Phylogenomic Characterization of a Novel *Corynebacterium* Species Associated with Fatal Diphtheritic Stomatitis in Endangered Yellow-Eyed Penguins

Sarah C. Saunderson,[a] Imen Nouioui,[b,c] Anne C. Midwinter,[d] David A. Wilkinson,[d] Melanie J. Young,[e,f,g] Kate M. McInnes,[e] Jim Watts,[g] Vartul Sangal[h]

[a]Department of Microbiology and Immunology, School of Biomedical Sciences, University of Otago, Dunedin, New Zealand
[b]Leibniz-Institut DSMZ–Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany
[c]School of Natural and Environmental Sciences, Faculty of Science, Agriculture, and Engineering, Newcastle University, Newcastle upon Tyne, United Kingdom
[d]Molecular Epidemiology and Veterinary Public Health Laboratory (mEpiLab), Infectious Disease Research Centre, School of Veterinary Science, Massey University, Palmerston North, New Zealand
[e]Science and Capability, Department of Conservation, Wellington, New Zealand
[f]Department of Zoology, University of Otago, Dunedin, New Zealand
[g]Coastal Otago District Office, Department of Conservation, Dunedin, New Zealand
[h]Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, United Kingdom

ABSTRACT Yellow-eyed penguins, *Megadyptes antipodes*, are an endangered species that are endemic to New Zealand. Outbreaks of diphtheritic stomatitis have caused significant mortality for this species, especially among young chicks. In this study, we isolated 16 *Corynebacterium* sp. isolates from the oral cavities of 2- to 14-day-old chicks at a range of infection stages and sequenced the genomes to understand their virulence mechanisms. Phylogenomic and matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) characterization indicate that these strains belong to a novel *Corynebacterium* species. A simple multiplex PCR-based diagnostic assay has been developed to identify these strains rapidly and reliably. Similar to other corynebacteria, genomic islands and prophages introduced significant diversity among these strains that has potentially led to minor functional variations between the two lineages. Despite the presence of multiple corynebacterial virulence genes and a *spaDEF*-type pilus gene cluster among these strains, the survival rate was much higher in *Galleria mellonella* larvae than in those inoculated with *Corynebacterium ulcerans* NZRM 818 and *Corynebacterium pseudotuberculosis* NZRM 3004. Therefore, these strains are opportunistic pathogens causing high mortality among young penguin chicks due to a less-developed immune system.

IMPORTANCE Yellow-eyed penguins, *Megadyptes antipodes*, are endangered species with a sharp decline in the numbers of breeding pairs over the last 2 decades. Diphtheritic stomatitis, characterized by a thick fibrinopurulent exudate in the oral cavities and symptoms, including inanition and significant weight loss, is responsible for significant mortality among the young chicks. These chicks are treated with antibiotics, amoxicillin-clavulanic acid or enrofloxacin, but do not always recover from the infection. The pathogen causing these infections and the mechanism of pathogenesis are unclear. This study has identified a novel *Corynebacterium* species to be associated with diphtheritic stomatitis in yellow-eyed penguins with potential virulence genes that are likely involved in pathogenesis. Importantly, a gene encoding an exotoxin, phospholipase D, is present among these strains. The inactivated form of this enzyme could potentially be used as an effective vaccine to protect these penguins from infection.

Yellow-eyed penguins (*Megadyptes antipodes*), or hoiho (te reo Māori), are one of the world's rarest penguin species and are endemic to New Zealand and its outlying islands (1). Over the last 2 decades, the numbers of breeding pairs have significantly declined due to a range of factors, including climate change (2), direct and indirect conflicts with commercial fisheries (3, 4), starvation events, and unexplained adult mortalities (5). Underpinning this decline in adult and juvenile numbers are mass mortality events affecting chick production (5–7), including starvation and oral infections known as diphtheritic stomatitis (5). Diphtheritic stomatitis is responsible for high mortality rates among 1- to 28-day-old chicks, which is characterized by a thick cream-colored fibrinopurulent exudate in the oral cavities and symptoms, including inanition and significant weight loss leading to death (7, 8). Corynebacterial strains, identified as *Corynebacterium amycolatum*, were previously isolated from the oral lesions of infected chicks (8). However, similar strains were also isolated from the oral cavity of healthy adult penguins (8). While the presence of avipoxvirus and an unidentified virus-like agent have been detected among the infected chicks (8), their roles with regard to infection and mortality remain unclear.

First described by Lehmann and Neumann in 1896, the genus *Corynebacterium* comprises >130 species and 11 subspecies (9, 10). *Corynebacterium* strains are Gram-stain-positive rod- or club-shaped bacteria characterized by genomic G+C content between 46 and 74 mol% (9). Several corynebacterial species are of biotechnological, medical, or veterinary importance (9, 10). In this study, we report the findings on novel *Corynebacterium* strains isolated from lesions derived from the oral cavities of 2- to 14-day-old yellow-eyed penguin chicks. The phylogenomic characterization identified two distinct lineages with potential virulence genes, which may be associated with diphtheritic stomatitis.

RESULTS

Clinical information related to bacterial isolates. Yellow-eyed penguin chicks (2 to 14 days old) at a range of diphtheritic stomatitis infection stages were identified within breeding regions of the Otago Peninsula (Table 1; Fig. 1; Fig. S1). Lesion exudate was swabbed in 10 chicks presenting with oral lesions consistent with the disease syndrome. The oral cavity of five chicks with no visible lesions were also swabbed from nests that were classed as infected due to a previous or an ongoing infection. In particular, two chicks, YEP-20 and YEP-50, previously possessed lesions that were debrided less than a week before the sampling. One chick, YEP-52, shared a nest with a twin chick that had a previous lesion debridement recorded. All three chicks developed lesions within 1-week postsampling. No previous lesions were observed in the twin chicks (YEP-5 and YEP-6) from the same nest; however, the former was found deceased of unknown causes 3 days postsampling, while the latter chick developed lesions 22 days later. The lesion status and nest details for one sample (YEP-12) are not recorded.

Small circular bacterial white-cream-color colonies appeared on the blood agar plates after 16 h of incubation at 37°C (Fig. S2). These isolates were Gram-positive coccobacilli, and a BLAST search of the nucleotide sequence of the 16S rRNA gene extracted from these genomes in GenBank (https://blast.ncbi.nlm.nih.gov/) revealed 98.7% similarity with a partial sequence (92% coverage) of 16S rRNA gene of *Corynebacterium ciconiae* strain BS13^{T}. However, these strains could not be assigned to any known species by matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) analysis, indicating that they are potentially novel *Corynebacterium* species.

Phylogenomic diversity between *Corynebacterium* sp. isolates. The size of genome assemblies of all 16 isolates from yellow-eyed penguins varied from 2.4 to 2.5 Mb with a GC content of 62.5 to 62.9 mol% and 54- to 99-fold average coverage

TABLE 1 Details of genome assemblies of *Corynebacterium* sp. isolates from penguins

Strain[a]	Breeding area	Chick ID	Swab site	Lineage	Approx coverage (×)	N_{50} (bp)	No. of contigs	Total size (Mb)	GC content (%)	No. of CDS[b]	No. of tRNAs	No. of tmRNAs
3B	Highcliff	YEP-3	Lesion	1	64	241,725	16	2.45	62.9	2,138	52	1
5A†	Highcliff	YEP-5	Oral cavity	1	82	217,541	15	2.40	62.9	2,069	52	1
6A†	Highcliff	YEP-6	Oral cavity	1	75	241,723	16	2.45	62.9	2,132	52	1
19B	Midsection	YEP-19	Lesion	1	91	294,621	14	2.45	62.9	2,131	52	1
48B¶	Okia	YEP-48	Lesion	1	54	174,460	19	2.40	62.9	2,071	52	1
49B¶	Okia	YEP-49	Lesion	1	68	199,773	17	2.40	62.9	2,071	52	1
50A*	Okia	YEP-50	Oral cavity	1	88	241,730	16	2.40	62.9	2,070	52	1
51B*	Okia	YEP-51	Lesion	1	72	280,374	12	2.40	62.9	2,070	52	1
52A	Okia	YEP-52	Oral cavity	1	82	241,719	18	2.40	62.9	2,067	52	1
71B	Highcliff	YEP-71	Lesion	1	68	252,028	15	2.45	62.9	2,130	52	1
73A	Highcliff	YEP-73	Lesion	1	82	240,796	15	2.40	62.9	2,069	52	1
7B	Highcliff	YEP-7	Lesion	2	59	215,111	17	2.46	62.7	2,106	51	1
12B	Unknown	YEP-12	Unknown	2	89	289,014	16	2.46	62.7	2,107	51	1
20A	A1 section	YEP-20	Oral cavity	2	84	254,507	17	2.50	62.7	2,072	51	1
74A	A1 section	YEP-74	Lesion	2	99	437,349	11	2.46	62.7	2,105	51	1
11A	Highcliff	YEP-11	Lesion	2	63	135,658	24	2.40	62.5	2,075	52	1

[a]Strains isolated from the same nests are labeled with the same symbols.
[b]CDS, coding DNA sequence.

(Table 1). The draft assemblies were annotated with 2,067 to 2,138 genes, 51 to 52 tRNAs, and 1 transfer-messenger RNA (tmRNA) gene (Table 1). Whole-genome sequences of these strains have been submitted to GenBank (accession numbers PQMG00000000 to PQMV00000000; Table S1).

The pan-genome contains 2,272 genes, with 1,962 shared by all of the strains, and 310 genes belonged to the accessory genome. The phylogenetic tree from the core genome separated the strains into two distinct lineages. These isolates were collected from yellow-eyed penguin chicks from four nest sites (Table 1; Fig. S1). Lineage 1 encompassed 11 of the 16 isolates from three nest sites, whereas the remaining five isolates from two sites grouped in lineage 2 strains (Fig. 2; Table 1). Interestingly, all the isolates collected from yellow-eyed penguins at the Victory Beach (Okia Reserve) and most of the isolates from Highcliff clustered in lineage 1, whereas both the isolates from the A1 section were restricted to lineage 2. This indicates a potential association between the lineages and the sampling sites; however, more samples need to be analyzed to establish such an association.

