ACTA UNIVERSITATIS AGRICULTURAE SUECIAE

VETERINARIA 102

Phylogeny, Diversity, Detection: Multiple Uses of 16S rRNA Genes in Veterinary Bacteriology

Malin Heldtander Königsson

SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES

Phylogeny, Diversity, Detection:
Multiple Uses of 16S rRNA Genes in Veterinary Bacteriology

Malin Heldtander Königsson

Akademisk avhandling som för vinnande av filosofie doktorsexamen kommer att offentligen försvaras i Tvåans föreläsningssal, Klinikcentrum, Ultuna, fredagen den 27 april, kl 13.00.

Abstract

Ribosomal RNA in general and the 16S unit in particular has proven very useful for phylogenetic as well as identification purposes in microbiology. Among the advantages in using the 16S rRNA molecule and its genes can be mentioned the fact that it is present and has the same function in all self-replicating cells, and that regions with different variability, from universal to hypervariable, occur and are homologous between organisms. This thesis, based on five scientific publications, describes the use of 16S rRNA genes as a tool for the establishment of phylogenies and the investigation of intraspecific variation in several mycoplasma species and as target region for a species-specific PCR for detection of *R. salmoninarum*, the causative agent of bacterial kidney disease.

The phylogeny of four goat mycoplasmas and three seal mycoplasmas was established by sequence analysis of the 16S rRNA genes of the type strains of the respective species. All three seal mycoplasmas and two goat mycoplasmas belonged to different clusters of the hominis group, and the other two goat mycoplasmas belonged to the *M. mycoides* cluster of the spiroplasma group.

M. capripneumoniae causes contagious pleuropneumonia in goats, a severe disease that is responsible for great economic losses in Africa and Asia. The 16S rRNA genes in *M. capripneumoniae* have an unusually high number of nucleotide differences between the two operons within a strain, as well as between the homologous operons of different strains. The molecular evolution and epidemiology of 12 *M. capripneumoniae* strains from three neighbouring African countries were investigated by 16S rDNA sequence analysis.

M. agalactiae and *M. bovis*, the causative agents of mastitis and respiratory problems in small ruminants and cattle, respectively, also have high degrees of variability in their 16S rRNA genes. The intraspecific variation in 17 and 8 strains of the respective species was investigated. No distinct evolutionary patterns could be distinguished, despite a high degree of variation. Some conclusions could still be drawn from the data, particularly for *M. agalactiae*, where the non-European strains shared three characteristic nucleotides, and European strains from the same or neighbouring countries were very similar.

The 16S rRNA gene sequences from 8 strains of *R. salmoninarum* were analysed, and sensitive and specific PCR primers and a set of oligonucleotides for DNA preparation by sequence capture were constructed. An internal standard, a mimic molecule, was developed to identify false negative results due to inhibition of the amplification reaction.

This work demonstrates that the 16S rRNA still is a powerful tool in many areas of microbiology. It is often a suitable target in detection by PCR for diagnostic purposes and it can be used for studies of genetic diversity and sometimes even for molecular epidemiology. As for molecular phylogeny, the 16S rRNA might well be the most powerful tool, at least so far, simply because of its versatility.

Key words:16S rRNA; Evolution; Genetic diversity; *Mollicutes*; Mycoplasma; Phylogeny; PCR; *Renibacterium salmoninarum*

Phylogeny, Diversity, Detection: Multiple Uses of 16S rRNA Genes in Veterinary Bacteriology

Malin Heldtander Königsson
Department of Veterinary Microbiology,
Swedish University of Agricultural Sciences
and
Department of Bacteriology, National Veterinary Institute,
Uppsala, Sweden

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2001

Acta Universitatis Agriculturae Sueciae
Veterinaria 102

ISSN 1401-6257
ISBN 91-576-5925-7
© 2001 Malin Heldtander Königsson, Uppsala
Tryck: SLU Service/Repro, Uppsala 2001

Abstract

Heldtander Königsson, M. 2001. Phylogeny, diversity, detection: multiple uses of 16S rRNA genes in veterinary bacteriology. Doctor's dissertation.
ISSN 1401-6257, ISBN 91-576-5925-7,

Ribosomal RNA in general and the 16S unit in particular has proven very useful for phylogenetic as well as identification purposes in microbiology. Among the advantages in using the 16S rRNA molecule and its genes can be mentioned the fact that it is present and has the same function in all selfreplicating cells, and that regions with different variability, from universal to hypervariable, occur and are homologous between org-anisms. This thesis, based on five scientific publications, describes the use of 16S rRNA genes as a tool for the establishment of phylogenies and the investigation of intraspecific variation in several mycoplasma species and as target region for a species-specific PCR for detection of *R. salmoninarum*, the causative agent of bacterial kidney disease.

The phylogeny of four goat mycoplasmas and three seal mycoplasmas was established by sequence analysis of the 16S rRNA genes of the type strains of the respective species. All three seal mycoplasmas and two goat mycoplasmas belonged to different clusters of the hominis group, and the other two goat mycoplasmas belonged to the *M. mycoides* cluster of the spiroplasma group.

M. capripneumoniae causes contagious pleuropneumonia in goats, a severe disease that is responsible for great economic losses in Africa and Asia. The 16S rRNA genes in *M. capripneumoniae* have an unusually high number of nucleotide differences between the two operons within a strain, as well as between the homologous operons of different strains. The molecular evolution and epidemiology of 12 *M. capripneumoniae* strains from three neighbouring African countries were investigated by 16S rDNA sequence analysis.

M. agalactiae and *M. bovis*, the causative agents of mastitis and respiratory problems in small ruminants and cattle, respectively, also have high degrees of variability in their 16S rRNA genes. The intraspecific variation in 17 and 8 strains of the respective species was investigated. No distinct evolutionary patterns could be distinguished, despite a high degree of variation. Some conclusions could still be drawn from the data, particularly for *M. agalactiae*, where the non-European strains shared three characteristic nucleotides, and European strains from the same or neighbouring countries were very similar.

The 16S rRNA gene sequences from 8 strains of *R. salmoninarum* were analysed, and sensitive and specific PCR primers and a set of oligonucleotides for DNA preparation by sequence capture were constructed. An internal standard, a mimic molecule, was developed to identify false negative results due to inhibition of the amplification reaction.

This work demonstrates that the 16S rRNA still is a powerful tool in many areas of microbiology. It is often a suitable target in detection by PCR for diagnostic purposes and it can be used for studies of genetic diversity and sometimes even for molecular epidemiology. As for molecular phylogeny, the 16S rRNA might well be the most powerful tool, at least so far, simply because of its versatility.

Key words:16S rRNA; Evolution; Genetic diversity; *Mollicutes*; Mycoplasma; Phylogeny; PCR; *Renibacterium salmoninarum*

Author's address: Malin Heldtander Königsson, Department of Bacteriology, National Veterinary Institute, SE-751 89 Uppsala, Sweden.

Fylogeni, diversitet, detektion: flera sätt att använda 16S rRNA gener i veterinärbakteriologi.

Sammanfattning

Ribosomalt RNA i allmänhet och 16S enheten i synnerhet har visat sig mycket användbar i många sammanhang inom mikrobiologin. Bland de främsta fördelarna med 16S rRNA generna är att de förekommer hos alla självreplikerande organismer och har områden med växlande grad av variation, från universella till hypervariabla. Målregioner för PCR-primerar kan därför väljas beroende på om man vill att alla eller bara en enda art av bakterier ska amplifieras. 16S rRNA sekvensen återspeglar också på ett bra sätt hela genomet varför molekylen och dess gen är den mest använda för fylogenetiska studier av framför allt bakterier. Denna avhandling, som är baserad på fem vetenskapliga artiklar, beskriver hur 16S rRNA generna har använts för att fastställa fylogenier och bestämma inomartsvariation hos ett antal olika mykoplasmer—en stor grupp bakterier som orsakar sjukdomar hos djur, växter och människor—samt att utveckla ett PCR-system för detektion av laxpatogenen *Renibacterium salmoninarum*.

Fylogenin hos fyra getmykoplasmer och tre sälmykoplasmer bestämdes genom sekvensanalys av 16S rRNA-generna från typstammarna av respektive art. Tidigare fylogenetiska studier av mykoplasmer har delat in dem i 5 grupper och ett tjugotal kluster. Alla tre sälmykoplasmerna och två getmykoplasmer placerade sig i den s.k. hominis-gruppen, men i olika kluster och subkluster. Resterande två getmykoplasmer visade sig oerhört närbesläktade och hemmahörande i *M. mycoides*-klustret i spiroplasmagruppen.

Mycoplasma capricolum subsp. *capripneumoniae* orsakar elakartad lungsjuka hos getter, en sjukdom som medför oerhörda ekonomiska förluster i Afrika och Asien. 16S rRNA-generna hos *M. capripneumoniae* är unika i att de har ett ovanligt stort antal sekvensskillnader, både mellan operonen inom samma isolat och mellan motsvarande operon från olika stammar. Variationen uppvisar ändå en stabilitet genom att 11 polymorfismer är gemensamma för alla analyserade stammar av arten och att en polymorfism som uppträtt en gång verkar vara stabil. Denna rikliga variation användes för att hos 12 stammar av arten studera epidemiologin inom tre angränsande afrikanska länder och för att studera den molekylära evolutionen hos arten.

Mycoplasma agalactiae och *Mycoplasma bovis*, som orsakar mastiter och luftvägsproblem hos getter och får respektive nötkreatur, har också en förhållandevis hög grad av variation i sina respektive 16S rRNA gener. Inomartsvariationen hos 17 respektive 8 stammar av de båda arterna utreddes med delsyftet att undersöka om den molekylära evolutionen skulle kunna fastställas på liknande sätt som för *M capripneumoniae*. Trots en stor variation mellan stammarna kunde inga distinkta evolutionära mönster upptäckas. Dock kunde vissa slutsatser dras, främst för *M. agalactiae*, där alla icke-europeiska stammar delade tre karaktäristiska nukleotider och stammar från samma eller angränsande europeiska länder uppvisade stora likheter.

Generna för 16S rRNA från 8 stammar av *R. salmoninarum* sekvenserades och artspecifika fluorescerande PCR-primerar konstruerades liksom primerar för DNA-preparation med "sequence capture". En intern standard, en så kallad MIMIC, gjordes också för att kontrollera att inga resultat var falskt negativa på grund av att inhiberande faktorer hindrat amplifieringsreaktionen.

Det här arbetet visar att 16S rRNA fortfarande är ett kraftfullt verktyg inom mikrobiologin. Det är ofta en lämplig målregion för PCR-detektion i diagnostiskt syfte och kan användas för studier av genetisk diversitet och ibland till och med för molekylär epidemiologi. Vad gäller molekylär fylogeni kanske 16S rRNA är det kraftfullaste verktyg som hittats hittills, bara för att det är så allmänt användbart.

Contents

Appendix

This thesis is based on the following five papers, which will be referred to by their Roman numerals.

I. Heldtander, M., Pettersson, B., Tully, J. G., and Johansson, K.-E., 1998. Sequence analysis of the 16S rRNA genes and phylogeny of four recently described goat mycoplasmas *Mycoplasma adleri*, *Mycoplasma auris*, *Mycoplasma cottewii*, and *Mycoplasma yeatsii*. International Journal of Systematic Bacteriology 48: 263–268.

II. Heldtander Königsson, M., Pettersson, B., and Johansson, K.-E., 2001. Phylogeny of the seal mycoplasmas *Mycoplasma phocae* corrig., *Mycoplasma phocicerebrale* corrig. and *Mycoplasma phocirhinis* corrig. based on sequence analysis of 16S rDNA. International Journal of Systematic and Evolutionary Microbiology, in press.

III. Heldtander, M., Wesonga, H., Bölske, G., Pettersson, B., and Johansson, K.-E., 2001. Genetic diversity and evolution of *Mycoplasma capricolum* subsp. *capripneumoniae* strains from Eastern Africa assessed by 16S rDNA sequence analysis. Veterinary Microbiology 78: 13–28.

IV. Heldtander Königsson, M., Bölske, G., and Johansson, K.-E. Intraspecific variation in the 16S rRNA gene sequences of *Mycoplasma agalactiae* and *Mycoplasma bovis* strains. Submitted.

V. Heldtander Königsson, M., Ballagi-Pordány, A., Jansson, E., and Johansson, K.-E. Detection of *Renibacterium salmoninarum* in tissue samples by sequence capture and fluorescent PCR based on the 16S rRNA gene. Manuscript.

Offprints are published by kind permission of the journals concerned.

Abbreviations

AFLP	amplified fragment length polymorphism
bp	base pair
CBPP	contagious bovine pleuropneumonia
CCPP	contagious caprine pleuropneumonia
DNA	deoxyribonucleic acid
E.	*Escherichia*
ELISA	enzyme linked immunosorbent assay
J-C	Jukes-Cantor
M.	*Mycoplasma*
ML	Maximum likelihood
MLST	multi locus sequence typing
M. capripneumoniae	*Mycoplasma capricolum* subsp. *capripneumoniae*
MP	Maximum parsimony
N-J	Neighbor-joining
nt	nucleotide
PCR	polymerase chain reaction
PFGE	pulsed field gel electrophoresis
pm pattern	polymorphism pattern
R.	*Renibacterium*
RAPD	randomly amplified polymorphic DNA
rDNA	gene or segment of a gene coding for rRNA
REA	restriction enzyme analysis
rep-PCR	repetitive element PCR
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rrn	ribosomal RNA operon
rRNA	ribosomal RNA
S	the Svedberg unit
S-	semi-conserved region
tRNA	transfer RNA
U-	universal region
V-	variable region

Voici mon secret. Il est très simple: on ne voit bien qu'avec le coeur. L'essentiel est invisible pour les yeux.

Antoine de Saint-Exupéry

Introduction

The number of described prokaryotic species today, less than 5,000, might not seem very impressive at first glance, but if we consider that it is estimated that >99% of the observable species cannot be cultivated by standard methods (Hugenholtz *et al.*, 1998) and that an organism has to be cultivated to be recognised as a species it is easy to realise that their impact on the world is far greater than we understand at present. Estimates that suggest that the biomass of all micro-organisms may constitute >50% of the total protoplastic biomass on Earth further confirm that realisation. The bacterial impact on for example plants and animals, does not only involve the parasitic form where they cause many infectious diseases. Organelles like chloroplasts and mitochondria have evolved from symbiotic relationships between bacteria and eucaryotic cells where the bacterium has been engulfed inside the eukaryotic cell. Microbiologists have existed for several hundred years and have always wanted to understand, detect, describe and classify micro-organisms. The Dutch botanist Antonie van Leeuwenhoek was the first person to observe micro-organisms, or, as he described them, "very little animalcules", in his home-made microscope in 1674. The first attempt at a systematic arrangement of the micro-organisms, based on morphological observation in the microscope, was made by Otto Müller in the late 18th century. In 1872 and 1878 the first isolations in pure culture were done which opened new perspectives on phenotypic descriptions and classifications based on tests developed for the distinguishing between bacteria. Different systems for the classification of bacteria were suggested at the beginning of the 20th century. Bergey's Manual (1923) was the first attempt at a common standard for the identification and classification of bacteria. The development of molecular methods was the beginning of a new era in microbiology and a vast number of new possibilities for classification arose. For classification to be useful, it has to be as simple and natural as possible. This implies that it should reflect the evolutionary relationships between organisms. The 16S rRNA genes have, for reasons discussed below, been found very suitable for this purpose.

