bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Identification of a chromosomally-encoded sucrose operon-like gene cluster**
2 **in *Vibrio parahaemolyticus* strain PH05 isolated from Negros Island, Philippines**
3
4
5 Czarina Anne E. De Mesa[a,c*], Remilyn M. Mendoza[a,c*], Edgar C. Amar[d], Leobert D. de la Peña[d]
6 and Cynthia P. Saloma[a,b#]
7
8
9 [a]Philippine Genome Center, University of the Philippines, A. Ma. Regidor Street, National
10 Science Complex, University of the Philippines Diliman, 1101, Quezon City, Metro Manila,
11 Philippines 1101
12
13 [b]National Institute of Molecular Biology and Biotechnology, University of the Philippines
14 Diliman, National Science Complex, College of Science University of the Philippines Diliman,
15 Quezon City, Metro Manila, Philippines 1101
16
17 [c]Science Education Institute, Department of Science and Technology, DOST Compound,
18 General Santos Avenue, Bicutan, Taguig City, Metro Manila Philippines 1631
19
20 [d]Southeast Asian Fisheries Development Center Aquaculture Department, Tigbauan, Iloilo
21 Philippines 5021
22
23
24
25 Contact information:
26
27 Czarina Anne E. De Mesa
28 cedemesa@up.edu.ph ORCID ID: https://orcid.org/0000-0003-3894-8617
29 Remilyn M. Mendoza
30 rmmendoza7@up.edu.ph ORCID ID: https://orcid.org/0000-0003-4937-118X
31 Edgar C. Amar
32 eamar@seafdec.org.ph
33 Leobert D. de la Peña
34 leobertd@seafdec.org.ph
35
36 #These authors contributed equally to this work
37
38
39
40 *Author for correspondence:
41 Cynthia P. Saloma Ph.D ORCID ID: https://orcid.org/0000-0002-3351-3465
42 email: cpsaloma@up.edu.ph / cpsaloma@gmail.com
43 telephone number: 898-18500 Loc 4703-04
44
45

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

46 **Identification of a chromosomally-encoded sucrose operon-like gene cluster**
47 **in *Vibrio parahaemolyticus* strain PH05 isolated from Negros Island, Philippines**
48
49
50 **ABSTRACT**
51

52 The ability of bacteria to metabolize a wide variety of carbon sources has been known to aid in
53 their ability for efficient colonization. *Vibrio parahaemolyticus*, a known aquatic pathogen has
54 been reported to have the ability to metabolize a number of carbohydrates including D-glucose,
55 D-galactose, L-arabinose, D-mannose, and D-ribose to name a few. Classical isolation of *V.*
56 *parahaemolyticus* from other members of the family *Vibrionaceae* relies on its carbon
57 utilization pattern. Conventionally, *V. parahaemolyticus* lacks the ability to utilize sucrose and
58 this has been the basis for its isolation using the Thiosulfate-citrate-bile salts-sucrose (TCBS)
59 agar. Reports of *V. parahaemolyticus* having the ability to utilize sucrose have been presented
60 yet there is paucity of information and detailed study on this phenotype. In this study, we
61 report the *V. parahaemolyticus* strain PH05 that has the ability to metabolize sucrose.
62 Phenotypic and genotypic characterization of this *V. parahaemolyticus* strain isolated from
63 Negros Island, Philippines, revealed that *V. parahaemolyticus* strain PH05 is atypical appearing
64 yellow on TCBS agar plates. It is capable of utilizing sucrose, unlike the majority of *V.*
65 *parahaemolyticus* isolates. Genome analyses of this strain revealed the presence of a
66 chromosomally encoded sucrose operon-like gene cluster encoded in chromosome 2 with the
67 following sucrose-utilization associated genes: scrY, ccpA, treP, scrK, and scrB genes coding for
68 sucrose porin, catabolite control protein A, PTS System sucrose-specific EIIBC component,
69 fructokinase, and sucrose-6-phosphate hydrolase. The mode of transmission of these genes to
70 *V. parahaemolyticus* strain PH05 is still unknown. However, the presence of insertion

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

71 sequences (IS) and phage elements in the same chromosome suggests horizontal gene transfer
72 events. Taken together, our results point to the possibility that acquired sucrose utilization
73 genes may contribute to the fitness of *V. parahaemolyticus* strain PH05 in the environment.

74

75 Keywords: *Vibrio parahaemolyticus*, atypical, sucrose, gene cluster, operon
76 Abbreviations: TCBS- Thiosulfate-citrate-bile salts-sucrose, NAN- Nutrient Agar with 1.5% NaCl,
77 NBN- nutrient broth with 1.5% NaCl
78

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

79 **INTRODUCTION**
80
81

82 *Vibrio parahaemolyticus* is a gram-negative halophilic bacterium usually associated with
83 sediments, shellfish, and zooplankton in marine and estuarine environments [5, 52]. Tran and
84 colleagues [59] identified a toxic strain of *V. parahaemolyticus* as the leading cause of acute
85 hepatopancreatic necrosis disease (AHPND) in shrimp. Moreover, clinical strains of *V.*
86 *parahaemolyticus* were identified to cause gastroenteritis accompanied by abdominal cramps
87 and watery diarrhea when raw or undercooked seafood contaminated with *V.*
88 *parahaemolyticus* is consumed [5, 16].

89 *Vibrio* species in general metabolize D-glucose anaerobically via the Embden-Meyerhof
90 pathway for energy production [36]. It has been established that the ability of bacteria to utilize
91 a wide variety of unique carbohydrates and use them efficiently, aid in the effective
92 colonization of bacteria in new niches [39]. In Freter's nutrient-niche hypothesis, it is proposed
93 that a bacterial species must be able to utilize a limiting nutrient better than another bacterial
94 species for successful colonization [15]. *Vibrio parahaemolyticus* which is known to have
95 adapted and evolved to colonize and establish niches in the intestines of humans and marine
96 organisms such as fish and shrimps, has been reported to utilize 71 different carbon sources.
97 Among these carbon sources are N-acetyl-D-glucosamine, mucus sugars such as L-arabinose, D-
98 galactose, D-gluconate, D-glucuronate, D-mannose, D-glucosamine, and D-ribose [21, 39, 60].
99 Although sucrose-utilizing *V. parahaemolyticus* strains have been reported [26, 42], no
100 genotypic data on this phenotype have been available.

101

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

102

103 The ability to metabolize sucrose is a variable trait in the family *Vibrionaceae*, and this
104 has been the basis of isolation using the Thiosulfate-citrate-bile salts-sucrose (TCBS) agar [1].
105 TCBS is a selective and differential culture medium used in the selective isolation and
106 cultivation of *Vibrio* species [36, 56]. Essential components of TCBS medium include yeast
107 extract and biological peptone, that serve as the source for nitrogen and amino acids; sodium
108 thiosulfate and sodium citrate, selective agents that render the medium alkaline, inhibiting the
109 growth of coliforms; Ox bile, which inhibits the growth of gram-positive bacteria; sucrose,
110 carbohydrate source; sodium chloride for optimum growth of *Vibrio* spp; ferric citrate, detects
111 production of hydrogen sulfide from sodium thiosulfate utilization; bromothymol blue, a pH
112 indicator; and agar as a solidifying agent. Conventionally, colonies of *Vibrio parahaemolyticus*
113 strains in TCBS agar appear blue to green and do not ferment sucrose [56]; hence no acid
114 production would entail changing TCBS agar from green to yellow. The US Food and Drug
115 Administration uses this biochemical reaction to identify strains of *V. parahaemolyticus* [23].

116 Sucrose, composed of one glucose unit and one fructose unit, is the most abundant
117 disaccharide in the environment due to its origin in higher plant tissues [40]. Metabolism of
118 sucrose involves the catalytic activities of sucrose-6-phosphate hydrolases and sucrose
119 phosphorylases.

120 This study reports an atypical *V. parahaemolyticus* strain PH05 isolated in Negros,
121 Philippines, under the Philippine Shrimp Pathogenomics Program (PSPP) of the Philippine
122 Genome Center. Bioinformatics analyses of its draft whole-genome sequence revealed the

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

123 presence of a sucrose operon-like gene cluster that has not been reported previously in *V.*

124 *parahaemolyticus* strains.

