The ISME Journal
October 2013, Volume 7, Pages 1985-1996
http://dx.doi.org/10.1038/ismej.2013.90

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

David Goudenège[1], Yannick Labreuche[1], Evelyne Krin[2,3], Dominique Ansquer[4], Sophie Mangenot[5], Alexandra Calteau[5,6], Claudine Médigue[5,6], Didier Mazel[2,3], Martin F Polz[7] and Frédérique Le Roux[1]

[1] Equipe Émergente Ifremer-UPMC Génomique des Vibrio, Roscoff, France
[2]Département Génomes et Génétique, Institut Pasteur, Unité Plasticité du Génome Bactérien, Paris, France
[3] CNRS UMR3525,Paris, France
[4] Département Lagons, Ecosystèmes et Aquaculture Durable en Nouvelle-Calédonie, IFREMER, Nouméa, New Caledonia
[5] CEA, Institut de Génomique, Genoscope, Evry, France
[6] CEA, Institut de Génomique, Génoscope & CNRS UMR8030 Laboratoire de Génomique Comparative, Evry, France
[7]Parsons Laboratory for Environmental Science and Engineering, MIT, Cambridge, MA, USA

*: Corresponding author : Frédérique Le Roux, email address : frederique.le-roux@sb-roscoff.fr

Abstract:

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

Keywords: horizontal gene transfer ; nigritoxin; phylogeny ; shrimp; virulence; vibrio

INTRODUCTION

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

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

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

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

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

Here, we further explore the genomic diversity and evolutionary history of pathogen-containing lineages of *V. nigripulchritudo*. We ask what types of genetic elements (i) differentiate each lineage and (ii) are shared among the HP strains within and across lineages. To address these questions, we performed high throughput sequencing (HTS)-based comparative genome analysis. Our phylogenetic analyses suggest that recent differential adaptation of lineages was followed by horizontal acquisition of genetic material, which distinguishes pathotypes within clades. A unique region carried by a large plasmid is specific to all HP strains and contains a gene encoding a putative toxin. Expression of this gene in a NP strain results in production of toxic culture supernatant.

MATERIALS AND METHODS

Bacterial strains, plasmids and culture conditions

The *V. nigripulchritudo* strains used for the genomic analyses are described in Table 1. Other bacterial strains and plasmids are described in Tables S1 and S2. *V. nigripulchritudo* strains were grown in Luria-Bertani (LB) or LB-agar (LBA) + NaCl 0.5M, marine broth (MB) or marine agar (MA) at 30°C. *E. coli* strains were grown in LB or on LBA at 37°C.

Spectinomycin (100 μg/ml), thymidine (0.3 mM) and diaminopimelate (0.3 mM) were added as supplements when necessary. Induction of the P_{BAD} promoter was achieved by the addition of 0.2% L-arabinose to the growth media, and conversely, was repressed by the addition of 1% D-glucose.

Genome sequencing, assembly and annotation

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

***In silico* analyses**

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

was created to perform comparative genomic and phylogenomic analyses. Orthologous proteins were defined as reciprocal best hit proteins with 80% MaxLrap and a minimum of 30% and 60% identity cutoff was used for intra- and inter-species analysis, respectively (Daubin et al 2002). The nucleic acid sequences were aligned using Muscle (Edgar 2004) and filtered by Gblocks (Castresana 2000). Phylogenetic trees were built using the parallel version of PhyML applied to Maximum-likelihood algorithm and GTR model as parameters (NNIs, γ4, invariant site) (Guindon et al 2010). Reliability was assessed by the bootstrap method with 100 replicates. Tree topologies were analyzed using the ETE2 Python library (Huerta-Cepas et al 2010).

Vector construction and mutagenesis

Alleles carrying an internal deletion were generated *in vitro* using a two-step PCR construction method (supplementary methods and Table S3) and cloned into an R6K γ-*ori*-based suicide vector that encodes the *ccdB* toxin gene under the control of an arabinose-inducible and glucose-repressible promoter, P_{BAD} (Le Roux et al 2007). Matings between *E. coli* and *V. nigripulchritudo* were performed at 30°C as described previously (Le Roux et al 2011). Selection of the plasmid-borne drug marker ($Spec^R$) resulted in integration of the entire plasmid in the chromosome by a single crossover. Elimination of the plasmid backbone resulting from a second recombination step was selected by arabinose induction of the *ccdB* toxin gene. Mutants were screened by PCR.

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

Transmission electron microscopy

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

Experimental challenges

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

RESULTS

General features of the *V. nigripulchritudo* genomes

The genome of strain SFn1 was closed and annotated manually. It consists of two circular chromosomes of 4.1 (chromosome 1) and 2.2 (chromosome 2) Mb with an average percent G+C content of 45.9 and 45.5, respectively (Table 1). As described previously, SFn1 also carries two plasmids of 250 (pA_{SFn1}) and 11 (pB_{SFn1}) kb (Le Roux et al 2011). There is a total of 5653 predicted coding DNA sequences (CDSs), 3618 being located on chromosome 1,

2035 on chromosome 2, 189 on pA_{SFn1} and 9 on pB_{SFn1}. Chromosome 1 and 2 contain 8 and 0 rRNA, and 97 and 6 tRNA genes, respectively. The genome sequences of the other 14 strains were partially assembled, with contigs number per strain ranging from 123 to 208. The approximate genome sizes ranged from 6.1 to 6.4 Mb (Table 1).

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

Genes differentiating *Vibrio nigripulchritudo* from other *Vibrionaceae*

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

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

A phenotypic characteristic of all *V. nigripulchritudo* strains is the formation of brown pigmented colonies. We identified two genes encoding a laccase (VIBNI_B0280) and a phosphotyrosinase (VIBNI_B1404), which might be involved in melanin production (Nosanchuk and Casadevall 2003). The deletion of each gene in the SFn1 strain decreased the colony pigmentation (Fig.1). However, mutants from the two genes reverted to full pigmentation over time, suggesting possible functional redundancy and physiological readjustments. Several attempts to create a double mutant were unsuccessful, suggesting that the presence of these two genes may be essential for SFn1. It is noteworthy that the lack of a second usable resistance marker prevents demonstration that the double mutant can only be constructed when one of the two genes is provided in trans. We are currently exploring a larger panel of antibiotic resistance genes to allow the development of such strategy in the future.

Polyketide synthases (PKS-type I) have also been implicated in melanin production (Kroken et al 2003). Amongst the putative PKS-NRPS identified in SFn1 (Table S4), only the plasmid gene VIBNI_pA0055 was annotated as a PKS-type I. However, curing pA_{SFn1} plasmid from

SFn1 did not result in a decrease of colony pigmentation (Le Roux et al 2011). In addition, strains SFn118 and ATCC27043^T do not contain the large plasmid but still produce a brown pigment. Taken together these data exclude a role of this gene in melanin production.

A cluster containing 4 *reb* genes (VIBNI_A3101, 3105, 3106 and 3107), potentially involved in the formation of refractile bodies (R-bodies) was identified in all *V. nigripulchritudo* strains (Fig. S3). R-bodies are coiled proteinaceous ribbons produced by Paramecium endosymbionts to which they confer a killer phenotype (Pond et al 1989). A similar syntenic group was also found in the genome sequences of *V. coralliilyticus* (ATCC_BAA_450) and *Marinomonas mediterranea* (MMB-1). Within the cluster, 6 other genes encoding proteins of unknown function, are also conserved between these bacteria. R-bodies were not observed by electron microscopy in cultured strains of *V. nigripulchritudo* and *V. coralliilyticus* (not shown). However, when this cluster of genes (VIBNI_A3108 to 3099) was cloned under the arabinose-inducible promoter P_{BAD} in *E. coli,* TEM observations revealed the presence of R-body-like structures only in the induced cells (Fig. 2) and resemble closely R-bodies of *Caedibacter taeniospiralis* and *M. mediterranea* (Schrallhammer et al 2012).

Recent emergence of 3 lineages within *V. nigripulchritudo*

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

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

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

A common, highly recombinogenic plasmid

We identified variants of the pA_{SFn1} plasmid in A_{MP}, B_{HP}, M strains and SOn1, raising the possibility that these plasmids are extensively circulating among the *V. nigripulchritudo* strains and convey some sort of selective advantage to their hosts. The sequence comparison of these replicons revealed modules that differentiate the plasmids by geography, clade and pathogenicity (Fig. 4). These plasmids contain a conserved core, consisting of 40 genes, that are localized within 14 modules (C1-14, Fig. 4). Core genes encode a replication/segregation system (C1), a polyketide synthase cluster (C3) and a conjugation machinery (C6). A phylogenetic tree based on a concatenated alignment of all 40 nucleic acid sequences split

V. nigripulchritudo plasmids into 5 clades with 100% bootstrap support (Fig. S5). Plasmid-containing hosts, which were grouped into clade A by chromosomal markers, split into two clades termed A_{HP} and A_{MP}, recapitulating different levels of pathogenicity (i.e., highly and moderately pathogenic). In the plasmid-based comparison, these two clades are a sister to clade B_{HP} and more distantly related to clade M (Fig. S5).

