Proteomics and Multilocus Sequence Analysis confirm the intraspecific variability of *Vibrio tapetis*

Sabela Balboa[1], José Bermudez-Crespo[2], Cristina Gianzo[2], José L. López[2†], and Jesús L. Romalde[1*]

[1]Departamento de Microbiología y Parasitología, and [2]Departamento de Genética. CIBUS. Universidad de Santiago de Compostela. Campus Sur s/n. 15782. Santiago de Compostela. Spain.

Running title: Study of the variability of *V. tapetis* by 2D-PAGE and MLSA

Submitted to: **FEMS Microbiology Letters**, August 2011.

Revised Ms FEMSLE -11-05-0439

† In loving memory of J.L. López who died of cancer during the course of this work.

[*]Corresponding author:

Telf.: (+34) 981 563100 ext. 16961

Fax: (+34) 981 528085

E-mail: jesus.romalde@usc.es

For Peer Review

ABSTRACT

Vibrio tapetis is the etiological agent of brown ring disease (BRD) in clams. Phenotypic, antigenic and genetic variability has been demonstrated, being established three groups associated with host origin. In this work we analyze the variability of representative strains of these 3 groups, CECT 4600^T and GR0202RD, isolated from Manila clam and carpet-shell clam respectively, and HH6087 isolated from halibut, on the basis of the whole proteome analysis by 2D-PAGE and MLSA. A quantitative analysis of the proteome showed that match coefficient rendered a similarity of 79% between the clam isolates while fish isolate showed similarities lower than 70%. A preliminary mass spectrometry (MS) assay allowed the identification of 27 proteins including 50S ribosomal protein L9, riboflavin synthase β subunit, ribose-phosphate pyrophosphokinase and succinyl-CoA synthase α subunit, among others. MLSA approach rendered similar results, showing the clam isolates a similarity of 99.4 %, higher than those observed between the fish isolate and both clam strains (98.2%). The topology of the Maximum Parsimony tree, obtained from 2D-PAGE analysis, and the phylogenetic tree, constructed with the Maximum Likelihood algorithm from concatenated sequences of 16S rRNA gene and five housekeeping genes (*atp*A, *pyr*H, *rec*A, *rpo*A, and *rpo*D), was very similar, confirming the closer relationship between both clam isolates.

Keywords: *Vibrio tapetis*, 2D-PAGE, MLSA, intraspecific variability.

ScholarOne Support 1-434/964-4100

INTRODUCTION

Bacterial species show an enormous variability at biochemical, serological and genetic levels. Characterizing strains of a given species is important for understanding the ecology and competitiveness in environmental samples and in epidemiological studies. *Vibrio* species are extensively distributed on marine environments being associated with a wide range of marine organisms and some are pathogenic to humans (Thompson *et al.,* 2006; Beaz-Hidalgo *et al*., 2010). Genotyping strategies such as restriction fragment length polymorphism (RFLP) and pulse field gel electrophoresis (PFGE) were traditionally used for epidemiological analysis of *Vibrio* isolates (Castro *et al*., 1997; Romalde *et al*., 2002). PCR typing methods have also been widely used among this bacterial genus such as Randomly amplified polymorphic DNA analysis (RAPD) and Repetitive-sequence-based polymerase chain reaction (Rep-PCR) based on both polymorphic, repetitive extragenic palindromic sequences (REP) and enterobacterial repetitive intergenic consensus (ERIC) (Rodríguez *et al*., 2006). More recently, amplified fragment length polymorphism (AFLP) and multilocus sequence analysis (MLSA) have allowed a more precise identification of *Vibrio* species (Beaz-Hidalgo *et al*., 2008; 2010; and references therein). MLSA is a typing method based on the concatenation of sequences of several housekeeping genes (Maiden, 2006). This technique has been successfully used on epidemiological studies (González-Escalona *et al*., 2008), as well as in species delineation (Stackebrandt *et al*., 2002; Stackebrandt & Ebers 2006), and the study of relationships between related groups (Thompson *et al*., 2008; Pascual *et a*l., 2010). This approach has demonstrated to be very useful and accurate for vibrios (Thompson *et al.,* 2005, 2009; Sawabe *et al*., 2007; Beaz-Hidalgo *et al*., 2010), being very congruent with results of DNA-DNA hybridization (Thompson *et al.*2009).
Proteomics could complement and extend the nucleic acid analytical technologies being an experimental link between the expressed product and the genome (Lester & Hubband, 2002; Philips & Boygo, 2005; Norbeck *et al*., 2006; Cash, 2009; Zhang *et al*., 2010). 2D-PAGE has been successfully applied for the discrimination of closely related isolates as demonstrated in several studies (Cash *et al*., 1995; Dumas *et al*., 2008), revealing even more degree of variability than with DNA-DNA hybridization, since protein content reflects dynamic changes produced in the cells as a response changes in the environment (Andersen *et al*., 1984; Cash, 2009; Zhang *et al*., 2010).
Vibrio tapetis is the causative agent of an epizootic infection described in adult clams called Brown Ring Disease (Borrego *et al*., 1996). First studies indicated that strains of this pathogen constituted a homogenous group. However, as new strains were isolated from different hosts, including

