Draft Genome Sequences of *Erysipelothrix* sp. Strains Isolated from Stranded Septic Bottlenose Dolphins in Alabama, USA

Kahylin Nesbitt,[a,b,c] Jennifer Bloodgood,[b,c,d] Megan M. Mullis,[b,d] Alissa C. Deming,[b,c*] Kathleen Colegrove,[e] Brandi Kiel Reese[b,d]

ᵃBiology Department, Oakwood University, Huntsville, Alabama, USA
ᵇDauphin Island Sea Lab, Dauphin Island, Alabama, USA
ᶜAlabama Marine Mammal Stranding Network, Dauphin Island, Alabama, USA
ᵈSchool of Marine and Environmental Sciences, University of South Alabama, Mobile, Alabama, USA
ᵉZoological Pathology Program, College of Veterinary Medicine, University of Illinois, Brookfield, Illinois, USA

ABSTRACT   We report the genome sequences of two *Erysipelothrix* isolates from fatal cases of sepsis in bottlenose dolphins (*Tursiops truncatus*). The genomes were found to be most closely related to *Erysipelothrix rhusiopathiae* and *Erysipelothrix piscisicarius*. This information expands our knowledge of the genetic characteristics of this pathogen, which can affect free-ranging marine mammals.

In February and March 2020, two bottlenose dolphin (*Tursiops truncatus*) carcasses (identified as 10DISL and 19DISL, respectively) were stranded off Orange Beach, Alabama. Necropsies were conducted by the Alabama Marine Mammal Stranding Network at Dauphin Island Sea Lab according to standard protocols, and paired tissue samples were collected for histological evaluation by the University of Illinois Zoological Pathology Program and for −80℃ archiving. Findings indicated that both animals died of bacterial sepsis. Gram-stained tissue sections showed intracellular, rod-shaped, Gram-positive bacteria consistent with *Erysipelothrix* spp. *Erysipelothrix* may cause disease in multiple species, including humans. Sources of infection in dolphins may include environmental contamination of wounds and ingestion of fish harboring the bacteria in their mucus layer (1).

Samples of cerebrum of carcass 10DISL and spleen of carcass 19DISL, chosen because of intracellular bacteria identified during histopathology, were shipped to the National Veterinary Services Laboratories, where they were inoculated onto blood agar (R01202; Remel Products), chocolate agar (R01302; Remel Products), and MacConkey agar made in-house. Plates were incubated at 37℃ in 5% CO_2 for 18 to 24 h. Single colonies of different colony types were streaked on blood agar and incubated under the same conditions. The original plates were also reincubated. Identification was performed using a Bruker Biotyper on an Autoflex Speed matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometer. DNA was extracted using the Promega Maxwell RSC whole-blood DNA kit. Whole-genome sequencing was performed according to established protocols using 2 × 250-bp paired-end chemistry and the Nextera XT library preparation kit on an Illumina MiSeq system, which generated 1.38 million reads for 10DISL and 1.36 million reads for 19DISL. Sequences were merged with FLASH v.2.2.0 (2), trimmed with Trim Galore v.0.6.5 (3), assembled with SPAdes v.3.15.3 (4), and annotated with Prokka v.1.11.1 within PATRIC v.3.6.12 comprehensive genome analysis (5–8). A phylogenetic tree was created in ezTree v.0.1 (9) using Prodigal (10). The amino acid sequences were compared to Pfam hidden Markov model profiles (11) using HMMER3 (12). Pfam profiles identified once were concatenated and aligned using MUSCLE (13). The alignment was used to construct a maximum likelihood tree (FastTree, JTT model, with 1,000 bootstraps) (14). Default parameters were used for all software. Average nucleotide identity (ANI) values were calculated using the JSpeciesWS webserver (15).

