Phylogenetic Analysis of the Family *Moraxellacea*:

Comparison of Phylogenetic Tree Topology and Identification of Individual Species by Primer Design Using Different Molecular Markers

Nhi Luu

Translational Life Science Technology

University of Maryland, Baltimore County

BTEC395: Translational Bioinformatics

Dr. Jeffrey Robinson

November 7th, 2022

Table of Contents

Abstract

The current phylogenetic classification of the family *Moraxellacea* is mainly based on the use of conventional housekeeping gene 16S ribosomal RNA (rRNA) due to its high conversed, and variable regions within their genetic profiles. However, there are other gene markers (*rpoB*, *gyrB*, *16S-23S ITS*) that can also be utilized to determine the phylogenetic relationship, providing positive identification of individual species, similarly to *16S rRNA*. Overall, the constructed phylograms demonstrated that *16S rRNA*, *rpoB*, and *gyrB* can be safely used as molecular markers to investigate the relationship, and distinction between closely related species. On the other hand, *16S-23S ITS* region showed minor discrepancies in tree topology compared to the other phylograms. The utilization of *16S-23S ITS* region in phylogenetic analysis is reliable but preferred to be in conjunction with other established molecular markers. Due to its highly variable areas, the *16S-23S ITS* is a useful gene in creation of primer pairs to distinguish individual species from others in the same family. The other marker genes, *rpoB*, and *gyrB* also had a good number of variable regions, making it easy to design primers for, in comparison to the *16S rRNA* region with low polymorphic profile.

Introduction

Microbiomes, or community of micro-organisms, such as bacteria, fungi, and viruses are found everywhere in different habitat, from environmental soils, to the ocean, in the atmospheres, as well as inside animals and the human bodies. These microbes combined play a vital role in human biological processes and digestive systems, such as breaking down necessary nutrients and producing important vitamins and enzymes (Wang et al., 2017). Although most bacteria are beneficial to the human health, they are also associated with human diseases due to their opportunistic pathogenicity. Over the year, there are more and more bacteria that are being classified from commensals to pathogens, causing infections, and diseases related to liver, respiratory tract, and more. While some opportunistic pathogens are well studied and classified, such as *Staphylococcus aureus*, *Escherichia coli* and *Pseudomonas aeruginosa*, some other species with emerging health importance are not being well addressed until recent years, like the family *Moraxellacea*.

The *Moraxellacea* family belongs in the order *Pseudomonadales*, comprising of three recognized genera which are *Acinetobacter*, *Moraxella*, and *Psychrobacter*, as well as other proposed genera. They are mostly harmless and commensal gram-negative bacteria, having different natural habitats, either residing as part of human microbiota, or free-living in the environmental soils, water, and food (Teixeira & Merquior, 2014). For instance, the soil and water living *Acinetobacter* species, such as *Acinetobacter baumanii* and *Acinetobacter calcoaceticus* is the main cause for infections in the blood, urinary tract, skin, pneumonia and more. Similarly, the *Moraxella* species, especially *Moraxella catarrhalis*, an exclusive human pathogen for respiratory tract infections, and *Moraxella bovis* on fatal cattle diseases. Lastly, the *Psychrobacter* species, living in the ocean, or cold saline environment, are the cause for infections in marine animals and food spoilage (Teixeira & Merquior, 2014).

Due to the difference in habitat among these genera, the species within this family went through a series of changes and reclassification since the early appearance of the type genus *Moraxella* in 1939, proposed by Lwoff (Rossau et al., 1991). In 1971, the *Acinetobacter* genus was classified into the same family with *Moraxella* genus, which was *Neisseriaceae*. In 1991, due to a drastic change in genetic composition within the family, the genera *Acinetobacter* and *Moraxella*, along with *Psychrobacter* (created in 1986) were re-classified into the family *Moraxellacea*e, using DNA-rRNA hybridizations (Rossau et al., 1991). As the concept of genetic marker and polymerase chain reaction progressed, a phylogenetic analysis of the family *Moraxellacea* was published in 1998, using 16S ribosomal RNA (*rRNA*) genetic marker (Pettersson et al., 1998). Thus, *16S rRNA* became a widely used genetic marker for phylogenetic tree constructions, and the current taxonomic classification scheme.

With *Acinetobacter* and *Moraxella* species having similar growth condition and habitat, it was expected for these species to be more closely related to each other than *Psychrobacter*. However, according to results obtained in Welter et al. (2021), *Acinetobacter* species were categorized into its own clade, and *Moraxella* and *Psychrobacter* species were more closely related to each other. Even though the relationships between the genera of family *Moraxellacea* have been studied, the phylogenetic classification still lacks the comparison of *16S rRNA* with other molecular markers. In fact, current analysis of phylogenetic relationship and identification among species mostly utilizes *16S rRNA* genes due to their highly conserved, and variable regions. In contrast, Reller et al. (2007) stated that *16S rRNA* genes were unable to distinguish all species within closely related genomic species, leading to misidentification (Reller et al., 2007; La Scola et al., 2006). Thus, other molecular marker genes with better polymorphic profile, such as the housekeeping genes like *rpoB*, and *gyrB*, or *16S-23S ITS* should be investigated.

