PhD THESIS

Miguel Carda Diéguez

V i b r i o   v u l n i f i c u s

from water *to* host

VNIVERSITAT D\G VALÈNCIA

SUPERVISORS

Carmen Amaro González
Francisco Rodríguez Valera

PHD PROGRAM OF BIOTECHNOLOGY

VALENCIA 2016

La Dra. Carmen Amaro González, catedrática del departamento de Microbiología y Ecología y directora del grupo de Patógenos Acuáticos del Estructura de Recerca Interdisciplinar en Biotecnología i Biomedicina (ERI BIOTECMED), ambos de la Universidad de Valencia, certifica que D. Miguel Carda Diéguez ha realizado bajo su dirección el trabajo titulado "*Vibrio vulnificus*: **from water to host**" y autoriza la lectura y defensa de la misma para optar al grado de Doctor en Ciencias Biológicas por la Universidad de Valencia.

Y para que así conste a los efectos oportunos, firma la presente en Valencia en Noviembre de 2016.

Fdo. Carmen Amaro González

El Dr. Francisco Eduardo Rodríguez Valera, catedrático del departamento de Producción Vegetal y Microbiología de la Universidad Miguel Hernández de Alicante, certifica que D. Miguel Carda Diéguez ha realizado bajo su dirección el trabajo titulado "*Vibrio vulnificus*: **from water to host**" y autoriza la lectura y defensa de la misma para optar al grado de Doctor en Ciencias Biológicas por la Universidad de Valencia.

Y para que así conste a los efectos oportunos, firma la presente en Alicante en Noviembre de 2016.

Fdo. Francisco Eduardo Rodríguez Valera

AKNOWLEDGMENTS/AGRADECIMIENTOS

En primer lugar, agradecer a toda mi familia por el apoyo moral y económico. En especial a mis padres sin los cuales no sería quien soy. Gracias a ellos he podido estudiar y formarme como biólogo y persona. Gracias por enseñarme a ser tenaz y a que no hay dibujo del mono de Aladdin que no se pueda dibujar y que no hay más impedimento que uno mismo frente a las barreras del día a día.

Mis más sinceros agradecimientos a mis directores Carmen y Paco por enseñarme tantas y tantas cosas y dedicar su tiempo a hacer de mi un microbiólogo con capacidad crítica. Estoy muy feliz por haber conocido este mundo y es gracias a que me disteis la oportunidad, sobretodo Carmen. Por supuesto a mis compañeros de trabajo que han de soportar los malos humos que tiene uno cuando día tras día no salen los experimentos. De estos, quiero destacar a David Pajuelo, compañero y gran amigo sin el cual no puedo ni imaginar cómo hubiese sido tanto tiempo en el laboratorio. Esos paseos hasta el Lasmi eran sin duda de los mejores momentos del día. Además, Amparo, Paco Roig, Celia, Eva, Carla, Paco Mex, Duarte, etc. gracias a todos por cada día y cada momento compartido conmigo.

Cuando empecé en el laboratorio de Carmen quería dedicarme a la investigación aplicada y no me veía haciendo ciencia básica así que empecé con Belén que me enseñó a trabajar con y para las empresas. Con ella aprendí desde el primer momento a trabajar para obtener siempre el máximo nivel. Gracias por eso también.

De mi estancia en San Juan de Alicante, por supuesto a Paco por darme la oportunidad de investigar en su laboratorio y aprender todas las aplicaciones y posibilidades de la bioinformática. Allí aprendí a trabajar con grandes investigadores y personas que hicieron de mi estancia no solo una gran experiencia científica sino también personal. De todas las grandes personas que compartieron sus conocimientos y tiempo conmigo fueron muy especiales Rohit, Carol y Ana. Los tres dedicaron tiempo a enseñarme y Paco, Rohit y Carol además compartieron su tiempo libre, lo cual les agradeceré siempre.

Mi segunda estancia fue en Boston…so…it was an honour for me to work with someone like Dr. Waldor, sorry, Matt! I understood how a laboratory with international recognition has to work. Those three months have probably been the best of my professional life. Mike and Troy helped me to make three month of work worth like a year. Thank you so so so so much! The intense three months were especially hard because of the distance from home and family and Matt Gerding made me feel like home. Those Thursdays in Penguins watching the NBA are part of the best memories that I took from Boston. Thanks to you and everyone in the laboratory.

En general, gracias a cualquier persona del departamento de Microbiología y Ecología y de la Universidad de Valencia, Miguel Hernández o Brigham and Women's Hospital que sienta que en algún momento me ayudó. También a la gente de la Consellería que me ayudó a pescar las anguilas.

Gracias también a mis amigos, tanto de Villarreal como a la COMP porque a veces es difícil desconectar de este trabajo pero con ellos es muy muy fácil. En especial a Fabián y Christian que me han ayudado en la impresión y diseño de la tesis, respectivamente. También a Natalia y Dani, compañeros de la carrera y del doctorado que siempre han estado ahí para compartir penas y alegrías.

Cómo no, gracias a mi mejor amiga y pareja que me ha hecho mejor persona cada día que hemos compartido y que sabe sacar lo mejor de mí. Gracias por aguantar…bueno, todos los momentos que has tenido que aguantar por tener un novio becario-científico. Gracias por entender los sábados, domingos, festivos, estancias, etc. que esta tesis te ha robado. Te quiero.

ABSTRACT

Vibrio vulnificus is an aquatic pathogen autochthonous from temperate, tropical and subtropical ecosystems where it lives either as a sessile cell, forming biofilms on biotic/abiotic surfaces, or as a free-swimming cell. From these locations, the pathogen can occasionally infect humans and fish (mainly farmed-eels) causing a disease named **vibriosis**. The most severe form of human and fish vibriosis is associated with the pathogen's ability to spread from the infection site to the bloodstream and multiply, process known as **invasion**.

Before invasion, the pathogen has to **colonize** the mucosal host surface, process that involves not only bacterial attachment/adhesion but also resistance to mucosal immunity, commensal microbiota (competitors) and bacterial predators (mainly amoeba and phages). Recently, Amaro and cols. obtained evidence that supports that mucin, the main protein in mucus, can activate horizontal gene transference (HGT) in *V. vulnificus*, which could lead to the emergence of new virulent clones in natural mucosal environments.

The objective of this thesis was to study the colonization and invasion processes under the global perspective that allow the "omic" technologies. In the first chapter, we focused our attention on a selected host for the pathogen in the aquatic environment, the eel, and analyzed its microbiome by using metagenomics. We describe for the first time the microbiome of the skin mucus of wild- and farmed-eels and compared it with that of the water. We discovered that mucus concentrates, selectively, bacteria present in water (mainly vibrios if present) and identified the genes involved in a successful colonization process, most of which could be considered virulence genes. Then, we developed a protocol to identify MGE and prophages in the metagenomes and described a series of putative ICEs, pathogenicity islands and prophages some of which contained virulence and antibiotic resistance genes. Finally, we were able to describe multiple lytic phages in skin-mucus, which could be considered as a part of the mucosal immunity.

The second chapter of this thesis is focused on the invasion, and, in particular, in the ability of this pathogen to grow in blood. The objective was to discover all the pathogen's genes involved in growth in human blood by applying the "omic" approach known as transposon insertion sequencing (TIS). TIS is a powerful method that couples high-density transposon mutagenesis with next-generation sequencing to comprehensively assess the fitness of thousands of transposon mutants across a genome. TIS considers that the essentiality of a gene in a condition should be proportional to its presence in the library. Accordingly, genes present in both control (LB-1 or inactivated HS) and tested (fresh HS) libraries are not essential or neutral, and genes present only in the control library are essential for growth in blood and, consequently, for invasion. We identified the genes, obtained the corresponding mutants and complemented strains and performed a series of experiments *in vitro* and *in vivo* to corroborate the essentiality of the gene and its function (if possible). We identified a series of genes involved in the CPS formation and others with unknown function as essential for human invasion. Moreover, a series of functions have been proposed as important for growth in human blood.

ABREBIATIONS

Abs$_x$	Absorbance at x nm
ANI	Average nucleotide identity
APA	Alkaline Peptone Water
ApE	A plasmid Editor
ARTIST	Analysis of high-Resolution Transposon-Insertion Sequences Technique
ASW	Artificial sea water
BACON	Bacteroidetes-Associated Carbohydrate-binding Often N-terminal
BALB/c	Bagg Albino
BAM	bacteriophage adherence to mucus
BCA	Bicinchoninic acid assay
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
CDC	Centers for Disease Control and Prevention
CDD	Conserved Domain Database
cDNA	complementary DNA
cfu	colony forming units
CON-ARTIST	Conditionally essential loci analysis
cpMGE	contigs with putative MGE
CPS	Capsule
CR	Complement resistant
CRISPR	clustered regularly interspaced short palindromic repeats
DE	Domain Essential
dH$_2$O	Distilled water
DNApol	DNA polymerase
dNTPs	Deoxynucleotide triphosphate
E	Essential
EL-ARTIST	Essential Loci analysis
FAO	Food and Agriculture Organization
FC	Fold change
Fe$_y^x$	farmed eels fished in water with y g/l of salinity and pH x
GC	guanine cytosine
GO	Gene ontology
GOS	Global Ocean Sampling
HGT	Horizontal gene transfer
HUP	hexanucleotide usage pattern
HS	Human serum

ICE	integrative and conjugative elements
ICTV	International Committee on Taxonomy of Viruses
IHS	Inactivated Human Serum
Ig	Immunoglobulin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	Kilobase
KDa	Kilo Dalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria Bertani broth
LD_{50}	Lethal dose 50
LPS	Lipopolysaccharide
LSU	Large subunit 16S gene
Mb	Megabase
MGE	Mobile genetic elements
MG-RAST	Metagenomic Rapid Annotations using Subsystems Technology
NCBI	National Center for Biotechnology Information
NE	No essential or Neutral
NQR	Na+-translocating NADH-ubiquinone reductase
NR	non-redundant
OM	Outer membrane
Op	Opaque
ORF	Open Reading Frame
OTU	Operational taxonomic units
PAE	Planta Acuarios y Experimentación
PBS	Phosphate buffered saline
PCA	principal component analysis
PCR	Polymerase chain reaction
PHYLIP	PHYLogeny Inference Package
PVDF	polyvinylidene difluoride
qPCR	Quantitative Polymerase chain reaction
qRT-PCR	Quantitative reverse-transcription Polymerase chain reaction
RDP	Ribosomal Database Protein
RNApol	RNA polymerase
RPKG	reads/kilobase/Gigabases
rRNA	ribosomal RNA
RTX	repeat in toxin
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate- Polyacrilamide gel electrophoresis
sgMGE	signature genes for MGE
TA	thymine adenine

TCBS	thiosulfate-citrate-bile salts-sucrose
TIS	transposon insertion sequencing
Tr	Translucent
tRNA	transfer RNA
TSA	trypticase soy agar
TUP	tetranucleotide usage pattern
TXSS	Type X secretion system
Vc	*Vibrio cholerae*
Vp	*Vibrio parahaemolyticus*
Vv	*Vibrio vulnificus*
VVM	*Vibrio vulnificus* medium
WE_y^x	wild eels fished in water with y g/l of salinity and pH x

CONCEPTS

Microbiota: Microorganisms from an environment that are taxonomically identified by 16S rRNA or other bioinformatic analysis.

Metagenome: the genes and genomes of the microbiota, including plasmids, mobile genetic elements and virus. In the same context, the viral content will be designated as virome.

Estuarine-water: a mixture of fresh-water draining from land and salty seawater (salinity between 3 and 35 g/l)

Natural fresh-water: waters from rivers, lakes, natural ponds etc. without supply from seawater (salinity below 1 g/l).

Colonization: Ability of a pathogen to attach to a tissue, resist the local immunity (including host innate immunity and biological interactions with commensal microbiota) and multiply on the tissue.

Invasion: Ability of a pathogen to spread from an epithelium to the blood stream, resist the bactericidal and bacteriological action of the blood's innate immunity (including phagocytosis, complement and transferrin) and multiply in blood.

Primary pathogen: A pathogen that causes a disease in non-immune healthy individuals.

Secondary pathogen: A pathogen that only causes diseases in an immune-depressed host.

Opportunistic pathogen: A pathogen that causes disease whenever it has an opportunity: i.e. a drop in host's immune defenses or an access to internal tissues through an injury, or both.

Obligate pathogen: A pathogen that depends on a host, in which it can cause or not a disease, for its survival as a species

Accidental pathogen: A pathogen that does not depend on a host, in which it can cause or not a disease, for its survival as a species

Nutritional immunity: The term "nutritional immunity" includes all the strategies displayed by the host in order to make nutrients inaccessible for bacterial growth. Nutrients would include from trace elements like iron, zinc and

manganese (usually sequestered by proteins) up to limiting carbon and/or nitrogen sources.

INDEX

INTRODUCTION

1. *Vibrio vulnificus*

1.1. Ecology

V. vulnificus is an aquatic bacterium that inhabits temperate, tropical and subtropical ecosystems where it lives either as a planktonic (free-swimming) or as a sessile bacterium, forming biofilms on different biological (fish, zooplankton, algae..) and non-biological surfaces (sediments…) (Bisharat *et al*., 1999; Marco-Noales *et al*., 2001; Neiman *et al*., 2011; Morris, 2003; Arias *et al*., 1999; Hor *et al*., 1995; Oliver *et al*., 1983; Oliver, 2005). The survival of this species in water is dependent on salinity (optimum 0.5-1 %NaCl) and temperature (optimum 25ºC). This bacterium enters into viable but not cultivable state at salinities below 0.1 %NaCl and temperatures below 4ºC (Randa *et al*., 2004; Marco-Noales *et al*., 1999).

V. vulnificus is worldwide recognized as a human and animal pathogen of health importance, especially in tropical countries located in the warmest areas of the planet (Tao *et al*., 2012; Kaspar and Tamplin, 1993, http://www.floridahealth.gov/diseases-and-conditions/vibrio-infections/vibrio-vulnificus/). However, there is cumulative evidence showing that *V. vulnificus* geographical distribution is expanding to new areas such as Northern European countries and the Atlantic coast of the United States and that this expansion is mainly due to global warming (Vezzulli *et al*., 2016). In parallel, the number of reported animal and human vibriosis in these areas, especially in the hottest months, has dramatically increased (Roux *et al*., 2015; Baker-Austin *et al*., 2010). For instance, Baker-Austin highlighted that more than 60 cases of human infections occurred in the Baltic Sea in 2006, one of the warmest summers in Baltic Sea history (Baker-Austin *et al*., 2012). Increase in life expectancy of people, HIV expansion, increase in raw seafood consumption and probably improved detection methods have been proposed as alternative or complementary explanations for this detected increase in *V. vulnificus* incidence.

1.2. Biotypes

INTRODUCTION

V. vulnificus is a highly heterogeneous species that has been subdivided into three biotypes on the basis of differences in host-range, phenotypic traits (indole, ornithine decarboxylase activity, growth at 42 ºC and acid production from mannitol, sorbitol and cellobiose) and geographic distribution (Tison *et al*., 1982; Bisharat *et al*., 1999). Biotype 1 is distributed worldwide and comprises the strains isolated from environmental and human samples that are positive for all the above mentioned phenotypic tests; Biotype 2 is also worldwide distributed and comprises all the fish virulent strains, which are mostly negative for the mentioned tests except cellobiose fermentation; and biotype 3 that is geographically restricted to Israel and comprises strains potentially virulent for humans that are mostly positive for the mentioned tests with the exception of cellobiose fermentation. Several serotyping systems have been described for this species (Simonson and Siebeling, 1986; Martin and Siebeling, 1991) . According to the serotyping scheme of Biosca *et al*. (1996), biotype 1 is serologically very diverse while biotype 2 only contains three different serovars (serovars E, A and I) and biotype 3 with a unique serovar (serovar O).

1.3. Vibriosis

The diseases caused by *V. vulnificus* are collectively known as vibriosis. Human vibriosis can be presented either as severe wound infections after contact with contaminated seawater (vibriosis type I) or as gastroenteritis after raw seafood ingestion (vibriosis type II). Both types of vibriosis have in common that they can lead to death by sepsis, especially in patients with high iron levels in blood, mainly due to hepatic or hematological disorders (hemochromatosis, hepatitis etc). Vibriosis type I is frequently contracted as a result of recreational swimming, fishing injuries, or seafood handling (Horseman and Surani, 2011). The infection in risk patients progresses rapidly to cellulitis, ecchymosis, and bullae, which can further progress to necrotizing fasciitis at the site of infection with a mortality rate around 25% (Bowdre *et al*., 1981). The antibacterial agents used to stop the infection and its spread are doxycycline, ceftazidime, and clindamycin (Klontz *et al*., 1988; Chiang and Chuang, 2003). Intensive and prompt wound care is essential. Surgical debridement, incision and drainage of

22

abscesses and, sometimes, amputation have proved to increase survival odds. For all these reasons, this species is popularly known in the USA as the "**marine flesh-eating bacterium**" (**Figure 1**). Vibriosis type II in risk patients can result, directly, in primary sepsis, a severe, fulminant systemic infection. Characteristics of this disease include fever, chills, nausea, hypotensive septic shock, and the formation of secondary lesions on the extremities of patients (Horseman and Surani, 2011). This primary septicemia is the most lethal infection caused by *V. vulnificus*, with an average mortality rate exceeding 50% (Horseman and Surani, 2011). Most times, antibiotic treatment is ineffective and patients have to be admitted in the intensive care. The main known reservoir for vibriosis type I is seawater and for vibriosis type II is the oyster.

Figure 1. Human vibriosis. Left leg of a patient soon after being admitted to the emergency room showing redness and swelling. Right, after multiple surgeries to remove affected tissue and weeks of ongoing wound care. Photo courtesy of Louisiana Sportsman.

Biotype 2causes **vibriosis** in several species of teleost (**fish vibriosis**), especially under captivity in fish farms. The main hosts for fish vibriosis are the **eel** species, *Anguilla anguilla* and *A. japonica* (Tison *et al*., 1982; Bisharat *et al*., 1999; Heidelberg *et al*., 2002). Fish vibriosis is a primary hemorrhagic septicaemia that occurs in farms as outbreaks of epizootics of high mortality. Because of the high density of fish in aquaculture facilities, the entry of a single pathogen can spread and infect rapidly, causing enormous costs for the company and sometimes bankruptcy. Remarkably, several cases of human vibriosis infections through diseased eel handling have also been reported in Europe suggesting that this biotype is a zoonotic agent (Roux *et al*., 2015). Aquaculture has grown in the last years and has been established in the market as an essential supplier of fish and seafood for consumption (**Figure 2**). Studies at global scale

suggest that aquaculture would pass fishing as the main supplier in the next decade (Food and Agriculture Organisation, 2014). For all these reasons this pathogen will be of concern also for aquaculture in the next years.
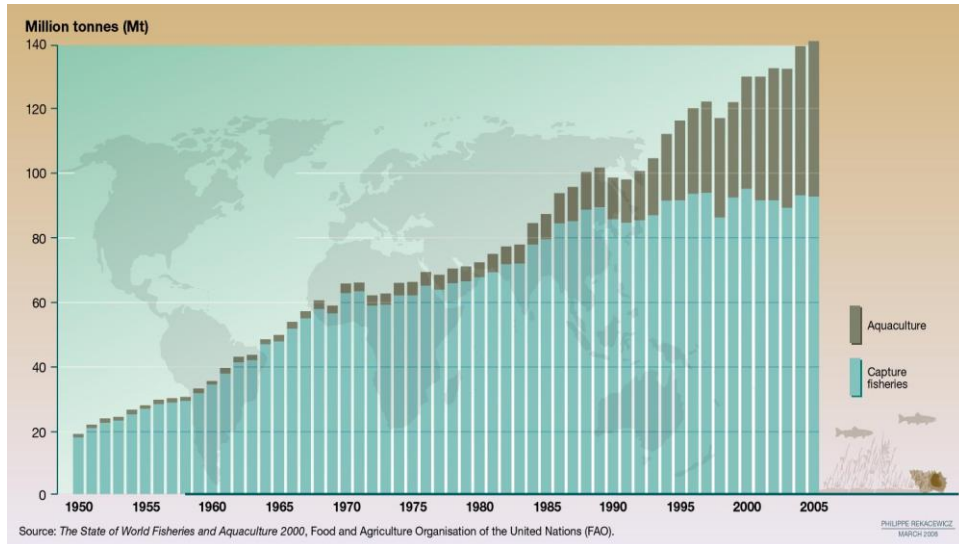
Figure 2. Millions of tonnes produced by aquaculture and fishing. Figure from FAO 2014.

1.4. Phylogeny

Phylogenetic studies based on MSA (multilocus sequence assay) of several loci have revealed the evolutionary history of *V. vulnificus*. These studies suggest that the species is divided in three linages that are not related with the three described biotypes (Sanjuán *et al*., 2011) (**Figure 3**). Sanjuán *et al* (2011) proved *pilF* to be a useful gene marker to differentiate between potentially human virulent and avirulent strains (Morrison *et al*., 2014). Unpublished results from Amaro's lab based on a phylogenomic study suggest that biotype 2 would have emerged several times by the acquisition of a virulence plasmid by horizontal gene transfer (HGT) and further recombination. The plasmid can be transferred in the laboratory among strains aided by a second plasmid that is conjugative and widespread in the species (Lee *et al*., 2008). Thus, according to all these studies, this biotype should be considered as a pathovar (proposed name pv. *piscis*) that would group all the strains with the ability to cause vibriosis in fish. Additional HGT events seem to have been produced along the evolution of *V. vulnificus* (unpublished).

24

Figure 3. Maximum likelihood phylogenetic tree of *V .vulnificus* specie. 7 concatenated loci of 115 isolates were aligned and maximum likelihood was used to build the three. Black, biotype 1 isolates; blue, biotype 2 serovar E isolates; red, biotype 2 non-serovar E isolates; green, biotype 3 isolates; #, human isolates; §, diseased fish isolates. Branches where recombination events involving the indicated loci might have occurred are indicated by arrows. The numbers at the nodes represent the percentage values given by bootstrap analysis of 1.000 replicates. Image from Sanjuán *et al*., (2011).
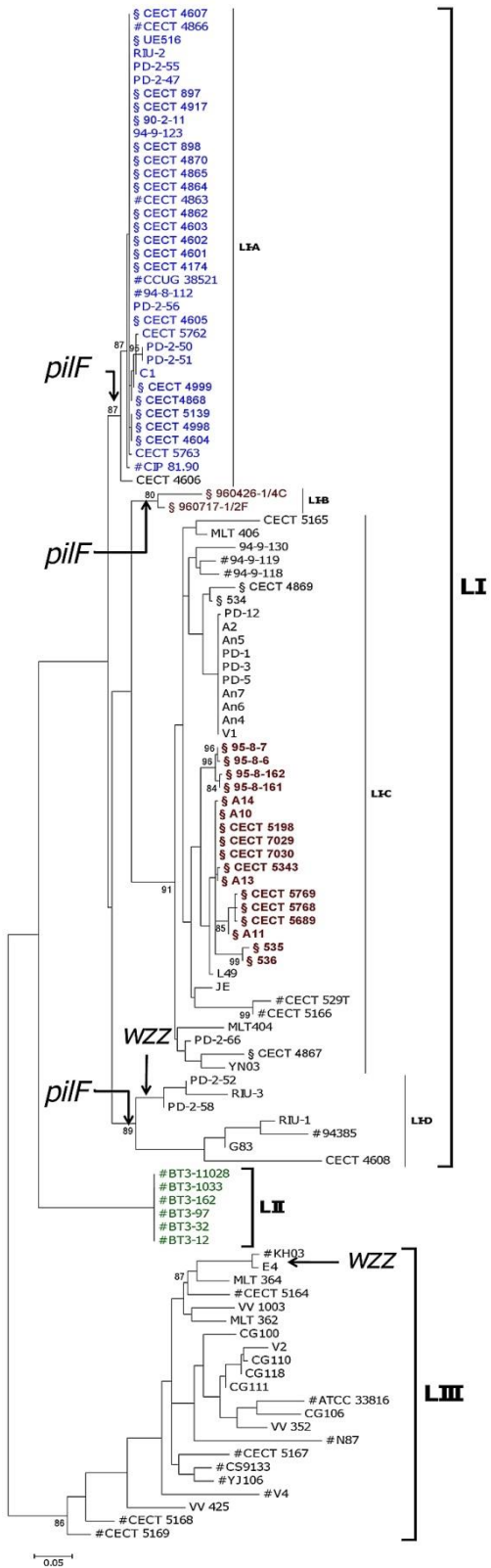
1.5. Genomes

Nowadays, there are 9 complete sequenced genomes in this species and 31 incomplete. All of them have a GC% of 46, similar to other vibrios. The first strain whose genome was sequenced and annotated was the strain YJ016 (Chen, 2003), a blood isolate of biotype 1 from a Taiwanese patient infected after eating raw oysters (**Figure 4**). *V. vulnificus*, like many other vibrios, has two chromosomes, chromosome 1 with a size of ~3.2 Mb and 2 with a size of ~1.8 Mb.

The number of genes in *V. vulnificus* is higher than in *V.cholerae* and similar to *V. parahaemolyticus,* another important foodborne pathogen. The higher number in ORF in *V. vulnificus* has been explained as a consequence of gene duplication, recombination, and HGT (Chen, 2003). Genomic comparisons proved that gene organization is more conserved in chromosome 1 than in chromosome 2, and that chromosome 2 is the one that has acquired most of the new genetic information.

Figure 4. Circular representation of chromosomes in *V. vulnificus* YJ016. From the outside inward: The first and second circles show predicted protein-coding regions on the plus and minus strands, by role, according to the color code of the COG functional categories. The third circle shows percentage G+C in relation to mean G+C for the chromosome. The fourth circle shows GC skew. The fifth circle shows super-integron (pink), type IV pilus genes (blue), capsular polysaccharide biosynthesis genes (green), iron acquisition genes (brown), extracellular enzyme and toxin genes (sky), and RTX genes (red). The sixth and seventh circles are tRNAs and rRNAs, respectively. Figure fromChen 2003.
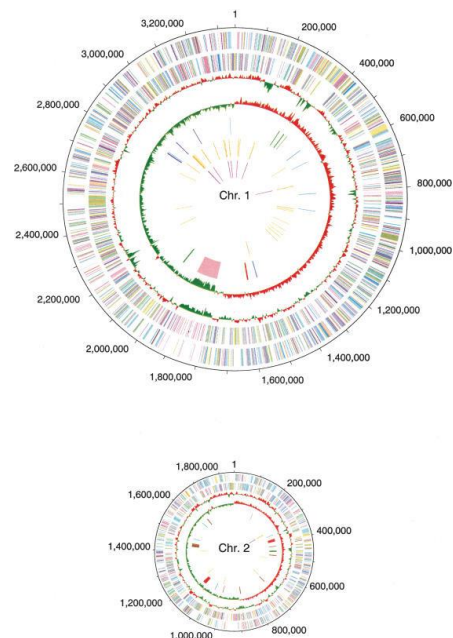
INTRODUCTION

On the other side, several plasmids have also been found in some strains. These are very important in some cases since they can encode new functions, some of them essential for virulence. For example, pR99 (and related plasmids collectively called pVvbt2) was proved to be essential for biotype 2 to infect eels (Lee *et al*., 2008; Valiente *et al*., 2008). Amaro & colleagues have made a great effort in order to unravel the role of this plasmid in the virulence for eels. According to their studies, this plasmid would be essential to resist immune system of eels and related teleost (Lee *et al*., 2008). The resistance would be mainly due the outer membrane protein (OMP) codified by *vep07*, whose deletion makes the strain sensitive to the fish but not to the human innate immune system (Lee *et al*., 2008). Moreover, a copy of *rtxA* gene was found in the sequence of pVvBt2. This gene encodes a MARTX (Multifunctional, Autoprocessing Repeats-in-Toxin) unique within the species in domain structure (MARTX type III). A detailed study of this toxin proved the role of MARTX type III in the interaction of *V. vulnificus* with a wide range of eukaryotic cells (**Figure 5**) (Lee *et al*., 2013). From amoebae to professional phagocytes, this toxin seems to enhance depolymerization of actin and loss of cell structure but not pore formation (Fullner and Mekalanos, 2000). Despite *vep07*, another gene in the plasmid has been the focus of attention. *vep20*, a gene that encodes a fish transferrin receptor involved in colonization and invasion of the eels (Pajuelo *et al*., 2015).

Moreover, a 48-53 kb conjugative plasmid was also sequenced in YJ016 (pYJ016) and CECT4602 although no putative role in virulence was determined for either (Chen, 2003; Lee *et al*., 2008).
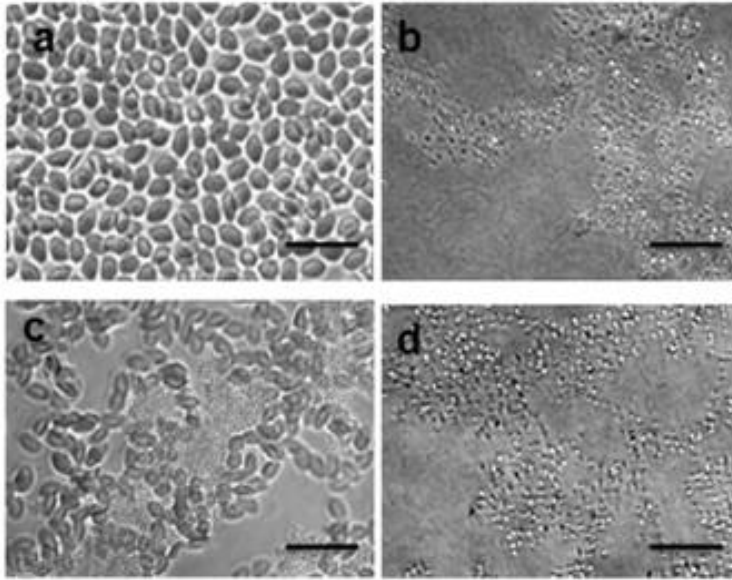
Figure 5. Cytotoxic effect of MARTX from *V. vulnificus* on eukaryotic cells. Microscopic observation of wild-eel erythrocytes infected with CECT4999 (b), CT285 (c) and CT316 (d) at 60 min post infection. Control (a), non-infected wild-eel erythrocytes. Bar, 100 mm. Figure from Lee *et al*. 2012.

2. *Vibrio vs* eels

V. vulnificus is an aquatic bacterium that survives as planktonic form or associated to biotic and abiotic surfaces. **Figure 6** summarizes the life cycle proposed for *V. vulnificus* by Pajuelo *et al.* (2016). According to this model, iron availability both within the host and between hosts would be one of the key signals for virulence and survival gene expression (Pajuelo *et al*., 2016). Thus, a change in iron concentration in the external environment could trigger transition from the free-living to the host-associated way of life through a cascade of transcriptional changes in gene expression such as those involved in capsule (CPS) and flagellum biosynthesis, chemotaxis, etc. (**Figure 6**).

On the other hand, it has been established that mucin and chitin act as chemoattractant for *V. vulnificus,* thus facilitating colonization of chitinous and mucosal surfaces in the environment (Colwell *et al*., 1977; Valiente *et al*., 2008). Chitin is the major component of the exoskeletons of arthropods such as crustaceans (e.g., crabs, lobsters and shrimps) and copepods, and mucin is the major glycoprotein in animal mucous surfaces. Interestingly, chitin can be sensed

by vibrios and then activate the DNA-uptaking machinery in *V. cholerae* and *V. vulnificus* (Valiente *et al*., 2008b). Further, unpublished results from Amaro's lab suggest that mucin, could act as chitin, activating DNA interchange. In summary, several components of the animal surface could trigger both attachment to and bacterial DNA exchange, facilitating survival and giving rise to the emergence of new genetic variants in *Vibrio*.
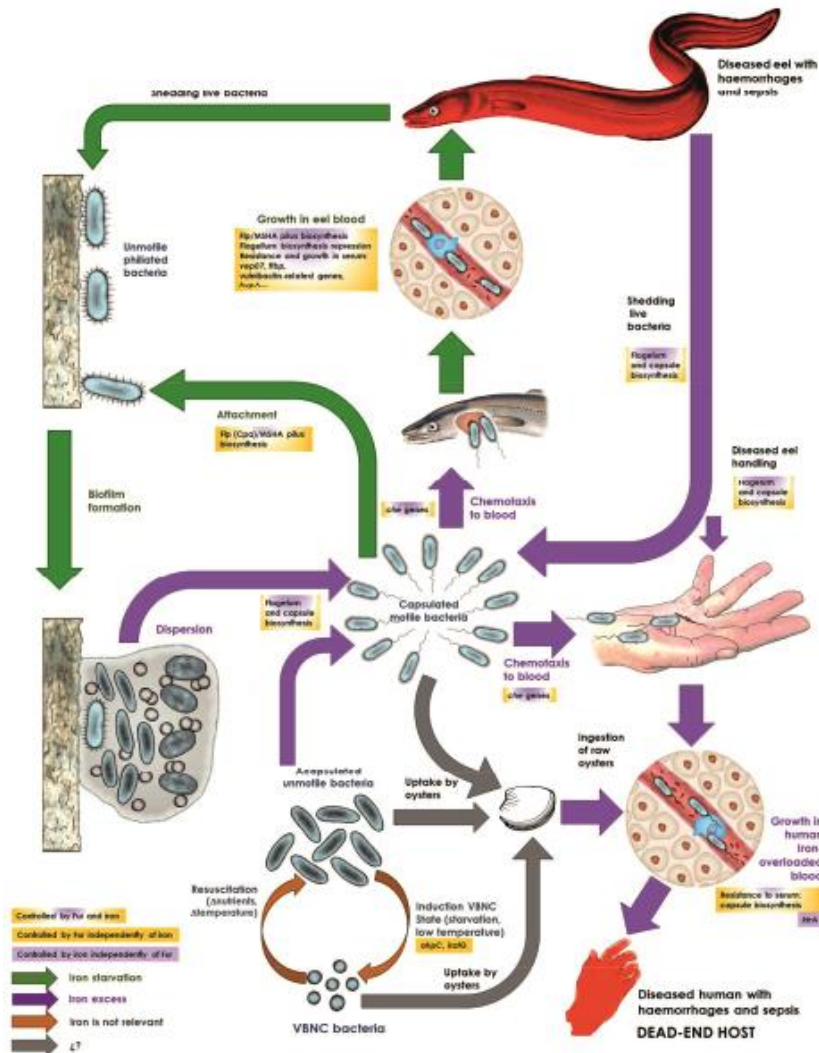
Figure 6. Picture taken from Pajuelo *et al*. 2016. (Iron and Fur and the life cycle of the zoonotic pathogen *V. vulnificus*) This figure summarizes the role of iron and Fur in the life cycle of *V. vulnificus*. Only the strains that possess the virulence plasmid pVvbt2 (Bt2) could invade successfully the eel blood and cause death by septicemia. The rest of the strains or biotypes would colonize mucosal surfaces of the eel or other fish species without invading them.

Eels are hosts for the biotype 2 of *V. vulnificus*. These isolates harbor a virulence plasmid encoding the ability to resist the eel innate immunity and cause eel vibriosis. In fact, it has been proposed that biotype 2 probably emerged in the fish farming environment, a habitat rich in nutrients, by acquiring the virulence plasmid from an unknown donor (Sanjuán *et al*., 2011). Eels do not present macroscopic scales on their surfaces. Instead, they are covered by a thick mucus layer rich in antimicrobial compounds that is presumably an iron-poor environment due to the iron-sequestering proteins secreted by the mucosal component of the eel innate immune system (Nielsen and Esteve-Gassent, 2006). Taken all the above into account, motile *Vibrio* would be attracted by the mucous surface of eels and there, the iron-restriction conditions would activate the mechanisms of attachment and DNA exchange. Only those isolates harboring the plasmid would infect eels and invade blood while the rest could survive in the eel surface where they would have to compete with the natural eel microbiota and face to bacteriostatic and bactericidal defenses of the animal.

2.1. The Eel

The European eel (*Anguilla anguilla*) is an euryhaline catadromous fish with a peculiar life cycle (**Figure 7**). It begins in the Sargasso Sea, where spawning takes place, and ends in the European rivers and lakes (6000 km away), where they become adults. The travel from Sargasso Sea lasts between 1-3 years which eels spend drifting with the Gulf Stream as larvae. Once they reach the continental fresh- and brackish-waters they spend most of their life in the "yellow eel" growth phase. During this time, eels will grow by eating insects, plankton, and small fish. Upon reaching sexual maturity, the eels, now in the "silver eels" phase, return to Sargasso Sea to spawn and die, usually between March and July (Tsukamoto, 2006). Eels can live up to 50 years and reach almost 1.5 meters of length, usually females reaching the biggest sizes. In 2010, Greenpeace International added the European eel, Japanese eel, and American eel to its

seafood red list. The intense fishing of eels and their habitat destructions are the main reasons for the decrease of population to truly dangerous levels.

There are many gaps in the knowledge of the eel life cycle that need to be filled, especially those related with its reproduction. Specifically, eels are not cultivated in fishfarms due to the fact that their reproduction in captivity is not possible. Because of this, glass-eels and elvers (young eels) are fished and transported to fish farm facilities where they are bred until they reach the adequate size for human consumption or repopulation.
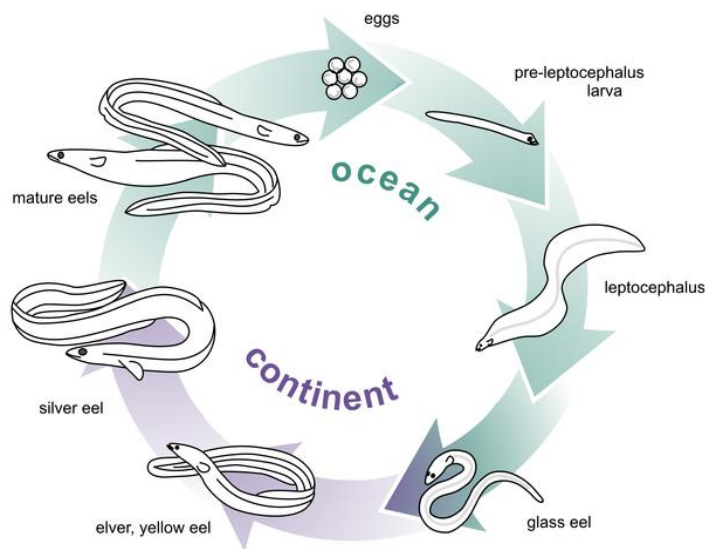
Figure 7. The life cycle of European eel. Figure from Scientific American **http://www.scientificamerican.com/**.

Besides vibriosis caused by *V. vulnificus* biotype 2, which can produce huge losses for the aquaculture industry, other bacteria (*Pseudomonas anguilliseptica*, *Edwarsiella tarda, Aeromonas veronii….*) and several parasites can also affect the viability of eels in nature and farms (López-Romalde *et al*., 2003; Wang *et al*., 2009; Joh *et al*., 2013). However, the main reason for eel decline is the human activities, mainly the man-made structures, such as dams and weirs. These structures complicate the advance of young eels and sexually mature eels and arrive to continental waters or the return to the Sargasso Sea, respectively.

INTRODUCTION

Scientists have focused their research on recovering the European eel population through its reproduction in captivity. Unfortunately, eels are supposedly only able to reproduce after having swum a distance of 6,500 km and scientists have not been able to simulate migration in the laboratory. A year ago, a group of Spanish scientist from Polytechnic University of Valencia was able to fecundate, breed and hatch European eel. The viability of these larva still need to be tested but this is a major advance in this field that could avoid the loss of this species (Asturiano *et al*., 2016).

2.2. Host colonization and invasion: serum survival

Once *V. vulnificus* has attached and colonized the mucosal surface, the next step would be to grow and invade it. To do it, the pathogen has to resist the bacteriostatic and bactericidal action of the innate immunity that acts at mucosal level: microcidal peptides, iron restriction, immunoglobulins, lysozyme…. According with the model proposed by Pajuelo *et al.* (2006)*, iron restriction would activate the expression of multiple genes involved in colonization such as a series of O-antigen biosynthetic genes, siderophore biosynthetic genes, resistance genes to oxidative and nitrosative stresses, etc. From this location, only the bacteria able to resist the bactericidal and bacteriostatic action of serum would be also able to survive in blood, spread to internal organs and causes death by sepsis. In the case of eels, the virulence plasmid encodes for a resistance system to these innate immune me*c*hanisms, which is under control of iron: the system is over-expressed under iron restriction. This resistance system together with the short generation time in eel serum (around 30 min at 28ºC) makes *V. vulnificus* biotype 2 extremely dangerous for the eel farming industry.

Pajuelo *et al*., (2016) results suggested that the CPS, the main bacterial structure involved in resistance to human serum (HS)(Amaro *et al*., 1994), is only over-expressed under iron excess conditions, meaning that if the pathogen arrives to human blood trough a wound exposed to the fish-farming environment, the pathogen would be unable to resist and would die. This circumstance explains why the severity of human vibriosis is correlated with high iron levels in blood, conditions in which a CPS would be optimally expressed. It has not been

32

described how a CPS protects the bacterium from being recognized but some hypotheses suggest that a change in charge of the bacterium would accomplish the objective (Roberts, 1996).

2.2.1. The complement system

The complement system is a part of the immune system that enhances the ability of antibodies and phagocytic cells to destroy pathogens, promoting inflammation at local level. The complement system consists of around 20 small proteins, liberated in the blood as pro-proteases. Once the first element is activated, this acts as a protease that activates the following element, and this event is produced successively. The final result of this cascade of activations is the opsonization and/or destruction of the pathogen as well as the activation of inflammation.

There have been described three pathways of complement activation (**Figure 8**). The classical pathway is activated by the binding of an antibody (or C1q) to a surface component, the MB-lectin pathway by the binding of a lectin to its target on the surface (for most encapsulated bacteria is a mannan-binding lectin) and, finally, the alternative pathway by the direct union of C3* (C3 activated by spontaneous hydrolysis) to LPS and other PAMPS (Pathogen Associated Molecular Patterns).

CLASSICAL PATHWAY	LECTIN PATHWAY	ALTERNATIVE PATHWAY
Specific union to pathogen surface	Mannose-binding lectin union to pathogen surface	Pathogen creates a local activation of complement

COMPLEMENT ACTIVATION

C3b covalent union to pathogen surface

Recruitment of inflammatory cells	Opsonization of pathogens	Perforation of pathogen surface

Death of pathogen
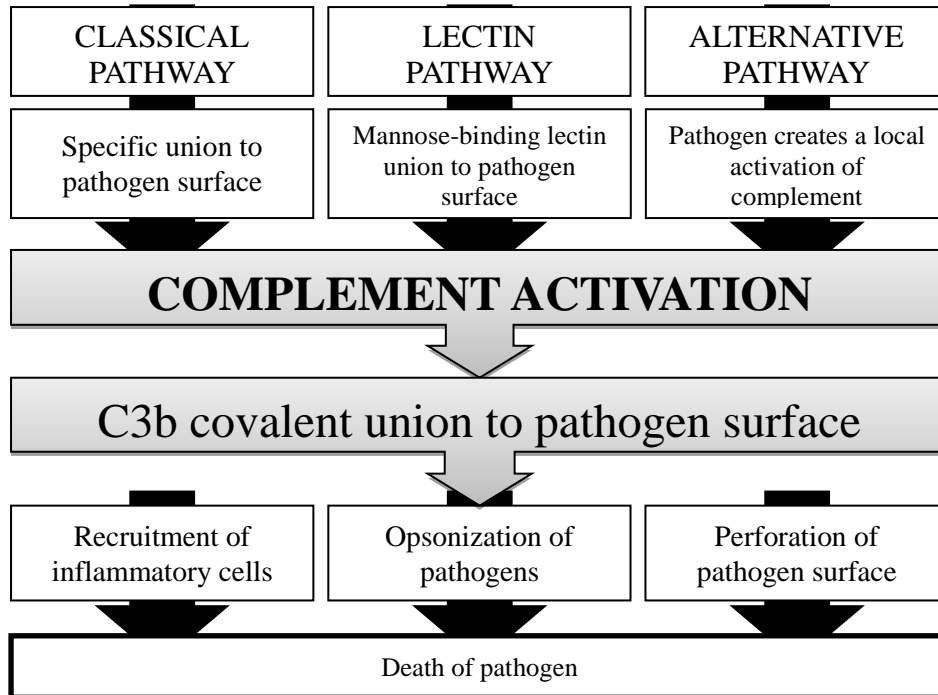
Figure 8. Schematic representation of the complement cascade.

2.2.2. The capsule of *V. vulnificus*

The capsule (CPS) is essential for *V. vulnificus* pathogenicity (Jones and Oliver, 2009) since acapsulated mutants are defective in resistance to HS and phagocytosis by macrophages (Tamplin *et al*., 1983, 1985; Kashimoto *et al*., 2003). There is no proved hypothesis to explain the molecular mechanism by which CPS protects against the bactericidal action of complement and phagocytosis. Interestingly, a spontaneous mechanism of phase variation between capsulated and acapsulated cells has been described in *V. vulnificus* (Starks *et al*., 2000) (**Figure 9**). Usually, bacterial phase variation is a reversible mutation that alters a structure on the surface (CPS, pili, flagellum, etc.) (Henderson *et al*., 1999). Capsulated cells can be distinguished easily from acapsulated ones on agar plates because they develop opaque (Op) colonies while acapsulated ones develop translucent (Tr) colonies.

The CPS operon is in chromosome 1 of *V. vulnificus* strains and has a similar structure than the one extensively described in *E. coli*. CPS components are

biosynthesized in the cytosol and transported to the outer membrane (OM). Biosynthesis genes are known to be the most variable in the operon while transport are very conserved between Enterobacteria (Chatzidaki-Livanis *et al*., 2006). High variation in operon structure and polysaccharide composition has been found among strains of the same species. Based on gene organization and CPS biochemistry strains can be classified in 4 groups. According to these values, 88% of *V. vulnificus* strains can be classified in group 1 due to the presence of *wza-wzb-wzc*(Chatzidaki-Livanis *et al.*, 2006). These three genes are highly conserved between *Enterobacteriaceae* and are involved in transport of CPS subunits. Wza, located in the OM, multimerizes in order to form a ring-like structure essential for secretion of polysaccharides across the OM(Dong *et al.*, 2006). Wzb (cognate phosphatase) and Wzc (tyrosine autokinase) are next to this gene and seem to be involved in assembly of subunits in the OM(Drummelsmith and Whitfield, 1999; Wugeditsch *et al.*, 2001).

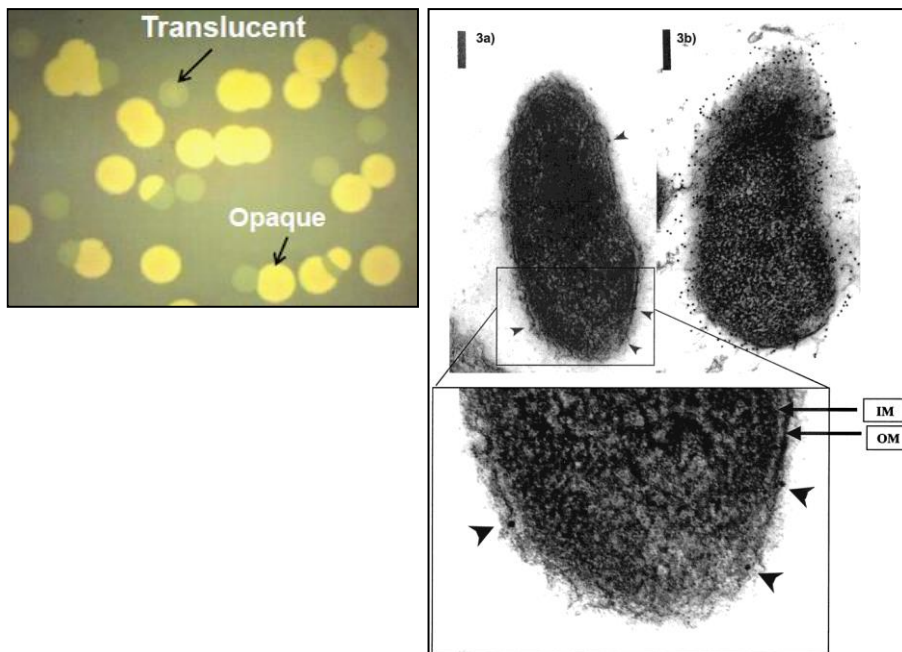

Figure 9. CPS and acapsulated phenotypes. A picture of the translucent and opaque colonies is on the left while Immunoelectron micrograph of *V. vulnificus* wza mutant is on the right. Figure from Wright 2001(Wright *et al*., 2001)**.**

3.   High-throughput sequencing in microbiology

For many years, the sequencing of DNA was a tedious process that needed of various persons and time to sequence one short DNA sequence. But now, we can sequence millions of DNA fragments in a short time and cheaper than before and practically all work is done by machines. Thanks to this, the number of deposited genomes in databases has increased day by day in all phyla. Sequencing and comparing all these genomes has shown to scientist incredible findings. For example, comparing two strains of the same bacterial specie, like *Escherichia coli,* showed that the shared genomic information was smaller than that shared by the human and the nematode *Caenorhabditis elegans* or the insect *Drosophila melanogaster*. This discovery changed completely the way to see bacteria and pathogens and opened a lot of questions. To answer some of these questions high-throughput sequencing has not stopped to improve. The number of techniques that has emerged from the DNA sequencing and its applications would take another thesis. In this work, we took advantage of two of those applications to study aspects of *V. vulnificus* biology which never have been analyzed like this before.

3.1. Metagenomics

Traditional culture-based microbiology applied to the description of the microbial composition of a determined environment produces an important bias in favor of culturable bacteria (Marshall, 1986). The first genomic technique developed to overcome this limitation was to clone specific genes (often 16S rRNA gene) to produce a profile of diversity from a natural sample (Marshall, 1986). The application of this technique, usually by pyrosequencing, revealed that the vast majority of microbial biodiversity had been missed by cultivation-based methods. Although 16S sequencing has limited taxonomic resolution, has allowed important discoveries such as the description of a new species *Staphylococcus dentisani,* from individuals without caries (unpublished results from A. Mira). Later, as the price of DNA sequencing decreased, metagenomics was developed. Metagenomics is the study of all the genetic material recovered

directly from an environmental sample and, therefore, it allows microbial ecology to be investigated at a much greater scale and detail than ever.

MG-RAST, one of the most used metagenomic databases, is holding 258.039 metagenomes, 880 billion sequences and all information that entails (Meyer *et al*., 2008). Understand how microbes interact among them, which microorganisms are more abundant, who are able to colonize a specific environment, how they change the environment, and how the environment changes the microbiota are interesting questions that now are closer to be answered thanks to metagenomics.

3.1.1. Microbiota

Using metagenomics microbiologists have unraveled the microbial composition of several environments such as oceans, lakes, rivers etc and have discovered and described hundreds of new species, genera and, even, a handful of new classes and phyla of microorganisms (Eloe-Fadrosh *et al*., 2016; Martin-Cuadrado *et al*., 2015; Mehrshad *et al*., 2016). Microbiota of intestine of multiple mammals (including humans), fish, insects, etc. have been metagenomically studied and described. In the case of humans, the microbial composition of different organs and tissues, the differences in microbial composition between gender, races, countries, life stages or between sick and healthy individuals have also been metagenomically analyzed (Pflughoeft and Versalovic, 2012; Parfrey and Knight, 2012; Schwartz *et al*., 2012; Bolnick *et al*., 2014; Yatsunenko *et al*., 2012; Schloissnig *et al*., 2013). From all these data, it has been estimated that the microbial population weighs around 1.5 kg in a regular human body and has been calculated that we have 6 microorganisms per human cell. Although recent data has suggested a decreased of this ratio to 1.3 to every one human cell (Sender *et al*., 2016).

Metagenomics has also interesting applications in medicine. Thus, the association of microbiome with several complex diseases is being investigated at high-resolution; i.e. type 2 diabetes, obesity, liver cirrhosis, colorectal cancer and rheumatoid arthritis (Karlsson *et al*., 2013; Jin and Flavell, 2013; Marchesi *et al*.,

2011; Zhang *et al*., 2015). Transplantation of fecal microbiota in patients with intestinal diseases is now a reality. This treatment has shown positive results to restore the intestinal microbiota after a dysbiosis (Petrof *et al*., 2013). Moreover, metagenomic results suggest a huge spread of antibiotic resistance genes over the planet. The expansion has such a magnitude that is not clear if human activities can be the only cause. In any case, the extensive use of antibiotics is not recommended and alternative ways are being studied like treatments to enhance and promote "healthy microbiota". For that, knowledge about how microbiota should be in a healthy individual or how it changes during an illness and more questions are being answered. The importance human microbiota has arrived to space since effects of space travelling over the microbiota of astronauts is under investigation (Voorhies and Lorenzi, 2016).

Less effort has been made to study the microbiota associated to non-human animals. Most of the studies in terrestrial and aquatic animals have been focused on the intestinal microbiota analysis in particular in farmed animals and has been related to production improvement (McAllister *et al*., 2015; Kasiraj *et al*., 2016; Barry *et al*., 2012). Particular interest has the study of the microbiota associated to healthy and diseased coral reef (coral bleaching) for its ecological impact worldwide (Badhai *et al*., 2016; Vega Thurber *et al*., 2009; Wu *et al*., 2012; Hennersdorf *et al*., 2016).

Regarding free-living microbiota, metagenomics has allowed its characterization all over the planet. Not only the microbiota from different oceans, rivers and lakes were analyzed but also the changes produced by temperature, oxygen or depth. Ghai *et al*., (2010) studied the peculiarities of a community in a zone of maximal photosynthetic activity, the deep chlorophyll maximum. Reproducible and robust genomic differences between surface and deep-sea species like *Pelagibacter* have been shown (Konstantinidis *et al*., 2009). Studies in the poles have also been performed in order to unravel key microbial drivers in these endangered environments (de Pascale *et al*., 2012; Wilkins *et al*., 2013). Polluted and "clean" lakes have been compared to understand the effect of humans in antibiotic resistance genes scattering (Bengtsson-Palme *et al*., 2014). More

examples could be added from a great diversity of environments and categories. Actually, because of the huge abundance of data deposited in databases some studies have been published using only public metagenomic datasets (Dutilh *et al*., 2014).

3.1.2. Functional metagenomics

These studies are not limited to identify presence or absence of determined taxa but additionally include identification of microbial functions that are enriched or depleted. Metagenomic studies determined that oral microbiota in individuals without caries had over-represented functions related to bacteriocin production. Moreover, mix-acid fermentation genes were over-represented in individuals with caries which explained its ability to decrease pH and produce the disease. As we have said, it has been estimated that relation human cell: microorganism is 1: 6. However, when the genetic diversity was assessed it was discovered that this proportion increase considerably and 99% of genes in our body are in microbes (Qin *et al*., 2010). In free-living environments, like oceans, metagenomic studies have unraveled some key milestones in microbial metabolisms (Ghai *et al*., 2013; Tseng and Tang, 2014). For example, detection of proteorhodopsin's widespread distribution in marine environments showed a new phototrophic strategy in marine ecosystems which explained how bacteria can be so abundant in the nutrient-limited open ocean (Venter *et al*., 2004). Proteorhodopsin is functionally a light-driven proton pump that fuels cellular processes by generating a proton motive force. Another example, a study suggested that *Prochlorococcus* significantly contribute to phosphonate utilization in global surface ocean waters by investigating the GOS (global ocean sampling) metagenomes (Feingersch *et al*., 2012).

In other fields, like animal farming, this has also become an interesting aspect to study. Agricultural and aquacultural studies are now focusing on intestinal microbiota in order to increase productivity. For example, studies are also trying to determinate if changes in the microbiota can produce an increase in fish size. Different diets are administrated and changes in microbiota are determined by metagenomics in order to see microorganisms that potentiate the growth of the

host or improve the immune system and made it more resistant to infections. This brings us to other of the recent discoveries made by scientist using metagenomics: the protection granted by bacteriophages in mucosal host-associated surfaces (Barr *et al*., 2013a, 2013b). Metagenomic studies in host-associated environments have suggested a putative role of phages to protect host from pathogens. Due to this, studies to use these phages as defense against pathogens (phage therapy) are increasingly frequent.

3.1.3. Viral metagenomics

It has been estimated that there is 10^{31} viruses on Earth which makes these the most abundant biological components (Breitbart and Rohwer, 2005). Despite this incredible abundance viral populations have been less studied than prokaryotic and most of the studies are focused on pathogenic viruses that infect eukaryotes. However, the role of viruses in the stability or instability of a microbiota seems to be extremely important and metagenomic studies have started to focus on this fraction of the microbiota. It is known in the microbiologist community that viruses have a very important role in the evolution of prokaryotes since it is one of the main elements for mobilization of genetic information (Suttle, 2007). Moreover, the high prevalence in some environments like oceans also suggests a paper of viruses in global carbon and energy cycling (Suttle, 2007).

Regarding geographic distribution, metagenomic studies indicate that viruses and bacteria have a weak geographic association at strain level but are global along the temperate and tropical waters around the Earth (Gonzaga and Rodriguez-Valera, 2013; Swan *et al*., 2013). Rodríguez-Valera *et al*. (2014) have shown that phages from vastly separated ocean sites are 97 % similar. They argue that even though oceans are continuous and water is mixed at some point, this is a process that takes thousands of years (Bollmann *et al*., 2010).

In metagenomics, special attention has been focused on bacteriophages (or phages), viruses that infects prokaryotes. Because of the greater abundance of prokaryotes comparing to eukaryotes the major part of viruses are bacteriophages. Phages are obligate parasites that besides introducing new

information in cells can regulate bacterial population. There have been some models that have tried to explain how prokaryotes and phages interact *in vivo* (Red-Queen hypothesis or Zoetrope). However, metagenomics has promoted the emergence of a new possible explanation: Constant-Diversity model (Rodriguez-Valera *et al*., 2009, 2014). This model is based on high diversity of concurrent clonal lineages within a single bacterial or archaeal population in an environment (**Figure 10**). The range of exploitable resources would increase thanks to this high diversity in clones lineages and smaller number of species. This diversity optimizes resource utilization by increasing the range of ecological niches that can be exploited. However, since prokaryotes reproduce clonally (producing exact copies of themselves) there is the risk that one clone could be more successful and increase its presence in the population. In this case, this clone would be more efficiently predated upon by the population of phages adapted to its receptors (Lotka-Volterra predator/prey equilibrium) and equilibrium would be restored.