16S rRNA

The ribosome

The ribosome is present in all self-replicating cells and forms the heart of the protein-synthesizing machinery. In bacteria, the ribosome is composed of two subunits, a small and a large one, referred to as the 30S and the 50S subunit, respectively. The S refers to the Svedberg unit, which is proportional to the sedimentation velocity in an ultracentrifuge and reflects the size and shape of the molecule. Both subunits contain ribosomal proteins and ribosomal RNA (rRNA). The large subunit is composed of 36 proteins and two rRNAs: the 5S rRNA of 120 nucleotides (nt) and the 23S rRNA of 2,900 nt. The small subunit is built out

of 21 proteins positioned around the 16S rRNA, which is about 1,500 nt in length.

In general it is said that the rRNAs are synthesised together as a long precursor RNA containing all three forms of rRNAs. The precursor molecule is then processed to give the mature rRNA species, and the individual rRNAs are cut from the molecule with specific nucleases. However, several reports about separate rRNAs indicate that this linked arrangement of the genes and their transcription is not the universal requisite it was once thought to be (Mattsson, 1993).

Still, the typical bacterial rRNA operon has the following organization:

5' – 16S rRNA – (spacer tRNA) – 23S rRNA – 5S rRNA – (trailer tRNA) 3'

The number of spacer and trailer tRNAs, if present at all, varies with species and so does the number of rRNA operons. As a general rule the fast-growing bacteria have a larger number of rRNA operons than the slow-growing ones. Ribosomal RNA operons are designated *rrnA*, *rrnB*, and so on. In species with more than one rRNA operon it is not uncommon to find occasional sequence differences, referred to as microheterogeneities or polymorphisms, between the operons of a single species. Similarly one may find sequence differences between the homologous rRNA operons in different strains of the same species. The two types of sequence variations are collectively referred to as intraspecific variation (Clayton *et al.*, 1995).

The folding of the rRNAs into secondary structures depends on the ability of the bases to form hydrogen bonds and the information for this is contained in the sequence. The secondary structure of rRNA is very conserved (Fig. 1), something that is achieved by compensatory base changes, i.e. a base substitution in one strand of a putative helix should be compensated by a complementary base change in the other strand. This finding has proven very useful for establishing secondary structures of rRNAs. The primary structure in rRNA molecules, on the other hand, even though more conserved than the average gene, includes regions of different evolutionary variability. The different regions in the 16S rRNA molecule range from extremely low to extremely high variability and are referred to as universal (U-) regions and variable (V-) regions, respectively (Gray *et al.*, 1984). Scattered between the U- and V-regions are semi-conserved (S-) regions of intermediate variability. A prokaryotic 16S rRNA contains 8 universal regions denoted U1–U8, from the 5'-terminus, and 9 variable regions (V1–V9).

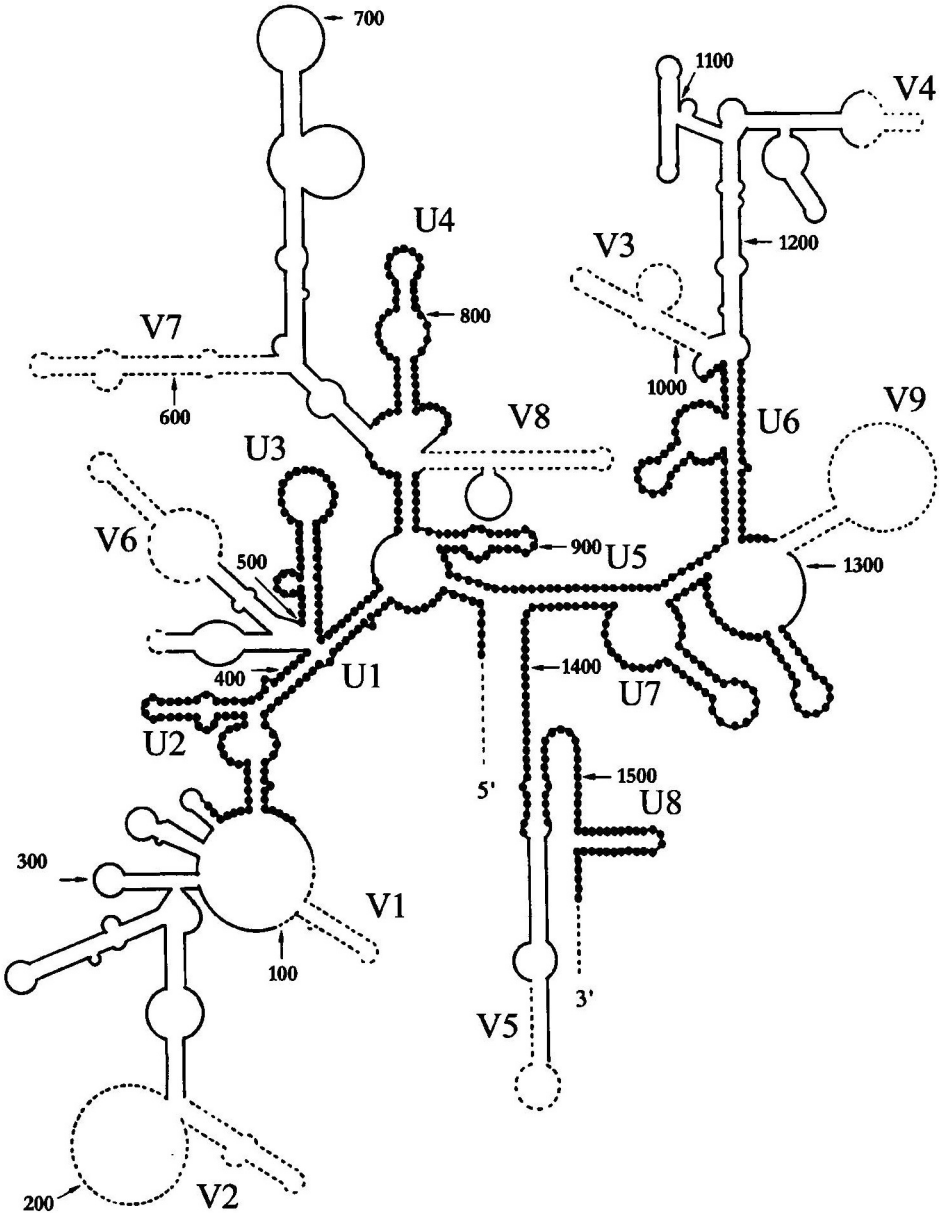

Figure 1. Secondary structure model of 16S rRNA from *E. coli*. Variable (V) regions are denoted by ---, semi-conserved (S) regions by — and universal (U) regions by -•-•-. The position of each 100th nucleotide is indicated with an arrow. Based on the model published by Woese *et al.* (1983). Adapted from Gray *et al.* (1984) and Johansson (1993).

Advantages and limitations of different uses of the 16S rRNA

The detection of micro-organisms in different sample materials is traditionally done by microbiological culture and although this is often still the most efficient way, there is a number of reasons why alternative methods are needed—fastidious organisms, slow growth, non culturable organisms, etc. The use of DNA probes and PCR has been and still is very common for this purpose, and the 16S rRNA or the 16S rRNA genes are often a suitable targets. The regions with different evolutionary variability allow detection and differentiation of micro-organisms on different taxonomic levels with the same target molecule.

Typing of micro-organisms with molecular methods has increased over the last ten years and a number of methods have been developed, one of which is typing by sequencing. Sequence typing can be based on different genes and 16S rRNA is fairly common.

Taxonomy was originally based on phenotypic data, such as serological and biochemical characteristics, but with the development of molecular methods the aspect of molecular phylogeny based on different genes has to be taken into account. By far the most common gene for phylogenetic characterisation of bacteria is the 16S rRNA gene. (Olsen, 1988; Olsen *et al.*, 1994; Olsen and Woese, 1993; Woese, 1987).

It is no coincidence that the 16S rRNA has become so widely used for the above mentioned purposes. The advantages of the molecule are many and have been noticed by many researchers, and they exceed by far its limitations. The fact that it is not only present in all self-replicating organisms, but also has the same function in all cells, increases the chance that a correct phylogenetic tree can be inferred. The regions of different variability include several advantages. Not only do they allow different levels of differentiation within the same molecule, but the universal regions also provide excellent PCR primer target sites that facilitate sequencing of the genes for phylogeny or species identification purposes. The universal regions are also a good help when aligning the sequences for phylogenetic inference. The vast number of 16S rRNA sequences already deposited in public databases, also provides a helpful tool for both species identification and phylogenetic analyses. Another advantage is that the 16S rRNA is not subject to any horizontal gene transfer. This is true in general, although evidence of possible exceptions from this rule has been presented (Wang and Zhang, 2000).

Despite the general excellence of the 16S rRNA, there are a few limitations that have to be considered. The variation in the molecule is usually not sufficient to distinguish between isolates of the same species, and even on the species level the degree of variation may sometimes be too low to separate closely related species by a 16S rRNA based PCR. This problem might, however, at least in some cases be solved by the addition of a restriction enzyme step to the

procedure. It is also interesting to note that there are examples of variation in the 16S rRNA genes, not only sufficient for distinction at the species level, but also high enough to use for epidemiological studies (Pettersson *et al.*, 1998). In phylogenetic discussions it is sometimes argued that one gene is not enough to reflect the true phylogeny, and that it would, therefore, be better to use a group of genes.

Streptomycin resistance

The first signs that changes in the ribosomes could be involved in antibiotic resistance came during the early 70's. It was shown that ribosomes with undermethylated RNA are resistant to kasugamycin, and that overmethylation of 23S rRNA is involved in MLS (macrolides, lincosamides, streptogramin type B compounds) resistance in *Staphylococcus aureus* (Cundliffe, 1990). Since then, several single mutations in both 16S and 23S rRNA have been found to be associated with resistance to a variety of antibiotics (Cundliffe, 1990; Prammananan *et al.*, 1998).

Resistance to streptomycin appeared immediately when this bactericidal antibiotic was used clinically (Finken *et al.*, 1993). Recently it was shown that transfer of antibiotic-resistance genes in soil populations of streptomycetes occurs (Wiener *et al.*, 1998). Mutational changes in the 16S rRNA of *Escherichia coli*, *Chlamydomonas* chloroplasts, Tobacco chloroplasts, and *Mycoplasma capricolum* subsp. *capripneumoniae*, have been shown to be associated with streptomycin resistance (Cundliffe, 1990; Montandon *et al.*, 1986). Alterations in protein S12 in the ribosomes is another mechanism that has been found to cause streptomycin resistance in bacteria (Finken *et al.*, 1993; Snyder and Champness, 1997).

Phylogeny

Molecular and bacterial phylogeny

The extensive use of molecular data in phylogenetic studies is a consequence of the development of various techniques in molecular biology since the late 1950's. During the 1960's and 1970's the dominating techniques were protein sequencing and, for closely related species or within populations, protein electrophoresis and DNA-DNA hybridisation. From the beginning of the 1980's, DNA sequence data from different genes is by far the most used.

The idea that the evolutionary history of a cell is documented in its macro molecules was presented by Zuckerkandl and Pauling in 1965 (Zuckerkandl and Pauling, 1965). Over the years, it has been concluded that most of the many qualitative differences between genotypic and phenotypic data favours the use of sequence data for phylogenetic studies (Olsen and Woese, 1993). For example, molecular character states, in contrast to many phenotypic characteristics, are

precisely defined and the number of definable characters is much higher at the genotypic level than at the phenotypic level. The sequence of a gene changes more often than the phenotypic outcome of the same gene, since many sequence changes do not affect the function of the gene. For the same reason, the genotypic evolution is continuous, while the phenotypic is not. The fact that similar phenotypes can be produced by many different genotypic changes also supports the use of genotypic data. The shortcoming of genotypic data is, however, the homoplasy resulting from the constant mutation that any individual site in a gene is subject to, and which the applied inference methods must be able to cope with.

About ten years after the pioneer work of Zuckerkandl and Pauling, Woese and Fox (1977) proposed the use of 16S rRNA sequences for phylogenetic studies, and the sequence analysis of 16S rRNA has since then become a standard method for phylogenetic studies of micro-organisms (Olsen and Woese, 1993; Woese, 1987).

Very little was known about bacterial phylogeny and evolution before the rRNAs were used for this purpose. The few things actually stated also tended to be wrong, simply proving the fact that neither morphology nor physiology of bacteria is very useful in defining the bacterial phylogeny (Woese, 1987). Based on rRNA data Woese and colleagues in 1990 proposed to change the existing classification of life—the two domains eukaryota and prokaryota—into a classification of three primary domains—*Archea*, *Bacteria* and *Eucarya* (Fig. 2) (Woese *et al.*, 1990). This classification is now widely accepted, though not undisputed (Mayr, 1998). With time, thanks to genome sequencing, more genes have been used for phylogeny of micro-organisms and the trees produced from data generated from these genes are not always in agreement with the trees obtained with rRNA data, which is one argument posed against the rRNA tree of life (Brown and Doolittle, 1997). This incongruence is explained by Woese as the result of rRNA totally lacking horizontal gene transfer, while most other genes suffer from it to at least some extent (Woese, 2000). The different views on classifying life into two or three domains can be explained by different perspectives. If we focus on multicellular organisms, what we can see and the evolution of the last billion years, the classification into two domains might seem logical. When looking at the first three billion years of evolution, trying to find the universal ancestor and its immediate descendants and using molecules and genes to do it the three domain solution seems much more likely to reflect the true genealogic relations (Woese, 1998a; Woese, 1998b).

16S rDNA sequencing

The first rRNA sequences were obtained in fragments by the oligonucleotide cataloguing method (Fox *et al.*, 1977). Sequence fragments generated by digestion with T1 ribonuclease produced a catalogue that was characteristic for each organism and could be used for phylogenetic comparison.

