



HAL
open science

Genomic characterisation of a novel species of *Erysipelothrix* associated with mortalities among endangered seabirds

Jiadong Zhong, Matej Medvecký, Jérémy Tornos, Augustin Clessin, Hubert Gantelet, Amandine Gamble, Taya L Forde, Thierry Boulinier

► **To cite this version:**

Jiadong Zhong, Matej Medvecký, Jérémy Tornos, Augustin Clessin, Hubert Gantelet, et al.. Genomic characterisation of a novel species of *Erysipelothrix* associated with mortalities among endangered seabirds. 2022. hal-03852500

HAL Id: hal-03852500

<https://hal.science/hal-03852500>

Preprint submitted on 15 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Genomic characterisation of a novel species of *Erysipelothrix* associated with mortalities**
2 **among endangered seabirds**

3

4 Jiadong Zhong¹, Matej Medvecký^{1,2}, Jérémy Tornos^{3,4}, Augustin Clessin^{3,5}, Hubert Gantelet⁴,
5 Amandine Gamble^{1,6}, Taya L. Forde^{1¶*}, Thierry Boulinier^{3¶}

6

7

8 ¹ Institute of Biodiversity, Animal Health & Comparative Medicine, University of Glasgow,
9 Glasgow, G12 8QQ United Kingdom

10 ² Department of Chemistry, Faculty of Science, University of Hradec Kralove, 500 03 Hradec
11 Kralove, Czech Republic

12 ³ CEFE, UMR 5175, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

13 ⁴ Ceva Biovac, Beaucauzé, France

14 ⁵ École Normale Supérieure de Lyon, 69342 Lyon Cedex 07, France

15 ⁶ Department of Ecology and Evolutionary Biology, University of California, Los Angeles, USA

16

17

18 ¶ These authors contributed equally to this work.

19

20 *Corresponding author:

21 E-mail: taya.forde@glasgow.ac.uk (TLF)

22

23

24 Key words: *Erysipelothrix* spp.; Yellow-nosed albatross; Comparative genomics; Conservation

25

26

27 **Repositories:** Reads for all samples sequenced in this study are available on European

28 Nucleotide Archive Sequence Read Archive (SRA) under accession number PRJEB50151.

29 Complete annotated genomes for isolates A18Y020d^T (type strain) and A18Y016a are available
30 under accession numbers GCA_940143175 and GCA_940143155, respectively.

31 **Abstract**

32 Infectious diseases threaten endangered species, particularly in small isolated populations.
33 Seabird populations on the remote Amsterdam Island in the Indian Ocean have been in decline
34 for the past three decades, with avian cholera caused by *Pasteurella multocida* proposed as the
35 primary driver. However, *Erysipelothrix* spp. has also been sporadically detected from
36 albatrosses on Amsterdam Island and may be contributing to some of the observed mortality. In
37 this study, we genomically characterised 16 *Erysipelothrix* spp. isolates obtained from three
38 Indian yellow-nosed albatross chick carcasses in 2019. Two isolates were sequenced using both
39 Illumina short-read and MinION long-read approaches, which – following hybrid assembly –
40 resulted in closed circular genomes. Mapping of Illumina reads from the remaining isolates to
41 one of these new reference genomes revealed that all 16 isolates were closely related, with a
42 maximum of 13 nucleotide differences distinguishing any pair of isolates. The nucleotide
43 diversity of isolates obtained from the same or different carcasses was similar, suggesting all
44 three chicks were likely infected from a common source. These genomes were compared with a
45 global collection of genomes from *E. rhusiopathiae* and other species from the same genus. The
46 isolates from albatrosses were phylogenetically distinct, sharing a most recent common ancestor
47 with *E. rhusiopathiae*. Based on phylogenomic analysis and standard thresholds for average
48 nucleotide identity and digital DNA-DNA hybridisation, these isolates represent a novel
49 *Erysipelothrix* species, for which we propose the name *Erysipelothrix amsterdamensis* sp. nov.
50 The type strain is *E. amsterdamensis* A18Y020d^T. The implications of this bacterium for
51 albatross conservation will require further study.

52 Introduction

53 Infectious diseases are among the major threats to the conservation of endangered species,
54 especially at local scales [1]. This is particularly true for populations in remote areas, where
55 anthropogenic activity has historically been limited but is increasing, such as with commercial
56 tourists to the Antarctic [2, 3]. Concerns have been raised about disease-associated declines
57 among albatross populations and how this could influence their long-term sustainability [4].
58 Avian cholera, caused by the bacterium *Pasteurella multocida*, has been causing recurrent chick
59 mortalities among two endangered albatross species on the remote Amsterdam Island in the
60 southern Indian Ocean since at least the mid-1990s when it was first detected [5–7]. Both the
61 Indian yellow-nosed albatross (*Thalassarche carteri*) and the sooty albatross (*Phoebastria fusca*)
62 have been affected, and the endemic endangered Amsterdam albatross (*Diomedea*
63 *amsterdamensis*) is also at risk [5]. The bacterium *Erysipelothrix rhusiopathiae* has also been
64 detected by PCR in swabs from live birds of these three albatross species and from the
65 endangered northern rockhopper penguin (*Eudyptes moseleyi*) on Amsterdam Island [6], and it
66 has been isolated concomitantly from mortalities among the Indian yellow-nosed albatross [5].
67 During ongoing population health monitoring conducted to infer transmission dynamics and
68 management opportunities [8–10], our team isolated *Erysipelothrix* spp. (identified at the time as
69 *E. rhusiopathiae*) in 2019 from the carcasses of three Indian yellow-nosed albatross chicks on
70 the island, strengthening the hypothesis that this bacterium could also be contributing to albatross
71 mortalities and decreased breeding success.

72 *E. rhusiopathiae* is a small, Gram-positive bacillus. While best known as an opportunistic
73 pathogen of pigs and poultry [11, 12], it affects a wide range of species [13] and in recent years it
74 has been implicated in large-scale wildlife mortalities [14]. It was also responsible for
75 jeopardising the translocation success of several critically endangered kakapo – a flightless
76 parrot – between New Zealand and offshore islands [15], and has caused the death of other
77 endangered birds in that area [16]. Erysipelas in wild birds is a septicemic disease [12], and
78 deaths have been reported in a range of wild terrestrial birds and seabirds worldwide [15].
79 Mortalities occur sporadically, and tend to be limited to a small number of individuals, although
80 cases are likely under-recognised and under-reported [17]. Rare mass mortality events in wild
81 birds have also been observed [18].

