Environmental Microbiology
November 2015, Volume 17, Issue 11, Pages 4189-4199
http://dx.doi.org/10.1111/1462-2920.12699
http://archimer.ifremer.fr/doc/00226/33685/

A single regulatory gene is sufficient to alter *Vibrio aestuarianus* pathogenicity in oysters

Goudenège David [1, 2, 3], Travers Marie-Agnes [4], Lemire Astrid [1, 2, 3], Petton Bruno [1], Haffner Philippe [4], Labreuche Yannick [1, 2, 3], Tourbiez Delphine [4], Mangenot Sophie [5], Calteau Alexandra [5, 6], Mazel Didier [7, 8], Nicolas Jean-Louis [1], Jacq Annick [9], Le Roux Frederique [1, 2, 3, *]

[1] Ifremer, UMR LEMAR (UBO/CNRS/IRD/Ifremer) 6539, 11 presqu'île du Vivier, F-29840 Argenton-en-Landunvez; France
[2] Sorbonne Universités, UPMC Univ Paris 06, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, F-29688, Roscoff Cedex; France
[3] CNRS UMR 8227, Integrative Biology of Marine Models, F-29688, Roscoff Cedex; France
[4] Ifremer, SG2M-LGPMM, Laboratoire de Génétique et Pathologie des Mollusques Marins Avenue de Mus de Loup, F-17390 La Tremblade; France
[5] Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Direction des Sciences du Vivant (DSV), Institut de Génomique (IG), 2 rue Gaston Crémieux, CP 5706, F-91057 Evry cedex; France
[6] CNRS UMR 8030, Laboratoire d'Analyse Bioinformatiques en Génomique et Métabolisme (LABGeM), F-91057 Evry cedex; France
[7] Institut Pasteur, Unité Plasticité du Génome Bactérien, Département Génomes et Génétique, F-75015 Paris; France
[8] CNRS UMR3525, F-75015 Paris; France
[9] Institut de Génétique et Microbiologie, UMR8621, CNRS-Université Paris-Sud, F-91405 Orsay Cedex; France

* Corresponding author : Frédérique Le Roux, tel 33 2 98 29 56 47 ; Fax 33 2 98 29 83 84 ; email address : frederique.le-roux@sb-roscoff.fr

Abstract :

Oyster diseases caused by pathogenic vibrios pose a major challenge to the sustainability of oyster farming. In France, since 2012 a disease affecting specifically adult oysters has been associated with the presence of *Vibrio aestuarianus*. Here, by combining genome comparison, phylogenetic analyses and high-throughput infections of strains isolated before or during the recent outbreaks, we show that virulent strains cluster into two *V. aestuarianus* lineages independently of the sampling dates. The bacterial lethal dose was not different between strains isolated before or after 2012. Hence, the emergence of a new highly virulent clonal strain is unlikely. Each lineage comprises nearly identical strains, the majority of them being virulent, suggesting that within these phylogenetically coherent virulent lineages a few strains have lost their pathogenicity. Comparative genomics allowed the identification of a single frameshift in a non-virulent strain. This mutation affects the *varS* gene that codes for a signal transduction histidine-protein kinase. Genetic analyses confirmed that *varS* is necessary for infection of oysters and for a secreted metalloprotease expression. For the first time in a *Vibrio* species, we show here that VarS is a key factor of pathogenicity.

INTRODUCTION

The development of aquaculture has been the source of anthropogenic changes on a massive scale, characterized by displacements of aquatic animals from their natural habitats, farming under high stocking density and exposition to environmental stresses. At the same time, over-exploitation of some species and anthropogenic stress on aquatic ecosystems have placed pressure on wild populations, providing opportunities for the emergence of an expanding array of new diseases (Harvell et al., 1999). This can be illustrated by the outbreaks of *Crassostrea gigas* oyster's diseases over the past decade threatening the long-term survival of commercial and natural stocks (Renault, 2011).

A disease affecting oyster spat since 2008 has been linked to the presence of an oyster herpes virus (OsHV-1 µvar) (Segarra et al., 2010) and to vibrio populations related to the *Splendidus* clade (Lemire et al, *in press*). In addition, over the last four years, the number of reported cases of adult mortalities associated with the presence of *Vibrio aestuarianus* has increased considerably (Garnier et al., 2007; Vezzulli et al., 2014). Interestingly, during the 2008-2012 period, this bacterial species had been rarely isolated from moribund oysters, suggesting the possible (re)-emergence of *V. aestuarianus* as an oyster pathogen.

Emergent infectious diseases can arise from genomic modifications of an infectious agent (Keim and Wagner, 2009). Such genomic alterations frequently result in the optimization of virulence genes, the acquisition of resistance cassettes, changes in prevalence, and/or adaptation to a new host. Thus, understanding an emerging disease requires investigating the pathogen and its evolution at the gene and genome level. If the recent emergence of *V.*

aestuarianus-caused diseases is due to a new virulent clonal strain, one can expect differences in the lethal doses between strains isolated during the recent outbreaks (heightened virulence) and the ones sampled a decade ago. Furthermore, one can hope that sequencing closely related isolates with contrasting virulence status and performing whole genome comparative analyses would lead to the identification of genomic modification(s) correlating with increased virulence.

Indeed, whole genome comparative analyses combined with mutagenesis of candidate genes have been successfully used to identify virulence determinants of vibrio pathogenic to oysters such as *V. crassostreae* (Lemire et al., *in press*) or shrimp such as *V. nigripulchritudo* (Goudenege et al., 2013). However such a strategy requires the sequencing of several closely related strains with contrasting virulence and is still limited to genetically tractable strains. To date attempts to perform reverse genetics in *V. aestuarianus* (strain 01_032) have proven unsuccessful (Labreuche et al., 2010) limiting the investigation of virulence mechanisms in this species.

In the present study, we explored the virulence potential and genome diversity of *V. aestuarianus* isolates. We asked whether the recent adult mortality outbreaks are due to the emergence of a specific genotype. To address this question, we performed a high throughput sequencing (HTS)-based comparative genome analysis of 14 *V. aestuarianus* strains isolated before or during the recent outbreaks along with bacterial lethal dose determination by experimental challenges. We then took advantage of the near identity of some strains with contrasting virulence properties to identify key factor(s) of *V. aestuarianus* pathogenicity by comparative and functional genomics.

RESULTS AND DISCUSSION

V. aestuarianus comprises virulent and non-virulent strains

To explore the virulence potential and genome diversity of *V. aestuarianus*, strains isolated from diseased oysters before (four strains, named 01_xxx to 07_xxx, the first two numbers corresponding to the year of isolation, *i.e.* 01 for 2001) and during the recent mortality outbreaks (six strains, named 12_xxx) were selected (Table 1). In addition, four strains isolated from healthy oysters, cockles or zooplankton and not linked to mortality events were added to our analysis (Table 1).

The strains were first injected into specific-pathogen-free (SPF) standardized oysters (Petton et al., 2013) at 10^7 CFU/animal, a bacterial concentration previously used in experimental infections (Garnier et al., 2007). At 6 days post injection, 10/14 strains caused mortality rates >80% (Fig. 1, black bars). We subsequently injected lower bacterial concentrations to oysters (ranging from 10^6 to 10^2 CFU/animal). Surprisingly, when injected at 10^2 CFU /animal, the strain 02_041 isolated in 2002 and six strains isolated in 2012 were still able to cause >80% mortality (Fig.1, grey bars).

