



**HAL**  
open science

## Comparative genomics of pathogenic lineages of *Vibrio nigripulchritudo* identifies virulence-associated traits

David Goudenège, Yannick Labreuche, Evelyne Krin, Dominique Ansquer, Sophie Mangenot, Alexandra Calteau, Claudine Médigue, Didier Mazel, Martin F Polz, Frédérique Le Roux

### ► To cite this version:

David Goudenège, Yannick Labreuche, Evelyne Krin, Dominique Ansquer, Sophie Mangenot, et al.. Comparative genomics of pathogenic lineages of *Vibrio nigripulchritudo* identifies virulence-associated traits. *The International Society of Microbiological Ecology Journal*, 2013, Integrated genomics and post-genomics approaches in microbial ecology, 7 (10), pp.1985-1996. 10.1038/ismej.2013.90 . pasteur-01699814

**HAL Id: pasteur-01699814**

**<https://pasteur.hal.science/pasteur-01699814v1>**

Submitted on 2 Feb 2018

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

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

1 **Population genomics of pathogenic lineages of *Vibrio nigripulchritudo* pinpoints**  
2 **virulence specific traits**

3

4 **Running Title:** *Vibrio nigripulchritudo* comparative genomics

5

6 David Goudenège<sup>1</sup>, Yannick Labreuche<sup>1</sup>, Evelyne Krin<sup>3,4</sup>, Dominique Ansquer<sup>2</sup>, Sophie  
7 Mangenot<sup>5</sup>, Alexandra Calteau<sup>5,6</sup>, Claudine Médigue<sup>5,6</sup>, Didier Mazel<sup>3,4</sup>, Martin F. Polz<sup>7</sup> and  
8 Frédérique Le Roux<sup>1</sup>

9

10 <sup>1</sup>Equipe émergente Ifremer-UPMC « Génomique des Vibrio », FR2424, Station Biologique  
11 de Roscoff, Place Georges Tessier, 29688 Roscoff cedex, France.

12 <sup>2</sup> Département Lagons, Ecosystèmes et Aquaculture Durables en Nouvelle-Calédonie,  
13 IFREMER, BP 2059, 98846 Nouméa cedex, New Caledonia.

14 <sup>3</sup> Institut Pasteur, Unité Plasticité du Génome Bactérien, Département Génomes et Génétique,  
15 75015 Paris, France

16 <sup>4</sup> CNRS UMR3525, F-75015 Paris, France

17 <sup>5</sup> CEA, Institut de Génomique, Genoscope, 91057 Evry cedex, France

18 <sup>6</sup> CEA, Institut de Génomique, Genoscope & CNRS UMR8030 Laboratoire de Génomique  
19 Comparative, 91057 Evry cedex, France

20 <sup>7</sup> Parsons Lab for Environmental Science and Engineering, MIT, Cambridge, MA 02139,  
21 USA.

22 **Corresponding author:** Frédérique Le Roux

23 Equipe émergente Ifremer-UPMC « Génomique des Vibrio », FR2424, Station Biologique de  
24 Roscoff, Place Georges Tessier, 29688 Roscoff cedex, France.

25 Tel 33 2 98 29 56 47, Fax 33 2 98 29 83 84, frederique.le-roux@sb-roscoff.fr

26

27 **ABSTRACT**

28

29 *Vibrio nigripulchritudo* is an emerging pathogen of farmed shrimp in New Caledonia and  
30 other regions in the Indo-Pacific. The molecular determinants of *V. nigripulchritudo*  
31 pathogenicity are unknown; however, molecular epidemiological studies have suggested that  
32 pathogenicity is linked to particular lineages. Here, we performed high throughput  
33 sequencing-based comparative genome analysis of 16 *V. nigripulchritudo* strains to explore  
34 the genomic diversity and evolutionary history of pathogen-containing lineages and to  
35 identify pathogen-specific genetic elements. Our phylogenetic analysis revealed three  
36 pathogen containing *V. nigripulchritudo* clades, including two clades previously identified  
37 from New Caledonia and one novel clade comprising putatively pathogenic isolates from  
38 septicemic shrimp in Madagascar. The similar genetic distance between the 3 clades indicates  
39 that they have diverged from an ancestral population roughly at the same time and  
40 recombination analysis indicates that these genomes have, in the past, shared a common gene  
41 pool and exchanged genes. Because each contemporary lineage is comprised of nearly  
42 identical strains, comparative genomics allowed differentiation of genetic elements specific to  
43 shrimp pathogenesis of varying severity. Notably, only a large plasmid present in all highly  
44 pathogenic strains encodes a toxin. Although less/non-pathogenic strains contain related  
45 plasmids, these are differentiated by a putative toxin locus. Expression of this gene by a non-  
46 pathogenic *V. nigripulchritudo* strain resulted in production of toxic culture supernatant,  
47 normally an exclusive feature of highly pathogenic strains. Thus, this protein, here termed  
48 “nigritoxin”, is implicated to an extent that remains to be precisely determined in the toxicity  
49 of *V. nigripulchritudo*.

50

51 **Key words:** horizontal gene transfer / nigritoxin / phylogeny / shrimp / virulence / vibrio

52

53 **INTRODUCTION**

54

55 The family *Vibrionaceae* comprises a diverse group of bacteria that are widespread within  
56 marine environments (Thompson et al 2004). It encompasses the well-studied human  
57 pathogen *Vibrio cholerae* as well as some less thoroughly characterized species that are  
58 responsible for infections in aquatic animals (Austin 2010). The diseases due to these  
59 organisms have serious environmental and economic consequences. It is now widely accepted  
60 that intensive aquaculture practices are accompanied by considerable risks both for the  
61 amplification of existing pathogenic strains and for the spread of novel virulence determinants  
62 from these pathogens into existing commensal bacteria (Waterfield et al 2004).

63

64 *Vibrio nigripulchritudo* was first isolated in New Caledonia in 1995 from diseased shrimp  
65 (*Litopenaeus stylirostris*) affected by a winter vibriosis called Syndrome 93 (Goarant et al  
66 2006b). Initially, it was only reported in two adjoining farms and hence considered as a  
67 geographically restricted phenomenon. However, in 1997, *V. nigripulchritudo* was isolated  
68 from moribund shrimp collected from a farm some 50 km away from the original location,  
69 albeit this time during summer. Because these conditions did not fit the Syndrome 93  
70 epidemiology, which typically occurs when water temperature is below 25°C, this new  
71 disease was named “Summer Syndrome” (Goarant et al 2006a). Mass mortalities of other  
72 penaeid shrimp (*Marsupenaeus japonicus* and *Penaeus monodon*) ascribed to  
73 *V. nigripulchritudo* infections have also been reported in Japan (Sakai et al 2007) and  
74 Madagascar (E. Chung pers.com.), suggesting that these bacterial pathogens may affect wider  
75 areas in the Indo-Pacific.