RNA polymerase beta subunit (*rpoB*), like its name, is one of the housekeeping genes that encodes for the β subunit in RNA polymerase. Several studies have utilized *rpoB* for identification of species because at least 1 copy of the gene is present in bacterial genome, providing better resolution than *16S rRNA* (Reller et al., 2007). Similarly, gyrase B (*gyrB*), a gene which encodes for the β subunit in DNA gyrase, shows better discrimination among close species with indistinguishable *16S rRNA* sequences (Yamamoto & Harayama, 1996). The operon that codes for ribosomal RNAs in bacteria contains the subunits for 16S, 23S, and 5S genes in that order from 5' to 3' direction, with 2 intergenic spacer regions (ITS) that separate each subunit (Tokajian et al., 2016). The 16S-23S rRNA intergenic transcribed spacer (*16S-23S ITS*) is the first spacing region between 16S and 23S that has the genetic information for $tRNA^{Ile}$ and $tRNA^{Ala}$ (see Supplement Figure 1 for details). Thus, the *16S-23S ITS* is also a potential molecular marker due to its variable content among species within the spacer at the genus level (Volokhov et al., 2012).

This paper aims to compare the phylogenetic relationship using different molecular genetic markers *rpoB*, *gyrB*, and *16S-23S ITS*, with the conventional 16S rDNA in phylogenetic trees construct and analysis for 10 species belong to the family *Moraxellacea* (Supplement Table 1), and *P.aeruginosa* in the same order as an outgroup. Furthermore, specific primer design using the alignments of different molecular markers genes will also be utilized for accurate identification of individual from species within the same family.

Methods

Bacterial Gene Accessions

The nucleotide sequences for 11 species in used this paper were obtained from NCBI GenBank or BioProject using BlastN algorithm. For *16S rRNA* sequences, the partial sequences of this gene were readily available through the RefSeq Targeted Loci Project on bacterial rRNA genes (Accession: PRJNA224725). For *rpoB*, and *gyrB* genes, complete coding sequences (CDS) were either obtained directly from the GenBank's record, or manually extracted from whole genome sequence using gene graphic search tool (Supplement Figure 2). The whole genome shotgun sequences were available through the Refseq Prokaryotic Genome Annotation Project (Accession: PRJNA224116). Similarly, the complete sequences for the gene *16S-23S ITS* were available on GenBank database, except for *Psychrobacter* species. For this genus, manual extraction of the ITS region was done by analyzing on NCBI gene graphic tool and cropped the spacer with partial *16S* and *23S rRNA*). For details on the region, see Supplement Figure 1. All sequences for each gene were retrieved in the forward direction (plus strand), with similar base pairs in length. The name of species, GenBank accession numbers, and extraction region is listed in Supplement Table 1.

Sequence Analysis and Phylogenetic Tree Construction

Nucleotide sequences of the 11 species for each molecular marker gene were aligned using ClustalX2 software (http://www.clustal.org/clustal2/). The alignment process was done automatically with the default settings for multiple alignment mode using complete alignment option. The option included pairwise alignment for sequence comparison, and dendrogram as output (Larkin et al., 2007). The multiple sequence alignments were trimmed with Jalview software (https://www.jalview.org/) for both ends to ensure a better phylogenetic construction (Waterhouse et al., 2009). Four phylogenetic trees for each molecular marker were constructed from the edited alignment and dendrogram output, using both Neighbor Joining (NJ) and Bootstrap algorithms. The

obtained phylogenetic trees were edited with iTOL software (https://itol.embl.de/), an editing tool to generate better representation of species topology, and calculated branch distance. In iTOL, branch distance was automatically calculated, which indicates the genetic change by calculating the nucleotide substitutions per site on aligned sequences (Letunic & Bork, 2021).

Primer Design

For the general purpose of this paper, primer pairs of each molecular marker were only created for the *Acinetobacter baumannii* species, rather than all listed species. Primer sequences were generated through manual selection that utilized several online tools. First, the alignments were observed and color coded by percentage identity option available in Jalview. The primer sequences were then manually selected based on the criteria for polymerase chain reaction (PCR) primer design, listed in Addgene (2019). The selected sequences for the sense strand were converted to their reverse complement, using genscript tool in order to create PCR-ready primers. The reverse sequences were obtained from reverse complement tool to convert the manually selected sequences for the sense strand (GenScript - Reverse Complement, n.d.). The primer statistics of the selected sequence were compared against the Oligo Evaluator tool to ensure the primer pairs have appropriate properties when running in PCR (OligoEvaluatorTM - Sequence Analysis Tool, n.d.). Primers were selected within appropriate range for melting temperature (Tm), %GC content, and 3' self-complementarity at low to moderate using the mentioned tools.