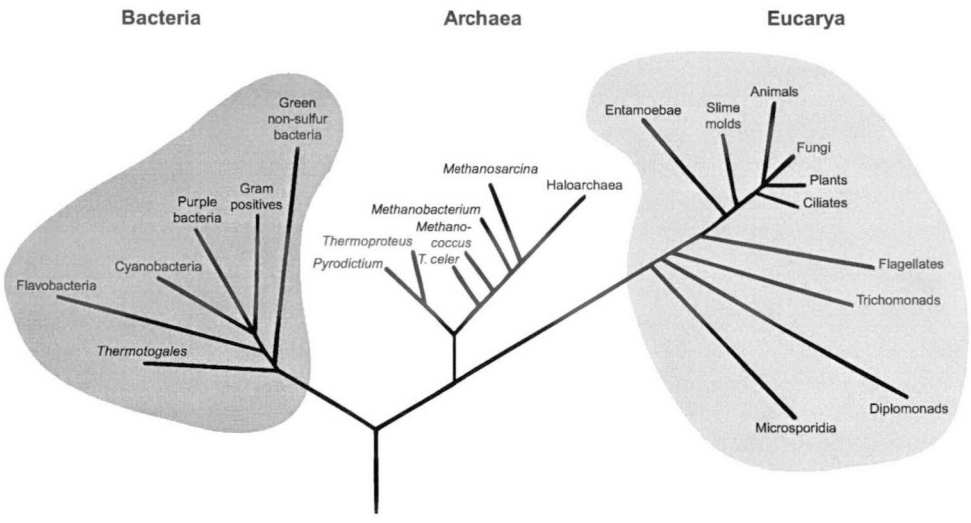

Figure 2. Universal phylogenetic tree, with branching orders based upon 16S rRNA sequence comparisons. Adapted from Olsen and Woese 1993.

Direct sequencing of rRNA without prior cloning became possible when the chain termination method by Sanger (Sanger *et al.*, 1977) was adapted for 16S rRNA sequencing (Lane *et al.*, 1985). The use of reverse transcriptase and DNA primers complementary to conserved regions of the 16S rRNA was a major breakthrough, but drawbacks, such as the frequent stretches of ambiguous positions and the disability to sequence the other strand, not to mention that the method is very time-consuming and cannot be used for uncultivable organisms, were still to be solved. The introduction of PCR (Mullis and Faloona, 1987; Saiki *et al.*, 1985) turned out to be the solution to at least some of these problems. Primers complementary to the conserved regions of the 16S rRNA genes have been published (Böttger, 1989; Pettersson, 1997), and the amplicons may either be cloned prior to sequencing or be sequenced directly. Direct sequencing of both strands of the amplicon can be done either by solid-phase sequencing or by cycle sequencing. The solid-phase method has been adapted to automation (Hultman *et al.*, 1989; Hultman *et al.*, 1991; Pettersson *et al.*, 1994a; Pettersson, 1997), and the use of single strands minimizes the background and increases the read length. The cycle sequencing method, on the other hand, is more rapid and less laborious when performed manually.

Phylogenetic inference

Aligning the sequences is the first step in phylogenetic analysis. An alignment is obtained by inserting gaps that correspond to insertions or deletions, in order to obtain positional homology between the sequences. Manual alignments are still considered the most accurate, since, despite the vast number of computer programs available to help in the matter, no computational program exists that automatically aligns the sequences correctly. More than 18,000 sequences of 16S

and 16S-like rRNA are accessible from public databases on the World Wide Web. GenBank (http://www.ncbi.nlm.nih.gov/) is probably the one containing the highest number of sequences, but rRNA sequences can be downloaded from the Ribosomal Database Project (http://www.cme.msu.edu/RDP/html/index.html) and the rRNA WWW server (http://rrna.uia.ac.be/), in a pre-aligned format, which lessens the manual alignment work somewhat.

There are several methods to infer a phylogenetic tree from a given sequence alignment. Three of the more common are Maximum Likelihood (ML), Maximum Parsimony (MP) and distance matrix methods (Li and Graur, 1991). Most important when using the ML approach to phylogenetic inference is the choice of substitution model. ML evaluates the probability that a chosen evolutionary model has given rise to the sequences observed in the data set. The method is based on a site-by-site analysis of the sequences and the phylogeny is inferred by finding the highest likelihood, i.e. the tree with a log-likelihood value closest to zero. The main advantages of ML are that the method usually copes better with distant sequences within a data set than other methods, and also that it is the method least affected by sampling error. The drawback is that it is both computer-intensive and time-consuming.

Several distance matrix methods exist but perhaps the most commonly used is the Neighbor-joining (N-J) method (Saitou and Nei, 1987). In N-J, branch lengths are calculated from a matrix of distance values between all pairs of taxa in the data set. A substitution model that corrects for multiple changes at single locations is chosen, and the tree is constructed with the criteria that the total length of its branches should be minimized. A N-J tree will never be exact, but is usually a good estimate of a reasonably true tree. The method is very popular, as the algorithm is fast, even when dealing with large data sets.

In MP, or Minimum Evolution, each position is treated individually, but only sites that are informative, in the sense that they favour some trees over the others, are considered (Li and Graur, 1991). The principle of the method is that the tree that needs the fewest changes to explain the observed sequence data is the correct one. The tree can be obtained by different tree searching methods and the choice here is dependent on the number of taxa analysed. Exhaustive search and Branch and bound are two exact search algorithms but unfortunately it is not possible to use them for data sets bigger than 10 and 20 taxa, respectively. For larger data sets a heuristic search, like Stepwise addition and Star decomposition, has to be applied. Unfortunately the shortest tree is not guaranteed to be found with a heuristic search.

A number of different models for nucleotide substitutions, with different amounts of free parameters that can be estimated from the data, have been developed over the years. The two most commonly used in bacterial phylogeny are the Jukes-Cantor (J-C) one-parameter model (Jukes and Cantor, 1969) and the

Kimura two-parameter model (Kimura, 1980). Both methods assume equal frequencies of all nucleotides, but the J-C model also assumes equality in nucleotide substitutions, while the Kimura model compensates for the fact that transitions are usually more frequent than transversions.

Bootstrap analysis can be used to estimate confidence limits for phylogenies (Felsenstein, 1985; Swofford *et al.*, 1996). Random sampling with replacement from the original data set is performed until a new data set containing the original number of data points is obtained. The procedure is repeated several times (100–1,000) and high bootstrap values usually, but not always, indicate stability of a clade.

There are also several program packages, either commercial or that can be downloaded as freeware from the Internet, available for phylogenetic inference. Only a few of the most widely used ones will be mentioned here.

PAUP, by David Swofford, was originally based only on parsimony analyses but now also includes ML and distance-based optimality criteria. Version 4.0 is to be released very soon, but since the release has already been postponed a couple of times, the beta version is for sale in the meantime. The PAUP user's manual is known as one of the best books on phylogenetic inference for both beginners and more experienced researchers.

PHYLIP, created by Joseph Felsenstein, is a free package of programs for inferring phylogenies, and includes most varieties of methods to infer phylogenetic trees. It can be downloaded from the PHYLIP homepage at http://evolution.genetics.washington.edu/phylip.html. A new version of PHYLIP, 3.6, has also been announced to be released soon.

BioEdit is a biological sequence-alignment editor that can be found at http://jwbrown.mbio.ncsu.edu/RNaseP/info/programs/BIOEDIT/bioedit.html, a website created by Tom Hall at the North Carolina State University. BioEdit provides a means of aligning sequences as well as inferring phylogenies. The program is free and includes all of the most common methods used to create phylogenetic trees.

Phylogeny based on 16S rRNA contra other genes

From time to time it is suggested that some gene or another reflects organismal phylogeny better than the 16S rRNA. Some of the more frequent ones are for example the 23S rRNA (Ludwig and Schleifer, 1994), the heat shock proteins, like hsp 70 (Sulaiman *et al.*, 2000) and the tuf gene (Baldauf *et al.*, 1996; Schneider *et al.*, 1997). All these genes have undoubtedly the advantage that they, at least among one or a few groups of organisms, have a higher resolving power than the 16S rRNA. On the other hand, most of them also share a couple of disadvantages, the greatest being the significantly lower number of sequences

available from these genes or proteins. Another problem in finding alternative phylogenetic markers that has been revealed by comparative analyses of the currently available genome sequences, is that the majority of the genes are not ubiquitously present. The number of possible alternative genes is, therefore, somewhat limited. Most genes are also subject to horizontal gene transfer to a greater extent than the 16S rRNA, a fact that also explains some of the discrepancies found between 16S phylogenies and phylogenies based on other genes or proteins (Woese, 2000).

In all, the agreement between tree topologies derived from 16S rRNA data and macromolecules with different functions is good enough to support the conclusion that the 16S rRNA trees very likely reflect the organismal phylogeny, at least globally (Ludwig and Schleifer, 1999).

Diversity

The term diversity is, at least in a biological context, definitely broad enough to need an explanation of what is meant by it in any given discussion. In bacteriology, the term is frequently used on the species, subspecies, or strain level, where the diversity discussed below species level is defined as microdiversity (Schloter *et al.*, 2000). Bacterial diversity is mostly caused by genetic variation, generated either by point mutations in the genome, reshuffling of segments of the genomic sequences, or horizontal gene transfer between organisms (Arber, 2000). In the era of whole-genome sequencing, it has also been revealed that there is a certain amount of diversity between the genomes within a single bacterial species (Lan and Reeves, 2000). Another form of diversity that is discussed within the bacterial domain is the division-level diversity. In 1987, the bacterial domain was described to contain 12 divisions (Woese, 1987), but today it is estimated that the number must be at least 40 (Hugenholtz *et al.*, 1998). In this work the diversity discussed concerns the 16S rRNA genes, and in part refers to differences between (closely related) species, but mostly to variation within a species, i.e. between strains, or even between the rRNA operons within a single strain.

Diversity in 16S rRNA

The 16S rRNA with its regions of different variability was for a long time considered to be completely stable within species. Complaints regarding the lack of differences between closely related and/or recently diverged species are often heard. Despite these facts examples of diversity within species on the 16S level have been reported (Clayton *et al.*, 1995; Martínez-Murcia *et al.*, 1999), and such variability has even been used for molecular epidemiology (Pettersson *et al.*, 1998).

In mycoplasmas

The mycoplasmas (micro-organisms belonging to the class *Mollicutes*) as a group have higher variability in the 16S rRNA genes than bacteria in general (Woese *et al.*, 1985; Woese, 1987). Perhaps the most striking example of this is *M. capripneumoniae*, with 11 polymorphisms in common for all strains analysed (Pettersson *et al.*, 1998), and in addition to that between 1 and 6 sequence differences more per isolate. Most of the ruminant mycoplasmas have two rRNA operons, and the two species *M. agalactiae* and *M. bovis* have also demonstrated tendencies to unusually high degrees of variability in the 16S rRNA genes (Subramaniam *et al.*, 1998).

In *Renibacterium salmoninarum*

Renibacterium salmoninarum was originally thought to harbour only one rRNA operon (Grayson *et al.*, 1999), which would seem natural in this slowgrowing bacterium. Recently, however, it was shown that the species actually contains two 16S rRNA operons, but that they are identical. One evidence for this complete identity is that no polymorphic positions are detected in the 16S rRNA sequences of the species. The rRNA genes of *R. salmoninarum* seems to be particularly stable, since strains differing widely in both date of isolation and geographical origin are identical or almost identical (Grayson *et al.*, 2000a).

The species concept

The species concept is something that is not only discussed among biologists of all disciplines, but also preoccupies many philosophers. For eukaryotes, there are at least 22 different concepts in use and consequently many arguments about which is the best one. In the world of microbiology, a species, although no official definition exists, is generally considered to be "*a group of strains that show a 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*" (Colwell *et al.*, 1995).

Classification of micro-organisms is done for several reasons. It is a good way to summarise the information about an organism and it is compulsory if we want to construct an identification system to be able to recognise new isolates of the species. The latter is, of course, especially important for diagnostic purposes of pathogenic species. Classification also plays an important role in tracing the origin and evolution of organisms.

Prokaryote taxonomists generally agree that the most reliable classification is achieved by the so-called polyphasic approach. This approach uses all genotypic (all data derived from nucleic acids), phenotypic (chemotaxonomic markers and expressed features, such as morphology, physiology and serology), and phylogenetic information when classifying species (Vandamme *et al.*, 1996). Recently, however, this approach to the species concept has been criticized as

being too conservative and lacking congruency with the concept delineated for higher organisms (Ward, 1998; Whitman *et al.*, 1998). For most of us, however, it is evident that comparing species concepts for eukaryotes and prokaryotes is difficult, because the basic notions of what constitutes a species would be necessarily different for higher organisms than for bacteria (Rosselló-Mora and Amann, 2001). One striking difference between the two groups is, for example, the sexual reproduction of the higher organisms versus the reproduction by binary fission or asexual spores, which is the most common in bacteria (although sexual reproduction by conjugation is an important way in which antibiotic resistance is spread among bacteria).

Apart from the discussion about which characters that are most suitable to be regarded in this description of new species, there is another fact that has to be considered in the matter. For an organism to be recognised as a species today, it has to be cultured. Since it has been estimated that >99% of the observable microbial species in nature is not cultivable by standard methods, it seems quite clear that the actual number of existing bacterial species is underestimated (Amann *et al.*, 1995). It has also been calculated that for an accurate description of a new species preferably 25 but at least 10 strains are needed (Sneath, 1977). Despite this fact, the trend seems to be going in the opposite direction. Of 156 new species described in 1999, 75% were based on one or two strains only, as opposed to 32% of the 44 newly described species in 1989 (Rosselló-Mora and Amann, 2001). New micro-organisms that cannot be cultivated may be described as *Candidatus* species if a sufficient number of characteristics have been reported (Murray and Stackebrandt, 1995).

A new name and a definition of the currently practised species concept were recently proposed by Rosselló et al (2001). They suggest that a phylo-phenetic species is "*a monophyletic and genomical coherent cluster of individual organisms that show a high degree of overall similarity with respect to many independent characteristics, and is diagnosable by a discriminative phenotypic property.*" The concept is based on numerical analysis of independently co-varying characters as well as phylogeny and genotypic cluster definition.

Molecular epidemiology

Typing and subtyping of micro-organisms is important epidemiologically for determining the source of infection, recognizing particularly virulent strains of organisms, and monitoring vaccination programs. To accomplish this, a number of different molecular methods have been invented over the years, with different advantages and drawbacks. The method of choice depends on available equipment and resources, and also on the exact epidemiological question at each occasion. To investigate the epidemiology of an infection on short term within a limited area, for example within a hospital during a few weeks or months, is quite different from analysing the possible common ancestry of the causative agent of a

disease that has appeared on different continents during several decades and therefore different demands are put on the typing methods.

Most of the methods used are based on examination of the whole chromosome. Pulsed field gel electrophoresis (PFGE) is often seen as the gold standard of molecular typing. Agarose plugs containing whole bacteria are obtained by combining bacterial isolates with molten agarose in small moulds. After treatment by in situ enzyme lysis and infrequent cutting restriction enzyme, electrophoresis with changed polarity of the current at predetermined intervals, i.e. pulsed field, allows clear separation of the large DNA fragments that can range from 10 to 800 kb (Schwartz and Cantor, 1984). The main drawback of the method is that it is time consuming. The procedure takes 2 to 3 days to complete.

