The pan-genome of Splendidus clade species in the family *Vibrionaceae*: insights into evolution, adaptation, and pathogenicity

Chunqi Jiang[1], Hisae Kasai[2], Sayaka Mino[1], Jesús L. Romalde[3], Tomoo Sawabe[1*]

[1]Laboratory of Microbiology, Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Japan
[2]Laboratory of Marine Biotechnology and Microbiology, Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Japan
[3]Departamento de Microbiología y Parasitología, CRETUS & CIBUS-Facultad de Biología. Universidade de Santiago de Compostela, Spain

***Corresponding Author**

Tomoo Sawabe
Laboratory of Microbiology, Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Japan.
E-mail: sawabe@fish.hokudai.ac.jp
Telephone/fax number: +81-138-40-5569

Running title:

Pan-genomic Analyses of Splendidus Clade

Originality-Significance Statement:

The Splendidus clade is the most diverse clade in the family *Vibrionaceae*. They have become known as pathogens in a wide range of marine vertebrates and invertebrates both in natural or aquaculture settings, but the genome backbones to be pathogens based on complete genome sequences are not yet known even in this era of advances in genomics. We obtained the complete genomes of all type strains in the Splendidus clade for the first time and performed pangenome analyses, which provides new insights into the evolution, environmental adaptation, and pathogenicity of this clade species. Fundamental pathogenicity of the Splendidus clade was revealed by the presence of virulence essential genes, such as *hppD* and an EPS type II secretion system (T2SS), in which T2SS may also

be involved in the different ecological niches of this clade. A type III secretion system (T3SS) which contained the highest number of virulence factors (VFs) was found only in *Vibrio chagasii*.

Summary

The Splendidus clade is the largest clade in *Vibrionaceae*, and its members are often related to mortality of marine animals with huge economic losses. The molecular bases of their pathogenicity and virulence, however, remain largely unknown. In particular, the complete genome sequences of the Splendidus clade species are rarely registered, which is one of the obstacles to predict core and/or unique genes responsible to their adaptation and pathogenicity, and to perform a fine scale meta-transcriptome during bacterial infection to their hosts. In this study, we obtained the complete genomes of all type strains in the Splendidus clade and revealed that 1) different genome sizes (4.4-5.9 Mb) with *V. lentus* the biggest and most of them had several big plasmids, likely because of the different features on mobilome elements, 2) the Splendidus clade consists of 19 species except *V. cortegadensis*, and 3 sub-clades (SC) were defined with the 15 most closely related members as SC1; 3) different carbohydrate degradation preferences may be the result of environmental adaptation, 4) a broad prediction of virulence factors (VFs) revealed core and species unique VF genes.

Keywords:

Vibrio, Splendidus clade, complete genome sequences, evolution, virulence factors, pangenome

INTRODUCTION

Vibrios are widespread in aquatic environments such as estuaries, coastal waters, and sediments, where they are frequently associated with both wild and reared marine organisms, including fish, molluscs, crustaceans, rotifers, and corals, in commensal or pathogenic relationships (Thompson *et al.*, 2004; Gomez-Gil *et al.*, 2014; Romalde *et al.*, 2014; Zhang and Li, 2021). As of Feb 2022, there were 197 validly published species in the family *Vibrionaceae*, including 136 ones in

the genus *Vibrio* described on the List of Prokaryotic names with Standing in Nomenclature (LPSN) (Parte *et al.*, 2020). *Vibrio splendidus* was originally described by Beijerinck in 1900 as a ubiquitous luminous marine bacterium (Baumann *et al.*, 1980; Thompson *et al.*, 2001, 2005; Le Roux *et al.*, 2009; Sawabe *et al.*, 2009). *V. splendidus* is distributed worldwide and is the dominant *Vibrio* species in marine environments, which also shows a remarkable genetic diversity (Thompson, *et al.*, 2003b; Le Roux *et al.*, 2009; Diéguez *et al.*, 2011). The ecologically and genetically diverse populations of *V. splendidus* may serve as useful models to study their environmental adaptations (Hunt *et al.*, 2008; Le Roux *et al.*, 2009). In addition, *V. splendidus* has become known as an opportunistic pathogen associated with various incidents of marine invertebrate mortalities, particularly in the Pacific Oyster *Crassostrea gigas* (Kehlet-Delgado *et al.*, 2020; Oyanedel *et al.*, 2020; Liang *et al.*, 2022).

V. splendidus belongs to the Splendidus clade, which is the biggest clade in the family *Vibrionaceae* (Sawabe *et al.*, 2007, 2013; Jiang *et al.*, 2022). The Splendidus clade consists of over 16 described species in the recent study of "Vibrio Clade 3.0" (Jiang *et al*., 2022): *V. atlanticus* (Diéguez *et al.*, 2011), *V. celticus* (Beaz-Hidalgo *et al.*, 2010), *V. chagasii* (Thompson *et al.*, 2003b), *V. coralliirubri* (Poli *et al.*, 2018), *V. crassostreae* (Faury *et al.*, 2004), *V. cyclitrophicus* (Anonymous, 2001; Hedlund and Staley, 2001), *V. echinoideorum* (Hira *et al.*, 2019), *V. fortis* (Thompson *et al.*, 2003a), *V. gallaecicus* (Beaz-Hidalgo *et al.*, 2009), *V. gigantis* (Roux *et al.*, 2005), *V. kanaloae* (Thompson *et al.*, 2003b), *V. lentus* (Macián *et al.*, 2001), *V. splendidus* (Baumann *et al.*, 1980), *V. tasmaniensis* (Thompson *et al.*, 2003c), *V. toranzoniae* (Lasa *et al.*, 2013), and an as yet unvalidated species "*V. profundi*" (Zhang *et al.*, 2019). Another species, *V. hemicentroti* CECT 8714^{T} has been identified as a later heterotypic synonym of *V. splendidus* NCCB 53037^{T} based on whole genome sequence indices (Kim *et al.*, 2013; Lucena *et al.*, 2017). However, due to the lack of fine scale complete genome sequences, the robust affiliations of several potential members to the Splendidus

74 clade have not been fixed yet (Jiang *et al.*, 2022), such as *V. artabrorum* (Diéguez *et al.*, 2011), *V.*
75 *pelagius* (Baumann, 1981) or *V. pomeroyi* (Thompson *et al.*, 2003b). In addition, *V.*
76 *cortegadensis* (Lasa *et al.*, 2014) was positioned between the Splendidus and Anguillarum clades by
77 the means of 16S rRNA gene sequence analysis and multilocus sequence analysis, with the closest
78 neighbours *V. tapetis*, *V. pomeroyi*, and *V. crassostreae*, but the accurate molecular phylogenetic
79 position remains unknown.

80 Meanwhile, most species in the Splendidus clade had shown an association with pathogenicity
81 (**Table 1**), particularly marine animals cultured in hatcheries, which has led to huge losses in the
82 aquaculture industry (Dubert *et al.*, 2017). Recently, genomic sequencing of these pathogens has
83 provided us with insights into the genetic mechanisms of virulence in some of these systems (Kehlet-
84 Delgado *et al.*, 2020). For example, according to the comparative genomic analysis of strains of *V.*
85 *toranzoniae*, differential features of iron acquisition systems and capsule synthesis genes were the
86 main reason that could explain the differences in the pathogenicity in fish (Lasa *et al.*, 2017). Bruto *et*
87 *al.* (2018) reported two ancestral traits which are necessary for virulence within the diverse
88 Splendidus clade, including an exported conserved protein (R5.7) and a
89 multifunctionalautoprocessing repeats-in-toxin (MARTX) cluster. Species-specific mechanisms of
90 cytotoxicity in *V. crassostreae* (R5.7) and *V. tasmaniensis* (Type 6 Secretion System effectors) have
91 also been demonstrated (Rubio *et al.*, 2019). Oyanedel *et al.* (2020) found that a MARTX cluster and
92 a type-six secretion system (T6SS) were necessary for virulence of *V. splendidus* in oysters. However,
93 the unexplored diversity and massive exchange of virulence factors within the Splendidus clade have
94 been indicated by Nasfi *et al.* (2015) using an epidemiological survey. We still could not find any
95 better conclusion on their genome backbone of the Splendidus clade species to be animal pathogens.
96 In spite of the abundance, species diversification, and pathogenicity of the Splendidus clade species,
97 there has been almost no complete genome available for them until now, in particular, no complete

genomes of type strains are available, which limits our understanding of the evolutionary history, ecology, ecogenomics and host-microbe interaction including disease processes of these species. Therefore, the aims of this study were 1) to obtain the complete genome sequences of type strains of Splendidus clade and related species, 2) to define the true membership of the Splendidus clade, 3) to better elucidate the evolutionary history of these species, and 4) to provide insights into hostmicrobial interactions, particularly in pathogenicity.

RESULTS AND DISCUSSION

Complete Genomes and True Members of the Splendidus Clade

Complete genomes of 18 type strains in the Splendidus clade and one related species were obtained. Genomes of all species consisted of two chromosomes and some of them (11 out of 19) possess one to three plasmids (**Table 2**). The genome sizes of Chromosome 1 (Chr. 1) ranged from 3,104,862 to 3,869,574 bp, and those of Chromosome 2 (Chr. 2) ranged from 1,338,470 to 2,060,851 bp. *V. lentus* LMG 21034^T had the biggest genome while *V. artabrorum* CECT 7226^T had the smallest. The biggest number of plasmids were identified in *V. chagasii* LMG 21353^T and *V. gigantis* LMG 22741^T, and the biggest plasmid (395,604 bp) was identified in *V. pelagius* ATCC 25916^T. Range of GC content was 41.4-44.8%. Numbers of CDS, rRNA, and tRNA were 3,850-5,137, 31-46, and 113-139, respectively.

After obtaining complete genomes, the concatenated split network of family *Vibrionaceae* (including 195 species) using 8 housekeeping genes (8-HKGs) was updated: three of the newly included species (*V. artabrorum* CECT 7226^T, *V. pelagius* ATCC 25916^T, and *V. pomeroyi* LMG 20537^T) were clustered with the 16 species of the Splendidus clade proposed in Jiang *et al*. (2022), which now represents the biggest clade (19 species) in the family *Vibrionaceae* (**Figure 1A**). These 19 species cover all species mentioned in recent research (Pérez-Cataluña *et al.*, 2016; Hira *et al.*,

122 2019; Zhang and Li, 2021) and include the as yet unvalidated, "*V. profundi*" (Zhang *et al.*, 2019).
123 However, "*V. profundi*" was excluded from the subsequent analyses in this study due to the lack of
124 both complete genome and available strains. *V. cortegadensis* CECT 7227^{T} was affiliated to the
125 orphan clade, which *Vibrio* genomospecies F6 strain FF-238 belongs. The topology was also
126 confirmed using the Maximum Likelihood (ML) algorism using the same sequence set used for
127 MLSA (**Figure 1B**). ANI value between *V. cortegadensis* CECT 7227^{T} and *Vibrio* genomosp. F6 str.
128 FF-238 (98.1%) exceeded the delineation boundary of species (95-96%), which indicates that they
129 are likely to be the same species (**Figure S1**).
130 In addition, compared to the genome size and GC content relationships among over 190
131 *Vibrionaceae* species (**Figure 2A**), the Splendidus clade species showed a rather stable GC content
132 (around 44%, with the exception of *V. gallaecicus*) but a bigger variation in genome sizes (4.4-5.9
133 Mb). Plasmids harboring strains tend to have bigger chromosomes (**Figure 2B**), but that genome size
134 variation was likely to be caused by sum of gene regions classified as "Mobiolome" (see
135 "Environmental Adaptation" section described below).
136
137 **Evolutionary Relationships in the Splendidus Clade**
138 The Splendidus clade has not been considered to be a well-defined group because of
139 phenotypic and genetic diversity (Thompson *et al.*, 2004, 2005). In addition, it has been mentioned
140 that 8-HKGs MLSA is a reliable and effective tool for delineating new species and clades in
141 *Vibrionaceae*, because of the higher gene resolution (Jiang *et al.*, 2022). According to the 8-HKGs
142 MLSA of 195 *Vibrionaceae*, the Splendidus clade showed two main evolutionary directions; *V. fortis*
143 and *V. pelagius* in one direction with "*V. profundi*", and the major branch including *V. splendidus* in
144 the other (**Figure 3A**). *V. gallaecicus* was deeply branched in the major. Genome sizes of *V.*

