Vagococcus fluvialis isolation and sequencing from urine of healthy cattle

Silvia Giannattasio-Ferraz,[1] Adriana Ene,[2] Laura Maskeri,[2] André Penido Oliveira,[3] Edel F. Barbosa-Stancioli,[1] and Catherine Putonti[2,4,5,][*]

[1]Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil
[2]Bioinformatics Program, Loyola University Chicago, Chicago, IL 60660, USA
[3]Empresa de Pesquisa Agropecuária de Minas Gerais—EPAMIG, Uberaba, MG 1170-495, Brazil
[4]Department of Biology, Loyola University Chicago, Chicago, Chicago, IL 60660, USA
[5]Department of Microbiology and Immunology, Stritch School of Medicine, Loyola University Chicago, Maywood, IL 60153, USA

*Corresponding author: cputonti@luc.edu

Abstract

While the gram-positive bacterium *Vagococcus fluvialis* has been isolated from the environment as well as fish, birds, and mammals, very little is known about the species. *V. fluvialis* is believed to be a probiotic in fishes. However, within mammals, it is more frequently isolated from infectious tissue, including on rare occasions human and livestock lesions. Prior to the study described here, *V. fluvialis* had never been found in healthy bovine animals. Here, we present the complete genomes of *V. fluvialis* UFMG-H6, UFMG-H6B, and UFMG-H7, novel strains isolated from urine samples from healthy bovine females. These are the first genomes of mammalian isolates and the first description of *V. fluvialis* from urine. The genomes did not encode for any known virulence genes, suggesting that they may be commensal members of the urine microbiota.

Keywords: *Vagococcus fluvialis*; cattle microbiota; urinary tract bacteria

Introduction

Vagococcus fluvialis is a gram-positive, catalase negative bacterium first isolated in 1974 from chicken feces and river water. While these first isolates were originally identified as members of the *Lactococcus* genus, 16S rRNA gene sequencing led to their classification as a new genus: *Vagococcus* (Collins *et al.* 1989). This species was detected in mammals in 1994, isolated from lesions in pigs, horses, cats, and cattle (Pot *et al.* 1994). *V. fluvialis* was first associated with human infections in 1997 (from blood culture, peritoneal fluid, and wound), however, these isolates differed from those collected from pigs (of an unknown clinical source) both phenotypically and by their SDS-PAGE profile (Teixeira *et al.* 1997). So far, only two other papers have described the circulation of the species in humans, one isolate from a root-filled tooth associated with periradicular lesions and the other an isolate from a rare infective endocarditis (Al-Ahmad *et al.* 2008; Jadhav and Pai 2019). With a distinct lipid pattern that has a high concentration of d-alanylcardiolipin (Fischer and Arneth-Seifert 1998), *V. fluvialis* was also described as a potential probiotic to fishes *in vitro* and *in vivo* with an immunomodulatory effect to the host and the protection against *Vibrio anguillarum*, an important fish pathogen (Román *et al.* 2012; Sorroza *et al.* 2012; Román *et al.* 2013).

There is a great lack of knowledge regarding this species. Relatively few genome assemblies are available for the genus, and prior to the study described here, only three draft *V. fluvialis* genomes were publicly available in GenBank. Here, we describe three new strains of *V. fluvialis* isolated from urine from healthy Gyr heifers: *V. fluvialis* UFMG-H6, UFMG-H6B, and UFMG-H7. Furthermore, our analysis found that only two of the previously published draft genomes are in fact *V. fluvialis*.

Materials and methods

The sample collection was occurred in May 2019 from a Brazilian herd composed of pure-by-origin Gyr cattle at the Agricultural Research Company of Minas Gerais State (EPAMIG). The reproduction in this livestock is realized by fixed time artificial insemination, without the presence of bulls. All of the experiments were performed according to relevant guidelines and were previously approved by the Ethics Committee in Animal Experimentation of the Universidade Federal de Minas Gerais, Brazil (CEUA/UFMG—40/2019).