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

Horizontal gene acquisition by HP strains within each lineage

Although all HP strains belong to near clonal lineages, none of these lineages exclusively consist of HP strains. We therefore performed comparative genomic analyses to identify HP-specific genes by first identifying genes present in all A_{HP} but not A_{MP} genomes and in all B_{HP} but not B_{NP} genomes. These data were then used to define a set of genes common to all HP strains across lineages where the plasmids were major contributors to the HP-specific gene pool.

As noted above (Table 1), the 3 A_{HP} strains carry two plasmids of 250 and 11 kb, previously demonstrated to be necessary for virulence in SFn1 (Le Roux et al 2011), while the 3 A_{MP} strains contain a replicon of 260 kb closely related to but distinct from the 250 kb plasmid.

The plasmid in A_{MP} strains lacks 4 modules that are specific to HP strains, (HP1-4), and instead carries 3 modules that are not present in the A_{HP} large plasmid (MP1-3) (Fig. 4). A_{MP}-specific genes encode several ABC transporters, metabolic proteins, and several proteins of unknown function. Four modules are A_{HP}-specific and encode a galactose utilization operon (HP1), a siderophore ABC transporter, a peptidase (HP2) and genes involved in purine metabolism (HP3) (Fig. S6). In module HP4, the gene VIBNI_pA0182 encodes a protein annotated as a "putative toxin" (described in detail below). Additionally, 25 A_{HP}-specific genes were located within 3 chromosomal islands (Fig.4). The first and second islands originated from the integration of a phage within the chromosomal thymidine synthase gene (*thyA*) (Fig. 5a) and a tRNA gene (Fig. 5b), respectively. These 2 islands encode common phage-related proteins (e.g., integrase, helicase, relaxase and restriction endonuclease system) (Fig. 4 CDS in blue), as well as other proteins of unknown function. The third island is an operon encoding an ABC transporter and a probable hemolysin translocator (HlyD) but does not contain a hemolytic toxin HlyA gene (Fig. 5c).

We next identified genes shared by all B_{HP} genomes but absent from the only B_{NP} genome (POn4). This analysis yielded more genes than the A_{HP} vs A_{MP} comparison, in part because the large (200 kb) plasmid present in all B_{HP} strains is absent in POn4 (Fig. 4). The four A_{HP} modules described above are also found in the B_{HP} plasmids and one module is specific to B_{HP} plasmids (Fig. 4, module B). In addition, 130 chromosomal genes are present in all B_{HP} strains but absent from POn4. Over 75% of these genes are located within two large genomic islands (Fig. 5d, e). The first island, integrated at a tRNA gene, appears to correspond to a prophage as it contains Cro regulators and phage E structural genes (Fig. 5d). The second island has the typical structure of an integrative conjugative element (ICE) (Fig. 5e). The 3

B_{HP} strains share 52 core genes required for ICE function and differ at several recombinational hot spots (Fig. 5e, HS1 to 5,) (Wozniak and Waldor 2010).

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

Nigritoxin, a new toxin sufficient to kill shrimp

As noted above, the HP4 module contains a gene (VIBNI_pA0182) annotated as a putative toxin. The VIBNI_pA0182 gene (2274 bp) encodes a 757 aa polypeptide corresponding to a putative protein with a theoretical molecular mass of 82,9 kDa. Using InterProScan, PFAM and Figfam we were unable to identify any particular functional domain within the putative protein. However, BlastP analysis revealed that aa 28 to 373 of the VIBNI_pA0182 gene product shares about 30% identity with a region of a putative toxin named Afp18 described in *Serratia entomophila* (aa 978 to 1329) (Hurst et al 2004) and *Yersinia ruckeri* (aa 937 to 1286) (Fig S7).

Numerous attempts to disrupt specific genes, including but not limited to VIBNI_pA0182, within the large plasmid of SFn1 were unsuccessful. Previous work suggests that genetic tools that have been successfully used to mutagenize *V. nigripulchritudo* chromosomal and pB_{SFn1} genes (R6K-based suicide vector) interfere in an unknown fashion with pA_{SFn1} (Le Roux et

al., 2011). Therefore, to assess whether the putative toxin is linked to *V. nigripulchritudo*-associated exotoxicity, a plasmid containing VIBNI_pA0182 under the control of a P_{BAD} promoter was introduced in the avirulent SFn118 strain and the supernatant toxicity was assayed by injection into shrimp. Supernatant from SFn118 containing a control plasmid, which carries a gene encoding GFP, was also tested. For this assay, strains were cultivated in the presence of 1% glucose or 0.2% arabinose in order to either repress or activate the P_{BAD} promoter.

Strikingly, one day post-injection, 100% mortality occurred in shrimp injected with supernatant prepared from P_{BAD}-VIBNI_pA0182 transconjuguants cultivated in the presence of arabinose, whereas no significant mortality was observed when animals were challenged with supernatant prepared from transconjuguants grown in the presence of glucose (Fig. 6). No significant mortality was observed in shrimp injected with P_{BAD}-GFP transconjuguant supernatants cultivated in arabinose or in glucose. Additionally, no mortality was observed following injection of supernatants prepared from *E. coli* carrying the same plasmids (P_{BAD}-VIBNI_pA0182 or P_{BAD}- gfp) in the presence of arabinose or glucose and 100% mortality occurred in shrimp injected with supernatant prepared from the HP strain SFn1. These data indicate that VIBNI_pA0182 contributes to the supernatant toxicity associated with HP strains. Culture supernatants were analyzed using SDS-PAGE electrophoresis in an attempt to confirm that supernatant toxicity correlates with the amount of the overproduced protein. However, the complexity of the protein profiles prevented the identification of a single protein (data not shown).

DISCUSSION

We performed HTS-based comparative genome analysis of 16 *V. nigripulchritudo* strains to explore the genomic diversity and evolutionary history of pathogen containing lineages within the species and investigate virulence determinants. Phylogenetic analysis based on the *V. nigripulchritudo* core genome suggests a recent emergence of 3 lineages (A, B and M) containing all disease associated HP strains but not exclusively so. HP specific genes acquired by HGT were evidenced in comparative genomic analyses. These genes include a toxin homolog, VIBNI_pA0182, which is encoded by the large plasmid present in all A_{HP}, B_{HP}, and M strains. Heterologous expression of VIBNI_pA0182 conferred a toxic phenotype following bacterial supernatant injection to shrimp, demonstrating its contribution to the supernatant toxicity associated with *V. nigripulchritudo*, a specific feature of HP strains.

Among genes unique to *V. nigripulchritudo* genomes or shared with a few other *Vibrio* species, two genes encoding a laccase and a phosphotyrosinase could be involved in melanin production (Nosanchuk and Casadevall 2003). Melanins have been associated with enhanced virulence in many microorganisms including (i) possession of antioxidant properties that can attenuate macrophage superoxide production, (ii) interference with the action of endogenous antimicrobial peptides and (iii) facilitation of ferrous iron uptake through a specific transport system. Whether melanin production plays a role in *V. nigripulchritudo* virulence and/or is essential for this bacterium are interesting issues that require a complete inactivation of the pigment synthesis, presently hampered by the lack of available resistance markers to allow multiple gene knock outs in these multiresistant strains.

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

Several genes related to antibiotic synthesis, particularly PKS-NRPS, are shared by all sequenced *V. nigripulchritudo* and specific to this species. These enzymes produce polyketides, a large class of secondary metabolites with a wealth of pharmacologically important activities, including antimicrobial (Staunton and Weissman 2001). PKS have recently been implicated in mediation of social interactions among *Vibrio* populations where the antibiotics are produced in a highly specific manner by few members of the population while all others are resistant, giving rise to the idea that they may play a role in interference competition between populations (Cordero et al 2012).

In New Caledonia, *V. nigripulchritudo* strains have been isolated from different sources (water, sediment, crustaceans) and were previously shown by MLST to be genetically diverse (Goarant et al 2006b). However, disease associated isolates were grouped into two fairly

monomorphic clades, termed A and B. In the present study, a phylogenetic analysis based on the concatenated *V. nigripulchritudo* core genome showed that isolates from Madagascar form a new clade (termed M), which is a sister to clade A and more distantly related to clade B. However, this phylogeny varies when individual genes are analyzed suggesting that past recombination has shaped the diversity within these clades. Within clades A, B and M the strains seem nearly clonal suggesting that the expansion of these lineages is relatively recent and independent from each other since neither mutation nor recombination has introduced much diversity into the core genome differentiating the contemporary strains. Importantly, because clades A and B also contain NP and MP strains, it was possible to identify genes and genetic elements acquired in the evolution of pathogenesis.