Several PCR based typing methods that amplify either specific DNA sequences, e.g. locus-specific restriction fragment length polymorphism (RFLP) (Welsh and McClelland, 1990; Williams *et al.*, 1990), repetitive element (rep)-PCR (Versalovic *et al.*, 1991), or random fragments, e.g. randomly amplified polymorphic DNA (RAPD)-PCR and amplified fragment length polymorphism (AFLP) (Janssen *et al.*, 1996) have been developed. The discrimination power is generally good for these methods, but slightly lower for locus-specific RFLP. The interlab reproducibility is not very high for RAPD-PCR and rep-PCR, which can definitely be considered a drawback. All of the methods are considerably faster than PFGE, with the possible exception of AFLP that takes two days to perform.

In 1998, a new method for analysis of bacterial isolates, called multilocus sequence typing (MLST), was presented (Maiden *et al.*, 1998). About 450 bp from 7 different house-keeping genes are sequenced. Each sequence is called an allele and the seven alleles from one strain constitute an allele profile. Every specific allele profile represents a sequence type. The relationship between isolates is shown as dendrograms based on pairwise differences between allele profiles. The advantages of this method is that it shows a high degree of variation since it is based on as much as seven genes, but is still stable, since the chosen genes are house-keeping genes. This implies that the method is useful for both short- and long-term epidemiological questions. The results are unambiguous and easy to assemble in a common database that can be accessed on the World Wide Web for interested researchers.

Molecular subtyping based on sequencing of one single gene or region has also been proposed from time to time. The most frequent arguments against it has been that one gene usually does not provide enough variability, especially since the sequence must be flanked by highly conserved regions suitable as primer target sites. Even the 16S rRNA genes, with their regions of different variability, does not usually present the amount of variation between strains that is needed for epidemiological typing, although at least one exception has been found (see below and Pettersson *et al.*, 1998).

Detection and identification

Traditionally, detection of bacteria has been done by isolation and cultivation. Methods for bacterial identification are commonly based on nutritional, biochemical and physiological characters, or on analyses of profiles and composition of chemotaxonomic markers, such as cellular fatty acids, mycolic acids, polar lipids and polyamines (Busse *et al.*, 1996). Even though traditional microbiological culturing in many cases is the most efficient way to detect bacteria, there are several reason why new, molecular methods have been introduced. Depending on the target organism cultivation can be difficult, take an unreasonably long time or even be impossible. The longer it takes to grow a bacterium, the greater the risk of overgrowth by fast-growing competitors, even when using selective media. There are also organisms that are too dangerous to grow or handle in any viable form because of their pathogenicity. There are also certain limitations in the discriminating power of traditional typing methods. All these facts explain the development of the molecular methods that have appeared in the last decades.

PCR

Of all basic molecular biology techniques developed during the last decades of the 20^{th,} century PCR is probably the one that has had the greatest impact. Thousands of articles about PCR have been published since the original reports in 1985 and 1987 (Mullis and Faloona, 1987; Saiki *et al.*, 1985), plus many reports on new PCR-based methods (Nygren, 2000). The PCR technique was quickly adopted for the diagnosis of infectious diseases. The strength of PCR is the exponential increase of the original copy number of the desired target. This is obtained by the repeated synthesis of the target DNA, where the newly synthesised fragments act as templates in the following cycles. However, this high sensitivity also constitutes the greatest risk of the method. If anything but the desired target is amplified it will increase in the same exponential way, and because of this it is important to be very careful when designing primers, and also when preparing the samples (Belak and Ballagi-Pordany, 1993). A set of PCR primers should preferably be designed to target a region of the template DNA that is specific enough not to amplify any other DNA that might be found in the prepared sample. This usually requires a region with a certain extent of variability. On the other hand, the region has to be stable enough to be constant in all strains of the actual species.

Pros and cons of different methods

It is often argued that the reproducibility of the molecular microbiology diagnostic tests between different laboratories is low or not reliable. This might be true to a certain extent, but it is also noteworthy that similar reliability tests are seldom or never done on traditional microbiological testing procedures, and on the rare occasion it is done, it turns out that the results of these conventional

methods differ greatly between laboratories as well. (Springer *et al.*, 1996; van Belkum, 2000).

Sample preparation

The sample preparation for PCR must be done very carefully. Cross-contamination between samples as well as contamination from amplicons is always a risk, and this has to be tested by frequent blanks throughout the handling. Sample preparation and analyses of PCR products should preferably be conducted in separate rooms (Belak and Ballagi-Pordany, 1993). Inhibitors that prevent the reaction are another hazard especially when the samples derive from faeces, kidneys, or other complex biological material. In this case the outcome might be false negative results, which is just as bad as the possibly false positive results. A positive internal control can be used to avoid this (Ballagi-Pordany and Belak, 1996). The concentration of the sample can also be a problem. If it is a low grade infection, and thus very few microbial organisms are present, a large amount of tissue has to be used for sample preparation to increase the chance of capturing the microbe, but still very little material can be used as a template in PCR, which means that the template has to be concentrated before use. Concentration of the sample is most commonly obtained by using matrix binding DNA or magnetic beads (Olsvik *et al.*, 1994). In general, the sample preparation has become the bottleneck or rate-limiting step in the analysis, since the introduction of rapid-cycling PCR instruments and real time PCR.

Sensitivity in theory and practice

In a discussion concerning the respective sensitivity of PCR and traditional bacteriological culturing, PCR is usually considered the more sensitive method of the two. This might be true in theory, especially if the PCR in question is based on targets that occur in multiple copies in each organism, like rRNAs and insertion elements. In reality, however, there is a number of considerations to take into account that complicate the situation.

Another fact to be considered is that PCR, unlike cultivation, also detects DNA from non-viable organisms. This can be an advantage when handling bacteria sensitive to transport and storage, but is a drawback when the PCR detects residual nucleic acid after initiation of effective therapy. Bacterial culture would not give a positive result in any of these two cases.

The mollicutes

The mycoplasmas, or mollicutes, is a group of organisms that are closely related to the Gram-positive bacteria. Based on their permanent lack of a cell wall, tendency to form fried-egg-like colonies on solid media, ability to pass through membrane filters of 450 nm and 220 nm in pore size, low G+C content of the genome and minute genome sizes of 0.58 to 2.2 Mbp (International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of Mollicutes., 1995),

they are arranged into a separate class called the *Mollicutes* (*mollis*, soft; *cutis*, skin, in Latin). Mollicutes are found in humans, mammals, reptiles, fish, arthropods and plants (not to mention that they contaminate cell cultures), and the list of host species is continuously increasing. The mollicutes are usually regarded as host-specific even though several exceptions have been reported (Nicholas *et al.*, 1998; Razin, 1992). Their general host and tissue specificity probably reflects their nutritionally exacting nature and obligate parasitic mode of life (Razin, 1992). The lists of new hosts and new mollicutes is only restricted by the willingness of mycoplasmologists to look for new species and, unfortunately, also because it has become increasingly difficult to describe a new species, for reasons discussed below. The first mycoplasma was isolated and described only at the end of the 19th century, even though the disease that it causes, contagious bovine pleuropneumonia (CBPP), had been known since the end of the 17th century (Nocard and Roux, 1898). In 1955, it was given the name *Mycoplasma mycoides* and became the type species for all mollicutes, however the name was in time changed to *Mycoplasma mycoides* subsp. *mycoides,* small colony type (SC).

One of the first whole genomes to be sequenced was the 580 kb genome of *Mycoplasma genitalium* (Fraser *et al.*, 1995). At the time, it was the smallest existing bacterial genome known, but today one of the uncultivable phytoplasmas, the Bermuda grass white leaf phytoplasma, has been found to have an even smaller genome of only 530 kb (Marcone *et al.*, 1999). The minute genome sizes of the mollicutes make them very interesting as study organisms for comparative analyses and in the attempts to artificially create the minimal

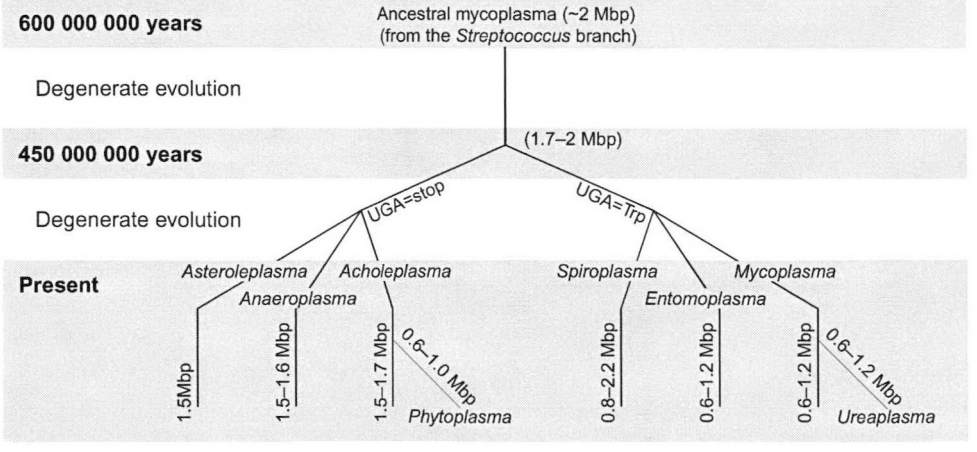

Figure 3. Hypothetical scheme for macroevolution of the *Mollicutes*. Members of the genus *Mesoplasma* and of the *Mycoplasma mycoides* cluster are also included in the *Entomoplasma* branch.

genome needed for an organism (Razin, 1997). The choice of *M. genitalium* as one of the first species for whole genome sequencing was, therefore, a logical one.

The special nature of the mollicutes has caused a lot of controversy and confusion around this group of bacteria along the years. Until the 1930s, when the true nature of viruses was defined, they were considered to be viruses, probably mostly because of their ability to pass through filters that no other bacteria could. Another misconception has been that they are no more than stable L-phase variants, i.e. wall-less forms, of common bacteria. This discussion was put to an end, however, when the first genomic analysis data could rule out any relationship between the mollicutes and stable L-forms of present-day bacteria. Over the years, it has also been discussed whether the mollicutes are relics from the primitive bacteria preceding modern bacteria, thus placing them at the root of the phylogenetic tree, or the product of degenerative evolution from walled bacteria. The most probable hypothesis presented so far, states that the mollicutes have evolved by genome reduction from a common ancestor with a genome of about 2,000 kb that derived from the *Streptococcus* phylogenetic branch and existed 600 million years ago. About 450 million years ago, the mollicute branch split into two major branches, one of which is recognised by its use of UGA as the codon for tryptophane (Trp) while the other uses UGA as stop codon like most other bacteria. One of these branches led to the *Anaeroplasma*, the *Asteroleplasma*, the *Acholeplasma*, and later the *Phytoplasma* branches, while the other split into the *Spiroplasma* branch, the *Entomoplasma* branch and the *Mycoplasma* branch, from which the *Ureaplasma* branch later arose (Fig. 3) (Maniloff, 1992; Razin *et al.*, 1998).

Today, the class *Mollicutes* consists of about 200 species that are taxonomically divided into four orders, five families and eight validly described genera (Tully *et al.*, 1993) (Table 1). A ninth group, the phytoplasmas, has been described (Seemüller *et al.*, 1998). Many species within the phytoplasma group have been described as *Candidatus* species but because these organisms cannot be cultivated, their taxonomic status has not been properly established, although they are phylogenetically related to the acholeplasmas. In 1989, the phylogeny of the mollicutes based on 16S rRNA was established and revealed five distinct groups, each including several clusters and subclusters (Weisburg *et al.*, 1989) (Fig. 4). Today, the mollicutes are still divided into the same five major groups, but the number of clusters and subclusters within some of the groups has increased considerably (Pettersson *et al.*, 1996a; Pettersson *et al.*, 2000; Pettersson *et al.*, 2001)

Many of the mycoplasmas isolated from different animals are pathogenic and of great concern in veterinary medicine. In ruminants, *M. mycoides* subsp. *mycoides* SC, *M. mycoides* subsp. *mycoides* LC, *M. agalactiae*, *M. bovis*, *M. capripneumoniae*, *M. capricolum* subsp. *capricolum*, *M. alkalescens*,

Table 1. Taxonomy of the class *Mollicutes*.

Classification[a]	No. of Species[b]	Habitat
Order I, *Mycoplasmatales*		
Family I, *Mycoplasmataceae*		
Genus I, *Mycoplasma*	102	Humans, animals
Genus II, *Ureaplasma*	6	Humans, animals
Order II, *Entomoplasmatales*		
Family I, *Entomoplasmataceae*		
Genus I, *Entomoplasma*	6	Insects, plants
Genus II, *Mesoplasma*	12	Insects, plants
Family II, *Spiroplasmataceae*		
Genus I, *Spiroplasma*	34	Insects, plants
Order III, *Acholeplasmatales*		
Family I, *Acholeplasmataceae*		
Genus I, *Acholeplasma*	13	Animals, insects, plants
Order IV, *Anaeroplasmatales*		
Family I, *Anaeroplasmataceae*		
Genus I, *Anaeroplasma*	4	Bovine/ovine rumen
Genus II, *Asteroleplasma*	1	Bovine/ovine rumen

[a]The phytoplasmas form a sister lineage to the acholeplasmas, but they have not been recognised as a genus because they cannot be cultivated.
[b]Number of validly described species within each genus.

M. bovigenitalium, M. bovirhinis, M. bovoculi, M. californicum and *M. ovipneumoniae* are examples of species causing different severe diseases (DaMassa *et al.*, 1992; Hale *et al.*, 1962; MacOwan, 1984; MacOwan and Minette, 1976; Ross, 1993; Simecka *et al.*, 1992).