These results allowed us to classify seven strains (12_063, 12_128a, 12_130, 12_142, 12_055, 12_016a and 02_041) as highly virulent (Vir+) (*i.e.* inducing >50% mortalities at 10^2 CFU/animal), and four strains (12_122, 11_U17, 11_KB19 and 01_151) as non-virulent (Vir-) (*i.e.* inducing <50% mortalities at 10^7 CFU/animal). Three strains (07_115, 01_308, 01_032) were defined as intermediate (*i.e.* pathogenic only at 10^7 CFU/animal).

General features of the *V. aestuarianus* genomes

The genome of strain 02_041 was assembled to near completion (8 contigs in total) and manually annotated. It consists of two circular chromosomes of 2.98 (chromosome 1; 4 contigs) and 1.21 Mb (chromosome 2, 4 contigs) with an average GC content of 43.11 and 42.16% respectively (Table 1; Fig.S1). Chromosomes 1 and 2 contain 7 and 0 rRNA operons, 74 and 10 tRNA genes, respectively. However because the genome is not fully assembled, some rRNA and tRNA genes may have been missed.

The genome sequences of the 13 other strains were partially assembled, with contig numbers per strain ranging from 38 to 732 and approximate genome sizes ranging from 4.2 to 4.99 Mb compared to 4.19 Mb for strain 02_041 (Table 1). The difficulty to achieve a better genome assembly may be attributed to i) a high number of transposition elements (184 transposase genes in the strain 02_041); ii) the large size of the chromosomal integron (Mazel et al., 1998) (94 cassettes in the chromosome 2 of strain 02_041) (Fig. S1).

Genes differentiating *V. aestuarianus* from other *Vibrionaceae*

A phylogenetic analysis based on concatenated nucleic acid sequences derived from 50 shared genes from 223 *Vibrionaceae* genome sequences including 14 *V. aestuarianus* strains and using *Shewanella baltica* as an outgroup demonstrated the cohesive genotypic structure of *V. aestuarianus* with relatively little diversity among genomes (Fig. S2). The clade *V. aestuarianus* is sister to a clade that contains two species previously associated with farmed fish diseases, *V. ordalii* and *V. anguillarum* (Austin, 2011). Our analyses confirmed that *V. aestuarianus*, *V. ordalii* and *V. anguillarum* are grouped in the *Anguillarum* clade (Sawabe et al., 2013).

Intraspecific genomic comparisons identified 2866 genes that are shared by all sequenced *V. aestuarianus* strains (Fig. S1) of which only 40 genes were found in ≤5 other *Vibrionaceae* genomes (Table S1). Among these *V. aestuarianus*-specific genes, we identified a cluster of genes homologous to the Toxin co-regulated (Tcp) pilus biosynthesis cluster encoded by a pathogenicity island in *V. cholerae* that is necessary for intestine colonization (Davis and Waldor, 2003). However in the strain 02_041, the *tcp* gene cluster is interrupted by a transposon, and genes encoding the accessory colonization factors (*acf*) are absent (Fig.S3) suggesting that this *tcp* like cluster may play a distinct role, if any, in *V. aestuarianus*.

Within *V. aestuarianus,* two lineages A and B contain a majority of Vir+ strains

The phylogenetic relationships based on the core genome of the *V. aestuarianus* strains included in this study were investigated (Fig.2). The main outcome of this analysis was the grouping of 6/7 Vir+ isolates into a clade A, which also contains one Vir-, and two intermediate strains. Clade A is a sister of Clade B containing one Vir+ and one intermediate strain. Both clades A and B show very little intra-clade diversity (>99 % average nucleotide identity – ANI– value) (Konstantinidis and Tiedje, 2005). Inter-clade diversity was also low as determined by the ANI value calculation (>98.4 %) and by the number of clade-specific genes (≃180 genes, essentially in a clade B-specific phage). Vir- strains isolated from oysters in Spain, zooplankton in Italy or cockles in Brittany were found to be more diverse.

As a consequence of the low inter-clade diversity, the genes commonly used for multilocus sequence analysis (*hsp60, pyrH, atpA, gyrB, recA, topA*) did not allow the separation of clades A and B with a high bootstrap value. Thus, we compared the phylogenetic relationships of each core gene (2866 trees) and identified 55 genes allowing the placement of

175 isolates in clade A or B with a high bootstrap value. Among them, a gene encoding a putative
176 D-lactate dehydrogenase (VIBAEv3_A30718) was selected to explore the genetic structure of
177 *V. aestuarianus* using a larger collection of strains (n=116) isolated from diseased animals
178 (Table S2). Phylogenetic analyses reveal that 87/116 (75%) and 29/116 (25%) of these strains
179 belong to clade A and B, respectively (Fig.3). When injected intramuscularly to oysters at 10^2
180 CFU /animal, 81/87 (93%) and 23/29 (79%) strains from respectively clade A and B were
181 classified as Vir+ (Fig.3). The remaining strains (indicated with an asterisk in Fig.3) were
182 defined as intermediate (i.e. inducing >50% mortalities when injected at 10^7 CFU/animal,
183 M.A. Travers, pers.com.). The dominance of clade A and Vir+ strains (belonging to either
184 clade A or B) was observed during the summer mortality events and the more recent
185 outbreaks, whatever the age of the diseased oysters (> or <12months). Altogether these data
186 demonstrate that strains belonging to *V. aestuarianus* and isolated from diseased oysters can
187 be grouped into two lineages containing a majority of Vir+ strains. However, since we did not
188 observe any correlation between *V. aestuarianus* lethal dose, genotype and isolation date, the
189 hypothesis of the emergence of a new virulent clonal strain is unlikely.
190
191 An alternative hypothesis is that physiological alteration(s) of the oysters leading to an
192 increased sensitivity to *V. aestuarianus* may explain the increased number of reported cases of
193 mortalities. Such physiological disorders may result from environmental factors (acquired
194 sensitivity), genetic trade-offs (innate sensitivity) or a combination of both. Experimental
195 infections using wild stock of "naive" oysters that have never experienced the spat disease or
196 selected lineages resistant to one/several infectious agents may help in testing this hypothesis.
197 Finally the identification of habitat(s) and a spatio-temporal survey of *V. aestuarianus* should
198 help in understanding the ecological parameters that modulate virulence, persistence and/or
199 prevalence of this pathogen.

Non-virulent strains have undergone genetic modification(s)

Phylogenetic analysis of whole genomes revealed that virulent strains are grouped into two *V. aestuarianus* lineages, containing nearly identical strains. As each lineage contains a majority of highly virulent strains, we hypothesized that their common ancestor was virulent, and that a few modern strains might have undergone genetic modification(s) leading to loss of pathogenicity. We therefore performed comparative genomic analyses to identify these genetic modification(s).