We previously demonstrated that it is possible to discriminate A_{HP} from A_{MP} strains based on the pA_{SFn1} plasmid genotype (Le Roux et al 2011). Here, we show that a plasmid sharing a core set of genes with pA_{SFn1} occurs in all pathogenic strains. These plasmids also contain distinct modules that permit plasmid classification based on geography, clade and pathogenicity. However, analysis of all genes individually shows, like in the case of the chromosomes, discordant phylogeny suggesting this plasmid is a highly recombinogenic evolutionary mosaic. Interestingly, a few genes cluster A_{HP} and B_{HP}, suggesting recombination events between plasmids carried by strains from different lineages. Further evidence of such genetic exchanges is the presence of 4 modules found only in A_{HP} and B_{HP}.

Within module HP4, a putative toxin-encoding gene, VIBNI_pA0182, was identified. This gene encodes a protein with a theoretical MW of 82.9 kDa, which is consistent with the molecular mass of the toxic proteinaceous compounds previously detected in SFn1 culture supernatant and shown to be dependent upon the presence of pA_{SFn1} (Le Roux et al 2011). In

the present study, we demonstrated that the expression of VIBNI_pA0182 is sufficient to induce supernatant toxicity in the NP strain SFn118. We propose to name the product of VIBNI_pA0182 “nigritoxin”. Althought we do not yet know the pathotype of disease associated strains from Madagascar, the fact that (i) the supernatant of Mada3020 was demonstrated to be toxic for *L. stylirostris* and (ii) the nigritoxin gene was also present in strains belonging to clade M strongly suggest that disease associated isolates from Madagascar are HP strains. Altogether, our results demonstrate that the nigritoxin is a key factor of *V. nigripulchritudo* shrimp pathogenesis shared by different lineages.

The nigritoxin is similar to a putative toxin described in *Serratia entomophila* (Afp18), which causes amber disease in the grass grub *Costelytra zealandica* (Coleoptera: Scarabaeidae) (Hurst et al 2004). Larval disease symptoms include cessation of feeding, clearance of the gut, amber coloration, and eventual death. A 155-kb plasmid, pADAP, is essential for production of amber disease symptoms. Within this plasmid, a defective prophage (afp) carries the antifeeding genes (Hurst et al 2004). It has been postulated that the first 16 CDSs of the *afp* cluster comprise the carriage region that forms an R-type pyocin structure, which functions as the delivery system for the putative toxin Afp18. As the expression of Afp18 in *E. coli* failed to cause the antifeeding pathotype, the authors suggested that the R-type pyocin is mandatory for the transport of the toxin to the target site (Hurst et al 2007). Here, aside from the nigritoxin gene, *afp* encoding phage tail-like structures could not be identified in the HP genomes. Expression of the nigritoxin in *E. coli* failed to confer a supernatant toxicity but was sufficient to induce a toxic phenotype in SFn118. As SFn118 does not contain any of the plasmids, ICE or prophage described as HP-specific traits, the secretion of the nigritoxin may be regulated by *V. nigripulchritudo* core genes, but the partners of this process are still unknown. On the other hand, we cannot rule out the possibility that the recombinant

nigritoxin may accumulate within *E. coli* cells as insoluble inclusion bodies, known to be devoid of biological activities. In the future, mutagenesis of the nigritoxin gene will allow identification of the domain(s) of the protein required for processing, export and toxic effects.

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

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

Acknowledgements :

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

Supplementary information is available at ISMEJ's website

The manuscript has been seen and approved by all of the authors. There is no conflict of interest. The material represents an original result and has not been submitted for publication elsewhere.

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Strains	Origin	Pathogenicity	Clade	Chromosomes (Mb)	Plasmid (kb)	CDSs chr	CDSs pA_	CDSs pB_	Accession Numbers
SFn1	Summer syndrome, moribund *L. stylirostris*, March 2000, New Caledonia	HP	A	4,11+2,21	250+11	5653 (3618+2035)	189	9	FO203526 VIBNI_A (Chr1), FO203527 VIBNI_B (Chr2), NC_015156 (pA), NC_010733 (pB)
SFn27	Summer syndrome, sediment pore water, diseased pond, December 2000, New Caledonia	HP	A	6,22	250+11	5655	191	9	CAOB01000001-CAOB01000122
SFn135	Summer syndrome, growout pond water, diseased pond, December 2002, New caledonia	HP	A	6,22	250+11	5672	191	9	CAOD01000001-CAOD01000125
Wn13	Moribund *L. stylirostris* hemoculture, November 2002, New Caledonia	MP	A	6,18	260	5611	197	-	CAOG01000001-CAOG01000119
ENn2	Moribund *L. stylirostris* hemoculture, June 2000, New Caledonia	MP	A	6,2	260	5631	199	-	CANV01000001-CANV01000117
BLFn1	Moribund *L. stylirostris* hemoculture, March 2001, New Caledonia	MP	A	6,16	260	5578	196	-	CANU01000001-CANU01000120
FTn2	Healthy *L. stylirostris*, healthy pond, healthy farm, March 2003, New Caledonia	HP	B	6,43	200	6080	147	-	CANW01000001-CANW01000168
SO65	Syndrome 93, moribund *L. stylirostris* hemoculture, May 1995, New Caledonia	HP	B	6,17	200	5768	151	-	CAOE01000001-CAOE01000181
AM115	Syndrome 93, moribund *L. stylirostris* hemoculture, May 1995, New Caledonia	HP	B	6,17	200	5927	149	-	CANT01000001-CANT01000208
POn4	Healthy *L. stylirostris* hemoculture, healthy pond, healthy farm, April 2002, New Caledonia	NP	B	6,07	-	5652	-	-	CAOA01000001-CAOA01000149
SOn1	Moribund *L. stylirostris* hemoculture, July 2001, New Caledonia	NP		6,44	approx190	6117	231	-	CAOF01000001-CAOF01000202
SFn118	Growout pond water, diseased pond, April 2002, New Caledonia	NP		6,11	-	5752	-	-	CAOC01000001-CAOC01000235
ATCC27043T	Seawater, chitin enrichment, USA	NP		6,11	-	5770	-	-	NZ-AFWJ00000000
Mada3021	Mud crab, September 2007, Madagascar	not determined	M	6,14	160	5775	127	-	CANX01000001-CANX01000136
Mada3020	Moribund *P. monodon* hemoculture, April 2007, Madagascar	not determined	M	6,15	160	5758	127	-	CANY01000001-CANY01000140
Mada3029	Moribund *P. monodon* hemoculture, Pond B11, September 2007, Madagascar	not determined	M	6,15	160	5768	127	-	CANZ01000001-CANZ01000131

LEGENDS TO FIGURES

Figure 1: Colony pigmentation of *V. nigripulchritudo* strains SFn1 (a), Δ *VIBNI_B0280* (laccase) (b) and Δ *VIBNI_B1404* (phosphotyrosinase) (c) derivatives grown on marine agar supplemented with 1% glycerol.

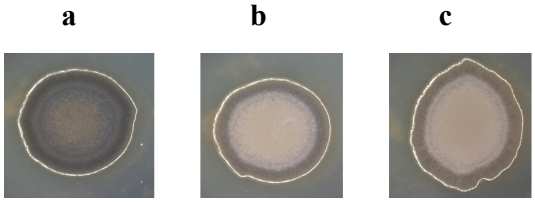


Figure 1

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).

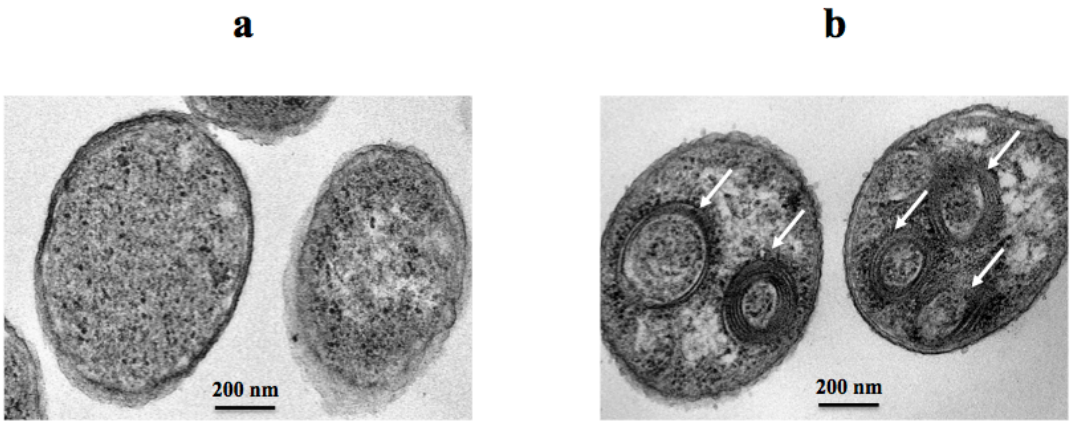


Figure 2

Figure 3: Phylogenetic analysis based on concatenated alignments of nucleic acid sequences of 4421 core genes from 16 *V. nigripulchritudo* strains (ATCC 27043^T 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.

