

Europäische Akademie

zur Erforschung von Folgen
wissenschaftlich-technischer Entwicklungen
Bad Neuenahr-Ahrweiler GmbH

Direktor:

Professor Dr. Carl Friedrich Gethmann

The Species Concept in Prokaryotic Taxonomy

Contribution to the Project „Biodiversity“

by

Ramon Rosselló Mora and Rudolf Amann

August 2000

Europäische Akademie

zur Erforschung von Folgen
wissenschaftlich-technischer Entwicklungen
Bad Neuenahr-Ahrweiler GmbH

Direktor:

Professor Dr. Carl Friedrich Gethmann

The Species Concept in Prokaryotic Taxonomy

Contribution to the Project „Biodiversity“

by

Ramon Rosselló Mora and Rudolf Amann

August 2000

Publisher:

Europäische Akademie

zur Erforschung von Folgen
wissenschaftlich-technischer Entwicklungen
Bad Neuenahr-Ahrweiler GmbH

Postfach 14 60, D-53459 Bad Neuenahr-Ahrweiler
Telefon: ++49 - (0)2641 - 973 - 300, Telefax -320
e-mail: europaeische.akademie@dlr.de

Director:

Professor Dr. Carl Friedrich Gethmann (V.i.S.d.P.)

ISSN 1435-487 X

Editing:

Dagmar Uhl, M. A.

Print:

Druckerei Martin Warlich, Bad Neuenahr-Ahrweiler

PREFACE

In 1997 the Europäische Akademie initiated an interdisciplinary research project dealing with the scientific foundation and the social relevance of biodiversity. The interdisciplinary working-group brought together scientists from several biological and non-biological disciplines such as zoology, botany, microbiology, genetics, palaeontology, jurisprudence, economy and philosophy of science. During the term of the project, several short-studies were carried out that dealt with specific problems of biodiversity research respectively. The results of these studies were integrated in the main-study, but for sake of comprehensibility and lucidity many details of the underlying specific problems had to be omitted. Nevertheless, these problems are important; so, the Europäische Akademie decided to publish the respective short-studies in its "Graue Reihe" starting with Nr. 5 "Zur Wissenschaftstheorie der Genetik. Materialien zum Genbegriff" and continued with Nr. 11 "Biodiversitätsforschung in Deutschland. Potentiale und Perspektiven" and Nr. 15 "Biodiversität als Problem der Naturethik. Literaturreview und Biobibliographie".

Due to the fact, that the "species" is assumed to be a significant biological unit for the description of biodiversity, the working-group focused on the methodological and procedural problems of modern species concepts. The results of the working-groups constitutive symposium on this issue in 1998 were published in *Theory in Biosciences*, Vol. 117 (1998) 3.

Rossello-Moras work on "The species concept in prokaryotic taxonomy" presents a comprehensive overview of the history of species concepts in microbiology, the procedures that are applied in modern research and finally the empirical as well as methodological shortcoming of some attempts to identify species among prokaryotes. The history of prokaryote classification itself shows that the development of species characterisation is tightly connected with the development of microbiological techniques. This is true also for the current approaches, which are based on genetic analysis. However, the genetic ana-

lysis alone does not provide sufficient knowledge for the exact circumscription of a prokaryotic species. Because as many additional phenotypic characters as possible have to be taken into account, the prokaryotic species concept can be characterised as a phenetic-polythetic concept. From a methodological point of view the procedures that are applied for the circumscription of a prokaryotic species provide an excellent example for a "pragmatic" approach. Because "species" are determined referring to explicit goals a "natural system" of organisms is not a necessary prerequisite. The usefulness of the resulting sets of prokaryotes for scientific as well as medical purposes is the most relevant criterion for the validity of the respective systematisation.

It may seem that the topic of this short-study is rather untypical for the immediate purposes of classical technology assessment, it should be regarded as a contribution to the understanding of the startling complexity we have to face, when dealing with biodiversity concepts. However, specific scientific knowledge, such as microbiological considerations in this case, can provide a basis and a starting point for competent and successful technology assessment.

Mathias Gutmann and Peter Janich

Marburg, August 2000

CONTENTS

| | | |
|-----------|--|----|
| 1- | Introduction | 7 |
| 2- | A short history of prokaryote classification | 10 |
| 3- | Methods and parameters used in prokaryotic classification | 16 |
| | 3.1- Genotypic methods | 17 |
| | 3.1.1- DNA base ratio | 18 |
| | 3.1.2- DNA-DNA similarity | 19 |
| | 3.1.3- Ribosomal RNA analysis | 22 |
| | 3.1.4- DNA-based typing methods | 25 |
| | 3.2- Phenotypic methods | 27 |
| | 3.2.1- Classical phenotypic analyses | 28 |
| | 3.2.2- Numerical taxonomy applied to phenotypic analyses | 29 |
| | 3.2.3- Chemotaxonomy | 31 |
| | 3.2.4- Phenotype typing methods | 32 |
| | 3.2.5- Identification keys and diagnostic tables | 33 |
| | 3.2.6- Microbial identification systems | 33 |
| 4- | Prokaryotic species concept | 35 |
| | 4.1- The concept | 36 |
| | 4.2- The borders | 38 |
| | 4.3- The 16S rRNA sequence data | 39 |
| | 4.4- Intraspecific subdivisions | 40 |
| | 4.5- Prokaryotic sex | 42 |
| | 4.6- Pure cultures and the culture collections | 44 |
| | 4.7- The species concept to be adopted for prokaryotes | 47 |
| | 4.8- Towards an improvement of the species circumscription | 50 |
| 5- | Conclusions | 54 |
| 6- | Glossary | 55 |
| | Bibliography | 57 |

The species concept in prokaryotic taxonomy

'Species would appear to have become a concept that dare not speak its name' Goodfellow et al. 1997 [26]

1- Introduction

Antonie van Leeuwenhoek first observed bacteria just around three hundred years ago, setting then the starting point of Microbiology. Since then, Microbiology as a science has developed enormously being one of the most advanced disciplines among the natural sciences. Prokaryotes, or cells without nucleus, have turned to be a very important group of living organisms for the everyday routine studies on basic and applied research like medicine, biotechnology and environmental biology.

It is surprisingly true that prokaryotes, nearly all invisible to the human eye, constitute an essential component of the earth's biota. They catalyze unique and indispensable transformations in the biogeochemical cycles of the biosphere, produce important components of the earth's atmosphere, and represent a large proportion of life's genetic diversity. In contrast to textbook knowledge, the estimated $4-6 \times 10^{30}$ prokaryotic cells existing on the earth (Table 1) might even constitute >50% of protoplasmic biomass (excluding most of the plant biomass that is made up of extra-cellular material like cell walls and structural polymers [110]).

Table 1. Number and biomass of prokaryotes in the world (data obtained from [110]).

| Environment | N° of prokaryotic cells $\times 10^{28}$ | g of C $\times 10^{15}$ |
|------------------------|--|-------------------------|
| Aquatic habitats | 12 | 2.2 |
| Oceanic subsurface | 355 | 303 |
| Soil | 26 | 26 |
| Terrestrial subsurface | 25-250 | 22-215 |
| Total | 415-640 | 353-546 |
| Plants* | – | 562 |

* Plant carbon as a sum of protoplasmic biomass, structural polymers and cell wall material.

It is also interesting to note that the majority of the prokaryotes may be located in oceanic and terrestrial subsurface environments. These habitats have been paid less attention than those of the earth's surface. Thus, research in microbiology might have dealt with the smaller part of the total Earth's prokaryotic community. This might explain in part why the total number of recognized prokaryote species does seemingly not account for a significant proportion of the total Earth's biodiversity (Table 2). To the date less than 5000 prokaryotes have been described with validly approved names. This number is linked to the isolation and characterization of microorganisms in pure cultures, an indispensable requisite for the recognition of prokaryotic species [26].

Table 2. Estimates of proportions of species of major groups of organisms contributing to the total of living biological diversity [12].

| Organism | Percent of contribution |
|---------------------|-------------------------|
| Vertebrates | 0.4 |
| Insects | 64.4 |
| Other arthropods | 7.7 |
| Other invertebrates | 6.7 |
| Plants | 2.4 |
| Algae | 1.6 |
| Protozoa | 1.6 |
| Fungi | 8.0 |
| Bacteria | 3.2 |
| Viruses | 4.0 |

The present last two decades saw the development of molecular techniques which application lead to quite unexpected discoveries like the recognition of the three domain tree of life [109], or the discovery of a very large potential prokaryote diversity made up of hitherto uncultured microorganisms [66, 67, 97, 106]. Both discoveries deeply impacted on prokaryote systematics and ecology, especially on the former one which is still under hot debate [30, 58, 111]. Several attempts to estimate the amount of species living on Earth have been made, and it seems that the

amount of recorded species is directly related to the effort and interest of the scientists [54]. Molecular techniques, most notably those based on 16S rRNA, which are directed to analyze community composition of environmental samples indicate that the hitherto classified prokaryotic species account for a very small portion of the real prokaryote diversity [11]. The contribution of prokaryotes to the total Earth's diversity has obviously been drastically underestimated.

The species concept for prokaryotes has been developed in parallel to the design of laboratory techniques that permitted the retrieval of useful information. Original species concepts based on morphological traits demonstrated to be wrongly tailored. Thus, its improvement has been fed back by the use of new information units (e.g. chemotaxonomic markers, DNA properties, rRNA sequences...). To the date, there is an agreement among the prokaryote taxonomists that the current species circumscription, although not perfect, is acceptable and pragmatic, and covers the primary goals of taxonomy like e.g. a rapid and reliable identification of strains [92]. There is, however, a strong criticism to the presently used prokaryotic species concept by non-taxonomists. For some of them, the current concept is too conservative leading to an underestimation of the real prokaryote diversity [19, 110]. They consider it a significant disadvantage that it is different from the concepts designed for eukaryotes [94, 105]. The current discussion about the adequacy of species concepts is not restricted to microbiologists, but there is also an open discussion among eukaryote taxonomists [12]. It indeed is tempting to install a more universal concept that covers the different major groups of organisms making the species units comparable. In this respect, it should be helpful to analyze in more detail the current prokaryote species concepts, its history and potential future. This is the aim of this study.

2- A short history of prokaryote classification

Prokaryote classification is the youngest, and more dynamic among the different classifications of living organisms. The reason is the small size of the prokaryotes, normally impossible to see with a naked eye, that hampered their discovery until a few centuries ago. Additionally, the development of a reliable prokaryote classification has been made difficult by the relative morphological simplicity of the prokaryotes. The lack of useful fossil register, together with the difficulties to retrieve helpful information from these small organisms have contributed to the instability of the prokaryote classification system. The development of new techniques directed to the understanding of those useful phenotypic and genotypic traits of microorganisms was the rate-determining step towards a reliable taxonomic schema for prokaryotes.

Most microscopists of the 17th and 18th centuries did not try to classify the infusion animalcules that they observed and often described so meticulously [48]. In these early years of bacteriology, supporters of the theory of spontaneous generation claimed that bacteria did not have ancestors and developed *de novo* from inanimate matter. Even after this theory was refuted, bacteria were treated as if there might be only a single species which could develop a great variety of shapes (pleomorphism). This period lasted essentially until the development of pure culture techniques [45]. Early attempts of microbial classifications were solely based on morphological observations. Otto Müller was the first to attempt a systematic arrangement of these microorganisms, however he did not make a clear distinction between what we now call protozoa and bacteria. In his last work published in 1786, he did create two form genera, *Monas* and *Vibrio*, which encompassed bacteria and differentiated the punctiform and the elongated types [48]. Several decades after, in 1838 Christian Ehrenberg extended Müller's nomenclature and added the helical bacteria. He described several genera and species that cannot now be recognized, but some of the species designations are still in use (e.g. *Spirochaeta plicatilis* and *Spirillum volutans*). Subsequent workers devised simpler classifications, but all these early bacterial systematists based their arrangements upon microscopic morphology, and assumed constancy of form at a time

when these theories of spontaneous generation and pleomorphism were still widely held and the germ theory of disease had yet to be proven [48]. In the 1870's Ferdinand Cohn still supported the idea that bacterial forms were constant irrespective of environmental conditions, but he already recognized the existence of a wide diversity of bacteria. He considered them to be unrelated to fungi, with close affinities with blue-green algae, a view that is now commonly accepted based on ultrastructural and molecular data. He arranged bacteria in six form genera, but appreciated that the physiologies, products and pathogenicities of similar-shaped organisms might differ. Robert Koch provided the germ theory of disease with his studies on *Bacillus anthracis*, and later concluded that the different forms of pathogenic bacteria must be regarded as different morphotypes of single distinct and constant species [48].

An important step for the development of microbiology was the accomplishment of the isolation of organisms in pure cultures. In 1872, Cohn's co-worker Joseph Schroeter had cultivated pure colonies of chromogenic bacteria, and in 1878 Joseph Lister obtained a pure culture of a milk-souring organism by dilution. Koch developed the technique of cultivation on solidified gelatin media, which was subsequently replaced by agar, and this is one of the still most common techniques for cultivation of microorganisms in pure cultures [48]. The achievement of the pure culture technique, marked the starting point of the 'golden age of microbiology'. With it one could retrieve direct information of the organisms avoiding information overlap due to the coexistence of several organisms with different physiological and/or genomic properties. Pure culture availability enforced the development of many tests for distinguishing bacteria which allowed the phenotypic description of the organisms, and formed the basis for their classification. During the first two decades of the present century, a large number of new species were classified by the many microbiological laboratories that were emerging. Physiological characters took a predominant role for bacterial classifications and replaced those morphological traits that so often led to confusion [48].

In 1923, Bergey's Manual of Determinative Bacteriology was published providing a modern identification key for bacteria. At the time

there was no official classification of bacteria [95], but this manual and the subsequent editions became the authoritative reference works on bacterial classification [48]. During the first half of the present century, several classification schemes appeared that were particularly significant [9], and led to the improvement of the bacterial nomenclature (i.e. the classification of Hauduroy et al. [32]; the sixth edition of the Bergey's manual of determinative bacteriology [6]; and the monograph by Zhdanov [114]). This publications provided a frame for unification of criteria among microbiologists and avoided nomenclatural problems that were relatively frequent in the early years of bacterial classification. At this point of time, the lack of scientific knowledge exchange among microbiologists, together with a strong tendency towards special purpose classification (artificial schemes in which one or a few properties of the organism are given undue prominence; [27]) were responsible for a confused nomenclatural standing of a quite large number of distinct bacteria. A good case example is *Pseudomonas stutzeri*, that during the first half of this century appeared in the literature at least under seven different names, often simultaneously [101].