ScholarOne Support 1-434/964-4100

For Peer Review

different mollusk and fish species, some variability on the basis of their antigenic, phenotypic and genotypic characteristics has been demonstrated, leading to the description of three main groups within this species that correlate with the type of host (Castro *et al.* 1996, 1997; Romalde *et al.* 2002; Rodriguez *et al.* 2006).

In this work, a proteomic method, 2D-PAGE, was used to study the intraspecific variability of representative strains of the three groups described for *V. tapetis*, as well as an additional indication of their phylogenetic relationship. The results obtained were compared with those of MLSA, a well established genetic technique to infere bacterial phylogeny.

MATERIAL AND METHODS

Bacterial strains

Representative strains of the three previously described groups of *Vibrio tapetis* (Rodríguez *et al.*, 2006) with different phenotypical, serological and genetic profiles as well as different host origin were used on this study: CECT 4600^T, type strain of the species isolated from Manila clam (*Ruditapes philippinarum*), GR0202RDRD obtained from Carpet shell clam (*R. decussatus*) and HH6087 isolated from Halibut (*Hipoglossus hipoglossus*)(Borrego *et al.*, 1996; Novoa *et al.*, 1998; Reid *et al.*, 2003). The bacteria were routinely aerobically grown on Marine Agar (MA) (Pronadisa, Spain) at 15ºC for 72h. Stock cultures were maintained frozen at -80°C in Marine Broth (MB) (Pronadisa) supplemented with 15% glycerol (v/v).

2D-Electrophoresis

Growth conditions

Bacterial inocula were prepared resuspending each strain to achieve 10^9 cells/ml adjusting the bacterial suspension to an OD=1 (580 nm) using a Lambda2 UV/VIS Spectophotometer (Perkin-Elmer, Germany). One flask per strain, containing 1 liter of sterile MB, was inoculated to achieve a final concentration of 10^5 cells/ml and were aerobically incubated in a Innova 4340 rotary shaker (70 rpm) (New Brunswick Scientific, USA) at 15ºC for 72 hours.

Sample preparation

Bacteria were harvested and washed with Tris-buffered sucrose (10 mmol Tris, 250 mmol sucrose pH 7). Then, samples were lyophilised. Proteins were extracted by suspending 40 mg of lyophilised bacteria in 1 mL standard lysis buffer (7M urea, 2M thiourea, 4% CHAPS (3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate) and 65 mM DTT (Dithiothreitol) during 3 hours at 27 ºC.

ScholarOne Support 1-434/964-4100

For Peer Review

and sonicating (3 cycles of 10 pulses). Next, samples were centrifuged at 14000 rpm for 30 min and supernatants were collected and subjected to protein precipitation using the Clean-up kit (GE Healthcare, Sweden). After suspension of the pellet in 1 mL of lysis buffer, protein concentration were measured with CB-X protein assay kit (Gbiosciences, USA). Finally, samples were stored at -80 ºC prior to use.

