Chapter 2

Phylogeny of *Pasteurellaceae*

Bożena M. Korczak and Peter Kuhnert

Institute of Veterinary Bacteriology, Vetsuisse Faculty University of Bern, Länggass-Str. 122, 3001 Bern, Switzerland

Corresponding author: bozena.korczak@vbi.unibe.ch

Running title: Phylogeny of *Pasteurellaceae*

Abstract

This chapter provides an overview on the DNA based phylogeny of the family *Pasteurellaceae* and the genetic relatedness between taxa taking into account the various gene targets and approaches applied in the literature. The classical 16S rRNA gene based phylogeny as well as phylogenies based on house-keeping genes are described. Moreover, strength and weakness of the different trees and their topology are discussed based on the phylogenetic groups resolved. The data should help to get a clearer picture on the recent, current and future classification and also provide information to genetic characterization of members of the family. The history of phylogeny applied to the family as well as the phylogenetic history of the family is thereby presented. In this way it is the story of the search for the optimal phylogenetic marker without giving a final conclusive suggestion but it is also a resource for choosing the appropriate gene target(s) for people investigating the phylogeny of groups of *Pasteurellaceae*.

Introduction

Phylogeny as a tool for delineating genealogic association or evolution has become increasingly useful for investigating the relationship of microorganisms. In bacteriology it is an indispensable approach for proper classification of taxa as well as for identification and diagnosis. DNA-DNA hybridization studies as a rough estimate of genetic similarities between species have been performed for a long time with *Pasteurellaceae* and helped establishing the family (Mutters et al., 1985; Pohl, 1981). However, the first report investigating phylogeny within *Pasteurellaceae* based on nucleic acid sequences dates back 20 years. At that time direct sequencing of RNA was easier than sequencing DNA and the template was in this case already abundantly present, namely the 16S ribosomal RNA (16S rRNA). This molecule was introduced to phylogeny by the well known work of Woese (1987) allowing the delineation of phylogenetic relationships by sequence comparison of 16S rRNA or its genes, respectively. In the first publication applying this promising new approach, Chuba et al. (1988) used only seven species in their study. However, they included the type species of the three genera ***H****aemophilus*, ***A****ctinobacillus* and ***P****asteurella* (usually referred to as "**HAP**") which for a long time were the only recognized genera within *Pasteurellaceae*.

Moreover, by inclusion of at that time [*Actinobacillus*] *actinomycetemcomitans* (now *Aggregatibacter actinomycetemcomitans*) and [*Haemophilus*] *aphrophilus* (now *Aggregatibacter aphrophilus*) actually members of a fourth genus were also investigated. Furthermore, two additional representatives from the genus *Actinobacillus*, namely *Actinobacillus equuli* and *Actinobacillus hominis* were analyzed. Results from this limited study, not only due to the small number of strains but also based on the short sequences of 450 base pairs (bp) generated, already indicated that [*A.*] *actinomycetemcomitans* was wrongly attributed to this genus. The three other Actinobacilli were found closely related and *A. lignieresii*, *A. equuli* and *A. hominis* are nowadays recognized as true *Actinobacillus* species (Kuhnert et al., 2007). Later rRNA-DNA hybridization showed the heterogeneity of the family with at least seven rRNA branches (De Ley et al., 1990). This problematic within the "HAP" group was then shown very impressively by the two publications of Dewhirst et al. (1992; 1993) analyzing a total of 70 strains of *Pasteurellaceae* based on 16S rRNA sequences thereby expanding the initial phylogenetic study of Chuba et al. (1988) significantly. The branching was extremely complex and there was no close relation between branching and taxonomic position of strains, what just represented the many misclassified species. Nevertheless, it took some time until this finding became common knowledge and revision of the family became accepted. For some years it seemed to be a taboo to break the old "HAPpy" family apart and suggest new genera. With the exception of the species *Lonepinella koalarum* (Osawa et al., 1995) representing a new genus (which however was hardly noticed) we almost missed the opportunity to do so in the last millennium. But then the paper of Angen and coworkers appeared in 1999 suggesting the genus *Mannheimia* including the species *M. haemolytica*, *M. glucosida*, *M. varigena*, *M. granulomatis* and *M. ruminalis* (Angen et al., 1999). Fortunately enough, this investigation was not only based on phylogeny using 16S rRNA gene but was a thorough poylphasic study which leaved no doubt for the necessity to create a new genus, even though its acceptance in the scientific community was not unanimous.

Since then a cascade of changes including reclassification as well as description of many new taxa within *Pasteurellaceae* based on 16S rRNA gene sequence phylogeny but also in respect to phenotypic characterisation has begun. In addition to the 16S rRNA gene a few house-keeping genes have been applied for phylogenetic purposes. The family has undergone important changes in particular at the genus level and in a very short time the number of genera has increased to nowadays thirteen. *Phocoenobacter uteri* representing also a new

genus came to the scene in 2000 (Foster et al., 2000). Then in 2003 strains previously referred to as [*Haemophilus*] *somnus*, [*Haemophilus*] *agni* or '*Histophilus ovis'* were analysed by 16S rRNA and *rpoB* gene sequencing. This together with earlier investigations by DNA-DNA reassociation showed that all these organisms are very closely related to each other and were classified within the same species *Histophilus somni* currently being the only member of the genus (Angen et al., 2003). In the same year investigation of genetic relation of bacteria previously classified as the avian [*Pasteurella*] *haemolytica*-'*Actinobacillus salpingitidis*'-[*Pasteurella*] *anatis* complex ended in the description of the new genus *Gallibacterium* with *G. anatis* and at least one genomospecies (Christensen et al., 2003a). In 2004, using 16S rRNA gene sequence comparison of previously phenotypically characterized strains (Bisgaard et al., 1999) resulted in assignment of the new genus *Volucribacter* which comprises the two species *V. psittacicida* and *V. amazonae* (Christensen et al., 2004a). Later that year organisms from airway infection of horses were investigated in a complex phylogenetic analysis, including 16S rRNA, *rpoB*, and *infB* gene sequences which in a polyphasic approach provided powerful evidence for the new genus *Nicoletella* currently consisting of only one species *N. semolina* (Kuhnert et al., 2004). Phylogeny based on 16S rRNA gene sequences was also useful for description of the genus *Avibacterium* in 2005 comprising *Av. paragallinarum*, *Av. gallinarum*, *Av. avium* and *Av. volantium* (Blackall et al., 2005). In 2006 accurate investigation of genetic relatedness between [*Actinobacillus*] *actinomycetemcomitants*, [*Haemophilus*] *aphrophilus*, [*Haemophilus*] *paraphrophilus* and [*Haemophilus*] *segnis* and a few other members of the family *Pasteurellaceae* provided arguments that these species should be placed to the new genus *Aggregatibacter* and they have been renamed to *Agg. actinomycetemcomitants*, *Agg. aphrophilus* (combined from former species [*H.*] *aphrophilus* and [*H.*] *paraphrophilus*) and *Agg. segnis* (Norskov-Lauritsen and Kilian, 2006). Finally, the so far last genus until the end of 2007 was *Bibersteinia* resulting from the reclassification of [*Pasteurella*] *trehalosi* which was supported by phylogenetic analysis of 16S rRNA gene sequences (Blackall et al., 2007).

Bioinformatic tools for phylogeny

There is one topic worth mentioning before we move on to get a closer look at the phylogeny of the *Pasteurellaceae*, which is: how can phylogenetic relationship properly be

reconstructed? In the style of the quote attributed to Winston Churchill that "*the only statistics you can trust are those you falsified yourself*" one is tempted to say "*the only phylogenetic tree you can trust is the one you constructed yourself*". In a simplified view phylogenetic reconstruction is a mathematical approach using sophisticated algorithms to reflect as much as possible biological principles, in this case mutational changes representing the time span between separation of two or more organisms. In as how far these models can in fact represent the historical flow and evolution as well as account for selection changes (if any) and other influences on mutational rate is of course a case for dispute. This is certainly one reason why there are various approaches and algorithms nowadays available and phylogeny has developed into a very complex if not complicated scientific field. The authors really feel out of proper competence to claim, that they have accounted for all the unknowns and that they have chosen the most appropriate approach for the phylogenetic reconstructions shown in this chapter. Rather, we try to give a simple representation of the phylogenetic relationships within the family without going into too much details e.g. about branching, distances and evolutionary aspects. The aim is more to show the big context. Own experience learnt us that the choice of algorithms, the numbers of species included as of course the quality of the "raw data" i.e. the DNA sequence (including the question how diverse copies of 16S rRNA genes within a species should be handled) can all influence the tree structure. This should in no way be a pretext or argument that more appropriate analyses can not be done, and in some cases indeed very nice conclusions can be drawn from intensive phylogenetic comparisons, but this was not our intension for this chapter.

Basically, there are three particular families of methods used for phylogenetic reconstruction which are distance matrix, maximum parsimony and maximum likelihood. Maximum parsimony is based on the minimum number of base changes which are needed to build the tree. Maximum likelihood considers various parameters which might influence phylogenetic reconstruction, as e.g. identical sites between sequences based on multiple mutations or transition/transversion bias. This is less well handled by the distance based trees but corrections like Jukes-Cantor can still to some extent count for this. A tree is then built from the distance matrix by e.g. the most common used method based on neighbour-joining (NJ) (Saitou and Nei, 1987). Maximum likelihood is considered the most appropriate approach, however parameters have to be set accordingly and the calculation takes a lot of computational resources especially with large trees.

Bootstrap analysis should be included with all methods as measures of robustness of nodes. Ideally, trees are built based on different methods and consistency between results indicates the trustfulness of the phylogenetic relationships.

In our hands Bionumerics from Applied Maths NV (Sint-Martens-Latem, Belgium) proved to be very useful for phylogenetic analyses since it is a comprehensive database program including several modules. Thereby, not only DNA sequences can be handled but also phenotypic traits or gel pictures (fragment analysis) as well as other markers. Moreover it includes many algorithms for phylogenetic analyses and is therefore very flexible. However, the programme is rather expensive and for that reason not always affordable.
Alternatively, there are several programs publicly available for phylogenetic analysis, as e.g. MEGA4 (Tamura et al., 2007), ClustalX (Thompson et al., 1997) or PHYLIP. Comprehensive lists of available software can be found on the web, as e.g. evolution.genetics.washington.edu/phylip/software.html.

Phylogeny based on 16S rRNA genes

The 16S rRNA gene has massively improved our understanding of microbial evolution. Since the 16S rRNA itself has a function in the ribosome and therefore in translation from DNA to protein it is vital and present in all so far known prokaryotes. The 16S rRNA is a linear molecule which however shows base pairing with itself resulting in linear, double-strand and hairpin regions (Gutell et al., 1994). The structural and functional constraints results in highly conserved as well as more variable regions within the gene. There are several advantages to that and to the 16S rRNA gene in general: (i) the information gained from the sequence allows delineating distant relations from more conserved regions as well as closer ones from the variable parts of the gene (ii) it allows to define universal primers from conserved regions for amplification and sequencing of the gene (iii) the size of about 1.5 kb is reasonable for data generation and handling and (iv) a large amount of data for 16S rRNA gene from a broad variety is available for phylogenetic comparison of numerous bacterial species from a large number of families. The 16S rRNA gene based phylogenetic analysis developed to a gold standard for genetic comparison of bacteria and its analysis has become a prerequisite for the description of new species and is therefore one of the pillars in taxonomy (Stackebrandt et al.,

2002). One of the conclusions from 16S rRNA gene sequence comparisons is that strains which show more than 3% divergence in their 16S rRNA gene sequence should be regarded separate species as determined by comparison to the reference method DNA-DNA hybridisation, whereas strains that are less than 3% divergent may or may not be members of different species and DNA-DNA hybridisation or equivalent methods have to be applied (Stackebrandt and Goebel, 1994). Therefore, 3% 16S rRNA divergence is applied as a conservative criterion for demarcating species.

Currently the 16S rRNA gene sequence-based phylogeny is recognised as most useful for investigation of relatedness also within the family *Pasteurellaceae*. Since the first descriptions of 16S rRNA sequences from *Pasteurellaceae* (Chuba et al., 1988) many increasingly large 16S rRNA gene based trees have been published (Dewhirst et al., 1992; Dewhirst et al., 1993; Olsen et al., 2005; Christensen et al., 2004b; Korczak et al., 2004; Angen et al., 2003). The 16S rRNA genes of all named and several unnamed species-like taxa have been determined and are available on e.g. GenBank (Benson et al., 2007). One important issue having some influence on phylogenetic analysis is the copy number of rRNA operons present in bacterial genomes. Exploring the genome sequences of some *Pasteurellaceae* results in three copies of 16S rRNA gene for *A. pleuropneumoniae* L20 (NC_009053), five for *Hist. somni* 129Pt (NC_008309), and six for *H. influenzae* KW20 Rd (NC_000907), [*Haemophilus*] *ducreyi* 35000HP (NC_002940), '*M. succiniciproducens*' MBEL55E (NC_006300) and *P. multocida* Pm70 (NC_002663). The degree of allelic heterogeneity varies between taxa and in some cases 16S rRNA gene sequences could differ up to several percent. It has been shown that at least for gamma-proteobacteria, the selection of a single 16S rRNA allele is sufficient for inferring phylogenies, and that orthologous alleles from the same strain would have generated similar phylogenetic trees, however, presence of heterogeneous alleles which appear in the sequence as ambiguous bases makes the analysis difficult (Olivier et al., 2005). This has especially become problematic since direct sequencing of PCR products amplified from genomic DNA became routinely used. Own experience using our protocol (Kuhnert et al., 2002) have shown that in certain cases the number of ambiguous positions is very high. In the case of [*Haemophilus*] *felis* we were even not able to properly determine the 16S rRNA gene sequence in our lab, probably due to an experimental artefact, since the 16S rRNA gene sequence of the type strain is available in GenBank. The ambiguous positions mounted up to over 3% within the type strain of [*H.*] *felis* (P. Kuhnert, unpublished data). A similar situation was observed with the 16S rRNA gene sequence of the type strain of *L. koalarum* in our

hands. Cloning of the corresponding 16S rRNA genes might be the only way to definitely answer the question if this is a PCR artefact or really representing the intra-strain variability (Acinas et al., 2004). Divergence of rRNA operons has been found more often in *Archaea* but also in *Bacteria* and its implication for phylogeny has been discussed (Acinas et al., 2005; Olivier et al., 2005). The problem within the family *Pasteurellaceae* in this context should not be overestimated but it is worth pointing out in the view of new genera and species that certainly will be described in the future.

Taking the above mentioned aspects into account we constructed an updated 16S rRNA gene based phylogenetic tree of the *Pasteurellaceae* (Fig. 1). It is based on a Jukes-Cantor corrected distance matrix and on Neighbour-Joining tree formation (Saitou and Nei, 1987) using sequences of type strains representing nearly all currently described species and a few candidates for new species as well as some genome sequences. The tree topology and branching shows that phylogenetically many members of *Pasteurellaceae* are unrelated to the taxa to which they have been previously classified mainly based on their phenotypic features. The groups and monotypic taxa correspond to ones formerly defined by others (Christensen et al., 2004b; Korczak et al., 2004; Olsen et al., 2005). The tree shows about 30 groups, in general representing the current and putative new genera, including monophyletic branches. These groups and their members as well as the monotypic taxa are also listed in Table 4 of Chapter 1. Generally the tree shows clearly that there are still many misclassified species which need a final proper classification.

Similarities of 16S rRNA gene sequences within the family are more than 89% and within a genus it is generally more than 93%. However, these values are an approximation and can not be taken for absolute. The may also change with more taxa of the family will be described.

