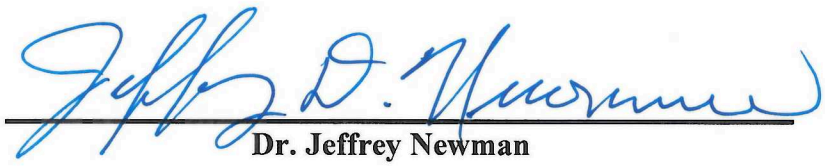
Novel Moraxella species isolated from clinical cases of ocular infection

Presented to the Faculty of Lycoming College in partial fulfillment of the requirements for Departmental Honors in Biology

By
Emily Konopka
Lycoming College
April 30, 2021

Approved by:



Dr. Jeffrey Newman

Dr. David Andrew

Dr. Jeremy Ramsey

Dr. Sarah Holstein

Table of Contents

Abstract

In a previous study to more precisely identify ocular isolates of *Moraxella*, isolates from cases of keratitis, conjunctivitis, and endophthalmitis were identified. In this present study, four of the Gram-negative diplobacilli isolates from patients affected with keratitis (inflammation of the cornea) at the University of Pittsburgh Medical Center (UPMC), Pittsburgh, PA, between 1997-2017 are described. These organisms had partial 16SrRNA gene sequences that clustered into one group upon initial evaluation. Once full-length 16S rRNA gene sequences were obtained, sequence comparisons were performed with BLAST, and creation of a neighbor-joining phylogenetic tree on MEGA7 revealed the isolates were most closely associated with *Moraxella lacunata* CCUG 57757A, *Moraxella lacunata* CCUG 4441^T, and *Moraxella equi* NCTC 11012^T. The genomes of the UPMC isolates were sequenced with Illumina NextSeq 550 and uploaded to the RAST website for annotation. Further comparative genomics were performed using metrics including rpoB sequence similarity, estimated DNA-DNA Hybridization (eDDH), Average Amino Acid Identity (AAI), and Average Nucleotide Identity (ANI) to identify the closest relatives and compare them at the species and genus levels. Based on ANI and AAI analyses, we propose the isolates K1664, K127, K1630, and K2450 represent four strains of a novel *Moraxella* species with K1664^T as the type strain. ANI values were compared between K1664^T and the three other isolates, K127 (98.23%), K1630 (99.78%), and K2450 (99.77%), all of which had ANI values greater than the 95% threshold, indicating they should be classified within the same species. ANI comparison also revealed K1664^T was highly similar to *M. lacunata* CCUG 57757A (99.53%), but only 94.81% similar to *M. lacunata* CCUG 4441^T, suggesting that CCUG 57757A should be reclassified into the novel species with K1664^T. The name *Moraxella kowalskii* sp. nov. is proposed for the new species.

Background

Moraxella is a gram negative, aerobic, diplococcus, ocular bacterial pathogen that has been isolated from cases of conjunctivitis, keratitis, and endophthalmitis. Many species are parasites of the mucous membranes of humans and have been linked to a number of infections in mammals. *Moraxella* was first isolated in 1896 and 1897 by Morax and Axenfield from patients with angular blepharitis and became known as *M. lacunata* (Whitcher et al. 2006). Angular blepharitis is one of the most common ocular conditions and manifests as inflammation of the eyelids and in rare cases may lead to altered eyelid morphology and visual defects resulting from keratopathy and corneal ulceration (Putnam 2016). *M. lacunata* is a commensal bacterium of human epithelia that is rarely associated with disease, but has been isolated from cases of conjunctivitis (Embers et al. 2011). Most cases of bacterial conjunctivitis are self-limiting and require no additional treatment, but *Moraxella* conjunctivitis may persist for weeks to months and may be misdiagnosed as inclusion conjunctivitis caused by *Chlamydia* (Epling 2012; Dawson et al. 1983). Bacterial keratitis is a serious ocular disease that *Moraxella* can cause in patients who are immunocompromised and may lead to loss of vision in severe cases (Durrani et al. 2019). Endophthalmitis is another infection that is caused by *Moraxella* characterized by severe inflammation of the interior of the eye. Despite the few species that have been characterized, the breadth of *Moraxella* variants and host range is quite extensive and may grow with further research (Embers et al. 2011).

Moraxella species are often difficult to grow in the laboratory due to their fastidious growth requirements, but they do grow on blood agar and tryptic soy broth agar (TSBA). In a clinical setting, difficulties in classifying *Moraxella* species based on phenotypic analyses often

results in general identification in the genus *Moraxella*. Genetic relationships have been studied extensively by genetic transformation, DNA-DNA hybridization, and 16S rDNA analysis, although the classification system for *Moraxella* species is still incomplete and evolving (Buchman et al. 1991). In the most recent publication analyzing *M. nonliquefaciens* and *M. osloensis*, the objective was to precisely identify ocular isolates of *Moraxella* to include species level differentiation (LaCroce et al. 2019). This study identified *Moraxella* from keratitis, conjunctivitis, and endophthalmitis specimens using DNA sequencing with vancomycin susceptibility, MALDI-TOF MS, and Biolog GenIII plates. Of these isolates, four were unnamed strains designated K127, K1630, K1664, and K2450. After partial DNA sequence analysis, six closely related reference species were identified. Neighbor-joining trees were generated from partial 16s rRNA sequences of the *Moraxella* isolates (LaCroce et al. 2019). From these trees, *M. bovis*, *M. bovoculi*, *M. caprae*, *M. equi*, and *M. lacunata* were identified as the most closely related species of the strains. Based on the branching pattern in the tree, the four novel isolates were suspected to be novel, but further analysis was needed to support this hypothesis.

Classical methods including 16S rRNA sequence similarity have been previously used for taxonomic classifications of novel species. However, because 16S rRNA is highly conserved, it only offers low resolution for classification at the species level. Therefore, in combination with 16S rRNA sequence similarity, we employed higher-resolution comparative genomic metrics for the classification of these clinically important *Moraxella* isolates. These metrics include estimated DNA-DNA Hybridization (eDDH), Average Nucleotide Identity (ANI), and Average Amino Acid Identity (AAI) for reliable taxonomic comparison. Sequence comparison of the β-subunit of bacterial RNA polymerase (rpoB) is also used for identification of the closest relative because it is less highly conserved than the 16S rRNA gene and therefore offers higher

resolution. The objective of the present study was to apply these metrics to classify these isolates, and it was concluded that they represent a novel *Moraxella* species with K1664^T as the type strain. A type strain functions as the representative of a species and is the strain used when the species is first described.

Methods

Isolation and Ecology

Moraxella sp. K127, *M*. sp. K1664, M. sp. K2450, and M. sp. K1630 are four of nine total unnamed isolates recorded between 1993 and 2017 from patients affected by keratitis at the University of Pittsburgh Medical Center (UPMC) in Pittsburgh, PA. The isolates were initially collected and cultured as described in LaCroce et al. 2019. Once these isolates were obtained from UPMC, they were routinely grown on tryptic soy broth agar (TSBA), Columbia Blood agar, or BUG+blood agar (Biolog Universal Growth agar with 5% sheep blood) and grown at 37^{o}C for 48 hours. Well isolated colonies were selected and re-streaked to ensure culture purity. The pure cultures were used to create permanent stocks maintained in Microbank bead vials (ProLab Diagnostics) at -80^{o}C.

16S rRNA PCR and Gel Electrophoresis

To confirm the identity of novel isolates, the 16S rRNA gene was amplified by polymerase chain reaction (PCR) and the DNA sequence was analyzed. The universal primers used for PCR were 27f (5' – AGAGTTTGATCMTGGCTCAG) and 1492r (5'-TACGGYTA CCTTGTTACGACTT-3'). Genomic DNA was isolated from cultures grown for 48 hours at 37^{o}C using the Qiagen Blood and Tissue Kit and quantified using a Qubit fluorometer

(ThermoFisher). Primers 27f and 1492r described by Lane (1991) were used to amplify the 16S rRNA gene via PCR of strains K1664^T, K127, K1630, and K2450, as described previously (Strahan et al., 2011).

PCR products were then analyzed on 1% agarose gel (1x Tris acetate-EDTA (TAE) buffer and 75 ng/mL ethidium bromide). The gel was run for approximately 30 minutes at 100 volts. DNA concentrations for each sample were estimated by comparing each PCR product to the brightness of bands in the 2 Log Ladder (New England Biolabs). PCR products were purified with ExoSAP (ThermoFisher) and sequenced with primers 27f and 1492r (Lane 1991) by Genewiz (Plainfield NJ USA) by DNA Sanger sequencing with funding from the Joanne and Arthur Haberberger Fellowship Program and Pennsylvania Academy of Science Undergraduate Research Fellowship. Reads were assembled using the CAP3 Contig Assembly Program available at http://doua.prabi.fr/software/cap3 (Huang & Madan, 1999). The almost full-length 16S rRNA gene sequences was used during an initial BLAST search (NCBI) (Zhang et al. 2000) against RefSeq genomes,16s rRNA and non-redundant nucleotide databases to identify 16s rRNA sequences of closely related species, all of which were <99.5% similar.

Genome Sequencing, Assembly, and GenBank Upload

The genomes of K1664^T, K127, K1630, and K2450 had not been sequenced previously. Whole genome shotgun sequencing using the Illumina NextSeq 550 (2x150bp) platform was performed by the University of Pittsburgh Microbial Genome Sequencing Center (MiGS) with funding from the Joanne and Arthur Haberberger Fellowship Program and Pennsylvania Academy of Science Undergraduate Research Fellowship. Genome sequences were assembled using SPAdes in PATRIC (Davic et al. 2019) and sequencing statistics were evaluated using the

online CheckM bioinformatics tool to ensure good-quality sequences (Parks et al. 2015). The assembly was annotated with the Rapid Annotation using Subsystems Technology (RAST) web service http://rast.nmpdr.org/ (Aziz et al. 2008; Overbeek et al., 2014). After assembly and annotation, the genomes were deposited to GenBank and sequencing reads were uploaded to the Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra).

Construction of Overall Genome Relatedness Indices

After the genomes were sequenced, the complete 16s rRNA sequences of the four strains were retrieved using the SEED viewer (Overbeek et al. 2014) and BLAST searches were performed against the RefSeq genome sequence database and the 16S ribosomal RNA sequence database for Bacteria and Archaea type strains. This was performed to identify the closest relatives based on full-length 16S rRNA sequence similarity. The most similar sequences were downloaded and MEGA 7 was used to align them and create a neighbor-joining phylogenetic tree (Kumar et al. 2016).

The RNA polymerase beta-subunit (rpoB) genes of the four strains were also retrieved in the SEED viewer and used in NCBI BLAST searches against the RefSeq database to identify similar sequences for phylogenetic analysis. Phylogenic comparisons with the rpoB gene provide higher resolution when identifying closest relatives as it is less-highly conserved than the 16s rRNA sequence. The most similar sequences were downloaded, aligned, and used to make a maximum-likelihood tree in MEGA7.