Microbiologists have been searching for a classification system that is produced by an orderly arrangement of organisms on the basis of their relationships. A phylogenetic classification of prokaryotes has already been attempted by Orla-Jensen in 1909. He argued based on the assumption that the early earth was devoid of organic substrates that the first bacteria on earth must have been autotrophic, and postulated that the methane oxidizers were the most primitive of all. Later in 1936, Kluver and Van Niel proposed that cocci were the most primitive bacteria because they had the simplest morphological form; rods developed from elongated of fused cocci, vibrios evolved from rods, spirilla evolved from vibrios, etc. [45]. In the seventh edition of the Bergey's Manual in 1957, and in the absence of usable fossil record, a quasi-evolutionary approach to arrangement, as used for plants and animals, was adopted. The photoautotrophs were regarded as the most primitive forms and the reduced, intracellular pathogens of the genus *Rickettsia*, to which the viruses were tentatively attached, the most advanced [48]. These schemes were purely

speculative and today most microbiologists would agree that the basic premises are probably incorrect. However, these schemes did provide logical unifying frameworks in which bacteria could be placed. Bacterial taxonomists conscientiously avoided the term 'related' and used 'similar' instead [45]. In the eighth edition of the Bergey's Manual, which was published in 1974, it was acknowledged that such a pseudo-phylogenetic approach was no longer justifiable and consequently the bacterial genera, sometimes grouped in families where thought useful, were arranged in 19 independent groups. Each of these groups was based upon a few readily determined characters indicated in a vernacular name such as 'Gram-positive cocci'; hence, no evolutionary relationships were implied within and among the groups. Instead, a tentative subdivision of the Kingdom *Prokaryotae* was given [48].

As the variety of methods for characterizing bacteria increased, so bacterial taxonomists suffered more and more from the lack of quantitative approaches to classification. In the late 1950s, numerical taxonomy was developed in parallel to the onset of the computer age as a part of multivariate analyses. Its aim was to devise a consistent set of methods for classification of organisms. Much of the impetus for the development of numerical taxonomy in bacteriology came from the problem of handling the large tables of data on physiological, biochemical, and other properties of numerous strains. There was thus a need for an objective method of taxonomic analyses, whose first aim was to sort individual strains of bacteria into homogeneous groups, conventionally species, and which would also assist in arrangement of species into genera and higher groupings [87]. The period of numerical taxonomy coincided with the rise of chemotaxonomy - the application of modern biochemical analytical techniques, principally chromatographic and electrophoretic separation methods, to the study of distributions of specific chemical constituents such as amino acids, proteins, sugars and lipids in bacteria [48].

During the early 1960s a new kind of information became available that suggested that relatedness among bacteria might be objectively determinable after all. This was based on the realization that it is the information coded in the DNA of an organism that makes an organism

what it is. An increasing knowledge of the properties of DNA and the development of molecular biology techniques supported the idea that bacteria might best be classified by comparing their genomes [45]. Initially, overall base compositions of DNAs (mol% G+C values) were used to compare bacterial genomes. Bacteria whose mol% G+C values differed markedly from one to another were obviously not of the same species. However, the single values obtained by the analysis of DNA base compositions allowed only very superficial comparisons of genomes. A much more precise method of comparison was needed, based on comparison of DNA nucleotide sequences. Derived from the physico-chemical properties of the DNA, DNA-DNA hybridization techniques were developed. A great practical advantage of this methods was that the continua that often occurred between groups defined by phenotypic characteristics usually failed to occur in DNA hybridization experiments; organisms tended to be either closely related or not [45]. DNA-DNA hybridization consequently became the standard technique for the circumscription of bacterial species. However, as these experiments were used increasingly in bacterial classification, some microbiologists worried that the data might merely be sets of figures with little practical value. If these data were obtained as an end in themselves, this certainly would be true [45]. Their practicality depends on subsequently determining those phenotypic characters that can be used to describe a DNA similarity group, and also on determining which phenotypic characters are the ones that could be used to identify new isolates easily, rapidly, and reliably [38]. Indeed, the Committee on Reconciliation of Approaches to Bacterial Systematics [107] recommended a bacterial species classification not to be recognized unless an diagnostic phenotypic property is provided.

In the late 1970s a remarkable breakthrough in attempts to determine relationships between distantly related bacteria was achieved by cataloging and posterior sequence analysis of the primary structure of ribosomal ribonucleic acids (rRNA). The rRNA sequences were shown to be very useful molecular marker for phylogenetic analyses [49]. Initially, partial catalogues of oligonucleotide sequences were generated, but as the sequencing techniques developed, nearly complete sequences of the 5S,

16S and 23S rRNA have been used to infer phylogenies [49]. Among the three rRNA molecules, 16S rRNA has been most widely studied. Thus for it most information is available [51]. The era of the 16S rRNA sequencing brought spectacular new information such as the recognition of *Archaeobacteria* or *Archaea* as an independent cellular lineage [109], and important rearrangements of the prokaryotic classification schema [65]. Ribosomal RNA sequencing has become to be one of the routine analysis in most of the microbiology laboratories, and although not required by the polyphasic approach, this information is provided in the description of most of the newly classified bacterial species. It is also becoming increasingly popular to propose new bacterial species using data generated from 16S rRNA sequencing studies [92]. Unfortunately, the resolution power of the 16S rRNA is often insufficient to guarantee correct delineation of bacterial species [23, 53]. Furthermore, its validity as a marker for phylogenetic studies is being questioned [13, 28a, 30].

Nevertheless, at the present the vast majority of bacterial taxonomists is convinced that 16S rRNA sequence analysis provides a stable and quite satisfactory framework for prokaryotic classification. However, it is also widely accepted that an adequate classification of prokaryotes, specially of the lower taxonomic ranks such as species, will only result from a polyphasic approach combining as many different techniques as possible, including a fine-tuning of the circumscription that takes into account the differing properties of the different bacterial groups [26].

3- Methods and parameters used in prokaryotic classification

Nowadays prokaryote taxonomists agree that a reliable classification can only be achieved by the exploration of the internal diversity of taxa by a wide range of techniques. For that it is necessary to investigate, as extensively as possible, the two sources of information: genotype, phenotype. Genotypic information is derived from the nucleic acids (DNA and RNA) present in the cell, whereas phenotypic information is derived from proteins and their functions, different chemotaxonomic markers and a wide range of other expressed features. Both sources of information need to be investigated and the lack of any of them might result in the rejection of the proposed classification. Especially, the description of both phenotype and genotype is strictly necessary for the delineation of new species in prokaryotes [107], and just recently, because of the ease and low costs of nucleic acid sequencing, comparative sequence analysis became an accepted, and in the near future essential trait for the species classification. Today, the accepted species classification will be only achieved by the recognition of the genomic distances and borders among the closest classified taxons (DNA-DNA similarity), and of those phenotypic traits that are exclusive and serve as diagnostic of the taxon (phenotypic property).

Altogether, the practice of retrieving as much information as possible by the use of a wide range of different methods is what prokaryote taxonomists call a "Polyphasic Approach" [103]. In this regard, prokaryote systematics has undergone spectacular changes in recent years by taking full advantage of developments in chemistry, molecular biology and computer science to improve the understanding of the relationships between microorganisms and the underlying genetic mechanisms on which they are based [27]. A relatively large set of techniques are being used as a routine for prokaryote classification. However, it is of primary importance to understand at which level these methods carry information, to realize their technical complexity, and the amount of time and money required. The resolution power of each technique is directly related to the kind of information that it retrieves, and the correct use of this information is essential to guarantee the adequate classification of a taxon. The taxonomic information level of some of these techniques is illustrated in Fig. 1.

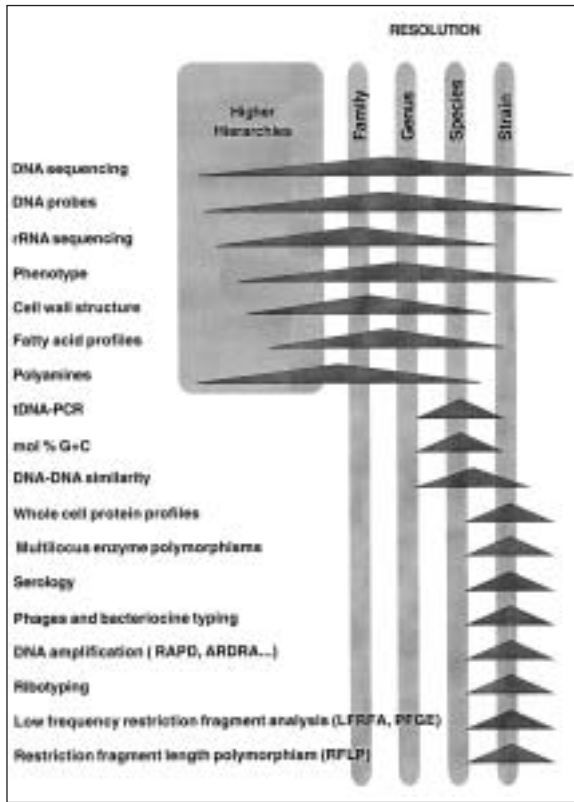


Figure 1:

Taxonomic resolution of some of the currently used techniques. Abbreviations are defined in the glossary.

The list of methods given comprises the major categories of taxonomic techniques required to study bacteria at different taxonomic levels, however, it is not meant to be complete.

3.1- Genotypic methods

Genotypic methods are those that are directed toward DNA or RNA molecules. Undoubtedly, these methods presently dominate modern taxonomic studies not only as consequence of technological progress, but primarily because of the present view that classification should reflect most closely the natural relationships as they are encoded in the DNA [103]. Unlike other cell constituents used for chemotaxonomy (see next section), only the amounts and not the composition of RNA and chromosomal

DNA are affected by growth conditions. Furthermore, the nucleic acids are universally distributed and they alone can be used as standards for wide-ranging comparisons [48].

The most complete sources of information are of course the complete bacterial genomes (Fig 1). As large-scale sequencing of complete genomes is not yet feasible at the present, several alternative approaches are taken. They include estimating the mean overall base composition of DNA, comparing genomic similarities by DNA-DNA pairing studies, generating unique sets of DNA fragments by digestion with restriction endonucleases (LFRFA, PFGE, RFLP...), sequence comparisons of selected genes, DNA-rRNA hybridization, and sequencing of rRNA [48].

3.1.1- DNA base ratio (mol % G+C; G+C content; G+C%)

The primary structure of DNA results from the linear succession of the four nucleotide bases adenine (A), thymine (T), guanine (G) and cytosine (C), and this succession determines the genetic information of an organism genome. Because of the double-stranded nature of the DNA, where both strands complement with base pairing G-C and A-T, the ratios G/C and A/T remain constant. However, the relative ratio $[G+C]/[A+T]$ varies from genome to genome. The base ratio of a DNA molecule is generally described as relative abundance of the pair G+C, thus it is also commonly called G+C content. Then, the DNA base ratio is calculated in percentage of G+C: $[G+C]/[A+T+C+G] \times 100$. This was the first nucleic acid technology applied to prokaryote systematics [47], and initially proved to be a useful and routine way of distinguishing between phenotypically similar and genotypically different strains [27]. It is usually one of the characteristics required for minimum descriptions of species and genera.

Among the prokaryotes G+C contents vary within the range of 20 to 80 mol% [96]. The greater the disparity in G+C mol% between two organisms the less closely related they are; theoretically, DNA molecules with differences of greater than 20 to 30 mol% can have virtually no

sequences in common [48]. Empirically it has been noted that organisms showing more than 10 mol% should not be assigned to the same genus and that 5 mol% is the maximum range permissible within a species. However, firm guidelines have yet to be set for the range of variation, but a range of 15 mol% can be taken as a strong indication for heterogeneity within a genus [26]. In this regard it should be noted that although differences in mol% are taxonomically useful for separating groups, similarities in base compositions do not necessarily indicate close relationships because the determinations do not take the linear sequences of bases in the DNA molecules into account; the criterion can only be used negatively.

3.1.2- DNA-DNA similarity (DNA-DNA pairing; DNA-DNA homology; DNA-DNA relatedness)

The measure of the whole genome DNA-DNA similarity is today still the standard technique for the species delineation. The rationale for using this parameter to set the borders of the species circumscription originates from the results of numerous studies, in which a high degree of correlation was found between DNA similarity and phenotypic similarity (i.e. chemotaxonomic, serological...[92]).

A unique property of DNA and RNA is the ability for reassociation or hybridization. The complementary strands of DNA, once denatured, can, under appropriate experimental conditions, reassociate to reform native duplex structures. The specific pairings are between the base pairs A-T and G-C, and the overall pairing of the nucleic acid fragments is dependent upon similar linear arrangements of these bases along the DNA. Under standardized conditions, DNAs from different organisms reassociate depending on the similarity of their nucleotide sequences, thereby allowing quantification of the degree of relatedness, usually expressed as per cent similarity or homology. It is important to note that the term homology has been replaced for the term similarity due to the inaccuracy of its use. There is no linear correlation between actual sequence identities and hybridization values; the latter gives only relative similarity values between genomes. There are several methodologies to measure DNA-DNA relatedness, but all of them rely on the same principle (Fig.

2). DNAs of two different organisms are mixed and denatured to give a solution of a mixture of single stranded DNA molecules (ssDNA). Under controlled experimental conditions DNA reassociation occurs and results in hybrid molecules: the higher the genetic similarity of the two organisms, the more nucleotide base sequences they have in common, and the more hybrid formation (hybridization) will occur. The comparison between the results obtained with the mixture of DNAs and pure reference DNA (homologous DNA) yields a degree of relatedness.

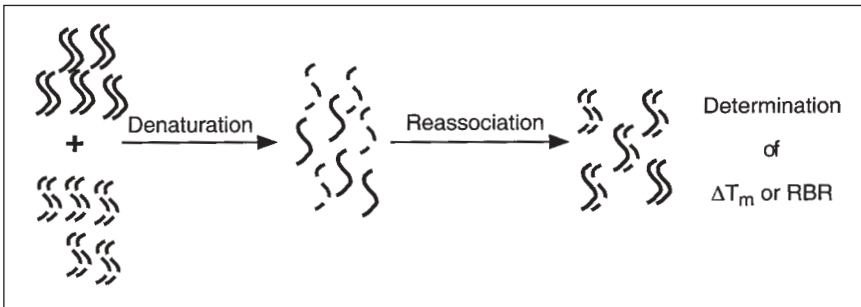


Figure 2: DNA-DNA reassociation assay

There are two main parameters that are used to measure the degree of relatedness: the relative binding ratio (RBR) and the difference in thermal denaturation midpoint (ΔT_m). Although both parameters result from the measure of different features, they are correlated and can be independently used for the species circumscription [39]. In a mixture of two different DNAs, after the hybridization procedure, the amount of double stranded units is dependent on the degree of identity between both DNAs. In this case double stranded DNA occurs in less extent as it would happen in the homologous mixture. RBR reflects the relative amount of hybrid formed in comparison to the homologous DNA, for which 100% reassociation is expected.