Renibacterium salmoninarum

R. salmoninarum is a slow-growing, highly fastidious and strongly Gram-positive bacterium that is non-motile, non-acidfast and does not produce spores. The G+C value is 55.5%, and the species displays a variety of unique features such as the cell-wall polysaccharide composition and an unusually large number of glycoproteins. *R. salmoninarum* is the only species within its genus and based on 16S rRNA analysis, it has been placed in the *Arthrobacter-Micrococcus* subline of the actinomycetes, most closely related to the *Arthrobacter* species. The species is the causative agent of bacterial kidney disease (BKD), a disease first

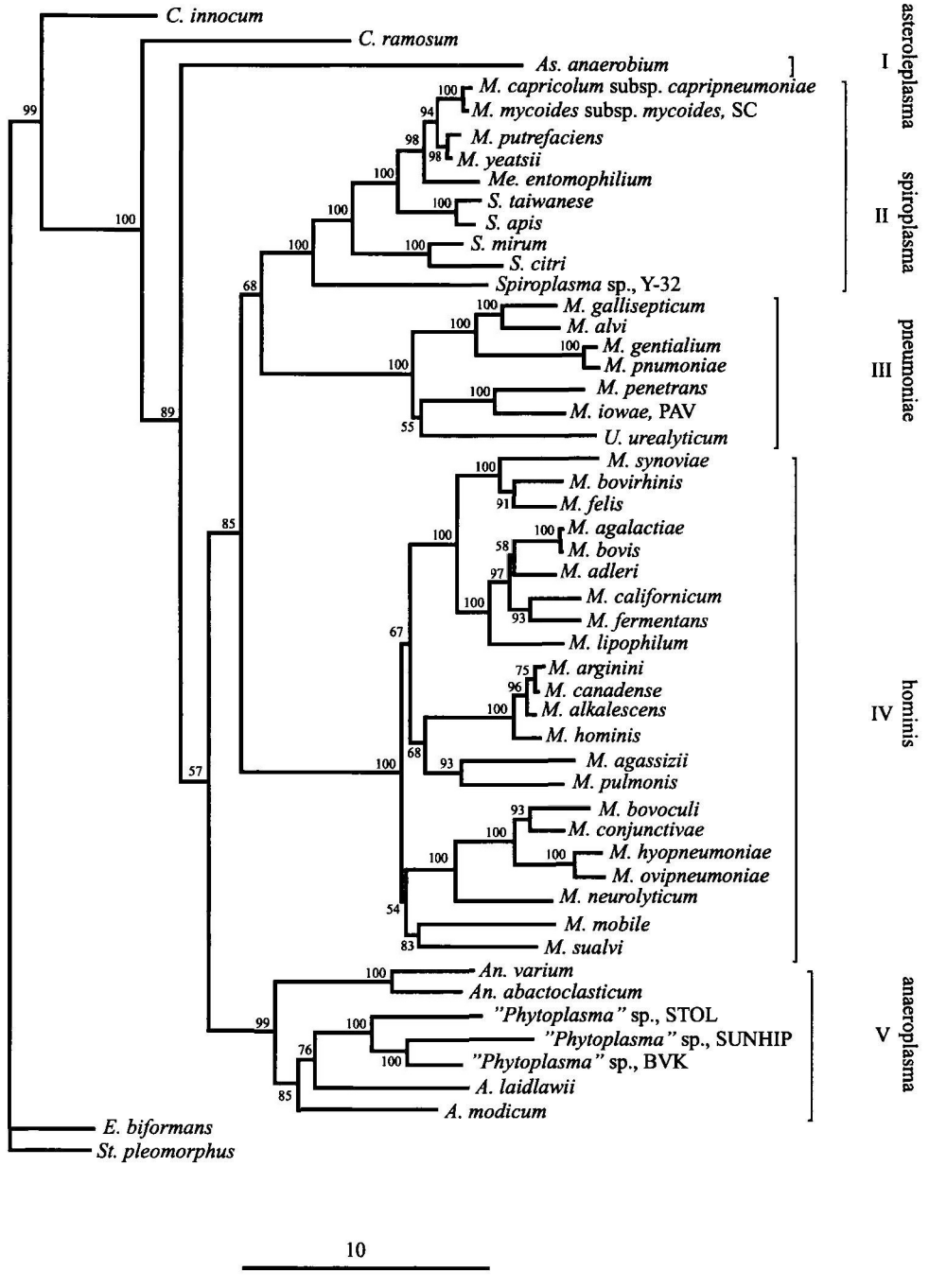

Figure 4. Phylogenetic tree based on 16S rRNA sequences, representing all groups of mycoplasmas. Representatives of the closely related genera *Clostridium* and *Eubacterium* are also included in the tree. *Streptococcus pleomorphus* and *Eubacterium biformans* represent outgroups. Adapted from Johansson *et al.* (1998).

detected in the river Dee in Scotland and, therefore, first described under the name Dee disease in 1930. This disease is today of major concern in all wild and cultured salmonid fish throughout the world (Evenden *et al.*, 1993; Fryer and Lannan, 1993). It has long been known that the disease is spread through both horizontal and vertical transmission, although the route of transmission and the molecular mechanisms involved in the pathogenicity of the bacterium are still mostly unresolved. One factor thought to largely contribute to the pathogenicity of *R. salmoninarum* is the ability to survive and replicate within the phagocytic cells of the host. This ability allows the bacterium to avoid not only the defence mechanisms of the host, but also the effects of anti-microbials used for therapy. It has also been noted that the bacterium can survive up to 21 days in faeces and sediment collected from a fresh-water environment.

Typing of different isolates of *R. salmoninarum* has proven very difficult due to an extraordinary uniformity among isolates of the pathogen, which has hampered the knowledge of the epizootiology of the disease. Several different typing methods, such as the 16S-23S intergenic spacer, 23S-5S intergenic spacer, exact tandem repeat locus and RAPD, have been evaluated, but found insufficient for distinguishing between some of the isolates from unrelated sources (Grayson *et al.*, 1999; Grayson *et al.*, 2000a; Grayson *et al.*, 2000b). Length variations between the tRNA intergenic spacer regions was recently investigated and revealed to be the most sensitive typing method so far (Alexander *et al.*, 2001).

Immunological techniques, such as immunofluorescence techniques (Bullock and Stuckey, 1975; Elliott and Barila, 1987) and enzyme-linked immunosorbent assays (ELISAs) (Gudmundsdóttir *et al.*, 1993; Jansson *et al.*, 1996; Pascho and Mulcahy, 1987), have been developed for diagnosis of the disease, as cultivation is not the best alternative with this fastidious and slowgrowing bacterium (Benediktsdóttir *et al.*, 1991). Complementary techniques are, however, still requested for confirmation, as there are reports of false positive serological reactions (Armstrong *et al.*, 1989; Austin *et al.*, 1985).

Aims of the present investigations

To establish the phylogeny of four goat and three seal mollicutes of the genus *Mycoplasma*, based on 16S rDNA sequence analysis. This aim forms a part of the extensive project to sequence the 16S rRNA genes of all mollicutes of the genus *Mycoplasma*, in order to use 16S rDNA sequencing as a tool for the description of new species within this genus in the future.

To investigate the extent of the intraspecific variation in the 16S rDNA sequences of 13 strains of *M. capripneumoniae* from a limited geographical area and to determine whether the resulting polymorphisms could be used as epidemiological markers.

To investigate the intraspecific variation in the 16S rRNA genes of the species *M. agalactiae* and *M. bovis*. The variation is interesting both as a possible tool for epidemiological studies and to evaluate the reliability of the existing detection systems, based on PCR of the 16S rRNA genes of these two species.

To construct and evaluate a species-specific detection system for *R. salmoninarum* based on fluorescent PCR targeting the 16S rRNA gene. The aim is to design a system which is useful for diagnosis of BKD.

Comments on materials and methods used

A more detailed description of the various materials and methods used in this study are given in papers I–V.

Bacterial strains (I–V)

For all phylogenetic studies (I and II), the type strains of the respective species were cultured in appropriate media and DNA was isolated by conventional phenol-chloroform preparation.

The field isolates used for sequence analysis in studies I and III–V were identified by standard biochemical tests, grown in appropriate media and then treated in the same way as the type strains. To guarantee pure cultures, the mycoplasma field strains were cloned five times prior to DNA isolation, if this was not previously done

16S rDNA sequencing (I–V)

Almost complete (96%) sequences of the 16S rRNA genes were obtained by solid phase sequencing (I and III) or cycle sequencing (II–V) of PCR products amplified from genomic DNA. The amplicons were generated by semi-nested PCR or single PCR with primer sets complementary to the universal regions U1/U8 (first/single PCR) and to U1/U5 and U2/U8 (semi-nested PCR), as defined by Gray et al (Gray *et al.*, 1984). All sequences were subjected to bi-directional rDNA sequencing. Single strands for solid-phase sequencing (I and III) were obtained by magnetic separation with paramagnetic beads and sequencing reactions were carried out according to the Sanger chain-termination method (Sanger *et al.*, 1977) with T7 DNA polymerase and fluorescently labelled sequencing primers (Pettersson, 1997). Cycle sequencing (II–V) with ThermoSequenase (Amersham Pharmacia Biotech, Sweden) was carried out according to the manufacturer's recommendations with the same set of fluorescently labelled primers as above. An ALFexpress™ (Amersham Pharmacia Biotech, Sweden) automated sequencer was used to read the sequences.

The shift from solid-phase to cycle sequencing was done mainly for two reasons. First of all, the cycle sequencing is a lot less laborious than the solid-phase method, unless an automated solid-phase system is available. For manual handling the solid phase procedure takes several hours and follows after two separate PCR reactions i.e. the first and the semi-nested amplification. The cycle sequencing reaction, on the other hand, is set up in about 30 minutes, the product from one single PCR can be used as template, and the actual reaction takes about 45 minutes in a thermo-cycler. The second main reason for changing to cycle

sequencing has to do with the increased risk for carry over contamination that always follows a nested or semi-nested PCR, i.e. any PCR using PCR products as template (Belak and Ballagi-Pordany, 1993). In our experience the cycle sequencing results in sequences with as high resolution and quality as the solid-phase technique and is, therefore, to be preferred in a laboratory without an automated solid-phase system.

In paper III, the sequences of the 16S rRNA genes of both operons of the studied strains were determined. The sequences of the *rrnA* operon were deduced by subtracting the sequence obtained with the *rrnB* operon-specific primer pair from the sequence obtained with the general primer pair (Johansson *et al.*, 1998; Pettersson *et al.*, 1998).

Phylogenetic analysis (I–II)

The sequences were aligned manually using the Genetic Data Environment software (GDE) (Smith, 1992). Gaps, and in paper I also ambiguities, were removed from the final data sets. Distance matrices corrected for multiple base changes at single locations by the one parameter model of Jukes & Cantor were constructed and phylogenetic trees were computed by the neighbor-joining program NEIGHBOR (Saitou and Nei, 1987) included in the phylogenetic program package PHYLIP, version 3.51c (Felsenstein, 1993). Bootstrap analyses were performed in sets of 500 (paper I) or 1,000 (paper II) resamplings with the SEQBOOT program.

PCR detection of *R. salmoninarum* (V)

Two sets of primers based on the 16S rRNA genes were constructed for the specific detection system developed for *R. salmoninarum*. One set was used for DNA isolation by sequence capture, which is the reason why both primers have a biotin handle in the 5' end. After hybridisation under agitation at 60°C for 3 h the samples were immobilized onto paramagnetic beads at room temperature for 2 h and then washed twice before the beads with the captured DNA segment were diluted in 20 µl of distilled water. Half of the capture product was used as template in PCR. DNA preparation with sequence capture is a good method when working with samples likely to contain substances that can inhibit the amplification reaction.

The second primer set was used for species-specific PCR detection of *R. salmoninarum*. One of the primers was Cy5-labelled to allow detection on the automated sequencer ALFexpress™ (Amersham Pharmacia Biotech) and evaluation with ALFwin Fragment Analyser 1.01. This increased the sensitivity of the PCR by at least one order of magnitude compared to conventional ethidium bromide staining after agarose gel electrophoresis.

A mimic molecule was constructed to be used as an internal control in the system. The mimic was produced by PCR of a segment of the human β-actin gene (GenBank accession no X00351) inserted into a pUC18 plasmid (Ballagi-Pordany and Belak, 1996; Englund *et al.*, 1999). The use of an internal control ensures that no false negative results are reported, since the lack of a mimic product in any sample after PCR will indicate that the amplification reaction has been inhibited.

Screening and selection for streptomycin resistance (III–IV)

The occurrence of either a thymidine or an adenine in the position corresponding to position 912 in *Escherichia coli* has been associated with streptomycin resistance in several bacteria (Cundliffe, 1990; Finken *et al.*, 1993; Montandon *et al.*, 1986). Among the almost 40 *M. capripneumoniae* strains analysed so far, a C→T mutation in this exact position has taken place at two occasions independently, which is remarkable. All the analysed strains of *M. agalactiae* and *M. bovis* harboured a thymidine in this position. This stimulated us to find out what the tendencies regarding streptomycin resistance were within these strains and also whether it is possible to select for streptomycin resistance in susceptible *M. capripneumoniae* strains by applying a selection pressure. The investigations were not intended to be complete antimicrobial drug-resistance investigations with exact MIC values, but merely a simple way to give indications of the tendencies concerning streptomycin resistance in these three species.

Five *M. agalactiae*, three *M. bovis* and seven *M. capripneumoniae* strains were tested for streptomycin susceptibility. The strains were cultured on F-plates and FP-plates, respectively, in the presence of a 6 mm PDM Antibiotic Sensitivity Disc II with 30 mg of streptomycin (AB Biodisk, Solna, Sweden), and the diameter of the growth inhibition zone was measured after 3–5 days, if present.

Four *M. capripneumoniae* strains that were found susceptible to streptomycin were grown on FP-plates for several passages, and susceptibility to streptomycin was tested during each passage with a PDM Antibiotic Sensitivity Disc II with 30 mg of streptomycin. Colonies growing on the border of the inhibition zone were used for the next passage. The clones were passaged until streptomycin resistance was obtained. Resistant strains were cultured as above, DNA was prepared by conventional phenol-chloroform extraction and sequence determination was performed as above.

Results and discussion

Phylogeny of the genus *Mycoplasma* (I–II)

The genus *Mycoplasma* is by far the largest of the *Mollicute* genera and includes 103 validly described species, 3 subspecies and 5 *Candidatus* species today. The number is constantly increasing, but slower than expected because it has become very laborious to describe a new species according to the minimum standard document (International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of Mollicutes., 1995). This document regulates how a new species is to be placed in the right genus by biochemical tests and based on growth conditions. Then the potentially new species has to be tested serologically against all species of the genus it belongs to. This is all very well for most genera since they usually include 5–15 and at the most 34 species, but for the genus *Mycoplasma* it involves over a hundred tests, and both performing the actual tests, but even more keeping the antisera in stock, is very difficult. The 16S rRNA has proven very useful for bacterial phylogeny in general (Olsen and Woese, 1993), and the agreement with mycoplasma taxonomy in the mycoplasmas is particularly good (Maniloff, 1992; Weisburg *et al.*, 1989; Woese *et al.*, 1980). It has, therefore, been suggested that 16S rDNA sequencing should be introduced as a tool for the description of new species within the genus *Mycoplasma*. For this to be possible, the 16S rRNA gene sequences from all described mollicutes of the genus *Mycoplasma* must be analysed. This has been done in a collaboration project between the Royal Institute of Technology and the National Veterinary Institute that was finished only last year (Johansson *et al.*, 2000) (Pettersson *et al.*, 2000; Pettersson *et al.*, 2001), although some of the sequences remain to be deposited in GenBank. The two phylogeny projects (I–II) in this work forms a part of that study.

The 16S rRNA genes from a total of 45 species were analysed in the study (some re-sequenced though). The results showed that the 103 species of the genus *Mycoplasma* were dispersed in 4 of the 5 phylogenetic *Mollicute* groups in about 20 clusters, some of them newly described from the study (Heldtander *et al.*, 1998; Pettersson *et al.*, 1994b; Pettersson *et al.*, 1996b; Pettersson *et al.*, 1998; Pettersson *et al.*, 2000; Pettersson *et al.*, 2001).