Lineage 1 is a relatively homogenous in comparison to lineage 2, where one isolate, 11A, is clearly distinct from the other isolates (Fig. 2). A total of 51 genes were specific to lineage 1, 28 genes with predicted functions and 23 encoding hypothetical proteins (Table S2). Similarly, 61 genes were unique to lineage 2 with 32 encoding hypothetical proteins (Table S2). Some lineage-specific genes with known functions encode cellular enzymes; e.g., the gene encoding β-glucosidase A is only present in lineage 1 isolates (Table S2). This gene is involved in carbohydrate metabolism, particularly hydrolyzing short-chain oligosaccharides and cellobiose (11). Similarly, a gene encoding acetaldehyde dehydrogenase is specific to lineage 2 which is absent among lineage 1 strains (Table S2). Acetaldehyde dehydrogenase protects cells from cytotoxicity by oxidizing acetaldehyde to acetyl-CoA (12). Such variations potentially indicate minor variations in metabolic activities between the two lineages.

Genome plasticity. Genomic islands (GIs) and prophages are often the sources of genome plasticity among pathogenic *Corynebacterium* species (13–16). A total of 15 GIs were identified among 16 *Corynebacterium* isolates from yellow-eyed penguins (Table S2). Five of the GIs (GI-1, GI-3, GI-6, GI-7, and GI-10) were conserved among all the strains. GI-1 encompasses genes involved in iron transport, including an anguibactin system regulator that plays an important role in virulence in vibrios (17, 18). The siderophore anguibactin system is located on a virulence plasmid in the fish pathogen *Vibrio anguillarum* (18). It is possible that these *Corynebacterium* strains have horizontally acquired this system from other marine pathogens. GI-1 also contains genes
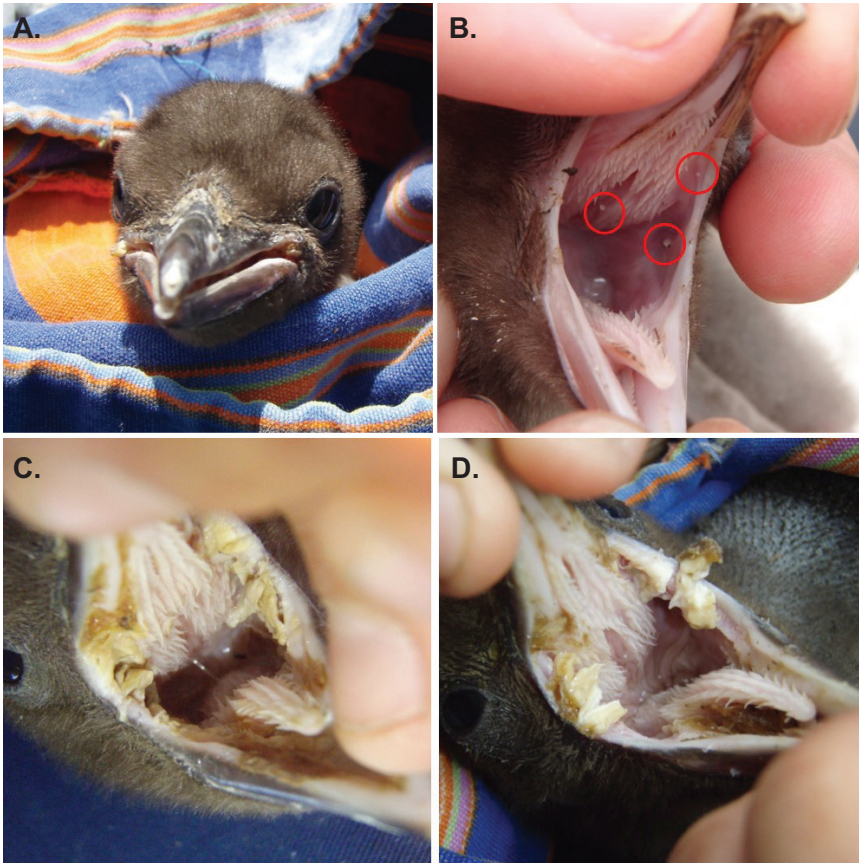
FIG 1 Fibrinopurulent diphtheria lesions in yellow-eyed penguin chicks. (A) Lesions at the commissures of the beak, which prevent closure of the oral cavity. (B) Red circles denote minor "pinpoint" lesions in the gape and choanal slit that are firmly adhered to the mucosa before overgrowth of fibrinopurulent lesion exudate and mucosa ulceration. (C and D) Extensive fibrinopurulent diphtheria lesions in the oral cavity of infected yellow-eyed penguin chicks, including both sides of the gape and under the tongue. Photos: M. J. Young.

encoding polyketide synthase, PksR, and phthiocerol synthesis polyketide synthase type I PpsE (Table S2) that are involved in diverse metabolic activities, including biosynthesis of mycolic acids (19). Similarly, genes on GI-3 encoded proteins that are involved in various cellular activities, including transcriptional regulator ClgR, membrane chaperone PspA/IM30 family protein, energy-coupling factor transporter transmembrane protein EcfT, biotin transport ATP-binding protein BioM, biotin transporter BioY, a hypothetical protein, and recombinase A (Table S2). GI-6 included genes encoding enzymes for carbohydrate metabolism, whereas GI-7 and GI-10 carried Cas (CRISPR-associated) genes (Table S2).

Most of the genes on lineage 1-specific GI-2 and GI-8 encoded hypothetical proteins, whereas GI-4 carries the genes involved in DNA metabolism, including a type-III restriction modification system (Table S2). This locus is present between the genes encoding bifunctional protein PutA and a hypothetical protein (locus tags in strain 3B, hoi102_00282 and hoi102_00288, respectively) that are also conserved among lineage 2 isolates, except for the latter gene, which is missing in strain 11A. A small genomic island with four genes, GI-5, followed the gene encoding hypothetical protein in four lineage 2 strains, with two genes encoding ascorbate-specific phosphotransferase enzyme IIA component and ascorbate-specific permease IIC component UlaA, respectively, and the remaining two encoding hypothetical proteins (Table S2). Similarly, GI-13 carried some Cas genes that were specific to lineage 2 strains except for strain 11A.

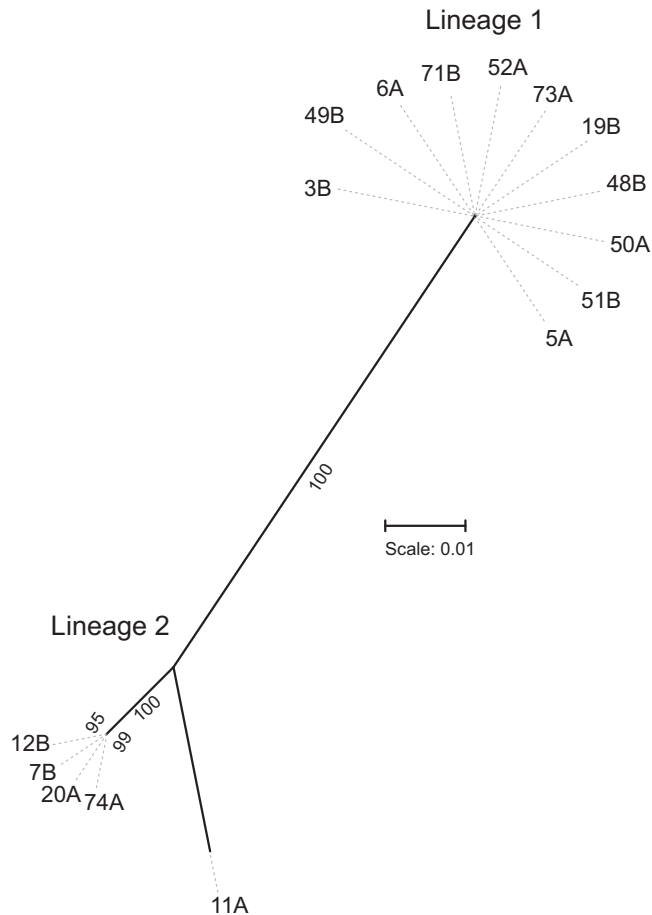A majority of genes on lineage 2-specific GI-11 and GI-14 encoded hypothetical

Lineage 1

FIG 2 A maximum-likelihood tree from the core genome of *Corynebacterium* isolates from yellow-eyed penguins. The scale bar represents nucleotide substitutions per nucleotide site.

proteins, whereas GI-12 contained several genes involved in oxidation of aromatic compounds, including the subunits of 3-phenylpropionate/cinnamic acid dioxygenase, ferredoxin-NAD(+) reductase subunit of benzene 1,2-dioxygenase, biphenyl dioxygenase subunit beta, cis-2,3-dihydrobiphenyl-2,3-diol dehydrogenase, and a long-chain fatty acid-CoA ligase, FadD13 (Table S2). GI-11 has been identified as a prophage (partial; Φ3) that has integrated between *attL* and *attR* sites in the genome.

GI-15 was only present in strain 11A, which included two genes encoding adenylate and guanylate cyclase catalytic domain proteins and two hypothetical proteins. GI-9, a large genomic island carrying a number of genes coding for phage-related proteins or hypothetical proteins was present in four lineage 1 isolates (3B, 6A, 19B, and 71B) and strain 11A of lineage 2. Indeed, a large part of this island has also been identified as an incomplete prophage, Φ2 (Table S2).