Results and Discussions

Phylogenetic Analysis Using Different Molecular Markers

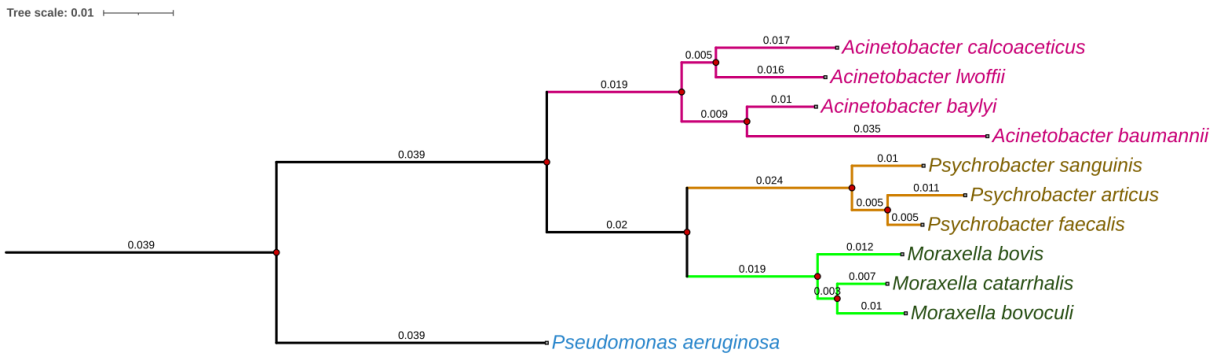


Figure 1. Phylogram of *16S rRNA* genes. Phylogram showing phylogenetic relationship of 10 species in the family *Moraxellacea* using *16S rRNA* genes with P. aeruginosa serves as an outgroup. The numerical values depict branch length based on evolutionary distance (nucleotide substitution per site) with noted distance scale. Different genus on each clade of the phylogram (*Acinetobacter*, *Psychrobacter*, and *Moraxella*) were highlighted accordingly.

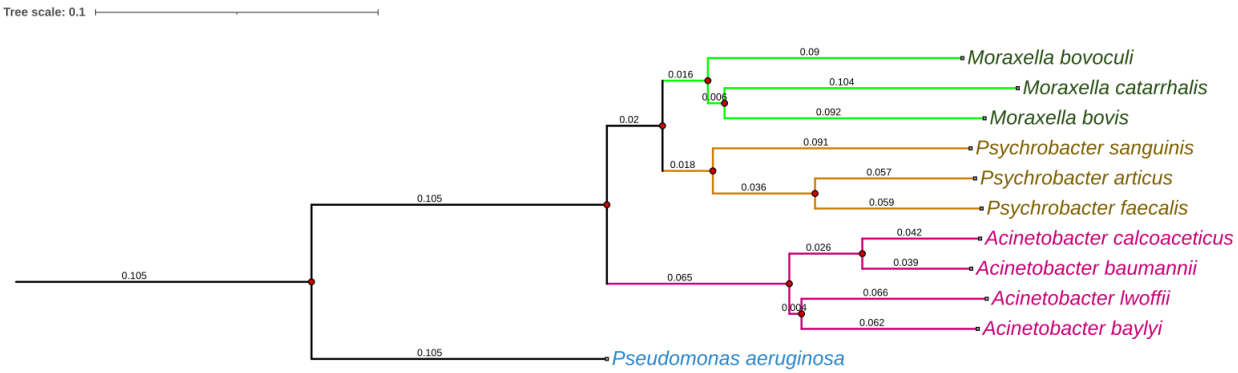


Figure 2. Phylogram of *rpoB* genes. Phylogram showing phylogenetic relationship based on *rpoB* nucleotide sequences for the 10 species of *Moraxella* family, with *P. aeruginosa* as outgroup. The numerical values depict branch length based on evolutionary distance (nucleotide substitution per site) with noted distance scale. Different genus on each clade of the phylogram (*Acinetobacter*, *Psychrobacter*, and *Moraxella*) were highlighted accordingly.

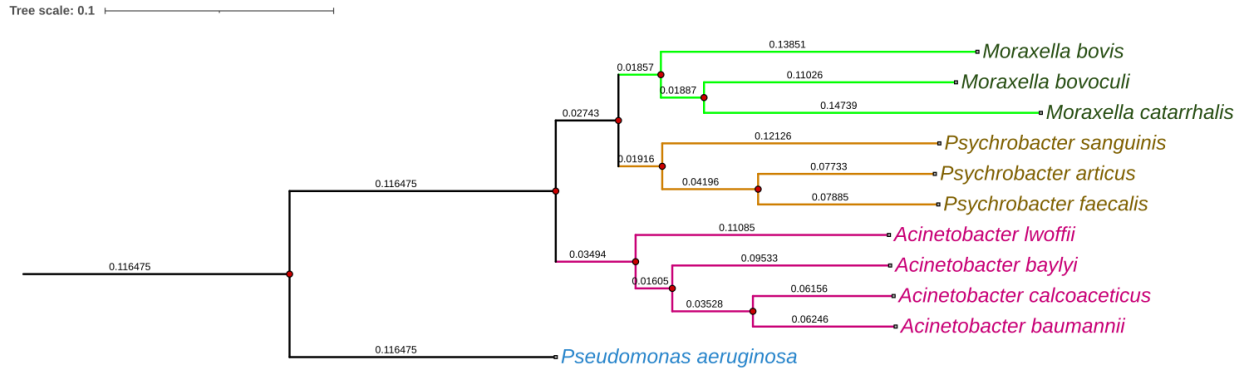


Figure 3. Phylogram of *gyrB* genes. Phylogram showing phylogenetic relationship of the subjected species in the *Moraxellacea* family based on *gyrB* nucleotide sequences. The numerical values depict branch length based on evolutionary distance (nucleotide substitution per site) with noted distance scale. Different genus on each clade of the phylogram (*Acinetobacter*, *Psychrobacter*, and *Moraxella*) were highlighted accordingly.