Phylogeny of goat and seal mycoplasmas

The 16S rDNA sequences were determined and the molecular phylogeny established for four goat and three seal mycoplasmas. Three of the goat mycoplasmas, *Mycoplasma auris*, *Mycoplasma cottewii* and *Mycoplasma yeatsii*, were isolated from the external ear canal of goats in Australia (DaMassa *et al.*, 1994) and the fourth one, *Mycoplasma adleri* was isolated from a goat's abscessed ankle in Maryland, USA (Del Giudice *et al.*, 1995). Of the three seal mycoplasmas, two, *Mycoplasma phocicerebrale* and *Mycoplasma phocirhinis*

(formerly *Mycoplasma phocacerebrale* and *Mycoplasma phocarhinis*, respectively), were isolated from pus from a seal lung and from the brain of a necropsied seal, respectively, during the seal epidemic in the Baltic Sea and in the North Sea during 1988 (Giebel *et al.*, 1991). The third seal mycoplasma, *Mycoplasma phocae* (formerly *Mycoplasmas phocidae*), was isolated from the respiratory tract of one of the more than 400 harbor seals that died during the virus epidemic along the New England coast in 1979 and 1980 (Ruhnke and Madoff, 1992). None of these mycoplasma species have been recognised as primary pathogens, but might possibly be involved in the production of pathological changes observed.

M. cottewii and *M. yeatsii* were found to belong to the *Mycoplasma mycoides* cluster of the spiroplasma group. They are like *M. putrefaciens*, however, not included in the classical *M. mycoides* cluster. The other two goat mycoplasmas and all three seal mycoplasmas were all parts of the hominis group but in different clusters and subclusters. *M. auris*, *M. phocicerebrale* and *M. phocae* clustered in the *M. hominis* cluster. *M. auris* and *M. phocicerebrale* were both comprised in the well defined *Mycoplasma alkalescens* subcluster, while *M. phocae* was positioned somewhere outside the *M. alkalescens* subcluster where the internal nodes of the *M. hominis* cluster are associated with rather weak bootstrap values. One goat and one seal mycoplasma, *M. adleri* and *M. phocirhinis*, were positioned in the newly defined *Mycoplasma bovis* cluster (Pettersson *et al.*, 2001). *M. phocirhinis* belonged to the *M. bovigenitalium* subcluster and *M. adleri* to the *M. felifaucium* subcluster. When the phylogeny of *M. adleri* was originally established in 1997, neither the *M. bovis* cluster nor any of its subclusters were yet characterised, for obvious reasons, and *M. adleri* was therefore said to belong to the *M. lipophilum* cluster. This reorganisation of the phylogeny of the mycoplasmas is a good example that a phylogeny is never the absolute truth but merely a suggestion on how to classify the species already analysed. The phylogeny will keep changing as new species are added.

Molecular diversity in some mycoplasmas of ruminants (III–IV)

M. capripneumoniae, the causative agent of contagious caprine pleuropneumonia CCPP (Hutcheon, 1881; Hutcheon, 1889; Thomas, 1893; MacOwan and Minette, 1976), represents a special case among the mycoplasmas in that it has an unusually high number of polymorphisms, and also a greater than average variability between strains of the species. When 20 strains of *M. capripneumoniae* were sequenced (Pettersson *et al.*, 1998), it was found that all strains had eleven polymorphisms in common and in addition to that, five to ten polymorphisms that were unique to one or a group of strains. Based on the detected polymorphisms, two lines of descent (line I and line II) could be traced and 11 polymorphism patterns (pm patterns) were identified. It was also found that these features might be used as epidemiological markers. Here the 16S rDNA sequences of both operons from 13 field strains of *M. capripneumoniae* from the

neighbouring African countries Kenya, Ethiopia and Tanzania were determined. Among the strains, four new and unique pm patterns reflecting the intraspecific variations within this species were found. Two of the patterns included sequence-length differences between the operons. One was characterised by a two-nucleotide insert and the other by a three-nucleotide deletion.

The 30 evolutionary events, which have been found in the 16S rRNA genes of the *M. capripneumoniae* strains, were used for the construction of an updated evolutionary tree based on comparative analysis. Insertion, deletion, or substitution of a nucleotide in a certain position was considered as one event. The insertion or deletion of two or three consecutive nucleotides was also regarded as one event. Compared to the corresponding tree constructed by Pettersson et al. (Pettersson *et al.*, 1998), this tree had six new branches, including the two obtained by the selection for streptomycin resistance. Note that the ancestor is putative, and that a strain which has only these 11 polymorphisms that are in common for all strains, has not yet been found.

The results from the investigation clearly show that the 16S rDNA polymorphisms can be used as epidemiological markers for CCPP in smaller geographical areas and also to study the molecular evolution within this species. It also poses the question why this species has such an excess of polymorphisms and why some of them seem stable, i.e. the 11 that are in common for all strains examined so far, while others are unique for one strain or a smaller group of strains. One possible explanation is the lack of RecA-mediated gene conversion, a mechanism that has been postulated to be involved in concerted evolution of genes in gene families (Abdulkarim and Hughes, 1996; Cilia *et al.*, 1996). It has not been investigated whether the RecA protein is absent or deficient in *M. capripneumoniae*, although it has been found in all mycoplasmas analysed so far (Razin *et al.*, 1998). The 11 polymorphisms present in all strains and evenly distributed over the 16S rRNA genes also imply that in contrast to *E. coli* (Martínez-Murcia et al., 1999), there are no segments in the 16S rRNA genes which are particularly prone to sequence homogenisation.

M. agalactiae and *M. bovis* are two closely related species (Askaa and Ernø, 1976; Edward and Freundt, 1973; Hale *et al.*, 1962) that cause diseases in goats and sheep (Lefévre, 1996) and cattle (Ross, 1993; Simecka *et al.*, 1992), respectively. Like *M. capripneumoniae* and many other mycoplasmas of ruminants, they have been found to have two rRNA operons (Christiansen and Ernø, 1990; Pettersson *et al.*, 1996a). The type strains of the two species display intraspecific variations in the form of seven and three polymorphisms, respectively, and some variation between strains have been reported (Subramaniam *et al.*, 1998). In order to determine the extent of the intraspecific variation and to find out if the polymorphisms could be used for the construction of an evolutionary tree and as epidemiological markers for mycoplasmoses caused by *M. agalactiae* or *M. bovis*, the 16S rRNA genes from of 16 and 7

strains, respectively, of the two species were sequenced and analysed. A relatively high degree of variability was discovered in both species, with regard to as well within as between strain variation. Despite this observation, however, no distinct evolutionary patterns could be recognised, which indicates a functional system for gene conversion in these species, in contrast to *M. capripneumoniae*. Some conclusions could still be drawn from the data, particularly for *M. agalactiae*, where all non-European strains shared three characteristic nucleotides, and European strains from the same or neighbouring countries were very similar.

Streptomycin resistance associated with mutations in the 16S rRNA gene?

Of the 7 *M. capripneumoniae* strains tested for streptomycin susceptibility only two were found to be resistant. Both of these strains harboured a T/C polymorphism in the position corresponding to position 912 of the 16S rRNA gene of *E. coli*, caused by a C→T mutation in the *rrnA* operon. This mutation has been shown to be associated with streptomycin resistance in several bacteria. The five other strains, in which this polymorphism was not found, were all streptomycin susceptible. The three *M. bovis* strains that were tested for streptomycin susceptibility displayed resistance to this antibiotic. The five *M. agalactiae* strains tested were found to be at least partially susceptible to streptomycin with inhibition zones of 10–16 mm. With three of the strains, however, some break through colonies were observed.

Interestingly, there is most probably a difference in how the two streptomycin resistant *M. capripneumoniae* strains acquired their resistance. One of the strains is the 102nd laboratory passage of a strain isolated in Gabés, which does not have the $899_{T/C}$ polymorphism (Pettersson et al., 1998) and is sensitive to streptomycin. The resistant strain must, therefore, have become streptomycin resistant in the laboratory, probably because it was cultured in the presence of streptomycin, but this has not been possible to confirm. The other resistant *M. capripneumoniae* strain, on the other, hand was isolated only a few years ago and has not been subjected to any passaging in the laboratory (except in connection with the isolation). Consequently, it must have acquired the streptomycin resistance in the field, probably when a diseased goat was treated with a mixture of penicillin and streptomycin, which is common in Ethiopia where it was isolated. This hypothesis could, however, not be confirmed in this particular case.

It is also interesting to note that one of the tested *M. agalactiae* strains, which seems to be an intermediate between *M. agalactiae* and *M. bovis* on the 16S rRNA level, showed less susceptibility to streptomycin than the other tested *M. agalactiae* strains, and it exhibited the smallest inhibition zone.

Selection for streptomycin resistance was successful for all four *M. capripneumonaie* strains included in the experiment. Two of the strains obtained resistance to streptomycin after 3 passages, and the other two strains became

streptomycin resistant after 5 passages. Sequencing of the region containing the position equivalent to *E. coli* 912 revealed that none of the clones had, in either of the operons, adapted the C→T transition mutation that has been reported to cause streptomycin resistance in for example *E. coli* and *M. tuberculosis*. However, one of the clones had a C→A transversion mutation in this position in *rrnA* and this substitution has been reported to cause streptomycin resistance in Tobacco chloroplasts (Cundliffe, 1990). None of the clones had the A→C transversion mutation in the position equivalent to position 523 in *E. coli* that has also been shown to cause streptomycin resistance.

This investigation concerning streptomycin resistance, although less extensive and sophisticated, clearly indicates that position 912 can be associated to streptomycin resistance in these mycoplasma species, but that it is obviously not the only mechanism involved or responsible. Alterations in protein S12 in the ribosomes is another plausible explanation, since it has been found in other bacteria (Finken *et al.*, 1993; Snyder and Champness, 1997).

Detection of *R. salmoninarum* (V)

Sequence analysis of the 16S rRNA genes from eight strains of *R. salmoninarum* isolated between 1973 and 1996 from various geographical localities in Canada, England, Iceland, Sweden and the USA, revealed that the sequences were identical but for one single position in one of the Canadian strains. This result is consistent with the report that *R. salmoninarum* contains two indistinguishable rRNA operons (Grayson *et al.*, 2000a). Two regions suitable as target sites for species-specific PCR primers were identified within the genes. In order to obtain maximum sensitivity without using nested PCR, primers for a fluorescent PCR were constructed for the detection of *R. salmoninarum*. Because of the many inhibiting substances present in kidney tissue, oligonucleotides for DNA isolation by sequence capture were also constructed, as well as a mimic molecule to be used as an internal control. The PCR was found to be specific for *R. salmoninarum* and allowed the detection of DNA equivalent to 1 to 10 genomes per reaction. The specificity of the PCR was tested with 11 bacterial strains representing the species most closely related to *R. salmoninarum* or species likely to be found in salmonid fish (see paper V). No PCR product was obtained with any DNA other than from *R. salmoninarum*. The DNA preparation by sequence capture offers a reliable method for analysing infected kidneys because it reduces the inhibiting substances from the tissue. The detection limit for sequence capture of kidney tissue samples that were negative in culture, spiked with *R. salmoninarum* DNA and followed by PCR, was 5 pg of DNA. The use of the mimic molecule in the system assures that no false negative results are reported. For detection with the automated sequencer, 100 mimic molecules per reaction were needed. The mimic did not have any negative effect on the amplification of the *R. salmoninarum* DNA.

The sequence capture fluorescent-PCR system reported here will in the future be tested on ovarian fluid and blood samples. If successful, this would provide a means to test broodstock without sacrificing the fish.

Concluding remarks and future perspectives

In this work, the 16S rRNA genes have been successfully used as tools to establish the phylogeny of seven mycoplasmas from goats and seals, to investigate the intraspecific variation among the strains of three different mycoplasma species, and to develop a PCR system for the detection of *R. salmoninarum*.

In a time when science is developing at an ever-increasing speed, the 16S rRNA has evolved from being a new and promising tool, via being a standardised method for detection and characterisation of bacteria, to a method by some scientists dismissed as ancient, limited and practically useless, which in my opinion is a foregone conclusion.

This work demonstrates that 16S rRNA still is a powerful tool in many areas of microbiology. It might not be the universal answer to all molecular questions and problems in microbiology, as was hoped in the beginning, but it is very useful in several different ways. The 16S rRNA is often a suitable target in detection by PCR for diagnostic purposes, and it can be used for studies of genetic diversity and sometimes even for molecular epidemiology. As for molecular phylogeny, the 16S rRNA might well be the most powerful tool, at least so far, simply because of its versatility.

The fact that the 16S rRNA is present in all self-replicating organisms coupled with the vast amount of sequence data already documented and available in public databases are two advantages that are hard to match. The use of one or a selection of other genes, which is argued by many scientists, will definitely be useful for defined groups of organisms especially when a high degree of resolution is needed. However, if the alternative genes are not present in all micro-organisms, a comparison outside this defined group will be impossible. Even if a group of new genes that are universally present and provide a higher degree of resolution can be found, it will take considerable time to gather as much information from these genes as the information available for 16S rRNA.

The increasing number of whole microbial genome sequences will certainly be a help in the search for new suitable phylogenetic and diagnostic markers. In the longer perspective, it might even be possible to compare complete genomes for phylogenetic purposes. However, there are quite a few problems to be solved first. An improved technique will be required for faster generation of whole genome data. A means of determining which segments to exclude because they are not relevant, as well as enhanced aligning methods that can cope with whole genomes, will also be needed.

It has been suggested that 16S rDNA sequence analysis should be mandatory in the description of new bacterial species (Fredericks and Relman, 1996; Johansson *et al.*, 2000; Rosselló-Mora and Amann, 2001). This would, in my opinion, be practical as well as helpful in general, and in the case of mollicutes of the genus *Mycoplasma* probably vital if we want new species of this genus to be described in the future.

All of the above reasons clearly indicates that the role of 16S rRNA in microbiology is far from insignificant, and that its importance will increase rather than decrease, in the near future.

References

Abdulkarim, F., and Hughes, D. 1996. Homologous recombination between the *tuf* genes of *Salmonella typhimurium*. *Journal of Molecular Biology 260*, 506-522.

Alexander, S. M., Grayson, T. H., Chambers, E. M., Cooper, L. F., Barker, G. A., and Gilpin, M. L. 2001. Variation in the spacer regions separating tRNA genes in *Renibacterium salmoninarum* distinguishes recent clinical isolates from the same location. *Journal of Clinical Microbiology 39*, 119-128.

Amann, R. I., Ludwig, W., and Schleifer, K.-H. 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological Reviews 59*, 143-169.

Arber, W. 2000. Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS Microbiology Reviews 24*, 1-7.

