. Both the lack of key phenotypic traits and the low resolution of the 16S rRNA gene as phylogenetic marker, make the identification and classification of new isolates belonging to this genus difficult.

The 16S rRNA gene sequence comparison has been demonstrated to be poorly discriminative for closely related strains [7, 32]. In addition, the species definition for bacteria still relies on the use of DNA-DNA hybridization (DDH) as a main parameter to reveal genomic coherence. DDH studies have been recommended to be performed to circumscribe prokaryotic species whenever 16S rRNA gene sequence similarities lay above 97% [18]. Although genomic comparisons based on DDH offers high taxonomic resolution, the technique itself also has important drawbacks as it is time-consuming, labor-intensive and the results may vary depending of the DDH method employed.

In the case of two type strains *Marichromatium gracile* DSM 203^T and *Marichromatium purpuratum* DSM 1591^T DDH did not help in deciding whether the two strains belong to the same species. The genomic comparison based on DDH analysis showed a relatedness of 71% [33] and this value is near the commonly accepted boundary limit of 70% for species delineation [43]. However, this conflicts with the low degree of phenotypic differences found between both type strains. Altogether the results could indicate that these two *Marichromatium* species are actually heterotypic synonyms.

As an alternative to DDH, the comparative sequence analysis of protein-encoding genes had also been widely explored. For example, Takewaki et al. [37] used the *dnaJ* gene for establishing phylogenetic relationship among 19 *Mycobaterium* species, and more recently Wang et al. [42] used the *gyrB* gene for taxonomic analysis for the *Bacillus subtilis* group. However, the benefit of using a single gene as phylogenetic marker is debatable since genes in bacterial chromosomes can be

subject to horizontal gene transfer (HGT), duplication, and homologous recombination [6, 8, 28].

In this context, the analysis of several protein-coding genes, a technique known as Multilocus Sequence Analysis (MLSA) has been regarded as a promising method for bacterial systematics [35]. Simultaneous analysis of multiple gene loci, typically seven, may help to reduce bias caused by the use of a single gene [13, 34]. In addition, the use of fast evolving protein-coding genes may offer some advantages over the highly conserved and poor species-level discriminative 16S rRNA gene for resolving intra-and intergeneric relationships among closely related bacterial species [9, 19]. Furthermore, the translated amino acid sequences can be used in evolutionary studies. For example, this method has supported the identification and species circumscription of closely related strains of *Vibrio* species. [39, 40, 41] and for some *Burkholderia* species [10].

Since type strains of the genus *Marichromatium* seem to be closely related both genomically (DDH) and phylogenetically (16S rRNA), here we report a first standard MLSA scheme for the identification, and as complementary tool for species circumscription, within the genus *Marichromatium*. The present study is based on the analysis of nine *Marichromatium* strains. It covers almost all strains available from the DSMZ culture collection. Besides, the MLSA approach was used to identify two new isolates, and emend the phylogenetic placement of strain *Allochromatium vinosum* DSM 185.

Material and Methods

Strains and sources.

The strains used in this study and details concerning geographical coordinates are listed in Table S1 (supplementary material). We also included in this study the strain *Allochromatium vinosum* DSM 185. In the context of this study, this strain was identified as member of the genus *Marichromatium* by its 16S rRNA gene sequence. Therefore, we designate this isolate as *Marichromatium sp*. strain WS-2008.

Cultivation.

All examined strains were cultivated in anoxic Pfennig medium [29]. Cultures were purified by repeated passages through agar shake dilution series [29]. The resulting single colonies were cultured in fresh anoxic Pfennig's medium with phototrophic growth conditions according to [29, 33].

DNA isolation and molecular analyses.

Chromosomal DNA was isolated following the method of Marmur [25]. The G+C determination for all strains was made by the method of Marmur & Doty [26]. Genomic DNA to be used for DNA-DNA reassociation experiments was diluted to a concentration of 0.4 mg/ml in 0.1X SSC and 1 ml aliquots were sheared by sonication. Fragment sizes, determined by elecrophoresis on 1% agarose gel, should be in the range of 200 to 800 bp as maximum. Denaturation was performed filling 1.5 ml glass vials with 500 µl of single and equally mixed DNA samples in a boiling water bath for at least 15 min. DNA-DNA hybridization was performed following the renaturation kinetic method described by [4] using a Beckman DU 640 model UV spectrophotometer equipped with a thermostated multicell changer and a temperature controller.

Comparative 16S rRNA gene analysis.

Phylogenetic affiliation of all examined strains in this study was first checked by comparative 16S rRNA gene sequence analysis using the method outlined in the previous work [33].

MLSA

Primer sequences used for amplification of universally present housekeeping genes and metabolic genes specific for the purple sulfur bacteria are listed in supplementary Table S2. For PCR amplifications genomic DNA was diluted to a final concentration of 30 ng/µl; PCR parameters used in this study are listed in Table S2. Six genes encoding photosynthetic reaction centre M (*pufM*), sulfate thioesterase thiohydrolase (*soxB*), translation elongation factor (*fusA*), recombinase A (*recA)*, chaperonin HSP 60 (*dnaK*)*,* and DNA gyrase subunit B (*gyrB*) as well as the 16S rRNA gene, and the 16S-23S internal transcriber spacer (ITS) were sequenced. Since specific primers and parameters for amplification of the genes *fusA*, *recA, gyrB, and dnaK* for purple sulfur bacteria were not available, touchdown-PCR (TD-PCR) amplification using degenerate primers was done according to [5, 15]. Shortly, TD-PCR parameters were: initial denaturation at 95°C (4 min), 10 cycles of denaturation at 95° (1 min), annealing temperature progressively lowered from 60°C to 50°C by 1°C every cycle (1 min), elongation step at 72°C (2 min), followed by 15 additional cycles with annealing temperature at 50°C, plus 10 min of final elongation time at 72°C. The PCR reaction was made in a final volume of 25 µl containing: 100

pmol of each primer (1 µl), 1 mM of total dNTPs (2.5 µl), 1 x PCR buffer (2.5 µl), and 1 U (0.1 µl) of taq-polymerase. After reaction, the size of the PCR product was checked by electrophoresis on 1% agarose. PCR products showing a single clear band were purified with QIA quick purification kit (Qiagen, Hilden, Germany) and sequenced using the ABI PRISM BigDye sequencing kit (Applied biosystems, Foster City, CA). Prior to sequencing, the resulting PCR cycle sequencing products were washed by passage through water-swelled Sephadex G-50 columns (Amersham Bioscience, Upsala Sweden) and collected by centrifugation on Nunclon 96 micro well plates (Nunclon™). Sequencing was made using an ABI PRISM 3100 genetic analyser (Applied Biosystems). Accession numbers of all sequences reported in this study are listed in Table S3 (supplementary online material).

Multilocus phylogenetic comparison

The sequences were aligned with Clustal X [38]. The unmatched ends of each gene sequence alignment were deleted in order to create a homogeneous matrix and then translated into their respective amino acid sequence by using the program BioEdit [12]. The presence of hypervariable regions within each aligned gene sequences were evaluated with the program Gblocks [3]. Sequence alignments are presented as supplementary material.

Phylogenies for each of the six single protein-coding gene trees based on the N-J method were made by using the program MEGA 3.1 [22] using the JTT model of protein evolution [20]. The same individual and the concatenated gene sequences that also included the 16S rRNA gene and ITS region sequences (totalizing 3685 bp) were also analysed by maximum likelihood method using PhyML [11] with the substitution model HKY85. The robustness of the resulting tree topologies was evaluated by 100 bootstrap replications.

In order to evaluate the bootstrap support as a function of the concatenated sequences, sets of 3 to 7 genes were randomly chosen and concatenated. For each concatenated reconstructed tree, the support of the clade was calculated by 100 bootstraps with the program PhyML [11].

In addition, reticulate network trees for the concatenated sequences were constructed by split decomposition analysis as implemented in the software Splits Tree version 4.0 [16] using Hamming distances.

Results

16S rRNA gene and ITS region sequences comparison

A total of nine *Marichromatium* strains, including four type strains and two new isolates were analyzed in this study. As can be seen in Table 1, the two additional new strains and the wrongly placed strain WS-2008 (formerly *Allochomatium vinosum* DSM 185) included in this study showed high 16S rRNA gene sequence similarities ranging from 99.5-100% to those of the type strains *M. gracile* DSM 203^T, and *M. purpuratum* DSM 1591^T. The 16S rRNA sequence similarities for all examined *Marichromatium* strains ranged from 98-100% (mean 99.3%) (Table 1 + S4). Sequence similarities of the 16S-23S ITS regions were in the range of 91.8-100% (mean 96.4%) (Supplementary Table S4).

DDH comparison

DDH values among strains of *Marichromatium* originating from Europe and South America ranged from 69-98% (Table 1), revealing that these strains have also closely related genomes. Lower values ranging from 45-79% were observed when the two type strains from India, *M. indicum* JA100^T and *M. bheemlicum* JA124^T, were hybridized with other *Marichromatium* strains (Table 1). Inter-strains DDH values above 70% agreed very well with corresponding 16S rRNA sequence similarity values >99% with only few exceptions, and DDH values below 70% were in general in concordance with the 16S rRNA gene sequence similarities below 99%.

Multi-gene comparison

Six protein-coding genes were analysed independently. Nucleotide sequences similarities of the genes *gyrB* and *fusA* ranged between 89.5-99.8% (mean 95.1%) and 90.0-100% (mean 94.5%), respectively (Supplementary Table S4). Amino acid sequence identities (AAIs) of gyrB and fusA were between 93.1-100% (mean 97.3%) and 88.1-100% (mean 93.8%), respectively. The genes *pufM*, *recA*, *soxB,* and *dnaK* yielded on average nucleotide sequence similarities of 96.6%, 97.2%, 97.6%, and 98.6% and AAIs of 99.4%, 99.8%, 99.2%, and 98.6%, respectively. The average nucleotide sequence similarities for all protein-coding genes were lower than those obtained for the 16S rRNA gene sequences (Supplementary Table S4).

Table 1. Cross-comparison of nine *Marichromatium* strains based on their DDH relatedness, 16S rRNA similarity, concatenated gene sequence similarity and amino acid identity (AAI).

	(1) DSM 1591^T	(2) DSM 1711	(3) DSM 203^T	(4) DSM 21765	(5) Strain CLA37	(6) Strain MP1	(7) Strain WS-2008	(8) JA100^T	(9) JA124^T
(1) DSM 1591^T		97 ± 2.91	75 ± 1.64	71 ± 3.07	78 ± 0.87	75 ± 4.86	72 ± 1.01	65 ± 6.16	57 ± 4.45
(2) DSM 1711	**99.85** 98.97 99.09		85 ± 1.65	86 ± 1.44	69 ± 0.47	70 ± 5.42	ND	77 ± 2.04	49 ± 2.21
(3) DSM 203^T	**99.55** 97.08 98.52	**99.70** 96.93 99.29		81 ± 5.55	98 ± 1.87	80 ± 3.16	77 ± 2.49	76 ± 2.97	60 ± 6.84
4) DSM 21765	**99.55** 97.16 98.63	**99.70** 97.01 98.41	**100** 99.77 99.88		74 ± 2.35	81 ± 0.53	72 ± 1.71	63 ± 4.82	65 ± 0.91
(5) Strain CLA37	**99.55** 97.31 98.86	**99.70** 97.05 98.63	**100** 99.35 99.65	**100** 99.43 99.77		83 ± 4.83	74 ± 3.66	72 ± 3.33	59 ± 6.38
(6) Strain MP1	**99.47** 97.20 98.29	**99.62** 96.78 97.61	**99.92** 98.26 98.86	**99.92** 98.33 98.97	**99.92** 98.29 98.97		76 ± 2.47	63 ± 3.03	45 ± 3.14
(7) Strain WS-2008	**99.47** 96.97 98.63	**99.62** 96.93 98.63	**99.92** 98.86 99.43	**99.92** 98.94 99.54	**99.92** 99.13 99.54	**99.85** 98.18 98.75		79 ± 2.66	47 ± 1.86
(8) JA100^T	**99.40** 95.20 97.73	**99.55** 95.08 97.05	**99.25** 95.65 97.39	**99.25** 95.80 97.50	**99.25** 95.80 97.73	**99.17** 95.95 97.16	**99.17** 95.69 97.50		66 ± 2.21
(9) JA124^T	98.13 95.20 96.93	98.13 94.63 96.25	98.13 95.72 96.93	98.13 95.72 96.93	98.13 95.61 97.16	98.05 95.91 97.05	98.05 95.35 96.71	98.20 94.33 96.48	

The values on the upper right side are DDH relatedness (%) determined in twice independent experiments ± 1 SD. The values on the lower left side are the correspondent 16S rRNA sequence similarities (%), upper row (pairwise similaritiy >99% are shown in bold); concatenated nucleotide sequence similarity, middle row; corresponding amino acid sequence identity, lower row. ND = Not determined. Asterisks indicate pair strains showing high percentages of DDH relatedness with the corresponding 16S rRNA similarity, concatenated sequence similarity and AAI.

Multi-gene phylogeny

The phylogenetic trees constructed with the almost full-length sequences of the 16S rRNA and ITS showed that the strains isolated from India clustered separately from the strains isolated from Europe and South America (Supplementary Fig. S1+S2). The same topology was found for the phylogenetic trees generated with the *gyrB* and *recA* gene sequences (Supplementary Fig. S3+S4). However, conflicting topologies were observed when the trees were constructed using the genes,*fusA, dnaK, pufM*, and *soxB.* The phylogenetic tree based on the *pufM* gene has shown that the two okenone-containing strains, *M. purpuratum* DSM 1591^T and DSM 1711, cluster separately from those *Marichromatium* strains that show spirilloxanthin as main pigment (Supplementary Fig. S5). In contrast, both strains appear in separate clusters when the tree is constructed with the *soxB* gene (Supplementary Fig. S6). On the other hand, the position of type strain *M. bheemlicum* JA124^T seems to be unstable considering that it clusters together with *M. gracile* strains in the *soxB* gene-based tree (Fig. S6), whereas the same strain grouped together with *M. purpuratum* strains in the *fusA* gene-constructed tree (Fig. S7). On the other hand, *M. indicum* JA100^T appears within the cluster formed by both *M. purpuratum* strains and the spirilloxanthin containing *M. gracile* strain WS-2008 in the *dnaK* gene-based tree (Fig. S8).

In MLSA the conflicting phylogenies of individual genes are overcome by the analysis of concatenated sequences. For this purpose, we concatenated the nucleotide sequences of the six protein-coding genes with the 16S rRNA and the ITS region sequences resulting in a total of 4331 bp. Comparative sequence analysis of the concatenated sequences resulted in the allocation of the two *M. purpuratum* strains DSM 1591^T and DSM 1711in a separate cluster (designed here as *Marichromatium purpuratum* phylogroup II) which is supported by a bootstrap value of 100%. Another cluster referred to as *Marichromatium gracile* phylogroup I is formed by the spirilloxanthin-containing strains *M. gracile* DSM 203^T, *M. gracile* DSM 21765, the two new strains CLA37 and MP1, and WS-2008 (formerly *A. vinosum* DSM 185). This cluster is supported by a high bootstrap value of 90%. The two *Marichromatium* type strains JA100^T and JA124^T originating from India form two independent lineages that are well separated from phylogroups I and II (Fig. 1 and S9).

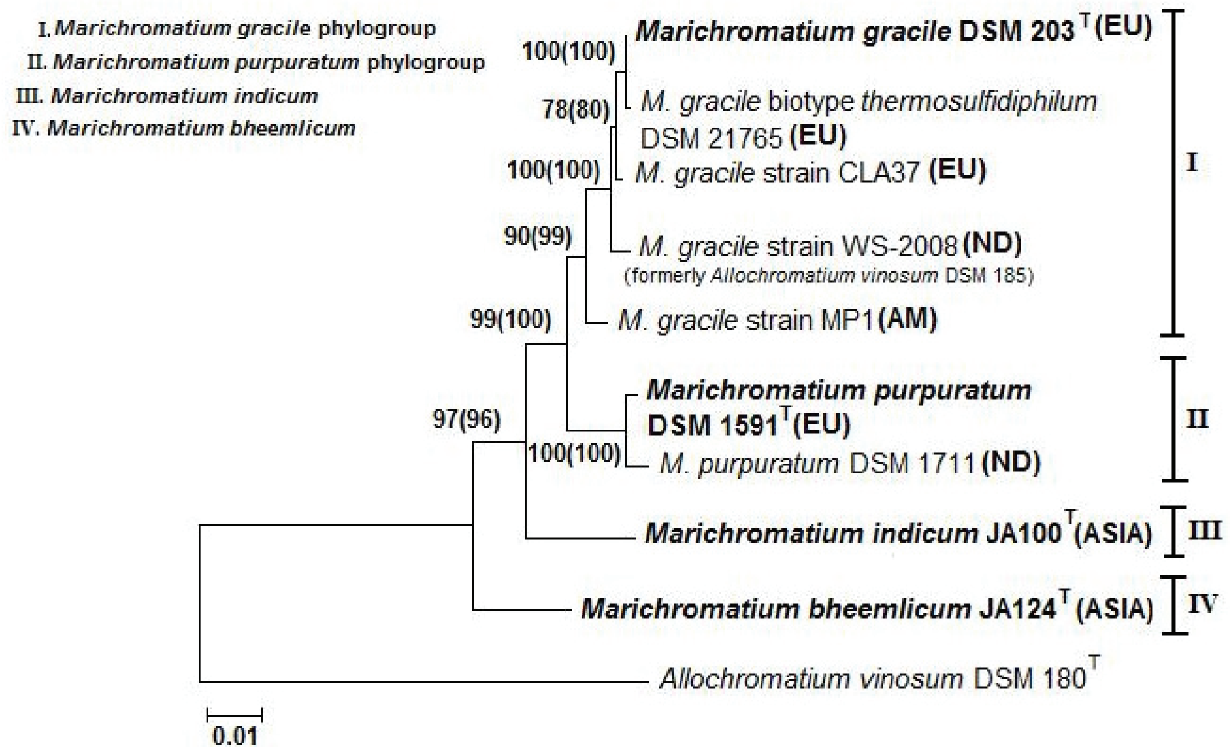


Figure 1. Phylogenetic tree reconstructed based on the maximum likelihood method using the concatenated nucleotide sequences (4331 bp) of the 16S rRNA, *gyrB*, *recA, fusA, dnaK, pufM, soxB,* and ITS region of nine *Marichromatium* strains. Distance estimation was obtained by the model of HKY as implemented in the program PhyML and distance method according to Jukes and Cantor [21]. Bootstrap support in percentage was based on 100 replications (JC distance method in parentheses). Bar indicates 1% of estimated sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup. (EU), strains originated from Europe; (AM), strain originated from South America; (ASIA), type strains originated from India; (ND), no data. Geographical coordinates are indicated in Table S1.

The topologies of phylogenetic trees reconstructed considering only the six protein-coding genes (2646 bp) (Supplementary Fig. S10) and the translated corresponding amino acid sequences (data not shown) were consistent with the topologies of the phylogenetic trees based on the 16S rRNA gene and the ITS region sequences. In all cases, random combination of genes tested showed that phylogroups I and II cluster separately from each other, while type species originating from India appear as independent branches far separated from both phylogroups. The same feature was observed when genes producing conflicting topologies (e.g *fusA, dnaK, pufM*, and *soxB*) are concatenated. For example, the type strain *M. bheemlicum* JA124^T appeared to form a basal branch within cluster I (data not shown). However, this apparent incongruence was solved with the inclusion of one of the following genes:

16S rRNA, *gyrB, recA*, or the ITS sequence. The influence of each gene and the ITS sequence over the final tree topology was evaluated by comparing the average bootstrap value of seven nodes from trees constructed with different gene combinations. The inclusion of only one gene (e.g. *recA*, 16S rRNA or *gyrB*) produced an increasing average bootstrap value of 59, 76 and 88%, respectively, while the inclusion of the ITS sequence alone produced an average value of 78%. Similar to a previous study [34], the stability of tree topologies increased in parallel with the number of concatenated genes (Fig. S11). In our case, a set of seven genes guaranteed a stable topology, which is in agreement with the minimal number of genes recommended for MLSA [9, 24].

Discussion

The slow evolutionary rate and relative lack of highly resolving informative characters of the 16S rRNA gene has been well documented [19, 31]. Due to these inherent structural characteristics phylogenetic trees reconstruction based on 16S rRNA sequences can not resolve the taxonomic position of closely related bacterial species.

The present study addresses the genus *Marichromatium* in which all strains except the type strain *M. bheemlicum* JA124^T share >99% of 16S rRNA sequence similarity. Therefore, they constitute what has been described as a microdiverse ribotype, i.e., a cluster with less than 1% of sequence divergence [1]. In this context, we have carried out a multi-gene based phylogenetic analysis. This was even more important since among the examined *Marichromatium* strains remarkable phenotypic differences were only found for a newly reported strain of *M. gracile* biotype *thermosulfidiphilum* DSM 21765 [33]. Notably, this new biotype exhibited a higher optimum growth temperature of 40°C and a higher sulfide tolerance of up to 16 mM than the other *Marichromatium* type strains [33]. However, in spite of its unusual phenotype, also this new biotype shares a 16S rRNA gene sequence similarity of 99.8% and DDH value of 81% with the type strain *M. gracile* DSM 203^T (Table 1).