For eDDH calculations, the Genome-to-Genome Distance Calculator from the Leibniz Institute (Meier-Kolthoff et al. 2013) was used. FASTA files for the genome of K1664^{T} and the closest relatives were uploaded to the online tool and resulting eDDH values were retrieved.

Whole genome sequence-based comparison metrics were calculated to compare K1664^T to the closest relatives on the species and genus level. The assembled genome was annotated with the Rapid Annotation with Subsystem Technology (RAST) website (Aziz et al. 2008) for data analysis, though the annotation available online at NCBI was produced by the NCBI Prokaryotic Genome Annotation Pipeline after genome sequences were deposited. Overall genome relatedness indices (OGRI) (Chun et al. 2014) were calculated relative to the genomes of *M. lacunata* CCUG 57757A, *M. lacunata* CCUG 4441^T, *M. equi* NCTC 11012^T, *M. bovis* CCUG 2133^T, *M. caprae* NCTC 12877^T, and *M. nonliquefaciens* CCUG 348^T available in NCBI (RefSeq). Average nucleotide identity (ANI) values were calculated using OrthoANI Tool (Lee et al. 2016; Yoon et al.2017), Average amino acid identity (AAI) values were determined from the bidirectional best hits (BBH) identified using the SEED Viewer Sequence-Based Comparison Tool (SVSBCT) and a custom calculator (http://lycofs01.lycoming.edu/~newman/AAI.html).

Construction of Venn Diagram

A Venn diagram was generated to display the number of unique and shared genes among five organisms. Two Venn diagrams were constructed, one comparing the novel isolates to *M. lacunata* DSM 18052^T, and another comparing *M. equi* DSM 18027^T, *M. lacunata* DSM 18052^T, *M. caprae* DSM 19149^T, and *M. bovis* DSM 6328^T to *M. sp*. K1664^T. To generate these diagrams, reciprocal sequence-based comparisons were performed on RAST to obtain shared and unique bi-direction best hits. Each .tsv file generated was downloaded and the data was transferred to the Venn Diagram Data Generator available at http://lycofs01.lycoming.edu/

~newman/CurrentResearch.html. Data generated in the "Gene Counts" tab for each of the five organisms was then pasted in the Venn Diagram Output Converter file to produce gene counts that were then transferred to the Venn Diagram template file.

Genome Features

Following annotation with RAST, the protein-coding sequences (CDS) of K1664^T, K127, K1630, and K2450 were compared with those from the closest related type species, *M. lacunata* CCUG 4441^T. These CDS with bidirectional best hits were identified by Seed Viewer Sequence-Based Comparison Tool and were used to generate a Venn diagram using custom calculators available at http://lycofs01.lycoming.edu/~newman/CurrentResearch.html.

Phenotypic Tests

Due to the fastidious growth of *M. lacunata* DSM 18052^T, typical tube and plate growth to test for carbon-utilization and enzymatic phenotypes of the organisms were not performed. Unless stated otherwise, phenotypic tests were performed as described in Reddy et al. 2007. API ZYM and API 20NE test strips (bioMérieux) and Biolog GenIII 96 well plates (Biolog) were used according to the manufacturers' instructions. Strains K1664^T, K127, K1630, and K2450 were routinely grown on BUG+Blood agar for 2 days at 37ºC prior to the phenotypic tests. Biolog GenIII was performed using Biolog's protocol C2, which suspends bacteria at higher cell density than usual and is designed for organisms with fastidious growth. Hemolysis and temperature growth experiments were performed on BUG+Blood agar (Biolog Universal Growth agar with 5% sheep blood).

Fatty Acid Methyl Ester (FAME) Analysis Methodology

Fatty acid methyl esters (FAMEs) were prepared from cells grown on BUG+Blood agar for 24 h at 37°C using the standard method (Sasser 2009). FAMEs were analyzed on an Agilent 6850 Gas Chromatograph using method RTSBA6 of the Sherlock Microbial Identification System, version 6.1 (MIDI Inc.).

Results

16 rRNA gDNA and PCR product Concentrations

Prior to Sanger and genomic sequencing, gDNA and PCR product concentrations were determined for preparation of final sequencing samples (Table 1). Cells were grown on TSBA plates at 37°C for 48 hours. Concentrations were estimated by comparing the brightness of PCR samples to the 2 Log Ladder (Figure 1). Samples were prepared at a concentration of 5 ng/µL according to the guidelines from GeneWiz. A dull band was seen in the negative control lane (Figure 2, lane 2).

Moraxella sp.	**[gDNA] (ng/µL)**	**[Estimated PCR product](ng/µL)**
K1664^{T}	108	120
K1630	26.1	57
K2450	21.2	57
K127	25.9	100

Table 1: 16 rRNA gDNA and PCR product Concentrations

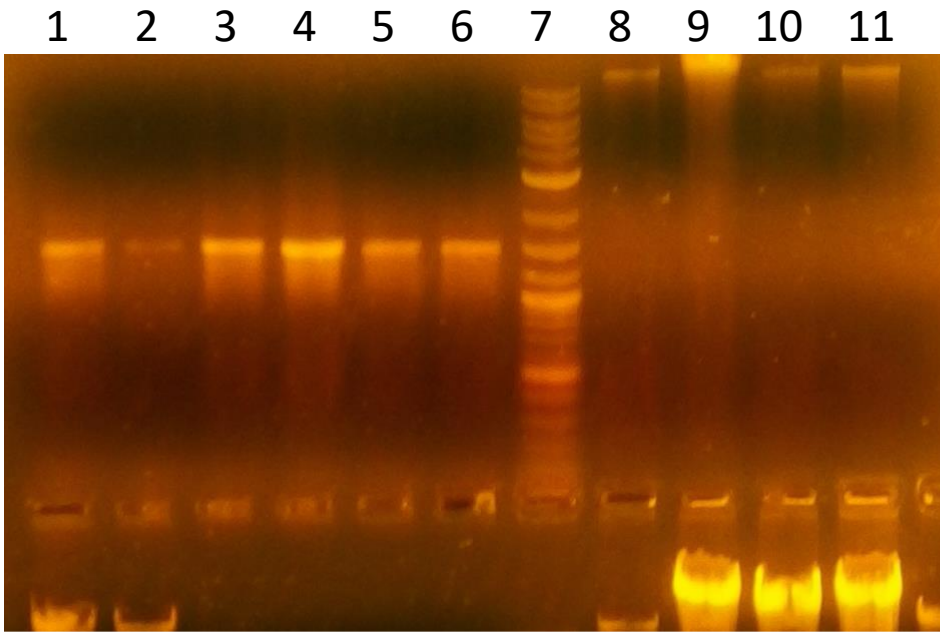


Figure 1. Gel electrophoresis of gDNA and 16S rDNA PCR products. Lanes 3-6 contain 16S rDNA and lanes 8-11 contain gDNA. Lane 1, positive control; lane 2, negative control; lane 3 and 8, K127; lanes 4 and 9, K1664^{T}; lanes 5 and 10, K2450; lanes 6 and 11, K1630.

Once 16S rDNA sequences were completed by GeneWiz and received, consensus sequences of the gene were used in BLAST searches to confirm identities of the novels based on their similarity to one another and other *Moraxella* species. Once these identities were confirmed, the gDNA samples were sent for genome sequencing at MiGS. Sequences were then assembled, annotated, and deposited into NCBI for each novel isolate. Sequencing chromatograms for the 16S rDNA sequences were evaluated to ensure good sequence quality (Figure 2). All sequences showed good base calls indicated by an absence of overlap and high signal intensity.

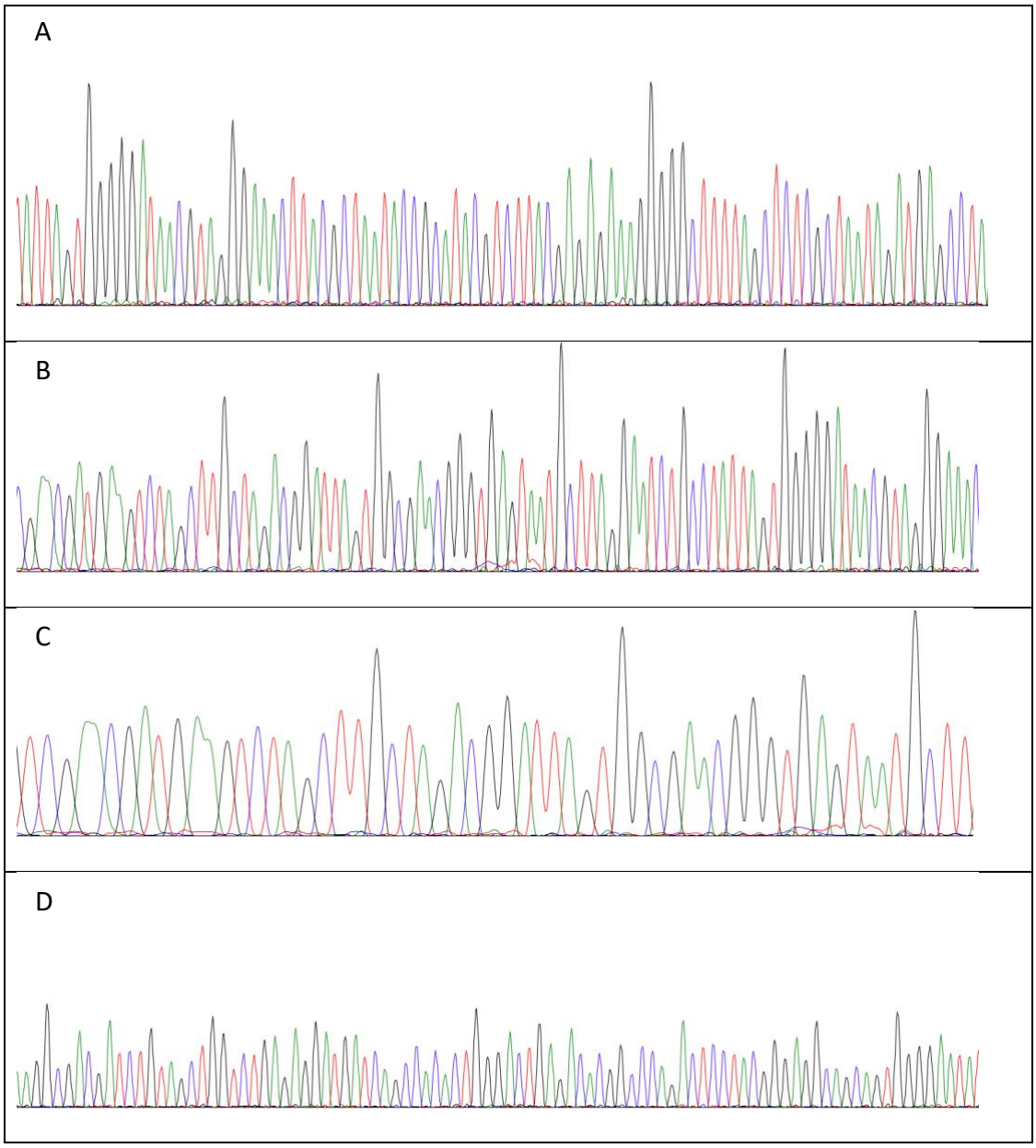


Figure 2. Sanger sequencing chromatograms of the 16S rRNA gene. A. K1664; B. K1630; C. K127; D. K2450.