125

126

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

127 # MATERIALS AND METHODS

128 **Isolation of *V. parahaemolyticus* strains**

129 *Vibrio* isolates from the Philippine Shrimp Pathogenomics Program (PSPP) collection of
130 the Philippine Genome Center were isolated in 2014-2016. The hepatopancreas of five (5)
131 shrimp were dissected and placed in a tear-resistant sterile plastic bag for homogenization
132 using normal saline solution (NSS). Serial dilutions from 10^{-1} to 10^{-4} were spread plated in
133 Nutrient Agar (*Pronadisa*) with 1.5% NaCl (NAN), Thiosulfate Citrate Bile Salts Sucrose (TCBS)
134 agar (*Pronadisa*)**,** and *Vibrio* chromogenic agar (*Pronadisa*) in duplicates. The presence of
135 luminous bacteria was noted, and dominant colonies were isolated after 18-24h. Pure cultures
136 were ascertained before preparing 20% glycerol stock and storage at -80 ^{o}C biofreezer for
137 biobanking.

138 *Vibrio* spp. from PSPP isolates initially identified as *V. parahaemolyticus* were randomly
139 selected for revival. From glycerol stock, a 100 µl aliquot was used for inoculation in 5 ml
140 nutrient broth with 1.5% sodium chloride (NBN), and tubes were then incubated overnight at
141 37 ^{o}C. Turbid broths were then streaked on TCBS agar plates and incubated at 37 ^{o}C for the
142 selective growth of *Vibrio* species. Single colonies from these TCBS plates were streaked on
143 NAN for further Gram-staining procedure, colony morphology observation, and biochemical
144 characterization.

145

146 **Genome Assembly and Annotation**

147 The Philippine Shrimp Pathogenomics Program of the Philippine Genome Center
148 previously sequenced more than 400 bacterial pathogens of shrimp as previously described

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

149 [35]. All analyses were done using default settings unless otherwise stated. The whole genome
150 sequence of *V. parahaemolyticus* strain PH05 was initially pre-processed using *Fastp v0.20.0* [8]
151 and assembled using *SPAdes genome assembler v3.11.1* [32]. Assembled contigs were assessed
152 using *QUAST v5.0.2* [30] and filtered using *SeqKit* [50] such that contigs with less than 1000
153 bases were omitted. Assembled and filtered contigs were then annotated using *Prokka v1.13.7*
154 [48], Rapid Annotation using Subsystem Technology (RAST) [4, 6, 34], and *eggNOgmapper* [18].
155 The presence of plasmids was queried through *PlasmidSpades* [32], *plasmidFinder* [9], and
156 *PLSDB* database [16]. A reference-based assembly of contigs was done using *chromosomer* with
157 a ratio threshold of 1 [55]. Identification of IS elements, transposons, and phage elements was
158 made by comparing the draft genome against the sequences found in *ISfinder* [51], *ICEfinder*
159 [28], and *PHASTER* [3].
160
161 **Confirmation of the identity of *V. parahaemolyticus* strain PH05**
162 Taxonomic affiliation as *Vibrio parahaemolyticus* was confirmed by Average Nucleotide
163 Identity (ANI) calculation implemented in *pyani* v0.2.9 [38] using default settings against the
164 representative genomes of *Vibrio parahaemolyticus*, *Vibrio harveyi*, *Vibrio campbellii* and *Vibrio*
165 *alginolyticus*. The *16s rRNA* gene of *V. parahaemolyticus* strain PH05 was also extracted from
166 the whole genome sequence and was subjected to BLASTn (wwww.blast.ncbi.nlm.nih.gov)
167 analysis.
168
169 **Gene Finding**

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

170 Genes related to sucrose utilization such as: sucrose porin, catabolite control proteins,
171 PTS System sucrose/threhalose-specific, fructokinase, sucrose-6-phosphate hydrolase, and
172 invertase [40] were searched from the annotated genomes of *V. parahaemolyticus* strain PH05,
173 as well in *V. parahaemolyticus* PH 1339 [35] and *V. parahaemolyticus* RIMD2210633 [59] which
174 served as the controls. The genomes of *V. parahaemolyticus* strain PH05 and *V.*
175 *parahaemolyticus* strain PH1339 were compared using Mauve [12]

176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

202
RESULTS AND DISCUSSION
203
204
205
206
V. parahaemolyticus strain PH05 showed an atypical phenotype in TCBS agar
207

208 *V. parahaemolyticus* is associated with food-borne diseases after seafood consumption
209 and recently, with acute hepatopancreatic necrosis disease (AHPND), a devastating infection of
210 cultured shrimp [27]. Isolation of *V. parahaemolyticus* commonly relies on the use of TCBS, a
211 medium selective against gram-positive bacteria and differentiates sucrose-utilizing *V. cholerae*
212 from non-sucrose utilizing *Vibrio* like *V. parahaemolyticus* [33]. However, the colonies of the
213 putative *V. parahaemolyticus* strain PH05 on TCBS agar were round with an entire margin and
214 notably colored yellow, indicating acid production from sucrose utilization (Figure 1). This
215 characteristic, while highly atypical in *V. parahaemolyticus,* had been previously reported [26,
216 42].

217
218
219
220
221
222
223
224
225
226

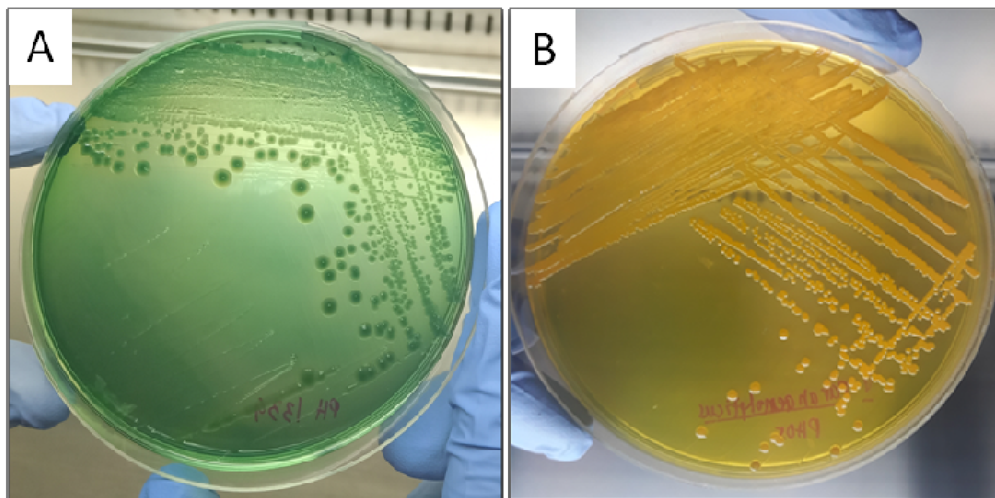


Figure 1. *Plate cultures (24-h) of V. parahaemolyticus strain PH1339 and the atypical V. parahaemolyticus strain PH05 in TCBS agar.* (A) *Vibrio parahaemolyticus* strain PH1339 shows the typical green to bluish-green colonies in TCBS agar. (B) Yellow colonies of putative *Vibrio parahaemolyticus* strain PH05 in TCBS agar indicative of sucrose utilization.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

227 The biochemical characteristics of putative *V. parahaemolyticus* strain PH05 initially
228 obtained from API 20NE analysis are presented in Table 1. Inputting this given biochemical
229 profile and the ability of the putative *V. parahaemolyticus* strain PH05 to utilize sucrose in the
230 ABIS online laboratory tool for bacterial identification database [11] resulted in the
231 identification of the putative *V. parahaemolyticus* to *Aeromonas media* (97.2%). However, since
232 this biochemical identification is not enough and not conclusive, we proceeded to the
233 verification of the identity of the putative *V. parahaemolyticus* strain PH05 through molecular
234 methods.