145 *gallaecicus*, *V. fortis* and *V. pelagius*, and the others averaged 5 Mb, but lower and higher GC content
146 was observed in *V. gallaecicus*, and *V. fortis* and *V. pelagius* (**Figure 2B**). With regards to phylogeny
147 and genome features, 3 sub-clades (SCs) are proposed in the Splendidus clade; 15 species including
148 *V. splendidus* (sub-clade 1, SC1), *V. fortis* and *V. pelagius* (SC2), and *V. gallaecicus* (singleton, SC3).
149 This sub-clade proposal is likely to be supported by the ANI and AAI matrix using complete genomes
150 (**Figure 3B and 3C**).
151 It is assumed that ancestory of the Splendidus clade had 5 Mb in size and 45% GC content
152 based on these complete genome comparisons, but during its evolution, GC content of *V. gallaecicus*
153 (SC3) reduced and genome sizes of SC1 species varied. It has been reported that two genome
154 sequences of *Prochlorococcus* lacking the DNA repair enzyme 6-0-methylguanineDNA
155 methyltransferase had very low GC content (Giovannoni *et al.*, 2005), however, this enzyme was
156 encoded in the all genomes of the Splendidus clade species, which indicates that other factors are the
157 cause of reduction of GC content in *V. gallaecicus*. Selection has been repeatedly proposed as the
158 major mechanism to drive toward low genomic GC content in free-living marine bacterioplankton
159 (Giovannoni *et al.*, 2014; Luo *et al.*, 2015). Luo *et al*. suggested that selection maintains the low GC
160 content of SAR11 in the marine population. Therefore, the reduction of GC content in *V. gallaecicus*
161 might also be caused by selection from habitat transition.
162

163 **Genomic Comparison in the Splendidus Clade**

164 In order to understand the evolutionary processes of the Splendidus clade, synteny profiling
165 was performed compared against *V. lentus* LMG 21034^{T}, which had the biggest genome size and
166 most
167 CDSs, using BLASTn. Genes on Chromosome 1 (Chr. 1) were more conserved than those on
168 Chromosome 2 (Chr. 2), while genes on plasmids were very diverse **(Figure 4A)**. The same results can

169 be inferred from the subsequent split pan-genomes of the Splendidus clade. A total of 11 out 19 species
170 possessed one to several plasmids, of which size ranged 1,858 to 332,195 bp **(Figure 2B and Table 2)**.
171 Unfortunately, not only simple gene annotations but also split-pan-genome of those plasmids (see the
172 section Pan-Genome Analysis) did not reaveal any conserved features, which means further detail
173 analyses how those plasmids affect genome plasticity and/or pathogenicity of the Splendidus calde
174 species.

175 It is known that mobile elements can enhance gene insertion and deletion (Vale *et al.*, 2022),
176 synteny profiling for inter and intra sub-clades was also performed to check the sub-clade definition
177 (**Figure 4B**). *V. splendidus* in SC1 showed a rather different genome composition compared to SC2 (*V.*
178 *fortis* and *V. pelagius*) and SC3 (*V. gallaecicus*), this might explain the greater number of genomic
179 islands (GEIs) that were predicted among the members of SC1 than SC2 and SC3 (**Figure 4C**), which
180 is a significant part of the horizontal gene transfer (HGT) events (Dobrindt *et al.*, 2004). Intra sub-clade
181 profiling in SC1 further revealed the insertion/deletion events among them (**Figure 4B**). ABC
182 transporter complex FhuCDB involved in Fe_3^+-hydroxamate import, responsible for energy coupling
183 to the transport system (Schultz-Hauser *et al.*, 1992), and genes *ddhABCD* for the conversion of
184 glucose-1-phosphate to CDP-4-keto-3,6-dideoxy-D-glucose (Pacinelli *et al.*, 2002), were absent in *V.*
185 *echinoideorum* compared to *V. lentus*. While, the aerobactin-producing *iucABCD-iutA* operon was
186 inserted in the genome of *V. kanaloae* compared to *V. toranzoniae*, aerobactin is a citrate-hydroxamate
187 siderophore that is important for the virulence of pathogenic enteric bacteria (Li *et al.*, 2021).

188

189 **Pan-Genome Analysis in the Splendidus Clade**

190 The complete genomes of all Splendidus clade type strains were used for pan-genome analysis
191 using Anvi'o v7. In the Splendidus clade pan-genome (**Figure 5A**), a total of 13,013 gene clusters
192 (GCs) with 83,187 genes were defined, in which 2,391 GCs with 44,619 genes (54%) were recognized
193 in the core-genome (2,198 GCs with 39,564 genes were recognized as the single-copy core-genome),

194 and 4,308 GCs with 32,001 genes (38%) were recognized in the accessory-genome. The remaining
195 genes (8%) were recognized as species-unique genes, among which, *V. chagasii* possessed the highest
196 number of unique genes (770), nearly 5 times that of the lowest. In addition, gene cluster analysis
197 showed that the highest number (4865) of GCs were identified in *V. gigantis*, while the lowest (3622)
198 were identified in *V. artabrorum*. A set of geometrically perfect but functionally diverse single-copy
199 core genes (27-BetterSCGs) was selected (**Table S1**) using a custom setting (--min-
200 geometrichomogeneity-index 1, --max-functional- homogeneity-index 0.9) for phylogenomic tree
201 reconstruction, the result of which also support the proposal of three sub-clades. The ANI boundary for
202 Splendidus clade and sub-clade was found to be around 77.5% and 84.0%, respectively (**Figure 5A**).
203 Furthermore, to identify the distributions and locations of these genes, split pan-genomes were
204 also performed using the sequences of Chr. 1, Chr. 2, and plasmids separately (**Figure 5B**). Most of the
205 genes (67%) in the pan-genome of Chr. 1 while only 29% in Chr. 2 were identified as core genes, and
206 no core genes were found in the pan-genome of plasmids, which may be responsible for the genome
207 diversity among them. Compared to the recent pan-genome analyses of the Halioticoli clade using 10
208 complete genomes (Jiang *et al*., 2022), core genes and specific genes on both chromosomes were
209 decreased but accessory genes increased in the Splendidus clade, especially in Chr. 2 **(Figure S3)**.
210 Meanwhile, a majority of the genes in plasmids were identified as species-unique genes, occupying
211 62% pan-genome.
212
213 **Environmental Adaptation of the Splendidus Clade**
214 Along with the pan-genome analysis, both functional classification and metabolism
215 reconstruction were performed for complete genomes and specific genomes of each Splendidus clade
216 species on the basis of the Clusters of Orthologous Genes (COGs) database and KEGG Orthologs
217 (KOs), respectively. In general, COG functional distribution among the Splendidus clade was the
218 same, but with different numbers because of their genome size variation (**Figure S4A**). However, one

219 exception was found in the category of "Mobilome: prophages, transposons (X)" marked in purple.
220 The "mobilome" is defined as genes including integrative and conjugative elements (ICEs), plasmids,
221 insertion sequences (IS), transposons, prophages, integrons, and other genomic islands, which play a
222 significant role in HGT (Frost *et al.*, 2005; Huang *et al.*, 2016). The number of this function varies
223 dramatically within the clade, from 15 in *V. gallaecicus* (0%) to 265 in *V. celticus* (5%), being more
224 abundant in the SC1, which may explain more plasmids and HGT events in this sub-clade, eventually
225 leading to their genome diversity of them. In addition, the diversification of mobilome related genes
226 in Splendidus clade also provides possible further evidence in the recently evaluated phage-bacteria
227 interactions in the context of natural diversity (Kauffman *et al.*, 2022), but further analyses are
228 needed.

229 The functional distributions from specific genomes differed (**Figure S4B**). The highest
230 number of uncharacterized genes was detected in the *V. chagasii* specific genome, while the *V.*
231 *gallaecicus* specific genome had the most abundant gene sets in the characterized functional
232 estimation. Most of the specific genes were assigned into the upper functional category of
233 "INFORMATION STORAGE AND PROCESSING" including the lower layer function of
234 "Transcription (K)", and "Replication, recombination and repair (L)"; and the upper functional
235 category of "CELLULAR PROCESSES AND SIGNALING" including the lower layer function of
236 "Defense mechanisms (V)", "Signal transduction mechanisms (T)", "Cell wall/membrane/envelope
237 biogenesis (M)", and "Mobilome: prophages, transposons (X)". A significant number (113) of "T"
238 was detected in *V. gallaecicus* specific genome, 8 folds the average number in others (14); the second
239 significant number (60) of "X" was detected in *V. chagasii* specific genome, seven-fold the average
240 number in others (9).

241 According to the KEGG metabolism reconstruction, the whole genomes of the Splendidus
242 clade shared similar numbers and distributions of metabolisms but with several differences (**Figure**

243 **S5A**). First, the numbers of genes encoded for "Other carbohydrate metabolism (A2, abbreviation
244 used in this study)" show a wide range within the clade, from 50 (*V. gallaecicus*, SC1) to 101 (*V.*
245 *celticus* and *V. artabrorum*), a one-fold difference, which may be the result of adaptations for different
246 living environments. In detail, differences were mainly found in the ability of Dgalacturonate
247 degradation (M00061 and M00631), galactose degradation (M00632), and ascorbate degradation
248 (M00550). Second, a significant abundance of the gene set involved in "Pathogenicity (K1)" with
249 M00542 and M00850 modules was detected in *V. chagasii*, 1.6 to 3.5 times greater than other species.
250 Finally, several genes encoding M00660 module "Plant pathogenicity (K3)" were only identified in *V.*
251 *chagasii*, indicating its potential pathogenicity towards marine plants. In the results of KEGG
252 metabolism reconstruction for each specific genome (**Figure S5B**), the numbers are few, but
253 distribution was diverse. The specific genome of *V. chagasii* possessed the highest number (40) of
254 reconstructed metabolism but 50% of them were associated with categories "K1" and "K3",
255 suggesting that this characterization of plant pathogenicity is likely unique among the Splendidus
256 clade species. Based on the KEGG annotation, they were recognized as a type III secretion protein
257 (T3SP) cluster C, F, J, L, Q, R, S, T, U, and V (K03219, K03221-K03230), with an ATP synthase
258 (K03224), as well as a zona occludens toxin (K10954), consistent with the specific T3SS related
259 genes identified in the following virulence factor prediction.