It has been shown that in some cases the 16S rRNA-based phylogeny does not support results provided by DNA-DNA hybridisation studies. Based on the study of Dewhirst (1992) *Actinobacillus capsulatus* was suggested to be excluded from the genus *Actinobacillus*, although DNA-DNA reassociation values showed that this taxon was correctly named (Escande et al., 1984; Mutters et al., 1989). Moreover, phylogenetic analysis based on *rpoB*, *infB* and *recN* unambiguously clarified that this species belongs to the genus *Actinobacillus* (Kuhnert et al., 2007). This finding indicates that even such stable markers and phylogenetic gold standards like the 16S rRNA gene might be subjected to lateral transfer (Yap et al., 1999). Furthermore, it shows that not only the 16S rRNA gene in some cases might be too conserved and therefore unable to differentiate between some closely related but ecological

distinct groups of bacteria (Fox et al., 1992), but also that it has to be complemented by other phylogenetic approaches based on different genes. A few such "phylogenes", i.e. genes suitable for phylogenetic analyses, have been applied for the *Pasteurellaceae* or at least for some groups of the family.

Phylogeny based on selected house-keeping genes

For a number of reasons house-keeping genes proved to be very promising targets for clarifying phylogenetic relationships within the various bacteria. These genes encode proteins which are essential for the cell metabolism and physiology, and therefore should be present in all species and comparable between different organisms. This kind of genes normally evolve at a higher rate than the 16S rRNA gene thereby providing a better resolution for differentiation of closely related species (Lawrence et al., 1991). It has been a big challenge to find powerful approaches being suited for this purpose since not every gene is of the equal value in investigation of phylogenetic relationships. The ideal phylogene has to fulfil certain criteria: (i) changes in the sequence should occur randomly, (ii) rates of changes should be constant in order to properly represent evolutionary distances, and (iii) the size of the sequence has to be large enough to provide sufficient information (Woese, 1987). Whole genome comparisons showed that there are 30-40 candidate genes for that purpose (Ludwig and Schleifer, 1999; Zeigler, 2003). Several such some house-keeping genes have been applied to investigate the phylogeny within certain genera or even the entire family of *Pasteurellaceae*.

Using amino acid sequences instead of nucleic acid sequences could be an advantage for comparison of far related groups, as e.g. loosely related genera within the family. The nucleic acid level includes all sequence variations which could be too high to properly resolve these phylogenies whereas the amino acid sequence level reduces diversity by neglecting silent substitutions and third codon positions. In our analyses we concentrated on the DNA sequence-based phylogeny of *Pasteurellaceae*. Nevertheless, the phylogeny using amino acid sequences could provide more useful information for certain targets.

Phylogeny based on infB

The single copy *infB* gene encodes the translation initiation factor 2 (IF2) that is involved in the initiation of protein synthesis in prokaryotes. The structure of the protein IF2 is highly conserved within several bacteria particularly in the domains responsible for binding and hydrolysis of GTP. However, this part contains also variable segments that are suitable for differentiation between groups of bacteria (Steffensen et al., 1997; Hedegaard et al., 1999). The length of the gene can vary between species and within the *Pasteurellaceae* it ranges between 2463 bp for *H. influenzae* KW20 Rd to 2526 bp for *A. pleuropneumoniae* L20.

Partial *infB* sequences have been used to investigate the phylogeny of the genus *Haemophilus* (Hedegaard et al., 2001). Fragments of approximately 650 nucleotides were amplified from strains representing at the time the genus and 453 bp long sequences covering the mentioned GTP-binding domain of IF2 were unambiguously determined and used for analysis of relationships within the group (Table 1). In an another study Christensen et al. (2004b) applied the same sequencing approach but used the deduced partial protein sequences for comparative phylogenetic analysis of 36 taxa of *Pasteurellaceae* representing nine genera. In order to obtain more information on the *infB* gene we developed a general PCR based sequencing approach for the family covering approximately 1360 bp of the gene. This was used in the description of the genus *Nicoletella* where it was phylogenetically compared to the type strains of nine established genera as well as type strains of species of the genus *Actinobacillus* (Kuhnert et al., 2004). The same approach has later been used for deeper understanding of genetic relationship within strains of the phenotypically closely related and therefore often misidentified [*Actinobacillus*] *rossi* and [*Pasteurella*] *aerogenes* (Mayor et al., 2006).

A comprehensive tree derived from the 1360 bp long sequences of 50 species representing all thirteen currently named genera as well as candidates for potential new taxa within the family has been constructed for illustration of the phylogeny provided by *infB* (Fig. 2). The tree is in good agreement with the 16S rRNA gene-based tree and a good separation of genera and groups was obtained with the *infB* phylogeny. Similarities of partial *infB* sequences within the current family were more than 73% and in general more than 83% within the same genus. However, these values are an approximation and might also change while new taxa will be described.

Phylogeny based on rpoB

The *rpoB* gene fulfils the main expectations for being a good phylogene. It encodes the β-subunit of the DNA-dependent RNA polymerase, the highly conserved molecule that catalyzes transcription of DNA into RNA. It is found among the *Bacteria* (Ovchinnikov et al., 1981) as well as among the *Archeae* (Klenk and Zillig, 1994). So far, it has been found in only one copy in all investigated bacterial genomes and the size within members of the family *Pasteurellaceae* ranges from 4029 bp for *A. pleuropneumoniae* L20, [*H.*] *ducreyi* 35000HP, *Hist. somni* 129Pt, *'M. succiniciproducens'* MBEL55E and *P. multocida* Pm70 to 4032 bp for *H. influenzae* KW20 Rd. Similar to the 16S rRNA gene the *rpoB* sequence contains conserved and variable regions (Lisitsyn et al., 1988; Palenik, 1992). This enables to select PCR and sequencing primers in conserved regions and interspacing variable regions provide very useful information about diversity of investigated strains at DNA level. The *rpoB* has been successfully applied as a molecular chronometer for investigation of phylogenetic relations within different groups of organisms (Klenk and Zillig, 1994; Nolte, 1995; Taillardat-Bisch et al., 2003; Rowland et al., 1993; Mollet et al., 1997).

A comprehensive analysis of the use of *rpoB* for the phylogeny of *Pasteurellaceae* was published by Korczak et al. (2004). The universal approach for the family allowed sequence determination of about 560 bp from all 72 included strains by use of a single pair of primers (Table 1). The fragment covers the most variable polypeptide region 4 of the DNA-dependent β-subunit RNA polymerase in *Escherichia coli* and despite its limited size proved to be highly informative for phylogenetic analysis of *Pasteurellaceae*. Figure 3 gives an updated *rpoB*-based tree. In general *rpoB* provided higher resolution than 16S rRNA gene, however, it was also not sufficient enough for differentiation of very closely related species and subspecies. Currently estimated similarities of partial *rpoB* sequences within the family were more than 77% and generally more than 87% within a genus. Again, these values are an approximation and may also change with more taxa of the family being described.

There was a good correlation between topologies in 16S rRNA and *rpoB*-based tree even though some discrepancies were observed. Phylogeny based on *rpoB* also correlated well with whole genome DNA-DNA hybridization results and appeared to be valuable for dissection of genera within the family. Partial *rpoB* gene sequences were further used for clarification of the taxonomic position of strains identified as *Hist. somni* previously known as '*H. somnus'*, '*H. agni'*, or '*Hist. ovis'* (Angen et al., 2003), including eight *Hist. somni* strains and the type

strains of at this time existing genera. A sequence analysis of the *rpoB* fragments showed that strains identified as *Hist. somni* were clearly separated from the other genera within the family and should be referred to a new genus. It has also been shown that phylogeny obtained from parsimony analysis of the deduced *rpoB* partial protein sequence of representative taxa of the *Pasteurellaceae* allows clear discrimination of different genera in the family (Christensen et al., 2004b).

Phylogeny based on sodA

The gene encoding manganese-dependent superoxide dismutase (*sodA*) is a further target that has been used for a better understanding of the phylogeny of *Pasteurellaceae*. The gene is rather small in size, ranging from 630 to 666 bp in investigated members of *Pasteurellaceae*. These differences in sequence length are the result of the presence of extra nucleotides corresponding to supplemental blocks of codons in the deduced amino acid sequences. Interestingly, additional amino acids define signatures which are specific for species of certain genera. The first study on *sodA* in particular comprised species belonging to the genus *Pasteurella* and related species (Gautier et al., 2005). One year later this was complemented by a study giving a more detailed phylogenetic insight into the genus *Haemophilus* (Cattoir et al., 2006). In these studies *sodA* fragments of 449 to 473 bp were generated from different species using conserved primers for PCR and sequencing (Poyart et al., 1998); Table 1). In the analysis of the genera *Pasteurella*, *Gallibacterium* and *Mannheimia* the *sodA* provided higher resolution than 16S rRNA gene with a mean identity of 83.5% compared to 98.2% with the latter (Gautier et al., 2005). In the study on the genus *Haemophilus* the mean identity value of *sodA* achieved 74.9% and 93.5% for 16S rRNA (Cattoir et al., 2006).

Figure 4 shows the phylogeny based on the currently available *sodA* sequences from *Pasteurellaceae*. Comparison of the similarity values of *sodA* between the 44 analysed *Pasteurellaceae* resulted in more than 64% sequence similarity within the family and more than 80% within genera. In the individual studies (Cattoir et al., 2006; Gautier et al., 2005), the *sodA* based tree topology was in good agreement with the corresponding phylogeny provided by the16S rRNA gene, however there were a few discrepancies observed in the full tree (Fig. 4). Clustering confirmed previous observations that certain taxa should be reclassified due to their phylogenetic position. However, [*Haemophilus*] *paracuniculus* interfered with the genus *Mannheimia*, members of the genus *Haemophilus sensu stricto* were

split in two and [*H.*] *felis* interfered with the genus *Pasteurella sensu stricto*. The full tree contains much more entries than the individual studies what might have led to this changed topology. Moreover, due to the high discriminatory power of the *sodA* gene resolution might be too high in order to delineate larger phylogenetic distances within the family and complementary analysis of amino acid sequences might be indicated in this case. Nevertheless, the sequences were shown to be useful for diagnostic purposes and have a good resolution at the species level. Therefore it certainly has a great potential to help resolving classification within *Pasteurellaceae*, but further analysis extending sequencing of *sodA* to the other members of the family will be necessary to see if it will also be useful for the phylogeny at the family level.

Phylogeny based on atpD

The *atpD* gene encodes the β-subunit of the ATP synthase. Its role is to couple the electrochemical potential difference for H^+ across the inner membrane to synthesis of ATP from ADP and P_i. The enzyme is composed of the two subunits F_o and F_1. F_o traverses the membrane facilitating proton transport and F_1 is located on the cytoplasmic side of the membrane responsible for ATPase activity. F_1 is composed of the five polypeptides α, β, γ, and ε. The gene *atpD* is responsible for the β-subunit functioning in nucleotide binding and catalysis. Its ubiquitous distribution within *Bacteria* and high sequence conservation makes it suitable for phylogenetic analysis (Walker et al., 1984). Within *Pasteurellaceae* from which data are available the gene is 1374 bp long. Sequences covering nearly whole length of *atpD* were generated and used for investigation of genetic diversity of *P. multocida* fowl cholera isolates (Petersen et al., 2001). By sequence analysis of *atpD* it was possible to separate *P. multocida* subsp. *septica* from the other two subspecies, but similar to 16S rRNA gene-based analysis, *P. multocida* subsp. *multocida* and *P. multocida* subsp. *gallicida* could not be separated using *atpD* showing close relatedness of these two subspecies. More information about usability of the gene for clarification of phylogeny within *Pasteurellaceae* was provided by an analysis comprising a number of different representatives of the family (Christensen et al., 2004b). The phylogeny based on analysis of deduced protein sequences from DNA indicated that *atpD* might be useful for recognition of genera within the family.

The *atpD* sequence of only a limited number of representatives of the family is available and primers used for determining them are listed in Table 1. Figure 5 shows the phylogenetic tree

based on *atpD* from 31 taxa including ten genera and a few representatives of potential new taxa. The tree is in good agreement with the 16S rRNA gene based phylogeny but better resolves some groups as e.g. *Bibersteinia* from *Mannheimia*. Therefore *atpD* is certainly also a promising phylogenetic marker for the family and more data including more representatives should be gained. Comparison of the sequence similarity of *atpD* between the limited number of analysed strains resulted in more than 78% within the family and more than 89% within genera.

Phylogeny based on *recN*, *thdF* and *rpoA* representing the whole genome sequence

DNA relatedness of taxa as determined by DNA-DNA hybridization gives an overall estimate about phylogenetic relationship between species and is a taxonomic marker (Wayne et al., 1987). However, the method is time-consuming requiring cross-comparisons of new taxa with established ones, technically challenging and sometimes not possible to perform with certain taxa. The reproducibility is questionable when comparing results generated in different laboratories under different conditions. Moreover, the method only has a resolution at the species level with 85% being the threshold for *Pasteurellaceae* species (compared to the generally used 70%) and therefore it is not useful to delineate relationship at the genus level (Mutters et al., 1989; Stackebrandt and Goebel, 1994). Whole genome sequencing provides the most precise information about genetic content of organisms and would allow cross-comparison with available genome sequences. However, this might be an option for the future since generating sequences from single strains is still expensive, comparison of the data using nowadays available bioinformatic tools is time-consuming and simple whole genome analysis processes are not commonly available. An alternative to whole genome sequence comparison was recently demonstrated by the work of Zeigler (2003) showing that comparison of only three genes *recN* (encoding DNA repair protein), *thdF* (encoding ATPase) and *rpoA* (encoding the α subunit of the RNA polymerase) between species has the potential to predict the whole-genome relatedness of bacteria thereby presenting a very promising solution for comprehensive phylogenetic studies of bacteria. Even more, only the *recN* gene might be sufficient for that purpose, however with less confidence. Whole-genome relatedness for two related bacterial strains expressed as a similarity value (SI) can be calculated from the similarities of the corresponding marker genes using the specific formulas

SI_{genome}=- 1.88+0.52(SI*recN*)+1.78(SI*thdF*)+0.52(SI*rpoA*) and SI_{genome}=-1.30+2.25(SI*recN*), respectively.

This approach was taken to establish a general sequence determination method for *Pasteurellaceae* based on 43 strains representing 37 species and subspecies including type strain of all the currently 13 genera from the family (Kuhnert and Korczak, 2006). Calculated genome similarity values based on *recN*, *thdF* and *rpoA* as well as of only *recN* correlated with available published DNA-DNA reassociation results. Moreover, it allowed deducing the %GC-content of species from the three genes, a further marker used in taxonomy. Like that it was for the first time possible to carry out "whole-genome" cross-comparisons of species and genera on a large scale and analysis of data allowed establishing threshold values for species (SI genome around 0.85) and genus (SI genome around 0.4) either based on all three genes or *recN* alone. The similarity values provide a broad view on genetic relatedness between species belonging to the family and showed clear allocation of species representing the same genus in particular groups. The approach was since used to clarify the taxonomic position of *A. capsulatus* as well as for description of the new species *Av. endocarditidis* (Bisgaard et al., 2007; Kuhnert et al., 2007).

DNA-DNA hybridization values as well as the similarity values calculated on the three genes also provide some phylogenetic information. However, the gene sequences of the three genes *recN*, *thdF* and *rpoA* can further be used to construct phylogenetic trees, thereby enhancing resolution gained from the sequence data compared to calculation of similarity values alone. Figure 6 shows a combined tree based on the distance matrices of the three genes. The tree very nicely resolves all the genera and monophyletic groups. Moreover it also clearly distinguishes species within a genus. In the previous study we included 16S rRNA gene, *rpoB*, and *infB* in a multilocus sequence analysis (MLSA) and constructed a combined tree based on all six genes (Kuhnert and Korczak, 2006). The congruence between trees obtained from each of the six genes calculated by Pearson correlation showed that topologies of individual trees obtained from the *recN*, *thdF* and *rpoA* genes most properly represents the combined tree whereby the *recN* tree showed the highest similarity to the combined tree (Kuhnert and Korczak, 2006). Congruence values with the combined tree were 97% for *recN*, 95% for *thdF*, 93% for *rpoB*, 92% for *rpoA* and *infB* and only 83% with 16S rRNA gene. This is further evidence that these genes and in particular *recN* are highly representative for the entire genome of taxa.