235

236 **Table 1** Biochemical characteristics of putative *V. parahaemolyticus* strain PH05

Reactions/Enzymes	
Reduction of Nitrates to nitrite and nitrite to nitrogen	+
Indole production	+
Glucose fermentation	+
Arginine dihydrolase	-
Urease	-
Hydrolysis (B-glucosidase) Esculin	+
Hydrolysis (protease) gelatin	+
B-galactosidase	+
Assimilation:	
Glucose	+
Arabinose	+
Mannose	+
Mannitol	+
N-acetyl glucosamine	-
D-maltose	+
Potassium gluconate	+
Capric acid	-
Adipic acid	-
Malic acid	+
Trisodium citrate	-
Phenylacetic acid	-
Oxidase	+

237
238

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

239 **Identification of of *Vibrio parahaemolyticus* strain PH05**

240 Raw reads of the whole genome sequence of the putative *V. parahaemolyticus* strain
241 PH05 were assembled and annotated. The draft genome assembly of the putative *Vibrio*
242 *parahaemolyticus* strain PH05 resulted in 26 contigs with a total length of 4954268 bp and G + C
243 content of 45.37%. The N50 of the assembly was 729047 bp. Annotation with *Prokka*
244 *v1.13.7* [48] revealed 4442 coding sequences (CDS), 110 tRNA, 43 miscellaneous RNA, and one
245 tmRNA.

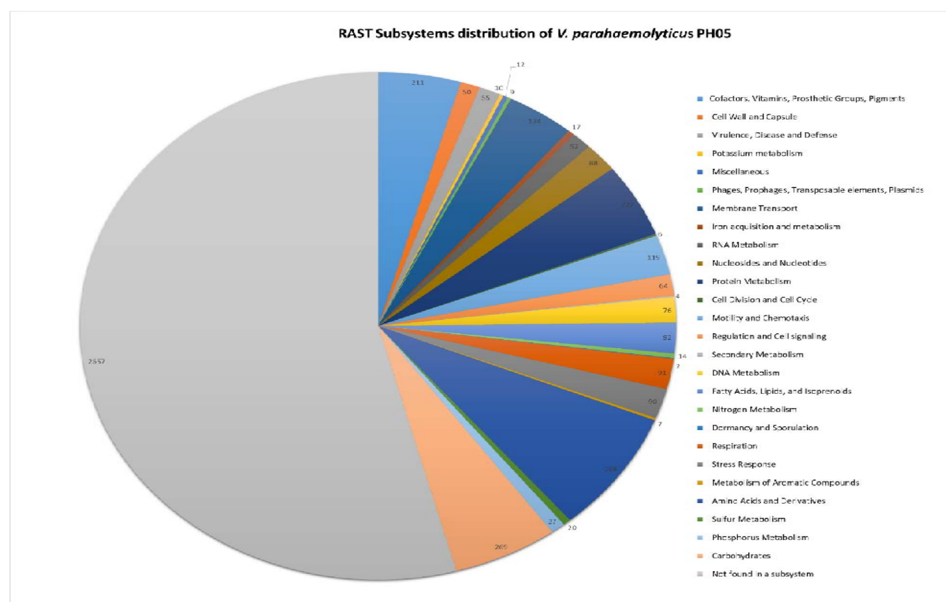
246 The python module *pyani* v0.2.9 [32] was used to calculate the ANI using the genome
247 sequence of *V. parahaemolyticus* strain PH05 against the representative genomes of *Vibrio*
248 *parahaemolyticus*, *Vibrio harveyi*, *Vibrio campbellii,* and *Vibrio alginolyticus.* Blast [8] and
249 MUMmer [29] were both used in the calculation of ANI similarity percentages. The ANIb
250 (98.43%) and ANIm (98.55%) percentage identities against the genome sequence of *V.*
251 *parahaemolyticus* RIMD2210633 confirmed the identity of *V. parahaemolyticus* strain PH05
252 (see Table 2 & Figure 2, supplementary material). Although there is a difference in values of
253 ANIb and ANIm, the 95-96% values of each method correspond to the 60-70% similarity cut-off
254 point for bacterial species delineation of DNA- DNA hybridization method [2, 24]. The result of
255 ANI calculations supports its initial identification using Blastn analysis of the partial sequence of
256 *16s rRNA* gene (See Table 1 Supplementary Material). However, with only ~60% of the *16s rRNA*
257 sequence recovered, its utility for accurate strain identification is limited [20]. Furthermore,
258 highly similar *Vibrio* species (such as *V. cholerae* and *V. mimicus*) cannot be differentiated with
259 their 16s rRNA gene sequences alone. [31]. The importance of using ANI in bacterial
260 identification is especially highlighted in this study because traditional protocols (TCBS culture,

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

261 Biochemical tests) we used together with *16s rRNA* gene were insufficient in identifying the
262 sucrose-fermenting *V. parahaemolyticus* strain PH05.
263
264 **Presence of a sucrose operon-like gene cluster in chromosome 2 of *Vibrio parahaemolyticus***
265 **strain PH05**
266
267 Annotation by RAST identified 4716 genes in V. *parahaemolyticus* PH05, to which 2557 (54%)
268 were not categorized to any subsystem and 2159 (46%) genes were classified under different
269 RAST subsystems (Figure 2). Subsystems are defined as functional roles that together
270 implement a specific biological process or structural complex [4, 6, 34]. Among the notable
271 subsystems, 55 genes were identified related to Virulence, Disease, and Defense; Phages,
272 Prophages, Transposable elements, and Plasmids subsystem; while in the carbohydrates
273 subsystem, 269 genes were identified including sucrose utilization genes.
274 To determine if the sucrose utilization gene is unique in *V. parahaemolyticus* strain
275 PH05, The annotated genomes of *V. parahaemolyticus* strain PH05, *V. parahaemolyticus* strain
276 PH1339, and *V. parahaemolyticus* RIMD2210633 were both searched for genes related to
277 sucrose metabolism such as sucrose-6-phosphate hydrolase, sucrose porin, PTS System-
278 sucrose/trehalose-specific, fructokinase, sucrose permease, and sucrose phosphorylase [39]. It
279 is noteworthy that the genomes of the *V. parahaemolyticus* strains scanned all possessed genes
280 encoding ATP-dependent fructokinase, catabolite control protein, and trehalose-specific PTS
281 system. In contrast, only *V. parahaemolyticus* strain PH05 genome contained genes coding for a
282 sucrose porin and sucrose-6-phosphate hydrolase (Table 2). Visualization in *Artemis* [43]
283 revealed that *V. parahaemolyticus* strain PH05 has a sucrose operon-like gene cluster (Figure 3).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

284 This gene cluster indicated the ability to utilize sucrose as a carbon source through PTS-
285 dependent sucrose transport and catabolism.



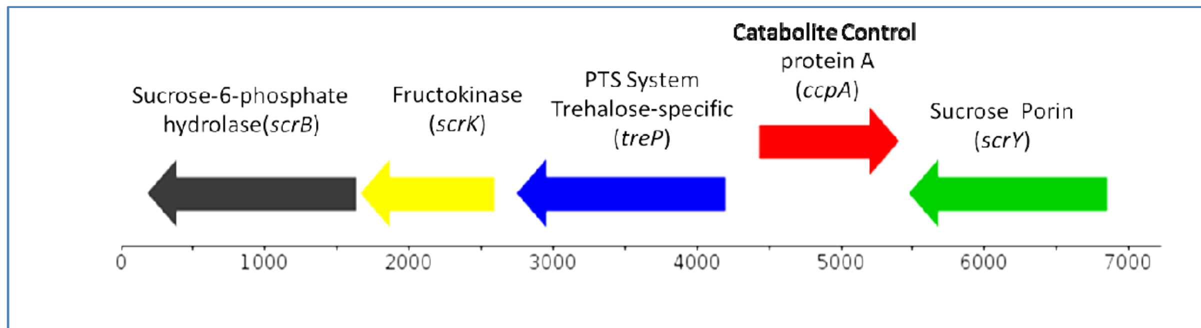
286 **Figure 2.** Distribution of subsystems in *V. parahaemolyticus* strain PH05 identified by RAST.
287 Genes involved in sucrose utilization were identified and classified under the carbohydrates
288 subsystems
289
290 **Table 2.** Sucrose-utilization-associated genes found in the annotated genome of *Vibrio*
291 *parahaemolyticus* strain PH05.