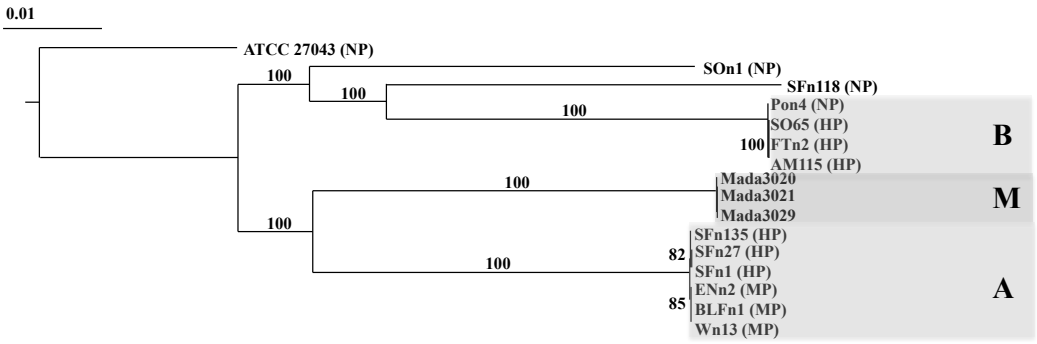


Figure 3

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].

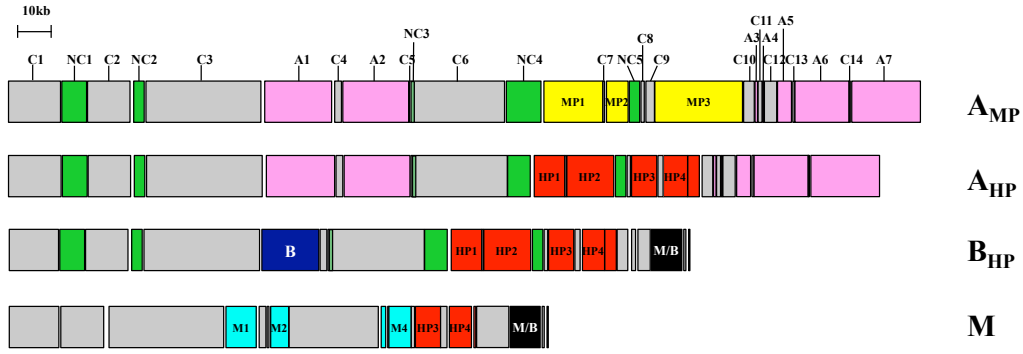

10kb
C1
NC1
C2
NC2
C3
A1
C4
A2
NC3
C5
C6
NC4
C7
NC5
C8
C9
C10
C11
A3
A4
A5
C12
C13
A6
C14
A7
MP1
MP2
MP3
HP1
HP2
HP3
HP4
B
M1
M2
M4
M/B
A_MP
A_HP
B_HP
M

Figure 4

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.

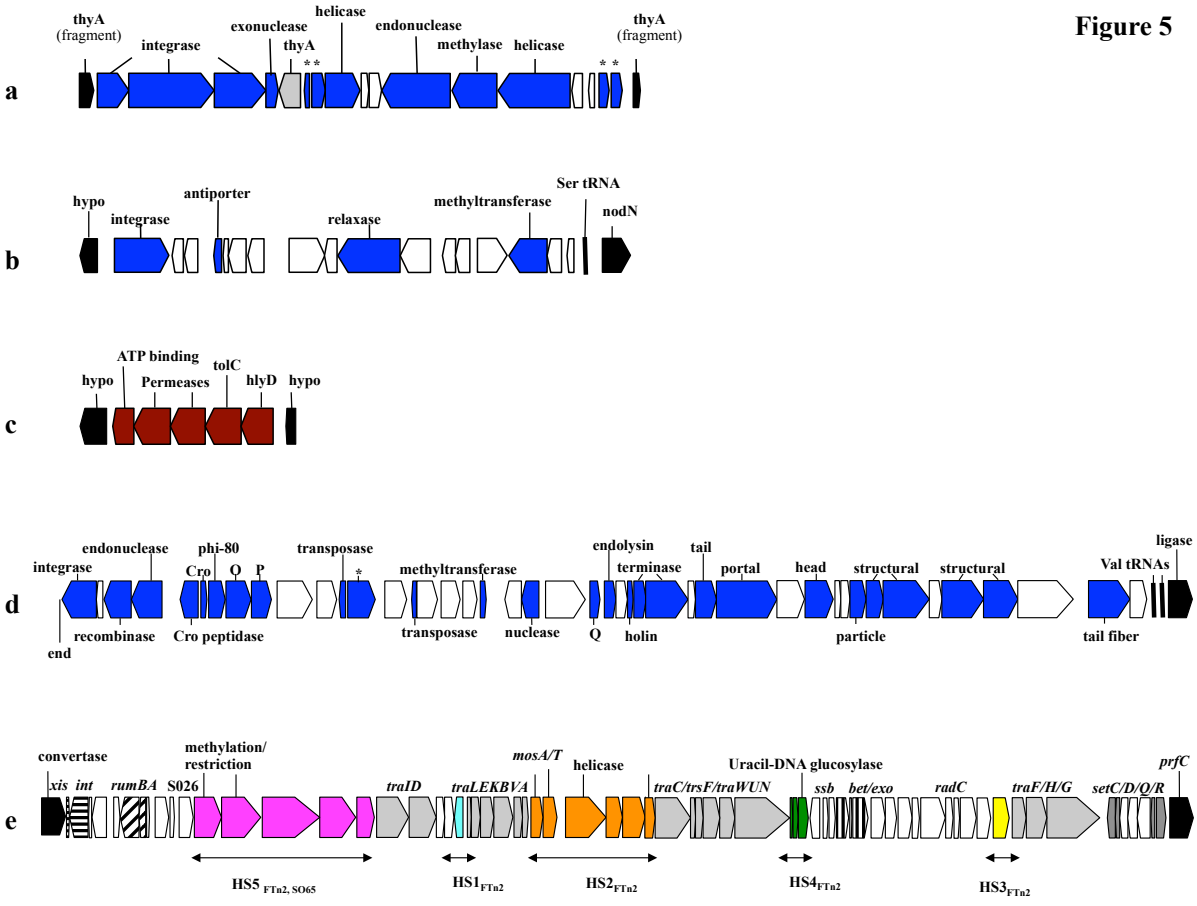


Figure 5

Figure 6 : Shrimp mortality in response to injection of bacterial culture supernatants. Shrimp, *L. stylirostris* (n=5), were injected with 100 µL 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 transconjuguant of SFn118 cultivated in the presence of arabinose or glucose; from the P_{BAD}-GFP transconjuguant 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.

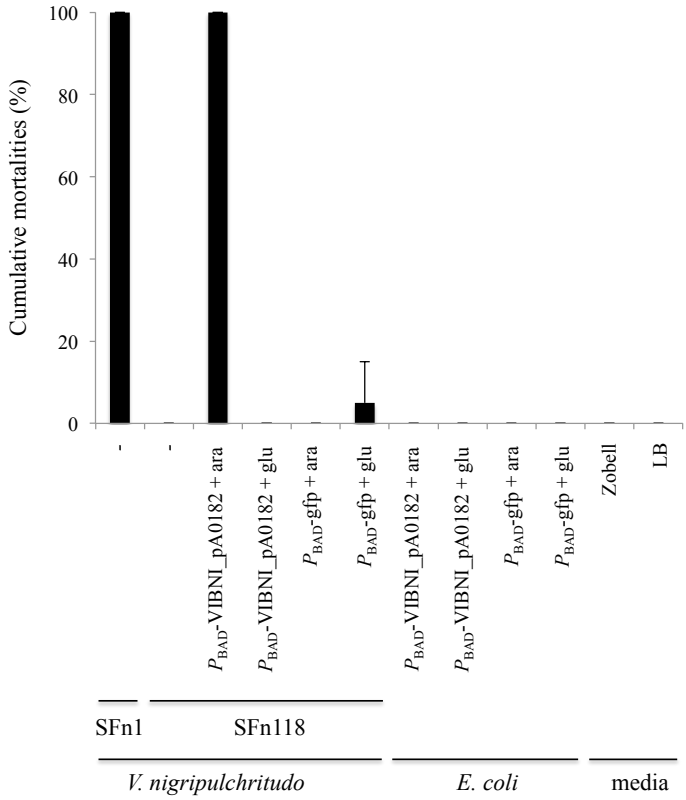


Figure 6

SUPPLEMENTARY METHODS

Polymerase chain reaction (PCR)

PCR performed for plasmid assembly were done in 50 µl volumes using the Herculase DNA polymerase (Agilent) following the manufacturer's instructions. Other PCR were performed in 50 µl volumes using the GoTaq (Promega) polymerase according to the manufacturer's instructions. Primers are listed in Table S3.