82 In this study, we sought to genomically characterise the *Erysipelothrix* strains isolated from
83 Indian yellow-nosed albatross on Amsterdam Island to identify potential sources, and to assess
84 the relationship between the multiple observed cases on the island.

85

86 **Methods**

87 Between December 2018 and March 2019, 27 chick and 2 adult carcasses of Indian yellow-
88 nosed albatross were necropsied in the field on Amsterdam Island (-37.797135, 77.571521) as
89 part of long-term population health monitoring (Appendix S1, Figure S1, Table S1). Samples
90 from multiple organs (liver, lung, brain, and heart) were collected from the carcasses for
91 bacteriological and histological analysis. Isolates of *Erysipelothrix* spp. were identified from
92 three of the chick carcasses, all sampled between the 13th and the 16th of January, 2019. The
93 bacterial growth in the field (Appendix S1) showed *Erysipelothrix* spp. exclusively, suggesting
94 that this bacterium could have been the cause of death. Between four and six isolates were
95 obtained from each carcass (n = 16 total). These were initially identified in the laboratory as *E.*
96 *rhusiopathiae* based on colony morphology on horse blood agar and MALDI-TOF, with the
97 exception of strain A18Y019b, which was classified as *Erysipelothrix* spp. DNA was extracted
98 from sub-cultured isolates (loop of colonies) using the GenElute Bacterial Genomic DNA Kit
99 (Sigma-Aldrich) as per the manufacturer's instructions. Extracted DNA was tested with a probe-
100 based qPCR targeting *E. rhusiopathiae* [19], then sequenced at MicrobesNG (Birmingham, UK).
101 Libraries were prepared using the Nextera XT v2 kit (Illumina) and sequenced on the Illumina
102 HiSeq platform, generating 250 base pair paired-end reads. Long-read MinION sequencing
103 (Oxford Nanopore) was performed at the University of Glasgow for two isolates (A18Y016a and
104 A18Y020d), with library preparation done using the rapid barcoding kit (SQK-RBK004).
105 Illumina reads were adapter and quality trimmed using Trimmomatic [20] and assembled *de*
106 *novo* using SPAdes [21] by MicrobesNG. Assembled genomes were evaluated using the online
107 version of Quality Assessment Tool for Genome Assemblies (QUAST)
108 (<http://cab.cc.spbu.ru/quast/>) [22]. Unless otherwise indicated, the remaining bioinformatics
109 analyses were performed within the CLIMB computing platform for microbial genomics [23].
110 To generate circularised (closed) reference genomes, the Illumina and MinION sequence data
111 were jointly assembled by Unicycler v0.4.4 [24] applying the 'normal' (default) hybrid mode.
112 For isolate A18Y016a, this did not result in a closed genome, so subsequently Canu v1.7 [25]

113 was used to perform an initial assembly using the long-read MinION data, and Pilon v1.22 [26]
114 applied to polish the assembly using Illumina reads, including three iterations. Tablet
115 v1.21.02.08 [27] was used to visually check the distance between forward and reverse reads
116 across the start and end positions of the linearised chromosome to confirm the genomes' circular
117 nature.
118 Prokka v1.13 [28] and BLAST2GO [29] were used to predict coding sequences (CDS), tRNA,
119 rRNA and tmRNA. The general feature format (.gff) files obtained by the two programmes were
120 merged to build a complete .gff annotation file. The PHASTER server (<https://phaster.ca/>) [30]
121 was used to identify and annotate prophage sequences. plasmidFinder server v2.1
122 (<https://cge.cbs.dtu.dk/services/PlasmidFinder/>) [31] was used to search for plasmid sequence
123 data in the genomes against the replicon database.
124 BWA-MEM [32] with default settings was used to map Illumina reads of each albatross isolate
125 against the A18Y020d closed genome, and the mpileup command in SAMtools v1.13 [33] was
126 used to generate summary text files of coverages across the reference genome. The pileup2snp
127 command in Varscan v2.4.4 [34] was used to generate a list of high-quality single nucleotide
128 polymorphisms (SNPs) that were filtered by applying parameters '--min-coverage' of 10, '--min-
129 var-freq' of 0.8 and '--min-avg-qual' of 20. A custom R script was developed to generate a
130 multi-fasta alignment file based on concatenated SNPs. Roary v3.12.0 [35], with default
131 parameters setting, was used to construct a core gene alignment for a wider phylogenetic
132 analysis, i.e. to place the albatross isolates within the context of global *Erysipelothrix* spp.
133 genomes. This included 106 genomes: a) the 16 isolates from this study; b) 86 previously
134 sequenced *E. rhusiopathiae* isolates, representing a diversity of geographic and host origins and
135 spanning the clades described for this species (Clade 1 (n = 7), Clade 2 (n = 14), intermediate
136 clade (n = 8), Clade 3 (n = 57)) [36]; c) two genomes of *Erysipelothrix piscisarius* (formerly
137 referred to as *E. sp. strain 2*) [37, 38]: the type strain 15TAL0474^T (GCA_003931795.1) [39],
138 and EsS2-7-Brazil (GCA_016617655.1). *E. piscisarius* is the most closely-related species to *E.*
139 *rhusiopathiae* currently known; d) a recently deposited *Erysipelothrix* genome, reported as a new
140 species, *Erysipelothrix sp. Poltava* (GCA_023221615.1); and e) *Erysipelothrix tonsillarum* strain
141 DSM14972^T (GCA_000373785.1), which was used as an outgroup, since this was previously
142 reported to be ancestral to these other species [36]. The SNP alignment and core gene alignment
143 files were used to estimate phylogenetic trees for the 16 albatross isolates and the full set of