260 Adaptive radiations have been considered to be important drivers for environmental fitness,
261 and a recent adaptive radiation leading to fine-scale ecophysiological differentiation in the
262 degradation of an algal glycan in marine microbes including some Splendidus clade populations has
263 been described, and differentiated alginate degradation pathways were observed among populations
264 (Hehemann *et al.*, 2016). Four types of polysaccharide lyase (PL) families were detected in most
265 specie of the Splendiudus clade, but still clear rules in those genes distribution were not elucidated
266 yet (**Figure S6**). In fact, there were no significant correlations between PL gene numbers and genome

267 size, which indicates such portion of alginate metabolism reported previously is unlikely to be the
268 major causes of genome size variation in the SC1. Increased number of complete genomes in the
269 Splendidus clade strains could provide better insights in the ecology and evolution of those recently
270 radiated group in the family *Vibrionaceae* (Sawabe *et al*., 2007).
271
272 **Fundamentals of Splendidus Clade Pathogenicity**
273 The pathogenicity of bacteria is regulated by a complicated system composed of a variety of
274 virulence factors, such as adherence, phagocytosis, chemotaxis, iron uptake, toxin, quorum sensing,
275 and secretion system (Chen *et al.*, 2016; Liu *et al.*, 2016). The broad predictions of VFs for the
276 Splendidus clade were obtained (**Table S3**) based on the complete genome sequences, which also
277 revealed the core (**Table 3**) and unique VF genes (**Figure 6**). In general, the highest number (136) of
278 vibrio VFs was predicted on *V. chagasii,* while the number in other members averaged 100 (**Figure**
279 **S2B**). Meanwhile, several core VFs were found common to the Splendidus clade, such as
280 mannosesensitive hemagglutinin (MSHA type IV pilus) *mshGHIJLMN*, type IV pilus (*pilBCD*), 55
281 kinds of flagella factors, and an extracellular protein secretion (EPS) type II secretion system (T2SS).
282 The T2SS is likely to be involved in the virulence mechanism and environmental fitness of the
283 Splendidus clade species. On the other hand, genes unique to subclades and species were identified,
284 e.g. absence of *flaC* in SC1, presence of a set of T3SS-related genes exclusively in *V. chagasii*. The
285 details are described as follows.
286
287 **Adherence**
288 Bacterial adhesion is a crucial step in the early stages of infection (Zhang and Li, 2021).
289 Mannose-sensitive hemagglutinin (MSHA) *msh* and Type IV pilus *pil* genes were widely distributed in
290 the Splendidus clade, while Type IVB pilus *tcp* and accessory colonization factors (ACF) *acf* genes

291 were virtually absent (**Figure S7**). Both *msh* and *pil* genes are required for biofilm formation and
292 environmental persistence in *V. cholerae* and *V. parahaemolyticus* (Floyd *et al.*, 2020; Sun *et al.*, 2022).
293 The mannose-sensitive hemagglutinin (MSHA) pilus, plays no role in pathogenicity but does so in
294 biofilm formation, and promotes the interactions between *V. cholerae* El Tor and mussel hemolymph
295 in the hemolymph serum, the efficiency of adherence and association with hemocytes is about twofold
296 more than its mutant without MSHA (Watnick *et al.*, 1999; Zampini *et al.*, 2003). Type IV pilus (T4P),
297 which evokes the immune response of hosts, is ubiquitous on the surfaces of Gram-negative bacteria
298 (Craig *et al.*, 2004), while toxin-coregulated pilus (Type IVB pilus) is used in *V. cholerae* to colonize
299 the human intestine with ACFs, causing cholera, a severe diarrheal disease (Li *et al.*, 2008).
300 The involvements of T4Ps in bacterial colonization in *V. tasmaniensis* LGP32 and *V. crassostreae* J29
301 have also been reported recently (Rubio *et al.*, 2019).
302

303 **Antiphagocytosis**

304 Although vibrio pathogens are commonly regarded as extracellular pathogens, an increasing
305 number of isolates have been found to be capable of invading cells (Zhang and Li, 2021). Capsular
306 polysaccharide (CPS) constitutes the outermost surface of the bacterial cell and is the main virulence
307 factor for antiphagocytosis (Morais *et al.*, 2018). The *cpsABCDEFGHIJ* genes, part of which are
308 common in many vibrio pathogens, such as *V. harveyi* (Bramhachari and Dubey, 2006), *V. alginolyticus*
309 (Muralidharan and Jayachandran, 2003), *V. parahaemolyticus* (Enos-Berlage and McCarter, 2000), and
310 *V. vulnificus* (Lee *et al.*, 2013), are absent in all the type strains of the Splendidus clade species and the
311 well-studied pathogen, *V. atlanticus* LGP32. It is likely that this *cps* family is not involved in the
312 pathogenicity of the Splendidus clade. However, other kinds of CPS were detected in the Splendidus
313 clade, such as *rml*, *wbf*, *wec*, and *wz* related genes, in particular, *wz* related genes (*wza*, *wzb*, and *wzc*)
314 were most abundant, 16 of the 18 members possessed all of them.

315

316 **Chemotaxis and Motility**

317 Chemotaxis plays an important role in infection and disease since chemotaxis signaling
318 pathways are widely distributed among pathogenic bacteria, meanwhile, recent research suggests that
319 chemotaxis is crucial in the early stages of infection in different pathogens (Matilla and Krell, 2018).
320 Results showed that kinds of flagella factors were shared among the clade, including 7, 1, 6, 14, 4,
321 16, 3, and 4 related genes in the *che*, *fil*, *fla*, *flg*, *flh*, *fli*, *flr*, and *mot* clusters, respectively. They are
322 mainly involved in bacterial chemotaxis (**Figure S8**) and flagellar assembly (**Figure S9**) (Terashima
323 *et al.*, 2008; Haiko and Westerlund-Wikström, 2013). One exception is flagellin (*flaC*) which was
324 only detected in the members of SC2 and SC3, but absent in those of SC1 (**Table S3**). The protein of
325 FlaC has not been discussed much in *Vibrio*, but the ability of binding epithelial cells and the
326 influence of cell invasion in *Campylobacter jejuni* TGH9011, a food-born pathogen, has been
327 reported (Song *et al.*, 2004).

328

329 **Iron Uptake**

330 Iron uptake systems (IUSs) are an essential part for disease infection, they are primarily
331 regulated by a ferric uptake regulator called Fur in response to iron availability, along with their own
332 specific regulators (Payne *et al.*, 2016; Li and Ma, 2017; Shin, 2021). Significant differences in the iron
333 acquisition systems between genomes of *V. toranzoniae* strain CECT 7225^{T} (no virulence) and R17
334 (virulence to fish) have been reported (Lasa *et al.*, 2017). In this study, numbers of Furs could be
335 identified in the core and accessory genomes of the Splendidus clade with several specific regulators
336 with each located in plasmids, for example, vibriobactin *vibABCE* in *V. echinoideorum*. This *vib* cluster
337 was also present in *V. cholerae* O1 biovar El, *V. vulnificus* CMCP6, and *V. anguillarum* ATCC 68554.

338

339 **Toxins**

340 Bacterial toxins are the major virulence factors that affect the functions of host cells and control
341 the vital processes of living organisms so they can facilitate microbial infection, they are one of the
342 most important virulence factors that determine whether an infection will succeed or fail (Sarkar *et al.*,
343 2021). The hemolysin toxins Vah and MARTX have been characterized as mainly responsible for the
344 hemolytic and cytotoxic activity of fish pathogens, which cause erythrocyte lysis in the host cells (Frans
345 *et al.*, 2011). Besides, studies have shown that a MARTX toxin cluster (*rtxACHBDE*) was necessary
346 for the virulence of *V. splendidus* in oysters (Bruto *et al.*, 2018; Oyanedel *et al.*, 2020). However, a
347 *rtxABCD* cluster was only identified in *V. echinoideorum*, and hemolysin/cytolysin *vvhA* were absent
348 from all type strains of the Splendidus clade.

349

350 **Quorum Sensing**

351 It is well-known that the outcome of the interaction between the host and bacterium is heavily
352 influenced by the bacterial population size. This intercellular communication is also known as the
353 "Quorum sensing (QS)" system, which is regulated by small diffusible signal molecules called
354 autoinducers (de Kievit and Iglewski, 2000; Whitehead *et al.*, 2001; Winzer and Williams, 2001).
355 Acyl-homoserine lactones (AHL) are one of the most common autoinducers in Gram-negative
356 bacteria, and there are three QS systems generally present in vibrios, for example, VanM/N,

357 VanS/PQ, and VanI/R quorum sensing systems in *V. anguillarum* serotype O1 (Milton, 2006);
358 LuxM/LuxN, LuxS/LuxPQ, and LuxCqsA/LuxCqsS systems in *V. crassostreae* J2-9 (Lemire *et al.*,
359 2015); and LuxM/LuxN, LuxR/LuxI, and LuxS/LuxPQ systems in *V. tasmaniensis* LGP32 (Tait *et al.*,
360 2010). A three-channel QS system was also found in *V. harveyi*, a well-recognized and serious
361 pathogen in fish and invertebrates (Defoirdt *et al.*, 2008; Yang *et al.*, 2011; Zhang *et al.*, 2020). They
362 were mediated by the harveyi autoinducer 1 (HAI-1), autoinducer 2 (AI-2), and cholerae autoinducer

363 1 (CAI-1), respectively. These autoinducers are detected at the cell surface by the LuxN, LuxQ and
364 CqsS two-component receptor proteins, respectively; and the periplasmic protein (LuxP) is required
365 for the detection of AI-2 by LuxQ (Defoirdt *et al.*, 2008). According to the prediction in VFDB, AI-2
366 (*luxS*) was widespread in all species of the Splendidus clade while CAI-1 (*cqsA*) was predicted in 15
367 out of 18 members. Based on the KOfam annotation in the pangenome, the LuxS/LuxPQ-like (AI-2
368 mediated) system was widespread among the clade, while LuxCqsA/LuxCqsS-like (CAI-1 mediated)
369 system was absent in *V. gallaecicus* and *V. artabrorum* (**Figure 7A**), both of which were considered
370 as environmental non-pathogenic species today (Romalde *et al.*, 2014). Meanwhile, according to the
371 *in vivo* mutant experiments, both AI-2 and CAI-1 were likely necessary for the virulence of *V.*
372 *harveyi* in brine shrimp (Defoirdt *et al.*, 2005, 2008). It seems to be the same way in virulence of the
373 Splendidus clade pathogens.
374
375 **Secretion System**
376 Secretion systems (SSs) in bacterial pathogens are responsible for the secretion of various
377 proteins and toxins which contribute towards promoting bacterial virulence, six different SSs have been
378 identified in Gram-negative bacteria, type I secretion system (T1SS) to Type VI system (T6SS) (Sarkar
379 *et al.*, 2021; Zhang and Li, 2021). Many pathogens use dedicated SSs to secrete proteins involved in
380 virulence from the cytosol of the bacteria into host cells or the host environment, but T3SS and T5SS
381 were found to be less associated with virulence (Green and Mecsas, 2016). Four types of SSs were
382 found present in the Splendidus clade.
383 An EPS T2SS system, consisting of 12 EPS proteins and one putative secretin GspD was
384 conserved in the core genome of Splendidus clade species (**Figure S10**). The type II secretion pathway
385 is regarded as one of the major virulence mechanisms in bacterial infection, it has been found in
386 numerous bacterial species, including several extracellular pathogens, such as human (*V. cholerae* and

387 *Pseudomonas aeruginosa*), fish (*Aeromonas hydrophila*), and plant (*Erwinia carotovora*) pathogens
388 (Sandkvist, 2001). Meanwhile, T2SS is also considered a major survival mechanism for environmental
389 species due to the degradative enzymes secreted (Johnson *et al.*, 2014). This kind of strategy for
390 maintaining fitness in different ecological niches in *V. cholerae* has been discussed previously (Sikora,
391 2013). Therefore, the conserved T2SS system in the Splendidus clade may be involved in the virulence
392 and different ecological niches of this clade.