The results of the present MLSA strongly support the established taxonomy of the genus *Marichromatium*. The four lineages indicated by MLSA are in perfect

agreement with the four validly established species *M. gracile*, *M. purpuratum*, M. *indicum,* and *M. bheemlicum* [2, 17, 23].

The use of MLSA has been recommended for the re-evaluation of well-studied species and for the identification of closely related strains [14]. In this study, the two new marine isolates CLA37 and MP1 could unambiguously be identified as strains of *M. gracile*. The source of strain CLA37 was the sediment of a coastal lagoon connected to the Adriatic Sea by a channel, whereas MP1 was retrieved from mangrove sediments of the northern coast of Perú. As nowadays indicated in the genus name, most strains of the genus *Marichromatium* seem to be of marine origin although the first strain described – *Marichromatium gracile*, formerly *Chromatium gracile* Streszewski 1913 - was obtained from a freshwater pond in Cracow [36].

Our study confirms that MLSA can be useful in bacterial taxonomy as complementary tool for defining species and for identification of new strains [30]. As other MLSA studies [27] our comparative analyses of concatenated sequences of strains of the genus *Marichromatium* have improved the confidence level of species designation. Moreover, new isolates of *Marichromatium* can now be easily identified based on sequencing of a relative small number of protein-coding genes; since this approach refers to sequences we submitted to public databases the repeated isolation of DNA from multiple reference strains required for DDH can be avoided.

We have here successfully applied MLSA to the genus *Marichromatium*. We propose that species within this genus still could be defined as groups of strains that share >97.5% similarity of the concatenated nucleotide sequences of *gyrB*, *recA*, *fusA*, *dnaK*, *pufM*, *soxB,* ITS, and 16S rRNA and >99% amino acid identity in the concatenated pufM, soxB, fusA, recA, dnaK, and gyrB. Based on this suggestion type strains *M. purpuratum* DSM 1591^{T} and *M. gracile* DSM 203^{T} could still maintain their separate species status although they share DDH values of >70%.

Acknowledgments

We are grateful to Dr. Ch. Sasikala from Jawaharlan Nehru Technological University, India for providing the original type strains of *M. indicum* JA100^T and *M. bheemlicum* JA124^T, Ms. Clelia Dona from the Department of Marine Microbiology of the University of Bremen for providing the *Marichromatium gracile* strain CLA37, and Daniel Santillano for proof reading the manuscript. Financial support of the University of Bremen and the Max Planck Society is gratefully acknowledged. RRM acknowledges the Ministerio de Ciencias e Innovación for the research support in the projects CE-CSD2007-0005 and CLG2006-12714-C02-02. WS was a PhD student of the International Max Planck Research School of Marine Microbiology (MarMic).

References

[1] S. G. Acinas, V. Klepac-Ceraj, D. E. Hunt, C. Pharino, I. Ceraj, D. L. Distel, M. F. Polz, Fine-scale phylogenetic architecture of a complex bacterial community, Nature 430 (2004) 551-554.

[2] K. Arunasri, Ch. Sasikala, V. C. Ramana, J. Süling, J. F. Imhoff, *Marichromatium indicum* sp. nov., a new purple sulfur *Gammaproteobacterium* from mangrove soil of Goa, India, Int. J. Syst. Evol. Microbiol. 55 (2005) 673-679.

[3] J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis, Mol. Biol. Evol. 17 (2000) 540-552.

[4] J. De Ley, H. Catoir, A. Reynaerts, The quantitative measurement of DNA hybridization from renaturation rates, Eur. J. Biochem. 12 (1970) 133-142.

[5] R. H. Don, P. T. Cox, B. J. Wainwright, K. Baker, J. S. Mattick, Touchdown PCR to circumvent spurious priming during gene amplification, Nucleic Acid Res 19 (1991) 4008.

[6] E. J. Feil, E. C. Holmes, D. E. Bessen, M. S. Chan, N. P. Day, M. C. Enright, R. Goldstein, D. W. Hood, A. Kalia, C. E. Moore, J. Zhou, B. G. Spratt, Recombination within natural populations of pathogenic bacteria: Short-term empirical estimates and long-term phylogenetic consequences, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 182-187.

[7] G. E. Fox, J. D. Wisotzkey, P. Jurtshuk, How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity, Int. J. Syst. Bacteriol. 42 (1992) 166-170.

[8] C. Fraser, W. P. Hanage, B. G. Spratt, Recombination and the nature of bacterial speciation, Science 315 (2007) 476-480.

[9] D. Gevers, F. M. Cohan, J. F. Lawrence, B. G. Spratt, T. Coenye, E. J. Feil, E. Stackebrandt, Y. Van de Peer, P. Vandamme, F. L. Thompson, J. Swings, Re-evaluating prokaryotic species, Nature Rev. 3 (2005) 733-739.

[10] M. B. Glass, A. G. Steigertwalt, J. G. Jordan, P. P. Wilkins, J. E. Gee, *Burkholderia oklahomensis* sp. nov., a *Burkholderia pseudomallei*-like species formerly known as the Oklahoma strains of *Pseudomonas pseudomallei,* Int. J. Syst. Evol. Microbiol. 56 (2006) 2171-2176.

[11] S. Guindon, O. Gascuel, A simple, fast and accurate algorithm to estimate large phylogenies by maximal likelihood, Syst. Biol. 52 (2003) 696-704.

[12] T. A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp. Ser. 41 (1999) 95-98

[13] W. P. Hanage, C. Fraser, B. Spratt, Fuzzy species among recombinogenic bacteria, BMC soil. 3 (2005) 6.

[14] W. P. Hanage, C. Fraser, B. Spratt, Sequences, sequence clusters and bacterial species, Phil. Trans. R. Soc. B. 361 (2006) 1917-1927.

[15] X. Q. Huang, S. Cloutier, Hemi-nested touchdown PCR combined with primer-template mismatch PCR for rapid isolation and sequencing of low molecular weight glutenin subunit gene family from a hexaploid wheat BAC library, BMC genetics 8 (2007) 18.

[16] D. H. Huson, D. Bryant, Application of phylogenetic networks in evolutionary studies, Mol. Biol. Evol. 23 (2006) 254-267.

[17] J. F. Imhoff, J. Süling, R. Petri, Phylogenetic relationships among *Chromatiaceae*, their taxonomic reclassification and description of the new genera *Allochromatium, Halochromatium, Isochromatium, Marichromatium, Thiococcus, Thiohalocapsa and Thermochromatium,* Int. J. Syst. Bacteriol. 48 (1998) 1129-1143.

[18] J. F. Imhoff, P. Caumette, Recommended standards for the description of new species of anoxygenic phototrophic bacteria, Int. J. Syst. Evol. Microbiol. 54 (2004) 1415-1421.

[19] P. R. Jensen, C. Mafnas, Biogeography of the marine actinomycete *Salinospora*. Environ. Microbiol, 8 (2006) 1881-1888.

[20] D. T. Jones, W. R. Taylor, J. M. Thornton, The rapid generation of mutation data matrices from protein sequences, Comput. Appl. Biosci. 8 (1992) 275-282.

[21] T. H. Jukes, C. R. Cantor, Evolution of protein molecules, in: H. N. Munro (Ed.), Mammalian Protein metabolism , Academic Press, New York, 1969, pp. 21-132.

[22] S. Kumar, K. Tamura, M. Nei, MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment, Brief. Bioinform. 5 (2004) 150-163.

[23] P. A. Kumar, T. S. Jyothsna, T. N. R. Srinivas, Ch. Sasikala, Ch. V. Ramana, J. F. Imhoff, *Marichromatium bheemlicum* sp. nov., a non-diazotrophic, photosynthetic gammaproteobacterium from a marine aquaculture pond, Int. J. Syst. Evol. Microbiol. 57 (2007) 1261-1265.

[24] M. C. Maiden, J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, B. G. Spratt, Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms, Proc. Natl. Acad. Sci. USA 95 (1998) 3140-3145.

[25] J. Marmur, A procedure for the isolation of desoxyribonucleic acid from micro-organisms, J. Mol. Biol. 3 (1961) 208-218.

[26] J. Marmur, P. Doty, Determination of the base composition of desoxyribonucleic acid from its thermal denaturation temperature, J. Mol. Biol. 5 (1962) 109-118.

[27] S. Mignard, J.-P Flandrois, A seven-gene, multilocus, genus-wide approach to the phylogeny of mycobacteria using supertrees, Int J. Syst. Evol. Microbiol. 58 (2008) 1432-1441.

[28] H. Ochman, J. G. Lawrence, E. A. Groisman, Lateral gene transfer and the nature of bacterial innovation, Nature 405 (2000) 299-305.

[29] N. Pfennig, H. G. Trüper, The family *Chromatiaceae*, in: A. Ballows, H. G. Trüper, M. Dworkin, W. Harder, K. H. Schleifer (eds.), The Prokaryotes, Springer, Berlin, Heidelberg, New York, 1992 pp. 3200-3221.

[30] D. Richter, D. Postic, N. Sertour, I. Livey, F.-R. Matuschka, G. Baranton, Delineation of *Borrelia burgdorferi sensu lato* species by multiclocus sequence analysis and confirmation of the delineation of *Borrelia spielmanii* sp. nov, Int. J. syst. Evol. Microbiol. 56 (2006) 873-881.

[31] T. Rogall, J. Wolters, T. Flohr, E. C. Böttger, Towards a phylogeny and definition of species at the molecular level within the genus *Mycobacterium,* Int. J. Syst. Bacteriol. 40 (1990) 323-330.

[32] R. Rosselló-Mora, R. Amann, The species concept for prokaryotes, FEMS Microbiol. Rev. 25 (2001) 39-67.

[33] W. Serrano, R. Amann, U. Fischer, A new moderately thermophilic and high sulfide tolerant biotype of *Marichromatium gracile*, isolated from tidal sediments of the German Wadden Sea: *Marichromatium gracile* biotype *thermosulfidiphilum,* Syst. Appl. Microbiol. 32 (2009) 1-7.

[34] V. Soria-Carrasco, M. Valens-Vadell, A. Peña, J. Antón, R. Amann, J. Castresana, R. Rosselló-Mora, Phylogenetic position of *Salinibacter ruber* based on concatenated protein alignments, Syst. Appl. Microbiol. 30 (2007) 171-179..

[35] E. Stackebrandt, W. Frederiksen, G. M. Garrity, P. A. D. Grimont, P. Kämpfer, M. C. Maiden, X. Nesme, R. Rosselló-Mora, J. Swing, H. G. Trüper, L. Vauterin, A. C. Ward, W. B. Whitman, Report of the *Ad Hoc* committee for the re-evaluation of the species definition in bacteriology, Int. J. Syst. Evol. Microbiol. 52 (2002) 1043-1047.

[36] B. Strzeszewski, Beiträge zur Kenntnis der Schwefelflora und der Umgebung von Krakau. Bull. Acad. Sci. Cracovie Ser. B (1913) 309-334.

[37] S. Takewaki, K. Okusumi, I. Manabe, M. Tanimura, K. Miyamura, K. Nakahara, Y. Yazaki, A. Ohkubo, R. Nagai, Nucleotide sequence comparison of the mycobacterial *dnaJ* gene and PCR-restriction fragment length polymorphism analysis for identification of mycobacterial species, Int. J. Syt. Bacteriol. 44 (1994) 159-166.

[38] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, D. G. Higgins, The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876-4882.

[39] F. L. Thompson, D. Gevers, C. C. Thompson, P. Dawyndt, S. Naser, B. Hoste, C. B. Munn, J. Swings, Phylogeny and molecular identification of Vibrios on the basis of multilocus sequence analysis, Appl. Environ. Microbiol. 71 (2005) 5107-5115.

[40] F. L. Thompson, B. Gomez-Gil, A. T. Ribeiro-Vasconcelos, T. Sawabe, Multilocus sequence analysis reveals that *Vibrio harveyi* and *V. campbellii* are distinct species, Appl. Environ. Microbiol. 73 (2007) 4279-4285.

[41] C. C. Thompson, F. L. Thompson, A. C. P. Vicente, Identification of *Vibrio cholerae* and *Vibrio mimicus* by multilocus sequence analysis (MLSA), Int. J. Syst. Evol. Microbiol. 58 (2008) 617-621.

[42] L.-T. Wang, F.-L. Lee, C.-J. Tai, H. Kasai, Comparison of *gyrB* gene sequences, 16S rRNA gene sequences and DNA-DNA hybridization in the *Bacillus subtilis* group, Int. J. Syst. Evol. Microbiol. 57 (2007) 1846-1850.

[43] L. G. Wayne, D. J. Brenner, R. R. Colwell, P. A. D. Grimont, O. Kandler, L. Krichevsky, L. H. Moore, W. C. Moore, R. G. E. Murray, E. Stackebrandt, M. P. Starr, H. G. Trüper, Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematics, Int. J. Syst. Bacteriol. 37 (1987) 463-464.

Supplementary tables

Table S1. Geographical coordinates of the *Marichromatium* strains used in this study

Locations	Strain/species name	Latitude	Longitude
Europe [a]	*Marichromatium gracile* DSM 203^{T}*	NR	NR
Europe [b]	*Marichromatium purpuratum* DSM 1591^{T}*	NR	NR
Europe [c]	*Marichromatium purpuratum* DSM 1711*	NR	NR
Europe [d]	*M. gracile* biotype *thermosulfidiphilum* DSM 21765	55°02'N	08°25'E
Europe [e]	*Marichromatium gracile* strain CLA37	45°41'N	13°07'E
Europe [f]	*Marichromatium gracile* strain WS-2008*	NR	NR
Asia [g]	*Marichromatium indicum* JA100^{T}	NR	NR
Asia [h]	*Marichromatium bheemlicum* JA124^{T}	17°51'N	83°25'E
America [i]	*Marichromatium gracile* strain MP1	03°29'S	80°23'W

[a] Originally isolated from streams around Cracow, Poland [36]. [b] Originally isolated near Split harbour, Adriatic sea, Croatia [2]. [c] Unknown origin. [d] Isolated from sediments samples from North Wadden Sea, Germany. [e] Isolated from sediments samples from Marano lagoon, northern Italy.
[f] Unknown origin, DSMZ deposited as *Allochromatium vinosum* DSM 185 = NCBI data base the 16S rDNA sequence is deposited as *Marichromatium sp.* strain WS-2008. [g] Originally isolated from mangroves sediments, nearby Goa, India. [h] Isolated from an aquaculture pond, India. [i] Original isolate from mangrove sediments northern coast of Perú.
* Strains obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ).

Table S2. Primer sequences used for the PCR amplification and sequencing of eight loci used in this study.

Locus	Probably function of gene (or protein)	Sequenced amplicon size (bp)	Primer	Primer sequence (5'-3') [a]	Annealing temperature [b]	Reference
16S rRNA	Protein synthesis	1340	GM3	AGAGTTTGATCMTGGC	44	Muyzer et al. [4]
			GM4	TACCTTGTTACGACTT	44	Muyzer et al. [4]
ITS	rRNA intergenic spacer	345	ITSf	GTGAAGTCGTAACAAGG	44	This study
			ITSr	CGTCACTTGACCATATCC	44	This study
pufM	Reaction center M	225	pufM.557F	CGCACCTGGACTGGA	55	Achenbach et al. [1]
			pufM.750R	CCCATGGTCCAGCGCCAGAA	55	Achenbach et al. [1]
soxB	Sulfate thioesterase/sulfate thiohydrolase	411	soxBF	GACGGTGGTGATACCTG	55	Hensen et al. [2]
			soxBR	CATGTCGCCGCCCTGCTG	55	Hensen et al. [2]
			soxinR	GACGTCCATGCCGTTGT	50	This study
			soxinF	ACAACGGCATGGACGTC	50	This study
recA	Recombinase A	366	recABDUP1	CCCGAGTCCTCCGGNAARACNAC [a]	TCHDW (52)	Santos and Ochman [5]
			recABGDN2	CGTTGCCGCCGGKNGTNRYYTC [a]	TCHDW (52)	Santos and Ochman [5]
			recintAF	ACCGGYGAGCAGGCGCT	55	This study
			recintAR	AGCGCCTGCTCRCCGGT	55	This study
fusA	Translation elongation factor G	330	fusAF	CATCGGCATCATGGCNCAYATHGA [a]	TCHDW (50)	Santos and Ochman [5]
			fusAR	CAGCATCGGCTGCAYNCCYTTRTT [a]	TCHDW (50)	Santos and Ochman [5]
			fusinF	GTCAACAAGATGGACCGCATG	55	This study
			fusinR	CATGCGGTCCATCTTGTTGAC	55	This study
dnaK	Chaperonin, HSP 60	741	KF1	ATYGGNATYGAYCTNGGNAC	55	Segal and Ron [6]
			KR1	GCNACNACYTCRTCNGGRTT	55	Segal and Ron [6]
			INT1	ATCATCGCCTACACCAACG	55	This study
			INT2	TGGCCACCGACCAGGAT	55	This study
gyrB	DNA gyrase subunit B	573	gyrBBAUP2	GCGGAAGCGGCCNGSNATGTA [a]	TCHDW (50)	Santos and Ochman [5]
			gyrBBndn1	CCGTCCACGTCGGCRTCNGYCAT [a]	TCHDW (50)	Santos and Ochman [5]

[a] Degenerate primer sequences are underlined, these primers were amplified by touchdown PCR.
[b] same annealing temperature for amplification and sequencing, in parentheses sequencing annealing temperature when amplification was made by touchdown PCR.

Table S3. Strains numbers and GenBank sequence accession number.

Marichromatium strains	GenBank accession number							
	16S rRNA	ITS	*pufM*	*soxB*	*fusA*	*recA*	*dnaK*	*gyrB*
M. purpuratum DSM 1591^T	EU850807	FJ376570	FJ376395	FJ376409	FJ376421	FJ376422	FJ376559	FJ164071
M. purpuratum DSM 1711	FJ376391	FJ376571	FJ376397	FJ376406	FJ376417	FJ376423	FJ376558	FJ376576
M. gracile DSM 203^T	EU850806	FJ376574	FJ376402	FJ376410	FJ376415	FJ376426	FJ376563	FJ164073
M. gracile biotype *thermosulfidiphilum*	EU377479	FJ376573	FJ376398	FJ376412	FJ376416	FJ376424	FJ376561	FJ164072
M. gracile strain CLA37	FJ376391	FJ376575	FJ376399	FJ376408	FJ376414	FJ376425	FJ376562	FJ376577
M. gracile strain WS-2008	GU289907	FJ376567	FJ376400	FJ376405	FJ376413	FJ376427	FJ376466	GU289908
M. indicum JA 100^T	AJ543328	FJ376569	FJ376401	FJ376407	FJ376420	FJ376430	FJ376560	FJ164075
M. bheemlicum JA 124^T	AM180952	FJ376568	FJ376402	FJ376411	FJ376419	FJ376428	FJ376565	FJ164074
M. gracile strains MP1	FJ376393	FJ376572	FJ376403	FJ376404	FJ376418	FJ376429	FJ376564	FJ376578

Table S4. Range and average sequence similarity and amino acid identity of the six protein-coding genes, the 16S rRNA gene and ITS region sequences.

Genes /ITS region	Sequence similarity (%)	Average similarity (%)	Amino acid identity (AAI) %	Average (AAI) %
pufM	92.4-100	96.6	98.6-100	99.4
soxB	95.3-100	97.6	96.3-100	99.2
fusA	90-100	94.5	88.1-100	93.8
recA	93.4-100	97.2	99.1-100	99.8
dnaK	96.8-99.7	98.6	96.3-100	98.6
gyrB	89.5-99.8	95.1	93.1-100	97.3
16S rRNA	98-100	99.3	NA	NA
ITS	91.8-100	96.4	NA	NA

NA= no applicable

Supplementary figures

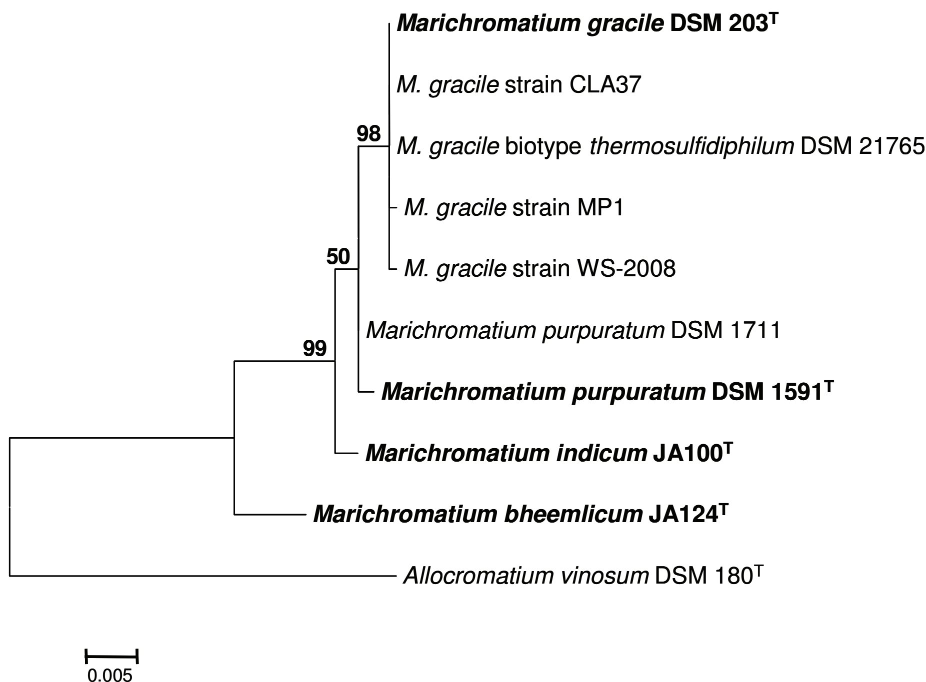


Figure S1. Maximum likelihood based phylogenetic tree using 16S rRNA gene sequences. Distances were estimated by the model HKY as implemented in the program PhyML. Bootstrap percentages (≥50) after 100 simulations are shown. Bar indicates 0.5% of sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup.