144 *Erysipelothrix* spp. genomes, respectively, using RAxML v8.2.4 [40] with Maximum-Likelihood
145 (ML) method implementing a GTR-GAMMA nucleotide substitution model, selected after using
146 MrModeltest v2.4 [41].
147 To determine the relatedness of the albatross isolates from this study with other less well-
148 described *Erysipelothrix* species available in the Genome Taxonomy Database
149 (<https://gtdb.ecogenomic.org/>), a separate phylogenetic tree based on the *rpoB* gene (beta subunit
150 of RNA polymerase) was estimated. This gene has been found to be a more suitable marker for
151 *Erysipelothrix* species delineation than the widely used 16S rRNA gene [37]. Nucleotide
152 sequences of *rpoB* genes were extracted from whole-genome assemblies of representative
153 samples (one per species, n = 16 total), and aligned using L-INS-i method as implemented in
154 MAFFT v7.313 [42]. Tree topology was then inferred by RAxML v8.2.4, implementing the GTR
155 model of nucleotide substitution; robustness of the tree was assessed by performing 1000
156 bootstrap replicates. For completeness, the 16S rRNA gene sequence of the newly sequenced
157 strain A18Y020d was extracted through blastn alignment and compared with 16S gene
158 sequences from *E. rhusiopathiae* strain ATCC 19414^T (NR_040837.1), *E. piscisicarius* strain
159 15TAL0474^T (NR_170392.1), and *E. tonsillarum* strain DSM 14972^T (NR_040871.1).
160 Two different genome relatedness indices were calculated in order to compare the *Erysipelothrix*
161 spp. isolated from albatrosses with previously characterised *Erysipelothrix* species and determine
162 the species assignment. Pairwise nucleotide-level comparisons were made using average
163 nucleotide identity (ANI) Calculator for OrthoANIu [43] (<https://www.ezbiocloud.net/tools/ani>),
164 and digital DNA-DNA hybridisation (dDDH) was performed using GGDC [44, 45]
165 (<https://ggdc.dsmz.de>), using Formula 2. This formula is independent of genome length and is
166 thus robust to the use of incomplete draft genome assemblies, as well as differences in gene
167 content.
168 The output of Roary was used in Scoary v1.6.16 [46] to conduct a pan-GWAS (genome-wide
169 association study) in order to identify any genes that might be unique to the albatross isolates,
170 where the binary trait considered was ‘albatross origin’ yes/no. NCBI blastn searches were
171 conducted on the Scoary results (i.e. genes identified as unique to the albatross isolates that were
172 absent in other *Erysipelothrix* spp.) to verify their absence in other bacterial species. Primer sets
173 typically used to amplify *Erysipelothrix* spp. [47] and *E. rhusiopathiae* specifically [19] were
174 evaluated *in silico* against the complete assemblies using Geneious v11.0.5 [48] to check for any

175 mismatches. Finally, to further characterise the *Erysipelothrix* spp. isolates from albatross, blastn
176 searches of the *de novo* Illumina assemblies were conducted to determine *in silico* the serotype
177 [49] and whether any *spa* genes (A, B or C) [50] were present, using methods previously
178 described [51].

179

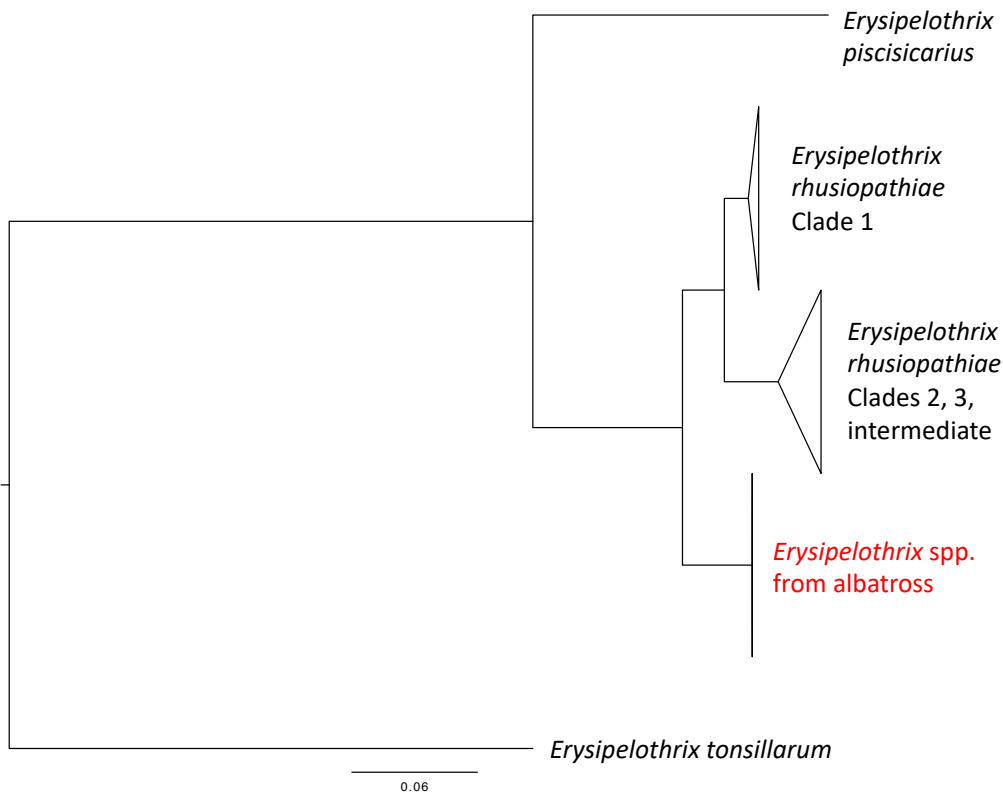
180 **Results**

181 Sixteen isolates identified as *Erysipelothrix* spp. were recovered from the carcasses of three
182 albatross chicks. Fifteen of these were initially identified as *E. rhusiopathiae* based on colony
183 morphology and MALDI-TOF with 99.9% confidence, and the other strain was identified as an
184 *Erysipelothrix* sp. Phenotypic characteristics determined during strain isolation are given in the
185 species description below. All 16 isolates were genome sequenced on the Illumina platform,
186 which yielded a high average depth of coverage of at least 58X for all genomes (Table S1). *De*
187 *nov*o genome assemblies comprised between 52 and 58 contigs per genome, with a total
188 assembly length ranging from 1.80 to 1.83 megabases. Complete circular genomes were obtained
189 for isolates A18Y020d^T (Figure S2; assembly GCA_940143175) – designated the type strain –
190 and A18Y016a (assembly GCA_940143155), with lengths of 1,908,712 bp and 1,910,750 bp,
191 respectively. This is approximately 120 kilobases (kb) longer than the Fujisawa *E. rhusiopathiae*
192 reference genome, corresponding to ~100 additional CDS (Table S2). Like the Fujisawa *E.*
193 *rhusiopathiae* genome, both genomes from albatrosses had a GC content of 36.6%. Preliminary
194 annotation of the two albatross-derived genomes using Prokka showed that this method predicted
195 a large number of hypothetical proteins, amounting to approximately one-third of predicted CDS.
196 Blast2GO facilitated the annotation of just over 1500 further CDS in each genome.

197 Using PHASTER, an intact prophage region comprising 48 genes was detected in both genomes
198 (Figure S3, Table S2). Based on global alignment using MUSCLE, this showed 43.6%
199 nucleotide identity with the incomplete phage sequence in the *E. rhusiopathiae* Fujisawa
200 reference genome (gene loci ERH_0581 to ERH_0629). No plasmids were detected in either
201 genome.