Sucrose-utilization associated genes	Annotation	Strain PH 05	Strain PH1339	Strain RIMD2210633
scrB	Sucrose-6-phosphate hydrolase	+	-	-
scrY	Sucrose Porin	+	-	-
scrK	Fructokinase	+	-	-
pfkA	ATP-dependent 6-phosphofructokinase isozyme 1	+	+	+
treP	PTS system trehalose-specific EIIBC component	+	-	-
treB	PTS system trehalose-specific EIIBC component	+	+	+
ccpA	Catabolite control Protein A	+	+	+
cra	Catabolite Repressor/ Activator	+	+	+

292

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

293



294

Figure 3. *Sucrose operon-like gene cluster found in V. parahaemolyticus strain PH05. V. parahaemolyticus* strain PH05 is atypical as it possesses a putative sucrose operon that consists of genes coding for sucrose-6-phosphate hydrolase, fructokinase, and PTS System which are the key players in a PTS-dependent sucrose transport and metabolism. A sucrose specific porin was also evidently found localized with the putative sucrose operon, as well as the catabolite control protein A that belongs to the LacI-GalR family of transcriptional regulators that has been reported to be often associated with sucrose-utilization genes. All the genes in the cluster are in the reverse stand except for the ccpA which is found on the forward strand.

295

296 In PTS-dependent sucrose transport and catabolism, external sucrose is transported and
297 phosphorylated by the PTS system's three-enzyme complex, giving off phosphorylated sucrose,
298 sucrose-6-phosphate, to the cell. This is hydrolyzed by sucrose-6-phosphate hydrolase, giving
299 off glucose-6-phosphate that can directly enter the glycolytic pathway; and fructose that is
300 phosphorylated by fructokinase to fructose-6-phosphate that also enters the glycolytic pathway
301 [40]

302 The amino acid sequence of the sucrose-6-phosphate hydrolase (*scrB*) of *V.*
303 *parahaemolyticus* strain PH05 contains the signature pentapeptide β-fructosidase motif
304 (NDPNG) (Supplementary material, Figure 2), which is found to be part of a large 14-amino acid
305 motif known as the sucrose-binding box wherein the aspartate residue act as the active site
306 [40,19]. Meanwhile, sucrose porin (*scrY*) is an example of a specific porin that act as channels

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

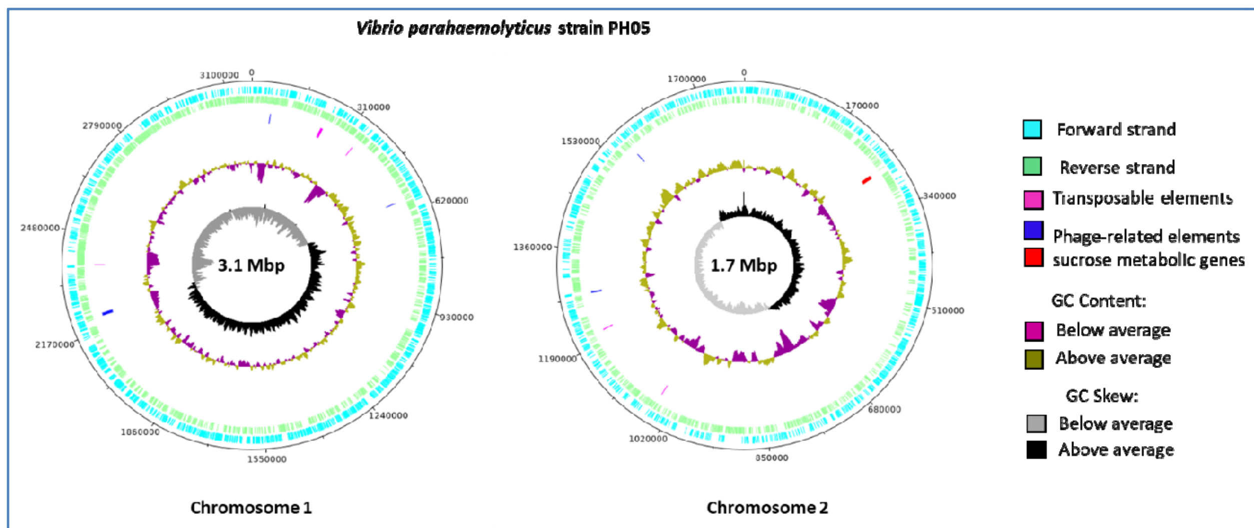
307 across the membrane [53]. Passage of sucrose through the sucrose porin has been
308 hypothesized to initially begin with sucrose molecules diffusing from the external solution to a
309 "trapping zone" in the sucrose porin. Then sucrose molecules slide along a greasy slide; enter
310 and pass the binding region to enter the cell [14, 53].

311 Blastx analysis of the *treP* and *ccpA* genes revealed that the *treP* gene has 99.79%
312 identity to the PTS system sucrose-specific IIB component (Glc family) /PTS system sucrose-
313 specific IIC component (Glc family) of *V. parahaemolyticus* in the database; while the *ccpA* gene
314 codes for a LacI family transcriptional regulator, catabolite control protein A, that is 98.48%
315 identical to that of *V. parahaemolyticus* in NCBI database. LacI-GalR family of transcriptional
316 regulators has been mentioned to be often associated with sucrose-utilization genes [40],
317 inducing or repressing the structural genes in a substrate-specific manner for better allocation
318 or utilization of energy by the cell.

319 It was initially hypothesized that the sucrose gene cluster found in *Vibrio*
320 *parahaemolyticus* strain PH05 was extra-chromosomally encoded since there were reports of
321 conjugative plasmid pUR400 found in *Salmonella* and *E. coli* [61, 44, 45] capable of transfer of
322 sucrose utilization genes. To confirm this, putative plasmid sequences were assembled using
323 *PlasmidSpades* [32], and plasmid-related elements were searched in *plasmidFinder* [9] and the
324 *PLSDB* database [16]. However, no plasmids were detected in the draft genome. It was
325 therefore highly probable that the sucrose genes were chromosomally encoded. The draft
326 genome was then scaffolded into two chromosomes using the tool *chromosomer* [55], which
327 uses genome similarities with a reference genome. The reference genome *V. parahaemolyticus*
328 RIMD2210633 was first used as a reference, but only the first chromosome was assembled (~3

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

329 Mb). Unexpectedly, we found a complete genome of *V. parahaemolyticus* strain HA2, which
330 possessed chromosomally encoded scr genes since it appeared in the result of the blastn and
331 blastx analyses of the sucrose porin and sucrose-6-phosphate hydrolase. The result of assembly
332 using *V. parahaemolyticus* strain HA2 as reference resulted in two chromosomes (Figure 4)
333 where the scr genes were found to be encoded in the smaller chromosome 2 (~1.7 Mb). The
334 assembly of the draft genome into the two chromosomes points to the high similarity between
335 the genome of *V. parahaemolyticus* strain HA2 and *V. parahaemolyticus* strain PH05, which are
336 both aquaculture isolates.



337
338 **Figure 4.** *Genome organization of V. parahaemolyticus strain PH05*. The two (2)
339 chromosomes of *V. parahaemolyticus* strain PH05 (mapped against *V. parahaemolyticus*
340 HA2) showing the genome composition and the location of the sucrose operon like gene
341 cluster in the chromosome 2, as well as relevant elements for HGT. The figure was generated using DNA plotter of the visualization tool Artemis [43]

342

343 We then compared, the segment of the chromosome of *V. parahaemolyticus* strain
344 PH05 containing the sucrose-operon-like cluster to that of *V. parahaemolyticus* HA2 and *V.*
345 *parahaemolyticus* strain PH1339. Based on genome comparison (Figure 5), it appeared likely
346 that there is an insertion of sucrose metabolic genes in the chromosome of *V.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

347 *parahaemolyticus* strain PH05 as well as in *V. parahaemolyticus* strain HA2 that is not found in
348 *V. parahaemolyticus* strain PH1339. This is supported by the presence of genes flanking the
349 sucrose operon-like gene cluster in the chromosomes of both *V. parahaemolyticus* strains PH05
350 and HA2. The apparent shift of the genes and absence of the *yyaP* gene in the chromosome
351 segment containing the sucrose operon-like gene cluster in *V. parahaemolyticus* strain PH05
352 was also evident. *V. parahaemolyticus* strain PH1339, similar to strain PH05, was isolated in
353 *Penaeus vannamei*.