TABLE 1 Assembly metrics for the two isolates

Strain identification	Isolate name	BioSample accession no.	No. of contigs	G+C content (%)	N_{50} (bp)	L_{50}	Genome length (bp)	No. of protein-coding sequences	No. of hypothetical proteins	Sequence coverage (×)
Erysipelothrix sp.	10DISL	SAMN23594042	96	36.46	72,029	7	1,718,626	1,225	450	160
Erysipelothrix sp.	19DISL	SAMN23594043	112	36.31	42,844	13	1,807,374	1,229	546	146.9

Isolate 10DISL contained 96 contigs, with a G+C content of 34.46%; the N_{50} value was 72,029 bp, the L_{50} value was 7, and the genome length was 1,718,626 bp, with 1,225 protein-coding sequences, 450 hypothetical proteins, and 160× sequence coverage (Table 1). Isolate 10DISL included genes for metabolism processes (160 genes), protein processing (184 genes), energy processes (93 genes), DNA processing (66 genes), RNA processing (40 genes), stress response, defense, and virulence (53 genes), and cellular processes (51 genes). Isolate 19DISL contained 112 contigs, with a G+C content of 36.31%; the N_{50} value was 42,844 bp, the L_{50} value was 13, and the genome length was 1,807,374 bp, with 1,229 protein-coding sequences, 546 hypothetical proteins, and 146.9× sequence coverage. Isolate 19DISL included genes for metabolism processes (160 genes), protein processing (184 genes), energy processes (94 genes), DNA processing (67 genes), RNA processing (40 genes), stress response, defense, and virulence (53 genes), and cellular processes (52 genes). All annotated genes in these categories were verified as homologues using the annotation output from PATRIC, thus indicating that the gene contents across these two isolates are highly conserved. The genomes from isolates 10DISL and 19DISL were related to both *Erysipelothrix rhusiopathiae* and *Erysipelothrix piscisicarius* with 86% ANI and were related to *Erysipelothrix tonsillarum* with 79% ANI (Fig. 1). The isolates from this study were related to each other with 99.6% ANI, indicating that the two genomes represented a novel species.

Data availability. These whole-genome assemblies have been deposited in GenBank under the accession numbers listed in Table 1. The assembled sequences for each isolate were deposited in the Sequence Read Archive (SRA) under BioSample accession numbers SAMN23594042 (10DISL) and SAMN23594043 (19DISL).