In clade B, 49 genes localized in seven genomic regions were present in the Vir+ strain 12_063 but not in the intermediate strain 01_308 (Table S3). These regions encode common phage-related proteins (*e.g.,* integrase, helicase, relaxase and restriction endonuclease system) as well as other proteins of unknown function. However, none of these genes were found in the Vir+ strains from clade A. Finally, a frameshift was observed in 13 genes of strain 01_308, the majority of them coding for proteins of unknown function. However it should be noted that comparative genomic analyses within this clade B were hampered by the small number of sequenced strains (one Vir+ and one intermediate) and by genome fragmentation.

In clade A, we could not identify any genes specific of the Vir+ strains. However, in the Vir- strain 01_151, we detected a frameshift in three genes encoding respectively an exported protein of unknown function (VIBAEv3_A31414 in strain 02_041), a putative acetyltransferase (VIBAEv3_A10934) and a membrane protein of unknown function (VIBAEv3_A20116). Interestingly, a single frameshift was identified in the intermediate strain 07_115 in a gene that codes for a signal transduction histidine-protein kinase (VarS) (Lenz et al., 2005). The *varS* gene was found intact in the 13 others strains. The *varS* gene

(VIBAEv3_A30043 in strain 02_041) codes for a protein of 925 amino acids (aa) and contains six domains (Fig.4A): an uncharacterized signal transduction histidine kinase domain (DUF2222), a cytoplasmic helical linker and methyl-accepting protein domain (HAMP), a phosphoacceptor domain (HisKA), an ATPase domain (HATPase_c), a response regulator receiver domain (response reg) and a histidine-containing phosphotransfer domain (HPt). In the strain 07_115 the deletion of one nucleotide results in a stop codon, generating a 677 aa protein that lacks the response reg and HPt domains (Fig. 4A).

Disruption of *varS* is sufficient to alter *V. aestuarianus* pathogenicity

The sensory system VarS/VarA (VarS being the sensor histidine-kinase and VarA the response regulator) has been implicated in the pathogenicity of a variety of Gram-negative bacteria, including among others, *Escherichia coli* (BarA/UvrY), *Salmonella typhimurium* (BarA/SirA), and *Pseudomonas aeruginosa* (GacS/GacA) (Chavez et al., 2010; Gooderham and Hancock, 2009; Jones, 2005; Timmermans and Van Melderen, 2010). Hence, we assessed the importance of *varS* for *V. aestuarianus* virulence using a previously described genetic approach relying on a suicide vector, which can be transferred by conjugation to potentially any *Vibrio* strain (Le Roux et al., 2007). However, a dramatic difference in DNA delivery (10^{-4} to 10^{-6} transconjuguant per recipient cells) and allelic exchange efficiency (0 to 10^{-8} integration per recipients) was observed between nearly clonal strains. These data highlight the limitations of genetic methods when working with environmental non-model strains. Limitations can occur at several levels from the DNA delivery inside the cells to the allelic exchange efficiency and the availability of selective genes.

We obtained a successful integration of the suicide plasmid by a single crossover in only one out of seven virulent strains (12_016a), showing intra-specific variation. After the second

recombination event leading to plasmid excision, 30% of the colonies carried the *varS* deletion (strains 12_016a_Δ*varS*). For two isolates selected randomly, this deletion did not impair bacterial growth in culture media, but resulted in a dramatic decrease in mortality rates induced after bacteria injection in oysters (Fig. 4B, lanes 3 and 4 compared to lane 1). Upon constitutive expression of *varS in trans* from a replicative plasmid, the virulence of the mutant 12_016a_Δ*varS* was partially restored (Fig. 4B, lane 5 compared to 3) and this of the intermediate strain 07_115 was increased (Fig. 4B, lane 6 compared to 2). These complementation experiments confirmed that *varS* is necessary to 12_016a pathogenicity and that the frameshift in *varS* is also involved in the 07_115 virulence attenuation.

The metalloprotease Vam production and/or secretion is regulated by VarS

The two-component regulatory system VarA/S has been involved in the regulation of the secreted hemagglutinin/metalloprotease gene *hapA* in *V. cholerae* (Jang et al., 2010). Here, the protease activity measured in the extracellular products (ECPs) of the 12_016a_Δ*varS* mutants (Fig. 5B, lanes 2 and 3) was found to be 3 times lower than that of the wild type virulent strain 12_016a (Fig. 5B, lane 1) and in the range of the intermediate wild-type strain 07_115 (Fig. 5B, lane 4). The SDS-PAGE protein profiles of the ECPs prepared from 07_115 and two independent clones of 12_016a_Δ*varS* were found to be very similar and significantly different from this of 12_016a (Fig.5A). A band (25-35 kDa) found more intense in 12_016a (Fig.5A, lane 1) was excised from the gel, analyzed by μLC-ESI MS/MS and demonstrated to correspond to a peptide derived from the Vam metalloprotease (VIBAEv3_B10595 in strain 02_041) an homologue of HA/P, the *hapA* gene product.

The Vam metalloprotease of the *V. aestuarianus* strain 01_032 has been previously demonstrated to be lethal to *C. gigas* oysters (Labreuche et al., 2010). The expression of this

gene by a non-toxicogenic *Vibrio* strain (*V. tasmaniensis* LMG20012^T) induced the same immunosuppressant effects on hemocytes as those observed for *V. aestuarianus* ECPs showing that this protein is sufficient to induce immunosuppression in oysters (Labreuche et al., 2006). However the formal demonstration of the predicted, or supposed, role of a candidate gene requires a gene deletion strategy. Several attempts to generate a Δ*vam* mutant were unsuccessful (100% wild type reversion after the second recombination event), preventing the drawing of a definitive conclusion about the direct role of Vam in virulence. This suggests that the presence of this gene is essential in this strain in our culture conditions. It is important to note that the lack of a second usable resistance marker prevented the demonstration that a *vam* mutant could be constructed when the gene was provided *in trans*. We are currently exploring a larger panel of antibiotic resistance genes to allow the development of such strategy in the future.

Our results show that VarS is a key regulator of *V. aestuarianus* virulence and Vam secretion and/or activity and/or production. Several studies have examined the contribution to virulence of various *Vibrio* metalloproteases in animal experimental models (Finkelstein et al., 1992; Jeong et al., 2000; Le Roux et al., 2007; Milton et al., 1992; Shao and Hor, 2000) but no conclusive evidence about the role of proteases in virulence was found, since mutants deficient in secreted proteases showed comparable virulence levels to their parental strains. There are only a few examples of toxins (such as diphtheria or tetanus), which act as single determinants to produce disease. Microbial pathogenesis is often multifactorial, and pathogens use several biochemical mechanisms operating in concert to produce infections and diseases (Finlay and Falkow, 1997). For instance, the HA/P metalloprotease from *V. cholerae* was reported to activate proteolytically both the El Tor cytolysin/haemolysin (Nagamune et al., 1996) and the cholera toxin CT, an ADP-ribosylating enterotoxin inducing a highly

secretory diarrhea (Booth et al., 1984). Research is now ongoing to identify other genes that are regulated by VarS (at the transcriptional and post-transcriptional levels) and the protein targets that are processed by Vam in the ECP fraction.

Finally, due to the near identity of strains within clades A and B and existence of ~10% of Vir- isolates, *V. aestuarianus* appears as a great model to investigate by comparative genomic the genetic modification(s) leading to loss of pathogenicity and identify new virulence candidate genes and regulators. In the future, each of these genes will be deleted to investigate their potential respective role in virulence.