First dimension

Isoelectrofocusing (IEF) was performed using a Protean IEF cell (Bio-Rad, USA) and 24 cm pH 4-7 IPG strips (GE Healthcare). For each sample, 400 µg of protein were resuspended in 390 µl of rehydration buffer (7M urea, 2M thiuorea, 4% CHAPS, 0.6 % DTT, 1% IPG buffer 4-7 and bromophenol blue traces). IEF was carried out at 20 ºC and according to the following steps: active rehydration (50V) for 12h, 250 V for 30 min, 500 V for 1h, 1000 V for 1h, 4000 V for 2h, 8000 V for 2 h and 10000 V to achieve 65 kVh.

Second dimension

Prior to run the second dimension, strips were equilibrated at room temperature for 15 min with an equilibration solution (6M urea, 50mM Tris-HCl pH 8.8, 30% glycerol, 2% SDS (sodium dodecyl sulfate) with addition of 1% DTT for 15 min in the same solution supplemented with of 2.5% of iodoacetamide. Strips were placed on top of a 21 x 26 cm 12,5% polyacrilamide gel and fixed with sealing solution (25mM Tris, 192 mM glycine, 0.1% SDS, 0.5% agarose, 0.01% bromophenol blue). Second dimension was performed according to Laemmli (1970), in an EttanDalt-Six electrophoretic system (GE Healthcare) at 30º C overnight (5mA/gel for 1 h, 10mA/per 1 h, followed by 16 mA/gel until bromophenol blue dye reached the bottom of the gel). Three independent cultures as well as protein extractions and 2D-PAGE were performed to assess the reproducibility of the experiment.

Protein visualization and image analysis

Gels were stained with Coomassie Brilliant Blue (CBB). CBB staining was carried out according to Neuhoff *et al*. (1988) with minor modifications and scanned in a densitometer, Microtek 9800XL (Microtek, USA), at 300 dpi resolution. Gels were stored in vacumm-sealed plastic bags at 4ºC. PDQuest Advance software version 8.0 (Bio-Rad) were used for spot detection, spot quantitation and reproducibility study.

In-gel digestion and peptide extraction

Protein spots chosen for mass spectrometric analysis (MS) were excised from the gels and manually digested. The gel pieces were rinsed thrice with AmBic buffer (50 mM ammonium bicarbonate in 50% HPLC grade methanol (Scharlau, Spain) and once with 10 mM dithiothreitol (Sigma-Aldrich,

For Peer Review

USA). The gel pieces were rinsed twice with AmBic buffer and dried in a SpeedVac before alkylation with 55 mM iodoacetamide (Sigma- Aldrich) in 50 mM ammonium bicarbonate. Once again, the gel pieces were rinsed with HPLC grade AmBic buffer (Scharlau), before being dehydrated by addition of HPLC grade acetonitrile (Scharlau) and dried in a SpeedVac. Modified porcine trypsin (Promega, USA) was added to the dry gel pieces at a final concentration of 20 ng/µl in 20 mM ammonium bicarbonate, incubating them at 37ºC for 16 h. Peptides were extracted thrice by 20 min incubation in 40 µl of 60% acetonitrile in 0.5% HCOOH (formic acid). The resulting peptide extracts were pooled, concentrated in a SpeedVac and stored at -20ºC.