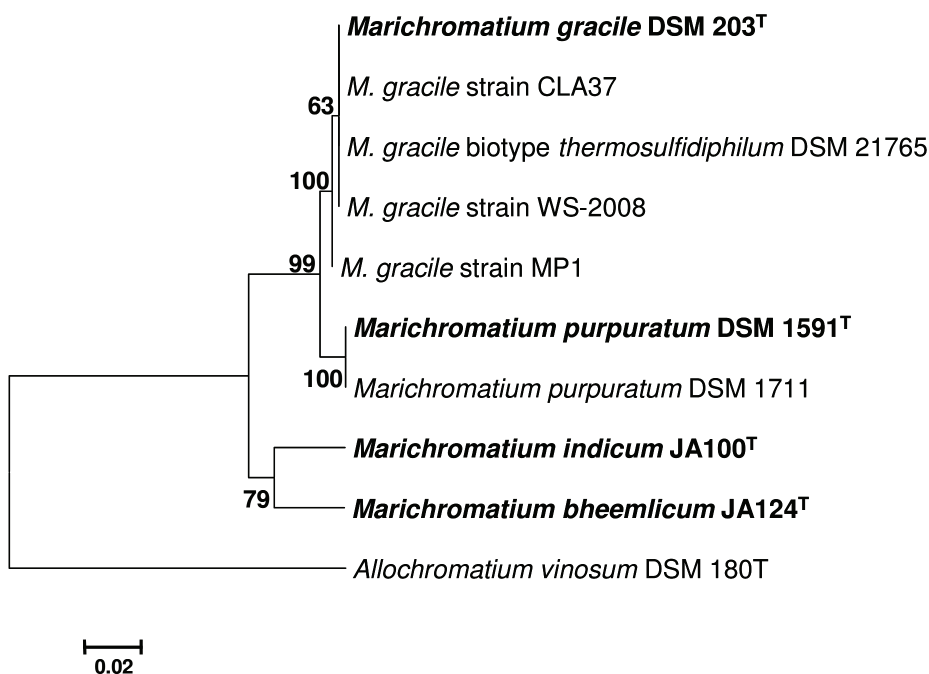


Figure S2. Maximum likelihood based phylogenetic tree using internal transcriber spacer (ITS) region sequences. Distances were estimated by the model HKY as implemented in the program PhyML. Bootstrap percentages (>50) after 100 simulations are shown. Bar indicates 2% of sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup.

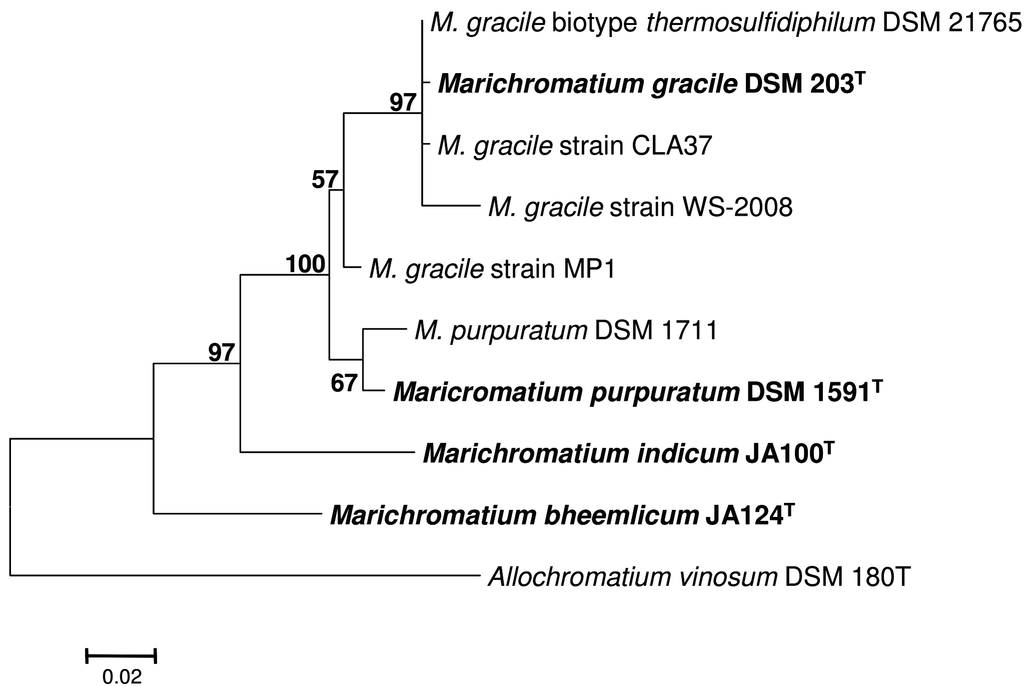


Figure S3. Maximum likelihood based phylogenetic tree using the DNA gyrase B (*gyrB*) gene sequences. Distances were estimated with the JTT model of protein evolution as implemented in the program Mega 3.1. Bootstrap percentages (>50) after 100 simulations are shown. Bar indicates 2% of sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup.

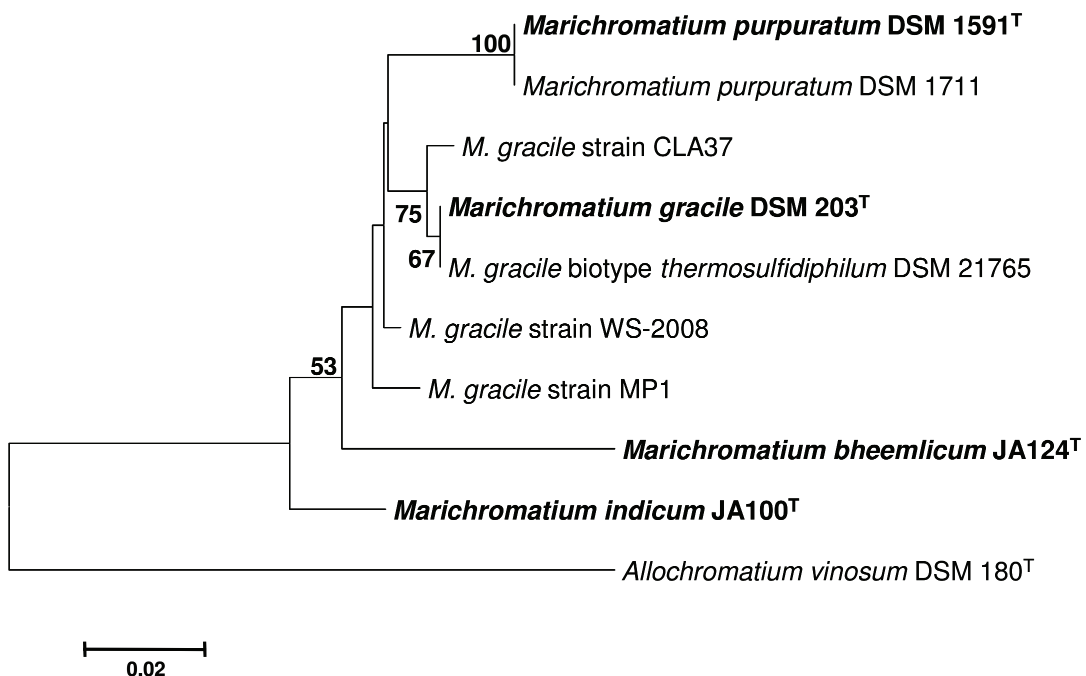


Figure S4. Maximum likelihood based phylogenetic tree using the recombinase A (*recA*) gene sequences. Distances were estimated with the JTT model of protein evolution as implemented in the program Mega 3.1. Bootstrap percentages (>50) after 100 simulations are shown. Bar indicates 2% of sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup.

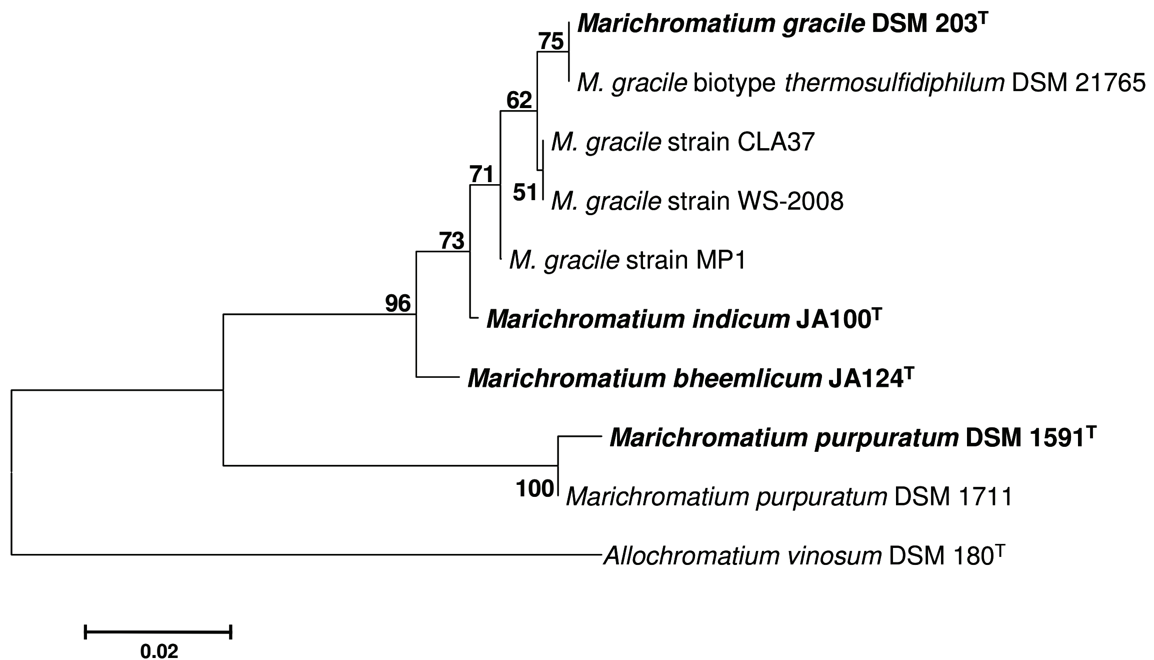


Figure S5. N-J based phylogenetic tree using reaction center *pufM* gene sequences. Distances were estimated with the JTT model of protein evolution as implemented in the program Mega 3.1. Bootstrap percentages (>50) after 100 simulations are shown. Bar indicates 2% of sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup.

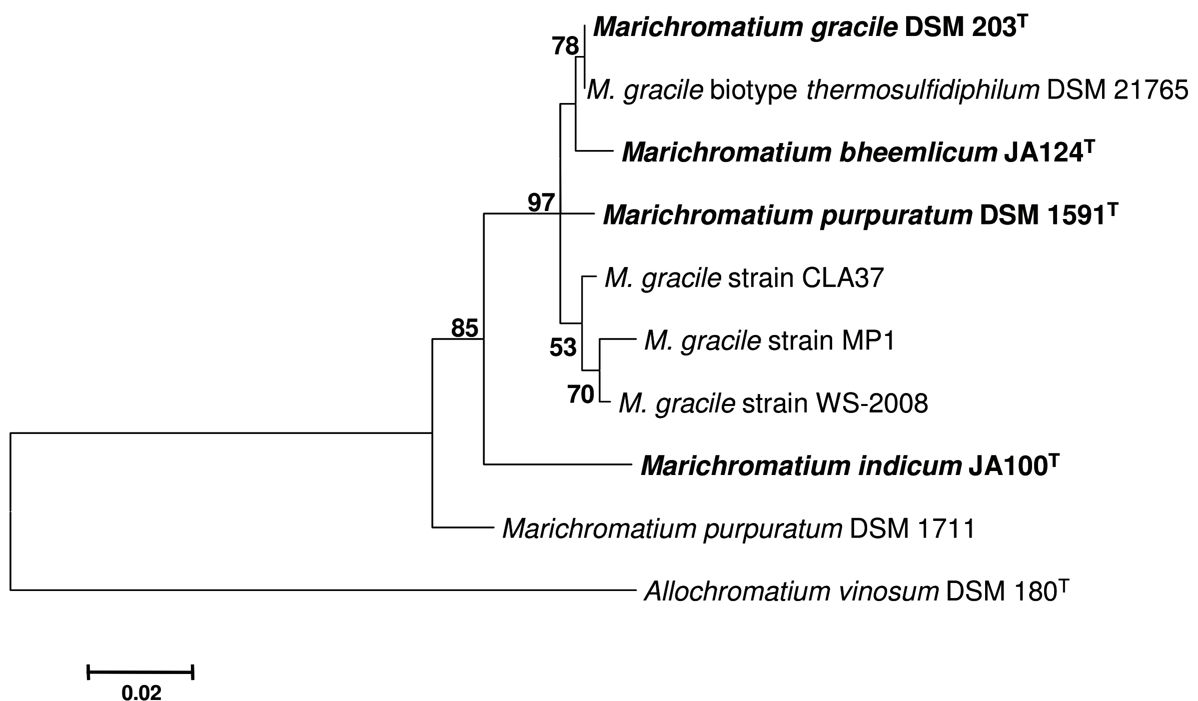


Figure S6. N-J based phylogenetic tree using the sulphate thioesterase/thiohydrolase *soxB* gene sequences. Distances were estimated with the JTT model of protein evolution as implemented in the program Mega 3.1. Bootstrap percentages (>50) after 100 simulations are shown. Bar indicates 2% of sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup.

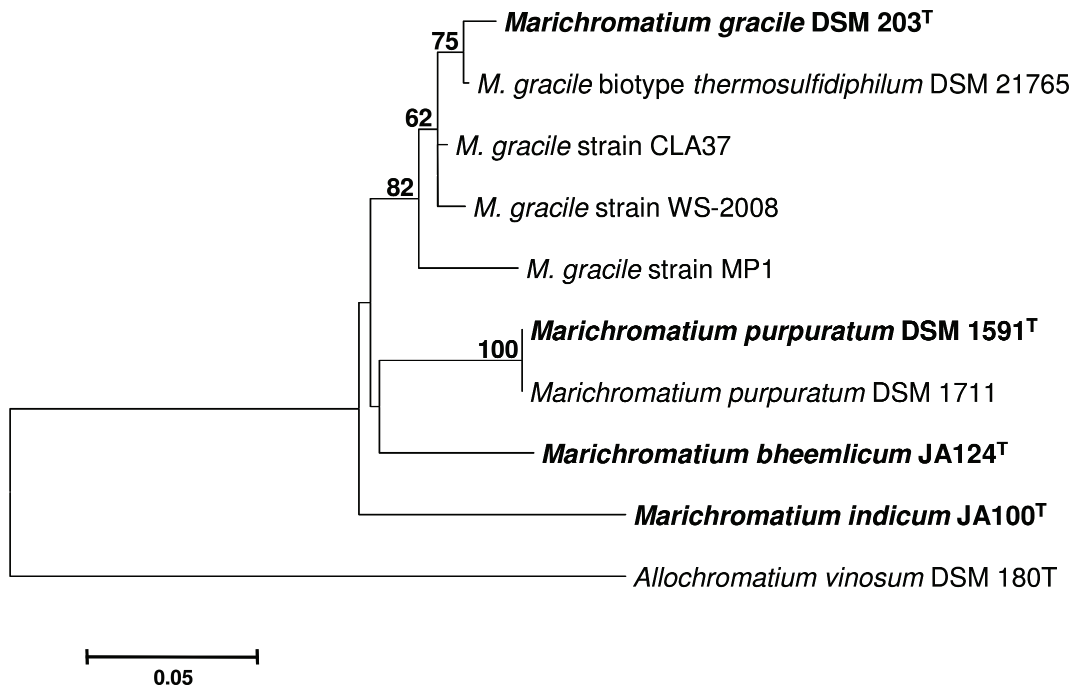


Figure S7. N-J based phylogenetic tree using the translation elongation factor G (*fusA*) gene sequences. Distances were estimated with the JTT model of protein evolution as implemented in the program Mega 3.1. Bootstrap percentages (>50) after 100 simulations are shown. Bar indicates 5% of sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup

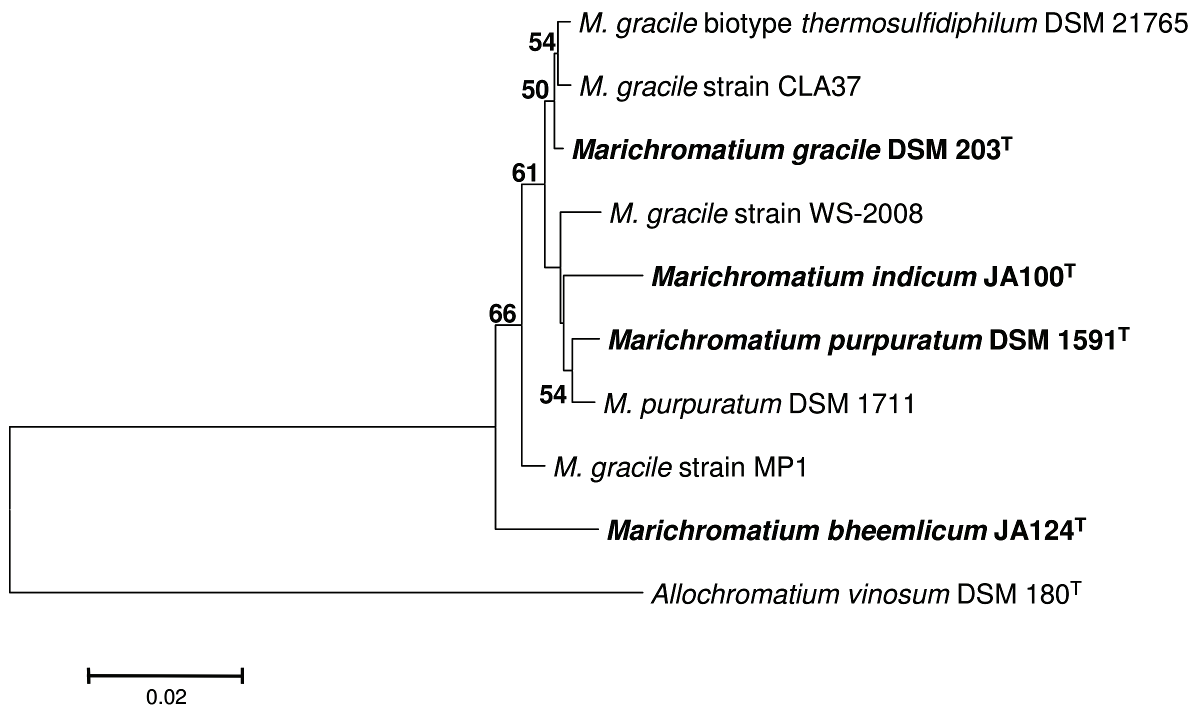


Figure S8. N-J based phylogenetic tree using the chaperonin, HSP60 (*dnaK*) gene sequences. Distances were estimated with the JTT model of protein evolution as implemented in the program Mega 3.1. Bootstrap percentages (>50) after 100 simulations are shown. Bar indicates 2% of sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup.