Armstrong, R. D., Martin, S. W., Evelyn, T. P., Hicks, B., Dorward, W. J., and Ferguson, H. W. 1989. A field evaluation of an indirect fluorescent antibody-based broodstock screening test used to control the vertical transmission of *Renibacterium salmoninarium* in Chinook salmon (*Oncorhynchus tshawytscha*). *Canadian Journal of Veterinary Research 53*, 385-389.

Askaa, G., and Ernø, H. 1976. Elevation of *Mycoplasma agalactiae* subsp. *bovis* to species rank: *Mycoplasma bovis* (Hale et al.) comb. nov. *International Journal of Systematic Bacteriology 26*, 323-325.

Austin, B., Bucke, D., Feist, S., and Rayment, J. 1985. A false positive reaction in the indirect fluorescent antibody test for *Renibacterium salmoninarum* with a "coryneform" organism. *Bulletin of European Association of Fish Pathology 5*, 8-9.

Baldauf, S. L., Palmer, J. D., and Doolittle, W. F. 1996. The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. *Proceedings of the National Academy of Sciences of the USA 93, 7749-7754.*

Ballagi-Pordany, A., and Belak, S. 1996. The use of mimics as internal standards to avoid false negatives in diagnostic PCR. *Molecular and Cellular Probes 10*, 159-164.

Belak, S., and Ballagi-Pordany, A. 1993. Experiences on the application of the polymerase chain reaction in a diagnostic laboratory. *Molecular and Cellular Probes 7*, 241-248.

Benediktsdóttir, E., Helgason, S., and Gudmundsdóttir, S. 1991. Incubation time for the cultivation of *Renibacterium salmoninarum* from Atlantic salmon, *Salmo salar* L., broodfish. *Journal of Fish Diseases 14*, 102.

Brown, J. R., and Doolittle, W. F. 1997. Archaea and the prokaryote-to-eukaryote transition. *Microbiology and Molecular Biology Reviews 61*, 456-502.

Bullock, G. L., and Stuckey, H. M. 1975. Fluorescent antibody identification and detection of the *Corynebacterium* causing kidney disease of salmonids. *Journal of the Fisheries and Research Board of Canada 32*, 2224-2227.

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

Böttger, E. C. 1989. Rapid determination of ribosomal RNA sequences by direct sequencing of enzymatically amplified DNA. *FEMS Microbiology Letters 65*, 171-176.

Christiansen, G., and Ernø, H. 1990. RFLP in rRNA genes of *Mycoplasma capricolum*, the caprine F38-like group and the bovine serogroup 7. *Zentralblatt für Bakteriologie, Supplement 20*, 479-488.

Cilia, V., Lafay, B., and Christen, R. 1996. Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. *Molecular Biology and Evolution 13*, 451-461.

Clayton, R. A., Sutton, G., Hinkle Jr, P. S., Bult, C., and Fields, C. 1995. Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. *International Journal of Systematic Bacteriology 45*, 595-599.

Colwell, R. R, Clayton, R. A., Ortiz-Conde, B. A., Jacobs, D., andRussek-Cohen, E. 1995. The microbial species concept and biodiversity. In: D. Allsopp, R. R. Colwell, D. L. Hawksworth (Eds.), *Microbial diversity and ecosystem function,* CAB International. Oxon, pp. 3-15.

Cundliffe, E. 1990. Recognition sites for antibiotics within rRNA. In: W. E. Hill, A. Dahlberg, R. A. Garret, P. B. Moore, D. Schlessinger, J. R. Warner (Eds.), *The ribosome: structure, function, and evolution,* American Society for Microbiology. Washington, D.C., pp. 479-490.

DaMassa, A. J., Tully, J. G., Rose, D. L., Pitcher, D., Leach, R. H., and Cottew, G. S. 1994. *Mycoplasma auris* sp. nov., *Mycoplasma cottewii* sp. nov., and *Mycoplasma yeatsii* sp. nov., new sterol-requiring mollicutes from the external ear canals of goats. *International Journal of Systematic Bacteriology 44*, 479-484.

DaMassa, A. J., Wakenell, P. S., and Brooks, D. L. 1992. Mycoplasmas of goats and sheep. *Journal of Veterinary Diagnostic Investigations 4*, 101-113.

Del Giudice, R. A., Rose, D. L., and Tully, J. G. 1995. *Mycoplasma adleri* sp. nov., an isolate from a goat. *International Journal of Systematic Bacteriology 45*, 29-31.

Edward, D. G. ff., and Freundt, E. A. 1973. Type strains of species of the order *Mycoplasmatales*, including the designation of neotypes for *Mycoplasma mycoides* subsp. *mycoides*, *Mycoplasma agalactiae* subsp. *agalactiae*, and *Mycoplasma arthritidis*. *International Journal of Systematic Bacteriology 23*, 55-61.

Elliott, D. G., and Barila, T. Y. 1987. Membrane filtration - fluorescent antibody staining procedure for detecting and quantifying *Renibacterium salmoninarum* in coelomic fluid of chinook salmon (*Oncorhynchus tshawytscha*). *Canadian Journal of Fisheries and Aquatic Sciences 44*, 206-210.

Englund, S., Ballagi-Pordany, A., Bölske, G., and Johansson, K.-E. 1999. Single PCR and nested PCR with a mimic molecule for detection of *Mycobacterium avium* subsp. *paratuberculosis. Diagnostic Microbiology and Infectious Disease 33*, 163-171.

Evenden, A. J., Grayson, T. H., Gilpin, M. L., and Munn, C. B. 1993. *Renibacterium salmoninarum* and bacterial kidney disease - the unfinished jigsaw. *Annual Review of Fish Diseases 3*, 87-104.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution 39*, 783-791.

Felsenstein, J., 1993. PHYLIP: Phylogeny inference package (version 3.51c). Seattle, Washington, University of Washington.

Finken, M., Kirschner, P., Meier, A., Wrede, A., and Böttger, E. C. 1993. Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. *Molecular Microbiology 9*, 1239-1246.

Fox, G. E., Pechman, K. R., and Woese, C. R. 1977. Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to procaryotic systematics. *International Journal of Systematic Bacteriology 27*, 57.

Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., and . 1995. The minimal gene complement of *Mycoplasma genitalium. Science 270*, 397-403.

Fredericks, D. N., and Relman, D. A. 1996. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. *Clinical Microbiology Reviews 9*, 18-33.

Fryer, J. L., and Lannan, C. N. 1993. The history and current status of *Renibacterium salmoninarum*, the causative agent of bacterial kidney disease in Pacific salmon. *Fisheries Research 17*, 15-33.

Giebel, J., Meier, J., Binder, A., Flossdorf, J., Poveda, J. B., Schmidt, R., and Kirchhoff, H. 1991. *Mycoplasma phocarhinis* sp. nov. and *Mycoplasma phocacerebrale* sp. nov., two new species from harbor seals (*Phoca vitulina* L.). *International Journal of Systematic Bacteriology 41*, 39-44.

Gray, M. W., Sankoff, D., and Cedergren, R. J. 1984. On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Research 12*, 5837-5852.

Grayson, T. H., Alexander, S. M., Cooper, L. F., and Gilpin, M. L. 2000a. *Renibacterium salmoninarum* isolates from different sources possess two highly conserved copies of the rRNA operon. *Antonie Van Leeuwenhoek 78*, 51-61.

Grayson, T. H., Atienzar, F. A., Alexander, S. M., Cooper, L. F., and Gilpin, M. L. 2000b. Molecular diversity of *Renibacterium salmoninarum* isolates determined by randomly amplified polymorphic DNA analysis. *Applied and Environmental Microbiology 66*, 435-438.

Grayson, T. H., Cooper, L. F., Atienzar, F. A., Knowles, M. R., and Gilpin, M. L. 1999. Molecular differentiation of *Renibacterium salmoninarum* isolates from worldwide locations. *Applied and Environmental Microbiology 65*, 961-968.

Gudmundsdóttir, S., Benediktsdóttir, E., and Helgason, S. 1993. Detection of *Renibacterium salmoninarum* in salmonid kidney samples: a comparison using double-sandwich ELISA and isolation on selective medium. *Journal of Fish Diseases 16*, 185-195.

Hale, H. H., Helmboldt, C. F., Plastridge, W. N., and Stula, E. F. 1962. Bovine mastitis caused by a *Mycoplasma* species. *Cornell Veterinary Medicine 52*, 582-591.

Heldtander, M., Pettersson, B., Tully, J. G., and Johansson, K.-E. 1998. Sequences of the 16S rRNA genes and phylogeny of the goat mycoplasmas *Mycoplasma adleri, Mycoplasma auris, Mycoplasma cottewii* and *Mycoplasma yeatsii. International Journal of Systematic Bacteriology 48*, 263-268.

Hugenholtz, P., Goebel, B. M., and Pace, N. R. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of Bacteriology 180*, 4765-4774.

Hultman, T., Bergh, S., Moks, T., and Uhlén, M. 1991. Bidirectional solid-phase sequencing of *in vitro*-amplified plasmid DNA. *BioTechniques 10*, 84-93.

Hultman, T., Ståhl, S., Hornes, E., and Uhlén, M. 1989. Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. *Nucleic Acids Research 17*, 4937-4946.

Hutcheon, D. 1881. Contagious pleuropneumoniae in Angora goats. *Veterinary Journal 13*, 171-180.

Hutcheon, D. 1889. Contagious pleuropneumonia in goats at Cape Colony, South Africa. *Veterinary Journal 29*, 399-404.

International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of Mollicutes. 1995. Revised minimum standards for description of new species of the class *Mollicutes* (division *Tenericutes*). *International Journal of Systematic Bacteriology 45*, 605-612.

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M., and Kersters, K. 1996. Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology 142*, 1881-1893.

Jansson, E., Hongslo, T., Höglund, J., and Ljungberg, O. 1996. Comparative evaluation of bacterial culture and two ELISA techniques for the detection of *Renibacterium salmoninarum* antigens in salmonid kidney tissues. *Diseases of Aquatic Organisms 27*, 197-206.

Johansson,K.-E. 1993. Detection and identification of mycoplasmas with diagnostic DNA probes complementary to ribosomal RNA. In: I. Kahane, A. Adoni (Eds.), Rapid Diagnosis of Mycoplasmas, Plenum Publishing Corporation, New York, pp. 139-154.

Johansson, K.-E., Heldtander, M. U. K, andPettersson, B. 1998. Characterization of mycoplasmas by PCR and sequence analysis with universal 16S rDNA primers. In: R. J. Miles, R. A. J. Nicholas (Eds.), *Methods in Molecular Biology; Mycoplasma protocols, vol. 104,* Humana Press. Totowa, NJ, pp. 145-165.

Johansson, K.-E., Pettersson, B., Heldtander, M., Bölske, G., and Tully, J. G. 2000. Phylogeny of the genus *Mycoplasma* based on 16S rDNA sequence analysis. *13th International Congress of IOM 13*, 66.

Jukes, T. H., and Cantor, C. R. 1969. Evolution of protein molecules. In: H. N. Munro (Eds.), *Mammalian Protein Metabolism,* Academic Press. New York, NY, pp. 21-132.

Kamla, V., Heinrich, B., and Hadding, U. 1996. Phylogeny based on elongation factor Tu reflects the phenotypic features of mycoplasmas better than that based on 16S rRNA. *Gene 171*, 83-87.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution 16*, 111-120.

Lan, R., and Reeves, P. R. 2000. Intraspecies variation in bacterial genomes: the need for a species genome concept. *Trends in Microbiology 8*, 396-401.

Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L., and Pace, N. R. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences of the USA 82*, 6955-6959.

Lefévre, P. C. 1996. Contagious agalactiae. In: M. Truszczynsky, J. Pearson, S. Edwards (Eds.), *The OIE manual of standards for diagnostic tests and vaccines,* Office International des Epizooties. Paris, pp. 363-368.

Li, W.-H., and Graur, D. 1991. Molecular phylogeny. *Fundamentals of molecular evolution,* Sinauer Associates, Inc. Sunderland, Massachusetts, pp. 99-135.

Ludwig, W., and Schleifer, K.-H. 1999. Phylogeny of *Bacteria* beyond the 16S rRNA standard. *ASM News 65*, 752-757.

Ludwig, W., and Schleifer, K. H. 1994. Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiology Reviews 15*, 155-173.

MacOwan, K. J. 1984. Role of mycoplasma strain F38 in contagious caprine pleuropneumonia. *Israel Journal of Medical Sciences 20*, 979-981.

MacOwan, K. J., and Minette, J. E. 1976. A mycoplasma from acute contagious caprine pleuropneumoniae in Kenya. *Tropical Animal Health and Production Journal 8*, 91-95.

Maiden, M. C., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., and Spratt, B. G. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the USA 95*, 3140-3145.

Maniloff, J. 1992. Phylogeny of Mycoplasmas. In: J. Maniloff, R. N. McElhaney, L. R. Finch, J. B. Baseman (Eds.), *Mycoplasmas: Molecular Pathology and Pathogenesis,* American Society for Microbiology. Washington, D.C., pp. 549-559.

Marcone, C., Neimark, H., Ragazzoni, A., Lauer, U., and Seemüller, E. 1999. Chromosome sizes of phytoplasmas composing major phylogenetic groups and subgroups. *Phytopathology 89*, 805-810.

Martínez-Murcia, A. J., Antón, A. I., and Rodríguez-Valera, F. 1999. Patterns of sequence variation in two regions of the 16S rRNA multigene family of *Escherichia coli*. *International Journal of Systematic Bacteriology 49*, 601-610.

Mattsson, J. G., 1993. *Stable RNA sequences as a tool for DNA probes and phylogenetic studies.* Swedish University of Agriculturual Sciences, Uppsala, Sweden.

Mayr, E. 1998. Two empires or three? *Proceedings of the National Academy of Sciences of the USA 95* , 9720-9723.

Montandon, P. E., Wagner, R., and Stutz, E. 1986. *E. coli* ribosomes with a C912 to U base change in the 16S rRNA are streptomycin resistant. *The EMBO Journal 5*, 3705-3708.

Mullis, K. B., and Faloona, F. A. 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods in Enzymology 155*, 335-350.

Murray, R. G., and Stackebrandt, E. 1995. Taxonomic note: implementation of the provisional status *Candidatus* for incompletely described procaryotes. *International Journal of Systematic Bacteriology 45*, 186-187.

Nicholas, R. A., Greig, A., Baker, S. E., Ayling, R. D., Heldtander, M., Johansson, K.-E., Houshaymi, B. M., and Miles, R. J. 1998. Isolation of *Mycoplasma fermentans* from a sheep. *Veterinary Record 142*, 220-221.

Nocard, E., and Roux, E. R. 1898. Le microbe de la péripneumonie. *Annales de l'Institut Pasteur 12*, 240-262.

Nygren, M., 2000. *Molecular diagnostics of infectious diseases*. Royal Institute of Technology, Stockholm, Sweden.

Olsen, G. J. 1988. Phylogenetic analysis using ribosomal RNA. *Methods in Enzymology 164*, 793-812.