76

77 A collection of *V. nigripulchritudo* isolates obtained from a variety of sites in New Caledonia  
78 over an 8-year period was previously characterized by both multilocus sequence typing  
79 (MLST) and experimental infections of *L. stylirostris* (Goarant et al 2006a, Goarant et al  
80 2006b). Although these environmental isolates were genotypically diverse, all strains isolated  
81 from shrimp affected by Summer Syndrome and Syndrome 93 grouped into two distinct  
82 clades (A and B, respectively). The virulence of representative strains was assessed using  
83 three complementary experimental infection models (Le Roux et al 2011). First, when shrimp  
84 were intramuscularly injected with live bacteria, strains could be operationally classified into  
85 highly pathogenic (HP), moderately pathogenic (MP) and non pathogenic (NP) variants based  
86 on >80%, 20-80% and <20% observed host mortality, respectively. Second, when shrimp  
87 were transiently immersed into water inoculated with different isolates, only HP strains were  
88 able to induce mortality. Finally when shrimp were intramuscularly injected with culture  
89 supernatants, those from HP strains but not MP strains caused death (exotoxicity). Strains  
90 from clade A were found to be either MP or HP (Reynaud et al 2008), while most strains  
91 assigned to clade B were HP (Goarant et al 2006b). Non-pathogenic isolates (NP) were  
92 genetically diverse. Hence, these results suggested that virulence is specific to clades A and  
93 B.

94

95 To gain better insight into the genetic and genomic underpinnings of the disease, we  
96 developed an HP strain from Summer syndrome clade A (designated SFn1) as a model. This  
97 strain carries two plasmids, pA<sub>SFn1</sub> and pB<sub>SFn1</sub> (formerly pA1066 and pB1067), of 250 and 11  
98 kb, respectively (Le Roux et al 2011, Reynaud et al 2008). By infecting shrimp with  
99 derivatives of SFn1 from which pA<sub>SFn1</sub> and/or pB<sub>SFn1</sub> had been lost, we found that both  
100 plasmids are necessary for full virulence. In the immersion model, only the wild type strain  
101 containing both pA<sub>SFn1</sub> and pB<sub>SFn1</sub> was virulent. In bacterial injection assays, SFn1 was more

102 virulent than a mutant containing pA<sub>SFn1</sub> alone, while a mutant containing only pB<sub>SFn1</sub> was  
103 avirulent. When supernatants were injected, toxicity was dependent upon the presence of  
104 pA<sub>SFn1</sub> only and attributed to a large, heat-sensitive factor. It is likely that differences between  
105 results obtained with the three challenge protocols reflect the different host-imposed barriers  
106 that are encountered by the bacteria. They may also indicate that there are multiple pathways  
107 by which the bacterium impairs shrimp viability.

108

109 Here, we further explore the genomic diversity and evolutionary history of pathogen-  
110 containing lineages of *V. nigripulchritudo*. We ask what types of genetic elements (i)  
111 differentiate each lineage and (ii) are shared among the HP strains within and across lineages.

112 To address these questions, we performed high throughput sequencing (HTS)-based  
113 comparative genome analysis. Our phylogenetic analyses suggest that recent differential  
114 adaptation of lineages was followed by horizontal acquisition of genetic material, which  
115 distinguishes pathotypes within clades. A unique region carried by a large plasmid is specific  
116 to all HP strains and contains a gene encoding a putative toxin. Expression of this gene in a  
117 NP strain results in production of toxic culture supernatant.

118

119

120

## 121 **MATERIALS AND METHODS**

122

### 123 **Bacterial strains, plasmids and culture conditions**

124 The *V. nigripulchritudo* strains used for the genomic analyses are described in Table 1. Other  
125 bacterial strains and plasmids are described in Tables S1 and S2. *V. nigripulchritudo* strains  
126 were grown in Luria-Bertani (LB) or LB-agar (LBA) + NaCl 0.5M, marine broth (MB) or

127 marine agar (MA) at 30°C. *E. coli* strains were grown in LB or on LBA at 37°C.  
128 Spectinomycin (100 µg/ml), thymidine (0.3 mM) and diaminopimelate (0.3 mM) were added  
129 as supplements when necessary. Induction of the P<sub>BAD</sub> promoter was achieved by the addition  
130 of 0.2% L-arabinose to the growth media, and conversely, was repressed by the addition of  
131 1% D-glucose.

132

### 133 **Genome sequencing, assembly and annotation**

134 The complete genome sequence of SFn1 strain was obtained from 1) a Sanger library  
135 sequencing leading to a 4-fold coverage; 2) a 454 single read library sequencing leading to a  
136 16-fold coverage (supplementary methods). The 14 other *V. nigripulchritudo* strains were  
137 sequenced using the Illumina HiSeq2000 technology with ~50-fold coverage. Contigs were  
138 assembled de novo using Velvet (Zerbino and Birney 2008) and genome assembly was  
139 improved by contig mapping against the SFn1 reference genome. Computational prediction of  
140 coding sequences (CDS) together with functional assignments were performed using the  
141 automated annotation pipeline implemented in the MicroScope platform (Vallenet et al 2009).  
142 Antismash and NapDoS softwares were used to identify and annotate secondary metabolite  
143 biosynthesis gene clusters in the *V. nigripulchritudo* genomes (Medema et al 2011, Ziemert et  
144 al 2012). Since the partial genome of *V. nigripulchritudo* type strain (ATCC 27043<sup>T</sup>) is  
145 available in Genbank (Hoffmann et al 2012), it was also included in this study.

### 146 ***In silico* analyses**

147 To investigate the core and flexible genomes, an all-versus-all BlastP search was performed  
148 using genomic sequences of 134 organisms: 118 *Vibrionaceae* (21 completes, 97 whole  
149 genome shotguns, including *V. nigripulchritudo* ATCC 27043<sup>T</sup>) and 1 *Shewanella baltica*  
150 (strain OS155, complete genome) available on Genbank and the 15 *V. nigripulchritudo*  
151 sequenced in the present study (Table 1). A dedicated precomputing repository (marshalling)

152 was created to perform comparative genomic and phylogenomic analyses. Orthologous  
153 proteins were defined as reciprocal best hit proteins with 80% MaxLrap and a minimum of  
154 30% and 60% identity cutoff was used for intra- and inter-species analysis, respectively  
155 (Daubin et al 2002). The nucleic acid sequences were aligned using Muscle (Edgar 2004) and  
156 filtered by Gblocks (Castresana 2000). Phylogenetic trees were built using the parallel version  
157 of PhyML applied to Maximum-likelihood algorithm and GTR model as parameters (NNIs,  
158  $\gamma_4$ , invariant site) (Guindon et al 2010). Reliability was assessed by the bootstrap method with  
159 100 replicates. Tree topologies were analyzed using the ETE2 Python library (Huerta-Cepas  
160 et al 2010).

161

## 162 **Vector construction and mutagenesis**

163 Alleles carrying an internal deletion were generated *in vitro* using a two-step PCR  
164 construction method (supplementary methods) and cloned into an R6K  $\gamma$ -*ori*-based suicide  
165 vector that encodes the *ccdB* toxin gene under the control of an arabinose-inducible and  
166 glucose-repressible promoter,  $P_{BAD}$  (Le Roux et al 2007). Matings between *E. coli* and *V.*  
167 *nigripulchritudo* were performed at 30°C as described previously (Le Roux et al 2011).  
168 Selection of the plasmid-borne drug marker (Spec<sup>R</sup>) resulted in integration of the entire  
169 plasmid in the chromosome by a single crossover. Elimination of the plasmid backbone  
170 resulting from a second recombination step was selected by arabinose induction of the *ccdB*  
171 toxin gene. Mutants were screened by PCR.