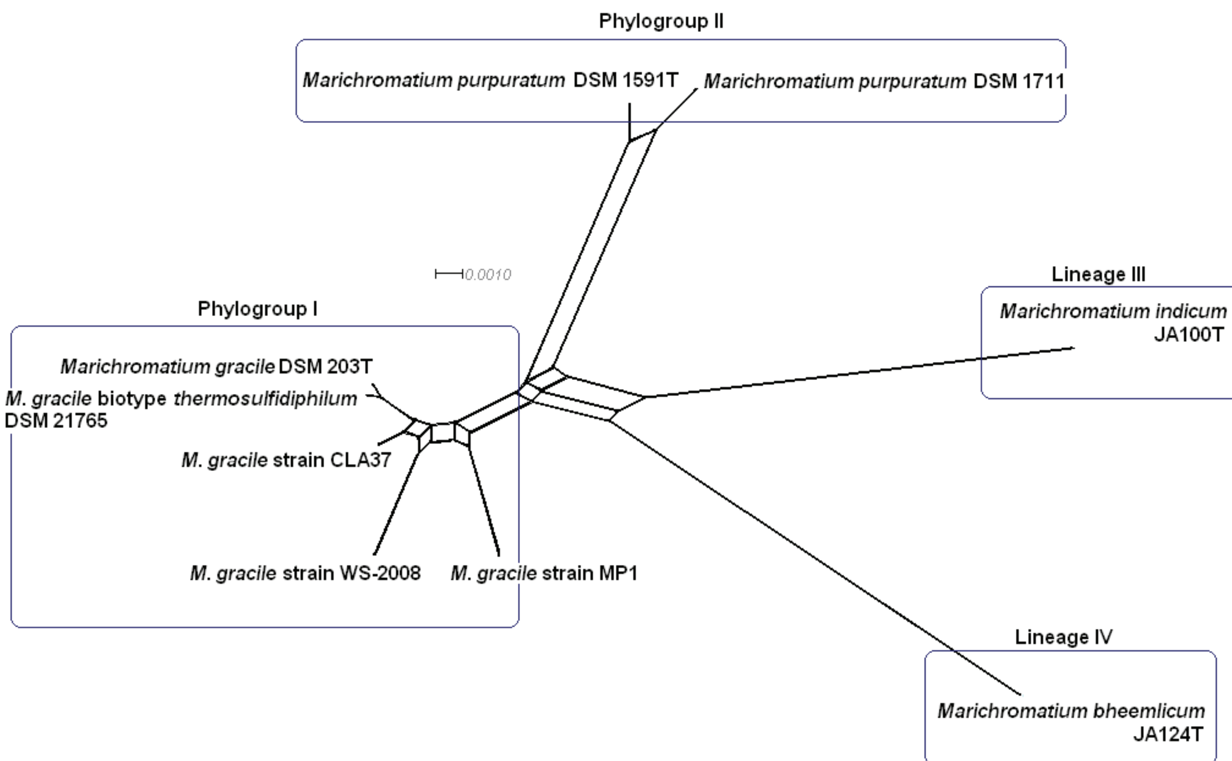


Figure S9. Concatenated split network tree constructed on the basis of the 16S rRNA, ITS, *pufM, soxB, fusA, recA,* and *dnaK* gene sequences (4331 bp). The network-like tree based on the concatenated sequences separate the nine *Marichromatium* strains in four phylotypes. Distances were corrected using the Jukes and Cantor method [21]. The scale bar indicates 0.1% sequence difference.

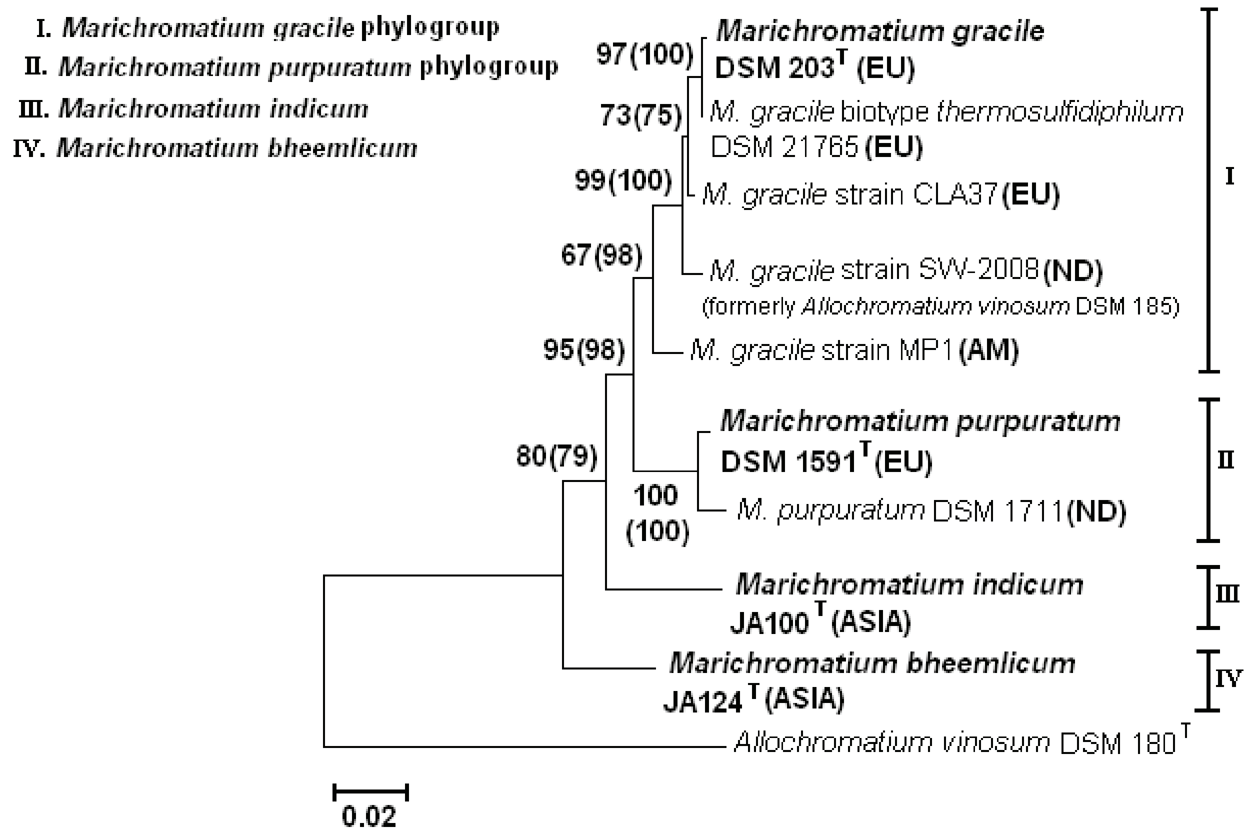


Figure S10. Phylogenetic tree reconstructed based on the concatenation of the protein-coding genes: *gyrB, recA, fusA, dnaK, pufM,* and *soxB* of nine *Marichromatium* strains totalizing (2646 bp). The tree was constructed by maximum likelihood method with the JTT model of protein evolution as implemented in the program PhyML and distance method according to Jukes and Cantor [21]. Bootstrap support in percentage was based on 100 replications (JC distance method in parentheses). Bar indicates 2% of estimated sequence divergence. *Allochromatium vinosum* DSM 180^T was used as outgroup. (EU), strains originated from Europe; (AM), strain originated from South America; (ASIA), type strains originated from India; (ND), no data. Geographical coordinates are indicated in Table S1.

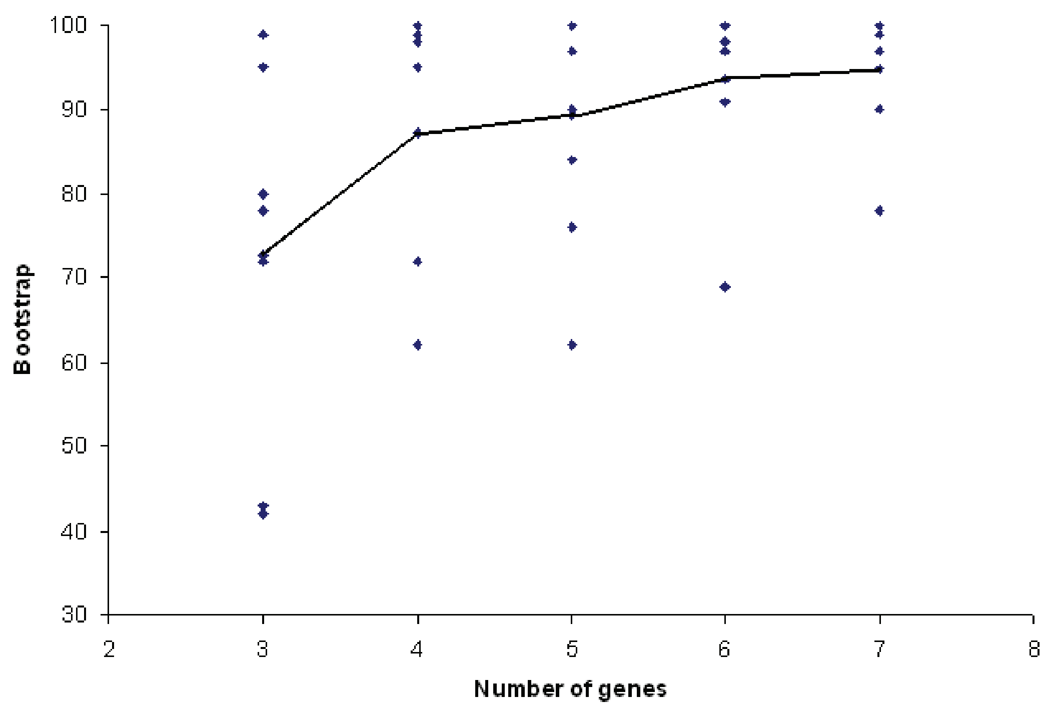


Figure S11. Bootstrap support of independent phylogenetic trees constructed with different number of genes of nine *Marichromatium* strains and *Allochromatium vinosum* DSM 180^T. The diamonds represent the bootstrap support for each node in the tree. The line connects the average bootstrap value of seven nodes for each data set.

Supplementary literature

[1] L. A. Achenbach, J. Carey, M. T. Madigan, Photosynthetic and phylogenetic primers for detection of anoxygenic phototrophs in natural environments. Appl. Environ. Microbiol. 67 (2001) 2922-2926.

[2] D. Hensen, D. Sperling, H. G. Trüper, D. C. Brune, C. Dahl, Thiosulphate oxidation in the phototrophic sulphur bacterium *Allochromatium vinosum*. Mol. Microbiol. 63 (2006) 794-810.

[3] J. F. Imhoff, H. G. Trüper, *Chromatium purpuratum* sp. nov. a new species of the *Chromatiaceae*. Zentbl. Bakteriol. I Abt. Orig., Reihe C1 (1980) 61-69.

[4] G. Muyzer, E. C. De Waal, A. G. Uitterlinden, Profiling of complex microbial population by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. 59 (1993) 695-700.

[5] R. Santos, H. Ochman, Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. Environ. Microbiol. 6 (2004) 754-759.

[6] G. Segal, E. Z. Ron, The *dnaKJ* operon of *Agrobacterium tumefaciens*: Transcriptional analysis and evidence for a new heat shock promoter. J. Bacteriol. 177 (1995) 5952-5958.

Figure S12. Multi-gene sequence alignment of concatenated nucleotide sequences (4331 bp) of the 16S rRNA, *gyrB*, *recA*, fusA, *dnaK*, *pufM*, and *soxB* genes and the 16S-23S Internal Transcribed Spacer (ITS) region.

M_pur_1591
M_pur_1711
M_grac_203_
M_grac_21765
M_grac_MP1
M_grac_CLA37
M_grac_WS-2008
M_indicum_JA100
M_bheem_JA124
A_vinosum_180

```
                         1130      1140      1150      1160      1170      1180      1190      1200      1210      1220      1230      1240      1250      1260
                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
M_pur_1591      TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTCGCAAAGCGGCGACGTGGAGCTAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
M_pur_1711      TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCAAAGCGGCGACGTGGAGCTAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
M_grac_203_     TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCAAAGCGGCGACGTGGAGCTAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
M_grac_21765    TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCAAAGCGGCGACGTGGAGCTAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
M_grac_MP1      TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCAAAGCGGCGACGTGGAGCTAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
M_grac_CLA37 .  TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCAAAGCGGCGACGTGGAGCTAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
M_grac_WS-2008  TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCAAAGCGGCGACGTGGAGCTAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
M_indicum_JA100 TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTTGCAAAGCGGCGACGTGGAGCTAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
M_bheem_JA124   TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGGGTCGCAAAGCCGCGAGGTGGAGCCAATCTCAGAAAACCGGTCGTAGTCCGGATTGCAGTCTGCAACTCGACTGCATGAA
A_vinosum_180   TCAAGTCATCATGGCCCTTATGGGCTGGGCTACACACGTGCTACAATGGTCGGTACAGAGCGTCGCGAAGCCGCGAGGTGGAGCCAATCGCACAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAA

                         1270      1280      1290      1300      1310      1320      1330      1340      1350      1360      1370      1380      1390      1400
                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
M_pur_1591      GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCGGGAGTGCGCATCACCCTCAGCGAAGAGGCGACGGGACGCAGCGATCACTTCGAGTAC
M_pur_1711      GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCGGGAGTGCGCATCACCCTCAGCGAAGAGGCGACGGGACGCAGCGATCACTTCGAGTAC
M_grac_203_     GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCGGGAGTGCGCATCACCCTCAGCGAGGAGGCGACGGGGCGCAGCGACCACTTCGAGTAC
M_grac_21765    GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCGGGAGTGCGCATCACCCTCAGCGAGGAGGCGACGGGGCGCAGCGACCACTTCGAGTAC
M_grac_MP1      GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCGGGGGTGCGCATCACCCTCAGCGAGGAGGCGACGGGACGCAGCGATCACTTCGAGTAC
M_grac_CLA37 .  GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCGGGAGTGCGCATCACCCTCAGCGAGGAGGCGACGGGGCGCAGCGACCACTTCGAGTAC
M_grac_WS-2008  GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCGGGAGTGCGCATCACCCTCAGCGAGGAGGCGACGGGGCGCAGCGACCACTTCGAATAT
M_indicum_JA100 GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCCGGGGTCCGGATCACCCTCGCCGAGGAGGCGACGGGACGCAGCGATCACTTCGAGTAT
M_bheem_JA124   GTCGGAATCGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCTTTGTACACACCGCCCGTCACACCATCGGGGGTGCGCATCATCCTCAGCGAGGAGGCGACCGGGCGCAGCGATCACTTCGAATAC
A_vinosum_180   GTCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATCGGGCGTGCGCATCGCCCTGAGCGACGAAGCCAGCGGGCGGCACGATGTGTTCGAGTAC

                         1410      1420      1430      1440      1450      1460      1470      1480      1490      1500      1510      1520      1530      1540
                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
M_pur_1591      CAGGGCGGCATCCGCGCCTTCGTCGAGCACCTCAACAGCAACAAGGAACCGATCCACCCGACGGTGATGCACTTCAGCGCCGAGCGCCAGGACATCTCCGTCGAACTGGCGATCCAGTGGAACAGTGGGTATCAGGAGAC
M_pur_1711      CAGGGCGGCATCCGCGCCTTCGTCGAGCACCTCAACAGCAACAAGGAACCGATCCACCCGACGGTGATGCACTTCAGCGCCGAGCGCCAGGACATCTCCGTCGAACTGGCGATCCAGTGGAACAGTGGGTATCAGGAGAC
M_grac_203_     CAGGGTGGCATCCGCGCCTTCGTCGAGCACCTCAACAGCAACAAGGAGCCGATCCACCCGACGGTGATGCACTTCAGCGCCGAGCGCCAGGACATCTCCGTCGAACTGGCGATCCAGTGGAACAGTGGGTATCAGGAGAC
M_grac_21765    CAGGGTGGCATCCGCGCCTTCGTCGAGCACCTCAACAGCAACAAGGAGCCGATCCACCCGACGGTGATGCACTTCAGCGCCGAGCGCCAGGACATCTCCGTCGAACTGGCGATCCAGTGGAACAGTGGGTATCAGGAGAC
M_grac_MP1      CAGGGTGGCATCCGCGCCTTCGTCGAGCACCTCAACAGCAACAAGGAGCCGATCCACCCGACGGTGATGCACTTCAGCGCCGAGCGCCAGGACATCTCCGTCGAACTGGCGATCCAGTGGAACAGTGGGTATCAGGAGAC
M_grac_CLA37 .  CAGGGTGGCATCCGCGCCTTCGTCGAGCACCTCAACAGCAACAAGGAGCCGATCCACCCGACGGTGATGCACTTCAGCGCCGAGCGCCAGGACATCTCCGTCGAACTGGCGATCCAGTGGAACAGTGGGTATCAGGAGAC
M_grac_WS-2008  CAGGGTGGCATCCGCGCCTTCGTCGAGCACCTCAACAGCAACAAGGAGCCGATCCACCCGACGGTGATGCACTTCAGCGCCGAGCGCCAGGACATCTCCGTCGAACTGGCGATCCAGTGGAACAGTGGGTATCAGGAGAC
M_indicum_JA100 CAGGGCGGCATCCGCGCCTTCGTCGAGCATCTCAACAGCAACAAGGAGCCGATCCACCCGACGGTGATGCACTTCAGCGCCGAGCGCCAGGACATCGCCGTGGAGCTGGCCATCCAGTGGAACAGCGGGTATCAGGAGAC
M_bheem_JA124   CAGGGCGGCATCCGCGCCTTCGTCGAACACCTCAACGGCAACAAGGAACCGATCCATCCGACGGTGATGCACTTCACCGCCGAGCGCCAGGACATCACGGTCGAGCTGGCGATCCAGTGGAACAGCGGGTATCAGGAGAC
A_vinosum_180   GAGGGCGGCATCCGCGCCTTCGTCGAGAATCTCAACCAGAACAAGGAGCCGATCCACCCGACGGTCGTCTATTTCAGCACCGAGCGTCAGGACATCGGCGTCGAGCTGGCGATCCAGTGGAACAACAGCTATCAGGAAAC

                         1550      1560      1570      1580      1590      1600      1610      1620      1630      1640      1650      1660      1670      1680
                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
M_pur_1591      CATCTTCTGCTACACCAACACCATTCCGCAGCGCGACGGCGGCACCCACCTCGCGGGGCTGCGCGCCGCACTCACCCGCACCCTCAACAGCTATATCGAGCAAGAGGCGCTGGACCGCAGCCAGAAGGTGCGCGCCATCG
M_pur_1711      CATCTTCTGCTACACCAACACCATTCCGCAGCGCGACGGCGGCACCCACCTCGCGGGGCTGCGCGCCGCACTCACCCGCACCCTCAACAGCTATATCGAGCAAGAGGCGCTGGACCGCAGCCAGAAGGTGCGCGCCATCG
M_grac_203_     CATCTTCTGCTACACCAACACCATCCCGCAGCGCGACGGCGGCACCCACCTCGCGGGGCTGCGCGCCGCGCTCACCCGCACCCTCAACAGCTATATCGAGCAAGAGGCGCTGGACCGCAGCCAGAAGGTGCGCGCCATCG
M_grac_21765    CATCTTCTGCTACACCAACACCATCCCGCAGCGCGACGGCGGCACCCACCTCGCGGGGCTGCGCGCCGCGCTCACCCGCACCCTCAACAGCTATATCGAGCAAGAGGCGCTGGACCGCAGCCAGAAGGTGCGCGCCATCG
M_grac_MP1      CATCTTCTGCTACACCAACACCATCCCGCAGCGCGACGGCGGCACCCACCTCGCGGGGCTGCGCGCCGCACTCACCCGCACCCTCAACAGCTATATCGAGCAAGAGGCGCTGGACCGCAGCCAGAAGGTGCGCGCCATCG
M_grac_CLA37 .  CATCTTCTGCTACACCAACACCATCCCGCAGCGCGACGGCGGCACCCACCTCGCGGGGCTGCGCGCCGCGCTCACCCGCACCCTCAACAGCTATATCGAGCAAGAGGCGCTGGACCGCAGCCAGAAGGTGCGCGCCATCG
M_grac_WS-2008  CATCTTCTGCTACACCAACACCATTCCGCAGCGCGACGGCGGCACCCACCTCGCGGGGCTGCGCGCCGCGCTCACCCGCACTCTCAACAGCTATATCGAGCAAGAGGCGCTGGACCGCAGCCAAAAGGTGCGCGCCATCG
M_indicum_JA100 CATCTTCTGCTACACCAACACCATCCCGCAGCGTGACGGCGGCACCCACCTCGCCGGGCTGCGCGCCGCGCTCACCCGCACCCTCAACAGCTATATCGAACAGGAGGGGCTGGAGCGCAGCCAGAAGGTACGCGCCATCG
M_bheem_JA124   CATCTTCTGTTACACCAACACCATCCCGCAGCGCGACGGCGGCACCCATCTCGCCGGGCTACGCGCTGCACTCACCCGCACCCTCAACAACTACATCGAGCATGAGGGATTGGACCGCAGCCAGAAGGTGCGTGCCATCG
A_vinosum_180   CATCTTCTGCTACACCAACACCATCCCGCAGCGTGACGGCGGCACCCATCTGGCCGGTCTGCGCGCAGCACTCACGCGCACGCTCAACGGCTACATCGAGCGCGAGGGACTGGACAAGAACCAGAAGATCCGCCCGATCG
```

M_pur_1591
M_pur_1711
M_grac_203_
M_grac_21765
M_grac_MP1
M_grac_CLA37 .
M_grac_WS-2008
M_indicum_JA100
M_bheem_JA124
A_vinosum_180

Chapter 4

The genus *Allochromatium* (*Chromatiales Chromatiaceae*) revisited: a study on its intragenic structure based on Multilocus Sequence Analysis (MLSA) versus DNA-DNA hybridization (DDH).

Wilbert Serrano[1, 2, *], Rudolf Amann[2] and Ulrich Fischer[1]

Manuscript in preparation

Abbreviations: TD-PCR, touchdown-PCR; DDH, DNA-DNA hybridization; MLSA, multilocus sequence analysis; (π), nucleotide diversity; (H), genetic diversity.

The GenBank/EMBL/DDBJ accession numbers for the *Marichromatium* strains sequences determined in this study are listed in Table 3.