For sampling, the animal's vulva was washed with soap and distilled water. Mid-stream urine was collected using a sterile 50 ml conical tube. The material was frozen and kept at –20°C until processing 48 hours later. Then 2 ml of aliquots were made, centrifuged and spread onto Lysogeny Broth (LB) agar plates. These plates were incubated overnight at 37°C and all of the individual colonies were picked and regrown in LB under the same conditions (overnight at 37°C). This process of plating and liquid

growth was repeated at least 3 times to obtain pure colonies. Single colonies were then picked and grown in liquid LB media overnight at 37°C, under agitation.

DNA was extracted using the Qiagen DNeasy UltraClean microbial kit (Qiagen, Hilden, Germany) according to manufacturer's instruction and quantified using a Qubit fluorometer. All of the isolates were identified by 16S rRNA gene sequencing using the 63F/1387R primer pair; sequencing was performed by Genewiz (New Brunswick, NJ) using each primer individually. The resulting sequences were queried against NCBI's 16S ribosomal sequence database via BLAST. Three isolates from two different animals were identified as *V. fluvialis*. Next, the DNA was sent to the Microbial Genomic Sequencing Center (MiGS) (Pittsburg, PA) for whole genome sequencing. The DNA was fragmented using an Illumina tagmentation enzyme and the indices were attached using PCR. Sequencing was performed using the Illumina NextSeq 550 platform producing 1.75 (UFMG-H6), 1.39 (UFMG-H6B), and 1.78 (UFMG-H7) million pairs of 150 nucleotide sequences. Next, the raw reads were trimmed using Sickle v1.33 (https://github.com/najoshi/sickle) and assembled using SPAdes v3.13.0 with the "only-assembler" option for k values of 55, 77, 99, and 127 (Bankevich *et al.* 2012). After assembly, the contigs were inspected using Geneious Prime (Biomatters Ltd., Auckland, New Zealand), removing contigs less than 500 nucleotides in length. The genome coverage for each assembly was calculated using BBMap v38.47 (https://sourceforge.net/projects/bbmap/). Annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.11 (Tatusova *et al.* 2016). Default parameters were used for each of the software tools, unless previously noted.

The draft genome assemblies were also examined for plasmid presence, resistance genes, secondary metabolites, prophage sequences, and the CRISPR/Cas system. To check if the strains have a plasmid, the Center for Genomic Epidemiology's (CGE) tool PlasmidFinder v2.1 was used with the following parameters: 60% minimum coverage, 90% minimum identity threshold, and gram-positive database (Carattoli *et al.* 2014). To complement this analysis, raw reads were also uploaded to the webtool PLACNETw to confirm the presence/absence of plasmids (Vielva *et al.* 2017). Assemblies were screened for antibiotic resistance genes using the CGE tool ResFinder v4.0, using default settings (Bortolaia *et al.* 2020). Secondary metabolites were predicted using the antiSMASH tool, with the strict detection parameter specified (Blin *et al.* 2019). The tool IslandViewer 4 was used to predict genomic islands (Bertelli *et al.* 2017). Each assembly as well as each publicly available *V. fluvialis* assembly was uploaded to the tool PHASTER for prophage prediction (Arndt *et al.* 2016). PHASTER predicts incomplete, questionable, and intact prophage regions. Intact nucleotide sequences were queried against NCBI nr/nt viral database for taxonomic classification. The genomes were also screened to detect CRISPR/Cas arrays using the webtool CRISPRCasFinder (Couvin *et al.* 2018).

The 16S rRNA sequences for each of the bovine *Vagococcus* genomes were extracted from the assembled genomes and queried against NCBI's rRNA_typestrains/16S_ribosomal_RNA database to retrieve 16S rRNA gene sequences from other *Vagococcus* species. The 16S rRNA sequence was also retrieved for *V. fluvialis* DSM 5731, as it is not represented in this database, and *E. faecium* DSM20477, to serve as an outgroup. The sequences were aligned using MAFFT v7.388 (Katoh and Standley 2013) a phylogenetic tree was derived using the FastTree v2.1.11 (Price *et al.* 2010) plug-in though Geneious Prime. The tree was visualized using iTOL v5 (Letunic and Bork 2016).