393 A T3SS related region was found exclusively in *V. chagasii* (LOCUS_5670 to LOCUS_6080)
394 (**Figure 7B**), it seems related to the plant pathogenicity according to the KEGG annotation. The same
395 genes were also found in the pathogens *V. parahaemolyticus* RIMD 2210633, *V. harveyi*
396 FDAARGOS_107, and *V. alginolyticus* FDAARGOS_110. Although T3SS was found to be less
397 associated with virulence (Green and Mecsas, 2016), a recent study reported that T3SS effector
398 proteins,
399 Val1686 and Val1680 from *V. alginolyticus*, were responsible for T3SS-mediated death of fish cells
400 (Zhao *et al.*, 2018). The VirB/D system, a model of T4SS, and five components of T6SS (Hcp, DotU,
401 VasA, VasK, TssA1) were found in the genomes of some of the members, in which Hcp protein was
402 involved in the immune system evasion and biofilm formation in *A. hydrophila* (Rasmussen-Ivey *et al.*,
403 2016). Rubio *et al*. (2019) showed that T6SS plays a critical role in the success of vibrio infections, and
404 Oyanedel *et al.* (2020) also showed the necessity of T6SS in the virulence of *V. splendidus* towards
405 oysters, but more details of these SSs need to be explored at a future date.

406

407 **Other VFs**

408 A hemolysin 4-hydroxyphenylpyruvate dioxygenase (4-hppD), encoded by *hppD* gene, has
409 been demonstrated to be related to virulence of *V. splendidus* in sea cucumbers (Liang *et al.*, 2016)
410 and oysters (Liang *et al.*, 2022). It was found highly conserved in the genomes of the Splendidus
411 clade types. Metalloprotease (Vsm) is another VF in *Vibrio* pathogens, metalloprotease activities

412 seem to be a common feature of pathogenic bacteria strains associated with mortality episodes of
413 *Crassostrea gigas* reared in France (Saulnier *et al.*, 2010; Zhang and Li, 2021). The known
414 metalloprotease encoding gene *vsm*, a major determinant of toxicity for extracellular products, was
415 reported in *V. splendidus* LGP32, which was re-named to be as *V. atlanticus* LGP32 (Binesse *et al.*,
416 2008; Zhang *et al.*, 2016), but was not found in the prediction for genomes of type strains of the
417 Splendidus clade. However, an *in silico* homology search showed that they were identified as
418 extracellular zinc metalloprotease (Hap), secreted by *V. cholerae* O1, in the part of genomes.
419

420 **CONCLUSION**

421 In summary, the complete genomes of Splendidus clade type strains were successfully obtained,
422 with a wide range of genome sizes and more numerous and bigger plasmids than other vibrios. Based
423 on the complete genomes, the most recent taxonomic analysis using MLSA of 8-HKGs with 195
424 *Vibrionaceae*, resulted in a total of 19 robust members in the Splendidus clade, which confirms it as
425 the largest clade in the family *Vibrionaceae*. Furthermore, the phylogenetic analyses revealed three sub-
426 clades in the Splendidus clade. COG and KEGG annotations were estimated for each genome of the
427 Splendidus clade. Overall, the same function and metabolism structures were shared among members
428 of the clade but with different carbohydrate degradation preferences, and several T3SS related proteins
429 were abundant exclusively on the plasmid of *V. chagasii*. Complete prediction of virulence factors
430 suggested that a T2SS system may be involved in the virulence mechanism and environmental fitness
431 of the Splendidus clade, meanwhile, AI-2 and CAI-1 quorum sensing systems are likely necessary for
432 the virulence of Splendidus clade pathogens. These results are useful in gaining a better knowledge of
433 the evolutionary history, environmental relationships, and pathological processes of the Splendidus
434 clade species, but further experimental evidence is needed.
435

436 **EXPERIMENTAL PROCEDURES**

437 **Strains, Culture and Genome Collection**

438 A total of 19 type strains of *V. splendidus* related species **(Table 2)** were obtained from the
439 CECT (Spanish Collection of Type Cultures), BCCM/LMG Bacteria collection (Belgian Coordinated
440 Collections of Microorganisms) and our laboratory collection. They were cultured on ZoBell 2216E
441 broth overnight at 25℃ with shaking for DNA extraction. The genomes of "Vibrio Clade 3.0" (Jiang
442 *et al.*, 2022) were retrieved for the classification of Splendidus clade, older draft genomes of this
443 clade were replaced by the genomes obtained in this study.

444

445 **DNA Extraction, Library Preparation and Sequencing**

446 Bacteria DNA was extracted using the Wizard genomic DNA purification kit (Promega,
447 Madison, WI, USA) following the manufacturer's instructions. The Nanopore sequencing library was
448 prepared using the Rapid Barcoding Kit (SQK-RBK004) and loaded onto the MinION device under
449 MinKNOW v3.6.0 (Oxford Nanopore Technologies, Oxford, UK) for sequencing. ONT raw reads
450 (fasta5 files) were basecalled by Guppy 3.2.8. The Illumina DNA library was constructed using
451 Nextera XT DNA Library Preparation Kit (Illumina) and sequenced with the Illumina MiSeq
452 platform (300 bp length) following the manufacturer's instructions. The adaptor sequences were
453 removed using the platanus trim function in Platanus_B (Kajitani *et al.*, 2020).

454

455 **Genome Assembly and Annotation**

456 *De novo* assemblies were performed on all strains using hybrid assembly. Most of the
457 complete genomes were obtained by Unicycler 0.4.7 or 0.4.8 (Wick *et al.*, 2017) with some
458 exceptions. Draft assembly of *V. chagasii* LMG 21353^{T} was created by Flye 2.8.3 (Kolmogorov *et*
459 *al.*, 2019) with genomeSize=5 m using Nanopore long reads, then sequences were corrected with
460 Racon 1.4.20 (Vaser *et al.*, 2017) and Medaka 1.0.1 (Oxford Nanopore Technologies Ltd.), finally

461 polished by Pilon 1.24 (Walker *et al.*, 2014)using Illumina short reads. The genome of *V. celticus*
462 CECT 7224^{T} was assembled by Canu 1.6 (Koren *et al.*, 2017) and polished by Pilon 1.24, then *dnaA*
463 gene and overlaps were checked and edited using *in silico* MolecularCloning ver. 7 (In Silico
464 Biology, Inc., Yokohama, Japan). The resulting complete genomes were checked with CheckM v1.1.3
465 (Parks *et al.*, 2015) and annotated with the DDBJ Fast Annotation and Submission Tool (DFAST)
466 v0.2.7 (Tanizawa *et al.*, 2018). All sequences used in this study are publically available under
467 DDBJ/EMBL/GenBank accession number of AP025458-AP025515 (Table 2).
468
469 **Molecular Phylogeny**
470 The molecular phylogenetic analysis was performed using multilocus sequence analysis
471 (MLSA) according to Sawabe *et al*. (2013). Split decomposition analysis using the concatenated
472 sequence was performed using SplitsTree 4.14.8 with a neighbor net drawing and a Jukes-Cantor
473 correction (Jiang *et al.*, 2022). The sequences were aligned with MUSCLE (Edgar, 2004), and the
474 phylogenetic tree was constructed using MEGA-X v10.1.8 (Kumar *et al.*, 2018) with 1,000 bootstraps
475 using Maximum Likelihood (ML) method and General Time Reversible model (Nei and Kumar,
476 2000).
477
478 **Genome Taxonomy**
479 Average Nucleotide Identity (ANI) values were calculated using Orthologous Average
480 Nucleotide Identity Tool version 0.93.1 (Lee *et al.*, 2016) and Genome-based distance matrix
481 calculator (ANI-Matrix), Kostas lab (Rodriguez-R and Konstantinidis, 2016). Average Amino Acid
482 Identity (AAI) values were calculated using Genome-based distance matrix calculator (AAI-Matrix)
483 (Rodriguez-R and Konstantinidis, 2016). *In silico* DNA-DNA hybridization (DDH) values were
484 estimated using a Genome-to-Genome Distance Calculator 2.1 (GGDC) (Meier-Kolthoff *et al.*,

485 2013). Data was visualized with ComplexHeatmap ver. 2.2.0 (Gu *et al.*, 2016). Genomic comparison
486 was performed with Circular Genome Viewer (CGView) using BLAST (Stothard and Wishart, 2005)
487 and visualized using Proksee (https://proksee.ca/).
488

489 **Pan-Genome Analyses**

490 Pan-genome analysis was performed using the Anvi'o program ver. 7 (Eren *et al.*, 2015).
491 Firstly, each genome sequence file was converted to an anvi'o contigs database (anvi-gen-
492 contigsdatabase) using Prodigal (Hyatt *et al.*, 2010), these contigs databases were decorated with hits
493 from HMM models (anvi-run-hmms). Meanwhile, gene annotation was performed using Clusters of
494 Orthologous Groups 2020 (COG20) (Galperin *et al.*, 2021) for function annotation (anvi-run-
495 ncbicogs), and KOfam (a customized HMM database of KEGG Orthologs (KOs)) (Aramaki *et al.*,
496 2020) for metabolism and pathway annotation (anvi-run-kegg-kofams). Then, an anvi'o genome
497 storage was generated using the prepared contigs databases (anvi-gen-genomes-storage), and next, the
498 pangenome was analyzed using NCBI's blastp for amino acid sequence similarity search and the
499 MCL algorithm (Van Dongen and Abreu-Goodger, 2012) for cluster identification (anvi-pan-
500 genome). In addition, Average Nucleotide Identity (ANI) values were calculated using the PyANI
501 with ANIb method (anvi-compute-genome-similarity) (Pritchard *et al.*, 2016). Finally, it was
502 visualized, decorated, and summarized manually (anvi-display-pan, anvi-summarize). Core genes
503 were filtered and extracted in fasta files for further analysis (anvi-get-sequences-for-gene-clusters).
504

505 **Genomic Islands (GEIs) and Virulence Factors (VFs) Prediction**

506 Genomic island (GEI), which is a important signof a HGT event (Dobrindt *et al*., 2004) was
507 predicted by IslandViewer4 (Bertelli *et al.*, 2017) using IslandPick, IslandPath-DIMOB, and
508 SIGIHMM methods, predictions supported by at least one method were used in this study.

509 Virulence factors (VFs) were predicted using VFanalyzer (Liu *et al.*, 2019), a comparative
510 pathogenomics-based VF analysis pipeline, on the basis of the virulence factor database (VFDB,
511 http://www.mgc.ac.cn/VFs/) (Chen *et al.*, 2016). Reference vibrio pathogens in VFDB were used for
512 comparison: *V. cholerae* O1 biovar El Tor str. N16961, *V. cholerae* O395, *V. fischeri* ES114, *V.*
513 *harveyi* ATCC BAA-1116, *V. parahaemolyticus* RIMD 2210633, *V. vulnificus* CMCP6, and *V.*
514 *vulnificus* YJ016.
515

516 **Acknowledgements**

517 This study was partly supported by MEXT KAKEN 19H03041.

518 **Conflicts of Interest**

519 The authors declare no conflicts of interest.

520 **Data availability**

521 The genomic sequences used in this study are publicaly available under DDBJ/EMBL/GenBank
522 accession number of AP025458-AP025515.