202 Pairwise identity of the 16S rRNA gene sequence from strain A18Y020d^T with that of *E.*
203 *rhusiopathiae* strain ATCC 19414^T was 99.9%; only two SNP differences were present across
204 1479 nucleotides (at positions 472 and 473; TC instead of CT). Only three further nucleotide
205 differences distinguished these sequences from the 16S sequence of *E. tonsillarum* strain DSM

206 14972^T (99.8% pairwise identity). Eight gaps were found in the alignment of the 16S sequence
207 from *E. piscisarius* strain 15TAL0474^T compared with the others, which also shared one of the
208 SNPs from A18Y020d^T at position 472, and had one additional nucleotide difference. The
209 phylogenetic tree (Figure 1), based on 427 conserved core genes, showed that the 16
210 *Erysipelothrix* spp. isolates from albatrosses form a monophyletic clade distinct from previously
211 characterised *Erysipelothrix* spp. genomes, which shares its most recent common ancestor with
212 *E. rhusiopathiae*. Based on the *rpoB* gene phylogeny, other *Erysipelothrix* species discovered to
213 date are more phylogenetically distant (Figure S4), with the exception of the recently deposited
214 sequence for *Erysipelothrix* sp. Poltava; based on *rpoB* analysis, ANI/dDDH (Figure S5) and
215 phylogenomic analyses (Figure S6), strain Poltava is concluded to belong to *E. rhusiopathiae*
216 Clade 2.



217
218 **Figure 1. Maximum likelihood phylogenetic tree of *Erysipelothrix* spp.** Isolates (n = 16) from
219 Indian yellow-nosed albatross chicks on Amsterdam Island (red text) were compared with other
220 representative *Erysipelothrix* genomes. The tree was estimated in RAxML using a multiple
221 sequence alignment of conserved (core) genes identified using Roary, implementing a GTR-
222 GAMMA model, and visualised using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>). *E.*

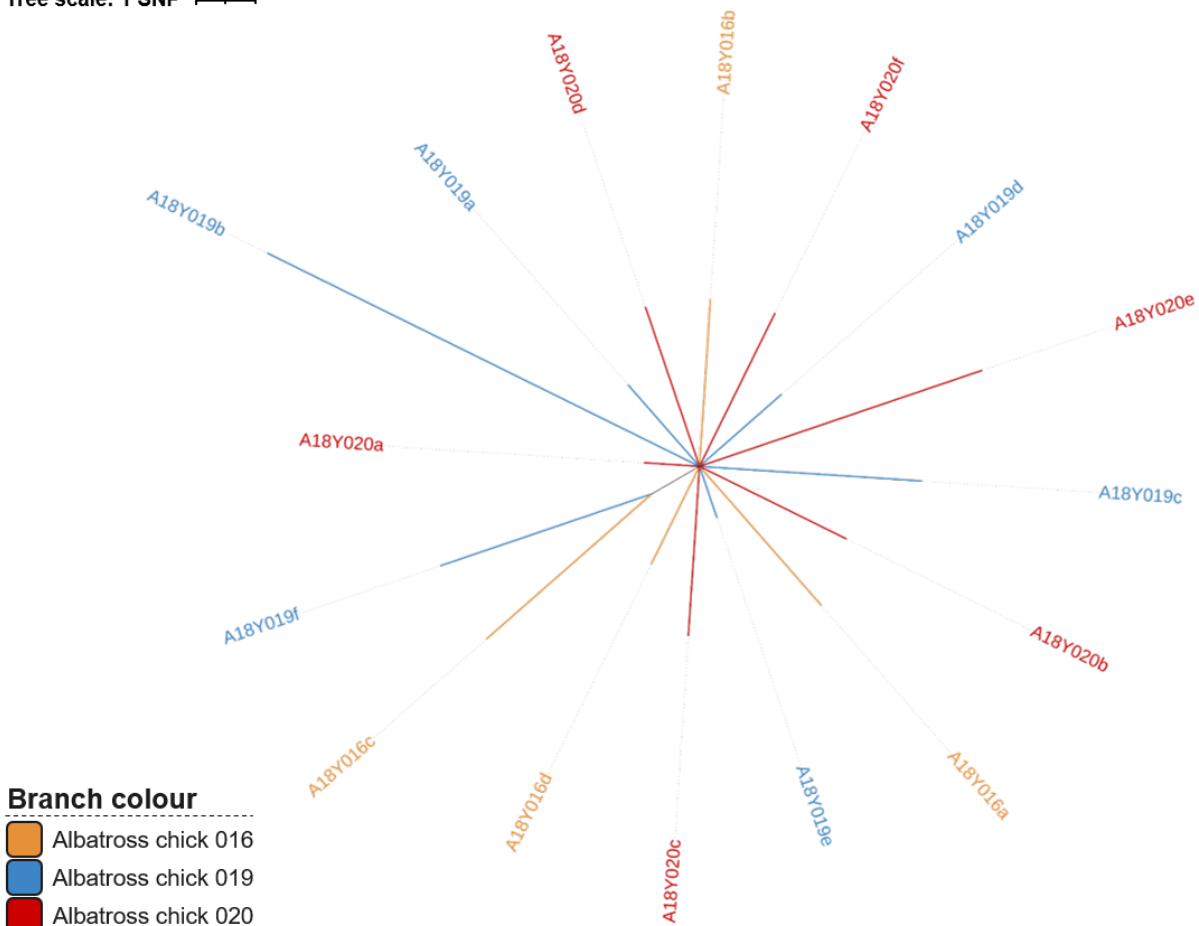
223 *rhusiopathiae* clades are shown collapsed. Bootstrap values for all major nodes (those shown)
224 were 100%. *E. tonsillarum* strain DSM14972^T was used as an outgroup for rooting.

225

226 Few SNP differences were observed among the albatross isolates. A total of 53 high-quality
227 unique SNPs were detected across all 16 genomes when mapping against the closed reference
228 genome (A18Y020d^T; Figure 2). All genomes were unique, including where multiple isolates
229 were sequenced from the same carcass (Table S1). The maximum pairwise SNP difference
230 separating two isolates was 13. *In silico* serotyping based on a genomic polysaccharide
231 biosynthetic locus [49] determined these isolates belong to serotype 1b. All genomes had
232 identical sequences of the *spaA* gene, belonging to Group 2 [51]. These all had one unique amino
233 acid residue at position 288 in comparison with any previously described sequences (serine
234 instead of alanine).