Cloning of deleted alleles into a suicide vector

Alleles carrying an internal deletion were generated *in vitro* using a two-step PCR construction method as described previously (Binesse et al 2008). Using genomic DNA of the strain SFn1, independent PCR amplifications of the regions (500 bp) encompassing the laccase (VIBNI_B0280) and the phosphotyrosinase (VIBNI_B1404) genes were performed using two primer pairs: (lac-1, -2) and (lac-3, -4), and (Pty-1, -2) and (Pty-3, -4) (Table S3). After gel purification, 100 ng of the two PCR products were mixed for each allele and a final PCR amplification was carried out using the most external primer pairs: (lac-1, -4) and (Pty-1, -4). After *EcoR1* digestion, the generated fragment was cloned in pSW8742T (Table S2). *E. coli* Π3813 and β3914 were used as a plasmid host for cloning and conjugation, respectively (Le Roux et al 2007).

Cloning the coding DNA sequences (CDSs) under the control of an inductible promoter in a replicative plasmid

Primers rebF and rebR, 0182-F and 0182-R or GFP-F and GFP-R (Table S3) were used to amplify the reb cluster (VIBNI_A3109 to 3099, Fig. S3), the nigritoxin and GFP genes, respectively. The PCR products were digested using restriction enzymes (*ApaI* and *XhoI*) and

ligated into a P15A-*ori*-based replicative vector (*araC* ;P_{BAD} ; ori_{p15A}; $oriT_{RP4}$; [Sp^R]) (Bartolome et al 1991, Le Roux et al 2011). *E. coli* TOP10 and β3914 were used as a plasmid host for cloning and conjugation, respectively. The results of all ligations were confirmed by plasmid sequencing.

Conjugation

Overnight cultures of donor and recipient were diluted at 1:100 in culture media without antibiotic and grown at 30°C to an OD_{600nm} of 0.3. The different conjugation experiments were done by a filter mating procedure described previously (Le Roux et al 2011) with a donor/recipient ratio of 1ml/10ml. Conjugations were performed overnight on filters incubated on LBA + NaCl 0.5N + diaminopimelic acid (DAP) plates at 30°C. Counter-selection of *ΔdapA* donor was done by plating on a medium devoid of DAP, supplemented with spectinomycin and 1% glucose. For mutagenesis, $Spec^R$ resistant colonies were isolated, grown in LB + NaCl 0.5N up to late logarithmic phase and spread on plate containing 0.2% arabinose. Mutants were screened by PCR using primers 1+4 flanking the different genes targeted.

Genome sequencing

The complete genome sequence of SFn1 strain was obtained using two sequencing technologies: 1) A Sanger library was constructed after mechanical shearing of DNA and cloning of 10 kpb fragments into pCNS (pSU18 derived). Plasmids were purified and end-sequenced using dye-terminator chemistry on ABI3730 sequencers leading to a 4-fold coverage. 2) A 454 single read library was constructed and sequenced to a 16-fold coverage. The reads obtained using the two technologies were assembled using Newbler

(www.roche.com). Then, primer walks, PCRs and transposon bombs were performed to finish the sequence of the *V. nigripulchritudo* reference genome.

Structural and functional genomes annotation

Computational prediction of coding sequences (CDSs) and other genome features (RNA encoding genes, ribosome binding sites, signal sequences, etc...), together with functional assignments were performed using the automated annotation pipeline implemented in the MicroScope platform (Vallenet et al., 2009). An extensive manual curation of the genes, which includes correction of the start codon positions and of the functional assignments, was performed. This expert procedure was supported by functional analysis results [e.g., InterPro, FigFam, PRIAM, COGs (Clusters of Orthologous Groups), PsortB] which can be queried using an exploration interface, and by synteny groups computation visualized in cartographic maps to facilitate genome comparison.

SUPPLEMENTARY TABLES

Table S1: Strains used and constructed in this study

Strain	*Description*	*Reference*
Π3813	B462 *ΔthyA::(erm-pir116)* [Erm ^{R}]	Le Roux *et al.,* 2007
β3914	β2163 *gyrA462, zei298::Tn10* [Km^{R} Em^{R} Tc^{R}]	Le Roux *et al.,* 2007
TOP10	(F–) mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (Str^{R}) endA1 nupG	Invitrogen
GV2	SFn1 *Δ VIBNI_B1404* (phosphotyrosinase)	This study
GV3	SFn1 *Δ VIBNI_B0280* (laccase)	This study
GV319	TOP10 + plasmid P_{BAD}*gfp*	This study
GV321	TOP10 + plasmid P_{BAD}*nigritoxin*	This study
GV330	SFn118 + plasmid P_{BAD}*gfp*	This study
GV331	SFn118 + plasmid P_{BAD}*nigritoxin*	This study
GV691	TOP10 + plasmid P_{BAD}*reb*	This study

Table S2: Plasmids used and constructed in this study

Plasmid	*Description*	*Reference*
pSW8742T	*oriV*$_{R6K\gamma}$; *oriT*$_{RP4}$; *araC*-P$_{BAD}$*ccdB;* [Spec^{R}]	Le Roux *et al.,* 2007
pSWδPty	pSW8742T; *Δ VIBNI_B1404*	This study
pSWδlac	pSW8742T; *Δ VIBNI_B0280*	This study
P$_{BAD}$*gfp*	*oriV*$_{p15A}$; *oriT*$_{RP4}$; *araC*-P$_{BAD}$*gfp;* [Spec^{R}]	This study
P$_{BAD}$*nigritoxin*	*oriV*$_{p15A}$; *oriT*$_{RP4}$; *araC*-P$_{BAD}$*VIBNI_pA0182;* [Spec^{R}]	This study
P$_{BAD}$*reb*	*oriV*$_{p15A}$; *oriT*$_{RP4}$; *araC*-P$_{BAD}$ *VIBNI_A3109 to 3099 ;* [Spec^{R}]	This study

Tables S3: Primers used in this study

Primer	*Sequence*
lac-1	GCCCGAATTCATGGACATACAAAGACGCAAGC
lac-2	GTCTGATGCGACCGATTCGAGTATGTTCTTCATCAAATTCAACG
lac-3	CGTTGAATTTGATGAAGAACATACTCGAATCGGTCGCATCAGAC
lac-4	GCCCGAATTCTTATACGACTTCAATGTAGCCC
Pty-1	GCCCGAATTCATGACTACCGTTGCAAAACTTAACTACC
Pty-2	CGTAATGGTATCCACGGTGCAACTCCTCGTCTCCGGACAAACC
Pty-3	GGTTTGTCCGGAGACGAGGAGTTGCACCGTGGATACCATTACG
Pty-4	GCCCGAATTCTTAGCCCTTTGCAAATGTGCG
GFP-F	GCCCGGGCCCATGAGTAAAGGAGAAGAACTTTTC
GFP-R	GCCCCTCGAGCTATTTGTATAGTTCATCCATGCC
0182-F	GCCCGGGCCCATGTCTCTACCCTCAAACCCC
0182-R	GCCCCTCGAGCTAAACTGATGATGAAGCCCC
rebF	GCCCGGGCCCATGGAAATCGATATAGAAAAAATGTTGCG
rebR	GCCCCTCGAGTCAATCAAACTTCTTGGCAACTTCC

Table S4: Predicted NRPS/PKS gene clusters in *Vibrio nigripulchritudo* SFn1 and their presence in other *Vibrio* genomes.