[1] Zentrum für Umweltforschung und nachhaltige Technologien, Fachbereich Biologie/Chemie, Abteilung Marine Mikrobiologie, Universität Bremen, D-28359 Bremen, Germany
[2] Max Planck-Institute for marine Microbiology, Celsiusstraße 1, D-28359 Bremen, Germany
* Corresponding author, email: wserrano@mpi-bremen.de

Abstract. In the present study, the taxonomic status of the anoxygenic photosynthetic bacteria of the genus *Allochromatium* is revisited. The inter- and intraspecies relationship of seven *Allochromatium* strains, including a set of well described type species, were examined by DNA-DNA hybridization (DDH) and multilocus sequence analysis (MLSA) using segments of eight protein-coding genes. The re-sequencing of the 16S rRNA, the internal transcriber spacer (ITS), multi-gene analysis and DDH comparison indicated that both type species *Allochromatium vinosum* DSM 180^{T} and *Allochromatium minutissimum* DSM 1376^{T} are closely related each other forming an independent phylogroup together with the strains *A. vinosum* DSM 183 and DSM 1686. The internal comparison of members of the *A. vinosum* phylogroup showed values of DDH relatedness above 80% and concatenated sequence similarities (4427 bp) above 97%. DDH comparison and MLSA scheme also allowed the separation of new isolate strain AX1YPE and *Allochromatium sp*. strain BH-2 as separate lineages. Strain BH-2 was first classified as a member of the species *A. vinosum* based on DDH comparison. However, this strain showed the lowest similarity values of the 16S rRNA gene and concatenated sequences, as well as amino acid identity (AAI) when compared to other *Allochromatium* strains, suggesting that strain BH-2 could be a news species.

Introduction

Anoxygenic photosynthetic purple sulfur bacteria represent an important group of prokaryotes which play a significant role in the sulfur cycle in anoxic environments where organic matter is degrading and light is available. Early classification of purple sulfur bacteria was mainly based on phenotypic criteria. As example, the ability to store elemental sulfur - a by product of sulfide oxidation - inside or outside cells has allowed the separation of the former family *Chromatiaceae* in two new families *Ectothiorhodospiraceae* and *Chromatiaceae* (Imhoff, 1984). With the advent of the molecular techniques and the use of the 16S rRNA gene as phylogenetic marker, the separation of both families - which now represents truly separate phylogenetic lineages - has been confirmed (Imhoff *et al*., 1998). Description of several genera which includes representative species from both marine and fresh-water ecosystems within the family *Chromatiaceae* was also possible by using the 16S rRNA approach. The genus *Allochromatium* (formerly *Chromatium*) has been described as a new genus within the *Chromatiaceae* (Imhoff *et al*., 1998). This genus comprises four validly described type strains: *Allochromatium vinosum* DSM 180^T, *A. minutissimum* DSM 1376^T, *A. warmingii* DSM 173^T (Imhoff *et al*., 1998), and the recently reported new species *A. renukae* DSM 18317 (Kumar *et al*., 2008). Among them, only the species *A. warmingii* DSM 173^T is significantly different in its cell shape and size from those of the other “small” *Allochromatium* species Imhoff *et al*. (1998).

The type strain *A. vinosum* DSM 180^T is currently subject of intensive research concerning mainly sulfur metabolism (Reinartz, 1998; Hensen *et al*., 2006) and determination of the molecular structure of the light harvesting photosynthetic pigments (Kereiche *et al*., 2008). However, despite the importance of knowing such essential molecular mechanisms, little is known about the real taxonomic status of the species *A. vinosum.* This species comprises several closely related strains which have been deposited in different culture collections. On the basis of their morphology and physiology, many of these strains were tentatively identified as members of the species *A. vinosum.* Interestingly, several studies have revealed that the separated type species *A. vinosum* DSM 180^T and *A. minutissimum* DSM 1376^T are very similar too. Although both species pair exhibit minimal phenotypic differences e.g. cell shape, the analysis of their genomic banding pattern (PFGE) and their phylogenetic banding pattern (riboprinting) clearly indicate that both separated type strains may represent one single species (Mas-Castella *et al*., 1996; Pavon & Gaju, 1997).

The process of describing new bacterial isolates and arranging them into taxonomic hierachy is part of a strictly regulated framework. Stackebrandt *et al.* (2002) have reported a general rule for the description of new bacterial species. A revised recommendation intended to serve as guideline for description of new species of anoxygenic phototrophic bacteria was also published by Imhoff & Caumette (2004). In a general sense, these recommended rules underline the importance of including phenotypic, chemotaxonomic, phylogenetic and genomic comparison in a polyphasic approach. Central for the delineation of species is genomic comparison by DNA-DNA hybridization (DDH) (Stackebrandt *et al.*, 1999; Imhoff & Caumette, 2004).

In addition, the use of multilocus sequence analysis (MLSA) has been proposed as a good alternative method (Imhoff & Caumette, 2004). This method has been extensively used for defining species in medical and environmental microbiology, e.g. *Burkholderia* (Godoy *et al.*, 2003; Glass *et al.*, 2006) and *Vibrios* (Thompson *et al.*, 2005, 2007; Sawabe *et al.*, 2007). MLSA has several advantages in comparison to the classical taxonomy, it is transferable between laboratories and a multi-gene phylogeny offers more information than a single gene-based phylogeny.

In this study, MLSA has been applied for identification and classification of a new purple sulfur bacterium isolate strain AX1YPE as well as for the taxonomic re-evaluation of *Allochromatium sp.* strain BH-2.

Material and methods

Type strains of the genus *Allochromatium* were obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ). In addition, a new isolate retrieved from soft marine sediments from central coast of Peru and designated here as *Allochromatium sp.* strain AX1YPE was also examined (the list of all strains included in this study and the corresponding geographic coordinates is given in Table S1, supplementary online material). *Allochromatium* strains were cultivated at 30℃ in anoxic Pfennig´s medium (Pfennig & Trüper, 1992). Cultures were purified using traditional cultivation methods (agar-shake dilution series) and the resulting single colonies were grown in completely filled screw-capped bottles containing Pfennig's medium (Pfennig & Trüper, 1992). Phototrophic growth conditions were set up according to Pfennig & Trüper (1992) and as was reported previously (Serrano *et al.*, 2009). NaCl concentrations were adjusted for each strain according to the

recommendation of the DSMZ. The new strain AX1YPE was grown in Pfennig's medium containing 3% of NaCl.

Genomic DNA was extracted according to the method of Marmur (1961) with modifications. For the purpose of the present study, the extracted genomic DNA was dissolved in 0.1x of SSC (sodium saline buffer). High concentrated genomic DNA 0.4 mg/ml were prepared for DDH experiments. The DDH experiments were made by renaturation rates according to the method of DeLey (1970), using a Beckman DU-640 spectrophotometer equipped with thermostated multicell changer and a temperature controller. The G+C mol% was determined by thermal-denaturation according to the method of Marmur & Doty (1962).

Eight protein-coding genes: photosynthetic reaction center M (*pufM*), sulphate thioesterase/thiohydrolase (*soxB*), 1-5 bisphosphate carboxylase oxygenase (*cbbL-1*), heat shock co-chaperonin (*groES*), elongation factor G (*fusA*), transcription repair coupling factor (*mfd*), recombinase A (*recA*), DNA gyrase (*gyrB*), and the 16S rRNA gene as well as the internal transcriber spacer (ITS) region sequences were PCR amplified and sequenced. Primer sequences retrieved from the literature, those designed in our study, and the PCR parameters for amplification and sequencing these gene loci are all listed in (Table S2, online supplementary material).

The gene sequences were deposited in the GeneBank databases with accession numbers as following: (FJ812038 to FJ812041), 16S rRNA gene; (FJ823965 to FJ823971), internal transcriber spacer (ITS); (FJ812042 to FJ812048), *pufM* genes; (FJ812049 to FJ912055), *soxB* genes; (FJ812056 to FJ812062), *cbbL*-1 genes; (FJ812070 to FJ812076), *groES* genes; (FJ812063 to FJ812069), *fusA* genes; (FJ823946 to FJ823952), *mfd* genes; (FJ823953 to FJ823958), *recA* genes, and (FJ823959 to FJ823964), *gyrB* genes.

The sequenced housekeeping genes were aligned by using the ClustalX program (Thompson *et al.*, 1997). Aligned sequences were trimmed at the end by using BioEdit software package; this program was also used for translating the nucleotide sequences into their corresponding amino acid sequences. Gene diversity indices, segregation sites (SG), allelic number (N), haplotype diversity (H), and nucleotide diversity (π) were analysed by using the software package DnaSP 4.0 (Rozas *et al.*, 2003). Test of neutrality based on Tajima's D with (0=neutral) and the ratios of the non-synonymous (d*N*) and synonymous (d*S*) substitutions per site were also

determined by using the software DnaSP 4.0. Phylogenetic analyses for each independent gene alignment as well as for the resulting concatenated sequences (4357 bp) were reconstructed by maximum likelihood using the program PhyML (Guindon & Gascuel, 2003) with substitution models HKY85 as was implemented in the program. The network split tree was reconstructed by using the program Split Tree 4.0 (Husson & Bryant, 2006).

Results and discussions

Based on the isolation of strain AX1YPE an extensive research within the genus *Allochromatium* was done. A total of twelve culture collection available *Allochromatium* strains were obtained (Table S2, supplementary online material). From them only nine strains were successfully cultivated. Phylogenetic affiliation on these cultivated *Allochromatium* strains were explored by sequencing their 16S rRNA gene and comparing them with those sequences deposited on the Genebank databases. Based on this analysis, we found that the strains designated as *Allochromatium vinosum* DSM 182 and *A. vinosum* DSM 185 affiliate within the genus *Thiorhodococcus* and *Marichromatium* respectively, consequently, these strains were excluded from further analysis.

Surprisingly, the re-sequenced 16S rRNA genes of the species pair *A. vinosum* 180^T and *A. minutissimum* DSM 1376^T were identical each other (Table1), contrasting with those sequences of the same species deposited in the GeneBank databases which exhibited a high number of nucleotide ambiguities (we could correct those ambiguities by using both new sequences in our current analysis). On the other hand, the re-sequenced 16S rRNA gene of the species *A. renukae* and the strain BH-2 matched exactly to those sequences from the same organisms deposited in the GeneBank data bases. In the present study, all new 16S rRNA sequences were deposited in the GeneBank databases with accession numbers as was listed above.

Pairwise cross-comparison among the seven *Allochromatium* strains based on their genomic DDH relatedness, similarities of the 16S rRNA gene, ITS region and multilocus concatenated sequences as well as the amino acid identity (AAI) are summarized in Table 1.

As was expected, intra-group cross-comparison among *A. vinosum* strains DSM 180^T, DSM 183 and DSM 1686 showed high 16S rRNA sequence similarity ranging from 99.76 to 99.84% (Table 1). As reported before, However, unexpectedly pairwise

comparison of both separated type species *A. vinosum* DSM 180^T and *A. minutissimum* DSM 1376^T had identical 16S rRNA gene sequences (100% similarity) and they were related by 98.89, 97.94 and 97.07% similarities to the strains AX1YPE, *A. renukae* DSM 18713, and strain BH-2, respectively (Table 1).

Table 1: Cross-comparison of seven *Allochromatium* strains based on their DNA-DNA relatedness,16S rRNA gene, ITS region, and concatenated nucleotide sequences similarity as well as their corresponding amino acid sequence identity (AAI).

Allochromatium strains compared	Parameters compared				
	DDH (%) *	16S rRNA sequence similarity (%)**	ITS sequence similarity (%)	Concatenated sequence similarity (%)	Average amino acid identity (AAI %)
A. vinosum DSM180 vs. *A. minutissimum* DSM 1376	**84 ± 4.9**	**100.0**	**100**	**98.1**	**99.4**
A. vinosum DSM 180 vs. strain AX1YPE	59 ± 5.4	98.9	95.2	96.3	98.9
A. vinosum DSM 180 *vs Allochromatium sp.* strain BH-2	81 ± 1.4	97.1	87.1	91.9	96.6
A. vinosum DSM 180 *vs. A .renukae* DSM 18713	50 ± 8.1	97.9	89.3	96.2	98.2
A. vinosum DSM 180 *vs. A. vinosum* DSM 183	**89 ± 3.7**	**99.8**	**100**	**98.8**	**99.4**
A. vinosum DSM 180 *vs. A. vinosum* DSM 1686	**94 ± 1.9**	**99.8**	**100**	**99.5**	**99.6**
A. minutissimum DSM 1376 *vs*. strain AX1YPE	64 ± 2.4	98.9	95.2	96.4	98.7
A. minutissimum DSM 1376 *vs. Allochromatium sp*. strain BH-2	78 ± 2.9	97.1	87.1	91.9	96.5
A. minutissimum DSM 1376 *vs. A. renukae* DSM 18713	49 ± 2.5	97.9	89.3	95.9	98.2
A. minutissimum DSM 1376 vs. *A. vinosum* DSM 183	**87 ± 2.2**	**99.8**	**100**	**98.0**	**99.5**
A. minutissimum DSM 1376 *vs. A. vinosum* DSM 1686	**92 ± 9.7**	**99.8**	**100**	**98.3**	**99.5**
Strain AX1YPE vs. *Allochromatium sp.* strain BH-2	54 ± 5.9	96.2	87.1	92.1	96.0
Strain AX1YPE vs *A. renukae* DSM 18713	61 ± 3.6	96.9	88.8	95.7	97.2

*= The reassociation values are the means ± 1 SD of two independent determinations.
**= Pairwise similarities calculated by comparing 1267 bp.

It is well known that comparing the 16S rRNA gene sequences do not allow far separation of closely related strains (Fox *et al*, 1992; Rossello-Mora and Amann, 2001). However, a polymorphic region found between the 16S-23S, known as internal transcriber spacer (ITS) has been reported to be more discriminative at intraspecies level than the 16S rRNA gene Gürtler & Stanisich (1996). In the present study, we used the ITS sequences in order to confirm the 16S rRNA-based phylogeny. Our results indicate that all cross-compared *A. vinosum* strains DSM 180^T, DSM 183, DSM 1686 and *A. minutissimum* DSM 1376^T showed completely identical ITS sequence with similarities of 100%, while these sequences compared to the strain AX1YPE, strain BH-2 and *A. renukae* DSM 18713 showed similarities of 95.23, 87.14 and 89.28, respectively (Table1). Based on the phylogentic analysis of the 16S rRNA gene and ITS region sequences, *A. vinosum* strains DSM 180^T, DSM 183, DSM 1686, and *A. minutissimum* DSM 1376^T appears grouped in an independent cluster (phylogroup) clearly separated from the other examined *Allochromatium* strains (Fig. 1+2).

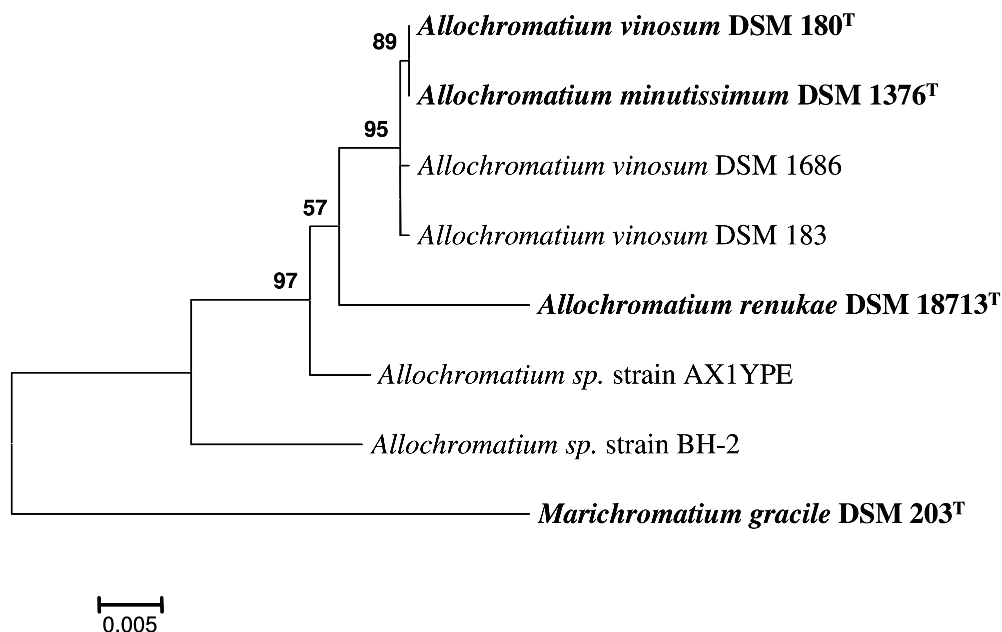


Figure 1. Phylogenetic tree reconstruction of the 16S rRNA gene sequences based on maximum likelihood as implemented in the program PhyML. Bootstrap percentages values > 50 is shown, after 100 repetitions. Bar indicates 0.5% of sequence divergence. Gene sequence of *Marichromatium gracile* DSM 203^T was used as outgroup.

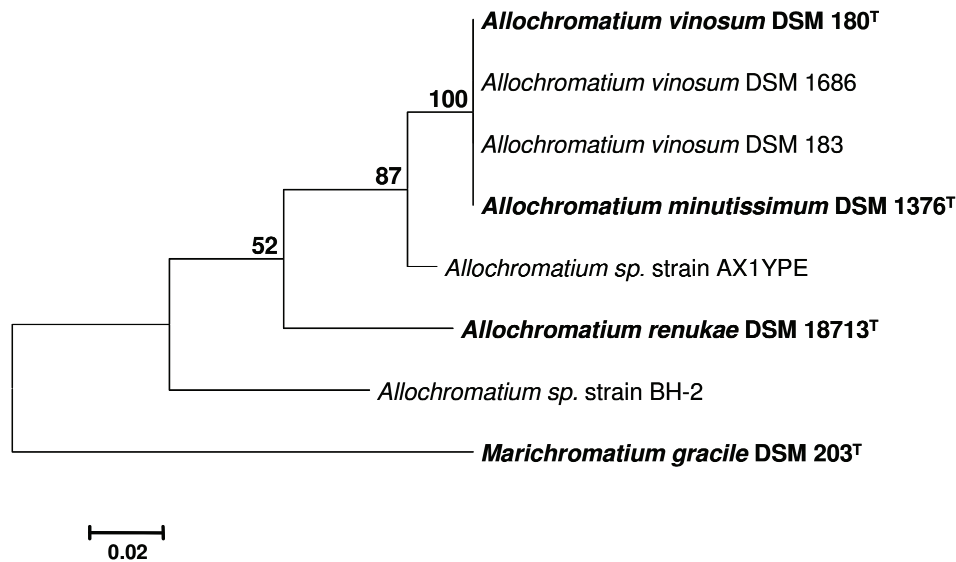


Figure 2. Phylogenetic tree reconstruction of the internal transcriber spacer ITS region sequences based on maximum likelihood as implemented in the program PhyML. Bootstrap percentages values > 50 is shown, after 100 repetitions. Bar indicates 2% of sequence divergence. Gene sequence of *Marichromatium gracile* DSM 203^T was used as outgroup

The genomic DDH-based pairwise comparison between two paired strains (*A.vinosum* DSM 180^T, versus *A. vinosum* DSM 183) and (*A. vinosum* DSM 180^T, versus *A. vinosum* DSM 1686) yielded a genomic relatedness of 89 and 94% respectively, however, unexpectedly the genomic comparison between pair species *A. vinosum* DSM 180^T and *A. minutissimum* DSM 1376^T also showed a higher degree of DDH of 84%. Similarly, DDH values above 85% were also found when the type strain *A. minutissimum* DSM 1376^T was compared to *A. vinosum* DSM 183 and *A. vinosum* DSM 1686 (Table 1). Thus, the DDH homology results together with those already found for the 16S rRNA and ITS similarity may suggest that all these strains belong to one single species. However, high DDH values of 81% (between *A. vinosum* DSM 180^T and strain BH-2) and 78% (between *A. minutissimum* DSM 1376^T and strain BH-2) were also found (Table1). The characterization of *Allochromatium sp.* strains BH-2 indicates that this strain possesses phenotypic properties not shared by other described *Allochromatium* species. However, strain BH-2 was not reported as a new species because its high DDH relatedness (87.6%) with respect to the type strain *A.*

vinosum DSM 180^T, although their corresponding 16S rRNA similarity of 97.2% was remarkably low (Herbert *et al*., 2005).

In order to solve this incongruence we developed a MLSA scheme, which includes the analysis of eight protein-coding genes (Table S1, supplementary material online). The use of a multigene approach offers the possibility to overcome possible bias caused by analysing a single gene e.g. recombination (Gevers *et al*., 2005).

The analysis of all examined loci and the resulting genetic diversity indices are shown in (Table 2). As can be seen, the length of the resulting fragment sequences varied from 213 bp (for *pufM*) to 840 bp (for *gyrB*). The examined loci exhibited base composition G+C mol% in the range of 62.19% (for *groES* gene) to 67.51% (for *cbbL* gene) with an average value of 63.62%. The average nucleotide diversity (π) among eight loci was 0.044, and they ranged from 0.024 (for *pufM*) to 0.064 (for *soxB*) (Table 2). The average gene diversity (H) was 0.923, while the average ratio of d*N*/d*S* was 0.074 (Table 2).

Table 2. Genetic diversity indices.