Analysis of bacterial genomes sequenced to date has estimated that approximately 65% contain prophages (Snipen *et al*., 2009); this suggests that phages must be considered part of the bacterial pan-genome. Isolation from pure culture of the host was a requirement to study phages until metagenomics expanded. Sequencing phages by high-throughput sequencing technologies made easier the study but the small amount of DNA and the difficulty to assemble and annotate the reads also slows down the process. Filtering strategies are used to separate bacterial from the viral population when the sample allows it. Sometimes, ultracentrifugation is used to isolate the viral population or to concentrate a sample after a filtering protocol. Serendipitously, studies of microbiota by metagenomic approaches proved to be a valid methodology to sequence phages (Dutilh *et al*., 2014; Martín-Cuadrado *et al*., 2007). When phages start replicating the concatenated viral DNA eventually become larger than bacterial DNA present in the cell. Because of that, sequencing bacterial populations has highlighted a great number of replicating phages. Moreover, some approaches have been developed in order to discern which host is being infected by the replicating phages in bacterial populations. One of the most

interesting and valid protocols is the uses of clustered regularly interspaced short palindromic repeats (CRISPR). This system was initially described by Mojica *et al*. (Mojica *et al*., 2000) and consists in a prokaryotic immune system that confers resistance to external genetic elements such as plasmids and phages. The system has evolved to cut and save in the bacterial genome parts of the foreign DNA. This way, bacteria would recognize the organisms that contain the sequence saved. These sequences are in a specific order that allow scientist to find it in the genome and to compare it to replicating phages in order to discover its host. Microbiologists have taken advantage of the CRISPR system not only to understand the putative hosts of phages but also for DNA editing.
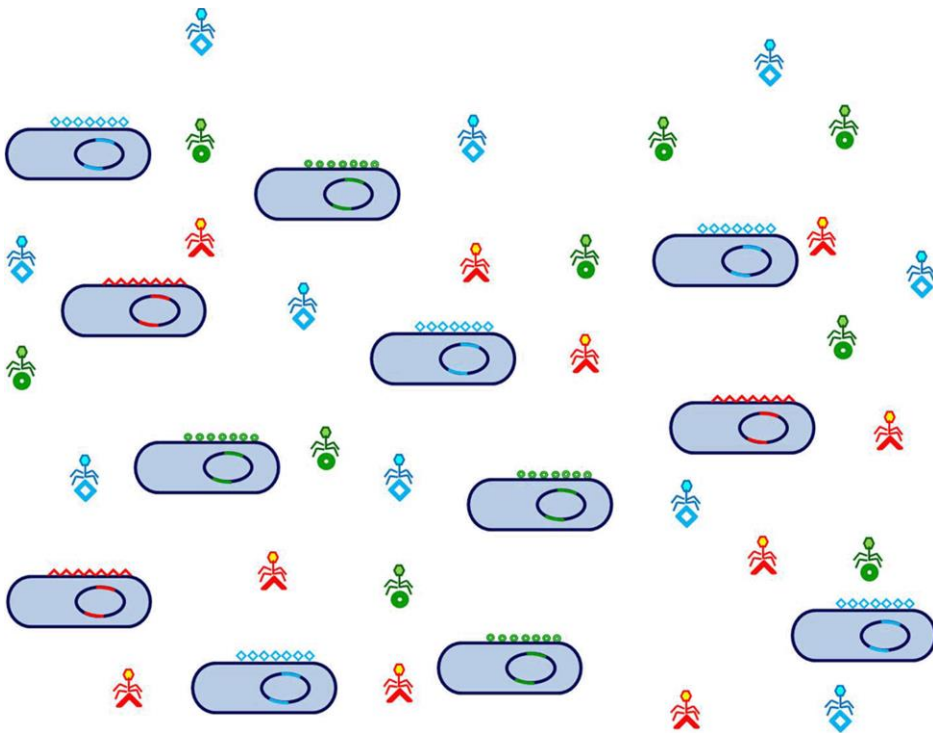
Figure 10. Costant-Diversity Model. Figure from Rodriguez-Valera and Ussery(2012).

Furthermore, metagenomic studies on viral populations have also tried to explain the role of this part of the microbiota in host-associated environments. This is the case of Dutilh *et al*. (2014), which proved the existence of a stable viral populations in the human gut microbiome by comparing several metagenomics datasets from this niche (Dutilh *et al*., 2014).The maintenance of phage

sequences in mucus-associated microbiota from gut samples suggested an attachment and an important role of this population in the intestinal cavity. Barr *et al*. proved phages as part of a non-host-derived immunity which was originated from attachment to mucosal surface and protection of hosts against pathogens (BAM) (Barr *et al*., 2013a, 2013b)(**Figure 11**).
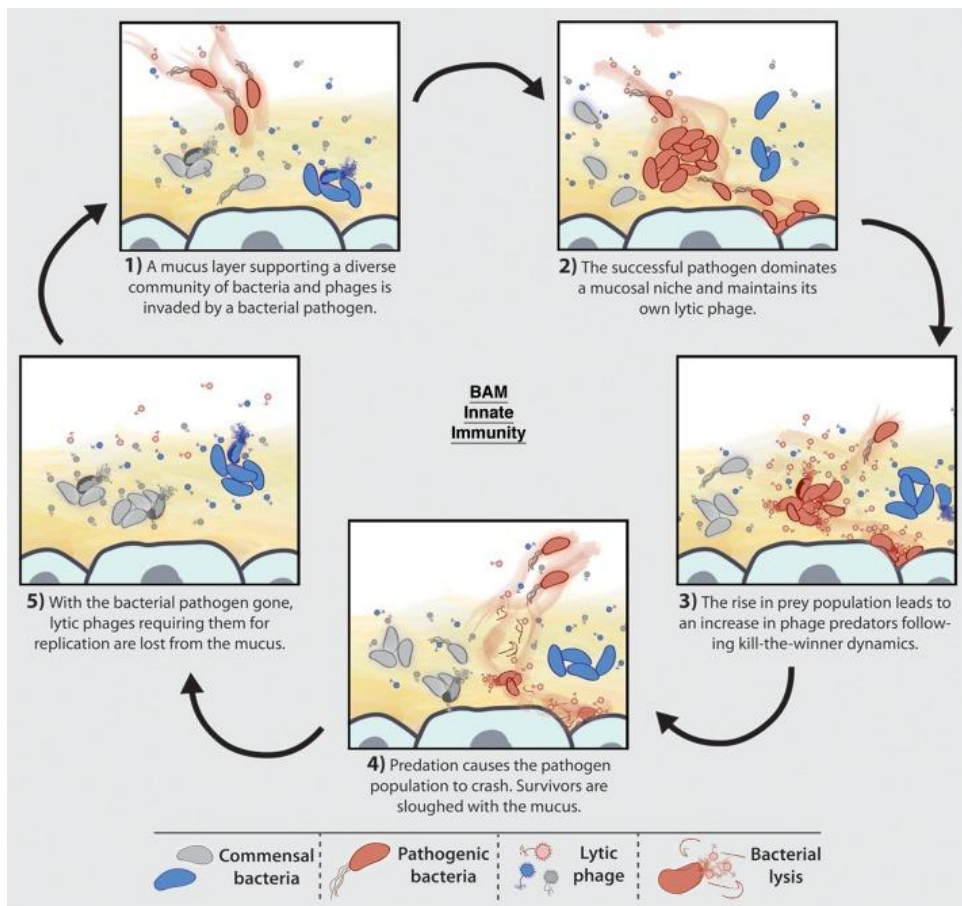
Figure 11. BAM model. Figure from Barr *et al*., 2013b.

3.1.4. Mobile genetic elements

The thousands of bacterial genomes sequenced (8644 according to NCBI genome list on 3rd of August of 2016) have expanded our view of the gene sharing between microorganisms and the pan-genome (Lapierre and Gogarten, 2009). Genomic comparisons have shown a great diversity in gene content between

close related strains (Muzzi and Donati, 2011). Nowadays, bacterial genomes are seen like variable entities that lose and gain genetic information continuously. For instance, *Escherichia coli* genome comparisons have shown a flexible genome of a typical strain to be of 1.4 Mb (25% and more than 1300 genes) many of which are considered important for virulence (Perna *et al*., 2001). A significant fraction of this flexible genome is codified in mobile genetic elements (MGE) like plasmids, phages, integrative conjugative elements (ICE), transposons, integrones and pathogenic islands. Each one of those elements can have genes that give benefits to the host. At evolutionary level this has a great importance. There are some examples of MGE that have provided important traits for pathogens. This is the case of cholerae toxin which was acquired by the integration of a phage in the genome of *Vibrio cholerae*. Without this toxin *V. cholerae* is not able to produce cholera. As explained previously, the resistance to the immune system of eels is codified in a plasmid in *V. vulnificus* biotype 2 without which it is unable to infect the host. (Lee *et al*., 2008).

3.2. Transposon insertion sequencing

Besides genomics and metagenomics, scientists have made an effort to develop new methodologies combining high throughput DNA sequencing and other techniques. For example, RNAseq combines RNA extraction and cDNA transcription with sequencing to perform transcriptomic studies for all genes in a genome. Now, we can study the transcriptomic response in the host and in the pathogen during an infection with dual RNAseq (Westermann *et al*., 2012). Furthermore, the combination of transposon mutagenesis and high-throughput DNA sequencing, called transposon insertion sequencing (TIS) or Tnseq, has allowed researchers to discern between genes essential and neutral to grow in a determined situation.

Independently, several variations of this technique have been developed, for example, transposon insertion site sequencing (Tnseq), insertion sequencing (INSeq), high-throughput insertion tracking by deep sequencing (HITS) and transposon directed insertion site sequencing (TraDIS) (Langridge *et al*., 2009; Gawronski *et al*., 2009; van Opijnen *et al*., 2009). All of them share the same

fundament: create a library using transposon mutagenesis in which ideally all genomic loci would be disrupted (**Figure 12**). Mutations of genes essential to grow in the medium where the library is done would not be present in the library. This means that if a gene is essential to grow in the control condition would not be tested in the condition of interest. Then, transposon libraries can be grown under specific conditions so mutants that grow less in this condition would be attenuated while mutants unaffected would grow fine. In order to map where the transposon has been inserted we would sequence these positions. How the library is processed with the aim to sequence only transposon disruptions is how the different techniques vary. Thanks to the high- throughput DNA sequencing, we can localize and count the transposon inserted at genome-wide level and therefore compare transposon abundance/presence in all the genes. Comparing abundances of transposon inserted inside a gene and between conditions would allow us to classify genes as neutral (high proportion of gene disrupted) and under-represented (low or 0 % of gene disrupted). In principle, transposons inserted in genes essential to grow on the tested condition would not be recovered in the output while neutral genes would be overrepresented. To identify genes underrepresented or "essential" loci and being able to differentiate them from "conditional essential" loci and "neutral" loci has been proved to require a series of controls and normalization process. Waldor and collaborators have set up this protocol for *V. cholerae*(Chao *et al*., 2016, 2013). Finally, the function of the selected genes can be accomplished by means of molecular biology techniques combined with *in vivo* and *in vitro* assays. While the main benefit of this technique is obvious (the capacity to study the necessity of genes at genome-wide level) the main drawback is not that clear: the bottlenecks.
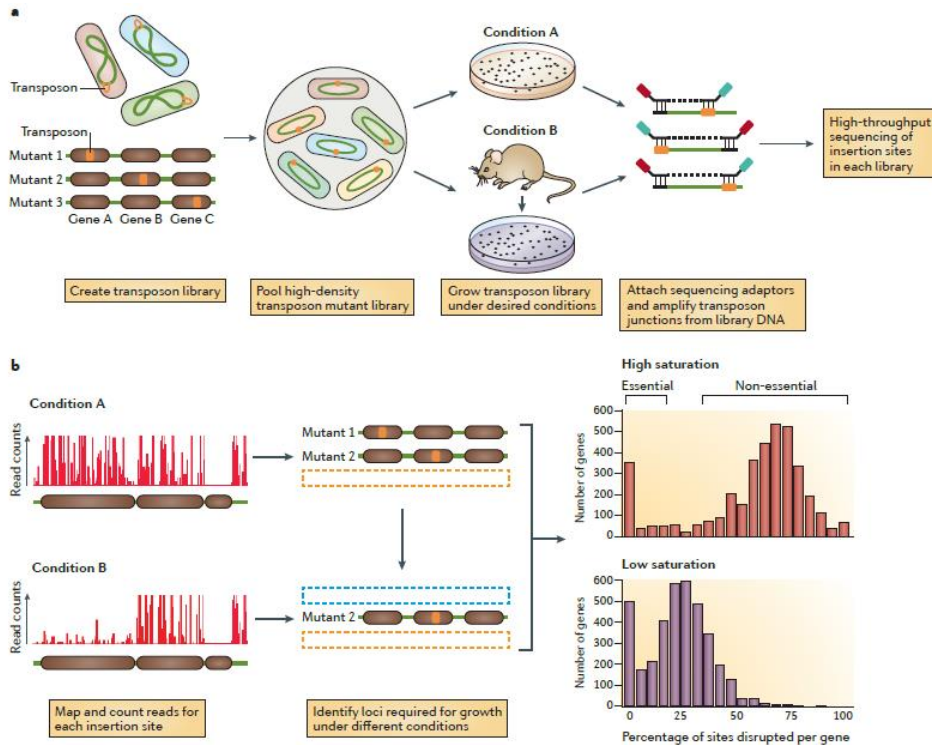
Figure 12. TIS basic protocol

Some of the conditions that scientist would like to test using this technique make an effect on bacteria that is uncontrollable. For example, it is known that 99% of bacteria inoculated in mice during *V. cholerae* infection are lost during the pass through the stomach. The low pH and adverse conditions in this organ destroy many cells. These bottlenecks need to be taken in account and make necessary the confirmation of results by mutation.

INTRODUCTION

CHAPTER 1

Vibrio vs Host: colonization and eel microbiota

1. Abstract

In this first chapter we have assessed the first of the steps *V. vulnificus* and many other opportunistic pathogens need to start the infection: the colonization of the host surface. The first objective was to know what is the natural microbiota of the skin-mucus of one known hosts of *V. vulnificus*, the European eel (*Anguilla anguilla*). Once determined, secondary objectives were to answer to the following questions: i) Does *V. vulnificus* (or its known mobilome) belong to this microbiota and if so, ii) what would be their putative competitors and predators (phages), iii) Does this microbiota interchange genetic information using mobile genetic elements and which. In order to do that, we took advantage of high-throughput DNA sequencing techniques to fully describe the microbial community that is attached to skin-mucus of wild-eels. We complemented our study by analyzing the variation in microbiota between water (free-living) and mucous (attached), farmed and wild-eels and among wild-eels captured at different points in the Mediterranean coast.

Gammaproteobacteria were the dominant class in the eel mucous with a representation of genera depending on water salinity where eels were captured. In general, bacterial species found in the eel skin-mucus were absent in the surrounding water, suggesting that mucous attracts, selects and concentrates water microorganisms. Thus, the genus *Vibrio* was the most abundant in the mucous of eels from water of higher salinity water (7-10 g/l) while a mixture of genera (*Pseudomonas, Aeromonas, Achromobacter, Sphingobium, Shewanella, Stenotrophomonas* or *Acinetobacter)* were predominant in water of lower salinity (1-3 g/l). Further, the most represented *Vibrio* species in mucosal samples from higher salinities was *V. anguillarum* and the most abundant at lower salinities was *V. cholerae* and close species (*V. metoecus*). *V. vulnificus* was present in the samples with salinity ≥1 g/l. Finally, changes in the way of life from free in nature to captured in farms, determined a complete change in the microbiota's composition.

Interestingly, we found a great abundance of replicating phage genomes (concatemers) in all the samples. They were assembled in four complete genomes

of three Myovirus and one Podovirus. We also found evidences that ΦKZ and Podovirus phages could be part of the resident microbiota associated to the eel mucosal surface and persist on them over time. Moreover, the viral abundance estimated by epiflorescent counts and by metagenomic recruitment from eel mucosa was higher than that of the surrounding water. These phages in the animal's mucus could act as agents controlling bacterial populations, including pathogenic species, providing a kind of innate immunity.

Finally, we unraveled the presence of mobile genetic elements (flexible metagenome) containing virulence and antibiotic resistance genes by using a strategy of in silico analysis developed in the present Thesis. We found prophages, ICEs, transposons etc. as well as and a series of evidences suggesting that bacteria where exchanging DNA in the skin-mucus.

In summary, this is the first description of the microbiota attached to the surface of a teleost. The environment that *V. vulnificus* finds when it tries to colonize the European eel has been studied.

2. Introduction

V. vulnificus is an aquatic flagellated bacterium that can live in the environment either associated to biotic/abiotic surfaces or as a free-swimming microorganism. As a free-living cell, *V. vulnificus* uses the flagellum to swim and find nutrients, which are more abundant on the water/surface interface. Supposedly, vibrio cells "sense" the nutrients and are attracted towards them (positive chemotaxis). Once on the surface, the pathogen can loss the motile appendices, attach to, multiply and colonize it by forming a biofilm. Thus, variate from one state (mobile) to the other (sessile) supposes a transcriptomic and proteomic major change in the bacterium.

Most of the studies performed in *V. vulnificus* have been focused on how this bacterium causes vibriosis in humans but not on how the bacterium survives outside humans, in the environment. Thus, no study has investigated the interaction of *V. vulnificus* with intermediate-hosts or reservoirs *in vivo*, which is extremely important to understand the entire life cycle of this pathogen. Analyze the microbiota of the putative *Vibrio*-reservoirs would be the first step to clear up this scenario.

With this objective, we selected a host that could naturally carry vibrio, the European eel. Previously, Marco-Noales *et al*. (2001) demonstrated that *V. vulnificus* can survive after artificial waterborne infection of eels by forming biofilms on their skin surfaces and, later, Valiente *et al*. (2008) proved that the eel skin-mucus attracts the bacterium facilitating gill colonization(Marco-Noales and Milán, 2001; Valiente *et al*., 2008b). *V. vulnificus*, as a natural inhabitant of estuarine environments, could attach to the mucosal surface of wild-eels and, even, arrive to the fish-farms with colonized eels. To colonize the wild-eel, the pathogen would have been able to compete with the natural eel microbiota, resist the mucosal eel defenses and multiply.

The European eel is an ancient and catadromous fish species that can inhabit multiple aquatic environments from the open sea to rivers or lakes. The life cycle of eels is fascinating, and is only partially known. Adult eels supposedly spawn

in the Sargasso Sea (hypothetically at a depth of hundred meters), migrate as young larvae towards Europe with the Gulf Stream, arrive 1-3 years later, and suffer three metamorphoses while colonizing ponds, lagoons or lakes, they grow and become sexually mature(Tsukamoto, 2006). Then, adult eels migrate back to the Sargasso Sea more than 6000 km against the Gulf Stream and spawn, closing their life cycle(Tsukamoto, 2006). In addition, the European eel is a species of commercial interest in several European and Asiatic countries, which has promoted their culture (or growth because reproduction in captivity is not possible) in intensive farms. With this objective, eels are captured in nature (at glass-eel or elver stage) and are grown in farms at physicochemical conditions that accelerate their growth: fresh-water or brackish water at 25-28 °C with high animal densities in tanks (Gousset, 1990). These conditions are favorable for proliferation of those vibrios present in tank's water or in fish. The European eel is also a critically endangered species due to multiple anthropogenic factors including habitat destruction, pollution, diseases and overfishing (http://www.iucnredlist.org). For all these reasons, to know the natural microbiota that travels with eels constitutes an interesting objective in itself.

Eels lack of macroscopic scales to protect its surface (Aida *et al*., 2003). To protect themselves, eels secrete a thick layer of mucus that contains multiple microcidal and microstatic compounds(Nielsen and Esteve-Gassent, 2006). According to Barr and co-workers, an additional protective function for the mucus would be to trap bacteriophages in order to detect and destroy invading bacteria before they reach to the epithelium (Barr *et al*., 2013a, 2013b). In consequence, *V. vulnificus* could colonize the mucus only if it resists mucosal defenses, including phages. Once colonized, eels could be fished and arrive to farms, where the pathogen would encounter the ideal conditions to multiply. From all vibrios colonizing eels, only those possessing the virulence plasmid pVvBt2 would be able invade blood and to cause fish vibriosis.

Taken all this into account, the first objective of this thesis was to know what is the natural microbiota of the skin-mucus of wild-eels. Once determined, secondary objectives were to answer to the following questions: i) Does *V.*

vulnificus (or its known mobilome) belong to this microbiota and if so, ii) what would be their putative competitors and predators (phages) iii) Does this microbiota interchange genetic information using mobile genetic elements and which. In order to do that, we took advantage of high-throughput DNA sequencing techniques to fully describe the microbial community that is attached to skin-mucus of wild-eels. We complemented our study by analyzing the variation in microbiota between water (free-living) and mucous (attached), farmed and wild-eels and among wild-eels captured at different points in the Mediterranean coast.

3. Hypothesis and objectives

We sampled the mucus from wild-eels captured in three Nature Parks located close to the Mediterranean Sea: Albufera Lake, Ebro Delta and Prado Cabanes. Additionally, skin-mucus from eels farmed in aquaculture facilities was also sampled. Finally, water microbiota from one of the wet lands (Ebro Delta) was also analyzed. All these microbiotas were sequenced and analyzed in order to answer and complete the next objectives:

3.1. Describe the microbiota of the skin-mucus of European eel

    3.1.1. Bacterial population attached to wild-eels

        3.1.1.1. Is *Vibrio* abundant in the surface of wild-eels?

        3.1.1.2. Does the skin-mucus microbiota of wild-eels vary with changes in the surrounding-water?

    3.1.2. Microbiota attached *vs* free-living

    3.1.3. Bacterial population attached to farmed-eels

        3.1.3.1. Is the skin-mucus microbiota of wild-eels different to that of farmed-eels?

        3.1.3.2. And if so, does the changes affect to bacterial pathogens?

3.2. Investigate the putative enrichment of bacteriophages in the mucus of an aquatic organisms

    3.2.1. Describe the viral population attached to eels

        3.2.1.1. Test Barr *et al*. hypothesis

3.3. Flexible-metagenome

    3.3.1. Develop a protocol to detect putative MGEs in a metagenome

        3.3.1.1. Annotate and classify the MGEs found in the skin-mucus metagenome

        3.3.1.2. Are virulence genes concentrated in eel' skin-mucus?

        3.3.1.3. Are evidences of GME transference in eel' skin-mucus?

4. Material and methods

4.1. Sampling and DNA isolation

The selected habitats for eel sampling as well as the positional coordinates are presented in **Figure 13**. Three Nature Parks in the Mediterranean Spanish coast were chosen to fish wild-eels (250-300 gr):   Albufera de Valencia (Valencia),Ebro Delta (Alfacada pond), two samples a year apart (2013-2014), and Prat de Cabanes-Torreblanca (Prado Cabanes)(Castellón). Moreover, a volume of 25 l of water from Ebro Delta (2014) was also analyzed. Albufera Nature Park is located on the Gulf of Valencia coast (Valencian Community, Spain) and contains a shallow (1 m depth on average) coastal lagoon that is the largest fresh-water lagoon in Spain. Because of the increased human activities in its densely populated surroundings, Albufera lagoon collapsed as a macrophyte-dominated lagoon and turned into a highly hypertrophic ecosystem with very dense phytoplankton populations primarily dominated by *Cyanobacteria* (Ghai *et al*., 2012). Ebro Delta Nature Park, one of the biggest Mediterranean wetlands, comprises the mouth of the Ebro river and surrounding areas (320 km^2) and is located in south of Catalonia (Spain).Prat de Cabanes-Torreblanca Nature Park is a wetland located in the coastline of the municipalities of Cabanes and Torreblanca (Castellón, in the Valencian Community). We also sampled about 10,000 wild-eels (0.33 g each exemplar) captured from various rivers of the Atlantic coast of Spain (Galicia)(**Figure 13**). Wild-eels were transported in a fresh-water tank to Valencia, where they were maintained for 24h prior to be processed (**Figure 13**). Finally, 20 farmed-eels (FE) of 250 gr from a local farm were also sampled.

To get the mucus samples, the fish from each location were put together in fishbowls with 1 l of sterile PBS, for 15-20 min (**Figure 14**). Wild-eels were returned without damage to their habitats and farmed-eels to their respective tanks. Then, the secreted mucus was recovered in sterile recipients, stored at 4ºC, transported to the lab and filtrated through 5 and 1 microns filters, used sequentially. Finally, the prokaryotic biomass was recovered on 0.22 microns filters, which were treated with 1 mg/ml lysozyme and 0.2 mg/ml proteinase K

(final concentrations). Nucleic acids were extracted with phenol/chloroform/isoamyl alcohol and chloroform/isoamyl alcohol, and DNA integrity was checked by agarose gel electrophoresis (Ghai *et al*., 2012).
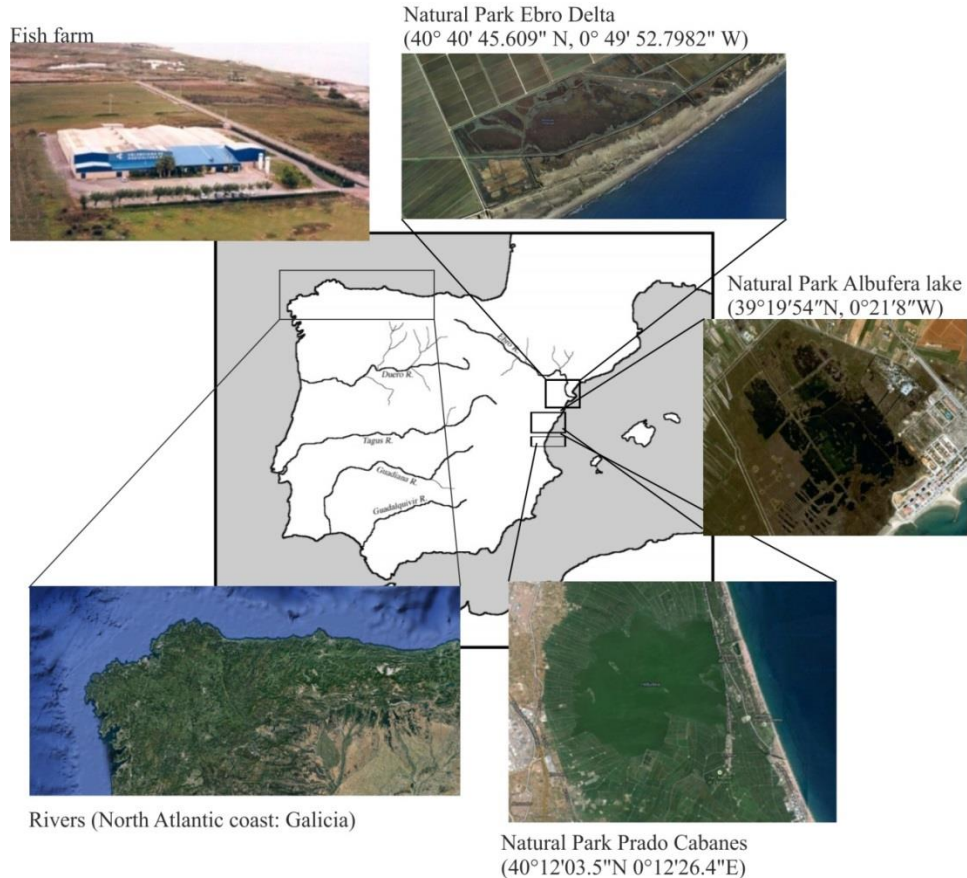
Figure 13. Sampling point in the Spanish coast for wild and farmed eels. Mucus from adult farmed eels was sampled in aquaculture facility. Wild eels were fished using nets.

4.2. Sequencing and assembly.

DNA samples were sequenced either by using a FLX sequencer (454 Roche) with Titanium chemistry (Centro Superior de Investigación en Salud Pública [CSISP, Valencia, Spain]or an Illumina HiSeq 2000 sequencer with pair-end technology (Macrogen [Seoul, Korea]). Assembly of 454 metagenomic reads was performed using Geneious Pro 5.4 with a minimum overlap of 50 nucleotides, 95% identity and allowing 1% of mismatches per read, while Illumina reads were assembled using Velvet (k-mer 51) (Zerbino and Birney, 2008). All the

metagenomes were deposited in NCBI SRA with the following accession codes: SRR1578065, SRR1578068, SRR1578098, SRR1580820, SRR1580821 and SRR1580823. Only assembled contigs bigger than 1kb were considered for the following analysis.

Fishing

Skin mucus sampling

Filtration & DNA extraction

Figure 14.Skin mucus microbiota sampling methodology. Nets were recovered 24 h after installation (fishing) and eels were deposited in fishbowls with PBS 1x sterile during 20 min (Skin mucus sampling). The resulting solution was filtered to concentrate the microbial population and DNA was extracted and sequenced (Filtration & DNA extraction).

4.3. Sequence analysis and annotation of the assembled contigs.

The GC content of the metagenomes was determined using the "geecee" program from EMBOSS package (Rice *et al*., 2000). BioEdit software was used to manipulate the sequences (Hall, 1999). The KEGG and SEED database were used to analyse metabolic pathways and functional classification of the proteins (Kanehisa *et al*., 2004) that matched the reads. To allow the interactive visualisation of genomic fragment comparisons, Artemis Comparison Tool ACTv.8 (Carver *et al*., 2005) was used. The assembled contigs were annotated using Prodigal (Hyatt *et al*., 2010; Zerbino and Birney, 2008) and MG-RAST pipeline. Annotation was refined manually using HHpred (Söding, 2005)

4.4. Community structure using all reads and ribosomal RNA (rRNA).

For taxonomy, the entire datasets were compared using BLASTX or BLASTP (Altschul *et al*., 1997) from the NCBI NR database (cut off expectation (*E*) value 10^{-5}, minimum length 50bp and minimum similarity of 95%) and analyzed using MEGAN with the "Percent Identity Filter" active (Huson *et al*., 2011). rRNA genes were identified by comparing the data sets against the RDP database (Cole *et al*., 2009). All reads that matched an rRNA sequence with an identity 95% and an alignment length of 100 bases against either the RDP or the LSU database were extracted. The best hit with a taxonomic affiliation was considered a reasonable closest attempt to classify the rRNA sequences.

4.5. Functional classification of reads in SEED categories

In order to categorize the reads obtained we used the pipeline from MG-RAST. This program automatically annotates reads and then classifies them in SEED categories. We download the results in .xls format and compare abundance using Microsoft Office Excel manually.

4.6. General characterization of phage genomes (Virome)

The largest phage contigs were identified manually after annotation of contigs. The rest of them were fished using these contigs on a BLASTN search (Altschul

et al., 1997). Moreover, CRISPR approach (see below) was used to identify prophages. Control phage genomes were downloaded from NCBI and were classified at family level (if it was necessary) using ICTV (International Committee on Taxonomy of Viruses). The identified viromes were compared with control phage genomes by using previously described protocol (Mizuno *et al*., 2013). Comparisons were done using BLASTX and BLOSUM45 matrix (Altschul *et al*., 1997). Minimum 30% sequence identity, 30aa length and a maximum e-value of 10^{-3}were considered to filter the results. Phylogenetic tree of selected genomes were constructed using PHYLIP package (Felsenstein, 1993).

To find out if the largest contigs were closed, direct terminal repeats for phages were determined using Uniprot UGENE program (Okonechnikov *et al*., 2012). BLASTP search was used to identify the genes of interest to build the trees and the Ig-like domains. *Pseudomonas* phage phiKZ was used as the query in the ΦKZ genus specific searches. Finally, phylogenetic analyses using a group of genes (terminase large subunit, RNA and DNA polymerase [RNApol and DNApol], ribonucleoside diphosphate reductase alpha chain [*nrdA*] and phosphate starvation-induced protein [*phoH*]) allowed the taxonomic classification of these contigs. Trees were done using FastTree with 100 bootstraps (Price *et al*., 2009). All genes and genomes were download from Genbank, Pfam or Uniprot database (Benson *et al*., 2005; Consortium, 2014; Finn *et al*., 2006).

4.7. ANI.

Average nucleotide identity (ANI) analyses were used to assign the genetic distance between the most closely related strains of the contigs to the same contigs.

4.8. *Vibrio* isolation and identification.

A volume of 1 ml of water or mucus from each one of the samples was inoculated into 4 ml of the *Vibrio* enrichment medium, Alkaline Peptone Water

(APA; 1% peptone extract supplemented with 1 % [wt/vol] NaCl at pH 8.6), and incubated for 12 h at 28 ºC with agitation (150 rpm). Then, volumes of 0.1 ml of a 1:10000 dilution were spread on plates containing the *Vibrio* selective media, TCBS (thiosulfate-citrate-bile salts-sucrose) (Conda 1074), and VVM (*Vibrio vulnificus* medium) (Cerdà-Cuéllar *et al*., 2000) agar and on the general medium TSA-1 (trypticase soy agar supplemented with 1 % [wt/vol] NaCl). Plates were incubated 24h at 28ºC. Suspected colonies were purified on TSA-1, isolated and lyophilized at -80ºC in LB-1 (Luria Bertani broth1 % [wt/vol] NaCl) supplemented with 20% (vol/vol) glycerol. The isolates were phenotypically identified with API 20E kit (bioMerieux) according to manufacturer's instructions. Bacterial suspensions in PBS were used as inocula. Examination of the strips was conducted after 24h using the API Database https://apiweb.biomerieux.com/. In parallel, genomic DNA was extracted from suspected colonies and amplified using PCR targeting rDNA by using universal primers to identify bacteria. The amplicons were sequenced and identified in the Genomic section of the SCSIE (Servicios Centrales de Soporte a la Investigación Experimental) from the Universidad of Valencia (Spain). To identify suspected *V. vulnificus,* primers vvhA-F (5′- CGCCACCCACTTTCGGGCC-3′) and vvhA-R (5′-CC GCGGTACAG GTTGGCGC-3′) were used to amplify hemolysin gene corresponding exclusively to *V. vulnificus*(Sanjuan and Amaro, 2007).

Selected strains were sequenced using Illumina (Macrogen [Seoul, Korea]), Genomic DNA was extracted using the Wizard Genomic DNA extraction kit (Promega) and assembly and annotation was followed as explained in 4.2 and 4.3.

4.9. Genomic comparisons

The contigs assembled from sequencing the strain M12v were compared with the deposited *V. metoecus* and *V. cholerae* O1 ElTor N16961 using GCviewer Comparison tool (Grant *et al*., 2012). Default values were used for BLAST atlas comparison.

4.10. MGE detection.

Several simultaneous approached were used to detect MGE in contigs >10 kb: by one side, to detect intra-contig variations in GC%, taxonomical annotation and hexanucleotide usage pattern (HUP) and by the other side to look for MGE-signature genes (sgMGE) by using BLASTP. BLASTN and ISfinder and ISbrowser were used to identify plasmids and pathogenic islands deposited in databases (Siguier, 2006; Kichenaradja *et al*., 2010). We considered as sgMGE, integrases, transposases, conjugative elements and phage or viral proteins. HUP values were calculated by using *compseq* program from EMBOSS package. Microsoft Excel tools helped us to differentiate genes annotated to the same taxon and to visualise the GC variation in the contigs.

The workflow diagram finally selected to detect MGE in a metagenome is showed in **Figure 15**. This consisted in selecting contigs by PCA (own program) using HUP, followed by detection of sgMGE and further analysis of the annotation and GC percentage. Contigs that were part of the variable genome (30S, 50S or lipopolysaccharide) clustered with contigs with putative MGE (cpMGE) and were considered false positives. Intra-contig changes in the taxonomic annotation plus presence of an integrase located just after a tRNA or several sgMGE were considered to certify the presence of a MGE in a cpMGE.

Three metagenomes generated in the present work (WE_3^8, WE_{10}^8 and WE_{10}^8W) and six metagenomes from various origins that were downloaded from EBI metagenomics or MG-RAST were used tested for the protocol (**Table 1**) (Hunter *et al*., 2014; Meyer *et al*., 2008).
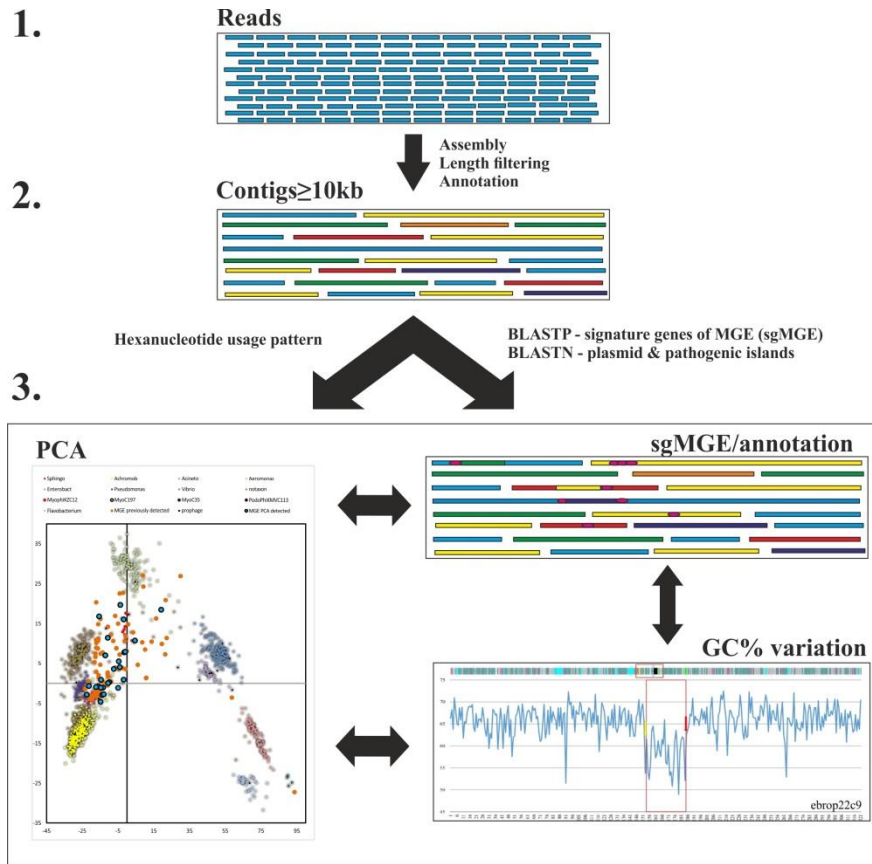
Figure 15. Mobile genetic elements (MGE) detection workflow diagram. The steps necessaries to detect contigs with a putative MGE are schematized in this figure. The presence of MGE in assembled contigs bigger than 10 kb can be assessed using 3 methods. We can detect those contigs looking for genes that usually are present in MGE (sgMGE) or known MGE (plasmids and pathogenic islands) by BLASTP and BLASTN, respectively. The analysis of contigs distribution in a PCA analysis according to its hexanucleotide usage pattern (HUP) can also help us to detect these interesting contigs. Finally, if the GC variation in a contig that we suspect that has a MGE can confirm our suspicions.

Table 1. Metagenomes used to detect MGE.

Name	Sample	Sequencing technology	Nº contigs >10 kb	Access number
WE$_3$[8]	Host-associated (mucus)	Illumina	1321	SRA185006, SRA185006
WE$_{10}$[8]	Host-associated (mucus)	Illumina	391	SRR1586370
DCM	Aquatic (Mediterranean)	Illumina	720	SRR037008
HumanDiabetes	Host-associated (gut)	Illumina	63	SRR341581
Mouse	Host-associated (feces)	Illumina	210	4535626.3, 4535627.3
SeaUrchin	Host-associated (gut)	Illumina	330	ERR895166
Moose	Host-associated (gut)	Illumina	587	ERS624611
Indian	Aquatic (Lake)	Illumina	211	ERS433966
Swedish	Aquatic (Lake)	Illumina	138	ERS433967

4.11.   Identification of bacterial hosts for phages

In order to identify the potential host for the phages found, CRISPR approach and tetranucleotide usage pattern (TUP) were attempted (Ogilvie *et al*., 2013; Stern *et al*., 2012). On one hand, three different approaches were used to search for putative CRISPR cassettes. (1) The protocol described by Stern *et al*. (2012) that consists in inferring the spacers using known CRISPR repeats and use these spacers to fish phage assembled contigs (Stern *et al*., 2012); to filter false positive spacers, a minimum of 75 bp read length and 80% query coverage were used and to consider a spacer in a contig, 100% identity and coverage were used. (2) Spacers were searched directly from assembled contigs using spacers from CRISPR database; only matches over 80% coverage and 90% identity were considered. (3) CRISPRfinder tool was used to find CRISPR cassettes in assembled contigs.

One the other hand, spacers detected from different approaches were used to find viral contigs and putative hosts for those contigs. BLASTN search was done considering 100% identity as a filter. Finally, TETRA 1.0 was used to identify contigs with similar TUP (Teeling *et al*., 2004). Only contigs bigger than 10 kb and values of 0.6 or over were considered for analysis.

4.12.   Abundance of virus and bacteria in the metagenomes

In order to compare the abundance of viral and bacterial populations in our datasets we counted the number of reads recruited to viral and bacterial concatenated contigs. The number of reads was calculated using BLASTN, considering a minimum identity of 95 and a maximum e-value of 10^{-3} for filtering the results. The number of reads recruited against the genomes assembled was normalized per the size of the genome or concatenated (kb) and the dataset (Gb).

The abundance of those phages in other niches was analyzed comparing by TBLASTN the reads of marine (GOS, Albufera, Sargasso sea, Tampa bay and Mediterranean bathypelagic habitat) and animal associated (mouse, termite,

canine, cow, coral and human) metagenomes against the viral proteins isolate in our data. Metagenomes were download from MG-RAST server (Meyer *et al*., 2008). A minimum identity of 10^{-5} e-value was considered for filtering the results. MG-RAST ID: 4440414.3, 4440440.3, 4440439.3, 4440413.3, 4440424.3, 4440422.3, 4440412.3, 4440411.3, 4440066.3, 4440062.3, 4440055.3, 4440056.3, 4440065.3, 4440063.3, 4440059.3, 4440064.3, 4483775.3, 4450680.3, 4450678.3, 4440284.3, 4440283.3, 4440285.3, 4440286.3, 4440102.3, 4444702.3, 4444703.3, 4444165.3, 4444164.3, 4440373.3, 4440375.3, 4440379.3, 4440377.3, 4440374.3, 4440381.3, 4440376.3, 4440378.3, 4440371.3, 4440370.3, 4440380.3, 4440372.3, 4447454.3, 4447455.3, 4447456.3, 4447457.3, 4447446.3, 4447447.3, 4447448.3, 4447449.3, 4441025.3, 4442464.3, 4441625.3, 4441625.4 4441627.3, 4441623.3, 4441624.3, 4441621.3, 4441622.3, 4441629.3, 4441628.3, 4441626.3, 4440330.3, 4440951.3, 4472804.3, 4472821.3, 4473347.3, 4473348.3, 4473365.3, 4473372.3, 4473378.3, 4473389.3, 4473411.3, 4473417.3, 4473438.3 and 4478542.3.

We also used VIROME database to search in the uploaded viromes for annotated ORFs to any of the viral genus found in our metagenomes (ΦKZ, ΦKMV, FelixO1like and Φ16). The number of ORFs was count using the online available tools in the VIROME website (Wommack *et al*., 2012).

4.13. Phage counts by microscopy

Samples from surrounding water and skin-mucus of eels farmed in tanks at 22 °C and 1% salinity in facilities at University of Valencia (Planta de Acuarios de Experimentación, PAE) were collected. Samples were maintained on ice, sonicated in 3 pulses during 4 seconds. Then, 50 and 3 ml of water and mucus, respectively, were directly filtered per 0.02-μm Anodisc polycarbonate filter (Whatman). Filters were stained with SYBR Green 5x, washed and visualized using epifluorescence microscope. For each sample, 25-30 images were analyzed using ImageJ (Schneider *et al*., 2012). Counts of bacteria and virus-like particles per milliliter were made using a previous protocol described by Patel *et al*. (2007).

5. Results

We obtained the metagenome of skin-mucus(SM-metagenome) from wild-eels and farmed-eels as well as water (W-metagenome) from Ebro Delta. The main physicochemical characteristics of the water of the sampled habitats are shown in **Table 2**. The generalities of each metagenome are shown in **Table 3**. Several contigs nearly identical to parts of known bacterial genomes were obtained; i.e. an 893 kb contig that is >95% identical along its entire length to a part of the genome of *Pseudomonas aeruginosa PA14* (originally isolated from a human patient [Lee *et al*., 2006]) (**Figure 16**). Assuming that contig's length is inversely proportional to bacterial diversity, the results of the assembly suggest that the skin-mucus is a poorly diverse niche. The comparison of alpha-diversity indexes of SM- and W-metagenome support this suggestion (see **Table 3** and 5.1.2.).

Table 2. Sampling points, types of samples and designation as well as main physico-chemical parameters of water of the sampled environments.

Sampling point*	Sample type	Sample designation**	Salinity (g/l)	pH	T (ºC)
Fishfarm close to the Mediterranean sea		$FE_4^{5.3}$	4	5.3	24
Rivers (North Atlantic coast: Galicia)		$WE_{\leq 1}^{7}$	≤1***	7	7
Natural Park: Albufera Lake	Skin mucus from eels	$WE_1^{9.5}$	1	9.5	19
Natural Park: Prado Cabanes		WE_7^{8}	7	8	20
Natural Park: Ebro Delta		WE_3^{8}	3	8	19
		WE_{10}^{8}	10	8	20
	Water	$WE_{10}^{8}W$	10	8	20

*. Coordenates of each sampling point are indicated in Supplementary Table 1.
**. Super- and sub-indexes indicate pH and salinity values, respectively.
***. The salinity of the rivers where eels were fished was ≤ 1 g/l.

Table 3.Sampled points and sequencing. pH/ salinity values are added to abbreviate names of samples as super/subindexes.

	FE$_{5.3}^{4}$ [**]	**WE$_{\leq1}^{7}$** [**]	**WE$_{1}^{9.5}$** [**]	**WE$_{7}^{8}$**	**WE$_{3}^{8}$**	**WE$_{10}^{8}$**	**WE$_{10}^{8}$W**
Dataset size (Mb)	70	49.8	45.3	1300	3939	3622	959
Number of sequences	101780	91574	70510	12325876	25287346	46064452	12490702
Average length (bp)	634,6	486,1	587,5	87.4	88.25	78.6	76.9
Number of contigs >1Kb (Mb)	6074 (8.2)	8217 (16.6)	1609 (20.8)	1016 (23.08)	5050 (45.13)	12069 (35.8)	2608 (6)
Average contig length (bp)	1354	2027.7	1291.5	2272	8938	2970.2	2302
Number of contigs >10Kb (Mb)	0	0	0	7 (0.3)	699 (33.26)	415 (11.2)	27 (0.56)
Average contig length (bp)	0	0	0	43310.6	47590.1	27052	20993
Viral contigs (Mb)	218 (0.30)	67 (0.12)	79 (0.11)	1 (0.22)	33 (0.43)	181 (0.86)	127 (0.6)
Average contig length (bp)	1399.6	1890	1413.8	220068	13066.2	4786.9	4770.9
Viral contigs > 10Kb (Mb)	0	0	0	1	12 (0.36)	21 (0.49)	14 (0.39)
Average contig length (bp)	0	0	0	220068	29921.8	23795.6	27867.9
alpha diversity	132.73	401.08*	101.46	157.54	50.99	107.72	422.59

* Rarefaction curve did not reach a plateau
**Sequenced using 454. We used Illumina for the rest
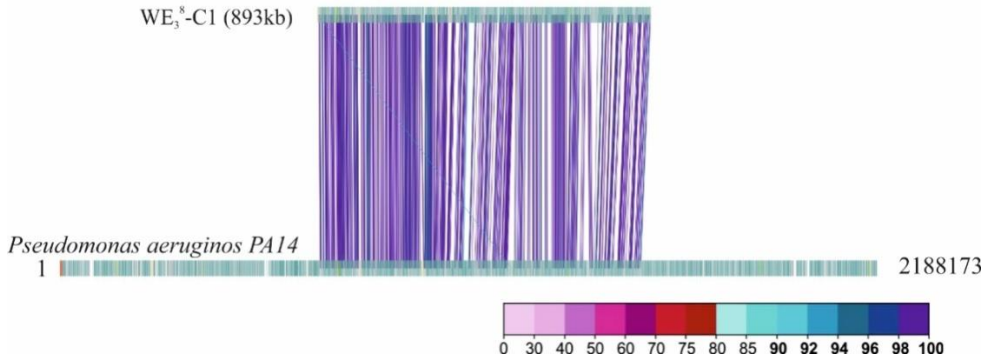
Figure 16. The largest bacterial contig assembled in this study. This contig from WE_3^8 metagenome aligned to a fragment of almost 900 kb in the chromosome 1 of *P. aeruginosa* PA14 using TBLASTX.

5.1. Superficial microbiota of the European eel

This is the first time that the microbiota of the mucosal surface of a wild-fish species is studied in depth and compared with that of the same species but in captivity. The European eel is an euryhaline and catadromous fish that lives in seawater (open sea and coast), brackish-water (mainly estuarine-waters) and fresh-water (lakes, rivers, ponds) along its life cycle. Thus, we decided to sample eels fished in water of different salinities and characteristics (**Table 2**): natural estuarine-waters (Ebro Delta and Prado Cabanes), natural fresh-waters (Atlantic rivers and Albufera Lake) and an artificial environment, an eel-farm that uses brackish-water to grow the eels (4 gr/l salinity). In addition, because eels are listed in the European Red List of endangered species (http://ec.europa.eu/environment/nature/conservation/species/redlist/downloads/European_freshwater_fishes.pdf), we decided to sample superficial mucus and return the sampled individuals to nature.

5.1.1. Description of microbiota in skin-mucus of wild-eels

The GC percentage is a useful tool to look for differences in bacterial composition among metagenomes. **Figure 17** shows GC profiles of the skin-mucus (SM) metagenomes from wild-eels, their comparison showed a direct relationship with water salinity (**Figure 17A**). Thus, eel mucus from low-salinity

waters (below 1 g/l) was unimodal with a peak around 65% GC, that from intermediate salinity (3 g/l) was bimodal with an additional peak around 45% GC, and that from high-salinity (7-10 g/l) was again unimodal but with a peak around 45% GC. (**Figure 17Aand B**).
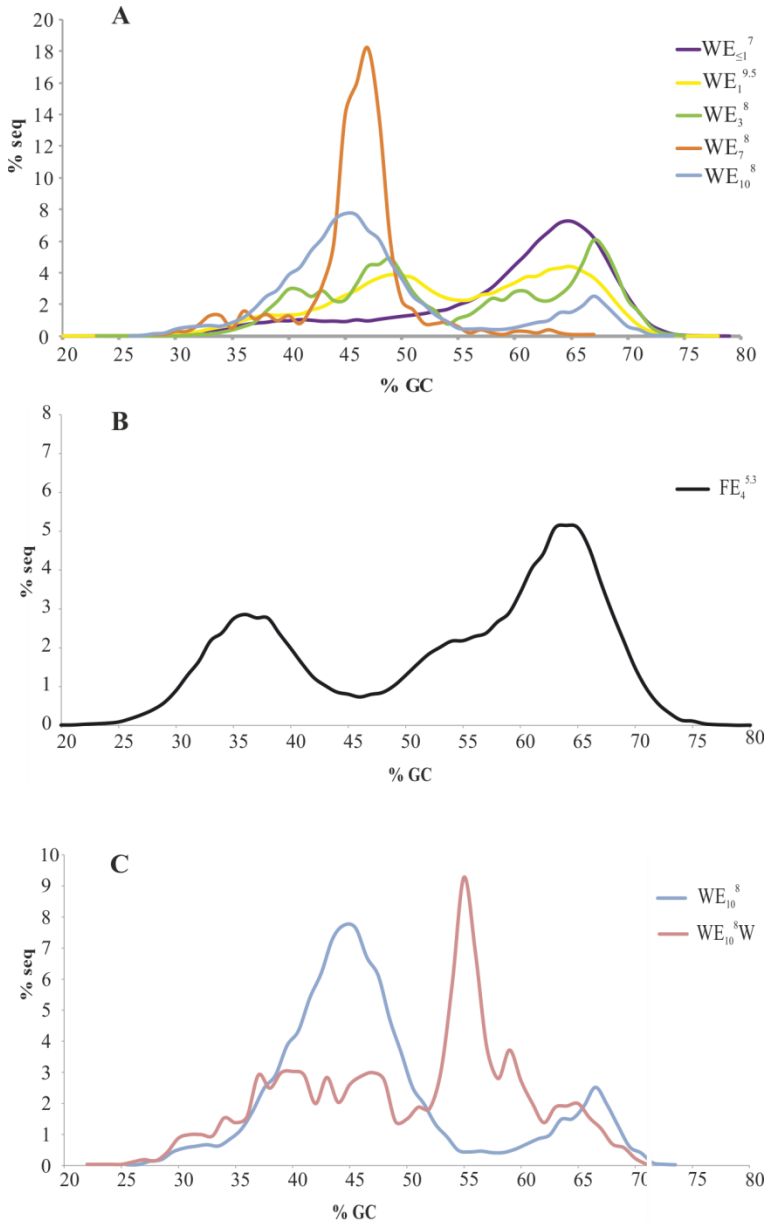
Figure 17. GC profiles of the SM- and W-metagenomes. Percentage of sequences (% seq) depending on the GC value (% GC). A, SM-metagenomes from wild-eels; B, SM-metagenomes from farmed-eels; C, SM and W-metagenomes from Ebro Delta.

Then, we analyzed the bacterial classes present in the SM-metagenomes by using
16S rRNA gene reads. We found that SM-microbiota was dominated by
Gammaproteobacteria, whose proportion in the metagenomes ranged between 30
% (wild-eels from river) and 95% (wild-eels from estuarine-water), again
depending on water salinity (**Figure 18**).   Flavobacteria (0.6% to 28%),
Betaproteobacteria (2.6% to 26%) and Alphaproteobacteria (5 to 24%) were
other phyla dominant in the SM-microbiota (**Figure18**).
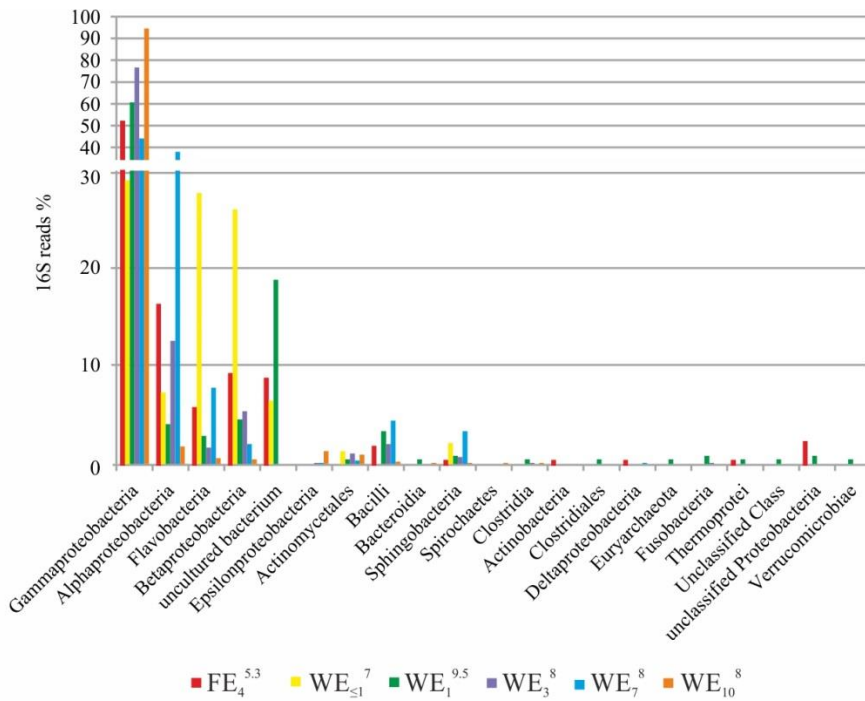
Figure 18. Main classes of bacteria found in the metagenomes using 16S reads. A, SM-
metagenomes from wild-eels; B, SM-metagenomes from farmed-eels; C, W-metagenome *vs* SM-
metagenome from the same habitat. Reads were searched for 16S gene and then blasted against
Ribosomal Database Protein (RDP).

The differences in CG profiles were more clearly correlated with composition in
bacterial genus. Figures 19 and 20 show the main bacterial genus detected in the
metagenomes by using two different approaches, 16S reads (**Figure 19**) and
annotation of the identified genes against the database (**Figure 20**). Clearly,
Vibrio was the dominant genus in the wild-eel's metagenomes from estuarine-
waters (46 % to 93.5% of the total) and decreased significantly in that from lake-
and river-waters (9.8% in water of 3 g/l salinity; 0-2.3% in waters ≤ 1 g/l
salinity) (**Figures 19 and 20**). In these last samples, *Pseudomonas* (22 to 38 %),

*Stenotrophomonas (*1 to 14 %*)* and *Achromobacter* (2 to 20%) were the most abundant common genera (**Figure 19 and 20**). In addition to these genera, *Acinetobacter*, *Stenotrophomonas*, *Aeromonas* and *Shewanella* completed the SM-microbiota of wild-eels (**Figures 19 and 20**). Finally, the annotation of the reads against reference genomes supported 16S results except for *Ensifer* genus in WE_7^8 metagenome where its presence was reduced drastically using this approach (**Figures 19 and 20**). This was probably caused by the presence of 16S reads in RDP database but the absence of genomes from *Ensifer* strains in the NCBI at the time the metagenomes were annotated.
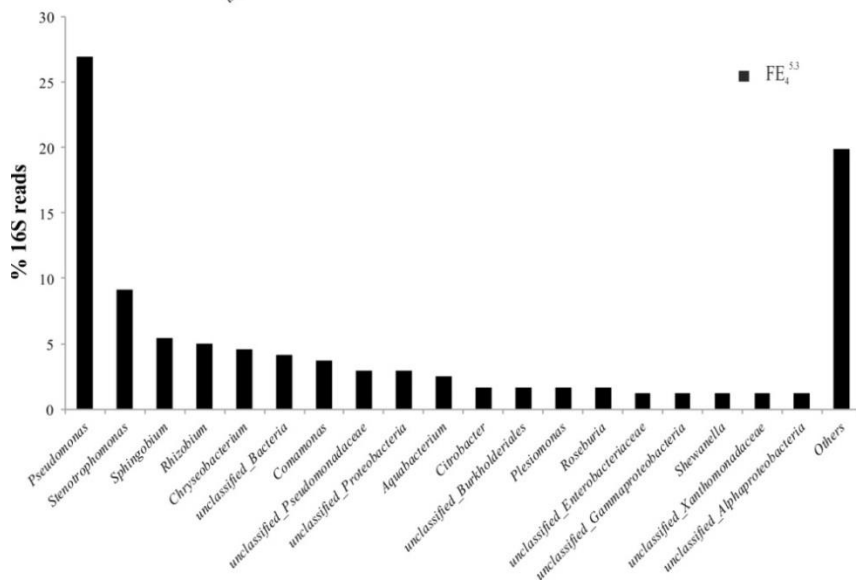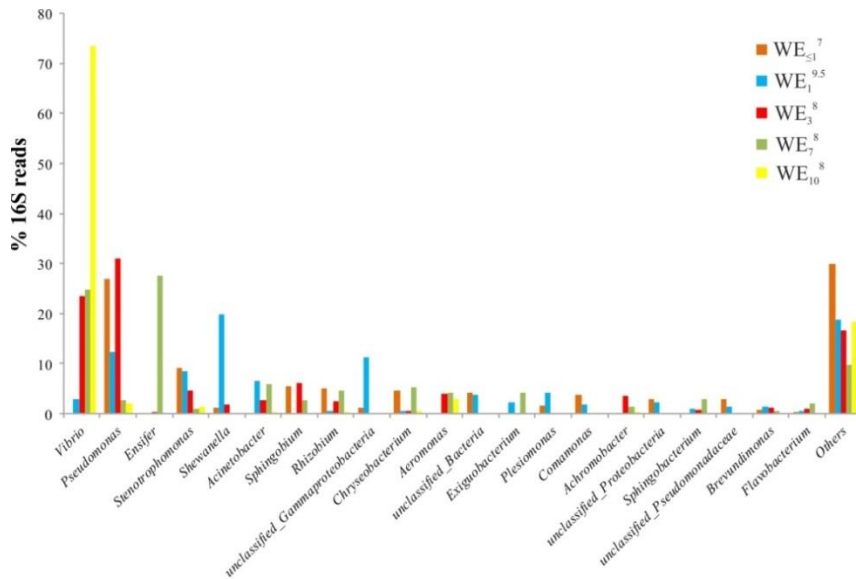
Figure 19. Main bacterial genera found in the metagenomes using 16S genes sequences. The 16S genes were compared against Ribosomal Database Protein. The most abundant genus for wild-(A) and (B) farmed-eels are represented.
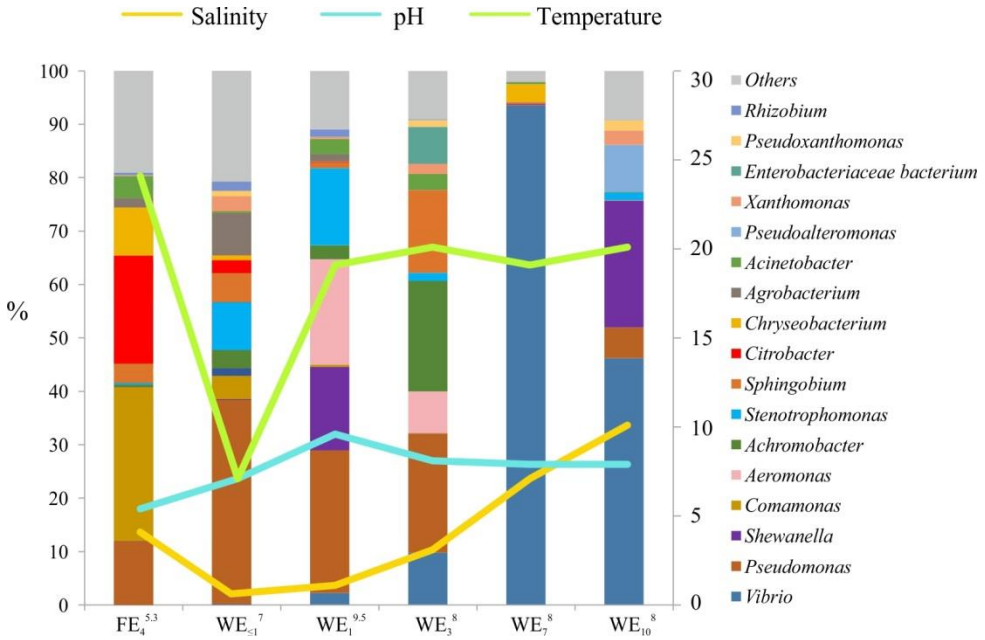
Figure 20. Main bacterial genera *vs* physico-chemical parameters of water. The proportion of the genus present in the sequenced metagenomes was assessed using all genes. The annotation system depended on the sequencing technology: Illumina, based on assembled contigs (BLASTP); 454, based on reads (BLASTX). Salinity, temperature and pH values are represented according to the scale on the right.