MATERIALS AND METHODS

Strains and culture conditions. The strains used for genomic analyses are described in Table 1. Other bacterial strains are described in Table S2 and S4. *Vibrio* isolates were grown in Zobell or Zobell agar, Luria-Bertani (LB) or LB-agar (LBA) + NaCl 0.5M, at 20°C. *Escherichia coli* strains were grown in LB or on LBA at 37°C. Chloramphenicol (5 to 25µg/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 medium, and conversely, repression was obtained by the addition of 1% D-glucose.

Genome sequencing, assembly and annotation. The complete genome sequence of 02_041 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 a 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. aestuarianus* reference genome. The 13 other *V. aestuarianus* strains were sequenced using the Illumina HiSeq2000 technology with a ~50-fold coverage. Contigs were assembled *de novo* using Velvet (Zerbino and Birney, 2008) and genome assembly was improved by contig mapping against the 02_041 reference genome. Computational prediction of coding sequences 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., 2013). 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 group computation visualized by cartographic maps to facilitate genome comparison.

***In silico* analyses.** To investigate the core and flexible genomes, an all-versus-all BlastP search was performed using genomic sequences of 209 *Vibrionaceae* and *Shewanella baltica* (strain OS155) available in Genbank and 14 *V. aestuarianus* 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 BMGE

(Criscuolo and Gribaldo, 2010). Phylogenetic trees were built using the parallel version of PhyML applied to Maximum-likelihood algorithm and GTR model as parameters (Guindon et al., 2010). Reliability was assessed by the bootstrap method with 100 replicates.

Vector construction. Cloning was performed using the Gibson assembly method according to the manufacturer's instructions (New England Biolabs, NEB). For the *varS* deletion, two independent PCR amplifications of the regions (500 bp) encompassing the *varS* gene were performed using two primer pairs (Δ*varS*-1+2 and Δ*varS*-3+4) (Table S4). An inside out PCR was performed using pSW7848T suicide vector DNA (Val et al., 2012) and primer pair (*pSW-F* and *pSW-R*) (Table S4). For the cloning of *varS* gene under a P_{LAC} promoter in a pMRB plasmid, a replicative plasmid found to be stable in *Vibrionaceae* (Le Roux et al., 2011), two independent PCR amplifications of the gene and plasmid were performed using the primers *varS-F*+R and *pMRB-F+R* reciprocally. After purification and quantification, 100 ng of each PCR product were mixed with the Gibson assembly Master Mix and incubated for 60 minutes at 50°C. Samples were diluted at 1/3 before *E. coli* transformation by the reaction product. Clones were controlled by digestion with restriction enzyme and sequencing using the primers described in Table S4. Strains Π3813 and β3914 were used as a plasmid host for cloning and conjugation, respectively (Le Roux et al., 2007). Plasmids and strains used and established in the present study are presented in Table S4.

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., 2007) 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. Selection of exconjugants and counter-selection of the Δ*dapA* donor was done by plating on a medium devoid of DAP, supplemented with chloramphenicol and 1% glucose. Cm^R resistant colonies were grown in LB + NaCl 0.5N up to late logarithmic phase and spread on plates containing 0.2% arabinose. Mutants were screened by PCR using primers Δ*varS*-1+4 (Table S4).

Extracellular products analyses. Bacterial extracellular products (ECPs) were produced by the cellophane overlay method as described previously (Le Roux et al., 2007). The protein concentration of the ECPs was measured by the method of Bradford with bovine serum albumin as the standard and normalized (BioRad). Protease activity was measured by the azocasein procedure as described previously (Miyoshi et al., 1987).

After concentration by ultrafiltration (Centricon® 10 Kda), twenty micrograms of crude ECPs were analyzed on a 4-15% Mini-PROTEAN® TGX Precast Gels. The differentially expressed protein band was manually excised from the gel, in-gel digested using trypsin and subjected to MS and MS/MS analyses for protein identification, following previously described protocols (Bernay et al., 2006).

Production of "pathogen free" oysters

Oysters (18 to 36 months; n=40) collected in the Fouras Bay (Marennes- Oléron, France) were transferred to the Ifremer facility located at Argenton (Brittany, France) for maturation conditioning as described previously (Petton et al., 2013). After gamete stripping and fertilization, obtained larvae, then spat, were reared under controlled conditions up to 12-13 months. PCR detection of oyster herpes virus was performed to confirm the negative status of oysters (Petton et al., 2013). *Vibrio* isolation on selective culture medium (Thiosulfate-citrate-bile salts-sucrose agar) confirmed a low *Vibrio* prevalence (~10 CFU/gr tissues).

Virulence studies using oysters. Bacteria were grown under constant agitation at 20°C for 24 h in Zobell. One hundred microliters of the diluted culture (10^7 to 10^2 CFU) were injected intramuscularly to anaesthetize SPF oysters (12-13 months old, 1.5 g, s.d. 0.2). The bacterial concentration was confirmed by conventional dilution plating on Zobell agar. After injection, the oysters were transferred to aquaria (10 oysters per aquarium) containing 2.5 liter of aerated 5 µm-filtered and UV-treated seawater at 20°C and kept under static conditions for 6 days. Each bacterial treatment was performed in duplicates and mortality was recorded daily.

ACKNOWLEDGEMENTS

We warmly thank Dr Maurizio Labbate (University of Technology, Sydney, Australia), Dr Maxime Bruto (Station Biologique de Roscoff) and Dr Marianne Alunno-Bruscia (Ifremer Argenton) for critically reading the manuscript. We also thank Pr Joël Henry and Dr Céline Zatylny-Gaudin (Université de Caen Basse-Normandie, Institut de Biologie Fondamentale et Appliquée) for kindly performing mass spectrometry analyses. We thank Dr Carla Pruzzo (University of Genoa, Italy) and Dr Ana Roque (IRTA, Spain) for providing U_17 and KB19 strains. We acknowledge the staff of the station Ifremer Argenton, La Tremblade, Bouin, particularly Max Nourry, the ABIMS (SBR Roscoff) and LABGeM (Evry) plateforms for technical support. The present study has been supported by the DPMA (Convention DPMA 2013- IFREMER 12/1210320/NYF), the ANR blanc (11-BSV7-023-01 « VIBRIOGEN », DG funding), the EMBRC France (AL funding) and the ANR bioadapt (13-ADAP-0007-01« OPOPOP»).

Author contributions

DG, MAT and AL contributed equally to this work.

426

427 **REFERENCES**

428 Austin, B. (2011). Taxonomy of bacterial fish pathogens. Vet Res 42, 20.

429 Bernay, B., Baudy-Floc'h, M., Gagnon, J., and Henry, J. (2006). Ovarian jelly-peptides
430 (OJPs), a new family of regulatory peptides identified in the cephalopod *Sepia*
431 *officinalis*. Peptides 27, 1259-1268.

432 Booth, B.A., Boesman-Finkelstein, M., and Finkelstein, R.A. (1984). *Vibrio cholerae*
433 hemagglutinin/protease nicks cholera enterotoxin. Infect Immun 45, 558-560.

434 Chavez, R.G., Alvarez, A.F., Romeo, T., and Georgellis, D. (2010). The physiological
435 stimulus for the BarA sensor kinase. Journal of bacteriology 192, 2009-2012.