172 The *gfp* and nigratoxin genes were PCR amplified and cloned under the control of a  $P_{BAD}$   
173 promoter in a P15A-*ori*-based replicative vector then transferred from *E. coli* to SFn118 by  
174 conjugation (supplementary methods). The same replicative vector was used to clone and  
175 express the *reb* genes cluster in *E. coli*.

176



177 **Transmission electron microscopy**

178 Transmission electron microscopy (TEM) was used to explore R-body production by *V.*  
179 *nigripulchritudo* SFn1 and *E. coli* carrying the reb gene cluster expression vector (GV691).  
180 Expression of recombinant R-bodies was carried out by incubating the GV691 culture at 37°C  
181 until OD<sub>600nm</sub> of 0.4, then adding arabinose and shifting the culture to 25°C for another 18 h  
182 of incubation. The bacteria were fixed, post fixed, dehydrated, embedded and ultracut as  
183 described previously (Schrallhammer et al 2012).

184

185 **Experimental challenges**

186 Bacterial supernatants were prepared from bacterial cultures grown overnight in MB +/-  
187 glucose or arabinose. Specimens of *L. stylirostris* were injected with either 100 µl of filtered  
188 (0.22 µm) supernatant or sterile culture medium as a control following a previously described  
189 procedure (Le Roux et al 2011). Experiments were conducted in triplicate (5 shrimp/tank,  
190 n=15). Survival was monitored for 24 hours, preliminary trials having established that no  
191 significant mortality occurred after this time. The experiment was repeated three times.

192

193

194

195 **RESULTS**  
196

197 **General features of the *V. nigripulchritudo* genomes**

198 The genome of strain SFn1 was closed and annotated manually. It consists of two circular  
199 chromosomes of 4.1 (chromosome 1) and 2.2 (chromosome 2) Mb with an average percent  
200 G+C content of 45.9 and 45.5, respectively (Table 1). As described previously, SFn1 also  
201 carries two plasmids of 250 (pA<sub>SFn1</sub>) and 11 (pB<sub>SFn1</sub>) kb (Le Roux et al 2011). There is a total  
202 of 5653 predicted coding DNA sequences (CDSs), 3618 being located on chromosome 1,  
203 2035 on chromosome 2, 189 on pA<sub>SFn1</sub> and 9 on pB<sub>SFn1</sub>. Chromosome 1 and 2 contain 8 and 0  
204 rRNA, and 97 and 6 tRNA genes, respectively. The genome sequences of the other 14 strains  
205 were partially assembled, with contigs number per strain ranging from 123 to 208. The  
206 approximate genome sizes ranged from 6.1 to 6.4 Mb (Table 1).

207  
208 Comparative genomics revealed pathotype consistent distribution of two plasmids. Like SFn1,  
209 the two other highly pathogenic strains belonging to clade A (A<sub>HP</sub>) carry two plasmids of 250  
210 and 11 kb while the 3 moderately pathogenic strains belonging to clade A (A<sub>MP</sub>) carry only a  
211 large plasmid of 260 kb. Within clade B, the 3 highly pathogenic strains (B<sub>HP</sub>) carry a large  
212 plasmid of 200 kb that is not present in the non pathogenic B<sub>NP</sub> strain. A plasmid was also  
213 evident in the NP strain SOn1; however, it appears only distantly related to the other plasmids  
214 and due to higher sequence divergence, it was not possible to complete its assembly using a  
215 PCR strategy based on the fully assembled genome SFn1. The NP strains SFn118 and  
216 ATCC27043<sup>T</sup> do not carry any plasmid. Finally, the 3 strains isolated in Madagascar carry a  
217 plasmid of 160 kb. The pathogenicity of these strains to the native shrimp *P. monodon* has not  
218 yet been assessed, due to the lack of experimental bio-secure facilities in Madagascar.  
219 However, the culture supernatant of Mada3020 was demonstrated to be toxic for *L.*  
220 *stylirostris* (unpublished results).

221

## 222 Genes differentiating *Vibrio nigripulchritudo* from other *Vibrionaceae*

223 Phylogenetic analysis of concatenated nucleic acid sequences derived from 122 shared  
224 proteins from 133 *Vibrionaceae* genome sequences including *V. nigripulchritudo* confirmed  
225 the cohesive genotypic structure of *V. nigripulchritudo* with relatively little diversity among  
226 genomes and only distant relationship to other, currently characterized *Vibrio* species (Fig.  
227 S1).

228

229 Intraspecific genomic comparisons revealed that 4421 proteins are shared by all sequenced *V.*  
230 *nigripulchritudo* strains. Amongst these proteins, 620 were unique to *V. nigripulchritudo*  
231 genomes (317 on chromosome 1 and 303 on chromosome 2), whereas 486 (242 on  
232 chromosome 1 and 244 on chromosome 2) are shared between 1 to 5 strains of other  
233 *Vibrionaceae* species (Fig. S2). Similarly to other vibrios, chromosome 1 carries a higher  
234 proportion of the core genes whereas chromosome 2 mainly carries accessory genes (Reen et  
235 al 2006).

236

237 A phenotypic characteristic of all *V. nigripulchritudo* strains is the formation of brown  
238 pigmented colonies. We identified two genes encoding a laccase (VIBNI\_B0280) and a  
239 phosphotyrosinase (VIBNI\_B1404), which might be involved in melanin  
240 production (Nosanchuk and Casadevall 2003). The deletion of each gene in the SFn1 strain  
241 decreased the colony pigmentation (Fig.1). However, mutants from the two genes reverted to  
242 full pigmentation over time, suggesting possible functional redundancy and physiological  
243 readjustments. Several attempts to create a double mutant were unsuccessful, suggesting that  
244 the presence of these two genes may be essential for SFn1. It is noteworthy that the lack of a  
245 second usable resistance marker prevents demonstration that the double mutant can only be

246 constructed when one of the two genes is provided in trans. We are currently exploring a  
247 larger panel of antibiotic resistance genes to allow the development of such strategy in the  
248 future.

249

250 Polyketide synthases (PKS-type I) have also been implicated in melanin production (Kroken  
251 et al 2003). Amongst the putative PKS-NRPS identified in SFn1 (Table S4), only the plasmid  
252 gene VIBNI\_pA0055 was annotated as a PKS-type I. However, curing pA<sub>SFn1</sub> plasmid from  
253 SFn1 did not result in a decrease of colony pigmentation (Le Roux et al 2011). In addition,  
254 strains SFn118 and ATCC27043<sup>T</sup> do not contain the large plasmid but still produce a brown  
255 pigment. Taken together these data exclude a role of this gene in melanin production.