Table 1 Genome statistics for *V. fluvialis* bovine isolates

|	UFMG-H6	UFMG-H6B	UFMG-H7
Length (bp)	2,679,177	2,858,425	2,993,433
# Contigs	41	45	43
Genome coverage (x)	182	38	165
N_{50} score (bp)	135,292	112,454	131,353
GC content (%)	33.08	44.68	32.93
# Coding genes	2,626	2,792	2,906
# tRNAs	50	49	52

The average nucleotide identity (ANI) of *V. fluvialis* UFMG-H6, UFMG-H6B, and UFMG H7 was computed using the tool JSpeciesWS (Richter *et al.* 2016). Each draft assembly was compared to each other, the three publicly available draft genomes of *V. fluvialis* in the NCBI Assembly database, and all other *Vagococcus* strains available through JSpeciesWS. The three *V. fluvialis* draft genomes include: *V. fluvialis* NCDO 2497 (Accession no. GCA_003987575.1) and *V. fluvialis* DSM 5731 (Accession no. GCA_003337315.1), both isolated from chicken feces, and *V. fluvialis* bH819 (Accession no. GCA_900163795.1), which was isolated from cheese. ANI values were visualized using Python.

Based upon our *Vagococcus* ANI analysis, we found that *V. fluvialis* bH819 had an ANI of 75.6% to the two other previously published genomes for the species. This is well below the 95% ANI threshold expected for strains of the same species (Ramasamy *et al.* 2014). (This strain had been excluded from our 16S analysis as it did not have a complete 16S rRNA sequence.) The ANI analysis prompted our further investigation of this genome. Relative to other *Vagococcus* species in the JSpeciesWS database, *V. fluvialis* bH819 has an ANI value ranging between 67% and 72%. We also compared it to *V. carniphilus* (NZ_CP060720), which has a marginally larger ANI value of 76.37%. The bH819 genome is not a bad assembly; the genome is assembled in 22 contigs with 379x coverage. Thus, we do not believe that the low ANI values detected are due to assembly issues. Using the Similar Genome Finder through Patric v3.6.7 (Davis *et al.* 2020), the closest relative identified was *Vagococcus* sp. strain UBA11317 (Accession no. GCA_03535935.1). ANI was calculated comparing this genome to *V. fluvialis* bH8019 identifying an ANI value of 99.96, suggesting that *Vagococcus* sp. strain UBA11317 and *V. fluvialis* bH8019 may represent a new species of *Vagococcus*. For our subsequent analysis of the three new bovine strains presented here, we removed *V. fluvialis* bH819 from consideration.

V. fluvialis UFMG-H6, UFMG-H6B, and UFMG H7 genomes were compared to the two ANI-confirmed, publicly available *V. fluvialis* genomes using the tool Anvi'o v6.2 (Eren *et al.* 2015). The pangenome was determined using the anvi-pan-genome function, using the NCBI BLAST option, setting the default minbit threshold of 0.5 and an MCL inflation value of 10. Single copy number core genes were identified and aligned using Anvi'o. The core genes were used to reconstruct a phylogenetic tree using the FastTree v2.1.11 (Price *et al.* 2010) plug-in though Geneious Prime and visualized using iTOL v5 (Letunic and Bork 2016).