Replicon	*Type*	*Gene*	*Length*	*Domain composition*	*I*
Chromosome A	Hybrid cluster	VIBNI_A1885	8760	A-PCP-C-A-PCP-KS-KR-PCP-TE	
		VIBNI_A1886	11820	C-A-PCP-C-A-PCP-C-C-A-PCP-TE	
		VIBNI_A1887	5034	A+PCP+C+A+PCP	
	Type III PKS	VIBNI_A2441	1044		
	NRPS	VIBNI_A3716	5628	C-A-PCP-TE	
		VIBNI_A3717	3375	C-A-PCP	
		VIBNI_A3718	8991	C-A-nMT-PCP-E-C-A-PCP	
Chromosome B	NRPS	VIBNI_B0988	10392	C-A-PCP-E-C-A-PCP-E-C	
Plasmid pA	Hybrid	VIBNI_pA0054	11826	C-A-ACP-KS-AT-KR-ACP-TE-KS-DH-KR-ACP	
	Type I PKS	VIBNI_pA0055	5988	KS-DH-KR-ACP-KS-AT	

[a] (absent from SFn118 and ATCC27043^T); [b] (all from clade A); [c] (absent from Son1)

SUPPLEMENTARY FIGURES

Fig. S1: Phylogenetic analysis based on concatenated alignments of nucleic acid sequences of 122 core genes from 133 *Vibrionaceae* strains and *Shewanella baltica* as an 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. Number at each node represents the percentage value given by bootstrap analysis of 100 replicates.

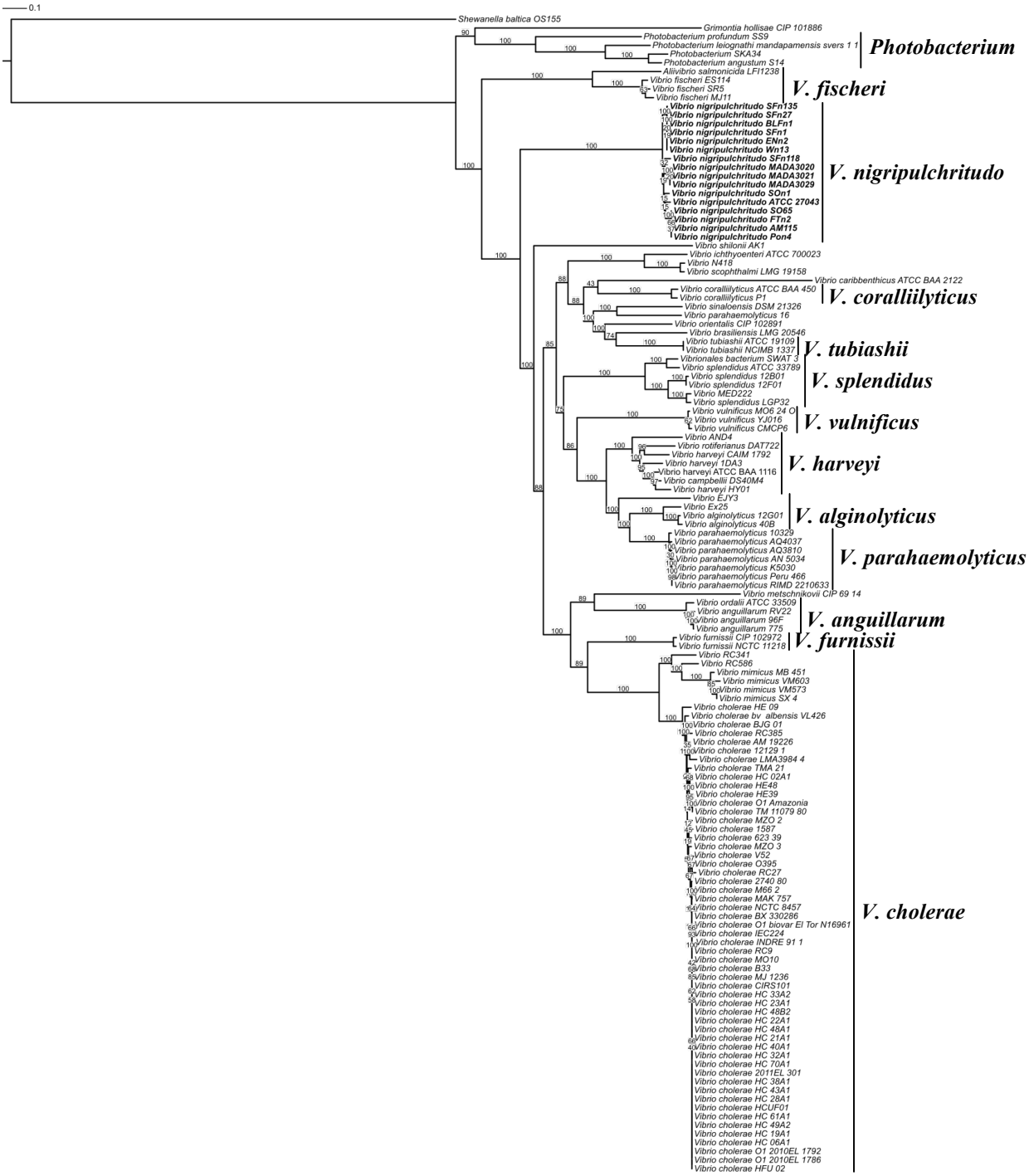

0.1
Shewanella baltica OS155
Grimontia hollisae CIP 101886
Photobacterium profundum SS9
Photobacterium leiognathi mandapamensis svers 1 1
Photobacterium SKA34
Photobacterium angustum S14
Photobacterium
Aliivibrio salmonicida LFI1238
Vibrio fischeri ES114
Vibrio fischeri SR5
Vibrio fischeri MJ11
V. fischeri
Vibrio nigripulchritudo SFn135
Vibrio nigripulchritudo SFn27
Vibrio nigripulchritudo BLFn1
Vibrio nigripulchritudo SFn1
Vibrio nigripulchritudo ENn2
Vibrio nigripulchritudo Wn13
Vibrio nigripulchritudo SFn118
Vibrio nigripulchritudo MADA3020
Vibrio nigripulchritudo MADA3021
Vibrio nigripulchritudo MADA3029
Vibrio nigripulchritudo SOn1
Vibrio nigripulchritudo ATCC 27043
Vibrio nigripulchritudo SO65
Vibrio nigripulchritudo FTn2
Vibrio nigripulchritudo AM115
Vibrio nigripulchritudo Pon4
V. nigripulchritudo
Vibrio shilonii AK1
Vibrio ichthyoenteri ATCC 700023
Vibrio N418
Vibrio scophthalmi LMG 19158
Vibrio caribbenthicus ATCC BAA 2122
Vibrio coralliilyticus ATCC BAA 450
Vibrio coralliilyticus P1
V. coralliilyticus
Vibrio sinaloensis DSM 21326
Vibrio parahaemolyticus 16
Vibrio orientalis CIP 102891
Vibrio brasiliensis LMG 20546
Vibrio tubiashii ATCC 19109
Vibrio tubiashii NCIMB 1337
V. tubiashii
Vibrionales bacterium SWAT 3
Vibrio splendidus ATCC 33789
Vibrio splendidus 12B01
Vibrio splendidus 12F01
Vibrio MED222
Vibrio splendidus LGP32
V. splendidus
Vibrio vulnificus MO6 24 O
Vibrio vulnificus YJ016
Vibrio vulnificus CMCP6
V. vulnificus
Vibrio AND4
Vibrio rotiferianus DAT722
Vibrio harveyi CAIM 1792
Vibrio harveyi 1DA3
Vibrio harveyi ATCC BAA 1116
Vibrio campbellii DS40M4
Vibrio harveyi HY01
V. harveyi
Vibrio EJY3
Vibrio Ex25
Vibrio alginolyticus 12G01
Vibrio alginolyticus 40B
V. alginolyticus
Vibrio parahaemolyticus 10329
Vibrio parahaemolyticus AQ4037
Vibrio parahaemolyticus AQ3810
Vibrio parahaemolyticus AN 5034
Vibrio parahaemolyticus K5030
Vibrio parahaemolyticus Peru 466
Vibrio parahaemolyticus RIMD 2210633
V. parahaemolyticus
Vibrio metschnikovii CIP 69 14
Vibrio ordalii ATCC 33509
Vibrio anguillarum RV22
Vibrio anguillarum 96F
Vibrio anguillarum 775
V. anguillarum
Vibrio furnissii CIP 102972
Vibrio furnissii NCTC 11218
V. furnissii
Vibrio RC341
Vibrio RC586
Vibrio mimicus MB 451
Vibrio mimicus VM603
Vibrio mimicus VM573
Vibrio mimicus SX 4
Vibrio cholerae HE 09
Vibrio cholerae bv. albensis VL426
Vibrio cholerae BJG 01
Vibrio cholerae RC385
Vibrio cholerae AM 19226
Vibrio cholerae 12129 1
Vibrio cholerae LMA3984 4
Vibrio cholerae TMA 21
Vibrio cholerae HC 02A1
Vibrio cholerae HE48
Vibrio cholerae HE39
Vibrio cholerae O1 Amazonia
Vibrio cholerae TM 11079 80
Vibrio cholerae MZO 2
Vibrio cholerae 1587
Vibrio cholerae 623 39
Vibrio cholerae MZO 3
Vibrio cholerae V52
Vibrio cholerae O395
Vibrio cholerae RC27
Vibrio cholerae 2740 80
Vibrio cholerae M66 2
Vibrio cholerae MAK 757
Vibrio cholerae NCTC 8457
Vibrio cholerae BX 330286
Vibrio cholerae O1 biovar El Tor N16961
Vibrio cholerae IEC224
Vibrio cholerae INDRE 91 1
Vibrio cholerae RC9
Vibrio cholerae MO10
Vibrio cholerae B33
Vibrio cholerae MJ 1236
Vibrio cholerae CIRS101
Vibrio cholerae HC 33A2
Vibrio cholerae HC 23A1
Vibrio cholerae HC 48B2
Vibrio cholerae HC 22A1
Vibrio cholerae HC 48A1
Vibrio cholerae HC 21A1
Vibrio cholerae HC 40A1
Vibrio cholerae HC 32A1
Vibrio cholerae HC 70A1
Vibrio cholerae 2011EL 301
Vibrio cholerae HC 38A1
Vibrio cholerae HC 43A1
Vibrio cholerae HC 28A1
Vibrio cholerae HCUF01
Vibrio cholerae HC 61A1
Vibrio cholerae HC 49A2
Vibrio cholerae HC 19A1
Vibrio cholerae HC 06A1
Vibrio cholerae O1 2010EL 1792
Vibrio cholerae O1 2010EL 1786
Vibrio cholerae HFU 02
V. cholerae