In addition, a 6.2- to 6.3-Kb conserved region was predicted as a potential incomplete prophage that carried the genes encoding sporulation initiation inhibitor protein (ParA family), chromosome-partitioning protein ParB, a hypothetical protein, *N*-acetylmuramoyl-L-alanine amidase LytC precursor, thioredoxin, thioredoxin reductase, and extracytoplasmic function (ECF) RNA polymerase sigma factor SigM (Φ1; Table S2). ParA and ParB proteins are involved in DNA separation and partitioning and are commonly present on bacterial chromosomes, plasmids, and bacteriophages (20–22). Thioredoxin and thioredoxin reductase are involved in a number of cellular activities, including response to oxidative stress (23). In bacteriophages, thioredoxin is important for the phage assembly (24). LytC is an

TABLE 2 Corynebacterial virulence-associated genes identified among penguin isolates

Gene	Function	Loci in strain 3B[a]
cwlH	Cell wall-associated hydrolase	hoi102_01530; hoi102_01529
rpfI	Peptidoglycan endopeptidase (RipA)	hoi102_00347
rpfA	Resuscitation-promoting factor A	hoi102_02179
rpfB	Resuscitation-promoting factor B	hoi102_01747
nrpS2	Nonribosomal peptide synthetase	hoi102_00519
dtsR1	Acyl/propionyl-CoA carboxylase b-subunit	hoi102_01834
dtsR2	Acyl/propionyl-CoA carboxylase b-subunit	hoi102_01833
accD3	Acyl/propionyl-CoA carboxylase b-subunit	hoi102_01113
	Secreted protein/subtilisin E precursor	hoi102_00965

[a]The locus tags of only one penguin isolate are mentioned, but these genes are present among all 16 isolates across both lineages.

autolysin responsible for hydrolysis of *N*-acetylmuramoyl-L-alanine linkage in peptidoglycan associated with virulence in *S. pneumoniae* (25). Overall, recombination has played a key role in introducing diversity among these corynebacterial strains, and the proteins encoded on these GIs are likely responsible for variations in cellular and metabolic activities both between and within each lineage.

Virulence characteristics. Surface pili are important for adhesion and invasion of the host cells (26–28). These pili are encoded by *spa* (sortase-mediated pilus assembly) gene clusters, and the number of pilus gene clusters varies between one and three among pathogenic corynebacteria (13–15, 28, 29). One pilus gene cluster has been observed among the genomes of *Corynebacterium* isolates from yellow-eyed penguins (*hoi102_00558* to *hoi102_00561* in strain 3B). The genes *hoi102_00558* and *hoi102_00560* encode 443-amino acid (aa)-long and 510-aa-long fimbrial subunits, respectively, and the *hoi102_00561* gene encodes a cell wall anchor domain-containing protein (891 aa). *hoi102_00559* is a class C sortase that is responsible for assembling the pilus. An additional class E sortase is present elsewhere in the genome (*hoi102_01169*), which is also conserved among all isolates.

The gene encoding diphtheria-like toxin is absent among these isolates; however, two genes encode phospholipase YtpA, and one gene encoding phospholipase D precursor may be responsible for virulence characteristics of these strains (30). Phospholipase D is a key virulence factor in *Corynebacterium pseudotuberculosis* and *Corynebacterium ulcerans* (29, 31, 32). However, the predicted phospholipase D precursor among penguin isolates (*hoi102_01010* in strain 3B) is relatively larger in size (577 aa) than those present in *C. pseudotuberculosis* and *C. ulcerans* (307 aa). A search in the Conserved Domain Database (33) confirmed that *hoi102_01010* belongs to the PhoD super family (activity, phosphodiesterase/alkaline phosphatase D; accession number cl26056). A BLAST search of the gene in the UniProt proteome database (https://www.uniprot.org/) showed 54% similarities with the phospholipase D genes of *Corynebacterium geronticis*, *Corynebacterium pelargi*, and *Corynebacterium pseudopelargi*. Phospholipase D is a member of the PhoD family, a diverse superfamily of enzymes that cleaves the phosphodiester bond linking the head group and diacyl glycerol moieties in phospholipids and is also involved in wide range of cellular processes (34–37).

A protein BLAST search of the virulence-associated genes from pathogenic corynebacteria identified two copies of cell wall-associated hydrolase, a gene encoding peptidoglycan endopeptidase RipA protein (*rpfI*), and genes encoding resuscitation-promoting factors, RpfA and RpfB, among all the isolates from yellow-eyed penguins (Table 2). These proteins are important for cell division and cell surface organization and likely play an important role in adhesion to the host cells and internalization of the pathogen (29, 38–42). In addition, genes encoding corynebacterial/nonribosomal peptide synthetase (linear gramicidin synthase subunit D) and subunits of acyl/propionyl-CoA carboxylase (DtsR1, DtsR2, and AccD3) are also identified in all yellow-eyed penguin isolates (Table 2). These genes were found to be upregulated during the macrophage infection by *C. pseudotuberculosis* (43). Moreover, *dtsR1*, *dtsR2*, and *accD3* genes are involved in

fatty acid and mycolic acid biosynthesis, which contribute to the virulence characteristics (44, 45). The virulence genes also include subtilisin, a protease potentially involved in hypoxic response during the colonization of the respiratory tract (46).

The virulence potential of isolates from yellow-eyed penguins was tested in a *Galleria mellonella* model using *C. ulcerans* NZRM 818 and *C. pseudotuberculosis* NZRM 3004 as positive controls and nonpathogenic *Corynebacterium* sp. strain NZRM 2522 and phosphate-buffered saline (PBS) as negative controls. All the larvae injected with *C. ulcerans* and *C. pseudotuberculosis* died within 48 h (>90% within 24 h; Table S3). In contrast, only one of the 30 caterpillars inoculated with strain NZRM 2522, test strains 3B and 6A, and two larvae inoculated with PBS died during the experiment. Similarly, no mortality was recorded for other isolates from penguins except for strains 11A and 71B, where 1/10 larvae and 3/10 larvae died, respectively (Table S3). However, only *Bacillus thuringensis* strains were isolated from these dead caterpillars. Therefore, *Corynebacterium* isolates from yellow-eyed penguins appears to be nonpathogenic in the caterpillar challenge model.

Taxogenomic characteristics. A maximum-likelihood (ML) tree constructed from corynebacterial 16S rRNA sequences obtained from GenBank distinctly separated isolates from penguins that were closely related to *C. ciconiae* strain BS13^T (DSM 44920^T; Fig. 3). Similarly, these novel isolates formed a distinct clade close to *C. ciconiae* DSM 44920^T in the ML tree constructed from 184 core genes, including representative *Corynebacterium* strains from GenBank (Fig. 4; Table S1). Principal-component analysis of MALDI-TOF spectra clustered these isolates distinctly from *C. ciconiae* and *Corynebacterium trachiae* (Fig. 5). Representative averaged MALDI-TOF spectra showed highly discriminant peaks between *C. ciconiae*, *C. trachiae*, and the penguin isolates (Fig. S3). Peaks at *m/z* values of 617, 6,484, and 11,224 showed positive associations with lineage 1 and lineage 2 strains with only subtle differences that could reliably distinguish them. A phenotypic characterization of representative strains from both the lineages, 3B and 7B, also revealed similar biochemical properties except for minor variations (Table 3). Strain 7B was able to produce *N*-acetyl-β-glucosaminidase and was able to utilize D-mannose, D-maltose, and malic acid that were not detected in strain 3B (Table 3). Strain 7B was also weakly positive for adipic acid fermentation.

Mean fragmented BLAST scores (FBS values) between the penguin isolates and *C. ciconiae* were 30.5 ± 0.9, which is significantly higher than the values between the penguin isolates and other corynebacterial genomes (1.9 ± 0.7; Data Set S1A). Average BLAST-based nucleotide identities (ANIb values) between the penguin isolates and *C. ciconiae* strains were 82.6 ± 0.1, again higher than those observed between the novel isolates and other corynebacterial strains (69.8 ± 1.1; Data Set S1B). Digital DNA-DNA hybridization (dDDH) values between the genomic sequences also suggest isolates from yellow-eyed penguins to be a novel species (Data Set S1C).

The dDDH values between strains within lineage 1 and lineage 2 were >99.7% and 83.3%, respectively. However, the pairwise dDDH values between individuals of the two lineages were slightly lower than the suggested cutoff value of 70% and varied between 63.2 and 66.7% (Fig. 6A; Data Set S1D), indicating that these lineages may represent two different species. However, average FBS and ANIb values between the two lineages are 83.6 ± 1.2 (81.2 to 85.5%; Fig. 6B; Data Set S1A) and 95.8 ± 0.2 (95.5 to 96.0%; Fig. 6C; Data Set S1B), respectively. Average amino acid identities (AAI) between the two lineages were also above the species threshold and varied between 97.4 and 97.7% (97.5 ± 0.8; Fig. 6D; Data Set S1E). These FBS, ANIb, and AAI values are consistent with lineage 1 and lineage 2 strains belonging to the same species (47, 48).