436 Criscuolo, A., and Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with Entropy):
437 a new software for selection of phylogenetic informative regions from multiple
438 sequence alignments. BMC Evol Biol 10, 210.

439 Daubin, V., Gouy, M., and Perriere, G. (2002). A phylogenomic approach to bacterial
440 phylogeny: evidence of a core of genes sharing a common history. Genome research 12,
441 1080-1090.

442 Davis, B.M., and Waldor, M.K. (2003). Filamentous phages linked to virulence of *Vibrio*
443 *cholerae*. Current opinion in microbiology 6, 35-42.

444 Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high
445 throughput. Nucleic acids research 32, 1792-1797.

446 Finkelstein, R.A., Boesman-Finkelstein, M., Chang, Y., and Hase, C.C. (1992). *Vibrio*
447 *cholerae* hemagglutinin/protease, colonial variation, virulence, and detachment. Infect
448 Immun 60, 472-478.

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

451 Garnier, M., Labreuche, Y., Garcia, C., Robert, M., and Nicolas, J.L. (2007). Evidence for the
452 involvement of pathogenic bacteria in summer mortalities of the Pacific oyster
453 *Crassostrea gigas*. Microbial ecology 53, 187-196.
454 Gooderham, W.J., and Hancock, R.E. (2009). Regulation of virulence and antibiotic
455 resistance by two-component regulatory systems in *Pseudomonas aeruginosa.* FEMS
456 Microbiol Rev 33, 279-294.
457 Goudenege, D., Labreuche, Y., Krin, E., Ansquer, D., Mangenot, S., Calteau, A., Medigue,
458 C., Mazel, D., Polz, M.F., and Le Roux, F. (2013). Comparative genomics of
459 pathogenic lineages of *Vibrio nigripulchritudo* identifies virulence-associated traits. The
460 ISME journal 7, 1985-1996.
461 Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010).
462 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing
463 the performance of PhyML 3.0. Syst Biol 59, 307-321.
464 Harvell, C.D., Kim, K., Burkholder, J.M., Colwell, R.R., Epstein, P.R., Grimes, D.J.,
465 Hofmann, E.E., Lipp, E.K., Osterhaus, A.D., Overstreet, R.M.*, et al.* (1999). Emerging
466 marine diseases--climate links and anthropogenic factors. Science 285, 1505-1510.
467 Jang, J., Jung, K.T., Yoo, C.K., and Rhie, G.E. (2010). Regulation of hemagglutinin/protease
468 expression by the VarS/VarA-CsrA/B/C/D system in *Vibrio cholerae*. Microb Pathog
469 48, 245-250.
470 Jeong, K.C., Jeong, H.S., Rhee, J.H., Lee, S.E., Chung, S.S., Starks, A.M., Escudero, G.M.,
471 Gulig, P.A., and Choi, S.H. (2000). Construction and phenotypic evaluation of a *Vibrio*
472 *vulnificus vvpE* mutant for elastolytic protease. Infect Immun 68, 5096-5106.
473 Jones, B.D. (2005). *Salmonella* invasion gene regulation: a story of environmental awareness.
474 Journal of microbiology 43 Spec No, 110-117.

475 Keim, P.S., and Wagner, D.M. (2009). Humans and evolutionary and ecological forces
476 shaped the phylogeography of recently emerged diseases. Nat Rev Microbiol 7, 813-
477 821.
478 Konstantinidis, K.T., and Tiedje, J.M. (2005). Genomic insights that advance the species
479 definition for prokaryotes. Proc Natl Acad Sci U S A 102, 2567-2572.
480 Labreuche, Y., Le Roux, F., Henry, J., Zatylny, C., Huvet, A., Lambert, C., Soudant, P.,
481 Mazel, D., and Nicolas, J.L. (2010). *Vibrio aestuarianus* zinc metalloprotease causes
482 lethality in the Pacific oyster *Crassostrea gigas* and impairs the host cellular immune
483 defenses. Fish Shellfish Immunol 29, 753-758.
484 Labreuche, Y., Soudant, P., Goncalves, M., Lambert, C., and Nicolas, J.L. (2006). Effects of
485 extracellular products from the pathogenic *Vibrio aestuarianus* strain 01/32 on lethality
486 and cellular immune responses of the oyster *Crassostrea gigas*. Dev Comp Immunol 30,
487 367-379.
488 Le Roux, F., Binesse, J., Saulnier, D., and Mazel, D. (2007). Construction of a *Vibrio*
489 *splendidus* mutant lacking the metalloprotease gene *vsm* by use of a novel
490 counterselectable suicide vector. Applied and environmental microbiology 73, 777-784.
491 Le Roux, F., Davis, B.M., and Waldor, M.K. (2011). Conserved small RNAs govern
492 replication and incompatibility of a diverse new plasmid family from marine bacteria.
493 Nucleic acids research 39, 1004-1013.
494 Lenz, D.H., Miller, M.B., Zhu, J., Kulkarni, R.V., and Bassler, B.L. (2005). CsrA and three
495 redundant small RNAs regulate quorum sensing in *Vibrio cholerae.* Molecular
496 microbiology 58, 1186-1202.
497 Mazel, D., Dychinco, B., Webb, V.A., and Davies, J. (1998). A distinctive class of integron in
498 the *Vibrio cholerae* genome. Science 280, 605-608.

499 Milton, D.L., Norqvist, A., and Wolf-Watz, H. (1992). Cloning of a metalloprotease gene
500 involved in the virulence mechanism of *Vibrio anguillarum.* Journal of bacteriology
501 174, 7235-7244.
502 Miyoshi, N., Shimizu, C., Miyoshi, S., and Shinoda, S. (1987). Purification and
503 characterization of *Vibrio vulnificus* protease. Microbiol Immunol 31, 13-25.
504 Nagamune, K., Yamamoto, K., Naka, A., Matsuyama, J., Miwatani, T., and Honda, T. (1996).
505 *In vitro* proteolytic processing and activation of the recombinant precursor of El Tor
506 cytolysin/hemolysin (pro-HlyA) of *Vibrio cholerae* by soluble hemagglutinin/protease
507 of *V. cholerae,* trypsin, and other proteases. Infect Immun 64, 4655-4658.
508 Petton, B., Pernet, F., Robert, R., and Boudry, P. (2013). Temperature influence on pathogen
509 transmission and subsequent mortalities in juvenile Pacific oysters *Crassostrea gigas*.
510 Aquacult Environ Interact 3, 257-273.
511 Renault, T. (2011). A review of mortality outbreaks in the Pacific oyster, Crassostrea gigas,
512 reported since 2008 in various European Union Member States and the related
513 implementation of Council Directive 2008/88/EC. Bulletin OIE 4, 51-52.
514 Sawabe, T., Ogura, Y., Matsumura, Y., Feng, G., Amin, A.R., Mino, S., Nakagawa, S.,
515 Kumar, R., Fukui, Y., Satomi, M.*, et al.* (2013). Updating the *Vibrio* clades defined by
516 multilocus sequence phylogeny: proposal of eight new clades, and the description of
517 *Vibrio tritonius* sp. nov. Front Microbiol 4, 414.
518 Segarra, A., Pepin, J.F., Arzul, I., Morga, B., Faury, N., and Renault, T. (2010). Detection and
519 description of a particular *Ostreid herpesvirus* 1 genotype associated with massive
520 mortality outbreaks of Pacific oysters, *Crassostrea gigas,* in France in 2008. Virus Res
521 153, 92-99.
522 Shao, C.P., and Hor, L.I. (2000). Metalloprotease is not essential for *Vibrio vulnificus*
523 virulence in mice. Infect Immun 68, 3569-3573.