Interestingly, the rough tree structure in Fig. 6 is congruent with the one described by Redfield et al. (2006). There a phylogenetic analysis of amino acid sequences of 12 well-conserved genes (*gapdH*, *lepA*, *ffh*, *serS*, *secD*, *dapA*, *ruvB*, *xerC*, *ispB*, *secA*, *crp*, and *dnaJ*) deduced from 8 *Pasteurellaceae* genome sequences was carried out (see also Fig. 3 Chapter 3). The two subclades represented by *H. influenzae* and *A. pleuropneumoniae* can be recognized in the *recN-thdF-rpoA* consensus tree. They are also specified by the DNA uptake signal sequences specific for the two large phylogenetic groups (Chapter 3). It will be interesting to see whether branches formed by the genus *Gallibacterium* and *Phocoenobacter* are also defined by specific DNA uptake signal sequences. One could speculate that also the Testudinis cluster, which forms a clear outgroup in all trees where this group was included (Fig. 1-5) has its specific uptake signal sequences.

Phylogeny based on multilocus sequence typing (MLST)

Multilocus sequence typing (MLST) is used to explore very close phylogenetic relationships and is mainly applied for defining genotypes within a species (Maiden et al., 1998). The technique is based on amplification and sequencing of internal fragments of about 500 bp from 6 to 8 house-keeping genes. Every unique sequence of each gene is given a unique allele number, and the combination of allele numbers of all gene loci defines the allelic profile of a strain and is given as a sequence type (ST). In addition to the designation of ST, sequences can also be used for phylogenetic analysis. It furthermore allows assessing the recombination frequency within species by looking at the congruence of the individual trees.

MLST has been applied for only few members of the family *Pasteurellaceae* mainly for *Haemophilus*. Meats et al. (2003) used fragments of *adk* (adenylate kinase), *atpG* (ATP synthase F1 γ subunit), *frdB* (fumarate reductase iron-sulfur protein), *fucK* (fuculokinase), *mdh* (malate dehydrogenase), *pgi* (glucose-6-phosphate isomerase) and *recA* (RecA protein) for characterisation of encapsulated and non-capsulated *H. influenzae* strains. For phylogenetic purpose sequences of all seven gene fragments were concatenated in a sequence of totally 3057 bp and a distance matrix based tree was constructed. The study allowed reconstructing phylogenetic relationships between different lineages of *H. influenzae*. Two major groups of encapsulated *H. influenzae* could be discerned from the strains analysed, whereas the non-capsulated isolates showed little structuring in the tree, probably as a result

of higher recombination frequency in the latter. The MLST typing scheme for *H. influenzae* is available on the internet under www.mlst.net.

The three genes *adk*, *pgi*, *recA* from the work of Meats et al. (2003) and additionally the *infB* were used in a multilocus sequence analysis of the genus *Haemophilus* resulting in the description of a new species *H. pittmaniae* (Norskov-Lauritsen et al., 2005). This study also clearly showed the advantage of using house-keeping genes rather than the 16S rRNA gene for resolving the phylogenetic relatedness within the genus *Haemophilus*.

A third study was done on [*Haemophilus*]. *parasuis* based on the genes *mdh*, *6pgd* (6-phosphogluconate dehydrogenase), *atpD*, *frdB*, *g3pd* (glyceraldehyde-3-phosphate dehydrogenase), *infB* and *rpoB* (Olvera et al., 2006). This investigation revealed that this species undergoes recombination to a significant extent and that it is very heterogeneous with many sequence types and no clonal structure. Nevertheless, two major lineages could be discerned.

An MLST scheme based on the genes *adk*, *aroA* (3-phosphoshikimate 1-carboxyvinyl transferase), *deoD* (purine-nucleoside phosphorylase), *gdhA* (glutamate dehydrogenase), *g6pd* (glucose-6-phosphate 1-dehydrogenase), *mdh* and *pgi* was also developed for *P. multocida* and applied on bovine strains (Davies et al., 2004). The study revealed that MLST gives only limited information on the phylogeny within the isolates from cattle, however it could still be a valuable tool for phylogenetic analysis of *P. multocida* from different sources.

MLST schemes have also been established for *Agg. actinomycetemcomitans* based on 6 of the genes used for *H. influenzae*. One study investigated the intragenomic recombination in the highly toxigenic clone JP2 (Eriksen et al., 2005). Whereas MLST showed almost no discrimination and target genes showed only minor sequence difference between the five strains analysed, they surprisingly found high recombination of rRNA genes. In a second study Haubek et al. (2007) looked at more geographically, ethnically and timely separated JP2 isolates and found a clonal population structure with evolutionary lineages corresponding to serotypes.

Phylogeny of the genera

In this section data gained from the various phylogenetic marker genes described in the previous part are summarized to get a closer look into the phylogeny between and within the

genera. Generally there is good congruence between the trees and not much discussion about the structure of recently described genera as well as of course the monophyletic genera. Discrepancies between the topology of trees are most probably due to recombination events for certain genes in certain species. This is one argument more to include as many genes as possible in a phylogenetic comparison of species.

The genus *Actinobacillus sensu stricto* is composed of the species *A. lignieresii*, *A. pleuropneumoniae*, *A. suis*, *A. ureae*, *A. hominis*, *A. arthritidis*, *A. capsulatus,* the two subspecies of *A. equuli* and two genomospecies (Table 4 of Chapter 1). These species form a distinct cluster in the 16S rRNA gene, *infB*, *rpoB* and the *recN-thdF-rpoA* based trees (Fig. 1-3 and 6) with the exception of *A. capsulatus* in the 16S rRNA gene tree where it clusters clearly outside. This was the reason why the taxonomic position of *A. capsulatus* within the genus was questioned. However, as mentioned previously in a recent polyphasic study including 23 strains it was shown, that this species should definitely be left in the genus (Kuhnert et al., 2007), confirming earlier DNA-DNA reassociation studies (Escande et al., 1984; Mutters et al., 1989). Furthermore, the type species *A. lignieresii* and *A. pleuropneumoniae* are very closely related in all trees where we have the corresponding sequence data available. Based on the genome similarity values of more than 0.9 of *A. lignieresii* to various *A. pleuropneumoniae* serotypes (Kuhnert and Korczak, 2006), these two species could phylogenetically be regarded as the same species.
Related to the genus *Actinobacillus* is the branch composed of the [*A.*] *minor* type strain and the '*A. porcitonsillarum*' reference strain. These are phenotypically very closely related and the main difference is in haemolysis (Moller et al., 1996; Gottschalk et al., 2003). The two form a distinct branch in the 16S rRNA, *infB, rpoB* and *recN-thdF-rpoA* trees, separated from but close to the cluster of *Actinobacillus*. The distinct branching, as well as the genome similarity value of only 0.79 between the two and less than 0.4 to other *Actinobacillus* species (Kuhnert and Korczak, 2006) favour the argument that [*A.*] *minor* and '*A. porcitonsillarum*' might represent separate species that form a new genus, adding that the species '*A. porcitonsillarum*' still awaits formal description. Strains belonging to this putative species pose a diagnostic problem since they can be misidentified as *A. pleuropneumoniae*, but can actually be regarded as normal flora of the pig (Gottschalk et al., 2003). All other species currently assigned to the genus *Actinobacillus* are clearly misclassified based on their phylogenetic position in all trees.

The genus *Aggregatibacter* has only recently been described and contains the three species *Agg. actinomycetemcomitans*, *Agg. aphrophilus* and *Agg. segnis* (Norskov-Lauritsen and Kilian, 2006). A monophyletic branching of the three species was shown in this paper, however only 6 other taxa of *Pasteurellaceae* and *E. coli* as an outgroup were included in the analysis. Monophyly of the genus is also seen in all protein encoding gene-based trees where more taxa were included (Fig. 2-5) as well as in the *recN-thdF-rpoA* consensus tree (Fig. 6). The structure in the 16S rRNA gene tree is somewhat less clear since *Agg. actinomycetemcomitans* clusters with *Haemophilus*. However, when including one or more additional sequences from *Agg. actinomycetemcomitans* strains the genus forms also a monophyletic cluster (data not shown). This phenomenon can often be seen and shows once more that phylogenetic trees are dynamic structures with changing topology and relationships when including new and additional data (Ludwig and Schleifer, 1999).

The genus *Avibacterium* is also one of the recently established genera and includes the reclassified species *Av. avium*, *Av. gallinarum, Av. paragallinarum* and *Av. volantium* (Blackall et al., 2005) as well as the new species *A. endocarditidis* (Bisgaard et al., 2007). The monophyletic nature of the genus is reflected in the 16S rRNA gene tree as well as in the other phylogenetic analyses.

The genus *Gallibacterium* currently comprises the species *G. anatis* and genomospecies 1 (Christensen et al., 2003a) and will be expanded by new species formed by representatives of Bisgaard Taxon 2 and 3 (H. Christensen, personal communication). In all phylogenetic trees they form a monophyletic branch. Remarkable is that *Gallibacterium* together with *Avibacterium* and *Volucribacter* form the 'avian' cluster 18 of Olsen (Olsen et al., 2005) in the 16S rRNA tree based on neighbor-joining (Fig. 1) as well as based on maximum-likelihood analysis (Christensen et al., 2004b). However, this 'avian' cluster is not reflected in the other trees where *Volucribacter* as well as the other two genera *Avibacterium* and *Gallibacterium* form distinct and separated clusters.

The genus *Haemophilus* contains many misclassified species what is based on the long taxonomic history this genus has. Growth requirements, i.e. the need for haemin (X factor) and β-NAD (V factor) were the basis to assign species to this genus. Phylogeny impressively

shows that the phenotypic criteria for classification are unrelated to genetic similarity of the species. In the herein presented analyses only a limited number of species from this genus were included, but the work of Hedegaard and collaborators showed that phylogeny of *infB* as well as DNA-DNA hybridizations suggest that based on genetic relationship the genus *sensu stricto* should be formed by *H. influenzae*, *H. aegyptius*, *H. haemolyticus*, *H. parainfluenzae* and *H. pittmaniae* (Hedegaard et al., 2001; Norskov-Lauritsen et al., 2005). This is also in agreement with the 16S rRNA gene phylogeny, however it stays in contrast to the branching observed in the *sodA* tree, where *H. influenzae*, *H. aegyptius*, *H. haemolyticus* are on a quite distant branch to *H. parainfluenzae* and *H. pittmaniae* (Fig. 4). The presence of two lineages can also be seen in the 16S rRNA gene tree, however the *Haemophilus sensu stricto* are on the same branch in this case.

The genus *Mannheimia* was established on a polyphasic approach and the five species can be found on the same branch in all phylogenetic analyses. *M. haemolytica* and *M. glucosida* are always closely related what is most clearly seen in the *recN-thdF-rpoA* tree. The genome similarity value of 0.88 (Kuhnert and Korczak, 2006) is also quite high and indicates the close genetic relationship between the two, which phylogenetically could be regarded the same species or at least represent subspecies of the same species. Related to the two is also *M. ruminalis* whereas *M. granulomatis* and *M. varigena* seem to form a different lineage within the genus as seen from all trees where data from all five species were available. From all trees it becomes obvious that '*M. succiniciproducens*' is wrongly assigned to the genus and shows that classification on simple and limited 16S rRNA gene analysis as done in this case can lead to wrong conclusions (Lee et al., 2002). The *recN-thdF-rpoA* analysis impressively shows this misclassification, since '*M. succiniciproducens*' is on a different and most distant cluster to the genus *Mannheimia* (Fig. 6).

Similar to the other two old "HAP" genera the genus *Pasteurella* needs revision, since it still includes many wrongly classified species despite the description of former members as *Mannheimia, Gallibacterium, Avibacterium* and *Bibersteinia*. What is regarded as *sensu stricto* includes the *P. multocida* subspecies, *P. canis*, *P. stomatis* and *P. dagmatis*. Good congruence for the phylogenetic clustering of these species is seen in the trees. Based on 16S rRNA gene and *rpoB* analyses it was speculated that the *P. multocida* subsp. *septica* could be regarded a single species, since it separates form the other two subspecies (Kuhnert et al.,

2000; Davies, 2004) what was already discussed in its original description by Mutters et al. (1985). However, based on the *recN-thdF-rpoA* tree topology and genome similarity values of about 0.87 it would make sense to keep the subspecies concept (Kuhnert and Korczak, 2006).

The genus *Volucribacter*, previously known as a group of organisms originating from psittacine birds called Bisgaard taxon 33 (Bisgaard et al., 1999), currently comprises two species *V. psittacicida* and *V. amazonae*. These two species appear to be genetically closely related to each other with 16S rRNA gene sequence similarity ranging from 99.5% to 99.8% what makes their differentiation using this marker hardly possible. However, the DNA-DNA reassociation value between type strains of *V. psittacicida* and *V. amazonae* was only 47%, clearly separating the two species from each other. Investigation of the genetic relationships between members of the taxon and the other known *Pasteurellaceae* based on 16S rRNA sequences provided clear evidence that the genus *Volucribacter* forms a distinct monophyletic group within the family *Pasteurellaceae* (Christensen et al., 2004a) and the highest 16S rRNA gene sequence similarity outside the genus it shares with Bisgaard taxon 34 strain 69 (94.6%) and with the type strain of *Av. avium* (94.5%) (Christensen et al., 2003b; Blackall et al., 2005; Christensen et al., 2003a).

The other named genera are all defined by a single species and form monophyletic taxa in most of the trees. *Bibersteinia* is the newest genus within the *Pasteurellaceae*, which nowadays comprises only one species namely *B. trehalosi*. This organism was assigned to the family some time ago. In 1990 Sneath and Stevens (Sneath and Stevens, 1990) described a new species as [*Pasteurella*] *trehalosi* that had been classified as a part of the [*Pasteurella*] *haemolytica* complex (Angen et al., 1999). The close phylogenetic relationship to the genus *Mannheimia* is reflected in all the trees (Fig. 1-6). However, in all trees *B. trehalosi* clearly separates from the *Mannheimia* species and previous 16S rRNA gene analysis indicated that the four recognized serovars within the species are closely related showing at least 98.7% similarity (Davies et al., 1996).

The genus *Lonepinella* represented by *L. koalarum* was described based on seven strains representing three biovars (Osawa et al., 1995). The 16S rRNA gene sequences of the seven strains showed some variability going to more than 2%. The monotypic nature of the genus is well represented in the phylogenetic trees.

Nicoletella semolina was shown to be a phylogenetically very homogenous species with very low intraspecies variability of 0.3%, 0.4% and 0.2% in the analysed *infB*, *rpoB* and 16S rRNA gene sequences, respectively (Kuhnert et al., 2004). The genus is always located on a distinct branch in the trees with little association to other genera.

The description of *Phocoenobacter uteri* is only based on a single isolate and there is no report in the literature on the isolation of additional strains of this species (Foster et al., 2000). Nevertheless, also this genus seems to form a monotypic branch in all the analysed trees. Interestingly, [*P.*] *skyensis* in the 16S rRNA gene tree collocates on the same branch and this distinct lineage was also shown by Birkbeck et al. (2002). Since we do not have comparable sequence data for the two species for the other genes, we can only speculate that the two might be members of the same genus.

Strains reclassified as *Hist. somni* were also shown to be rather homogenous with only 0.5% 16S rRNA gene divergence in 19 isolates from geographically different parts of the world (Angen et al., 2003). Analysis of *rpoB* resulted in two subclusters within the monotypic main cluster (Angen et al., 2003). The monotypic nature of *Histophilus* is also seen in the *recN-thdF-rpoA* based tree (Fig. 6).

Putative new genera

As outlined a few times in this chapter as well as in others, new genera and species will be described within *Pasteurellaceae*. Phylogenetic analyses indicate which of the named species are candidates to form a new genus and will help to classify unnamed (e.g. the Bisgaard taxa) or newly isolated taxa. In the following we list a few such groups based on the herein presented phylogenetic investigations. Some of these groups are rather clearly defined others less, and more thorough investigations for description of these putative new genera are certainly needed. However, it becomes obvious that the number of genera within the family can easily increase to over 30. In this context it should be remembered that good taxonomic practise includes more than a single strain for description of new taxa and we recommend to refer to the "minimal standards" recently published for taxonomy of *Pasteurellaceae* (Christensen et al., 2001; Christensen et al., 2007).