235

Tree scale: 1 SNP ———



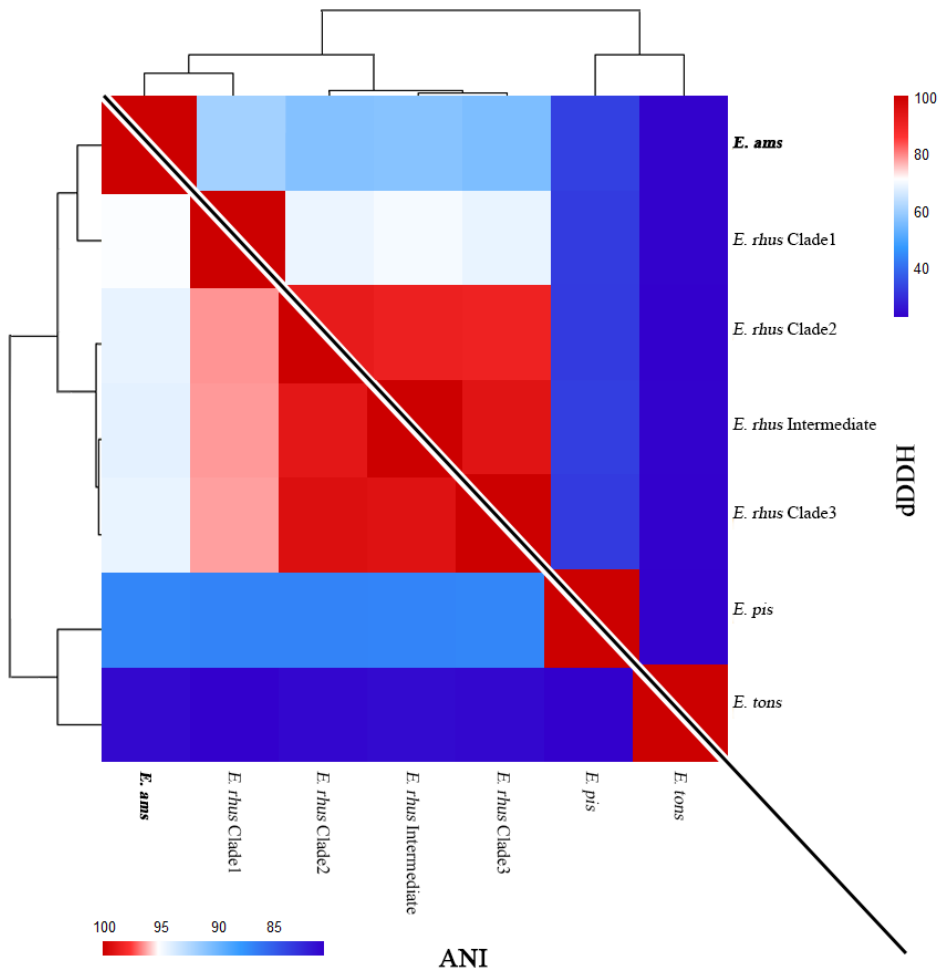
236

237

238 **Figure 2. Unrooted maximum likelihood tree of 16 *Erysipelothrix* spp. isolates from Indian**
239 **yellow-nosed albatross from Amsterdam Island.** The tree was based on high-quality
240 concatenated SNPs detected through mapping to the circularised genome of A18Y020d^T. Isolates
241 originated from three different carcasses, which are distinguished by different branch colours (as
242 shown in the legend). Tree was plotted with Interactive Tree Of Life (iTOL) version 6.3.2 [52].

243

244 The ANI scores comparing the representative *Erysipelothrix* spp. type strain from albatross
245 against *E. rhusiopathiae* were 95.0% with *E. rhusiopathiae* Clade 1, and just above 94% for
246 representatives from the other clades, whereas they were 86.9% with *E. piscisicarius* and 80.9%
247 with *E. tonsillarum* (Figure 3). The dDDH scores with *E. rhusiopathiae* Clades 1, 2 and 3 were
248 60.0%, 56.3% and 55.5%, respectively, while they were 32.7% with *E. piscisicarius* and 22.8%
249 with *E. tonsillarum*. The accepted values for categorising isolates as the same species are >95%
250 similarity for ANI [53, 54] and >70% for dDDH [45]. Thus, based on these standard metrics for
251 genomic comparison, and the phylogenomic analysis, we conclude that these isolates belong to a
252 novel species of the genus *Erysipelothrix*.



253

254 **Figure 3. Genome relatedness of *Erysipelothrix* spp. from albatross to previously-described**

255 ***Erysipelothrix* species. (A) The ANI scores comparing the *Erysipelothrix* spp. from albatross**

256 **(represented by A18Y020d^T) against other *Erysipelothrix* spp. were all below the threshold for**

257 **being considered a common species (>95%). (B) The dDDH scores comparing the *Erysipelothrix***

258 **spp. from albatross against other *Erysipelothrix* spp. were also well below the accepted threshold**

259 **for being considered a single species (>70%). Genomes included for comparison: *E.***

260 ***rhusiopathiae* Clade 1 (P-92), Clade 2 (P-190), intermediate (Fujisawa), and Clade 3 (F793); *E.***

261 ***piscisicarius* (15TAL0474^T); and *E. tonsillarum* (DSM14972^T).**

262

263 Two hundred and two genes unique to the albatross-derived *Erysipelothrix* genomes (present in

264 all 16 vs. absent in all 90 other genomes) were found using Scoary. However, after conducting

265 blastn searches, the majority of these genes had significant hits to genes in *E. rhusiopathiae* or

266 bacteria within the family *Erysipelotrichaceae*, or more rarely, with other bacteria. Only six
267 genes were found with no hits or hits with low query cover ($\leq 12\%$); a further six genes had
268 $<75\%$ nucleotide identity with the highest scoring blastn hit (Table S3).

269 The *Erysipelothrix* spp. DNA from all albatross isolates produced typical amplification curves
270 using the probe-based qPCR by Pal et al. [19] designed to detect *E. rhusiopathiae*. This is despite
271 the fact that the forward primer and probe sequences had two mismatches each (Figure S7);
272 based on *in silico* alignment, the 165 bp segment being amplified corresponded to positions
273 1,211,721 to 1,221,885 in the A18Y020d^T genome. The primers designed by Makino et al. [47]
274 for *Erysipelothrix* species more broadly were a perfect match, and based on *in silico* alignment
275 would be expected to amplify the targeted 407 bp segment of the 16S rRNA gene (positions
276 1,221,445 to 1,221,851 in the A18Y020d genome).