Fig.S2: Circular representation of the *V. nigripulchritudo* SFn1 genome with comparative genomic data. From the outside inward: the 1st circle (orange bars) shows CDSs that are unique to *V. nigripulchritudo* genomes, the 2d circle (green bars) shows CDSs that are shared between 1 to 5 strains of other *Vibrionaceae* species. The 3rd circle (blue bars) corresponds to the CDSs that are common to all sequenced *V. nigripulchritudo*.

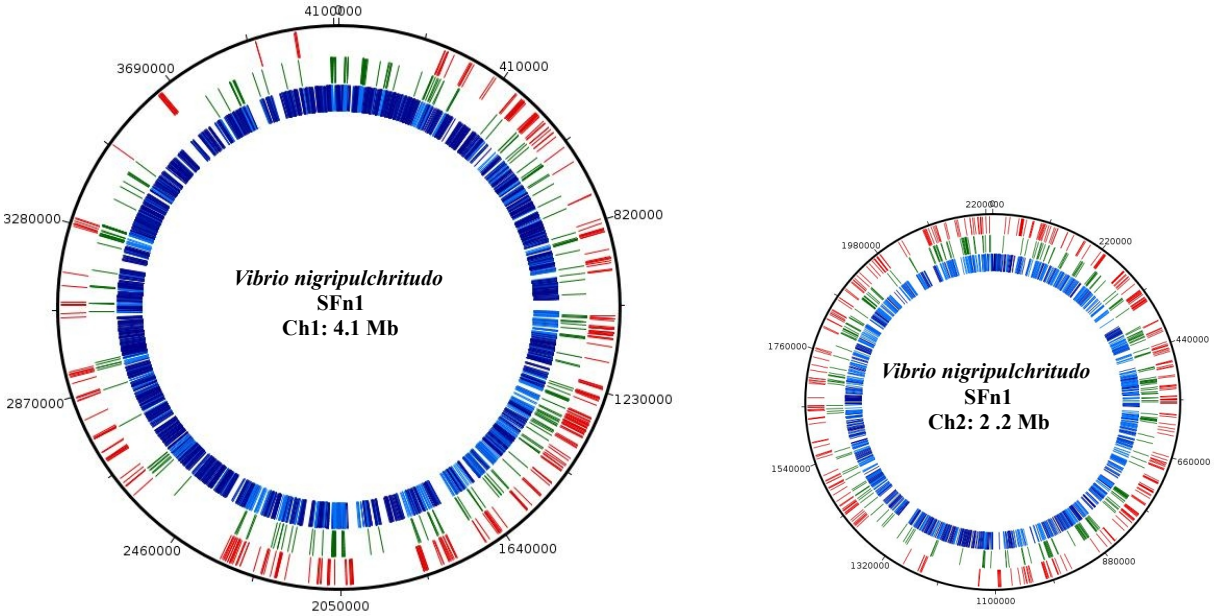
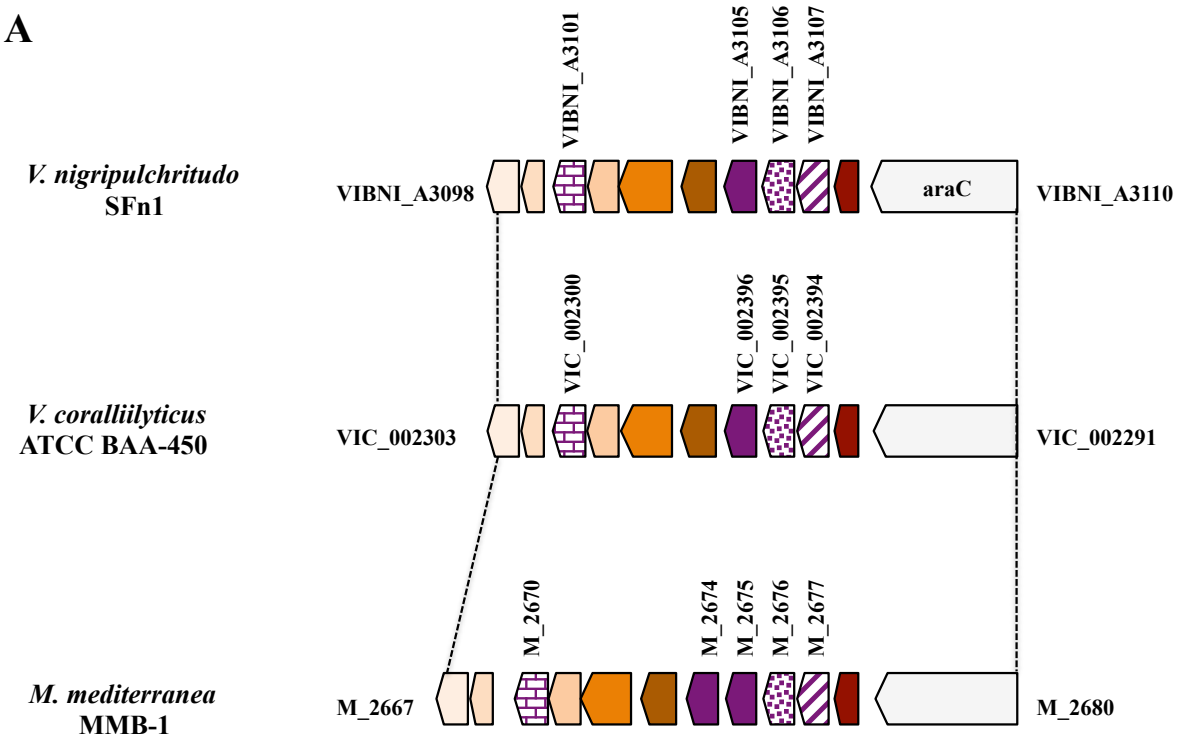


Fig.S3: Organization and phylogeny of the *reb* gene cluster in *Vibrio nigripulchritudo*, *V. coralliilyticus* and *Marinomonas mediterranea.* (A) The 4 *reb* genes (purple) are localized within a syntenic group shared by the three species. (B) Phylogenetic tree of various Reb proteins. The tree was built by the Neighbor-joining method based on sequences aligned using Muscle. Branch lengths are drawn to scale and proportional to the number of amino acid changes. Number at each node represents the percentage value given by bootstrap analysis of 1000 replicates. GenBank accession numbers of *V. coralliilyticus* and *M. mediterranea* genome sequences are NZ_ACZN00000000 and NC_015276 respectively.