Mass spectrometric analysis and protein identification

A combination of matrix-assisted laser-desorption-ionization time-of-flight mass spectrometry (MALDI-TOF) (MS) and MALDI-TOF/TOF (MS/MS) was used for protein identification according to the following procedure. Dried samples were dissolved in 4 µl of 0.5% formic acid. Equal volumes (0.5 µl) of peptide and matrix solution, consisted of 3 mg CHCA (α-Cyano-4-hidroxycinnamic acid) dissolved in 1 ml of 50% acetonitrile in 0.1% TFA (trifluoroacetic acid), were deposited using the thin layer method, onto a 384 Opti-TOF MALDI plate (Applied Biosystems, USA). Mass spectrometric data were obtained in an automated analysis loop using 4800 MALDI-TOF/TOF analyzer (Applied Biosystems). MS spectra were acquired in reflectron positive-ion mode with a Nd:YAG, 355 nm wavelength laser, averaging 1000 laser shots and at least three trypsin autolysis peaks were used as internal calibration. All MS/MS spectra were performed by selection the precursors with a relative resolution of 300 full width at half maximum (FWHM) and metastable suppression. Automated analysis of mass data was achieved using the 4000 Series Explorer Software V3.5. Peptide mass fingerprinting (PMF) and peptide fragmentation spectra data of each sample were combined through the GPS Explorer Software v3.6 using Mascot software v2.1. (Matrix Science) to search against a non-identical protein database (NCBInr release data 20100526), with 30 ppm precursor tolerance, 0.35 Da MS/MS (analysis of the tandem mass) fragment tolerance, CAM (carbamydomethyl cystein) as fixed modification, oxidized methionine as variable modification and allowing 1 missed cleavage. All spectra and database results were manually inspected in detail using the previous software. Protein scores greater than 56 were accepted as statistically significant ($P < 0.05$), considering positive the identification when protein score CI (confidence Interval) was above 98%. In case of MS/MS spectra, total ion score CI was above 95%.

For Peer Review

Statistical analysis

Similarity percentages between *V. tapetis* isolates were calculated on the basis of protein profile similarities calculated between pairs of isolates using the Simple Matching co-efficient (Sneath and Sokal, 1973).

Phylogenetic analysis based on proteins

BioNumerics 5.1 2D software (Applied-Maths, USA) was used to construct a Maximum Parismony tree based on the different protein content of *V. tapetis* isolates.

Phylogenetic analysis based on MLSA

Genomic DNA extraction and amplification of the 16S rRNA gene was performed as previously described (Beaz-Hidalgo *et al.*, 2008). Sequences for five protein-coding housekeeping genes *atp*A (α subunit of ATPase), *pyr*H (uridyl monophosphate kinase), *rec*A (recombinase A), *rpo*A (α subunit of RNA polymerase), and *rpo*D (RNA polymerase sigma factor) were performed according to Thompson *et al.* (2004, 2005, 2007) and Pascual *et al.* (2010). Sequencing reactions were performed with the GenomeLab DTCS-Quick Start Kit (Beckman Coulter, Ireland). Sequence data analysis was performed with DNAstar Seqman program (Lasergene, USA). Percentage of similarity of concatenated sequence of genes was calculated using DNAstar Megalingn program (Lasergene). For maximum-likelihood (ML) analysis, the optimal model of nucleotide substitution was estimated throug the program jmodeltest 0.1.1 (Posada, 2008) using the Akaike Information Criterion (AIC). The ML estimation was implemented in PHYML (Gindon & Gascuel, 2003), using GTR model as recommended by jmodeltest 0.1.1. Bootstrap analyses were performed using 1000 replications.

RESULTS

Automatic software analysis was used to determine the protein spot number in each 2D gel. Total amount of protein loaded on each gel was the same for the three strains. However, there were important differences in the number of spots among the three strains. Thus, average spot number detected was 729 spots (± 13, standard deviation), 681 spots (± 2) and 556 (± 6) for CECT 4600^{T}, GR0202RD and HH6087, respectively. Figure 1 shows standard protein pattern of each strain separated by 2-DE using a 24 cm, linear pH 4-7 IPG strip in the first dimension. Technical variation of 2-DE was determined by correlation coefficients between gel replicates of each strain. These correlation coefficients showed high degree of congruence for each strain, being all of them about

ScholarOne Support 1-434/964-4100

For Peer Review

0.90 (0.91 for CECT 4600^T and GR0202RD, 0.85 for HH6087), clearly showing high reproducibility (Fig. 1).

Visual inspection of gels showed that the majority of proteins detected were localised in the acidic part of the pH range studied and they also showed similar or different protein profile depending on a specific molecular weight region (Fig. 2). Thus, high molecular weight region is very similar between all strains while low molecular weight is more similar between CECT 4600^T and GR0202RD strains than between CECT 4600^T and HH6087 strains. Besides, protein profile of middle region seems to be specific for each strain although some common spots are also present in this region (Fig. 2).