277

278 Discussion

279 Through genomic comparisons, we have determined that the *Erysipelothrix* strains isolated from
280 Indian yellow-nosed albatross chick carcasses on Amsterdam Island belong to a previously
281 undescribed species, which we propose be named *Erysipelothrix amsterdamensis* sp. nov. This
282 reflects its geographic origin, as well as the co-occurrence of the iconic and endemic Amsterdam
283 albatross, *D. amsterdamensis*. Anthropogenic sources are often suspected when novel pathogens
284 are detected in new locations. While livestock was originally suggested as a possible source of
285 the *Erysipelothrix* spp. found on Amsterdam Island [5], we were unable to find such a link, since
286 this particular *Erysipelothrix* species has not been previously documented elsewhere. While no
287 earlier *Erysipelothrix* spp. isolates from this island were sequenced, using traditional phenotypic
288 methods, initial investigations found that they belonged to serotype 1b [5]. Our finding of the
289 same serotype using *in silico* approaches may suggest that the same species/strain has been in
290 circulation since at least the mid-1990s. Previous PCR amplification of *Erysipelothrix* spp.
291 isolates from Amsterdam Island was performed using the primers by Makino et al. [47] targeting
292 the 16S rRNA gene and which should amplify sequence from all members of the genus
293 *Erysipelothrix*. We also found that the primers and probe described by Pal et al. [19] targeting a
294 noncoding region 3' to the 5S rRNA gene for detection of *E. rhusiopathiae* were able to amplify
295 DNA from *E. amsterdamensis*; care should therefore be taken when interpreting results using
296 this primer/probe set on *Erysipelothrix* spp. isolated from rare sources. Development of a PCR

297 protocol that distinguishes between these two species would be valuable. The genes found to be
298 unique to *E. amsterdamensis* in this study (Table S3) would be good initial candidate targets.
299 Moreover, the two closed genomes generated during this study will facilitate further comparative
300 genomic studies within the genus *Erysipelothrix*. Our comparisons of the 16S rRNA and *rpoB*
301 gene sequences from different *Erysipelothrix* species highlight the limited nucleotide diversity in
302 the 16S rRNA gene, and confirm previous suggestions that the *rpoB* gene is a more suitable
303 marker for *Erysipelothrix* species delineation [37].

304 SNP analysis of 16 *E. amsterdamensis* isolates from the three albatross chicks showed that they
305 are highly related. Given that the number of SNPs distinguishing isolates of the same carcass was
306 similar to that among isolates from different carcasses – as well as the close proximity in space
307 and time of the mortality events – this strongly suggests that the three chicks were infected from
308 a common source. Whether this small number of SNPs arose during infection within the host, or
309 was already present in the environment, remains unknown; a molecular clock for *Erysipelothrix*
310 has yet to be established [36]. While we cannot rule out the possibility that mutations occurred
311 while the strains were maintained in conservation medium (several months unfrozen), we feel it
312 is unlikely that sufficient replication occurred to explain the number of variants observed within
313 and among these isolates. Further sampling of the albatross colony would help to determine
314 whether this *Erysipelothrix* species was a new introduction to Amsterdam Island, and whether
315 other strains are in circulation. Moreover, sampling from islands at comparable latitudes would
316 help to elucidate its geographic distribution.

317 Interestingly, *E. amsterdamensis* harbours a *spaA* gene, which codes for what is commonly
318 described as one of the most critical proteins for immunogenicity and virulence of *E.*
319 *rhusiopathiae* [55–57]. This gene has been found to occur in all Clade 2, Clade 3 and
320 intermediate clade *E. rhusiopathiae* genomes described to date [51], whereas *E. rhusiopathiae*
321 Clade 1 carries a *spaB* gene [36], and *E. piscisarius* carries a *spaC* gene [39]. That *E.*
322 *amsterdamensis* was isolated in association with chick mortalities in the absence of other
323 bacteria suggests it was the likely cause of death, and the presence of the *spaA* gene lends further
324 support to the likely virulence of this species. However, further investigation, including
325 challenge studies, will be necessary for such characterisation.

326 *E. rhusiopathiae* is typically considered an opportunistic pathogen, often manifesting clinically
327 in stressed individuals or populations, for example following the kakapo translocation [15].

328 Further studies should explore whether *Erysipelothrix*-associated mortalities detected in other
329 subantarctic islands could have been caused by the new species described here, or other novel
330 taxa, and which set of potential host species of those relatively simple island communities may
331 be involved in the epidemiological dynamics.

332

333 **Description of *Erysipelothrix amsterdamensis* sp. nov**

334 *Erysipelothrix amsterdamensis* sp. nov. (am.ster.dam.en'sis. N.L. fem. *amsterdamensis* referring
335 to Amsterdam island, from which the first strains to be characterised were isolated).

336 Cells are Gram-positive, rod-shaped, and non-motile. Small pin-point sized, round, flat grey
337 colonies with smooth contours are observed on 5% horse blood agar after 24 h growth at 37°C at
338 normal atmospheric conditions. Catalase negative. The G+C content of the genomic DNA of the
339 type strain is 36.6%. Whole genome sequencing indicates isolates carry the *Erysipelothrix spaA*
340 gene and belong to *Erysipelothrix* serotype 1b. Members of the species can be distinguished
341 from other members of the genus *Erysipelothrix* based on phylogenomic analysis and genomic
342 metrics.

343 The type strain, A18Y020d^T, was isolated from an Indian yellow-nosed albatross chick carcass
344 on Amsterdam Island. The whole genome of the type strain A18Y020d is available in GenBank
345 (GCA_940143175).

346 **Author Statements**

347 *Author contributions*

348 Formal analysis: JZ. Investigation: JT, AC, HG, AG, TLF, TB. Methodology (Field): TB, AC,
349 JT, AG; (Laboratory): HG, TB, AC, JT, AG; (Bioinformatics): JZ, MM, TLF. Supervision: MM,
350 TLF, TB. Visualisation: JZ, MM. Writing – original draft: JZ, TLF. Writing – review & editing:
351 all authors.

352 *Conflicts of interest*

353 The authors declare no conflict of interest.

354 *Funding information*

355 MM was supported by an Academy of Medical Sciences Springboard award, with contribution
356 from the Wellcome Trust and Global Challenges Research Fund (SBF005\1023, to TLF). JT
357 received support from CEVA and ANRT for a CIFRE PhD fellowship and project France relance
358 TVACALBA. AG was supported by the US National Science Foundation (DEB-1557022) and
359 the Strategic Environmental Research and Development Program (RC2635). TLF was supported
360 by a Biotechnology and Biological Sciences Research Council Discovery Fellowship
361 (BB/R012075/1). The research was also supported by the French Polar Institute (IPEV
362 ECOPATH-1151), ANR ECOPATHS (ANR-21-CE35-0016), Zone Atelier Antarctique et Terres
363 Australes (ZATA) and OSU OREME ECOPOP. The funders had no role in study design, data
364 collection and analysis, or preparation of the manuscript.