Data availability

All sequencing data are available in the NCBI Assembly database (www.ncbi.nlm.nih.gov/assembly) and short read archive (SRA) (www.ncbi.nlm.nih.gov/sra). The assemblies for *V. fluvialis* UFMG-H6, UFMG-H6B, and UFMG-H7 can be accessed with the following accession numbers: GCA_012102095.1, GCA_012102 505.1, and GCA_012102415.1, respectively. The raw reads can be accessed with the accession numbers SRR11455641, SRR11
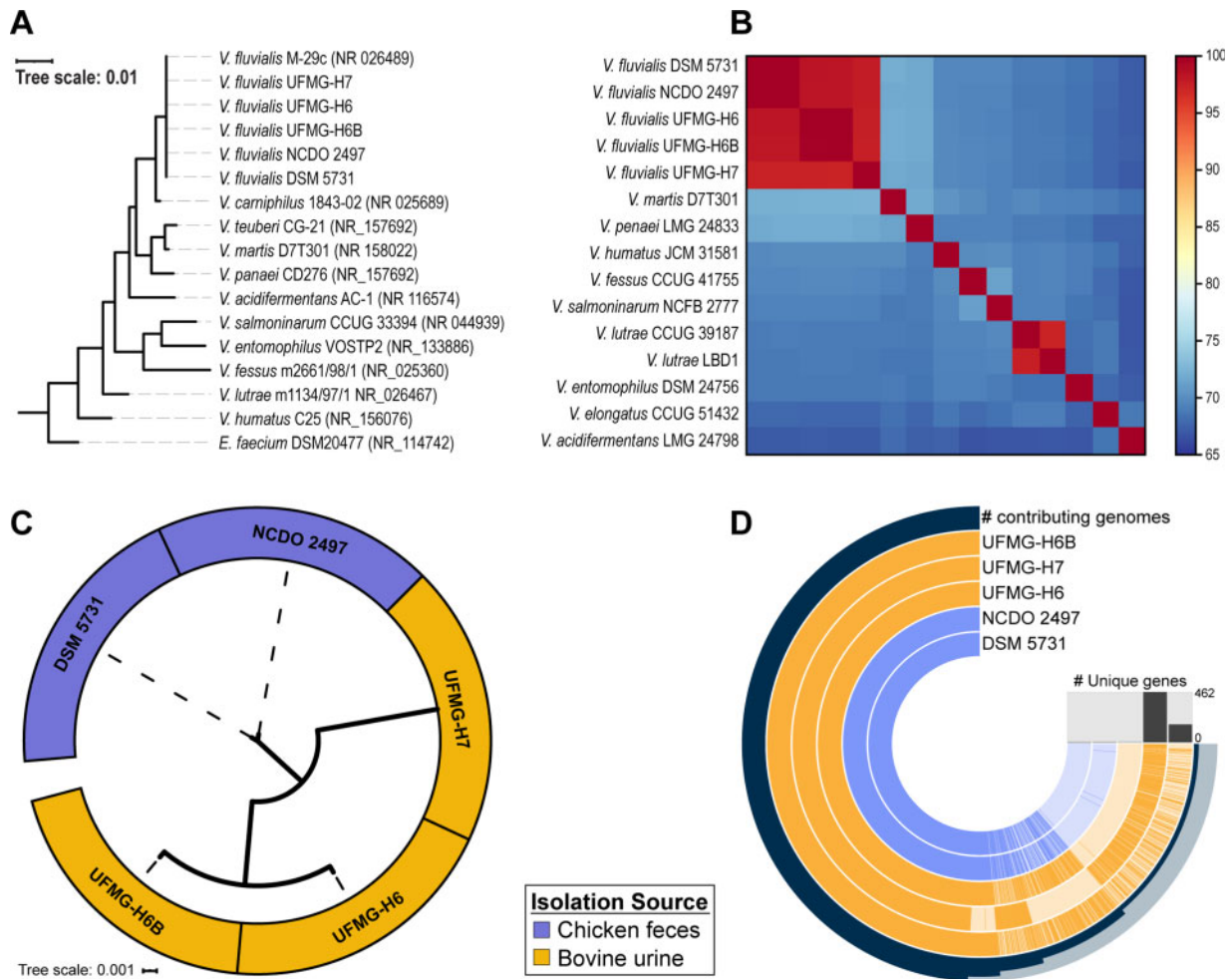
Figure 1 Comparison of five *V. fluvialis* draft genome sequences. (A) Phylogenetic tree of *Vagococcus* species based on the 16S rRNA sequence. (B) ANI comparison of *V. fluvialis* strains and other *Vagococcus* species. (C) Phylogenetic tree of *V. fluvialis* strains based on the core genome. (D) Pangenome of *V. fluvialis* strains. The three bovine urine isolates are shown in gold. Each gold and blue ring corresponds to a genome. Each ray corresponds with the presence (dark) or absence (light) of a given gene with the outer ring indicating the number of genomes that contain the particular gene. Genes were ordered according to this number. The histogram indicates the number of genes that are unique to each particular genome.

455850, and SRR11455640 for UFMG H-6, UFMG-H6B, and UFMG-H7, respectively.

Results and discussion

The genome statistics for the three *Vagococcus fluvialis* strains are shown in Table 1. None of the strains were found to contain a plasmid or to encode for antibiotic resistance genes. antiSMASH identified a putative bacteriocin encoded by *V. fluvialis* UFMG-H7. The presence of this putative bacteriocin may be related to the previously described probiotic potential of this specie (Román *et al.* 2012, 2013; Sorroza *et al.* 2012), although this hypothesis needs to be explored further.