Genome Sequencing and Statistics:

Following genomic sequencing and assembly, the quality of the sequences was assessed to ensure accuracy of data gathered in whole genome sequence-based comparison metrics. SPAdes on PATRIC was used to gather some statistics including genome size, the total number of contiguous sequences (contigs), and average coverage. CheckM on KBase was used to

examine completeness and contamination of each sample. Completeness and contamination are estimated using marker genes specific to a genome's deduced lineage within a reference genome tree. The data from these sequencing statistics can be found in Table 2.

Moraxella sp.	K1664^{T}	K1630	K2450	K127
Genome size (Mbp)	2.62	2.62	2.66	2.68
Contigs	125	136	153	116
Average coverage	174x	208x	219x	228x
Completeness (%)	97.7	97.7	98.1	97.8
Contamination (%)	0.82	0.82	0.82	0.83
GC content (%)	44.4	44.3	44.3	44.4

Table 2. Illumina NextSeq 550 sequencing statistics for K1664^{T}, K1630, K2450, and K127.

The genome sizes for all genomes were nearly identical, averaging at approximately 2.645 Kbp. Total number of contigs ranged from 116-153. Average coverage ranged from 174x-228x for all 4 organisms. The targeted average coverage was ~100x, therefore this sequencing run exceeded this goal. Completeness was assessed to determine the percent of the genome that was successfully sequenced. The completeness values in Table 2 range from 97.7-98.1% showing that all four organisms had nearly complete assemblies. The contamination was <1% for all four organisms.

Once all four genomes were assembled, annotated, and assessed for good quality, 16S rRNA and genome sequences were deposited into the 16S ribosomal RNA GenBank and Sequence Read Archive (SRA) databases. The accessions for these deposits as well as for the assembly are displayed in Table 3.

Moraxella sp.	16S rRNA	Assembly	Raw Reads (SRA)
Moraxella sp. K1664	MW622078	GCF_015223165.1	SRR12884805
Moraxella sp. K1630	MW622076	GCF_015223135.1	SRR12884635
Moraxella sp. K2450	MW622074	GCF_015223155.1	SRR12884712
Moraxella sp. K127	MW622079	GCF_015223095.1	SRR12884543

Table 3. Accessions for 16S rRNA and genome sequence and assembly deposits.

16S rRNA and rpoB Phylogeny to Identify Closest Relatives

Following genome sequencing, assembly, and annotation, the full-length 16S rRNA and rpoB genes were gathered from K1664^T and BLAST searches were performed to identify the closest relatives. Following these BLAST searches, sequences of the most closely related species were downloaded, aligned in MEGA7, and used to construct phylogenetic tree. The 16S rDNA sequence alignments were used to make a neighbor-joining tree which illustrated the evolutionary relatedness of the novel organisms to the closest relatives, but bootstrap values were relatively low (Figure 3). In the rpoB maximum-likelihood tree (Figure 4), the K1664^T sequence clustered most closely with sequences from *M. equi* (NCTC 11012T= DSM 18027T), *M. lacunata* (CCUG 4441T= DSM 18052T), *M. caprae* (NCTC 12877T= DSM 19149T), and *M. bovis* (CCUG 2133T= DSM 6328T). These strains were purchased from the German Culture Collection (DSMZ).

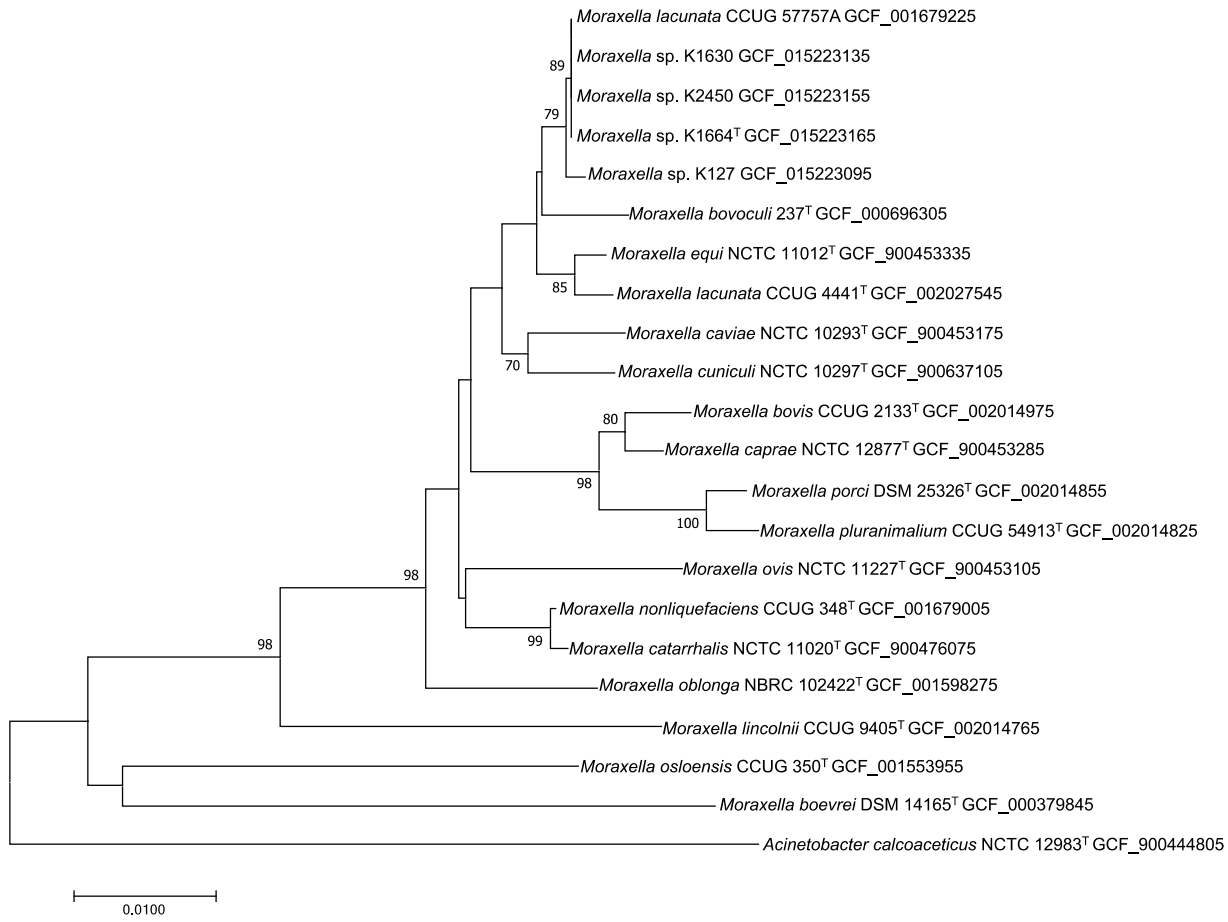


Figure 3. 16S rRNA Sequence Similarity Neighbor-Joining Tree. Bootstrap values below 70 are not shown. *Acinetobacter calcoaceticus* NCTC 12983^{T} was used as an outgroup. The evolutionary distances were computed using the Kimura 2-parameter method (Kimura 1980) and are in the units of the number of base substitutions per site. The tree was made using 22 nucleotide sequences. After gaps were removed, the final dataset contained 1,242 bases

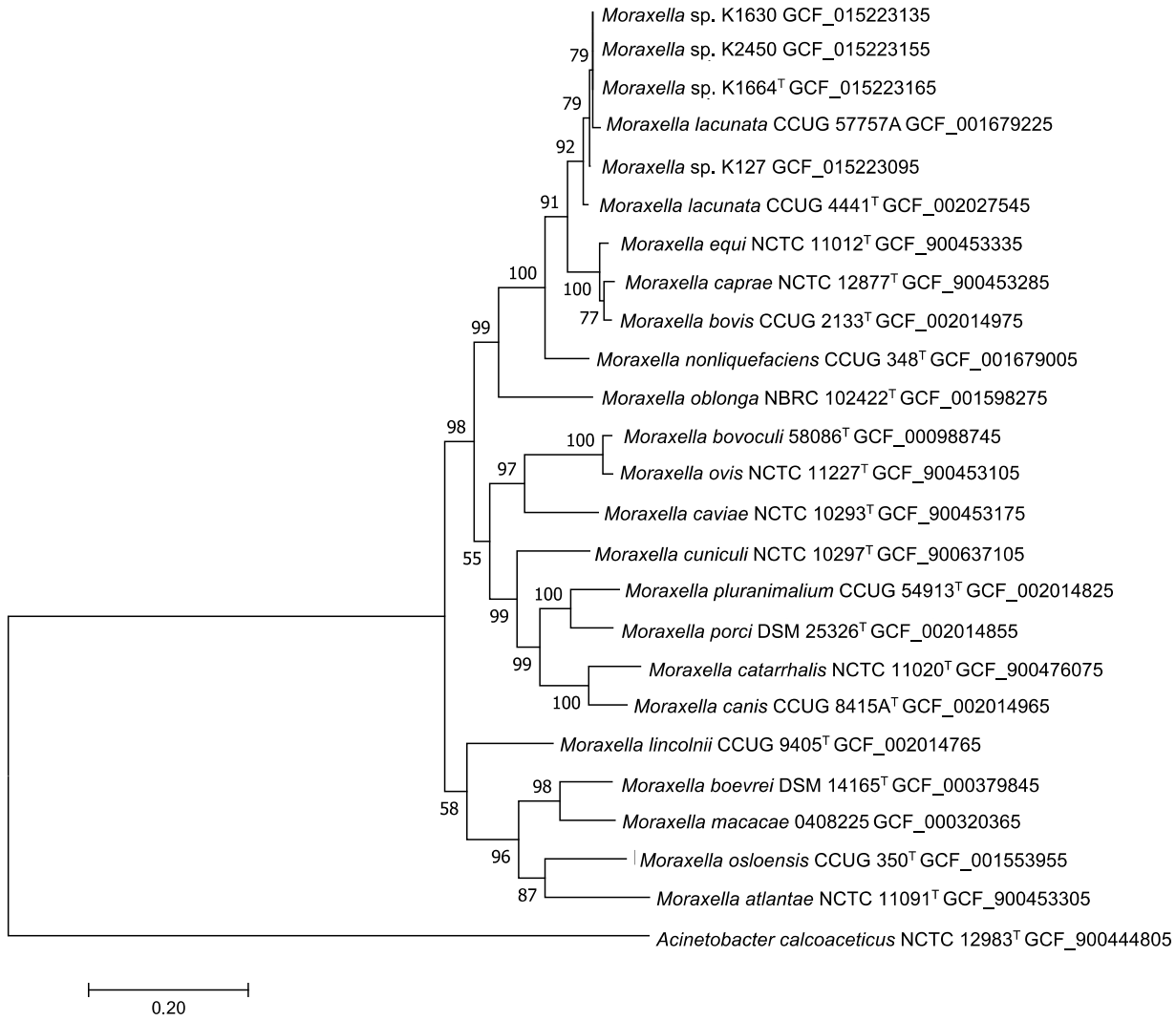


Figure 4. *rpoB* Sequence Similarity Maximum-Likelihood Tree.
Evolutionary history was deduced using the Maximum Likelihood method on the Kimura 2-parameter model (Kimura 1980). Branch lengths are measured with the number of substitutions per site. The tree was made using 25 nucleotide sequences. All positions with gaps were removes and the final dataset contained 3,983 bases.