523 **References**

524 Anonymous (2001) Notification that new names and new combinations have appeared in volume 51,
525 part 1, of the IJSEM. *Int J Syst Evol Microbiol* **51**: 269–269.
526 Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., and Ogata, H. (2020)
527 KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score
528 threshold. *Bioinformatics* **36**: 2251–2252.
529 Baumann, P., Baumann, L., Bang, S., and Woolkalis, M. (1980) Reevaluation of the taxonomy of
530 *Vibrio*, *Beneckea*, and *Photobacterium*: abolition of the genus *Beneckea*. *Curr Microbiol* **4**: 127–
531 132.
532 Beaz-Hidalgo, R., Diéguez, A.L., Cleenwerck, I., Balboa, S., Doce, A., de Vos, P., and Romalde, J.L.
533 (2010) *Vibrio celticus* sp. nov., a new Vibrio species belonging to the Splendidus clade with
534 pathogenic potential for clams. *Syst Appl Microbiol* **33**: 311–315.
535 Beaz-Hidalgo, R., Doce, A., Pascual, J., Toranzo, A.E., and Romalde, J.L. (2009) *Vibrio gallaecicus*
536 sp. nov. isolated from cultured clams in north-western Spain. *Syst Appl Microbiol* **32**: 111–117.
537 Bertelli, C., Laird, M.R., Williams, K.P., Lau, B.Y., Hoad, G., Winsor, G.L., and Brinkman, F.S.
538 (2017) IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic*
539 *Acids Res* **45**: W30–W35.

Binesse, J., Delsert, C., Saulnier, D., Champomier-Vergès, M.-C., Zagorec, M., Munier-Lehmann, H., et al. (2008) Metalloprotease Vsm is the major determinant of toxicity for extracellular products of *Vibrio splendidus*. *Appl Environ Microbiol* **74**: 7108–7117.

Bramhachari, P.V. and Dubey, S.K. (2006) Isolation and characterization of exopolysaccharide produced by *Vibrio harveyi* strain VB23. *Lett Appl Microbiol* **43**: 571–577.

Bruto, M., Labreuche, Y., James, A., Piel, D., Chenivesse, S., Petton, B., et al. (2018) Ancestral gene acquisition as the key to virulence potential in environmental *Vibrio* populations. *ISME J* **12**: 2954–2966.

Chen, L., Zheng, D., Liu, B., Yang, J., and Jin, Q. (2016) VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res* **44**: D694–D697.

Craig, L., Pique, M.E., and Tainer, J.A. (2004) Type IV pilus structure and bacterial pathogenicity. *Nat Rev Microbiol* **2**: 363–378.

Defoirdt, T., Boon, N., Sorgeloos, P., Verstraete, W., and Bossier, P. (2008) Quorum sensing and quorum quenching in *Vibrio harveyi*: Lessons learned from in *vivo* work. *ISME J* **2**: 19–26.

Defoirdt, T., Bossier, P., Sorgeloos, P., and Verstraete, W. (2005) The impact of mutations in the quorum sensing systems of *Aeromonas hydrophila*, *Vibrio anguillarum* and *Vibrio harveyi* on their virulence towards gnotobiotically cultured *Artemia franciscana*. *Environ Microbiol* **7**: 1239–1247.

Diéguez, A.L., Beaz-Hidalgo, R., Cleenwerck, I., Balboa, S., de Vos, P., and Romalde, J.L. (2011) *Vibrio atlanticus* sp. nov. and *Vibrio artabrorum* sp. nov., isolated from the clams *Ruditapes philippinarum* and *Ruditapes decussatus*. *Int J Syst Evol Microbiol* **61**: 2406–2411.

Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2**: 414–424.

Van Dongen, S. and Abreu-Goodger, C. (2012) Using MCL to extract clusters from networks., pp. 281–295.

Dubert, J., Barja, J.L., and Romalde, J.L. (2017) New insights into pathogenic Vibrios affecting bivalves in hatcheries: Present and future prospects. *Front Microbiol* **8**: 762.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

Enos-Berlage, J.L. and McCarter, L.L. (2000) Relation of capsular polysaccharide production and colonial cell organization to colony morphology in *Vibrio parahaemolyticus*. *J Bacteriol* **182**: 5513–5520.

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., and Delmont, T.O. (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**: e1319.

Faury, N., Saulnier, D., Thompson, F.L., Gay, M., Swings, J., and Roux, F. Le (2004) *Vibrio crassostreae* sp. nov., isolated from the haemolymph of oysters (*Crassostrea gigas*). *Int J Syst Evol Microbiol* **54**: 2137–2140.

Floyd, K.A., Lee, C.K., Xian, W., Nametalla, M., Valentine, A., Crair, B., et al. (2020) c-di-GMP modulates type IV MSHA pilus retraction and surface attachment in *Vibrio cholerae*. *Nat Commun* **11**: 1549.

Frans, I., Michiels, C.W., Bossier, P., Willems, K.A., Lievens, B., and Rediers, H. (2011) *Vibrio anguillarum* as a fish pathogen: virulence factors, diagnosis and prevention. *J Fish Dis* **34**: 643–661.

Frost, L.S., Leplae, R., Summers, A.O., and Toussaint, A. (2005) Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **3**: 722–732.

585 Galperin, M.Y., Wolf, Y.I., Makarova, K.S., Vera Alvarez, R., Landsman, D., and Koonin, E. V (2021)
586 COG database update: focus on microbial diversity, model organisms, and widespread
587 pathogens. *Nucleic Acids Res* **49**: D274–D281.
588 Giovannoni, S.J., Cameron Thrash, J., and Temperton, B. (2014) Implications of streamlining theory
589 for microbial ecology. *ISME J* **8**: 1553–1565.
590 Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., et al. (2005) Genome
591 streamlining in a cosmopolitan oceanic bacterium. *Science (80-)* **309**: 1242–1245.
592 Gomez-Gil, B., Thompson, C.C., Matsumura, Y., Sawabe, Toko, Iida, T., Christen, R., et al. (2014)
593 The Famlily *Vibrionaceae*. In *The Prokaryotes*. Berlin, Heidelberg: Springer Berlin Heidelberg,
594 pp. 659–747.
595 Green, E.R. and Mecsas, J. (2016) Bacterial secretion systems: An overview. *Microbiol Spectr* **4**:.
596 Gu, Z., Eils, R., and Schlesner, M. (2016) Complex heatmaps reveal patterns and correlations in
597 multidimensional genomic data. *Bioinformatics* **32**: 2847–2849.
598 Haiko, J. and Westerlund-Wikström, B. (2013) The role of the bacterial flagellum in qdhesion and
599 virulence. *Biology (Basel)* **2**: 1242–1267.
600 Hedlund, B.P. and Staley, J.T. (2001) *Vibrio cyclotrophicus* sp. nov., a polycyclic aromatic
601 hydrocarbon (PAH)-degrading marine bacterium. *Int J Syst Evol Microbiol* **51**: 61–66.
602 Hehemann, J.-H., Arevalo, P., Datta, M.S., Yu, X., Corzett, C.H., Henschel, A., et al. (2016) Adaptive
603 radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes.
604 *Nat Commun* **7**: 12860.
605 Hira, J., Bentdal, S., Devold, H., Stensvåg, K., and Landfald, B. (2019) *Vibrio echinoideorum* sp.
606 nov., isolated from an epidermal lesion on the test of a green sea urchin (Strongylocentrotus
607 droebachiensis). *Int J Syst Evol Microbiol* **69**: 2277–2282.
608 Huang, J., Ma, J., Shang, K., Hu, X., Liang, Y., Li, D., et al. (2016) Evolution and diversity of the
609 antimicrobial resistance associated mobilome in *Streptococcus suis*: A probable mobile genetic
610 elements reservoir for other *Streptococci*. *Front Cell Infect Microbiol* **6**:.
611 Hunt, D.E., David, L.A., Gevers, D., Preheim, S.P., Alm, E.J., and Polz, M.F. (2008) Resource
612 partitioning and sympatric differentiation among closely related bacterioplankton. *Science (80-)*
613 **320**: 1081–1085.
614 Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal:
615 prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*
616 **11**: 119.
617 Jiang, C., Tanaka, M., Nishikawa, S., Mino, S., Romalde, J.L., Thompson, F.L., et al. (2022) Vibrio
618 Clade 3.0: New *Vibrionaceae* evolutionary units using genome-based approach. *Curr Microbiol* **79**:
619 1–15.
620 Johnson, T.L., Fong, J.C., Rule, C., Rogers, A., Yildiz, F.H., and Sandkvist, M. (2014) The type II
621 secretion system delivers matrix proteins for biofilm formation by *Vibrio cholerae*. *J Bacteriol*
622 **196**: 4245–4252.
623 Kajitani, R., Yoshimura, D., Ogura, Y., Gotoh, Y., Hayashi, T., and Itoh, T. (2020) Platanus_B: an
624 accurate *de novo* assembler for bacterial genomes using an iterative error-removal process. *DNA*
625 *Res* **27**: dsaa014.
626 Kauffman, K.M., Chang, W.K., Brown, J.M., Hussain, F.A., Yang, J., Polz, M.F., and Kelly, L.
627 (2022) Resolving the structure of phage–bacteria interactions in the context of natural diversity.
628 *Nat Commun* **13**: 372.

Kehlet-Delgado, H., Häse, C.C., and Mueller, R.S. (2020) Comparative genomic analysis of Vibrios yields insights into genes associated with virulence towards *C. gigas* larvae. *BMC Genomics* **21**: 599.

de Kievit, T.R. and Iglewski, B.H. (2000) Bacterial quorum sensing in pathogenic relationships. *Infect Immun* **68**: 4839–4849.

Kim, D., Baik, K.S., Hwang, Y.S., Choi, J.-S., Kwon, J., and Seong, C.N. (2013) *Vibrio hemicentroti* sp. nov., an alginate lyase-producing bacterium, isolated from the gut microflora of sea urchin (*Hemicentrotus pulcherrimus*). *Int J Syst Evol Microbiol* **63**: 3697–3703.

Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019) Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* **37**: 540–546.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017) Canu: scalable and accurate long-read assembly via adaptive k -mer weighting and repeat separation. *Genome Res* **27**: 722–736.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* **35**: 1547–1549.

Lasa, A., Diéguez, A.L., and Romalde, J.L. (2014) *Vibrio cortegadensis* sp. nov., isolated from clams. *Antonie Van Leeuwenhoek* **105**: 335–341.

Lasa, A., Diéguez, A.L., and Romalde, J.L. (2013) *Vibrio toranzoniae* sp. nov., a new member of the Splendidus clade in the genus *Vibrio*. *Syst Appl Microbiol* **36**: 96–100.

Lasa, A., Gibas, C.J., and Romalde, J.L. (2017) Comparative genomic analysis of two *Vibrio toranzoniae* strains with different virulence capacity reveals clues on its pathogenicity for fish. *Front Microbiol* **8**:.

Lee, I., Ouk Kim, Y., Park, S.-C., and Chun, J. (2016) OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* **66**: 1100–1103.

Lee, K.-J., Kim, J.-A., Hwang, W., Park, S.-J., and Lee, K.-H. (2013) Role of capsular polysaccharide (CPS) in biofilm formation and regulation of CPS production by quorumsensing in *Vibrio vulnificus*. *Mol Microbiol* **90**: 841–857.

Lemire, A., Goudenège, D., Versigny, T., Petton, B., Calteau, A., Labreuche, Y., and Le Roux, F. (2015) Populations, not clones, are the unit of vibrio pathogenesis in naturally infected oysters. *ISME J* **9**: 1523–1531.

Li, C., Pan, D., Li, M., Wang, Y., Song, L., Yu, D., et al. (2021) Aerobactin-mediated iron acquisition enhances biofilm formation, oxidative stress resistance, and virulence of *Yersinia pseudotuberculosis*. *Front Microbiol* **12**: 699913.