256

257 A cluster containing 4 *reb* genes (VIBNI\_A3101, 3105, 3106 and 3107), potentially involved  
258 in the formation of refractile bodies (R-bodies) was identified in all *V. nigripulchritudo* strains  
259 (Fig. S3). R-bodies are coiled proteinaceous ribbons produced by Paramecium endosymbionts  
260 to which they confer a killer phenotype (Pond et al 1989). A similar syntenic group was also  
261 found in the genome sequences of *V. coralliilyticus* (ATCC\_BAA\_450) and *Marinomonas*  
262 *mediterranea* (MMB-1). Within the cluster, 6 other genes encoding proteins of unknown  
263 function, are also conserved between these bacteria. R-bodies were not observed by electron  
264 microscopy in cultured strains of *V. nigripulchritudo* and *V. coralliilyticus* (not shown).  
265 However, when this cluster of genes (VIBNI\_A3108 to 3099) was cloned under the  
266 arabinose-inducible promoter P<sub>BAD</sub> in *E. coli*, TEM observations revealed the presence of R-  
267 body-like structures only in the induced cells (Fig. 2) and resemble closely R-bodies of  
268 *Caedibacter taeniospiralis* and *M. mediterranea* (Schrallhammer et al 2012).

269

270 **Recent emergence of 3 lineages within *V. nigripulchritudo***

271 The phylogenetic relationship based on the core genome of strains included in this study, was  
272 investigated (Fig.3). The main outcome of this analysis was the grouping of all HP isolates,  
273 along with some MP and NP strains, into three clades with very little intra-clade diversity.  
274 Clades A and B contain all the HP strains along with some MP and NP isolates confirming  
275 previous MLST analysis (Goarant et al 2006b) while all isolates from Madagascar form a  
276 distinct clade (termed M). This clade is a sister to clade A while clade B isolates are also more  
277 closely related to the non-pathogenic (NP) isolates SOn1 and SFn118 (Fig. 3).

278

279 To investigate evidence for past recombination among the clades, we compared the  
280 phylogenetic relationships of each core genes (Fig. S4). This confirmed the placement of  
281 isolates into the three clades for 81% of the genes with >75% bootstrap values; however, the  
282 relationship among the 3 clades remains unresolved in 44% of the cases while in 19, 17 and  
283 20%, clade affiliation switches between AM, BM or AB, suggesting that recombination has  
284 shuffled genes between clades.

285

286 Because phylogenetic analysis indicated that the core genome of each clade contains very  
287 little diversity (Fig. 3), we further investigated the percent nucleotide identities of each gene  
288 among all strains within each clade. Within clades A, B and M, the extent of gene identities  
289 ranges from 99.5-100, 98.6-100, 99.8-100%. Overall, this suggests that the clades are of near  
290 clonal and relatively recent origin since neither mutation nor recombination has introduced  
291 much diversity into their core genomes since the emergence from a common ancestral  
292 population.

293

294 **A common, highly recombinogenic plasmid**

295 We identified variants of the pAS<sub>Fn1</sub> plasmid in A<sub>MP</sub>, B<sub>HP</sub>, M strains and SOn1, raising the  
296 possibility that these plasmids are extensively circulating among the *V. nigripulchritudo*  
297 strains and convey some sort of selective advantage to their hosts. The sequence comparison  
298 of these replicons revealed modules that differentiate the plasmids by geography, clade and  
299 pathogenicity (Fig. 4). These plasmids contain a conserved core, consisting of 40 genes, that  
300 are localized within 14 modules (C1-14, Fig. 4). Core genes encode a replication/segregation  
301 system (C1), a polyketide synthase cluster (C3) and a conjugation machinery (C6). A  
302 phylogenetic tree based on a concatenated alignment of all 40 nucleic acid sequences split  
303 *V. nigripulchritudo* plasmids into 5 clades with 100% bootstrap support (Fig. S5). Plasmid-  
304 containing hosts, which were grouped into clade A by chromosomal markers, split into two  
305 clades termed A<sub>HP</sub> and A<sub>MP</sub>, recapitulating different levels of pathogenicity (i.e., highly and  
306 moderately pathogenic). In the plasmid-based comparison, these two clades are a sister to  
307 clade B<sub>HP</sub> and more distantly related to clade M (Fig. S5).

308

309 To determine whether individual gene histories are consistent with the concatenated  
310 phylogeny, analyses were performed for each of the 40 gene sequences of the core plasmid  
311 using SOn1 as an outgroup (not shown). A total of 18/40 trees split the strains in 3 clades with  
312 >75% bootstrap values; 22% show a trifurcation of A, M and B clades, whereas in 0, 17 and  
313 22% of the trees AM, BM, or AB were sister clades, respectively. In addition, 3 topologies  
314 (17, 11 and 11%) cluster all HP strains (A<sub>HP</sub> and B<sub>HP</sub>). Overall, these results suggest that this  
315 plasmid is a highly recombinogenic evolutionary mosaic.

316

### 317 **Horizontal gene acquisition by HP strains within each lineage**

318 Although all HP strains belong to near clonal lineages, none of these lineages exclusively  
319 consist of HP strains. We therefore performed comparative genomic analyses to identify HP-

320 specific genes by first identifying genes present in all  $A_{HP}$  but not  $A_{MP}$  genomes and in all  $B_{HP}$   
321 but not  $B_{NP}$  genomes. These data were then used to define a set of genes common to all HP  
322 strains across lineages where the plasmids were major contributors to the HP-specific gene  
323 pool.

324

325 As noted above (Table 1), the 3  $A_{HP}$  strains carry two plasmids of 250 and 11 kb, previously  
326 demonstrated to be necessary for virulence in SFn1 (Le Roux et al 2011), while the 3  $A_{MP}$   
327 strains contain a replicon of 260 kb closely related to but distinct from the 250 kb plasmid.  
328 The plasmid in  $A_{MP}$  strains lacks 4 modules that are specific to HP strains, (HP1-4), and  
329 instead carries 3 modules that are not present in the  $A_{HP}$  large plasmid (MP1-3) (Fig. 4).  $A_{MP}$ -  
330 specific genes encode several ABC transporters, metabolic proteins, and several proteins of  
331 unknown function. Four modules are  $A_{HP}$ -specific and encode a galactose utilization operon  
332 (HP1), a siderophore ABC transporter, a peptidase (HP2) and genes involved in purine  
333 metabolism (HP3) (Fig. S6). In module HP4, the gene VIBNI\_pA0182 encodes a protein  
334 annotated as a “putative toxin” (described in detail below). Additionally, 25  $A_{HP}$ -specific  
335 genes were located within 3 chromosomal islands (Fig.4). The first and second islands  
336 originated from the integration of a phage within the chromosomal thymidine synthase gene  
337 (*thyA*) (Fig. 5a) and a tRNA gene (Fig. 5b), respectively. These 2 islands encode common  
338 phage-related proteins (e.g., integrase, helicase, relaxase and restriction endonuclease system)  
339 (Fig. 4 CDS in blue), as well as other proteins of unknown function. The third island is an  
340 operon encoding an ABC transporter and a probable hemolysin translocator (HlyD) but does  
341 not contain a hemolytic toxin HlyA gene (Fig. 5c).