A distinct branch is always formed by the [*P.*] *testudinis* and Bisgaard taxon 14. This Testudinis branch of the phylogenetic trees is certainly underrepresented by isolates analysed

in this chapter as well as in the literature in general. What in diagnostics is identified as [*P.*] *testudinis* is probably a phylogenetically very heterogenous group of isolates as 16S rRNA gene and *rpoB* sequence analyses indicate (P. Kuhnert and H. Christensen, unpublished data). It can therefore be assumed, that at least one new genus and several new species can be delineated from such strains in the future.

The Delphinicola cluster formed by the sea mammal isolates [*Actinobacillus*] *scotiae* and [*Actinobacillus*] *delphinicola* shows also deep branching in the 16S rRNA, *infB* and *rpoB* trees and could therefore form a distinct new genus (Fig. 1-3).

The wrongly classified '*M. succiniciproducens*' was included in all herein constructed trees and forms a monophyletic branch in each. There is currently only one isolate reported which is deposited at the Culture Collection of the University of Goteborg (CCUG) but due to patent rights not publicly available (Lee et al., 2002). This makes further taxonomic investigations of this putative new genus difficult even though a lot of genomic and proteomic data are available (Hong et al., 2004; Lee et al., 2006). Another species isolated from bovine rumen, [*Actinobacillus*] *succinogenes* clusters close to '*M. succiniciproducens*', however the branches are rather deep, making it unlikely that they belong to the same genus. The same holds true for Bisgaard taxon 10, however only 16S rRNA gene data is available in this case.

A Rodent cluster is formed by [*Pasteurella*] *pneumotropica*, Bisgaard Taxon 17 and [*A.*] *muris*. It is however difficult to assess if this corresponds to one genus or more and further data are needed to get a more precise picture.

The Seminis cluster formed of the species [*Actinobacillus*] *seminis*, [*P.*] *aerogenes* and [*Pasteurella*] *mairii* also forms a distinct branch in the 16S rRNA, the *infB* and the *rpoB* tree, suggesting a genus rank of this group as has been suggested earlier (Christensen et al., 2005). More phenotypic and genotypic data is however needed on the [*A.*] *seminis* in order to define this genus and separate the species within. Interestingly, the RTX toxin operon *pax*, supposed to be involved in abortion cases, can be found in this phylogenetic cluster (Mayor et al., 2006).

The closely related species [*Haemophilus*] *parahaemolyticus* and [*Haemophilus*] *paraphrohaemolyticus* form the Parahaemolyticus group, a distinct branch as seen in the 16S rRNA gene and *sodA* trees.

The transfer of [*A.*] *minor* and '*A. porcitonsillarum*' to a separate genus was discussed before. This is especially obvious when looking at the *recN-thdF-rpoA* tree, where they cluster distantly from the genus *Actinobacillus*.

The Parasuis group consists of the two species [*H.*] *parasuis* and [*Actinobacillus*] *indolicus* forming a monophyletic branch in the 16S rRNA, *infB*, *rpoB* and *recN-thdF-rpoA* trees. Interestingly the two type strains of the species showed a significant genome similarity values of 0.6 suggesting they in fact belong to the same genus (Kuhnert and Korczak, 2006).
The three species [*H.*] *ducreyi*, [*Pasteurella*] *caballi* and [*Pasteurella*] *bettyae* are on the same branch in the 16S rRNA gene tree, but clearly separated by distance. The fact that they cluster separately in the *rpoB*, *sodA*, *atpD* trees could indicate that they represent separate genera.
[*H.*] *felis* and [*Haemophilus*] *haemoglobinophilus* each forms a separate branch in the 16S rRNA, *rpoB* and *sodA* trees, suggesting they represent two genera.
The phylogenetic position of Bisgaard taxon 7 in the 16S rRNA, *rpoB* and *atpD* trees also suggest that it forms a new genus. Of course analysis of more strains representing this taxon is necessary to support this. The same holds true for Bisgaard taxon 5.
Difficult to interpret are the data for [*Pasteurella*] *langaaensis*, [*Actinobacillus*] *rossii* and [*Actinobacillus*] *porcinus*, which are related in the *rpoB* sequence but the latter being separated in the 16S rRNA gene tree. Data are not conclusive at the moment.
Also limited data is available for [*H.*] *paracuniculus*, which nevertheless forms a monotypic branch in the *sodA* tree and forms a distinct cluster with *A. capsulatus* in 16S rRNA gene tree.

Conclusions

Based on the different phylogenetic approaches presented and the various genes they rely on one wonders which would be the perfect phylogene. However, it is obvious that each of the genes has its advantages but will also result in difficult to interprete relationship in certain cases. The 16S rRNA gene phylogeny will certainly continue to be a *conditio sine qua non* for the investigation and description of new taxa. It might be useful to delineate the large phylogenetic relations over bacterial families very nicely and in many cases it is also resolving species relationship. However, this gene has also many drawbacks, stands in some cases in contradiction to the phylogeny based on alternative targets and should therefore not be taken as the sole reference gene for phylogenetic analysis in *Pasteurellaceae*. It will be interesting to see, if this holds true for other bacterial families as well. House-keeping genes proved very useful for investigating the phylogeny within *Pasteurellaceae*, and from the ones described in this chapter (*infB*, *rpoB*, *sodA*, *atpD*) each will probably be equally applicable. In

general the inclusion of several genes for phylogenetic analysis is in any case favourable and will give more precise ideas on the genetic relationship of taxa. In this respect the most promising candidates would be genes representing whole genome similarities like *recN*, *thdF* and *rpoA*. If, based on current knowledge, any single gene should be recommended as "the phylogene" for *Pasteurellaceae* it might be the *recN*.

Figure captions

Fig. 1 Phylogenetic tree based on the 16S rRNA gene from members of *Pasteurellaceae*. *E. coli* was included as an outgroup to root the tree. The neighbour-joining tree was built from a Jukes-Cantor corrected distance matrix in Bionumerics v.4.6. Bootstrap values of 500 trees are indicated as % confidence values for the branching. The distance bar on top represents % differences.

Fig. 2 Phylogenetic tree based on the *infB* gene from members of *Pasteurellaceae*. *E. coli* was included as an outgroup to root the tree. The neighbour-joining tree was built from a Jukes-Cantor corrected distance matrix in Bionumerics v.4.6. Bootstrap values of 500 trees are indicated as % confidence values for the branching. The distance bar on top represents % differences. * Gene ID, Oralgen www.oralgen.lanl.gov

Fig. 3 Phylogenetic tree based on the *rpoB* gene from members of *Pasteurellaceae*. *E. coli* was included as an outgroup to root the tree. The neighbour-joining tree was built from a Jukes-Cantor corrected distance matrix in Bionumerics v.4.6. Bootstrap values of 500 trees are indicated as % confidence values for the branching. The distance bar on top represents % differences. * Gene ID, Oralgen www.oralgen.lanl.gov

Fig. 4 Phylogenetic tree based on the *sodA* gene from members of *Pasteurellaceae*. *E. coli* was included as an outgroup to root the tree. The neighbour-joining tree was built from a Jukes-Cantor corrected distance matrix in Bionumerics v.4.6. Bootstrap values of 500 trees are indicated as % confidence values for the branching. The distance bar on top represents % differences.

Fig. 5 Phylogenetic tree based on the *atpD* gene from members of *Pasteurellaceae*. ***E. coli*** was included as an outgroup to root the tree. The neighbour-joining tree was built from a Jukes-Cantor corrected distance matrix in Bionumerics v.4.6. Bootstrap values of 500 trees are indicated as % confidence values for the branching. The distance bar on top represents % differences. * Gene ID, Oralgen www.oralgen.lanl.gov

Fig. 6 Consensus tree based on the *recN*, *thdF* and *rpoA* genes from members of *Pasteurellaceae*. ***E. coli*** was included as an outgroup to root the tree. Jukes-Cantor correction was applied for the distance matrix and neighbor-joining for tree construction in Bionumerics v.4.6. Cophenetic correlations are given, indicating the reliability of the branching compared to the actual genetic relatedness of the taxa. The distance bar on top represents % differences.

Reference list

Acinas,S.G., Marcelino,L.A., Klepac-Ceraj,V., and Polz,M.F. (2004). Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. J. Bacteriol. *186*, 2629-2635.

Acinas,S.G., Sarma-Rupavtarm,R., Klepac-Ceraj,V., and Polz,M.F. (2005). PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. Appl. Environ. Microbiol. *71*, 8966-8969.

Angen,O., Ahrens,P., Kuhnert,P., Christensen,H., and Mutters,R. (2003). Proposal of *Histophilus somni* gen. nov., sp. nov. for the three species incertae sedis '*Haemophilus somnus*', '*Haemophilus agni*' and '*Histophilus ovis*'. Int. J. Syst. Evol. Microbiol. *53*, 1449-1456.

Angen,O., Mutters,R., Caugant,D.A., Olsen,J.E., and Bisgaard,M. (1999). Taxonomic relationships of the [*Pasteurella*] *haemolytica* complex as evaluated by DNA-DNA hybridizations and 16S rRNA sequencing with proposal of *Mannheimia haemolytica* gen. nov., comb. nov., *Mannheimia granulomatis* comb. nov., *Mannheimia glucosida* sp. nov., *Mannheimia ruminalis* sp. nov. and *Mannheimia varigena* sp. nov. Int. J. Syst. Bacteriol. *49*, 67-86.

Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J., Ostell,J., and Wheeler,D.L. (2007). GenBank. Nucleic Acids Res. *35*, D21-D25.

Birkbeck,T.H., Laidler,L.A., Grant,A.N., and Cox,D.I. (2002). *Pasteurella skyensis* sp. nov., isolated from Atlantic salmon (*Salmo salar* L.). Int. J. Syst. Evol. Microbiol. *52*, 699-704.

Bisgaard,M., Christensen,J.P., Bojesen,A.M., and Christensen,H. (2007). *Avibacterium endocarditidis* sp. nov., isolated from valvular endocarditis in chickens. Int. J. Syst. Evol. Microbiol. *57*, 1729-1734.

Bisgaard,M., Hinz,K.-H., Petersen,K.D., and Christensen,J.P. (1999). Identification of members of the *Pasteurellaceae* isolated from birds and characterization of two new taxa isolated from psittacine birds. Avian Pathol. *28*, 369-377.

Blackall,P.J., Bojesen,A.M., Christensen,H., and Bisgaard,M. (2007). Reclassification of [*Pasteurella*] *trehalosi* as *Bibersteinia trehalosi* gen. nov., comb. nov. Int. J. Syst. Evol. Microbiol. *57*, 666-674.

Blackall,P.J., Christensen,H., Beckenham,T., Blackall,L.L., and Bisgaard,M. (2005). Reclassification of *Pasteurella gallinarum*, [*Haemophilus*] *paragallinarum*, *Pasteurella avium* and *Pasteurella volantium* as *Avibacterium gallinarum* gen. nov., comb. nov., *Avibacterium paragallinarum* comb. nov., *Avibacterium avium* comb. nov. and *Avibacterium volantium* comb. nov. Int. J. Syst. Evol. Microbiol. *55*, 353-362.

Cattoir,V., Lemenand,O., Avril,J.L., and Gaillot,O. (2006). The *sodA* gene as a target for phylogenetic dissection of the genus *Haemophilus* and accurate identification of human clinical isolates. Int. J. Med. Microbiol. *296*, 531-540.

Christensen,H., Bisgaard,M., Aalbaek,B., and Olsen,J.E. (2004a). Reclassification of Bisgaard taxon 33 with proposal of *Volucribacter psittacicida* gen. nov., sp. nov. and *Volucribacter amazonae* sp. nov. as new members of *Pasteurellaceae* . Int. J. Syst. Evol. Microbiol. *54*, 813-818.

Christensen,H., Bisgaard,M., Bojesen,A.M., Mutters,R., and Olsen,J.E. (2003a). Genetic relationships among avian isolates classified as *Pasteurella haemolytica*, '*Actinobacillus salpingitidis*' or *Pasteurella anatis* with proposal of *Gallibacterium anatis* gen. nov., comb. nov. and description of additional genomospecies within *Gallibacterium* gen. nov. Int. J Syst. Evol. Microbiol. *53*, 275-287.

Christensen,H., Bisgaard,M., Frederiksen,W., Mutters,R., Kuhnert,P., and Olsen,J.E. (2001). Is characterization of a single isolate sufficient for valid publication of a new genus or species? Proposal to modify recommendation 30b of the Bacteriological Code (1990 Revision). Int. J Syst. Evol. Microbiol. *51*, 2221-2225.

Christensen,H., Foster,G., Christensen,J.P., Pennycott,T., Olsen,J.E., and Bisgaard,M. (2003b). Phylogenetic analysis by 16S rDNA gene sequence comparison of avian taxa of Bisgaard and characterization and description of two new taxa of *Pasteurellaceae*. J. Appl. Microbiol. *95*, 354-363.

Christensen,H., Kuhnert,P., Bisgaard,M., Mutters,R., Dziva,F., and Olsen,J.E. (2005). Emended description of porcine [*Pasteurella*] *aerogenes*, [*Pasteurella*] *mairii* and [*Actinobacillus*] *rossii*. Int. J. Syst. Evol. Microbiol. *55*, 209-223.

Christensen,H., Kuhnert,P., Busse,H.J., Frederiksen,W.C., and Bisgaard,M. (2007). Proposed minimal standards for the description of genera, species and subspecies of the *Pasteurellaceae*. Int. J. Syst. Evol. Microbiol. *57*, 166-178.

Christensen,H., Kuhnert,P., Olsen,J.E., and Bisgaard,M. (2004b). Comparative phylogenies of the housekeeping genes *atpD*, *infB* and *rpoB* and the 16S rRNA gene within the *Pasteurellaceae*. Int. J. Syst. Evol. Microbiol. *54*, 1601-1609.

Chuba,P.J., Bock,R., Graf,G., Adam,T., and Gobel,U. (1988). Comparison of 16S rRNA sequences from the family *Pasteurellaceae*: phylogenetic relatedness by cluster analysis. J. Gen. Microbiol. *134*, 1923-1930.

Davies,R.L. (2004). Genetic diversity among *Pasteurella multocida* strains of avian, bovine, ovine and porcine origin from England and Wales by comparative sequence analysis of the 16S rRNA gene. Microbiology *150*, 4199-4210.

Davies,R.L., MacCorquodale,R., and Reilly,S. (2004). Characterisation of bovine strains of *Pasteurella multocida* and comparison with isolates of avian, ovine and porcine origin. Vet. Microbiol. *99*, 145-158.

Davies,R.L., Paster,B.J., and Dewhirst,F.E. (1996). Phylogenetic relationship and diversity within the *Pasteurella haemolytica* complex based on 16S rRNA sequence comparison and outer membrane protein and lipopolysaccharide analysis. Int. J. Syst. Bacteriol. *46*, 736-744.

De Ley,J., Mannheim,W., Mutters,R., Piechulla,K., Tytgat,R., Segers,P., Bisgaard,M., Frederiksen,W., Hinz,K.H., and Vanhoucke,M. (1990). Inter- and intrafamilial similarities of rRNA cistrons of the *Pasteurellaceae*. Int. J. Syst. Bacteriol. *40*, 126-137.

Dewhirst,F.E., Paster,B.J., Olsen,I., and Fraser,G.J. (1992). Phylogeny of 54 representative strains of species in the family *Pasteurellaceae* as determined by comparison of 16S rRNA sequences. J. Bacteriol. *174*, 2002-2013.

Dewhirst,F.E., Paster,B.J., Olsen,I., and Fraser,G.J. (1993). Phylogeny of the *Pasteurellaceae* as determined by comparison of 16S ribosomal ribonucleic acid sequences. Zentralbl. Bakteriol. *279*, 35-44.