ΔT_m is a reflection of the thermal stability of the DNA duplexes. Double stranded DNA denaturation is mainly dependent on three factors: the G+C mol%, the ionic strength of the solution in which the DNA is dissolved, and the temperature. G+C mol% is a constant parameter character-

istic of each DNA, and the ionic strength of the hybridization solution is normally kept constant. Thus, the single variable parameter is the temperature. In a denaturation kinetic curve of a double stranded DNA, the temperature at which 50% of the DNA strands are already denatured is called melting temperature or thermal denaturation midpoint (T_m). The heteroduplexes of two non-homologous DNAs account for lower number of paired bases than those from the homologous, that means less amount of hydrogen bonds. Thus, the duplexes are less stable and in a denaturation kinetics the T_m is reached at lower temperatures (Fig. 3). In this case the parameter used for measuring the DNA-DNA relatedness, ΔT_m is the difference between the homologous DNA T_m and the heterologous DNA T_m .

An advantage of ΔT_m upon RBR, is that the former parameter is comparable with independence of the method used for hybridization and without transformation of the data. However, RBR results are subjected to differences that are related to the hybridization technique used, and this has to be taken into account when comparing results obtained with different hybridization protocols [29].

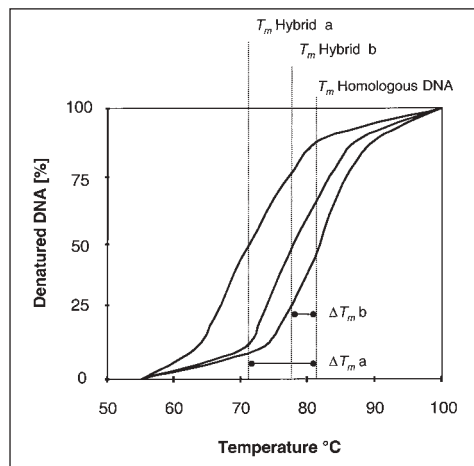


Figure 3: Thermal denaturation curves of a homologous DNA and two hybrid DNAs.

Derived from the observation of numerous studies with well defined prokaryotic species, it is currently accepted that values of 70% or higher RBR, or 5°C of ΔT_m or lower, are generally accepted borders for the species circumscription [107]. However, it is important to realize that

DNA relatedness values do not reflect the actual degree of sequence similarity at the level of the primary structure. Indeed, it has been estimated that prokaryotic DNA heteroduplexes will not be formed, even under non-stringent conditions, unless the DNA strands show at least 80% sequence complementarity. Thus, depending on the sequence similarity of the reassociating single strands, a difference of about 20% of sequence identity is spread between 0% and 100% DNA-DNA similarity [27]. Additionally, DNA-DNA similarity studies are time-consuming and they are hampered by the fact that they rely on pair-wise comparisons, which directly causes that experiments are normally undergone with a relative small set of organisms. Therefore, unless the reference strains chosen are representative of the constituent species incorrect conclusions could easily be drawn [31].

Despite of the problems, the advantages of DNA-DNA similarity analyses outweigh their limitations, and it is an attractive measure as it can be applied to all cultivable prokaryotes irrespective of their growth requirements. They provide a unified concept for the delineation of bacterial species, and among other properties, these studies can also be used to detect and identify unknown isolates [26].

3.1.3- Ribosomal RNA analysis (rRNA)

Techniques involving the analysis of ribosomal RNA (rRNA) or of the genes coding rRNA (rDNA) have on the last 25 years revolutionized the prokaryotic taxonomy. The conclusions drawn from these studies are based on the assumption that the rRNA genes are highly conserved because of the fundamental role of the ribosome in protein synthesis. Ribosomal RNAs are molecules with universal, constant and highly constrained functions that were established at early stages in evolution and that are not affected by changes in the organism's environment. Therefore, and because they are large molecules containing considerable genetic information, they have been chosen as the molecular chronometers at least in the prokaryotic world [113]. Two more assumptions are basic for the validity of this approach, namely that lateral gene transfer has not occurred between rRNA genes, and that the amount of evolution or dissimilarity between

rRNA sequences of a given pair of organisms is representative of the variation shown by the corresponding genomes [26]. If this holds true, the variations in the rRNA primary structures among the prokaryotes will reflect evolutionary distances among the organisms.

The three rRNAs are classified by their sedimentation rates as 23S, 16S and 5S, which have chain lengths of about 3300, 1650 and 120 nucleotides, respectively, and can be separated by gradient centrifugation or electrophoresis. Until recently direct and complete sequencing of the larger rRNA molecules was not feasible as a routine approach. Instead, sequence data were analyzed indirectly by DNA-rRNA hybridization or partial sequences were obtained semi-directly by oligonucleotide cataloguing [48]. Initially, for prokaryotes, complete sequencing of 5S rRNAs was used for phylogenetic inferences. However, this technique waned in favor of 16S rRNA sequencing for various reasons, including the greater information content of the larger molecule. Nowadays, almost complete sequencing of the latter is a routine. The 23S rRNA molecule is a larger information unit than the 16S, thus with higher resolution power for phylogenetic studies [50]. However, because still technical problems for sequencing due to its gene length, the sequencing has not been as popular as for the 16S and the amount of sequences in the databases is much smaller.

Meanwhile, the 16S rRNA approach is one of the most widely used standard techniques in microbiology. Consequently, a comprehensive sequence data set (18.000 entries in 1999) is available in generally accessible databases. The phylogenetic analysis of these data provides the basis for an ongoing evaluation and restructuring of the current bacterial systematics accompanied by emendation, reclassification and hence renaming of bacterial taxa. It is also widely accepted to apply the rRNA technology as an integrated part of a polyphasic approach for new descriptions of bacterial species or higher taxa [50]. The congruence of 16S rRNA-based reconstructions of phylogenetic trees with those based on alternative molecules, such as 23S rRNA, ATPase subunits, elongation factors and RNA polymerases has been tested and resulted on very similar tree topologies [50].

An important feature of the 16S rRNA molecule in its use as universal standard parameter for phylogenetic inferences is the relative ease of sequence alignment [20]. Alignment is the first critical step of sequence-based phylogenetic analyses. Given that positions with a common ancestry have to be compared for reliable phylogenetic conclusions, homologous positions have to be arranged in common columns in correct alignment [50]. All rRNA molecules share a common secondary and higher-order structure. Many of these structure elements are identical or similar with respect to their position within the molecule as well as number and position of paired bases, or internal and terminal loops, while the primary structures differ [50]. The conservation of helical elements is maintained independently of primary structure conservation by compensating base changes at positions involved in base pairing. In this regard, checking a primary structure for its potential of higher-structure formation usually helps to improve the primary structure alignment. However, there remain variable regions which still cannot be unambiguously aligned and it is the subjective decision of the researcher how to arrange the data [50]. In the case of the rRNAs an additional fact facilitates the arrangement of data. There are comprehensive databases of aligned sequences accessible to the public. These alignments have been established and are maintained by specialists. The databases contain secondary structure information, and can be used as a guide to inserting new sequence data [50].

What is the best method for inferring phylogenetic relationships from sequence data? Computer simulations and experiments have revealed that all methods fail when the assumptions upon which they are based are badly violated [20]. There are three major approaches for tree reconstruction: distance matrix, maximum parsimony, and maximum likelihood methods. The two last approaches are based on models of evolution and, in general, operate by selecting trees which maximize the congruency of tree topology and the measured data under the criteria of the model [50]. Given that the treeing methods are based on different models and treat the data in different ways, a perfect match of the tree topologies cannot be expected. It has to be taken into consideration that the models only partially reflect reality. Thus for a final reconstruction of a tree, the

use of the different methods together with the calculations based on several data subsets is strongly recommended. In many cases the separation of clusters or subtrees is stable whereas their relative branching orders differs with the applicable alternative treeing methods. In such cases a fairly acceptable compromise is to use a consensus tree which shows detailed branching patterns, where stable topologies emerged, and multifurcations where inconsistencies or uncertainties could not be resolved [50].

16S rRNA sequencing and comparison analyses have demonstrated high resolution power for measuring the degree of relatedness between organisms above the species level. However, as more sequence information becomes available, it is evident that the resolution power of 16S rRNA sequences is limited when closely related organisms are being inspected [26, 92]. Thus, as it is discussed in a later chapter, there can be no bacterial species definition based solely on sequence similarity of rRNAs or their genes. Absolute values for delineating species cannot be set based on the low resolution power at this level [26]. However, the rRNA sequencing approach has additional advantages e.g. for the exploration of the uncultured prokaryote species diversity [20], and design of phylogenetic probes for a cultivation-independent monitoring of microbial communities [3].

3.1.4- DNA-based typing methods (DNA fingerprinting)

DNA-based typing methods generally allow to detect intraspecific diversity, e.g. subdivisions within species into a number of distinct types. They are an alternative to the classical phenotypic analyses [103]. One can differentiate between two basic techniques: (i) the first-generation typing methods, based in whole genome restriction fragment analysis, (ii) and the polymerase chain reaction (PCR) methods, based on the amplification of genome fragments. All these techniques rely on the separation and visualization of DNA fragments by gel electrophoresis.

The first-generation DNA-based methods consist on the generation of fragments of the whole genome by using restriction enzymes. Common

restriction enzymes recognize specific combinations of four to six bases. Due to the size of the bacterial genome (between 0.6 to 9.5 megabases, Mb; [22]) the digestion with common enzymes results in a complex mixture of DNA fragments of different sizes that, in most of the cases, is difficult to analyze. However, the number of DNA fragments can be reduced by selecting restriction enzymes which only rarely cut DNA, recognizing a specific combination of six to eight bases. The technique is referred to as low-frequency restriction fragment analysis (LFRFA). The fragments, however, are too large to be separated by conventional agarose gel electrophoresis, and therefore pulse-field-gel-electrophoresis (PFGE) has to be used [103]. The result is a more or less complex electrophoretic pattern for each strain. The comparison of patterns by numerical analysis leads to the establishment of similarity groups within the species. This technique, combined with Southern blot hybridization, yields data on the genome size and organization, that in the future could be of relevance for a comprehensive description of any organism [69]. The number and distribution of ribosomal RNA operons through the genome might be a simple, but perhaps taxonomically useful feature of the genome that can be revealed by this technique [25].

Alternatively, the complex DNA patterns generated by regular restriction digestion with common enzymes can be transferred to a membrane and then hybridized with a labeled probe, which allows to reveal the hybridized fragments. A typical example of one of these developments is the ribotyping method, which uses rRNA as a probe [103].

The introduction of the polymerase chain reaction (PCR) methodology into the microbiology laboratory has opened a vast array of applications. Among others, a battery of different typing methods was developed. Different methods in which short arbitrary sequences were used as primers in the PCR assay were described: oligonucleotides of about 20 bases are used in arbitrary primed PCR (AP-PCR); oligonucleotides of about 10 bases are used in randomly amplified polymorphic DNA analysis (RAPD) [103]. Those are only few of the many examples of the application of PCR to the typing of bacteria. However, all of them are only applicable to the understanding of the intraspecific diversity. These techniques are not sui-

table for the circumscription of prokaryotic species as well as for higher taxonomic units, and for this reason their use in prokaryotic taxonomy is rather limited.

3.2- Phenotypic methods

The phenotype is the observable expression of the genotype. Until the new molecular methodologies were available for the prokaryote taxonomists, the classification was exclusively based on the observation of the morphology, physiology and growth conditions of the organisms. These investigations were directly linked to the use of pure cultures, and the laboratory capabilities to cultivate the organisms and analyze their properties. Thus, the classification schemes were biased towards aerobic heterotrophic microorganisms for which an extensive retrieval of information was easy. One of the disadvantages of analyzing the phenotype is that the whole information potential of a prokaryotic genome is never expressed. The gene expression is directly related to the environmental conditions (e.g. growth conditions in the laboratory). Prokaryotic phenotype can not be based on the simple observation of the organism. The prokaryotes lack complex morphological features. Most do not show life cycles with different morphological stages, and lack an ontogenetic development. Thus, the analysis of the prokaryotic phenotype mostly relies on the development of experimental techniques that evidences direct or indirectly different phenotypic properties. In this respect, most of the phenotypic analyses are biased towards the observation of enzyme activities, substrate utilisation profiles, and growth conditions. One of the main basic problems is that unless we know the genes that are responsible for the phenotypic traits that we observe, we can not distinguish between autapomorphic and synapomorphic characters. Prokaryote taxonomy has traditionally overseen this problem. However, due to the development of the molecular techniques, and the establishment of a reliable phylogenetic schema, taxonomists start to regard the autapomorphic or synapomorphic nature of the phenotypic traits. Soon, complete genome sequencing will shed light on the homology of phenotypic traits.

Four main problems must be considered when planning a phenotypic study of bacteria. The first two are that analysis based upon different character sets might show poor agreement (congruence), and that, as mentioned above, the phenotype represents a very small part of each organism's genome [48]. These problems are combated basing a general purpose classification on a large number of characters from a wide phenotypic range. The other two problems are concerned with analysis. Many methods exist for the calculation of similarities and for the arrangement of strains according to these similarities, and the application of different combinations of methods to a set of data can lead to a wide variety of interpretations. Fortunately, bacteriologists tend to restrict themselves to relatively few methods (those available in the computer programs) and discrepancies between the results of different approaches may give useful insights into stabilities and other aspects of the relationships implied [48].

Phenotypic analysis is the most tedious task in the classification of microorganisms. It requires much time and skill, and the techniques should be standardized to avoid subjective observations. An important aspect when analyzing the phenotype of a prokaryotic species is that the strains should be chosen to represent the known diversity and environmental niches of the group being studied [48]. It is most important to include recent isolates because classification based entirely upon 'museum pieces', which have been maintained in major culture collections and other laboratories for many years and have become adapted to laboratory conditions, lay poor foundations for diagnostic schemes intended to identify isolates from the real world [48]. In classifying prokaryotes it is desirable to use an orderly approach based on common sense, and on the use of the tests that are pertinent [84].

3.2.1- Classical phenotypic analyses

The classical or traditional phenotypic tests are used in identification schemes in the majority of microbiology laboratories. They constitute the basis for the formal description of taxa, from species and subspecies up to genus and family. While genotypic data alone are sufficient to allo-

cate taxa in a phylogenetic tree and to draw the major borderlines in classification systems, the consistency of phenotypic and genotypic characters is required to generate useful classification systems [103]. The classical phenotypic characteristics of bacteria comprise morphological, physiological, and biochemical features. Individually, many of these characteristics have been shown to be irrelevant as parameters for genetic relatedness, yet as a whole, they provide descriptive information enabling us to recognize taxa. The morphology of a bacterium includes both cellular (shape, endospore, flagella, inclusion bodies, Gram staining...) and colonial characters (color, dimensions, form...). The physiological and biochemical features include data on growth at different temperatures, pH values, salt concentrations, or atmospheric conditions, growth in the presence of various substances such as antimicrobial agents, and data on the presence or activity of various enzymes, metabolization of compounds, etc. Reproducibility of results within and between laboratories is a major problem which can be addressed with highly standardized procedures [103].