524 Timmermans, J., and Van Melderen, L. (2010). Post-transcriptional global regulation by CsrA
525 in bacteria. Cellular and molecular life sciences : CMLS 67, 2897-2908.
526 Val, M.E., Skovgaard, O., Ducos-Galand, M., Bland, M.J., and Mazel, D. (2012). Genome
527 engineering in *Vibrio cholerae*: a feasible approach to address biological issues. PLoS
528 Genet 8, e1002472.
529 Vallenet, D., Belda, E., Calteau, A., Cruveiller, S., Engelen, S., Lajus, A., Le Fevre, F.,
530 Longin, C., Mornico, D., Roche, D.*, et al.* (2013). MicroScope--an integrated microbial
531 resource for the curation and comparative analysis of genomic and metabolic data.
532 Nucleic acids research 41, D636-647.
533 Vezzulli, L., Pezzati, E., Stauder, M., Stagnaro, L., Venier, P., and Pruzzo, C. (2014). Aquatic
534 ecology of the oyster pathogens *Vibrio splendidus* and *Vibrio aestuarianus.*
535 Environmental microbiology.
536 Zerbino, D.R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly
537 using de Bruijn graphs. Genome research *18*, 821-829.
538
539
540

TITLES AND LEGENDS TO FIGURES

Figure 1: Oyster mortality in response to experimental infection by *V. aestuarianus* strains selected for genome sequencing. 10^7 CFU (black bar) or 10^2 CFU (grey bar) of the tested strain were injected intramuscularly into oysters (n= 10, in duplicate). Cumulative mortality (%) was assessed after 6 days. Strains were classified as virulent (Vir+) (i.e. inducing >50% mortalities at 10^2 CFU /animal), non-virulent (Vir-) (i.e. inducing <50% mortalities at 10^7 CFU /animal) or intermediate (i.e. pathogenic only at 10^7 CFU /animal).

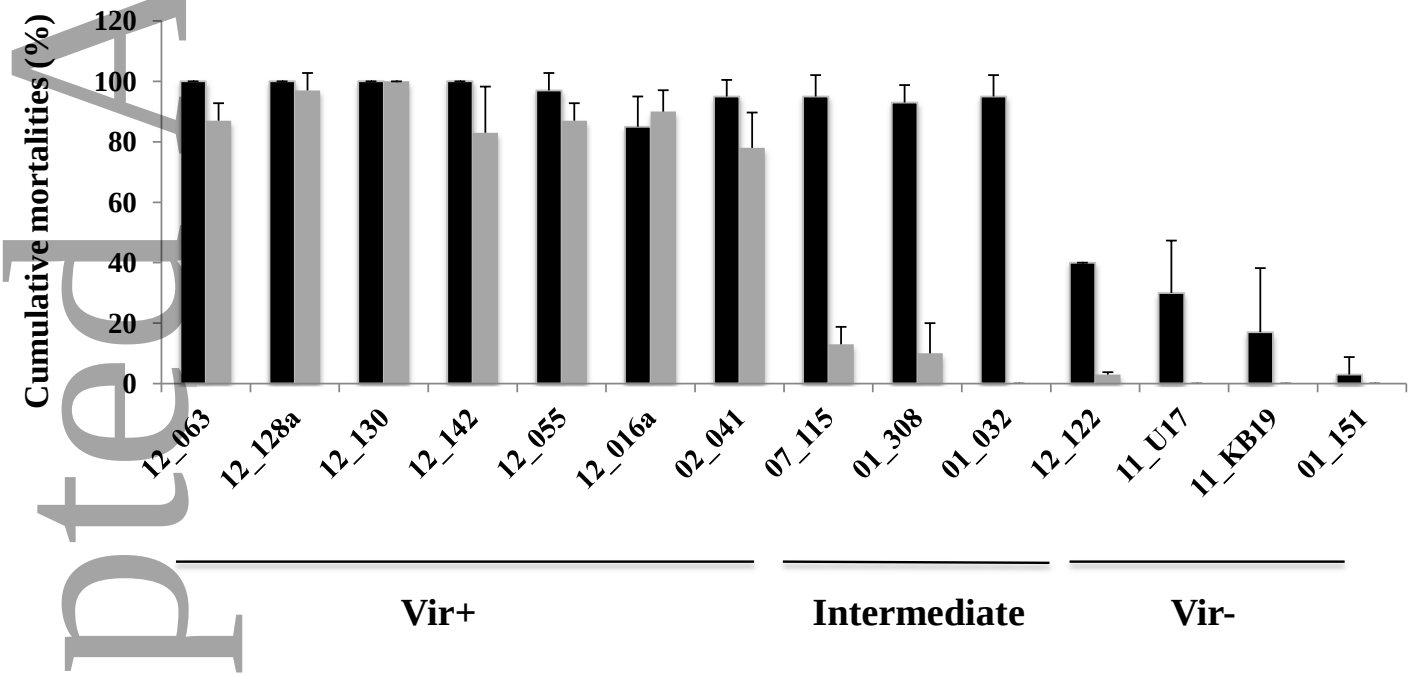
Figure 2: Phylogenetic analysis based on concatenated alignments of nucleic acid sequences of 2866 core genes from 14 *V. aestuarianus* strains and KB19 as an outgroup. The tree was built by the Maximum-Likelihood method based on a sequence alignment generated by Muscle. Branch lengths are drawn to scale 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 *V. aestuarianus* strain (Vir+: virulent; Vir-: non virulent; int: intermediate) is indicated in parentheses.

Figure 3: *V. aestuarianus* isolates phylogeny analysis of partial D-lactate dehydrogenase gene sequences and virulence status. The tree was built by the Maximum-Likelihood method based on a sequence alignment generated by 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. The black bars indicate the % of mortalities occurring at 6 days post-injection (10^2 CFU/animal). The strains indicated with