Eriksen,K.T., Haubek,D., and Poulsen,K. (2005). Intragenomic recombination in the highly leukotoxic JP2 clone of Actinobacillus actinomycetemcomitans. Microbiology *151*, 3371-3379.

Escande,F., Grimont,P.A.D., and Bercovier,H. (1984). Deoxyribonucleic acid relatedness among strains of *Actinobacillus* spp. and *Pasteurella ureae*. Int. J. Syst. Bacteriol. *34*, 309-315.

Foster,G., Ross,H.M., Malnick,H., Willems,A., Hutson,R.A., Reid,R.J., and Collins,M.D. (2000). *Phocoenobacter uteri* gen. nov., sp. nov., a new member of the family *Pasteurellaceae* Pohl (1979) 1981 isolated from a harbour porpoise (*Phocoena phocoena*). Int. J Syst. Evol. Microbiol *50*, 135-139.

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

Gautier,A.L., Dubois,D., Escande,F., Avril,J.L., Trieu-Cuot,P., and Gaillot,O. (2005). Rapid and accurate identification of human isolates of *Pasteurella* and related species by sequencing the *sodA* gene. J. Clin. Microbiol. *43*, 2307-2314.

Gottschalk,M., Broes,A., Mittal,K.R., Kobisch,M., Kuhnert,P., Lebrun,A., and Frey,J. (2003). Non-pathogenic *Actinobacillus* isolates antigenically and biochemically similar to *Actinobacillus pleuropneumoniae*: a novel species? Vet. Microbiol. *92*, 87-101.

Gutell,R.R., Larsen,N., and Woese,C.R. (1994). Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol. Rev. *58*, 10-26.

Haubek,D., Poulsen,K., and Kilian,M. (2007). Microevolution and patterns of dissemination of the JP2 clone of *Aggregatibacter* (*Actinobacillus*) *actinomycetemcomitans*. Infect. Immun. *75*, 3080-3088.

Hedegaard,J., Okkels,H., Bruun,B., Kilian,M., Mortensen,K.K., and Norskov-Lauritsen,N. (2001). Phylogeny of the genus *Haemophilus* as determined by comparison of partial *infB* sequences. Microbiology *147*, 2599-2609.

Hedegaard,J., Steffensen,S.A., Norskov-Lauritsen,N., Mortensen,K.K., and Sperling-Petersen,H.U. (1999). Identification of *Enterobacteriaceae* by partial sequencing of the gene encoding translation initiation factor 2. Int. J. Syst. Bacteriol. *49*, 1531-1538.

Hong,S.H., Kim,J.S., Lee,S.Y., In,Y.H., Choi,S.S., Rih,J.K., Kim,C.H., Jeong,H., Hur,C.G., and Kim,J.J. (2004). The genome sequence of the capnophilic rumen bacterium *Mannheimia succiniciproducens*. Nat. Biotechnol. *22*, 1275-1281.

Klenk,H.P. and Zillig,W. (1994). DNA-dependent RNA polymerase subunit B as a tool for phylogenetic reconstructions: branching topology of the archaeal domain. J. Mol. Evol. *38*, 420-432.

Korczak,B., Christensen,H., Emler,S., Frey,J., and Kuhnert,P. (2004). Phylogeny of the family *Pasteurellaceae* based on *rpoB* sequences. Int. J. Syst. Evol. Microbiol. *54*, 1393-1399.

Kuhnert,P., Boerlin,P., Emler,S., Krawinkler,M., and Frey,J. (2000). Phylogenetic analysis of *Pasteurella multocida* subspecies and molecular identification of feline *P. multocida* subsp. *septica* by 16S rRNA gene sequencing. Int. J. Med. Microbiol. *290*, 599-604.

Kuhnert,P., Frey,J., Lang,N.P., and Mayfield,L. (2002). A phylogenetic analysis of *Prevotella nigrescens*, *Prevotella intermedia* and *Porphyromonas gingivalis* field strains reveals a clear species clustering. Int. J. Syst. Evol. Microbiol. *52*, 1391-1395.

Kuhnert,P., Korczak,B., Falsen,E., Straub,R., Hoops,A., Boerlin,P., Frey,J., and Mutters,R. (2004). *Nicoletella semolina* gen.nov., sp.nov., a new member of *Pasteurellaceae* isolated from horses with airway disease. J. Clin. Microbiol. *42*, 5542-5548.

Kuhnert,P. and Korczak,B.M. (2006). Prediction of whole genome DNA-DNA similarity, determination of G+C content and phylogenetic analysis within the family *Pasteurellaceae* by multilocus sequence analysis (MLSA). Microbiology *152*, 2537-2548.

Kuhnert,P., Korczak,B.M., Christensen,H., and Bisgaard,M. (2007). Emended description of *Actinobacillus capsulatus* Arseculeratne 1962, 38AL. Int. J. Syst. Evol. Microbiol. *57*, 625-632.

Lawrence,J.G., Hartl,D.L., and Ochman,H. (1991). Molecular considerations in the evolution of bacterial genes. J. Mol. Evol. *33*, 241-250.

Lee,J.W., Lee,S.Y., Song,H., and Yoo,J.S. (2006). The proteome of *Mannheimia succiniciproducens*, a capnophilic rumen bacterium. Proteomics *6*, 3550-3566.

Lee,P.C., Lee,S.Y., Hong,S.H., and Chang,H.N. (2002). Isolation and characterization of a new succinic acid-producing bacterium, *Mannheimia succiniciproducens* MBEL55E, from bovine rumen. Appl. Microbiol. Biotechnol. *58*, 663-668.

Lisitsyn,N.A., Monastyrskaya,G.S., and Sverdlov,E.D. (1988). Genes coding for RNA polymerase β subunit in bacteria. Structure/function analysis. Eur. J. Biochem. *177*, 363-369.

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

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. Proc. Natl. Acad. Sci. U. S. A *95*, 3140-3145.

Mayor,D., Korczak,B.M., Christensen,H., Bisgaard,M., Frey,J., and Kuhnert,P. (2006). Distribution of RTX toxin genes in strains of [*Actinobacillus*] *rossii* and [*Pasteurella*] *mairii*. Vet. Microbiol. *116*, 194-201.

Meats,E., Feil,E.J., Stringer,S., Cody,A.J., Goldstein,R., Kroll,J.S., Popovic,T., and Spratt,B.G. (2003). Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. J. Clin. Microbiol. *41*, 1623-1636.

Moller,K., Fussing,V., Grimont,P.A.D., Paster,B.J., Dewhirst,F.E., and Kilian,M. (1996). *Actinobacillus minor* sp. nov., *Actinobacillus porcinus* sp. nov., and *Actnobacillus indolicus* sp. nov., three new V factor-dependent species from the respiratory tract of pigs. Int. J. Syst. Bacteriol. *46*, 951-956.

Mollet,C., Drancourt,M., and Raoult,D. (1997). *rpoB* sequence analysis as a novel basis for bacterial identification. Mol. Microbiol. *26*, 1005-1011.

Mutters,R., Ihm,P., Pohl,S., Frederiksen,W., and Mannheim,W. (1985). Reclassification of the genus *Pasteurella* Trevisan 1887 on the basis of deoxyribonucleic acid homology, with proposals for the new species *Pasteurella dagmatis*, *Pasteurella canis*, *Pasteurella stomatis*, *Pasteurella anatis*, and *Pastuerella langaa*. Int. J. Syst. Bacteriol. *35*, 309-322.

Mutters,R., Mannheim,W., and Bisgaard,M. (1989). Taxonomy of the Group. In *Pasteurella* and Pasteurellosis, C.Adlam and J.M.Rutter, eds. (London: Academic Press), pp. 3-34.

Nolte,O. (1995). Nucleotide sequence and genetic varability of a part of the *rpoB* gene encoding the second largest subunit of DNA-directed RNA polymerase of *Neisseria meningitidis*. Med. Microbiol. Lett. *4*, 59-67.

Norskov-Lauritsen,N., Bruun,B., and Kilian,M. (2005). Multilocus sequence phylogenetic study of the genus *Haemophilus* with description of *Haemophilus pittmaniae* sp. nov. Int. J. Syst. Evol. Microbiol. *55*, 449-456.

Norskov-Lauritsen,N. and Kilian,M. (2006). Reclassification of *Actinobacillus actinomycetemcomitans, Haemophilus aphrophilus, Haemophilus paraphrophilus* and *Haemophilus segnis* as *Aggregatibacter actinomycetemcomitans* gen. nov., comb. nov., *Aggregatibacter aphrophilus* comb. nov. and *Aggregatibacter segnis* comb. nov., and emended description of *Aggregatibacter aphrophilus* to include V factor-dependent and V factor-independent isolates. Int. J. Syst. Evol. Microbiol. *56*, 2135-2146.

Olivier,A., Lee,H.Y., and Cote,J.C. (2005). Study of the heterogeneity of 16S rRNA genes in gamma-proteobacteria: implications for phylogenetic analysis. J. Gen. Appl. Microbiol *51*, 395-405.

Olsen,I., Dewhirst,F.E., Paster,B.J., and Busse,H.J. (2005). Family *Pasteurellaceae*. In Bergey's Manual of Systematic Bacteriology, R.Garrity, ed. (New York: Springer).

Olvera,A., Cerda-Cuellar,M., and Aragon,V. (2006). Study of the population structure of *Haemophilus parasuis* by multilocus sequence typing. Microbiology *152*, 3683-3690.

Osawa,R., Rainey,F.A., Fujisawa,T., Lang,E., Busse,H.J., Walsh,T., and Stackebrandt,E. (1995). *Lonepinella koalarum* gen. nov., sp. nov., a new tannin-protein complex degrading bacterium. Syst. Appl. Microbiol. *18*, 368-373.

Ovchinnikov,Y.A., Monastyrskaya,G.S., Gubanov,V.V., Guryev,S.O., Chertov,O.Y., Modyanov,N.N., Grinkevich,V.A., Makarova,I.A., Marchenko,T.V., Polovnikova,I.N., Lipkin,V.M., and Sverdlov,E.D. (1981). The primary structure of *Escherichia coli* RNA polymerase. Nucleotide sequence of the *rpoB* gene and amino-acid sequence of the β-subunit. Eur. J. Biochem. *116*, 621-629.

Palenik,B. (1992). Polymerase evolution and organism evolution. Curr. Opin. Genet. Dev. *2*, 931-936.

Petersen,K.D., Christensen,H., Bisgaard,M., and Olsen,J.E. (2001). Genetic diversity of *Pasteurella multocida* fowl cholera isolates as demonstrated by ribotyping and 16S rRNA and partial *atpD* sequence comparisons. Microbiology *147*, 2739-2748.

Pohl,S. (1981). DNA relatedness among members of *Haemophilus*, *Pasteurella* and *Actinobacillus*. In *Haemophilus*, *Pasteurella* and *Actinobacillus*, M.Kilian, W.Frederiksen, and E.L.Biberstein, eds. (London: Academic Press), p. 253.

Poyart,C., Quesne,G., Coulon,S., Berche,P., and Trieu-Cuot,P. (1998). Identification of streptococci to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase. J. Clin. Microbiol. *36*, 41-47.

Redfield,R.J., Findlay,W.A., Bosse,J., Kroll,J.S., Cameron,A.D., and Nash,J.H. (2006). Evolution of competence and DNA uptake specificity in the *Pasteurellaceae*. BMC Evol. Biol. *6*, 82.

Rowland,G.C., Aboshkiwa,M., and Coleman,G. (1993). Comparative sequence analysis and predicted phylogeny of the DNA-dependent RNA polymerase β subunits of *Staphylococcus aureus* and other eubacteria. Biochem. Soc. Trans. *21*, 40S-41S.

Saitou,N. and Nei,M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. *4*, 406-425.

Sneath,P.H. and Stevens,M. (1990). *Actinobacillus rossii* sp. nov., *Actinobacillus seminis* sp. nov., nom. rev., *Pasteurella bettii* sp. nov., *Pasteurella lymphangitidis* sp. nov., *Pasteurella mairi* sp. nov., and *Pasteurella trehalosi* sp. nov. Int. J. Syst. Bacteriol. *40*, 148-153.

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

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

Steffensen,S.A., Poulsen,A.B., Mortensen,K.K., and Sperling-Petersen,H.U. (1997). *E. coli* translation initiation factor IF2 - an extremely conserved protein. Comparative sequence analysis of the *infB* gene in clinical isolates of *E. coli*. FEBS Lett. *419*, 281-284.

Taillardat-Bisch,A.V., Raoult,D., and Drancourt,M. (2003). RNA polymerase beta-subunit-based phylogeny of *Ehrlichia* spp., *Anaplasma* spp., *Neorickettsia* spp. and *Wolbachia pipientis*. Int. J Syst. Evol. Microbiol. *53*, 455-458.

Tamura,K., Dudley,J., Nei,M., and Kumar,S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. *24*, 1596-1599.

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

Walker,J.E., Saraste,M., and Gay,N.J. (1984). The *unc* operon. Nucleotide sequence, regulation and structure of ATP-synthase. Biochim. Biophys. Acta *768*, 164-200.

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

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

Yap,W.H., Zhang,Z., and Wang,Y. (1999). Distinct types of rRNA operons exist in the genome of the actinomycete *Thermomonospora chromogena* and evidence for horizontal transfer of an entire rRNA operon. J. Bacteriol. *181*, 5201-5209.

Zeigler,D.R. (2003). Gene sequences useful for predicting relatedness of whole genomes in bacteria. Int. J. Syst. Evol. Microbiol. *53*, 1893-1900.

Table 1. Primers for amplification and sequencing of target genes used for phylogenetic investigations within the family *Pasteurellaceae*.