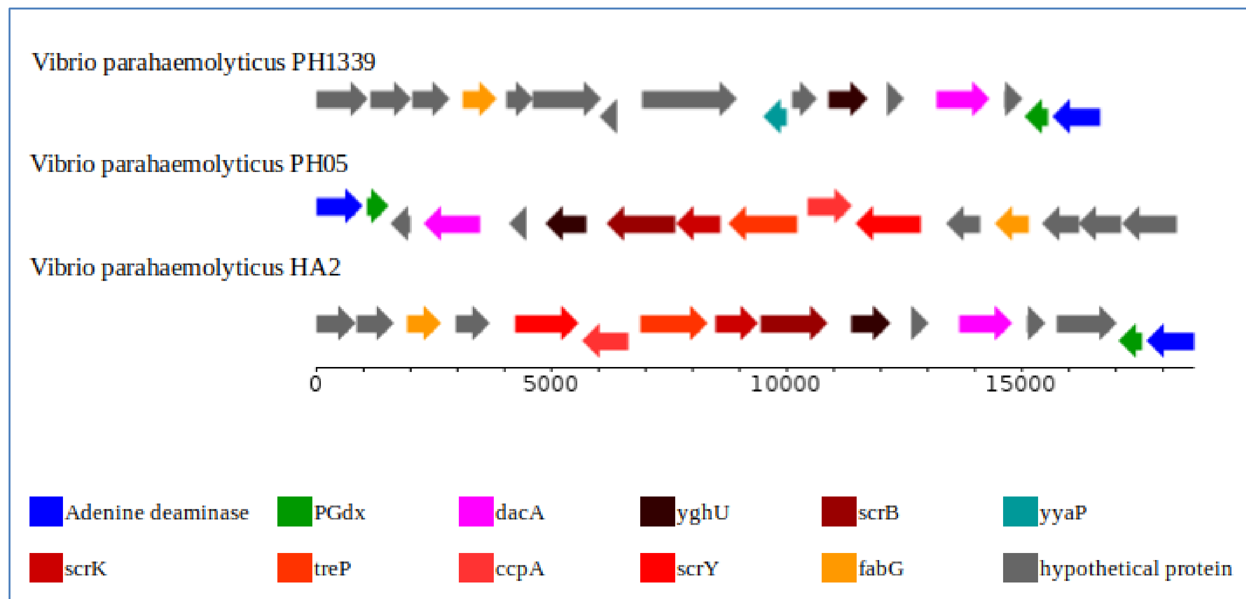


Figure 5. *Linear representation of the Chromosome 2 segment of V. parahaemolyticus strain PH05 harboring the sucrose operon-like gene cluster.* Highly similar gene architecture of *V. parahaemolyticus* strain PH05 was mapped with that of *V. parahaemolyticus* strain HA2 (Accession: SAMN07680340). Evident insertion of the sucrose metabolic gene cluster in *V. parahaemolyticus* strain PH05 was also observed when compared with *V. parahaemolyticus* strain PH1339 (Accession: SRR9937595). Genome comparison (supplementary material Figure 3 and 4) was visualized using Mauve [12].

354

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

355 To further investigate the mechanism of insertion of sucrose-utilization genes in
356 chromosome 2 of *V. parahaemolyticus* strain PH05, we looked at several databases and
357 annotation tools. *Prokka* [48] was able to annotate a few transposons that were all found in
358 chromosome 1. Amino acids from the genome were functionally annotated using eggNog
359 mapper [18] and revealed some hypothetical proteins in chromosome 2 as transposases.
360 ISfinder [51] was used to identify insertion sequences usually associated with transposons.
361 Phage elements were also identified through *Prokka* [48] and *PHASTER* [3]. However, these
362 were found to be localized far from the scr gene cluster. Taken together, the scr gene cluster
363 was probably transferred through a natural transformation event since *Vibrio* species are
364 known to be naturally competent [54]. For instance, Chitin-induced competence in *V. cholerae*
365 allowed it to acquire gene clusters that altered its serogroup [7].

366 On the other hand, foreign DNA can be integrated into the chromosome by illegitimate
367 recombination in which even short homologous regions facilitate the integration of a foreign
368 DNA [14]. In the mauve comparisons, there were homology sites with the *V. parahaemolyticus*
369 strain HA2 with the sequences encoding the scr gene cluster of *V. parahaemolyticus* strain PH05
370 (Figure 3 in the supplementary). These sites were probably used to facilitate the integration of
371 the scr gene clusters during its period of competence.

372 The ability to metabolize a wide array of carbon sources confers an adaptive advantage
373 to the organism in an environment where glucose is limiting. Sucrose, a disaccharide consisting
374 of one glucose and one fructose unit, is the most abundant disaccharide owing to its origin in
375 higher plant tissues [40, 49], and not all microorganisms possess the machinery to metabolize
376 this carbohydrate.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

377 One way by which *V. parahaemolyticus* is differentiated from other *Vibrio* species such
378 as *V. cholerae* and *V. alginolyticus* is through their difference in colony morphology in TCBS agar
379 plates. It has been established that *V. cholerae* and *V. alginolyticus* are sucrose fermenters,
380 hence, the yellow colonies that they produce in TCBS agar plates, while *V. parahaemolyticus* is
381 known to be unable to ferment sucrose and exhibits green colonies in TCBS agar. However,
382 previous studies have reported yellow *V. parahaemolyticus* isolates from oysters and mussels
383 [42] and from the diarrheal stool of a patient [26]. In the present study, we report an atypical *V.*
384 *parahaemolyticus* strain PH05 where a sucrose operon-like gene cluster was found encoded in
385 its chromosome 2.

386 *Vibrio parahaemolyticus* strain PH05 was isolated from *Penaeus vannamei* from a
387 shrimp pond in Negros Island, Philippines. Its ability to utilize sucrose may be a form of
388 metabolic adaptation to thrive in that specific environment. In some aquaculture ponds,
389 molasses (which is high in sucrose) was used as an organic fertilizer [47] and as a growth
390 promoter of heterotrophic bacteria to control the accumulation of nitrogenous wastes [46]. In
391 the study of Tendencia, Bosma, & Sorio [57], molasses was used to promote the growth of
392 "yellow *Vibrios*" to act against pathogenic *Vibrio,* rendering shrimp less susceptible to
393 Whitespot Syndrome Virus (WSSV). Antagonistic action of these yellow *Vibrio* against
394 pathogenic green *Vibrio* species such as *V. parahaemolyticus* and *V. harveyi* can be attributed to
395 bacteriocin production [58]. However, the ability to metabolize sucrose in an environment with
396 adequate sucrose concentrations may have also contributed to their ability to compete with
397 non-sucrose fermenting *Vibrio* species. The ability of *V. parahaemolyticus* strain PH05 to utilize

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

398 sucrose may add to its overall fitness given that in its environment, sucrose is a limiting nutrient
399 [39].

400 In *Streptococcus pneumoniae*, a gram-positive bacterial pathogen, two sucrose-
401 metabolizing systems, the *sus* and *scr* systems, were reported and described by Iyer and Camilli
402 [19] to play niche-specific roles in virulence. Sucrose is a major carbon source used by *S.*
403 *pneumoniae* during colonization and infection. Their study found that the uptake and
404 catabolism of sucrose have essential roles in colonization of murine nasopharynx to which the
405 *scr* system is primarily involved, and an accessory role in the infection of the lungs with the *sus*
406 system as a key player. Moreover, in another study by Shelburne and colleagues [49], the
407 catabolite control protein A (*ccpA*) ortholog in Group A *Streptococcus* (GAS) was found to
408 influence the transcript levels of many carbohydrate utilization genes and several GAS virulence
409 factors, including the cytolysin streptolysin S, and speB (cysteine protease). Their data
410 demonstrated that the ccpA of GAS regulates virulence factors' production depending on the
411 nutrient availability in the environment. In a nutrient-rich environment, the production of
412 streptolysin S is repressed, while in a nutrient-limiting environment, the production of SpeB is
413 increased. The interconnection between carbohydrate metabolism and pathogenesis in the
414 gram-negative *V. parahaemolyticus* stain PH05 is an exciting aspect that can be explored and
415 studied in the future.