PCR-based diagnostic assay. Two genes specific to lineage 1 encoding a hypothetical protein and α/β hydrolase family protein, respectively, and two genes specific to lineage 2, both encoding hypothetical proteins (Table S4), did not show significant similarities with other *Corynebacterium* genomes and were selected for the diagnostics assay. These loci were amplified in a multiplex PCR in all isolates from yellow-eyed penguins, including two strains of each from *C. diphtheriae* and *C. ulcerans*. All 11 lineage 1 strains produced two amplicons of 219 bp and 274 bp, while all lineage 2 strains
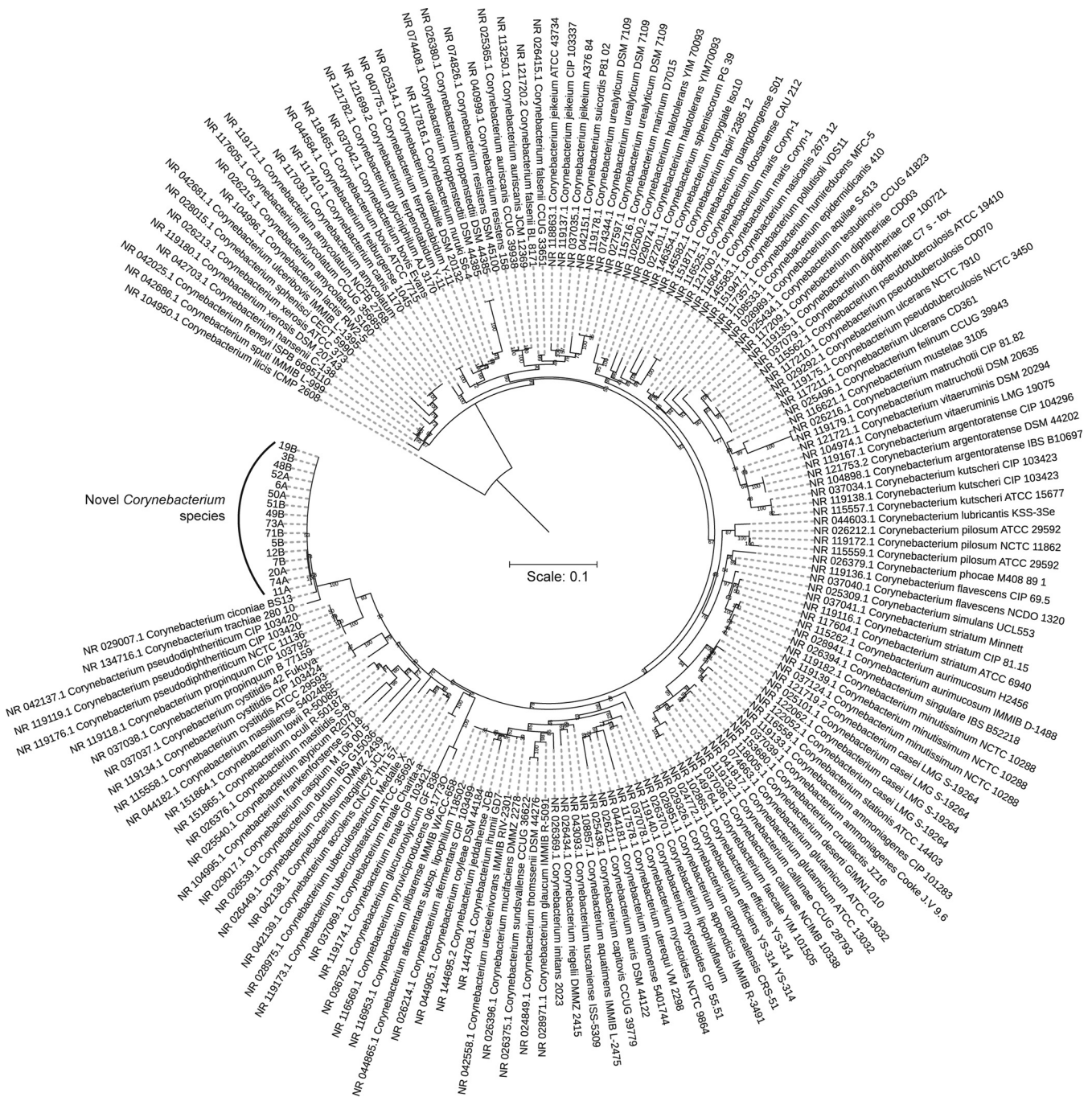
FIG 3 A maximum-likelihood tree from the nucleotide sequences of 16S rRNA genes of *Corynebacterium* strains obtained from GenBank. The scale bar represents nucleotide substitutions per nucleotide site.

produced amplicons of 425 bp and 486 bp (Fig. S4). None of the four loci were amplified among *C. diphtheriae* or *C. ulcerans* strains.

DISCUSSION

Corynebacterium sp. isolates from yellow-eyed penguins belong to a novel species. *Corynebacterium* is a diverse genus that encompasses species of industrial, veterinary, or medical importance (9). In this study, we identified a novel *Corynebacterium* species isolated from yellow-eyed penguins in New Zealand. Phylogenetic analyses of rRNA gene sequences and the core genome revealed these strains to be quite distinct from other corynebacteria, with *C. ciconiae* being the closest relative (Fig. 3 and 4).
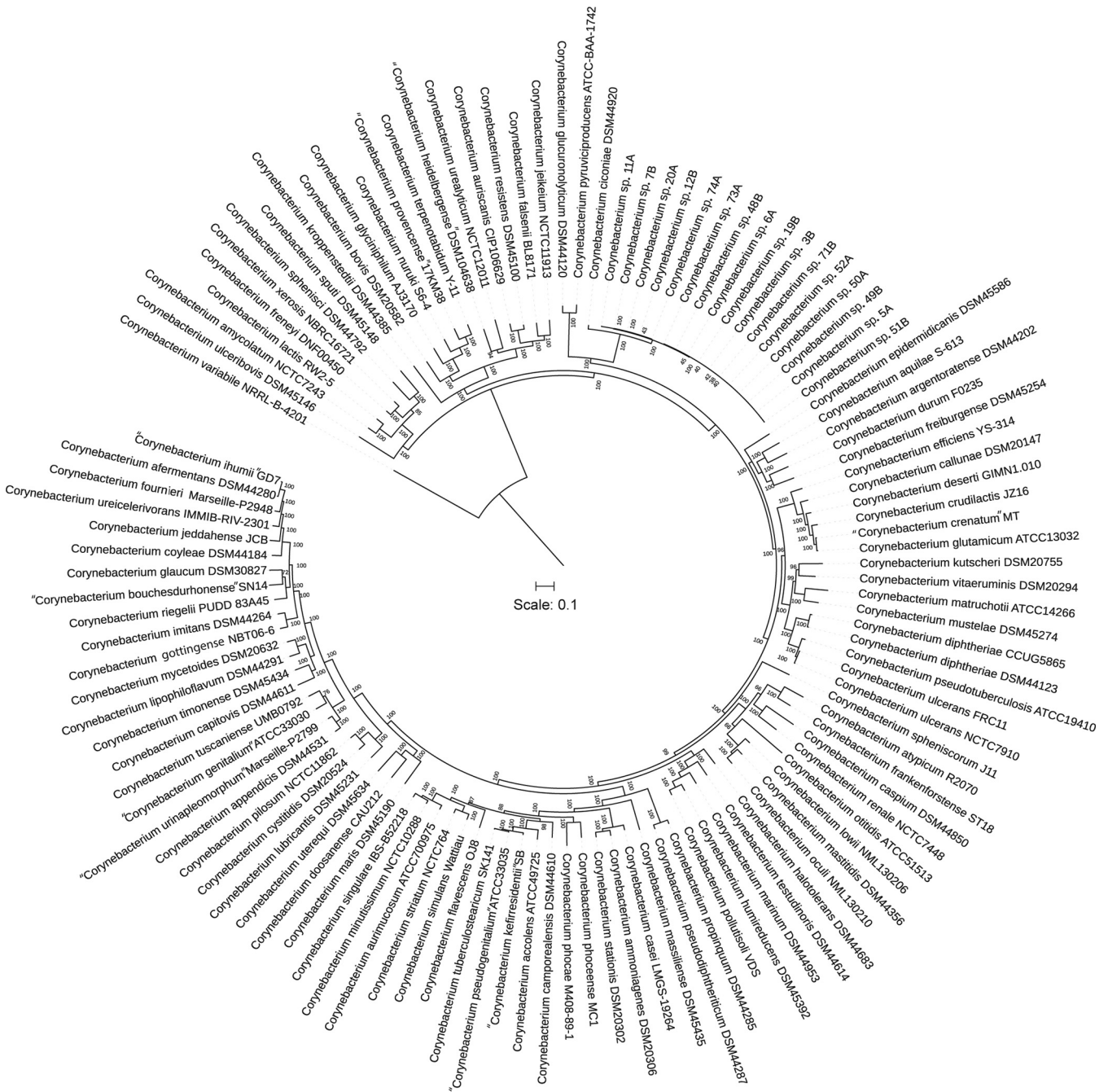
FIG 4 A maximum-likelihood tree from the alignment of 184 core proteins of 117 corynebacterial genomes (Table S1). The scale bar represents amino-acid substitutions per site. Some species designations in the genomic data obtained from GenBank are pending validation. These species designations are mentioned in quotation marks.

MALDI-TOF spectral analyses distinctly separated these strains from the type strains of *C. ciconiae* and *C. trachiae* (Fig. 5 and Fig. S3), and genome-based taxonomic analyses, FBS (48), and ANIb values (47–50), also indicated these *Corynebacterium* strains from yellow-eyed penguins to be a novel species (Data Set S1A-C).

Phylogenomic analyses of these strains revealed two lineages within this species (Fig. 2), which is consistent with the ANIb and AAI values between individuals of both the lineages (species threshold, ≥95%; Fig. 6C and D; Data Set S1B and S1E). However, dDDH values were slightly below the threshold of ≥70% (Fig. 5A; Data Set S1D) (51, 52). The taxogenomic thresholds should be considered with flexibility in combination
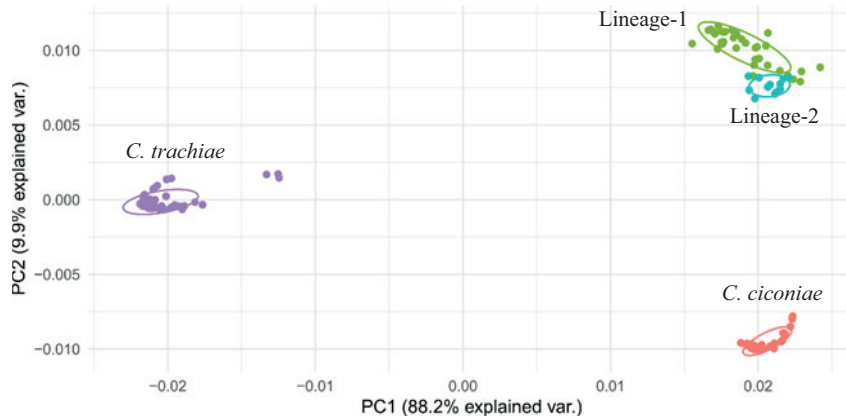
FIG 5 Principal-component analysis of MALDI-TOF peak intensity profiles.

with other phylogenetic and biological properties to define species (53). The gene content is highly conserved between the two lineages (Table S2) with only minor differences in the biochemical properties (Table 3) and in MALDI-TOF spectra (Fig. 5; Fig. S3). The strains of both the lineages are associated with diphtheritic stomatitis with similar clinical manifestations and possess the same set of virulence-associated genes (Table 3). Therefore, these lineages could potentially be assigned to a single species. A manuscript is currently being prepared with a detailed description of this novel *Corynebacterium* species, including a proposal to name lineage 1 and lineage 2 as subspecies. Two type strains, 3B^T and 7B^T, representatives of lineages 1 and 2, have been deposited at the DSMZ German collection of microorganisms and cell cultures (strains DSM 111184 and DSM 111183) and at the New Zealand Reference Culture Collection (strains NZRM 4755 and NZRM 4756), respectively.