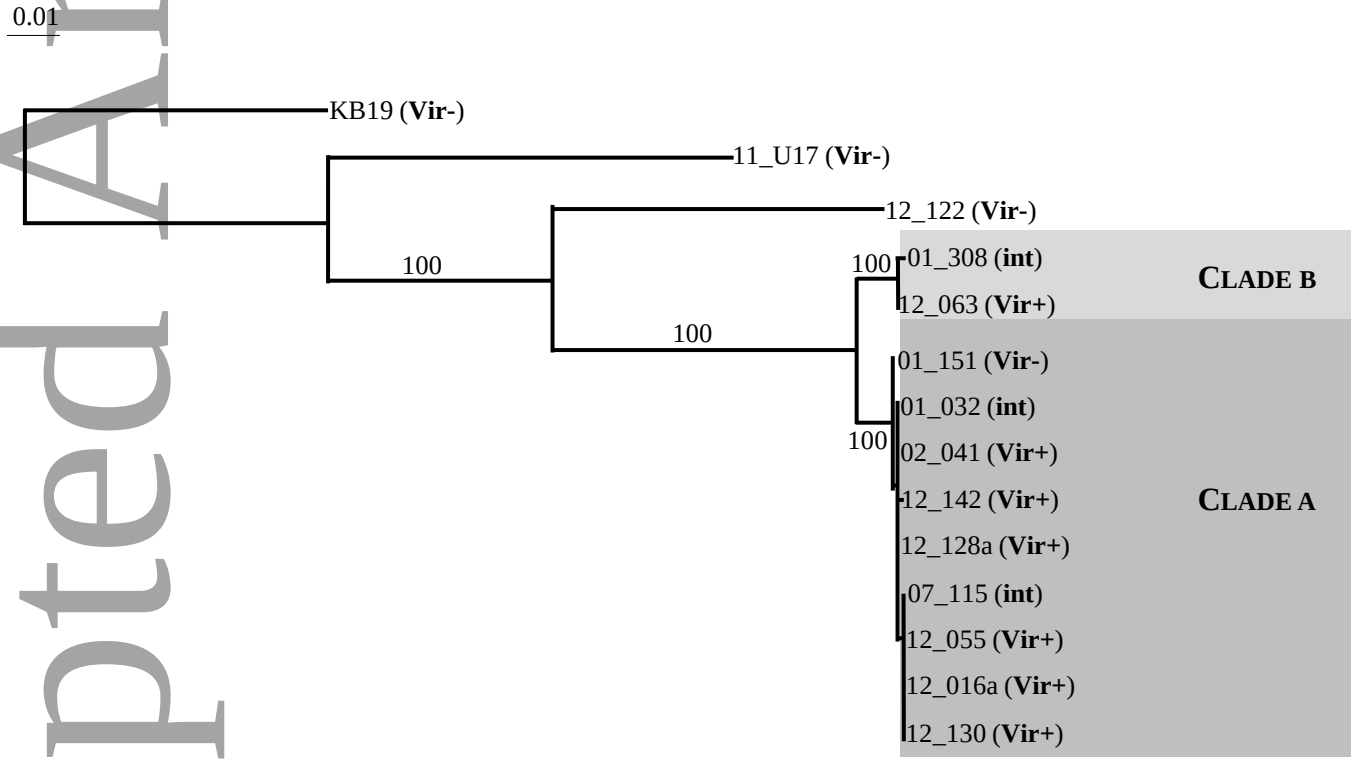
an asterisk were defined as intermediate, *i.e.* inducing >50% mortalities when injected at 10^7 CFU/animal.

Figure 4: Role of *varS* in *Vibrio aestuarianus* pathogenicity**. A-** Schematic representation of the functional domains of VarS identified in the Vir+ strain 02_041 and the truncated protein resulting from a frameshift in the intermediate strain 07_115. DUF2222 corresponds to an uncharacterized signal transduction histidine kinase domain; HAMP, a cytoplasmic helical linker and methyl-accepting protein domain; HisKA, a phosphoacceptor domain; HATPase_c, an ATPase domain; Response reg, a response regulator receiver domain; HPt, an histidine-containing phosphotransfer domain. **B-** Experimental infection of wild type *V. aestuarianus,* Δ*varS* mutants and complemented Δ*varS* mutants. 10^2 CFU of the tested strains (lane 1: strain 12_016a wild type; lane 2: strain 07_115 wild type; lanes 3 and 4: GV1124 and 1125, two independent clones of 12_016a_ Δ*varS*; lane 5: GV1124 i.e. 12_016a_ Δ*varS* carrying an expression vector for *varS*, pMRB-P_{LAC}*varS*; lane 6: 07_115 carrying pMRB-P_{LAC}*varS)* was intramuscularly injected into oysters (n= 20, in duplicate). Mortality (%) was assessed after 6 days.

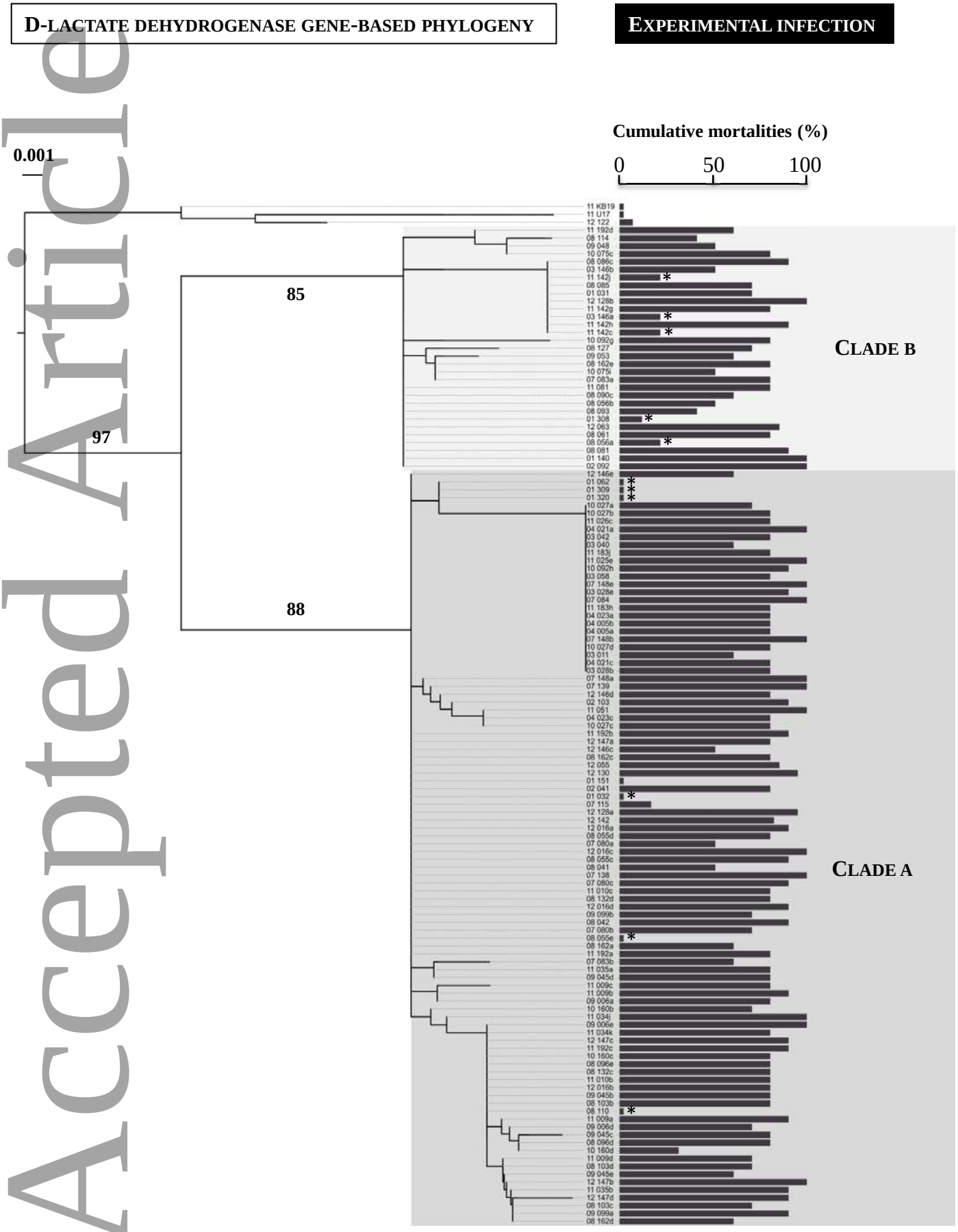
Figure 5: Role of *varS* in *Vibrio aestuarianus* metalloprotease expression**. A-** Extracellular product analysis by Coomassie blue stained 10% SDS-PAGE gel (lane 1: strain 12_016a wild type; lane 2 and 3: GV1124 and 1125, two distinct clones of 12_016a_ Δ*varS*; lane 4: 07_115 wild type). Arrow indicates the Vam metalloprotease identified by MS/MS. **B-** Proteolytic activities of ECPs was determined by an azocasein assay as described in Materials and Methods (absorbance at 440 nm).