Housekeeping genes: *16S rRNA*, *rpoB*, *gyrB*

The phylogenetic trees were constructed using 4 molecular markers *16S rRNA*, *rpoB*, *gyrB*, and *16S-23S ITS* sequences (figure 1, 2, 3, and 4, respectively) for 10 species in the *Moraxella* family with *P. aeruginosa* as an outgroup. The phylograms for *16S rRNA*, *rpoB*, *gyrB* showed better topology compared to *16S-23S ITS*. In figure 1, 2, and 3, the 10 species of the family *Moraxellacea* were grouped together into 1 branch separated from the outgroup *P. aeruginosa*. From the *Moraxella* family branch, the first node separated into 2 sister clades, with all

listed *Acinetobacter* species in 1 clade, and the other 2 genera into the other clade. The other clade further branched into the *Moraxella* and *Psychrobacter* clades. Thus, the 3 figures clearly showed that *Moraxella* and *Psychrobacter* species are more closely related to each other than *Acinetobacter* species. This was a similar result to the phylogenetic tree in the study conducted by Welter at al. (2021), where each genus was grouped together in its own clade, with *Acinetobacter* species in one clade, *Moraxella* and *Psychrobacter* species in their own clade that shared a common ancestor.

The *Psychrobacter* clade (yellow-coded) had the most stable node among all the genes, where *P. articus* and *P. faecalis* had a closer evolutionary relationship than *P. sanguinis*. For the other 2 clades, the nodes were rather changeable with branching scheme that grouped sister species differently. For instance, figure 1 (*16S rRNA*), showed similar topology with figure 3 (*gyrB*) for the *Moraxella* clade which *M. catarrhalis* and *M. bovoculi* are sister species, and *M. bovis* at basal. Nevertheless, this relationship was different from the phylogeny construct in Welter et al. (2021), where *M. bovis* and *M. bovoculi* are sister species. This sister relationship was also different from what was observed in figure 2 (*rpoB*). For *Acinetobacter* clade (pink-coded), the phylograms of these 3 genes showed different topology for sister species and basal members. However, figure 2 and 3 showed that *A. calcoaceticus* and *A. baumanii* were sister species, and this relationship was also observed in several research articles (Welter et al., 2021; Yamamoto & Harayama, 1996; La Scola et al., 2006).

Among these 3 figures, *16S rRNA* had the shorter branch lengths compared to the other 2 genes. As branch distance described the nucleotide substitutions per site, it showed that *rpoB*, and *gyrB* have more genetic diversity in its sequences, compared to *16S rRNA*. Nevertheless, any of these genes could be safely used in phylogenetic analysis to inspect the evolutionary relationship of bacterial species, with emphasis on *rpoB* or *gyrB* genes for more closely related species.

16S-23S ITS region

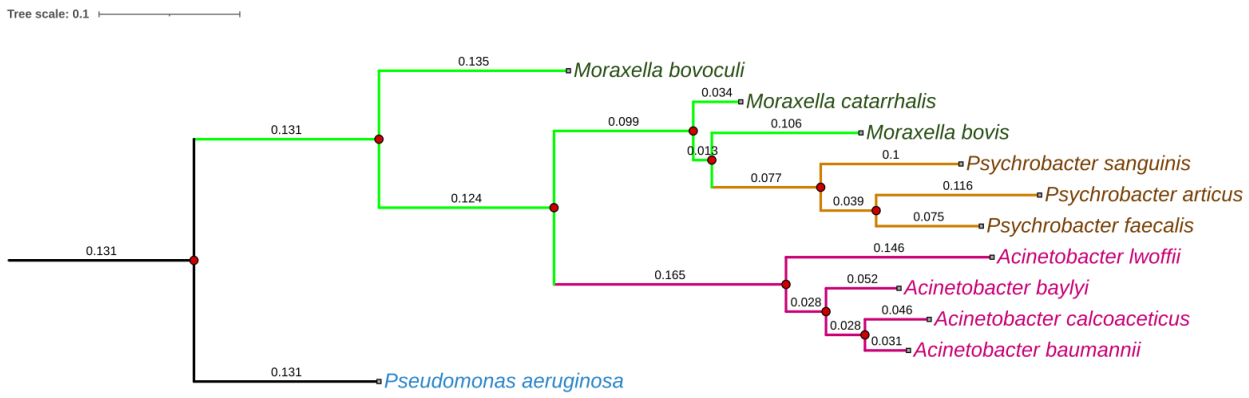


Figure 4. Phylogram of *16S-23S ITS* genes. Phylogram showing phylogenetic relationship between species of family *Moraxellacea*e, using *16S-23S ITS* region of nucleotide sequences. The numerical values depict branch length based on evolutionary distance with noted distance scale. Different genus on each clade of the phylogram (*Acinetobacter*, *Psychrobacter*, and *Moraxella*) were highlighted accordingly.

The use of ITS region on phylogenetic analysis was not as established as the 3 housekeeping genes previously described above. Although there has been increased number of studies to utilize this region as molecular markers, there was still limited number of sequences available on NCBI for various species in the family *Moraxellacea*. Thus, the *16S-23S ITS* sequences obtained for this paper was selected manually by cropping the middle region with partial end of 16S, and partial start of 23S (Supplement Figure 1). Due to this manual extraction, this could result in slight alignment errors, causing phylogenetic discrepancies compared to the other phylograms discussed above. There was also a lot of indels embedded in the aligned sequences, but this can be a good region for positive identification of individual species, thus the indels were kept for next steps. Regardless, the phylogeny construct for this gene still showed some similarities with the phylograms constructed from other molecular markers.