365

366 *Ethical approval*

367 The experimental design was approved by the French Regional Animal Experimentation Ethical
368 Committee n°036 (Ministry of Research permit #10257-2018011712301381) and by the Comité
369 de l'Environnement Polaire (A-2018-123 and A 2018-139 for 2018-2019).

370 *Acknowledgements*

371 We thank Nicolas Keck, Karine Lemberger, Christophe Barbraud, Karine Delord and Iain
372 Sutcliffe for discussions related to the work. Support is acknowledged from Réserve Naturelle
373 Nationale des Terres Australes Françaises, notably from Célia Lesage. We thank technical

374 support agents at Ceva Biovac for help with laboratory work. This paper is a contribution of
375 IPEV project ECOPATH-1151 to the Plan National d'Action Albatros d'Amsterdam.

376

377

378 **References**

- 379 1. **Heard MJ, Smith KF, Ripp K, Berger M, Chen J, et al.** Increased threat of disease as
380 species move towards extinction. *Conserv Biol* 2013;27:1378–1388.
- 381 2. **Grimaldi W, Jabour J, Woehler EJ.** Considerations for minimising the spread of infectious
382 disease in Antarctic seabirds and seals. *Polar Record* 2011;47:56–66.
- 383 3. **Cerdà-Cuéllar M, Moré E, Ayats T, Aguilera M, Muñoz-González S, et al.** Do humans
384 spread zoonotic enteric bacteria in Antarctica? *Sci Total Environ* 2019;654:190–196.
- 385 4. **Uhart MM, Gallo L, Quintana F.** Review of diseases (pathogen isolation, direct recovery
386 and antibodies) in albatrosses and large petrels worldwide. *Bird Conservation International*
387 2018;28:169–196.
- 388 5. **Weimerskirch H.** Diseases threaten Southern Ocean albatrosses. *Polar Biol* 2004;27:374–
389 379.
- 390 6. **Jaeger A, Lebarbenchon C, Bourret V, Bastien M, Lagadec E, et al.** Avian cholera
391 outbreaks threaten seabird species on Amsterdam Island. *PLOS ONE* 2018;13:e0197291.
- 392 7. **Jaeger A, Gamble A, Lagadec E, Lebarbenchon C, Bourret V, et al.** Impact of Annual
393 Bacterial Epizootics on Albatross Population on a Remote Island. *EcoHealth* 2020;17:194–
394 202.
- 395 8. **Bourret V, Gamble A, Tornos J, Jaeger A, Delord K, et al.** Vaccination protects
396 endangered albatross chicks against avian cholera. *Conservation Letters* 2018;11:e12443.
- 397 9. **Gamble A, Garnier R, Jaeger A, Gantelet H, Thibault E, et al.** Exposure of breeding
398 albatrosses to the agent of avian cholera: dynamics of antibody levels and ecological
399 implications. *Oecologia* 2019;189:939–949.
- 400 10. **Gamble A, Bazire R, Delord K, Barbraud C, Jaeger A, et al.** Predator and scavenger
401 movements among and within endangered seabird colonies: Opportunities for pathogen
402 spread. *Journal of Applied Ecology* 2020;57:367–378.
- 403 11. **Opriessnig T, Coutinho TA.** Erysipelas. In: *Diseases of Swine*. John Wiley & Sons, Ltd. pp.
404 835–843.
- 405 12. **Bricker JM, Saif YM.** Erysipelas. In: Swayne DE (editor). *Diseases of Poultry*. Somerset,
406 NJ: John Wiley & Sons; 2013. pp. 986–993.

- 407 13. **Brooke CJ, Riley TV.** Erysipelothrix rhusiopathiae: bacteriology, epidemiology and clinical
408 manifestations of an occupational pathogen. *J Med Microbiol* 1999;48:789–799.
- 409 14. **Kutz S, Bollinger T, Branigan M, Checkley S, Davison T, et al.** Erysipelothrix
410 rhusiopathiae associated with recent widespread muskox mortalities in the Canadian Arctic.
411 *Can Vet J* 2015;56:560–563.
- 412 15. **Gartrell BD, Alley MR, Mack H, Donald J, McInnes K, et al.** Erysipelas in the critically
413 endangered kakapo (*Strigops habroptilus*). *Avian Pathol* 2005;34:383–387.
- 414 16. **Jayasinghe M, Midwinter A, Roe W, Vallee E, Bolwell C, et al.** Seabirds as possible
415 reservoirs of Erysipelothrix rhusiopathiae on islands used for conservation translocations in
416 New Zealand. *J Wildl Dis* 2021;57:534–542.
- 417 17. **Wolcott MJ.** Erysipelas. In: Thomas NJ, Hunter DB, Atkinson CT (editors). *Infectious*
418 *Diseases of Wild Birds*. Blackwell Publishing Professional. pp. 332–340.
- 419 18. **Jensen WI, Cotter SE.** An outbreak of erysipelas in eared grebes (*Podiceps nigricollis*). *J*
420 *Wildl Dis* 1976;12:583–586.
- 421 19. **Pal N, Bender JS, Opriessnig T.** Rapid detection and differentiation of Erysipelothrix spp.
422 by a novel multiplex real-time PCR assay. *J Appl Microbiol* 2010;108:1083–1093.
- 423 20. **Bolger AM, Lohse M, Usadel B.** Trimmomatic: a flexible trimmer for Illumina sequence
424 data. *Bioinformatics* 2014;30:2114–2120.
- 425 21. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, et al.** SPAdes: a new
426 genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*
427 2012;19:455–477.
- 428 22. **Gurevich A, Saveliev V, Vyahhi N, Tesler G.** QUAST: quality assessment tool for genome
429 assemblies. *Bioinformatics* 2013;29:1072–1075.
- 430 23. **Connor TR, Loman NJ, Thompson S, Smith A, Southgate J, et al.** CLIMB (the Cloud
431 Infrastructure for Microbial Bioinformatics): an online resource for the medical
432 microbiology community. *Microbial Genomics*, 2016;2:e000086.
- 433 24. **Wick RR, Judd LM, Gorrie CL, Holt KE.** Unicycler: Resolving bacterial genome
434 assemblies from short and long sequencing reads. *PLoS Computational Biology*
435 2017;13:e1005595.
- 436 25. **Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, et al.** Canu: scalable and
437 accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome*
438 *Res* 2017;27:722–736.
- 439 26. **Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, et al.** Pilon: an integrated tool for
440 comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*
441 2014;9:e112963.