16S rRNA and Genome-Based Comparisons:

The Overall Genome Relatedness Indices (OGRI) table (Table 4) displays the relatedness of *M. sp.* K1664^T to K1630, K2450, K127, *M. lacunata* CCUG 4441^T, *M. equi* NCTC 11012^T, *M. bovis* CCUG 2133^T, and *M. caprae* NCTC 12877^T based on 16S rRNA, Average Nucleotide

Identity (ANI), estimated DNA-DNA hybridization (eDDH), and Average Amino Acid Identity (AAI) sequence-based similarities. These comparisons offer insight into whether the isolates are novel compared to the closest relatives. The highest OGRI values obtained were all below the species thresholds of 95-96.5% for ANI, 70% for eDDH, and 95% for AAI, confirming that K1664^T is a novel species. Because the 16S rRNA sequence comparisons of the four strains were all >99.5%, K1664^T was designated as the representative strain for the new species being proposed.

Organism	Accession	16S rRNA (%)	ANI (%)	eDDH (%)	AAI (%)
Moraxella sp. K1664					
Moraxella sp. K1630	GCF_015223135.1	100	99.8	99.8	99.8
Moraxella sp. K2450	GCF_015223155.1	100	99.8	99.5	99.8
Moraxella lacunata CCUG 57757A	GCF_001679225.1	100	99.5	99.3	99.4
Moraxella sp. K127	GCF_015223095.1	99.86	98.2	93.5	98.1
Moraxella lacunata CCUG 4441	GCF_002027545.1	99.09	94.8	70.9	95.0
Moraxella caprae NCTC12877	GCF_900453285.1	98.64	90.2	50.5	90.4
Moraxella equi NCTC11012	GCF_900453335.1	99.44	90.1	49.0	90.4
Moraxella bovis CCUG 2133	GCF_002014975.1	97.07	90.0	49.7	90.1
Moraxella nonliquefaciens CCUG 348	GCF_001679005.1	98.39	86.3	37.6	88.1

Table 4. Overall Genome Relatedness Indices. Percent similarities relative to strain K1664^T are shown for the 16S rRNA gene, estimated DNA-DNA hybridization (eDDH), average nucleotide identity (ANI) and average amino acid identity (AAI).

The most similar 16S rDNA sequences to K1664^T from the RefSeq database were *M. lacunata* CCUG 57757A (100%), *M. lacunata* CCUG 4441^T (99.09%), *M. equi* NCTC 11012^T (99.44%), *M. nonliquefaciens* CCUG 348^T (98.39%), *M. caprae* NCTC 12877^T (98.64%), and *M. bovis* CCUG 2133^T (97.07%). ANI revealed K1630, K2450, and *M. lacunata* CCUG 57757A were all above 99% whereas K127 was 98.2% similar. Of the most similar ANI values to validly named type species K1664^T was most similar to *M. lacunata* CCUG 4441^T (94.8%). Other ANI comparisons include *M. equi* NCTC 11012^T (90.1%), *M. caprae* NCTC 12877^T (90.2%), and *M.*

bovis CCUG 2133^T (90.0%). All ANI comparisons are displayed as a heat map in Figure 3. All AAI values were above the ~70% genus threshold relative to the type species *M. lacunata* CCUG 4441^T confirming K1664^T, K1630, K2450, and K127 belong in the *Moraxella* genus.

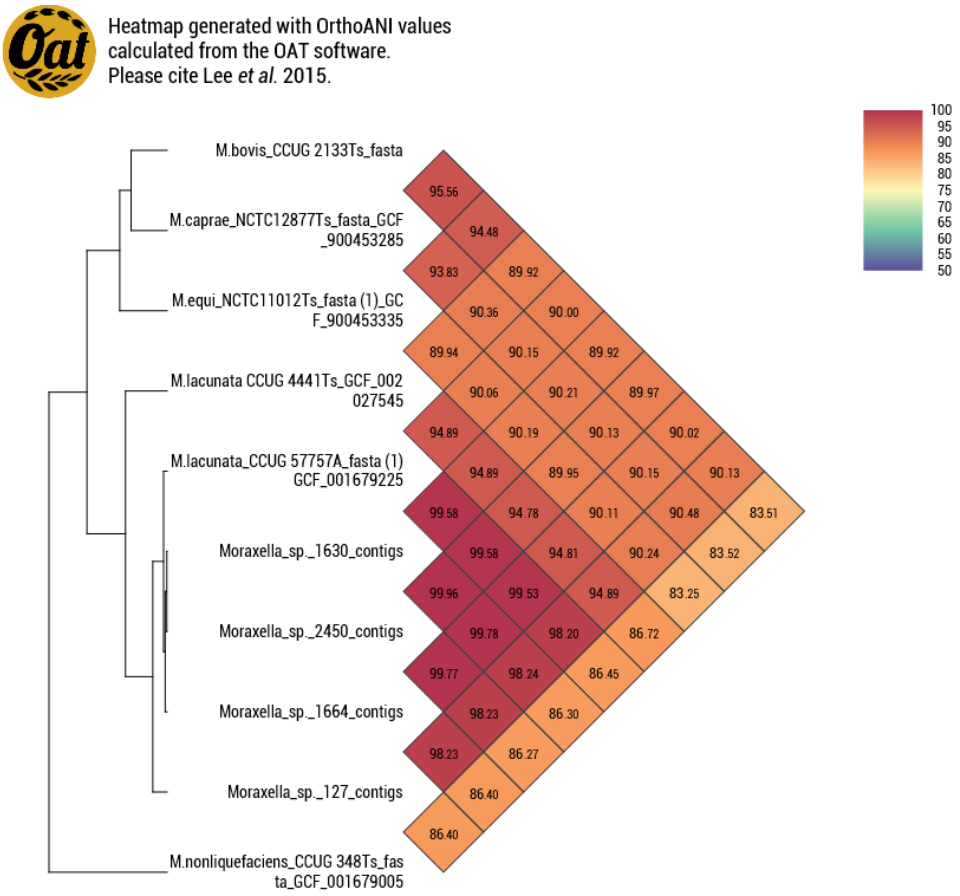


Figure 3: Average Nucleotide Identity among organisms in this study

Venn Diagram of Unique and Shared Genes

The two Venn diagrams generated are illustrated in Figures 4 and 5. From comparison of the four novel isolates to *M. lacunata* CCUG 4441^T, 370 genes were shared among K1664^T, K1630, K2450, and K127 (Figure 4). In total, 2,112 genes are shared among all five organisms.

M. sp. K127 showed the highest count of unique genes among the four novel strains, consistent with its greater variation seen in 16S rRNA, *rpoB*, and ANI similarities. Strain K1664^{T} was compared to the four closest relatives, and all five organisms shared 1,631 genes (Figure 5). Overall, 365 CDS were unique to K1664^{T}.

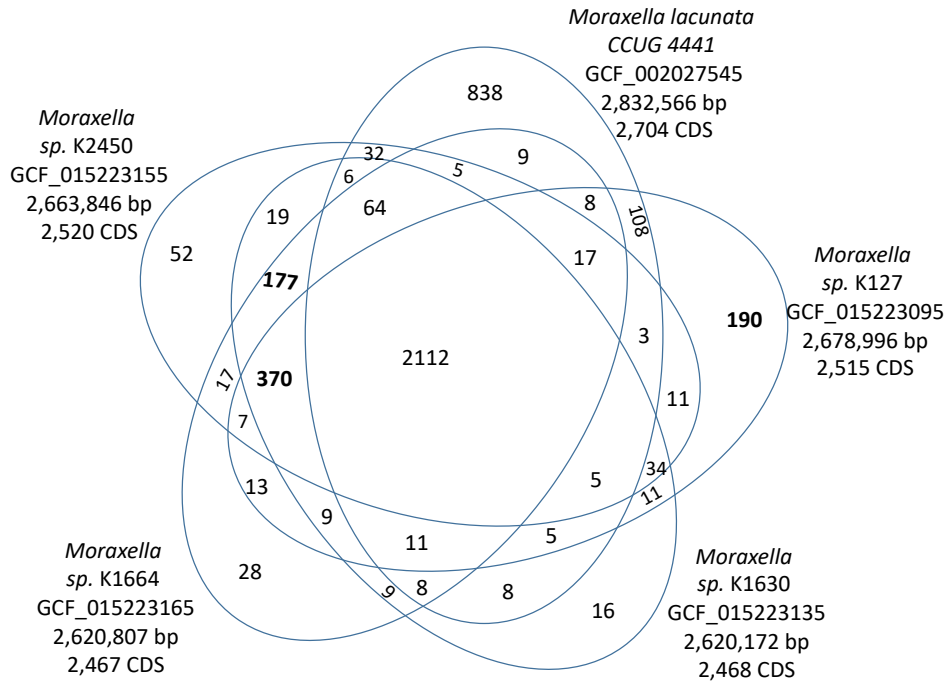


Figure 4. Venn diagram of *Moraxella sp.* K1664^{T}. Organisms chosen for this Venn diagram where K1630, K2450, K127, and the closest relative *M. lacunata* CCUG 4441^{T}. The organisms were chosen to identify shared novel genes between the four isolates not found in the closest relative.