Locus	Size bp	SG	N	H	π	Tajima'D	d*N*/d*S*	G+C(mol%)
pufM	213	16	4	0.714	0.024	-1.090*	0.000	63.3
soxB	507	84	7	1.000	0.064	- 0.315*	0.198	64.3
cbbL	492	39	6	0.952	0.033	0.167*	0.197	67.5
groES	729	66	6	0.952	0.032	- 0.736*	0.048	62.8
fusA	555	77	6	0.952	0.052	- 0.397*	0.063	63.1
mfd	222	33	6	0.952	0.057	- 0.280*	0.015	62.1
recA [b]	372	42	5	0.933	0.046	- 0.414*	0.016	63.4
gyrB [b]	840	97	5	0.933	0.050	- 1.340*	0.058	63.4
average				**0.923**	**0.044**		**0.074**	**63.6**

SG=number of segregating sites; N=number of alleles; H=gene diversity; π =Nucleotide diversity; D=Tajima test values not significantly different from zero (P>0.10). [b] calculations were made based on the analysis of six *Allochromatium* strains.

Notably, some genetic diversity indices vary remarkably when these groups were analysed independently. As example, the calculated average nucleotide diversity (π) of 0.013 for the *A. vinosum* phylogroup was almost three times lower than the overall average (π) of (0.044), while the three independent lines constituted by the strains AX1YPE, BH-2 and the species *A. renukae* showed a higher (π) value of 0.063. Similarly, the gene diversity (H) for the *A. vinosum* phylogroup was H=0.770 contrasting with the average gene diversity of the entire data set of H=0.923. Thus, both the low nucleotide diversity (π) and gene diversity (H) observed for the *A. vinosum* phylogroup reinforces our assumption that all these strains are closely

related, and consequently their taxonomic status should be revised as was suggested by Mas-Castella *et al*. (1996) and Pavon & Gaju (1997). In addition, our finding also is supported by the fact that the DDH relatedness among strains within this subgroup was high >80%, clearly above the accepted threshold limit of 70% for delineation of species (Wayne et al., 1987). While, the unclear taxonomic status of *A. vinosum* strain BH-2 as was determined by its high DDH and low 16S rRNA gene similarities compared to the other *Allochromatium* strains can be not resolved by using these standard tools. However, such classification would be possible when the MLSA scheme is taken in consideration. As can be seen in Table1, strain BH-2 showed the lowest similarity values in its concatenated gene sequence and average amino acid identity in comparison to other *Allochromatium* strains, which was also in line with the low similarity values obtained in the 16S rRNA and ITS genes sequence comparison (Table 1).

The constructed concatenated phylogenetic tree showed that the *A. vinosum* subgroup formed a separate cluster from the other examined *Allochromatium* strains, this separation is supported by a bootstrap value of 100%, while strain BH-2 appears so far related to the other strains (Fig. 3 and Fig. S1, see supplementary online material). These resulting concatenated tree topologies were also in agreement the same as observed for the trees constructed with the 16S rRNA and ITS (Fig 1+2).

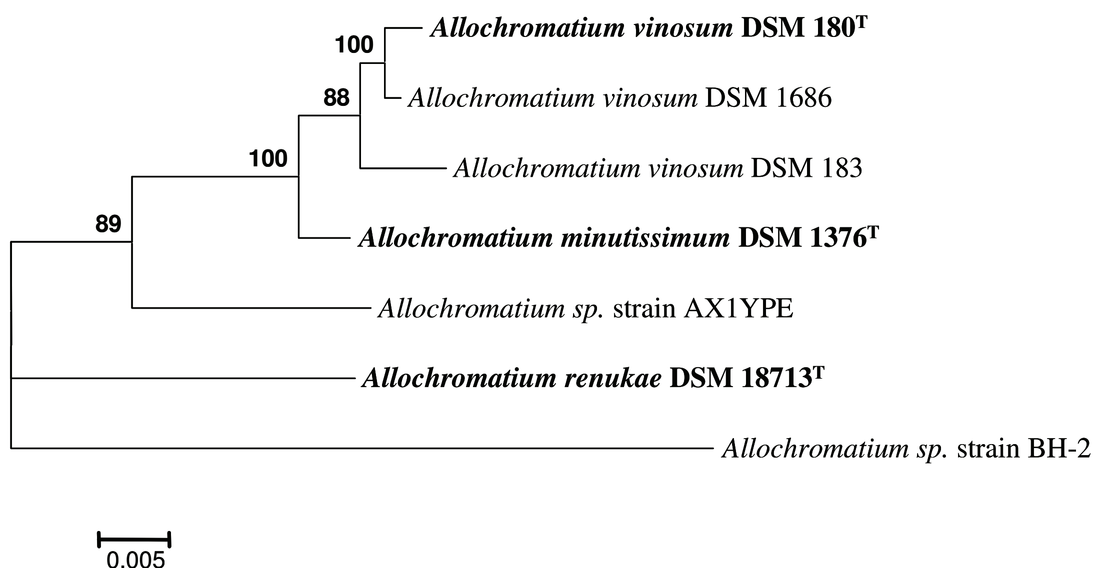


Figure 3. Phylogenetic tree reconstruction based on the concatenation of six protein-coding genes, 16S rRNA gene and ITS region sequences (4427 bp) of seven *Allochromatium* strains. The tree is based on the maximum likelihood algorithm as implemented in the program PhyML. Bootstraps values >50 after 100 repetitions is shown. Bar represent 0.5% of sequence divergence.

The data generated in this study could be helpful for differentiating closely related strains of photosynthetic bacteria of the genus *Allochromatium*. As we demonstrated here, there is no reason to maintain the species *A. vinosum* DSM 180^T and *A. minutissimum* DSM 1376^T as separated entities since both type strains were similar in many features. In addition, the separation of the strain BH-2 as a new species might be possible under the MLSA scheme; this strain showed the lowest similarity on its 16S rRNA, ITS, concatenated multigene sequences and the amino acid sequence identity (AAI).

Acknowledgement.

Support of the Bremen University and the Max Planck Society is gratefully acknowledged. W. S. is a member of the International Max Planck Research School of Marine Microbiology (MarMic).

Literature

De Ley, J., Catoir, H. & Reynaerts, A. (1970). The quantitative measurement of DNA hybridization from renaturation rates. Eur. J. Biochem. 12, 133-142.

Fox, G. E., Wisotzkey, J. D. & Jurtshuk, P. Jr. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int. J. Syst. Bacteriol. 42, 166-170.

Glass, **M. B., Steigertwalt, A.G., Jordan, J. G., Wilkins, P.P. & Gee, J. E. (2006).** *Burkholderia oklahomensis* sp. nov., a *Burkholderia pseudomallei*-like species formerly known as the Oklahoma strains of *Pseudomonas pseudomallei*. Int. J. Syst. Evol. Microbiol. 56, 2171-2176.

Gevers, D., Cohan, F. M., Laurence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F. & Swings, J. (2005). Re-evaluating prokaryotic species. Nature Rev. 3, 733-739.

Godoy, **D., Randle, G., Simpson, A. J., Aanensen, D. M., Pitt, T. L., Kinoshita, R. & Spratt, B. G. (2003)**. Multilocus sequence typing and evolutionary relationship among the causative agents of meliodosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. J. Clin. Microbiol. 41, 2068-2079.

Guindon, S. & Gascuel O. (2003). A simple, fast and accurate algorithm to estimate large phylogenies by maximal likelihood. Syst. Biol. 52, 696-704.

Gürtler, V. & Stanisich, V. A. (1996). New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. Microbiology. 142, 3–16.

Hensen, D., Sperling, D., Trüper, H. G., Brune, D. C. & Dahl, C. (2006). Thiosulphate oxidation in the phototrophic sulphur bacterium *Allochromatium vinosum*. Mol. Microbiol. 62, 794-810.

Herbert, R. A., Ranchou-Peyruse, A., Duran, R., Guyoneaud, R. & Schwabe, S. (2005). Characterization of purple sulfur bacteria from the South Andros Black Hole cave system: highlights taxonomic problems for ecological studies among the genera *Allochromatium* and *Thiocapsa.* Environ. Microbiol. 7, 1260-1268.

Huson, D. H. & Bryant, D. (2006). Applications of phylogenetic networks in evolutionary studies. Mol. Biol. Ecol. 23: 254-267.

Imhoff, J. F. (1984). Reassignment of the genus *Ectothiorhodospira* Pelsh 1936 to a new Family, *Ectothiorhodospiraceae* fam. nov., and emended description of the *Chromatiaceae* Bavendamm 1924. Int. J. Syt. Bacteriol. 34, 338-339.

Imhoff, F. I., Suling, J. & Petri, R. (1998). Phylogenetic relationship among the *Chromatiaceae*, their taxonomic reclassification and description of the new genera *Allochromatium*, *Halochromatium, Isochromatium, Marichromatium, Thiococcus*, *Thiohalocapsa* and *Thermochromatium*. Int. J. Syst. Bacteriol. 48, 1129-1143.

Imhoff, J.F. & Caumette, P. (2004). Recommended standards for the description of new species of anoxygenic phototrophic bacteria. Int. J. Syst. Evol. Microbiol. 54, 1415-1421.

Kereiche, S., Bourinet, L., Keegstra, W., Arteni, A., Verbavatz, J., Boekema , E., Robert, B. & Gall, A. (2008). The peripheral light-harvesting complexes from purple sulfur bacteria have different 'ring' sizes. FEBS Microbiol. Lett. 582, 3650-3656.

Kumar, S., Tamura K. & Nei, M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150-163.

Kumar, P. A., Srinivas, T. N. R., Sasikala, Ch. & Ramana, Ch. V. (2008). *Allochromatium renukae* sp. nov. Int. J. Syst. Evol. Microbiol. 58, 404-407.

Martinez-Murcia, A. J., Benlloch, S. & Collins, M. D. (1992). Phylogenetic interrelationships of members of the genera *Aeromonas* and *Plesiomonas* as

determined by 16S ribosomal DNA sequencing: Lack of congruence with results of the DNA-DNA hybridizations. Int. J. Syst. Bacteriol. 42, 412-421.

Mas-Castella, J., Guerrero, R. & De Jonckheere, J. F. (1996). High degree of similarity between *Chromatium vinosum* and *Chromatium minutissimum* as revealed by riboprinting. Int. J. Syst. Bacteriol. 46, 922-925.

Marmur, J. (1961). A procedure for the isolation of desoxyribonucleic acid from micro-organisms. J. Mol. Biol. 3, 208-218.

Marmur, J., & Doty, P. (1962). Determination of the base composition of desoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. 5, 109-118.

Muyzer, G., De Waal, E. C. & Uitterlinden, A. G. (1993). Profiling of complex microbial population by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. 59, 695-700.

Pavon, **V. & Gaju, N. (1997).** Genomic relationship of five species of the genus *Chromatium* by analysis of large restriction fragments (macrorestriction analysis) using Pulsed-Field Gel Electrophoresis. Int. J. Syst. Bacteriol. 47, 586-589.

Pfennig, N. & Trüper, H. G. (1992). The family Chromatiaceae, In: Balows A., H. G. Trüper, M. Dworkin, W. Harder, K. H. Schleifer (eds) The Prokaryotes. Springer, Berlin, Heidelberg, New York, pp 3200-3221.

Reinartz, **M., Tschäpe, J., Brüser, T., Trüper H. G. & Dahl, C. (1998).** Sulfide oxidation in the phototrophic sulfur bacterium *Chromatium vinosum*. Arch. Microbiol. 170, 59-68.

Rossello-Mora, **R. & Amann, R. (2001)**. The species concept for prokaryotes. FEMS Microbiol. Rev. 25, 39-67.

Rozas J., Sánchez-Del Barrio J. C., Messeguer, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Serrano, W., Amann R. & Fischer, U. (2009). A new moderately thermophilic and high sulfide tolerant biotype of *Marichromatium gracile*, isolated from tidal sediments of the German Wadden Sea: *Marichromatium gracile* biotype *thermosulfidiphilum*. Syst. Appl. Microbiol. 32, 1-7.

Thompson, J. D., Gibson , T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876-4882.

Thompson, F. L., Gevers D., Thompson C. C., Dawyndt P., Naser S., Hoste B., Munn C.B. & Swings J. (2005). Phylogeny and molecular identification of Vibrios on the basis of multilocus sequence analysis. Appl. Environ. Microbiol. 71, 5107-5115.

Thompson, F. L., Gomez-Gil B., Ribeiro-Vasconcelos A. T. & Sawabe, T. (2007). Multilocus sequence analysis reveals that *Vibrio harveyi* and *V. campbellii* are distinct species. Appl. Environ. Microbiol. 73, 4279-4285.

Sawabe, **T., Kita-Tsukamoto K. & Thomson, F.L. (2007).** Inferring the evolutionary history of *Vibrios* by means of multilocus sequence analysis. J. Bacteriol. 189, 7932-7936.

Stackebrandt, E., Frederiksen W., Garrity G. M., Grimont P. A. D., Kämpfer P., Maiden M. C. J., Nesme X., Rossello-Mora R., Swings J., Trüper H. G., Vauterin L., Ward A. C. & Whitman, W. B. (2002). Report of the Ad Hoc committee for the re-evaluation of the species definition in bacteriology. Int. J. Syst. Evol. Microbiol. 52, 1043-1047.

Wayne, L. G. , Brener, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichesvsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. & Trüper, H. G. (1987). International committee on sytematic bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int. J. Syst. Bacteriol. 37, 463-464.

Supplementary tables and figures

Table S1. Primer sequences used for the PCR amplification and sequencing of ten loci used in this study.

Locus	Probably function of gene (or protein)	Sequenced amplicon size (bp)	Primer	Primer sequence (5'-3') [a]	Annealing temperature [b]	Reference
16S rRNA	Protein synthesis	>1400	GM3	AGAGTTTGATCMTGGC	44	Muyzer *et al*. (1993)
			GM4	TACCTTGTTACGACTT	44	Muyzer *et al*. (1993)
ITS	rRNA intergenic spacer	422	ITSf	GTGAAGTCGTAACAAGG	44	This study
			ITSr	CGTCACTTGACCATATCC	44	This study
pufM	Reaction center M	225	pufM.557F	CGCACCTGGACTGGA	55	Achenbach *et al*. (2001)
			pufM.750R	CCCATGGTCCAGCGCCAGAA	55	Achenbach *et al*. (2001)
soxB	Sulfate thioesterase/sulfate	507	soxBF	GACGGTGGTGATACCTG	55	Hensen *et al*. (2006)
	thiohydrolase		soxBR	CATGTCGCCGCCCTGCTG	55	Hensen *et al*. (2006)
			soxinR	GACGTCCATGCCGTTGT	50	This study
cbbL-1	1-5 Bisphosphate carboxylase	492	cbbLWF	ATCATGCACGACTACAT	44	This study
	oxygenase		cbblWR	CAGGTCTCCATGGCGA	44	This study
groES	Chaperonin groES like family	729	groESWF	GCAGATGGTCAAGGAAGT	44	This study
			groESWR	CGACAGGCCGACTTCTT	44	This study
fusA	Translation elongation factor G	330	fusAF	CATCGGCATCATGGCNCAYATHGA [a]	TCHDW (50)	Santos & Ochman (2004)
			fusAR	CAGCATCGGCTGCAYNCCYTTRTT [a]	TCHDW (50)	Santos & Ochman (2004)
			fusinF	GTCAACAAGATGGACCGCATG	55	This study
			fusinR	CATGCGGTCCATCTTGTTGAC	55	This study
mfd	Transcription repair coupling	222	recGF	GGGCGACGTGGGCDSNGGNAARAC	TCHDW (55)	Santos & Ochman (2004)
	factor.		recGR	GGGTCCGGGGGATNGGNGTNGC	TCHDW (55)	Santos & Ochman (2004)
	Helicase II superfamily		mfinF	ACCGCTTCGCCGACTGGC	55	This study
			mfinR	GCCAGTCGGCGAAGCGGT	55	This study
recA	Recombinase A	372	recABDUP1	CCCGAGTCCTCCGGNAARACNAC [a]	TCHDW (52)	Santos & Ochman (2004)
			recABGDN2	CGTTGCCGCCGGKNGTNRYYTC [a]	TCHDW (52)	Santos & Ochman (2004)
			recintAF	ACCGGYGAGCAGGCGCT	55	This study
			recintAR	AGCGCCTGCTCRCCGGT	55	This study
gyrB	DNA gyrase subunit B	840	gyrBBAUP2	GCGGAAGCGGCCNGSNATGTA [a]	TCHDW (50)	Santos & Ochman (2004)
			gyrBBndn1	CCGTCCACGTCGGCRTCNGYCAT [a]	TCHDW (50)	Santos & Ochman (2004)

[a] Degenerates primers sequences are underlined, these primers were amplified by touchdown PCR. [b] same annealing temperature for amplification and sequencing, in parentheses sequencing annealing temperature when amplification was made by touchdown PCR.

Table S2. Geographical coordinates of the *Allochromatium* strains used in this study

Locations	Strain/species name	Latitude	Longitude
NR	*Allochromatium vinosum* DSM 180[T]	NR	NR
NR	*Allochromatium minutissimum* DSM 1376[T]	NR	NR
America [a]	*Allochromatium sp.* strain AX1YPE	12°02'S	72°13'W
America [b]	*Allochromatium sp.* strain BH-2	23°58'N	77°42'W
Asia [c]	*Allochromatium renukae* DSM 18713[T]	16°54'N	82°14'E
NR	*Alochromatium vinosum* DSM 183	NR	NR
NR	*Allochromatium vinosum* DSM 1686	NR	NR
NR	*Allochromatium vinosum* DSM 182 [d]	NR	NR
NR	*Allochromatium vinosum* DSM 185 [e]	NR	NR
NR	*Allochromatium vinosum* DSM 186*	NR	NR
NR	*Allochromatium vinosum* DSM 189*	NR	NR
NR	*Allochromatium vinosum* DSM 195*	NR	NR

[a] Originally isolated from soft marine sediments, Callao, Perú. [b] Originally isolated from South Andros Black Hole, Bahamas. [c] Originally isolated from brackish water pond, Kakinada India. [d] Based on the 16S rRNA sequence comparison affiliates to the genus *Thiorhodococcus*.
[e] Based on 16S rRNA comparison affiliates to the genus *Marichromatium*. NR= not reported
* Strains which failed to grow in fresh Pfennig's medium.

Table S3. Range of the sequence similarity and amino acid identity of protein-coding genes and sequences similarity of 16S rRNA gene and ITS region.

Genes /ITS region	Sequence similarity (%)	Average similarity (%)	Amino acid identity (AAI) %	Average (AAI) %
pufM	92.9-100	97.5	100	100
soxB	89.5-100	93.7	88.7-99.4	94.2
cbbL-1	94.7-100	96.6	96.3-100	98.4
groES	93.5-100	96.7	96.7-100	98.8
fusA	90.8-100	95.0	95.7-100	98.5
mfd	89.6-100	94.4	97.3-100	99.2
recA	90.8-100	94.0	96.7-100	98.7
gyrB	89.0-100	95.2	91.8-100	96.7
16S rRNA	97.0-100	98.05	NA	NA
ITS	87.1-100	93.8	NA	NA

NA= no applicable

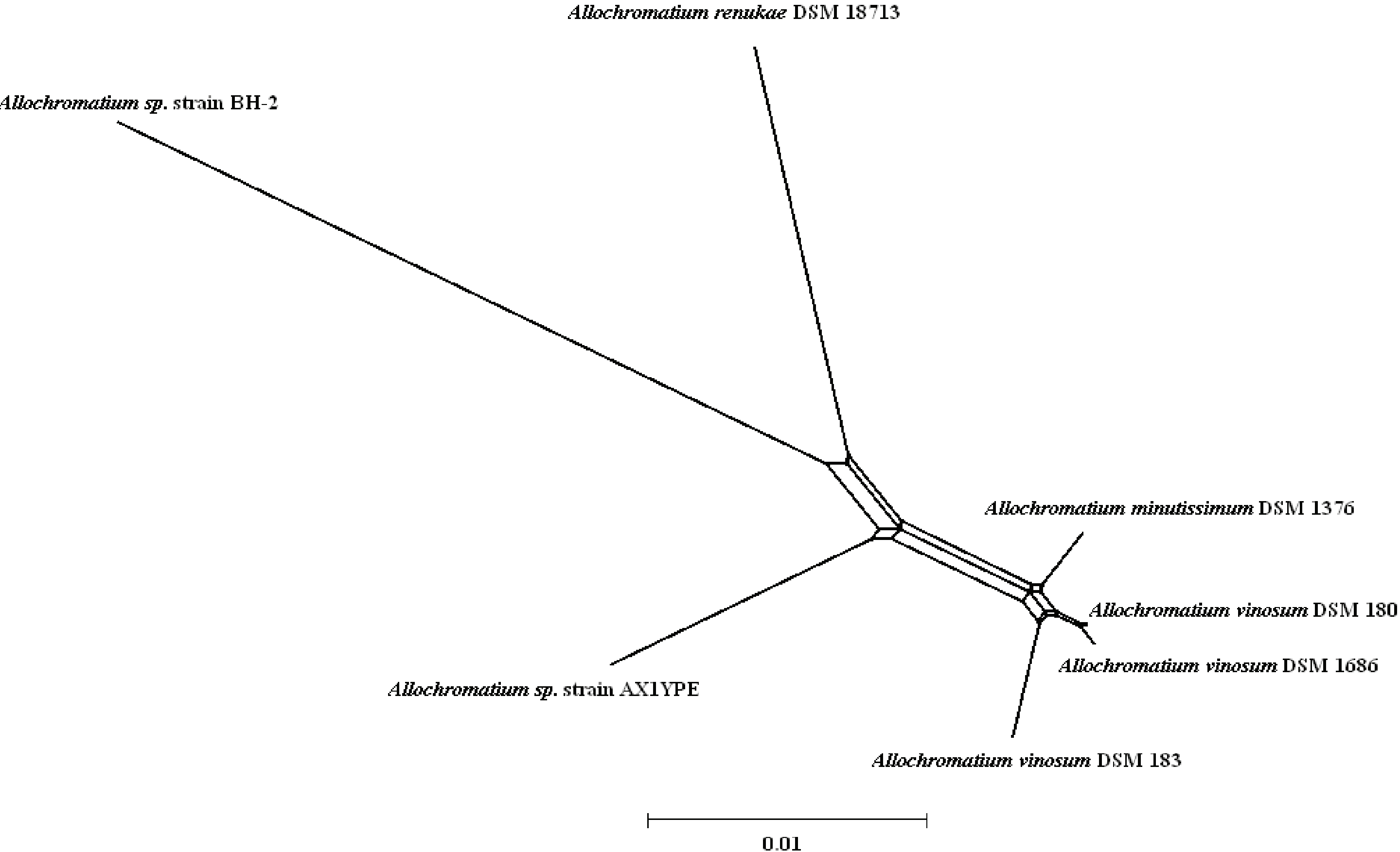


Figure S1 . Concatenated split network tree based on eight gene loci. The *pufM*, *soxB*, *cbbL-1*, *groES*, *fusA*, *recG*, ITS and 16S rRNA gene sequences (4427) from seven *Allochromatium* strains were concatenated and the split network tree was generated using the Split Tree4 program. Correction of the sequences similarities were made by using the method of Jukes and Cantor. Bar indicates 1% of sequence divergence.