Target gene	PCR primers[a]		Product length[c]	Sequencing primers		References
	Name	Sequence[b]		Name	Sequence[b]	
16S rRNA gene	16SUNI-L	5‘–AGAGTTTGATCATGGCTCAG–3‘	1400	16SRNAI-S	5‘–CTACGGGAGGCAGCAGTGGGG–3‘	Kuhnert *et al*., 2002
	16SUNI-R	5‘–GTGTGACGGGCGGTGTGTAC–3‘		16SRNA1-S	5’–CTACGGGAGGCAGCAGTGTAG–3‘	
				16SRNAII-S	5‘–GTGTAGCGGTGAAATGCGTAG–3‘	
				16SRNA2-S	5‘–GTGTAGGGGTAAAATCCGTAG–3‘	
				16SRNAIV-S	5‘–GGTTAAGTCCCGCAACGAGCGC–3‘	
				16SRNA4-S	5‘–GCTTAAGTGCCATAACGAGCGC–3‘	
				16SRNAV-S	5‘–CCCCACTGCTGCCTCCCGTAG–3‘	
				16SRNAVI-S	5‘–CTACGCATTTCACCGCTACAC–3‘	
				16SRNA6-S	5‘–CTACGGATTTTACCCCTACAC–3‘	
				16SRNAVIII-S	5‘–GCGCTCGTTGCGGGACTTAACC–3‘	
				16SRNA8-S	5‘–GCGCTCGTTATGGCACTTAAGC–3‘	
infB	1186F	5‘–ATYATGGGHCAYGTHGAYCAYGGHAARAC–3‘	650	1748R	5‘–GTAGCAACCGGACCACGACCTTTAT–3‘	Hedegaard *et al*., 1999
	1833R	5‘–TATCCGACGCCGAACTCCGRTTNCGCATNGCNCGNAYNCGNCC–3‘		or 1745R	5‘–GCAACCGRICCICTICCTTTRTC–3‘	
	infB-L	5‘–ATGGGNCACGTTGACCACGGTAAAAC–3‘	1400	infB-L	5‘–ATGGGNCACGTTGACCACGGTAAAAC–3‘	Kuhnert *et al*. 2004
	infB-R	5‘–CCGATACCACATTCCATACC–3‘		infB-R	5‘–CCGATACCACATTCCATACC–3‘	
				infB-1	5‘–CGTGAYGAGAARAAAGCACGTGAAG–3‘	
				infB-2	5‘–CTTCACGTGCTTTYTTCTCRTCACG–3‘	
	infB-R1	5‘–TCGTTGTAGTTTTTCACGCCGATACC–3‘		infB-R1	5‘–TCGTTGTAGTTTTTCACGCCGATACC–3‘	Mayor *et al*. 2006
rpoB	Pasrpob-L	5‘–GCAGTGAAAGARTTCTTTGGTTC–3‘	560	Pasrpob-L	5‘–GCAGTGAAAGARTTCTTTGGTTC–3‘	Korczak *et al*. 2004
	Rrpob-R	5‘–GTTGCATGTTNGNACCCAT–3‘		Rrpob-R	5‘–GTTGCATGTTNGNACCCAT–3‘	
sodA	d1	5‘–CCITAYICITAYGAYGCIYTIGARCC–3‘	510	d1	5‘–CCITAYICITAYGAYGCIYTIGARCC–3‘	Poyart *et al*., 1998, Gautier *et al*., 2005, Cattoir *et al*., 2006
	d2	5‘–ARRTARTAIGCRTGYTCCCAIACRTC–3‘		d2	5‘–ARRTARTAIGCRTGYTCCCAIACRTC–3‘	
atpD	pmkatpf3w[d]	5’–GCTCGTCAAGCAAGTATYAC–3’	1430	pmkatpf3w[d]	5’–GCTCGTCAAGCAAGTATYAC–3’	Petersen *et al.*, 2001
	pmkatpr3w	5‘–CTAAYACTTCRTCGATKGWAC–3‘		pmkatpr3w	5‘–CTAAYACTTCRTCGATKGWAC–3‘	
				pmkatp4wf	5’–CGCAATGGGATCWTCWG–3’	
				pmkatpf5	5’–RAGYTATGAAGAACAAKC–3’	
				pmkatpf7	5’–TAYRAYGTTGCHCGTGGYG–3’	
				pmkatp4wr	5’–GWACCRTTAAATACTTCMGC–3’	
				pmkatpr5	5‘–GTTCTTSACCMACRAC–3‘	
rpoA	rpoA-L	5‘–TCTGTRACAGAATTTTTAAARCC–3‘	1000	rpoA-L	5‘–TCTGTRACAGAATTTTTAAARCC–3‘	Kuhnert & Korczak, 2006
	rpoA-R	5‘–TTGCNGGNGGCCARTTTTCAAGG–3‘		rpoA-R	5‘–TTGCNGGNGGCCARTTTTCAAGG–3‘	
recN	recN-L	5‘–CAACTYACTATYAATMATTTTGC–3‘	1400	recN-L	5‘–CAACTYACTATYAATMATTTTGC–3‘	
	recN-R	5‘–CTAATGCCYACRTCYACTTCATC–3‘		recN-R	5‘–CTAATGCCYACRTCYACTTCATC–3‘	

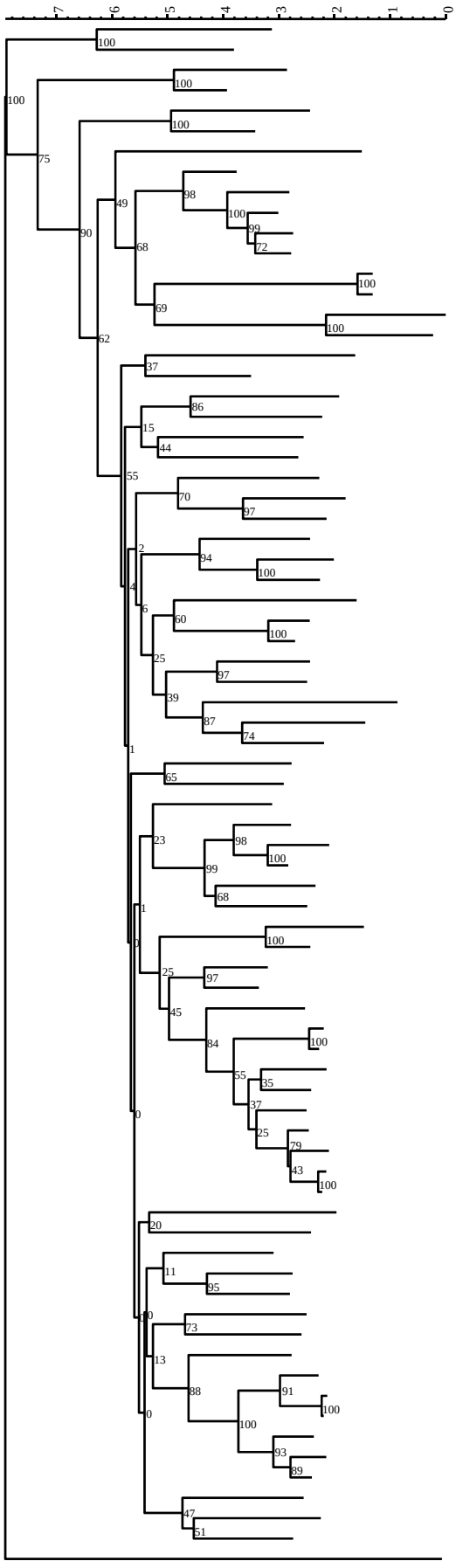
Table 1 (continued)

Target gene	PCR primers[a]		Product length[c]	Sequencing primers		References
	Name	Sequence[b]		Name	Sequence[b]	
recN	recN-L2	5‘–CATTTAACGGTTAATAATTTTGC–3‘		recN-L2	5‘–CATTTAACGGTTAATAATTTTGC–3‘	Kuhnert & Korczak, 2006
	recN-R2	5‘–CTAATYCCMACATCNACYTCATC–3‘		recN-R2	5‘–CTAATYCCMACATCNACYTCATC–3‘	
	recN_Vp-R	5‘–CTGATCCCTACATCAATTTCATC–3‘		recN_Vp-R	5‘–CTGATCCCTACATCAATTTCATC–3‘	
				recN_Ga-1	5‘–GCTATTTAGATGAATTGGTTGAGG–3‘	
				recN_Ga-2	5‘–CCTCAACCAATTCATCTAAATAGC–3‘	
thdF	thdF_first-L2	5‘–ATGAAAGANACNATYGTNGCWCARGC–3‘	1400	thdF_first-L2	5‘–ATGAAAGANACNATYGTNGCWCARGC–3‘	
	thdF_first-R2	5‘–TGTTTATTTNCCRATRCARAANGARC–3‘		thdF_first-R2	5‘–TGTTTATTTNCCRATRCARAANGARC–3‘	
	thdF-L2	5‘–AAAGANACNATYGTNGCWCARGC–3‘		thdF-L2	5‘–AAAGANACNATYGTNGCWCARGC–3‘	
	thdF-R2	5‘–TTATTTNCCRATRCARAANGARC–3‘		thdF-R2	5‘–TTATTTNCCRATRCARAANGARC–3‘	
	thdF_MP-L	5’–AAAGAMACCATTGTTGCWCAAGC–3’		thdF_MP-L	5’–AAAGAMACCATTGTTGCWCAAGC–3’	
				thdF-1	5‘–GGAATGAAAGTCGTGATTGCAGG–3‘	
				thdF-2	5‘–CCTGCAATCACGACTTTCATTCC–3‘	
				thdF-3	5‘–GGATGAAAGTGGTAATTGC–3‘	
				thdF-4	5’–GCAATTACCACTTTCATCC–3’	
adk	adk-up	5’–GGTGCACCGGGTGCAGGTAA–3’	620	adk-up	5’–GGTGCACCGGGTGCAGGTAA–3’	Meats *et al*. 2003
	adk-dn	5’–CCTAAGATTTTATCTAACTC–3’		adk-dn	5’–CCTAAGATTTTATCTAACTC–3’	
atpG	atpG-up	5‘–ATGGCAGGTGCAAAAGAGAT–3‘	560	atpG-up	5‘–ATGGCAGGTGCAAAAGAGAT–3‘	
	atpG-dn	5‘–TTGTACAACAGGCTTTTGCG–3‘		atpG-dn	5‘–TTGTACAACAGGCTTTTGCG–3‘	
frdB	frdB-up	5‘–CTTATCGTTGGTCTTGCCGT–3‘	580	frdB-up	5‘–CTTATCGTTGGTCTTGCCGT–3‘	
	frdB-dn	5‘–TTGGCACTTTCCACTTTTCC–3‘		frdB-dn	5‘–TTGGCACTTTCCACTTTTCC–3‘	
fucK	fucK-up	5‘–ACCACTTTCGGCGTGGATGG–3‘	560	fucK-up	5‘–ACCACTTTCGGCGTGGATGG–3‘	
	fucK-dn	5‘–AAGATTTCCCAGGTGCCAGA–3‘		fucK-dn	5‘–AAGATTTCCCAGGTGCCAGA–3‘	
mdh	mdh-up	5‘–TCATTGTATGATATTGCCCC–3‘	550	mdh-up	5‘–TCATTGTATGATATTGCCCC–3‘	
	mdh-dn	5‘–ACTTCTGTACCTGCATTTTG–3‘		mdh-dn	5‘–ACTTCTGTACCTGCATTTTG–3‘	
pgi	pgi-up	5‘–GGTGAAAAAATCAATCGTAC–3‘	590	pgi-up	5‘–GGTGAAAAAATCAATCGTAC–3‘	
	pgi-dn	5‘–ATTGAAAGACCAATAGCTGA–3‘		pgi-dn	5‘–ATTGAAAGACCAATAGCTGA–3‘	
recA	recA-up	5’–ATGGCAACTCAAGAAGAAAA–3’	620	recA-up	5’–ATGGCAACTCAAGAAGAAAA–3’	
	recA-dn	5’–TTACCAAACATCACGCCTAT–3’		recA-dn	5’–TTACCAAACATCACGCCTAT–3’	
adk	adk.34f	5‘–GGIAAAGGIACWCARGCICARTT–3‘	580	adk.34f	5‘–GGIAAAGGIACWCARGCICARTT–3‘	Nørskov-Lauritsen *et al*. 2005
	adk.610r	5‘–CTTCCACTTTTTKYGTMCCGTC–3‘		adk.610r	5‘–CTTCCACTTTTTKYGTMCCGTC–3‘	
pgi	pgi.838f	5’–GATGGIAAAGAYGTIATGCC–3’	490	pgi.838f	5’–GATGGIAAAGAYGTIATGCC–3’	
	pgi.1331r	5’–GCTGACCAYAAIGARTAACG–3’		pgi.1331r	5’–GCTGACCAYAAIGARTAACG–3’	

Table 1 (continued)

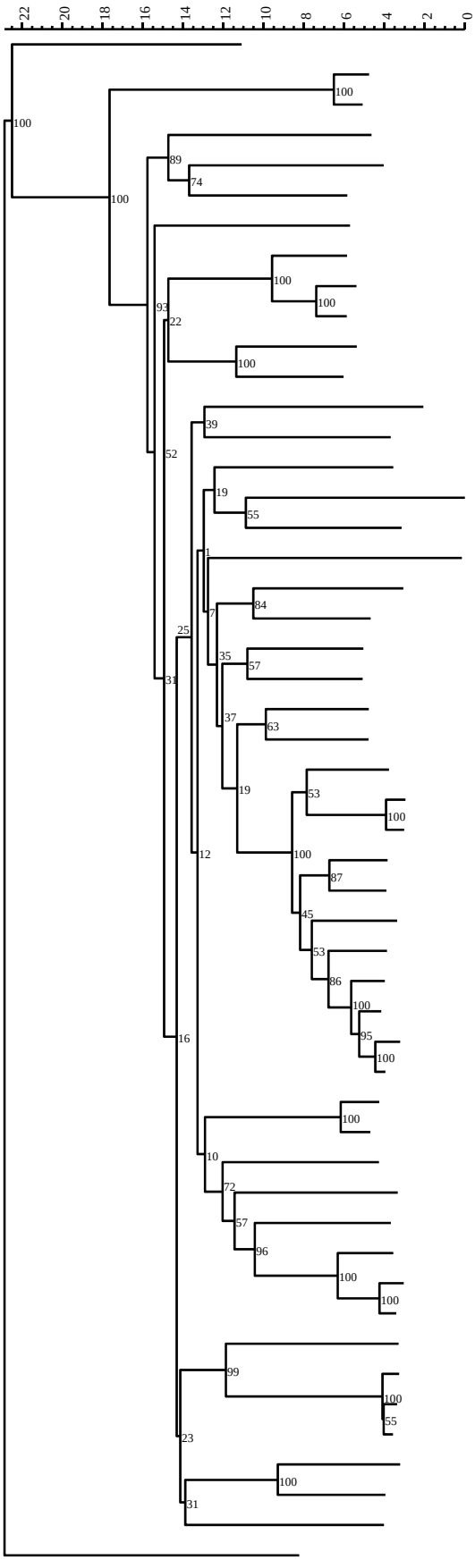
Target gene	PCR primers[a]		Product length[c]	Sequencing primers		References
	Name	Sequence[b]		Name	Sequence[b]	
recA	recA.54f recA.617r	5'–GARAAACAATTTGGKAAAGGC–3' 5'–TTACCRAACATMACRCCIAT–3'	560	recA.54f recA.617r	5'–GARAAACAATTTGGKAAAGGC–3' 5'–TTACCRAACATMACRCCIAT–3'	Nørskov-Lauritsen *et al.* 2005
6pgd	6pgdF 6pgdR	5'–TTATTACCGCACTTAGAAG–3' 5'–CGTTGATCTTTGAATGAAGA–3'	630	6pgdF 6pgdR	5'–TTATTACCGCACTTAGAAG–3' 5'–CGTTGATCTTTGAATGAAGA–3'	Olvera *et al.*, 2006
atpD	atpDF atpDR	5'–CAAGATGCAGTACCAAAAGTTTA–3' 5'–ACGACCTTCATCACGGAAT–3'	610	atpDF atpDR	5'–CAAGATGCAGTACCAAAAGTTTA–3' 5'–ACGACCTTCATCACGGAAT–3'	
frdB	frdBF frdBR	5'–CATATCGTTGGTCTTGCCGT–3' 5'–TTGGCACTTTCGATCTTACCTT–3'	580	frdBF frdBR	5'–CATATCGTTGGTCTTGCCGT–3' 5'–TTGGCACTTTCGATCTTACCTT–3'	
g3pd	3gpdF 3gpdR	5'–GGTCAAGACATCGTTTCTAAC–3' 5'–TCTAATACTTTGTTTGAGTAACC–3'	570	3gpdF 3gpdR	5'–GGTCAAGACATCGTTTCTAAC–3' 5'–TCTAATACTTTGTTTGAGTAACC–3'	
rpoB	rpoBF rpoBR	5'–TCACAACTTTCICAATTTATG–3' 5'–ACAGAAACCACTTGTTGCG–3'	470	rpoBF rpoBR	5'–TCACAACTTTCICAATTTATG–3' 5'–ACAGAAACCACTTGTTGCG–3'	
adk	adk Forward adk Reverse	5'–AAGGBACWCAAGCVCAAT–3' 5'–CACTTTTTKYGTMCCGTC–3'	570	adk Forward adk Reverse	5'–AAGGBACWCAAGCVCAAT–3' 5'–CACTTTTTKYGTMCCGTC–3'	Davies *et al.*, 2004
aroA	aroA Forward aroA Reverse	5'–TTTACCGGGCTCCAAAAG–3' 5'–CTTTTACGCGCCAGTTAT–3'	990	aroA Forward aroA Reverse	5'–TTTACCGGGCTCCAAAAG–3' 5'–TGCATCATCTTAAGGGTG–3'	
deoD	deoD Forward deoD Reverse	5'–GTGCATTTGCCGATGTTG-3' 5'–TGGTGTTGTTTGTTCGTG–3'	620	deoD Forward deoD Reverse	5'–GTGCATTTGCCGATGTTG–3' 5'–TGGTGTTGTTTGTTCGTG–3'	
gdhA	gdhA Forward gdhA Reverse	5'–CTTAGTTGAACCTGAACG–3' 5'–CTTGACCTTCAATCGTGC–3'	1100	gdhA Forward gdhA Reverse	5'–CGCGTTAACCACATTACC–3' 5'–CCCTTCAGCCACTAATTG–3'	
g6pd	g6pd Forward g6pd Reverse	5'–CHGGYGAYYTMACTYATCG–3' 5'–TTTBGCGATBARTTTRTCRGC–3'	1400	g6pd Forward g6pd Reverse	5'–GATGCTGCCGATTATGG–3' 5'–CAAGACTTTTGCCACTTC–3'	
mdh	mdh Forward mdh Reverse	5'–AAGTTGCWGTWYTAGGTG–3' 5'–CCTAATTCAATATCYGCACG–3'	910	mdh Forward mdh Reverse	5'–TGTCCAAAAGCTTGTGTG–3' 5'–CCTAATTCAATATCYGCACG–3'	
pgi	pgi Forward pgi Reverse	5'–GCCGTGGTTGGTTGATGG–3' 5'–TTGGCTTGGCGCGATGAA–3'	900	pgi Forward pgi Reverse	5'–TGATTTCTGGTGAATGGA–3' 5'–GGAAATACGCTGCAAAAC–3'	