Olsen, G. J., and Woese, C. R. 1993. Ribosomal RNA: a key to phylogeny. *The FASEB Journal 7*, 113-123.

Olsen, G. J., Woese, C. R., and Overbeek, R. 1994. Minireview: The winds of (evolutionary) change: breathing new life into microbiology. *Journal of Bacteriology 176*, 1-6.

Olsvik, O., Popovic, T., Skjerve, E., Cudjoe, K. S., Hornes, E., Ugelstad, J., and Uhlen, M. 1994. Magnetic separation techniques in diagnostic microbiology. *Clinical Microbiology Reviews 7*, 43-54.

Pascho, R. J., and Mulcahy, D. 1987. Enzyme-linked immunosorbent assay for a soluble antigen of *Renibacterium salmoninarum*, the causative agent of salmonid bacterial kidney disease. *Canadian Journal of Fisheries and Aquatic Sciences 44*, 183-191.

Pettersson, B., 1997. Direct solid-phase 16S rDNA sequencing: a tool in bacterial phylogeny. Royal Institute of Technology, Stockholm, Sweden.

Pettersson, B., Bölske, G., Thiaucourt, F., Uhlén, M., and Johansson, K.-E. 1998. Molecular evolution of *Mycoplasma capricolum* subsp. *capripneumoniae* strains, based on polymorphisms in the 16S rRNA genes. *Journal of Bacteriology 180*, 2350-2358.

Pettersson, B., Johansson, K.-E., and Uhlén, M. 1994a. Direct automated solid phase sequencing of an in vitro amplified segment of 16S rDNA from bacteria. *Proceedings of the 6th European Congress on Biotechnology*359-362.

Pettersson, B., Johansson, K.-E., and Uhlén, M. 1994b. Sequence analysis of 16S rRNA from mycoplasmas by direct solid-phase DNA sequencing. *Applied and Environmental Microbiology 60*, 2456-2461.

Pettersson, B., Leitner, T., Ronaghi, M., Bölske, G., Uhlén, M., and Johansson, K.-E. 1996b. Phylogeny of the *Mycoplasma mycoides* cluster as determined by sequence analysis of the 16S rRNA genes from the two rRNA operons. *Journal of Bacteriology 178*, 4131-4142.

Pettersson, B., Tully, J. G., Bölske, G., and Johansson, K.-E. 2000. Updated phylogenetic description of the *Mycoplasma hominis* cluster (Weisburg *et al.* 1989) based on 16S rDNA sequences. *International Journal of Systematic and Evolutionary Microbiology 50*, 291-301.

Pettersson, B., Tully, J. G., Bölske, G., and Johansson, K.-E. 2001. Re-evaluation of the classical *Mycoplasma lipophilum* cluster (Weisburg et al. 1989) and description of two new clusters in the hominis group based on 16S rDNA sequences. *International Journal of Systematic and Evolutionary Microbiology 51*, 633-643.

Pettersson, B., Uhlén, M., and Johansson, K.-E. 1996a. Phylogeny of some mycoplasmas from ruminants based on 16S rRNA sequences and definition of a new cluster within the hominis group. *International Journal of Systematic Bacteriology 46*, 1093-1098.

Prammananan, T., Sander, P., Brown, B. A., Frischkorn, K., Onyi, G. O., Zhang, Y., Böttger, E. C., and Wallace, Jr. R. J. 1998. A single 16S ribosomal RNA substitution is responsible for resistance to amikacin and other 2-deoxystreptamine aminoglycosides in *Mycobacterium abscessus* and *Mycobacterium chelonae. The Journal of Infectious Diseases 177*, 1573-1581.

Razin, S. 1992. Mycoplasma taxonomy and ecology. In: J. Maniloff, R. N. McElhaney, L. R. Finch, J. B. Baseman (Eds.), *Mycoplasmas: molecular pathology and pathogenesis,* American Society for Microbiology. Washington, D.C., pp. 3-22.

Razin, S. 1997. The minimal cellular genome of mycoplasma. *Indian Journal of Biochemistry & Biophysics 34*, 124-130.

Razin, S., Yogev, D., and Naot, Y. 1998. Molecular biology and pathogenicity of mycoplasmas. *Microbiology and Molecular Biology Reviews 62*, 1094-1156.

Ross, R. F. 1993. Mycoplasma - animal pathogens. In: I. Kahane, A. Adoni (Eds.), *Rapid diagnosis of mycoplasmas,* Plenum Press. New York, NY, pp. 69-109.

Rosselló-Mora, R., and Amann, R. 2001. The species concept for prokaryotes. *FEMS Microbiology Reviews 25*, 39-67.

Ruhnke, H. L., and Madoff, S. 1992. *Mycoplasma phocidae* sp. nov., isolated from harbor seals (*Phoca vitulina* L.). *International Journal of Systematic Bacteriology 42*, 211-214.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science 230*, 1350-1354.

Saitou, N., and Nei, M. 1987. The neighbor-joining method:a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution 4*, 406-425.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the USA 74*, 5463-5467.

Schloter, M., Lebuhn, M., Heulin, T., and Hartmann, A. 2000. Ecology and evolution of bacterial microdiversity. *FEMS Microbiology Reviews 24*, 647-660.

Schneider, B., Gibb, K. S., and Seemuller, E. 1997. Sequence and RFLP analysis of the elongation factor Tu gene used in differentiation and classification of phytoplasmas. *Microbiology 143*, 3381-3389.

Schwartz, D. C., and Cantor, C. R. 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell 37*, 67-75.

Seemüller, E., Marcone, C., Lauer, U., Ragazzoni, A., and Göschl, M. 1998. Current status of molecular classification of the phytoplasmas. *Journal of Plant Pathology 80*, 3-26.

Simecka, J. W., Davis, J. K., Davidson, M. K., Ross, S. E., Städtlander, C. T. K. H., andCassell, G. H. 1992. Mycoplasma Diseases of Animals. In: J. Maniloff, R. N. McElhaney, L. R. Finch, J. B. Baseman (Eds.), *Mycoplasmas: Molecular biology and pathogenesis,* American Society for Microbiology. Washington, D.C., pp. 391-415.

Smith, S., 1992. GDE: Genetic data environment, version 2.2. Millipore Imaging Systems. Ann Arbor, MI, USA,

Sneath, P. H. A. 1977. The maintenance of large numbers of strains of microorganisms, and the implications for culture collections. *FEMS Microbiology Letters 1*, 333-334.

Snyder, L., and Champness, W. 1997. Mutations in bacteria. *Molecular genetics of bacteria,* ASM Press. Washington, D.C., pp. 75-104.

Springer, B., Stockman, L., Teschner, K., Roberts, G. D., and Böttger, E. C. 1996. Two-laboratory collaborative study on identification of mycobacteria: molecular versus phenotypic methods. *Journal of Clinical Microbiology 34*, 296-303.

Subramaniam, S., Bergonier, D., Poumarat, F., Capaul, S., Schlatter, Y., Nicolet, J., and Frey, J. 1998. Species identification of *Mycoplasma bovis* and *Mycoplasma agalactiae* based on the *uvr*C genes by PCR. *Molecular and Cellular Probes 12*, 161-169.

Sulaiman, I. M., Morgan, U. M., Thompson, R. C., Lal, A. A., and Xiao, L. 2000. Phylogenetic relationships of *Cryptosporidium* parasites based on the 70- kilodalton heat shock protein (HSP70) gene. *Applied and Environmental Microbiology 66*, 2385-2391.

Swofford, D. L., Olsen, G. J., Waddell, P. J., andHilllis, D. M. 1996. Phylogenetic inference. In: D. M. Hillis, C. Moritz, B. K. Mable (Eds.), *Molecular systematics,* Sinauer Associates, Inc. Sunderland, MA, pp. 407-514.

Thomas, P. 1893. Rapport medical sur la maladie des chèvres désignée par les Arabes sous le nom de Bou-Frida qui a sévi épizootiquement pendant l'hiver 1872-73 dans le cercle de Djelfa. In: A. Jourdan (Eds.), *Rapport medical sur le Bou-Frida,* Gouvernement Général civil de l'Algérie. Algiers, Algeria, pp. 1-35.

Tully, J. G., Bové, J. M., Laigret, F., and Whitcomb, R. F. 1993. Revised taxonomy of the class *Mollicutes*: Proposed elevation of a monophyletic cluster of arthropod-associated mollicutes to ordinal rank (*Entomoplasmatales* ord. nov.), with provision for familial rank to separate species with nonhelical morphology (*Entomoplasmataceae* fam. nov.) from helical species (*Spiroplasmataceae*), and emended descriptions of the order *Mycoplasmatales*, family *Mycoplasmataceae. International Journal of Systematic Bacteriology 43*, 378-385.

van Belkum, A. 2000. Are diagnostic microbiology tests reproducible? *Journal of Microbiological Methods 43*, 65-66.

Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K., and Swings, J. 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological Reviews 60*, 407-438.

Versalovic, J., Koeuth, T., and Lupski, J. R. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Research 19*, 6823-6831.

Wang, Y., and Zhang, Z. 2000. Comparative sequence analyses reveal frequent occurrence of short segments containing an abnormally high number of non-random base variations in bacterial rRNA genes. *Microbiology 146*, 2845-2854.

Ward, D. M. 1998. A natural species concept for prokaryotes. *Current Opinion in Microbiology 1*, 271-277.

Weisburg, W. G., Tully, J. G., Rose, D. L., Petzel, J. P., Oyaizu, H., Yang, D., Mandelco, L., Sechrest, J., Lawrence, T. G., Van Etten, J., Maniloff, J., and Woese, C. R. 1989. A phylogenetic analysis of the mycoplasmas: basis for their classification. *Journal of Bacteriology 171*, 6455-6467.

Welsh, J., and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research 18*, 7213-7218.

Whitman, W. B., Coleman, D. C., and Wiebe, W. J. 1998. Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences of the USA 95*, 6578-6583.

Wiener, P., Egan, S., Huddleston, A. S., and Wellington, E. M. 1998. Evidence for transfer of antibiotic-resistance genes in soil populations of streptomycetes. *Molecular Ecology 7*, 1205-1216.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research 18*, 6531-6535.

Woese, C. 1998a. The universal ancestor. *Proceedings of the National Academy of Sciences of the USA 95*, 6854-6859.

Woese, C. R. 1987. Bacterial evolution. *Microbiological Reviews 51*, 221-271.

Woese, C. R. 1998b. Default taxonomy: Ernst Mayr's view of the microbial world. *Proceedings of the National Academy of Sciences of the USA 95*, 11043-11046.

Woese, C. R. 2000. Interpreting the universal phylogenetic tree. *Proceedings of the National Academy of Sciences of the USA 97*, 8392-8396.

Woese, C. R., and Fox, G. E. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences of the USA 74*, 5088-5090.

Woese, C. R., Gutell, R., Gupta, R., and Noller, H. F. 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. *Microbiological Reviews 47*, 621-669.

Woese, C. R., Kandler, O., and Wheelis, M. L. 1990. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proceedings of the National Academy of Sciences of the USA 87*, 4576-4579.

Woese, C. R., Maniloff, J., and Zablen, L. B. 1980. Phylogenetic analysis of the mycoplasmas. *Proceedings of the National Academy of Sciences of the USA 77*, 494-498.

Woese, C. R., Stackebrandt, E., and Ludwig, W. 1985. What are mycoplasmas: The relationship of tempo and mode in bacterial evolution. *Journal of Molecular Evolution 21*, 305-316.

Zuckerkandl, E., and Pauling, L. 1965. Molecules as documents of evolutionary history. *Journal of Theoretical Biology 8*, 357-366.

Acknowledgements

This work was performed at the R&D section of the department of Bacteriology at the National Veterinary Institute. I would like to thank everyone who has aided, assisted, helped, encouraged and supported me during these years, especially:

Karl-Erik Johansson, Kaggen, my supervisor, for introducing me to the mycoplasmas, for constant enthusiasm and encouragement during all the projects, and of course for trying to keep my writing scientific (how do you like this sentence?;-)).

My associate supervisors, Andras Ballagi-Pordány, Bertil Pettersson and Göran Bölske for sharing with me the expertise they have in their respective fields and for knowledgeable and prompt help whenever I asked for it.

Eva Jansson, Hezron Wesonga and Joseph Tully, my co-authors, for interesting and fruitful collaborations.

The rest of the Bakt/FoU-group: Anja, my roommate and scientific travel companion, for sharing not only the mycoplasmas with me, but also the joy of yellow bikes, coca cola and "Sofia dansar go-go" in double stereo. Marianne for all the laughs, skilful technical assistance and for keeping track of everything and everyone—what would we do without you? Stina, my colleague and personal wailing wall, for all our discussions about science, project management and life in general, and for always cheering me up and telling me not to quit.

Everyone else using the FoU-labs, at present or in the past, most of all: Boel for being a good friend as well as colleague. Magda for constantly good spirits and for taking over the lawsonias—what a relief! Anna and Christian for knowing the things I don't and asking me questions that I sometimes actually can answer. Märit and Puck for all the good times with our friend ALF and for reminding me that there are other bacteria than mycoplasmas in the world.

The people in the mycoplasma-lab, Virginia, Stefan, Agneta and Karin, for growing an almost infinite number of mycoplasma strains for me, usually quicker than is really possible with these fastidious creatures.

The whole SmgB-group for nice seminars with interesting scientific discussions as well as nourishing chitchat. A special thanks to Jens, for the time we served together in "SACO-styrelsen" and for all the encouraging comments and anecdotes during the last weeks of writing.

All the staff at the department of bacteriology, former and present, for help in various ways and for making coffee and lunch breaks much more friendly.

All friends from the "real world", Hobba, Milla, Niklas, Pernilla, Nicklas, Caroline, Kristina, Ann-Sofie, Henning, Anna, Cilla, Annica, Tonka, Kerstin, Ulla, Catharina, Malin, Märy, Charlotta, and everyone that I forgot to mention, for friendship, laughs and shared worries, and for believing more in me than I do myself.

Wolfgang Amadeus Mozart for composing the most inspiring work I know, Requiem, which helped me write a good part of this thesis.

My family in law, including all aunts, uncles and cousins, for making me part of your family without a moment of hesitation.

My parents, for supporting me in whatever I do and for teaching me not to take life or anything else more seriously than necessary.

My sister, Johanna, for being my greatest supporter and ally as well as best friend. I'm sure that one day we'll find a way to move Uppsala and Vienna closer together!

Kalle, for everything great and small and most of all for loving me even on really bad days. Married or not, this is your thesis as much as mine.

Acta Universitatis Agriculturae Sueciae
presents doctoral theses from the Faculty of Agriculture, Landscape Architecture and Horticulture, the Faculty of Forestry and the Faculty of Veterinary Medicine at the Swedish University of Agricultural Sciences. Each faculty has its own subseries, Agraria, Silvestria and Veterinaria, respectively, with separate numbering.

ISSN 1401-6257
ISBN 91-576-5925-7