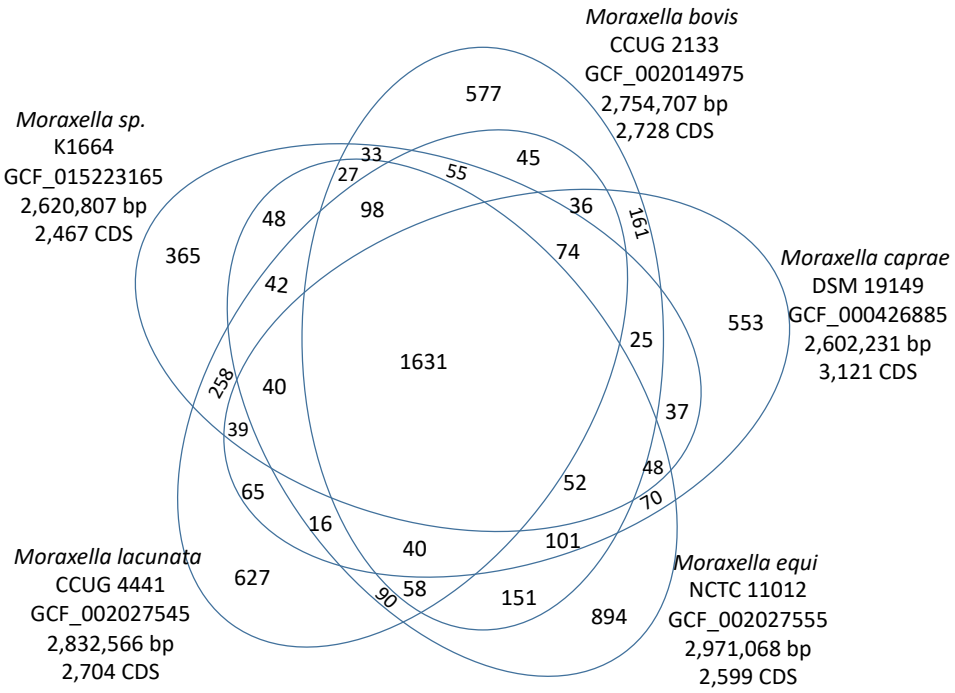


Figure 5. Venn diagram of *Moraxella sp.* K1664^{T}. Organisms chosen for this Venn diagram where the closest relatives to identify unique genes in K1664^{T} representing the proposed novel species.

Genome Features

While examining the CDS shared among these four novel strains but absent from *M. lacunata* CCUG 4441^{T}, several genes of interest were identified, some which may be related to its pathogenesis in keratitis. While many of the 370 CDS encode hypothetical proteins of unknown functions, some of those with predicted functions can be correlated with phenotypes, and specific phenotypic differences between the strains and the closest relatives. Of the 371 CDS, 6 had predicted functions and are listed in Table 5. Unique genes of interest included prolyl endopeptidase (gene #988 in Table 5) and Class A beta lactamase (gene #2674 in Table 5).

Of the 370 shared genes between M. sp. K1664^T, K1630, K2450, and K127, only 65 had predicted functions.			
Gene #	**Predicted Function**	**Gene #**	**Predicted Function**
29	Metallophosphoesterase	1099	Ferric iron ABC transporter, iron-binding protein
41	Uncharacterized MFS-type transporter	1100	Ferric iron ABC transporter, permease protein
45	Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)	1101	Ferric iron ABC transporter, ATP-binding protein
102	Cell division inhibitor Slr1223 (YfcH in EC), contains epimerase/dehydratase and DUF1731	1387	Mobile element protein
168	VgrG protein	1396	Type IV pilus biogenesis protein PilE
230	Sulfate transport system permease protein CysT	1554	Alkaline phosphatase (EC 3.1.3.1)
286	Transposase and inactivated derivative	1556	Alkaline phosphatase
292	Mobile element protein	1669	Similar to eukaryotic Peptidyl prolyl 4-hydroxylase, alpha subunit (EC 1.14.11.2)
294	Antitoxin HigA	1670	probable transcriptional regulator
295	Toxin HigB	1713	Cardiolipin synthase, bacterial type ClsA
305	Transposase and inactivated derivatives	1795	Transcriptional regulator, LysR family
364	Mobile element protein	1796	Dienelactone hydrolase and related enzymes
366	Alkaline phosphatase (EC 3.1.3.1)	1800	Alcohol dehydrogenase (EC 1.1.1.1)
370	Alkaline phosphatase (EC 3.1.3.1)	1893	Organic hydroperoxide resistance protein
418	Sodium/glutamate symporter	1895	Plasmid related protein
439	Putative transposase InsK for insertion sequence element IS150	1904	Tetratricopeptide repeat family protein
467	Biopolymer transport protein ExbD/TolR	1923	Esterase/lipase
468	MotA/TolQ/ExbB proton channel family protein	2059	IS1381, transposase OrfA
496	MOSC domain protein	2100	Domain often clustered or fused with uracil-DNA glycosylase
552	Isochorismatase (EC 3.3.2.1)	2118	DNA polymerase III subunits gamma and tau (EC 2.7.7.7)
652	Galactose/methyl galactoside ABC transport system, D-galactose-binding periplasmic protein	2126	Putative helicase
669	Mobile element protein	2134	Maebl
670	Uncharacterized MFS-type transporter	2147	Cytochrome C553 (soluble cytochrome f)
819	Carbon starvation protein A	2246	Ser/Thr protein phosphatase family protein
868	Uncharacterized protein COG3236	2524	RelB/StbD replicon stabilization protein (antitoxin to RelE/StbE)
963	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	2525	mRNA interferase RelE
967	MBL-fold metallo-hydrolase superfamily	2555	DNA replication protein
976	Pantothenate:Na+ symporter (TC 2.A.21.1.1)	2589	Type IV pilin PilA
988	Prolyl endopeptidase (EC 3.4.21.26)	2592	Mobile element protein
1029	Superfamily II DNA and RNA helicase	2598	Uncharacterized UPF0118 membrane protein
1044	VapB protein (antitoxin to VapC)	2674	Class A beta-lactamase (EC 3.5.2.6)
1045	VapC toxin protein	2687	Predicted signal-transduction protein containing cAMP-binding and CBS domains
1065	HrgA protein		

Table 5. Unique genes of strains K1664^T, K1630, K2450, and K127 that have predicted functions. The genes that do not have predicted functions are identified as "hypothetical protein" and were not included in this table.

Physiology and Chemotaxonomy

The short rod-shaped cells exhibited a negative Gram stain, did not produce endospores, and lacked motility. For phenotypic tests, cells were grown on BUG+Blood agar (Figure 6) at 37^{o}C for 24-48 hours (Figure 7). Several phenotypic differences from the reference strains were noted (Table 6).

Growth response with 71 different carbons sources and growth under 23 inhibitory conditions were tested on the Biolog GenIII plate. API ZYM and API 20NE tests were used to investigate enzymatic activities and carbon source metabolism. Beta-hemolysis and temperature tests were performed on BUG+Blood agar. Strains K1664^{T}, K1630, K2450, and K127 were uniquely resistant to penicillin, amoxicillin, ampicillin, and carbenicillin. The four novel isolates and *M. lacunata* DSM 18052^{T} were positive for nitrate reduction and alkaline phosphatase activity. Only *M. caprae* DSM 19149^{T} and *M. bovis* DSM 6328^{T} had beta-hemolysis activity. Like most other *Moraxella* species, the four isolates displayed positive results for catalase and oxidase (Table 6). The negative result for catalase in *M. bovis* DSM 6328^{T} was found to be a result of a single nucleotide deletion illustrated in Figure 7. Absence of beta-hemolysis in the novel isolates was found to be the result of a 6 nucleotide deletion illustrated in Figure 8.

The fatty acid methyl ester (FAME) profile revealed percentages (Table 7) that were typical of *Moraxella* species. The major FAMEs identified are $C_{18:1}$ ω9c, summed feature 3 corresponding to $C_{16:1}$ ω7c or $C_{16:1}$ ω6c, $C_{16:0}$, $C_{17:1}$ ω8c, $C_{18:0}$, and $C_{18:3}$ ω6c (6,9,12). Strain K127 had a much higher percentage of $C_{17:1}$ ω8c that the other novel isolates. Strains K1664^{T}, K1630, K2450, and K127 had a higher percentage of $C_{16:0}$ in comparison to the closely related type strains.

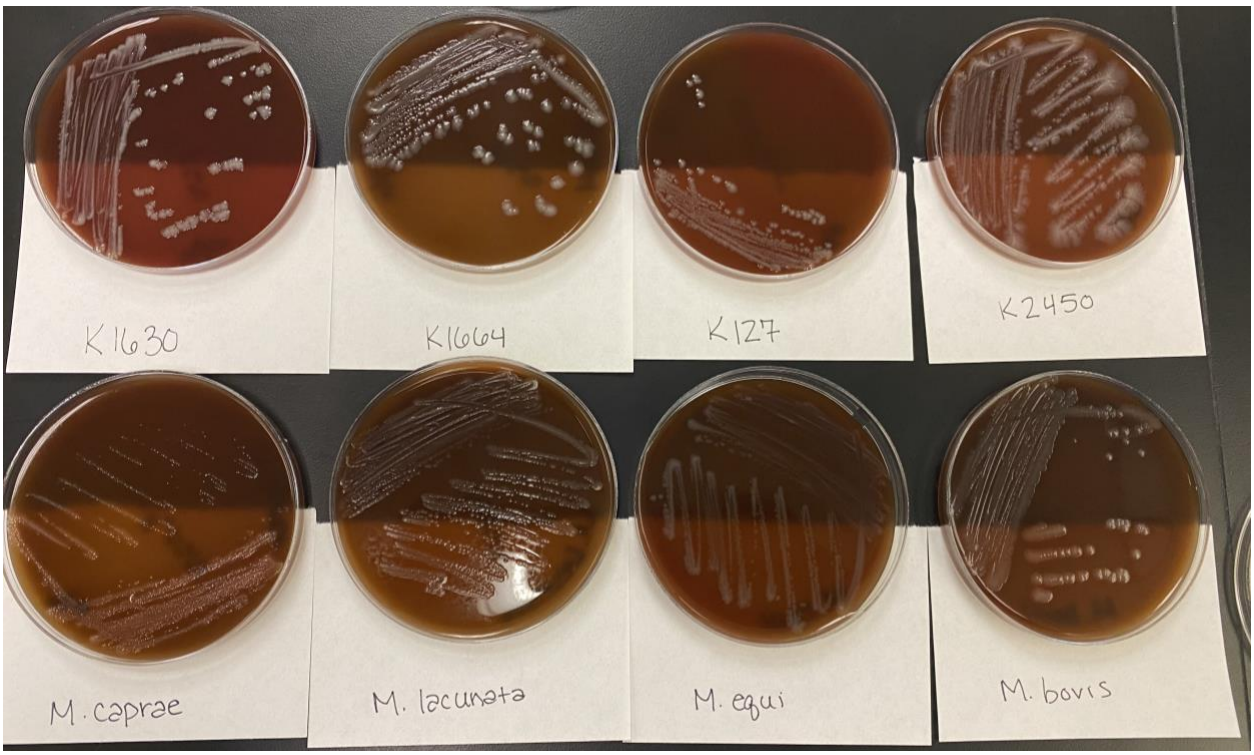


Figure 6. Streak plates on BUG+Blood agar incubated for 48 hours at 37ºC.