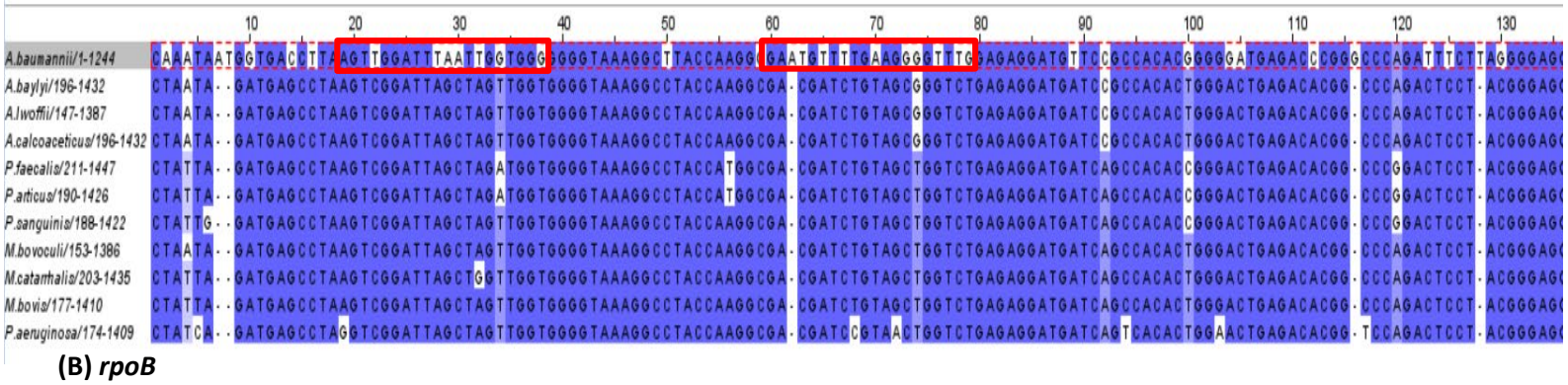
Compared to the previous figures, *Psychrobacter* and *Acinetobacter* species were grouped in their own clade (yellow, and pink, respectively), with *Psychrobacter* clade had the same branching with previous figures. The species of *Acinetobacter* clade displayed more nodes branching compared to the other housekeeping genes, which indicated more diversity in the gene sequences. However, the branch for sister species *A. calcoaceticus* and *A. baumannii* was still observed in figure 4, similarly to figure 2 and 3. The most distinctive difference from figure 4 to the other figures was that the *Moraxella* species did not form distinct clade. In figure 4, *M. bovoculi* was the basal member for the other 9 species, *M. catarhalis*, and *M. bovis* further branched out with separated nodes, and served as basal members for *Psychrobacter* clade. In fact, Welter at al. (2021) exhibited similar relationship where a few *Moraxella* species (not in this paper) were inside the *Psychrobacter* clade as basal members. Nevertheless, from previous phylogenetic scheme, *Acinetobacter* clade should share the same common ancestor with its sister clade consisting of the other 2 genera, but it was not observed in figure 4. Thus, among the 4 molecular markers, *16S-23S ITS* gene clearly encompassed more genetic diversity within species. Although the intergenic region appeared to be an acceptable marker, phylogenetic analysis of this gene should be accompanied by more published marker genes, like *16S rRNA* or *rpoB*.

Primer pairs design

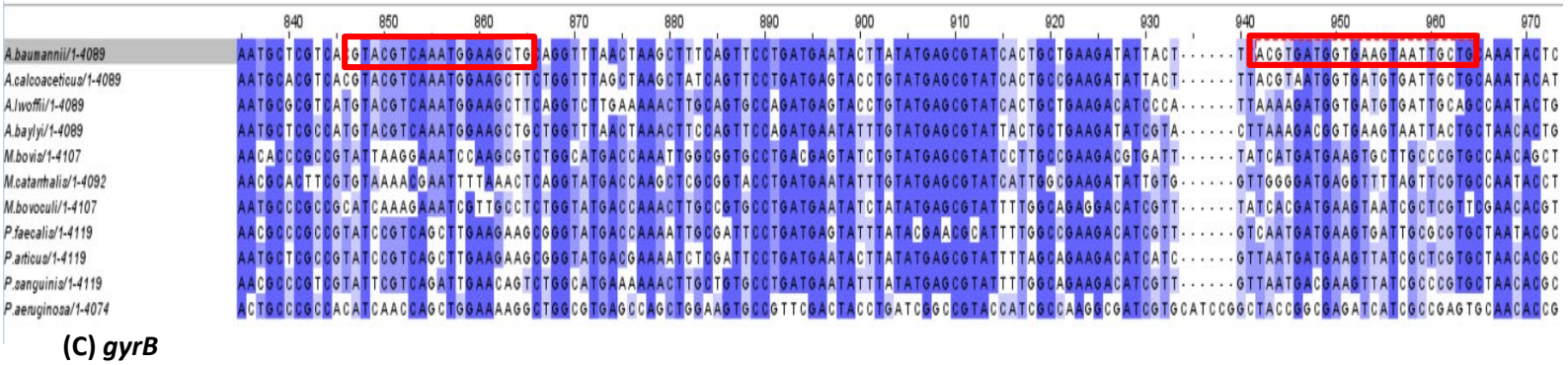
Gene/Location	Sequence (5'-3')	Length (b.p)	Tm (˚C)	GC content (%)
16S rRNA				
Forward	AGTTGGATTTAATTGGTGGG	20	60.1	40
Reverse	CAAACCCCTTCAAAACATTC	20	60.8	40
rpoB				
Forward	GTACGTCAAATGGAAGCTG	19	58.3	47.4
Reverse	CAGCAATTACTTCACCATC	19	55.5	42.1
gyrB				
Forward	AATATCCATTGCGTGTGATC	20	59.8	40
Reverse	GCGTGCTAGAATTTCAACAT	20	59.4	40
16S-23S ITS				
Forward	TCACGGTAATTAGTGTGATC	20	54.9	40
Reverse	GCTTAACCTAAACTCTTGAG	20	52.5	40

Table 1. Potential PCR primers for positive identification of *Acinetobacter baumannii* among closely related species in the family *Moraxellacea*. Primer-pairs were manually created based on the criteria listed in Addgene (2019) for PCR primers. The listed sequences are from 5' to 3', with forward sequence as primer or the antisense strand, and reverse for the sense strand.