Then, we identified the bacterial species in the metagenomes by counting the number of reads recruited against genomes from the most similar strain (**Figure 21**). First, we focused our attention on the genus *Vibrio* because it includes several pathogenic species for fish and humans. Again, a direct relation between the species attached to mucus of wild-eels and the water salinity was observed. Thus, *V. anguillarum*, *V. metoecus* (a recently described *V. cholerae*-closely related species)(Kirchberger *et al*., 2014) and *V. cholerae* were the most abundant *Vibrio* species in mucus from eels inhabiting waters of salinity 7-10 , 3 and 1 g/l, respectively (**Figure 21**).*V. fischeri*(currently *Aliivibrio fischeri*), *V. furnissii* and *V. vulnificus* were also part of the *Vibrio* species associated to eel mucus in the Nature Parks (**Figure 21**).All the mentioned species, except the symbiotic species *Al. fischeri*, are well known human and/or fish pathogens.

Among them, *V. anguillarum* and *V. vulnificus* are pathogenic fora wide range of teleost, including eels(Amaro and Biosca, 1996; Frans *et al*., 2011).
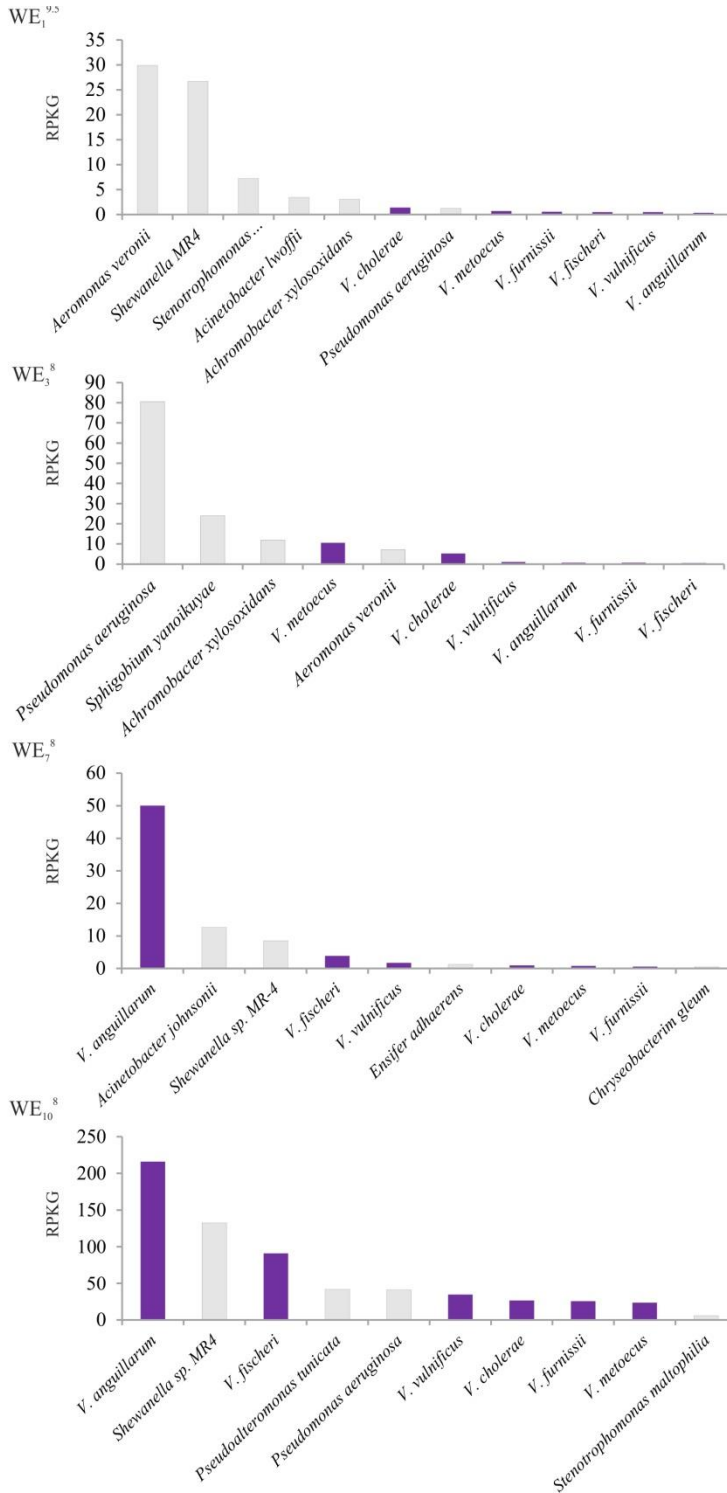
Figure 21.Microbiota composition of skin-mucus from wild-eels. The abundance of different species in the metagenome was calculated by counting the number of reads recruited against from each species. The result was normalized dividing by the size of the genome (kb) and the dataset (Gb) (RPKG). *Aliivibrio (Vibrio) fischeri* was considered inside the *Vibrio* genus.

In both species, fish virulence relays, partially, on a virulence plasmid; pJM1 of *V. anguillarum* encodes an iron-uptake system depending on the siderophore anguibactin while pVvbt2 of *V. vulnificus* encodes a partially described system of resistance to the eel innate immunity(Valiente *et al*., 2008b; Naka *et al*., 2011). We did not find any evidence of the presence of both plasmids (or their genes) in our metagenomes concluding that the strains present in the SM-microbiota were eel avirulent (for more details see 5.2.). Interestingly, the proportion of *V. vulnificus* in SM-metagenomes was the same in all the metagenomes (around 0.5-1) with the exception of WE_{10}^8 where it reached 34.64 (**Figure 21**). This finding supports the hypothesis that *V. vulnificus* could be part of the resident SM-microbiota, at least those from natural estuarine environments. Those eels carrying eel-pathogenic clones could be analysed with a low probability because any stress causing a drop in their defences would lead to the death of the animal.

V. metoecus was recently described as the closest relative of *V. cholerae* (Kirchberger *et al*., 2014)*)*. This species has been co-isolated with *V. cholerae* in coastal waters and found in clinical specimens in the United States (Orata *et al*., 2015). In this work we found both species together in only one of the SM-metagenomes, that corresponding to eels from Prado Cabanes, a natural brackish-water environment. Surprisingly, *V. metoecus* was the main *Vibrio* species found in this metagenome. Originally, *V. metoecus* was described from a non-virulent clone isolated from Chesapeak Bay (the USA) and designated as *Vibrio sp. RC341*(Haley *et al*., 2010a). In addition, we were successful in the isolation of *V. metoecus* from skin-mucus of wild-eels using culture techniques and we sequenced its genome using Illumina. We performed an ANI analysis by using the genome of this isolate as reference and found a percentage of 98.24%，not enough to assign our sequences to the isolate *RC341* but sufficient to assign them to the same species. According to this, we designated the sequenced strain as *Vibrio metoecus* M12v. The comparison of M12v-contigs with the published

genomes of *V. metoecus* (5 strains isolated from a brackish coastal pond on the US east coast, as well as four clinical)and *V. cholerae* highlighted that M12v shared virulence characteristics with *V. cholerae* that make this isolate different from all the *V. metoecus* strains so far sequenced and studied. Thus, M12v-strain presented in common with pandemic *V. cholerae,* 1 out of 10 CTX-phi-genes, the *zot* gene (Zonula Occludens toxin), 1 out of 3 RS1-genes (a satellite phage), and the loci for a toxin co-regulated pilus (TCP) biogenesis and for ElTor-RTX biosynthesis, modification and transport while lacked of the pathogenicity islands VSP-1 and -2, as the rest of *V. metoecus* strains (**Figure 22**). TCP is also involved in the mucus colonization of the small intestine by pandemic *V. cholerae*(Herrington, 1988), which suggests a general role for this pilus in the colonization of any mucosal surface, including mucosal fish surface. Other functions that were represented in M12v and *V. cholerae* El Tor but were absent in strains of *V. metoecus* were vibriobactin (siderophore) biosynthetic cluster, and genes for nickel transport, a phosphotransferase system and a monovalent cation/proton antiporter (F**igure 22 and 23**). Finally, downstream the TCP locus we found a type 6 secretion system (T6SS) absent in ElTor but previously sequenced in other *V. cholerae* strains (strains 1421-77 DA89). This result makes M12v isolate quite interesting because, apparently, is able to produce the receptor for phage, the TCP pilus, and contains in its genome evidences of having harbored both the phage and the satellite phage RS1. On the other hand, the Rtx toxin of *V. cholerae* has been related with protection against amoeba in the environment(Lee *et al.*, 2007). Its presence in our *V. metoecus* isolate is an evidence of selective pressure for this kind of genes in the environment. The presence of these genes, especially those involved in virulence (siderophore and TCP), suggests M12v would be more virulent than the rest of *V. metoecus* known strains. Moreover, we can assume these genes were acquired by HGT, meaning that *V. metoecus* avirulent strains can become pathogenic by acquiring *V. cholerae* virulence genes.

Finally, another species identified in the SM-metagenomes from wild-eels were also opportunistic pathogens; *Pseudomonas aeruginosa*, *Stenotrophomonas maltophilia*, *Achromobacter xylosoxidans*, *Aeromonas veronii,* this last being

considered as potential pathogen for eels (Frans *et al*., 2011; Yi *et al*., 2013). Other like *Sphingobium yanoikuyae* was also abundant the SM-microbiota of some samples and seemed to have an important role in genetic information interchange (see below). *Sph. yanoikuyae* was originally isolated from a clinical specimen and can degrade a variety of chemicals in the environment such as aromatic and chloroaromatic compounds, phenols like nonylphenol and pentachlorophenol, herbicides such as (RS)-2-(4-chloro-2-methylphenoxy) propionic acid and hexachlorocyclohexane, and polycyclic aromatic hydrocarbons. Pesticides are used in the surrounding park rice fields, which correlates with the presence of bacteria hypothetically able to degrade these compounds.
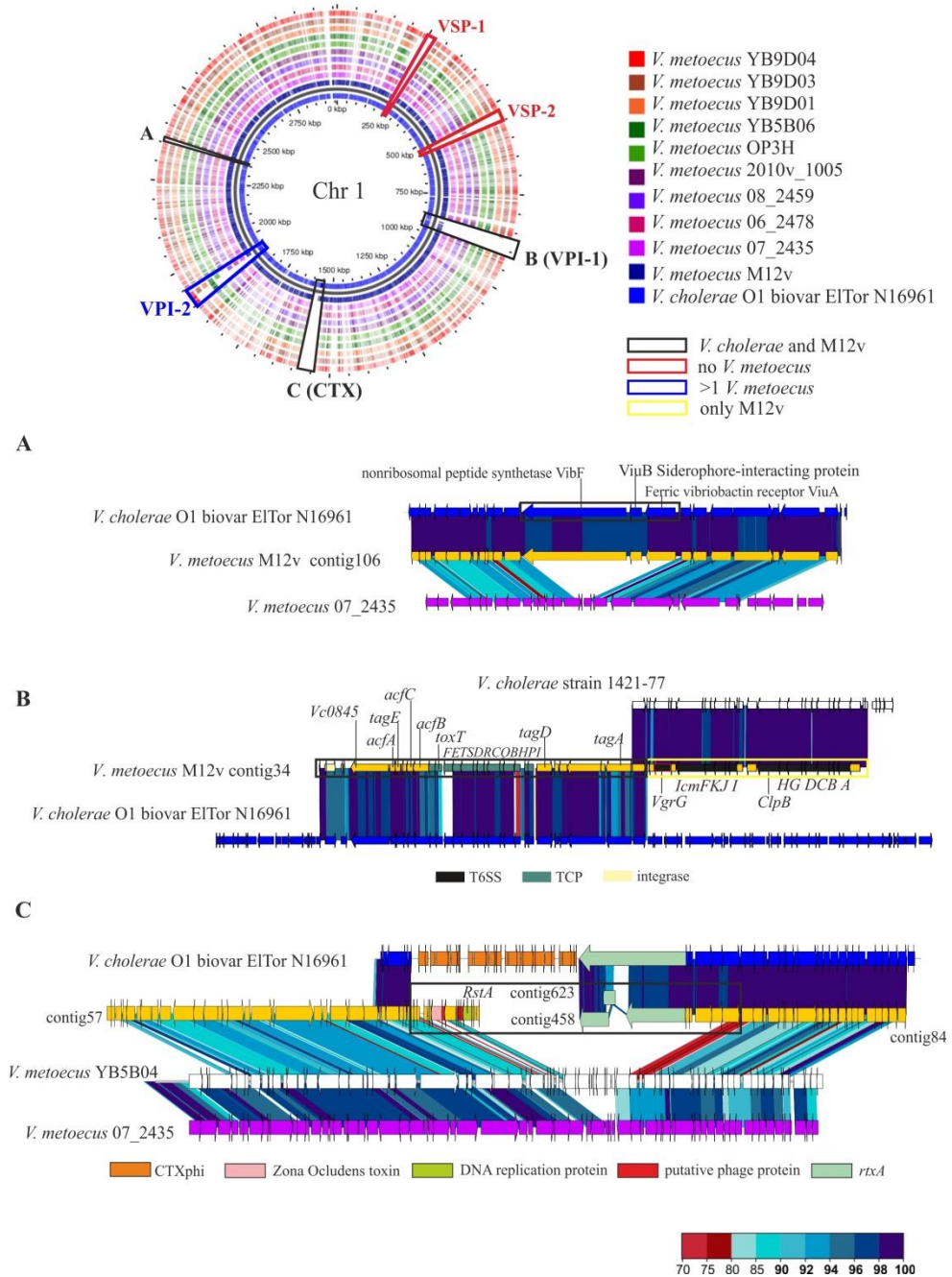
Figure 22. *V. metoecus* M12vBLAST atlas. The chromosome 1 of *V. cholerae* O1 biovar ElTor N16961 (reference) was compared with *V. metoecus* deposited in Genbank and the strain sequenced in this study. Hits of BLASTN analysis are shown using blocks that represent genes. Each ring belongs to one strain. Islands were differentiated according to the strains that share them. Thus, black boxes were used to highlight islands shared by *V. cholerae* and *V. metoecus* M12v while blue and red were used when more than 1 and none *V. metoecus* strain had the island, respectively. Islands represented only in M12v were plotted using BLASTX against reference and the most similar *V. metoecus* strain. *Vibrio* pathogenic island one, VPI-1; *Vibrio* pathogenic island two, VPI-2; CTXφ bacteriophage, CTX.
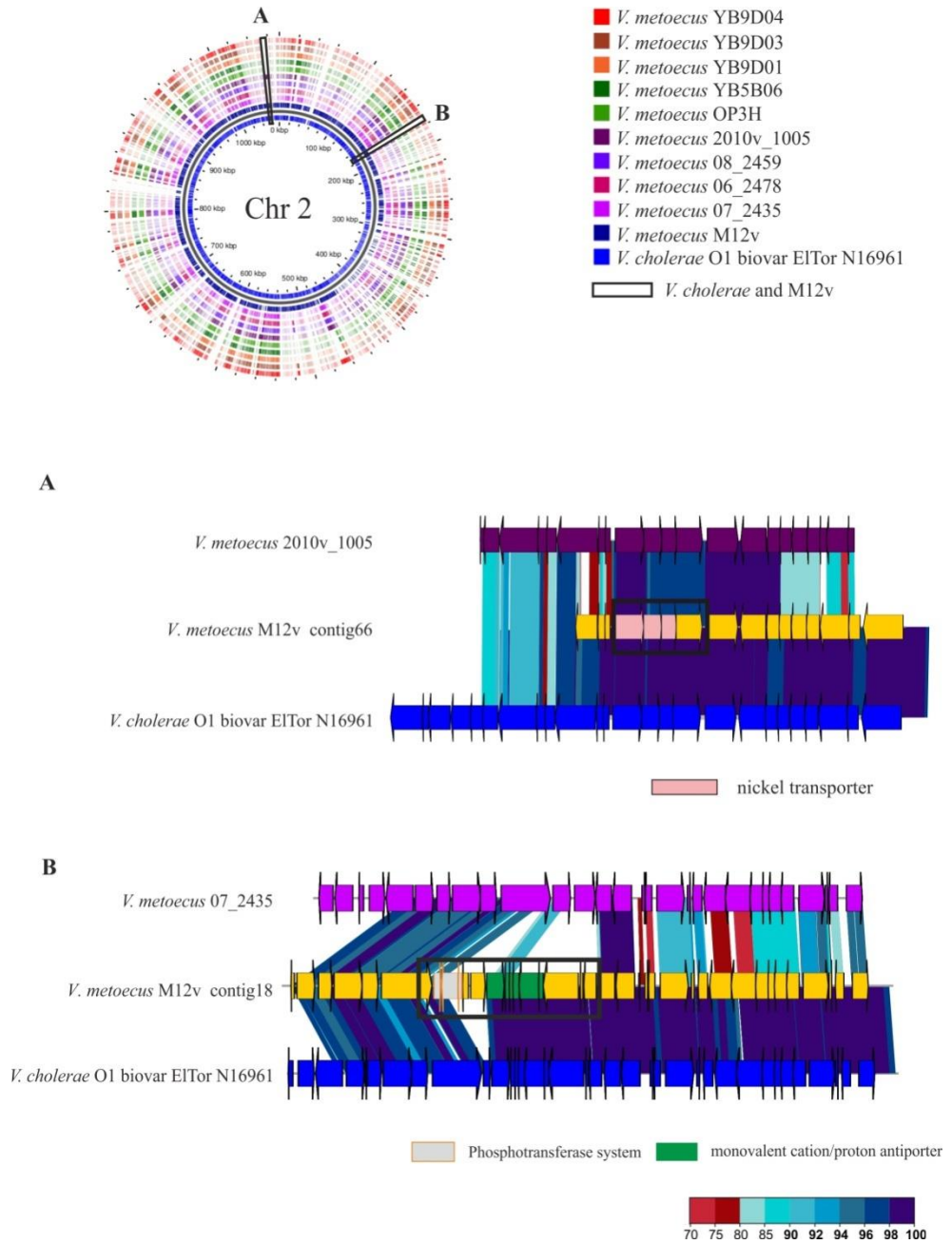
Figure 23. *V. metoecus* **M12vBLAST atlas.** The chromosome 2 of *V. cholerae* O1 biovar ElTor N16961 (reference) was compared with *V. metoecus* deposited in Genbank and the strain sequenced in this study. Hits of BLASTN analysis are shown using blocks that represent genes. Each ring belongs to one strain. Black boxes were used to highlight islands shared by *V. cholerae* and *V. metoecus* M12v which were plotted using BLASTX against reference and the most similar *V. metoecus* strain.

5.1.2.Mucus-attached *vs* free-living bacteria

We compared the attached-host microbiota with that of the surrounding water and found that they were completely different at all the metagenome's characterization levels. First, W-metagenome's GC profile was multimodal instead of bimodal with none of the maximum peaks corresponding to the two peaks observed in the SM-metagenome's GC profiles (**Figure 17C**). In addition, W-metagenome did not present a clear major bacterial class, being Actinobacteria (20%) and Gammaproteobacteria (17%) the dominant classes followed by Betaproteobacteria, Alphaproteobacteria, Cyanobacteria and Flavobacteria, present in similar percentages (9-13 %) (**Figure 24**).
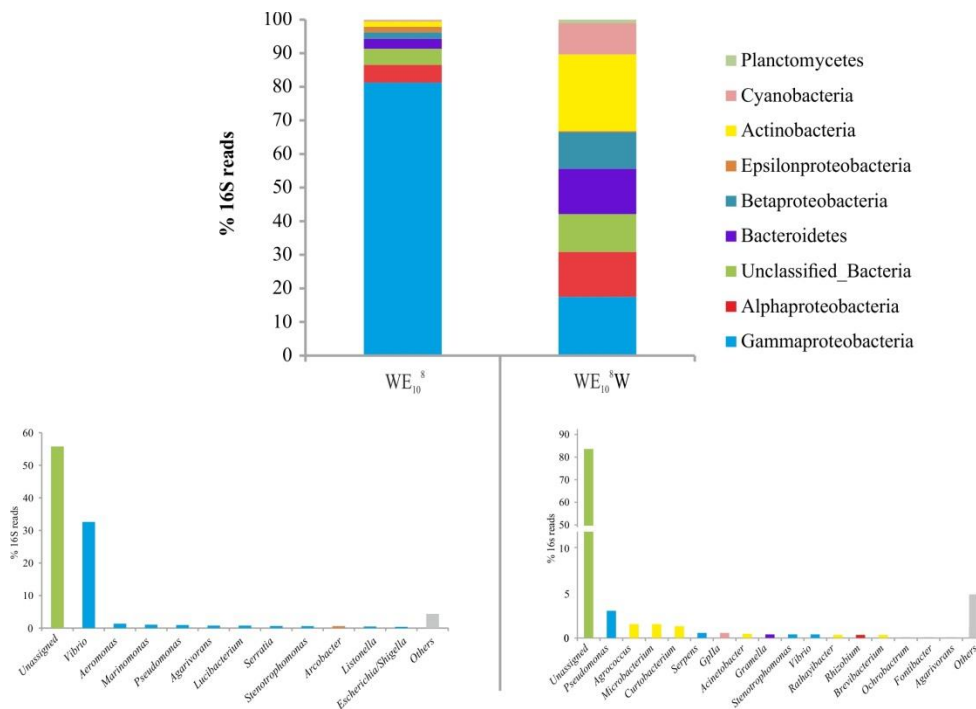
Figure 24. Differences between attached and free-living microbiota at class and genus level. 16S reads found in WE_{10}^8 and WE_{10}^8W metagenomes were compared to Ribosomal Database Protein in order to classify them taxonomically.

Using 16S, the percentage of reads that could not be assigned to any known genus was higher in W-metagenome (82%) than in SM-metagenome (55%) (**Figure 24**). Regarding the identified genus, *Pseudomonas*, *Agrococcus*, *Microbacterium* and *Curtobacterium* were the dominant genus in W-

metagenome, while *Vibrio* was a minor genus. Interestingly, *Pseudomonas* was present in the same proportions in both metagenomes (around 2%) while *Vibrio* was overrepresented in skin-mucus (32% *vs* 1% in water), which suggests that mucus concentrates the vibrios present in water. Finally, we identified the bacterial species present in W-metagenome by the annotation of the genes and calculate its relative abundance by recruiting of reads against concatenated contigs. Surprisingly, no known species was identified within the dominant genera, all of them probably belonging to the "nonassigned genus" group (**Figure 54**). The dominant identified OTUs (operational taxonomic units) were: *CandidatusAquiluna* (Actinobacteria), *Flavobacterium* sp., *unculteredMediterranean phages*, *Chroococcales* (Cyanobacteria) and *CandidatusPelagibacter*. It was interesting to find among them representatives of the most abundant uncultured bacteria in seawater (*Candidatus Aquiluna* and *Candidatus Pelagibacter*) as well as marine bacterial phages.

We tried to identify *Vibrio* species in W-metagenome to compare them with those found in skin-mucus since water was taken where eels were fished. According to 16S results *Vibrio sp.* LW004.09 (isolated from lobster in a hatchery) and *Vibrio cholerae* were present in both samples while SM-metagenome hosted many other vibrios. The fact that the most abundant species in eel-mucus, *V. anguillarum*, was not present in water, or was present in such a low abundance that none contig was assembled suggests, again, that eel-mucus concentrates the few *V. anguillarum* free-live isolates, which after attachment could multiply and increase its presence in mucus.

In parallel, we tried to isolate *Vibrio* spp. from water by using culture methods and we were able to isolate several putative *Vibrio*, one of which was identified as *V. vulnificus* by PCR using the primers for the identification of this species (Vvh primers). The isolate was sequenced by using Illumina methodology and its genome was also analyzed for the presence of virulence genes. We detected an *rtxA* gene in the genome of this strain. Rtx are toxins that present internal or external repeats (repeat in toxin). *V. vulnificus* produces various types of Rtx toxins all of them classified within the subfamily MARTX (Multifunctional-

autoprocessing-repeat-in-toxin). MARTXVv present two external modules that contain the repeated sequences and an internal module that contains different effector domains. Once secreted from the bacterium, the repeat regions of MARTXVv bind to eukaryotic membranes on target cells and form a pore that allows the trans-location of the central portion of the toxin across the membrane (Gavin and Satchell, 2015). This internal portion contains various effector domains, one of them with protease activity (CPD, cysteine protease domain), once activated, liberates the rest of the effector domains, which confer the cytotoxic functions of the toxin (Gavin and Satchell, 2015). MARTXVv of biotype 2 is both an important virulence factor for humans and fish and a weapon against the main predators of bacteria in biofilms, the amoeba (Lee *et al*., 2013) Eleven effector domains have been found in the various *Vibrio* MARTX toxins, the last one was described this year and designated ChaN (Haem-binding uptake). Precisely this domain is present in MARTX toxin of our isolate (**Figure 25**). In this figure, the rest of predicted functional domains are shown. This finding provides evidence that these new variants are continuously emerging in the environment probably by recombination either with *rtxA* genes carried on MGEs or with *rtxA* genes of other environmental vibrios.

Finally, neither pVvBt2 nor its main ORFs were detected in the genome of this isolate, confirming that probably was avirulent for eels due to its predicted inability to grow in eel blood. This result together with the absence of eel-pathogenic clones in skin-mucus suggests that the pathogenic clones are rare in the environment and are only detected after an amplification event associated to outbreaks in farms.
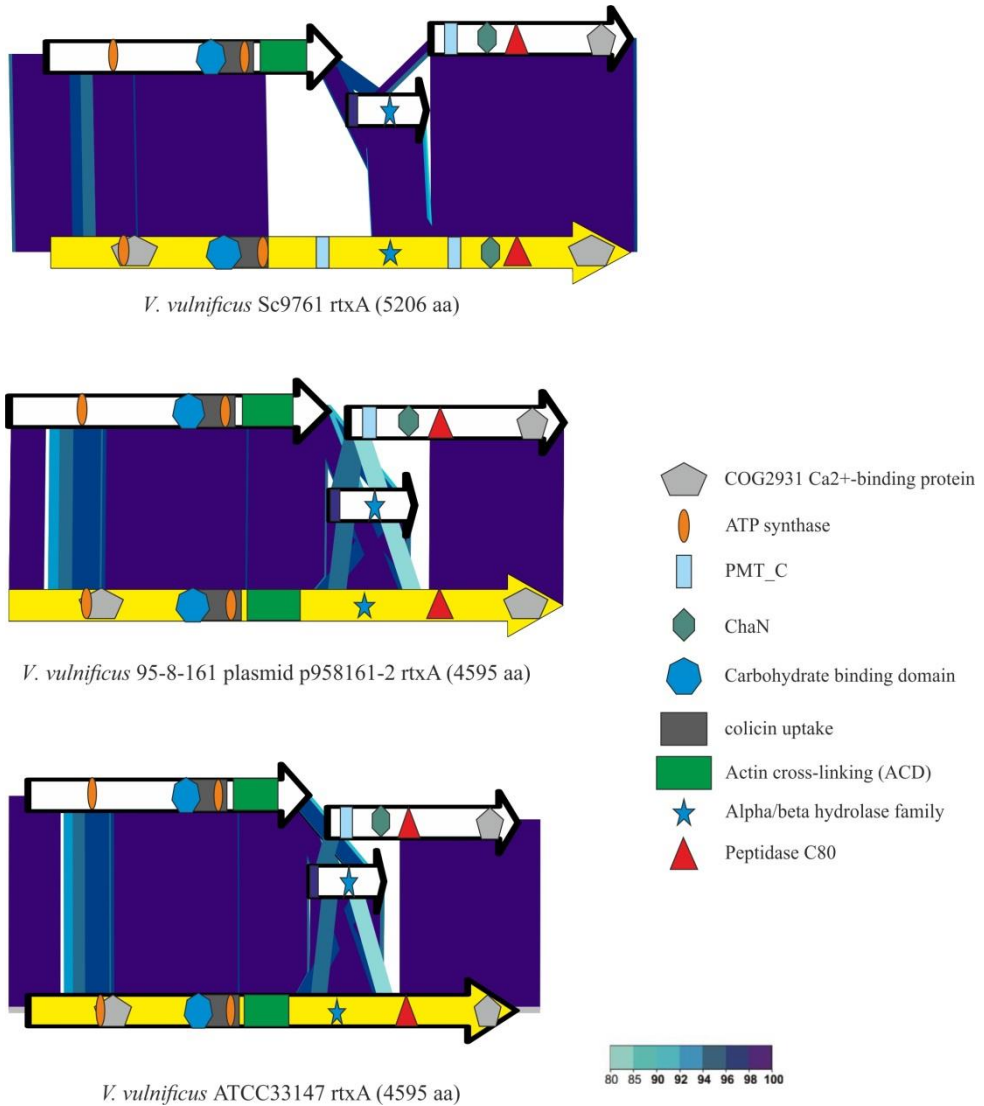
Figure 25. *rtxA* gene comparison. The gene rtxA sequenced in the *V. vulnificus* isolate in W-metagenome was compared to the most similar in Genbank using tBLASTX. Minimum identity and length alignment of 80 and 150 were used, respectively. White narrows represent the three contigs in which the *rtxA* was assembled. The yellow arrows are those deposited in Genbank.

5. 1. 3. Functional differences between mucus-attached and free-living microbiota

The differences in bacterial composition between host-attached and free-living bacteria clearly suggest that mucus acts as a selective niche that is enriched in the microorganisms able to resist the mucosal innate immunity of eels, compete

against commensal bacteria for the resources and avoid phage infection (see below). For all these, we suggest that only bacteria with specific capabilities will be able to attach, maintain and grow in skin-mucus. We hypothesized that these "specialised" bacteria are present in low concentration in water and that because of these "specialised capabilities" they attach to and multiply in skin-mucus. Supporting this hypothesis we found that alpha diversity indexes decreased 4 times in attached microbiota compared with water microbiota (**Table 3**).

To find out what selective pressures were acting in skin-mucus and how the successful colonizers overcome them, we analysed and compared the functional capabilities of SM- and W-microbiotas using MG-RAST. This pipeline is able to classify metagenomic reads in functional categories according to SEED hierarchical classification. At the highest classification level (level 1), cell-wall/capsule, membrane-transport, virulence/disease/defence, regulation/cell-signalling, iron-acquisition/metabolism, nitrogen/sulphur/potassium-metabolism and motility/chemotaxis were the over-represented categories in SM-metagenome and clustering-based-subsystems (unknown function), amino-acids/derivatives, protein metabolism, nucleosides/nucleotides, phages/prophages/transposable-elements/plasmids and photosynthesis were the over-represented ones in W-metagenome (**Figure 26**).

To unravel the specific functions behind these categories we went deep in the classification (level 3) by applying two filters, >10 times overrepresented and <0.02% present (**Figure 27A**) and >5 times overrepresented and >0.02% present (**Figure 27B**). We analysed in depth the category "cholera toxin" and those related with bacterial secretion systems. The specific genes for cholera toxin were not found in either of the metagenomes. Instead of it, we found in SM-microbiota those corresponding to the toxins Ace (accessory cholera toxin) and ZOT (zonula occludens toxin), both present in CTX phage of *V. cholerae*(Dalsgaard *et al*., 2001). This result suggests that *V. cholerae* clones present in SM-microbiota contain the phage but without the genes for the cholera toxin. Again, we found that pathogenic clones, in this case belonging to V. cholerae, are not abundant in the environment, even though considering that

mucus seems to concentrate vibrios. Related to secretion systems, all types were overrepresented in mucus samples, including T6SS which is known to be essential to compete against other bacteria (**Figure 28**) (Ma and Mekalanos, 2010; Hachani *et al*., 2011).
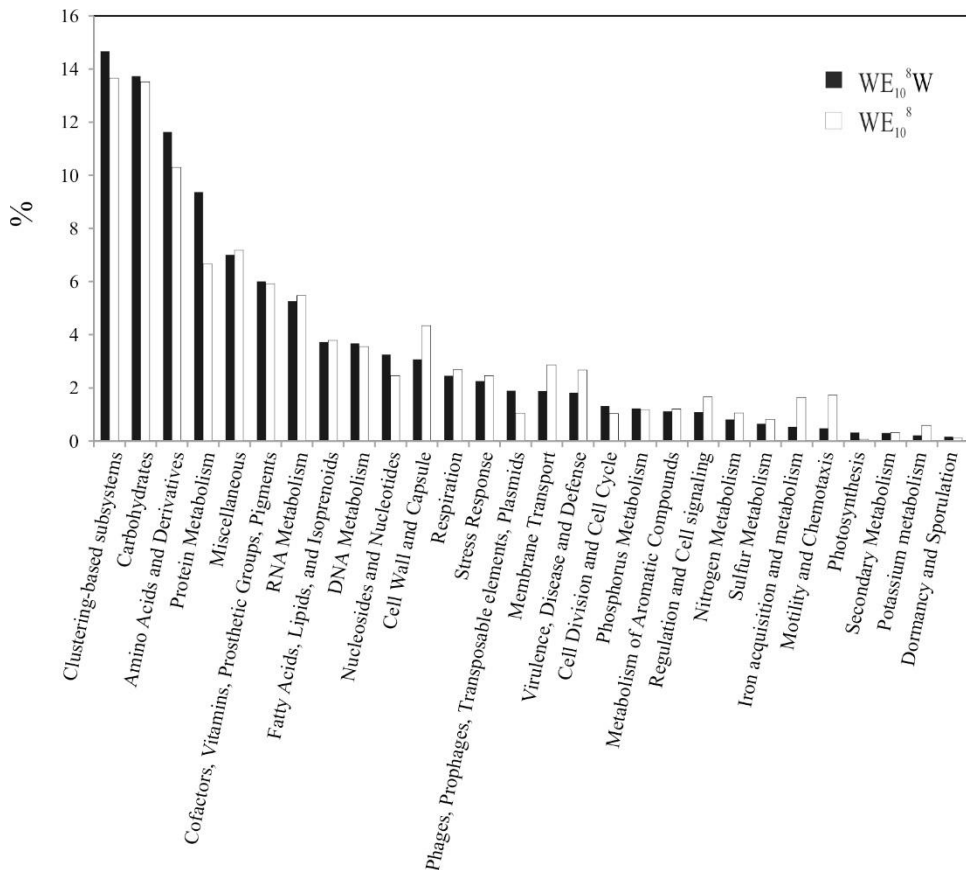
Figure 26. Functional capacities of microbiota attached to SM and free-living in W. The presence of each category at was measured by calculating the percentage of reads assigned from the whole metagenome annotated.

Figure 27. Functional comparison of microbiota attached to SM and W microbiota attached to SM and W at level 3. We calculate the difference in presence (%) between $WE_{10}{}^8$ (white) and $WE_{10}{}^8W$ (black). We consider a function overrepresented when its presence was more than 0.02% and 5 times more abundant or if it was 10 times more abundant when the percentage was lower than 0.02. A, genes overrepresented at least 10 times and low presence (<0.02 %); B, at least 5 times and percentage >0.02.

Figure 28. Membrane transport functional categories in SM and W-metagenomes. The abundance of functions categorized in "Membrane transport" was compared between WE_{10}^8 (white)and WE_{10}^8W (black) microbiotas. The functions inside "-" category were plotted in the upper box.

We analysed in depth the SM-metagenomes to look for other toxin genes. We found a paralog for *rtxA* gene, the main virulence factor for *V. vulnificus*, in an *Al. fischeri* contig, but without significant similarity with any of the *rtxA* genes described so far in *V. vulnificus*. In fact, this toxin shows the same structure than the *rtxA* toxin previously described in *Al. fischeri* (**Figure25**).We also found two genes for toxins in two *Pseudomonas* contigs: ExoU and ToxA which were 100 and 99 % identical to their respective copies in *P. aeruginosa* WH-SGI-V-07317 (see 5.2.1.)Both toxins are injected by T3SSsinto macrophages protecting bacteria from the immune system (Michalska and Wolf, 2015). ExoU is a potent phospholipase A2 that cleaves plasma membrane phospholipids causing rapid lysis of macrophages and other cell types(Sato and Frank, 2014) and ToxA is an ADP-ribosylating toxin that induces death by apoptosis after modifying a residue of diftamide of EF-2 (Sharma and FitzGerald, 2010). The targets for both toxins are universally conserved among eukaryotic organisms, including the amoeba. We hypothesize that alike the MARTXVv, both toxins could be used by commensal *Pseudomonas* to fight against amoeba in the biofilms, thus supporting our original hypothesis about the existence of environmental reservoirs for virulence genes.

We also found interesting to look for differences in antibiotic resistance genes between SM- and W-microbiotas (**Figure 29**). The same pipeline in MG-RAST classified a series of reads inside "Resistance to antibiotics and toxic compounds" (RATC) category. The comparison of abundance in attached and free-living microbiota showed that genes for cobalt-zinc-cadmium, multidrug efflux pumps, copper, aminoglycoside adenylyltransferases, chromium compounds, fosfomycin resistance, lysozyme inhibitors, MAR (Multiple Antibiotic Resistance) locus and bile hydrolysis were duplicated in SM-microbiota. Some of these categories could be linked with resistance to antibiotics (aminoglycoside adenylyltransferases, fosfomycin resistance), others with resistance to mucosal innate immunity (lysozyme inhibitors, bile hydrolysis...) and others with both (multidrug efflux pumps..) (Piddock, 2006). The higher number of genes for resistance to heavy metals in the attached microbiota than in free-living could be related with a higher concentration of these toxic compounds on the mucin layer. In fact, heavy-metals can bind to negative charges of the mucin of the mucus covering the eel epidermis, which could concentrate these metals (Volesky and Holan, 1995).

Resistance to fluoroquinolones was clearly dominant in W-metagenome (**Figure 29**). Fluoroquinolones are used extensively in both human and veterinary medicine, being considered important weapons against Gram-negative and Gram-positive bacteria(Van Doorslaer *et al*., 2014). These antibiotics can be excreted by human and animals as unchanged compounds into hospital or municipal sewage (Kaplan *et al*., 2013) resist the wastewater treatment plants and remain for decades in the environment acting on resident microbial communityand selecting for those harboring resistances (Weber *et al*., 1998; Córdova-Kreylos and Scow, 2007; Kümmerer, 2009). Three big Hospitals are located in the surroundings of the Ebro Delta, which could explain the high prevalence of fluoroquinolone resistance genes in W-metagenome.

Figure 29. Abundance of genes involved in resistance to antibiotic and toxic compounds. The abundance was calculated using MG-RAST pipeline and it was compared manually using Microsoft Excel. * at least 2 times more abundant in SM-microbiota than W; ** at least 2 times more abundant in water than in SM-microbiota; *** at least 2 times more abundant in farmed than in wild eels.

Based on all the above, we propose the following functions as essential for bacteria to attach and maintain in the skin-mucus of eels:

i)   biofilm production and regulation (exopolysaccharide-production, sigma-dependent-biofilm-formation, VieSAB-signal-transduction-system)

ii)   bacterial communication (quorum-sensing, autoinducer-2-transport/processing)

iii)   bacterial competence (bacteriocin-like-peptides, ABC-transporter-peptides, T6SS)

iv)   Adherence factors (colonization-factor-antigen-1, curli, accessory-colonization-factor, *Campylobacter*-adhesion)

v)   Resistance to innate immunity (lysozyme resistance, bile hydrolysis, multidrug efflux pumps, siderophores, hemin-uptake, *Pseudomonas* toxins, T3SS)

vi)   Resistance to predators in biofilms (*Pseudomonas* toxins)

vii)   Metals and drugs resistance (cobalt-zinc-cadmium, copper, multidrug efflux pumps)

In contrast to this, only two categories at level 3 were overrepresented in the W-metagenome in comparison to mucus-metagenome, *Staphylococcal phi-Mu50B-like prophages* and *Phage neck proteins* (data not shown), probably due to the abundance of phages in this sample.

5.1.3. Farmed *vs* Wild

Finally, we compared the taxonomic composition of the SM-microbiota from wild- and farmed- eels. For this comparison, we used the metagenomes $FE_4^{5.3}$ and WE_3^8 because of its similarity in water salinity. The GC profiles from both metagenomes were bimodal but the lowest GC peak was not the same; around 35% in farmed *vs* 45% in wild-eels (**Figure 17B**). Differences in microbial composition were also seen at class level since Alphaproteobacteria was present in higher levels in mucus from farmed eels (**Figure 18**). At genus level, *Vibrio,*

Aeromonas and *Shewanella* were not present in farmed-eels-metagenome while *Comamonas, Citrobacter* and *Chryseobacterium* were significantly abundant only when eels were grown in aquaculture facilities (**Figure 19 and 20**). As expected, the combination of salinity and low pH in the farm seems to be detrimental for *Vibrio*. Interestingly, farmed-eels maintained some of the genera found in the eels fished at wet lands as *Pseudomonas*, *Acinetobacter, Stenotrophomonas* and *Sphingobium* in significant proportions, according to 16S results. A deepest analysis on the microbiota attached to farmed-eels by studying the contigs assembled and recruiting reads against them showed that 52% of the contigs matched to *Comamonas testosteroni* (data not shown). The great abundance of this species in these conditions can be explained by its ability to resist low pH and its resistance to antibiotics (Oppermann *et al*., 1996; Ma *et al*., 2009). Currently, a study in culturable microbiota from farmed Zebrafish identified this species in gut contents (Cantas *et al*., 2012).

Antimicrobials are widely used in aquaculture facilities to control bacterial diseases (Defoirdt *et al*., 2011). To find out whether the continuous use of antimicrobials to control diseases outbreaks was biasing microbiota's genetic composition, we analyzed in depth the differences between farmed and wild-eel metagenomes as explained before. **Figure 28** shows the different resistance categories distributed by metagenome. On the contrary to that expected, farmed-metagenome didn't present more antimicrobial-related genes than wild-metagenome, with only one exception, resistance to zinc and cadmium.

Unexpectedly, methicillin- and vancomycin-resistance clearly were overrepresented in SM from wild-eels. Methicillin, a semisynthetic penicillinase was developed in response to the emergence and spread of penicillin resistance and introduced into the clinic in 1959. Methicillin-resistant *S. aureus* (MRSA) emerged throughout the 1960s in many countries, including those where methicillin was not available, and it is now ubiquitous worldwide (Jevons *et al*., 1963; Diekema *et al*., 2004). Multiple genetic mechanisms have been reported to underlay this resistance, among them the production of specific beta-lactamases or the modification of penicillin binding proteins (Arede *et al*., 2013; Zapun *et*

al., 2008). These genes were not related to *Staphylococcus*-contigs which suggests that methicillin resistance is ancient and widespread in the environment. Vancomycin is used as a first-line treatment for complicated skin infections, bloodstream infections, endocarditis, bone, joint infections, and meningitis caused by methicillin-resistant *Staphylococcus aureus* (Deresinski, 2009; Zar *et al*., 2007). Vancomycin-resistance genes were also found in SM-metagenomes from river and lake-eels, environments located in the North-Western and Eastern areas of Spain. Although these genes are usually found in *Staphylococcus* and *Enterococcus* clinical strains (Gardete and Tomasz, 2014) new variants of *van* genes have also been found in 30,000-year-old Beringian permafrost (D'Costa *et al*., 2011). Again, we can relate these resistances neither with *Staphylococcus*- nor *Enterococcus*-related contigs, which suggests that vancomycin resistance genes, as methicillin resistance genes, are ancient and widespread in the environment.

5.2. MGE in the attached microbiota: flexible metagenome

In order to study the fully pathogenic capacity of eel microbiota, we assessed an important repository of virulence genes: the flexible metagenome (or MGE present in a metagenome). The role of MGE in the emergence of new virulent variants/species is well documented(Waldor and Mekalanos, 1996; Valiente *et al*., 2008a).Moreover, to detect signals of DNA interchange events between the microorganisms attached to skin-mucus would clear the picture of microbial interactions *in vivo*. As we mentioned before, the main gene responsible for serum resistance during eel infection by *V. vulnificus* biotype 2, *vep07*, has only been found in a plasmid acquired by this species through horizontal transfer(Lee *et al*., 2008). First of all, we searched for pVvBt2 and *vep07*in our dataset by using BLASTP and BLASTN without any luck. We had the same result when we looked for other virulence plasmids deposited in Genbank. Since no methodology to find MGE in metagenomes was found in the specialised literature, we decided to develop a specific protocol or workflow (see 4.9).

5.2.1. The flexible metagenome in the superficial microbiota of *A. anguilla*

We analysed 6 host-associated and 3 aquatic metagenomes (**Table 6**) and we were able to find putative MGE in all of them. Usually, an insertion event entails a change in the nucleotide pattern that could be detected in a contig of enough size. Interestingly, the PCA analysis from HUP values showed contigs assigned to the same bacterial species with and without MGE forming clusters clearly distinguishable (**Figure 30 and 31**). This approach also allowed us to differentiate those contigs annotated as phages, including replicating phages and prophages (see 5.3. Replicating phages in the superficial microbiota of eels) (**Figure 32**). Regarding GC variation, the interchange of DNA material between GC-rich populations was easily detected since various genera found between 50 to 70 GC composition, like *Pseudomonas* and *Sphingobium*, apparently interchanged large quantities of DNA(>35 kb) (**Figure 33**). However, we could not find inserted sequences looking at the GC content in the samples majorly composed by *Vibrio*.

Using the workflow described, we were able to find 118 and 12 cpMGE in $WE_3{}^8$ and $WE_{10}{}^8$, respectively, and none in the rest of SM-metagenomes probably because very few contigs bigger than 10 kb were assembled (**Figure 34**). Once we identified the cpMGE in the metagenome, we re-annotated the genes inside using HHpred and, then, classify them according to its content. **Table 4** shows the confirmed contigs and the type of MGE that contained.

Figure 30. Representation of hexanucleotide usage pattern (HUP) using PCA of aquatic related metagenomes. HUP of contigs bigger 10 kb was analyzed for the different metagenomes. Contigs annotated to the most abundant genus and those that were considered putative mobile genetic elements were coloured differently. A, Indian lake; B, Swedish lake; C, Mediterranean Sea (DCM).
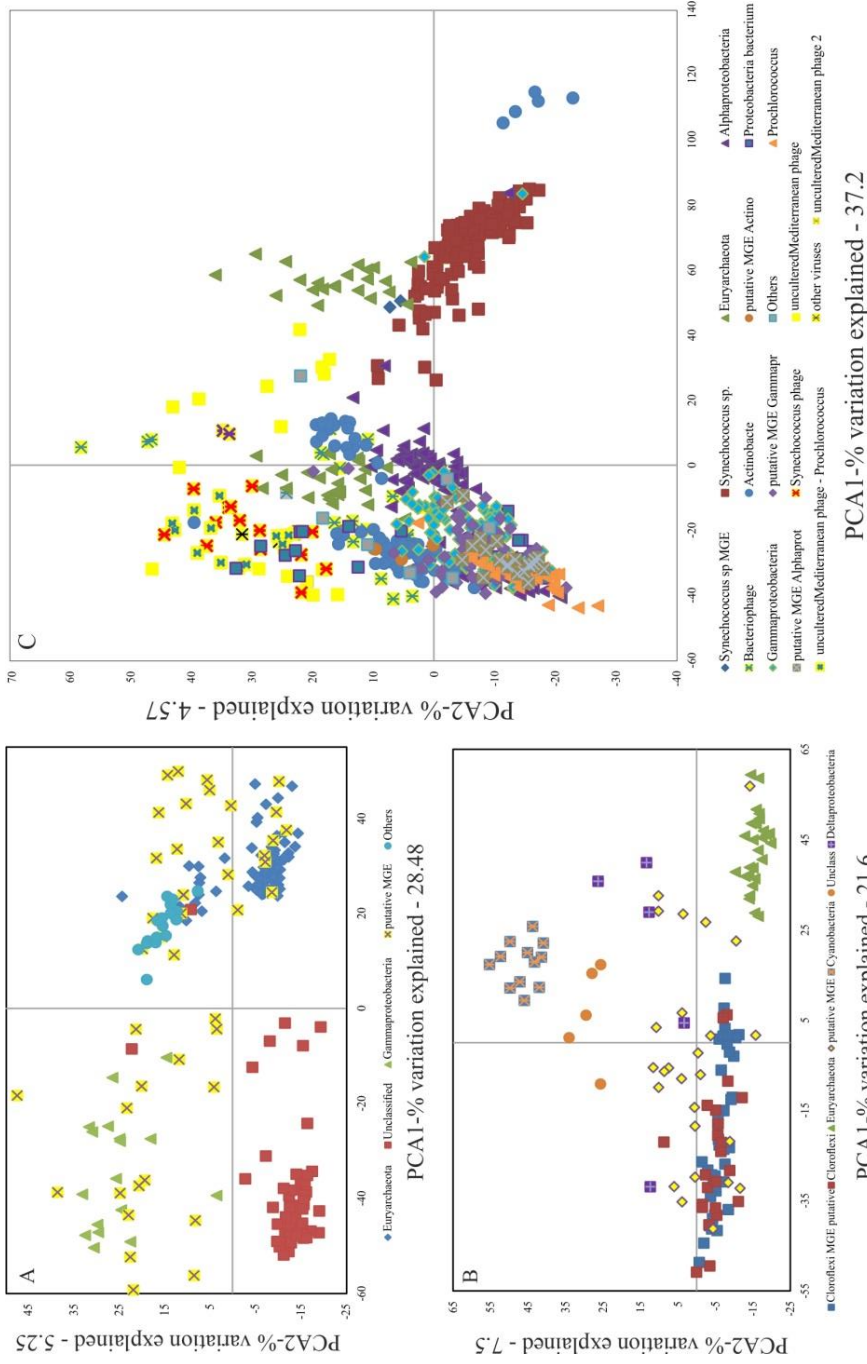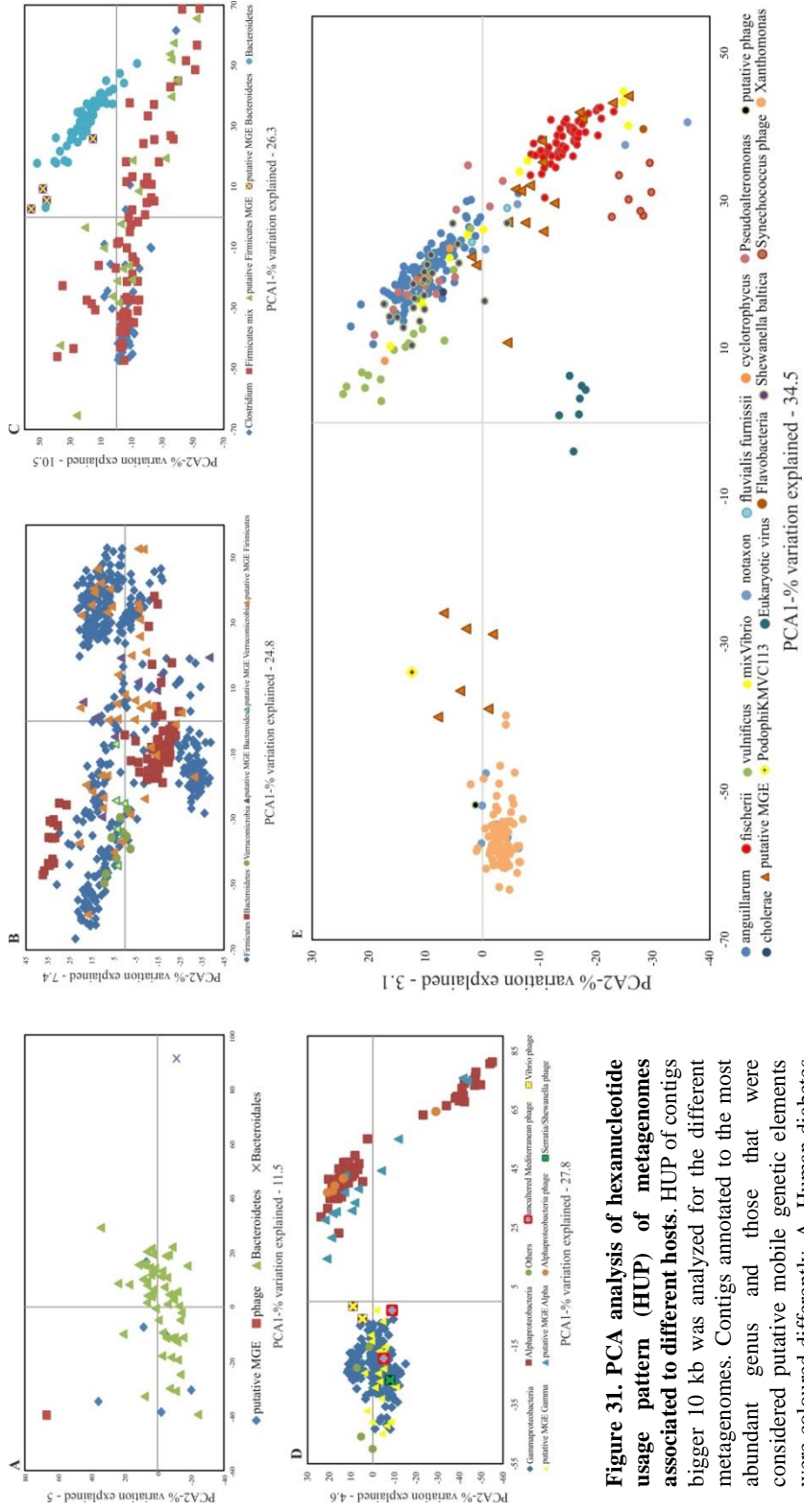
Figure 31. PCA analysis of hexanucleotide usage pattern (HUP) of metagenomes associated to different hosts. HUP of contigs bigger 10 kb was analyzed for the different metagenomes. Contigs annotated to the most abundant genus and those that were considered putative mobile genetic elements were coloured differently. A, Human diabetes type 2; B, Moose; C, Mouse feces; D, Sea urchin; E, Ebro14.