3.2.2- Numerical taxonomy applied to phenotypic analyses

Numerical taxonomy is also known as the computer-assisted classification [27]. It appeared when the variety of methods for characterizing bacteria increased and bacterial systematists suffered more and more from the lack of quantitative approaches to classification [48]. Sneath and Sokal [89] define numerical taxonomy as *'the grouping by numerical methods of taxonomic units into taxa on the basis of their character states'*. The methods require the conversion of information about taxonomic entities into numerical quantities. Numerical taxonomy principles, based on the concept of Adansonian taxonomy, state that maximum information content should be achieved: i.e., all possible characters should be studied for the strains, they should be weighted equally, and taxa should be defined on the basis of overall similarity according to the results of the analyses. Numerical taxonomy is often incorrectly used as a synonym for phenetic analyses of phenotypic data. However, this also applies to those

studies of genotypic data, for which cladistic analyses of gene sequences are the most known.

A numerical taxonomy study of phenotypic data involves five essential steps [15]:

a- *Selection of strains.* There are no rules controlling the range of diverseness among the strains to be examined. It is recommended that the set of strains should be as large as possible, which should include cultures of historical, pathological, or environmental importance. It is important to include reference strains (type strains) whose identity has been established for comparative purposes. As mentioned above, it is recommendable that the strains represent fresh isolates, so that little modifications have been occurred due to laboratory adaptations.

b- *Test selection.* Routine tests should represent a broad spectrum of the biological activities of the organism and include morphological, colonial, biochemical, nutritional, and physiological characters. Tests that are not highly reproducible, this includes some of those routinely employed in conventional bacterial taxonomy, should be avoided. An optimum number of tests for numerical taxonomy is considered to be in the range of 100 to 200. Standardization of treatment, inoculation, and incubation of the strains is also required.

c- *Data coding.* Generally data are given in a binary numerical format. Positive responses (plus) are coded as 1 and negative responses (minus) are coded as 0. Weighting of characters is usually avoided.

d- *Computer analysis.* Coded data is computerized by one of the several available programs. Among the different similarity coefficients available, bacteriologists generally employ either the simple matching coefficient (S_{SM}), or the Jaccard coefficient (S_J). The data analysis generate similarity matrices containing information about relationships among the strains. Subsequently, cluster analysis are performed to finally generate the dendrograms.

e- *Presentation and interpretation of results.* This last step is mainly dependent on the data set under study. Results can be presented in sor-

ted similarity matrices, as well as a dendrograms. These branched diagrams are partially informative because they are generated with the highest similarity values linking a pair of organisms, but gives an easy visual overview. One problem of dendrograms is that inexperienced researchers are not aware that they do not necessarily indicate phylogenetic relationships between strains (i.e., relationships based on ancestry of the organisms), and sometimes can led to misinterpretations.

Numerical taxonomy has supported the development of stable prokaryotic classifications, especially the determination of homogeneous groups that can be equated with taxospecies. Furthermore, the databases generated are essentially information storage and identification systems.

3.2.3- Chemotaxonomy

Phenotypic methods comprise all those that are not directed toward DNA or RNA, therefore, they also include the chemotaxonomic techniques. The term ‘chemotaxonomy’ refers to the application of analytical methods to collect information on various chemical constituents of the cell to classify bacteria. As the introduction of chemotaxonomy is generally considered one of the essential milestones in the development of modern bacterial classification, it is often treated as a separate unit in taxonomic reviews [103]. However, as the parameters measured are a direct reflection of the expression of the genetic information of an organism, they should be regarded as phenotype. As for the other phenotypic and the genotypic techniques, some of the chemotaxonomic methods have been widely applied on vast numbers of bacteria whereas others were so specific that their application was restricted to particular taxa [103].

Chemotaxonomy is concerned with the discontinuous distribution of specific chemicals, notably amino acids, lipids, proteins and sugars, and in this sense can be considered to provide good characters for classification and identification [27]. It is, however, important that the observed variation in chemical composition is the result of genetic differences and not due to variation in cultivation conditions. Therefore, it is usually

necessary to grow cultures under carefully standardized growth regimes before comparative chemotaxonomic work can be undertaken. Rigorously standardized cultivation conditions are particularly important in studies involving quantitative analyses of chemical data [27]. Several techniques are increasingly being used as routine in prokaryotic taxonomy, e.g.:

Cell wall composition. Generally used for the classification of Gram positive organisms. The peptidoglycan type and teichoic acids are analyzed.

Lipids. The composition and relative ratio of fatty acids (hydroxylated, non-hydroxylated, branched...), polar lipids (specially phospholipids), lipopolysaccharides, isoprenoid quinones (ubiquinones, menaquinones...), generally analyzed by chromatography are used successfully for discriminating among taxa of various ranks.

Polyamines. Polycationic compounds with important but unclear role in the prokaryotic cell. Its composition and relative ratio can be discriminative for taxa above genus.

3.2.4- Phenotype typing methods

Like in the section 1.4 (DNA-typing methods), these are techniques that are useful for establishing relationships within a prokaryote species, but generally lack resolution above this taxon [80, 82, 103]. There are several methods that have been successfully used for discriminating strains as well as for the understanding of the intraspecific variability: (i) serotyping, based on the presence of variability in the antigenic constituents of the cells (capsules, cell envelopes, flagella, fimbria...), (ii) electrophoretic protein profiles, based on the extraction of the proteins and separation on, normally, polyacrylamide gels (whole cell protein profiles, Gram negative outer membrane protein profiles, and multilocus enzyme electrophoresis...), (iii) Lipopolysaccharide electrophoretic profiles, where variations on the O-side chains are reflected in different ladder-like electrophoretic patterns, (iv) Pyrolysis mass spectrometry, Fourier trans-

formation infrared spectroscopy, and UV resonance Raman spectroscopy. These are sophisticated analytical techniques which examine the total chemical composition of bacterial cells. Because of the complexity of the analytical apparatus these techniques have so far only been used for particular groups of bacteria [27].

3.2.5- Identification keys and diagnostic tables

One of the goals of the phenotypic characterization is the construction of a framework for an accurate identification of organisms. This framework can consist in dichotomous identification keys where the identity of an isolate is being tested with an orderly, steplike series of questions. However, diagnostic tables are more common in microbiology. Diagnostic tables contain more information than the dichotomous keys and are much more successful as a determinative aid [99]. These tables are based upon the sharing of several (usually unweighted) characters, which are characteristic and identifies the taxon (the species phenotypic property; [107]). In diagnostic tables variable characters within the studied taxon are as well recorded, and this is a good indication of the intraspecific diversity. The success in the identification of new isolate to already established species is dependent on how accurate the description of the species was, and the accuracy is dependent on the size of the dataset analyzed. It is postulated that for a rather accurate description of a species, a minimum of 10 but better 25 strains should be studied [85]. However, in the majority of cases, new species and genera are described on the basis of only a few strains or even only one. Poor descriptions based on a small set of strains can lead to improper phenotypic circumscription of taxa, thus hindering the identification of new isolates.

3.2.6- Microbial identification systems

A significant contribution of industry specially to the clinical microbiology was the development of miniaturized identification systems

based on classical methods. Several systems are commercially available (e.g. API, Analytab Products, Plainview, N.Y.; Biolog, Biolog Inc., Hayward, Calif.; Vitek, Vitek Systems, Inc. Hazelwood, Mo.), mostly based on modifications of classical methods [18]. First-generation systems were addressed to the identification of members of the family *Enterobacteriaceae*, and consisted on miniaturized tubes containing individual substrates, multicompartiment tubes or plates with multiple substrates, and paper strips or disks impregnated with dehydrated substrates. This methods developed by the incorporation of highly sophisticated, computer-generated identification data bases tailored for each system [18]. However, these systems addressed mainly the identification of organisms with high medical importance. All of them consist of a relative short number of key tests useful to identify a particular group of microorganisms. Because of this special purpose development, miniaturized identification tests have shown to be successful for the identification of clinical isolates. However, this systems have been less successful for the identification of environmental isolates due to a lack of knowledge of the phenotypic diversity of the microorganisms in natural environments. Thus, this systems should be viewed cautiously when applied to samples that are not from clinical origin [43]. Miniaturized systems are commonly used for phenotype exploration in classification purposes. However, some taxonomists regret their use in classification because the reduced set of tests [70]. Although not commercial, miniaturized systems have been developed for classification purposes consisting of large numbers of physiological tests [41]. These systems have been used successfully for the examination of the physiological diversity of environmental isolates [41, 82].

4- Prokaryotic species concept

As detailed in the introduction, prokaryote taxonomy has mostly developed during the present century. The classification system as well as the Linnean nomenclature were adopted as an analogy to those established systems for eukaryotes, specially the Botanical Code [88]. The adopted system has been satisfactory for all levels of bacterial classification but except the species. Supraspecific classes, being regarded as abstract entities [102], can be compared to those classification systems established for eukaryotes. However, the concept of prokaryote species is different. This is because no universal concept exists [36, 55], and species are (regarded as individuals) practical entities [102] for which the requirements for their circumscription vary depending on the species concept adopted (biological, phenetic, evolutionary...; [55]). Unfortunately, through the history of prokaryote taxonomy, there has been paid much more attention to the taxa nomenclature [e.g. 9, 10, 17, 33, 86, 88, 98] than to the practical circumscription of the species concept applied to prokaryotes.

Today's prokaryotic species concept results from empirical improvements of what has been thought to be a unit. The circumscription of the species has been optimized through the development of microbiological methods that reveal both genotypic and phenotypic properties of prokaryotes, which can not be retrieved through simple observation. As it is mentioned in the introduction there is currently a phase of strong criticism to the current definition of the prokaryotic species among some non-taxonomists [19, 94, 105, 109]. The current concept is criticized as too conservative and ill-defined. On the other hand, many other microbiologists find the present concept acceptable [93]. We will argue in the following that analyzing the current state of the available techniques and knowledge about microorganisms, the current concept is the most practicable we can have at the moment. It also fulfills several important requirements for being a concept (e.g. the resulting classification schema is stable), and should be regarded as a not ideal but practical concept.

4.1- The concept

Early definitions of bacterial species were often based on monothetic groups (groups based on a unique set of features considered to be both sufficient and necessary for the groups so defined) described by subjectively selected sets of phenotypic properties [26]. This concept had severe limitations as strains which varied in key characters could not be accommodated in existing taxonomies. Moreover, such classifications often lacked uniformity as different criteria were frequently used by different investigators for the same group of organisms. Another consequence of this approach was that the rank of species became very unevenly defined over the whole range of bacterial variation. The number of species in a genus was influenced by the aims of the taxonomist, the extent to which the taxon had been studied, the criteria adopted to define the species and the ease by which the strains could be brought into pure culture. Members of environmentally and medically important genera have been underclassified and those in industrially significant taxa overclassified [26].

Until the discovery of the existence of the DNA as information-containing molecule, prokaryote classification was based solely on phenotypic characteristics. The development of the numerical taxonomy [89], where the individuals are treated as Operational Taxonomic Units being polythetic (they can be defined only in terms of statistically covarying characteristics), achieved a more objective circumscription of prokaryotic units. The discovery of the genetic information gave a new dimension to the species concept for microorganisms. Parameters like G+C content and overall DNA-DNA similarity have additionally been used for the circumscription of this basic unit in a more natural way, and as it is discussed below, these parameters enable at least a first rough insight into phylogenetic relationships. Thus, the species concept for prokaryotes evolved to be mostly phenetic or polythetic. This means that species are defined by a combination of characters, each of which may occur also outside the given class and may be absent in any member of the class [102].

There is no official definition of a species in microbiology. However, from a microbiologist point of view '*a microbial species is a concept*

represented by a group of strains, that contains freshly isolated strains, stock strains maintained in vitro for varying periods of time, and their variants (strains not identical with their parents in all characteristics), which have in common a set or pattern, of correlating stable properties that separates the group from other groups of strains' [28]. This definition does only apply to those prokaryotes which have been isolated in pure culture (essential for the classification of new prokaryotic species), and excludes those uncultured organisms which constitute the largest proportion of living prokaryotes. However, a prokaryote species is generally considered to be *'a group of strains that show high degree of overall similarity and differ considerably from related strain groups with respect to many independent characteristics'*, or *'a collection of strains showing a high degree of overall similarity, compared to other, related groups of strains'* [16].

There are, in the literature, at least three different species definitions that, to the date, tend to disappear due to the unification of criteria: (i) Taxospecies, defined as a group of organisms (strains, isolates) with mutually high phenotypic similarity that form an independent phenotypic cluster, (ii) genomic species as a group showing high DNA-DNA hybridization values, and (iii) nomenspecies as a group that bears binomial name [16]. The simultaneous occurrence of these three conceptual units have been definitively avoided by the agreement that a species classification can only be achieved by the integration of both phenotypic and genotypic parameters. Indeed, the committee on reconciliation of approaches to bacterial systematics [107] recommended *'that a distinct genospecies that cannot be differentiated from another genospecies on the basis of any known phenotypic property not be named until they can be differentiated by some phenotypic property'*.

It is nowadays accepted, among microbial taxonomists that a prokaryotic species should be classified after the analysis and comparison of as many parameters as possible, combining phenotypic and genotypic markers in what is known to be the Polyphasic Taxonomy [103]. The evaluation of several distinct independent phenotypic and genotypic characters have promoted a more unified approach to the objective delineation

of a bacterial species concept. The 'polyphasic species' has distinct advantages over traditional more descriptive species concepts, especially since it can be expected to yield well described species, a stable nomenclature and better identification systems. The polyphasic species concept is universally applicable. However, the details of the approach need to be tailored taking into account the differing behavioral properties of the members of taxonomically diverse genera [26]. Polyphasic species should be regarded as synonym of polythetic species or phenetic species. It has been noted that this concept of species is extremely useful for dealing with biological entities endowed with intrinsic variability, since it can accommodate individual members that lack one or other character considered typical of the class [102].

4.2- The borders

Unfortunately, there are no absolute borders for the circumscription of prokaryotic species, and this results to be a problem for those non-taxonomists who are searching for the identification of their new isolates belonging to hitherto unclassified species. As it is said above, a species classification can only be achieved through the analysis of several independent phenotypic and genotypic characters in a polyphasic approach. It is necessary to show that the group of strains under study form an independent and diagnosable unit within the established classification scheme. The species circumscription approach is a tedious task often underestimated in its importance.