[a] Annealing temperature is not given since it might vary due to different conditions and equipment in different laboratories.
[b] I=deoxyinosine; W= A or T; R= G or A; K= G or T; Y= C or T; M= A or C
[c] PCR product length might differ between species. Approximate size of amplicons is given.
[d] Primer aligns to *atpG*



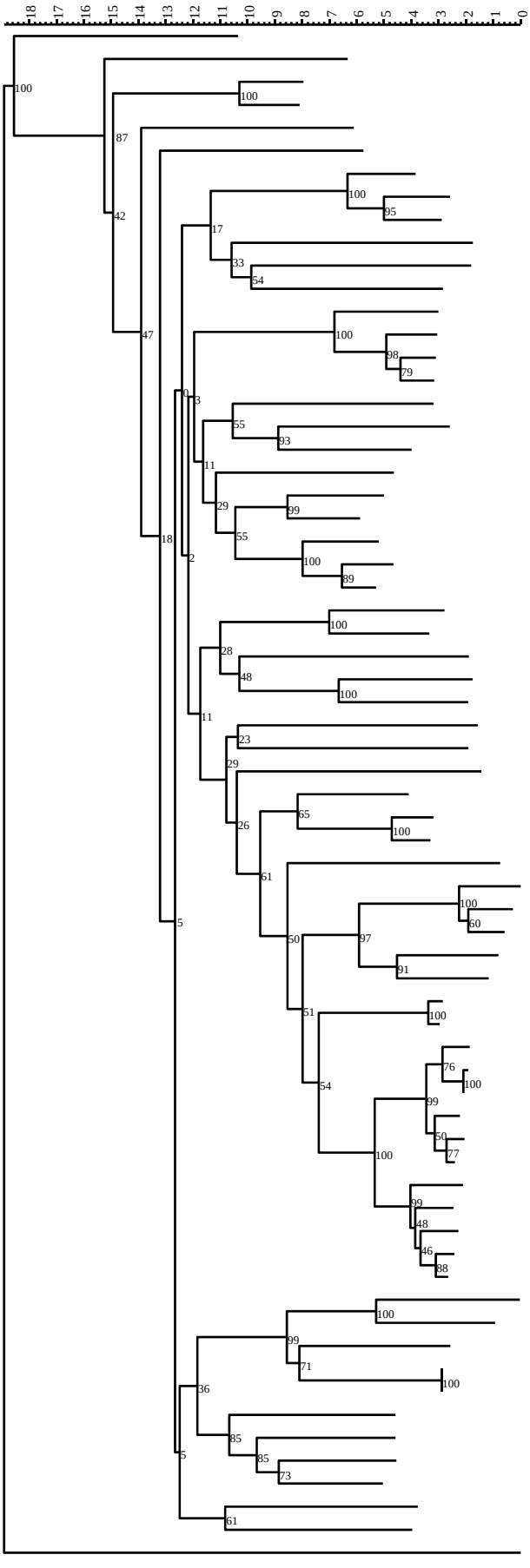
Bisgaard Taxon 14	CCUG16499	AY362901	**Testudinis**
[*Pasteurella*] *testudinis*	CCUG19802^T	AY362926	
[*Pasteurella*] *skyensis*	NCTC13204^T	AJ243202	***Phocoenobacter***
Phocoenobacter uteri	NCTC12872^T	X89379	
[*Actinobacillus*] *scotiae*	NCTC129 22^T	Y09653	**Delphinicola**
[*Actinobacillus*] *delphinicola*	NCTC12870^T	AY362889	
Histophilus somni	HS8025^T	AF549387	***Histophilus***
Avibacterium endocarditidis	20186H4H1^T	DQ465412	
Avibacterium paragallinarum	NCTC11296^T	AY498868	
Avibacterium avium	CCUG12833^T	AY362916	***Avibacterium***
Avibacterium gallinarum	NCTC11188^T	AY362921	
Avibacterium volantium	CCUG3713^T	AY362928	
Volucribacter amazonae	CCUG47537^T	AY216869	***Volucribacter***
Volucribacter psittacicida	CCUG47536^T	AY216868	
Gallibacterium genomospecies 1	CCM5974	AF228015	***Gallibacterium***
Gallibacterium anatis	CCUG15563^T	AF228001	
[*Pasteurella*] *langaaensis*	CCUG15566^T	AY362922	**Langaa**
[*Actinobacillus*] *rossii*	ATCC27072^T	AY362895	**Rossii**
[*Actinobacillus*] *succinogenes*	CCUG43843^T	AY362898	**Succinogenes**
Bisgaard Taxon 10	CCUG15572	AF024528	
Nicoletella semolina	CCUG43639^T	AY508816	***Nicoletella***
'*Mannheimia succiniciproducens*'	MBEL55E	NC_006300	**Succiniciproducens**
[*Actinobacillus*] *muris*	NCTC12432^T	AY362894	
Bisgaard Taxon 17	CCUG17206	AY362902	**Rodent**
[*Pasteurella*] *pneumotropica*	NCTC8141^T	AY362924	
[*Pasteurella*] *aerogenes*	ATCC27883^T	U66491	
[*Actinobacillus*] *seminis*	CCUG27187^T	AY362897	**Seminis**
[*Pasteurella*] *mairii*	CCUG27189^T	AY362923	
Haemophilus haemolyticus	NCTC10659^T	M75045	
Haemophilus aegyptius	CCUG25716^T	AY362905	
Haemophilus influenzae	Rd KW20	NC_000907	***Haemophilus sensu stricto***
Haemophilus parainfluenzae	CCUG12836^T	AY362908	
Haemophilus pittmaniae	CCUG48703^T	AJ290755	
Aggregatibacter actinomycetemcomitans	ATCC33384^T	M75039	
Aggregatibacter segnis	ATCC33393^T	M75043	***Aggregatibacter***
Aggregatibacter aphrophilus	CCUG3715^T	AY362906	
Bisgaard Taxon 7	CCUG24852	AY362904	**Taxon 7**
[*Haemophilus*] *haemoglobinophilus*	CCUG3714^T	AY362907	**Haemoglobinophilus**
Bibersteinia trehalosi	NCTC10370^T	AY362927	***Bibersteinia***
Mannheimia ruminalis	CCUG38470^T	AF053900	
Mannheimia haemolytica	NCTC9380^T	M75080	
Mannheimia glucosida	CCUG38457^T	AY362912	***Mannheimia***
Mannheimia granulomatis	ATCC49244^T	AY362913	
Mannheimia varigena	CCUG38462^T	AF053893	
[*Haemophilus*] *paraphrohaemolyticus*	NCTC10670^T	M75076	**Parahaemolyticus**
[*Haemophilus*] *parahaemolyticus*	NCTC8479T	AJ295746	
'*Actinobacillus porcitonsillarum*'	CCUG46996	AF486274	**Minor**
[*Actinobacillus*] *minor*	NM305	AY362893	
Actinobacillus genomospecies 1	CCUG22229	AF247722	
Actinobacillus lignieresii	NCTC4189^T	AY362892	
Actinobacillus pleuropneumoniae	4074^T	D30030	
Actinobacillus ureae	CCUG2139^T	AY362900	
Actinobacillus arthritidis	CCUG24862^T	AY362929	***Actinobacillus sensu stricto***
Actinobacillus genomospecies 2	CCUG15571	AF247717	
Actinobacillus equuli subsp. *haemolyticus*	CCUG19799^T	AF381187	
Actinobacillus hominis	NCTC11529^T	AY362890	
Actinobacillus suis	ATCC33415^T	AY362899	
Actinobacillus equuli subsp. *equuli*	ATCC19392^T	AF381186	
Lonepinella koalarum	ACM3666^T	Y17189	***Lonepinella***
Bisgaard Taxon 5	CCUG16493	AY362903	**Taxon 5**
[*Actinobacillus*] *porcinus*	CCUG38924^T	AY362896	**Porcinus**
[*Actinobacillus*] *indolicus*	46KC2	AY362891	**Parasuis**
[*Haemophilus*] *parasuis*	CCUG3712^T	AY362909	
[*Haemophilus*] *paracuniculus*	ATCC29986^T	M75061	**Paracuniculus**
Actinobacillus capsulatus	CCUG12396^T	AY362886	
[*Haemophilus*] *felis*	ATCC49733^T	AF224292	**Felis**
Pasteurella multocida subsp. *septica*	CCUG17977^T	AF294411	
Pasteurella multocida subsp. *multocida*	CCUG17976^T	AF294410	
Pasteurella multocida subsp. *gallicida*	CCUG17978^T	AF294412	***Pasteurella sensu stricto***
Pasteurella dagmatis	CCUG12397^T	AY362920	
Pasteurella canis	CCUG12400^T	AY362919	
Pasteurella stomatis	CCUG17979^T	AY362925	
[*Pasteurella*] *bettyae*	CCUG2042^T	AY362917	**Bettyae**
[*Haemophilus*] *ducreyi*	35000HP	NC_002940	**Ducreyi**
[*Pasteurella*] *caballi*	ATCC49197^T	AY362918	**Caballi**
Escherichia coli	MG1655	NC_00913	

Figure 1. 16S rRNA



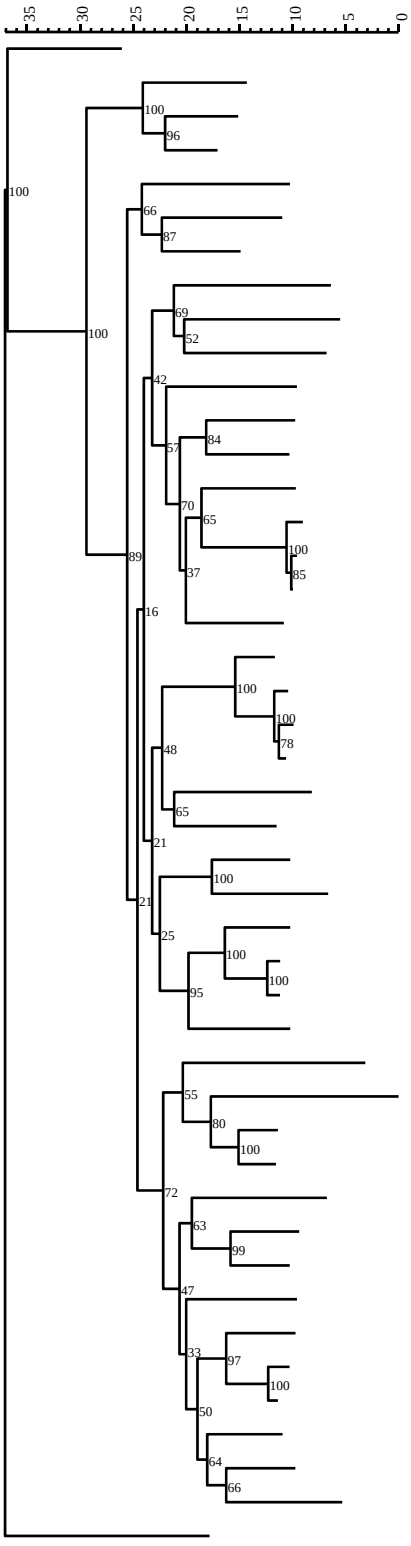
[*Pasteurella*] *testudinis*	CCUG19802^T	EU350942
Gallibacterium genomospecies 1	CCM5974	EU350939
Gallibacterium anatis	CCUG15563^T	AY508843
'*Mannheimia succiniciproducens*'	MBEL55E	NC_006300
[*Actinobacillus*] *succinogenes*	CCUG43843^T	EU350937
Aggregatibacter actinomycetemcomitans	HK1651	AA01922*
[*Actinobacillus*] *muris*	NCTC12432^T	EU350935
[*Pasteurella*] *aerogenes*	ATCC27883^T	DQ211791
[*Actinobacillus*] *seminis*	CCUG27187^T	DQ211793
[*Pasteurella*] *mairii*	CCUG27189^T	DQ211784
Avibacterium paragallinarum	NCTC11296^T	EU350938
Avibacterium gallinarum	NCTC11188^T	EF055560
Nicoletella semolina	CCUG43639^T	AY508835
Histophilus somni	HS8025^T	AY508848
Phocoenobacter uteri	NCTC12872^T	AY508849
[*Actinobacillus*] *scotiae*	CCUG46379^T	EU350936
[*Actinobacillus*] *delphinicola*	NCTC12870^T	EU350934
[*Haemophilus*] *ducreyi*	35000HP	NC_002940
[*Actinobacillus*] *rossii*	ATCC27072^T	DQ211763
[*Actinobacillus*] *porcinus*	NM319^T	EF059970
[*Actinobacillus*] *indolicus*	46KC2	EF059971
[*Haemophilus*] *parasuis*	CCUG3712^T	DQ410886
'*Actinobacillus porcitonsillarum*'	CCUG46996	EF055563
[*Actinobacillus*] *minor*	NM305	EF055562
Actinobacillus genomospecies 1	CCUG22229	AY508850
Actinobacillus pleuropneumoniae	4074^T	AY508851
Actinobacillus lignieresii	NCTC4189^T	AY508852
Actinobacillus arthritidis	CCUG24862^T	AY508854
Actinobacillus genomospecies 2	CCUG15571	AY508853
Actinobacillus capsulatus	CCUG12396^T	AY508855
Actinobacillus hominis	CCUG19800^T	AY508856
Actinobacillus ureae	CCUG2139^T	AY508860
Actinobacillus suis	ATCC33415^T	AY508857
Actinobacillus equuli subsp. *haemolyticus*	CCUG19799^T	AY508858
Actinobacillus equuli subsp. *equuli*	ATCC19392^T	AY508859
Haemophilus influenzae	Rd KW20	NC_000907
Haemophilus aegyptius	CCUG25716^T	EU350940
Bibersteinia trehalosi	NCTC10370^T	EF055561
Mannheimia granulomatis	ATCC49244^T	EF055564
Mannheimia varigena	CCUG38462^T	DQ410888
Mannheimia ruminalis	CCUG38470^T	DQ410889
Mannheimia haemolytica	NCTC9380^T	AY508847
Mannheimia glucosida	CCUG38457^T	DQ410887
Pasteurella canis	CCUG12400^T	EF055567
Pasteurella multocida subsp. *gallicida*	CCUG17978^T	EF055565
Pasteurella multocida subsp. *septica*	CCUG17977^T	EF055566
Pasteurella multocida subsp. *multocida*	CCUG17976^T	AY508844
Volucribacter psittacicida	CCUG47536^T	EF077221
Volucribacter amazonae	CCUG47537^T	EU350941
Lonepinella koalarum	ATCC700131^T	AY508846
Escherichia coli	MG1655	NC_00913