Table 4. Contigs with MGE detected using the methodology described

Genus	Contig (length kb)	MGE classification	MGE length (kb)	Virulence	Detected
Pseudomonas	Ebro13p22C1 (893.5)	Int (Cupriavidus)		chemotaxis Acriflavin resistance	GC
		Int		Penicilin	sgMGE
		Int		Bleomycin	sgMGE
		IS	0.6-0.9	ISPa2/4/5 exotoxin A	BLASTN
		IS		ExoU island B	BLASTN
	Ebro13p22C2 (555.3)			Virulence associated protein	GC
		Int (Pseudomonas)			sgMGE
Sphingobium	Ebro13p22C4 (520)	Int	6	Penicilin resistance	sgMGE
	Ebro13p22C11 (329)	Int (Pseudomonas)	47	toxin-antioxin	sgMGE
	Ebro13p22C9 (360.7)	prophage	22.5	Serine protease	sgMGE
	Ebrop1C93 (82)	ICE	≈29.4	arsenic resistance	sgMGE
	Ebro13p22C19 (208.8)	-	6.4	Cobalt/zinc/cadmium transporter	sgMGE
	Ebro13p22C22 (173.4)	Int	≥46	arsenic resistance	sgMGE
	Ebro13p22C25 (151.2)	ICE	≥70	siderophore	sgMGE
	Ebro13p22C29 (137.9)	-	≥23.5		sgMGE
	Ebro13p22C30 (134.4)	Int	≥55.8	Penicilin resistance	sgMGE
	Ebro13p22C38 (109.3)	Int	9.5	chitinase	sgMGE
	Ebro13p22C49 (79.6)	Int	≥17.2	Kanamycin resistance	sgMGE
Achromobacter	Ebro13p22C24 (160.3)	ICE	68.4	-	sgMGE
	Ebro13p22C37 (112.4)	Int	32.9	toxin-antioxin	sgMGE
	Ebro13p22C40 (112.4)	ICE	75.2	protease	sgMGE
	Ebro13p22C158 (32.2)	ICE	24	Multidrug resistace pump	sgMGE

	Ebro13p22C164 (31.8)	transposon	24.8	Arsenic & Mercuric resistance	sgMGE
	Ebro13p1C98 (78.9)	Int	34.6	toxin-antioxin	sgMGE
	Ebro13p1C111 (41.9)	ICE	32.4	-	sgMGE
	Ebro13p1C190 (41.9)	-	28	toxin-antioxin	sgMGE
Aeromonas	Ebro13p1C55 (129.8)	prophage	34.5	-	sgMGE
	Ebro14p22C8 (111.2)	-	2.8**	-	annotation
	Ebro14p22C14 (94.5)	transposon?	14	fimbrial operon	sgMGE
	Ebro14p22C32 (56)		9.2	-	annotation
	Ebro14p22C35 (52.8)	-	18	fimbrial operon	annotation
Vibrio	Ebro14p22C18 (76.9)	CRISPR	0.2	-	sgMGE
	Ebro14p22C144 (23.7)	-	-	-	sgMGE
	Ebro14p22C192 (18.6)	-	-	-	sgMGE
	Ebro14p22C278 (13.9)	-	-	-	sgMGE
-	Ebro13p1C57 (125.9)	ICE	125.9	Arsenic resistance	sgMGE
-	Ebro14p22C168 (20.4)	ICE	≥16	-	sgMGE
-	Ebro14p22C96 (31)	putative prophage	-	-	sgMGE
-	Ebro14p22C386 (10.6)	-	≥6	Arsenic resistance	sgMGE
-	Ebro14p22C174 (20)	-	-	RHS toxin	sgMGE
-	Ebro13p22C110	putative prophage	43.1	hemolysin	sgMGE

Islands = Isl
Integron= Int
GC =detected by GC
sgMGE= detected by BLASTP usign signature genes
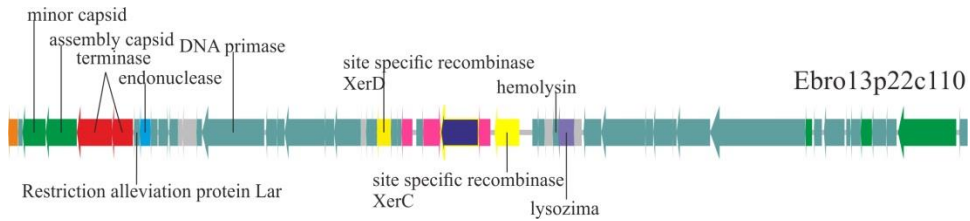
Figure 32. Phage identified using MGE detection protocol. The reannotation of a MGE detected by one of the methods described in the protocol classified it in a prophage. Genes were reannotated using HHpred.
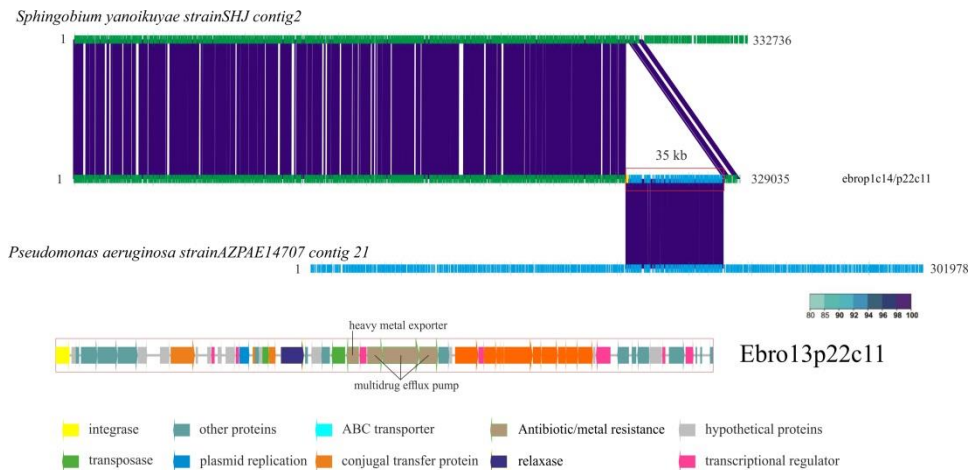
Figure 33. Interchange of information between different genera. The presence of an ICE was detected looking at the annotation and GC variation of a contig annotated to Sphingobium yanoikuyae. BLASTN was used to compare the contig against the genomes of the most similar strains in NCBI. The island was reannotated and classified as an ICE. Minimum identity of 80 was used to filter the BLASTN results.
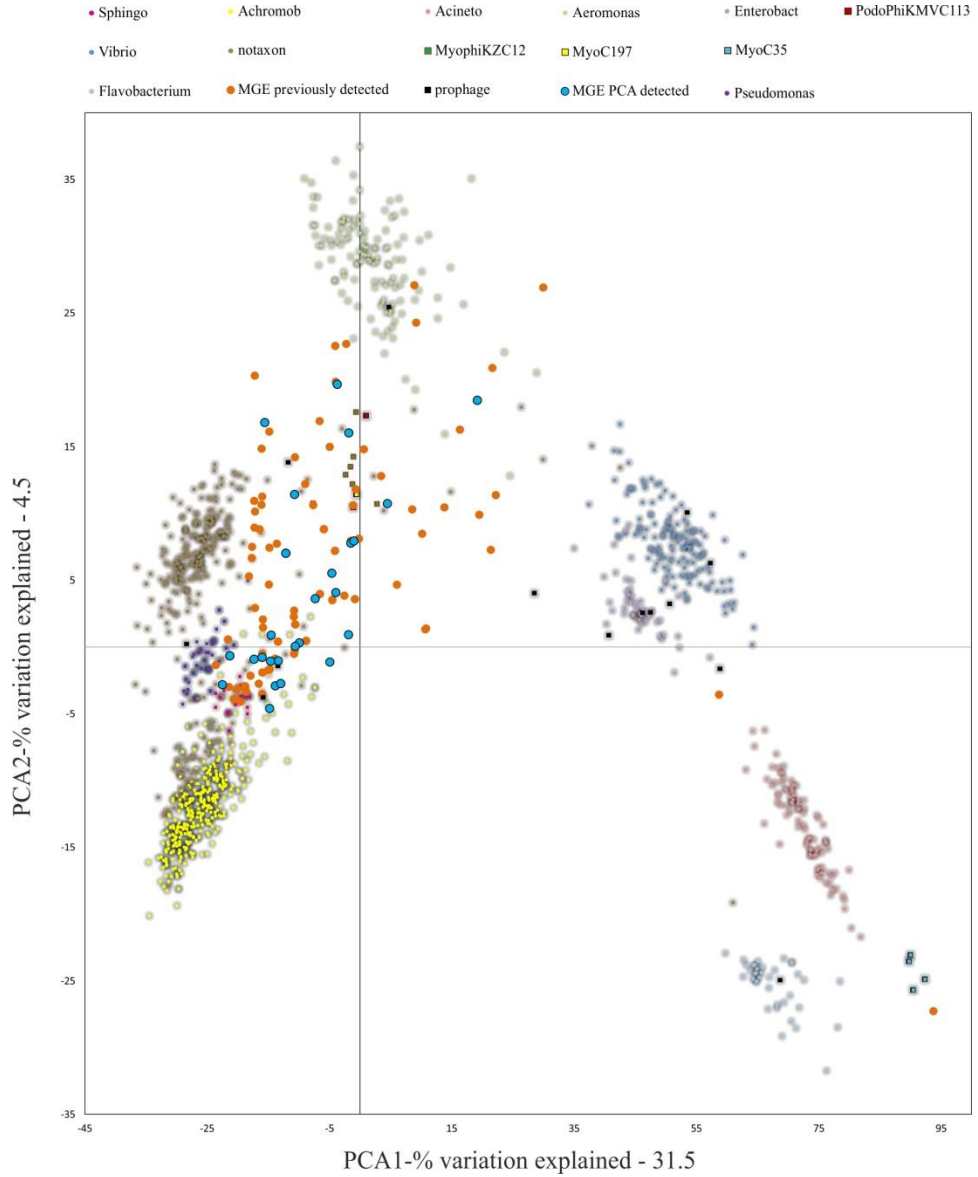
Figure 34. Hexanucleotide usage pattern (HUP) distribution of the attached microbiota to epidermal mucus of European eels in WE₃⁸. PCA was used to analyze the distribution of contigs bigger than 10 kb according to its HUP.

Only 30 of the MGE could be clearly classified: multiple insertion sequences (IS), 3; integrative and conjugative elements (ICE), 8; integrons, 12; CRISPR, 1; prophages, 4; and transposons, 2. These MGE were detected in the contigs assigned to the dominant genus *Vibrio, Pseudomonas, Sphingobium, Achromobacter, Enterobacterium* and *Aeromonas* **(Table 7, Figure 35 and 36)**.

The known predicted functions for the identified MGE were mostly related with colonization strategies: chemotaxis and fimbriae (attachment), RHS toxin and toxin/antitoxin systems (competence among bacteria)(Bliven and Maurelli) and siderophores (resistance to nutritional immunity). Resistance genes for penicillin, kanamycin, and chloramphenicol could be related more with resistance to bactericidal compounds liberated by competitors (fungi and bacteria) than to resistance to drugs used in clinical and liberated to the environment. For example, a chloramphenicol resistance gene was detected in a putative ICE from *Al. fischeri* contig in WE_{10}^8. To date, no reports have mentioned a resistance of *Al. fischeri* to this antibiotic. With regard to arsenic and mercuric, whose resistances were more abundant among SM- than W-microbiota, their presence in MGE would increase survival rate after horizontal transfer events.

We analysed in depth the genetic context of the two IS identified in *Pseudomonas*-contigs, since both were related with two of the toxins previously associated with one of the mucus-selective functions: resistance to innate immune system (phagocytosis); the toxins ToxA and ExoU. The genetic context of the gene *toxA* showed a series of IS upstream of the gene (**Figure 37**). Interestingly, it has been previously reported that the presence of various copies of IS upstream of *toxA* is an evidence of genetic rearrangement that allows phenotypic conversion from avirulent to virulent and vice versa in *Pseudomonas* aeruginosa (Sokol *et al*., 1994). More interestingly, the genetic context for *exoU* was compatible with a pathogenicity island (**Figure 37**). This result is quite interesting because it suggests that genes for toxins could be interchanged in the environment contributing to increase the resistance of the bacteria to the phagocytosis, a phenotype that could server later to survive inside hosts. We believe that these findings support, again, our original hypothesis about the

existence of environmental reservoir for virulence genes, some of which could be carried by MGE.
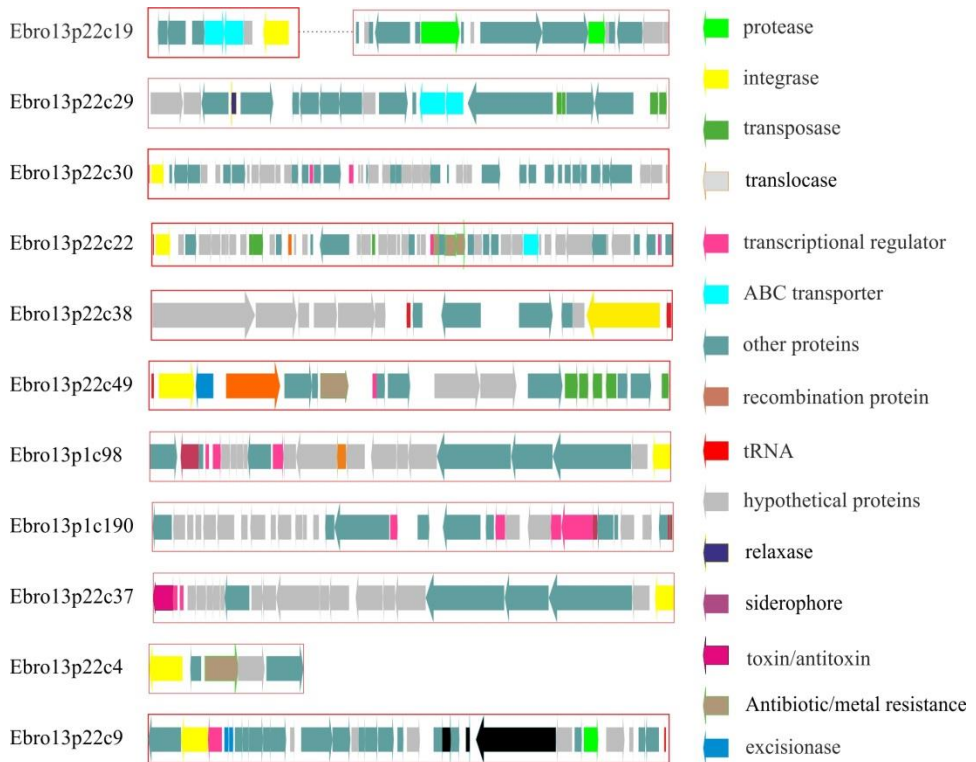
Figure 35. Representation of some contigs with MGE detected. Contigs bigger than 10 kb detected with PCA or using sgMGE were reannotated using Hhpred and the plotted. sgMGE and genes of interest that MGE had were coloured differently.
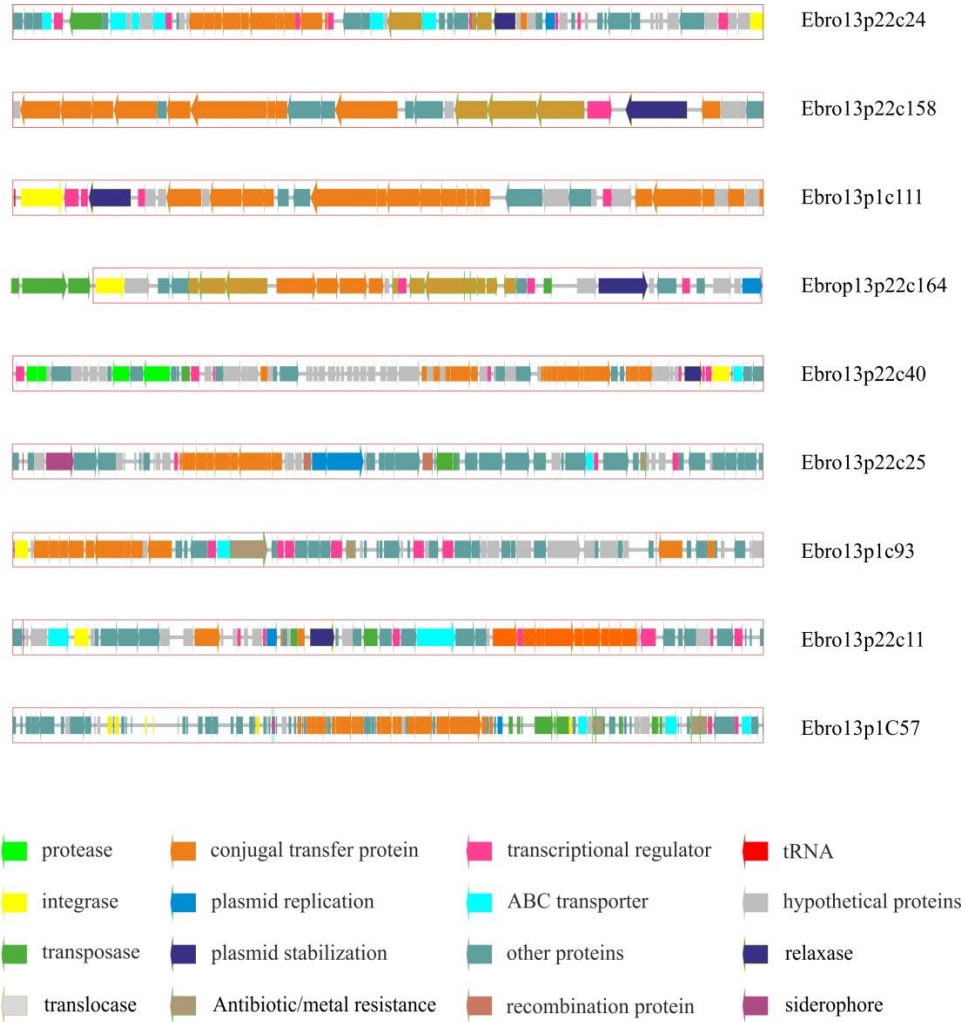
Figure 36. Integrative and conjugative elements (ICE) detected. Some of the MGE detected had sgMGE that allowed its classification as ICE. Interesting genes transported by MGE were highlighted.

As we have commented before, we were also able to find evidences of interchange of wide regions among bacteria. Thus, we found inserted sequences from different genus in *Sphingobium* and *Achromobacter* contigs. We focused our attention on one contig (**Figure 38**) which contains a putative ICE of significant size (47 kb) with a 100% identity and coverage to a contig of *Pseudomonas aeruginosa strain AZPAE14707* and *AZPAE14724*. These strains were isolated from the respiratory track of a patient in Greece and from an intra-

abdominal tract infection in Italy, respectively. Moreover, this ICE was highly similar to other *Pseudomonas* sequences from different ecosystems in the repository of NCBI (**Figure 38**). Interestingly, three genes encoding a multidrug efflux pump were inserted in this ICE. Other contigs that took our attention were $_{Ebro13p22}$C164 and $_{Ebro13p22}$C158, which were part of the same ICE in *Achromobacter* (**Figure 39**). Surprisingly, resistance to three different components (mercuric, arsenic and multidrug efflux pump) were codified in this MGE. The putative ICE (> 48 kb) separated in these contigs also hit *Pseudomonas* genomes in the databases including one that also matched the previous mentioned ($_{Ebro13p22}$C11). *Pseudomonas*, *Sphingobium* and *Achromobacter* have similar GC composition and their abundance in our metagenomes was also similar suggesting both are important factors for DNA interchanges in the environment.

The genes present on MGE are subject to continuous processes of evolution due to the selective forces acting in the environment. It seems that the mucous surface of eels is forcing MGE to transport genes related with microbial fitness in mucus such as those found in the present work: genes encoding resistance to innate defenses, toxics and natural antibiotics.
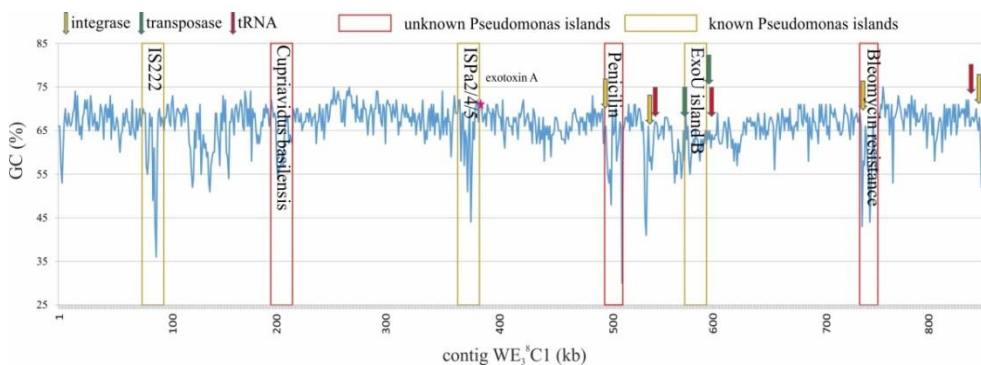
Figure 37.**MGE in a contig of *Pseudomonas***. The GC percentage of each ORF is represented in the graph for the entire contig. The integrases, transposases and tRNA are indicated using arrows and detected islands are highlighted using colored boxes.
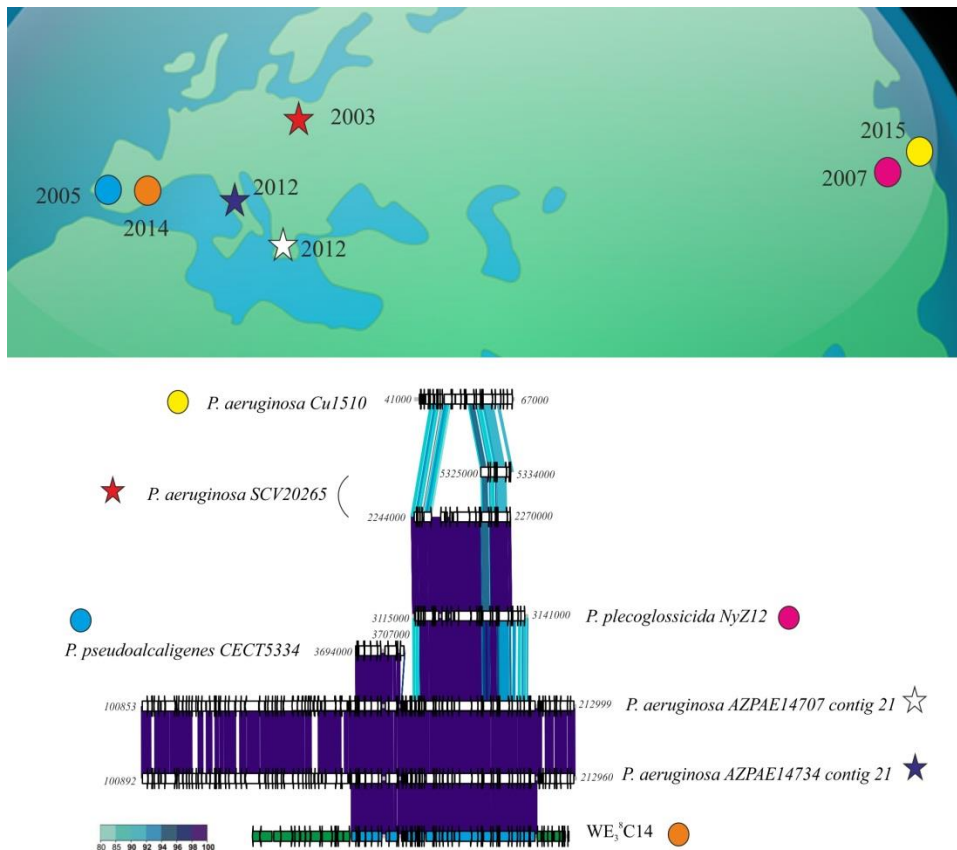
Figure 38. Distribution of a MGE detected between Pseudomonas strains. The ICE inserted in a Sphingobium yanoikuyae contig was compared to other Pseudomonas genomes deposited in NCBI using BLASTN. Minimum identity of 80 was used to filter the results. Localization and year of isolation of these strains is marked in the globe map. Strains isolated from patients were marked with a star and environmental isolates with a circle.

Figure 39. Long DNA sequences can be interchanged between genera with similar GC. A, MGE separated in two contigs assembled in our metagenome hit to Achromobacter and Pseudomonas genomes in databases after BLASTN. B, contigs were reannotated using HHpred and plotted. Interesting genes and parts annotated to different genera or origin were highlighted. Minimum identity of 80 was considered for filtering results.

5.3. Replicating phages in the superficial microbiota of eels

As we explained in the previous chapter, we found MGEs in all the analyzed metagenomes. Some of the detected MGEs were prophages, so we decided to find out if lytic or, even, replicating phages were also present in SM- and W-metagenomes. In addition, we were interested in finding out if the hypothesis of Barr *et al* (2013) about the role of bacteriophages as a substantial part of natural mucosal defenses against pathogens in humans could also be extended to mucus that covers aquatic animals such as eels.

5.3.1. Bacteriophage detection in metagenome datasets

As we have already mentioned the assembly of the Illumina datasets yielded long contigs. Annotation and analysis allowed the detection of 17 viral contigs >10 kb and 23 contigs <10kb, which were clearly related (>98 % nucleotide identity) to the longer contigs (**Figure 16**). We selected 5 of these contigs for further analysis. Long fragments of Caudovirales genomes have been detected in cellular metagenomes, probably due to the concatamer formation during lytic cycle, a natural process of genome amplification (Mizuno *et al*., 2013). In order to check if any of the assembled phages represented complete genomes, we performed all *vs* all comparisons of these viral contigs. We detected contigs overlapping in a circular fashion in the metagenome datasets from the 0.22 and 1 µm filters of the $WE_3{}^8$ (see **Figure 40**). This result suggests that all the genes in the genome had been captured and that these contigs represent complete phage genomes. Another method that may be useful in detecting the completeness of viral genomes is the presence of terminal repeats within the same contig. We identified a total of 4 complete phage genomes by these methods (See **Table 5** for details). Their analysis revealed that all correspond to tailed-bacteriophages.
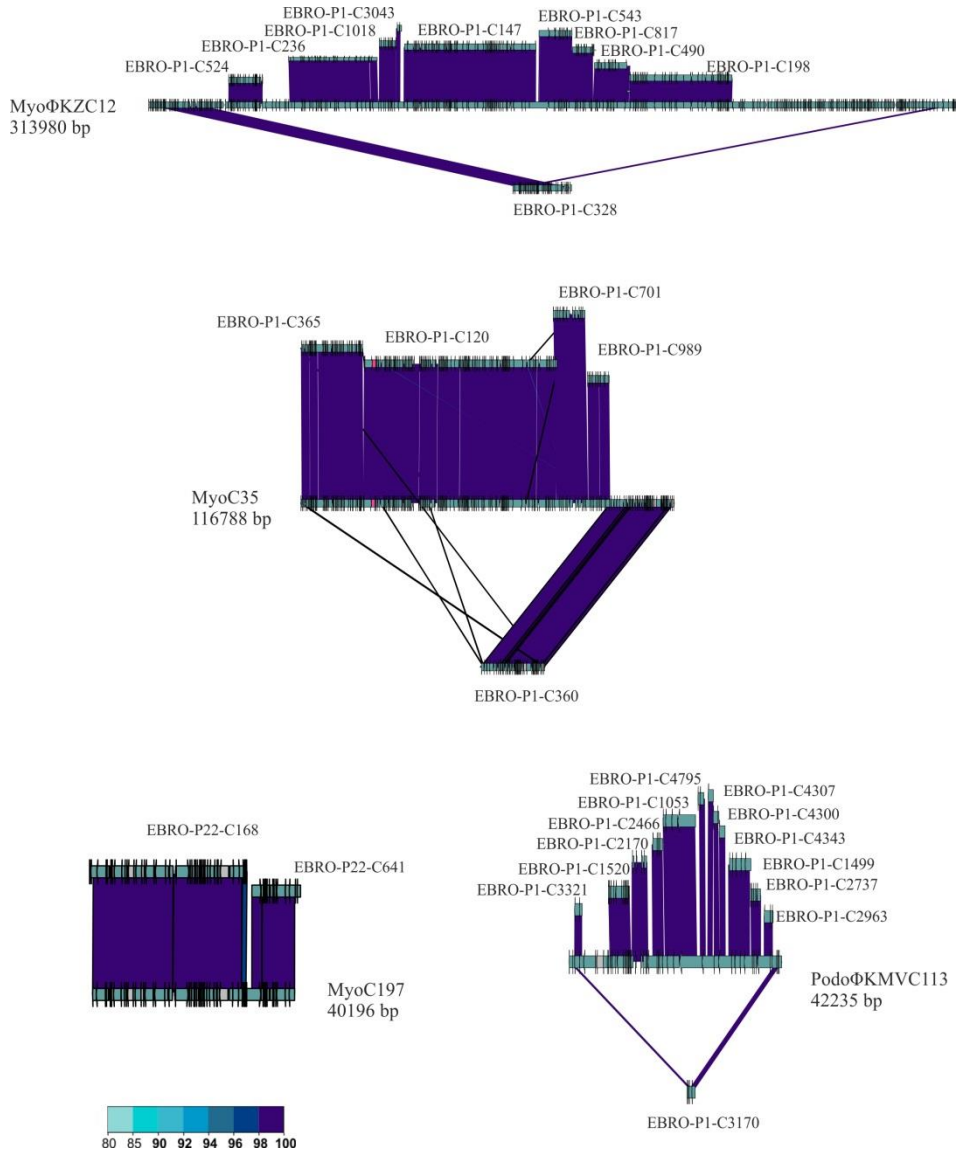
Figure 40. Viral contigs alignment. All viral contigs larger than 10 Kb were aligned using TBLASTX. The identity of the alignment is represented by a color code.

Additionally, nine putative prophages were also identified by the presence of host genes at one or both ends of the contig being their hosts *Stenotrophomonas*, *Achromobacter* and *Enterobacteriaceae bacterium 9_2_54FAA* (**Table 6 and Figure 41**). Since we had multiple metagenomes, we checked for redundant contigs and found several examples (**Figure 40**). Even samples taken one year apart yielded identical phage contigs indicating a remarkable resilience and

conservation (**Figure 42**). Both, phages and prophages were detected also by the MGE detection protocol.

Table 5 Complete sequenced phages properties.

	Contig	Length (pb)	Genus	%GC	ORF	%annotated ORF
Cabanes	MyoΦKZC1	220.117	ΦKZ	45,19	227	32,15
Ebro Delta	MyoΦKZC12	313.980	ΦKZ	58,22	281	79,60
	MyoC35	116.788	FelixO1like	34,72	92	43,81
	PodoΦKMVC113	42.235	phiKMV-like (Podoviridae)	59,28	28	54,90
	MyoC197	40.198*	unclassified Myoviridae	60,95	58	57

*incomplete genome

Table 6. Genomic properties of prophages sequenced

Contig	Prophage length (bp)	GC%	Host
ProStenoC191	41068	68.3	*Stenotrophomonas*
ProAchroC45	42197	63.4	*Achromobacter*
ProAchroC186	33297	63.4	
ProEnteroC6	30621	48.1	
ProEnteroC38	20823	46.79	
ProEnteroC171	34456	47.3	*Enterobacteriaceae bacterium*
ProEnteroC260	23048	49.6	
ProEnteroC286	16618	44.2	
ProEnteroC299	25725	53.1	

Figure 41. Sequenced prophages present in WE₃⁸.
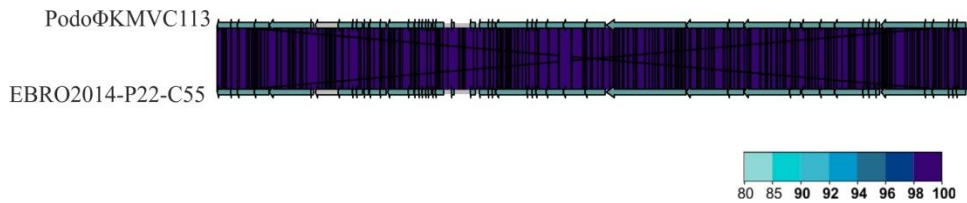
Figure 42. PodoΦKMVC113 in Ebro Delta samples(WE₃⁸ and WE₁₀⁸).

5.3.2.Myoviruses

Among tailed bacteriophages, Myoviruses are known because of their large sizes ranging from 11.6 to 358.6 kb. We found three Myoviruses that we named MyoФKZC1, MyoФKZC12 and MyoC35, with genome sizes of 220, 313 and 116 kb, respectively (**Table 5, Figure 43**, **44 and 45**). MyoФKZC12 and MyoФKZC1 shared a protein with a putative CAAX protease domain and MyoC35 presented a putative ubiquitin-ligase, and the alpha and beta subunits of a proteasome complex. By complete genome analysis, MyoФKZC12 and MyoФKZC1 were predicted to be part of the ФKZ genus (**Figure 46**). This genus is notorious for their genes involved in nucleotide metabolism (e.g., thymidylate synthase, thymidylate kinase, ribonucleoside diphosphate reductase subunit beta [*NrdB*] and alpha [*NrdA*], and dihydrofolate reductase, *RuvC* Holliday junction resolvasome) (Jang *et al*., 2013; Cornelissen *et al*., 2012) all of which were found in MyoФKZC12 and MyoФKZC1 genomes. Phylogenetic analysis using DNApol (**Figure 47**), terminase large subunit (**Figure 48**) and ribonucleoside diphosphate reductase alpha chain (*nrdA*) (**Figure 49**) genes also supported the taxonomic classification. The GC content of MyoФKZC1 was in the expected range and MyoФKZC12 reached the highest value ever published for this genus (58.22%) (**Table 7**). Genomic comparison between *Pseudomonas* phage ФKZ, PA7, and MyoФKZC12 and MyoФKZC1 revealed a low structural conservation together with a low nucleotide identity (**Figure 43B**). In fact, only some structural protein, DNApol, RNApol and terminase showed homology (30-60 % aminoacid identity). Previous studies have shown a high rate of divergence among members of this genus (Jang *et al*., 2013; Cornelissen *et al*., 2012). Representatives of this genus target a variety of Gram-negative bacteria such as *Pseudomonas*, *Vibrio*, *Yersinia*, *Cronobacter*, *Salmonella* and *Erwinia* (**Table 7**). We could not assign a putative host for these Myoviruses. Recent studies have shown that the eight RNA polymerase subunits from *Pseudomonas* phage ФKZ form 2 polymerases: virion (vRNAP) and non virion RNA polymerase (nvRNAP) (Ceyssens *et al*., 2014). Moreover, Ceyssens *et al.* (2014)suggested that all subunits are present in all phages of the genus and the transcription is

completely independent from the host. We have been able to identify some of them in MyoΦKZC12 and MyoΦKZC1 genomes. Further phylogenetic analysis using all RNApol subunits found in each ΦKZ genome showed a clustering of the eight different subunits in the tree (**Figure 50**).
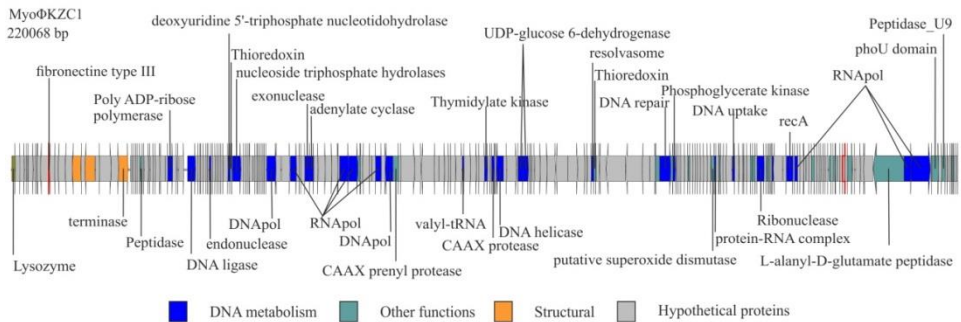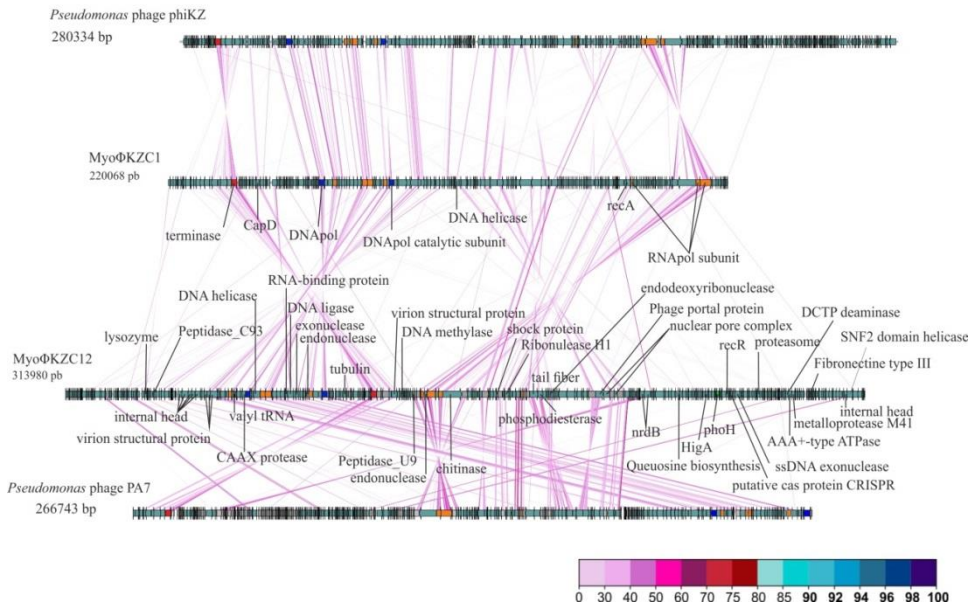
Figure 43. Genome representation of MyoΦKZC1 and genome comparison of ΦKZ. A, The complete annotation of MyoΦKZC1 is represented by arrows and different colours were used to differentiate functions. Ig-domain is highlighted using a red fringe. B, TBLASTX comparisons were done between MyoΦKZC12, MyoΦKZC1, *Pseudomonas* phage phiKZ and PA7
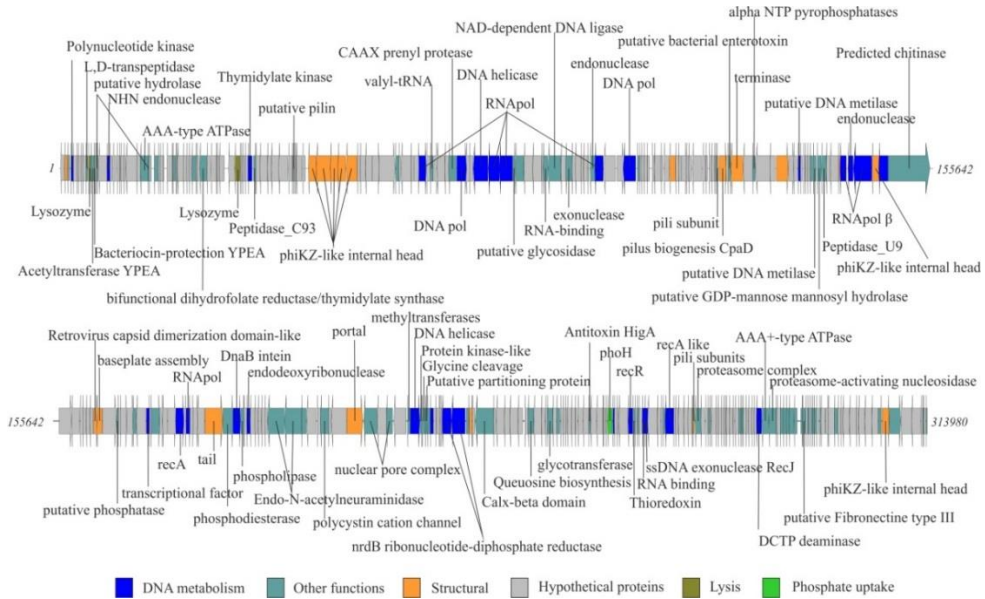
Figure 44. МуоΦKZC12, complete new phage genome. Complete genome of the longest phage is represented in two rows. Gene functions ORFs are colored according to the color code.
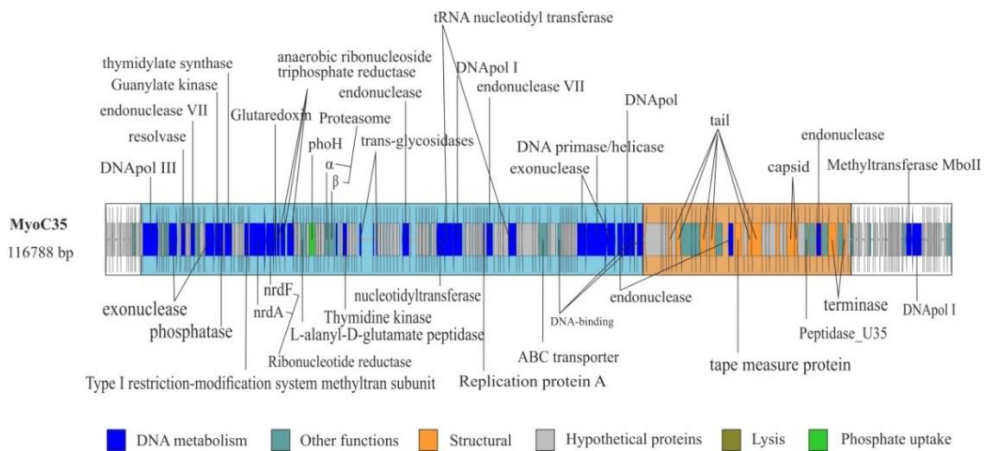
Figure 45. Complete genome of MyoC35. Functional domains were highlighted using boxes colored according the function of the genes. White box indicate no functional domain.
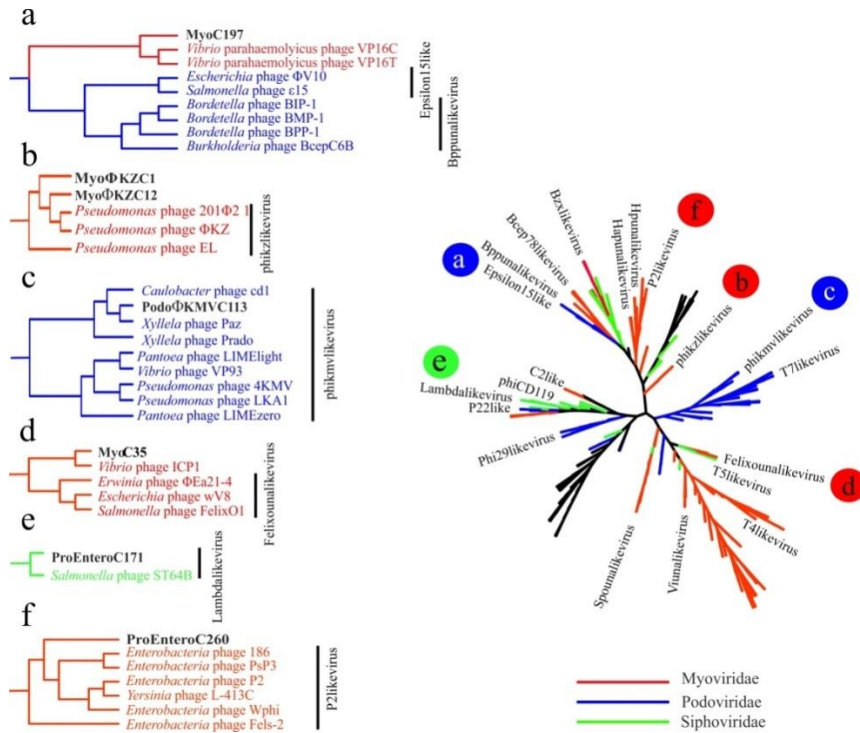
Figure 46. Complete genomes phylogeny. All-vs-all comparison using NCBI phages genomes and the ones sequenced in our metagenomes. The clustering was based on a sequence similarity derived metric. Colors were used to differentiate viral families. Branches in black indicate unclassified genomes. The branches in which the new complete genomes clustered are shown in detail on the left and marked with letters that correspond with the position in the tree.

Table 7. Genomic properties of ΦKZ members.

PhiKZlike	Length (bp)	%GC	ORF	tRNA	RNApol
Pseudomonas phage EL	211215	49.3	201	1	6
Pseudomonas phage phiKZ	280334	36.8	369	6	8
Pseudomonas phage 201phi2-1	316674	45.3	461	1	7
Pseudomonas phage PhiPA3	309208	47.43	375	3	5
Pseudomonas phage OBP	283757	43.5	309	4	8
Halocynthia phage JM-2012	167292	35.4	163	0	8
Erwinia phage phiEaH2	243050	51.28	262	11	8
Yersinia phage phiR1-37	262391	32.4	367	4	7
Cronobacter phage CR5	223989	50.1	231	0	7
Pseudomonas phage PA7	266743	36.96	337	7	8
Salmonella phage SPN3US	240413	48.54	264	2	6

MyoC35 (**Figure 45**, 116 kb) was classified as a FelixO1like phage by both whole genome comparison against NCBI (**Figure 46**) and clustering according to the terminase gene phylogenetic tree (**Figure 48**). In both analyses, MyoC35 clustered with *Cronobacter* phage vB_CsaM_GAP31 and *Vibrio* phage ICP1 although showed a low structural conservation and only DNApol, NrdA and some structural proteins were similar (<50% identity) (**Figure 51**). Those phages have not been assigned to any genus within the Myoviridae family. *Vibrio* phage ICP1 has been linked to FelixO1like genus by homology analysis using NrdA (Seed *et al*., 2011). Phylogenetic analysis using this gene clustered MyoC35 close to *Vibrio* phage ICP1 (**Figure 49**). However, the same analysis using DNApol clustered MyoC35 close to the Φ16 branch, a group recently added to NCBI which also appeared in this niche (see below) (**Figure 47**). *Vibrio* phage ICP1 is a very specific *Vibrio* phage which has been isolated from stool samples from a cholera patient (Seed *et al*., 2011). In WE$_3$[8]metagenome, only small contigs were annotated as *V. cholerae*, however *Vibrio sp. RC341* (closely related to *V.cholerae* and *V.mimicus* (Haley *et al*., 2010b)) was one of the most abundant organisms by number of contigs. They could be representatives of MyoC35 host.
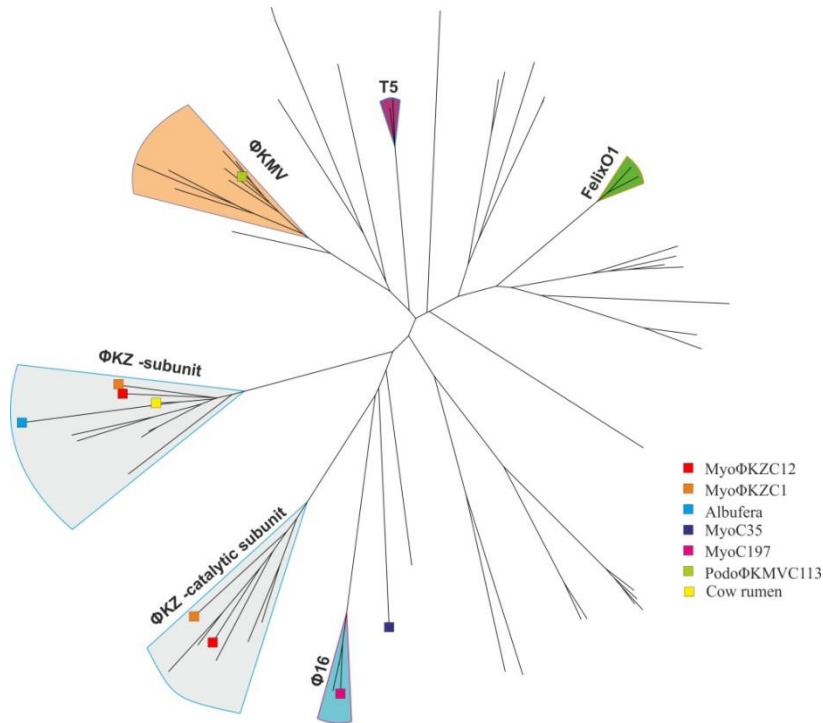
Figure 47. DNA polymerase phylogeny. Maximum-likelihood tree build using DNApol genes identified in the sequenced genomes, Albufera and Cow rumen metagenomes are from Genbank database. Clearly clustered genus and new identified DNApol are labelled. The two ΦKZ subunits are labelled using the same color.

Interestingly MyoΦKZC12 and MyoC35 genomes harbor a *phoH* gene, an auxiliary metabolic gene implicated in the regulation of phosphate uptake and metabolism under low-phosphate conditions (Hsieh and Wanner, 2010). The inclusion of this gene in the phylogenetic analysis supports the value of *phoH* as a signature gene for marine viruses (**Figure 52**)(Goldsmith *et al*., 2011).

MyoC197 was predicted to be an incomplete myovirus with mixed characteristics (**Figure 53**, 40.2 kb). On one hand, this phage clustered with two *Vibrio* phages (VP16C and VP16T) by several phylogenetic analyses (**Figure 46, 47 and 48**) although lacking the polypeptide deformylase gene and *vapE* (virulence associated protein), two typical genes of these *Vibrio* phages (Seguritan *et al*., 2003). On the other hand, MyoC197 contained several lambda phage related genes (genes for lambda head decoration protein D, major capsid protein E and

terminase large subunit GpA) and some Mu-like viruses related genes (genes for Mu-like prophage tail sheath protein, phage tail tube protein, FluMu protein Gp41, Mu-like prophage DNA circulation protein and an uncharacterized protein conserved in bacteria DUF2313). In fact, 25 of 58 genes in the contig had the highest similarity to *Vibrio* phage genes (50-70 % identity) (**Figure 54**). The *Vibrio* phages came from an environmental isolate of *V. parahaemolyticus* and are now classified as a new genus named Φ16 (Seguritan *et al*., 2003). Interestingly, MyoC197 was one of the most abundant viruses recruited from WE_3^8(**Figure 55E**). Contigs annotated as *Vibrio* in metagenome from eels fished in estuarine waters (WE_3^8) were assigned to *Vibrio sp. RC341*, *V. anguillarum*, *V. vulnificus* and *Al. fischeri* suggesting those species as putative hosts for MyoC197.
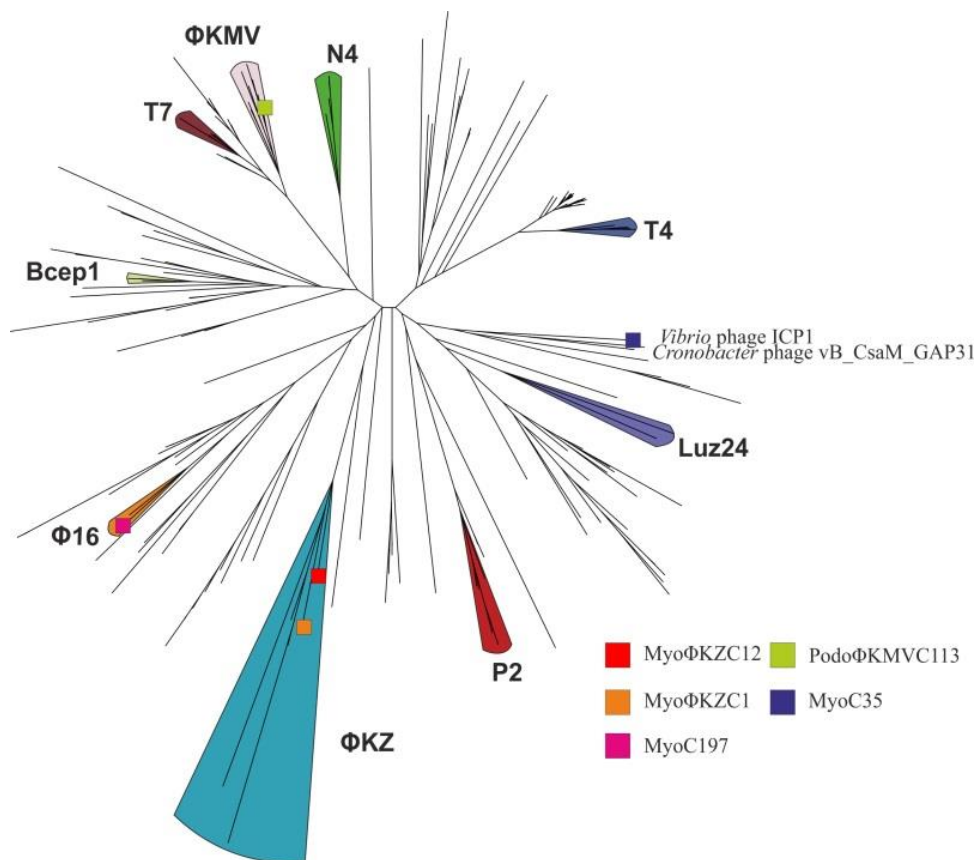
Figure 48.Terminase phylogeny. Large terminase subunits recovered from all genomes and others selected from Genbank were used to construct a maximum-likelihood tree. Branches were labelled when genus were perfectly clustered in it. The branches corresponding to our data are coloured.

Figure 49. Ribonucleoside diphosphate reductase alpha chain phylogeny. NrdA from the bacteriophages and the putative hosts were used for the analysis. Branches clustered in one genus were grouped using the genus name.
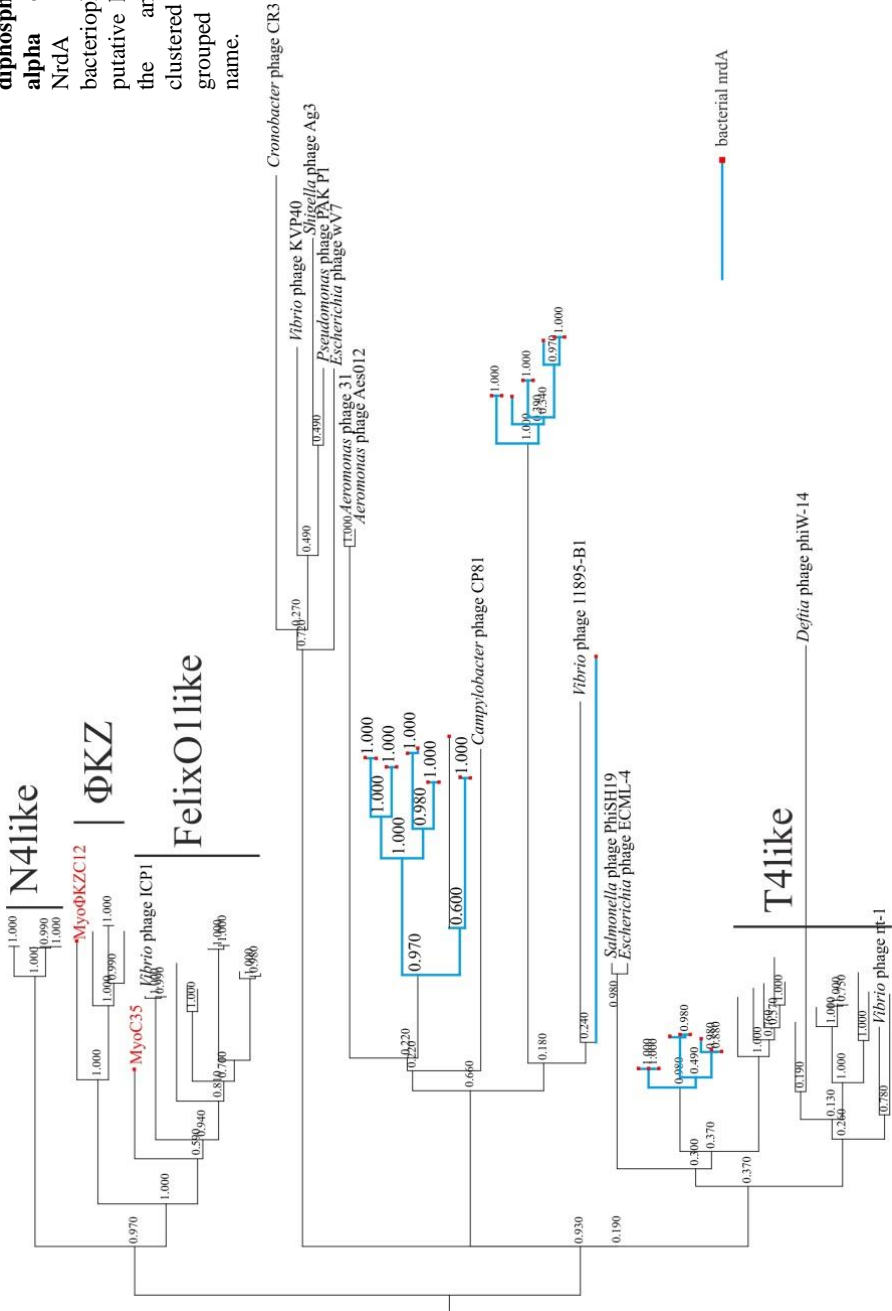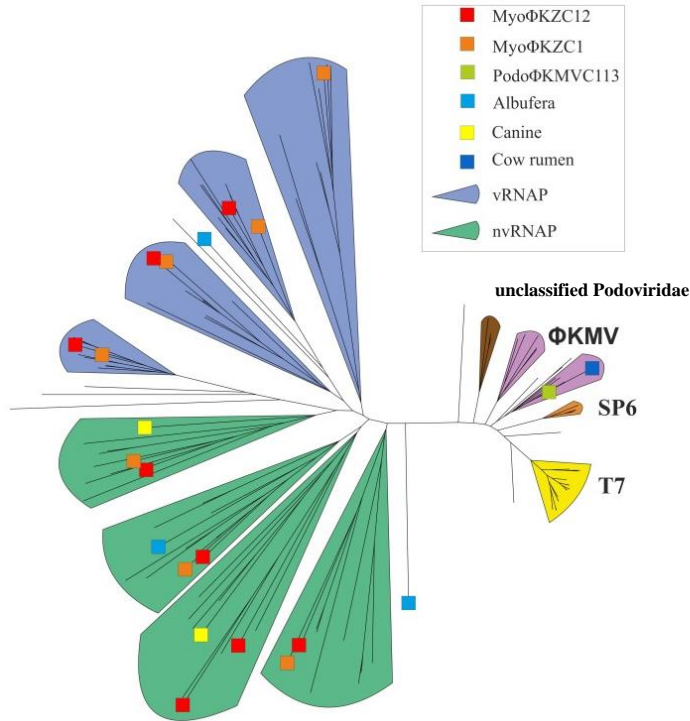
Figure 50. RNA polymerase phylogeny. Maximum-likelihood tree using RNApol from Genbank database, MyoΦKZC12, MyoΦKZC1and PodoΦKMVC113 genomes and from Albufera, canine feces and cow rumen metagenome. Branches in which genera have been clustered separately were coloured differently and RNApol from ΦKZ members were coloured accordingly to the subunit. RNApol genes from metagenome of eels were marked with coloured boxes.
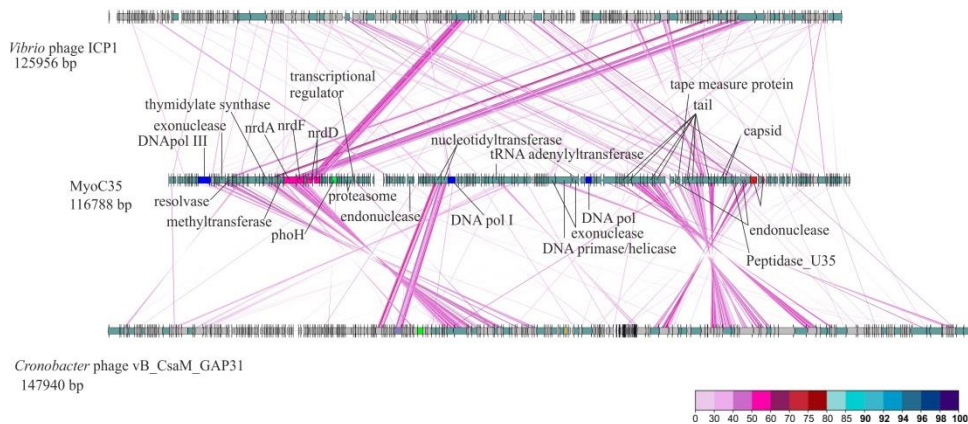
Figure 51. Genome comparison of MyoC35 and the two most similar complete phages from databases. MyoC35 was compared against *Vibrio* phage ICP1 and *Cronobacter* phage vB_CsaM_GAP31 using TBLASTX.
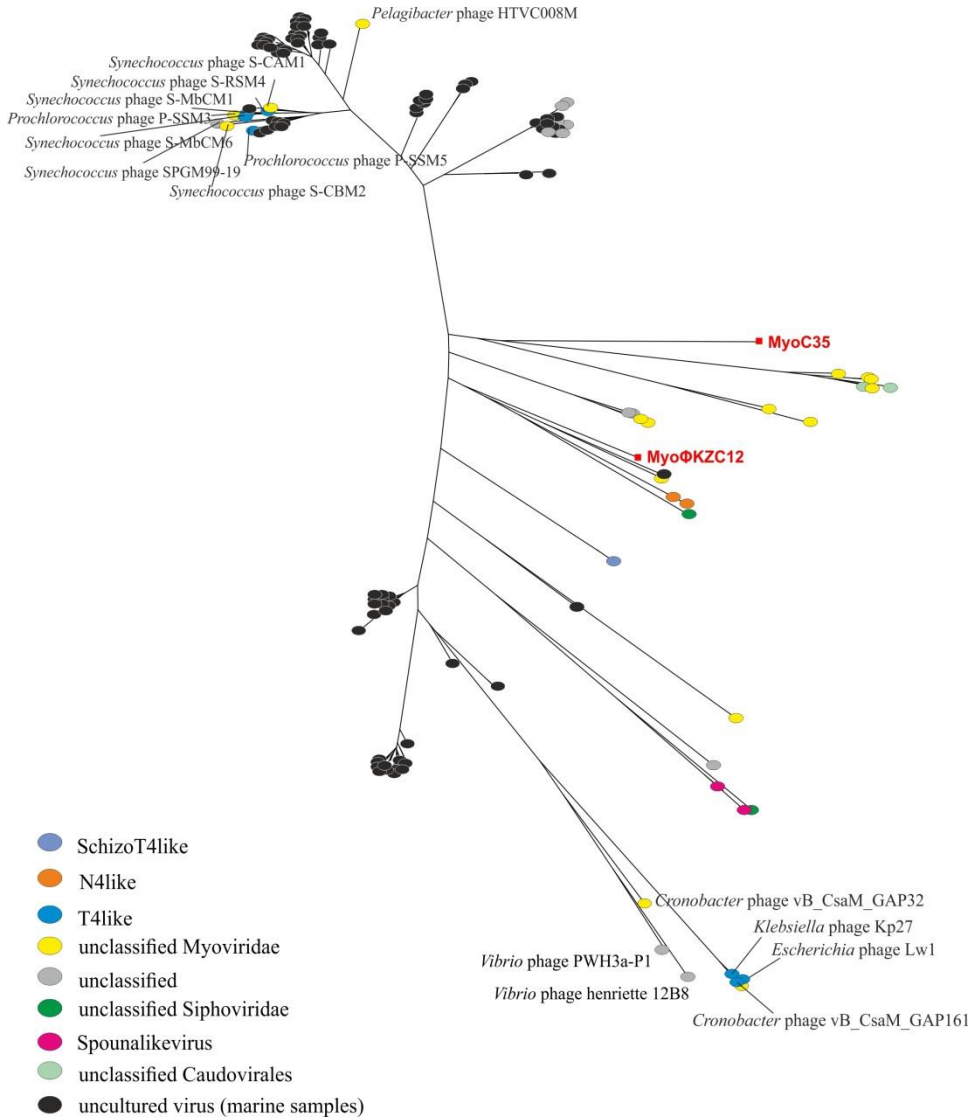
Figure 52. Phylogenetic tree showing the clustering of *phoH* genes of virus isolated from marine samples and from cultured phages. *phoH* genes identified in our genomes are label in red. The taxonomy of the rest is labeled with different colors. The names of phage genomes which clustered with virus from marine samples were added.