342

343 We next identified genes shared by all  $B_{HP}$  genomes but absent from the only  $B_{NP}$  genome  
344 (POn4). This analysis yielded more genes than the  $A_{HP}$  vs  $A_{MP}$  comparison, in part because

345 the large (200 kb) plasmid present in all B<sub>HP</sub> strains is absent in POn4 (Fig. 4). The four A<sub>HP</sub>  
346 modules described above are also found in the B<sub>HP</sub> plasmids and one module is specific to  
347 B<sub>HP</sub> plasmids (Fig. 4, module B). In addition, 130 chromosomal genes are present in all B<sub>HP</sub>  
348 strains but absent from POn4. Over 75% of these genes are located within two large genomic  
349 islands (Fig. 5d, e). The first island, integrated at a tRNA gene, appears to correspond to a  
350 prophage as it contains Cro regulators and phage E structural genes (Fig. 5d). The second  
351 island has the typical structure of an integrative conjugative element (ICE) (Fig. 5e). The 3  
352 B<sub>HP</sub> strains share 52 core genes required for ICE function and differ at several  
353 recombinational hot spots (Fig. 5e, HS1 to 5,) (Wozniak and Waldor 2010).

354

355 Finally, we determined which genes are present in all HP strains (A<sub>HP</sub> and B<sub>HP</sub>) but absent  
356 from the A<sub>MP</sub> and B<sub>NP</sub> strains. This showed that only the 4 plasmid modules (HP1 to 4)  
357 described above are present in and specific to all HP strains (Fig. 4 and S6). Interestingly, the  
358 module HP3 and a part of HP4 were also identified in the plasmid carried by clade M isolates  
359 (Fig. 4) and the culture supernatant of Mada3020 was demonstrated to be toxic for *L.*  
360 *stylirostris* (unpublished results). These results raise the possibility that the HP3 and/or HP4  
361 modules might be particularly important in the emergence of pathogenic strains.

362

### 363 **Nigritoxin, a new toxin sufficient to kill shrimp**

364 As noted above, the HP4 module contains a gene (VIBNI\_pA0182) annotated as a putative  
365 toxin. The VIBNI\_pA0182 gene (2274 bp) encodes a 757 aa polypeptide corresponding to a  
366 putative protein with a theoretical molecular mass of 82,9 kDa. Using InterProScan, PFAM  
367 and Figfam we were unable to identify any particular functional domain within the putative  
368 protein. However, BlastP analysis revealed that aa 28 to 373 of the VIBNI\_pA0182 gene  
369 product shares about 30% identity with a region of a putative toxin named Afp18 described in



370 *Serratia entomophila* (aa 978 to 1329) (Hurst et al 2004) and *Yersinia ruckeri* (aa 937 to  
371 1286) (Fig S7).

372

373 Numerous attempts to disrupt specific genes, including but not limited to VIBNI\_pA0182,  
374 within the large plasmid of SFn1 were unsuccessful. Previous work suggests that genetic tools  
375 that have been successfully used to mutagenize *V. nigripulchritudo* chromosomal and pB<sub>SFn1</sub>  
376 genes (R6K-based suicide vector) interfere in an unknown fashion with pA<sub>SFn1</sub> (Le Roux et  
377 al., 2011). Therefore, to assess whether the putative toxin is linked to *V. nigripulchritudo*-  
378 associated exotoxicity, a plasmid containing VIBNI\_pA0182 under the control of a P<sub>BAD</sub>  
379 promoter was introduced in the avirulent SFn118 strain and the supernatant toxicity was  
380 assayed by injection into shrimp. Supernatant from SFn118 containing a control plasmid,  
381 which carries a gene encoding GFP, was also tested. For this assay, strains were cultivated in  
382 the presence of 1% glucose or 0.2% arabinose in order to either repress or activate the P<sub>BAD</sub>  
383 promoter.

384

385 Strikingly, one day post-injection, 100% mortality occurred in shrimp injected with  
386 supernatant prepared from P<sub>BAD</sub>-VIBNI\_pA0182 transconjugants cultivated in the presence  
387 of arabinose, whereas no significant mortality was observed when animals were challenged  
388 with supernatant prepared from transconjugants grown in the presence of glucose (Fig. 6).  
389 No significant mortality was observed in shrimp injected with P<sub>BAD</sub>-GFP transconjugant  
390 supernatants cultivated in arabinose or in glucose. Additionally, no mortality was observed  
391 following injection of supernatants prepared from *E. coli* carrying the same plasmids (P<sub>BAD</sub>-  
392 VIBNI\_pA0182 or P<sub>BAD</sub>- gfp) in the presence of arabinose or glucose and 100% mortality  
393 occurred in shrimp injected with supernatant prepared from the HP strain SFn1. **These data**  
394 **indicate that VIBNI\_pA0182 contributes to the supernatant toxicity associated with HP**

395 **strains.** Culture supernatants were analyzed using SDS-PAGE electrophoresis in an attempt to  
396 confirm that supernatant toxicity correlates with the amount of the overproduced protein.  
397 However, the complexity of the protein profiles prevented the identification of a single  
398 protein (data not shown).

399

400

## 401 **DISCUSSION**

402

403 We performed HTS-based comparative genome analysis of 16 *V. nigripulchritudo* strains to  
404 explore the genomic diversity and evolutionary history of pathogen containing lineages within  
405 the species and investigate virulence determinants. Phylogenetic analysis based on the *V.*  
406 *nigripulchritudo* core genome suggests a recent emergence of 3 lineages (A, B and M)  
407 containing all **disease** associated HP strains but not exclusively so. HP specific genes acquired  
408 by HGT were evidenced in comparative genomic analyses. These genes include a toxin  
409 homolog, VIBNI\_pA0182, which is encoded by the large plasmid present in all A<sub>HP</sub>, B<sub>HP</sub>, and  
410 M strains. Heterologous expression of VIBNI\_pA0182 conferred a toxic phenotype following  
411 bacterial supernatant injection to shrimp, demonstrating its contribution to the supernatant  
412 toxicity associated with *V. nigripulchritudo*, a specific feature of HP strains.

413

414 Among genes unique to *V. nigripulchritudo* genomes or shared with a few other *Vibrio*  
415 species, two genes encoding a laccase and a phosphotyrosinase could be involved in melanin  
416 production (Nosanchuk and Casadevall 2003). Melanins have been associated with enhanced  
417 virulence in many microorganisms including (i) possession of antioxidant properties that can  
418 attenuate macrophage superoxide production, (ii) interference with the action of endogenous  
419 antimicrobial peptides and (iii) facilitation of ferrous iron uptake through a specific transport

420 system. Whether melanin production plays a role in *V. nigripulchritudo* virulence and/or is  
421 essential for this bacterium are interesting issues that require a complete inactivation of the  
422 pigment synthesis, presently hampered by the lack of available resistance markers to allow  
423 multiple gene knock outs in these multiresistant strains.

424

425 Another feature of *V. nigripulchritudo* genomes, shared with *V. coralliilyticus*, is the presence  
426 of a cluster of **reb** genes potentially involved in the formation of "R-bodies" (Pond et al  
427 1989). These cytoplasmic inclusions were first described in *Caedibacter taeniospirilis*, an  
428 obligate symbiont of paramecia to which they confer a killer phenotype. This enables them to  
429 kill symbiont-free competitors by an unknown mechanism. Recently a study demonstrated  
430 that Reb homologues are widely distributed in members of proteobacteria (Raymann et al  
431 2013). In the present study, although we did not observe R-body structures in cultured strains  
432 of *V. nigripulchritudo* or *V. coralliilyticus*, the induction of the expression of *V.*  
433 *nigripulchritudo* *reb* genes in *E. coli* was sufficient to produce R-bodies suggesting that  
434 specific conditions are required for *reb* gene expression in *Vibrio*. In the future it will be  
435 interesting to perform grazing tests to investigate R-bodies in *V. nigripulchritudo* or *V.*  
436 *coralliilyticus* and test their killing potential in **protists**.