Characteristic	1	2	3	4	5	6	7	8
Catalase	+	+	+	+	+	+	-	+
Beta Hemolysis	-	-	-	-	-	-	+	+
Cephalexin (30 ∝g)	I	I	I	I	S	S	S	S
Cephalosporin (30 ∝g)	I	I	I	I	S	S	S	S
Ampicillin (10 ∝g)	R	R	R	R	S	I	I	S
Penicillin G (6.5 µg)	R	R	R	R	S	S	S	S
Carbenicillin (100 µg)	R	R	R	R	S	R	S	S
Amoxicillin (20 µg)	R	R	R	R	S	I	S	S
API ZYM								
alkaline phosphatase	+	+	+	+	+	-	-	-
Acid phosphatase	-	-	-	+	-	-	-	+
API 20NE								
NO_3 reduction	+	+	+	+	+	-	-	-
Gelatin hydrolysis	+	+	+	+	-	+	-	-
Biolog GenIII Plate								
dextrin	-	w	-	w	-	-	-	-
pH 6	-	+	+	w	-	w	-	+
troleandomycin	+	-	w	-	-	-	-	-
rifamycin SV	+	-	+	+	-	-	-	-
L-alanine	-	-	-	+	-	-	-	-
L-histidine	+	+	+	+	-	-	+	+
L-serine	+	+	-	+	-	+	-	+
lincomycin	+	-	w	-	-	-	-	-
pectin	w	w	w	-	+	-	-	-
methyl pyruvate	w	+	-	-	-	w	-	-
citric acid	+	+	+	+	-	+	-	-
nalidixic acid	+	-	+	+	-	+	w	-
tween-40	+	+	w	+	-	+	-	+
b-hydroxy-D,L-butyric acid	-	-	-	-	-	-	+	+
acetoacetic acid	+	+	+	+	w	+	-	+

Table 6. Differential phenotypes between 1. K1664^T 2. K1630 3. K2450 4. K127 5. *M. lacunata* DSM 18052^T 6. *M. equi* DSM 18027^T 7. *M. bovis* DSM 6328^T and 8. *M. caprae* DSM 19149^T. + indicate positive results. – indicates negative results. w indicates weak positive results. S is antibiotic sensitivity. I is intermediate antibiotic sensitivity. R indicated resistance to an antibiotic.

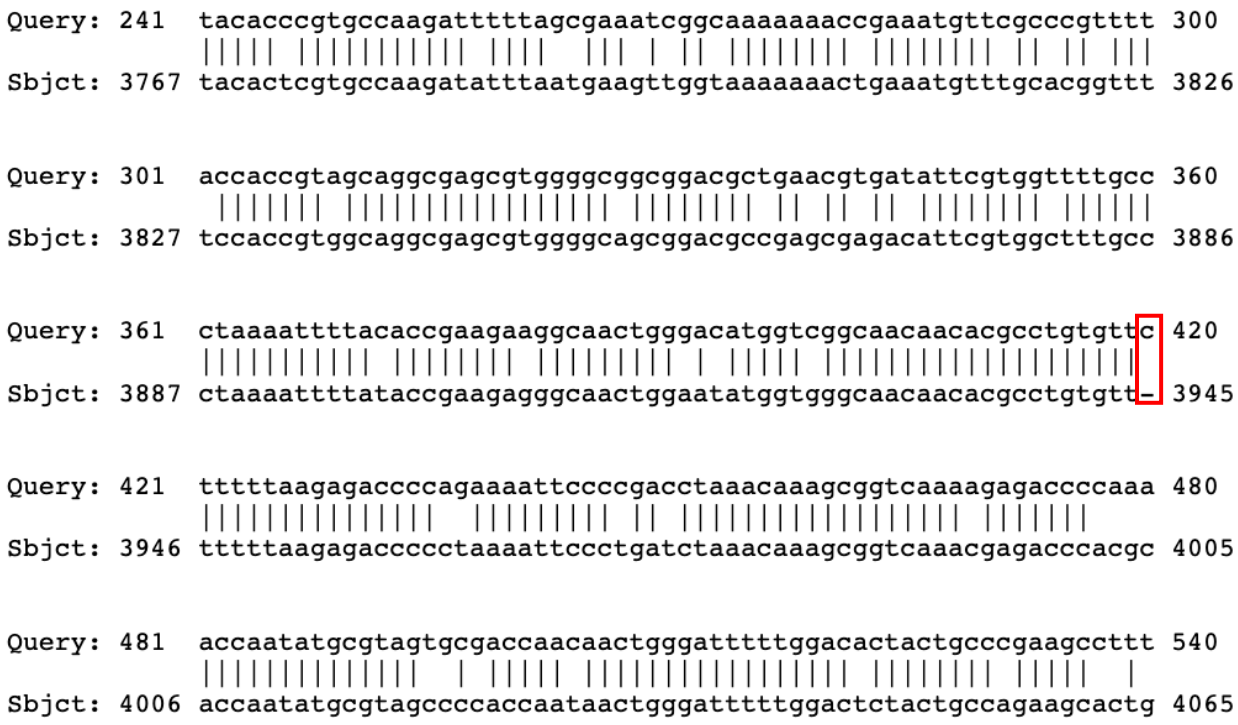


Figure 7. Catalase gene sequence alignment of K1664^{T} to *M. bovis* CCUG 2133^{T}. A frameshift deletion mutation is boxed in red.

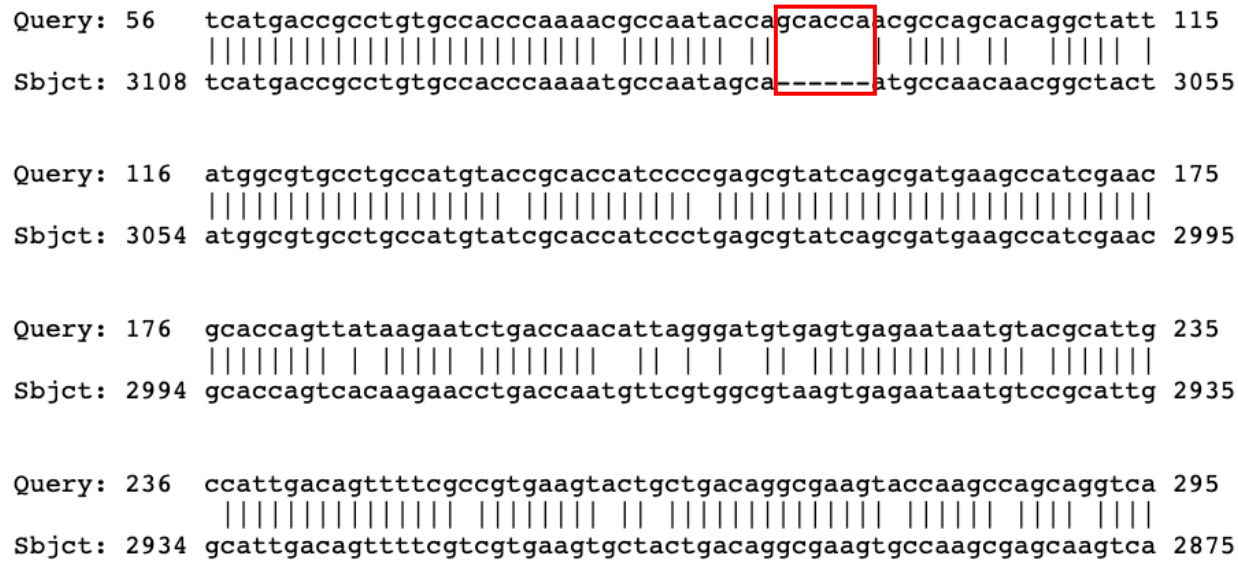


Figure 8. 21 kDa hemolysin precursor gene sequence alignment of *M. bovis* CCUG 2133^{T} to K1664^{T}. A 6 base pair deletion is boxed in red.

Fatty acid composition	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
$C_{10:0}$	----	TR	----	----	2.0	1.8	1.4	1.5
$C_{12:0}$	TR	TR	TR	TR	2.9	3.6	3.3	1.0
$C_{12:0}$ 3OH	----	----	----	----	2.0	3.5	2.7	4.3
iso-$C_{16:0}$	----	----	TR	TR	2.0	4.0	1.7	1.8
$C_{16:0}$ N alcohol	TR	1.2	TR	TR	1.3	TR	TR	----
$C_{16:0}$	**12.3**	**17.0**	**11.6**	8.3	6.2	6.0	4.9	5.8
iso-$C_{17:1}$	----	----	----	----	----	TR	1.2	1.2
$C_{17:1}$ *ω8c*	3.6	3.7	8.6	**18.0**	1.2	3.0	**10.8**	1.0
$C_{17:0}$	TR	TR	TR	TR	----	----	TR	1.2
$C_{18:3}$ *ω6c* (6,9,12)	1.4	1.5	1.4	TR	8.7	3.3	5.6	4.3
$C_{18:1}$ *w9c*	**53.3**	**36.0**	**37.1**	**40.4**	**49.5**	**28.0**	**47.6**	**37.9**
$C_{18:0}$	2.5	2.9	3.1	1.9	2.7	1.0	2.6	3.6
Summed Feature 3	**25.6**	**35.3**	**35.3**	**28.1**	**17.3**	**39.9**	**11.7**	**15.2**
Summed Feature 6	----	----	----	TR	TR	----	1.5	1.2

Table 7. Fatty acid composition of Moraxella sp. K1664^T compared with closely related type strains of the genus *Moraxella*.
Values represent percentages of the total fatty acid with those amounting to 10% or more in bold. - are those not detected. TR, traces, are fatty acids present but amounting to less than 1%. Strains were grown at 37°C on BUG+Blood plates for 24 hours prior to fatty acid analysis. Strains: 1. K1664^T; 2. K1630; 3. K2450; 4. K127; 5. *M. lacunata* DSM 18052^T; 6. *M. equi* DSM 18027^T; 7. *M. bovis* DSM 6328^T; 8. *M. caprae* DSM 19149^T.

Discussion

After receiving the four isolates from the University of Pittsburgh Medical Center, fresh streak plates provided pure cultures for gDNA extraction and sequencing of the genome and 16S rRNA gene. The PCR amplified 16S rDNA and gDNA had acceptable DNA concentrations, indicated by the intensity of each band (Figure 1), required for 16S rDNA and genome sequencing at GeneWiz and MiGS. Strain K1664^T had the highest concentrations of both gDNA and 16S rDNA indicated by band intensity from gel electrophoresis (Figure 1, lanes 4 and 9), and these results were consistent with the Qubit measurements outlined in Table 1. The positive

control (Figure 1, lane 1) had a band as expected, but the negative control (Figure 1, lane 2) also showed a light, unexpected band at the same migration distance as the positive control indicating some PCR product may have contaminated the negative control well. However, the band in the negative control lane was minute enough, with an estimated DNA concentration of <5ng/mL, that repeating the PCR and gel electrophoresis steps were not required. The 16S rDNA samples were amplified with 27f and 1492r primers, and bands on the gel migrated to the expected region of approximately 1,500 base pairs as indicated by the 2 Log Ladder. 16S rDNA sequencing statistics outlined in Table 2 indicated that the sequences were high-quality.