Chapter 5

Allochromatium humboldtianum sp. nov. , a novel *Allochromatium* species isolated from soft marine sediments from Peruvian coast.

Wilbert Serrano[1, 2, *], Jan Schrübbers[1], Rudolf Amann[2] and Ulrich Fischer[1]

Manuscript in preparation

[1] Zentrum für Umweltforschung und nachhaltige Technologien, Fachbereich Biologie/Chemie, Abteilung Marine Mikrobiologie, Universität Bremen, D-28359 Bremen, Germany
[2] Max Planck-Institute for marine Microbiology, Celsiusstraße 1, D-28359 Bremen, Germany
Corresponding author, email: wserrano@mpi-bremen.de

Abstract. — A new purple sulfur bacterium, strain AX1YPE, was isolated from marine sediments from shallow waters near Callao bay, Peru. Strain AX1YPE grows anaerobically, synthesizes bacteriochlorophyll *a* and carotenoid pigments of the spirilloxanthin series. Cells are Gram-negative rods and actively motile by a polar flagellum. Strain AX1YPE is able to grow photolithoautotrophically with sulfide and thiosulfate as electron donors. During growth on sulfide, sulfur globules are stored inside the cells. The new organism uses amonium salt, N_2, urea, and nitrite as nitrogen sources.The organism is able to grow optimally at 15 to 60 µE m^{-2} s^{-1} photosynthetically active radiation (PAR) photon flux density. Photoassimilation of acetate and other carbon sources was possible at these low light intensities. Analysis of the (16S) rDNA gene sequence indicates that strain AX1YPE clusters in a separate branch within the genus *Allochromatium* of the family *Chromatiaceae.* Strain AX1YPE showed a 16S rRNA gene sequence silmilarity of 98.9% with *A. vinosum* and *A. minutissimum,* 96% with *A. renukae,* and 95% with *A. warmingii*. DNA-DNA hybridization (DDH) values to the closest relatives *A. vinosum* and *A. minutissimum* were 59 and 64% respectively. Based on phenotypic, phylogenetic, and DDH studies we describe strain AX1YPE as a new species of the genus *Allochromatium* for which the name *Allochromatium humboldtianum* sp. nov. is proposed.

Introduction

Species of phototrophic bacteria are spread all over the world where light is available and they comprise a wide phylogenetically diverse group. Purple sulfur bacteria (PSB) are anoxygenic photosynthetic prokaryotes which grow anaerobically in the presence of light and use of H_2S as an electron donor (Pfennig and Trüper, 1992). The class *Gamma-proteobacteria* has the two families *Chromatiaceae* and *Ectothiorhodospiraceae* as representative phylogenetic branches of PSB (Imhoff et al., 1998). A common feature of the *Chromatiaceae* is the intracellular storage of sulfur globules while the members of the other family deposite this outside.

Phylogenetic relationship among members of the family *Chromatiaceae* is mainly based on the 16S rRNA gene sequence analysis and the presence of some remarkable phenotypic properties (see Imhoff *et al*., 1998; Kumar *et al*., 2008) To date, the genus *Allochromatium* comprises the four recognized type species *A.*

vinosum DSM 180, *A. minutissimum* DSM 1376, *A. warmingii* DSM 173 (see Imhoff et al., 1998) and the more recently described *A. renukae* DSM 17378 (Kumar et al., 2008).

In the present comunication, a new purple sulfur bacterium, strain AX1YPE, isolated from soft marine sediments of shalow waters from Callao, Peru, is fully described using a polyphasic approach which also includes DDH comparison. To our knowledge, strain AX1YPE is the fisrt species of a marine representative of the genus *Allochromatium.*

Material and Methods

During the course of the research cruise along the coast of Peru on board of IMARPE R/V Jose Olaya in April 2005, sediment samples from nearshore waters of the fixed Station 1 (12°02'S 72°13'W) were obtained at 48 m depth. Undisturbed sediment samples were taken with a plexigas multicorer sampler. Aliquotes of the top surface sediments, which also contain *Thioploca* mats, were used for first enrichment cultures in 50 ml screw capped bottles containig anoxic basal medium as described by Pfennig and Trüper (1981) and 3% of NaCl. The medium was supplemented with sulfide (4 mM), thiosulfate (2 mM), and acetate (2 mM) final concentrations respectively. Initial incubation on board was done at room temperature under continuous illumination of a 40W tungsten lamp. In laboratory, a pure culture was obtained after successive re-purifications steps using both liquid cultures and deep agar shake dilution series. The agar shake method was also used to test the sensibility to oxygen. Optimal salinity, carbon substrate assimilation, and utilization of different electron donors were tested using anoxic basal medium prepared as mentioned above. Assimilation of several carbon sources was carried out in the presence or absence of $NaHCO_3$. Nitrogen source utilization was tested by replacing ammonium chloride by other nitrogen sources. Diazotrophy was tested by cultivating the organism under N_2 atmosphere confirmed by repeated sub-culturing (three times). All experiments were performed using completely filled 15 ml Hungate tubes and screw cap bottles under continuous light exposition of 100 µE m^{-2} s^{-1} PAR at 30°C. Growth was measured turbidometrically at 650 nm in a Pharmacia Biotech, Novaspec II model UV spectrophotometer.
Since strain AX1YPE was isolated from a high primary productivity environment where sun light is strongly absorbed and scattered by suspended particles in the

upper 20 m of the water column. Different light intensities from 1 to 100 μE m^{-2} s^{-1} PAR were tested in order to study the ability of strain AX1YPE to perform photosynthesis under low light conditions.

Cell morphology and motility were observed under a Zeiss phase-contrast microscope. Intra-cytoplasmic structures and flagella were observed by transmission electron microscopy (TEM). The *in vivo* absorption spectra of cell suspensions in a sucrose gradient according to Pfennig & Trüper (1992) were recorded using a Beckman DU 640 model UV spectrophotometer.

Genomic DNA was extracted according to Marmur (1961). Purity of the extracted genomic DNA was checked spectrophotometrically and by gel electrophoresis. Determination of the mol % G + C content was made by the thermal denaturation method according to Marmur & Doty (1961). DNA-DNA hybridization was performed following the renaturation kinetic method according to De Ley *et al*. (1970) and as reported by Serrano *et al*. (2009).

PCR amplification of the 16S rRNA gene of strain AX1YPE was made according to Muyzer *et al*. (1993) and sequencing of the PCR product was performed as described by Sanger *et al*. (1977). Sequences were aligned with those deposited in the geneBank data bases showing <0.5% of nucleotides ambiguities as reccomended by Stackebrandt *et al*. (2002). Sequence alignment and phylogenetic tree reconstruction were carried out as described previously by Serrano *et al*. (2009). The final phylogentic tree was constructed by the maximum likelihood method using the program Phy-ML (Guindon & Gascuel, 2003). In addition, PCR based fingerprinting methods were also employed to study the banding pattern differences between strain AX1YPE and two other closely related *Allochromatium* type species.

Results and discussions.

Morphological features are shown in Figures 1 and 2. Strain AX1YPE is a rod shaped organism of 2 x 3 μm cell size (Fig. 1), posesses one polar flagellum (Fig. 2), and the fine ultra-structure analysis indicate the presence of an internal membrane system of the vesicular type (TEM pictures are in process). The colour of the cell suspension was brownish-red and the presence of Bchl *a* was evidenced by peaks of absorption maxima at 371, 590, 798, and 854 nm while carotenoid pigments seem to be of the spirilloxanthin series with absorption peaks at 485 and 547 nm (data not shown).

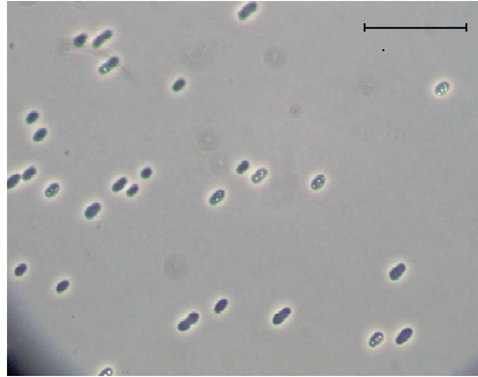


Figure 1. Phase-contrast photomicrograph of strain AX1YPE grown phototrophically with sulfide as electron donor, showing sulfur globules within the cells. Bar indicates 10 µm.

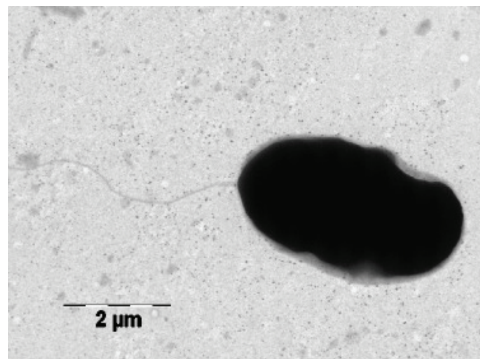


Figure 2. Electron microphotograph of a negative stained cell of strain AX1YPE showing a polar flagellum.

Strain AX1YPE is strictly anaerobic and is able to grow in a photo-lithoautotrophic fashion with sulfide as electron donor and CO_2. Best growth was observed when the organism grew photolithoheterothrophically with sulfide and thiosulfate as electron donors and pyruvate or acetate as carbon sources. The complete morphological, physiological, and molecular characteristics are summarized in Table 1. Sulfide was necessary for growth of strain AX1YPE under any growth condition tested, while sulfite, cysteine, and hydrogen did not support growth (data not shown). Although strain AX1YPE was isolated from marine environment, the organism was able to grow without NaCl, but tolerates up to 5% NaCl (optimum 3%). The pH for growth ranges from 6.5 to 8.5, with an optimum at 6.9-8.0. Good growth was observed at temperatures from 25-35°C (optimum 30°C). Strain AX1YPE does not require vitamins for growth. The DNA base composition of strain AX1YPE was 63.9 mol% G+C which is in the range for all reported *Allochromatium* strains (Table 1). However, it is remarkable that strain AX1YPE is able to grow at relative low light intensities and that this strain uses different nitrogen sources such as ammonium salts, glutamate,

nitrite, N_2 or urea. Both features have not been reported for other *Allochromatium* strains so far.

Since photosynthesis is a light dependent process, low light intensities may become a limiting factor. However, our results indicate that strain AX1YPE is well adapted to grow at low light intensities. Good growth was observed under photon fluxes between 15 to 60 µE m^{-2} s^{-1} PAR while at light intensities above 100 µE m^{-2} s^{-1} PAR growth is slow. These results give some hints about the importance of ecological adaptation of the new isolate to its environment.

Since the new strain shows 16S rRNA similarity of 98.9% to its closest relatives *A. vinosum* and *A. minutissimum* (Fig. 3), species identification was ensured additionally by using DNA-DNA hybridization studies as recommended by Imhoff & Caumette, (2004).

Table 1. Characteristics of different *Allochromatium* species

Character	*Allochromatium warmingii DSM 173**	*Allochromatium vinosum DSM 180**	*Allochromatium minutissimum DSM 1376**	*Allochromatium sp.* strain BH-2 *DSM 15591**	*Allochromatium renukae* JA136T*	*Allochromatium sp.* strain AX1YPE**
Cell shape	Rod	Rod	Rod	Rod	Avoid to rod	Ovoid to Rod
Cell size (width x length; µm)	3..5-4.0 x 5.0-11.0	2.0 x 2.5-6.0	1.0-1.2 x 2.0	1.8 – 2.5	2.0-2.5x3.0-5.0	2.0 x 3.0
Flagellation	ND	Single polar	Single polar	Single polar	ND	Single polar
Color of cell suspension	Purple-violet	Brown-red	Brown-red	Rose-red	Purple-violet	Brown-red
Carotenoides	Rhodopinal	Spirilloxanthin	Spirilloxanthin	Spirilloxanthin	Lycopene	Spirilloxanthin
NaCl optimum (%)	NR	0-1 %	none	2	None	3
Vitamin B12 requirement	+	-	-	-	-	-
Optimum temperature (°C)	25-30	25-35	25-35	25	28-32	25-35
Optimum pH	7.0 (6.5-7.3)	7.0-7.3 (6.5-7.6)	7.0-7.3 (6.5-7.6)	ND	7.2-8.0 (6.8-8.5)	7.0-7.5
Photolithoautotrophy	+	+	+	+	+	+
Photoheterotrophy	+	+	+	+	+	+
Nitrogen source utilized	Ammonium salts	Ammonium salts	Ammonium salts	Ammonium salts	Armonium salts, N_2	Ammonium salts, N_2 urea, glutamate, nitrite
Organic substrates utilized:						
Formate	-	+	+	+	-	+
Lactate	-	(+)	(+)	+	-	+
Fumarate	-	+	+	(+)	+	+
Malate	-	+	+	(+)	+	+
Pyruvate	NR	+	+	(+)	+	+
Glycolate	-	(+)	(-)	(-)	-	+
Fructose	NR	-	-	(-)	NR	-
Glucose	-	-	-	(-)	+	+
Molecular features						
DDH (%) relatedness	NA	59 ± 5.44	64 ± 2.38	54 ± 5.88	61 ± 3.59	100
16S rRNA (%) similarity	94.79	98.89	98.89	97.23	97.94	100
DNA G+C content (mol%)	55.1-60.2 (Tm)	64.3 (Tm)	63.7 (Tm)	65.7 (HPLC)	63.3 (HPLC)	63.9 (Tm)

*= Data taken from Pfennig & Trüper (1992); Herbert *et al*., (2005) and Kumar *et al*., (2008), organic substrates tested in this study are shown in parentheses. **= All organic substrates used in this study were tested during photolithoheterotrophic growth of strain AX1YPE with Na_2S x $9H_2O$ (5mM). NA= not applicable. NR= not reported.

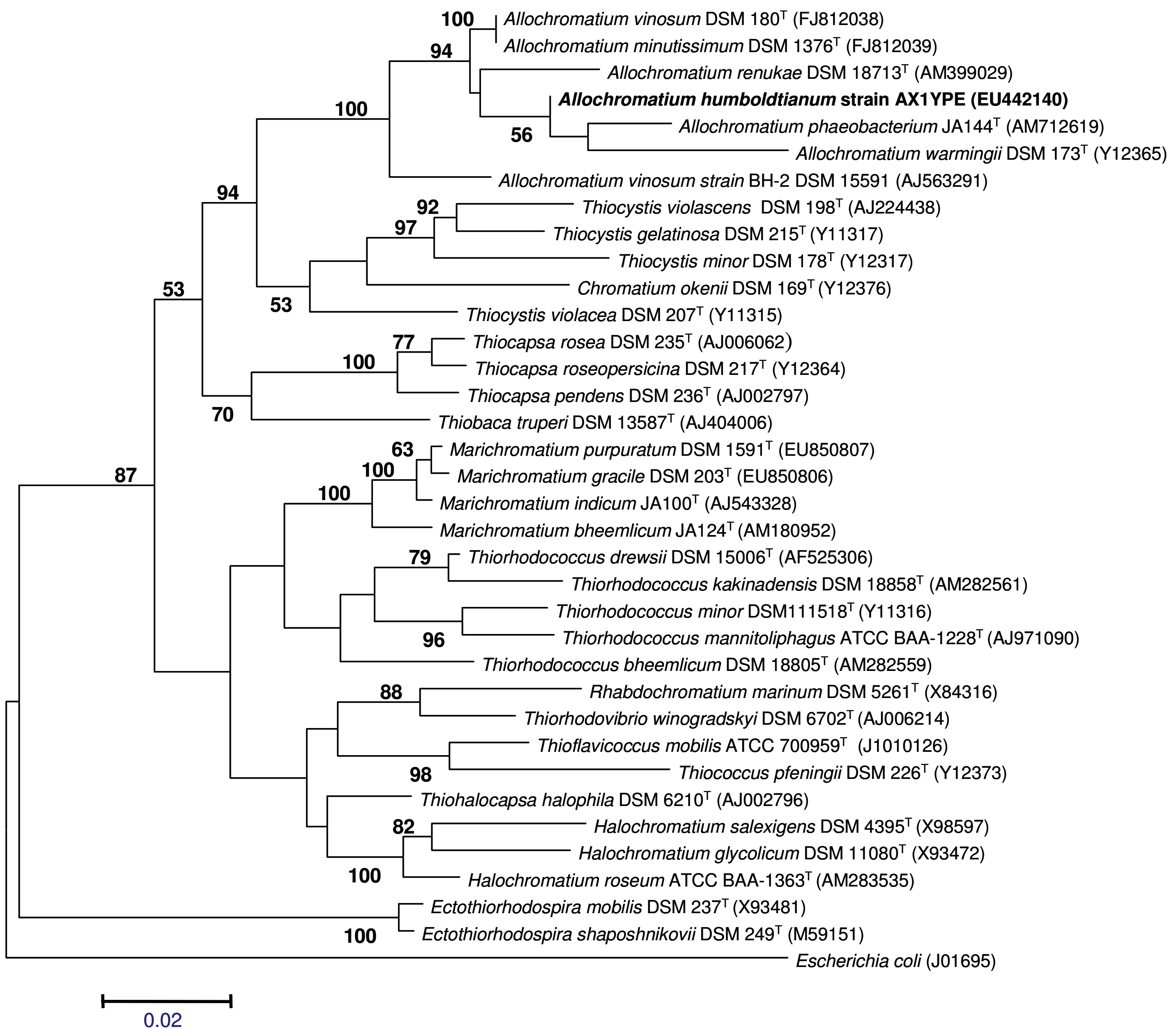


Figure 4: 16S rRNA-based tree reflecting the phylogenetic relationship of strain AX1YPE and reference *Chromatiaceae* species. The tree is the result of maximum likelihood method using the program PhyML of nearly full length 16S rRNA sequences. The bar indicates 2% of estimated sequence divergence. Accession numbers of the reference sequences are given in parentheses.

The DDH comparison gave values of 59 and 64% of relatedness to both type species *A. vinosum* and *A. minutissimum*. Furthermore, DDH studies clearly revealed that strain AX1YPE showed DDH values below 70% to all compared type strains (Table 1). In all cases, DDH values were below the accepted threshold limit for species delineation according to Wayne *et al*. (1987). Finally, it is important to remark that the presence of one or more relevant properties which are unique to strains could also be used as an important factor to consider when arranging new isolates as new species (Imhoff & Caumette, 2004). Thus, the ability to grow at low light intensities and the use of a variety of nitrogen source, clearly distinguish strain AX1YPE from other *Allochromatium* type strains. Therefore, strain AX1YPE is sufficiently distinct and should be considered as a new species within the genus *Allochromatium* for which the name *Allochromatium humbodltianum* sp. nov. is proposed.

Description of *Allochromatium humboldtianum* sp. nov.

Allochromatium humboldtianum (hum.bold-ti-a'num. N. L. neut. adj. humboldtianum, named after the famous German (Prussian) naturalist Alexander von Humboldt, who described the cold ocean current of the South Pacific, Humboldt current also named Peru current).