Both the lineages can easily be identified in a simple multiplex PCR (Fig. S4). A BLAST search of the entire NCBI nucleotide database did not return any significant similarities of these loci with any other genomes, confirming their suitability for diagnostics of these lineages. This will also eliminate the need to run time- and labor-intensive biochemical tests for identifying these strains in the laboratory, which may expedite the treatment of the affected yellow-eyed penguin chicks.

Opportunistic pathogens to young penguin chicks. Lineage 1 is more prevalent among the nests than lineage 2, the latter being more common to nests at the site A1 section (Table 1; Fig. S1). Also, the pairs of twin chicks (YEP-5 and YEP-6; YEP-48 and YEP-49; YEP-50 and YEP-51) harbored identical lineages, which suggests each nest is infected by a single prevalent lineage. Both the lineages are associated with diphtheritic stomatitis, which has a high infection and mortality rate among mainland New Zealand populations of neonate yellow-eyed penguin chicks (7, 8). Adhesion and invasion of the host cells are important for infection and are facilitated by surface pili (54). These pili are generally composed of major subunit (shaft pilin), minor subunit (base pilin), and tip pilin that are assembled by sortases (54, 55). Hoi102_00559 showed 49% sequence identity (89% query coverage) with the *srtC* gene, and Hoi102_00560 showed 31% identity (81% query coverage) with SpaD, a major pilin subunit, of *C. diphtheriae*. Hoi102_00558 is similar to the minor pilin subunit, SpaI (27% query coverage and 31% identity). The Hoi102_00561 protein is an LPXTG cell wall anchor domain-containing protein also identified as von Willebrand factor type A domain protein. The von Willebrand factor type A domain is also present in the tip protein of SpaDEF pili, SpaF; therefore, isolates from yellow-eyed penguins possess a *spaDEF*-type pilus gene cluster which is reported to interact with laryngeal and lung epithelial cells (54, 55).

Several corynebacterial virulence genes are present among these strains (Table 2), including two genes encoding phospholipase YtpA, which is reported to be involved

TABLE 3 Biochemical properties of representative strains using API ZYM and API 20NE systems[a]

Biochemical test	Isolate 3B	Isolate 7B
API ZYM		
Esterase (C4)	−	−
Esterase lipase (C8)	−	−
Lipase (C14)	+	+
Leucine aminopeptidase	+	+
Valine aminopeptidase	−	−
Cystine aminopeptidase	−	−
Trypsin	−	−
A-chymotrypsin	−	−
Acid phosphatase	−	−
Naphthol-AS-BI-phosphohydrolase	−	−
α-Galactosidase	+	+
β-Galactosidase	−	−
β-Glucuronidase	−	−
α-Glucosidase	−	−
β-Glucosidase	−	−
N-acetyl-β-glucosaminidase	−	+
α-Mannosidase	−	−
α-fucosidase		−
API 20NE		
Potassium nitrate	−	−
L-Tryptophane	−	−
D-Glucose	−	−
L-Arginine	−	−
Urea	−	−
Esculin ferric citrate	−	−
Gelatin (bovine origin)	−	−
4-Nitrophenyl-β-D-galactopyranoside	−	−
D-Glucose	+	+
L-Arabinose	−	−
D-Mannose	−	+
D-Mannitol	−	−
N-acetyl-glucosamine	−	−
D-Maltose	−	+
Potassium gluconate	−	−
Capric acid	−	−
Adipic acid	−	W
Malic acid	−	+
Trisodium citrate	−	−
Phenylacetic acid	−	−

[a]+, positive reaction; −, negative reaction; W, weak positive.

in the biosynthesis of bacilysocin, which competitively inhibits growth of other microorganisms (56). The gene (hoi102_01010) encoding phospholipase D may be the main virulence factor among these strains (Table S2), which is an exotoxin involved in the persistence and spread of the pathogen within the host cells through hydrolysis of sphingomyelin in endothelial membranes, which increases the vascular permeability (29, 57). However, mortality in *Galleria mellonella* larvae on inoculation of penguin strains was similar to that caused by nonpathogenic corynebacteria or PBS (Table S3). A protein BLAST search revealed that all corynebacterial virulence-associated genes reported among the penguin isolates are also present in the *C. ciconiae* genome (data not shown). *C. ciconiae* was isolated from the trachea of wild black storks, *Ciconia nigra* (58), and is not reported to be causing any fatal infections. Therefore, the novel strains from yellow-eyed penguins are likely opportunistic pathogens causing high mortality among young chicks, potentially due to weak immune systems.

Diphtheritic stomatitis in chicks and a potential prevention strategy. Yellow-eyed penguins are an endangered species that are threatened with extinction if current
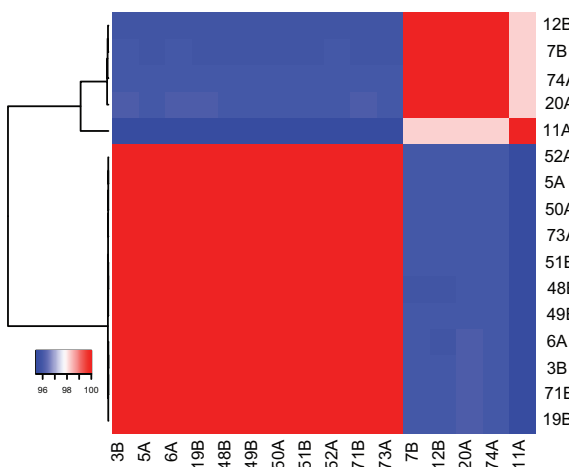
A. Digital DDH
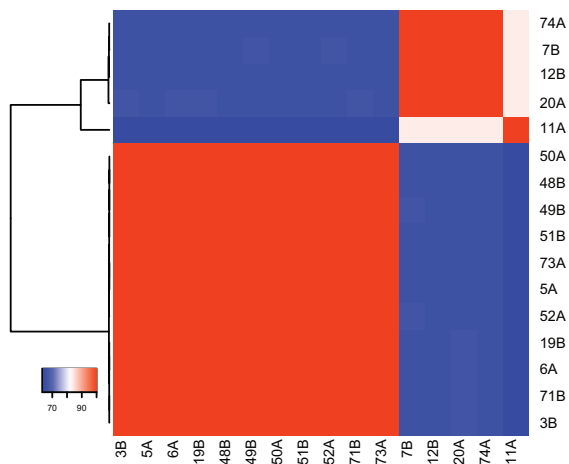
B. Fragmented BLAST Scores
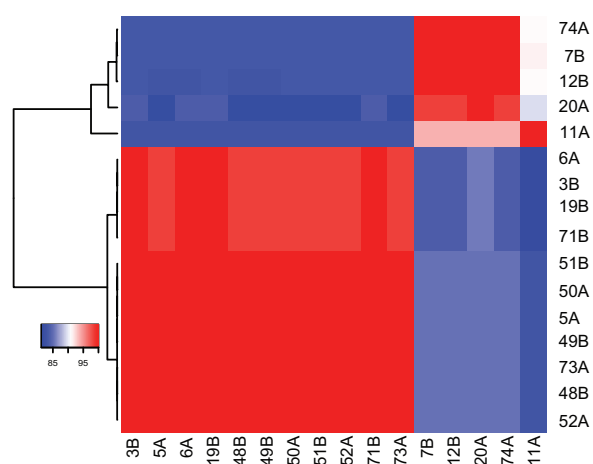
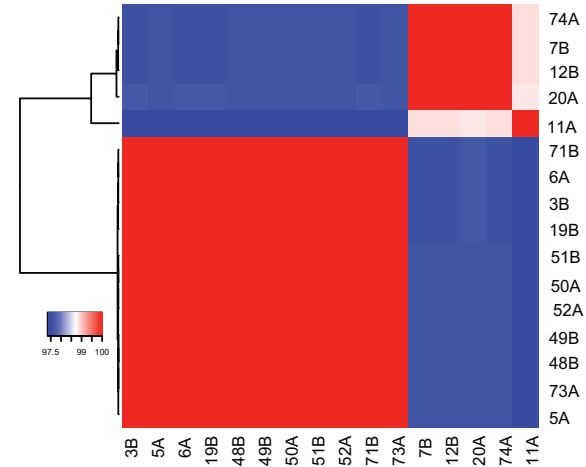C. Average Nucleotide Identities

D. Average Amino-acid Identities

FIG 6 Heat maps from the pairwise genomic matrices for *Corynebacterium* isolates from yellow-eyed penguins. (A) fragmented BLAST scores; (B) digital DNA-DNA hybridization values; (C) genomic average nucleotide identities; (D) average amino-acid identities.

population trends persist (2). Annual outbreaks of diphtheritic stomatitis in chicks results in widespread reproductive failure throughout their mainland distribution. An intensive treatment with a combination of amoxicillin-clavulanic acid or enrofloxacin does not always result in full recovery of the chicks, a problem compounded with the potential physical stress during multiple handling events and crushing injuries from adult penguins. An effective vaccine to protect yellow-eyed penguins from diphtheritic stomatitis that will also limit the direct handling of young chicks might help mitigate this population decline.