5.3.3. Podoviruses

According to the phylogenetic analysis from the complete genome, PodoΦKMVC113 (**Figure 53**, 42.2 kb) is a Podovirus from the *Autographivirinae* subfamily and a putative new member of the ΦKMVlike genus of T7 related phages infecting *P. aeruginosa* (40.7 – 44.5 kb) (**Figure 46**). RNApol, a hallmark gene for this genus, was found in PodoΦKMVC113, while

single-strand interruptions (nicks) also typical in this genus were not present (Kulakov *et al*., 2009). Annotation results showed several genes assigned to other phages such as *Caulobacter* phage Cd1 and *Xanthomonas* phages. Phylogenetic analysis using the terminase and RNApol genes clustered PodoΦKMVC113 with *Caulobacter* phage Cd1, *Xylella* phages Paz and Prado, all of them members of ΦKMVlike genus (**Figure 48 and 50**). As the other ΦKMV phages, PodoΦKMVC113 follows the structural organization based on three functional gene clusters encoded on the forward strand but with a rearrangement of lysis and structural domains (data not shown). A proteasome alpha subunit was also identified in this genome.
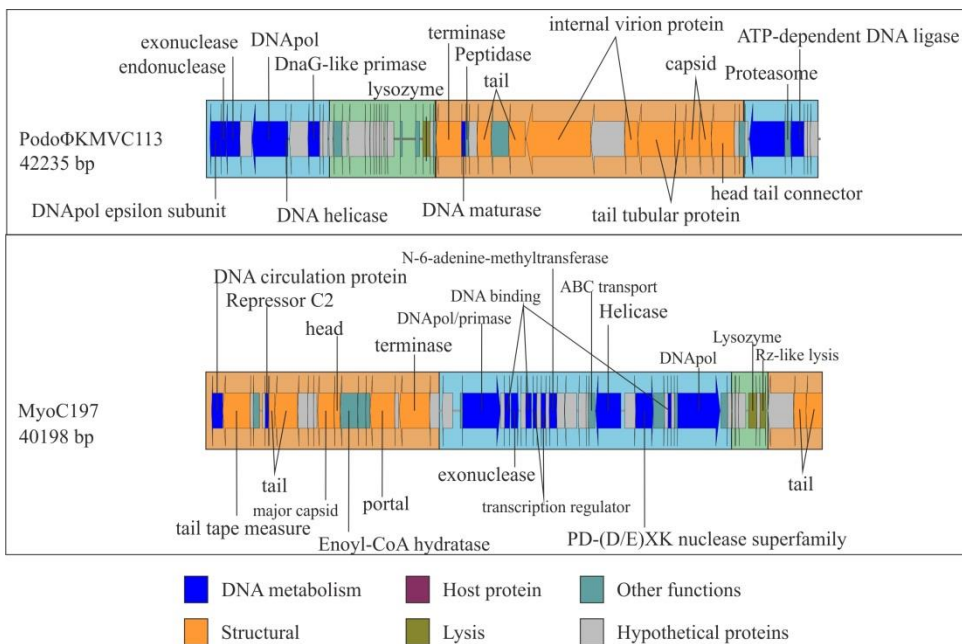
Figure 53. Genome representation of PodoΦKMVC113 and MyoC197. Domains were boxed according to the function of the genes present in it: orange, blue and green indicating structural, DNA metabolism and lysis domains, respectively. Ig-like domains are highlighted using a red fringe.

The segregation of the branch of PodoΦKMVC113, *Caulobacter* phage Cd1 and *Xylella* phages Prado and Paz suggests the formation of a new group very close to the ΦKMV genus (**Figure 46**). Interestingly, PodoΦKMVC113 was completely assembled in the same sampling point a year later, WE_{10}^8 (**Figure 40**)

(identity 100%). Although no contig similar to PodoФKMVC113 was assembled in the metagenome from the water, several reads were recruited against this genome (**Figure 55E**). It is probable that the complete genome would be assembled with larger sequencing efforts.
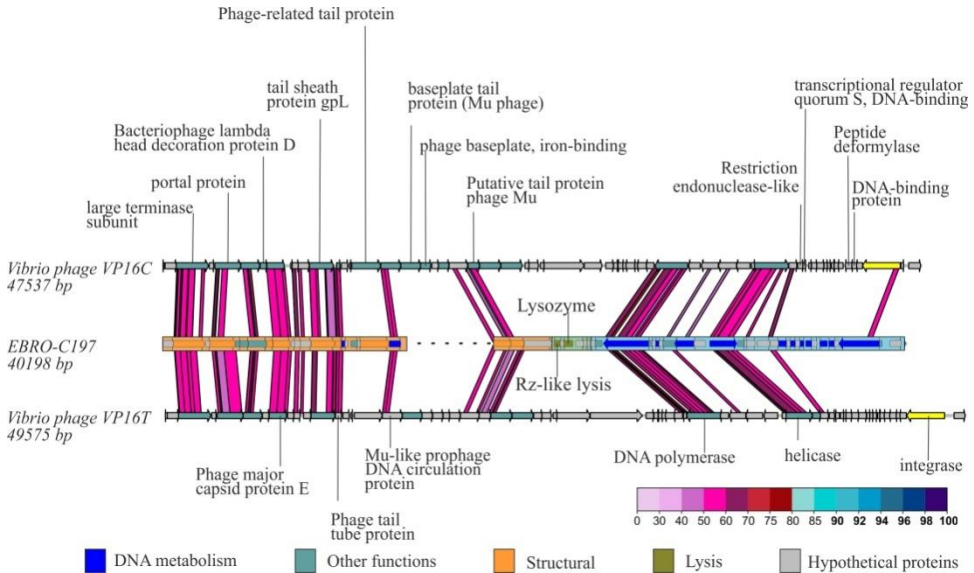
Figure 54. *Vibrio* phages comparison. MyoC197 was compared against *Vibrio* phage VP16T and VP16C using TBLASTX. MyoC197 was oriented in the same sense as the *Vibrio* phages used for comparison. The annotation of MyoC197 was coloured according to function.

5.3.4. Phage abundance in mucus and water and presence of similar phages in other habitats

In order to assess the relative abundance of these phages in the different metagenomes, the contigs were compared with the raw reads. In all cases, the viral contigs recruited more than bacterial contigs except for *Pseudomonas* and *Sphingobium* in the samples from Ebro Delta (WE_3^8 and WE_{10}^8) (**Figure 55A and B**). In WE_7^8, where the bacterial population was calculated to be composed by 60-80% of *Vibrio*, the number of reads recruited to MyoФKZC1 genome was higher than those recruited to bacterial representatives (**Figure 55C**). Moreover, when these proportions were compared with the water metagenome, the relative

abundance of phages and bacteria were reversed, suggesting that bacteriophages were retained or continuously produced in the skin-mucus (**Figure 55D**).

To test this apparently higher abundance of phages in mucus samples, we directly counted phages from mucus of farmed-eels and the surrounding water by epifluorescence. As expected, the abundance of phages in mucus samples ($3.14 \cdot 10^6$ virus-like particles/ml) was higher than in the surrounding water ($1.47 \cdot 10^5$ virus-like particles/ml) confirming that mucus concentrates phages present in the water. We also counted bacteria and compared the ratio virus/bacteria in the mucus and water samples. Bacteria were more abundant in mucus ($1.62 \cdot 10^6$ *vs* $1.49 \cdot 10^4$ bacteria-like particles per ml of mucus or water, respectively), suggesting that bacterial population was even more concentrated in the eel secretion. This led to a lower phage to bacteria ratio in the mucus (ca. 2:1) than in the water (ca. 10:1). Barr *et al.* (2013) found 4.4 higher ratios of phages to bacteria in the mucus than in the surrounding environment (on average) in a study that included one teleost surface mucus. It is possible that our method to collect mucus by passive release from the fish (rather than by suction device as done in the mentioned paper) might have affected the results.

The relative abundance of phages in their respective samples was analyzed by recruitment as well. On one hand, MyoC197 was the genome that recruited the most while PodoΦKMVC113, -C35 and -C12 reached similar but lower levels from $WE_3{}^8$ (**Figure 55E**). On the other hand, a year later, PodoΦKMVC113 was the only one that increased its recruitment level while the Myoviruses almost disappeared. Considering that the bacterial population changed in this last sample, probably as a reflection of a change in water salinity (3 to 10 g/L) (**Figure 55A and B**), the persistence of PodoΦKMVC113 and MyoC197 in the skin-mucus of eels in Ebro Delta for a year suggests that the host of these phages had remained in the skin-mucus while the host of the others Myoviruses did not. The bacterial genus that maintained and even raised its proportion in the skin-mucus through time was *Vibrio*. Therefore, this result suggests that *Vibrio* could be the host of PodoΦKMVC113 and MyoC197. Accordingly, the great reduction of the others Myoviruses mentioned could be related to a reduction in the host

population. The decreasing numbers of reads recruited by *Pseudomonas, Achromobacter* and *Sphingobium* make us consider them as putative hosts for these phages (**Figure 55A and B**). Finally, the only phage which recruited a significant number of reads in the metagenome from the water sample was PodoΦKMVC113.
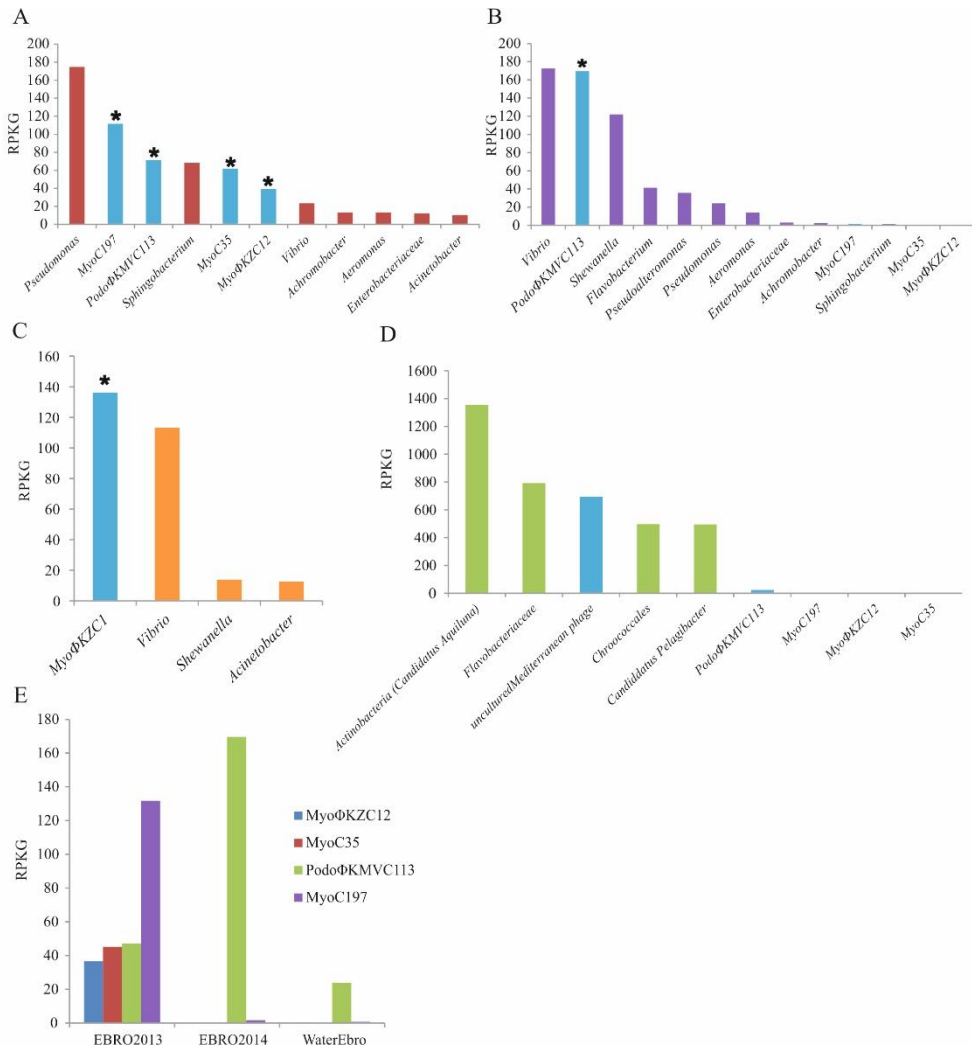
Figure 55. Bacterial and viral abundance in the metagenomes. The abundance of bacterial genera and viral genomes were calculated by normalizing the number of reads recruited in their respective metagenomes per the size of the concatenated contigs (Kb) and the gigabases of the dataset (RPKG). Bacteriophage counts were highlighted in blue. Complete genomes were marked with an asterisk. A, WE_3^8; B, WE_{10}^8; C, WE_7^8; D, WEW_{10}^8; E, Abundance of each phage genome in their respective metagenome; EBRO2013 and EBRO2014, metagenomes WE_3^8 and WE_{10}^8 respectively; WaterEbro metagenome, WEW_{10}^8.

Finally, metagenomes and metaviromes from marine and animal associated habitats were downloaded from MG-RAST in order to look for these phages in other niches. None of these datasets gave a number of matches high enough to consider that a similar virus was present in the searched metagenomes except for canine feces and cow rumen metagenomes in which reads similar to the ΦKZ and ΦKMV representatives were found. When reads identified as DNApol or RNApol were included to probe the presence of these genera in the rest of metagenomes, a single DNApol found in the cow rumen metagenome was similar to MyoΦKZC12 and MyoΦKZC1 and it clustered in the expected branch **(Figure 47)**. Finally, two RNApol genes were recovered similar to MyoΦKZC12 from a canine metagenome and one to PodoΦKMVC113 from a cow metagenome that clustered in the respective branch **(Figure 50)**. Moreover, we searched in the VIROME database for annotated open reading frames (ORFs) of the different viral genus found in skin-mucus. Members of ΦKMV, FelixO1likevirus and Φ16 were found in practically all marine, soil and host-associated viromes while ΦKZ genus was present in very low amounts in some marine viromes but was highly represented in host-associated ones, especially in the cow rumen virome. ΦKZ members have been isolated from bacterial pure cultures from a great diversity of environments: sewage, pond water, compost, soil, chicken feces, fresh- and marine-water, but this is the first time that they have been sequenced from an environmental habitat. It is noteworthy that 40 contigs from Albufera metagenome were annotated as ΦKZ (30-60% identity). Furthermore, DNApol and RNApol subunits were found in some of these contigs and included in the analysis confirming the annotation **(Figure 47 and 48)**. This turned ΦKZ in the only genus found in all three metagenomes from mucus of wild-eels.

6. Discussion

Vibrio vulnificus is an aquatic pathogen able to cause death by sepsis in humans and some fishes. All the strains of the species are considered potentially dangerous for humans while only those harboring the virulence plasmid pVvBt2, designated as biotype 2, can also be of concern in animal health. Farmed-fish are the most susceptible hosts for *V. vulnificus* biotype 2 infections while it is not known if this biotype causes vibriosis in nature. Our hypothesis is that wild-eels and, more precisely, their external mucous surface constitute an important reservoir for this species. According to this hypothesis, free-living vibrios should be attracted by the superficial mucus and colonize the eel surface where HGT event should be frequents. As vibrios colonize the mucous surface, they should not only be targeted by the mucosal immune system, but also should interact with commensal microbial species already present in the skin. But, the microbial composition of eel skin mucus is unknown.

Taken all these into account, the objective of this chapter was to describe the skin-mucus microbiome of wild-eels from different natural ecosystems and compare it with that of the surrounding water and farmed eels by using a metagenomic approach. Once described, we have focused our attention in disentangling not only the bacterial species composition and its functional capabilities but also in highlighting the mucus-mobilome (MGEs, including prophages) and –virome (phages). We expected to find *V. vulnificus* among the resident mucus microbiome and also evidences of HGT events affecting bacterial species in the microbiome.

We found that the skin-mucus microbiota from wild-eels was dominated by Gammaproteobacteria, and its composition in genera and species was variable, apparently, depending on water salinity where eels were captured. Thus, *Vibrio* monopolized the skin-mucus in the eels from estuarine waters (≥ 7 g/l salinity) while a mixture of *Pseudomonas, Stenotrophomonas*, *Achromobacter*, *Sphingobium*, *Aeromonas* and *Shewanella* predominated in that from lakes and rivers (≤ 3 g/l salinity). Interestingly, changes in water salinity not only determined *Vibrio* proportions in skin-mucus but also the *Vibrio* species that

formed part of the microbiome. Thus, skin-mucus microbiome was dominated by *V. anguillarum* in eels from estuary and by *V. cholerae* and *V. metoecus* in eels from lake. It is well known that *V. anguillarum* is more halophile than *V. cholerae* and its closest-species, *V. metoecus* (Bordas *et al.*, 1998; Singleton *et al.*, 1982a, 1982b), which could suggest that the composition in bacterial species of mucus microbiome was reflecting that of the eel-surrounding water. However, the most abundant vibrio species in mucus from eels in estuaries (*V. anguillarum*) was not present in water, at least undetectable for metagenomics, which also suggests that mucus probably concentrates selectively the most successful mucus-colonizers present in water. Interestingly, the proportion of *V. vulnificus* in SM-metagenomes was the same in all the metagenomes (around 0.5-1) and achieved a 35% in that from the estuarine water of highest salinity (10 g/l salinity). This finding supports the hypothesis that *V. vulnificus* is part of the resident SM-microbiota in eels from natural estuarine environments. We did not find evidences of the presence of the virulence plasmid in the metagenomes, suggesting that these resident vibrios do not belong to the eel-virulent biotype. According to this finding, biotype 2 is not abundant, at least in the sampled ecosystems. The most accepted hypothesis stablishes the origin of the biotype 2 in Asiatic fish-farms after conjugation events between an unknown donor and fish-commensal strains of *V. vulnificus*. This biotype successfully amplified after successive outbreaks and spread to the rest of the world, probably with imported animals.

The first vibriosis in farmed-eels occurred as outbreaks of high mortality in the 80's and the 90's (Amaro *et al.*, 1992; Biosca *et al.*, 1991). To control them, some eel-farmers decided to use fresh water instead of brackish water, but this measure resulted in the emergence of new serovars, less virulent than the original one (serovar E), but adapted to salinity between 0.2-0.4 % (Fouz *et al.*, 2006). After that, most of the intensive eel farms closed and the remaining combined low salinity with low pH, a measure that seems to have been efficient according to the results obtained in the present work: no vibrio, including *V. vulnificus,* was detected in SM from farmed eels. Moreover, these changes also reduced the presence of other members of the SM-microbiota of wild-eels and only

Pseudomonas, *Acinetobacter, Stenotrophomonas* and *Sphingobium* were also detected in the mucus of farmed-eels in significant proportions. In return, genus resistant to these conditions colonized the skin mucus of farmed eels, mainly: *Comamonas, Citrobacter* and *Chryseobacterium.* In particular, *C. testosteroni* seems to be able to resist perfectly these conditions and colonize the SM of farmed eels. Previous reports already proved the resistance of *C.testosteroni* to low pH and antibiotics (Oppermann *et al.*, 1996; Ma *et al.*, 2009). This result is quite interesting, because confirms that the outcome of infectious diseases in water can be modified by changing the external environment at which the fish's mucosal surfaces are exposed and, consequently, natural mucosal microbiome.

Although the captured wild-eels were apparently healthy, their SM contained a high proportion of pathogenic bacteria mainly belonging to *Vibrio* genus (*V. anguillarum, V.cholerae, V. vulnificus, etc.*).We suspected that mucus could concentrate aquatic bacteria with special abilities to attach to, to survive in, and to multiply on mucosal surfaces. To discover what abilities would have to display a successful-mucus colonizer, we compared the functionalities of the mucus- and water-microbiome. The results suggest that potential colonizers should be able to adhere, mainly by forming biofilms, compete (T6SS) and communicate with other bacteria, and resist the innate immunity (lysozyme, iron starvation), the predators (bacteriophages and amoeba) and the toxic effects of heavy metals and drugs. Besides these, some important virulence factors (RTX toxins, TCP pilus, antibiotic resistance, exotoxin A, toxin U) were also detected in the attached microbiota. Interestingly, most of these abilities are displayed by human pathogenic intestinal bacteria (Lee *et al.*, 2007; Fullner and Mekalanos, 2000; Sato and Frank, 2014; Michalska and Wolf, 2015; Waldor and Mekalanos, 1996). According to this, mucus could act selecting those vibrios from the total pool of vibrios present in water that carry specific virulence genes and constitute a reservoir not only for pathogenic bacteria but also for virulence genes. We obtained an interesting proof in favor of this hypothesis. We isolated and sequenced a strain of *V. metoecus*, the species closest to *V. cholerae* and found that our isolate contained genes in common with *V. cholerae* that were not present in the published genomes of *V. metoecus* from water or extraintestinal

infections: a vibriobactin biosynthetic cluster to overcome iron starvation (nutritional immunity) (Keating *et al.*, 2000); a gene for an ElTor-RTX-like toxin to fight against protozoa (predators) in biofilms (Lee *et al.*, 2007); a TCP pilus-biosynthetic cluster to colonize human intestine and receipt CTXphi phage (Waldor and Mekalanos, 1996); and, finally, a biosynthetic cluster for a T6SS to compete with other bacteria (Ma and Mekalanos, 2010). All these findings support the hypothesis that mucus is acting as a selective pressure in favor of the phenotypic traits needed to survive and proliferate in it, some of which are considered virulence factors for humans. This last observation also would suggest that the mucus could accelerate the evolution of intestinal pathogens as *V. cholerae*.

The acquisition of new genes by HGT has been determinant in the evolution of *V. cholerae* and *V. vulnificus* as pathogenic species; the plasmid pVvBt2, in case of *V. vulnificus,* and the phage CTXphi, in case of *V. cholerae*. As we have commented before, we did not find any pVvBt2-related gene in our metagenomes. However, we found genetic markers for the phage, mainly the toxin Zot. In fact, our isolate of *V. metoecus* presented in its genome evidences of having possessed the phage and a related satellite-phage. This made us suspect that mucus could also favor the genetic material interchange and contain MGE harboring virulence genes.

To answer this question we decided to investigate the presence of MGE in our metagenomes by using a new workflow. According to our results, the major interchange of genetic information among bacteria in mucus is done by using ICEs as vectors. But contrary to that expected most of the genes carried by ICEs and the rest of MGEs were heavy-metals- and drug-resistance genes. Although we could not confirm it, we also detected presumptive pathogenic islands, one of which carried a virulence factor related with resistance to phagocytosis in *Pseudomonas*(Michalska and Wolf, 2015; Sato and Frank, 2014; Sharma and FitzGerald, 2010). We also found evidences that MGE elements are only successfully fixed in bacterial genomes when the interchange was produced between bacteria with similar GC percentage.

Surprisingly, we found multiple phage-related contigs in our metagenomes that were analyzed in depth. The BAM model of Barr *et al*, (2013) proposes that mucosal surfaces concentrate phages present in the surrounding environment, which act as a part of the mucosal innate defences, controlling bacteria population. Contrary to that expected, we did not find that SM contained more phage than water. On the contrary, bacteria were relatively more abundant in our mucus samples than in water. This finding does not contradict the hypothesis of Barr *et al.* (2013). First, they sampled human intestinal mucus, a completely different environment, and second, the sampling way to get mucus (see 4.1.) probably did not retrieve the mucus layers most closely associated to the fish surface. In any case, we found a replicating viral community in this niche formed by Myovirus and Podovirus. In addition, we found evidences that ΦKZ genus and the Podovirus could be part of the resident microbiota associated to the eel mucosal surface.

Phages use Ig-like domains present in the capsids to attach to mucin. However, only a single putative Ig-domain was found in MyoΦKZC12 and a fibronectin type III domain in MyoΦKZC1 phage. Considering that the Barr study was performed in T4 group bacteriophages and MyoΦKZC12 and MyoΦKZC1 are members of a genus classified as T4-like, it is possible that attachment through Ig-like domains is specific of T4 bacteriophages and the rest use a different protein. When other Ig-like and carbohydrate-binding domains were searched in the sequenced phages, a C-type lectin domain was found in MyoΦKZC1 and MyoC35. Moreover, three and one Ig-like domains annotated in Pfam as invasion/intimin cell adhesion were found in MyoΦKZC12 and MyoC197, respectively. Finally, we found BACON domain (Bacteroidetes-Associated Carbohydrate-binding Often N-terminal) in a tail protein of MyoC197.

In summary, this is the first time that the microbiota associated to a wild aquatic organism is assessed. The wild-eel-mucus microbiota is especially rich in vibrios belonging to pathogenic species as well as in MGE and phages. We also found evidences that mucus could accelerate the evolution of intestinal vibrios by

selecting for traits that could make them successful in mucus intestinal colonization and by favouring genetic interchange.

CHAPTER 2

CHAPTER 2

Vibrio vulnificus: growth in human serum

1. Abstract

V. vulnificus is an extremely invasive pathogen able to spread from the infection site to the blood stream and cause death by sepsis in less than 24h. To be so successful in sepsis, this pathogen needs efficient systems to overcome the innate immunity and multiply rapidly in human blood. The molecular basis for growth in human serum have been studied in *V. vulnificus* basically applying a single gene approach. In this thesis we have applied for the first time a genome approach that combines transposition mutagenesis with high-throughput DNA sequencing, a technique known as transposon insertion sequencing or TIS (see 3.2). TIS has allowed us to discriminate which are the genes of *V. vulnificus* essential to grow in human serum from a genome wide perspective and at the same time, to separate those involved in resistance to complement killing from those essential to overcome nutritional immunity.

TIS results suggest that genes involved in capsule biosynthesis and transport are the only ones essential to resist complement killing in *V. vulnificus*. We found two new genes *VV0357* and *VV0358* located in the capsule locus, whose participation in the capsule formation was proven after getting the corresponding mutants. These two genes are exclusive for the strain selected for the study, YJ016, a strain especially virulent and resistant to human serum. On the other hand, we identify a series of genes putatively essentials to overcome nutritional immunity. Among them, we highlight genes related to i) iron-uptake (*hupA*, spermidine/putrescine transport system); ii) homeostasis under stressing conditions (*clp* and *lon*); iii) secretion of exoenxymes and toxins (T2SS); and iv) ionic gradient built up (Na+-translocating NADH-quinone reductase complex); Finally, we also identified genes putatively involved in resistance to microcidal peptides (LPS changes and *ompH*).

Finally, we compared our results with that obtained in other pathogens, in particular in other vibrios, and propose a list of genes essential for *Vibrio* to survive in artificial culture media.

2. Introduction

Vibrio vulnificus is an aquatic pathogen able to cause a disease named vibriosis in humans and eels (Amaro and Biosca, 1996). Human vibriosis is acquired either by ingestion of raw seafood (vibriosis type I) or by contact of a wound with brackish water/infected animals (vibriosis type II). Although fish vibriosis is also acquired by contact, a pre-existing wound is not a requirement since the pathogen can use intact epithelium (mainly that of gills) as a portal of entry into host (Marco-Nolaes *et al*., 2001). The ability of the pathogen to spread rapidly from the epithelium to the bloodstream is common to human and eel vibriosis. In fact, the incubation time for both vibriosis is quite rapid, averaging only 8-12 h, in case of eels, and 16-24 h in case of humans. For this reason, rapid antibiotic treatment of septicemic patient/animal is critical, with infected hosts not being treated before 3 days having a 100 % fatality rate (Jones and Oliver, 2009, unpublished results).

The capacity of *V. vulnificus* to cause a fast death by sepsis in the eel has been directly related with rapid growth in blood (Amaro *et al*., 2015). This ability partially relays on a plasmid, exclusively present in biotype 2 strains (pVvBt2) that encodes resistance to eel-complement killing as well as ability to use eel-transferrin as an iron source (Pajuelo *et al*., 2015). pVvBt2 also harbors a copy of the *rtxA1* locus, which is involved in non-host specific resistance to phagocytosis by neutrophils (Jones and Oliver, 2009). This combination of genes is so efficient that the pathogen is able to cause death by sepsis in healthy non-immune eels (Marco-Noales *et al*., 2001). For all these reasons, *V. vulnificus* is considered as a primary eel pathogen.

In case of the human, the invasion mechanisms are far away from be completely clarified, probably because of the high genetic variability of the strains able to cause human vibriosis (the three biotypes are able to infect humans). Thus, some of the most relevant virulence factors that have been proposed to be involved in human invasion by *V. vulnificus* are: a potassium pump involved in resistance to human serum (Chen *et al*., 2004) ; a capsule (CPS), a surface lipoprotein and a

flagellin, all of them involved in triggering an inadequate citoquine response (Lee *et al.*, 2010; Powell *et al.*, 1997; Chen *et al.*, 2004; Lo *et al.*, 2011; Yoshida *et al.*, 1985; Kuo *et al.*, 2015; Kim *et al.*, 2014; Duong-Nu *et al.*, 2015); iii) the toxin RtxA1 in resistance to phagocytosis (Lo *et al.*, 2011; Kuo *et al.*, 2015); and finally a sialic acid modified surface involved in mice invasion (Lubin *et al.*, 2015). Additionally, it is generally accepted that virulence for humans is directly depending on iron content in blood (Bullen *et al.*, 1991;Pajuelo *et al.*, 2016). This statement is based on epidemiological data: individuals with chronic liver disease with high iron levels in blood have been reported to be up 80 times more likely to develop primary sepsis than healthy individuals (Centers for Disease Control and Prevention (CDC), 1993). For all these reasons, *V. vulnificus* is considered as an opportunistic human pathogen (a secondary and accidental human pathogen).

The second chapter of this thesis is focused on clarifying the invasion strategies used by *V. vulnificus* to grow in human blood by applying the "omic" approach known as transposon insertion sequencing (TIS). TIS is a powerful method that couples high-density transposon mutagenesis with next-generation sequencing to comprehensively assess the fitness of thousands of transposon mutants across a genome. TIS considers the essentiality of a gene in a condition should be proportional to its presence in the library. According to this criterion, genes present in both control and tested libraries would be not essential or neutral, and genes present only in the control library would be essential for growth in the tested condition.

The specific objective was to discover all the genes from the pathogen involved in resistance to humoral innate immunity in blood. Humoral innate defenses in blood maintained *in vitro* mainly rely on the complement killing together with nutritional immunity. Both conditions can be simulated by growing the bacterium in fresh human serum and, in parallel, in heat-inactivated serum (56ºC, 30 min), a process that abolishes human complement action but not the bacteriostatic effect on growth of nutritional immunity. Thus, we first selected a strain after an initial screening of growth in fresh human serum, set up the growth conditions for TIS analysis and performed the experiments considering two tested-conditions,

growth in human serum (HS) and inactivated human serum (IHS) and compared between them and with growth in an artificial culture medium, LB. By comparing the libraries obtained in HS *vs* IHS conditions, we expected to see all the genes needed to resist human complement, by comparing IHS *vs* LB conditions, all the genes involved in resistance to nutritional immunity, and essential to persist and grow in HS. Then, we got a knockout mutant for selected genes and probed their functions in resistance to the humoral innate immunity by performing *in vitro* and *in vivo* tests.

3. Hypothesis and objectives

Our hypothesis is that virulence for humans in *V. vulnificus* is directly related with ability to resist the humoral innate immunity in blood. This hypothesis is based on the results obtained in the eel (Roux *et al*., 2015). To discover the involved genes at genome wide level, we first create a highly dense transposon library. Once this was correctly created, we grew it *in vitro* (LB) and HS active and inactive. Sequencing and comparing the results helped us to achieve the following objectives:

   3.1. Creating a Transposon Library *in vitro*
      3.1.1.   Essential genes in a general culture medium and comparison to other bacteria
   3.2. Essential genes in HS
      3.2.1.   Survival in HS
      3.2.2.   Growth in HS

4. Material and methods

4.1. Strains, plasmids and general culture conditions.

Strains and plasmids used in this chapter are listed in **table 8**. All the strains were routinely grown in LB (with shaking at 200 rpm) or on LBA (LB-agar) at 37 ºC for 18-24 h in agitation (170 rpm). A spontaneous resistance mutant (YJSm^r) was selected after growing the strain YJ016 on LBA-Sm (streptomycin, 500 µg/ml) and checking that the mutant- and the wild-type's growth curves in LB were statistically similar. YJSm^r was used to generate all the knockout mutants in the genes selected after TIS assays. Antibiotic concentrations used in the experiments for genetic modification were: Sm, 500 µg/ml; chloramphenicol (Cm), 20 µg/ml (*E.coli)* or 2 µg/ml (*V. vulnificus)*; and kanamycin (Km), 150 µg/ml. *E. coli* SM10 *lambda pir* carrying the Himar1 suicide transposon vector pYB742 (CmR) was used for transposon mutagenesis (Yamaichi *et al*., 2015). All strains were kept in LB+ 20% glycerol at -80 ºC.

4.2. Resistance and growth in HS.

Cells in exponential-phase of growth in LB were washed twice in PBS and were inoculated in HS (commercial HS, Sigma, H4522) or inactivated-HS (56 ºC, 30) by triplicate (10^5 cfu/ml). Bacterial counts on LBA plates were performed at 0, 2, 4 and 6 h post incubation at 37ºC. For that, serial ten-fold dilutions in PBS were performed in 96-well plates, and 10 µl of each dilution were spotted on LBA plates (drop plate count).

Table 8 Species, strains and plasmids used in the study

Specie, strain, plasmid	Characteristics/Genotypes	Isolation/Source	Country	Year of isolation
Vibrio vulnificus				
YJ016 (YJ)	Biotype 1; Serovar non-E, non-A, non-O	Human blood	Taiwan	1993
CECT5168 or CDC7184	Biotype 1; Serovar non-E, non-A, non-O	Human blood	USA	1977
CECT4999	Biotype 2, Serovar E	Diseased eel	Spain	1999
CECT5769	Biotype 2, Serovar A	Diseased eel	Spain	2004
12	Biotype 3, Serovar O	clinical	Israel	2002
11028	Biotype 3, Serovar O	clinical	Israel	1997
YJSm^r	YJ016 streptomycin resistant	This study		
YJSm^{r}Δ337	YJ016SmR VV0337-defective mutant	This study		
YJSm^{r}Δ357	YJ016SmR VV0357-defective mutant	This study		
YJSm^{r}Δ358	YJ016SmR VV0358-defective mutant	This study		
YJSm^{r}c357	Δ357 complemented strain	This study		
Escherichia coli				
DH5α	Cloning strain F$^-$ Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r$_k^-$, m$_k^+$) *pho*A *sup*E44*thi*-1 *gyr*A96 *rel*A1 λ$^-$	Invitrogen		
SM10 *λpir*	Cloning strain and transposon donor *thi thr leu tonA lacy supE recA::RP4-2-Tc::Mu λpir Kan*^{r}	(Simon *et al*., 1983)	USA	1983
Nova Blue	Cloning strain and protein expression			
Plasmids				
pGEM-T easy	T/A Cloning vector, Amp^r	Promega		
pDM4	suicide vector, deletion mutant construction	(Milton *et al*., 1996)		
pYB742	Transposon vector	Waldor		
pMMB207	Deletion mutant complementation	Addgene		
pET30a	expresion vector (His-tag)			

4.3. Transposon library construction

4.3.1. Transposon mutagenesis

All the libraries (two per tested or control condition) were constructed according to Chao *et al*. (2013) and Pritchard *et al*. (2014). Briefly, washed bacteria from 1 ml of an overnight culture either of YJSm^r or *E.coli* SM10 transposon donor (**Table 8**) were mixed in a final volume of 500 µl of fresh LB. Then, 50 µl-aliquots were spotted onto 0.45 µm-filters (Millipore, HAWP04700) placed on

LBA plates and were incubated for 6 hours. After conjugation, the filters were recovered and introduced in 50 ml-tubes containing 6 ml of LB (5 filters per tube), tubes were vigorously vortexed, and 3 ml were spread onto LBA+Sm+Km plates ($24.5x24.5cm^2$, Corning) and incubated for 24h. Km was added to agar medium to select the bacteria with transposon inserted.

Bacteria grown on the plates (approx. 1.5 million per plate) were scrapped with 12 ml of LB and introduced in 50 ml-falcon tubes. Then, 3 ml were used to obtain genomic DNA (see below) and the rest were kept with 20% glycerol at -80 ºC.

4.3.1.1. DNA extraction

Genomic DNA was extracted with the Wizard Genomic DNA extraction™ kit (Promega), using at least 10x volumes for all the lysis reagents. Final precipitation step for genomic DNA was performed as described by Pritchard *et al*., (2014). Briefly, cells were pelleted, resuspended in lysis buffer and incubated at 70-80 ºC with shaking until we saw lysis (approx.. 5 min). After an RNase A treatment (30-60 min at 37 ºC) followed by protein precipitation on ice (5 min), samples were centrifuged at maximum speed in 2 ml-eppendorfs, supernatants, recovered and, finally, pipetted into 15 ml-corning tubes. Then, DNA was precipitated with isopropanol, washed once with 70 and 100 % ethanol, dried for 15 min and resuspended in TE buffer (1 M Tris-HCl (pH 8.0) and 0.2 ml EDTA (0.5 M)) at -80ºC until use.

Plasmid DNA was extracted using Zyppy Plasmid Mini- (D4019) or Midiprep™ Kit (D4026). We follow manufacture protocol except for an extra centrifugation after last wash with an empty eppendorf at maximum speed and we used water for eluting DNA. Plasmid samples were kept at -20ºC until use.

4.3.1.2. DNA treatment

In order to map in the genome the transposon we need to enrich the extracted DNA in sequences lying downstream of the transposon insertion.

4.3.1.2.1.   Shearing and end repairing

5-10 µg of genomic DNA was sheared in fragments of ≈350 bp by acoustic disruption (M220 focused ultrasonicator, Covaris, Woburn, MA). Then, blunt ends were repaired with the Quick Blunting™ Kit (NEB, E1201L). Per sample:

5-10 µg	Sheared DNA
5.5 µl	10x Buffer
4 µl	dNTPs 1 mM
2 µl	Blunting enzyme mix
To 55 µl	dH$_2$O

4.3.1.2.2.   Tailing with Taq

Adaptor needs to be ligated to sheared DNA. In order to prepare DNA for efficient ligation with adaptors (T tail overhang), we tailed with adenine nucleotides to form an "A" tail overhang on the ends. Per sample:

32 µl	Blunted DNA
5 µl	10x PCR standard Buffer
10 µl	dATP 10 mM
3 µl	Taq polymerase

4.3.1.2.3.   Adapter and ligation

First, adaptors need to be created by ligating:

2.4 µl	100 µM NH$_2$ truncated fork primer 5'-TACCACGACCA-NH$_2$-3'
2.4 µl	100 µM index fork adapter 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGTCGTGGTAT-3'
0.2 µl	2 mM MgCl$_2$

Settings: 95 ℃ for 4 min, 95 ℃ for 1 min and repeat 75 times decreasing 1 ℃ every time (until 20 ℃). Then created adapters are ligated to tailed DNA purified. Per sample:

1.2 µg   Purified DNA sample

0.8 µl   Adapter mix

1.5 µl   10x T4 DNA ligase buffer

1 µl   T4 DNA ligase (Thermo Scientific, EL0014)

To 15 µl H_2O

Incubate at 16 ºC or room temperature overnight and then spike ligation reaction by adding the next reagents per sample and incubating 2h at same temperature:

1 µl   10x T4 DNA ligase buffer

1 µl   T4 DNA ligase

8 µl H_2O

4.3.1.2.4.   Amplification of transposon junctions for sequencing

A PCR is required to enrich sample on transposon associated sequence. First, transposon junctions are amplified with primers to the transposon (Himmer-3-out primer) and to the adaptor previously ligated (index R primer). This will amplify only in one direction (as synthesis in blocked by the NH_2group). Do 5 PCR tubes per sample:

50 µl   5x Phusion HF buffer

500 ng   Ligated DNA (100 ng/tube)

1.25 µl   100 µM Himmer-3-out primer

1.25 µl   100 µM index R primer 5'-GTGACTGGAGTTCAGACGTGTG-3'

6.25 µl   dNTPs 10 mM

To 250 µl dH_2O

1.25 µl   Phusion DNA polymerase (Thermo Scientific, F530S)

Aliquot 50 µl into PCR tubes and run with next settings:

1=95ºC 1 min
2= 98ºC 10 sec
3= 53 ºC 30 sec
4= 72 ºC 30 sec

5=go to 2, 29 times
6=72 ºC 10 min
7= 4ºC for ever

PCR reaction needs to be purified before second PCR in which barcodes and P5 and P7 attachment site for Illumina sequencing are added. Again, aliquot mix in 5 PCR tubes:

50 µl	5x Phusion HF buffer
500 ng	Purified PCR product from above
1.25 µl	100 µM 5' variable primer
1.25 µl	100 µM 3' Index barcode (vary per simple)
6.25 µl	dNTPs 10 mM
To 250 µl	dH$_2$O
1.25 µl	Phusion DNA polymerase

Use same settings with these changes:

3=55 ºC 30 sec
5=go to 2, 17 times

Finally, a selection of size fragments (200-400 bp) was done after running the purified sample in an electrophoresis gel at 2% agarose.

4.3.1.3. DNA sequencing

Number of DNA molecules was estimated using qPCR (SYBR) by triplicate with a standard curve. Efforts were made in order to sequence 10 million reads per sample on a MiSeq (Illumina, San Diego, CA).

4.4. Transposon libraries assays

An aliquot of the obtained library was diluted 1:30 in LB, HS and inactivated-HS by duplicated to obtain approximately 10^8 cfu/ml. How much the library must be diluted to obtain this concentration was calculated with spectrophotometer and proved by recounts. After 2h of incubation, cells were recovered by centrifugation, resuspended in 3 ml of LB and spread on LBA+Sm+Km plates (24.5x24.5cm^2, Corning). Plates were incubated at 37ºC overnight. Like we did with *in vitro* library, we recover the bacteria grown in the plates and 3 ml were used for DNA extraction and the rest kept with 20% glycerol at -80 ºC.

4.5. Differentiate *neutral* and *under-represented* genes

Previously described methodologies were used to detect insertion sites and to visualize (Artemis) and statistically analyse the results (ARTIST) (Pritchard *et al.*, 2014; Chao *et al.*, 2013). Briefly, >15 nt reads were mapped to the *V. vulnificus* YJ016 genome database composed of chromosome 1 (NC_005139.1) chromosome 2 (NC_005140.1) and plasmid pYJ016 (NC_005128.1), allowing for no mismatches. We discarded reads that did not align at any TA site in the genome and we randomly distributed the reads mapping at multiple TA sites. The number of reads at each TA site was tallied, datasets were normalized for origin proximity, and transposon-insertion profiles were depicted using Artemis. EL-ARTIST was used to analyse the two *V. vulnificus* chromosomes and the plasmid pYJ016, independently (Pritchard *et al.*, 2014). Chromosome 1 classifications were obtained using hidden Markov model analysis following usual values for sliding window (SW) training (10 TA sites; $P < 0.005$) while chromosome 2 classifications were obtained using SW analysis (15 TA sites; $P < 0.005$). For chromosome one and plasmid, genetic loci with fewer than 10 TA sites were retroactively excluded from classification. In chromosome two, 15 TA sites were used.

Because of the differences in TA % disrupted between libraries we used Mann-Whitney test in order to compare the TA abundances. Minimum 90% MWU test values, p-value<0.00001 and fold change (FC) bigger than 5 were considered for the analysis. The FC represents the number of times a gene is more represented in one library *vs* other. These criteria used will exclude the genes with a small number of TA sites.

4.6. Bioinformatic analysis of interesting genes

Genes listed as essential to grow or survive in HS were checked manually for annotation. Those who were annotated as *hypothetical protein* were reannotated using HHpred or BLASTP against Genbank. Moreover, domains were analysed in Conserved Domain Database (CDD) at the NCBI website. Protein Homology/analogy Recognition Engine V 2.0 (Phyre^2) pipeline was used for VV0358 (Kelley *et al.*, 2015). Primers were designed manually using basic

sequence edition programs, mainly ApE (A plasmid Editor v2.0.47;http://biologylabs.utah.edu/jorgensen/wayned/ape/).

4.7. Deletion mutants

Deletion plasmids for selected genes were generated by allelic exchange. Each of the mutants in **table 8** were constructed by ligating ≈500 bp PCR products generated from the primers in **Table 9**to the suicide vector, pDM4, using Gibson assembly (Gibson *et al*., 2009). PCR was performed with Phusion DNA polymerase in a final volume of 20 µl. Per sample:

4 µl	5x Phusion HF buffer
0.2 µl	50 mM $MgCl_2$
1 µl	10 mM Forward primer
1 µl	10 mM Reverse primer
1 µl	dNTPs 10 mM
12.6 µl	dH_2O
0.2 µl	Phusion DNA polymerase

Every PCR was run on a gel of agarose at 1% and stained with RedSafe Nucleic Acid Stanning Solution (20000x) (JH science, 21141). The constructed plasmid was introduced by electroporation in *E.coli* SM10 after cleaning the ligation using dialysis and then inserted by conjugation in YJSm[r]. Dialysis was realised by diluting 10 µl of the ligation with 5 µl of molecular water over a 0.02µm-filter (Millipore, VSWP02500) and incubating in a plate with 10 ml of miliQ water during 15 min. Positive colonies for constructed plasmid were checked by PCR, in this case with DreamTaq DNA polymerase (Thermo Scientific, EP0702). Per sample:

2.5µl	10x DreamTaq buffer
1 µl	10 mM Forward primer
1 µl	10 mM Reverse primer
1 µl	dNTPs 10 mM
19.4 µl	dH_2O

CHAPTER 2

<center>0.1 µl   Phusion DNA polymerase</center>

To carry out the conjugation, 40 µl of washed overnight cultures of each bacterium were mixed and spotted on 0.45 µm-filter (Millipore, HAWP04700). After 4h of incubation, cells were recovered in LB and spread on LBA+Cm+Sm. Finally, sucrose-based counter-selection was performed as described by Donnenberg and Kaper (Donnenberg and Kaper, 1991). Briefly, 3 random Cm+Sm resistant colonies were grown overnight, the culture was diluted 1:100 in LB with 10% sacarose and incubated again until it reached a value of Abs_{600}=0.5. Then, 100 µl of eachone of a serial ten-fold dilutions of the culture were spread on LBA plates. PCR and DNA sequencing were used to confirm the junctions of each in-frame deletion.

Table 9. Primers used in chapter 2

Primer	Restriction site	Sequence	Product size (bp)	Utilization
337-A-F	Sma1	acaatttgtggaatcccgggtagctgcaacgatgtgaatttatg	500	Allelic exchange
337-A-R	Sma1	tacggccattgtttttagttccatgtaagcc	500	Allelic exchange
337-B-F	Sma1	ctaaaaacaatggccgtaagaaaattttatagg	500	Allelic exchange
337-B-R	Sma1	aacctgagctctcccgggtcccactcctcctagaaatattc	500	Allelic exchange
357-A-F	Sma1	acaatttgtggaatcccgggctcggtgttgagaaagtatc	500	Allelic exchange
357-A-R	Sma1	gtttacgcacttgtaatatttgtttcatcagtaattaatc	500	Allelic exchange
357-B-F	Sma1	atattacaagtgcgtaaacaacaatgagtctc	500	Allelic exchange
357-B-R	Sma1	aacctgagctctcccgggccgtcatagccaagtgc	500	Allelic exchange
358-A-F	Sma1	gcggataacaatttgtggaatcccggttaaagagcc	500	Allelic exchange
358-A-R	Sma1	tagctgtaattatactaatttgagactcattgttgtttac	500	Allelic exchange
358-B-F	Sma1	atgagtctcaaattagtataattacagctaatgaatattc	500	Allelic exchange
358-B-R	Sma1	catgcgggtaacctgagctctcccgtccacaccacg	500	Allelic exchange
358-F-Nde1	Nde1	cgcgcatatgatgagtctcaaatta	1812	Protein recombination
358-R-Sal1	Sal1	cgcggtcgacgctgtaattatactc	1812	Protein recombination
358-F-Pst1	Pst1	cgcgctgcaggcggatcaattgcgtaaacaa	1812	complementation
358-R-BamH1	BamH1	cgcgggatccgtgctggtgaaaataaaga	1812	complementation
357-Fw-EcoR1	EcoR1	cgcggaattcgggttacgaaacgtatttgagag	2142	complementation
357-Rv-Xba1	Xba1	cgcgtctagaaagtcgattgcagtttgcgc	2142	complementation

4.8. Complementation of deletion mutants

Complementation of the mutants was done using pMMB207 (Cm^r). The 1.8 and 1.5 kb of genes *VV0357* and *VV0358* were amplified by PCR. PCR purified products and plasmid were digested using two different FastDigest enzymes (Thermo Scientific). Phosphates were eliminated from plasmid using FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific, EF0652) and both were cleaned directly from the tube (without separating step on electrophoresis gel). After overnight ligation using T4 ligase we cleaned the mix and transform *E. coli* DH5α cells. SM10 cells were then transformed with the plasmid inserted in DH5α cells by electroporation or by conjugation and then conjugated to the corresponding deletion mutant using the same protocol described before. Colonies were checked for the gene and tested for phenotype recovery.

4.9. Libraries comparison

We compare our data to previous published studies. The list of genes essential to grown in a rich media or to survive in the serum in the libraries was compared. To do that we had in account the number of copies of the gene since duplications made impossible to detect an essential gene using TIS. . Duplications means that no single transposon insertion mutant can disrupt both genes in the same cell, so one copy can always rescue/suppress the mutant. BLASTP tool was used to check for the presence and the number of copies of the genes from *V.cholerae* N16961, *V. parahaemolyticus* RIMD 2210633, *Moraxella catarrhalis* BBH18 and *E. coli* ST131 and K12 in *V. vulnificus* YJ016. We considered hits that had at least 35% of identity, covered 85% and the p-value was smaller than 10^{-5}.

4.10.   Mice Infection

Wild type and deletion mutants were grown to exponential phase as previously explained. Cells concentration was calculated by measuring the absorbance at 600 nm (ABS$_{600}$). Dilutions on PBS were done in order to inject on mice $1x10^6$ and $1x10^7$ cfu/ml. 4 BALB/c mice 6- to 8-week old were used per dose and

strain. Mice survival was checked after 24h and bacterial virulence calculated by LD_{50} (lethal doses 50%).

4.11.   Immunostaining of cell-associated polysaccharides

Crude fractions of cell-associated polysaccharides (LPS plus CPS) were obtained from overnight cultures in LB as described by Hitchcock and Brown(1983). LPS and CPS antigens were separated by SDS-PAGE (Laemmli, 1970) in discontinuous gels (4% stacking gel, 10% separating gel), transferred to a PVDF membrane (Bio-Rad) (Towbin *et al.*, 1979) and subjected to immunoblot analysis. The membranes were probed with YJ016-specific serum diluted 1:3000 and were developed following incubation with anti-rabbit IgG HRP-conjugated secondary antibody diluted 1:10000 (Sigma), using Immobilon Western Chemiluminescent HRP Substrate (Millipore).

4.12.   Real-Time PCR (RT-PCR)

The level expression of the genes was tested by RT-PCR. First, RNA was extracted from pelleted cells directly, or after two washes with LB if sample came from serum, using Direct-zol RNA MiniPrep kit (Zymo Research, R2052). DNA was eliminated with two treatments of TURBO DNase (Ambion, Life Technologies, AM2239). RNA was purified with RNA clean & concentrator kit (Zymo Research, R1018) or GeneJET RNA clean-up and concentration Micro kit (Thermo Scientific, K0841). cDNA was produced from 1 µg of purified RNA by M-MLV Reverse Transcriptase kit (Invitrogen) as described by the manufacturer.

Quantification of cDNA was performed with Power SYBR® Green PCR Master Mix (Applied Biosystems) by using the StepOne Plus RT-PCR System (Applied Biosystems). Reactions were carried out in a final volume of 20 µl, per sample:

1 µl	1 mM Forward primer (**Table 10**)
1 µl	1 mM Reverse primer
2 µl	cDNA
6 µl	DEPC H_2O
10 µl	2x Master Mix

The settings included 10 min of denaturalization at 95ºC followed by 40 cycles of 15 sec of denaturalization at 95ºC and 1 min of annealing and extension at 62ºC.

Table 10. Primers used for RT-PCR.

Primer	Sequence
337-F	GCCAAAGCGCAGCATTTTA
337-R	TGTATTCATAACGGGCGATCTG
357-F	CGGTACAATTCATTACCTAGCCAAT
357-R	TCCCTGCTGCGAAAACTTCT
358-F	CACGGATTACCGAGACTTTCG
358-R	CAAGCAGATGACTGCAAGAATAGG

4.13.   Heparinase activity assay

Heparinase activity was measured using three methodologies. First, the heparin degradation was measured using toluidine blue. Briefly, bacterium was incubated with PBS-1 with heparin during 6h at exponential phase. After that time, toluidine blue was added to the medium at X final concentration. Since this reagent reacts with heparin turning to pink, the absence of change was considered negative.

Second, the ability to cleave heparin was tested in blood. Fresh human blood was obtained from donors and maintained in special tubes with heparin. Bacteria tested were incubated in the blood and continuous observation were made to observe coagulation. Since heparin is an anti-coagulant, the coagulation of the blood was considered a positive result.

Finally, to test if the bacteria was able to cleave and use the heparin as a nutrient we grew the bacteria in minimal medium with heparin (Banga and Tripathi, 2009) and incubate the bacteria for 72 h to observe growth.

5. Results

5.1. Selection of the *V. vulnificus* strain

We tested the ability to resist and grow in HS of the *V. vulnificus* strains listed in **Table 8**. The strains were selected as representative of the biotypes and serovars of the species. The growth curves are shown in **Figure 56**. Only the two biotype 1 strains resisted the bactericidal and bacteriostatic effect of HS and effectively multiply after 6 h of incubation (**Figure 56**). The strain YJ016 showed the highest rate of growth in HS, which made it the best candidate to differentiate genes essential to resist the innate immunity in a short-time experiment (previous studies in the lab demonstrated that complement inactivates spontaneously after 6-8 h of incubation at 37ºC). YJ016 is a blood isolate of biotype 1 from a septicemic patient who eat raw oyster in Taiwan (Chen, 2003). This strain belong to the most virulent phylogenetic group (Sanjuán *et al*., 2011).
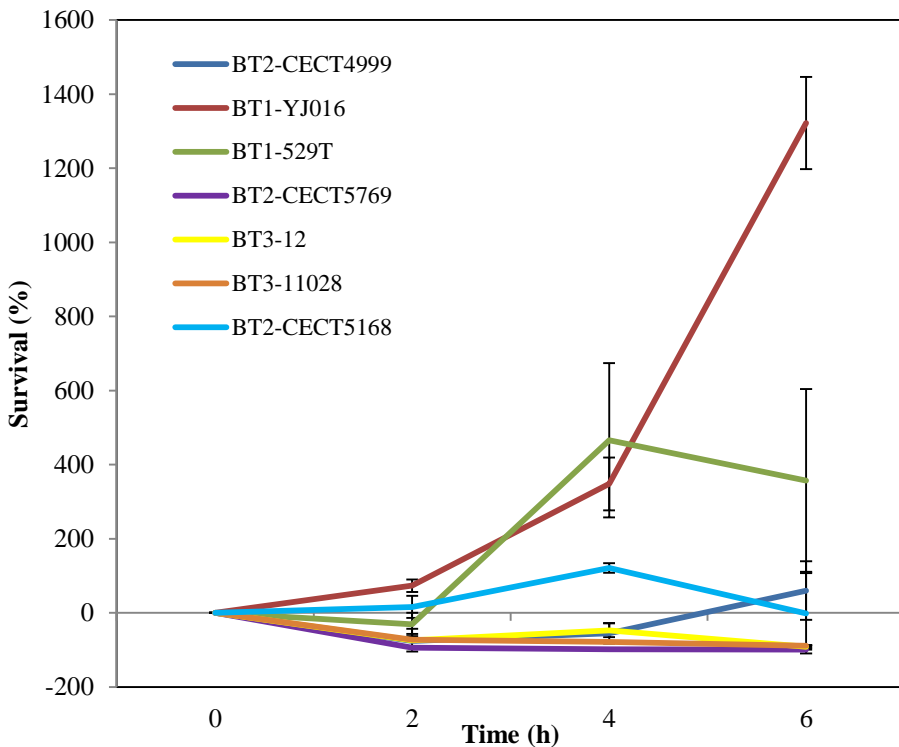
Figure 56. Resistance to HS of *V. vulnificus*. The initial bacterial population was considered "0" and survival at each time was calculated as a percentage. Bars indicate the standard deviation for three replicates.