The most accepted parameter to give a numerical and quasi absolute boundary for the species circumscription is the value of the overall DNA similarity. Values expressed as percentage of similarity or ΔT_m (see methods) are considered to be, in some extent, crude measures of genomic distances among microorganisms. Just because the results obtained are an indirect reflection of the genomic sequence similarity at the level of the primary structure [27], DNA reassociation approaches represent the best applicable procedure for the inference of phylogenetic relationships among closely related prokaryotes [107]. In this way, and based on nume-

rous studies in which a high degree of correlation was found between DNA similarity, and chemotaxonomic, genomic, serological and numerical similarity, the DNA reassociation has been used as the standard for species delineation [92]. Empirically it has been observed that most of the well defined prokaryotic species harbor strains which genomic similarities are above 70%. This fact made the committee on reconciliation of approaches to bacterial systematics to recommend to set the borders of the species circumscription in terms of DNA similarity. It is stated that '*a phylogenetic definition of species generally would include strains with approximately 70% or greater DNA-DNA relatedness and with 5°C or less ΔT_m* ' [107]. It is important to note that the values recommended by the Committee are not absolute numbers in defining the genomic boundary of a single prokaryotic species. In some cases, like *Pseudomonas stutzeri* [79], the recommended values might be too narrow to harbor all the strains of a single species [100, 103]. Additionally, the second most used genomic parameter in prokaryotic taxonomy, the G+C content gives also some numerical boundaries for the unit. Empirically it is observed that a single species does not usually contain strains whose G+C mol% exceed a range of 5%. Both parameters are to the date necessary for an adequate species classification.

4.3- The 16S rRNA sequence data

The development of the molecular techniques applied to bacterial systematics introduced a new parameter that is having an enormous influence into the prokaryotic classification, the 16S rRNA sequence. This technique enabled remarkable breakthroughs in attempts to determine relationships between distantly related bacteria [27]. Molecular sequencing is dominated by the possibility of drawing genealogical trees that represent lines of descent [112, 113]. The reconstruction of phylogenetic trees based on 16S rRNA has been tested by the comparison with other genes with similar information content (23S rRNA, elongation factor Tu, and ATPases; [50]) which, in some extent, resulted in similar phylogenies. Markers with different functional pressures report on different periods of evolutionary time. A fast evolving gene can only report on recent deve-

lopments whereas a conserved molecule fails in this respect but does well on ancient events [50]. In these trees, organisms represent terminal points in a genealogical tree, and it is important to note that due to the absence of an useful fossil register no time-scaled patterns of phylogenetic ancestry can be drawn.

16S rRNA analyses have opened the door for a more objective classification system among prokaryotes, especially for taxa above the species, for example a practicable border zone for a genus definition would be 95% of sequence similarity [50]. It has been observed that species having 70% or greater DNA similarity usually have more than 97% 16S rRNA sequence similarity [92]. This value could have been taken as an absolute boundary for the species circumscription. However, disagreements to these observations are starting to appear regularly in the literature indicating the lack of resolution power of this molecule at the species level. In this regard, it is possible to find different species with identical [76] or nearly identical 16S rRNA sequences [23, 53], a micro heterogeneity of the 16S rRNA genes within a single species [5, 37] or, in exceptional cases, single organisms with two or more 16S rRNA genes with relatively high sequence divergence [61, 64]. Thus, due to the highly conserved nature of the 16S rRNA, there is no linear correlation between DNA-DNA similarity percent and 16S rRNA similarity for closely related organisms [26, 92] and the bacterial species definition can not be solely based on sequence similarity of rRNAs. However, comparative analysis of 16S rRNA is a very good method for a first phylogenetic affiliation of both potentially novel and poorly classified organisms [26]. Due to the practical advantage of the 16S rRNA approach also for identification purposes, it is recommended to include the ribosomal sequence in new descriptions of prokaryotic species. Probably in the near future, it will be a necessary parameter, together with DNA-DNA similarity and G+C content, for any classification.

4.4- Intraspecific subdivisions

Subspecies is the lowest taxonomic rank that has official standing in nomenclature [88]. A species may be divided into two or more subspe-

cies based on minor but consistent phenotypic variations within the species or on genetically determined clusters of strains within the species [95]. However, many bacterial species are endowed with a relative internal heterogeneity that does not show consistency for a subspecific subdivision. Then, strains of a single species can sometimes be grouped in terms of some independent special characteristics. These groups or infra-subspecific subdivisions are not arranged in any order of rank, and may overlap one another [88]. For example, members of any species can be grouped in terms of biochemical or physiological properties (biovar or biotype); of pathogenic reactions (pathovar or pathotype); of reactions to bacteriophages (phagovar or phagotype); antigenic characteristics (serovar or serotype); and so on...

DNA reassociation experiments, however, produce intraspecific subdivisions that are often seen as potential species, but their independent classification is hampered by the lack of an diagnostic phenotypic property [107]. There is, in the literature, a nomenclatural confusion about the phenotypically similar but genotypically distinct groups, which have been referred to as genomic species, genospecies [107], DNA groups [37], genomospecies [7, 8], genomic groups [115], or genomovars [79, 100]. There is an open discussion about the adequacy of these terms used to designate these subdivisions [103]. Actually, each of the different intraspecific units defined with DNA similarity values can not be considered a single species *per se*. The polythetic nature of the currently accepted species concept does not allow the recognition of a species based on a single characteristic. This means that although the DNA reassociation values give a numerical boundary for the species circumscription, this unit can not be recognized unless there is an overall agreement of the distinct characters analyzed in a polyphasic study. Thus, the suffix '-species' is inadequate for this term as far as DNA similarity groups *per se* lack a specific standing. Similarly, the term 'group' is informal and has no nomenclatural standing, and among the suffixes, '-var' is recommended [88]. The term genomovar has been suggested to accommodate different DNA similarity groups within a nomenclature [79, 100]. This term has been positively accepted by some taxonomists [26, 104] because it indicates that a

genomic species are an integral part of a nomenclature and hence should not be overlooked in subsequent taxonomic work. It is suggested that genomovars encompassed in nomenclature should be numbered, and not named [79]. Ultimately, following the identification of further phenetic identifications the genomovars can be given a formal name.

It is empirically observed that the circumscription of the prokaryotic species should be more relaxed in absolute values of genomic similarity [100, 103]. The level of 70% binding and $5^{\circ}\text{C } \Delta T_m$ is very strict, and often phenotypic consistency would not be achieved if these recommendations were strictly applied. More relaxed boundaries for the species delineation would be a group of strains sharing 50 to 70% DNA reassociation and 5 to 7°C difference in thermal stability between the homologous and heterologous duplexes [100], which is a more realistic standard [103]. Allowing internal genomic heterogeneity (represented by genomovars) and more relaxed genomic boundaries, the circumscription of the prokaryotic species will have a even more conservative nature, but it may be a more pragmatic definition as far as it facilitates diagnosis.

4.5- Prokaryotic sex

Attempts to circumscribe the prokaryotic species in terms of genetic exchange, in what could be analogous to the 'biological species concept' (BSC, [56, 57]) have been made [77, 78]. However, this prokaryotic species concept has no direct theoretical analogy to the BSC developed for eukaryotes. It has been adapted to the particular way in which prokaryotes share gene pools. Prokaryotes do not have the same reproductive mechanisms as eukaryotes, i.e. meiosis, fertilization. They are haploid and reproduce asexually by binary fission in which the genetic material is passed vertically from mother to daughter cell. Thus the evolutive pattern would be only linked to the genome organization rearrangements and mutation rate. The latter one is considered to be one of the major sources of genetic diversity [110]. However, prokaryotes have several common systems for genetic exchange called horizontal gene transfer systems. The three best known genetic transfer systems are: (i) conjugation: transfer of

DNA from a prokaryotic donor to a recipient cell through cell-to-cell contact which is mediated by plasmids, (ii) transformation: uptake of free DNA from the environment through active or passive cell transport systems, and (iii) transduction: spread of prokaryotic genes by temperate bacteriophages [46]. Originally, it was assumed that these horizontal gene transfer systems were restricted to closely related organisms, and that the homologous recombination would follow the same patterns observed for eukaryotes [78]. However, genetic exchange in prokaryotes is less frequent but more promiscuous than that in eukaryotes [13]. In the sexual eukaryotes, populations that are separated by only 2% sequence divergence are frequently unable to exchange genes. Prokaryotes, in contrast, may undergo homologous recombination with related species that are up to 25% (and possibly more) divergent in the sequences of homologous genes; they can also accept and express new genes on plasmids from extremely divergent sources [14]. Prokaryote taxonomy has consequently, to date, not benefited greatly from studies involving genetic exchange of chromosomal material, and the BSC is far from being realized [40]. Indeed, there is no sense in applying the BSC for prokaryotes because it fails in one of the basic statements, the interbreeding discrimination [57].

Independent of the best species concept for microbiology, taxonomists are aware that the horizontal gene transfer may have important consequences for bacterial systematics [40]. It is not yet known what the real significance of the gene exchange among prokaryotes is, although it is seen to be a much rarer eventuality than it is in eukaryotes [13]. Actually, the most questioned aspect within prokaryotic systematics is the current phylogenetic reconstructions. Genealogical relationships are basically inferred after 16S rRNA sequence analysis, a single gene. And although the tree is consistent with other trees generated for different molecules [50], there is the question of whether microbiologists are dealing with a true organismal tree or just a gene tree [111]. The skepticism on the value of the 16S rRNA as a biological chronometer originates from the release of the first whole prokaryotic genome sequences. It seems that horizontal gene transfer might have played a more relevant role in evolution than it was supposed, and some incongruities are found with the current phylo-

genetic scheme [30, 72]. The answer on the extent of gene exchange among prokaryotes will have to await further whole genome sequences.

Fortunately, the prokaryotic species concept, as it is currently conceived, is remote from the problematic of gene exchange. It is true that the acquisition of genomic material through horizontal gene transfer, and/or the presence of extrachromosomal elements can lead to misclassifications or misidentifications of prokaryotes because of the influence on their phenotype [40]. However, such characters, specially the extrachromosomal elements, should be excluded from taxonomic studies once they are known to happen [40]. In spite of these problems, which may be minimized by computer-assisted extensive phenotypic analyses [40], the species concept for prokaryotes is appropriate because it is based on whole genome similarities [92]. As it is explained in the methods section, DNA reassociation values are an indirect expression of the real genome sequence identity. One can estimate that for example a 96% of sequence identity between two genomes would lead to 70% DNA similarity [92], or 4°C of ΔT_m [108]. Genome sequence variation within the range of values recommended for a single species can account for significant differences in the phenotype. However, despite genomic rearrangement caused by horizontal transfer and the presence of extrachromosomal elements, the primary structure of the majority of the genes are most likely not involved [92]. Changing the physical map will not markedly influence the extent to which DNA hybridizes, and even if the genetic changes affect one of the characters used in the phenotypic characterization of the species, the DNA similarity values will most likely not change to a measurable extent. The current microbiological species definition consequently is relatively insensitive to genetic rearrangement, gene amplification, mutation, and exchange of genetic material over a nonpredictable range of taxa [92].

4.6- Pure cultures and the culture collections

The prerequisite of the current prokaryotic species concept is the isolation of the microorganisms in pure cultures. Biochemical tests, genome analyses or chemical component analysis can not be performed unless

the organism under study is analyzed separately [70]. The isolation and maintenance of organisms in pure cultures require time and skill and remains one of the prime challenges for microbiologists. However, the classification of a prokaryotic species can not rely on a single strain or isolate, but an effort should be made to collect as many strains as possible to evidence the real intraspecific diversity. It is calculated that a number of 25 strains are necessary for an accurate description of a species, and that the lower tolerable limit would seem to be 10 strains [85]. This number is, unfortunately, seldom achieved in taxonomic publications. Looking through the two most renowned journals for prokaryotic taxonomy, i.e. the 'International Journal of Systematic Bacteriology', and 'Systematic and Applied Microbiology', most of the newly classified species are based on a single isolate, thus providing a relative poor phenetic description that can lead to misclassifications, or inability to assign new isolates to this species.

After the classification of a new species, it is important to assign one of the isolates to be the reference strain which other scientists may use for comparisons. This is named '*type strain and consist of living cultures of an organism which are descended from a strain designated as the nomenclatural type. The strain should have been maintained in pure culture and should agree closely in its characters with those in the original description*' [88]. Actually, it is also recommended to select reference strains for each infraspecific subdivision and preserve them, especially those belonging to different genomovars which might be classified as different species later on.

Another important aspect of the pure culture technique is that isolated strains need to be preserved and made available to the scientific community for information and comparison. Most of the microbiologists tend to build their own culture collections. However, reference collections of bacteria (culture collections) have been set up for the maintenance of large numbers of strains of microorganisms, as well as to have available those reference strains, specially type strains, that are necessary for any comparative work [24, 85]. In this regard, it is nowadays necessary to deposit the type strain of a species (or the reference strains of a geno-

var; [79]) in one of the publicly accessible reference culture collections [24].

It is currently impossible to retrieve sufficient information for taxonomic studies from uncultured organisms identified in the environmental samples. However, the development of the rRNA approach to explore uncultured prokaryotes in natural samples has given valuable insights into prokaryotic diversity [68, 20]. Unfortunately, as stated above, the rRNA approach lacks resolution at the species level. Also it is very difficult to infer physiological characteristics of an organism based only on its rRNA sequence without having it in pure culture. Phylogenetically related prokaryotes can have diverse physiologies whereas physiologically similar organisms can occur in different phylogenetic lineages [59]. A novel rRNA sequence isolated from nature therefore merely indicates that there is a currently unknown microorganism in the environment. The knowledge of the phylogenetic affiliation of the new microorganisms might, however, support attempts of cultivation since those can be based on growth conditions formulated for the closest cultivable relative [42]. Additionally, the rRNA approach allows the design of rRNA-targeted probes for unique sequence motifs of the unknown microorganism. These probes can subsequently be used for the in situ identification of the organism from which the sequence was retrieved [2, 3]. This approach permits, among many other applications, the monitoring of the isolation of new organisms [35], as well as the determination of the morphology, abundance, distribution in certain habitats and even indications on growth rates and physiological activities of uncultured organisms [75, 81]. These new methods allow for retrieval of some information on the uncultured organisms and for the recognition of their uniqueness within the hitherto established classification scheme. In this regard, the International Committee on Systematic Bacteriology implemented the category of *Candidatus* to record the properties of putative taxa of prokaryotes [60]. This category is used for ‘*describing prokaryotic entities for which more than a mere sequence is available but for which characteristics required for description according to the International Code of Nomenclature of Bacteria are lacking*’ [60]. Such descriptions should include not only the phylogenetic

information, but also information on morphological and ecophysiological features as far as they can be retrieved in situ, together with the natural environment in which the organism can be identified. It is important to note that *Candidatus* is not a rank but a provisional status, and efforts to isolate and characterize the members of the putative taxa should be made to enable their definite classification.