437

438 Several genes related to antibiotic synthesis, particularly PKS-NRPS, are shared by all  
439 sequenced *V. nigripulchritudo* and specific to this species. These enzymes produce  
440 polyketides, a large class of secondary metabolites with a wealth of pharmacologically  
441 important activities, including antimicrobial (Staunton and Weissman 2001). PKS have  
442 recently been implicated in mediation of social interactions among *Vibrio* populations where  
443 the antibiotics are produced in a highly specific manner by few members of the population

444 while all others are resistant, giving rise to the idea that they may play a role in interference  
445 competition between populations (Cordero et al 2012).

446

447 In New Caledonia, *V. nigripulchritudo* strains have been isolated from different sources  
448 (water, sediment, crustaceans) and were previously shown by MLST to be genetically diverse  
449 (Goarant et al 2006b). However, disease associated isolates were grouped into two fairly  
450 monomorphic clades, termed A and B. In the present study, a phylogenetic analysis based on  
451 the concatenated *V. nigripulchritudo* core genome showed that isolates from Madagascar  
452 form a new clade (termed M), which is a sister to clade A and more distantly related to clade  
453 B. However, this phylogeny varies when individual genes are analyzed suggesting that past  
454 recombination has shaped the diversity within these clades. Within clades A, B and M the  
455 strains seem nearly clonal suggesting that the expansion of these lineages is relatively recent  
456 and independent from each other since neither mutation nor recombination has introduced  
457 much diversity into the core genome differentiating the contemporary strains. Importantly,  
458 because clades A and B also contain NP and MP strains, it was possible to identify genes and  
459 genetic elements acquired in the evolution of pathogenesis.

460

461 We previously demonstrated that it is possible to discriminate A<sub>HP</sub> from A<sub>MP</sub> strains based on  
462 the pA<sub>SFn1</sub> plasmid genotype (Le Roux et al 2011). Here, we show that a plasmid sharing a  
463 core set of genes with pA<sub>SFn1</sub> occurs in all pathogenic strains. These plasmids also contain  
464 distinct modules that permit plasmid classification based on geography, clade and  
465 pathogenicity. However, analysis of all genes individually shows, like in the case of the  
466 chromosomes, discordant phylogeny suggesting this plasmid is a highly recombinogenic  
467 evolutionary mosaic. Interestingly, a few genes cluster A<sub>HP</sub> and B<sub>HP</sub>, suggesting

468 recombination events between plasmids carried by strains from different lineages. Further  
469 evidence of such genetic exchanges is the presence of 4 modules found only in A<sub>HP</sub> and B<sub>HP</sub>.

470

471 Within module HP4, a putative toxin-encoding gene, VIBNI\_pA0182, was identified. This  
472 gene encodes a protein with a theoretical MW of 82.9 kDa, which is consistent with the  
473 molecular mass of the toxic proteinaceous compounds previously detected in SFn1 culture  
474 supernatant and shown to be dependent upon the presence of pA<sub>SFn1</sub> (Le Roux et al 2011). In  
475 the present study, we demonstrated that the expression of VIBNI\_pA0182 is sufficient to  
476 induce supernatant toxicity in the NP strain SFn118. We propose to name the product of  
477 VIBNI\_pA0182 “nigritoxin”. Although we do not yet know the pathotype of disease  
478 associated strains from Madagascar, the fact that (i) the supernatant of Mada3020 was  
479 demonstrated to be toxic for *L. stylirostris* and (ii) the nigritoxin gene was also present in  
480 strains belonging to clade M strongly suggest that disease associated isolates from  
481 Madagascar are HP strains. Altogether, our results demonstrate that the nigritoxin is a key  
482 factor of *V. nigripulchritudo* shrimp pathogenesis shared by different lineages.

483

484 The nigritoxin is similar to a putative toxin described in *Serratia entomophila* (Afp18), which  
485 causes amber disease in the grass grub *Costelytra zealandica* (Coleoptera: Scarabaeidae)  
486 (Hurst et al 2004). Larval disease symptoms include cessation of feeding, clearance of the gut,  
487 amber coloration, and eventual death. A 155-kb plasmid, pADAP, is essential for production  
488 of amber disease symptoms. Within this plasmid, a defective prophage (afp) carries the  
489 antifeeding genes (Hurst et al 2004). It has been postulated that the first 16 CDSs of the *afp*  
490 cluster comprise the carriage region that forms an R-type pyocin structure, which functions as  
491 the delivery system for the putative toxin Afp18. As the expression of Afp18 in *E. coli* failed  
492 to cause the antifeeding pathotype, the authors suggested that the R-type pyocin is mandatory

493 for the transport of the toxin to the target site (Hurst et al 2007). Here, aside from the  
494 nigr toxin gene, *afp* encoding phage tail-like structures could not be identified in the HP  
495 genomes. Expression of the nigr toxin in *E. coli* failed to confer a supernatant toxicity but was  
496 sufficient to induce a toxic phenotype in SFn118. As SFn118 does not contain any of the  
497 plasmids, ICE or prophage described as HP-specific traits, the secretion of the nigr toxin may  
498 be regulated by *V. nigr pulchritudo* core genes, but the partners of this process are still  
499 unknown. On the other hand, we cannot rule out the possibility that the recombinant  
500 nigr toxin may accumulate within *E. coli* cells as insoluble inclusion bodies, known to be  
501 devoid of biological activities. In the future, mutagenesis of the nigr toxin gene will allow  
502 identification of the domain(s) of the protein required for processing, export and toxic effects.

503

504 Since microbial pathogenesis is often multifactorial (Finlay and Falkow 1997), it is likely that  
505 other genes than the nigr toxin are involved in the *V. nigr pulchritudo* virulence process. We  
506 previously demonstrated that pB<sub>SFn1</sub> is necessary for the expression of full virulence in SFn1  
507 (Le Roux et al 2011). Other mobile genetic elements (MGEs) discriminate the pathotypes  
508 within each lineage. In addition to pB<sub>SFn1</sub>, 3 chromosomal genomic islands are specific to A<sub>HP</sub>  
509 strains. One island encodes an ABC transporter and a probable hemolysin translocator. In *E.*  
510 *coli*, HlyD has been demonstrated to be essential for the secretion of the RTX hemolytic toxin  
511 HlyA (Pimenta et al 2005). Several putative RTX proteins have been identified in the *V.*  
512 *nigr pulchritudo* genomes but were not restricted to HP strains (not shown). In clade B, two  
513 large genomic islands, a prophage and an ICE, are specific to B<sub>HP</sub> strains. Because no CDSs  
514 annotated in these MGEs can be assigned clearly to a virulence factor, it remains to be seen  
515 whether these HP specific genetic elements are required for virulence.