A phospholipase D vaccine is an option for disease prevention. Vaccines targeting phospholipase have been tested against *Acinetobacter baumannii*, *Mycobacterium abscessus*, and *C. pseudotuberculosis* (59–61). Survival rates were improved in mice, with reduced pulmonary bacterial load and cytokine levels in the broncho-alveolar lavage fluid and the serum when they were immunized with histidine fusion proteins, small protein A (His-SmpA), and phospolipase D (His-PLD) following the challenge with *A. baumannii* (61). Similarly, immunization with plasmid DNA encoding the *M. abscessus* phospholipase C (MA-PLC) or with purified MA-PLC protein rapidly cleared the infection in the mouse model (60). Formalin-inactivated phospolipase D-rich culture

supernatants offered 95% protection among sheep challenged with *C. pseudotuberculosis* (59). Therefore, inactivated phospholipase D enzyme from these novel strains could potentially be used to immunize yellow-eyed penguins. Due to the early age at which yellow-eyed penguin chicks are affected by diphtheritic stomatitis, a passive immunization strategy would be pertinent, that is the immunization of female adults for maternal transfer of protective IgY into chicks via their egg yolk. This strategy has shown promise, with maternal IgY providing protective effects to young broiler chickens against *Clostridium perfringens*-induced necrotic enteritis (62).

In summary, *Corynebacterium* isolates associated with significant mortality among young chicks of yellow-eyed penguins belong to a novel species. All these strains are equipped with genes encoding the subunits and assembly of SpaDEF type surface pili that are important for adhesion and invasion of the host cells and produce phospholipase D, which is a well-characterized virulence factor. These strains are nonpathogenic to wax moth caterpillars and appear to be opportunistic pathogens for young penguin chicks, with high mortality rates.

MATERIALS AND METHODS

Bacterial isolates and microbiological characterization. In November 2014, swabs were collected from the oral cavities of 2- to 14-day-old yellow-eyed penguin chicks from nests from four breeding areas on the Otago Peninsula (Table 1; Fig. 1; Fig. S1). Swabs were transported using Transystem Amies transport medium with charcoal (Copan, Brescia, Italy) and were stored at 6 to 15°C for up to 24 h after field collection. They were cultured on Columbia horse blood agar plates (Fort Richard, Auckland, New Zealand) at 37°C for 16 to 20 h. Isolates were characterized using Gram staining and MALDI-TOF mass spectrometry as previously described (63). Phenotypic characteristics of two representative isolates, 3B and 7B, were determined using API ZYM and API 20 NE kits following the manufacturer's instructions (bioMérieux, France).

Genome sequencing and analyses. First, 5 ml brain heart infusion broth (Oxoid, UK) was inoculated with a single colony and was incubated overnight at 37°C for 16 h in a shaking incubator. Genomic DNA was extracted from 2 ml of this culture using the UltraClean microbial DNA isolation kit (MoBio, USA). The genomes were sequenced on a MiSeq instrument (Illumina, Inc., USA) and the paired-end reads were assembled using SPAdes 3.9.0 (64). Genome sequences were annotated using Prokka 1.11 (65) and were compared using Roary with an identity threshold of 70% (66, 67). A maximum-likelihood (ML) tree was constructed from the core genomic sequence alignment using IQ-TREE (68) with the best-fit substitution model GTR+I, 100,000 ultrafast bootstraps, and 100,000 approximate likelihood ratio test and Shimodaira-Hasegawa (SH-aLRT) tests. IslandViewer 4 was used to predict genomic islands (GIs) in the genomes using *Corynebacterium epidermidicanis* strain DSM 45586 as the reference (69). Any prediction inconsistencies, i.e., regions or parts of any region predicted as a GI in some strains but not in others, were not considered unless they were predicted in ≥50% strains. In the latter case, they were recorded as GIs. Potential prophage sequences were identified using PHASTER (70, 71). Any inconsistencies in the predicted prophage-associated regions were also resolved using the approach used for GIs.

The protein sequences of known virulence genes from pathogenic *Corynebacterium diphtheriae*, *Corynebacterium pseudotuberculosis*, and *Corynebacterium ulcerans* (13, 14, 29, 42) were searched within the data set using BLAST (72). BLAST hits with ≥40% query coverage and sequence identities with an E value of <10⁻⁵ were considered positive.

Characterization of virulence properties in *Galleria mellonella*. Thirty *Galleria mellonella* larvae (Biosuppliers, Auckland, New Zealand) were each inoculated with two randomly selected penguin isolates (3B and 6A), *C. ulcerans* NZRM 818 (positive control), PBS (negative control), and nonpathogenic *Corynebacterium* sp. strain NZRM 2522 (negative control) as previously described (73). A smaller number, 10 larvae each, were inoculated for 11 additional strains from penguins and *C. pseudotuberculosis* NZRM 3004 to test the robustness of the results. Larvae were incubated at 37°C and assessed daily for up to 5 days. Hemolymph of dead larvae was cultured on Columbia horse blood agar.

Taxogenomic analyses. 16S rRNA gene sequences were extracted from assembled genomes using the RNAmmer 1.2 server (74). Nucleotide sequences of 16S rRNA genes of all known *Corynebacterium* strains were obtained from GenBank (accessed on 10 May 2018). The sequences were aligned using MUSCLE (75), and the sites with the missing data were removed using MEGA X (76). A maximum-likelihood (ML) tree was calculated from the resulting alignment using a TPM3u+I+G4 substitution model and 100,000 ultrafast bootstraps and SH-aLRT tests using IQ-TREE (68).

Whole-genome sequences of *Corynebacterium* strains representing 99 species were obtained from GenBank (Table S1; accessed on 10 May 2018). All corynebacterial genome sequences were annotated using Prokka 1.11 (65) and were compared using BGPA 1.3 with the default settings (77). An ML tree was constructed from the concatenated sequence alignment of 184 core proteins after stripping the sites with missing data following the LG+F+I+G4 substitution model and 100,000 ultrafast bootstraps and SH-aLRT tests using IQ-TREE (68). The phylogenetic trees were visualized using iTOL (78) and were rerooted on the longest branch separating *Corynebacterium ilicis* ICMP 2608 for the 16S rRNA gene phylogeny and *Corynebacterium variabile* NRRL-B 4201 for the core genome phylogeny.

Pairwise FBS values were calculated using Gegenees (79) with a fragment size of 500 bp, and pairwise ANIb was calculated using JSpecies (80). dDDH values were calculated between one representative strain,

3B, and other *Corynebacterium* genomes using GGDC 2.1 (51, 52). Pairwise dDDH values were also calculated among the genomes of penguin isolates using the Genome-to-Genome Distance Calculator (GGDC) 2.1, and pairwise average amino acid identity (AAI) was calculated using the genome-based distance matrix calculator (81). Heat maps from the pairwise genomic matrices were created using Heatmapper (82).

MALDI-TOF spectral analysis. MALDI-TOF protein spectra were obtained in triplicate for all isolates using a previously described protein purification method (83) on a MALDI Biotyper Microflex LT/SH device using the MALDI Biotyper 8468 library (Bruker, USA). Principal-component analyses and the identification of discriminant peaks between species were performed using the MALDIquant (84) and sda (85) R packages. Control strains of *C. trachiae* and *C. ciconiae*, species that showed the closest genetic relationship to penguin isolates, were used for reference spectra.

Primer designing and PCR amplification. The pan-genomic nucleotide sequences were BLAST-searched (72) in the entire *Corynebacterium* genomic data set to identify genes that are unique to the strains isolated from yellow-eyed penguins. Primers were designed using the Web-based version of Primer3 (86) for two genes specific to each lineage (Table S3). The genes were amplified in a 25 μl multiplex PCR containing 200 μM deoxynucleoside triphosphate (dNTP), 5 μl of 5× Q5 reaction buffer, 1 unit of Q5 high-fidelity *Taq* polymerase (New England Biolabs, USA), 10 pmol of each primer, and ~50 ng of the template DNA. Thermal cycling conditions include an initial hold at 94°C for 5 min, 30 cycles of 94°C for 30 s, 55°C for 45 s, and 72°C 1 min, followed by final extension at 72°C for 10 min. The amplicons were separated by electrophoresis on 1.5% agarose gel.

Data availability. The whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession numbers PQMG00000000 to PQMV00000000 (BioProject PRJNA393261).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

DATA SET S1, XLSX file, 0.3 MB.
FIG S1, PDF file, 1.9 MB.
FIG S2, PDF file, 2.3 MB.
FIG S3, PDF file, 0.7 MB.
FIG S4, PDF file, 0.8 MB.
TABLE S1, PDF file, 0.1 MB.
TABLE S2, PDF file, 1.1 MB.
TABLE S3, PDF file, 0.01 MB.
TABLE S4, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Seddon PJ, van Heezik Y, Ellenberg U. 2013. Yellow-eyed penguin (*Megadyptes antipodes*), p 360. *In* Borboroglu PG, Boersma PD (ed), Penguins: natural history and conservation. University of Washington Press, Seattle, WA.
2. Mattern T, Meyer S, Ellenberg U, Houston DM, Darby JT, Young M, van Heezik Y, Seddon PJ. 2017. Quantifying climate change impacts emphasises the importance of managing regional threats in the endangered yellow-eyed penguin. PeerJ 5:e3272. https://doi.org/10.7717/peerj.3272.
3. Darby JT, Dawson SM. 2000. Bycatch of yellow-eyed penguins (*Megadyptes antipodes*) in gillnets in New Zealand waters 1979–1997. Biological Conservation 93:327–332. https://doi.org/10.1016/S0006-3207(99)00148-2.
4. King SD, Harper GA, Wright JB, McInnes JC, van der Lubbe JE, Dobbins ML, Murray SJ. 2012. Site-specific reproductive failure and decline of a population of the endangered yellow-eyed penguin: a case for foraging habitat quality. Mar Ecol Prog Ser 467:233–244. https://doi.org/10.3354/meps09969.
5. Gartrell B, Agnew D, Alley M, Carpenter T, Ha HJ, Howe L, Hunter S, McInnes K, Munday R, Roe W, Young M. 2017. Investigation of a mortality cluster in wild adult yellow-eyed penguins (*Megadyptes antipodes*) at

Otago Peninsula, New Zealand. Avian Pathol 46:278–288. https://doi.org/10.1080/03079457.2016.1264568.
6. Alley MR, Gartrell BD, Morgan KJ. 2005. Wildlife cases from Massey University November 2004–April 2005. Kokako 12:16.
7. Alley MR, Hunter SA, Howe LJ, Sijbranda DC, Gartrell BD. 2013. Severe mortalities in yellow-eyed penguins, *Megadyptes antipodes*, during the 2012–2013 breeding season. Kokako 20:20.
8. Alley MR, Suepaul RB, McKinlay B, Young MJ, Wang J, Morgan KJ, Hunter SA, Gartrell BD. 2017. Diphtheritic stomatitis in yellow-eyed penguins (*Megadyptes Antipodes*) in New Zealand. J Wildl Dis 53:102–110. https://doi.org/10.7589/2015-07-195.
9. Bernard KA, Funke G. 2015. *Corynebacterium*, p 1–70. *In* Whitman WB, Rainey F, Kämpfer P, Trujillo M, Chun J, DeVos P, Hedlund B, Dedysh S (ed), Bergey's manual of systematics of Archaea and Bacteria. John Wiley & Sons, Inc., New York, NY.
10. Sangal V, Burkovski A. 2020. Insights into old and new foes: pan-genomics of *Corynebacterium diphtheriae* and *Corynebacterium ulcerans*, p 81–100.