5.2. Generation of a high density transposon library

The correct application of TIS methodology implies the creation of a library of mutants for the tested strain. This library must cover the complete genome if we want to study all the genes. Regrettably, the generation of a high-density transposon library is not a simple process and has to be adapted for each bacterial species, strain and experimental conditions. After a series of preliminary experiments, we decided to incubate YJ016 for two hours in LB, HS and inactivated HS (IHS).

A spontaneous streptomycin resistance strain of YJ016 (YJSm^r) was used to generate a high-density transposon-insertion library with a Mariner-based transposon (see methods 4.3). This transposon inserts in TA positions in the genome (Chiang and Rubin, 2002; Yamaichi *et al*., 2014). Library was first generated in LB that would be used as a control to discriminate genes essential to grow in a general culture medium. A total of 159,709 different transposon mutants inserted in TA positions were recovered from the library (75.9 % of TA in genome) according to sequencing results. Considering that we were missing the genes essential to grow in LB, it seems likely that we had achieved the objective of creating a library that fully covered all the genome. Regarding the analysis of the sequenced library, TIS considers that the essentiality of a gene for a condition should be inversely proportional to its presence in the library. According to this criterion, the genes which mutation makes bacteria unable to grow should be considered *essential* whereas genes whose mutation does not affect bacterial fitness would be consider *not essential* or *neutral*. In order to compare these libraries and discern which genes were "*under-represented*" in these functions we needed to be able to map where the transposon was inserted. To do that, high-throughput sequencing and bioinformatics were applied. Considering the TA positions that could have been disrupted and the TA positions covered in our library, we classified genomic loci as "*under-represented*"/"*regional*" when the number of transposons sequenced in the loci were significantly low or "*neutral*", when reads were mapped in practically the total sequence of the gene. Further, we differentiated within *essential* genes two

categories on the basis of the location of reads, "*under-represented*" and "*regional*" when the reads were located in the full sequence or in a part of the gene, respectively. EL-Artist analysis was used to detect genomic loci significantly *under-represented* in TA disruptions (*Under-represented* or *regional*) *vs* the rest (*neutral*) (Pritchard *et al*., 2014). This program, developed by doctor Waldor's laboratory can be trained to detect genomic loci significantly *under-represented*, according the number of TA positions in the genome and in the library created. Using previously consensual filters, we found that the number of genes of the YJ016 strain (two chromosomes and one plasmid: a total of 4978 genes) *under-represented* and *regional* in LB was 497 (9%) and the 197 (4%), respectively (**Figure 57**).
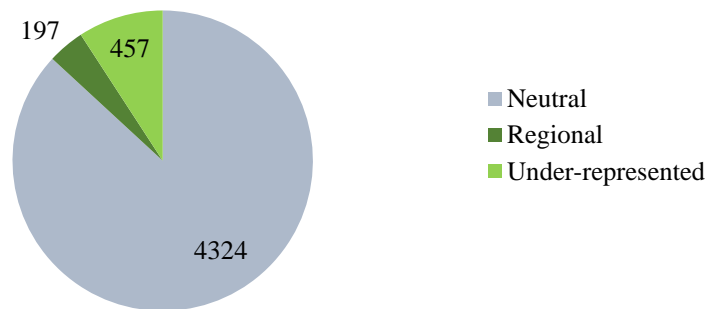
Figure 57. Genomic loci classification according to TIS results *in vitro*.

5.3. *Essential* genes to grow in LB and comparison to other proteobacteria

The default values in EL-ARTIST analysis were used for filtering results (sliding window 10 and p-value< 0.005). We detected 548 essential genes (408 *under-represented* and 140 *regional* genes) with more than 10 TA sites in chromosome I and plasmid and none in chromosome II. Then, we used a value for sliding window of 15 and a p-value of 0.005 together with manual (TA-percentage/gene) and visual (Artemis) analysis and we were able to identify 59 essential genes (16 *under-represented* and 43 *regional* genes with more than 15 TA sites) in chromosome II. In **Figure 58** the number of genes with different proportions of

disrupted TA-sites is plotted. We can see how the genes with a low percentage (*under-represented* or *regional* genes) are mostly localized in chromosome I. A similar distribution has been found by other researchers for other bacterial species with the same procedure (Chao *et al*., 2016; Hubbard *et al*., 2016; Palace *et al*., 2014).

A
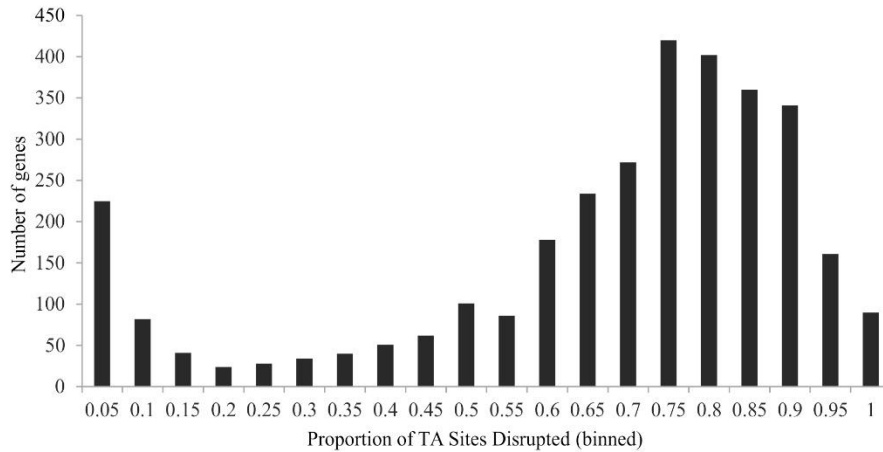
B
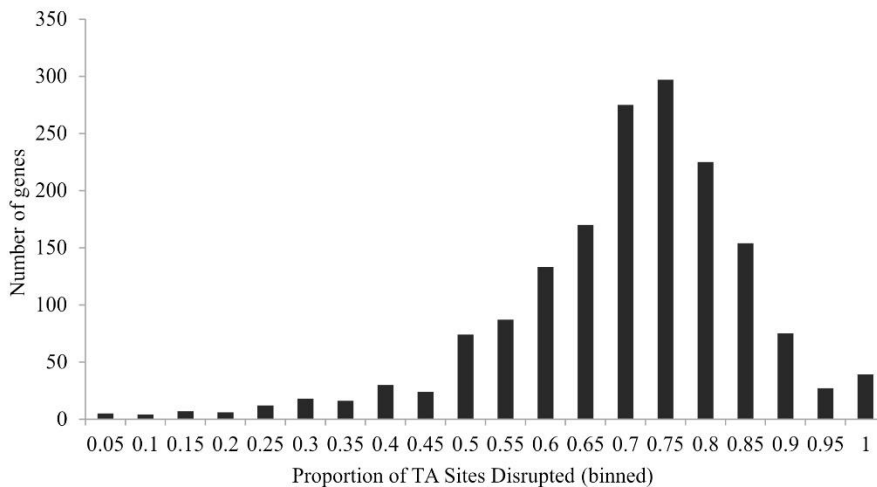
Figure 58. LB profiles. Distribution of proportions for genes with at least 10 TA sites in *V. vulnificus*. (A) TA percentage distribution in chromosome I. (B) TA percentage distribution in chromosome II. The number of reads in A and B that had a determined percentage of TA sites disrupted is represented in bars.

Surprisingly, we also found *under-represented* and *regional* genes in the plasmid pYJ016. This is a conjugative plasmid with no role either in virulence or in

growth in artificial media (Lien-I Hor, personal communication). The detected genes were: *VVP32* (*regional*), encoding a putative addiction module antitoxin, *VVP39* (*regional*), a putative PilT protein, *VVP66* (*under-represented*) a transcriptional repressor of the CopG family, and *VVP67* (*under-represented*), a hypothetical protein. The role for VVP32 as a part of a hypothetic toxin/antitoxin system would be the maintenance of the plasmid in the host cell while that of VVP66, as a putative member of a family of plasmid repressors, the control of the copy number of the plasmid. Then, to disrupt anyone of the two genes would be deleterious for the cell. With regard to the other two genes, we have no information about the reason for its essentiality to bacterial growth. In any case, we believe that having found these plasmid genes as essential for growth is a signal of the high feasibility and accuracy of the technique in discriminating those genes that are essential for growing in a general culture medium at genomic level.

Then, we compared these results to those previously obtained with the same technique for *V. cholerae* and *V. parahaemolyticus* by Chao *et al*. (2013) and Hubbard *et al*. (2016). As both studies did, we first compared the lists of *essential* genes for *V. vulnificus* YJ016 and *E. coli* K12 (Baba *et al*., 2006). We found discrepancies in 31 *under-represented* genes that were *essential* for *E. coli* and *neutral* for *V. vulnificus*. Taking into account that the list elaborated for *E. coli* was made by mutation of each gene and therefore is more robust, we analysed the discrepancies. 20 out of the 31 discrepant genes were duplicated in YJ016 genome, which would explain why they were not detected by TIS (mutation of one copy can be supressed by the other one). Thus, the discrepancies were reduced to 11 genes. A similar number of discrepancies were found by Chao *et al*. when compared the list of *essential* genes in *V. cholerae* and *E. coli*(Baba *et al*., 2006). Comparing all the *Vibrio* lists, we found common discrepancies such as genes involved in riboflavin biosynthesis and chromosome compaction, *neutral* for *Vibrio* and *essential* for *E. coli*. The neutrality of riboflavin genes in *V. cholerae* was proved by deletion of the genes by Chao *et al*., (2013). Further, the authors proposed a putative alternative pathway to transport exogenous riboflavin for this organism because mutants deficient in the

novo synthesis only could grow in minimal medium if it was supplemented with riboflavin. Even though no experiments were conducted to explain the essentiality of genes related to chromosome compaction, it is quite possible that alternative machineries could be codified in the genome of vibrios. Furthermore, *ftsX* was found not *essential* for both *V. vulnificus* and *V. cholerae*. The rest of discrepancies could be due to differences in the methodology used (TIS in *Vibrio vs* deletion mutants in *E. coli*) and the experimental conditions (i.e. differences in salt concentration in the culture medium).

Finally, the *essential* genes in LB for the three *Vibrio* species were compared. First, *V. vulnificus* YJ016 shares more *essential* genes with *V. parahaemolyticus* (356) than with *V. cholerae* (336), which could be related with the common habitat (marine waters and filtering organisms) and the phylogenetic proximity of both species (Lewis *et al*., 2011; Montiert *et al*., 2010) (**Figure 59**). However, the significance of this datum is unclear because *V. vulnificus* and *V. parahaemolyticus* also contain around 1000 more genes in their genomes than *V. cholerae* (4523-4831 *vs* 3504). Second, around 303 genes were essential for the three *Vibrio* species to grow in LB (*Vibrio-essential-genes*). **Table 11** and **12** show the unique-*essential*-genes for each species. These comparisons add value to the method and to previous results.
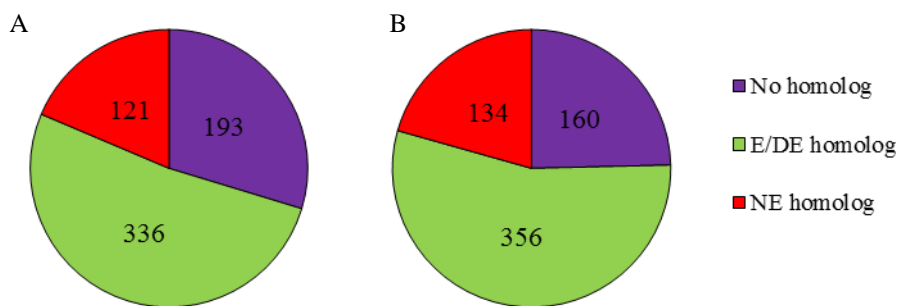
Figure 59. Comparison of essential genes to grow in LB in *V. cholerae, V. parahaemolyticus* sand *V. vulnificus*. The list of genes under-represented in *V. vulnificus* transposon library in LB were blasted to proteins from *V.cholerae* N16961 (A) and *V. parahaemolyticus* RIMD 2210633 (B). Homologs were categorized according to published studies in essential (E), domain essential (DE) and not essential (NE). Duplicated genes were considered as NE.

Table 11. List of essential genes for *V. vulnificus* (Vv) and *V. parahaemolyticus* (Vp) but not for *V. cholerae* (Vc). Essential genes for Vv were blasted against Vp and Vc and the results were plotted in the table. Genes with homology were indicated with the respective gene name and its EL-ARTIST category assigned in its study. These are: essential (E), domain essential (DE) and neutral or no-essential (NE). When a gene had less than TA sites was also indicated.

Vv locus	Annotation	Vc locus	Vc Artist	Vp locus	Vp Artist
VV0010	Uncharacterized protein	-	-	VP0010	DE (<10TAs)
VV0014	DNA gyrase subunit B (EC 5.99.1.3)	duplicated	-	VP0012	E
VV0165	Uncharacterized protein	-	-	VP0096	E
VV0391	50S ribosomal protein L18	-	-	VP0273	E
VV0910	LPS-assembly lipoprotein LptE	-	-	VP0726	E
VV1147	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.-)	duplicated	-	VP2157	E
VV1187	Electron transport complex subunit E	duplicated	-	VP2107	E
VV1190	Electron transport complex subunit C	-	-	VP2104	DE
VV1192	Electron transport complex subunit A	duplicated	-	VP2102	E
VV1257	Uncharacterized protein	-	-	VP1922	DE
VV1661	Uncharacterized protein	-	-	VP1591	E
VV1670	Uncharacterized protein	-	-	VP1530	DE
VV1768	Diadenosine tetraphosphate hydrolase	-	-	VP2571	E
VV1966	Integration host factor subunit alpha (IHF-alpha)	duplicated	-	VP1294	E
VV2357	Lipid A export ATP-binding/permease protein MsbA (EC 3.6.3.-)	duplicated	-	VP0982	E
VV2708	Translation initiation factor IF-2	-	-	VP2456	E
VV2727	Iron-sulfur cluster insertion protein ErpA	duplicated	-	VP2474	E
VV2762	78-dihydro-6-hydroxymethylpterin-pyrophosphokinase	duplicated	-	VP2505	E
VV2845	VisC protein	duplicated	-	VP2587	E
VV2953	Single-stranded DNA-binding protein (SSB)	-	-	VP2709	E
VV2982	30S ribosomal protein S18	-	-	VP2738	E
VV2984	30S ribosomal protein S6	-	-	VP2740	E
VV3125	Uncharacterized protein	-	-	VP2880	E
VV3160	50S ribosomal protein L7/L12	-	-	VP2923	E
VVA0360	DNA and RNA helicase	duplicated	-	VPA1468	E
VVA0590	Uncharacterized protein	-	-	VPA0719	DE
VVA0597	Flp pilus assembly protein TadB	-	-	VPA0725	DE
VVA0598	Flp pilus assembly protein TadC	-	-	VPA0726	DE
VVA0978	Uncharacterized protein conserved in bacteria	-	-	VPA1026	DE
VV0436	Flavin prenyltransferase UbiX (EC 2.5.1.129)	VC2541	NE	VP0574	E
VV0501	Transcriptional regulator of sugar metabolism	VC0486	NE	VP0479	E
VV0604	Penicillin-binding protein activator LpoA	VC0581	NE	VP0449	E
VV0913	Apolipoprotein N-acyltransferase (EC 2.3.1.-)	VC0958	NE	VP0729	E
VV1386	ABC-type antimicrobial peptide transport	VC1684	NE	VP1167	DE

Vv locus	Annotation	Vc	Vc Artist	Vp	Vp Artist
	system, ATPase component				
VV1550	Acylphosphatase (EC 3.6.1.7) (Acylphosphate phosphohydrolase)	*VC1355*	NE	*VP1619*	E
VV2218	Uncharacterized protein	*VC1743*	NE	*VP0376*	DE
VV2284	Holliday junction ATP-dependent DNA helicase RuvB (EC 3.6.4.12)	*VC1845*	NE	*VP1052*	DE
VV2298	Chromosome partition protein MukB (Structural maintenance of chromosome-related protein)	*VC1714*	NE	*VP0974*	E
VV2391	Bifunctional protein FolD	*VC1942*	NE	*VP0879*	E
VV2443	Uncharacterized protein	*VC1044*	NE	*VP2586*	E
VV2559	Bifunctional uridylyltransferase/uridylyl-removing enzyme (UTase/UR) (EC 2.7.7.59)	*VC2262*	NE	*VP0735*	E
VV2613	RecBCD enzyme subunit RecD (EC 3.1.11.5) (Exonuclease V subunit RecD)	*VC2319*	NE	*VP2373*	E
VV2614	RecBCD enzyme subunit RecB (EC 3.1.11.5) (Exonuclease V subunit RecB)	*VC2320*	NE	*VP2374*	E
VV2615	RecBCD enzyme subunit RecC (EC 3.1.11.5) (Exonuclease V subunit RecC)	*VC2322*	NE	*VP2375*	E
VV2876	Predicted integral membrane protein	*VC0459*	NE	*VP2987*	DE
VV2995	Penicillin-binding protein 1A	*VC2635*	NE	*VP2751*	E
VV3036	Uncharacterized protein	*VC0355*	NE	*VP2777*	DE
VV3050	cAMP-binding protein	*VC2614*	NE	*VP2793*	E
VV3134	3-dehydroquinate dehydratase (EC 4.2.1.10) (Type II DHQase)	*VC0297*	NE	*VP2879*	DE
VVA0072	Predicted membrane protein	*VCA0040*	NE	*VPA1624*	E
VVA0678	Uncharacterized protein	*VCA0874*	NE	*VP1888*	E
VVA0809	Response regulator	*VCA0256*	NE	*VPA0919*	E
VVA1080	Deoxycytidylate deaminase	*VCA0840*	NE	*VPA1192*	DE

Table 12. List of essential genes for *V. vulnificus* (Vv) and *V. cholerae* (Vc)but not for *V. parahaemolyticus* (Vp). Essentialgenes for Vv were blasted against Vp and Vc and the results were plotted in the table. Genes with homology were indicated with the respective gene name and its EL-ARTIST category assigned in its study. These are: essential (E), domain essential (DE) and neutral or no-essential (NE). When a gene had less than TA sites was also indicated.

Vv locus	Annotation	Vc	Vc Artist	Vp	Vp Artist
VV0452	Low specificity phosphatase	*VC2525*	DE	-	-
VV0797	Predicted transcriptional regulator	*VC0763*	E	-	-
VV1657	Thiol-disulfide isomerase and thioredoxin	*VC1246*	E	-	-
VV1944	Translation initiation factor IF-3	*VCA0288*	E	-	-
VV3135	Biotin carboxyl carrier protein	*VC2651*	DE	-	-
VV3228	Potassium uptake protein TrkA	*VC0296*	DE	-	-
VV0281	Lipid A biosynthesis lauroyltransferase (EC 2.3.1.241)	*VC0213*	E	duplicated	-
VV0588	DNA topoisomerase 4 subunit B (EC 5.99.1.3)	*VC2431*	E	duplicated	-

VV0615	UDP-N-acetylmuramate--L-alanine ligase (EC 6.3.2.8)	VC2400	E	duplicated	-
VV0905	Rod shape-determining protein RodA	VC0949	E	duplicated	-
VV2424	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta (EC 6.4.1.2)	VC1000	DE	duplicated	-
VV3136	Biotin carboxylase	VC0295	DE	duplicated	-
VV0462	1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)	VC2513	DE	VP0358	NE
VV0765	Uncharacterized protein	VC0751	E	VP0659	NE
VV0867	Phosphatidylglycerophosphatase A (EC 3.1.3.27)	VC0856	DE	VP0685	NE
VV0916	Phosphate starvation-inducible protein PhoH	VC0956	E	VP0732	NE
VV0981	Cell division protein ZipA	VC2214	E	VP0799	NE
VV1019	Negative modulator of initiation of replication	VC2096	E	VP0838	NE
VV1451	Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT)	VC1159	DE	VP1247	NE
VV1453	Transport ATP-binding protein CydC	VC1180	E	VP1249	NE
VV1454	Transport ATP-binding protein CydD	VC1181	DE	VP1250	NE
VV1191	Electron transport complex subunit B	VC1014	E	VP1448	NE
VV2709	Transcription termination/antitermination protein NusA	VC2329	E	VP1536	NE
VV1662	3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase (EC 4.2.1.59)	VC1483	E	VP1589	NE
VV1655	Cell division protein ZapC	VC1246	E	VP1596	NE
VV1651	Uncharacterized protein	VC1246	E	VP1603	NE
VV1553	Integral membrane protein	VC1246	E	VP1624	NE
VV3058	Tryptophan--tRNA ligase (EC 6.1.1.2) (TrpRS)	VC2623	E	VP2804	NE
VV3102	Elongation factor P (EF-P)	VC2651	DE	VP2845	NE
VV0178	Uncharacterized protein	VC0084	DE	VP3007	NE
VV1668	Signal transduction histidine kinase	VC1483	E	VPA0009	NE
VV0479	4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262)	VC0444	DE	VPA1460	NE
VV0896	Trehalose-6-phosphate hydrolase	VC0890	E	VPA1641	NE

Using Uniprot database, we classified *Vibrio-essential*-genes in functional categories (Gene Ontology) and compared them to the whole genomes of the three strains (**Figure 60**). Uniprot classifies the genes in three categories: biological process, cellular part and molecular function. The *Vibrio-essential-genes* clearly dominated in i) *signal transduction*, ii) *maintenance of stationary phase*, iii) *developmental*, iv) *reproduction,* and v) *Ftsz-dependent cytokinesis* processes; dominated in i) *intracellular organelle*, ii) *organelle* and iii)
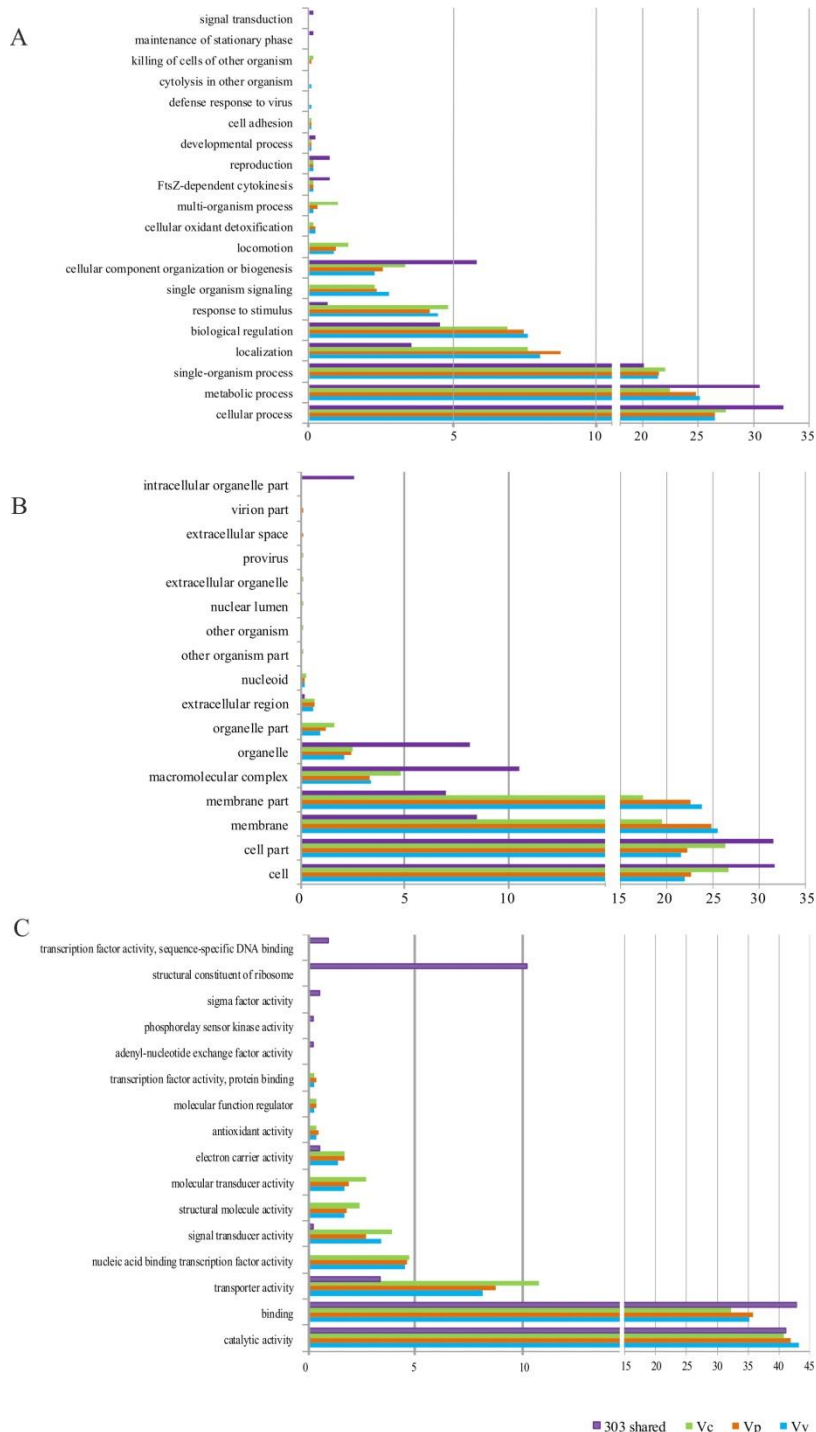
Figure 60. Comparison of Gene Ontology categories between complete genomes and genes essential to grow in LB. *V. vulnificus* (Vv)*, V. parahaemolyticus* (Vp) and *V. cholerae* (Vc) genes in the complete genomes and the genes essential in all of them to grow in LB (303 shared) were classified in Gene ontology categories using Uniprot. The percentage each category occupies in the corresponding group is represented in bars. A, biological process; B, cell part; C, molecular function.

macromolecular complex cellular-parts; and, finally dominated in i) *transcription factor*, *sequence specific DNA binding*, ii) *sigma factor,* iii) *phosphorelay sensor kinase,* and *iv) adenyl-nucleotide exchange factor* molecular-activities. Interestingly we also detected specific-*V. vulnificus essential* genes in the biological-processes, i) cytolysis in other organisms and ii) defence response to virus (**Figure 60**)

5.4. *Essential* genes to grow in HS and comparison to other studies

In order to unravel the genes of *V. vulnificus* responsible for resistance to complement killing we created two independent libraries. Approximately 10^8 mutants from *V. vulnificus* YJSm^r library were incubated, by duplicated, during 2h at 37ºC in IHS and HS and, then plated onto LBA plates. As we already mentioned, HS is bactericidal due to complement and bacteriostatic due to nutritional immunity (see introduction 2.2.1). Theoretically, we could discriminate the *essential* genes to resist complement from those *essential* to overcome nutritional immunity by comparing the mutants recovered after growing the libraries in IHS *vs* HS and ISH *vs* LB, respectively. While the EL-ARTIST compared the number of TA positions covered in the library grown in LB *vs* the number that putatively could have been disrupted, here, we compared the TA positions in two libraries grown in different conditions. Thus, we considered a gene as *essential* to resist human complement if it was under-represented in HS *vs* IHS, and a gene as *essential* to resist the nutritional immunity if it was under-represented in IHS *vs* LB.

Sequencing efforts were made to obtain around 6 million sequences per sample. The percentage of TA sites covered decreased from 75.9% in LB to 31% on average in HS and IHS. This could be due to sampling bias or an increase in the number of genes without inserted transposon. We assessed the frequency of TA sites sequenced per gene and compared them to the profile distribution in LB. The LB-profile in chromosome I and II was similar to the ones obtained by Pritchard *et al*. (2014) in*V.cholerae.* However, the frequency of genes with higher number of sequenced TA (*neutral*) was reduced and that of the genes with 0-0.05 TA sites (*under-represented*) was increased in serum. In addition, the

profile distribution changed in HS and IHS libraries and a leftward shift was observed, which suggested a bottleneck (**Figure61**). Despite this, the comparison of TA sites disrupted in HS *vs* IHS and in IHS *vs* LB resolved a list of 13 *under-represented* genes in HS (supposedly involved in resistance to complement) and 85 *under-represented* genes in IHS (supposedly involved in resistance to nutritional immunity), respectively (**Table 13**).
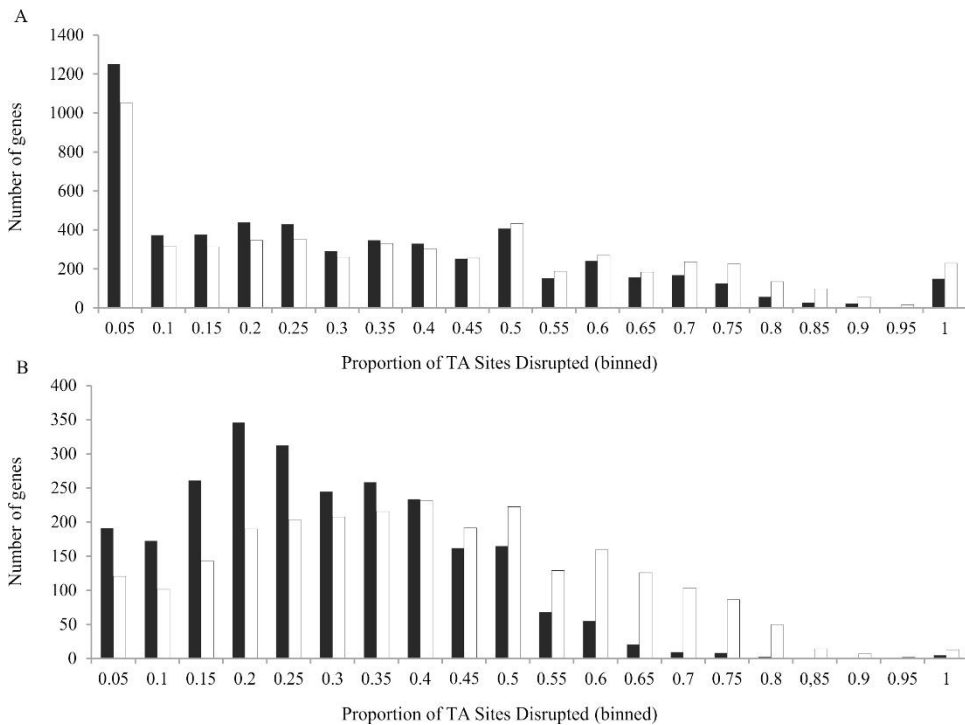
Figure 61. Growth in HS *vs* IHS. TA percentage distribution in chromosome I (A) and II (B) when libraries grew in HS (red) and IHS (green). The number of reads that had a determined percentage of TA sites disrupted is represented in bars.

Table 13. List of genes essential to resist complement killing in HS. The genes listed have been categorized as essential to grow on active HS. The presence and category in *E. coli* ST131 (Ec) or *M.catarrhalis* BBH18 (Mc) is indicated.

	Gene name	NCBI accession	Ec or Mc locus	Essential in Ec or Mc	Description
Chr1	*VV0337*	WP_043877362.1	Ec	-	*wza*, outer membrane protein
	VV0340	WP_011149313.1	Ec	-	*wzc*, tyrosine protein kinase
	VV0341	WP_043877039.1	Ec	-	*wecB*, UDP-N-acetylglucosamine 2-epimerase
	VV0342	WP_011149315.1	Ec	-	*wecC*, UDP-N-acetyl-D-mannosamine dehydrogenase
	VV0352	WP_011149325.1	Ec	-	mannose-1-phosphate guanylyltransferase
	VV0353	WP_011149326.1	-	-	phosphomannomutase
	VV0355	WP_011149328.1	-	-	imidazole glycerol phosphate synthase subunit *HisH*
	VV0357	WP_011149330.1	-	-	dehydrogenase
	VV0358	WP_011149331.1	-	-	hypothetical protein
	*VV0360**	WP_011149333.1	-	-	glycosyltransferase
	VV0364	WP_011149337.1	-	-	*capD*, nucleoside-diphosphate sugar epimerase
Chr2	*VVA0762*	WP_011152092.1	-	-	cytochrome C biogenesis/ divalent-cation tolerance protein CutA
	VVA0954	WP_011152252.1	-	-	anthranilate phosphoribosyltransferase (glycosyltransferase)

* p-value >10^5

5.4.1. *Essential* genes to resist complement in HS

The main humoral anti-bacterial effector present in serum is the complement, a series of proteins secreted by the liver as inactivated pro-proteins that are activated in cascade after bacterial recognition resulting in its destruction by lysis. As anyone would expect, bacteria have developed different mechanisms to avoid this recognition. In particular, some of them produce an external structure commonly known as capsule or CPS which improves the protection that the lipopolysaccharide (LPS) confers (see Introduction 2.2.2). The list of the genes found to be essential to resist complement killing is showed in **Table 13**. None of these genes was in the plasmid pYJ016. We could identify in the chromosome I

11 genes most of them presumptively involved in O-antigen/CPS biosynthesis and transport in *V. vulnificus*. Among them, *wza* (*VV0337*), *wzc* (*VV0340*), *wecB* (*VV0341*), *wecC* (*VV0342*) and *mannose-1-phosphate guanylyltransferase (VV0352)* have been demonstrated to have a role in biosynthesis/transport of O-antigen and/or CPS in *E. coli* and/or *Vibrio*(Dong *et al*., 2006; Chen *et al*., 2007; Stroeher *et al*., 1997; Meier-Dieter *et al*., 1990). With regard to *wzb* (*VV0339*), the cognate phosphatase of *wzc*, did not pass the filters in our assay, although it was very close (p-value 0.007, fold change 3.15). The small size of this gene could explain its exclusion from the list of *under-represented* genes. The function of *wza* in CPS/LPS biosynthesis in *V. vulnificus* have been already demonstrated (Fong *et al*., 2010; Chatzidaki-livanis *et al*., 2006).The rest of the identified genes could also be involved in LPS/CPS biosynthesis as they were predicted to participate in polysaccharide biosynthesis. In fact, all these genes were located in the same locus (**Figure 62**).
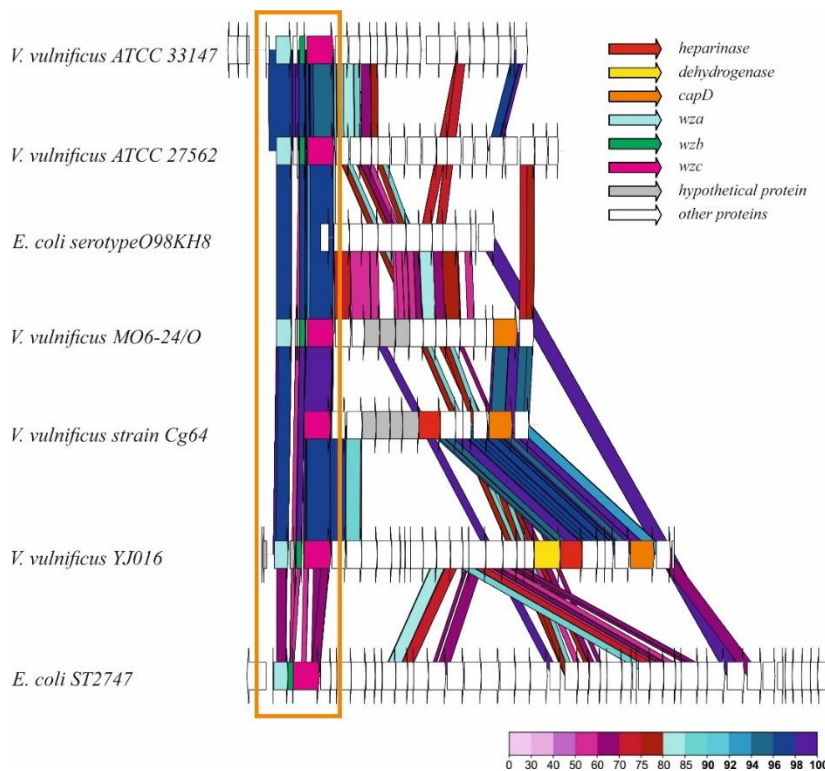
Figure 62. **Plot of CPS locus alignment from *V. vulnificus* and *E. coli*.** The top hits when *V. vulnificus* YJ016 CPS locus was blast against NCBI database were aligned using TBLASTX. Minimum identity and length of 50 and 100 was used for the alignment, respectively. The identity of the alignment is represented by a color code.

To validate our results obtained by TIS we selected one of the genes whose function in CPS production in *V. vulnificus* had been demonstrated, *wza* (Wright *et al*., 2001)and obtained a deletion mutant (Δ*wza* or Δ*VV0337*). The *wza*, *wzb* and *wzc* genes are highly conserved in *Vibrio* and also in Enterobacteria and the specific function of the three proteins, the translocation of CPS type 1 polysaccharides, has been demonstrated in *E. coli*(Wright *et al*., 2001; Drummelsmith and Whitfield, 2000).We found that the colonial morphology of Δ*VV0337* on agar plates was translucent while that of the wild-type strain was opaque (**Figure63**). It is well documented in *V. vulnificus* that encapsulated cells grow on agar plates as opaque colonies while unencapsulated or less-encapsulated grow as translucent colonies (Wright *et al*., 1990). Then, we extracted and immunostained cell's polysaccharides and found that the mutant, effectively, lacked CPS (**Figure 64**).
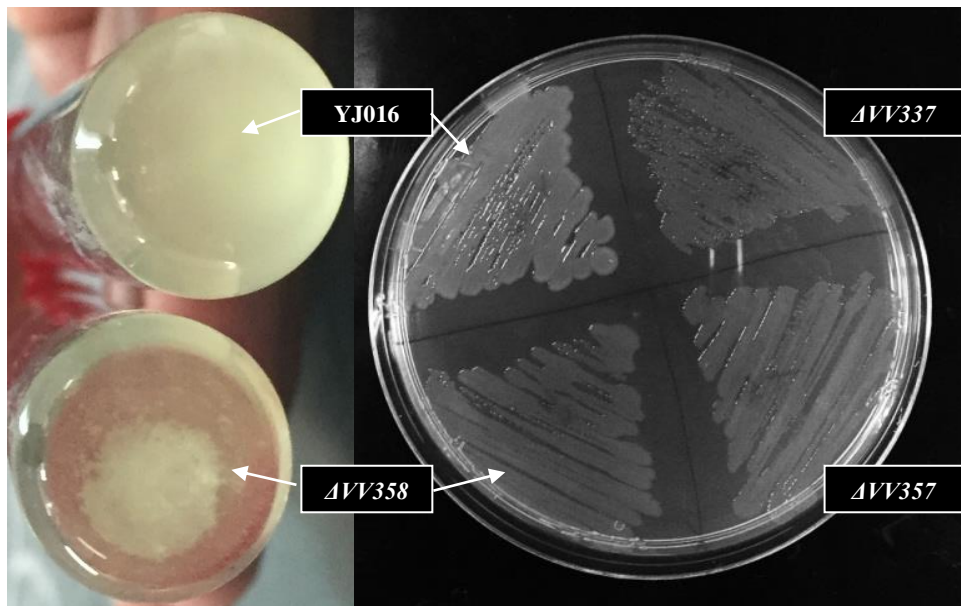
Figure 63. Phenotypes of YJ016 and their derivative mutants growing on culture media. Image in color shows the aspect of the bottom of LB tubes after 18 h of incubation at 37ºC. Notice that mutant cells aggregated in the bottom. Image in B&W shows the aspect on agar plates. Notice that the wild-type phenotype was opaque and the mutant-phenotype was translucent. Picture was taken using camera of iPhone 6.

Figure 64. CPS detection by immunostaining. Bacteria were grown in LB, polysaccharides were extracted with the protocol of Hichcok & Brown (1981), separated by SDS-PAGE, transferred to NC-sheets and immunostained with specific antibodies against YJ016.
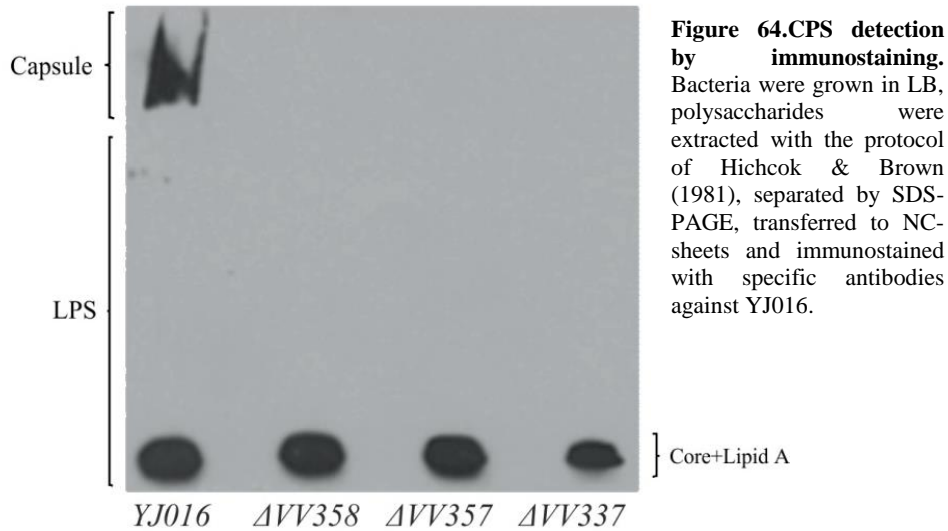
Further, we did not detect significant differences in the quantity of cell-polysaccharides between the mutant and the wild-type strain, which confirmed its role proposed for Wza_{Vv} to transport polysaccharides to the extracellular medium. Finally, the mutant was sensitive to human complement, which validated the results obtained using TIS (**Figure 65**).
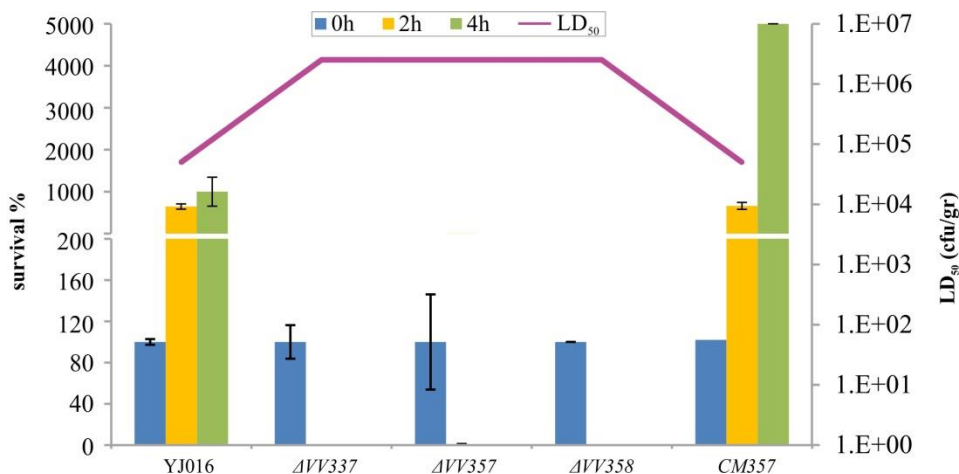
Figure 65. Bacterial growth in HS and virulence for mice (LD_{50} or doses lethal 50%) of *V. vulnificus* YJ06, deletion mutants and complemented. Bacteria were grown in LB and then inoculated in HS for 4 hours. Changes in bacterial population (cfu/ml) were measured after 0, 2 and 4 hours of serum incubation. Mice were infected as described in Material and Methods 4.10. and LD_{50} was calculated and expressed as cfu per gr of mice (BALB/c).

Interestingly, 4 additional genes in the list (**Table 13 and 14**)were not previously related with CPS or LPS biosynthesis: a *dehydrogenase* (VV0357), a *hypothetical protein* (VV0358) (**Figure 66**), a *divalent-cation tolerance protein CutA* (VVA0762) and a *glycosyltransferase* (VVA0954) (**Figure 67**). Then, we get deleted mutants in each one of these four genes. Surprisingly, Δ*VV357* and Δ*VV358*showedtranslucent colonies on agar plates similar to those of Δ*wza$_{Vv}$* (**Figure 63**). Encapsulated *V. vulnificus* cells can be additionally distinguished from unencapsulated (or less encapsulated ones) because they can aggregate in the bottom of the culture tubes (Phippen and Oliver, 2015). This aggregation was seen in the bottom of the tubes of the three mutants: Δ*wza$_{Vv}$*, Δ*VV0357* and Δ*VV0358* (**Figure 63**). Further, no CPS was observed in the polysaccharide's extracts from Δ*VV0357* and Δ*VV0358* after electrophoresis, transference to NC-filters and immunostaining (**Figure 64**). In consequence, deletion of each one of both genes produced a bacterium without CPS on its outer surface, a phenotype similar to that associated with a loss or a deficiency in CPS biosynthesis and/or transport. Not surprisingly, *VV0357* and *VV0358* genes were also included in the capsule locus (**Figure 62**). Then, we grew both mutants in HS and found them as sensitive to the bactericidal effect of complement as Δ*wza$_{Vv}$* (**Figure 65**). Further, we analysed the transcription of the three genes in LB and HS and found that they were constitutively expressed without significant differences among them (**Figure 68**).
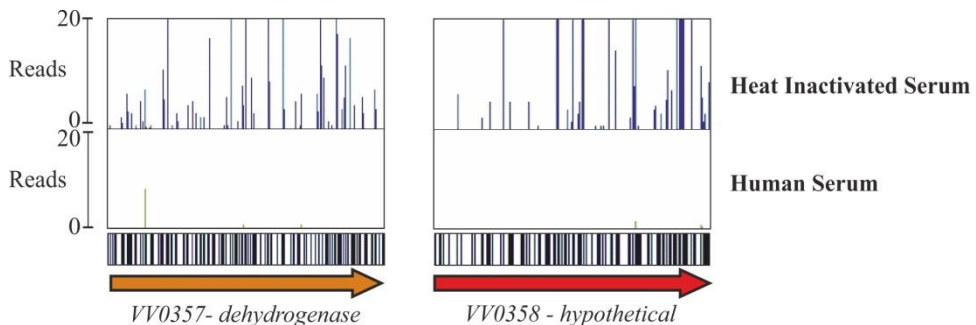
Figure 66. Transposon-insertion profiles of VV0357 and VV0358. The profiles obtained for two *essential* genes for *V. vulnificus* to resist complement killing is showed. The number of reads mapped to a TA site along the horizontal axis for the heat inactivated and activated HS library are indicated in the vertical axis.
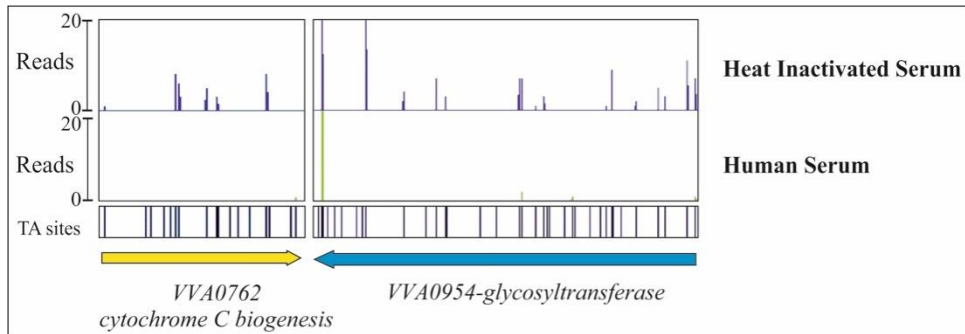
Figure 67. Reads mapped in the two genes from chromosome 2putatively *essential* for resistance to complement killing. The number of reads mapped to a TA site along the horizontal axis for the heat inactivated and activated HS library are indicated in the vertical axis. TA positions in each gene sequence are represented over the arrow with black lines.

Table 14. Fold change and p-values for genes putatively *essential* for serum survival that were apparently no related to CPS or LPS.

Locus	Avg p-value	FC
VV0357	7.5E-09	30.31
VV0358	2.9E-06	67.61
VVA0762	8.6E-04	50.08
VVA0954	1.2E-03	10.18
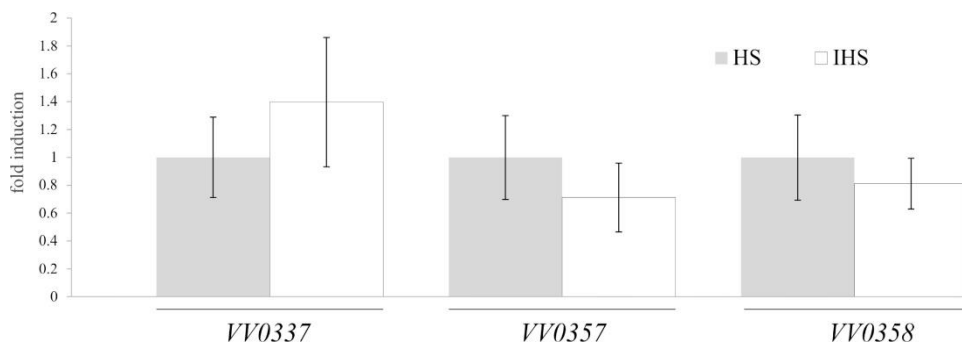
Figure 68. Expression level of genes essential to grow in HS. The differences between *VV0337* (*wza*), *VV0357* (*hydrogenase*)and *VV0358* (*hypothetical protein*) were calculated using RT-PCR. *recA* was used as housekeeping gene. The bars represent the standard deviation for three independent replicates.

Regarding the genes located in chromosome II (*VVA0762* and *VVA0954*), deletion of each one of both genes did not change either colony phenotype or growth in HS (data not shown).

174

Finally, we tested the virulence for mice of the three translucent mutants and the wild type strain and found a 100 times decrease in the LD_{50} for each one of the mutants, confirming the role of CPS in virulence. The complementation of $\Delta VV0357$ (CM357) reversed the phenotype and even increased the survival rate in HS in comparison to the wild type strain (**Figure 65**).

We analysed *in silico* the proteins VV0357, annotated as a putative dehydrogenase, and VV0358, annotated as a hypothetical protein. VV0357 was predicted to have a transmembrane domain which suggests a putative location in the membrane of the bacterium. VV0358 hit (100% identity) a protein annotated as a heparinase by using BLASTP against NR database. Further, a 100% of confidence against heparinase III protein (lyase) was retrieved by Phyre2 analysis. *heparinase*$_{Vv}$ was not present in other *V. vulnificus* genomes although it showed high similarity with other putative heparinise genes present in other *Vibrio* species, the closest one belonging to *V. parahaemolyticus*. By using CDD analysis (Marchler-Bauer *et al*., 2014), we detected two putative domains in most of the vibrio-heparinases: a heparinase II/III-like domain (pfam07940) and a alginate lyase domain (cl00179). Alginate and heparin sulphate lyases are involved in glycosaminoglycan degradation (GAGs). GAGs are highly negatively charged polysaccharides, formed from disaccharide repeating units. GAG lyases have been proposed to be used to degrade and utilize glycosaminoglycans as a source of carbon in the bacterium's natural environment(Wong *et al*., 2000). Heparinise III was initially described in *F. heparinum* (formerly *Pedobacter*) as a cytoplasmic protein which is secreted to periplasmic space upon induction with heparin. This protein is an enzyme that cleaves the glycosidic linkage between amino sugars and uronic acids in heparin and heparan sulfate (Dong *et al*., 2012; Su *et al*., 1996). No signal peptide for exportation or for membrane protein was detected in the putative *heparinase*$_{Vv}$ (Driessen and Nouwen, 2008). We tried to find out if *V. vulnificus* YJ016 was able to degrade heparin using three different protocols (toluidine blue staining, blood coagulation and growth in minimal medium with heparin as the sole carbon source) but we obtained negative results. More studies are needed to

determine the role of this protein in CPS biosynthesis and virulence in *V. vulnificus*.

5.5. Essential genes to overcome nutritional immunity in serum

V. vulnificus is an invasive pathogen able to grow rapidly in blood and cause death by sepsis in less than 24 hours (Jones and Oliver, 2009). Once the bacterium has spread to bloodstream, it needs not only to resist the bactericidal action of complement but also to overcome the nutritional immunity in order to multiply and cause disease. The term "nutritional immunity" includes all the strategies displayed by the host in order to make nutrients inaccessible for bacterial growth. Nutrients would include from trace elements like iron, zinc and manganese (usually sequestered by proteins) up to limiting carbon and/or nitrogen sources. In order to discover what nutrients "are sensed" as limiting in serum and what genes are involved in overcoming this limitation in *V. vulnificus*, we compared the libraries grown in IHS and LB. Con-ARTIST analysis defined 85 genes *under-represented* in IHS (**Table 15**). In this list, genes more than 50 times under-represented in IHS library but with a p-value <0.001 were included. These genes would be putatively essential to grow in serum but not in LB. Considering this, it is possible that we would miss genes necessaries to grow in both, LB and serum.

Table 15. List of essential genes for HS growth. The table represents all the genes that TIS analysis considered essential to grow on HS. The p-value and the fold change (FC) for each locus is in the table together with the annotation. Chromosome 1 (VV); Chromosome 2 (VVA).

LOCUS	p-value	FC	Annotation
VV0002	1.06E-06	18.27	tRNA modification GTPase MnmE
VV0040	8.58E-10	14.67	ATP-dependent DNA helicase Rep
VV0122	6.24E-11	14.57	glycerol-3-phosphate 1-O-acyltransferase
VV0214	1.21E-20	90.49	type II secretory pathway, component EpsD
VV0215	3.75E-09	20.45	type II secretory pathway, ATPase EpsE
VV0216	1.92E-08	33.63	type II secretory pathway, component EpsF
VV0221	2.44E-10	53.44	type II secretory pathway, component EpsK
VV0327	2.90E-08	5.75	ADP-L-glycero-D-mannoheptose-6-epimerase
VV0349	2.04E-11	19.31	GDP-mannose 4,6-dehydratase
VV0363	4.21E-10	20.54	aminotransferase
*VV0452 **	1.13E-04	51.27	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase
*VV0462 **	1.09E-04	144.2	1-acyl-sn-glycerol-3-phosphate acyltransferase
VV0586	8.17E-06	16.16	3,5-cyclic-nucleotide phosphodiesterase
VV0599	8.73E-07	26.46	stringent starvation protein A
VV0624	1.91E-10	97.71	carbamoylphosphate synthase small subunit

VV0625	5.88E-21	18.98	carbamoylphosphate synthase large subunit
VV0752	2.90E-06	5.56	inositol monophosphate family protein
VV0769	6.13E-19	382	outer membrane protein assembly factor BamB
VV0775	6.78E-11	21.37	inosine-5-monophosphate dehydrogenase
VV0776	1.13E-12	5.88	GMP synthase [glutamine-hydrolyzing]
VV0821	2.40E-06	97.59	tmRNA-binding protein
VV0977	2.81E-10	7.82	phosphoenolpyruvate-protein phosphotransferase
VV1030	2.25E-08	5.06	succinate dehydrogenase flavoprotein subunit
VV1032	1.99E-13	9.22	2-oxoglutarate dehydrogenase E1
VV1034	1.40E-06	6.96	succinyl-CoA ligase subunit beta
VV1104	1.10E-05	5.50	ATP-dependent Clp protease proteolytic subunit
VV1105	2.62E-06	6.09	ATP-dependent Clp protease ATP-binding subunit ClpX
VV1106	1.60E-05	71.10	ATP-dependent Lon protease, bacterial type
VV1262	3.13E-08	5.55	cell envelope biogenesis protein AsmA
VV1272	2.40E-06	60.21	fatty acid/phospholipid synthesis protein PlsX
VV1321 *	1.73E-04	135.6	AsnC family transcriptional regulator
VV1441	5.78E-06	5.02	phosphogluconate repressor HexR, RpiR family
VV1472 *	2.32E-05	86.91	exodeoxyribonuclease I
VV1542	4.90E-09	6.34	diguanylate cyclase
VV1652	1.76E-15	11.38	NAD-glutamate dehydrogenase
VV1659	5.95E-21	1644	membrane protein
VV1667	2.00E-06	10.85	Putative signal transduction protein; structural genomics
VV1669	9.64E-06	6.26	histidine kinase
VV1671	3.13E-10	8.15	cytochrome c oxidase, cbb3-type subunit I
VV1674	7.28E-09	14.34	cytochrome CBB3
VV1675 *	3.00E-04	50.97	putative analog of CcoH
VV1676	1.16E-10	10.50	ATPase P
VV1678	8.77E-07	17.75	cytochrome biogenesis protein
VV1681	1.51E-05	11.38	tRNA 2-thiocytidine(32) synthetase TtcA
VV1689	1.89E-06	8.55	putrescine/spermidine ABC transporter substrate-binding protein
VV1690	5.28E-06	10.56	spermidine/putrescine ABC transporter substrate-binding protein
VV1693	3.90E-08	11.39	chromosome segregation ATPase
VV1696	4.05E-10	20.03	hypothetical protein
VV1727	2.97E-06	5.85	Fructosamine kinase
VV1736	4.65E-06	7.60	hypothetical protein
VV1745	1.16E-05	7.15	hypothetical protein
VV1757	3.91E-08	5.82	adenylate and Guanylate cyclase catalytic domain protein
VV1764	8.75E-06	8.92	hypothetical protein
VV1770	9.81E-09	31.67	Predicted ATPase
VV1783	6.88E-07	6.21	hypothetical protein
VV1787	2.56E-08	10.43	hypothetical protein
VV1829 *	7.07E-05	212.4	acetyltransferase
VV1840	7.04E-07	5.58	transposase
VV1872	1.67E-05	14.82	hypothetical protein
VV1921	1.53E-08	7.59	microtubule-binding protein
VV1922	5.99E-06	8.60	hypothetical protein
VV1924	5.22E-06	10.76	hypothetical protein
VV2079	1.15E-05	11.97	fumarate hydratase, class I
VV2150	1.09E-09	25.49	putative Bacteriophage CI repressor helix-turn-helix domain
VV2274	6.62E-08	9.13	lipoprotein
VV2275	1.46E-12	30.83	protein TolB
VV2318	5.89E-06	12.05	3-phosphoshikimate 1-carboxyvinyltransferase
VV2325	1.38E-14	30.68	isocitrate dehydrogenase
VV2419	1.95E-05	6.76	glutamine phosphoribosylpyrophosphate amidotransferase

VV2429	1.93E-08	7.25	3-oxoacyl-(Acyl-carrier-protein)synthase
VV2548	1.73E-07	343.2	outer membrane protein OmpH
VV2585	9.10E-11	11.12	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrF
VV2587	2.66E-08	16.49	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrD
VV2588	2.52E-08	24.19	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrC
VV2589	8.96E-10	5.03	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrB
VV2590	5.81E-10	20.49	Na+-transporting NADH:ubiquinone oxidoreductase, subunit NqrA
VV2772	2.44E-12	20.22	pyruvate dehydrogenase, E2 component, dihydrolipoamide acetyltransferase
VV2773	9.02E-31	13.01	pyruvate dehydrogenase complex, dehydrogenase component
VV2781	9.83E-13	6.77	prepilin peptidase; type IV prepilin-like proteins leader peptide processing enzyme
VV2794	1.42E-07	6.37	glutamate--cysteine ligase
VV2836	1.93E-08	5.77	negative regulator of sigma E activity
VV3216	1.30E-09	22.04	coproporphyrinogen III oxidase
VV3262	2.56E-11	27.85	glucose inhibited division protein A
VVA0407	4.58E-06	5.21	dihydroorotase
VVA0781	6.93E-06	5.05	heme receptor

*p-value<0.001 and FC>50

Type 2 secretion system (T2SS). Interestingly, genes for subunits D, E, F and K from a T2SS were found in this list (**Figure 69**). Although the genes for the rest of subunits did not pass filtering, they got very close. This result suggests that T2SS could be essential for *V. vulnificus* to grow in HS. T2SS is well-known for its role in secretion of multiple exo-enzymes and toxins involved in pathogenesis (McCoy-Simandle *et al*., 2011; von Tils *et al*., 2012; Sandkvist, 2001; Sánchez and Holmgren, 2008). In fact, T2SS is utilized by *V. cholerae* to export 19 proteins including the cholera toxin (Korotkov *et al*., 2012) and in *V. vulnificus* to secrete two important virulence factors, the protease VvpE and the hemolysin VvhA (Hwang *et al*., 2011).Moreover, studies in opportunistic pathogens such as *Legionella pneumophila*, *V. cholerae* and *Yersinia enterocolitica* have suggested that T2SSs have an important function in both virulence for humans and survival in the environment (Söderberg *et al*., 2004; Cianciotto, 2009; McCoy-Simandle *et al*., 2011; Kirn *et al*., 2005; Stauder *et al*., 2012; Wong *et al*., 2012).None of the enzymes included in this list was a putative candidate to be secreted by this system. We hypothesized T2SS could export hydrolytic enzymes (such as VvpE) to serum that would allow bacteria to acquire nutrients from macromolecules

such as proteins, lipoproteins, lipids, glucolipids etc. Further studies are needed to prove it.
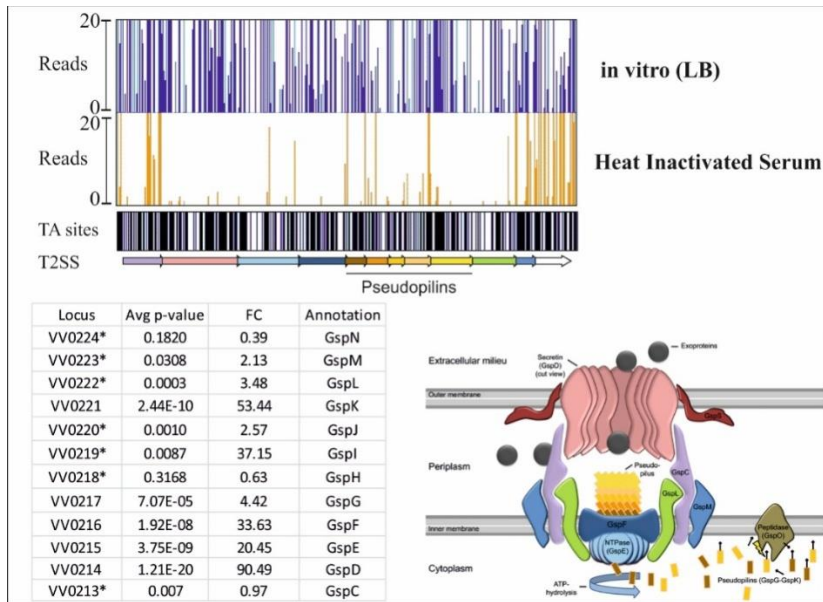
Figure 69. Type 2 secretion system (T2SS) in transposon libraries. The number of reads mapped in the TA positions in the T2SS locus is represented in the top figure. The different p-values, fold change (FC) and annotation name for each gene is represented in the table while its organization in the membrane is in the right. Disposition in the membrane is copied from von Tils *et al.* study on Yersinia. (von Tils *et al.*, 2012)
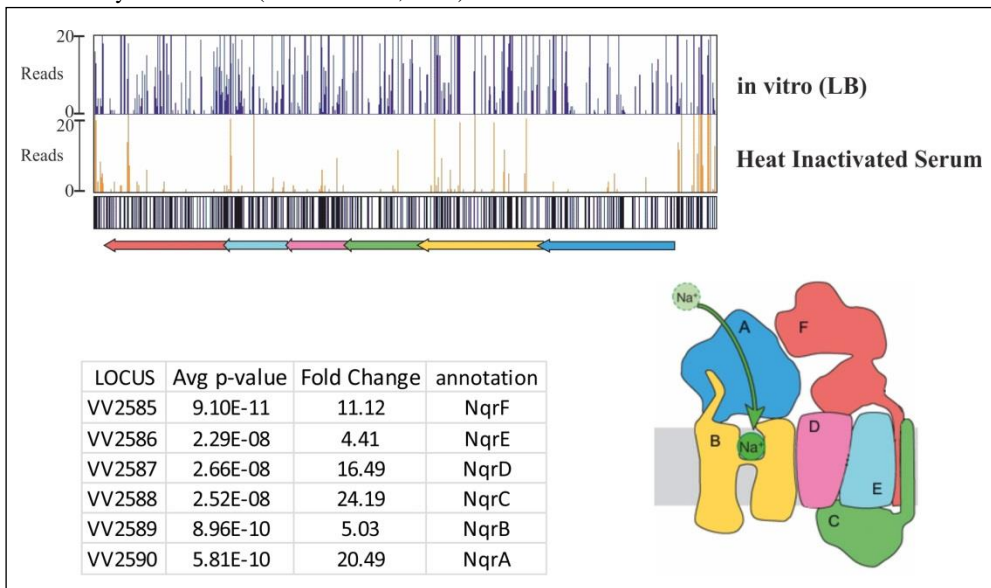
Figure 70. Na+-translocating NADH-ubiquinone reductase complex operon *in vitro* and heat inactivated HS. The abundance of transposon in the different loci in the genome is represented in the top figure. The genes forming the complex are in the table and in the representation obtained from Julia Steuber *et al.* (Steuber *et al.*, 2014).