416 **CONCLUSION**

417

418 *Vibrio parahaemolyticus* strain PH05 isolated from Negros Island, Philippines in the
419 summer of 2014, exhibits an atypical phenotype as it possesses the ability to utilize sucrose as

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

420 demonstrated by its growth in TCBS. This phenotype was supported by the presence of genes
421 associated with sucrose metabolism such as sucrose porin, sucrose-6-phosphate, fructokinase,
422 PTS system sucrose-specific IIC component (Glc family), and catabolite control protein A. These
423 genes are arranged as a sucrose operon-like gene cluster found in chromosome 2 (~1.7 Mb).
424 The presence of IS sequences, transposases, and phage elements suggests horizontal gene
425 transfer events in the chromosome with the scr gene cluster's acquisition a likely metabolic
426 adaptation to the environment possibly linked to its virulence and pathogenicity.
427
428
429
430
431

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

432 **CREDIT AUTHOR STATEMENT**

433 **De Mesa, CA** and **Mendoza, RM** were involved in Conceptualization, Methodology,
434 Investigation, Writing (original draft), and Visualization; **Amar, EA** and **De la Peña, LD** were
435 involved in the Conceptualization, Collection of Specimen, Biochemical Characterization and
436 Funding acquisition; **Saloma, CP** was involved in Conceptualization, Writing (Review and
437 Editing), Resources, Supervision and Funding acquisition.

438

439 **DECLARATION OF COMPETING INTEREST**

440 The authors declare no conflict of interest.

441

442 **ACKNOWLEDGMENT**

443 The authors would like to acknowledge the contributions of Ms. Sarah Penir and the
444 people who previously worked under the Philippine Shrimp Pathogenomics Program, the Core
445 Facility for Bioinformatics of Philippine Genome Center and the Negros Prawn Producers
446 Cooperative for their generosity and assistance.

447

448 **FUNDING**

449 This work was supported by a Department of Science and Technology, Philippine Council
450 for Agriculture, Aquatic and Natural Resources Research and Development (DOST-PCAARRD)
451 grant to CPS, EAA and LDD and a DOST Science Education Institute fellowship to CADM and
452 RMM.

453

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

454 **NUCLEOTIDE SEQUENCE DATA**

455 The Illumina reads of *V. parahaemolyticus* PH05 have been deposited under the accession no.
456 SRR13286682, and the genome assembly has been deposited at DDBJ/EMBL/Genbank under
457 the accession no. JAELVO000000000.

458 **LITERATURE CITED**

459
460 [1] Abushattal S., Vences A., Barca A.V., Osorio C.R. (2020) Diverse Horizontally-Acquired
461 Gene Clusters Confer Sucrose Utilization to Different Lineages of the Marine Pathogen
462 *Photobacterium damselae* subsp. damselae. *Genes*.11(11):1244.
463 https://doi.org/10.3390/genes11111244
464
465 [2] Arahal, D. R. (2014). Whole-genome analyses: Average nucleotide identity. In Methods
466 in Microbiology (Vol. 41, pp. 103–122). https://doi.org/10.1016/bs.mim.2014.07.002
467
468 [3] Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., and Wishart, D. S. (2016).
469 PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res,
470 44(W1), W16-21. https://doi.org/10.1093/nar/gkw387
471
472 [4] Aziz R.K., Bartels D., Best A.A., DeJongh M., Disz T., Edwards R.A., Formsma K., Gerdes S.,
473 Glass E.M., Kubal M., Meyer F., Olsen G.J., Olson R., Osterman A.L., Overbeek R.A.,
474 McNeil L.K., Paarmann D., Paczian T., Parrello B., Pusch G.D., Reich C., Stevens R.,
475 Vassieva O., Vonstein V., Wilke A., Zagnitko O. 2008. The RAST Server: Rapid
476 Annotations using Subsystems Technology. *BMC genomics.*
477 https://doi.org/10.1186/1471-2164-9-75
478
479 [5] Barbieri, E., Falzano, L., Fiorentini, C., Pianetti, A., Baffone, W., Fabbri, A., … Donelli, G.
480 (1999). Occurrence, diversity, and pathogenicity of halophilic Vibrio spp. and non-O1
481 *Vibrio cholerae* from estuarine waters along the Italian Adriatic coast. Appl. Environ.
482 Microbiol, 65(6), 2748–2753. https://doi.org/10.1128/aem.65.6.2748-2753.1999
483
484 [6] Brettin T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., Olson, R., Overbeek,
485 R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A., Stevens, R., Vonstein, V.,
486 Wattam, A.R., Xia, F. 2015. RASTtk: A modular and extensible implementation of the
487 RAST algorithm for building custom annotation pipelines and annotation batches of
488 genomes. *Sci Rep*. https://doi.org/10.1038/srep08365
489

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

490 [7] Blokesch, M., and Schoolnik, G. K. (2007). Serogroup conversion of *Vibrio cholerae* in
491 aquatic reservoirs. PLoS Pathog., 3(6), 0733–0742.
492 https://doi.org/10.1371/journal.ppat.0030081
493
494 [8] Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden,
495 T. L. (2009). BLAST+: architecture and applications. *BMC bioinformatics*, 10, 421.
496 https://doi.org/10.1186/1471-2105-10-421
497
498
499 [9] Carattoli, A., Zankari, E., Garciá-Fernández, A., Larsen, M. V., Lund, O., Villa, L., …
500 Hasman, H. (2014). In Silico detection and typing of plasmids using plasmidfinder and
501 plasmid multilocus sequence typing. Antimicrob. Agents Chemother., 58(7), 3895–3903.
502 https://doi.org/10.1128/AAC.02412-14
503
504 [10] Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ
505 preprocessor. Bioinformatics, 34(17), i884–i890.
506 https://doi.org/10.1093/bioinformatics/bty560
507
508 [11] Costin S., and S. Ionut. (2017). ABIS online – Advanced bacterial identification software,
509 an original tool for phenotypic bacterial identification. Accessed at:
510 http://www.tgw1916.net/bacteria_abis.html
511
512 [12] Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: Multiple
513 alignments of conserved genomic sequence with rearrangements. Genome Res, 14(7),
514 1394–1403. https://doi.org/10.1101/gr.2289704
515
516 [13] De Vries, J., and Wackernagel, W. (2002). Integration of foreign DNA during natural
517 transformation of Acinetobacter sp. by homology-facilitated illegitimate recombination.
518 P Natl Acad Sci USA, 99(4), 2094–2099. https://doi.org/10.1073/pnas.042263399
519
520 [14] Forst, D., Welte, W., Wacker, T., and Diederichs, K. (1998). Structure of the sucrose-
521 specific porin ScrY from Salmonella typhimurium and its complex with sucrose. Nat.
522 Struct. Mol. Biol. 5, 37–46. doi:10.1038/nsb0198-37
523
524 [15] Freter R. (1988) Mechanisms of bacterial colonization of the mucosal surfaces of the
525 gut. Virulence mechanisms of bacterial pathogens. Washington DC: *American Society of*
526 *Microbiology*; p. 45–60.
527
528 [16] Galata, V., Fehlmann, T., Backes, C., and Keller, A. (2019). PLSDB: a resource of complete
529 bacterial plasmids. Nucleic Acids Res, 47(D1), D195–D202.
530 https://doi.org/10.1093/nar/gky1050
531

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

532 [17] Honda, T., Iida, T., Akeda, Y., and Kodama, T. (2008). Sixty years of *Vibrio*
533 *parahaemolyticus* research. Microbe, 3, pp. 462–466.
534 https://doi.org/10.1128/microbe.3.462.1
535
536 [18] Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., … Bork,
537 P. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional
538 annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res, 44.
539 https://doi.org/10.1093/nar/gkv1248
540
541 [19] Iyer R. and A. Camilli. (2007). Sucrose metabolism contributes to in vivo fitness of
542 Streptococcus pneumonia. Mol Microbiol. 66(1): 1-13.
543
544 [20] Jenkins, C., Ling, C. L., Ciesielczuk, H. L., Lockwood, J., Hopkins, S., McHugh, T. D., …
545 Kibbler, C. C. (2012). Detection and identification of bacteria in clinical samples by 16S
546 rRNA gene sequencing: Comparison of two different approaches in clinical practice. J.
547 Med. Microbiol., 61(4), 483–488. https://doi.org/10.1099/jmm.0.030387-0
548
549 [21] Kalburge S.S., Carpenter M.R., Rozovsky S., Boyd E.F. (2017) Quorum sensing regulators
550 are required for metabolic fitness in *Vibrio parahaemolyticus*. *Infect Immun*.85(3)
551
552 [22] Kalia V.C., P. Kumar, R. Kumar, A. Mishra, and S. Koul. (2015). Genome-wide analysis for
553 rapid identification of Vibrio species. Indian J. Microbiol. 55(4):375–383. doi
554 https://dx.doi.org/10.1007%2Fs12088-015-0553-5
555
556 [23] Kaysner C.A. and A. De Paola Jr. 2004. Bacteriological Analytical Manual. US Food and
557 Drugs Administration. Accessed at: https://www.fda.gov/food/laboratory-methods-
558 food/bam-vibrio
559
560 [24] Kim, M., Oh, H. S., Park, S. C., and Chun, J. (2014). Towards a taxonomic coherence
561 between average nucleotide identity and 16S rRNA gene sequence similarity for species
562 demarcation of prokaryotes. Int J Syst Evol Microbiol, 64(PART 2), 346–351.
563 https://doi.org/10.1099/ijs.0.059774-0
564
565 [25] Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular
566 Evolutionary Genetics Analysis across Computing Platforms. Mol. Biol Evol. 35(6), 1547–
567 1549. https://doi.org/10.1093/molbev/msy096
568
569 [26] Lam S. (1985). A mucoid, sucrose-positive *Vibrio parahaemolyticus* strain isolated from
570 diarrhoeal stool. Singapore Med J. 26(6):479-81
571
572 [27] Letchumanan, V., Chan, K. G., and Lee, L. H. (2014). *Vibrio parahaemolyticus*: A review
573 on the pathogenesis, prevalence, and advance molecular identification techniques.
574 Front Microbiol, 5. https://doi.org/10.3389/fmicb.2014.00705
575