Sixty representative proteins (common and unique for each strain) of the three strains were selected and sequenced by MS but only 27 of these proteins were identified (Table 1). Only one of 6 common proteins present in the three strains was identified as 50S ribosomal protein L9. Of the eight proteins selected in common for CECT4600^T and GR0202RD, five were identified as Riboflavin synthase β subunit, Ribose-phosphate pyrophosphokinase, Succinyl-CoA synthase α subunit, 3-hydroxydecanoyl-(acyl carrier protein) dehydratase and peptidyl-prolyl cis-trans isomerase B (rotamase B). One and two common proteins between CECT4600^T and HH6087 and between RG0202RD and HH6087 respectively were selected but no identification for any of them was achieved. Regarding to the strain-specific proteins, 10 out 20 spots were identified, including intregrases, transcriptional regulators and ABC transporter. Interestingly, two proteins selected as unique for CECT4600^T and GR0202RD results to be the same a hypothetical protein pVT1_26.

Level of protein profile similarity within *Vibrio tapetis* was calculated between pairs of strains applying the simple matching co-efficient formula. Results of matching co-efficient showed similarity of 79% between CECT 4600^T and GR0202RD strains, 69% between CECT 4600^T and HH6087 strains and 60% between GR0202RD and HH6087 strains. These results were used to construct an un-rooted tree (Fig. 3), which showed that GR0202RD strain was clearly more closely related to CECT 4600^T than HH6087.

Fragments of the 16S rRNA gene (1531 bp) and five coding-protein housekeeping genes, *rpo*D (535 bp), *rpo*A (863 bp), *pyr*H (540 bp), *atp*A (1194 bp) and *rec*A (789 bp), were sequenced to yield a concatenated sequence of 4090 nucleotides, which corresponded to more than 80% of the coding regions of each gene. Similarity percentages between concatenated sequences of the three isolates were 99.4% between CECT 4600T^T and GR0202RD, 98.2% between CECT 4600T^T and

For Peer Review

HH6087 and 98.2% between GR0202RD and HH6087. These results indicate higher similarity between clam isolates (CECT 4600T^{T} and GR0202RD) than with the fish isolate (HH6087). This similarity can also be seen in the phylogentic tree generated by Maximum Likelihood (ML) in which clam isolates are evolutively closer than the fish isolate (Fig. 3). The independent analysis of the each gene rendered, in general, similar tree topologies, being *rpo*A and *rpo*D the poorest discriminative genes (data not shown).

DISCUSSION

Protein analysis based on gene expression have been based on the 2D-PAGE technique (Klose, 1975; O´Farrell, 1975) which has also proved to be a powerful typing tool with a high degree of discriminatory capacity in intraspecific studies as well as in interspecific analysis of closely related organisms such as *Listeria*, *Ferroplasma*, *Frankia* or *Mycoplasma* (Rodwell & Rodwell, 1978; Benson *et al*., 1984; Gormon & Phan-Thanh, 1995; Dopson *et al.* 2004; Dumas *et al*., 2008). However, its value to determine taxonomic positions of distant organisms is limited as demonstrated by Knight *et al.* (2004) who, in an *in silico* work performed with more than 100 eukaryotic and prokaryotic species, observed little or no relationship between 2D gels and phylogeny.

Automatic software analysis revealed differences in protein spot number, ranging from 729 spots for strain CECT 4600^{T} to 556 for strain HH6087. Similarity of protein profiles was higher between strains isolated from clam species (CECT 4600^{T} and GR0202RD), than between these strains and the fish isolate (HH6087). Spot number and the similarity percentages between the *V. tapetis* strains are in agreement with those reported in previuos studies for other bacterial species (Gormon & Phan-Thanh, 1995; Govorun *et al*., 2003; Dopson *et al*., 2004). It has been described that for this proteomic approach, as for all the phenotypic methods for classification of microorganisms, a crucial step is to grow the cells under the same conditions to be sure that the observed changes are due to strain differences rather than growth (Dopson *et al*., 2004). Cells employed in this study were grown in the same media and conditions, obtaining a high reproducibility (correlation coefficients of 0.9) for the different protein extractions and gels.