Following genome sequencing, several sequence-based comparisons were performed to identify the most closely related species to the novel isolates. The full-length 16S rDNA sequences were gathered from RAST and used in a series of BLAST searches. *M. lacunata* CCUG 57757A (100%) showed the highest 16S rRNA sequence similarity to K1664^T. The most closely related type strains were *M. lacunata* CCUG 4441^T (99.09%), *M. equi* NCTC 11012^T (99.44%), *M. nonliquefaciens* CCUG 348^T (98.39%), *M. caprae* NCTC 12877^T (98.64%), and *M. bovis* CCUG 2133^T (97.07%) (Table 4). Because K127 was more different from the other three isolates based on 16S rRNA, rpoB, and ANI sequence similarity, K1664^T was designated as the type strain for the new species being proposed. The 16S rRNA Neighbor-Joining tree (Figure 3) showed strains K1664^T, K1630, K2450, and K127 clustered most closely with *M. lacunata* CCUG 57757A. M. bovoculi 237^T also clustered closely with the four novel strains, and *M. equi* NCTC 11012^T, *M. lacunata* CCUG 4441^T, *M. bovis* CCUG 2133^T (97.07%), and *M. nonliquefaciens* CCUG 348^T clustered further away.

Relatedness was also assessed by *rpoB* sequence similarity to confirm the most closely related organisms. *rpoB* is a protein-encoding gene and is less highly conserved than 16S rRNA

genes, and therefore offered greater resolution. A Maximum-Likelihood tree was made after sequences were aligned on MEGA7 (Figure 4). The bootstrap values were noticeably higher than the 16S rRNA Neighbor-Joining tree, indicating how many times the same branching pattern was observed during each phylogenetic construction. Compared to the 16S rRNA tree, *M. lacunata* CCUG 57757A still clustered most closely with strain K1664^T, but this time, the most closely related type strains clustered more closely in the order *M. lacunata* CCUG 4441^T, *M. equi* NCTC 11012^T, *M. caprae* NCTC 12877^T, *M. bovis* CCUG 2133^T (97.07%), and *M. nonliquefaciens* CCUG 348^T. Based on this clustering, *M. lacunata* CCUG 4441^T, *M. equi* NCTC 11012^T, *M. caprae* NCTC 12877^T, and *M. bovis* CCUG 2133^T represented the four most closely related type strains and were ordered from the DSMZ for further phenotypic testing.

Whole-genome sequenced based comparisons including ANI and AAI were subsequently performed to determine if strain K1664^T and the related isolates were in fact a novel species within the *Moraxella* genus (Table 4). ANI values for strain K1664^T compared to K1630, K2450, K127, and *M. lacunata* CCUG 57757A were all >98% which is above the 95% species threshold, confirming that the four isolates are the same species. Typically when organisms are deposited to NCBI, they are given a name corresponding to the organism with the highest 16S rRNA similarity. In this case, the 16S rRNA sequence of *M. lacunata* CCUG 57757A is 99.11% similar to the type strain *M. lacunata* CCUG 4441^T, explaining its classification at a strain of *M. lacunata*. However, the ANI comparisons in Figure 3 show *M. lacunata* CCUG 57757A is only 94.89% similar to *M. lacunata* CCUG 4441^T, which is just below the species threshold, whereas in comparison to the four novel strains, it was >98% similar. Taken together, these data indicate that *M. lacunata* CCUG 57757A should be reclassified into the proposed species *Moraxella kowalskii* with K1664^T, K1630, K2450, and K127. AAI values for all closely related organisms

in comparison to the type species *M. lacunata* CCUG 4441^T were above the ~75% genus threshold, confirming all organisms fall within the genus *Moraxella*.

A total of 2,112 shared bidirectional best hits were identified between the four novels strains and *M. lacunata* CCUG 4441^T (Figure 4) and 1,631 shared bidirectional best hits were identified between strain K1664^T and the closest type strains (Figure 5) using a Venn diagram created from the SEED Viewer Sequence Based Comparison Tool http://lycofs01.lycoming.edu/~newman/CurrentResearch.html. Bidirectional best hits are pairs of proteins encoded by different genomes that are more similar to one another than either is to any other protein encoded by the other genome. The vast majority of the time, these proteins are orthologs, which means that the gene encoding them are derived from a gene in a common ancestor organism via speciation. Orthologs also typically serve the same function in different organisms. Ultimately, this tool identifies bidirectional best hits that represent shared orthologs between two genomes. These gene counts constitute the core genome. Strains K1664^T, K1630, K2450, and K127 shared 370 CDS that are unique to the proposed species *Moraxella kowalskii*.

Following construction of Venn diagrams, the 370 CDS unique to the novel isolates were evaluated further. A 2,577 bp CDS was annotated by RAST as encoding a prolyl endopeptidase (WP_172823638) (Figure 6). To ensure this CDS was not carried by *M. lacunata* CCUG 4441^T, the amino acid sequence was searched against the non-redundant protein database using BLASTP. This CDS was less than 42% identical to its closest BLASTP matches. Further research revealed prolyl endopeptidase is a requisite protease in the formation of *N*-acetyl proline-glycine-proline, a molecule associated with corneal ulceration from proteolytic destruction of collagen (Gaggar et al. 2008). Corneal ulceration is a major characteristic of

keratitis, and prolyl endopeptidase likely plays a major role in *M.* sp. K1664, K127, K1630, and K2450's pathogenesis of this disease.

A gene encoding a Class A beta-lactamase (WP_003657002) is present in all four novel isolates but absent in the four reference genomes. Beta-lactamases inactivated antibiotics by hydrolyzing the amide bond of the characteristic beta-lactam ring (Matagne et al. 1998). A BLASTP search of the non-redundant protein database revealed that *M. catarrhalis* shared the same protein (WP_152720806) This protein confers antibiotic resistance against beta-lactam antibiotics. This finding was supported by Kirby-Bauer antibiotic tests which revealed unique resistance to penicillin, ampicillin, amoxicillin, and carbenicillin among the four isolates, but not in the reference organisms. The novel strains also showed intermediate resistance to cephalexin, which also is a member of the beta-lactam family of antibiotics.

Fifteen phage-related genes, present on 2 separate contigs, were annotated in the strain K127 genome. Of these genes, four were unique to K127 and annotated as "Phage protein." A phage repressor protein cI, annotated in NCBI as a helix-turn-helix domain-containing protein (WP_194108658), in strain K127 had bidirectional best hits with *M. lacunata* CCUG 4441^{T}, *M. bovis* CCUG 2133^{T}, and *M. equi* NCTC 11012^{T}. This protein has a known role in controlling gene expression in temperate phages and controlling lysis and lysogeny by a transcriptional switch involving face-to-face promoters (Aravind et al. 2005). This suggests a lysogenic phage is integrated into the genomes of *M. sp.* K127, *M. lacunata* CCUG 4441^{T}, *M. bovis* CCUG 2133^{T}, and *M. equi* NCTC 11012^{T}. A CRISPR array with 16 repeat regions and 16 spacer regions is also present in strain K127. CRISPR arrays are comprised of two major function elements including 23-55 bp tandem short DNA repeats and 26-72 bp variable spacer sequences separating the conserved regions which combined provide a type of "immune" defense for bacteria against

viruses (Karimi et al. 2018). Therefore, this finding is consistent with the proposed presence of a prophage within the genome of K127, providing a mechanism of defense when the virus becomes active.

The four novel strains share a unique operon encoding three ferric iron ABC transporter proteins including an iron-binding protein (WP_083101857), a permease protein (WP_065255690), and ATP-binding protein (WP_065255691). This operon likely confers an advantageous mechanism of acquiring iron from host proteins and carrier molecules to support pathogenesis (Lewis et al. 2012). In addition, all eight strains contain three nearby coding sequences annotated in RAST as ferric vibriobactin proteins VctD (WP_194092626), VctG (WP_065256021), and Vct P (WP_065256023). Similar proteins have been previous found in other pathogenic bacteria such as *Vibrio cholerae* and have been speculated to provide a mechanism to evade the siderocalin-mediated mammalian innate immune response (Li et al. 2012). Such a system could thus play a role in the pathogenesis of these organisms.

Of the eight strains studied, the four novel isolates and *M. lacunata* DSM 18052^T were positive for nitrate reduction corresponding with the presence of a four gene cluster that encodes respiratory nitrate reductase chains (WP_065256882, WP_062499468, WP_065256878, WP_065256879). Only these five strains were also positive for alkaline phosphatase activity. Strains K1664^T, K127, K1630, K2450, and *M. equi* tested positive for gelatinase. The four novel isolates were uniquely resistant to ampicillin, penicillin, amoxicillin, and carbenicillin corresponding with a 942 bp CDS encoding a Class A beta-lactamase (WP_003657002). Of the eight strains, only *M. sp*. K127 and *M. bovis* DSM 6328^T were weakly positive for acid phosphatase activity on the API ZYM strip, a result consistent with previously reported phenotypic results of *M. bovis* (Juni and Bøvre 2015). *M. bovis* DSM 6328^T and *M. caprae* DSM

19149^{T} were the only strains positive for beta hemolysis on BUG+Blood and Columbia Blood agars.

Although beta-hemolysis was only observed in *M. bovis* DSM 6328^{T} and *M. caprae* DSM 19149^{T}, further analysis of the annotated genomes revealed strains K1664^{T}, K1630, K2450, and K127 also carried genes annotated at 21 kDa hemolysin precursor (WP_065256358). This gene is 936 bp in the genomes of the novel isolates, whereas in *M. bovis* DSM 6328^{T} and *M. caprae* DSM 19149^{T}, the sequence is 996 bp. After performing a nucleotide BLAST search of the gene sequence from *M. bovis* DSM 6328^{T} against the strain K1664^{T} genome, the results indicated a 6 base pair deletion that caused the deletion of a serine and threonine residue, likely rendering the hemolysin precursor non-functional, explaining the absence of hemolysis in these strains (Figure 8).

All eight strains produced greyish-white colonies that turned blue upon exposure to oxidase reagent, indicating positive oxidase tests. Of the strains studied, all were catalase positive except for *M. bovis* DSM 6328^{T}. This finding contrasts to previous descriptions of *M. bovis* strains which have reported the majority are catalase positive (Angelos et al. 2007). During further evaluation of catalase genes annotated on the SEED Viewer for *M. bovis* DSM 6328^{T}, coding sequences of 426 bp and 1056 bp were noted, while a 1536 bp CDS was present in strain K1664^{T}. Alignment of these regions (Figure 7) revealed a single base deletion frameshift mutation [3945delT] corresponding to a loss of function of catalase. This was supported by the absence of catalase activity in *M. bovis* DSM 6328^{T}.