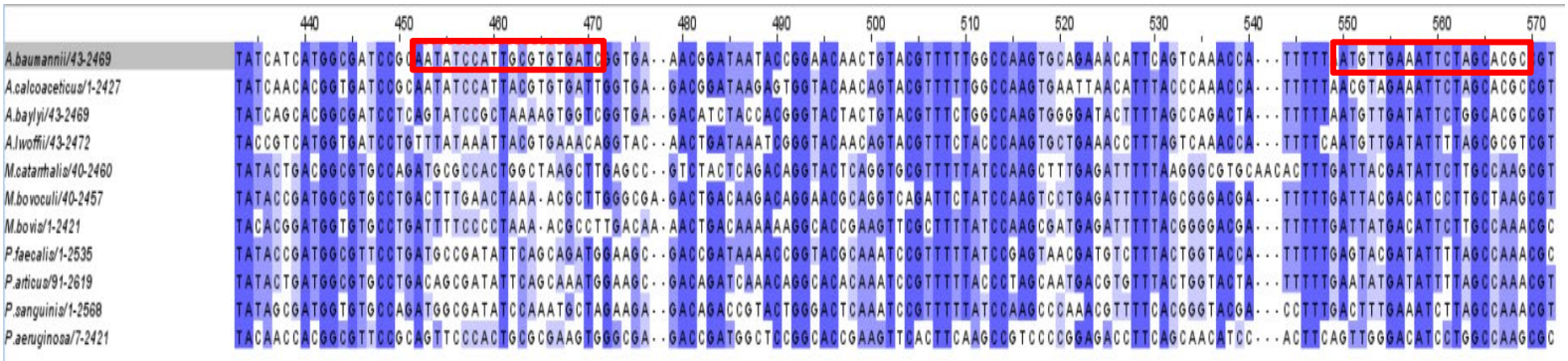
(A) *16S rRNA*



(B) *rpoB*



(C) *gyrB*



(D) *16S-23S ITS*

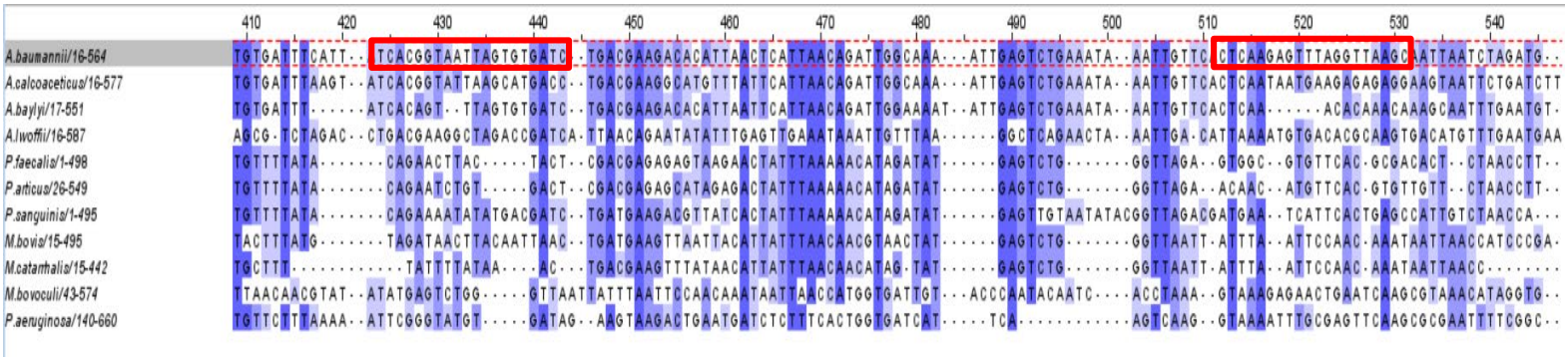


Figure 5. Multiple alignments of different molecular markers for 10 species of family *Moraxellacea*, and outgroup *P. aeruginosa*. The color-coded region was generated in Jalview, with dark purple indicated areas with high percent identity, light purple as moderate, and white as low to no identical bases in that site. Solid outlines in red are the primer-pairs for forward, and reverse sequence specific to *A. baumannii*, respectively. For the reverse sequence, reverse complement conversion must be done to generate the PCR-ready primers sequence listed in Table 1.

Following the bioinformatics pipelines, the final part of this phylogenetic analysis aimed to create potential primers to accurately distinguish individual species from species in the same family *Moraxellacea*. Thus, potential primers design was attempted for individual A. baumannii species. Table 1 listed the potential primers that could be used for PCR amplification that was specific to *A. baumannii*. Figure 5 showed the multiple alignments of each marker gene, with the listed forward and reverse primers unique to *A. baumanii* outlined in red. The manual selection was based on the percent identity among the alignments (color coded in purple, and white).