4.7- The species concept to be adopted for prokaryotes

Through the manuscript it is stated that the polythetic species concept for prokaryotes, based on many independent phenotypic and genotypic traits, is the best concept we can, to the date, conceive to harbor prokaryotic units. Lately strong criticisms to the polythetic concept were articulated [19, 94, 105, 110]. These scientists argue that the current concept is too conservative, and that there is a lack of congruency with the concept delineated for higher organisms. Indeed, it is very difficult to make comparisons of species concepts for both kinds of organisms, and as May stated, the basic notions about what constitutes a species would be necessarily different for vertebrates than for bacteria [54]. Comparisons among the different species units of the different living organisms can be done only after a universal species concept that applies to all organisms has been devised. This is, however, not easy. Eukaryote taxonomists are currently disagreeing about which is the adequate concept to be used among the > 20 concepts currently discussed [12, 21, 36, 52, 55, 63, 90]. There is, however, a general tendency that Mayr's BSC should be abandoned because of its lack of practicability. Among the species concepts [36, 55], there seem to be two candidates which are universally applicable and thus could serve for classification of all living organisms: the phenetic (polythetic) species concept and the evolutionary species concept.

(i) The phenetic species concept. The phenetic or polythetic [102] species concept (PSC) is a similarity concept based on statistically covarying characteristics which are not necessarily universal among the members of the taxa [36]. This is the concept that has empirically been adopted to circumscribe the prokaryotic species, which to the date has

demonstrated to be rather stable as well as operational. This concept has no theoretical commitment as it is considered theory-neutral or theory-free. A theoretical foundation has traditionally been seen as a valuable characteristic of a species concept [36, 55]. However, the significance of theory in a species concept is a controversial issue between philosophers and scientists. The more theoretically significant a concept is, the more difficult it is to apply [36]. Scientists, in general, see no real need to include their criteria of application in their theory-based concepts, as well as pragmatism in their concepts is seen as a valuable virtue. As it is analyzed by Hull [36], the PSC covers most of the primary requirements for being a concept, i.e. universality, monism and applicability, which are considered to be the most valuable characteristics for scientists. It is, to the date, the applied concept for the prokaryotic species, but also it has been recommended for the higher organisms [90], and although it is considered to be theory-free this is one of the most persuasive concepts to the date conceived [36].

(ii) The evolutionary species concept. The evolutionary species concept (ESC) has been considered the most theoretically committed of the species concepts [36, 55]. Since it is regarded to be the only one that can serve as a primary concept because it can accommodate all types of species known to the date [55]. The evolutionary species is '*an entity composed of organisms which maintains its identity from other such entities through time and over space, and which has its own independent evolutionary fate and historical tendencies*' [55]. Evolutionary species is a lineage concept which is explicitly temporal, treating these units as lineages extended in time (space-time worms; [36]). This concept, however, has no pragmatic significance for the prokaryotes when we analyze the current state of knowledge about this group of organisms. Among the prokaryotes we can not recognize an evolutionary fate nor historical tendencies just because of the lack of a useful fossil register. Morphological features, in general, have little information content in prokaryotes, and therefore, from the spare prokaryotic fossils we can not retrieve sufficient information on genotypic and phenotypic characteristics of the ancestors of the present prokaryotic species. The same is true for a prospective. Currently, the knowledge of evolutionary tempo and mode of evolutionary

changes of prokaryotes is rather incomplete. Different groups have been demonstrated to not evolve isochronically [92]. The predictions become even more difficult if we take into account the possibilities of horizontal gene transfers between distant groups. Thus to the date, the adoption of an ESC for prokaryotes is not yet possible.

(iii) Other concepts. Linked to the ESC there are two additional concepts which could easily serve as primary concepts for prokaryotes: the monophyletic (or autapomorphic) species concept and the diagnostic species concept [36]. Both of them are defined as phylogenetic (or genealogical) concepts with a minimal time-dimension. In spite of the lack of useful fossil register for prokaryotes, the modern molecular techniques permitted the establishment of genealogical trees among the prokaryotes through the 16S rRNA gene sequence analysis [50]. As it is discussed above, this phylogeny reconstruction is based on a single gene analysis, and its validity to represent the organismal phylogeny is today questioned. However, derived from the congruency of the reconstructed trees with similar slowly-evolving molecules [50], together with the coherency of the phenetically designed taxa with the phylogenetic tree [91], it is most likely that at least the local branches represent stable genealogical relationships.

The monophyletic species concept considers that '*a species is the least inclusive monophyletic group definable by at least one autapomorphy*' [36]. Disregarding the low resolution of the 16S rRNA analyses, we can recognize each of the species being a monophyletic group. The problem is, however, the recognition of which characters are really autapomorphies. The only possibility to recognize an autapomorphy among the members of a taxon is the availability of the gene sequence codifying for the character for all the members of the species. The sequences should show to be homologous, unique for the taxon, and excluded from horizontal gene transfer. This is, to the date a nearly impossible task and thus this concept is not operational for the prokaryotes.

The diagnostic species concept considers that '*a species is the smallest diagnosable cluster of individual organisms within which there*

is a parental pattern of ancestry and descent' [36]. We can recognize a pattern of ancestry among the prokaryote species in our phylogenetic tree, as well as the diagnosable units are circumscribed after the polyphasic approach. This concept would indeed serve as a primary concept for prokaryotes. There are, however, some problems derived from the phylogenetic analyses that are currently undergone. First of all the most common marker used, the 16S rRNA, lacks resolution at the species level and divergent polythetic units can be indistinguishable when using this approach. The second problem, less worrying, is whether we are really reconstructing an organismal phylogeny. The use of a more resolute molecule (perhaps the 23S rRNA) which is recognized to reconstruct organismal phylogeny, would lead the diagnostic species concept to be the most adequate among all.

To the date, taking into account the present knowledge of the prokaryotic properties and the available techniques for prokaryotic taxonomy, the polythetic (phenetic) species concept is the most recommended to be used. It harbors the most valuable properties of a species concept, universality, monism, and applicability. And its application have shown to give a rather stable, objective and predictable classification system.

4.8- Towards an improvement of the species circumscription

'The adequacy of characterization of a bacterium is a reflexion of time; it should be as full as modern techniques make possible. Unfortunately, one now regarded as adequate is likely, in ten years time, to be hopelessly inadequate!' Cowan, 1965 [17]. This statement has been the evolutionary fate of the taxonomy of the prokaryotes, and we even can predict that it will still hold true in the near future. And this is mainly because the development and improvement of the taxonomic classification for prokaryotes is linked to the development of modern molecular techniques. During the last years the number of new isolates as well as the amount of information useful for systematics have increased significantly. Also, the availability of information has been improved by the generation of diverse databases (16S rRNA, fatty acids, metabolic markers...) as

well as software packages to handle them [73, 103]. Among the different sources of information that are currently emerging, there is one for which it is difficult to predict how significant it will be for prokaryotic taxonomy: the complete genome sequence. Like the rRNA-based reconstructions of the prokaryotic phylogeny have changed dramatically the classification system, we can suppose that genomics will give us answers about naturalness or artificialness of the current classification system for prokaryotes.

So far little attention has been paid to the genome organization as an additional parameter for the species circumscription. A simplified, though less expensive and time-consuming, source of information on genome organization as compared to whole genome sequencing is the physical mapping of prokaryotic genomes [22]. From experiments based mainly on pulsed field gel electrophoresis (PFGE) of large genome fragments and Southern hybridization, one can retrieve information on genome size, plasmids, number of chromosomes and their topology, number and distribution of house-keeping genes like *rnm* operons, genome rearrangements... Actually, in a nearly forgotten publication, Krawiec [44] proposed the use of the sequence of chromosomal loci as an identifying characteristic of bacterial species. He argued that the chromosomal organization like the presence/absence of genes, are important selective feature and fundamental characters of an evolving population; that the structural organization of the genome demarcates functional units of the genome; and that the organization of a genome is directly linked to the niche that this organism has been adapted. Actually, he proposed that the sequence of loci establishes the identity of a species as well as preserves the identity by creating a barrier to the exchange of genes. There are currently too few prokaryotes whose genome has been analyzed to evaluate those assumptions. However, from these analyses one can retrieve valuable information for a polyphasic approach like genome size, and number and distribution of *rnm* operons. Indeed, both parameters have shown to be fairly conservative within a single species as it was observed for *Pseudomonas stutzeri* [25], and for *Helicobacter pylori* [1]. On the other hand, some incongruences have been observed for members of the species

Vibrio cholerae [62]. Further intraspecific studies are needed to evaluate the potential of this technique.

It would be desirable to recognize each of the prokaryotic species as a single monophyletic lineage within the phylogenetic tree. Due to the lack of resolution of the 16S rRNA sequence data for the identification of single bacterial species, there is a need for a more informative molecule with higher resolution power. The 23S rRNA would be a possible candidate, but for it still not enough data are available [50]. It has been suggested that a more resolute phylogenetic relationships among closely related species, or among members of the same species, can be drawn after the sequence analysis of the intergenic spacer regions (ISR) of the rRNA operons [4, 74], or sequence clustering based on protein-coding genes [71]. These techniques could be useful in the absence of a sufficiently resolute gene which is suitable for the universal organismal phylogeny.

Furthermore, phenotypic analyses should be more exhaustively performed in order to recognize the internal variability of a single species, as well as the phenotypic property which is exclusive of the species. The development of commercial kits has facilitated the task of identification of organisms. However, some researchers regret their use because of the restricted number of tests [70], which are commonly addressed to the identification of clinical isolates. It is actually true that these kits often fail in the identification of new environmental isolates, and that the number of tests performed are not high enough for any significant numerical taxonomic study. It would be desirable, for a serious taxonomic work, to use a large set of biochemical and physiological tests which have been shown to be useful for environmental isolates. The commercial availability of miniaturized physiological tests like the one developed by Kämpfer et al. [41], with more than 300 independent tests, would provide solid databases for the recognition of new prokaryotic species.

We are entering, however, the era of the genomics. Dozens of prokaryote genomes are currently sequenced or waiting to be sequenced (www.tigr.org). Microbiology will be able to access soon an enormous flood of sequences. Preliminary results show a conserved genome orga-

nization and sequence identity within a single species [1], and those strain-specific genes, which have probably been acquired by lateral gene transfer are mostly clustered in single hypervariable regions in what can be considered hot spots of recombination [34]. It is, indeed, too early to evaluate the input of genomics in prokaryotic taxonomy and specifically into the species concept. In ten years time we shall look back to the bacterial classification and the polythetic species concept of the late 90's considering Cowan's prophecy.

5- Conclusions

The current species concept for prokaryotes has been developed empirically in the 20th century in parallel to the improvement of microbiological techniques. Due to the inherent properties of the prokaryotes, natural relationships could only be drawn after the analysis of large number of independent covarying characters in what has been called the Polyphasic Approach. This approach produces a classification system made upon species circumscribed in terms of overall similarity, being treated as polythetic units. Thus, a prokaryotic species is mainly a phenetic-polythetic concept, that although it is considered theory-free, harbors the most valuable virtues of being a concept, i.e. universality, monism and applicability.

The practical circumscription of the prokaryotic species is made in terms of genomic similarity (DNA similarity), i.e. the extent to which two genomes can form hybrids. The absolute borders of each species should, however, be set taking into account as much additional phenotypic characters as possible that show overall agreement. Thus the classification of a prokaryotic species requires the recognition of a diagnostic phenotypic properties which guarantees the identification of any new isolate. There are no strict lines for the circumscription of a prokaryotic species, though the details of the approach need to be tailored taking into account the differing behavioral properties of the members of each diverse taxa.

Molecular techniques, specially the 16S rRNA sequencing, have given insights into the genealogical relationships among prokaryotes. Unfortunately, these approaches lack resolution at the level of prokaryotic species. It is possible that future developments of molecular techniques applied to prokaryotes may improve the concept of species. Finer genealogical relationships will be drawn, and the concept of prokaryotic species may develop towards a phylogenetic species concept (i.e. diagnostic species concept), which is more theoretically committed than the PSC. In contrast, due to the impossibility to infer evolutionary relationships among prokaryotes, it seems nearly impossible to formulate an evolutionary species concept for the members of this domain.

6- Glossary

Archaea: or Archaeabacteria, are a heterogeneous group of prokaryotes which differ markedly from other prokaryotes (Bacteria) in their 16S ribosomal RNA sequences and in other important characteristics of cellular composition.

ARDRA: Amplified rDNA Restriction Analysis.

Autapomorphy: in phylogenetics an autapomorphy denotes a homologous character common to all members of a single taxa and thought to be exclusive of the group.

Bacteria: a diverse group of prokaryotes which differ from the Archaea in their 16S ribosomal RNA sequences and in other important characteristics of cellular composition.

Homology: having a common ancestor.

Genotype: genetic constitution of an organism, which acting together with environmental factors determines phenotype.

LFRFA: Low Frequency Restriction Fragment Analysis.

Monism: philosophical theory that expresses that the plurality of the world can be explained in a single principle. In taxonomy, a monistic view is to think that a single level of organization exists across all organisms that deserves to be recognized as the species level.

Monothetic: a classification that determines group membership according to the states of just one or few characters, but it may use different characters at different stages of the process.

PCR: Polymerase Chain Reaction, a molecular biological technique that allows the enzymatic amplification of a defined region of DNA.

PFGE: Pulsed Field Gel Electrophoresis.

Phenetic: Adj. Applied to a classification based on overall similarity, as determined by equal weighting of all known characters. Since all observable characters are used, a phenetic classification will make use of molecular genetic data if these are available. In contrast to phyletic and phylogenetic, the term phenetic does not have any evolutionary implications, other than in the sense of showing the end product of evolution. Phenetic classifications may include phenotypic and genotypic characters.

Phenotype: the visible or otherwise measurable physical and biochemical characteristics of an organism, a result of the interaction of genotype and environment.

Polyphasic taxonomy: Term used to signify successive or simultaneous taxonomic studies of a group of organisms using an array of techniques designed to yield both molecular and phenotypic data.

Polythetic: a classification based on many characters, not all of which are necessarily shown by every member of the group.

Prokaryote: the etymology of the word indicates the absence of a true nucleus, separated from the cytoplasm by a nuclear membrane. Prokaryotes lack additional membrane-containing structures characteristic of eukaryotes (e.g. chloroplasts and mitochondria) and unlike the latter have ribosomes with a sedimentation constant of 70S. Introns are rare in prokaryotes and DNA is usually present in a single molecule. The prokaryotes comprise two domains, *Bacteria* and *Archaea*.