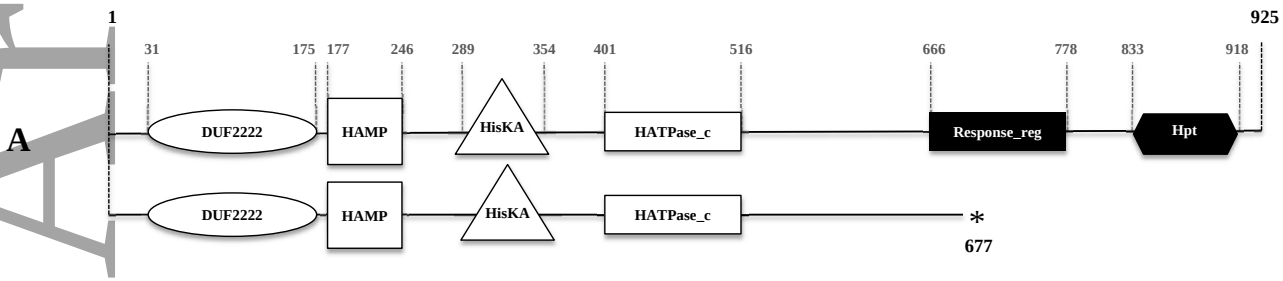
EML_12699_F1.pdf

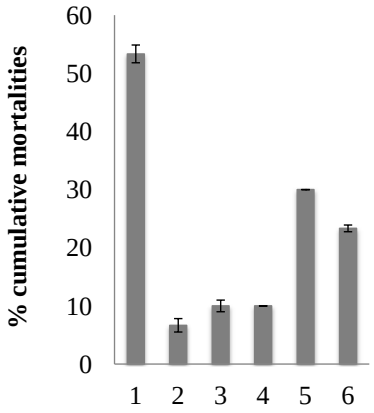


EMI_12699_F2.pdf

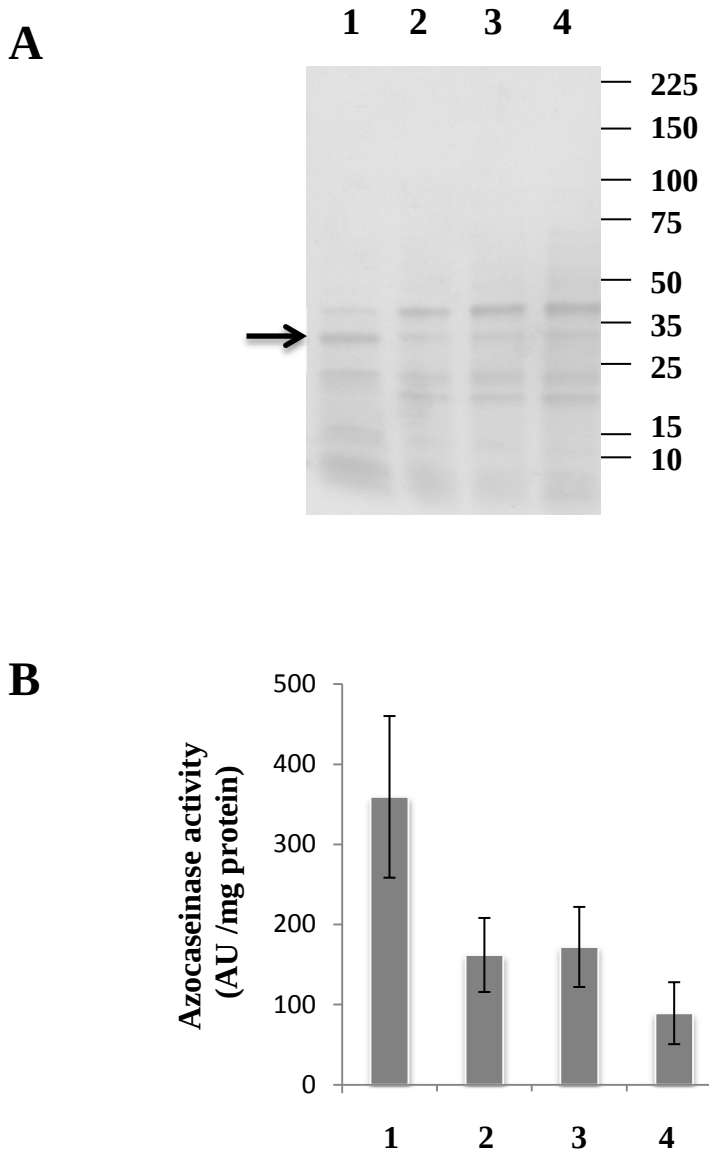


EMI_12699_F3.pdf





EMI_12699_F4.pdf



EMI_12699_F5.pdf

Context	Strain	Origin	Mortality on field	Contigs number	Genome size (Mb)	Cl
Mortality outbreaks 2001-02	01_032	Oyster, September 2001, Argenton, Brittany, France	yes	38	4.20	41
	01_151	Oyster, July 2001, La Trinité, Brittany, France	yes	73	4.36	43
	01_308	Oyster, August 2001, Normandy, France	yes	157	4.49	45
	02_041	***Oyster, 2002, Argenton, Brittany, France***	***yes***	***8***	***4.20***	***40***
Mortality outbreaks 2012	12_016a	Oyster, March 2012, La Tremblade, Charente Maritime, France	yes	52	4.25	42
	12_055	Oyster, June 2012, Agnas, Charente Maritime, France	yes	50	4.25	42
	12_063	Oyster, September 2012, Brest, Brittany, France	yes	141	4.51	45
	12_128a	Oyster, September 2012, Brittany, France	yes	65	4.24	42
	12_130	Oyster, September 2012, Agnas, Charente Maritime, France	yes	80	4.29	43
	12_142	Oyster, Octobre 2012, Normandy, France	yes	115	4.38	44
Other	07_115	Oyster, 2007, Brittany, France	no	44	4.24	42
	11_KB19	Oyster, March 2011, Fangar Bay, Spain	no	707	4.99	52
	11_U17	Zooplankton, May 2011, Goro lagoon, Italy	no	732	4.41	44
	12_122	Cockle, August 2012, Brittany, France	no	399	4.90	49

Table 1: Strains used in the study