It can be observed clearly that multiple alignments for *16S rRNA* (Fig. 5A) displayed higher percent identity (dark purple) compared to the *rpoB*, *gyrB*, and *16S-23S ITS* genes. It was also relatively difficult to select primer-pairs for *16S rRNA* due to the highly conserved regions with very high or very low GC content in variable regions. Furthermore, the *16S rRNA* alignment for *A. baumannii* contained several base differences so it was not as difficult to select primers in comparison to *Moraxella* and *Psychrobacter* species. Among these alignments, Fig. 5D (*16S-23S ITS*) displayed most variable regions with many sites that could be selected for potential primers, compared to *16S rRNA* alignments. Nevertheless, actual PCR analysis should be carried out to confirm the success of hypothetical primer-pairs design, whether the primers were viable, providing accurate and specific identification of the species

References

Addgene. (2019). Addgene: Protocol - How to Design Primers. Addgene.org. https://www.addgene.org/protocols/primer-design/

GenScript - Reverse Complement. (n.d.). GenScript. https://www.genscript.com/sms2/rev_comp.html

La Scola, B., Gundi, V. A. K. B., Khamis, A., & Raoult, D. (2006). Sequencing of the *rpoB* Gene and Flanking Spacers for Molecular Identification of *Acinetobacter* Species. *Journal of Clinical Microbiology*, *44*(3), 827–832. https://doi.org/10.1128/jcm.44.3.827-832.2006

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)*, *23*(21), 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Letunic, I., & Bork, P. (2021). Interactive Tree of Life (iTOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Research*, *49*(W1), W293–W296. https://doi.org/10.1093/nar/gkab301

OligoEvaluatorTM - Sequence Analysis Tool. (n.d.). Sigma-Aldrich Co. LLC. http://www.oligoevaluator.com/OligoCalcServlet

Pettersson, B., Kodjo, A., Ronaghi, M., Uhlen, M., & Tonjum, T. (1998). Phylogeny of The Family *Moraxellacea*e by 16S Rdna Sequence Analysis, With Special Emphasis On Differentiation of *Moraxella* Species. *International Journal of Systematic Bacteriology*, *48*(1), 75–89. https://doi.org/10.1099/00207713-48-1-75

Reller, L. B., Weinstein, M. P., & Petti, C. A. (2007). Detection and Identification of Microorganisms by Gene Amplification and Sequencing. *Clinical Infectious Diseases*, *44*(8), 1108–1114. https://doi.org/10.1086/512818

Rossau, R., Van Landschoot, A., Gillis, M., & De Ley, J. (1991). Taxonomy of *Moraxellacea*e fam. nov., a New Bacterial Family to Accommodate the Genera *Moraxella*, *Acinetobacter*, and *Psychrobacter* and Related Organisms. *International Journal of Systematic Bacteriology*, *41*(2), 310–319. https://doi.org/10.1099/00207713-41-2-310

Teixeira, L. M., & Merquior, V. L. C. (2014). The Family *Moraxellacea*e. *The Prokaryotes*, 443–476. https://doi.org/10.1007/978-3-642-38922-1_245

Tokajian, S., Issa, N., Salloum, T., Ibrahim, J., & Farah, M. (2016). 16S–23S rRNA Gene Intergenic Spacer Region Variability Helps Resolve Closely Related Sphingomonads. Frontiers in Microbiology, 7. https://doi.org/10.3389/fmicb.2016.00149

Volokhov, D. V., Simonyan, V., Davidson, M. K., & Chizhikov, V. E. (2012). RNA Polymerase Beta Subunit (*rpoB*) Gene and The 16S–23S rRNA Intergenic Transcribed Spacer Region (ITS) As Complementary Molecular Markers in Addition to The *16S rRNA* Gene for Phylogenetic Analysis and Identification of the Species of the Family *Mycoplasmataceae*. *Molecular Phylogenetics and Evolution*, *62*(1), 515–528. https://doi.org/10.1016/j.ympev.2011.11.002

Wang, B., Yao, M., Lv, L., Ling, Z., & Li, L. (2017). The Human Microbiota in Health and Disease. Engineering, 3(1), 71–82. https://doi.org/10.1016/j.eng.2017.01.008

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2--A Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics*, *25*(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Welter, D. K., Ruaud, A., Henseler, Z. M., De Jong, H. N., van Coeverden de Groot, P., Michaux, J., Gormezano, L., Waters, J. L., Youngblut, N. D., & Ley, R. E. (2021). Free-Living, Psychrotrophic Bacteria of the Genus *Psychrobacter* Are Descendants of Pathobionts. *MSystems*, *6*(2). https://doi.org/10.1128/msystems.00258-21

Yamamoto, S., & Harayama, S. (1996). Phylogenetic Analysis of *Acinetobacter* Strains Based on the Nucleotide Sequences of *gyrB* Genes and on the Amino Acid Sequences of Their Products. *International Journal of Systematic Bacteriology*, *46*(2), 506–511. https://doi.org/10.1099/00207713-46-2-506