Figure 2. *infB*



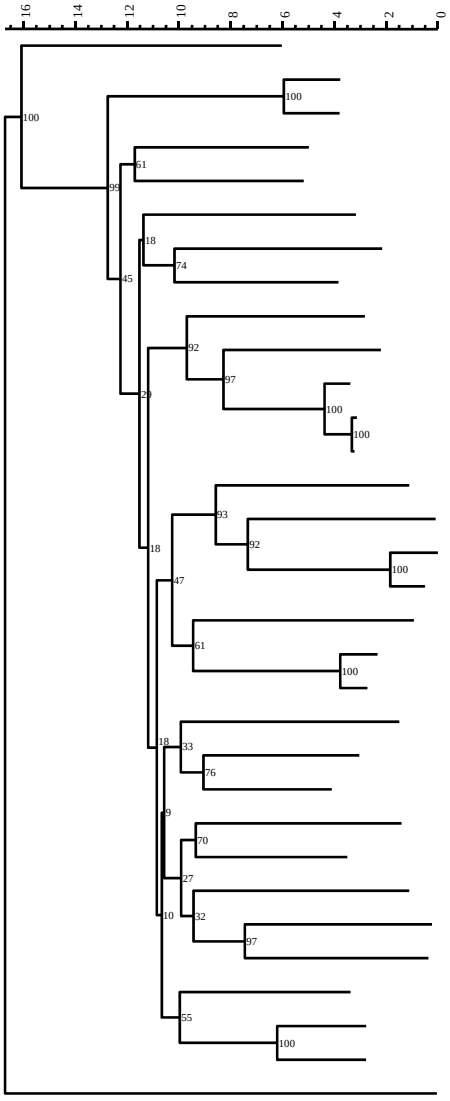
Bisgaard Taxon 14	CCUG16499	AY314043
[*Pasteurella*] *testudinis*	CCUG19802^{T}	AY362972
Aggregatibacter aphrophilus	CCUG3715^{T}	AY362953
Aggregatibacter actinomycetemcomitans	HK1651	AA00355*
Lonepinella koalarum	ATCC700131^{T}	AY170218
[*Haemophilus*] *felis*	ATCC49733^{T}	AY362954
[*Pasteurella*] *aerogenes*	ATCC27883^{T}	AY314040
[*Pasteurella*] *mairii*	CCUG27189^{T}	AY362968
[*Actinobacillus*] *seminis*	CCUG27187^{T}	AY362945
[*Actinobacillus*] *succinogenes*	CCUG43843^{T}	AY362946
'*Mannheimia succiniciproducens*'	MBEL55E	NC_006300
Bisgaard Taxon 5	CCUG16493	AY314039
Avibacterium paragallinarum	NCTC11296^{T}	EU350943
Avibacterium volantium	CCUG3713^{T}	AY362973
Avibacterium avium	CCUG12833^{T}	AY362965
Avibacterium gallinarum	NCTC11188^{T}	AY314041
[*Pasteurella*] *bettyae*	CCUG2042^{T}	AY314042
Histophilus somni	HS8025^{T}	AY170207
[*Haemophilus*] *haemoglobinophilus*	CCUG3714^{T}	AY362955
Pasteurella canis	CCUG12400^{T}	AY314038
Pasteurella dagmatis	CCUG12397^{T}	AY362966
Pasteurella stomatis	CCUG17979^{T}	AY362971
Pasteurella multocida subsp. *septica*	CCUG17977^{T}	AY362970
Pasteurella multocida subsp. *multocida*	CCUG17976^{T}	AY170216
Pasteurella multocida subsp. *gallicida*	CCUG17978^{T}	AY362969
Gallibacterium genomospecies 1	CCM5974	AY314033
Gallibacterium anatis	CCUG15563^{T}	AY314032
Bisgaard Taxon 7	CCUG24852	AY314030
Volucribacter psittacicida	CCUG47536^{T}	EF077220
Volucribacter amazonae	CCUG47537^{T}	EU350944
[*Actinobacillus*] *muris*	NCTC12432^{T}	AY314029
Nicoletella semolina	CCUG43639^{T}	AY508861
Phocoenobacter uteri	NCTC12872^{T}	AY170219
Bibersteinia trehalosi	NCTC10370^{T}	AY314028
[*Actinobacillus*] *indolicus*	46KC2	AY362938
[*Haemophilus*] *parasuis*	CCUG3712^{T}	AY314027
[*Haemophilus*] *ducreyi*	35000HP	NC_002940
Mannheimia ruminalis	CCUG38470^{T}	AY362961
Mannheimia haemolytica	ATCC33396^{T}	AY170217
Mannheimia glucosida	CCUG38457^{T}	AY362959
Mannheimia varigena	CCUG38462^{T}	AY362962
Mannheimia granulomatis	ATCC49244^{T}	AY362960
'*Actinobacillus porcitonsillarum*'	CCUG46996	AY362942
[*Actinobacillus*] *minor*	NM305	AY362939
Actinobacillus capsulatus	CCUG12396^{T}	AY362932
Actinobacillus hominis	NCTC11529^{T}	AY362937
Actinobacillus ureae	CCUG2139^{T}	AY362950
Actinobacillus suis	ATCC33415^{T}	AY362947
Actinobacillus equuli subsp. *equuli*	ATCC19392^{T}	AY362936
Actinobacillus equuli subsp. *haemolyticus*	CCUG19799^{T}	AY362935
Actinobacillus lignieresii	NCTC4189^{T}	AY170215
Actinobacillus pleuropneumoniae	4074^{T}	AY314026
Actinobacillus genomospecies 1	CCUG22229	AY362948
Actinobacillus genomospecies 2	CCUG15571	AY362951
Actinobacillus arthritidis	CCUG24862^{T}	AY362974
Bisgaard Taxon 17	CCUG17206	AY314035
[*Pasteurella*] *pneumotropica*	NCTC8141^{T}	AY314034
Haemophilus parainfluenzae	CCUG12836^{T}	AY362956
Haemophilus aegyptius	CCUG25716^{T}	AY362952
Haemophilus influenzae	Rd KW20	NC_000907
[*Actinobacillus*] *porcinus*	CCUG38924^{T}	AY362941
[*Actinobacillus*] *rossii*	ATCC27072^{T}	AY362940
[*Pasteurella*] *langaaensis*	CCUG15566^{T}	AY362967
[*Pasteurella*] *caballi*	ATCC49197^{T}	AY314036
[*Actinobacillus*] *delphinicola*	NCTC12870^{T}	AY314037
[*Actinobacillus*] *scotiae*	CCUG46379^{T}	AY362944
Escherichia coli	MG1655	NC_00913

Figure 3. *rpoB*



[*Pasteurella] testudinis*	CIP 82.117^{T}	AY702508
Aggregatibacter actinomycetemcomitans	CIP 52.106^{T}	DQ005612
Aggregatibacter aphrophilus	CIP 70.73^{T}	DQ005614
Aggregatibacter segnis	CIP 103292^{T}	DQ005625
Gallibacterium anatis	CIP 102679^{T}	AY702516
Haemophilus pittmaniae	CCUG48703^{T}	DQ267822
Haemophilus parainfluenzae	CIP 102513^{T}	AY702518
'*Mannheimia succiniciproducens*'	MBEL55E	NC_006300
[*Pasteurella*] *caballi*	CIP 103915^{T}	AY702495
[*Pasteurella*] *bettyae*	CIP 102636^{T}	AY702494
[*Pasteurella*] *langaaensis*	CIP 102678^{T}	AY702499
Pasteurella canis	CIP 103294^{T}	AY702496
Pasteurella stomatis	CIP 102680^{T}	AY702505
Pasteurella dagmatis	CIP103293^{T}	AY702497
Pasteurella multocida subsp. *septica*	CIP A125^{T}	AY702503
Pasteurella multocida subsp. *multocida*	CIP 103286^{T}	AY702502
Pasteurella multocida subsp. *gallicida*	CIP 103285^{T}	AY702501
[*Haemophilus*] *felis*	CIP 103890^{T}	DQ005616
Avibacterium avium	CIP 100919^{T}	AY702493
Avibacterium gallinarum	CIP 102676^{T}	AY702498
Avibacterium volantium	CIP 102677^{T}	AY702509
Avibacterium paragallinarum	CIP 103453^{T}	DQ005620
[*Pasteurella*] *mairii*	CIP 103902^{T}	AY702500
[*Pasteurella*] *aerogenes*	CIP 80.14^{T}	AY702492
Histophilus somni	129Pt	NC_008309
[*Haemophilus*] *haemoglobinophilus*	CIP 53.88^{T}	DQ005617
Haemophilus haemolyticus	CIP 103290^{T}	DQ005618
Haemophilus aegyptius	CIP 52.129^{T}	DQ005613
Haemophilus influenzae	Rd KW20	NC_000907
[*Pasteurella*] *pneumotropica*	CIP 66.16^{T}	AY702504
[*Pasteurella*] *skyensis*	CIP 107586^{T}	AY702506
[*Haemophilus*] *parasuis*	CIP 100918^{T}	DQ005624
[*Haemophilus*] *parahaemolyticus*	CIP 56.86^{T}	DQ005621
[*Haemophilus*] *paraphrohaemolyticus*	CIP 102512^{T}	DQ005622
[*Haemophilus*] *ducreyi*	35000HP	NC_002940
Actinobacillus pleuropneumoniae	CIP 100916^{T}	DQ005626
Actinobacillus ureae	CIP 62.8^{T}	AY702515
Bibersteinia trehalosi	CIP 103829^{T}	AY702507
Mannheimia ruminalis	CIP 106062^{T}	AY702513
Mannheimia haemolytica	CIP 103426^{T}	AY702512
Mannheimia glucosida	CIP 106063^{T}	AY702510
Mannheimia granulomatis	CIP 103433^{T}	AY702511
Mannheimia varigena	CIP 106064^{T}	AY702514
[*Haemophilus*] *paracuniculus*	CIP 107045^{T}	DQ005619
Escherichia coli	MG1655	NC_00913

Figure 4. *sodA*



Bisgaard Taxon 14	CCUG16499	AY296242
Gallibacterium anatis	F149^{T}	AY296227
Gallibacterium genomospecies 1	CCUG23139	AY296226
Bisgaard Taxon 5	7.3	AY296238
Aggregatibacter actinomycetemcomitans	HK1651	AA01482*
[*Actinobacillus*] *delphinicola*	CCUG41362^{T}	AY296240
Volucribacter psittacicida	JEO101	AY296228
Avibacterium gallinarum	NCTC11188^{T}	AY296246
Bisgaard Taxon 7	7.4	AY296239
Pasteurella canis	NCTC11621^{T}	AF326331
Pasteurella multocida subsp. *septica*	HIM746-6^{T}	AF326329
Pasteurella multocida subsp. *multocida*	NCTC10322^{T}	AF326326
Pasteurella multocida subsp. *gallicida*	HIM830-7^{T}	AF326327
Bibersteinia trehalosi	NCTC10370^{T}	AY296231
[*Haemophilus*] *parasuis*	CCUG3712^{T}	AY296235
Mannheimia haemolytica	NCTC9380^{T}	AY296230
Mannheimia glucosida	P737	AY296229
[*Haemophilus*] *ducreyi*	35000HP	NC_002940
Actinobacillus lignieresii	NCTC4189^{T}	AY296232
Actinobacillus pleuropneumoniae	4074^{T}	AACK01000012
[*Pasteurella*] *bettyae*	CCUG2042^{T}	AY296236
[*Pasteurella*] *caballi*	ATCC49197^{T}	AY296244
[*Actinobacillus*] *rossii*	CCUG24851^{T}	AY296243
Histophilus somni	129Pt	NC_008309
[*Pasteurella*] *aerogenes*	ATCC27883^{T}	AY296233
Lonepinella koalarum	ACM3666^{T}	AY296241
'*Mannheimia succiniciproducens*'	MBEL55E	NC_006300
[*Actinobacillus*] *succinogenes*	CCUG43843^{T}	AY296234
Haemophilus influenzae	Rd KW20	NC_000907
[*Pasteurella*] *pneumotropica*	NCTC8141^{T}	AY296245
Bisgaard Taxon 17	CCUG17206	AY296237
Escherichia coli	MG1655	NC_00913

Figure 5. *atpD*

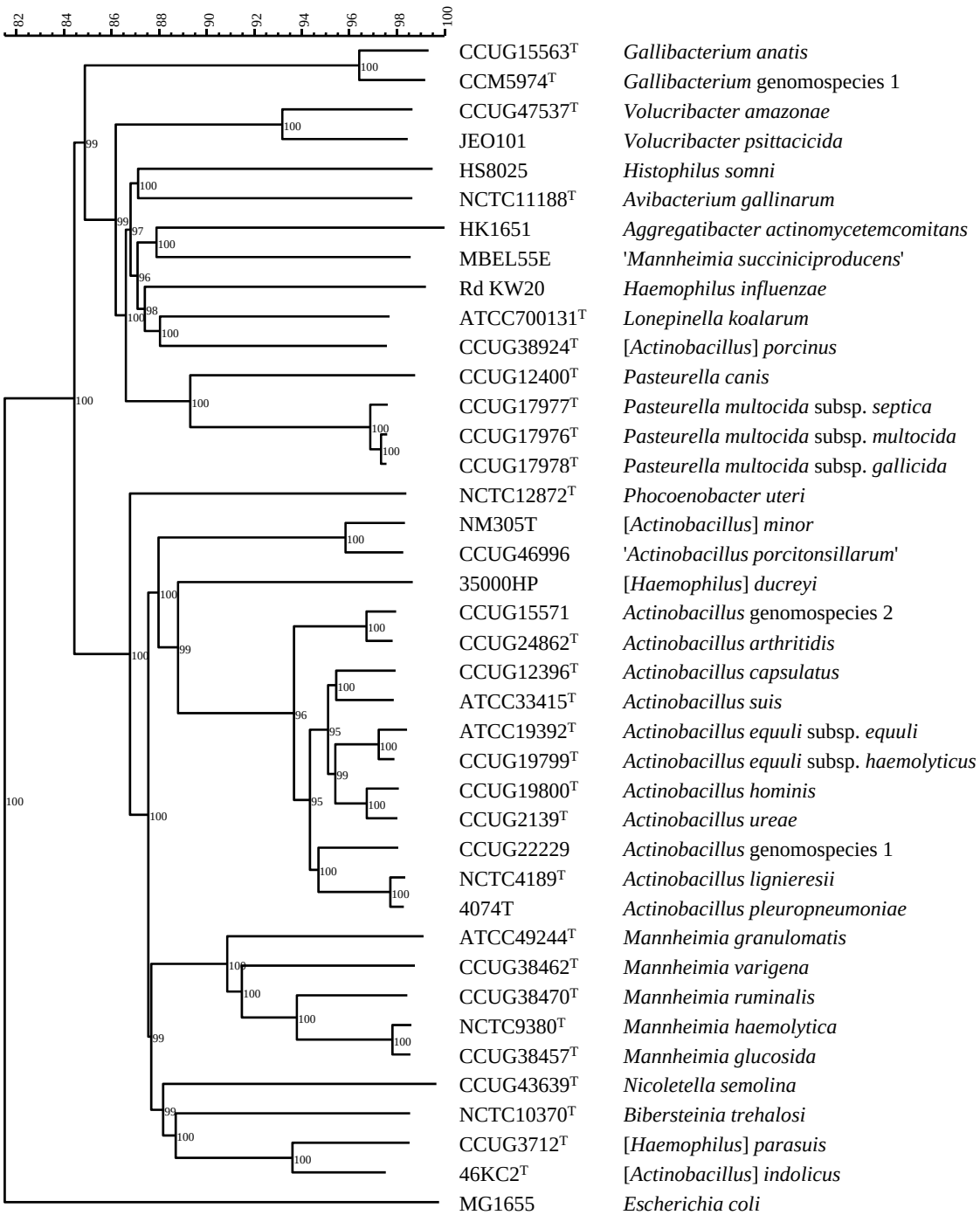


Figure 6. *recN_thdF_ rpoA*