M. lacunata DSM 18052^{T} did not produce an adequate response in the positive control well on the Biolog GenIII likely due to its fastidious growth requirements, and therefore will not be discussed. Strains K1664^{T}, K1630, K127, and K2450 showed unique resistance to rifamycin.

Previous studies have reported a link between mutations in bacterial *rpoB* genes often found in pathogenic clinical isolates and resistance to rifamycin, which may serve as an explanation for these results (Goldstein 2014). Several mutations in the *rpoB* genes of the novel isolates compared to the reference organisms were identified, but further analysis is required to elucidate which mutation could be responsible for this phenotype. Interestingly, strain K1664^T showed unique resistance to lincomycin (lincosamide) and troleandomycin (macrolide). Combined resistance to these two types of antibiotics have been previously reported where three possible mechanisms of resistance were discussed: ribosomal modification, efflux of antibiotic, and inactivation of antibiotic (Leclercq 2002). Further analysis of this unique resistance would provide useful as strain K1664^T has shown the greatest diversity of antibiotic resistance amongst the eight strains studied here.

The eight strains studied tested negative for the majority of the carbon-utilization assays on the BIOLOG GenIII plate, a finding consistent with other *Moraxella* species (Juni and Bøvre 2015). All strains studied here with the exception of *M. equi* DSM 18027^T were positive for L-histidine utilization. The negative result for this strain is likely due to the absence of a gene encoding histidine ammonia-lyase (WP_065256994).

Based on these results, strains K1664^T, K1630, K2450, and K127 represent a novel species with strain K1664^T and the type strain. Strain K1664^T is different enough from the most closely related species based on ANI to represent a novel species. Based on 16S rRNA, *rpoB*, and ANI similarities, *M. lacunata* CCUG 57757A should be reclassified into the novel species with the four novel strains. Several phenotypes unique to the novel strains have been presented and will be representative of the new species. The name *Moraxella kowalskii* sp. nov. is proposed for the new species. Future work will require the deposit of strains K1664^T and K127

to culture collections prior to submission of a manuscript to the International Journal of Systematic and Evolutionary Microbiology (IJSEM).

Description of Moraxella *kowalskii sp. nov*.

Colonies are circular with slightly irregular edges, greyish-white, smooth, and glossy. Cells are gram-stain negative, diplococcobacilli approximately 0.5 μm wide and 2.0 μm long. The major fatty acids are $C_{18:1}$ ω9c, Sum In Feature 3 ($C_{16:1}$ ω7c/ $C_{16:1}$ ω6c), $C_{16:0}$, $C_{17:1}$ ω8c, $C_{18:0}$, and $C_{18:3}$ ω6c (6,9,12). Growth occurs at 22–37^{O}C (optimum, 30-37^{o}C), pH 6.0-7.0 (optimum, pH 7.0) and with 1.0–4.0% NaCl (optimum, 1.0% NaCl), but no growth on BUG + blood occurs in the absence of oxygen. Exhibits catalase, oxidase and nitrate reduction activity. With the API ZYM strip, K1664^{T} is positive for alkaline phosphatase, C4 esterase, C8 esterase lipase, and leucine arylamidase but negative for C14 lipase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, acid phosphatase, naphthol-as-bi-phosphohydrolase, α-galactosidase, β-galactosidase, β-glucuronidase, α- glucosidase, β- glucosidase, n-acetyl-β-glucosaminidase, α- mannosidase, and α- fucosidase. With the API 20NE strip, K1664^{T} is positive for nitrate reduction and liquefaction of gelatin, but negative for hydrolysis of PNPG, production of acid from glucose, arginine dihydrolase, urease and assimilation of glucose, arabinose, mannose, mannitol, N-acetyl glucosamine, maltose, gluconate, caprate, adipate, malate, citrate, and phenylacetate. In the BIOLOG GenIII plate, a positive response is obtained in the wells containing 1% NaCl, 4% NaCl, 1% Na-lactate, D-fructose-6-PO_4, troleandomycin, rifamycin SV, L-glutamic acid, L-histidine, pectin, glucuronamide, vancomycin, tetrazolium violet, tetrazolium blue, methyl pyruvate, D-lactic acid methyl ester, L-lactic acid, citric acid, α-keto-glutaric acid, D-malic acid, L-malic acid, bromo-succinic acid, nalidixic acid, LiCl, tween-

40, γ-amino-butyric acid, α-hydroxy-butyric acid, acetoacetic acid, and acetic acid; negative responses are obtained in wells dextrin, D-maltose, D-trehalose, D-cellobiose, gentiobiose, sucrose, D-turanose, stachyose, pH 5, D-raffinose, α-D-lactose, D-melibiose, β-methyl-D-glucoside, D-salicin, N-acetyl-D-glucosamine, N-acetyl-β-D-mannosamine, N-acetyl-D-galactosamine, N-acetyl neuraminic acid, 8% NaCl, α-D-glucose, D-mannose, D-fructose, D-galactose, 3-methyl glucose, D-fucose, L-fucose, L-rhamnose, inosine, fusidic acid, D-serine, D-sorbitol, D-mannitol, D-arabitol, myo-inositol, glycerol, D-glucose-6-PO4, D-aspartic acid, D-serine, minocycline, gelatin, glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-pyroglutamic acid, guanidine HCl, niaproof 4, D-galacturonic acid, L-galacturonic acid lactone, D-gluconic acid, D-glucuronic acid, mucic acid, quinic acid, D-saccharic acid, p-hydroxy-phenylacetic acid, K-tellurite, β-hydroxy-D,L-butyric acid, α-keto-butyric acid, propionic acid, formic acid, aztreonam, Na-butyrate, and Na bromate. The type strain, K1664^{T}, was isolated from the eye of a patient with infectious keratitis.

Acknowledgments:

This study was funded by the Joanne and Arthur Haberberger Fellowship and the Pennsylvania Academy of Sciences Undergraduate Research Fellowship. These contributions supported critical parts of the project, and for that, we are very grateful. The Venn Diagram Generator and AAI Calculator used in this work were developed by Thomas Sontag, Andrew Gale, Jordan Krebs and Eileen Peluso.

References:

Angelos JA, Spinks PQ, Ball LM, George LW. Moraxella bovoculi sp. nov., isolated from calves with infectious bovine keratoconjunctivitis. *Int. J. Syst. Evol. Microbiol*. 2007;57:789-795.

Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM. The many faces of the helix-turn-helix domain: Transcription regulation and beyond. 2005;29:231–262. https://doi.org/10.1016/j.fmrre.2004.12.008

Aziz RK, Bartels D, Best A, DeJongh M, Disz T *et al.* The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics.* 2008;9:1–15.

Buchman AL, Pickett MJ. Moraxella atlantae bacteraemia in a patient with systemic lupus erythematosis. J Infect 1991;23:197-9 doi: 10.1016/0163-4453(91)92335-3

Chun, J and Rainey FA. Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. Int J Syst Evol Microbiol 64, 316-24 (2014).

Dawson, C. Follicular conjunctivitis. Conjunctivitis. In Clinical Ophthalmology External Disease; Duane, T.D., Ed.; Harper and Row: Philadelphia, PA, USA, 1983; pp. 1–19.

Durrani AF, Faith SC, Kowalski RP, Yu M, Romanowski E *et al*. Moraxella keratitis: Analysis of risk factors, clinical characteristics, management and treatment outcomes. Am J Ophthalmol 2019;197:17–22. doi: 10.1016/j.ajo.2018.08.055

Embers ME, Doyle LA, Whitehouse CA, Selby EB, Chappell M et al. Characterization of a Moraxella species that causes epistaxis in macaques. Vet Microbiol 2011;147:367–375. doi: 10.1016/j.vetmic.2010.06.029

Epling J. Bacterial conjunctivitis. BMJ Clin Evid 2012; 2012:0704.

Gaggar A, Jackson PL, Noerager BD, O'Reilly PJ, McQuaid DB *et al.* A novel proteolytic cascade generates an extracellular matrix-derived chemoattractant in chronic neutrophilic inflammation. *J Immunol* 2008;180: 5662–5669. doi: 10.4049/jimmunol.180.8.5662

Goldstein BP. Resistance to rifampicin: a review. J. Antibiot. 2014;67:625-630. doi:10.1038/ja.2014.107.

Juni E, Bøvre K. Moraxella. In Bergey's Manual of Systematics of Archaea and Bacteria; 2015. doi:https://doi.org/10.1002/9781118960608.gbm01204

Karimi Z, Ahmadi A, Najafi A, Ranjbar R. Bacterial CRISPR Regions: General Features and their Potential for Epidemiological Molecular Typing Studies. Open Microbiol J 2018;12:59-70. doi: 10.2174/1874285801812010059

LaCroce SJ, Wilson MN, Romanowski JE, Newman JD, Jhanji V *et al*. Moraxella nonliquefaciens and M. osloensis Are Important Moraxella Species That Cause Ocular Infections. Microorganisms 2019;7:163. doi: 10.3390/microorganisms7060163

Leclercq R. Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. Clin Infect Dis 2002;34:482-492. doi: 10.1086/324626

Lee I, Kim YO, Park SC, Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol.* 2016;66(2):1100–1103.

Matagne A, Lamotte-Brasseur J, Frère JM. Catalytic properties of class A beta-lactamases: efficiency and diversity. Biochem J 1998;330:581-598. doi: 10.1042/bj3300581

Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics.* 2013;14.

Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 2014;42(D1):206–14.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 2015;25:1043-1055. doi:10.1101/gr.186072.114

Pettersson B, Kodjo A, Ronaghi M, Uhlén M, Tønjum T. Phylogeny of the family Moraxellaceae by 16S rDNA sequence analysis, with special emphasis on differentiation of Moraxella species. Int J Syst Bacteriol 1998;48 Pt 1:75–89 doi:10.1099/00207713-48-1-75

Putnam CM. Diagnosis and management of blepharitis: an optometrist's perspective. Clin Optom 2016;8:71–78. doi: 10.2147/OPTO.S84795

Reddy CA, Beveridge TJ, Breznak JA, Marzluf G, Schmidt TM, Snyder LR, editors. Methods for General and Molecular Microbiology. 3rd ed. ASM Press; 2007.

Sasser M. Technical Note # 101 Microbial Identification by Gas Chromatograph ic Analysis of Fatty Acid Methyl Esters (GC-FAME). 2009;(May 1990):1–6.

Whitcher JP, Cevallos V, Moraxella CV. Moraxella, down but not out--the eye bug that won't go away. Br J Ophthalmol 2006;90:1215–1216. doi:10.1136/bjo.2006.100339

Yoon SH, Ha S min, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek.* 2017;110(10):1281–6.