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

576 [28] Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., … Ou, H.-Y. (2019). ICEberg 2.0: an updated
577 database of bacterial integrative and conjugative elements. Nucleic Acids Res, 47.
578 https://doi.org/10.1093/nar/gky1123
579
580 [29] Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., & Zimin, A. (2018).
581 MUMmer4: A fast and versatile genome alignment system. PLoS computational biology,
582 14(1), e1005944. https://doi.org/10.1371/journal.pcbi.1005944
583
584 [30] Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., and Gurevich, A. (2018). Versatile
585 genome assembly evaluation with QUAST-LG. Bioinformatics, 34(13), i142–i150.
586 https://doi.org/10.1093/bioinformatics/bty266
587
588 [31] Nhung, P. H., Shah, M. M., Ohkusu, K., Noda, M., Hata, H., Sun, X. S., … Ezaki, T. (2007).
589 The dnaJ gene as a novel phylogenetic marker for identification of Vibrio species. Syst
590 Appl Microbiol, 30(4), 309–315. https://doi.org/10.1016/j.syapm.2006.11.004
591
592 [32] Nurk S. A. Bankevich, D. Antipov, A. Gurevich, A. Korobeynikov, A. Lapidus, A. Prjibelsky,
593 A. Pyshkin, A. Sirotkin, Y. Sirotkin, R. Stepanauskas, J. McLean, R. Lasken, S.R.
594 Clingenpeel, T. Woyke, G. Tesler, M.A. Alekseyev, P.A. Pevzner (2013) Assembling
595 Genomes and Mini-metagenomes from Highly Chimeric Reads. In: Deng M., Jiang R., Sun
596 F., Zhang X. (eds) Research in Computational Molecular Biology. RECOMB 2013. Lecture
597 Notes in Computer Science, vol 7821. Springer, Berlin, Heidelberg
598
599 [33] Oliver, J. D. (2003). Chapter 17 Culture media for the isolation and enumeration of
600 pathogenic Vibrio species in foods and environmental samples. In Progress in Industrial
601 Microbiology, 37(C), 249–269. https://doi.org/10.1016/S0079-6352(03)80020-6
602
603 [34] Overbeek, R., Olson R., Pusch G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes,
604 S., Parrello, B., Shukla, M., Vonstein, V, Wattam, A.R., Xia, F., Stevens, R. 2014. The SEED
605 and The Rapid Annotation of microbial genomes using Subsystems Technology (RAST).
606 *Nucleic Acid Res*. https://doi.org/10.1093/nar/gkt1226
607
608 [35] Penir, S., dela Peña, L., Cabillon, N., Bilbao, A., Amar, E., and Saloma, C. 2019. Draft
609 Genome Sequence of *Vibrio parahaemolyticus* Strain PH1339, Which Causes Acute
610 Hepatopancreatic Necrosis Disease in Shrimp in the Philippines. Microbiol. Resour.
611 Announc.. 8 (46) e01020-19; DOI: 10.1128/MRA.01020-19
612
613 [36] Percival, S.L. & Williams, D.W. 2014. Microbiology of Waterborne Diseases. 2nd ed.
614 Academic Press. Pp 237-248. https://doi.org/10.1016/B978-0-12-415846-7.00012-3
615
616 [37] Pfeffer C. and J.D. Oliver. 2003. A comparison of thiosulphate-citrate-bile salts-sucrose
617 (TCBS) agar and thiosulphate-chloride-iodide (TCI) agar for the isolation of Vibrio species

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

618 from estuarine environments. Lett Appl Microbiol. 36 (3). doi.org/10.1046/j.1472-
619 765X.2003.01280.x
620
621 [38] Pritchard, L, R.H. Glover, S. Humphris, J.G. Elphinstone, and I.K. Toth. 2016. Genomics
622 and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant
623 pathogens. Anal. Methods. 8, 12-24.
624
625 [39] Regmi, A., and Boyd, E. F. (2019). Carbohydrate metabolic systems present on genomic
626 islands are lost and gained in *Vibrio parahaemolyticus*. BMC Microbiol, 19(1).
627 https://doi.org/10.1186/s12866-019-1487-6
628
629
630 [40] Reid S. J. and V.R. Abratt. 2005. Sucrose utilization in bacteria: genetic organization and
631 regulation. Appl Microbiol Biotechnol. 67, 312-321. DOI 10.1007/s00253-004-1885-y
632
633 [41] Ritsema, T., Hernández, L., Verhaar, A., Altenbach, D., Boller, T., Wiemken, A., and
634 Smeekens, S. (2006). Developing fructan-synthesizing capability in a plant invertase via
635 mutations in the sucrose-binding box. Plant J, 48(2), 228–237.
636 https://doi.org/10.1111/j.1365-313X.2006.02862.x
637
638 [42] Rojas, M. V. R., Matté, M. H., Dropa, M., da Silva, M. L., and Matté, G. R. (2011).
639 Caracterização de *Vibrio parahaemolyticus* isolados de ostras e mexilhões em São Paulo,
640 Brasil. Rev Inst Med Trop S Paulo, 53(4), 201–205. https://doi.org/10.1590/S0036-
641 46652011000400005
642
643 [43] Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., and Barrell,
644 B. (2000). Artemis: Sequence visualization and annotation. Bioinformatics, 16(10), 944–
645 945. https://doi.org/10.1093/bioinformatics/16.10.944
646
647 [44] Schmid, K., M. Schupfner, and R. Schmitt. (1982). Plasmid-mediated uptake and
648 metabolism of sucrose by Escherichia coli K-12. J. Bacteriol. 151:68-76.
649
650 [45] Schmid, K., R. Ebner, J. Altenbuchner, R. Schmitt, and J. W.Lengeler. (1988). Plasmid-
651 mediated sucrose metabolism in Escherichia coli K12: mapping of the scr genes of
652 pUR400. Mol. Microbiol. 2:1-8.
653
654 [46] Schneider, O., Sereti, V., Eding, E. H., and Verreth, J. A. J. (2006). Molasses as C source
655 for heterotrophic bacteria production on solid fish waste.
656 https://doi.org/10.1016/j.aquaculture.2006.08.053
657
658 [47] SEAFDEC Institute of Aquaculture. (1980). Fish Farming HANDBOOK: A handy reference
659 for farm broadcasters, information writers extension workers, fish farmers teachers, and
660 students. Retrieved from http://hdl.handle.net/10862/1490
661