516

517 Overall, it appears that three lineages containing shrimp pathogens have evolved  
518 independently and fairly recently within *V. nigripulchritudo*. Because these lineages contain  
519 strains with varying pathogenicity as well as non-pathogens, it is possible to reconstruct at  
520 least partially the genomic changes that have accompanied evolution of the HP pathotype. A  
521 key factor appears to have been the acquisition of the nigratoxin, which is located within a  
522 large, mobile plasmid suggesting a plausible avenue for rapid and parallel evolution from  
523 harmless to pathogenic.

524

#### 525 **Acknowledgements :**

526 The authors acknowledge Dr Eliane Chung and Iony Manitra Razanajatovo for providing the  
527 strains isolated in Madagascar, Sophie Le Panse for the TEM analysis, the staff of the station  
528 Ifremer Nouvelle Calédonie, the ABIMS and LABGeM platform for technical support, Dr  
529 Simonetta Gribaldo, Otto X. Cordero and Dr Cyrille Goarant for fruitful discussions and Dr  
530 Brigid Davis for critical reading of the manuscript. The present study has been supported by  
531 the Genoscope (appel d'offre 2007 and 2011) Region Bretagne (SAD Vibrigen 6633), the  
532 Institut Pasteur, the CNRS (UMR3525), the ANR blanc (11-BSV7-023-01 « VIBRIOGEN » )  
533 and Ifremer (DG post doctoral funding).

534

535 Supplementary information is available at ISMEJ's website

536

#### 537 **REFERENCES**

538

539 Austin B (2010). Vibrios as causal agents of zoonoses. *Vet Microbiol* **140**: 310-317.

540

541 Castresana J (2000). Selection of conserved blocks from multiple alignments for their use in  
542 phylogenetic analysis. *Mol Biol Evol* **17**: 540-552.

543

544 Cordero OX, Wildschutte H, Kirkup B, Proehl S, Ngo L, Hussain F *et al* (2012). Ecological  
545 populations of bacteria act as socially cohesive units of antibiotic production and resistance.  
546 *Science* **337**: 1228-1231.

547

548 Daubin V, Gouy M, Perriere G (2002). A phylogenomic approach to bacterial phylogeny:  
549 evidence of a core of genes sharing a common history. *Genome research* **12**: 1080-1090.

550

551 Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high  
552 throughput. *Nucleic acids research* **32**: 1792-1797.

553

554 Finlay BB, Falkow S (1997). Common themes in microbial pathogenicity revisited.  
555 *Microbiology and molecular biology reviews : MMBR* **61**: 136-169.

556

557 Goarant C, Ansquer D, Herlin J, Domalain F, Imbert F, de Decker S (2006a). "Summer  
558 Syndrome" in *Litopenaeus stylirostris* in New Caledonia: Pathology and epidemiology of the  
559 etiological agent, *Vibrio nigripulchritudo*. *Aquaculture* **253**: 105-113.

560

561 Goarant C, Reynaud Y, Ansquer D, de Decker S, Saulnier D, le Roux F (2006b). Molecular  
562 epidemiology of *Vibrio nigripulchritudo*, a pathogen of cultured penaeid shrimp (*Litopenaeus*  
563 *stylirostris*) in New Caledonia. *Systematic and applied microbiology* **29**: 570-580.

564



565 Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010). New  
566 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the  
567 performance of PhyML 3.0. *Syst Biol* **59**: 307-321.

568

569 Hoffmann M, Monday SR, Allard MW, Strain EA, Whittaker P, Naum M *et al* (2012). *Vibrio*  
570 *caribbeanicus* sp. nov., isolated from the marine sponge *Scleritoderma cyanea*. *International*  
571 *journal of systematic and evolutionary microbiology* **62**: 1736-1743.

572

573 Huerta-Cepas J, Dopazo J, Gabaldon T (2010). ETE: a python Environment for Tree  
574 Exploration. *BMC Bioinformatics* **11**: 24.

575

576 Hurst MR, Glare TR, Jackson TA (2004). Cloning *Serratia entomophila* antifeeding genes--a  
577 putative defective prophage active against the grass grub *Costelytra zealandica*. *Journal of*  
578 *bacteriology* **186**: 5116-5128.

579

580 Hurst MR, Beard SS, Jackson TA, Jones SM (2007). Isolation and characterization of the  
581 *Serratia entomophila* antifeeding prophage. *FEMS microbiology letters* **270**: 42-48.

582

583 Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG (2003). Phylogenomic analysis of  
584 type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc Natl Acad Sci*  
585 *U S A* **100**: 15670-15675.

586

587 Le Roux F, Binesse J, Saulnier D, Mazel D (2007). Construction of a *Vibrio splendidus*  
588 mutant lacking the metalloprotease gene *vsm* by use of a novel counterselectable suicide  
589 vector. *Applied and environmental microbiology* **73**: 777-784.

590

591 Le Roux F, Labreuche Y, Davis BM, Iqbal N, Mangenot S, Goarant C *et al* (2011). Virulence  
592 of an emerging pathogenic lineage of *Vibrio nigripulchritudo* is dependent on two plasmids.  
593 *Environmental microbiology* **13**: 296-306.

594

595 Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA *et al* (2011).  
596 antiSMASH: rapid identification, annotation and analysis of secondary metabolite  
597 biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic acids research*  
598 **39**: W339-346.

599

600 Nosanchuk JD, Casadevall A (2003). The contribution of melanin to microbial pathogenesis.  
601 *Cell Microbiol* **5**: 203-223.

602

603 Pimenta AL, Racher K, Jamieson L, Blight MA, Holland IB (2005). Mutations in HlyD, part  
604 of the type 1 translocator for hemolysin secretion, affect the folding of the secreted toxin.  
605 *Journal of bacteriology* **187**: 7471-7480.

606

607 Pond FR, Gibson I, Lalucat J, Quackenbush RL (1989). R-body-producing bacteria.  
608 *Microbiological reviews* **53**: 25-67.

609

610 Raymann K, Bobay LM, Doak TG, Lynch M, Gribaldo S (2013). A genomic survey of reb  
611 homologs suggests widespread occurrence of R-bodies in proteobacteria. *G3 (Bethesda)* **3**:  
612 505-516.

613

614 Reen FJ, Almagro-Moreno S, Ussery D, Boyd EF (2006). The genomic code: inferring  
615 Vibrionaceae niche specialization. *Nat Rev Microbiol* **4**: 697-704.  
616

617 Reynaud Y, Saulnier D, Mazel D, Goarant C, Le Roux F (2008). Correlation between  
618 detection of a plasmid and high-level virulence of *Vibrio nigripulchritudo*, a pathogen of the  
619 shrimp *Litopenaeus stylirostris*. *Applied and environmental microbiology* **74**: 3038-3047.  
620

621 Sakai T, Hirae T, Yuasa K, Kamaishi T, Matsuyama T, Miwa S *et al* (2007). Mass mortality  
622 of cultured kuruma prawn *Peneaus japonicus* caused by *Vibrio nigripulchritudo*. *Fish*  
623 *Pathology* **42**: 141-147.  
624

625 Satchell KJ (2007). MARTX, multifunctional autoprocessing repeats-in-toxin toxins. *Infect*  
626 *Immun* **75**: 5079-5084.  
627

628 Schrallhammer M, Galati S, Altenbuchner J, Schweikert M, Gortz HD, Petroni G (2012).  
629 Tracing the role of R-bodies in the killer trait: Absence of toxicity of R-body producing  
630 recombinant *E. coli* on paramecia. *Eur J Protistol.*  
631

632 Staunton J, Weissman KJ (2001). Polyketide biosynthesis: a millennium review. *Nat Prod*  
633 *Rep* **18**: 380-416.  
634

635 Thompson FL, Iida T, Swings J (2004). Biodiversity of vibrios. *Microbiol Mol Biol Rev* **68**:  
636 403-431, table of contents.  
637

638 Vallenet D, Engelen S, Mornico D, Cruveiller S, Fleury L, Lajus A *et al* (2009). MicroScope:  
639 a platform for microbial genome annotation and comparative genomics. *Database (Oxford)*  
640 **2009**: bap021.