Li, J., Lim, M.S., Li, S., Brock, M., Pique, M.E., Woods, V.L., and Craig, L. (2008) *Vibrio cholerae* toxin-coregulated pilus structure analyzed by hydrogen/deuterium exchange mass spectrometry. *Structure* **16**: 137–148.

Li, Y. and Ma, Q. (2017) Iron acquisition strategies of *Vibrio anguillarum*. *Front Cell Infect Microbiol* **7**: 342.

Liang, W., Zhang, C., Liu, N., Zhang, W., Han, Q., and Li, C. (2016) Cloning and characterization of *Vshppd*, a gene inducing haemolysis and immune response of *Apostichopus japonicus*. *Aquaculture* **464**: 246–252.

Liang, W., Zhang, W., and Li, C. (2022) *Vibrio splendidus* virulence to *Apostichopus japonicus* is mediated by hppD through glutamate metabolism and flagellum assembly. *Virulence* **13**: 458–470.

Liu, B., Zheng, D., Jin, Q., Chen, L., and Yang, J. (2019) VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res* **47**: D687–D692.

675 Liu, R., Chen, H., Zhang, R., Zhou, Z., Hou, Z., Gao, D., et al. (2016) Comparative transcriptome
676 analysis of *Vibrio splendidus* JZ6 reveals the mechanism of its pathogenicity at low
677 temperatures. *Appl Environ Microbiol* **82**: 2050–2061.
678 Lucena, T., Carmen Macián, M., Arahal, D.R., Rodrigo-Torres, L., and Pujalte, M.J. (2017) Whole
679 genome sequences reveal *Vibrio hemicentroti* Kim et al. 2013 as a later heterotypic synonym of
680 *Vibrio splendidus* (Beijerinck 1900) Baumann et al. 1981. *Int J Syst Evol Microbiol* **67**: 1669–
681 1671.
682 Luo, H., Thompson, L.R., Stingl, U., and Hughes, A.L. (2015) Selection maintains low genomic GC
683 content in marine SAR11 lineages. *Mol Biol Evol* **32**: 2738–2748.
684 Macián, M.C., Ludwig, W., Aznar, R., Grimont, P.A., Schleifer, K.H., Garay, E., and Pujalte, M.J.
685 (2001) *Vibrio lentus* sp. nov., isolated from Mediterranean oysters. *Int J Syst Evol Microbiol* **51**:
686 1449–1456.
687 Matilla, M.A. and Krell, T. (2018) The effect of bacterial chemotaxis on host infection and
688 pathogenicity. *FEMS Microbiol Rev* **42**: fux052.
689 Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.-P., and Göker, M. (2013) Genome sequence-based species
690 delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**:
691 60.
692 Milton, D.L. (2006) Quorum sensing in vibrios: Complexity for diversification. *Int J Med Microbiol*
693 **296**: 61–71.
694 Morais, V., Dee, V., and Suárez, N. (2018) Purification of capsular polysaccharides of *Streptococcus*
695 *pneumoniae*: Traditional and new methods. *Front Bioeng Biotechnol* **6**: 145.
696 Muralidharan, J. and Jayachandran, S. (2003) Physicochemical analyses of the exopolysaccharides
697 produced by a marine biofouling bacterium, *Vibrio alginolyticus*. *Process Biochem* **38**: 841–
698 847.
699 Nasfi, H., Travers, M.A., de Lorgeril, J., Habib, C., Sannie, T., Sorieul, L., et al. (2015) A European
700 epidemiological survey of *Vibrio splendidus* clade shows unexplored diversity and massive
701 exchange of virulence factors. *World J Microbiol Biotechnol* **31**: 461–475.
702 Nei, M. and Kumar, S. (2000) Molecular evolution and phylogenetics, Oxford university press.
703 Oyanedel, D., Labreuche, Y., Bruto, M., Amraoui, H., Robino, E., Haffner, P., et al. (2020) *Vibrio*
704 *splendidus* O-antigen structure: a trade-off between virulence to oysters and resistance to
705 grazers. Environ Microbiol 22: 4264–4278.
706 Pacinelli, E., Wang, L., and Reeves, P.R. (2002) Relationship of *Yersinia pseudotuberculosis* O
707 antigens IA, IIA, and IVB: the IIA gene cluster was derived from that of IVB. *Infect Immun* **70**:
708 3271–3276.
709 Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM:
710 assessing the quality of microbial genomes recovered from isolates, single cells, and
711 metagenomes. *Genome Res* **25**: 1043–1055.
712 Parte, A.C., Sardà Carbasse, J., Meier-Kolthoff, J.P., Reimer, L.C., and Göker, M. (2020) List of
713 Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol*
714 *Microbiol* **70**: 5607–5612.
715 Payne, S.M., Mey, A.R., and Wyckoff, E.E. (2016) *Vibrio* iron transport: Evolutionary adaptation to
716 life in multiple environments. *Microbiol Mol Biol Rev* **80**: 69–90.
717 Pérez-Cataluña, A., Lucena, T., Tarazona, E., Arahal, D.R., Macián, M.C., and Pujalte, M.J. (2016)
718 An MLSA approach for the taxonomic update of the Splendidus clade, a lineage containing
719 several fish and shellfish pathogenic *Vibrio* spp. *Syst Appl Microbiol* **39**: 361–369.

720 Poli, A., Romano, I., Mastascusa, V., Buono, L., Orlando, P., Nicolaus, B., et al. (2018) *Vibrio*
721 *coralliirubri* sp. nov., a new species isolated from mucus of red coral (*Corallium rubrum*)
722 collected at Procida island, Italy. *Antonie Van Leeuwenhoek* **111**: 1105–1115.
723 Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G., and Toth, I.K. (2016) Genomics and
724 taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal*
725 *Methods* **8**: 12–24.
726 Rasmussen-Ivey, C.R., Figueras, M.J., McGarey, D., and Liles, M.R. (2016) Virulence factors of
727 *Aeromonas hydrophila*: In the wake of reclassification. *Front Microbiol* **7**: 1337.
728 Rodriguez-R, L.M. and Konstantinidis, K.T. (2016) The enveomics collection: a toolbox for
729 specialized analyses of microbial genomes and metagenomes. *PeerJ Prepr* **4**: e1900v1.
730 Romalde, J.L., Dieguez, A.L., Lasa, A., and Balboa, S. (2014) New *Vibrio* species associated to
731 molluscan microbiota: A review. *Front Microbiol* **4**: 413.
732 Roux, F. Le, Goubet, A., Thompson, F.L., Faury, N., Gay, M., Swings, J., and Saulnier, D. (2005)
733 *Vibrio gigantis* sp. nov., isolated from the haemolymph of cultured oysters (*Crassostrea gigas*).
734 *Int J Syst Evol Microbiol* **55**: 2251–2255.
735 Le Roux, F., Zouine, M., Chakroun, N., Binesse, J., Saulnier, D., Bouchier, C., et al. (2009) Genome
736 sequence of *Vibrio splendidus* : an abundant planctonic marine species with a large genotypic
737 diversity. *Environ Microbiol* **11**: 1959–1970.
738 Rubio, T., Oyanedel, D., Labreuche, Y., Toulza, E., Luo, X., Bruto, M., et al. (2019) Species-specific
739 mechanisms of cytotoxicity toward immune cells determine the successful outcome of *Vibrio*
740 infections. *Proc Natl Acad Sci* **116**: 14238–14247.
741 Sandkvist, M. (2001) Type II secretion and pathogenesis. *Infect Immun* **69**: 3523–3535.
742 Sarkar, P., Issac, P.K., Raju, S. V., Elumalai, P., Arshad, A., and Arockiaraj, J. (2021) Pathogenic
743 bacterial toxins and virulence influences in cultivable fish. *Aquac Res* **52**: 2361–2376.
744 Saulnier, D., De Decker, S., Haffner, P., Cobret, L., Robert, M., and Garcia, C. (2010) A large-scale
745 epidemiological study to identify bacteria pathogenic to pacific oyster *Crassostrea gigas* and
746 correlation between virulence and metalloprotease-like activity. *Microb Ecol* **59**: 787–798.
747 Sawabe, T., Kita-Tsukamoto, K., and Thompson, F.L. (2007) Inferring the evolutionary history of
748 vibrios by means of multilocus sequence analysis. *J Bacteriol* **189**: 7932–7936.
749 Sawabe, T., Koizumi, S., Fukui, Y., Nakagawa, S., Ivanova, E.P., Kita-Tsukamoto, K., et al. (2009)
750 Mutation is the main driving force in the diversification of the Vibrio splendidus clade. *Microbes*
751 *Environ* **24**: 281–285.
752 Sawabe, Tomoo, Ogura, Y., Matsumura, Y., Feng, G., Amin, A.R., Mino, S., et al. (2013) Updating
753 the *Vibrio* clades defined by multilocus sequence phylogeny: proposal of eight new clades, and
754 the description of *Vibrio tritonius* sp. nov. *Front Microbiol* **4**: 414.
755 Schultz-Hauser, G., Köster, W., Schwarz, H., and Braun, V. (1992) Iron(III) hydroxamate transport in
756 *Escherichia coli* K-12: FhuB-mediated membrane association of the FhuC protein and negative
757 complementation of fhuC mutants. *J Bacteriol* **174**: 2305–2311.
758 Shin, S.-H. (2021) Regulation of iron-uptake systems in *Vibrio vulnificus*, a ferrophilic bacterium.
759 *Med Biol Sci Eng* **4**: 69–82.
760 Sikora, A.E. (2013) Proteins secreted via the type II secretion system: Smart strategies of *Vibrio*
761 *cholerae* to maintain fitness in different ecological niches. *PLoS Pathog* **9**: e1003126.
762 Song, Y.C., Jin, S., Louie, H., Ng, D., Lau, R., Zhang, Y., et al. (2004) FlaC, a protein of
763 *Campylobacter jejuni* TGH9011 (ATCC43431) secreted through the flagellar apparatus, binds
764 epithelial cells and influences cell invasion. *Mol Microbiol* **53**: 541–553.