Cells are Gram-negative rods, measuring about 2 x 3 µm and are motile by a polar flagellum. Culture suspension is brownish-red. Pigments include the bacteriochorophyll *a* and carotenoids of the spirilloxanthin series and the photosynthetic membrane system consisted of the vesicular type. No vitamin requirement for growth, strictly anaerobic. Growth modes are photolithoautotrophy, photolithoheterotrophy and chemomixotrophy. Strain AX1YPE is able to grow at low light intensities (16 - 60 µE m^{-2} s^{-1} PAR). Sulfide and thiosulfate are electron donors and elemental sulfur formed as intermediate oxidation product is stored inside the cells. Organic substrates photo-assimilated in the presence of sulfide without carbonate are: formate, lactate, fumarate, malate, pyruvate, and glucose. The organism is able to use different nitrogen compounds. Diazotrophic growth was also possible. Salinity range 0-5% NaCl (optimum 3%). Temperatures range from 25-35°C (optimum 30°C), pH range 6.5-8.5 (optimum 6.9-8.0). Strain AX1YPE tolerates significant levels of sulfide up to 8 mM. The G+C content of the DNA is 63.9 mol% (Tm). The new strain is AX1YPE (=DSM 21881^T)

Habitat: Strain AX1YPE was isolated from soft marine sediments of shallow water from Callao Bay, Perú.

Acknowledgement.

We thank the captain and the crew of the R/V "Jose Olaya", our scientific colleagues at Instituto del Mar del Peru (IMARPE), and all the scientific team on board at April 2005 cruise. Funding was provided by the Max Planck Society; W.S. was supported by a fellowship from the "Zentrale Kommission für Forschung und wissentschaflichen Nachwuchs" of Bremen University and the Max Planck Society. We are grateful to Hans G.Trüper, Universität Bonn, Germany, for suggesting the specific epithet. W. S. was also a member of the International Max Planck Research School of Marine Microbiology (MarMic).

References

De Ley, J., Catoir, H. & Reynaerts, A. (1970). The quantitative measurement of DNA hybridization from renaturation rates. Eur. J. Biochem. 12: 133-142.

Guindon, S. & Gascuel, O. (2003). A simple, fast and accurate algorithm to estimate large phylogenies by maximal likelihood. Syst. Biol. 52, 696-704.

Herbert, R. A., Ranchou-Peyruse, A., Duran, R., Guyoneaud, R. & Schwabe, S. (2005). Characterization of purple sulfur bacteria from the South Andros Black Hole cave system: highlights taxonomic problems for ecological studies among the genera *Allochromatium* and *Thiocapsa.* Environ. Microbiol. 7, 1260-1268.

Imhoff, J. F. & Caumette, P. (2004). Recommended standards for the description of new species of anoxygenic phototrophic bacteria. Int. J. Syst. Evol. Microbiol. 54: 1415-1421.

Imhoff, J. F., Süling, J. & Petri, R. (1998). Phylogenetic relationships among *Chromatiaceae*, their taxonomic reclassification and description of the new genera *Allochromatium, Halochromatium, Isochromatium, Marichromatium, Thiococcus, Thiohalocapsa and Thermochromatium*. Int. J. Syst. Bacteriol. 48:1129-1143.

Kumar, P. A., Srinivas, T. N. R., Sasikala, Ch. & Ramana, Ch. V. (2008). *Allochromatium renukae* sp. nov. Int. J. Syst. Evol. Microbiol. 58, 404-407.

Marmur, J. (1961). A procedure for the isolation of desoxyribonucleic acid from micro-organisms. J. Mol. Biol. 3, 208-218.

Marmur, J., & Doty, P. (1962). Determination of the base composition of desoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. 5, 109-118.

Muyzer, G., De Waal, E. C. & Uitterlinden, A. G. (1993). Profiling of complex microbial population by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. 59: 695-700.

Pfennig, N. & Trüper, H. G.. (1981). Isolation of members of the families *Chromatiaceae* and *Chlorobiaceae*. In: Starr, M. P., Stolp, H., Trüper, H.G., Balows, A., Schlegel H.G. (eds.), The Prokaryotes. Springer, Berlin, Heidelberg, New York, pp. 279-289.

Pfennig, N. & Trüper, H. G. (1992). The family *Chromatiaceae*. In: Balows, A., Trüper, H.G., Dworkin, M., Harder, W., Schleifer, K. H. (eds.), The Prokaryotes. Springer, Berlin, Heidelberg, New York, pp. 3200-3221.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminator inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.

Stackebrand, E., Frederiksen, E. W., Garrity, G. M., Grimont, P. A. D., Kämpfer, P., Maiden, M. C. J., Nesme, X., Rosello-Mora, R., Swings, J., Trüper, H. G., Vauterin, L., Ward, A. C. & Whitman, W. B. (2002). Report of the *Ad Hoc* committee for the re-evaluation of the species definition in bacteriology. Int. J. Syst. Evol. Microbiol. 52: 1043-1047.

Serrano, W., Amann R. & Fischer, U. (2009). A new moderately thermophilic and high sulfide tolerant biotype of *Marichromatium gracile*, isolated from tidal sediments of the German Wadden Sea: *Marichromatium gracile* biotype *thermosulfidiphilum*. Syst. Appl. Microbiol. 32, 1-7.

Wayne. L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. & Trüper, H. G. (1987). Report of the *Ad Hoc* comittee on reconciliation of approaches to bacterial systematics. Int. J. Syst. Bacteriol. 37: 463-464.

Supplementary figure

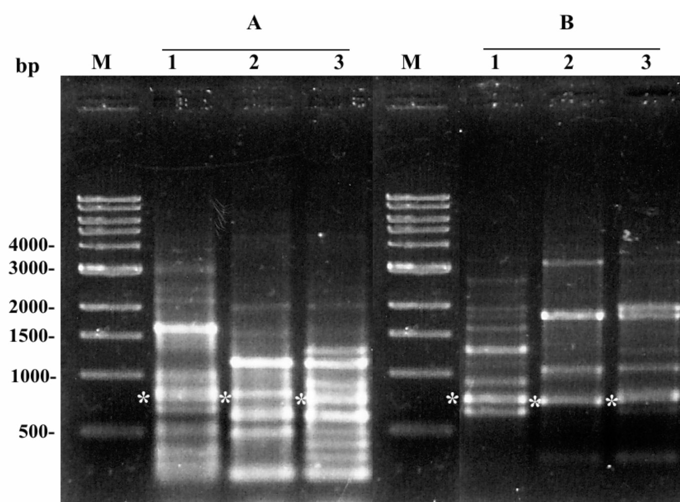


Fig. S1.Variation of the fingerprint patterns generated using the two different primer sets BOX-A1R (A) and RAPD (SMO3) (B) among three *Allochromatium* spp. Lanes 1, Strain AX1YPE; lanes 2 *Allochromatium vinosum* DSM 180^T; lanes 3, *Allochromatium minutissimum* DSM 1376^T. Lines M, 1kb ladder.

Appendices

Appendix 1

Microorganisms

28. PFENNIG'S MEDIUM I
(modified 1988, for purple sulfur bacteria)

Solution A:

$CaCl_2 \times 2\ H_2O$	1.25	g
KH_2PO_4	1.70	g
NH_4Cl	1.70	g
KCl	1.70	g
$MgSO_4$	2.50	g
Distilled water	4000.00	ml

(For marine or estuarine isolates add 100.0 g NaCl to this solution and increase the $MgSO_4 \times 7\ H_2O$ to 15.0 g).

Solution B:

Distilled water	860.00	ml

Autoclave in a cotton-stoppered Erlenmeyer flask and cool to room temperature under an atmosphere of N_2 in an anaerobic jar.

Solution C:

Vitamin B_{12} solution (0.002% in H_2O)	5.00	ml

Filter sterilize.

Solution D:

Trace element solution (SL-12 B)	5.00	ml

Autoclave at 121°C for 15 min.

Solution E:

$NaHCO_3$	7.50	g
H_2O	100.00	ml

Bubble with CO_2 and, after saturation, filter sterilize under CO_2 pressure into sterile, gas-tight, 100 ml screw-cap bottle.

Solution F:

$Na_2S \times 9\ H_2O$ (10 g in 100 ml)	20.00	ml

Prepare in a screw-cap bottle, bubble with N_2 to replace air, close tightly and autoclave.

Continued on next page

Microorganisms

Trace element solution SL-12 B:

Distilled water	1000.00	ml
Na_2-EDTA	3.00	g
$FeSO_4$ x 7 H_2O	1.10	g
$CoCl_2$ x 6 H_2O	190.00	mg
$MnCl_2$ x 2 H_2O	50.00	mg
$ZnCl_2$	42.00	mg
$NiCl_2$ x 6 H_2O	24.00	mg
Na_2MoO_4 x 2 H_2O	18.00	mg
H_3BO_3	300.00	mg
$CuCl_2$ x 2 H_2O	2.00	mg

Adjust pH to 6.0.

Autoclave solution A for 45 min. in 5-litre special bottle or flask (with four openings at the top) at 121°C, together with a teflon-coated magnetic bar. In this 5-litre bottle, two openings for tubes are in the central, silicon rubber stopper; a short, gas-inlet tube with a sterile cotton filter; and an outlet tube for medium, which reaches the bottom of the vessel at one end and has, at the other end, a silicon rubber tube with a pinch cock and a bell for aseptic dispensing of the medium into bottles. The other two openings have gas-tight screw caps; one of these openings is for the addition of sterile solutions and the other serves as a gas outlet.

After autoclaving cool solution A to room temperature under a N_2 atmosphere with a positive pressure of 0.05 - 0.1 atm (a manometer for low pressure will be required). Saturate the cold medium with CO_2 by magnetic stirring for 30 min. under a CO_2 atmosphere of 0.05 - 0.1 atm. Add solution B, C, D, E and F through one of the screw-cap openings against a stream of either N_2 gas or better, a mixture of 95% N_2 and 5% CO_2 while the medium is magnetically stirred.

Adjust the pH of the medium with sterile HCl or Na_2CO_3 solution (2 mol/liter each) to pH 7.3. Distribute the medium aseptically through the medium outlet tube into sterile, 100 ml bottles (with metal caps and autoclavable rubber seals) using the positive gas pressure (0.05 - 0.1 atm) of the N_2/CO_2 gas mixture: Leave a small air bubble in each bottle to meet possible pressure changes. The tightly sealed, screw-cap bottles can be stored for several weeks or months in the dark. During the first 24 h, the iron of the medium precipitates in the form of black flocks. No other sediment should arise in the otherwise clear medium. Incubate in the light using a tungsten lamp. Feed periodically with neutralized solution of sodium sulfide (see medium 27) to replenish sulfide and with other supplement solutions (see Ref. 3365).

Appendix 2

Marmur method for DNA extraction (modified fishing method)
Preparation and solutions:

1. Suspention buffer (50ml)
   - Tris HCl (pH 8) 20 mM final concentration
   - EDTA (pH 8) 5 mM final concentration
   - Sucrose 6 g.

   Complete to 50 ml with filtered distilled water.

2. Lysis buffer (50 ml)
   - Tris-HCl (pH 8) 100 mM final concentration
   - EDTA (pH 8) 40 mM final concentration
   - NaCl 0.5 M final concentration
   - SDS 4% final concentration

   Complete to 50 ml with filtered distilled water. Add 1 ml of β-mercapto-ethanol and 250 μl of proteinase K (20mg/ml) to this solution aliquote and store at -20°C, stable several months.

3. SSC-Buffer: 20x stock solution (3M NaCl, 0.3 M Na-citrate)

   - dissolve 175.3 g NaCl and 88.2 g of Na-citrate in 800 ml milliQ water
   - add milliQ water to 800 ml
   - adjust the pH to 7 and complete to 1L
   - autoclave

Appendix 3.

Protocol for the G+C mol% determination taken from theDepartment of Marine Microbiology of the University of Bremen.

Spectroscopic and Thermal Properties of Nucleic Acids

Spectroscopic properties and purity

Nucleic acids absorb ultraviolet radiation (light) in the wavelength range of 200 – 300 nm, due to the conjugated aromatic nature of the bases. The wavelength of maximum absorption by both DNA and RNA is 260 nm, which is distinct from the λ_{max} = 280 nm of protein. The absorption properties of nucleic acids can be used for detection, quantitation and assessment of purity.
The shape of the spectrum is the same for both RNA and DNA. Nucleic acids can be detected at concentrations as low as 2.5µg/ml. A solution of ds DNA with a concentration of 50 µg/ml has an A_{260} of 1.0. The corresponding value of RNA or ssDNA is approximately 2.5. UV absorbance can also be used to estimate the purity of DNA preparations.
A_{320} should be < 0.1 values, greater than this indicate particulates or other substances in the preparation. Most important is the determination of the ratio of absorbance at 260 nm and 280 nm (A_{260}/A_{280}). Pure dsDNA has an A_{260}/A_{280} around 1.8, and pure RNA one around 2.0. Protein with its λ_{max} = 280 nm has a 260/280 ratio < 1.0. Therefore, if a DNA solution has an A_{260}/A_{280} greater than 1.8, this suggests RNA contamination, whereas one clearly less than 1.8 suggests the presence of undesired proteins in the sample. Because many proteins do not contain high amounts of aromatic amino acids, a more suitable indicator of protein contamination is the A_{234}/A_{260} ratio. Nucleic acids have an absorbance minimum at 234 nm and the presence of proteins results in an increase in this ratio. Generally the ratio should be < 0.5.

Thermal denaturation

Increased temperature as well as a number of chemicals can bring about the denaturation of nucleic acids. The strands of dsDNA dissociate when the hydrogen bonds between the base pairs are loosened. This can be achieved by heating double-stranded nucleic acids.

The thermal behaviors of dsDNA and RNA are very different. If the temperature increases, the absorbance of an RNA sample gradually and erratically increases as the stacking of the bases in double-stranded regions is reduced. Shorter regions will denature before longer ones.
The thermal denaturation, also called melting of ds DNA is co-operative. The denaturation of the ends of the double-helix, and of more mobile AT-rich regions, will destabilize adjacent regions, leading to a progressive and concerted melting of the whole structure at a well-defined temperature corresponding to the midpoint of the smooth transition, and known as the melting temperature (T_m). The melting temperature is defined as that temperature at which half of the helical structure is lost. The melting of dsDNA can be easily detected by measuring the absorption at 260 nm. The melting is accompanied by a 40% increase in absorbance. T_m is a function of the G+C content, and ranges from 80°C to 100°C for long DNA molecules. It is obvious that large numbers of G-C base pairs increases the T_m of the DNA, while DNA with mainly A-T composition has a lower melting point.

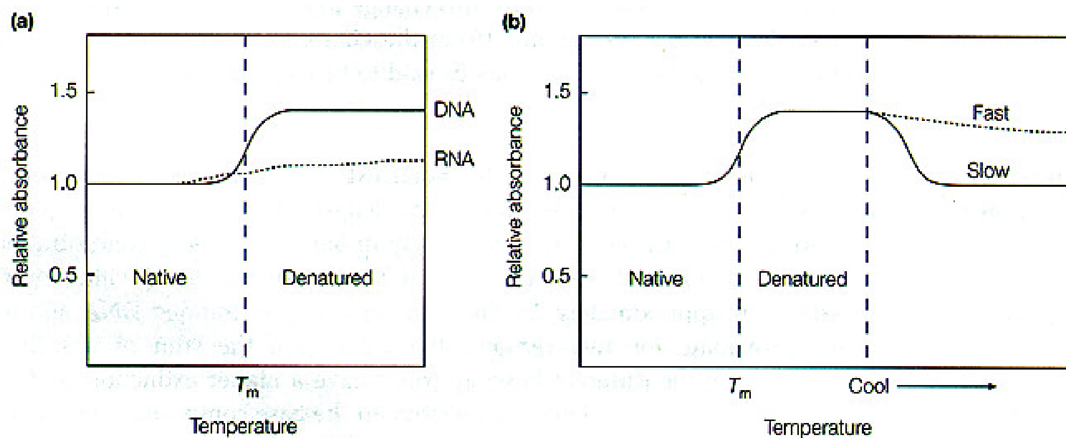


Fig. (a) Thermal denaturation of double-stranded DNA (co-operative) and RNA; (b) renaturation of DNA by fast and slow cooling. Turner, P.C. et al. Molecular Biology 2ndEd. BioScientific Publ.Oxford 2000

Determination of mol % G+C in DNA samples

It is necessary to use DNA of high molecular weight. Normally the quality of DNA is checked on an 1% agarose gel. A relatively sharp band should be visible with little or no smearing. Because of lack of time we ommit this procedure in this course.

But it is absolutely necessary to check the DNA for protein and RNA contamination, by determining the A_{260}/ A_{280} ratio. If this ratio is above 1.9, incubate the sample with RNase and repeat the extraction procedure. Contaminants can block or distort the light beam when the samples are scanned.

Using an appropiate aliquot, dilute each DNA sample in 0.1x SSC to give a final concentration of 20 µg/ml [A_{260}= 1.0 ≈ 50µg dsDNA/ml]. Determine the absorbance of each sample at 260 nm. Dilutions should be made with 0.1x SSC, which also serves as a blank. For the samples having an absorbance between 0.5 and 1.0 determine the absorption spectrum from 200 nm to 320 nm.

All DNA samples must be degassed in an exsiccator before preparing the T_m cell holder and the cuvettes, otherwise air bubbles may form in the cuvette, resulting in inaccurate readings.

Samples for Mol % G + C determinations will be diluted in 0.1x SSC for the spectrophotometry scans. The dissolved salt concentration in the SSC will affect the melting temperature (Tm) of the sample. It is important that the salt concentration of each sample is the same. If you are extracting DNA only for Mol % G + C determinations, dissolve the DNA in 0.1x SSC. DNA dissolved in low salt buffers are also useable if DNA concentrations are high enough to reduce carry over of salts with the buffer

Every spectrophotometer is different, so it will be necessary to check the sensitivity and accuracy of the scans using *E. coli* as the standard. *E. coli* DNA has a melting temperature of approximately 78 °C. Most of the published methods recommend relatively high concentrations of 20 µg/ml of DNA for the measurement. However, most of these papers, laboratory manuals etc. were written more than 20 years ago and the quality of the spectrophotometers has improved since then. It may be possible to use lower concentrations of DNA. A dilution series using the same *E. coli* sample is suggested to test the sensitivity of' your spectrophotometer.

The Beckman Spectrophotometer DU600 is already tested for sensitivity and accuracy using a dilution series (25, 12.5, 6.25, and 3.125 µg/µl) of *E. coli* strain K12 DNA. The lower limit appears to be 12.5 µg/ml for the Beckman Spectrophotometer DU 600.
You should chose a temperature window of 50 to 90 °C for your scans. If you suspect that you have low G + C DNA then start your scan at a lower temperature.

For the Beckman Spectrophotometer DU600 the following program is suggested:

- Start Temperature: 50° C
- Delay: 5 minutes
- Ramp rate: 1°C/min.
- Final Temperature: 90°C
- Read interval: 0.5° C
- No Matching
- No Smoothing

Do not fill the cuvettes completely. The liquid will expand when heated. 300 µl was optimal for the cuvettes. If the cuvettes are not nearly full, water condenses from the sample during analysis, causing an increase in sample concentration and absorption. Each of the cuvettes must have a solution in it, even if it is not being analyzed. Load the cuvettes in the following order:

Cuvette 1	0.1x SSC Blank
Cuvette 2	*E. coli DNA*
Cuvette 3	Sample DNA
Cuvette 4	Sample DNA
Cuvette 5	Sample DNA
Cuvette 6	Sample DNA

Run the scan. This will take a little more than an hour. When finished, calculate the T_m using the first derivative plot function.

To calculate the Mol % G + C of your sample, use the following equation from Marmur and Doty (1962)

$$\text{Mol \% G+C} = \frac{[T_m \text{ sample} + (90.5 - T_m \textit{ E.coli})] - 69.3}{0.41}$$

It is important to use the T_m for the *E. coli* that was scanned with your samples**,** not the general T_m of *E. coli* from the literature.

Acknowledgments

I would like to thank Prof. Ulrich Fischer and Prof. Rudolf Amann for their supervision during the development of my thesis and for giving me the opportunity to work in a nice cooperative interaction between the Department of Marine Microbiology at Bremen University and the Department of Molecular Ecology at MPI; Prof. Ramon Rosello-Mora and Dr. Karl-Heinz Blotevogel who kindly accepted to participate in the evaluation of this work; Jörg Wulf the commandant in chief of technicians of the molecular ecology group at MPI, whenever I needed a help in the lab. he was there, I extend this acknowledgement to other members of the molecular ecology group technicians, scientist, and colleagues PhD students for the nice working atmosphere, sorry if I do not mention a name in particular; my colleagues at the Department of Marine Microbiology in the UFT building; special thanks to Birgit Heyduck-Söller and Annina Hube my closest officemates, who shared with me a wonderful time, and for making my rare darkness days into sunny days, thanks a lot for your warm friendship; Tina Stickan, who cast her expert eye over my writings for catching "spling mstakes". Finally, special acknowledgment to my MarMic colleagues, classmates and "new MarMics", without them it wouldn`t have been as much fun.

Major funding for my fellowship was provided from the Zentrale Kommission für Forschung und wissenschaflichen Nachwuchs of Bremen University. Further support came also from the Max Planck Society. Last but not least, I acknowledge to those anonymous faculty members who evaluated me and decided to include me in the nascent MarMic School. This work was done between July 2005 and June 2009.