PHASTER results indicate the presence of an intact prophage in UFMG-H7; this prophage is 50.3 kb long, encoding 58 known proteins including terminase, integrase, tail, capsid, and portal proteins. PHASTER indicates that this prophage most closely resembles the siphovirus Bacillus phage BCJA1 (Accession no. NC_006557), and thus is likely a member of the *Siphoviridae* family. So far, this is the first predicted prophage infectious of *V. fluvialis*. This prophage does not closely resemble any previously characterized phage or prophage. Using discontiguous blast, only

modest similarity was detected (<15% query coverage and <70% sequence identity).

Only one of the bovine *V. fluvialis* strains encodes for the CRISPR/Cas system; *V. fluvialis* UFMG-H6B has two spacer arrays and Cas type I-E genes. Cas type I-E was first described in *E. coli* genomes where the deletion of *Cas1* gene implied in a sensitivity to DNA damage [see review (Makarova and Koonin 2015)]. This is the first description of Cas type I-E in a *Vagococcus* strain; the Cas type II system (type A and type C) has been detected in other species from the genera (Pourcel *et al.* 2020). IslandViewer analyses did not show any genetic islands or integrative elements for any of our *V. fluvialis* strains.

Our ANI analysis identified two publicly available *V. fluvialis* draft genomes in GenBank; an additional genome labeled as *V. fluvialis* is in fact not a member of this species (see *Materials and methods*), and was thus excluded from our analysis. 16S rRNA gene sequence and ANI analyses confirmed the species designation of *V. fluvialis* UFMG-6, UFMG-6B, and UFMG-7 (Figure 1, A and B, respectively). The core genome of these two confirmed *V. fluvialis* genomes and our three isolates was identified. It includes 3298 single copy number genes. The phylogenetic tree derived from this core genome shows that *V. fluvialis* UFMG-H7 is distinct

from the other two bovine isolates, which clade together (Figure 1C). *V. fluvialis* UFMG-H6 and UFMG-H6B were isolated from the same animal, but they represent two different strains circulating within the urinary microbiota; their core genome has 92.73% sequence identity. Furthermore, the bovine urinary isolates clade separately from the two previously deposited genomes, both isolates from chicken feces.

Although built with only five genomes, *Vagococcus fluvialis* appears to have an open pangenome (Figure 1D). Further sequencing of isolates is needed to see if this holds true. 1143 genes are found within the accessory genome of the pangenome. The bovine urinary genomes encode for numerous genes not found within the two previously deposited genomes. Furthermore, UFMG-H7 includes 462 genes unique from the genomes of the two chicken fecal isolates and the two other bovine urine isolates. *V. fluvialis* has, however, been isolated from several other animals and environmental niches. Thus, sequencing isolates from these diverse sources is necessary to uncover the genic diversity of the species.

This is the first report showing *V. fluvialis* circulation in the urinary tract. The role of this bacteria in the urine microbiota is thus far unknown. However, it is important to highlight that, although previously isolated from lesions in cattle (Pot *et al.* 1994), none of the strains described here encode for genes associated with virulence factors, suggesting that these isolates do not have a pathogenic potential. Further *in vitro* evaluations of *V. fluvialis* UFMG-H6, UFMG-H6B, and UFMG-H7 would be necessary to understand if these strains have a probiotic potential, potentially related to the putative bacteriocin identified, and to look at their role in the urinary microbial community.

Literature cited

Al-Ahmad A, Pelz K, Schirrmeister JF, Hellwig E, Pukall R. 2008. Characterization of the first oral *Vagococcus* isolate from a root-filled tooth with periradicular lesions. Curr Microbiol. 57: 235–238. doi: 10.1007/s00284-008-9182-0.

Arndt D, Grant JR, Marcu A, Sajed T, Pon A, *et al.* 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44:W16–W21. doi: 10.1093/nar/gkw387.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al.* 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19:455–477. doi: 10.1089/cmb.2012.0021.

Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Fraser University Research Computing Group S, *et al.* 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 45:W30–W35. doi: 10.1093/nar/gkx343.

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, *et al.* 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 47:W81–W87. doi: 10.1093/nar/gkz310.

Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, *et al.* 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 75:3491–3500. doi:10.1093/jac/dkaa345.

Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, *et al.* 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 58:3895–3903. doi:10.1128/AAC.02412-14.

Collins MD, Ash C, Farrow JAE, Wallbanks S, Williams AM. 1989. 16S Ribosomal ribonucleic acid sequence analyses of *Lactococci* and related taxa. Description of *Vagococcus fluvialis* gen. nov., sp. nov. J Appl Bacteriol. 67:453–460. doi:10.1111/j.1365-2672.1989.tb02516.x.

Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, *et al.* 2018. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res. 46:W246–W251. doi: 10.1093/nar/gky425.

Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, *et al.* 2020. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. Nucleic Acids Res. 48:D606–D612. doi: 10.1093/nar/gkz943.

Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, *et al.* 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ. 3:e1319.doi: 10.7717/peerj.1319.

Fischer W, Arneth-Seifert D. 1998. d-Alanylcardiolipin, a major component of the unique lipid pattern of *Vagococcus fluvialis*. J Bacteriol. 180:2950–2957.

Jadhav KP, Pai PG. 2019. A rare infective endocarditis caused by *Vagococcus fluvialis*. J Cardiol Cases. 20:129–131. doi:10.1016/j.jccase.2019.07.001.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol Biol Evol. 30:772–780. doi:10.1093/molbev/mst010.

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. doi:10.1093/nar/gkw290.

Makarova KS, Koonin EV. 2015. Annotation and classification of CRISPR-Cas systems. Methods Mol Biol. 1311:47–75. doi: 10.1007/978-1-4939-2687-9_4.

Pot B, Devriese LA, Hommez J, Miry C, Vandemeulebroecke K, *et al.* 1994. Characterization and identification of *Vagococcus fluvialis* strains isolated from domestic animals. J Appl Bacteriol. 77: 362–369. doi:10.1111/j.1365-2672.1994.tb03436.x.

Pourcel C, Touchon M, Villeriot N, Vernadet J-P, Couvin D, *et al.* 2020. CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers. Nucleic Acids Res. 48:D535–D544. doi:10.1093/nar/gkz915.

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS ONE. 5: e9490.doi:10.1371/journal.pone.0009490.

Ramasamy D, Mishra AK, Lagier J-C, Padhmanabhan R, Rossi M, *et al.* 2014. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol. 64:384–391. doi:10.1099/ijs.0.057091-0.

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. doi:10.1093/bioinformatics/btv681.

Román L, Real F, Padilla D, El Aamri F, Déniz S, *et al.* 2013. Cytokine expression in head-kidney leucocytes of European sea bass (*Dicentrarchus labrax* L.) after incubation with the probiotic *Vagococcus fluvialis* L-21. Fish Shellfish Immunol. 35:1329–1332. doi:10.1016/j.fsi.2013.07.036.

Román L, Real F, Sorroza L, Padilla D, Acosta B, *et al.* 2012. The *in vitro* effect of probiotic *Vagococcus fluvialis* on the innate immune parameters of *Sparus aurata* and *Dicentrarchus labrax*. Fish Shellfish Immunol. 33:1071–1075. doi:10.1016/j.fsi.2012.06.028.

Sorroza L, Padilla D, Acosta F, Román L, Grasso V, *et al.* 2012. Characterization of the probiotic strain *Vagococcus fluvialis* in the protection of European sea bass (*Dicentrarchus labrax*) against vibriosis by *Vibrio anguillarum*. Vet Microbiol. 155:369–373. doi:10.1016/j.vetmic.2011.09.013.

Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, *et al.* 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44:6614–6624. doi:10.1093/nar/gkw569.

Teixeira LM, Carvalho MG, Merquior VL, Steigerwalt AG, Brenner DJ, *et al.* 1997. Phenotypic and genotypic characterization of *Vagococcus fluvialis*, including strains isolated from human sources. J Clin Microbiol. 35:2778–2781.

Vielva L, de Toro M, Lanza VF, de la Cruz F. 2017. PLACNETw: a web-based tool for plasmid reconstruction from bacterial genomes. Bioinformatics. 33:3796–3798. doi:10.1093/bioinformatics/btx462.

Communicating editor: D. J. de Koning