Tree scale: 0.1 ⊢————·———┤

Erysipelothrix larvae (GCF 001545095.1)
Erysipelothrix inopinata (GCF 014396165.1)
Erysipelothrix tonsillarum DSM 14972 (GCF 000373785.1)
Erysipelothrix piscisicarius (GCF 003931795.1)
Erysipelothrix rhusiopathiae (GCF 009906265.1)
Erysipelothrix rhusiopathiae (GCF 003722215.1)
Erysipelothrix rhusiopathiae (GCF 003725505.1)
Erysipelothrix rhusiopathiae ATCC 19414(GCF 000160815.2)
Erysipelothrix rhusiopathiae (GCF 900637845.1)
Erysipelothrix rhusiopathiae (GCF 900448055.1)
19DISL
10DISL
Erysipelothrix rhusiopathiae SY1027(GCF 000404205.1)
Erysipelothrix rhusiopathiae (GCF 006384935.1)
Erysipelothrix rhusiopathiae str. Fujisawa(GCF 000270085.1)
Erysipelothrix rhusiopathiae (GCF 003226675.1)
Erysipelothrix rhusiopathiae (GCF 006384955.1)
Erysipelothrix rhusiopathiae (GCF 007725185.1)
Erysipelothrix rhusiopathiae (GCF 001723625.1)
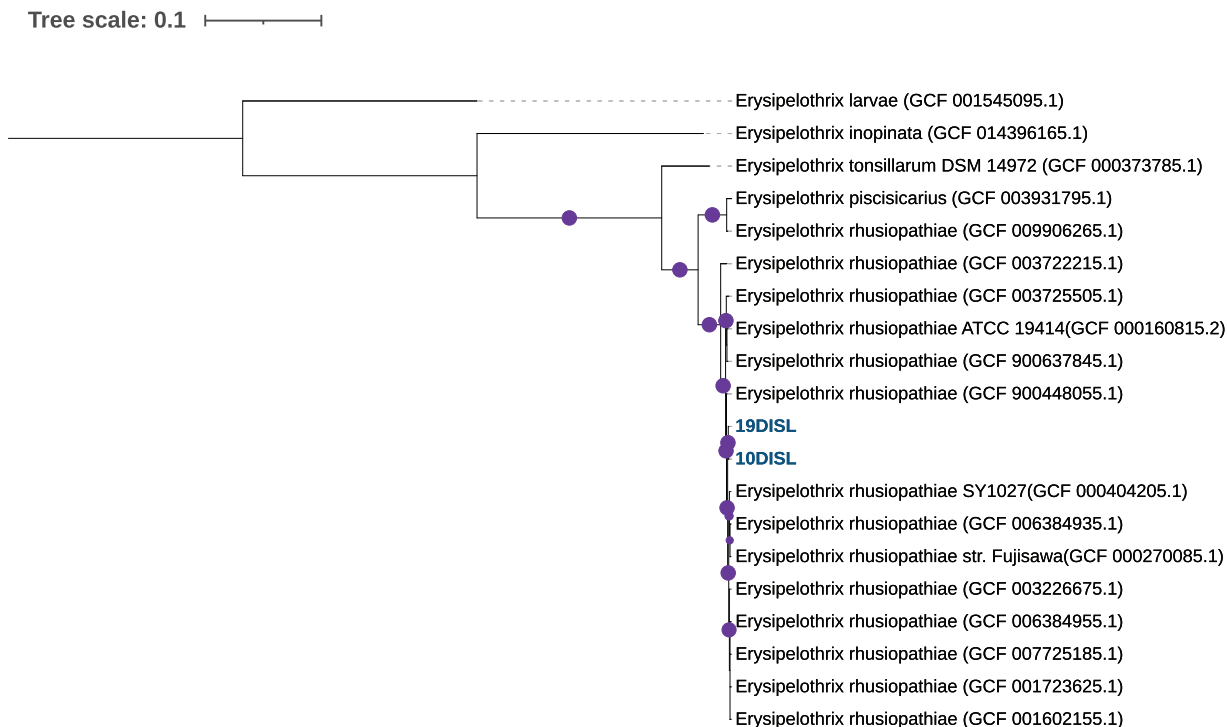Erysipelothrix rhusiopathiae (GCF 001602155.1)

FIG 1 Phylogenetic tree using single-copy marker genes for whole genomes to confirm bacterial species. The purple dots represent bootstrap values greater than 70%. Numbers in parentheses are GenBank accession numbers.

REFERENCES

1. Díaz-Delgado J, Arbelo M, Sierra E, Vela A, Domínguez M, Paz Y, Andrada M, Domínguez L, Fernández A. 2015. Fatal *Erysipelothrix rhusiopathiae* septicemia in two Atlantic dolphins (*Stenella frontalis* and *Tursiops truncatus*). Dis Aquat Organ 116:75–81. https://doi.org/10.3354/dao02900.
2. Magoc T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963. https://doi.org/10.1093/bioinformatics/btr507.
3. Krueger F. 2015. Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore.
4. 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.
5. Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, Butler RM, Chlenski P, Conrad N, Dickerman A, Dietrich EM, Gabbard JL, Gerdes S, Guard A, Kenyon RW, Machi D, Mao C, Murphy-Olson D, Nguyen M, Nordberg EK, Olsen GJ, Olson RD, Overbeek JC, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomas C, VanOeffelen M, Vonstein V, Warren AS, Xia F, Xie D, Yoo H, Stevens R. 2020. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. Nucleic Acids Res 48:D606–D612. https://doi.org/10.1093/nar/gkz943.
6. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
7. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206–D214. https://doi.org/10.1093/nar/gkt1226.
8. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.
9. Wu Y-W. 2018. ezTree: an automated pipeline for identifying phylogenetic marker genes and inferring evolutionary relationships among uncultivated prokaryotic draft genomes. BMC Genomics 19:921. https://doi.org/10.1186/s12864-017-4327-9.
10. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.
11. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285. https://doi.org/10.1093/nar/gkv1344.
12. Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference. Genome Inform 23:205–211. https://doi.org/10.1142/9781848165632_0019.
13. Edgar R. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.
14. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650. https://doi.org/10.1093/molbev/msp077.
15. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931. https://doi.org/10.1093/bioinformatics/btv681.