In Barh D, Soares S, Tiwari S, Azevedo V (ed), Pan-genomics: applications, challenges, and future prospects. Academic Press, San Diego, CA. https://doi.org/10.1016/B978-0-12-817076-2.00004-4.

11. Singh G, Verma AK, Kumar V. 2016. Catalytic properties, functional attributes and industrial applications of beta-glucosidases. 3 Biotech 6:3. https://doi.org/10.1007/s13205-015-0328-z.

12. Vasiliou V, Pappa A, Petersen DR. 2000. Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. Chem Biol Interact 129:1–19. https://doi.org/10.1016/S0009-2797(00)00211-8.

13. Trost E, Blom J, Soares Sde C, Huang IH, Al-Dilaimi A, Schroder J, Jaenicke S, Dorella FA, Rocha FS, Miyoshi A, Azevedo V, Schneider MP, Silva A, Camello TC, Sabbadini PS, Santos CS, Santos LS, Hirata R Jr, Mattos-Guaraldi AL, Efstratiou A, Schmitt MP, Ton-That H, Tauch A. 2012. Pangenomic study of *Corynebacterium diphtheriae* that provides insights into the genomic diversity of pathogenic isolates from cases of classical diphtheria, endocarditis, and pneumonia. J Bacteriol 194:3199–3215. https://doi.org/10.1128/JB.00183-12.

14. Grosse-Kock S, Kolodkina V, Schwalbe EC, Blom J, Burkovski A, Hoskisson PA, Brisse S, Smith D, Sutcliffe IC, Titov L, Sangal V. 2017. Genomic analysis of endemic clones of toxigenic and non-toxigenic *Corynebacterium diphtheriae* in Belarus during and after the major epidemic in 1990s. BMC Genomics 18:873. https://doi.org/10.1186/s12864-017-4276-3.

15. Subedi R, Kolodkina V, Sutcliffe IC, Simpson-Louredo L, Hirata R Jr, Titov L, Mattos-Guaraldi AL, Burkovski A, Sangal V. 2018. Genomic analyses reveal two distinct lineages of *Corynebacterium ulcerans* strains. New Microbes New Infect 25:7–13. https://doi.org/10.1016/j.nmni.2018.05.005.

16. Sangal V, Hoskisson PA. 2014. Corynephages: infections of the infectors, p 67–82. *In* Burkovski A (ed), *Corynebacterium diphtheriae* and related toxigenic species. Springer, Heidelberg, Germany. https://doi.org/10.1007/978-94-007-7624-1.

17. Wertheimer AM, Verweij W, Chen Q, Crosa LM, Nagasawa M, Tolmasky ME, Actis LA, Crosa JH. 1999. Characterization of the angR gene of *Vibrio anguillarum*: essential role in virulence. Infect Immun 67:6496–6509. https://doi.org/10.1128/IAI.67.12.6496-6509.1999.

18. Naka H, Liu M, Actis LA, Crosa JH. 2013. Plasmid- and chromosome-encoded siderophore anguibactin systems found in marine vibrios: biosynthesis, transport and evolution. Biometals 26:537–547. https://doi.org/10.1007/s10534-013-9629-z.

19. Ridley CP, Lee HY, Khosla C. 2008. Evolution of polyketide synthases in bacteria. Proc Natl Acad Sci U S A 105:4595–4600. https://doi.org/10.1073/pnas.0710107105.

20. Dedrick RM, Mavrich TN, Ng WL, Cervantes Reyes JC, Olm MR, Rush RE, Jacobs-Sera D, Russell DA, Hatfull GF. 2016. Function, expression, specificity, diversity and incompatibility of actinobacteriophage parABS systems. Mol Microbiol 101:625–644. https://doi.org/10.1111/mmi.13414.

21. Ravin NV, Rech J, Lane D. 2008. Extended function of plasmid partition genes: the Sop system of linear phage-plasmid N15 facilitates late gene expression. J Bacteriol 190:3538–3545. https://doi.org/10.1128/JB.01993-07.

22. Bignell C, Thomas CM. 2001. The bacterial ParA-ParB partitioning proteins. J Biotechnol 91:1–34. https://doi.org/10.1016/s0168-1656(01)00293-0.

23. Zeller T, Klug G. 2006. Thioredoxins in bacteria: functions in oxidative stress response and regulation of thioredoxin genes. Naturwissenschaften 93:259–266. https://doi.org/10.1007/s00114-006-0106-1.

24. Russel M, Model P. 1985. Thioredoxin is required for filamentous phage assembly. Proc Natl Acad Sci U S A 82:29–33. https://doi.org/10.1073/pnas.82.1.29.

25. Berry AM, Lock RA, Hansman D, Paton JC. 1989. Contribution of autolysin to virulence of *Streptococcus pneumoniae*. Infect Immun 57:2324–2330. https://doi.org/10.1128/IAI.57.8.2324-2330.1989.

26. Ott L, Holler M, Rheinlaender J, Schaffer TE, Hensel M, Burkovski A. 2010. Strain-specific differences in pili formation and the interaction of *Corynebacterium diphtheriae* with host cells. BMC Microbiol 10:257. https://doi.org/10.1186/1471-2180-10-257.

27. Mandlik A, Swierczynski A, Das A, Ton-That H. 2008. Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. Trends Microbiol 16:33–40. https://doi.org/10.1016/j.tim.2007.10.010.

28. Sangal V, Blom J, Sutcliffe IC, von Hunolstein C, Burkovski A, Hoskisson PA. 2015. Adherence and invasive properties of *Corynebacterium diphtheriae* strains correlates with the predicted membrane-associated and secreted proteome. BMC Genomics 16:765. https://doi.org/10.1186/s12864-015-1980-8.

29. Trost E, Al-Dilaimi A, Papavasiliou P, Schneider J, Viehoever P, Burkovski A, Soares SC, Almeida SS, Dorella FA, Miyoshi A, Azevedo V, Schneider MP, Silva A, Santos CS, Santos LS, Sabbadini P, Dias AA, Hirata R Jr, Mattos-Guaraldi AL, Tauch A. 2011. Comparative analysis of two complete *Corynebacterium ulcerans* genomes and detection of candidate virulence factors. BMC Genomics 12:383. https://doi.org/10.1186/1471-2164-12-383.

30. Flores-Diaz M, Monturiol-Gross L, Naylor C, Alape-Giron A, Flieger A. 2016. Bacterial sphingomyelinases and phospholipases as virulence factors. Microbiol Mol Biol Rev 80:597–628. https://doi.org/10.1128/MMBR.00082-15.

31. Oliveira A, Oliveira LC, Aburjaile F, Benevides L, Tiwari S, Jamal SB, Silva A, Figueiredo HCP, Ghosh P, Portela RW, De Carvalho Azevedo VA, Wattam AR. 2017. Insight of genus *Corynebacterium*: ascertaining the role of pathogenic and non-pathogenic species. Front Microbiol 8:1937. https://doi.org/10.3389/fmicb.2017.01937.

32. Viana MVC, Figueiredo H, Ramos R, Guimaraes LC, Pereira FL, Dorella FA, Selim SAK, Salaheldean M, Silva A, Wattam AR, Azevedo V. 2017. Comparative genomic analysis between *Corynebacterium pseudotuberculosis* strains isolated from buffalo. PLoS One 12:e0176347. https://doi.org/10.1371/journal.pone.0176347.

33. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2014. CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D236. https://doi.org/10.1093/nar/gku1221.

34. Zambonelli C, Roberts MF. 2003. An iron-dependent bacterial phospholipase D reminiscent of purple acid phosphatases. J Biol Chem 278:13706–13711. https://doi.org/10.1074/jbc.M210363200.

35. Rodriguez F, Lillington J, Johnson S, Timmel CR, Lea SM, Berks BC. 2014. Crystal structure of the *Bacillus subtilis* phosphodiesterase PhoD reveals an iron and calcium-containing active site. J Biol Chem 289:30889–30899. https://doi.org/10.1074/jbc.M114.604892.

36. Selvy PE, Lavieri RR, Lindsley CW, Brown HA. 2011. Phospholipase D: enzymology, functionality, and chemical modulation. Chem Rev 111:6064–6119. https://doi.org/10.1021/cr200296t.

37. Frohman MA. 2015. The phospholipase D superfamily as therapeutic targets. Trends Pharmacol Sci 36:137–144. https://doi.org/10.1016/j.tips.2015.01.001.

38. Martinelli DJ, Pavelka MS, Jr. 2016. The RipA and RipB peptidoglycan endopeptidases are individually nonessential to *Mycobacterium smegmatis*. J Bacteriol 198:1464–1475. https://doi.org/10.1128/JB.00059-16.