*Na+-translocating NADH-ubiquinone reductase complex(NQR).*NQR seemed to be also necessary to grow in serum but not in LB(**Figure 70**). This complex is a membrane pump widespread among pathogenic and non-pathogenic bacteria that consists of six subunits (NqrA, B, C, D, E and F) responsible for building up an electrochemical gradient in environments rich in sodium (Steuber *et al*., 2014; Hase *et al*., 2001; Zhou *et al*., 1999). The gradient can be used as an energetic supply for Na+/H+ antiporters, Na+-dependent flagellar motors, Na+-dependent multidrug efflux pumps, as well as Na+ symporters for alanine, glutamate, proline, serine, citrate, and inorganic phosphate(Steuber, 2001; Rich *et al*., 1995; Unemoto and Hayashi, 1993; Dibrov, 2005). In fact, this Na+ gradient drives substrate uptake, motility, and efflux of antibiotics in *V. cholerae* (Steuber *et al*., 2014). *V. vulnificus* is a brackish water bacterium and its sodium requirements for an optimal growth should be the same than that described for other marine vibrios, which use the NQR-system for active uptake of nutrients (Unemoto and Hayashi, 1993). Sodium content in LB used in this work was in the range of normal sodium content in blood, which apparently did not match with our results. However, HS also contains heat-resistant peptides, which can form channels on the bacterial outer and inner membranes (Skulachev, 1998). Apparently, the protein-motive force is disrupted in this condition and thanks to NQR complex bacteria would be able to grow in an hostile medium like HS (Skulachev, 1998).

Transport system for spermidine/putrescine. Two genes from the *potABC2D* locus involved in polyamine transport were categorized as under-represented in serum (**Table 15** and **16**). While the genes for two different subunits PotD were included in the list, the genes for subunits PotA, PotB and PotC as well as for VV1688 (a hypothetical protein that is not clear if it is part of the locus or not) had p- and fold change values out of the filters. According to our results, the two periplasmic subunits (*potD*) could be considered as essential for *V. vulnificus* to overcome nutritional immunity. Our study correlates with the transcriptomic study performed by Williams *et al*., (2014) who also found that *potABCD* was up-regulated in *V. vulnificus* when growing in HS. Recent experimental evidence has highlighted the importance of polyamines in the pathogenesis of bacteria, in particular in biofilm formation, escape from phagolysosomes, bacteriocin

production, toxin activity and protection from oxidative and acid stress (Shah and Swiatlo, 2008). Interestingly, vulnibactin, the siderophore produced by *V. vulnificus* under iron-starvation contains norspermidine in the core of the molecule (Okujo *et al*., 1994; Keating *et al*., 2000). We proposed that these polyamines taken from serum could be used as a central building block for siderophore biosynthesis by *V. vulnificus in vivo*.

Table 16. Locus involved in spermidine and putrescine translocation system

LOCUS	Annotation	Avg p-value	FC
VV1685	*potA*	1.7E-04	3.81
VV1686	*potB*	5.8E-04	2.91
VV1687	*potC*	1.2E-02	2.31
VV1688	Uncharacterized protein	4.7E-01	20.43
VV1689	*potD*	1.9E-06	8.55
VV1690	*potD*	5.3E-06	10.56

Heme receptor: A heme receptor (VVA0781) located in chromosome 2 and classified inside TonB-dependent heme/haemoglobin receptor family was present in the list. Iron-starvation forms part of the nutritional immunity against pathogens, in particular, against *V. vulnificus*. *V. vulnificus* expresses multiple systems to overcome iron-limitations, one relays on siderophore biosynthesis concomitant to the biosynthesis of the siderophore ferric-uptake complex and two outer membrane receptors for hemin and haemoglobin (Pajuelo *et al*., 2015; Parrow *et al*., 2013). Our results suggest that, from the two receptors, HupA is the most important to grow in HS from heme compounds, at least during the first two hours of incubation.

Purine and pyrimidine synthesis. 5 genes involved in purine and pyrimidine synthesis were part of the genes essential to grow in HS (**Table 15**). Previous studies have demonstrated the importance of some of these genes in growth in blood of *E. coli, Salmonella enterica* and *Bacillus anthracis*, suggesting that purines and pyrimidines are limiting nutrients for growth in serum (Samant *et al*., 2008).

Genes involved in resistance to stress conditions: We found in the list a series of genes that have been described as involved in resistance to stressing conditions, including resistance to microcidal peptides. We hypothesized that this could be its function in serum.

a) *ompH*. This gene encodes an outer membrane protein that is regulated by iron and has been related to resistance to microcide peptides in *V. vulnificus*(Alice *et al*., 2008).

b) Genes for Clp and Lon proteases. These genes have been related with homeostasis, DNA damage and stress survival and, in pathogenic bacteria, expression of the virulence genes (Lee *et al*., 2006b; Robertson *et al*., 2000). The elimination of oxidatively modified proteins performed by Clp and Lon proteases, is a crucial process in maintaining cellular homeostasis, especially during stress (Ngo *et al*., 2013).

c) Genes involved in membrane biogenesis in response to stressing conditions: genes encoding a negative regulator of sigma E, the cell envelope biogenesis protein AsmA, and the outer membrane protein assembly factor BamB have been related to membrane biogenesis in response to stress (Ahuja *et al*., 2009). Finally, some genes involved in LPS biosynthesis/modification (*glycerol-3-phosphate 1-O-acyltransferase, ADP-L-glycero-D-mannoheptose-6-epimerase, GDP-mannose 4,6-dehydratase, 3-deoxy-D-manno-octulosonate 8-phosphate phosphatase* and *1-acyl-sn-glycerol-3-phosphate acyltransferase)* (**Table 15**). Modifications of LPS have been reported to protect bacteria against the host peptides *in vivo*(Rosenfeld and Shai, 2006).

5.6. *V. vulnificus vs* other bacteria

Poor similarities were found when the list of genes essential to survive in HS was compared to those obtained with TIS in other bacteria(Phan *et al*., 2013; de Vries *et al*., 2014). On one side, none of the essentials genes for serum survival of YJ016 is present in *Moraxella catarrhalis* BBH18 (**Table 13**). In the same way,

none gene considered essential for *M. catarrhalis* to survive in HS was found in our list. From the 53 genes responsible for serum survival for *M. catharralis* BBH18 only 14 had homologous in *V. vulnificus* YJ016. Two of these 14 had two copies in the genome and 5 were essential to grow *in vitro*. None was catalogued as under-represented to grow in HS, either.

Even though 5 of the 12 essential genes for *V. vulnificus* YJ016 had an homologous copy in *E. coli* ST131 none of them was found essential in the Phan *et al* study (**Table13**) (Phan *et al*., 2013; de Vries *et al*., 2014). Since *E. coli* ST131 had homologous to *wza* and *wzc* someone would expect these genes were essential in this bacterium as they were in *V. vulnificus* YJ016. Moreover, 46 genes were listed as essential to *E. coli* ST131 to survive in HS. We found homologous in *V. vulnificus* YJ016for 26 genes, 4 of them with two or more copies and other 4 essentials to grow *in vitro*. Interestingly, 2 genes were categorized as under-represented to grow in serum. LPS biosynthesis (including both O-antigen biosynthesis and lipid A-core biosynthesis) was categorized as essential in *E. coli* ST131 while these genes were essential for LB-growth in *V. vulnificus* YJ016. Nevertheless, this could be due antigenic differences since other genes related to LPS biosynthesis/modification were found as essential for growing in HS in *V. vulnificus* (*wecB*, *wecC*, *mannose-1-phosphate guanylyltransferase*, *phosphomannomutase* and the *glycosyltransferase*).

The low similarities seen in these comparisons could be due not only to different bacterial strategies to grow in HS but also to different protocol. For example, in both studies HS was used diluted and was incubated 30 or 90min while we used undiluted HS and incubated 2 h(Phan *et al*., 2013; de Vries *et al*., 2014).

Finally, the results obtained when we studied the genes essential to grow in HS were compared to a study in *E.coli* K12(Samant *et al*., 2008). Samant *et al* investigate which genes were required for the growth in IHS with the gene-deletion library created for this strain (Keio list)(Baba *et al*., 2006). They found *de novo* purine and pyrimidine biosynthesis were critical pathways for *E. coli* growth in this condition. They inferred that a low concentration of nucleotide precursors is present in HS and consider this as the main barrier for pathogens to

grow in human blood. Similar results were seen in the same study with other two pathogens: *Salmonella enterica* (Gram-negative) and *Bacillus anthracis* (Gram-positive). The deletion mutants of almost all the genes listed reduced the growth rate of the three pathogens in HS. The comparison of the 85 genes obtained in our analysis with the 22 mutants listed by Samant as with potential growth defect in HS yield 6 genes that were common in both studies (**Table 17**). All of them are related to purine and pyrimidine synthesis, confirming results obtained by Samant *et al*. However, 2 genes had no homologs in *V. vulnificus* and the rest (14) were classified as neutral for growth in serum. When we look in for this 14 genes in the LB-library we saw 5 genes categorized as *essential* or *domain essential*. Moreover, T2SS, NQR or Clp protease complexes were not considered essential for *E.coli* K12 but were part of the list obtained in our study. Actually, only 48 of the 85 genes listed as essential to grow in HS in *V. vulnificus* YJ016 had a homolog in K12 strain in which NQR complex is not present. This discrepancy is probably caused by difference between the species biology and maybe are the cause that *V. vulnificus* is able to cause septicemia and death in a patient in a short period of time.

Table 17. Comparison of essential genes to grow in HS in *E. coli* K12 with *V. vulnificus* (Vv). Genes categories in Samant *et al* study were blasted (BLASTP) against *V. vulnificus* YJ016 genes and homologs were compared with those classified as essential to grow in serum. E, essential; NE, non-essential, -, no homolog found.

Gene names	*E. coli* K12 locus	Vv Locus	classification in IS	classification *in vitro*
carA	*b0032*	*VV0624*	E	NE
carB	*b0033*	*VV0625*	E	NE
gcvR[a]	*b2479*	*VV2505*	NE	NE
guaA	*b2507*	*VV0776*	E	NE
guaB	*b2508*	*VV0775*	E	NE
ihfB[a]	*b0912*	*VV1301/VV1107*	NE	DE/NE
lipA[a]	*b0628*	*VV0900*	NE	E
lysS[a]	*b2890*	*VV0669*	NE	E
nadB[a]	*b2574*	*VV2839*	NE	NE
panB[a]	*b0134*	*VV2763*	NE	E
panC[a]	*b0133*	*VV2764*	NE	E
purA	*b4177*	*VV3066/VVA1235*	NE	NE/NE
purC	*b2476*	*VV1468*	NE	NE

purD	*b4005*	*VV3143*	NE	NE
purE	*b0523*	*VV3218*	NE	NE
purF	*b2312*	*VV2419*	E	NE
purH	*b4006*	*VV3142*	NE	NE
purK	*b0522*	*VV3219*	NE	NE
purL	*b2557*	*VV0847*	NE	NE
purM	*b2499*	*VV2515*	NE	NE
pyrB	*b4245*	*VV2920*	NE	NE
pyrC	*b1062*	*VVA0407*	E	NE
pyrE	*b3642*	*VV0280*	NE	NE
pyrF	*b1281*	*VV1304*	NE	DE
rpoN[a]	*b3202*	*VV0448*	NE	NE
rseA[a]	*b2572*	*VV2837*	NE	NE
ydaS	*b1357*	-	-	-
ydaT	*b1358*	-	-	-

[a]Identified as false positives by testing

6. Discussion

One of the main advantages that high-throughput sequencing has provided to microbiologist is the depth at which we can study now microbes. In particular, the development of TIS has approached us to understand more precisely the contribution of each genetic locus to bacterial fitness from a genome-wide perspective. Comprehend which genes are essential for a bacterium to infect the host must be a priority in the study of pathogens. In *V. vulnificus*, one of the most dangerous marine opportunistic pathogens, is particularly interesting the rapidity to expand, cause a septicemia and death in humans, especially in those immunocompromised. Using TIS we have been able to detect some of the genes essential for this pathogen to develop such a rapid and dangerous infection.

The incubation of the transposon library in human serum with active and inactive complement allowed us to discover genes putatively essential for complement-resistance (CR) in *V. vulnificus*. These genes seemed to be involved in CPS synthesis and assembly because most of them were located in the "capsule locus". This locus contains conserved (exportation and assembly proteins) and variable (oligosaccharide-subunit biosynthesis) genes, and both were found in the list of CR-genes. Previous works in other pathogens has highlighted the importance of bacterial external envelopes in resistance to complement activity (Gemski *et al*., 1980; Marshall and Gunn, 2015; Keo *et al*.; Barnes *et al*., 2003; Hammerschmidt *et al*., 1994). The most common bacterial CR-strategies involve external envelopes and consist in i) preventing complement binding to their targets by presenting very long O-antigen chains and/or CPS, and ii) mimicking host tissues by incorporating sialic acid (Neu5Ac) into LPS and/or CPS. Early studies performed in *V. vulnificus* found a correlation between capsule and resistance to human serum by proving that mutants deficient in CPS transport were sensitive to human serum (Wright *et al*., 1990)(Wright *et al*.). A recent study performed by Lubin *et al.* in YJ016 also found that this bacterium can add sialic acid moieties to their LPS(Lubin *et al*., 2015). Although multiple capsular types have been described in *V. vulnificus* and most of them are unable to grow in undiluted human serum, the molecular basis for CR remains largely unknown.

Nevertheless, the most accepted hypothesis is that the negatively charged CPS of *V. vulnificus* prevents depot complement (Roberts, 1996).

We first get a mutant in one of the conserved genes in the capsule locus, *wza*. Our experiments demonstrated that the mutant-associated phenotype correlated with a loss of CPS (translucent colonies and bacterial aggregation) concomitantly with a loss of CR and a decrease in virulence degree for mice. Then, we focused our attention in 4 genes that were apparently unrelated with CPS biosynthesis. None of these genes corresponded to the ones previously identified as enzymes involved in LPS-sialization (Lubin *et al*., 2015). The 4 genes were automatically annotated and had never been studied before. Two of them were included in the capsule locus (*VV0357* and *VV0358*), which suggested their participation in the CPS formation, and the other 2 were located in chromosome II (*VVA0762* and *VVA0954*). We got ΔVV0357 and ΔVV0358 and studied and compared their phenotype with that of the wild-type strain. The deletion of each gene by separate produced a loss of capsule concomitantly with human complement sensitivity and a decrease in virulence degree. On the contrary, no phenotypic change was associated to the deletion of *VVA0762* or *VVA0954*. The short length of *VVA0762* (342 nt) could be the explanation of why this was considered *under-represented* in serum while *VVA0954* could represent a false positive, consequence of the putative bias during the library construction in HS and IHS.

A search in the database showed that *VV0357* (*dehydrogenase*) and *VV0358* (*hypothetical protein*) were practically exclusive of *V. vulnificus* YJ016. One of the domains in the protein VV0358 was homologous (100%) to the heparinase II/III domain present in a heparinase of probed activity against heparin from *Flavobacterium heparinum* (*Pedobacter*). Finding this function in the study was very surprising since no capacity to coagulate blood had been reported in *V. vulnificus* or other *Vibrio*. Regrettably, all the tests performed in this work to find this activity in our isolate were unsuccessful. BLAST results showed homologous genes in other vibrios and genus but looking at the "*heparinase*" domains from the hits and comparing to the original domain in *F. heparinum* we

suspect that "heparinases" automatically annotated in the databases are not true heparinises and propose a revision of these entries in the Genbank database.

A second domain with a putative alginate lyase activity was found in VV0358. Taking into account that alginate and heparan-sulfate are both proteoglycans we suspected that VV0358 could be involved in the CPS turnover. Further studies are in progress to test this hypothesis.

Heparinase from *F. heparinum* is transported to the periplasmic space where it performs its function. However, online tools did not predict signal peptides in VV0358 for exportation and therefore we cannot consider it as a periplasmic protein. In addition, no statistical significance was found in the prediction programs for transmembrane proteins. Interestingly, VV0357 was predicted to have a transmembrane domain which made us suspects that VV0358 could be attached to the inner or outer membrane through VV0357.

In summary, we found that the unique virulence factor that is unequivocally related with resistance to complement killing in *V. vulnificus* is the capsule. The selected strain is one of the most virulent for humans and the most efficient in growth in serum among all the tested strains. We hypothesized that VV0357 and VV0358 could be on the basis of such unusual efficiency.

On the other hand, interesting results were retrieved when HIS- and LB-libraries were compared. This pathogen had two main entrance ways in the host: by ingestion of contaminated raw food or through a wound. The second option avoids many difficult steps for the bacteria to reach bloodstream and if it resists innate immune activity the colonization of the host will only depend on its capacity to replicate and expand in the blood. In order to see which genes were necessary for this we differentiate genes essential for human serum growth and related this with resistance to microcidal peptides and nutritional immunity. We found 85 genes that could be essential for *V. vulnificus* to overcome nutritional immunity. From this list, 5 genes were involved in synthesis of purines and pyrimidines, which suggests that nucleotides deficiency is an important part of the nutritional immunity against bacteria. Similar results were obtained by

Samant *et al*., (2008) who highlighted the importance of nucleotides for the growth of *E. coli, Salmonella enterica and Bacillus anthracis* in human serum (Samant *et al*., 2008).

We also found genes related to surface protein complexes involved in exo-protein secretion (T2SS) and electrochemical gradient building (NQR). The genes that form these were not considered essential in previous studies with *E. coli* and *M. catarrhalis* but, as mentioned previously, the depth in which we can study with high-throughput sequencing now is increasing the knowledge obtained. It is also possible that these functions are essential for growth in human serum in *V. vulnificus* but neutral in these species.

T2SS is used by *V. vulnificus* to secrete the main exoprotease (VvpE) and the most abundant cytolysin (VvhA) to the extracellular medium (Hwang *et al*., 2011). VvpE hydrolyzes a great variety of proteins and glycoproteins (including human transferrin), which are abundant in serum (Miyoshi and Shinoda, 2000) and VvhA is a cytotoxin with hemolytic activity (Wright *et al*., 1985). It has been proposed that VvhA acts together with VvpE to provide bacteria not only with carbon and nitrogen sources but also with an iron source as heme (Veyrier and Cellier, 2015). Although both genes were not in the list, we hypothesize that T2SS should be essential for the export of both products (and probably others) to the bloodstream to provide nutrients for bacterial growth.

In relation to iron acquisition, we found the gene for a heme receptor of *V. vulnificus, hupA* and two genes for exogenous spermidine/putrescine transport (spermidine is the central block in the vulnibactin molecule) in the list of essential genes. The main micronutrient that forms part of the nutritional immunity is iron (Hood and Skaar, 2012). This metal is sequestered by transferrin in blood and forms part the heme group of hemoproteins. The virulence of *V. vulnificus* for humans is strongly correlated with iron content in blood (Hor *et al*., 2000). In fact, *V. vulnificus* expresses various systems to get it, one of them involving two heme/hemoglobin receptors and the other vulnibactin plus its transport system (Webster and Litwin, 2000). From the two genes for heme receptors, only *hupA* was present in our list. Precisely this gene is the

unique of the two involved in virulence for mice (Pajuelo *et al*., 2014). We hypothesized that *V. vulnificus* uses HupA to get iron in blood, at least during the first steps of infections, and that could utilize exogenous spermidine as a precursor for vulnibactin biosynthesis.

We also found as essential genes for grow in HS, those encoding a NQR system. These systems are used by bacteria to maintain the Na^+/proton gradient in certain natural environments (Häse and Barquera, 2001). Hubbard and Chao classified genes for NQR complexes as neutral in LB suggesting that they would not be essential for growth in a regular medium like LB in *Vibrio*(Hubbard *et al*., 2016; Chao *et al*., 2013). Since sodium content in LB and human serum are in the same range, we consider that NQR could help to maintain the electrostatic equilibrium in a stressing medium such as IHS. In fact, it has been demonstrated that NQR system maintains the proton-motive force when the bacterial membranes are altered (Steuber *et al*., 2014; Steuber, 2001). Since complement is inactivated by heating at 65ºC, we suspect that bactericidal peptides should be still active in IHS. This was supported by two facts: i) one of the OMP, OmpH, involved in resistance to bactericidal effects of microcidal peptides was identified in the list, and ii) several genes involved in LPS biosynthesis/modification that have been also related to resistance to microcidal peptides were also identified in the list(Rosenfeld and Shai, 2006).

Finally, we have mentioned several times the study of Chao *et al* and Hubbard *et al*. These assays also used TIS to investigate the genes essential in *Vibrio* species. In particular, *V. cholerae* and *V. parahaemolyticus* were studied for *in vitro* and infection conditions. One of the interesting results from our analysis was to determine if the three *Vibrio* species share a group of genes essential for *in vitro* growth. By comparing the genes categorized as under-represented to grow *in vitro* in the three *Vibrio* species we were able to discern 303 genes homologous and essential. After classify these 303 genes and the complete genomes of the three strains used in each study in GO categories we compared them and proposed a series of functions over-represented in the shared group. These functions are: *structural constituent of ribosome* and *binding* (molecular

function); *intracellular organelle part*, *organelle*, *macromolecular complex*, *cell part* and *cell* (cellular part); *cellular component organization biogenesis*, *metabolic process* and *cellular process* (biological process). Even further studies could reduce this list; these 303 genes could be the core-genome for growth in a regular medium in *Vibrio* genus (**Table 18**).

Table 18. List of genes that putatively form the core-genome for growth in a regular medium in *Vibrio* genus

Vv ID	annotation	Vv ID	annotation
VV0003	putative inner membrane protein translocase component YidC	VV1299	cytidylate kinase
VV0004	ribonuclease P	VV1300	30S ribosomal protein S1
VV0005	50S ribosomal protein L34	VV1303	tetratricopeptide repeat protein
VV0011	chromosomal replication initiation protein	VV1318	transcriptional regulator CysB
VV0012	DNA polymerase III subunit beta	VV1322	putative cell division protein FtsK-like
VV0020	glycyl-tRNA synthetase subunit beta	VV1323	outer-membrane lipoprotein carrier protein
VV0021	glycyl-tRNA synthetase subunit alpha	VV1325	seryl-tRNA synthetase
VV0024	sulfur transfer protein SirA	VV1341	adenylosuccinate lyase
VV0033	protoporphyrinogen oxidase	VV1343	tRNA-specific 2-thiouridylase MnmA
VV0080	uroporphyrinogen-III synthase	VV1346	DNA-binding protein H-NS
VV0081	porphobilinogen deaminase	VV1360	UDP-2,3-diacylglucosamine hydrolase
VV0087	diaminopimelate epimerase	VV1362	cysteinyl-tRNA synthetase
VV0112	signal recognition particle GTPase	VV1420	50S ribosomal protein L25
VV0115	RNA polymerase factor sigma-32	VV1442	glutamate decarboxylase
VV0121	4-hydroxybenzoate octaprenyltransferase	VV1482	hypothetical protein
VV0168	glutamate racemase	VV1551	dissimilatory sulfite reductase, gamma subunit
VV0177	ubiquinone/menaquinone biosynthesis methyltransferase	VV1942	50S ribosomal protein L20
VV0179	putative ubiquinone biosynthesis protein UbiB	VV1943	50S ribosomal protein L35
VV0185	delta-aminolevulinic acid dehydratase	VV1945	threonyl-tRNA synthetase
VV0186	DNA polymerase I	VV1970	phenylalanyl-tRNA synthetase subunit beta
VV0187	ribosome biogenesis GTP-binding protein YsxC	VV1971	phenylalanyl-tRNA synthetase subunit alpha
VV0241	bifunctional (p)ppGpp synthetase II/ guanosine-3,5-bis pyrophosphate 3-pyrophosphohydrolase	VV2090	asparaginyl-tRNA synthetase
VV0243	guanylate kinase	VV2121	fatty acid metabolism regulator
VV0283	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase	VV2288	aspartyl-tRNA synthetase
VV0286	50S ribosomal protein L28	VV2314	DNA topoisomerase I
VV0287	50S ribosomal protein L33	VV2321	translation initiation factor IF-1
VV0292	phosphopantetheine adenylyltransferase	VV2353	3-deoxy-manno-octulosonate cytidylyltransferase
VV0297	3-deoxy-D-manno-octulosonic-acid kinase	VV2355	tetraacyldisaccharide 4-kinase
VV0324	3-deoxy-D-manno-octulosonic-acid transferase	VV2361	outer membrane-specific lipoprotein transporter subunit LolE
VV0374	30S ribosomal protein S10	VV2363	putative ABC transporter integral membrane subunit
VV0375	50S ribosomal protein L3	VV2411	DNA polymerase III subunits gamma and tau
VV0376	50S ribosomal protein L4	VV2422	folylpolyglutamate synthase

VV0377	50S ribosomal protein L23	VV2500	succinyl-diaminopimelate desuccinylase
VV0378	50S ribosomal protein L2	VV2504	dihydrodipicolinate synthase
VV0379	30S ribosomal protein S19	VV2529	DNA polymerase III subunit epsilon
VV0380	50S ribosomal protein L22	VV2539	cell cycle protein MesJ
VV0381	30S ribosomal protein S3	VV2540	acetyl-CoA carboxylase carboxyltransferase subunit alpha
VV0382	50S ribosomal protein L16	VV2541	DNA polymerase III subunit alpha
VV0383	50S ribosomal protein L29	VV2544	lipid-A-disaccharide synthase
VV0384	30S ribosomal protein S17	VV2545	UDP-N-acetylglucosamine acyltransferase
VV0385	50S ribosomal protein L14	VV2546	(3R)-hydroxymyristoyl-ACP dehydratase
VV0386	50S ribosomal protein L24	VV2547	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase
VV0387	50S ribosomal protein L5	VV2549	outer membrane protein assembly factor YaeT
VV0388	30S ribosomal protein S14	VV2550	membrane-associated Zn-dependent protease 1
VV0389	30S ribosomal protein S8	VV2551	1-deoxy-D-xylulose 5-phosphate reductoisomerase
VV0390	50S ribosomal protein L6	VV2552	CDP-diglyceride synthetase
VV0392	30S ribosomal protein S5	VV2553	undecaprenyl pyrophosphate synthase
VV0393	50S ribosomal protein L30	VV2554	ribosome recycling factor
VV0394	50S ribosomal protein L15	VV2555	uridylate kinase
VV0395	preprotein translocase subunit SecY	VV2556	elongation factor Ts
VV0396	30S ribosomal protein S13	VV2557	30S ribosomal protein S2
VV0397	30S ribosomal protein S11	VV2558	methionine aminopeptidase
VV0398	30S ribosomal protein S4	VV2572	prolyl-tRNA synthetase
VV0399	DNA-directed RNA polymerase subunit alpha	VV2584	thiamin biosynthesis lipoprotein ApbE
VV0400	50S ribosomal protein L17	VV2622	tetrahydrodipicolinate N-succinyltransferase
VV0449	ABC-type transport system, ATPase component	VV2705	30S ribosomal protein S15
VV0450	hypothetical protein	VV2707	ribosome-binding factor A
VV0451	hypothetical protein	VV2710	ribosome maturation factor RimP
VV0453	putative polysialic acid capsule expression protein	VV2713	phosphomannomutase
VV0461	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	VV2714	dihydropteroate synthase
VV0469	geranylgeranyl pyrophosphate synthase	VV2715	ATP-dependent Zn protease
VV0470	50S ribosomal protein L21	VV2723	tyrosyl-tRNA synthetase
VV0471	50S ribosomal protein L27	VV2728	glutamate-1-semialdehyde aminotransferase
VV0472	GTPase ObgE	VV2763	3-methyl-2-oxobutanoate hydroxymethyltransferase
VV0475	dihydrofolate reductase	VV2764	pantoate--beta-alanine ligase
VV0478	dimethyladenosine transferase	VV2771	dihydrolipoamide dehydrogenase
VV0480	parvulin-like peptidyl-prolyl isomerase	VV2782	dephospho-CoA kinase
VV0481	organic solvent tolerance protein	VV2785	50S ribosomal protein L19
VV0502	glucosamine--fructose-6-phosphate aminotransferase	VV2786	tRNA (guanine-N(1)-)-methyltransferase
VV0561	RNA polymerase sigma factor RpoD	VV2787	16S rRNA-processing protein RimM
VV0562	DNA primase	VV2788	30S ribosomal protein S16
VV0564	30S ribosomal protein S21	VV2789	signal recognition particle GTPase
VV0565	metal-dependent protease	VV2801	carbon storage regulator
VV0568	dihydroneopterin aldolase	VV2803	alanyl-tRNA synthetase
VV0589	DNA topoisomerase IV subunit A	VV2814	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
VV0594	50S ribosomal protein L13	VV2816	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
VV0595	30S ribosomal protein S9	VV2817	cell division protein FtsB
VV0606	S-adenosyl-methyltransferase MraW	VV2818	phosphopyruvate hydratase
VV0607	cell division protein FtsL	VV2819	CTP synthetase
VV0608	cell division protein FtsI	VV2827	4-phosphopantetheinyl transferase
VV0609	UDP-N-acetylmuramoylalanyl-D-glutamate--2, 6-diaminopimelate	VV2828	pyridoxine 5-phosphate synthase

	ligase		
VV0610	UDP-N-acetylmuramyl pentapeptide synthase	VV2830	GTP-binding protein Era
VV0611	phospho-N-acetylmuramoyl-pentapeptide- transferase	VV2831	ribonuclease III
VV0612	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase	VV2832	signal peptidase I
VV0613	cell division membrane protein FtsW	VV2835	positive regulator of sigma E activity
VV0614	undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	VV2842	aminomethyltransferase
VV0616	cell division septal protein FtsQ	VV2846	2-octaprenyl-6-methoxyphenyl hydroxylase
VV0617	cell division protein FtsA	VV2857	fructose-bisphosphate aldolase
VV0618	cell division protein FtsZ	VV2858	phosphoglycerate kinase
VV0619	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase	VV2863	S-adenosylmethionine synthetase
VV0620	hypothetical protein	VV2870	Holliday junction resolvase-like protein
VV0621	preprotein translocase subunit SecA	VV2905	permease
VV0623	dihydrodipicolinate reductase	VV2906	permease
VV0669	lysyl-tRNA synthetase	VV2910	valyl-tRNA synthetase
VV0678	prolipoprotein diacylglyceryl transferase	VV2932	rod shape-determining protein MreD
VV0679	thymidylate synthase	VV2933	rod shape-determining protein MreC
VV0684	30S ribosomal protein S20	VV2934	rod shape-determining protein MreB
VV0685	mviN protein	VV2978	alanine racemase
VV0686	bifunctional riboflavin kinase/FMN adenylyltransferase	VV2979	replicative DNA helicase
VV0687	isoleucyl-tRNA synthetase	VV2981	50S ribosomal protein L9
VV0688	lipoprotein signal peptidase	VV2983	primosomal replication protein N
VV0690	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	VV2986	site-specific DNA methylase
VV0712	putative lipoprotein	VV3012	primosome assembly protein PriA
VV0755	cysteine desulfurase	VV3014	cell division protein
VV0756	scaffold protein	VV3025	triosephosphate isomerase
VV0758	co-chaperone HscB	VV3031	30S ribosomal protein S7
VV0760	ferredoxin	VV3032	30S ribosomal protein S12
VV0766	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	VV3033	protein TusB; tRNA 5-methylaminomethyl-2-thiouridine synthase
VV0767	histidyl-tRNA synthetase	VV3034	protein TusC
VV0770	GTP-binding protein EngA	VV3035	sulfur transfer complex subunit TusD
VV0819	hypothetical protein	VV3074	N-acetylmuramoyl-L-alanine amidase
VV0820	hypothetical protein	VV3075	ATPase or kinase
VV0824	inorganic polyphosphate/ATP-NAD kinase	VV3077	ribosome-associated GTPase
VV0825	molecular chaperone GrpE	VV3078	oligoribonuclease
VV0833	chaperone protein DnaJ	VV3079	phosphatidylserine decarboxylase
VV0868	1-deoxy-D-xylulose-5-phosphate synthase	VV3082	phosphoglyceromutase
VV0869	geranylgeranyl pyrophosphate synthase	VV3087	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase
VV0900	lipoyl synthase	VV3088	serine acetyltransferase
VV0901	lipoyltransferase	VV3110	6-phosphofructokinase
VV0902	hypothetical protein	VV3150	uroporphyrinogen decarboxylase
VV0906	penicillin-binding protein 2	VV3159	DNA-directed RNA polymerase subunit beta
VV0909	DNA polymerase III subunit delta	VV3161	50S ribosomal protein L10
VV0911	leucyl-tRNA synthetase	VV3162	50S ribosomal protein L1
VV0915	putative metalloprotease	VV3163	50S ribosomal protein L11
VV0918	2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase	VV3164	transcription antitermination protein NusG
VV0925	peptidyl-tRNA hydrolase	VV3165	preprotein translocase subunit SecE
VV0926	ribose-phosphate pyrophosphokinase	VV3169	biotin--protein ligase
VV0928	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	VV3170	UDP-N-acetylenolpyruvoylglucosamine reductase
VV0929	outer membrane lipoprotein LolB	VV3172	phosphatidylserine synthase

VV0930	glutamyl-tRNA reductase		VV3179	3-polyprenyl-4-hydroxybenzoate decarboxylase
VV0931	peptide chain release factor 1		VV3181	transcription termination factor Rho
VV0932	methylase of polypeptide chain release factor		VV3226	methionyl-tRNA formyltransferase
VV0935	2-dehydro-3-deoxyphosphooctonate aldolase		VV3249	bifunctional N-acetylglucosamine-1-phosphate uridyltransferase/glucosamine-1-phosphate acetyltransferase
VV0949	glutamyl-tRNA synthetase		VV3250	F0F1 ATP synthase subunit epsilon
VV0982	NAD-dependent DNA ligase LigA		VV3251	F0F1 ATP synthase subunit beta
VV1002	adenylate kinase		VV3252	F0F1 ATP synthase subunit gamma
VV1003	ferrochelatase		VV3253	F0F1 ATP synthase subunit alpha
VV1013	glutaminyl-tRNA synthetase		VV3254	F0F1 ATP synthase subunit delta
VV1044	arginyl-tRNA synthetase		VV3255	F0F1 ATP synthase subunit B
VV1049	inactive homolog of metal-dependent protease		VV3256	F0F1 ATP synthase subunit C
VV1188	electron transport complex protein RnfG		VV3257	F0F1 ATP synthase subunit A
VV1189	electron transport complex protein RnfD		VV3258	F0F1 ATP synthase subunit I
VV1235	phosphatidylglycerophosphate synthase		VVA0001	hypothetical protein
VV1243	ribonucleotide-diphosphate reductase subunit beta		VVA0410	NAD synthetase
VV1244	ribonucleotide-diphosphate reductase subunit alpha		VVA0411	nicotinic acid mononucleotide adenylyltransferase
VV1247	DNA gyrase, subunit A		VVA0578	ribosomal biogenesis GTPase
VV1259	methionyl-tRNA synthetase		VVA0785	D-alanyl-alanine synthetase A
VV1274	acyl carrier protein S-malonyltransferase		VVA0810	signal transduction histidine kinase
VV1276	acyl carrier protein		VVA1092	GTP cyclohydrolase I
VV1280	thymidylate kinase		VVA1648	pyridoxamine 5'-phosphate oxidase
VV1281	DNA polymerase III subunit delta		VVA1695	transcriptional regulator
			VVA1697	ParA family protein

CONCLUSIONS

Chapter one. *Vibrio* vs Host: colonization and eel microbiota

I. The skin-mucus microbiome of wild-eels is different from that of surrounding water and farmed-eels. Gammaproteobacteria is the most abundant bacterial class in the skin-mucus microbiome of wild- and farmed-eels but with different species/genus composition.

II. Water salinity is the main factor but not the only one in determining skin-mucus microbiome composition. Mucus concentrates selectively bacteria possessing specific capabilities from those present in water.

III. Functionalities of microbiome in eel skin-mucus differ significantly from those of water-microbiome and predict that successful colonizers have to contain in their genomes genes for i) attachment, mainly by forming biofilms, iii) competence and communication with other bacteria, and iii) resistance to innate immunity, predators (amoeba) and heavy metals/drugs.

IV. *Pseudomonas, Acinetobacter, Stenotrophomonas and Sphingobium* are part of the resident microbiota common to wild- and farmed-eels. *Vibrio* monopolizes the skin-mucus in the wild-eels from estuarine waters (≥ 7 g/l salinity) while a mixture of *Pseudomonas, Stenotrophomonas*, *Achromobacter*, *Sphingobium*, *Aeromonas* and *Shewanella* predominates in that from lakes and rivers (≤ 3 g/l salinity).

V. Mucus concentrates selectively those vibrios from the total pool present in water that carry specific virulence genes and constitutes a reservoir not only for pathogenic bacteria but also for virulence genes. *V. anguillarum* was the most abundant *Vibrio* in the skin-mucus from estuaries while in freshwater *V. cholerae/V. metoecu*s were the most abundant inside *Vibrio* genus.

VI. The combination of freshwater and low pH in fish farm's facilities efficiently eliminates pathogenic *Vibrio* and other pathogenic species from the eel mucosal surface. *Comamonas*, *Chryseobacterium* and

Citrobacter are the main genera attached to epidermal mucus of farmed-eels.

VII. Mucus could favor the genetic material interchange; it contains mobile genetic elements (mainly ICEs) harboring virulence and/or drug resistance genes. The interchanges seem to have been mainly produced among bacteria with similar GC.

VIII. The skin-mucus of eels is a complex media where bacteriophages are also present. Bacteriophages are in high concentration (x20 times more than bacteria) in skin-mucus of eels and are actively infecting bacterial hosts, supporting the BAM model described by Barr et al. (Barr et al., 2013a).Myoviruses and Podoviruses are the most abundant viruses and ΦKZ genus and the Podovirus are part of the resident microbiota associated to the eel mucosal surface.

Chapter two. *Vibrio vulnificus*: growth in human serum

I. Transposon insertion sequencing is an optimal approach to study the essential genes of *Vibrio vulnificus* to grow in a determined condition.

II. The essential genes for by *V. vulnificus* YJ016 to grow in a general medium, like LB, are 457 *under-represented* (*essential*) and 197 *regional* (*domain essential*).

III. *V. vulnificus, V. cholerae* and *V. anguillarum* share 303 genes *essential* to grow in LB which can be designated as the *Vibrio*-core-genome for growth in a general medium.

IV. The list of genes essential to resist human complement killing is composed of 10 genes. These are involved in capsule formation and include 2 new genes: *VV0357* and *VV0358*.

V. *VV0357* and *VV0358* are practically exclusive of *V. vulnificus* YJ016 and are included in the capsule locus. Its deletion makes the bacterium unencapsulated, sensitive to human complement killing and diminishes significantly its virulence for mice. Its levels of expression are in accordance with the rest of the capsule genes and do not vary in human serum *vs* LB.

VI.   Although *VV0358* is annotated as a *heparinase* we considered that the *heparinase$_{Vv}$* has not this activity and recommend the revision of the automatic annotated *heparinases* with a heparinase II/III like domain in databases.

VII.   *V. vulnificus* YJ016 also has to overcome the nutritional immunity in order to grow in human serum. We have found 85 genes that could be essential for *V. vulnificus* to overcome nutritional immunity.

VIII.   The type 2 secretion system, the Na+-translocating NADH-ubiquinone reductase complex, proteases, iron acquisition, nucleotide *de novo* synthesis and putrescine and spermidine translocation are some of the functions putatively responsible for the resistance to nutritional immunity and the rapid growth of *V. vulnificus* in human blood.

CONCLUSIONS

REFERENCES

Aida, K., Tsukamoto, K., and Yamauchi, K. eds. (2003). *Eel Biology*. Tokyo: Springer Japan doi:10.1007/978-4-431-65907-5.

Alice, A. F., Naka, H., and Crosa, J. H. (2008). Global gene expression as a function of the iron status of the bacterial cell: influence of differentially expressed genes in the virulence of the human pathogen *Vibrio vulnificus*. *Infect. Immun.* 76, 4019–37. doi:10.1128/IAI.00208-08.

Altschul, S. F., Madden, T. L., Schäffer, a a, Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–402.

Amaro, C., and Biosca, E. G. (1996). *Vibrio vulnificus* biotype 2, pathogenic for eels, is also an opportunistic pathogen for humans. *Appl. Environ. Microbiol.* 62, 1454–7.

Amaro, C., Biosca, E. G., Fouz, B., Toranzo, a. E., and Garay, E. (1994). Role of iron, capsule, and toxins in the pathogenicity of *Vibrio vulnificus* biotype 2 for mice. *Infect. Immun.* 62, 759–763.

Amaro, C., Esteve, C., Biosca, E. G., Fouz, B., and Toranzo., A. E. (1992). Comparative study of phenotypic and virulence properties in *Vibrio vulnificus* biotype 1 and 2 obtained from a European eel farm experiencing mortalities. *Dis. Aquat. Org* 13.

Arede, P., Ministro, J., and Oliveira, D. C. (2013). Redefining the Role of the -Lactamase Locus in Methicillin-Resistant *Staphylococcus aureus*: -Lactamase Regulators Disrupt the MecI-Mediated Strong Repression on mecA and Optimize the Phenotypic Expression of Resistance in Strains with Constitutive mecA Ex. *Antimicrob. Agents Chemother.* 57, 3037–3045. doi:10.1128/AAC.02621-12.

Arias, C. R., Macian, M. C., Aznar, R., Garay, E., and Pujalte, M. J. (1999). Low incidence of *Vibriovulnificus* among *Vibrio* isolates from sea water and shellfish of the western Mediterranean coast. *J. Appl. Microbiol.* 86, 125–134. doi:10.1046/j.1365-2672.1999.00641.x.

Asturiano, J. F., Sørensen, S. R., Pérez, L., Lauesen, P., and Tomkiewicz, J. (2016). First Production of Larvae Using Cryopreserved Sperm: Effects of

Preservation Temperature and Cryopreservation on European Eel Sperm Fertilization Capacity. *Reprod. Domest. Anim.* 51, 485–91. doi:10.1111/rda.12706.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2. doi:10.1038/msb4100050.

Badhai, J., Ghosh, T. S., and Das, S. K. (2016). Composition and Functional Characterization of Microbiome Associated with Mucus of the Coral Fungia echinata Collected from Andaman Sea. *Front. Microbiol.* 7, 936. doi:10.3389/fmicb.2016.00936.

Baker-Austin, C., Stockley, L., Rangdale, R., and Martinez-Urtaza, J. (2010). Environmental occurrence and clinical impact of *Vibriovulnificus* and *Vibrio* parahaemolyticus: a European perspective. *Environ. Microbiol. Rep.* 2, 7–18. doi:10.1111/j.1758-2229.2009.00096.x.

Baker-Austin, C., Trinanes, J. A., Taylor, N. G. H., Hartnell, R., Siitonen, A., and Martinez-Urtaza, J. (2012). Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat. Clim. Chang.* 3, 73–77. doi:10.1038/nclimate1628.

Banga, J., and Tripathi, C. K. M. (2009). Response surface methodology for optimization of medium components in submerged culture of *Aspergillus flavus* for enhanced heparinase production. *Lett. Appl. Microbiol.* 49, 204–209. doi:10.1111/j.1472-765X.2009.02640.x.

Barnes, A. C., Young, F. M., Horne, M. T., and Ellis, A. E. (2003). Streptococcus iniae: serological differences, presence of capsule and resistance to immune serum killing. *Dis. Aquat. Organ.* 53, 241–7. doi:10.3354/dao053241.

Barr, J. J., Auro, R., Furlan, M., Whiteson, K. L., Erb, M. L., Pogliano, J., Stotland, A., Wolkowicz, R., Cutting, A. S., Doran, K. S., et al. (2013a). Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc. Natl. Acad. Sci. U. S. A.* 110, 10771–6. doi:10.1073/pnas.1305923110.

Barr, J. J., Youle, M., and Rohwer, F. (2013b). Innate and acquired bacteriophage-mediated immunity. *Bacteriophage* 3, e25857. doi:10.4161/bact.25857.

Barry, K. A., Middelbos, I. S., Vester Boler, B. M., Dowd, S. E., Suchodolski, J. S., Henrissat, B., Coutinho, P. M., White, B. A., Fahey, G. C., and Swanson, K. S. (2012). Effects of dietary fiber on the feline gastrointestinal metagenome. *J. Proteome Res.* 11, 5924–33. doi:10.1021/pr3006809.

Bengtsson-Palme, J., Boulund, F., Fick, J., Kristiansson, E., and Larsson, D. G. J. (2014). Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India. *Front. Microbiol.* 5, 1–14. doi:10.3389/fmicb.2014.00648.

Benson, D. a, Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Wheeler, D. L. (2005). GenBank. *Nucleic Acids Res.* 33, D34-8. doi:10.1093/nar/gki063.

Biosca, E. G., Amaro, C., Esteve, C., Alcaide, E., and Garay., E. (1991). First record of *Vibriovulnificus* biotype 2 from diseased European eel *Anguilla anguilla*. *J. Fish Dis.* 14, 103–109.

Bisharat, N., Agmon, V., Finkelstein, R., Raz, R., Ben-Dror, G., Lerner, L., Soboh, S., Colodner, R., Cameron, D. N., Wykstra, D. L., et al. (1999). Clinical, epidemiological, and microbiological features of *Vibriovulnificus* biogroup 3 causing outbreaks of wound infection and bacteraemia in Israel. Israel *Vibrio* Study Group. *Lancet* 354, 1421–4.

Bliven, A. K., and Maurelli, A. T. "Evolution of Bacterial Pathogens Within the Human Host," in *Virulence Mechanisms of Bacterial Pathogens, Fifth Edition* (American Society of Microbiology), 3–13. doi:10.1128/microbiolspec.VMBF-0017-2015.

Bollmann, M., Bosch, T., Colijn, F., Ebinghaus, R., Körtzinger, A., Latif, M., Matthiessen, B., Oschlies, A., Petersen, S., Proelß, A., et al. (2010). Living With the Oceans. *World Ocean Rev. Living with Ocean.* 1, 236.

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Org, E., Parks, B., Lusis, A. J., Knight, R., Caporaso, J. G., and Svanbäck, R. (2014). Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nat. Commun.* 5. doi:10.1038/ncomms5500.

Bordas, M. A., Balebona, M. C., Rodriguez-Maroto, J. M., Borrego, J. J., and Morinigo, M. A. (1998). Chemotaxis of pathogenic *Vibrio* strains towards mucus surfaces of gilt-head sea bream (Sparus aurata L.). *Appl. Environ. Microbiol.* 64, 1573–5.

Bowdre, J. H., Poole, M. D., and Oliver, J. D. (1981). Edema and

hemoconcentration in mice experimentally infected with *Vibriovulnificus*. *Infect. Immun.* 32, 1193–9.

Breitbart, M., and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13, 278–284. doi:10.1016/j.tim.2005.04.003.

Bullen, J. J., Spalding, P. B., Ward, C. G., and Gutteridge, J. M. (1991). Hemochromatosis, iron and septicemia caused by *Vibriovulnificus*. *Arch. Intern. Med.* 151, 1606–9.

Cantas, L., Sørby, J. R. T., Aleström, P., and Sørum, H. (2012). Culturable gut microbiota diversity in zebrafish. *Zebrafish* 9, 26–37. doi:10.1089/zeb.2011.0712.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G., and Parkhill, J. (2005). ACT: the Artemis Comparison Tool. *Bioinformatics* 21, 3422–3. doi:10.1093/bioinformatics/bti553.

Centers for Disease Control and Prevention (CDC) (1993). *Vibriovulnificus* infections associated with raw oyster consumption--Florida, 1981-1992. *MMWR. Morb. Mortal. Wkly. Rep.* 42, 405–7.

Cerdà-Cuéllar, M., Jofre, J., and Blanch, A. R. (2000). A selective medium and a specific probe for detection of *Vibriovulnificus*. *Appl. Environ. Microbiol.* 66, 855–9.

Ceyssens, P.-J., Minakhin, L., Van den Bossche, A., Yakunina, M., Klimuk, E., Blasdel, B., De Smet, J., Noben, J.-P., Bläsi, U., Severinov, K., et al. (2014). Development of giant bacteriophage ΦKZ is independent of the host transcription apparatus. *J. Virol.* doi:10.1128/JVI.01347-14.

Chao, M. C., Abel, S., Davis, B. M., and Waldor, M. K. (2016). The design and analysis of transposon insertion sequencing experiments. *Nat. Rev. Microbiol.* 14, 119–128. doi:10.1038/nrmicro.2015.7.

Chao, M. C., Pritchard, J. R., Zhang, Y. J., Rubin, E. J., Livny, J., Davis, B. M., and Waldor, M. K. (2013). High-resolution definition of the *Vibriocholerae* essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data. *Nucleic Acids Res.* 41, 9033–48. doi:10.1093/nar/gkt654.

Chatzidaki-livanis, M., Jones, M. K., Anita, C., and Wright, A. C. (2006). Genetic Variation in the *Vibriovulnificus* Group 1 Capsular Polysaccharide

Operon Genetic Variation in the *Vibriovulnificus* Group 1 Capsular Polysaccharide Operon. *J. Bacteriol.* 188, 1987–1998. doi:10.1128/JB.188.5.1987.

Chatzidaki-Livanis, M., Jones, M. K., and Wright, A. C. (2006). Genetic variation in the *Vibriovulnificus* group 1 capsular polysaccharide operon. *J. Bacteriol.* 188, 1987–98. doi:10.1128/JB.188.5.1987-1998.2006.

Chen, C.-Y. (2003). Comparative Genome Analysis of *Vibriovulnificus*, a Marine Pathogen. *Genome Res.* 13, 2577–2587. doi:10.1101/gr.1295503.7.

Chen, Y., Bystricky, P., Adeyeye, J., Panigrahi, P., Ali, A., Johnson, J. A., Bush, C., Morris, J., and Stine, O. (2007). The capsule polysaccharide structure and biogenesis for non-O1 *Vibriocholerae* NRT36S: genes are embedded in the LPS region. *BMC Microbiol.* 7, 20. doi:10.1186/1471-2180-7-20.

Chen, Y.-C., Chuang, Y.-C., Chang, C.-C., Jeang, C.-L., and Chang, M.-C. (2004). A K+ yptake protein, TrkA, is required for serum, protamine, and polymyxin B resistance in *Vibriovulnificus*. *Infect. Immun.* 72, 629–36.

Chiang, S. L., and Rubin, E. J. (2002). Construction of a mariner-based transposon for epitope-tagging and genomic targeting. *Gene* 296, 179–85.

Chiang, S.-R., and Chuang, Y.-C. (2003). *Vibriovulnificus* infection: clinical manifestations, pathogenesis, and antimicrobial therapy. *J. Microbiol. Immunol. Infect.* 36, 81–8.

Cianciotto, N. P. (2009). Many substrates and functions of type II secretion: lessons learned from *Legionella pneumophila*. *Future Microbiol.* 4, 797–805. doi:10.2217/fmb.09.53.

Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., Kulam-Syed-Mohideen, a S., McGarrell, D. M., Marsh, T., Garrity, G. M., et al. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 37, D141-5. doi:10.1093/nar/gkn879.

Colwell, R. R., Kaper, J., and Joseph, S. W. (1977). *Vibriocholerae*, *Vibrioparahaemolyticus*, and other vibrios: occurrence and distribution in Chesapeake Bay. *Science* 198, 394–6.

Consortium, T. U. (2014). Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 42, D191-8. doi:10.1093/nar/gkt1140.

Córdova-Kreylos, A. L., and Scow, K. M. (2007). Effects of ciprofloxacin on salt

marsh sediment microbial communities. *ISME J.* 1, 585–95. doi:10.1038/ismej.2007.71.

Cornelissen, A., Hardies, S. C., Shaburova, O. V, Krylov, V. N., Mattheus, W., Kropinski, A. M., and Lavigne, R. (2012). Complete genome sequence of the giant virus OBP and comparative genome analysis of the diverse ΦKZ-related phages. *J. Virol.* 86, 1844–52. doi:10.1128/JVI.06330-11.

D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., et al. (2011). Antibiotic resistance is ancient. *Nature* 477, 457–461. doi:10.1038/nature10388.

Dalsgaard, A., Serichantalergs, O., Forslund, A., Lin, W., Mekalanos, J., Mintz, E., Shimada, T., and Wells, J. G. (2001). Clinical and Environmental Isolates of *Vibriocholerae* Serogroup O141 Carry the CTX Phage and the Genes Encoding the Toxin-Coregulated Pili. *J. Clin. Microbiol.* 39, 4086–4092. doi:10.1128/JCM.39.11.4086-4092.2001.

Defoirdt, T., Sorgeloos, P., and Bossier, P. (2011). Alternatives to antibiotics for the control of bacterial disease in aquaculture. *Curr. Opin. Microbiol.* 14, 251–258. doi:10.1016/j.mib.2011.03.004.

Deresinski, S. (2009). Vancomycin in Combination with Other Antibiotics for the Treatment of Serious Methicillin-Resistant *Staphylococcus aureus* Infections. *Clin. Infect. Dis.* 49, 1072–1079. doi:10.1086/605572.

Dibrov, P. (2005). The sodium cycle in vibrio *cholerae*: riddles in the dark. *Biochemistry. (Mosc).* 70, 150–3.