RAPD: Randomly Amplified Polymorphic DNA.

RFLP: Restriction Fragment Length Polymorphism.

Synapomorphy: in cladistic phylogenetics denotes a homologous character common to two or more taxa and thought to have originated in their most recent common ancestor.

Bibliography

1. **Alm, R.A., L.-S.L. Ling, D.T. Moir, B.L. King, E.D. Brown, P.C. Doig, D.R. Smith, B. Noonan, B.C. Guild, B.L. DeJonge, G. Carmel, P.J. Tummino, A. Caruso, M. Uria-Nickelsen, D.M. Mills, C. Ives, R. Gibson, D. Merberg, S.D. Mills, Q. Jiang, D.E. Taylor, G.F. Vovis, and T.J. Trust.** 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature*. **397**:176-180.
2. **Amann, R., W. Ludwig, and K.-H. Schleifer.** 1994. Identification of uncultured bacteria: a challenging task for molecular taxonomists. *ASM News*. **60**:360-365.
3. **Amann, R.I., W. Ludwig, and K.-H. Schleifer.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**:143-169.
4. **Anton, A.I., A.J. Martinez-Murcia, and F. Rodriguez-Valera.** 1998. Sequence diversity in the 16S-23S intergenic spacer region (ISR) of the rRNA operons in representatives of the *Escherichia coli* ECOR collection. *J. Mol. Evol.* **47**:62-72.
5. **Bennasar, A., R. Rossello-Mora, J. Lalucat, and E.R.B. Moore.** 1996. 16S rRNA gene sequence analysis relative to genomovars of *Pseudomonas stutzeri* and proposal of *Pseudomonas balearica* sp. nov. *Int. J. Syst. Bacteriol.* **46**:200-205.
6. **Breed, R.S., E.G.D. Murray, and A. Hitchens.** 1948. *Bergey's manual of determinative bacteriology* 6th Ed. William & Wilkins Co., Baltimore, Maryland.
7. **Brenner, D.J., P.A.D. Grimont, A.G. Steigerwalt, G.R. Fanning, E. Ageron, and C.F. Riddle.** 1993. Classification of citrobacteria by DNA hybridization: designation of *Citrobacter farmeri* sp. nov., *Citrobacter youngae* sp. nov., *Citrobacter braakii* sp. nov., *Citrobacter werkmanii* sp. nov., *Citrobacter sedlakii* sp. nov., and three unnamed *Citrobacter* genomospecies. *Int. J. Syst. Bacteriol.* **43**:645-658.

8. **Brenner, D.J., H.E. Müller, A.G. Steigerwalt, A.M. Whitney, C.M. O'Hara, and P. Kämpfer.** 1998. Two new *Rahnella* genospecies that cannot be phenotypically differentiated from *Rahnella aquatilis*. *Int. J. Syst. Bacteriol.* **48**:141-149.
9. **Buchanan, R.E.** 1955. Taxonomy. *Ann. Rev. Microbiol.* **9**:1-20.
10. **Buchanan, R.E., R.S. John-Brooks, and R.S. Breed.** 1948. International bacteriological code of nomenclature. *J. Bacteriol.* **55**:287-306.
11. **Bull, A.T., M. Goodfellow, and J.H. Slater.** 1992. Biodiversity as a source of innovation in biotechnology. *Annu. Rev. Microbiol.* **46**:219-252.
12. **Claridge, M.F., H.A. Dawah, and M.R. Wilson.** 1997. Practical approaches to species concepts for living organisms, p. 1-15. *In* M.F. Claridge, H.A. Dawah, and M.R. Wilson (ed.), *Species: the units of biodiversity*. Chapman & Hall, London.
13. **Cohan, F.M.** 1994. Genetic exchange and evolutionary divergence in prokaryotes. *Trends Ecol. Evol.* **9**:175-180.
14. **Cohan, F.M.** 1996. The role of genetic exchange in bacterial evolution. *ASM News* **62**:631-636.
15. **Colwell, R.R. and B. Austin.** 1981. Numerical taxonomy, p. 444-449. *In* T.P. Gerhardt, R.G.E. Murray, R.N. Costilow, E.W. Nester, W.A. Wood, N.R. Krieg, and G.B. Phillips (ed.), *Manual of methods for general bacteriology*. American Society for Microbiology, Washington D. C.
16. **Colwell, R.R., R.A. Clayton, B.A. Ortiz-Conde, D. Jacobs, and E. Russek-Cohen.** 1995. The microbial species concept and biodiversity, p. *In* D. Allsopp, R.R. Colwell, and D.L. Hawksworth (ed.), *Microbial diversity and ecosystem function*. CAB International, Oxon, UK.

17. **Cowan, S.T.** 1965. Principles and practice of bacterial taxonomy - a forward look. *J. Gen. Microbiol.* **39**:148-159.
18. **D'Amato, E.E., R.H. Taylor, J.C. Blannon, and D.J. Reasoner.** 1991. Substrate profile systems for the identification of bacteria and yeasts by rapid and automated approaches, p. 128-136. *In* A. Balows, W.J.J. Hausler, K.L. Herrmann, H.D. Isenberg, and H.J. Shadomy (ed.), *Manual of clinical microbiology*. American Society for Microbiology, Washington D.C.
19. **Dykhuisen, D.E.** 1998. Santa Rosalia revisited: why are there so many species of bacteria? *Ant. van Leeuwen.* **73**:25-33.
20. **Embley, T.M. and E. Stackebrandt.** 1997. Species in practice: exploring uncultured prokaryote diversity in natural samples, p. 61-81. *In* M.F. Claridge, H.A. Dawah, and M.R. Wilson (ed.), *Species: The units of biodiversity*. Chapman & Hall, London.
21. **Ereshefsky, M.** 1992. *The units of evolution. Essays on the nature of species*. The MIT Press. Cambridge, Massachusetts.
22. **Fonstein, M. and M. Haselkorn.** 1995. Physical mapping of bacterial genomes. *J. Bacteriol.* **177**:3361-3369.
23. **Fox, G.E., J.D. Wisotzkey, and P. Jurtschuk 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.
24. **Gibbons, E., P.H.A. Sneath, and S.P. Lapage.** 1989. Reference collections of bacteria - the need and requirements for type strains, p. 2325-2327. *In* S.T. Williams, M.E. Sharpe, and J.G. Holt (ed.), *Bergey's manual of systematic bacteriology*. Williams & Wilkins, Baltimore.
25. **Ginard, M., J. Lalucat, B. Tümmeler, and U. Römling.** 1997. Genome organization of *Pseudomonas stutzeri* and resulting taxonomic and evolutionary considerations. *Int. J. Syst. Bacteriol.* **47**:132-143.

26. **Goodfellow, M., G.P. Manfio, and J. Chun.** 1997. Towards a practical species concept for cultivable bacteria, p. 25-59. *In* M.F. Claridge and H.A. Dawah (ed.), *Species: the units of biodiversity*. Chapman & Hall, London.
27. **Goodfellow, M. and A.G. O'Donnell.** 1993. Roots of bacterial systematics, p. 3-54. *In* M. Goodfellow and A.G. O'Donnell (ed.), *Handbook of new bacterial systematics*. Academic Press Ltd., London.
28. **Gordon, R.E.** 1978. A species definition. *Int. J. Syst. Bacteriol.* **28**:605-607.
- 28a. **Gribaldo, S., V. Lumia, R. Creti, E. Conway de Macario, A. Sanangelantoni and P. Cammarano.** 1999. Discontinuous occurrence of the *hsp70* (*dnaK*) gene among *Archaea* and sequence features of HSP70 suggest a novel outlook on phylogenies inferred from this protein. *J. Bacteriol.* **181**:434-443.
29. **Grimont, P.A.D., M.Y. Popoff, F. Grimont, C. Coynault, and M. Lemelin.** 1980. Reproducibility and correlation study of three deoxyribonucleic acid hybridization procedures. *Curr. Microbiol.* **4**:325-330.
30. **Gupta, R.S.** 1998. What are archaebacteria: life's third domain or monoderm prokaryotes related to Gram-positive bacteria? A new proposal for the classification of prokaryotic organisms. *Mol. Microbiol.* **29**:695-707.
31. **Hartford, T. and P.H.A. Sneath.** 1990. Experimental error in DNA-DNA pairing: a survey of the literature. *J. Appl. Bacteriol.* **68**:527-542.
32. **Hauduroy, P., G. Ehringer, G., J. Magrou, A.R. Prevot, Rosset, and A. Urbain.** 1953. *Dictionnaire des bacteries pathogenes* 2nd Ed. Masson et Cie, Paris, France.

33. **Heise, H. and M.P. Starr.** 1946. Nomenifers: are they christened or classified? *System. Zool.* **17**:458-467.
34. **Holloway, B.W. and A.F. Morgan.** 1986. Genome organization in *Pseudomonas*. *Ann. Rev. Microbiol.* **40**:79-105.
35. **Huber, R., S. Burgraff, T. Mayer, S.M. Barns, P. Rossnagel, and K.O. Stetter.** 1995. Isolation of a hyperthermophilic archaeum predicted by in situ RNA analysis. *Nature.* **376**:57-58.
36. **Hull, D.L.** 1997. The ideal species concept-and why we can't get it, p. 357-380. *In* M.F. Claridge, H.A. Dawah, and M.R. Wilson (ed.), *Species: the units of biodiversity*. Chapman & Hall, London.
37. **Ibrahim, A., P. Gerner-Smidt, and W. Liesack.** 1997. Phylogenetic relationship of the twenty-one DNA groups of the genus *Acinetobacter* as revealed by 16S ribosomal DNA sequence analysis. *Int. J. Syst. Bacteriol.* **47**:437-841.
38. **Johnson, J.L.** 1973. Use of nucleic acid homologies in the taxonomy of anaerobic bacteria. *Int. J. Syst. Bacteriol.* **23**:308-315.
39. **Johnson, J.L.** 1989. Nucleic acids in bacterial classification, p. 2306-2309. *In* S.T. Williams, M.E. Sharpe, and J.G. Holt (ed.), *Bergey's manual of systematic bacteriology*. Williams & Wilkins, Baltimore.
40. **Jones, D.** 1989. Genetic methods, p. 2310-2312. *In* S.T. Williams, M.E. Sharpe, and J.G. Holt (ed.), *Bergey's manual of systematic bacteriology*. Williams & Wilkins, Baltimore.
41. **Kämpfer, P., R.M. Kroppenstedt, and W. Dott.** 1991. A numerical classification of the genera *Streptomyces* and *Streptoverticillium* using miniaturized physiological tests. *J. Gen. Microbiol.* **137**:1831-1891.

42. **Kane, M.D., L.K. Poulsen, and D.A. Stahl.** 1993. Monitoring the enrichment and isolation of sulfate-reducing bacteria by using oligonucleotide hybridization probes designed from environmentally derived 16S rRNA sequences. *Appl. Environ. Microbiol.* **59**:682-686.
43. **Klinger, J.M., R.P. Stowe, D.C. Obenhuber, T.O. Groves, S.K. Mishra, and D.L. Pierson.** 1992. Evaluation of the Biolog automated microbial identification system. *Appl. Environ. Microbiol.* **58**:2089-2092.
44. **Krawiec, S.** 1985. Concept of bacterial species. *Int. J. Syst. Bacteriol.* **35**:217-220.
45. **Krieg, N.R.** 1988. Bacterial classification: an overview. *Can. J. Microbiol.* **34**:536-540.
46. **Lanka, E. and W. Pansegrau.** 1999. Genetic exchange between microorganisms, p. 386-415. *In* J.W. Lengeler, G. Drews, and H.G. Schlegel (ed.), *Biology of the Prokaryotes*. Blackwell Science, Stuttgart.
47. **Lee, K.Y., R. Wahl, and E. Barbu.** 1956. Contenu en bases purique et pyrimidiques des acides deoxyribonucleiques des bacteries. *Ann. Institut Pasteur* **91**:212-224.
48. **Logan, N.A.** 1994. *Bacterial Systematics*. 1. Blackwell Scientific Publications, London.
49. **Ludwig, W. and K.-H. Schleifer.** 1994. Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiol. Rev.* **15**:155-173.
50. **Ludwig, W., O. Strunk, S. Klugbauer, N. Klugbauer, M. Weizenegger, J. Neumaier, M. Bachleitner, and K.-H. Schleifer.** 1998. Bacterial phylogeny based on comparative sequence analysis. *Electrophoresis* **19**:554-568.

51. **Maidak, B.L., G.J. Olsen, N. Larsen, R. Overbeek, M.J. McCaughey, and C.R. Woese.** 1987. The RDP (Ribosomal Database Project). *Nucleic Acids Res.* **25**:109-110.
52. **Mallet, M.** 1995. A species definition for the modern synthesis. *Trends Ecol. Evol.* **10**:294-299.
53. **Martinez-Murcia, A.J., S. Benlloch, and M.D. Collins.** 1992. Phylogenetic interrelationships of members of the genera *Aeromonas* and *Plesiomonas* as determined by 16S ribosomal DNA sequencing: lack of congruence with results of DNA-DNA hybridizations. *Int. J. Syst. Bacteriol.* **42**:412-421.
54. **May, R.M.** 1988. How many species are there on Earth? *Science.* **241**:1441-1449.
55. **Mayden, R.L.** 1997. A hierarchy of species concepts: the denouement in the saga of the species problem., p. 381-324. *In* M.F. Claridge, H.A. Dawah, and M.R. Wilson (ed.), *Species: the units of biodiversity.* Chapman & Hall, London.
56. **Mayr, E.** 1970. *Populations, species, and evolution.* The Belknap Press of Harvard University Press, Cambridge, Massachusetts.
57. **Mayr, E.** 1987. The ontological status of species: scientific progress and philosophical terminology. *Biol. Phil.* **2**:145-166.
58. **Mayr, E.** 1998. Two empires or three? *Proc. Natl. Acad. Sci. USA* **95**:9720-9723.
59. **Murray, R.G.E. and K.-H. Schleifer.** 1994. Taxonomic note: a proposal for recording the properties of putative taxa of prokaryotes. *Int. J. Syst. Bacteriol.* **44**:174-176.
60. **Murray, R.G.E. and E. Stackebrandt.** 1995. Taxonomic note: implementation of the provisional status *Candidatus* for incompletely described prokaryotes. *Int. J. Syst. Bacteriol.* **45**:186-187.