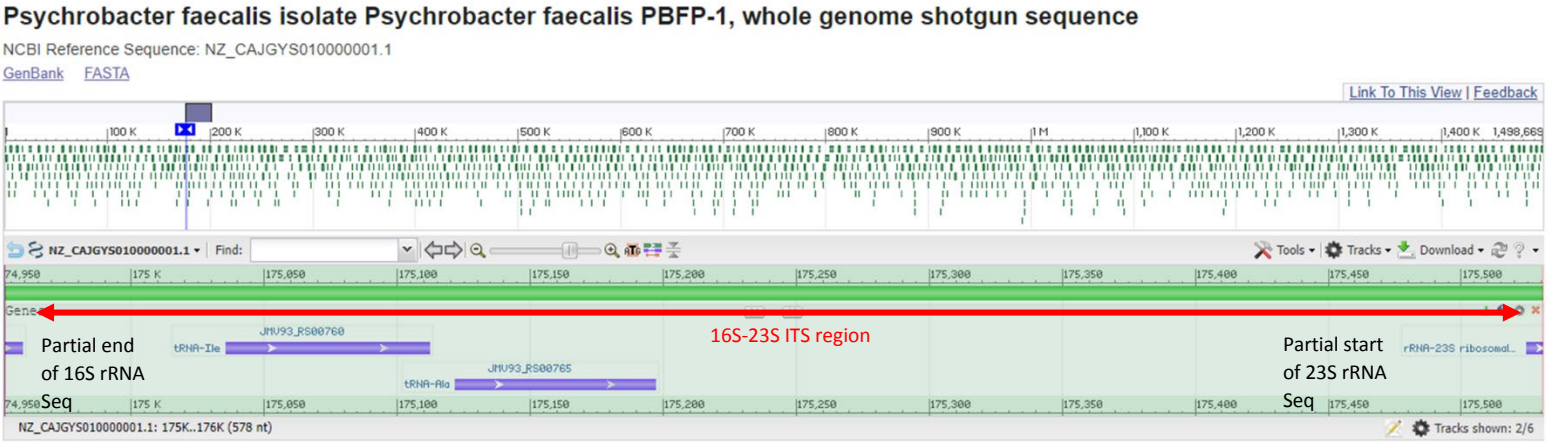
Supplemental Data

Supplement Table 1. Nucleotide sequence accession numbers of the 4 different genes for all 11 species used in this paper. Sequences were obtained from NCBI's GenBank or BioProject. The FASTA sequences can be looked up in NCBI database using the listed accession numbers. If the nucleotide sequences were not readily available on records, extraction regions from whole genome sequence were listed in the format from (start) base position to (end) base position for the extracted gene of the sequence. If the extraction region is noted as "complement", conversion was done to flip the template strand from negative to positive direction (plus strand). See the methods section for more details on manual sequence extraction. *16S* rRNA Accessions retrieved from RefSeq Targeted Loci Project on bacterial rRNA genes (Accession: PRJNA224725). Whole genome sequence retrieved from Refseq Prokaryotic Genome Annotation Project (Accession: PRJNA224116)

Species/Gene	Accession no.	Extraction region (from..to)
Acinetobacter baumannii		
16S rRNA	LN611374.1	
16S-23S rRNA Intergenic Spacer	AY601824	
rpoB	NZ_CP015121	REGION: complement(3634487..3638575)
gyrB	NZ_CP015121	REGION: 3794..6262
Acinetobacter calcoaceticus		
16S rRNA	FM210755.1	
16S-23S rRNA Intergenic Spacer	AY601821	
rpoB	DQ207474	
gyrB	NZ_CP070518	REGION: 1944402..1946870
Acinetobacter lwoffii		
16S rRNA	LS483013	
16S-23S rRNA Intergenic Spacer	AY601835	
rpoB	LC102685	
gyrB	LC102629.1	
Acinetobacter baylyi		
16S rRNA	EU604245	
16S-23S rRNA Intergenic Spacer	EF591586	
rpoB	LC102672	
gyrB	LC102616	
Moraxella bovis		
16S rRNA	AF005183.1	
16S-23S rRNA Intergenic Spacer	EU014535	
rpoB	CP030241	REGION: 2129624..2133730
gyrB	CP030241	REGION: complement(2787687..2790122)
Moraxella catarrhalis		
16S rRNA	NR_118775.1	
16S-23S rRNA Intergenic Spacer	EU014604	
rpoB	CP008804	REGION: complement(1766986..1771077)
gyrB	LC217749	
Moraxella bovoculi		
16S rRNA	KF637298	
16S-23S rRNA Intergenic Spacer	MN840837	
rpoB	CP011381.2	
gyrB	CP011381	
Psychrobacter faecalis		
16S rRNA	AJ421528	
16S-23S rRNA Intergenic Spacer	NZ_CAJGYS010000001	REGION: 174953..175530
rpoB	NZ_CAJGYS010000001	REGION: 267490..271608
gyrB	NZ_CAJGYS010000001	REGION: complement(1024273..1026807)
Psychrobacter sanguinis		

16S rRNA	NR_117833	
16S-23S rRNA Intergenic Spacer	CP085990	REGION: 1204154..1204732
rpoB	NZ_CAJHAX010000036	REGION: complement(19356..23474)
gyrB	NZ_CAJHAX010000024	REGION: 29307..31961
Psychrobacter articus		
16S rRNA	NR_042907	
16S-23S rRNA Intergenic Spacer	CP000082	REGION: 662076..662691
rpoB	NZ_CAJGZE010000062	REGION: complement(5394..9512)
gyrB	CP000082	REGION: 5176..7794
Pseudomonas aeruginosa		
16S rRNA	FJ972538	
16S-23S rRNA Intergenic Spacer	KX641439	
rpoB	CP007224	REGION: 713437..717510
gyrB	CP007224	REGION: 4275..6695

Supplement Figure 1. Sample NCBI Gene Graphics. Example of *16S-23S ITS* region that was selected manually for *P. faecalis* from whole genome sequence.



Supplement Figure 2. Sample NCBI Gene Graphics. Example of manual sequence extraction for *rpoB* genes from whole genome sequence on *M. bovis*. As a side note, all strand with negative direction were converted to positive.