Diekema, D. J., BootsMiller, B. J., Vaughn, T. E., Woolson, R. F., Yankey, J. W., Ernst, E. J., Flach, S. D., Ward, M. M., Franciscus, C. L. J., Pfaller, M. A., et al. (2004). Antimicrobial Resistance Trends and Outbreak Frequency in United States Hospitals. *Clin. Infect. Dis.* 38, 78–85. doi:10.1086/380457.

Dong, C., Beis, K., Nesper, J., Brunkan-Lamontagne, A. L., Clarke, B. R., Whitfield, C., and Naismith, J. H. (2006). Wza the translocon for *E. coli* capsular polysaccharides defines a new class of membrane protein. *Nature* 444, 226–229. doi:10.1038/nature05267.

Dong, W., Lu, W., McKeehan, W. L., Luo, Y., and Ye, S. (2012). Structural basis of heparan sulfate-specific degradation by heparinase III. *Protein Cell* 3, 950–61. doi:10.1007/s13238-012-2056-z.

Donnenberg, M. S., and Kaper, J. B. (1991). Construction of an eae deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect. Immun.* 59, 4310–4317.

Van Doorslaer, X., Dewulf, J., Van Langenhove, H., and Demeestere, K. (2014). Fluoroquinolone antibiotics: an emerging class of environmental micropollutants. *Sci. Total Environ.* 500–501, 250–69. doi:10.1016/j.scitotenv.2014.08.075.

Driessen, A. J. M., and Nouwen, N. (2008). Protein Translocation Across the Bacterial Cytoplasmic Membrane. *Annu. Rev. Biochem.* 77, 643–667. doi:10.1146/annurev.biochem.77.061606.160747.

Drummelsmith, J., and Whitfield, C. (1999). Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). *Mol. Microbiol.* 31, 1321–32.

Drummelsmith, J., and Whitfield, C. (2000). Translocation of group 1 capsular polysaccharide to the surface of *Escherichia coli* requires a multimeric complex in the outer membrane. *EMBO J.* 19, 57–66. doi:10.1093/emboj/19.1.57.

Duong-Nu, T.-M., Jeong, K., Hong, S. H., Nguyen, H.-V., Ngo, V.-H., Min, J.-J., Lee, S. E., and Rhee, J. H. (2015). All Three TonB Systems Are Required for *Vibriovulnificus* CMCP6 Tissue Invasiveness by Controlling Flagellum Expression. *Infect. Immun.* 84, 254–65. doi:10.1128/IAI.00821-15.

Dutilh, B. E., Cassman, N., McNair, K., Sanchez, S. E., Silva, G. G. Z., Boling, L., Barr, J. J., Speth, D. R., Seguritan, V., Aziz, R. K., et al. (2014). A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* 5, 1–11. doi:10.1038/ncomms5498.

Eloe-Fadrosh, E. A., Paez-Espino, D., Jarett, J., Dunfield, P. F., Hedlund, B. P., Dekas, A. E., Grasby, S. E., Brady, A. L., Dong, H., Briggs, B. R., et al. (2016). Global metagenomic survey reveals a new bacterial candidate phylum in geothermal springs. *Nat. Commun.* 7, 10476. doi:10.1038/ncomms10476.

Feingersch, R., Philosof, A., Mejuch, T., Glaser, F., Alalouf, O., Shoham, Y., and Béjà, O. (2012). Potential for phosphite and phosphonate utilization by *Prochlorococcus*. *ISME J.* 6, 827–834. doi:10.1038/ismej.2011.149.

Felsenstein, J. (1993). PHYLIP phylogeny inference package, version 3.5 c.

Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., et al. (2006). Pfam: clans, web tools and services. *Nucleic Acids Res.* 34, D247-51. doi:10.1093/nar/gkj149.

Fong, J. C. N., Syed, K. A., Klose, K. E., and Yildiz, F. H. (2010). Role of *Vibrio* polysaccharide (vps) genes in VPS production, biofilm formation and *Vibriocholerae* pathogenesis. *Microbiology* 156, 2757–2769. doi:10.1099/mic.0.040196-0.

Food and Agriculture Organisation (2014). World review of fisheries and aquaculture. *State World Fish. Aquac.*, 3–89. doi:E-ISBN 978-92-5-108276-8.

Fouz, B., Larsen, J. L., and Amaro, C. (2006). *Vibriovulnificus* serovar A: An emerging pathogen in European anguilliculture. *J. Fish Dis.* 29, 285–291. doi:10.1111/j.1365-2761.2006.00719.x.

Frans, I., Michiels, C. W., Bossier, P., Willems, K. A., Lievens, B., and Rediers, H. (2011). *Vibrioanguillarum* as a fish pathogen: virulence factors, diagnosis and prevention. *J. Fish Dis.* 34, 643–61. doi:10.1111/j.1365-2761.2011.01279.x.

Fullner, K. J., and Mekalanos, J. J. (2000). In vivo covalent cross-linking of cellular actin by the *Vibriocholerae* RTX toxin. *EMBO J.* 19, 5315–23. doi:10.1093/emboj/19.20.5315.

Gardete, S., and Tomasz, A. (2014). Mechanisms of vancomycin resistance in *Staphylococcus aureus*. *J. Clin. Invest.* 124, 2836–2840. doi:10.1172/JCI68834.

Gavin, H. E., and Satchell, K. J. F. (2015). MARTX toxins as effector delivery platforms. *Pathog. Dis.* 73, ftv092. doi:10.1093/femspd/ftv092.

Gawronski, J. D., Wong, S. M. S., Giannoukos, G., Ward, D. V., and Akerley, B. J. (2009). Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. *Proc. Natl. Acad. Sci.* 106, 16422–16427. doi:10.1073/pnas.0906627106.

Gemski, P., Cross, A. S., and Sadoff, J. C. (1980). K1 antigen-associated resistance to the bactericidal activity of serum. *FEMS Microbiol. Lett.* 9,

193–197.

Ghai, R., Hernandez, C. M., Picazo, A., Mizuno, C. M., Ininbergs, K., Díez, B., Valas, R., DuPont, C. L., McMahon, K. D., Camacho, A., et al. (2012). Metagenomes of Mediterranean coastal lagoons. *Sci. Rep.* 2, 490. doi:10.1038/srep00490.

Ghai, R., Martin-Cuadrado, A.-B., Molto, A. G., Heredia, I. G., Cabrera, R., Martin, J., Verdú, M., Deschamps, P., Moreira, D., López-García, P., et al. (2010). Metagenome of the Mediterranean deep chlorophyll maximum studied by direct and fosmid library 454 pyrosequencing. *ISME J.* 4, 1154–66. doi:10.1038/ismej.2010.44.

Ghai, R., Mizuno, C. M., Picazo, A., Camacho, A., and Rodriguez-Valera, F. (2013). Metagenomics uncovers a new group of low GC and ultra-small marine Actinobacteria. *Sci. Rep.* 3, 1–8. doi:10.1038/srep02471.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth* 6, 343–345.

Goldsmith, D. B., Crosti, G., Dwivedi, B., McDaniel, L. D., Varsani, A., Suttle, C. a, Weinbauer, M. G., Sandaa, R.-A., and Breitbart, M. (2011). Development of phoH as a novel signature gene for assessing marine phage diversity. *Appl. Environ. Microbiol.* 77, 7730–9. doi:10.1128/AEM.05531-11.

Gonzaga, A., and Rodriguez-valera, F. (2013). Genomic Diversity of " Deep Ecotype " *Alteromonas macleodii*. 5, 1220–1232. doi:10.1093/gbe/evt089.

Gousset, B. (1990). European eel (Anguilla anguilla L.) farming technologies in Europe and in Japan: Application of a comparative analysis. *Aquaculture* 87, 209–235. doi:10.1016/0044-8486(90)90060-Z.

Grant, J. R., Arantes, A. S., and Stothard, P. (2012). Comparing thousands of circular genomes using the CGView Comparison Tool. *BMC Genomics* 13, 202. doi:10.1186/1471-2164-13-202.

Hachani, A., Lossi, N. S., Hamilton, A., Jones, C., Bleves, S., Albesa-Jové, D., and Filloux, A. (2011). Type VI secretion system in *Pseudomonas* aeruginosa: secretion and multimerization of VgrG proteins. *J. Biol. Chem.* 286, 12317–27. doi:10.1074/jbc.M110.193045.

Haley, B. J., Grim, C. J., Hasan, N. A., Choi, S., Chun, J., Brettin, T. S., Bruce, D. C., Challacombe, J. F., Detter, J. C., Han, C. S., et al. (2010a). Comparative genomic analysis reveals evidence of two novel *Vibrio* species closely related to *V*. *cholerae.*

Haley, B. J., Grim, C. J., Hasan, N. a, Choi, S.-Y., Chun, J., Brettin, T. S., Bruce, D. C., Challacombe, J. F., Detter, J. C., Han, C. S., et al. (2010b). Comparative genomic analysis reveals evidence of two novel *Vibrio* species closely related to V. *cholerae*. *BMC Microbiol.* 10, 154. doi:10.1186/1471-2180-10-154.

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*

Hammerschmidt, S., Birkholz, C., Zähringer, U., Robertson, B. D., van Putten, J., Ebeling, O., and Frosch, M. (1994). Contribution of genes from the capsule gene complex (cps) to lipooligosaccharide biosynthesis and serum resistance in *Neisseria meningitidis*. *Mol. Microbiol.* 11, 885–96.

Häse, C. C., and Barquera, B. (2001). Role of sodium bioenergetics in *Vibriocholerae*. *Biochim. Biophys. Acta* 1505, 169–78.

Hase, C. C., Fedorova, N. D., Galperin, M. Y., and Dibrov, P. A. (2001). Sodium Ion Cycle in Bacterial Pathogens: Evidence from Cross-Genome Comparisons. *Microbiol. Mol. Biol. Rev.* 65, 353–370. doi:10.1128/MMBR.65.3.353-370.2001.

Heidelberg, J. F., Heidelberg, K. B., and Colwell, R. R. (2002). Bacteria of the gamma-subclass Proteobacteria associated with zooplankton in Chesapeake Bay. *Appl. Environ. Microbiol.* 68, 5498–507.

Henderson, I. R., Owen, P., and Nataro, J. P. (1999). Molecular switches--the ON and OFF of bacterial phase variation. *Mol. Microbiol.* 33, 919–32.

Hennersdorf, P., Mrotzek, G., Abdul-Aziz, M. A., and Saluz, H. P. (2016). Metagenomic analysis between free-living and cultured *Epinephelus fuscoguttatus* under different environmental conditions in Indonesian waters. *Mar. Pollut. Bull.* doi:10.1016/j.marpolbul.2016.05.009.

Herrington, D. A. (1988). Toxin, toxin-coregulated pili, and the toxR regulon are essential for *Vibriocholerae* pathogenesis in humans. *J. Exp. Med.* 168, 1487–1492. doi:10.1084/jem.168.4.1487.

Hitchcock, P. J., and Brown, T. M. (1983). Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. *J. Bacteriol.* 154, 269–77.

Hood, M. I., and Skaar, E. P. (2012). Nutritional immunity: transition metals at the pathogen–host interface. *Nat. Rev. Microbiol.* 10, 525–537. doi:10.1038/nrmicro2836.

Hor, L. I., Chang, Y. K., Chang, C. C., Lei, H. Y., and Ou, J. T. (2000). Mechanism of high susceptibility of iron-overloaded mouse to *Vibriovulnificus* infection. *Microbiol. Immunol.* 44, 871–8.

Hor, L.-I., Goo, C.-T., and Wan, L. (1995). Isolation and characterization of*Vibriovulnificus* inhabiting the marine environment of the southwestern area of Taiwan. *J. Biomed. Sci.* 2, 384–389. doi:10.1007/BF02255226.

Horseman, M. A., and Surani, S. (2011). A comprehensive review of *Vibriovulnificus*: an important cause of severe sepsis and skin and soft-tissue infection. *Int. J. Infect. Dis.* 15, e157–e166. doi:10.1016/j.ijid.2010.11.003.

Hsieh, Y.-J., and Wanner, B. L. (2010). Global regulation by the seven-component Pi signaling system. *Curr. Opin. Microbiol.* 13, 198–203. doi:10.1016/j.mib.2010.01.014.

Hubbard, T. P., Chao, M. C., Abel, S., Blondel, C. J., Abel zur Wiesch, P., Zhou, X., Davis, B. M., and Waldor, M. K. (2016). Genetic analysis of *Vibrio parahaemolyticus* intestinal colonization. *Proc. Natl. Acad. Sci.*, 201601718. doi:10.1073/pnas.1601718113.

Hunter, S., Corbett, M., Denise, H., Fraser, M., Gonzalez-Beltran, A., Hunter, C., Jones, P., Leinonen, R., McAnulla, C., Maguire, E., et al. (2014). EBI metagenomics--a new resource for the analysis and archiving of metagenomic data. *Nucleic Acids Res.* 42, D600–D606. doi:10.1093/nar/gkt961.

Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N., and Schuster, S. C. (2011). Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* 21, 1552–60. doi:10.1101/gr.120618.111.

Hwang, W., Lee, N. Y., Kim, J., Lee, M.-A., Kim, K.-S., Lee, K.-H., and Park, S.-J. (2011). Functional characterization of EpsC, a component of the type II secretion system, in the pathogenicity of *Vibriovulnificus*. *Infect. Immun.*

79, 4068–80. doi:10.1128/IAI.05351-11.

Hyatt, D., Chen, G.-L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119. doi:10.1186/1471-2105-11-119.

Jang, H. Bin, Fagutao, F. F., Nho, S. W., Park, S. Bin, Cha, I. S., Yu, J. E., Lee, J. S., Im, S. P., Aoki, T., and Jung, T. S. (2013). Phylogenomic network and comparative genomics reveal a diverged member of the ΦKZ-related group, marine vibrio phage ΦJM-2012. *J. Virol.* 87, 12866–78. doi:10.1128/JVI.02656-13.

Jevons, M. P., Coe, A. W., and Parker, M. T. (1963). Methicillin resistance in staphylococci. *Lancet (London, England)* 1, 904–7.

Jin, C., and Flavell, R. a (2013). Innate sensors of pathogen and stress: Linking inflammation to obesity. *J. Allergy Clin. Immunol.* 132, 287–94. doi:10.1016/j.jaci.2013.06.022.

Joh, S.-J., Ahn, E.-H., Lee, H.-J., Shin, G.-W., Kwon, J.-H., and Park, C.-G. (2013). Bacterial pathogens and flora isolated from farm-cultured eels (*Anguilla japonica*) and their environmental waters in Korean eel farms. *Vet. Microbiol.* 163, 190–5. doi:10.1016/j.vetmic.2012.11.004.

Jones, M. K., and Oliver, J. D. (2009). *Vibriovulnificus*: disease and pathogenesis. *Infect. Immun.* 77, 1723–33. doi:10.1128/IAI.01046-08.

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32, D277-80. doi:10.1093/nar/gkh063.

Kaplan, E., Ofek, M., Jurkevitch, E., and Cytryn, E. (2013). Characterization of fluoroquinolone resistance and qnr diversity in Enterobacteriaceae from municipal biosolids. *Front. Microbiol.* 4. doi:10.3389/fmicb.2013.00144.

Karlsson, F. H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C. J., Fagerberg, B., Nielsen, J., and Bäckhed, F. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498, 99–103. doi:10.1038/nature12198.

Kashimoto, T., Ueno, S., Hanajima, M., Hayashi, H., Akeda, Y., Miyoshi, S., Hongo, T., Honda, T., and Susa, N. (2003). *Vibriovulnificus* induces

macrophage apoptosis in vitro and in vivo. *Infect. Immun.* 71, 533–5.

Kasiraj, A. C., Harmoinen, J., Isaiah, A., Westermarck, E., Steiner, J. M., Spillmann, T., and Suchodolski, J. S. (2016). The effects of feeding and withholding food on the canine small intestinal microbiota. *FEMS Microbiol. Ecol.* 92. doi:10.1093/femsec/fiw085.

Kaspar, C. W., and Tamplin, M. L. (1993). Effects of temperature and salinity on the survival of *Vibriovulnificus* in seawater and shellfish. *Appl. Environ. Microbiol.* 59, 2425–9.

Keating, T. A., Marshall, C. G., and Walsh, C. T. (2000). Vibriobactin biosynthesis in *Vibriocholerae*: VibH is an amide synthase homologous to nonribosomal peptide synthetase condensation domains. *Biochemistry* 39, 15513–21.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi:10.1038/nprot.2015.053.

Keo, T., Collins, J., Kunwar, P., Blaser, M. J., and Iovine, N. M. *Campylobacter* capsule and lipooligosaccharide confer resistance to serum and cationic antimicrobials. *Virulence* 2, 30–40.

Kichenaradja, P., Siguier, P., Perochon, J., and Chandler, M. (2010). ISbrowser: an extension of ISfinder for visualizing insertion sequences in prokaryotic genomes. *Nucl. Acids Res.* 38, D62-68. doi:10.1093/nar/gkp947.

Kim, S. Y., Thanh, X. T. T., Jeong, K., Kim, S. B., Pan, S. O., Jung, C. H., Hong, S. H., Lee, S. E., and Rhee, J. H. (2014). Contribution of Six Flagellin Genes to the Flagellum Biogenesis of *Vibriovulnificus* and In Vivo Invasion. *Infect. Immun.* 82, 29–42. doi:10.1128/IAI.00654-13.

Kirchberger, P. C., Turnsek, M., Hunt, D. E., Haley, B. J., Colwell, R. R., Polz, M. F., Tarr, C. L., and Boucher, Y. (2014). *Vibrio* metoecus sp. nov., a close relative of *Vibriocholerae* isolated from coastal brackish ponds and clinical specimens. *Int. J. Syst. Evol. Microbiol.* 64, 3208–3214. doi:10.1099/ijs.0.060145-0.

Kirn, T. J., Jude, B. A., and Taylor, R. K. (2005). A colonization factor links *Vibriocholerae* environmental survival and human infection. *Nature* 438, 863–6. doi:10.1038/nature04249.

Klontz, K. C., Lieb, S., Schreiber, M., Janowski, H. T., Baldy, L. M., and Gunn, R. A. (1988). Syndromes of *Vibriovulnificus* infections. Clinical and epidemiologic features in Florida cases, 1981-1987. *Ann. Intern. Med.* 109, 318–23.

Konstantinidis, K. T., Braff, J., Karl, D. M., and DeLong, E. F. (2009). Comparative Metagenomic Analysis of a Microbial Community Residing at a Depth of 4,000 Meters at Station ALOHA in the North Pacific Subtropical Gyre. *Appl. Environ. Microbiol.* 75, 5345–5355. doi:10.1128/AEM.00473-09.

Korotkov, K. V., Sandkvist, M., and Hol, W. G. J. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat. Rev. Microbiol.* doi:10.1038/nrmicro2762.

Kulakov, L. a, Ksenzenko, V. N., Shlyapnikov, M. G., Kochetkov, V. V, Del Casale, A., Allen, C. C. R., Larkin, M. J., Ceyssens, P.-J., and Lavigne, R. (2009). Genomes of "phiKMV-like viruses" of *Pseudomonas aeruginosa* contain localized single-strand interruptions. *Virology* 391, 1–4. doi:10.1016/j.virol.2009.06.024.

Kümmerer, K. (2009). Antibiotics in the aquatic environment--a review--part I. *Chemosphere* 75, 417–34. doi:10.1016/j.chemosphere.2008.11.086.

Kuo, S.-Y., Chou, M.-C., Lee, S.-L., Wang, Y., Chen, C.-L., Lin, P.-T., and Lo, H.-R. (2015). *Vibriovulnificus* RtxA1 modulated calcium flux contributes reduced internalization in phagocytes. *Life Sci.* 132, 55–60. doi:10.1016/j.lfs.2015.03.027.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–5.

Langridge, G. C., Phan, M.-D., Turner, D. J., Perkins, T. T., Parts, L., Haase, J., Charles, I., Maskell, D. J., Peters, S. E., Dougan, G., et al. (2009). Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants. *Genome Res.* 19, 2308–16. doi:10.1101/gr.097097.109.

Lapierre, P., and Gogarten, J. P. (2009). Estimating the size of the bacterial pan-genome. *Trends Genet.* 25, 107–110. doi:10.1016/j.tig.2008.12.004.

Lee, B. C., Kim, M. S., Choi, S. H., and Kim, T. S. (2010). Involvement of capsular polysaccharide via a TLR2/NF-kappaB pathway in *Vibriovulnificus*-induced IL-8 secretion of human intestinal epithelial cells.

Int. J. Mol. Med. 25, 581–91.

Lee, C.-T., Amaro, C., Wu, K.-M., Valiente, E., Chang, Y.-F., Tsai, S.-F., Chang, C.-H., and Hor, L.-I. (2008). A common virulence plasmid in biotype 2 *Vibriovulnificus* and its dissemination aided by a conjugal plasmid. *J. Bacteriol.* 190, 1638–48. doi:10.1128/JB.01484-07.

Lee, C.-T., Pajuelo, D., Llorens, A., Chen, Y.-H., Leiro, J. M., Padrós, F., Hor, L.-I., and Amaro, C. (2013). MARTX of *Vibriovulnificus* biotype 2 is a virulence and survival factor. *Environ. Microbiol.* 15, 419–32. doi:10.1111/j.1462-2920.2012.02854.x.

Lee, D. G., Urbach, J. M., Wu, G., Liberati, N. T., Feinbaum, R. L., Miyata, S., Diggins, L. T., He, J., Saucier, M., Déziel, E., et al. (2006a). Genomic analysis reveals that Pseudomonas aeruginosa virulence is combinatorial. *Genome Biol.* 7, R90. doi:10.1186/gb-2006-7-10-r90.

Lee, I., Berdis, A. J., and Suzuki, C. K. (2006b). Recent developments in the mechanistic enzymology of the ATP-dependent Lon protease from *Escherichia coli*: highlights from kinetic studies. *Mol. Biosyst.* 2, 477–83. doi:10.1039/b609936j.

Lee, J. H., Kim, M. W., Kim, B. S., Kim, S. M., Lee, B. C., Kim, T. S., and Choi, S. H. (2007). Identification and characterization of the *Vibriovulnificus* rtxA essential for cytotoxicity in vitro and virulence in mice. *J. Microbiol.* 45, 146–52.

Lewis, A. L., Lubin, J. B., Argade, S., Naidu, N., Choudhury, B., and Boyd, E. F. (2011). Genomic and metabolic profiling of nonulosonic acids in Vibrionaceae reveal biochemical phenotypes of allelic divergence in *Vibriovulnificus*. *Appl. Environ. Microbiol.* 77, 5782–5793. doi:10.1128/AEM.00712-11.

Lo, H.-R., Lin, J.-H., Chen, Y.-H., Chen, C.-L., Shao, C.-P., Lai, Y.-C., and Hor, L.-I. (2011). RTX Toxin Enhances the Survival of *Vibriovulnificus* During Infection by Protecting the Organism From Phagocytosis. *J. Infect. Dis.* 203, 1866–1874. doi:10.1093/infdis/jir070.

López-Romalde, S., Magariños, B., Ravelo, C., Toranzo, A. E., and Romalde, J. L. (2003). Existence of two O-serotypes in the fish pathogen *Pseudomonas anguilliseptica*. *Vet. Microbiol.* 94, 325–33.

Lubin, J.-B., Lewis, W. G., Gilbert, N. M., Weimer, C. M., Almagro-Moreno, S.,

Boyd, E. F., and Lewis, A. L. (2015). Host-like carbohydrates promote bloodstream survival of *Vibriovulnificus* in vivo. *Infect. Immun.* 83, 3126–36. doi:10.1128/IAI.00345-15.

Ma, A. T., and Mekalanos, J. J. (2010). In vivo actin cross-linking induced by *Vibriocholerae* type VI secretion system is associated with intestinal inflammation. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4365–70. doi:10.1073/pnas.0915156107.

Ma, Y.-F., Zhang, Y., Zhang, J.-Y., Chen, D.-W., Zhu, Y., Zheng, H., Wang, S.-Y., Jiang, C.-Y., Zhao, G.-P., and Liu, S.-J. (2009). The complete genome of *Comamonas testosteroni* reveals its genetic adaptations to changing environments. *Appl. Environ. Microbiol.* 75, 6812–9. doi:10.1128/AEM.00933-09.

Marchesi, J. R., Dutilh, B. E., Hall, N., Peters, W. H. M., Roelofs, R., Boleij, A., and Tjalsma, H. (2011). Towards the Human Colorectal Cancer Microbiome. *PLoS One* 6, e20447. doi:10.1371/journal.pone.0020447.

Marchler-Bauer, a., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., et al. (2014). CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226. doi:10.1093/nar/gku1221.

Marco-Noales, E., Biosca, E. G., and Amaro, C. (1999). Effects of salinity and temperature on long-term survival of the eel pathogen *Vibriovulnificus* biotype 2 (serovar E). *Appl. Environ. Microbiol.* 65, 1117–26.

Marco-Noales, E., and Milán, M. (2001). Transmission to eels, portals of entry, and putative reservoirs of *Vibriovulnificus* serovar E (biotype 2). *Appl. ….* doi:10.1128/AEM.67.10.4717.

Marco-Noales, E., Milán, M., Fouz, B., Sanjuán, E., and Amaro, C. (2001). Transmission to eels, portals of entry, and putative reservoirs of *Vibriovulnificus* serovar E (biotype 2). *Appl. Environ. Microbiol.* 67, 4717–25.

Marshall, J. M., and Gunn, J. S. (2015). The O-Antigen Capsule of *Salmonella enterica Serovar Typhimurium* Facilitates Serum Resistance and Surface Expression of FliC. *Infect. Immun.* 83, 3946–3959. doi:10.1128/IAI.00634-15.

Marshall, K. C. ed. (1986). *Advances in Microbial Ecology*. Boston, MA:

Springer US doi:10.1007/978-1-4757-0611-6.

Martin, S. J., and Siebeling, R. J. (1991). Identification of *Vibriovulnificus* O serovars with antilipopolysaccharide monoclonal antibody. *J. Clin. Microbiol.* 29, 1684–8.

Martin-Cuadrado, A.-B., Garcia-Heredia, I., Moltó, A. G., López-Úbeda, R., Kimes, N., López-García, P., Moreira, D., and Rodriguez-Valera, F. (2015). A new class of marine Euryarchaeota group II from the Mediterranean deep chlorophyll maximum. *ISME J.* 9, 1619–34. doi:10.1038/ismej.2014.249.

Martín-Cuadrado, A.-B., López-García, P., Alba, J.-C., Moreira, D., Monticelli, L., Strittmatter, A., Gottschalk, G., and Rodríguez-Valera, F. (2007). Metagenomics of the Deep Mediterranean, a Warm Bathypelagic Habitat. *PLoS One* 2, e914. doi:10.1371/journal.pone.0000914.

McAllister, T. A., Meale, S. J., Valle, E., Guan, L. L., Zhou, M., Kelly, W. J., Henderson, G., Attwood, G. T., and Janssen, P. H. (2015). RUMINANT NUTRITION SYMPOSIUM: Use of genomics and transcriptomics to identify strategies to lower ruminal methanogenesis. *J. Anim. Sci.* 93, 1431–49. doi:10.2527/jas.2014-8329.

McCoy-Simandle, K., Stewart, C. R., Dao, J., DebRoy, S., Rossier, O., Bryce, P. J., and Cianciotto, N. P. (2011). *Legionella pneumophila* type II secretion dampens the cytokine response of infected macrophages and epithelia. *Infect. Immun.* 79, 1984–97. doi:10.1128/IAI.01077-10.

Mehrshad, M., Amoozegar, M. A., Ghai, R., Shahzadeh Fazeli, S. A., and Rodriguez-Valera, F. (2016). Genome Reconstruction from Metagenomic Data Sets Reveals Novel Microbes in the Brackish Waters of the Caspian Sea. *Appl. Environ. Microbiol.* 82, 1599–612. doi:10.1128/AEM.03381-15.

Meier-Dieter, U., Starman, R., Barr, K., Mayer, H., and Rick, P. D. (1990). Biosynthesis of enterobacterial common antigen in *Escherichia coli*. Biochemical characterization of Tn10 insertion mutants defective in enterobacterial common antigen synthesis. *J. Biol. Chem.* 265, 13490–7.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, a, Stevens, R., Wilke, a, et al. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9, 386. doi:10.1186/1471-2105-9-386.

Michalska, M., and Wolf, P. (2015). *Pseudomonas* Exotoxin A: optimized by evolution for effective killing. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00963.

Milton, D. L., O'Toole, R., Horstedt, P., and Wolf-Watz, H. (1996). Flagellin A is essential for the virulence of *Vibrioanguillarum. J. Bacteriol.* 178, 1310–9. doi:8631707.

Miyoshi, S., and Shinoda, S. (2000). Microbial metalloproteases and pathogenesis. *Microbes Infect.* 2, 91–8.

Mizuno, C. M., Rodriguez-Valera, F., Kimes, N. E., and Ghai, R. (2013). Expanding the marine virosphere using metagenomics. *PLoS Genet.* 9, e1003987. doi:10.1371/journal.pgen.1003987.

Mojica, F. J. M., Diez-Villasenor, C., Soria, E., and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* 36, 244–246. doi:10.1046/j.1365-2958.2000.01838.x.

Montiert, S., Suffredini, E., Ciccozzi, M., and Croci, L. (2010). Phylogenetic and evolutionary analysis of *Vibrio* parahaemolyticus and *Vibrioalginolyticus* isolates based on toxR gene sequence. *New Microbiol.* 33, 359–372.

Morris, J. G. (2003). Cholera and other types of vibriosis: a story of human pandemics and oysters on the half shell. *Clin. Infect. Dis.* 37, 272–80. doi:10.1086/375600.

Morrison, S. S., Pyzh, R., Jeon, M. S., Amaro, C., Roig, F. J., Baker-Austin, C., Oliver, J. D., and Gibas, C. J. (2014). Impact of analytic provenance in genome analysis. *BMC Genomics* 15 Suppl 8, S1. doi:10.1186/1471-2164-15-S8-S1.

Muzzi, A., and Donati, C. (2011). Population genetics and evolution of the pan-genome of *Streptococcus pneumoniae*. *Int. J. Med. Microbiol.* 301, 619–622. doi:10.1016/j.ijmm.2011.09.008.

Naka, H., Dias, G. M., Thompson, C. C., Dubay, C., Thompson, F. L., and Crosa, J. H. (2011). Complete genome sequence of the marine fish pathogen *Vibrio* anguillarum harboring the pJM1 virulence plasmid and genomic comparison with other virulent strains of *V. anguillarum* and *V. ordalii*. *Infect. Immun.* 79, 2889–900. doi:10.1128/IAI.05138-11.

Neiman, J., Guo, Y., and Rowe-Magnus, D. A. (2011). Chitin-Induced Carbotype Conversion in *Vibriovulnificus*. *Infect. Immun.* 79, 3195–3203. doi:10.1128/IAI.00158-11.

Ngo, J. K., Pomatto, L. C. D., and Davies, K. J. A. (2013). Upregulation of the mitochondrial Lon Protease allows adaptation to acute oxidative stress but dysregulation is associated with chronic stress, disease, and aging. *Redox Biol.* 1, 258–264. doi:10.1016/j.redox.2013.01.015.

Nielsen, M., and Esteve-Gassent, M. (2006). The eel immune system: present knowledge and the need for research. *J. Fish Dis.*, 65–78.

Ogilvie, L. a, Bowler, L. D., Caplin, J., Dedi, C., Diston, D., Cheek, E., Taylor, H., Ebdon, J. E., and Jones, B. V (2013). Genome signature-based dissection of human gut metagenomes to extract subliminal viral sequences. *Nat. Commun.* 4, 2420. doi:10.1038/ncomms3420.

Okonechnikov, K., Golosova, O., and Fursov, M. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28, 1166–7. doi:10.1093/bioinformatics/bts091.

Okujo, N., Saito, M., Yamamoto, S., Yoshida, T., Miyoshi, S., and Shinoda, S. (1994). Structure of vulnibactin, a new polyamine-containing siderophore from *Vibriovulnificus*. *Biometals* 7, 109–16.

Oliver, J. D. (2005). "*Vibriovulnificus*," in *Most* (American Society of Microbiology), 253–276. doi:10.1128/9781555815714.ch25.

Oliver, J. D., Warner, R. A., and Cleland, D. R. (1983). Distribution of *Vibriovulnificus* and other lactose-fermenting vibrios in the marine environment. *Appl. Environ. Microbiol.* 45, 985–98.

van Opijnen, T., Bodi, K. L., and Camilli, A. (2009). Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods* 6, 767–72. doi:10.1038/nmeth.1377.

Oppermann, U. C., Belai, I., and Maser, E. (1996). Antibiotic resistance and enhanced insecticide catabolism as consequences of steroid induction in the gram-negative bacterium Comamonas testosteroni. *J. Steroid Biochem. Mol. Biol.* 58, 217–23.

Orata, F. D., Kirchberger, P. C., Méheust, R., Barlow, E. J., Tarr, C. L., and Boucher, Y. (2015). The Dynamics of Genetic Interactions between *Vibrio*

metoecus and *Vibriocholerae* , Two Close Relatives Co-Occurring in the Environment. *Genome Biol. Evol.* 7, 2941–2954. doi:10.1093/gbe/evv193.

Pajuelo, D., Hernández-Cabanyero, C., Sanjuan, E., Lee, C.-T., Silva-Hernández, F. X., Hor, L.-I., MacKenzie, S., and Amaro, C. (2016). Iron and Fur in the life cycle of the zoonotic pathogen *Vibriovulnificus*. *Environ. Microbiol.* doi:10.1111/1462-2920.13424.

Pajuelo, D., Lee, C.-T., Roig, F. J., Hor, L.-I., and Amaro, C. (2015). Novel host-specific iron acquisition system in the zoonotic pathogen *Vibriovulnificus*. *Environ. Microbiol.* 17, 2076–89. doi:10.1111/1462-2920.12782.

Palace, S. G., Proulx, M. K., Lu, S., Baker, R. E., and Goguen, J. D. (2014). Genome-Wide Mutant Fitness Profiling Identifies Nutritional Requirements for Optimal Growth of Yersinia pestis in Deep Tissue. *MBio* 5, 1–12. doi:10.1128/mBio.01385-14.

Parfrey, L. W., and Knight, R. (2012). Spatial and temporal variability of the human microbiota. *Clin. Microbiol. Infect.* 18 Suppl 4, 8–11. doi:10.1111/j.1469-0691.2012.03861.x.

Parrow, N. L., Fleming, R. E., and Minnick, M. F. (2013). Sequestration and Scavenging of Iron in Infection. *Infect. Immun.* 81, 3503–3514. doi:10.1128/IAI.00602-13.

de Pascale, D., De Santi, C., Fu, J., and Landfald, B. (2012). The microbial diversity of Polar environments is a fertile ground for bioprospecting. *Mar. Genomics* 8, 15–22. doi:10.1016/j.margen.2012.04.004.

Patel, A., Noble, R. T., Steele, J. a, Schwalbach, M. S., Hewson, I., and Fuhrman, J. a (2007). Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. *Nat. Protoc.* 2, 269–76. doi:10.1038/nprot.2007.6.

Perna, N. T., Plunkett 3rd, G., Burland, V., Mau, B., Glasner, J. D., Rose, D. J., Mayhew, G. F., Evans, P. S., Gregor, J., Kirkpatrick, H. A., et al. (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 409, 529–533. doi:10.1038/35054089.

Petrof, E. O., Gloor, G. B., Vanner, S. J., Weese, S. J., Carter, D., Daigneault, M. C., Brown, E. M., Schroeter, K., and Allen-Vercoe, E. (2013). Stool substitute transplant therapy for the eradication of *Clostridium difficile* infection: "RePOOPulating" the gut. *Microbiome* 1, 3. doi:10.1186/2049-

2618-1-3.

Pflughoeft, K. J., and Versalovic, J. (2012). Human microbiome in health and disease. *Annu. Rev. Pathol.* 7, 99–122. doi:10.1146/annurev-pathol-011811-132421.

Phan, M.-D., Peters, K. M., Sarkar, S., Lukowski, S. W., Allsopp, L. P., Moriel, D. G., Achard, M. E. S., Totsika, M., Marshall, V. M., Upton, M., et al. (2013). The Serum Resistome of a Globally Disseminated Multidrug Resistant Uropathogenic *Escherichia coli* Clone. *PLoS Genet.* 9, e1003834. doi:10.1371/journal.pgen.1003834.

Phippen, B. L., and Oliver, J. D. (2015). Role of Anaerobiosis in Capsule Production and Biofilm Formation in *Vibriovulnificus*. *Infect. Immun.* 83, 551–559. doi:10.1128/IAI.02559-14.

Piddock, L. J. V (2006). Multidrug-resistance efflux pumps - not just for resistance. *Nat. Rev. Microbiol.* 4, 629–36. doi:10.1038/nrmicro1464.

Powell, J. L., Wright, A. C., Wasserman, S. S., Hone, D. M., and Morris, J. G. (1997). Release of tumor necrosis factor alpha in response to *Vibriovulnificus* capsular polysaccharide in in vivo and in vitro models. *Infect. Immun.* 65, 3713–8.

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–50. doi:10.1093/molbev/msp077.

Pritchard, J. R., Chao, M. C., Abel, S., Davis, B. M., Baranowski, C., Zhang, Y. J., Rubin, E. J., and Waldor, M. K. (2014). ARTIST: High-Resolution Genome-Wide Assessment of Fitness Using Transposon-Insertion Sequencing. *PLoS Genet.* 10, e1004782. doi:10.1371/journal.pgen.1004782.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi:10.1038/nature08821.

Randa, M. A., Polz, M. F., and Lim, E. (2004). Effects of temperature and salinity on *Vibriovulnificus* population dynamics as assessed by quantitative PCR. *Appl. Environ. Microbiol.* 70, 5469–76. doi:10.1128/AEM.70.9.5469-5476.2004.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. *Trends Genet.* 16, 2–3.

Rich, P. R., Meunier, B., and Ward, F. B. (1995). Predicted structure and possible ionmotive mechanism of the sodium-linked NADH-ubiquinone oxidoreductase of *Vibrio* alginolyticus. *FEBS Lett.* 375, 5–10. doi:10.1016/0014-5793(95)01164-A.

Roberts, I. S. (1996). The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annu. Rev. Microbiol.* 50, 285–315. doi:10.1146/annurev.micro.50.1.285.

Robertson, G. T., Kovach, M. E., Allen, C. A., Ficht, T. A., and Roop, R. M. (2000). The *Brucella abortus* Lon functions as a generalized stress response protease and is required for wild-type virulence in BALB/c mice. *Mol. Microbiol.* 35, 577–88.

Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pasić, L., Thingstad, T. F., Rohwer, F., and Mira, A. (2009). Explaining microbial population genomics through phage predation. *Nat. Rev. Microbiol.*7, 828–36. doi:10.1038/nrmicro2235.

Rodriguez-Valera, F., Mizuno, C. M., and Ghai, R. (2014). Tales from a thousand and one phages. *Bacteriophage* 4, e28265. doi:10.4161/bact.28265.

Rodriguez-Valera, F., and Ussery, D. W. (2012). Is the pan-genome also a pan-selectome? *F1000Research* 1, 16. doi:10.12688/f1000research.1-16.v1.

Roig, F. J., González-Candelas, F., and Amaro, C. (2011). Domain organization and evolution of multifunctional autoprocessing repeats-in-toxin (MARTX) toxin in vibrio *vulnificus*. *Appl. Environ. Microbiol.* 77, 657–668. doi:10.1128/AEM.01806-10.

Rosenfeld, Y., and Shai, Y. (2006). Lipopolysaccharide (Endotoxin)-host defense antibacterial peptides interactions: Role in bacterial resistance and prevention of sepsis. *Biochim. Biophys. Acta - Biomembr.* 1758, 1513–1522. doi:10.1016/j.bbamem.2006.05.017.

Roux, F. Le, Wegner, K. M., Baker-Austin, C., Vezzulli, L., Osorio, C. R., Amaro, C., Ritchie, J. M., Defoirdt, T., Destoumieux-Garzón, D., Blokesch, M., et al. (2015). The emergence of *Vibrio* pathogens in Europe: ecology, evolution, and pathogenesis (Paris, 11–12th March 2015). *Front. Microbiol.*

6. doi:10.3389/fmicb.2015.00830.

Samant, S., Lee, H., Ghassemi, M., Chen, J., Cook, J. L., Mankin, A. S., and Neyfakh, A. A. (2008). Nucleotide Biosynthesis Is Critical for Growth of Bacteria in Human Blood. *PLoS Pathog.* 4, e37. doi:10.1371/journal.ppat.0040037.

Sánchez, J., and Holmgren, J. (2008). Cholera toxin structure, gene regulation and pathophysiological and immunological aspects. *Cell. Mol. Life Sci.* 65, 1347–60. doi:10.1007/s00018-008-7496-5.

Sandkvist, M. (2001). Type II Secretion and Pathogenesis. *Infect. Immun.* 69, 3523–3535. doi:10.1128/IAI.69.6.3523-3535.2001.

Sanjuan, E., and Amaro, C. (2007). Multiplex PCR Assay for Detection of *Vibriovulnificus* Biotype 2 and Simultaneous Discrimination of Serovar E Strains. *Appl. Environ. Microbiol.* 73, 2029–2032. doi:10.1128/AEM.02320-06.

Sanjuán, E., González-Candelas, F., and Amaro, C. (2011). Polyphyletic origin of *Vibriovulnificus* biotype 2 as revealed by sequence-based analysis. *Appl. Environ. Microbiol.* 77, 688–95. doi:10.1128/AEM.01263-10.

Sato, H., and Frank, D. W. (2014). Intoxication of Host Cells by the T3SS Phospholipase ExoU: PI(4,5)P2-Associated, Cytoskeletal Collapse and Late Phase Membrane Blebbing. *PLoS One* 9, e103127. doi:10.1371/journal.pone.0103127.

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Mende, D. R., Kultima, J. R., Martin, J., et al. (2013). Genomic variation landscape of the human gut microbiome. *Nature* 493, 45–50. doi:10.1038/nature11711.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Meth* 9, 671–675.

Schwartz, S., Friedberg, I., Ivanov, I. V, Davidson, L. a, Goldsby, J. S., Dahl, D. B., Herman, D., Wang, M., Donovan, S. M., and Chapkin, R. S. (2012). A metagenomic study of diet-dependent interaction between gut microbiota and host in infants reveals differences in immune response. *Genome Biol.* 13, r32. doi:10.1186/gb-2012-13-4-r32.

Seed, K., Bodi, K., and Kropinski, A. (2011). Evidence of a dominant lineage of

Vibriocholerae-specific lytic bacteriophages shed by cholera patients over a 10-year period in Dhaka, Bangladesh. *MBio* 2, 1–9. doi:10.1128/mBio.00334-10.Editor.

Seguritan, V., Feng, I., and Rohwer, F. (2003). Genome sequences of two closely related *Vibrio* parahaemolyticus phages, VP16T and VP16C. *J. Bacteriol.* 185, 6434–6447. doi:10.1128/JB.185.21.6434.

Sender, R., Fuchs, S., and Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* 14, e1002533. doi:10.1371/journal.pbio.1002533.

Shah, P., and Swiatlo, E. (2008). A multifaceted role for polyamines in bacterial pathogens. *Mol. Microbiol.* 68, 4–16. doi:10.1111/j.1365-2958.2008.06126.x.

Sharma, A. K., and FitzGerald, D. (2010). *Pseudomonas* exotoxin kills Drosophila S2 cells via apoptosis. *Toxicon* 56, 1025–1034. doi:10.1016/j.toxicon.2010.07.007.

Siguier, P. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32–D36. doi:10.1093/nar/gkj014.

Simon, R., Priefer, U., and Pühler, A. (1983). A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* 1, 784–791. doi:10.1038/nbt1183-784.

Simonson, J., and Siebeling, R. J. (1986). Rapid serological identification of *Vibriovulnificus* by anti-H coagglutination. *Appl. Environ. Microbiol.* 52, 1299–304.

Singleton, F. L., Attwell, R., Jangi, S., and Colwell, R. R. (1982a). Effects of temperature and salinity on *Vibriocholerae* growth. *Appl. Environ. Microbiol.* 44, 1047–58.

Singleton, F. L., Attwell, R. W., Jangi, M. S., and Colwell, R. R. (1982b). Influence of salinity and organic nutrient concentration on survival and growth of *Vibriocholerae* in aquatic microcosms. *Appl. Environ. Microbiol.* 43, 1080–5.

Skulachev, V. P. (1998). Uncoupling: new approaches to an old problem of bioenergetics. *Biochim. Biophys. Acta* 1363, 100–24.

Snipen, L., Almøy, T., and Ussery, D. W. (2009). Microbial comparative pan-

genomics using binomial mixture models. *BMC Genomics* 10, 385. doi:10.1186/1471-2164-10-385.

Söderberg, M. A., Rossier, O., and Cianciotto, N. P. (2004). The type II protein secretion system of *Legionella pneumophila* promotes growth at low temperatures. *J. Bacteriol.* 186, 3712–20. doi:10.1128/JB.186.12.3712-3720.2004.

Söding, J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21, 951–60. doi:10.1093/bioinformatics/bti125.

Sokol, P. A., Luan, M. Z., Storey, D. G., and Thirukkumaran, P. (1994). Genetic rearrangement associated with in vivo mucoid conversion of *Pseudomonas aeruginosa* PAO is due to insertion elements. *J. Bacteriol.* 176, 553–62. doi:8300510.

Starks, A. M., Schoeb, T. R., Tamplin, M. L., Parveen, S., Doyle, T. J., Bomeisl, P. E., Escudero, G. M., and Gulig, P. A. (2000). Pathogenesis of infection by clinical and environmental strains of *Vibriovulnificus* in iron-dextran-treated mice. *Infect. Immun.* 68, 5785–93.

Stauder, M., Huq, A., Pezzati, E., Grim, C. J., Ramoino, P., Pane, L., Colwell, R. R., Pruzzo, C., and Vezzulli, L. (2012). Role of GbpA protein, an important virulence-related colonization factor, for *Vibriocholerae*'s survival in the aquatic environment. *Environ. Microbiol. Rep.* 4, 439–45. doi:10.1111/j.1758-2229.2012.00356.x.

Stern, A., Mick, E., Tirosh, I., Sagy, O., and Sorek, R. (2012). CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. *Genome Res.* 22, 1985–94. doi:10.1101/gr.138297.112.

Steuber, J. (2001). Na(+) translocation by bacterial NADH:quinone oxidoreductases: an extension to the complex-I family of primary redox pumps. *Biochim. Biophys. Acta* 1505, 45–56.

Steuber, J., Vohl, G., Casutt, M. S., Vorburger, T., Diederichs, K., and Fritz, G. (2014). Structure of the V. *cholerae* Na+-pumping NADH:quinone oxidoreductase. *Nature* 516, 62–67. doi:10.1038/nature14003.

Stroeher, U. H., Parasivam, G., Dredge, B. K., and Manning, P. A. (1997). Novel *Vibriocholerae* O139 genes involved in lipopolysaccharide biosynthesis. *J. Bacteriol.* 179, 2740–7.

Su, H., Blain, F., Musil, R. A., Zimmermann, J. J., Gu, K., and Bennett, D. C. (1996). Isolation and expression in *Escherichia coli* of hepB and hepC, genes coding for the glycosaminoglycan-degrading enzymes heparinase II and heparinase III, respectively, from Flavobacterium heparinum. *Appl. Environ. Microbiol.* 62, 2723–34.

Suttle, C. a (2007). Marine viruses--major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi:10.1038/nrmicro1750.

Swan, B. K., Tupper, B., Sczyrba, A., Lauro, F. M., Martinez-Garcia, M., González, J. M., Luo, H., Wright, J. J., Landry, Z. C., Hanson, N. W., et al. (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc. Natl. Acad. Sci. U. S. A.* 110, 11463–8. doi:10.1073/pnas.1304246110.

Tamplin, M. L., Specter, S., Rodrick, G. E., and Friedman, H. (1983). Differential complement activation and susceptibility to human serum bactericidal action by *Vibrio* species. *Infect. Immun.* 42, 1187–90.

Tamplin, M. L., Specter, S., Rodrick, G. E., and Friedman, H. (1985). *Vibriovulnificus* resists phagocytosis in the absence of serum opsonins. *Infect. Immun.* 49, 715–8.

Tao, Z., Larsen, A. M., Bullard, S. A., Wright, A. C., and Arias, C. R. (2012). Prevalence and Population Structure of *Vibriovulnificus* on Fishes from the Northern Gulf of Mexico. *Appl. Environ. Microbiol.* 78, 7611–7618. doi:10.1128/AEM.01646-12.

Teeling, H., Waldmann, J., Lombardot, T., Bauer, M., and Glöckner, F. O. (2004). TETRA: a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. *BMC Bioinformatics* 5, 163. doi:10.1186/1471-2105-5-163.

von Tils, D., Blädel, I., Schmidt, M. A., and Heusipp, G. (2012). Type II secretion in *Yersinia*-a secretion system for pathogenicity and environmental fitness. *Front. Cell. Infect. Microbiol.* 2, 160. doi:10.3389/fcimb.2012.00160.

Tison, D. L., Nishibuchi, M., Greenwood, J. D., and Seidler, R. J. (1982). *Vibriovulnificus* biogroup 2: new biogroup pathogenic for eels. *Appl. Environ. Microbiol.* 44, 640–6.

Tseng, C. H., and Tang, S. L. (2014). Marine microbial metagenomics: From

individual to the environment. *Int. J. Mol. Sci.* 15, 8878–8892. doi:10.3390/ijms15058878.

Tsukamoto, K. (2006). Oceanic biology: Spawning of eels near a seamount. *Nature* 439, 929–929. doi:10.1038/439929a.

Unemoto, T., and Hayashi, M. (1993). Na(+)-translocating NADH-quinone reductase of marine and halophilic bacteria. *J. Bioenerg. Biomembr.* 25, 385–91.

Valiente, E., Lee, C. T., Lamas, J., Hor, L., and Amaro, C. (2008a). Role of the virulence plasmid pR99 and the metalloprotease Vvp in resistance of *Vibriovulnificus* serovar E to eel innate immunity. *Fish Shellfish Immunol.* 24, 134–41. doi:10.1016/j.fsi.2007.10.007.

Valiente, E., Lee, C.-T., Hor, L.-I., Fouz, B., and Amaro, C. (2008b). Role of the metalloprotease Vvp and the virulence plasmid pR99 of *Vibriovulnificus* serovar E in surface colonization and fish virulence. *Environ. Microbiol.* 10, 328–38. doi:10.1111/j.1462-2920.2007.01454.x.

Vega Thurber, R., Willner-Hall, D., Rodriguez-Mueller, B., Desnues, C., Edwards, R. a, Angly, F., Dinsdale, E., Kelly, L., and Rohwer, F. (2009). Metagenomic analysis of stressed coral holobionts. *Environ. Microbiol.* 11, 2148–63. doi:10.1111/j.1462-2920.2009.01935.x.

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D., Paulsen, I., Nelson, K. E., Nelson, W., et al. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. 304, 66–74.

Veyrier, F. J., and Cellier, M. F. (2015). Metal economy in host-microbe interactions. *Front. Cell. Infect. Microbiol.* 4. doi:10.3389/fcimb.2014.00190.

Vezzulli, L., Grande, C., Reid, P. C., Hélaouët, P., Edwards, M., Höfle, M. G., Brettar, I., Colwell, R. R., and Pruzzo, C. (2016). Climate influence on *Vibrio* and associated human diseases during the past half-century in the coastal North Atlantic. *Proc. Natl. Acad. Sci.* 113, E5062–E5071. doi:10.1073/pnas.1609157113.

Volesky, B., and Holan, Z. R. (1995). Biosorption of heavy metals. *Biotechnol. Prog.* 11, 235–250. doi:10.1021/bp00033a001.

Voorhies, A. A., and Lorenzi, H. A. (2016). The Challenge of Maintaining a Healthy Microbiome during Long-Duration Space Missions. *Front. Astron. Sp. Sci.* 3, 1–7. doi:10.3389/fspas.2016.00023.

de Vries, S. P. W., Rademakers, R. J. A., van der Gaast-de Jongh, C. E., Eleveld, M. J., Hermans, P. W. M., and Bootsma, H. J. (2014). Deciphering the genetic basis of *Moraxella catarrhalis* complement resistance: a critical role for the disulphide bond formation system. *Mol. Microbiol.* 91, 522–37. doi:10.1111/mmi.12475.

Waldor, M. K., and Mekalanos, J. J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272, 1910–1914. doi:10.1126/science.272.5270.1910.

Wang, X., Wang, Q., Xiao, J., Liu, Q., Wu, H., Xu, L., and Zhang, Y. (2009). Edwardsiella tarda T6SS component evpP is regulated by esrB and iron, and plays essential roles in the invasion of fish. *Fish Shellfish Immunol.* 27, 469–77. doi:10.1016/j.fsi.2009.06.013.

Weber, W., Henkes, H., Möller, P., Bade, K., and Kühne, D. (1998). Toxic spongiform leucoencephalopathy after inhaling heroin vapour. *Eur. Radiol.* 8, 749–55. doi:10.1007/s003300050467.

Webster, A. C., and Litwin, C. M. (2000). Cloning and characterization of vuuA, a gene encoding the *Vibriovulnificus* ferric vulnibactin receptor. *Infect. Immun.* 68, 526–34.

Westermann, A. J., Gorski, S. A., and Vogel, J. (2012). Dual RNA-seq of pathogen and host. *Nat. Rev. Microbiol.* 10, 618–630. doi:10.1038/nrmicro2852.

Wilkins, D., Yau, S., Williams, T. J., Allen, M. A., Brown, M. V, DeMaere, M. Z., Lauro, F. M., and Cavicchioli, R. (2013). Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol. Rev.* 37, 303–35. doi:10.1111/1574-6976.12007.

Williams, T. C., Blackman, E. R., Morrison, S. S., Gibas, C. J., and Oliver, J. D. (2014). Transcriptome sequencing reveals the virulence and environmental genetic programs of *Vibriovulnificus* exposed to host and estuarine conditions. *PLoS One* 9, e114376. doi:10.1371/journal.pone.0114376.

Wommack, K. E., Bhavsar, J., Polson, S. W., Chen, J., Dumas, M., Srinivasiah, S., Furman, M., Jamindar, S., and Nasko, D. J. (2012). VIROME: a

standard operating procedure for analysis of viral metagenome sequences. *Stand. Genomic Sci.* 6, 427–39. doi:10.4056/sigs.2945050.

Wong, E., Vaaje-Kolstad, G., Ghosh, A., Hurtado-Guerrero, R., Konarev, P. V, Ibrahim, A. F. M., Svergun, D. I., Eijsink, V. G. H., Chatterjee, N. S., and van Aalten, D. M. F. (2012). The *Vibriocholerae* colonization factor GbpA possesses a modular structure that governs binding to different host surfaces. *PLoS Pathog.* 8, e1002373. doi:10.1371/journal.ppat.1002373.

Wong, T. Y., Preston, L. A., and Schiller, N. L. (2000). ALGINATE LYASE: review of major sources and enzyme characteristics, structure-function analysis, biological roles, and applications. *Annu. Rev. Microbiol.* 54, 289–340. doi:10.1146/annurev.micro.54.1.289.

Wright, A. C., Morris, J. G., Maneval, D. R., Richardson, K., and Kaper, J. B. (1985). Cloning of the cytotoxin-hemolysin gene of *Vibriovulnificus*. *Infect. Immun.* 50, 922–4.

Wright, A. C., Powell, J. L., Kaper, J. B., and Morris, J. G. (2001). Identification of a Group 1-Like Capsular Polysaccharide Operon for *Vibriovulnificus*. *Infect. Immun.* 69, 6893–6901. doi:10.1128/IAI.69.11.6893-6901.2001.

Wright, A. C., Simpson, L. M., Oliver, J. D., and Morris, J. G. (1990). Phenotypic evaluation of acapsular transposon mutants of *Vibriovulnificus*. *Infect. Immun.* 58, 1769–73.

Wu, S., Tian, J., Wang, G., Li, W., and Zou, H. (2012). Characterization of bacterial community in the stomach of yellow catfish (*Pelteobagrus fulvidraco*). *World J. Microbiol. Biotechnol.* 28, 2165–74. doi:10.1007/s11274-012-1022-5.

Wugeditsch, T., Paiment, A., Hocking, J., Drummelsmith, J., Forrester, C., and Whitfield, C. (2001). Phosphorylation of Wzc, a tyrosine autokinase, is essential for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *J. Biol. Chem.* 276, 2361–71. doi:10.1074/jbc.M009092200.

Yamaichi, Y., Chao, M. C., Sasabe, J., Clark, L., Davis, B. M., Yamamoto, N., Mori, H., Kurokawa, K., and Waldor, M. K. (2015). High-resolution genetic analysis of the requirements for horizontal transmission of the ESBL plasmid from *Escherichia coli* O104:H4. *Nucleic Acids Res.* 43, 348–360. doi:10.1093/nar/gku1262.

Yamaichi, Y., Chao, M. C., Sasabe, J., Clark, L., Davis, B. M., Yamamoto, N.,

Mori, H., Kurokawa, K., and Waldor, M. K. (2014). High-resolution genetic analysis of the requirements for horizontal transmission of the ESBL plasmid from *Escherichia coli* O104:H4. *Nucleic Acids Res.*, 1–13. doi:10.1093/nar/gku1262.

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., et al. (2012). Human gut microbiome viewed across age and geography. *Nature*. doi:10.1038/nature11053.

Yi, S.-W., You, M.-J., Cho, H.-S., Lee, C.-S., Kwon, J.-K., and Shin, G.-W. (2013). Molecular characterization of Aeromonas species isolated from farmed eels (Anguilla japonica). *Vet. Microbiol.* 164, 195–200. doi:10.1016/j.vetmic.2013.02.006.

Yoshida, S., Ogawa, M., and Mizuguchi, Y. (1985). Relation of capsular materials and colony opacity to virulence of *Vibriovulnificus*. *Infect. Immun.* 47, 446–51.

Zapun, A., Contreras-Martel, C., and Vernet, T. (2008). Penicillin-binding proteins and β-lactam resistance. *FEMS Microbiol. Rev.* 32, 361–385. doi:10.1111/j.1574-6976.2007.00095.x.

Zar, F. A., Bakkanagari, S. R., Moorthi, K. M. L. S. T., and Davis, M. B. (2007). A Comparison of Vancomycin and Metronidazole for the Treatment of *Clostridium difficile*-Associated Diarrhea, Stratified by Disease Severity. *Clin. Infect. Dis.* 45, 302–307. doi:10.1086/519265.

Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–9. doi:10.1101/gr.074492.107.

Zhang, X., Zhang, D., Jia, H., Feng, Q., Wang, D., Liang, D., Wu, X., Li, J., Tang, L., Li, Y., et al. (2015). The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat. Med.* 21, 895–905. doi:10.1038/nm.3914.

Zhou, W., Bertsova, Y. V, Feng, B., Tsatsos, P., Verkhovskaya, M. L., Gennis, R. B., Bogachev, A. V, and Barquera, B. (1999). Sequencing and preliminary characterization of the Na+-translocating NADH:ubiquinone oxidoreductase from *Vibrioharveyi*. *Biochemistry* 38, 16246–52.