61. **Mylvaganam, S. and P.P. Dennis.** 1992. Sequence heterogeneity between the two genes encoding 16S rRNA from the halophilic archaeobacterium *Haloarcula marismortui*. *Genetics*. **130**:399-410.
62. **Nandi, S., G. Khetawat, S. Sengupta, R. Majumder, S. Kar, R.K. Bhadra, S. Roychudhury, and J. Das.** 1997. Rearrangements in the genomes of *Vibrio cholerae* strains belonging to different serovars and biovars. *Int. J. Syst. Bacteriol.* **47**:858-862.
63. **Nelson, G.** 1989. Species and taxa: systematics and evolution, p. 60-81. *In* D. Otte and J.A. Endler (ed.), *Speciation and its consequences*. Sinauer Associates, Sunderland, Massachusetts.
64. **Nübel, U., B. Engelen, A. Felske, J. Snaidr, A. Wieshuber, R.I. Amann, W. Ludwig, and H. Backhaus.** 1996. Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. *J. Bacteriol.* **178**:5636-5643.
65. **Olsen, G.J., C.R. Woese, and R. Overbeek.** 1994. The winds of (evolutionary) change: breathing new life into microbiology. *J. Bacteriol.* **176**:1-6.
66. **Pace, N.R.** 1996. New perspective on the natural microbial world: molecular microbial ecology. *ASM News*. **62**:463-470.
67. **Pace, N.R.** 1997. A molecular view of microbial diversity and the biosphere. *Science*. **276**:734-740.
68. **Pace, N.R., D.A. Stahl, D.J. Lane, and G.J. Olsen.** 1985. Analyzing natural microbial populations by rRNA sequences. *ASM News*. **51**:4-12.
69. **Palleroni, N.J.** 1993. Structure of bacterial genome, p. 57-113. *In* M. Goodfellow and A.G. O'Donnell (ed.), *Handbook of new bacterial systematics*. Academic Press Ltd., London.

70. **Palleroni, N.J.** 1997. Prokaryotic diversity and the importance of culturing. *Ant. van Leuwen.* **72**:3-19.
71. **Palys, T., L.K. Nakamura, and F.M. Cohan.** 1997. Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int. J. Syst. Bacteriol.* **47**:1145-1156.
72. **Pennisi, E.** 1998. Genome data shake tree of life. *Science* **280**:672-674.
73. **Perez Canhos, V., G.P. Manfio, and L.D. Blaine.** 1993. Software tools and databases for bacterial systematics and their dissemination *via* global networks. *Ant. van Leeuwen.* **64**:205-229.
74. **Perez Luz, S., F. Rodriguez-Valera, R. Lan, and P.R. Reeves.** 1998. Variation of the ribosomal operon 16S-23S gene spacer region in representatives of *Salmonella enterica* subspecies. *J. Bacteriol.* **180**:2144-2151.
75. **Poulsen, L.K., G. Ballard, and D.A. Stahl.** 1993. Use of rRNA fluorescence in situ hybridization for measuring activity of single cells in young established biofilms. *Appl. Environ. Microbiol.* **59**:1354-1360.
76. **Probst, A., C. Hertel, L. Richter, L. Wassill, W. Ludwig, and W.P. Hammes.** 1998. *Staphylococcus condimenti* sp. nov., from soy sauce mash, and *Staphylococcus carnosus* (Schleifer and Fischer 1982) subsp. *utilis* subsp. nov. *Int. J. Syst. Bacteriol.* **48**:651-658.
77. **Ravin, A.W.** 1960. The origin of bacterial species: Genetic recombination and factors limitin it between bacterial populations. *Bact. Rev.* **24**:201-220.
78. **Ravin, A.W.** 1963. Experimental approaches to the study of bacterial phylogeny. *The American Naturalist.* **97**:307-318.

79. **Rossello, R., E. Garcia-Valdes, J. Lalucat, and J. Ursing.** 1991. Genotypic and phenotypic diversity of *Pseudomonas stutzeri*. System. Appl. Microbiol. **14**:150-157.
80. **Rossello, R., E. Garcia-Valdes, A.J.L. Macario, J. Lalucat, and E. Conway de Macario.** 1992. Antigenic diversity of *Pseudomonas stutzeri*. System. Appl. Microbiol. **15**:617-623.
81. **Rossello-Mora, R., B. Thamdrup, H. Schäfer, R. Weller, and R. Amann.** 1999. The response of the microbial community of marine sediments to organic carbon input under anaerobic conditions. System. Appl. Microbiol. in press.
82. **Rossello-Mora, R.A., J. Lalucat, W. Dott, and P. Kämpfer.** 1994. Biochemical and chemotaxonomic characterization of *Pseudomonas stutzeri* genomovars. J. Appl. Bacteriol. **76**:226-233.
83. **Rossello-Mora, R.A., J. Lalucat, and E. Garcia-Valdes.** 1994. Comparative biochemical and genetic analysis of naphthalene degradation among *Pseudomonas stutzeri* strains. Appl. Environ. Microbiol. **60**:966-972.
84. **Smibert, R.M. and N.R. Krieg.** 1994. Phenotypic characterization, p. 607-654. In P. Gerhardt, R.G.E. Murray, W.A. Wood, and N.R. Krieg (ed.), Methods for general and molecular bacteriology. American Society for Microbiology, Washington, D. C.
85. **Sneath, P.H.A.** 1977. The maintenance of large numbers of strains of microorganisms and the implications for culture collections. FEMS Microbiol. Lett. **1**:333-334.
86. **Sneath, P.H.A.** 1989. Bacterial nomenclature, p. 2317-2321. In S.T. Williams, M.E. Sharpe, and J.G. Holt (ed.), Bergey's manual of systematic bacteriology. Williams & Wilkins, Baltimore.

87. **Sneath, P.H.A.** 1989. Numerical taxonomy, p. 2303-2305. *In* S.T. Williams, M.E. Sharpe, and J.G. Holt (ed.), *Bergey's manual of systematic bacteriology*. Williams & Wilkins, Baltimore.
88. **Sneath, P.H.A.** 1992. International code of nomenclature of Bacteria. American Society for Microbiology, Washington, D. C.
89. **Sneath, P.H.A. and R.R. Sokal.** 1973. Numerical taxonomy. 1. W. H. Freeman and Company, San Francisco.
90. **Sokal, R.R. and C.T. J.** 1970. The biological species concept: a critical evaluation. *The American Naturalist*. **104**:127-153.
91. **Stackebrandt, E.** 1992. Unifying phylogeny and phenotypic diversity, p. 19-47. *In* A. Balows, H.G. Trüper, M. Dworkin, W. Harder, and K.-H. Schleifer (ed.), *The prokaryotes*. Second edition. Springer-Verlag, Berlin.
92. **Stackebrandt, E. and B.M. Goebel.** 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.
93. **Stahl, D.A.** 1996. Molecular approaches for the measurement of density, diversity and phylogeny, p. 102-114. *In* C.J. Hurst, G.R. Knudsen, M.J. McInerney, L.D. Stetzenbach, and M.V. Walter (ed.), *Manual of environmental microbiology*. ASM Press, Washington, DC.
94. **Staley, J.T.** 1997. Biodiversity: are microbial species threatened? *Curr. Opin. Biotechnol.* **8**:340-345.
95. **Staley, J.T. and N.R. Krieg.** 1989. Classification of procaryotic organisms: an overview, p. 2299-2302. *In* S.T. Williams, M.E. Sharpe, and J.G. Holt (ed.), *Bergey's manual of systematic bacteriology*. Williams & Wilkins, Baltimore.

96. **Tamaoka, J.** 1994. Determination of DNA base composition, p. 463-469. *In* M. Goodfellow and A.G. O'Donnell (ed.), Chemical methods in prokaryotic systematics. John Wiley and Sons, Chichester.
97. **Torsvik, V., R. Sørheim, and J. Goksøyr.** 1996. Total bacterial diversity in soil and sediment communities - a review. *J. Industr. Microbiol.* **17**:170-178.
98. **Trüper, H.G.** 1996. Help! Latin! How to avoid the most common mistakes while giving Latin names to newly discovered prokaryotes. *Microbiología SEM.* **12**:473-475.
99. **Trüper, H.G. and K.-H. Schleifer.** 1992. Prokaryote characterization and identification, p. 126-148. *In* A. Balows, H.G. Trüper, M. Dworkin, W. Harder, and K.-H. Schleifer (ed.), The prokaryotes. Second edition. Springer-Verlag, Berlin.
100. **Ursing, J.B., R.A. Rossello-Mora, E. Garcia-Valdes, and J. Lalucat.** 1995. Taxonomic note: a pragmatic approach to the nomenclature of phenotypically similar genomic groups. *Int. J. Syst. Bacteriol.* **45**:604.
101. **Van Niel, C.B. and M.B. Allen.** 1952. A note on *Pseudomonas stutzeri*. *J. Bacteriol.* **64**:413-422.
102. **Van Regenmortel, M.H.V.** 1997. Viral species, p. 17-24. *In* M.F. Claridge and H.A. Dawah (ed.), Species: the units of biodiversity. Chapman & Hall, London.
103. **Vandamme, P., B. Pot, M. Gillis, P. De Vos, and J. Swings.** 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol. Rev.* **60**:407-438.
104. **Vandamme, P., P. Segers, M. Ryll, J. Hommez, M. Vancanneyt, R. Coopman, R. De Baere, Y. Van de Peer, K. Kersters, R. De Wachter, and K.H. Hinz.** 1998. *Pelistega europaea* gen. nov., sp. nov., a bacterium associated with respiratory disease in pigeons: taxonomic structure and phylogenetic allocation. *Int. J. Syst. Bacteriol.* **48**:431-440.

105. **Ward, D.M.** 1998. A natural species concept for prokaryotes. *Curr. Opin. Microbiol.* **1**:271-277.
106. **Ward, D.M., R. Weller, and M.M. Bateson.** 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature.* **345**:63-65.
107. **Wayne, L.G., D.J. Brenner, R.R. Colwell, P.A.D. Grimont, O. Kandler, M.I. Krichevsky, L.H. Moore, W.E.C. Moore, R.G.E. Murray, E. Stackebrandt, M.P. Starr, and H.G. Trüper.** 1987. Report of the Ad Hoc Committee on reconciliation of approaches to Bacterial Systematics. *Int. J. Syst. Bacteriol.* **37**:463-464.
108. **Werman, S.D., M.S. Springer, and R.J. Britten.** 1996. Nucleic acids I: DNA-DNA hybridization, p. 169-203. *In* D.M. Hillis, C. Moritz, and B.K. Mable (ed.), *Molecular Systematics*. Sinauer Associates, Inc., Sunderland, USA.
109. **Wheelis, M.L., O. Kandler, and C.R. Woese.** 1992. On the nature of global classification. *Proc. Natl. Acad. Sci. USA* **89**:2930-2934.
110. **Whitman, W.B., D.C. Coleman, and W.J. Wiebe.** 1998. Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. USA* **95**:6578-6583.
111. **Woese, C.** 1998. The universal ancestor. *Proc. Natl. Acad. Sci. USA* **95**:6854-6859.
112. **Woese, C.R.** 1987. Bacterial Evolution. *Microbiol. Rev.* **51**:221-271.
113. **Woese, C.R.** 1992. Prokaryote systematics: the evolution of a science, p. 3-18. *In* A. Balows, H.G. Trüper, M. Dworkin, W. Harder, and K.-H. Schleifer (ed.), *The prokaryotes*. Second edition. Springer-Verlag, Berlin.

114. **Zhdanov, V.M.** 1953. Opređelitel virusov celovska i žitvotnykh (Guide to the human and animal viruses). Acad. Med. Sci. USSR, Moscow.
115. **Ziemke, F., M. Höfle, J. Lalucat, and R. Rossello-Mora.** 1998. Reclassification of *Shewanella putrefaciens* Owen's genomic group II as *Shewanella baltica* sp. nov. Int. J. Syst. Bacteriol. **48**:179-186.

This work has been revised by Hans G. Trüper, Brian Tindall and Peter Kämpfer whose contributions are acknowledged.

In der Grauen Reihe sind bisher erschienen:

- 1 Technikfolgenabschätzung: Konzeptionen im Überblick; Carl Friedrich Gethmann, Armin Grunwald, 9/96; 2. Aufl. 7/98
- 2 Umweltprobleme und globaler Wandel als Thema der Ethik in Deutschland; Carl Friedrich Gethmann, 9/96; 2. Aufl. 10/98
- 3 Sozialverträgliche Technikgestaltung: Kritik des deskriptivistischen Verständnisses, Armin Grunwald, 10/96
- 4 Technikfolgenbeurteilung der Erforschung und Entwicklung neuer Materialien. Perspektiven in der Verkehrstechnik. Endbericht zum Vorprojekt, Arbeitsgruppe Neue Materialien, 1/97
- 5 Zur Wissenschaftstheorie der Genetik. Materialien zum Genbegriff; Mathias Gutmann, Peter Janich, 4/97
- 6 Klimavorhersage und –vorsorge; Stephan Lingner, Carl Friedrich Gethmann, 7/97
- 7 Xenotransplantation. Ethische Fragen und Probleme; Jan P. Beckmann, 7/97
- 8 Perspektiven der Robotik. Überlegungen zur Ersetzbarkeit des Menschen; Michael Decker, 11/97
- 9 Philosophie in Rußland. Tendenzen und Perspektiven; Carl Friedrich Gethmann, Nikolaj Plotnikov, 5/98
- 10 Technikfolgenbeurteilung in Ländern Mittel- und Osteuropas; Gerhard Banse (Hrsg.); 6/98
- 11 Biodiversitätsforschung in Deutschland. Potentiale und Perspektiven; Mathias Gutmann, Wilhelm Barthlott (Hrsg.); 11/98
- 12 Biodiversität als Problem der Naturethik. Literaturreview und Bibliographie; Thorsten Galert, 12/98
- 13 Geistiges Eigentum und Copyright im multimedialen Zeitalter. Positionen, Probleme, Perspektiven; Gerhard Banse, Christian J. Langenbach (Hrsg.); 2/99

- 14 Materials Science in Europe; Karl-Michael Nigge; 3/99
- 15 Modelling Climate Change and its Economic Consequences. A review; Meinhard Schröder, Stephan Lingner (eds.); 6/99
- 16 Robotik. Einführung in eine interdisziplinäre Diskussion; Michael Decker (Hrsg.); 9/99
- 17 „Protection Profile“ – Ein industriepolitischer Ansatz zur Förderung des „neuen Datenschutzes“; Otto Ulrich; 11/99
- 18 Zur Umweltgefährdungsbewertung von Schadstoffen und Schadstoffkombinationen durch Reichweiten- und Persistenzanalyse; Ulrich Müller-Herold, Martin Scheringer; 12/99
- 19 Environmental Standards. Combined Exposures and their Effects on Human Beings and their Environment (Summary); Christian Streffer et al.; 01/00
- 20 Genetische Diagnostik und Versicherungsschutz. Die Situation in Deutschland; Felix Thiele (Hrsg.); 01/00
- 21 Entwicklung und Innovation; Michael Weingarten; 04/00