39. Hett EC, Chao MC, Deng LL, Rubin EJ. 2008. A mycobacterial enzyme essential for cell division synergizes with resuscitation-promoting factor. PLoS Pathog 4:e1000001. https://doi.org/10.1371/journal.ppat.1000001.

40. Ott L, Holler M, Gerlach RG, Hensel M, Rheinlaender J, Schaffer TE, Burkovski A. 2010. *Corynebacterium diphtheriae* invasion-associated protein (DIP1281) is involved in cell surface organization, adhesion and internalization in epithelial cells. BMC Microbiol 10:2. https://doi.org/10.1186/1471-2180-10-2.

41. Kolodkina V, Denisevich T, Titov L. 2011. Identification of *Corynebacterium diphtheriae* gene involved in adherence to epithelial cells. Infect Genet Evol 11:518–521. https://doi.org/10.1016/j.meegid.2010.11.004.

42. Trost E, Ott L, Schneider J, Schroder J, Jaenicke S, Goesmann A, Husemann P, Stoye J, Dorella FA, Rocha FS, Soares Sde C, D'Afonseca V, Miyoshi A, Ruiz J, Silva A, Azevedo V, Burkovski A, Guiso N, Join-Lambert OF, Kayal S, Tauch A. 2010. The complete genome sequence of *Corynebacterium pseudotuberculosis* FRC41 isolated from a 12-year-old girl with necrotizing lymphadenitis reveals insights into gene-regulatory networks contributing to virulence. BMC Genomics 11:728. https://doi.org/10.1186/1471-2164-11-728.

43. McKean S, Davies J, Moore R. 2005. Identification of macrophage induced genes of *Corynebacterium pseudotuberculosis* by differential fluorescence induction. Microbes Infect 7:1352–1363. https://doi.org/10.1016/j.micinf.2005.05.002.

44. Gande R, Dover LG, Krumbach K, Besra GS, Sahm H, Oikawa T, Eggeling L. 2007. The two carboxylases of *Corynebacterium glutamicum* essential for fatty acid and mycolic acid synthesis. J Bacteriol 189:5257–5264. https://doi.org/10.1128/JB.00254-07.

45. Marrakchi H, Laneelle MA, Daffe M. 2014. Mycolic acids: structures, biosynthesis, and beyond. Chem Biol 21:67–85. https://doi.org/10.1016/j.chembiol.2013.11.011.

46. Han Z, Kautto L, Meyer W, Chen SC, Nevalainen H. 2018. Growth and protease secretion of *Scedosporium aurantiacum* under conditions of hypoxia. Microbiol Res 216:23–29. https://doi.org/10.1016/j.micres.2018.08.003.

47. Konstantinidis KT, Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A 102:2567–2572. https://doi.org/10.1073/pnas.0409727102.

48. Sangal V, Goodfellow M, Jones AL, Schwalbe EC, Blom J, Hoskisson PA, Sutcliffe IC. 2016. Next-generation systematics: an innovative approach to

resolve the structure of complex prokaryotic taxa. Sci Rep 6:38392. https://doi.org/10.1038/srep38392.

49. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol 68:461–466. https://doi.org/10.1099/ijsem.0.002516.

50. Sangal V, Goodfellow M, Blom J, Tan GYA, Klenk HP, Sutcliffe IC. 2018. Revisiting the taxonomic status of the biomedically and industrially important genus *Amycolatopsis*, using a phylogenomic approach. Front Microbiol 9:2281. https://doi.org/10.3389/fmicb.2018.02281.

51. Auch AF, Klenk HP, Göker M. 2010. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Stand Genomic Sci 2:142–148. https://doi.org/10.4056/sigs.541628.

52. Auch AF, von Jan M, Klenk HP, Göker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134. https://doi.org/10.4056/sigs.531120.

53. Li X, Huang Y, Whitman WB. 2015. The relationship of the whole genome sequence identity to DNA hybridization varies between genera of prokaryotes. Antonie Van Leeuwenhoek 107:241–249. https://doi.org/10.1007/s10482-014-0322-1.

54. Reardon-Robinson ME, Ton-That H. 2014. Assembly and function of *Corynebacterium diphtheriae* pili, p 123–141. *In* Burkovski A (ed), *Corynebacterium diphtheriae* and related toxigenic species. Springer, Heidelberg, Germany. https://doi.org/10.1007/978-94-007-7624-1.

55. Mandlik A, Swierczynski A, Das A, Ton-That H. 2007. *Corynebacterium diphtheriae* employs specific minor pilins to target human pharyngeal epithelial cells. Mol Microbiol 64:111–124. https://doi.org/10.1111/j.1365-2958.2007.05630.x.

56. Tamehiro N, Okamoto-Hosoya Y, Okamoto S, Ubukata M, Hamada M, Naganawa H, Ochi K. 2002. Bacilysocin, a novel phospholipid antibiotic produced by *Bacillus subtilis* 168. Antimicrob Agents Chemother 46:315–320. https://doi.org/10.1128/aac.46.2.315-320.2002.

57. McKean SC, Davies JK, Moore RJ. 2007. Expression of phospholipase D, the major virulence factor of *Corynebacterium pseudotuberculosis*, is regulated by multiple environmental factors and plays a role in macrophage death. Microbiology 153:2203–2211. https://doi.org/10.1099/mic.0.2007/005926-0.

58. Fernandez-Garayzabal JF, Vela AI, Egido R, Hutson RA, Lanzarot MP, Fernandez-Garcia M, Collins MD. 2004. *Corynebacterium ciconiae* sp. nov., isolated from the trachea of black storks (*Ciconia nigra*). Int J Syst Evol Microbiol 54:2191–2195. https://doi.org/10.1099/ijs.0.63165-0.

59. Hodgson AL, Carter K, Tachedjian M, Krywult J, Corner LA, McColl M, Cameron A. 1999. Efficacy of an ovine caseous lymphadenitis vaccine formulated using a genetically inactive form of the *Corynebacterium pseudotuberculosis* phospholipase D. Vaccine 17:802–808. https://doi.org/10.1016/S0264-410X(98)00264-3.

60. Le Moigne V, Rottman M, Goulard C, Barteau B, Poncin I, Soismier N, Canaan S, Pitard B, Gaillard JL, Herrmann JL. 2015. Bacterial phospholipases C as vaccine candidate antigens against cystic fibrosis respiratory pathogens: the *Mycobacterium abscessus* model. Vaccine 33:2118–2124. https://doi.org/10.1016/j.vaccine.2015.03.030.

61. Li H, Tan H, Hu Y, Pan P, Su X, Hu C. 2017. Small protein A and phospholipase D immunization serves a protective role in a mouse pneumonia model of *Acinetobacter baumannii* infection. Mol Med Rep 16:1071–1078. https://doi.org/10.3892/mmr.2017.6688.

62. Keyburn AL, Portela RW, Ford ME, Bannam TL, Yan XX, Rood JI, Moore RJ. 2013. Maternal immunization with vaccines containing recombinant NetB toxin partially protects progeny chickens from necrotic enteritis. Vet Res 44:108. https://doi.org/10.1186/1297-9716-44-108.

63. Nisa S, Bercker C, Midwinter AC, Bruce I, Graham CF, Venter P, Bell A, French NP, Benschop J, Bailey KM, Wilkinson DA. 2019. Combining MALDI-TOF and genomics in the study of methicillin resistant and multidrug resistant *Staphylococcus pseudintermedius* in New Zealand. Sci Rep 9:1271. https://doi.org/10.1038/s41598-018-37503-9.

64. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

65. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

66. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421.

67. Tange O. 2011. GNU Parallel: the command-line power tool. ;login 36:42–47.

68. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300.

69. Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing G, Lau BY, Hoad G, Winsor GL, Brinkman FS. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res 45:W30–W35. https://doi.org/10.1093/nar/gkx343.

70. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. Nucleic Acids Res 39:W347–W352. https://doi.org/10.1093/nar/gkr485.

71. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387.

72. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.

73. Bojanic K, Acke E, Roe WD, Marshall JC, Cornelius AJ, Biggs PJ, Midwinter AC. 2020. Comparison of the pathogenic potential of *Campylobacter jejuni*, *C. upsaliensis* and *C. helveticus* and limitations of using larvae of *Galleria mellonella* as an infection model. Pathogens 9:713. https://doi.org/10.3390/pathogens9090713.

74. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.

75. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.

76. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096.

77. Chaudhari NM, Gupta VK, Dutta C. 2016. BPGA: an ultra-fast pan-genome analysis pipeline. Sci Rep 6:24373. https://doi.org/10.1038/srep24373.

78. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242–W245. https://doi.org/10.1093/nar/gkw290.

79. Agren J, Sundstrom A, Hafstrom T, Segerman B. 2012. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. PLoS One 7:e39107. https://doi.org/10.1371/journal.pone.0039107.

80. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126–19131. https://doi.org/10.1073/pnas.0906412106.

81. Rodriguez-R LM, Konstantinidis KT. 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. PeerJ Preprints 4:e1900v1. https://doi.org/10.7287/peerj.preprints.1900v1.

82. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS. 2016. Heatmapper: Web-enabled heat mapping for all. Nucleic Acids Res 44:W147–W153. https://doi.org/10.1093/nar/gkw419.

83. Koziel M, O'Doherty P, Vandamme P, Corcoran GD, Sleator RD, Lucey B. 2014. *Campylobacter corcagiensis sp. nov.*, isolated from faeces of captive lion-tailed macaques (*Macaca silenus*). Int J Syst Evol Microbiol 64:2878–2883. https://doi.org/10.1099/ijs.0.063867-0.

84. Gibb S, Strimmer K. 2012. MALDIquant: a versatile R package for the analysis of mass spectrometry data. Bioinformatics 28:2270–2271. https://doi.org/10.1093/bioinformatics/bts447.

85. Ahdesmaki M, Zuber V, Gibb S, Strimmer K. 2015. sda: shrinkage discriminant analysis and CAT score variable selection. https://CRAN.R-project.org/package=sda.

86. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3: new capabilities and interfaces. Nucleic Acids Res 40:e115. https://doi.org/10.1093/nar/gks596.