765 Stothard, P. and Wishart, D.S. (2005) Circular genome visualization and exploration using CGView.
766 *Bioinformatics* **21**: 537–539.
767 Sun, J., Li, X., Qiu, Y., Xue, X., Zhang, M., Yang, W., et al. (2022) Quorum sensing regulates
768 transcription of the pilin gene mshA1 of MSHA pilus in *Vibrio parahaemolyticus*. *Gene* **807**:
769 145961.
770 Tait, K., Hutchison, Z., Thompson, F.L., and Munn, C.B. (2010) Quorum sensing signal production
771 and inhibition by coral-associated vibrios. *Environ Microbiol Rep* **2**: 145–150.
772 Tanizawa, Y., Fujisawa, T., and Nakamura, Y. (2018) DFAST: a flexible prokaryotic genome
773 annotation pipeline for faster genome publication. *Bioinformatics* **34**: 1037–1039.
774 Terashima, H., Kojima, S., and Homma, M. (2008) Chapter 2 Flagellar motility in *Bacteria*., pp. 39–
775 85.
776 Thompson, F.L., Hoste, B., Vandemeulebroecke, K., and Swings, J. (2001) Genomic diversity
777 amongst *Vibrio* isolates from different sources determined by fluorescent amplified fragment
778 length polymorphism. *Syst Appl Microbiol* **24**: 520–538.
779 Thompson, F.L., Iida, T., and Swings, J. (2004) Biodiversity of Vibrios. *Microbiol Mol Biol Rev* **68**:
780 403–431.
781 Thompson, F. L., Thompson, C.C., Hoste, B., Vandemeulebroecke, K., Gullian, M., and Swings, J.
782 (2003a) *Vibrio fortis* sp. nov. and *Vibrio hepatarius* sp. nov., isolated from aquatic animals and
783 the marine environment. *Int J Syst Evol Microbiol* **53**: 1495–1501.
784 Thompson, F. L., Thompson, C.C., Li, Y., Gomez-Gil, B., Vandenberghe, J., Hoste, B., and Swings, J.
785 (2003b) *Vibrio kanaloae* sp. nov., *Vibrio pomeroyi* sp. nov. and *Vibrio chagasii* sp. nov., from sea
786 water and marine animals. *Int J Syst Evol Microbiol* **53**: 753–759.
787 Thompson, F.L., Thompson, C.C., and Swings, J. (2003c) *Vibrio tasmaniensis* sp. nov., isolated from
788 Atlantic Salmon (*Salmo salar L.*). *Syst Appl Microbiol* **26**: 65–69.
789 Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., et al. (2005)
790 Genotypic diversity within a natural coastal bacterioplankton population. *Science (80-)* **307**:
791 1311–1313.
792 Travers, M.-A., Boettcher Miller, K., Roque, A., and Friedman, C.S. (2015) Bacterial diseases in
793 marine bivalves. *J Invertebr Pathol* **131**: 11–31.
794 Vale, F.F., Lehours, P., and Yamaoka, Y. (2022) Editorial: The role of mobile genetic elements in
795 bacterial evolution and their adaptability. *Front Microbiol* **13**:.
796 Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017) Fast and accurate *de novo* genome assembly
797 from long uncorrected reads. *Genome Res* **27**: 737–746.
798 Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014) Pilon: An
799 integrated tool for comprehensive microbial variant detection and genome assembly
800 improvement. *PLoS One* **9**: e112963.
801 Watnick, P.I., Fullner, K.J., and Kolter, R. (1999) A role for the mannose-sensitive hemagglutinin in
802 biofilm formation by *Vibrio cholerae* El Tor. *J Bacteriol* **181**: 3606–3609.
803 Whitehead, N.A., Barnard, A.M.L., Slater, H., Simpson, N.J.L., and Salmond, G.P.C. (2001) Quorum-
804 sensing in Gram-negative bacteria. *FEMS Microbiol Rev* **25**: 365–404.
805 Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. (2017) Unicycler: Resolving bacterial genome
806 assemblies from short and long sequencing reads. *PLOS Comput Biol* **13**: e1005595.
807 Winzer, K. and Williams, P. (2001) Quorum sensing and the regulation of virulence gene expression
808 in pathogenic bacteria. *Int J Med Microbiol* **291**: 131–143.
809 Yang, Q., Han, Y., and Zhang, X.-H. (2011) Detection of quorum sensing signal molecules in the
810 family *Vibrionaceae*. *J Appl Microbiol* **110**: 1438–1448.

Zampini, M., Canesi, L., Betti, M., Ciacci, C., Tarsi, R., Gallo, G., and Pruzzo, C. (2003) Role for mannose-sensitive hemagglutinin in promoting interactions between *Vibrio cholerae* El Tor and mussel hemolymph. *Appl Environ Microbiol* **69**: 5711–5715.

Zhang, C., Liang, W., Zhang, W., and Li, C. (2016) Characterization of a metalloprotease involved in *Vibrio splendidus* infection in the sea cucumber, *Apostichopus japonicus*. *Microb Pathog* **101**: 96–103.

Zhang, N.-X., Zhang, D.-C., and Qiao, N.-H. (2019) *Vibrio profundi* sp. nov., isolated from a deepsea seamount. *Antonie Van Leeuwenhoek* **112**: 1603–1610.

Zhang, W. and Li, C. (2021) Virulence mechanisms of vibrios belonging to the Splendidus clade as aquaculture pathogens, from case studies and genome data. *Rev Aquac* raq.12555.

Zhang, X.H., He, X., and Austin, B. (2020) *Vibrio harveyi*: a serious pathogen of fish and invertebrates in mariculture. *Mar Life Sci Technol* **2**: 231–245.

Zhao, Z., Liu, J., Deng, Y., Huang, W., Ren, C., Call, D.R., and Hu, C. (2018) The *Vibrio alginolyticus* T3SS effectors, Val1686 and Val1680, induce cell rounding, apoptosis and lysis of fish epithelial cells. Virulence 9: 318–330.

825 **Table 1. List of *Vibrio splendidus* related species.** Bold species were newly confirmed in
826 Splendidus clade in this study than Jiang *et al*., 2022. +: contains strains pathogenic for marine
827 animals, NA: not available.

Species	Clade	Pathogenicity	Source	Place	16S rRNA accession	References
V. artabrorum	Splendidus	NA	Cultured clams, *Ruditapes philippinarum*	Galicia, north-western Spain	EF599164	Diéguez *et al.* 2011
V. atlanticus	Splendidus	+	Cultured clams, *Ruditapes philippinarum*	Galicia, north-western Spain	EF599163	Diéguez *et al.* 2011
V. celticus	Splendidus	+	Cultured clams, *Ruditapes philippinarum*	Galicia, north-western Spain	EF599162	Beaz-Hidalgo *et al.* 2011
V. chagasii	Splendidus	NA	Gut of turbot larvae, *Scophthalmus maximus*	Austevoll, Norway	MT757983	Thompson *et al.* 2003b
V. coralliirubri	Splendidus	NA	Mucus of red coral, *Corallium rubrum*	Procida island, Naples, Italy	HG942391	Poli *et al.* 2018
V. crassostreae	Splendidus	+	Haemolymph of diseased oyster, *Crassostreae gigas*	La Tremblade, France	EF094887	Faury *et al.* 2004
V. cyclitrophicus	Splendidus	NA	Creosote-contaminated marine sediments	Puget Sound, Washington, USA	DQ481610	Hedlund and Staley 2001
V. echinoideorum	Splendidus	NA	Epidermal lesion of green sea urchin, *Strongylocentrotus droebachiensis*	Northern Norway	MG788349	Hira *et al.* 2019
V. fortis	Splendidus	+	White shrimp larvae, *Litopenaeus vannamei*	Ecuador	AJ514916	Thompson et al. 2003a
V. gallaecicus	Splendidus	NA	Cultured Manila clams, *Ruditapes philippinarum*	Galicia, north-western Spain	EU541605	Beaz-Hidalgo *et al.* 2009
V. gigantis	Splendidus	+	Diseased oyster, *Crassostreae gigas*	La Tremblade, France	AJ582810	Le Roux *et al.* 2005
V. kanaloae	Splendidus	+	Diseased oyster larvae, *Ostrea edulis*	IFREMER, France	AJ316193	Thompson *et al.* 2003b
V. lentus	Splendidus	+	Cultivated oyster	Vinaroz, Spain	AJ278881	Macián *et al.* 2001
V. pelagius	Splendidus	+	Seawater enriched with succinate	Hawaii, off coast Oahu, USA	AJ293802	(Baumann *et al.* 1971) Baumann *et al.* 1981
V. pomeroyi	Splendidus	+	Bivalve larvae, *Nodipecten nodosus*	LCMM Florianópolis, southern Brazil	AJ491290	Thompson *et al.* 2003b
V. splendidus	Splendidus	+	Marine fish	-	AJ515229	(Beijerinck 1900) Baumann *et al.* 1981
V. tasmaniensis	Splendidus	+	Atlantic salmon, *Salmo salar*	Tasmania, Australia	AJ316192	Thompson *et al.* 2003c
V. toranzoniae	Splendidus	+	Cultured Manila clams, *Venerupis philippinarum*	Galicia, north-western Spain	HE978310	Lasa *et al.* 2013
"*V. profundi*"	Splendidus	NA	A deep-sea seamount	Near Yap Trench in the tropical western Pacific	KT900237	Zhang *et al.* 2019
V. cortegadensis	*V. genomosp*. F6	NA	Healthy Manila clam, *Venerupis philippinarum*	Galicia, north-western Spain	HF955037	Lasa *et al.* 2014

828

Table 2. General information of completed genomes for Splendidus clade related species. Taxa IDs are indicated as: 01-*V. artabrorum* CECT 7226^T, 02-*V. atlanticus* CECT 7223^T, 03-*V. celticus* CECT 7224^T, 04-*V. chagasii* LMG 21353^T, 05-*V. coralliirubri* DSM 27495^T, 06-*V. crassostreae* LMG 22240^T, 07-*V. cyclitrophicus* LMG 21359^T, 08-*V. echinoideorum* DSM 107264^T, 09-*V. fortis* LMG 21557^T, 10-*V. gallaecicus* CECT 7244^T, 11-*V. gigantis* LMG 22741^T, 12-*V. kanaloae* LMG 20539^T, 13-*V. lentus* LMG 21034^T, 14-*V. pelagius* ATCC 25916^T, 15-*V. pomeroyi* LMG 20537^T, 16-*V. splendidus* LMG 19031^T, 17-*V. tasmaniensis* LMG 20012^T, 18-*V. toranzoniae* CECT 7225^T, 19-*V. cortegadensis* CECT 7227^T. **IDs** 1-18 are Splendidus clade species. "**Abbr"** indicates the abbreviation names used in this study, "**Chr.**" indicates chromosome.

ID	Abbr	Chr. 1	Sequence size (bp) Chr. 2	Plasmids	Total size (bp)	GC content (%)	CDSs	Number 16S rRNA	tRNA	Accession number
01	Art	3,104,862	1,338,470		4,443,332	44.1	3,850	12	127	AP025458-AP025459
02	Atl	3,469,123	1,605,756	32,802	5,107,681	44.0	4,415	15	137	AP025460-AP025462
03	Cel	3,620,060	2,016,907		5,636,967	44.5	4,898	14	139	AP025463-AP025464
04	Cha	3,413,052	1,828,839	187,827; 81,098; 58,987	5,569,803	44.2	4,989	14	137	AP025465-AP025469
05	Cor	3,681,934	2,060,851		5,742,785	44.5	5,027	14	135	AP025470-AP025471
06	Cra	3,475,743	1,947,964	237,546; 143,128	5,804,381	44.4	5,052	13	139	AP025476-AP025479
07	Cyc	3,415,033	1,727,822	134,644	5,277,499	43.8	4,460	13	137	AP025480-AP025482
08	Ech	3,599,125	2,020,999	72,315; 45,679	5,738,118	43.7	4,935	15	137	AP025483-AP025486
09	For	3,262,072	1,687,370	332,195	5,281,637	44.7	4,507	13	134	AP025487-AP025489
10	Gal	3,292,782	1,833,873		5,126,655	41.4	4,432	15	130	AP025490-AP025491
11	Gig	3,575,998	2,018,505	225,164; 4,542; 1,858	5,826,067	44.2	5,063	14	138	AP025492-AP025496
12	Kan	3,174,145	1,478,351		4,652,496	43.9	3,989	12	130	AP025497-AP025498
13	Len	3,691,782	1,953,633	251,461; 3,959	5,900,835	44.0	5,137	10	124	AP025499-AP025502
14	Pel	3,187,009	1,381,362	395,604	4,963,975	44.8	4,252	13	135	AP025503-AP025505

15	Pom	3,635,862	2,023,719		5,659,581	44.6	4,820	14	138	AP025506-AP025507
16	Spl	3,869,574	2,009,543		5,879,117	44.0	5,086	14	133	AP025508-AP025509
17	Tas	3,226,267	1,501,469	149,865; 114,713	4,992,314	44.1	4,303	15	133	AP025510-AP025513
18	Tor	3,169,590	1,436,362		4,605,952	44.0	3,949	13	131	AP025514-AP025515
19	-	3,116,092	1,374,279	94,062; 84,136	4,668,569	42.5	4,069	14	113	AP025472-AP025475