641

642 Waterfield NR, Wren BW, Ffrench-Constant RH (2004). Invertebrates as a source of  
643 emerging human pathogens. *Nat Rev Microbiol* **2**: 833-841.

644

645 Wozniak RA, Waldor MK (2010). Integrative and conjugative elements: mosaic mobile  
646 genetic elements enabling dynamic lateral gene flow. *Nat Rev Microbiol* **8**: 552-563.

647

648 Zaitseva J, Jenewein S, Oswald C, Jumpertz T, Holland IB, Schmitt L (2005). A molecular  
649 understanding of the catalytic cycle of the nucleotide-binding domain of the ABC transporter  
650 HlyB. *Biochem Soc Trans* **33**: 990-995.

651

652 Zerbino DR, Birney E (2008). Velvet: algorithms for de novo short read assembly using de  
653 Bruijn graphs. *Genome research* **18**: 821-829.

654

655 Ziemert N, Podell S, Penn K, Badger JH, Allen E, Jensen PR (2012). The natural product  
656 domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary  
657 metabolite gene diversity. *PloS one* **7**: e34064.

658

659

**Table 1:** *Vibrio nigripulchritudo* strains included in the study

## LEGENDS TO FIGURES

**Figure 1:** Colony pigmentation of *V. nigripulchritudo* strains SFn1 (a),  $\Delta$  *VIBNI\_B0280* (laccase) (b) and  $\Delta$  *VIBNI\_B1404* (phosphotyrosinase) (c) derivatives grown on marine agar supplemented with 1% glycerol.

**Figure 2:** Transmission electron micrographs of bacterial cells. (a) not induced (b) induced expression of the *reb* genes cluster (*VIBNI\_A3108* to *3099*) cloned under the arabinose-inducible promoter  $P_{BAD}$  in *E. coli* (GV691) and producing several recombinant R-bodies (white arrows).

**Figure 3:** Phylogenetic analysis based on concatenated alignments of nucleic acid sequences of 4421 core genes from 16 *V. nigripulchritudo* strains (ATCC 27043<sup>T</sup> as outgroup). Trees were built by the Maximum-Likelihood method based on sequences aligned using Muscle. Branch lengths are drawn to scale and are proportional to the number of nucleotide changes. Numbers at each node represent the percentage value given by bootstrap analysis of 100 replicates. The pathotype of each strain, when tested, is indicated in parentheses.

**Figure 4:** Plasmid diversity among the different pathotypes. Modules are defined by gene(s) found in all plasmids (core C1 to 14 in gray); genes that are specific to strains isolated in New Caledonia (NC1 to 5, in green) or in Madagascar (M1 to 4, in turquoise); genes that are specific to clade A (A1 to 7, in pink), clade B (B in blue) or shared by clades M and B (M/B in black); genes that are specific to MP (MP1 to 3, in yellow) or specific to all HP (HP1 to 4,

in red). The SFn1 strain-specific gene labels are VIBNI\_: C1 [pA0002-pA0019], C2 [pA0028-pA0045], C3 [pA0051-pA0066], C4 [pA0088-pA0089], C5 [pA0113], C6 [pA0118-pA0134], C7 [pA0150], C8 [pA0169-pA0170], C9 [pA0178-pA0180], C10 [pA0198-pA0205], C11 [pA0207-pA0208], C12 [pA0210-pA0214], C13 [pA0222], C14 [pA0225], NC1 [pA0020-pA0027], NC2 [pA0047-pA0049], NC3 [pA0115-pA0117], NC4 [pA0135-pA0140], NC5 [pA0167-pA0168], A1 [pA0068-pA0087], A2 [pA0090-pA0112], A3 [pA0206], A4 [pA0209], A5 [pA0215-pA0221], A6 [pA0223-pA0224], A7 [pA0226-pA0239], HP1 [pA0141-pA0149], HP2 [pA0151-pA0166], HP3 [pA0171-pA0177], HP4 [pA0181-pA0196]. The BLFn1 strain-specific gene labels are: VIBNIBLFn1\_: MP1 [p0128-p0145], MP2 [p0147-pA0154], MP1 [p0164-p0188]. The strain AM115 gene labels were VIBNIAM115\_: B [p0058-p0066], M/B [p0155-p0165]. The strain Mada3020 gene labels were: VIBNIMada3020\_: M1 [p0052-p0057], M2 [p0061-p0069], M3 [p0088-p0089], M4 [p0091-p0099].

**Figure 5:** Structure of chromosomal genomic islands present in all  $A_{HP}$  but absent in all  $A_{MP}$  strains (**a** [VIBNI\_A0333-0352], **b** [VIBNI\_B1043-1063] and **c** [VIBNI\_B1464-1471]) or present in all  $B_{HP}$  and absent in the  $B_{NP}$  strain (**d** [VIBNIAM115\_840001-84152], **e** [VIBNIAM115\_1620066-1620139]). CDSs in blue indicate a phage origin (A, B and D). CDSs in brown indicate a putative Hly translocator complex (**c**). CDSs in white encode proteins with unknown function. CDSs in black indicate the MGE boundaries, labeled “hypo” when encoding a hypothetical protein. “End” indicates the end of a contig. Within the ICE, recombination hot spot (HS) genes are indicated in pink (HS5), turquoise (HS1), orange (HS2), green (HS4), and yellow (HS3). The HS contents are shared by one or two other  $B_{HP}$  strains (subscript). Regulator genes are indicated by a star.

**Figure 6 :** Shrimp mortality in response to injection of bacterial culture supernatants. Shrimp, *L. stylirostris* (n=5), were injected with 100  $\mu$ L of bacterial supernatants prepared from overnight cultures. Survival was assessed after 24 hours. Experiments were conducted three times. Shrimp were intramuscularly injected with: supernatants prepared from the *V. nigripulchritudo* HP strain SFn1 and avirulent strain SFn118; from the  $P_{BAD}$ -VIBNI\_pA0182 transconjugant of SFn118 cultivated in the presence of arabinose or glucose; from the  $P_{BAD}$ -GFP transconjugant of SFn118 cultivated in the presence of arabinose or glucose; from *E. coli* carrying the  $P_{BAD}$ -VIBNI\_pA0182 plasmid cultivated in the presence of arabinose or glucose; from *E. coli* carrying the  $P_{BAD}$ -GFP plasmid cultivated in the presence of arabinose or glucose; with medium Zobell or LB as a control.