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

662 [48] Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. Bioinformatics,
663 30(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153
664
665 [49] Shelburne S.A III., D. Keith, N. Horstmann, P. Sumby, M.T. Davenport, E.A.Graviss, R.G.
666 Brennan, and J.M. Musser. (2008). A direct link between carbohydrate utilization and
667 virulence in the major human pathogen group A Streptococcus. P. Natl. Acad. Sci. USA.
668 5(5), 1698-1703.
669
670 [50] Shen W, Le S, Li Y, Hu F. (2016). SeqKit: A Cross-Platform and Ultrafast Toolkit for
671 FASTA/Q File Manipulation. PLoS ONE 11(10):
672 e0163962. https://doi.org/10.1371/journal.pone.0163962
673
674 [51] Siguier, P. (2006). ISfinder: the reference centre for bacterial insertion sequences.
675 Nucleic Acids Res, 34(90001), D32–D36. https://doi.org/10.1093/nar/gkj014
676
677 [52] Su, Y.C. and C. Liu. (2007). *Vibrio parahaemolyticus*: a concern of seafood safety. Food
678 Microbiol. 549-558.
679
680 [53] Sun, L., Bertelshofer, F., Greiner, G., and Böckmann, R. A. (2016). Characteristics of
681 sucrose transport through the sucrose-specific porin scry studied by molecular dynamics
682 simulations. Front. Bioeng. Biotech, https://doi.org/10.3389/fbioe.2016.00009
683
684 [54] Sun, Y., Bernardy, E. E., Hammer, B. K., and Miyashiro, T. (2013). Competence and
685 natural transformation in Vibrios. Mol. Microbiol, 89, 583–595.
686 https://doi.org/10.1111/mmi.12307
687
688 [55] Tamazian, G., Dobrynin, P., Krasheninnikova, K., Komissarov, A., Koepfli, K. P., and
689 O'Brien, S. J. (2016). Chromosomer: A reference-based genome arrangement tool for
690 producing draft chromosome sequences. GigaScience, 5(1), 38.
691 https://doi.org/10.1186/s13742-016-0141-6
692
693 [56] Tankeshwar, A. (2013). Thiosulfate citrate bile salts sucrose (TCBS) agar: composition,
694 uses and colony characteristics. Culture Media used in Microbiology, Laboratory
695 Diagnosis of Bacterial Disease. Accessed at: https://microbeonline.com/thiosulfate-
696 citrate-bile-salts-sucrose-tcbs-agarcomposition-uses-colony-characteristics/
697
698 [57] Tendencia, E. A., Bosma, R. H., and Sorio, L. R. (2012). Effect of three innovative culture
699 systems on water quality and white spot syndrome virus (WSSV) viral load in WSSV-fed
700 Penaeus monodon cultured in indoor tanks. Aquaculture, 350–353, 169–174.
701 https://doi.org/10.1016/j.aquaculture.2012.04.001
702
703 [58] Thompson, J., Gregory, S., Plummer, S., Shields, R. J., and Rowley, A. F. (2010). An in
704 vitro and in vivo assessment of the potential of Vibrio spp. as probiotics for the Pacific
705 White Shrimp, Litopenaeus vannamei. J. Appl. Microbiol., 109(4), 1177–1187.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

706 https://doi.org/10.1111/j.1365-2672.2010.04743.x
707
708 [59] Tran L, L. Nunan, R.M. Redman, L.L. Mohney, C.R. Pantoja, K. Fitzsimmons, and D.V.
709 Lightner. (2013) Determination of the infectious nature of the agent of acute
710 hepatopancreatic necrosis syndrome affecting penaeid shrimp. Dis. Aquat. Organ.
711 105(1):45–55. DOI: 10.3354/dao02621
712
713 [60] Whitaker WB, Richards GP, Boyd EF. (2014) Loss of sigma factor RpoN increases
714 intestinal colonization of *Vibrio parahaemolyticus* in an adult mouse model. *Infect*
715 *Immun. 82(2):544–56*
716
717 [61] Wohlhieter, J. A., J. R. Lazere, N. J. Snellings, E. M. Johnson, R. M. Synenki, and L. S.
718 Baron. (1975). Characterization of transmissible genetic elements from sucrose-
719 fermenting Salmonella strains. J. Bacteriol. 122:401-4069.
720
721

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

722
723 **Supplementary Material**
724
725
726 **Table 1**. Blastn result of partial *16s rRNA* gene of *Vibrio parahaemolyticus* strain PH05 showing
727 100% identity with *Vibrio parahaemolyticus* strains in the database.
728

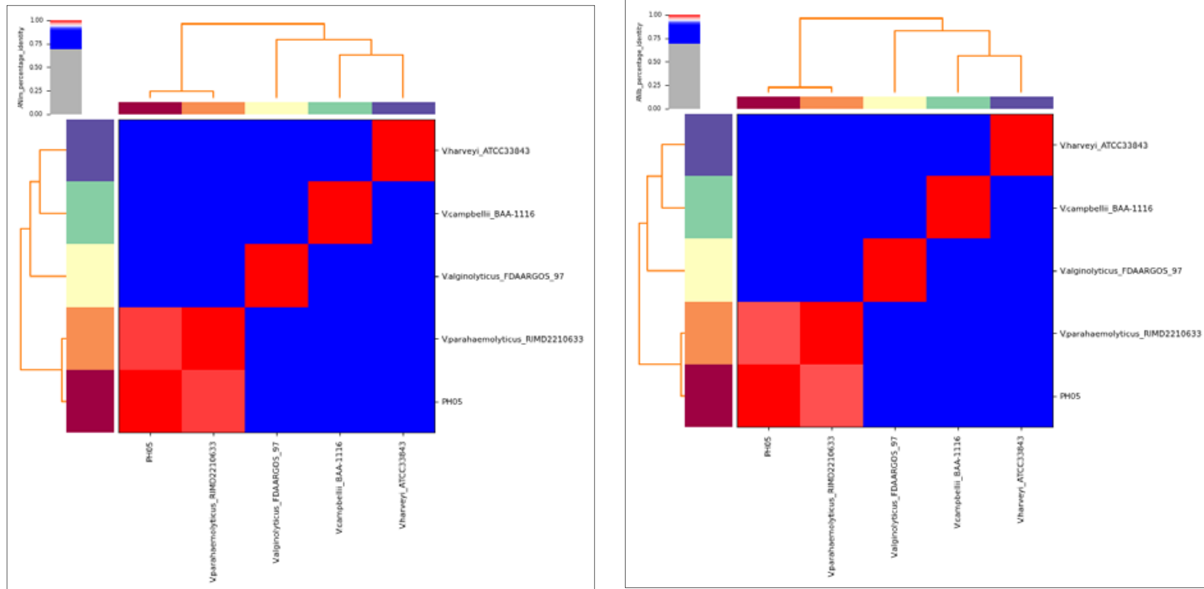
Description	Accession no.	Percent identity	E value	Max score	Total score	Query Cover
Vibrio parahaemolyticus strain 2012AW-0154 chromosome 1, complete sequence	CP035701.1	100%	0.0	1737	19031	100%
Vibrio parahaemolyticus strain FDAARGOS_662 chromosome 2, complete sequence	CP044071.1	100%	0.0	1737	19018	100%
Vibrio parahaemolyticus strain MVP1 chromosome 1, complete sequence	CP043421.1	100%	0.0	1737	18743	100%
Vibrio parahaemolyticus strain Vb0624 chromosome 1, complete sequence	CP041202.1	100%	0.0	1737	19081	100%
Vibrio parahaemolyticus strain 160807 chromosome 2, complete sequence	CP033142.1	100%	0.0	1737	1737	100%

729
730
731
732 **Table 2.** Percentage Identities of *V. parahaemolyticus* strain PH05 against *Vibrio* reference
733 genomes.

Species	strain	Genbank accession	Refseq accession	Percentage similarity (PH05)	
				ANIb	ANIm
Vibrio alginolyticus	FDAARGOS_97	GCA_001471275.2	GCF_001471275.2	83.47%	86.87%
Vibrio campbellii	BAA-1116	GCA_000464435.1	GCF_000464435.1	80.96%	85.18%
Vibrio harveyi	ATCC33843	GCA_000770115.2	GCF_000770115.1	80.69%	85.19%
V. parahaemolyticus	RIMD2210633	GCA_000196095.1	GCF_000196095.1	98.43%	98.55%

734
735
736

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.27.433172; this version posted February 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

737 

738

739 **Figure 1. Average Nucleotide Identity (ANI) heat maps.** The degree of similarity between
740 putative *Vibrio parahaemolyticus* strain PH05 and other Vibrio species was assessed using
741 ANI. The ANI heat maps showed higher percentage identity (98%, see Table 2) of *V.*
742 *parahaemolyticus* strain PH05 with that of *V. parahaemolyticus* RIMD 2210633,
743 confirming the identity of *V. parahaemolyticus* strain PH05. Heatmaps from both ANIb (left)
744 and ANim (right) calculations are identical due to small differences in their percentage values.
745