839 **Table 3. List of core virulence factor gene sets in the Splendidus clade**

VF Class	Virulence factors (number)	Related genes	Function	References
Adherence	Mannose-sensitive hemagglutinin (MSHA type IV pilus) (8)	*mshB*	Biofilm formation	Floyd *et al.*, 2020; Sun *et al.*, 2022
		mshG		
		mshH		
		mshI		
		mshJ		
		mshL		
		mshM		
		mshN		
	Type IV pilus (3)	*pilB*		
		pilC		
		pilD		
Secretion system	EPS type II secretion system (12)	*epsC*	Virulence mechanisms and environmental fitness	Sandkvist, 2001; Sikora, 2013; Johnson *et al*., 2014
		epsE		
		epsF		
		epsG		
		epsH		
		epsI		
		epsJ		
		epsK		
		epsL		
		epsM		
		epsN		
		gspD		
Chemotaxis and motility	Flagella (55)	*cheABRVWYZ*	Bacterial chemotaxis and flagellar assembly	Terashima *et al*., 2008; Haiko and Westerlund-Wikström, 2013
		filM		
		flaABDEGL		
		flgABCDEFGHI JKLMN		
		flhABFG		
		fliADEFGHIJKL NOPQRS		
		flrABC		
		motABXY		

840

841 **Figure captions (80 mm, 169 mm or 110 mm)**

842 **Figure 1.** Molecular phylogenetic analyses using network and bifurcating methods based on
843 concatenated eight housekeeping gene sequences (*ftsZ*, *gapA*, *gyrB*, *mreB*, *pyrH*, *recA*, *rpoA*, and
844 *topA*). A) The split network of a total of 195 *Vibrionaceae* species with *Escherichia coli* K-12
845 (ASM584v2) as outgroup. The eight housekeeping gene sequences were concatenated, and the
846 network was constructed using the SplitsTree4 ver. 4.14.8 with neighbor net drawing method and
847 Jukes-Cantor correction. B) The Maximum Likelihood (ML) rooted tree for the Splendidus clade,
848 constructed by MEGA-X v10.1.8 with 1000 bootstraps, *E. coli* K12 MG1655 and *V. cholerae* ATCC
849 14035^T as outgroups. Reference sequence data was obtained from Jiang *et al*. (2022).

850 **Figure 2.** Genome size and GC content relationship. A) Position of the 18 Splendidus clade species
851 in 193 *Vibrionaceae* species (8 genera). Data was obtained from Jiang *et al*. (2022). B) Species
852 affiliation and sub-clade (SC) of the Splendidus clade. Black borders in B represent the presence of
853 plasmids.

854 **Figure 3.** Sub-clade delineation in the Splendidus clade. A) Splendidus clade intra split network in
855 the broad network of 195 *Vibrionaceae* on the basis of eight housekeeping genes (8-HKGs).
856 Heatmap matrix of values from B) Average Nucleotide Identity (ANI) and C) Average Amino Acid
857 Identity (AAI) within the Splendidus clade. Values were calculated using genome-based distance
858 matrix calculator (ANI-Matrix/AAI-Matrix). Taxa abbreviations are indicated as Table 2.
859 Dendrogram showed the hierarchical clustering. Heatmap was visualized using ComplexHeatmap
860 ver. 2.2.0.

861 **Figure 4.** Genome synteny profiling of A) the Splendidus clade species against *V. lentus* LMG
862 21034^T, B) inter and intra sub-clades in the Splendidus clade with reference genome in bold. The
863 genomes were compared using CGView using BLASTn. The color indicates the presence of
864 homologous genes in each genome. C) The numbers of genomic islands (GEIs) predicted in each
865 Splendidus clade species. Taxa abbreviations are indicated in Table 2.

866 **Figure 5.** The pan-genome analyses using A) complete genome sequences and B) split genome
867 sequences of the Splendidus clade species. Circle bars represent the occurrence of gene clusters in
868 each genome. Gene cluster represents a group of homologues identified based on the amino acid
869 sequence similarity. Heatmap in the upper right corner of A represents ANI similarity between these
870 genomes, and the above phylogenetic tree was constructed using amino acid sequences of 27 better
871 single-copy core genes (SCGs) by embedded FastTree tool. Details of statistical information of
872 layers on the right were shown in Table S2. Taxa abbreviations are indicated as: Art-*V. artabrorum*
873 CECT 7226^T, Atl-*V. atlanticus* CECT 7223^T, Cel-*V. celticus* CECT 7224^T, Cha-*V. chagasii* LMG
874 21353^T,
875 Cor-*V. coralliirubri* DSM 27495^T, Cra-*V. crassostreae* LMG 22240^T, Cys-*V. cyclitrophicus* LMG
876 21359^T, Ech-*V. echinoideorum* DSM 107264^T, For-*V. fortis* LMG 21557^T, Gal-*V. gallaecicus* CECT
877 7244^T, Gig-*V. gigantis* LMG 22741^T, Kan-*V. kanaloae* LMG 20539^T, Len-*V. lentus* LMG 21034^T,
878 Pel-*V. pelagius* ATCC 25916^T, Pom-*V. pomeroyi* LMG 20537^T, Spl-*V. splendidus* LMG 19031^T, Tas-
879 *V. tasmaniensis* LMG 20012^T, Tor-*V. toranzoniae* CECT 7225^T.

880 Figure 6. Broad prediction of virulence factors (VFs) related genes for the Splendidus clade, the left
881 text represents VF class, only differences were shown here, detailed predictions and locations are

listed in Table S3. The upper phylogenetic tree was constructed by MEGA-X using nucleic acid sequences of the 27-BetterSCGs, and edited using FigTree v1.4.4. Taxa abbreviations are indicated as: Art-V. artabrorum CECT 7226T, Atl-V. atlanticus CECT 7223T, Cel-V. celticus CECT 7224T Cha-V. chagasii LMG 21353T, Cor-V. coralliirubri DSM 27495T, Cra-V. crassostreae LMG 22240T , Cys-*V. cyclitrophicus* LMG 21359^T, Ech-*V. echinoideorum* DSM 107264^T, For-*V. fortis* LMG 21557^T, Gal-*V. gallaecicus* CECT 7244^T, Gig-*V. gigantis* LMG 22741^T, Kan-*V. kanaloae* LMG 20539^T, Len-*V. lentus* LMG 21034^T, Pel-*V. pelagius* ATCC 25916^T, Pom-*V. pomeroyi* LMG 20537^T, Spl-*V. splendidus* LMG 19031^T, Tas-*V. tasmaniensis* LMG 20012^T, Tor-*V. toranzoniae* CECT 7225^T.

Figure 7. KEGG pathway reconstruction of A) "Quorum sensing" (vibrio part of map02024) and B) "Bacterial secretion system" (map03070) for Splendidus clade species using KEGG Mapper Reconstruct Tool. A clear illustration of B can be found in the supplemental files. Taxa IDs are indicated as: 01-*V. artabrorum* CECT 7226^T, 02-*V. atlanticus* CECT 7223^T, 03-*V. celticus* CECT 7224^T, 04-*V. chagasii* LMG 21353^T, 05-*V. coralliirubri* DSM 27495^T, 06-*V. crassostreae* LMG 22240^T, 07-*V. cyclitrophicus* LMG 21359^T, 08-*V. echinoideorum* DSM 107264^T, 09-*V. fortis* LMG 21557^T, 10-*V. gallaecicus* CECT 7244^T, 11-*V. gigantis* LMG 22741^T, 12-*V. kanaloae* LMG 20539^T, 13-*V. lentus* LMG 21034^T, 14-*V. pelagius* ATCC 25916^T, 15-*V. pomeroyi* LMG 20537^T, 16-*V. splendidus* LMG 19031^T, 17-*V. tasmaniensis* LMG 20012^T, 18-*V. toranzoniae* CECT 7225^T.

Figure S1. Heatmap generated with OrthoANI values. These values were calculated from the Orthologous Average Nucleotide Identity Tool version 0.93.1 using genomes of related species in Figure 1.

Figure S2. The total numbers of virulence factors (VFs) related genes predicted in each Splendidus clade species. Taxa abbreviations are indicated as: Art-*V. artabrorum* CECT 7226^T, Atl-*V. atlanticus* CECT 7223^T, Cel-*V. celticus* CECT 7224^T, Cha-*V. chagasii* LMG 21353^T, Cor-*V. coralliirubri* DSM 27495^T, Cra-*V. crassostreae* LMG 22240^T, Cys-*V. cyclitrophicus* LMG 21359^T, Ech-*V. echinoideorum* DSM 107264^T, For-*V. fortis* LMG 21557^T, Gal-*V. gallaecicus* CECT 7244^T, Gig-*V. gigantis* LMG 22741^T, Kan-*V. kanaloae* LMG 20539^T, Len-*V. lentus* LMG 21034^T, Pel-*V. pelagius* ATCC 25916^T, Pom-*V. pomeroyi* LMG 20537^T, Spl-*V. splendidus* LMG 19031^T, Tas-*V. tasmaniensis* LMG 20012^T, Tor-*V. toranzoniae* CECT 7225^T.

Figure S3. Identification of accessory, core and specific genomes in the split pangenomes of Chromosome 1, Chromosome 2, and Plasmids for the Splendidus and Halioticoli clade using Anvi'o7.

Figure S4. Numbers (right) and percentages (left) from COG functional classification for each Splendidus clade species in A) complete genomes and B) specific genomes. The abbreviations for the Splendidus clade species are represented in Table 2. Abbreviations for COG categories are indicated as: "Transcription (K)", "Replication, recombination and repair (L)", "Cell wall/membrane/envelope biogenesis (M)", "Signal transduction mechanisms (T)", "Defense mechanisms (V)", and "Mobilome: prophages, transposons (X)", the others are represented in Table S4.

Figure S5. Numbers (right) and percentages (left) from KEGG metabolism construction for each

922 Splendidus clade species in A) complete genomes and B) specific genomes. The abbreviations for
923 the Splendidus clade species are represented in Table 2. Abbreviations for KEGG categories are
924 shown as: “Other carbohydrate metabolism (A2)”, “Pathogenicity (K1)”, and “Plant pathogenicity
925 (K3)”, the others are shown in Table S5.

926 **Figure S6.** The maximum copy number of alginate lyase families within Splendidus clade. Lyases
927 were predicted using the dbCAN2 meta server (HMMdb v9) with HMMER, DIAMOND, and
928 eCAMI tools , and domains supported by more than two tools were used. The upper phylogenetic
929 tree was constructed by MEGA-X using nucleic acid sequences of the 27-BetterSCGs, and edited
930 using FigTree v1.4.4. Polysaccharide lyase (PL) families are represented by colored rectangles. ND
931 indicates that the lyases were not detected.
932 **Figure S7.** KEGG pathway reconstruction of “Biofilm formation” (map05111) for Splendidus clade
933 species using KEGG Mapper Reconstruct Tool. Taxa IDs are indicated as Figure 7.

934 **Figure S8.** KEGG pathway reconstruction of “Bacterial chemotaxis” (map02030) for Splendidus
935 clade species using KEGG Mapper Reconstruct Tool. Taxa IDs are indicated as Figure 7.

936 **Figure S9.** KEGG pathway reconstruction of “Flagellar assembly” (map02040) for Splendidus clade
937 species using KEGG Mapper Reconstruct Tool. Taxa IDs are indicated as Figure 7.

938 **Figure S10.** KEGG pathway reconstruction of “Bacterial secretion system” (map03070) for
939 Splendidus clade species using KEGG Mapper Reconstruct Tool. Taxa IDs are indicated as Figure 7.

940