- 442 27. **Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, et al.** Using Tablet for visual
443 exploration of second-generation sequencing data. *Brief Bioinform* 2013;14:193–202.
- 444 28. **Seemann T.** Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–
445 2069.
- 446 29. **Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, et al.** Blast2GO: a universal tool
447 for annotation, visualization and analysis in functional genomics research. *Bioinformatics*
448 2005;21:3674–3676.
- 449 30. **Arndt D, Grant JR, Marcu A, Sajed T, Pon A, et al.** PHASTER: a better, faster version of
450 the PHAST phage search tool. *Nucleic Acids Res* 2016;44:W16–W21.
- 451 31. **Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, et al.** In Silico
452 Detection and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus Sequence
453 Typing. *Antimicrobial Agents and Chemotherapy* 2014;58:3895–3903.
- 454 32. **Li H.** Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
455 *Oxford University Press*;arXiv:1303.3997.
- 456 33. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al.** The Sequence Alignment/Map
457 format and SAMtools. *Bioinformatics* 2009;25:2078–2079.
- 458 34. **Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, et al.** VarScan 2: somatic
459 mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*
460 2012;22:568–576.
- 461 35. **Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, et al.** Roary: rapid large-scale
462 prokaryote pan genome analysis. *Bioinformatics* 2015;31:3691–3693.
- 463 36. **Forde T, Biek R, Zadoks R, Workentine ML, De Buck J, et al.** Genomic analysis of the
464 multi-host pathogen *Erysipelothrix rhusiopathiae* reveals extensive recombination as well as
465 the existence of three generalist clades with wide geographic distribution. *BMC Genomics*
466 2016;17:461.
- 467 37. **Grazziotin AL, Vidal NM, Hoepers PG, Reis TFM, Mesa D, et al.** Comparative genomics
468 of a novel clade shed light on the evolution of the genus *Erysipelothrix* and characterise an
469 emerging species. *Sci Rep* 2021;11:3383.
- 470 38. **Grazziotin AL, Vidal NM, Hoepers PG, Reis TFM, Mesa D, et al.** Author Correction:
471 Comparative genomics of a novel clade shed light on the evolution of the genus
472 *Erysipelothrix* and characterise an emerging species. *Sci Rep* 2021;11:9861.
- 473 39. **Pomaranski EK, Griffin MJ, Camus AC, Armwood AR, Shelley J, et al.** Description of
474 *Erysipelothrix pisciscarius* sp. nov., an emergent fish pathogen, and assessment of virulence
475 using a tiger barb (*Puntigrus tetrazona*) infection model. *Int J Syst Evol Microbiol*
476 2020;70:857–867.

- 477 40. **Stamatakis A.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
478 phylogenies. *Bioinformatics* 2014;30:1312–1313.
- 479 41. **Nylander JAA.** *MrModeltest*. Uppsala University: Evolutionary Biology Centre; 2004.
- 480 42. **Katoh K, Standley DM.** MAFFT multiple sequence alignment software version 7:
481 improvements in performance and usability. *Mol Biol Evol* 2013;30:772–780.
- 482 43. **Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J.** A large-scale evaluation of algorithms to
483 calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110:1281–1286.
- 484 44. **Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M.** TYGS and LPSN: a
485 database tandem for fast and reliable genome-based classification and nomenclature of
486 prokaryotes. *Nucleic Acids Research* 2021;gkab902.
- 487 45. **Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M.** Genome sequence-based species
488 delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*
489 2013;14:1–14.
- 490 46. **Brynildsrud O, Bohlin J, Scheffer L, Eldholm V.** Rapid scoring of genes in microbial pan-
491 genome-wide association studies with Scoary. *Genome Biol* 2016;17:238.
- 492 47. **Makino S, Okada Y, Maruyama T, Ishikawa K, Takahashi T, et al.** Direct and rapid
493 detection of *Erysipelothrix rhusiopathiae* DNA in animals by PCR. *J Clin Microbiol*
494 1994;32:1526–1531.
- 495 48. **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, et al.** Geneious Basic: an
496 integrated and extendable desktop software platform for the organization and analysis of
497 sequence data. *Bioinformatics* 2012;28:1647–1649.
- 498 49. **Shiraiwa K, Ogawa Y, Nishikawa S, Eguchi M, Shimoji Y.** Identification of serovar 1a,
499 1b, 2, and 5 strains of *Erysipelothrix rhusiopathiae* by a conventional gel-based PCR. *Vet*
500 *Microbiol* 2018;225:101–104.
- 501 50. **To H, Nagai S.** Genetic and antigenic diversity of the surface protective antigen proteins of
502 *Erysipelothrix rhusiopathiae*. *Clin Vaccine Immunol* 2007;14:813–820.
- 503 51. **Forde TL, Kollanandi Ratheesh N, Harvey WT, Thomson JR, Williamson S, et al.**
504 Genomic and Immunogenic Protein Diversity of *Erysipelothrix rhusiopathiae* Isolated From
505 Pigs in Great Britain: Implications for Vaccine Protection. *Front Microbiol* 2020;11:418.
- 506 52. **Letunic I, Bork P.** Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree
507 display and annotation. *Nucleic Acids Res* 2021;49:W293–W296.
- 508 53. **Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S.** High throughput
509 ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*
510 2018;9:5114.

- 511 54. **Kim M, Oh H-S, Park S-C, Chun J 2014.** Towards a taxonomic coherence between
512 average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation
513 of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*;64:346–
514 351.
- 515 55. **Makino S, Yamamoto K, Murakami S, Shirahata T, Uemura K, et al.** Properties of
516 repeat domain found in a novel protective antigen, SpaA, of *Erysipelothrix rhusiopathiae*.
517 *Microb Pathog* 1998;25:101–109.
- 518 56. **Imada Y, Goji N, Ishikawa H, Kishima M, Sekizaki T.** Truncated surface protective
519 antigen (SpaA) of *Erysipelothrix rhusiopathiae* serotype 1a elicits protection against
520 challenge with serotypes 1a and 2b in pigs. *Infect Immun* 1999;67:4376–4382.
- 521 57. **Shimoji Y, Mori Y, Fischetti VA.** Immunological characterization of a protective antigen of
522 *Erysipelothrix rhusiopathiae*: identification of the region responsible for protective
523 immunity. *Infect Immun* 1999;67:1646–1651.
- 524