A
V. nigripulchritudo SFn1
VIBNI_A3098
VIBNI_A3101
VIBNI_A3105
VIBNI_A3106
VIBNI_A3107
araC
VIBNI_A3110
V. coralliilyticus ATCC BAA-450
VIC_002303
VIC_002300
VIC_002396
VIC_002395
VIC_002394
VIC_002291
M. mediterranea MMB-1
M_2667
M_2670
M_2674
M_2675
M_2676
M_2677
M_2680

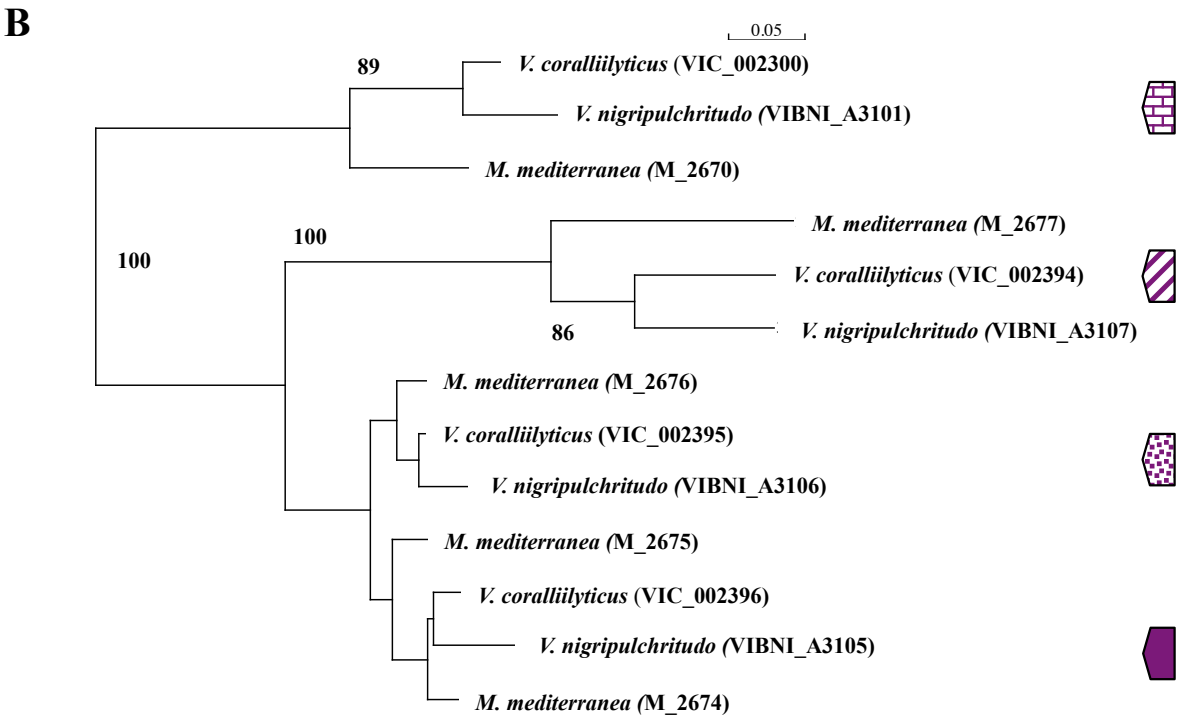

B
0.05
89
V. coralliilyticus (VIC_002300)
V. nigripulchritudo (VIBNI_A3101)
M. mediterranea (M_2670)
100
100
M. mediterranea (M_2677)
V. coralliilyticus (VIC_002394)
86
V. nigripulchritudo (VIBNI_A3107)
M. mediterranea (M_2676)
V. coralliilyticus (VIC_002395)
V. nigripulchritudo (VIBNI_A3106)
M. mediterranea (M_2675)
V. coralliilyticus (VIC_002396)
V. nigripulchritudo (VIBNI_A3105)
M. mediterranea (M_2674)

Fig.S4: Comparison of each core genes phylogenetic relationship. For each gene a tree was built by the Maximum-Likelihood method based on sequences aligned using Muscle (4421 trees). Each tree topology was analyzed using the ETE2 Python library: 81% of the trees led to the clustering of SFn1, SFn27, SFn135, ENn2, Wn13 and BLFn1 strains in clade A ; SO65, FTn2, AM115, POn4 strains in clade B ; Mada3020, Mada3021 and Mada3029 in clade M ; with a bootstrap value >75%, using ATCC 27043^T as an outgroup. The topologies obtained were classified based on AM, BM or AB being sister clades or ABM clades relationship being unresolved. Percentage of each topology is mentioned on the top of the figure.

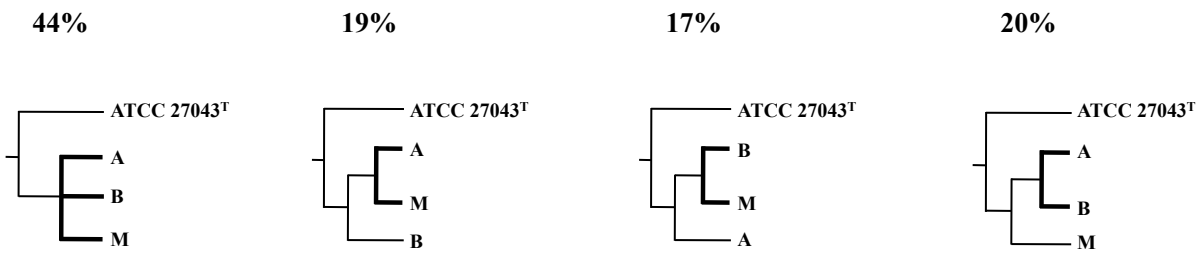


Fig.S5: Phylogenetic analysis based on concatenated alignments of nucleic acid sequences of 40 backbone genes from 13 plasmids (SOn1 as an outgroup). Trees were built by the Maximum-Likelihood method based on sequences aligned using Muscle. Branch lengths are drawn to scale and are proportional to the number of nucleotide changes. Number at each node represents the percentage value given by bootstrap analysis of 100 replicates.

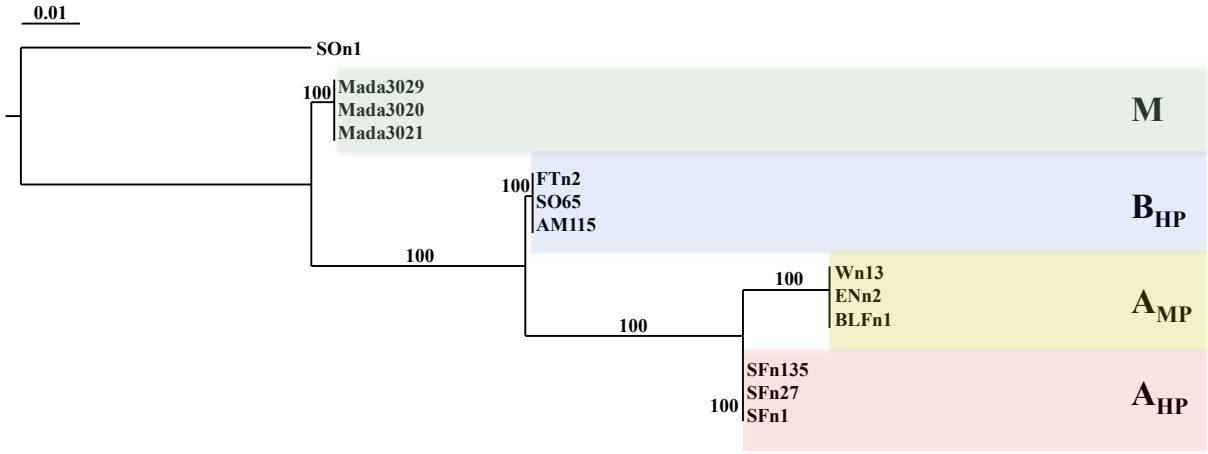


Fig.S6: Structure of the HP-specific modules carried by the large plasmid. Grey CDSs correspond to core genes. Green CDSs are found in all New Caledonian isolates. Hatched red CDSs belong to a Galactose operon (HP1 [VIBNI_pA0141-pA0149]), dotted red CDSs represent genes encoding a siderophore ABC transporter and a peptidase (HP2 [VIBNI_pA0151-pA0166]), squared red CDSs are implicated in purine metabolism (HP3 [VIBNI_pA0171-pA0177]), within the filled red gene cluster (HP4 [VIBNI_pA0181-pA0196]), the VIBNI_pA0182 encodes the nigritoxin. Regulator genes are indicated by a star.

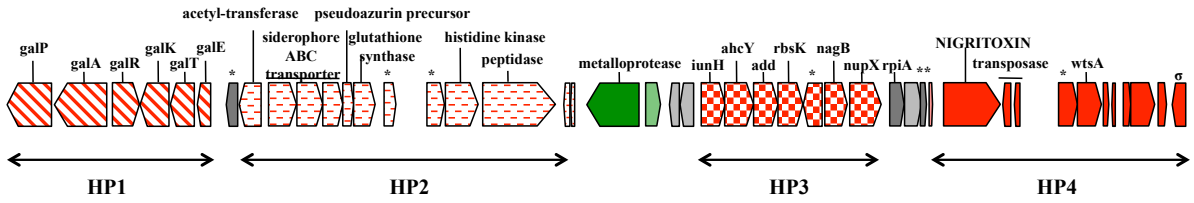


Fig.S7: Protein alignment of the VIBNI_pA0182 gene product, and the Afp18 putative toxin described in *Serratia entomophila* (Genbank acc. no: YP_026154) and *Yersinia ruckeri* *(*ZP_04616994)