The majority of proteins detected, regardless of the strain, were localised in the acidic part of the pH range studied. This finding agrees with results of other authors who observed in halophilic bacteria a predominance of proteins with low p*I* over high p*I* (Kiraga *et al*., 2007). The identified proteins

could be related with important functions in the cells, such as 50S ribosomal protein L9, metabolic pathways, including riboflavin synthase β subunit, ribose-phosphate pyrophosphokinase, and peptidyl-prolyl cis-trans isomerase B (rotamase B), as well as intregrases, transcriptional regulators and ABC transporter. The identification as the same hypothetical protein pVT1_26 of two nonequivalent spots in the profiles of strains CECT 4600^{T} and GR0202RD indicates the possibility of existence of charge variants of the same protein between isolates, as previously described for other bacterial species (Cash, 2009). The lack of a complete genome sequence for *V. tapetis* and, therefore, the unavailability of an appropriate database is reflected in our study where only 27 out of the 60 proteins sequenced by MS were identified, making necessary more studies to characterize the proteome of this pathogen.

In comparison with proteomics, genomic procedures as MLSA give the advantages that information is fairly stable, unaffected by the growth conditions of bacteria, and able to generate highly reproducible and portable data, which would allow the comparison of results between laboratories using the public online databases. MLSA has demonstrated to be a powerful, both intra- and interspecific, discriminative tool within the *Vibrio* genus (Thompson *et al.*, 2004, 2005, 2007, 2009; Pascual *et al.*, 2010). The choice of the protein encoding genes for the MLSA is the most important aspect in a correct MLSA analysis, since by the moment any set of housekeeping genes has proved to be universally applicable. This fact is particularly difficult in the case of a set of strains belonging to the same species or to tightly related taxa, due to the need of genes able to measure such low variability. In our case, each selected gene has been previously used for *Vibrio* species (Thompson *et al.*, 2004, 2005, 2007). and the results obtained were in agreement with those reached with genotyping methods (Castro *et al.*1996, 1997; Romalde *et al*. 2002; Rodríguez *et al*.2006)

Both methods, 2D-PAGE and MLSA, rendered trees with similar topology, appearing the clam isolates more closely related than its counterpart from fish. In addition, the relative branching order is clearly in agreement with the three genetic groups previously described on the basis of typing methods (Romalde *et al.,* 2002; Rodríguez *et al.*2006). The congruence between the results obtained in the phylogenetic study of housekeeping genes (conservative approach), and the analysis of the whole proteome of the isolates (dynamic approach), provides also an inter-validation of the techniques.

In conclusion, the proteomic approach using 2D-PAGE can be a useful complementary tool for the study of the intraspecific variability of *Vibrio tapetis*. In addition, it is a method that does not require prior information on the genome sequence and possesses the added value of describing gene

ScholarOne Support 1-434/964-4100

expression at protein level, which can furnish helpful information on host-pathogen interaction and pathogenic processes.

ACKNOWLEDGEMENTS

This work was partially supported by Grants AGL2006-13208-C02-01 and AGL2010-18438 from the Ministerio de Ciencia e Innovación (MICINN) (Spain). The kind donation of strains by Drs. JJ Borrego (University of Málaga, Spain) and TH Birkbeck (University of Galsgow, UK) is greatly acknowledge. SB and JBC thank the Ministerio de Ciencia Innovación (MICINN) and Xunta de Galicia (Spain) for research fellowships.

For Peer Review

1
2
3 339 **REFERENCES**
4
5 340 Andersen H, Christiansen G & Christiansen C (1984) Electrophoretic analysis of proteins from
6
7 341 *Mycoplasma capricolum* and related serotypes using extracts from intact cells and from
8
9
10 342 Minicells containing cloned *Mycoplasma* DNA. *J Gen Microbiol* **130**: 1409-1418.
11
12 343 Beaz-Hidalgo R, Cleenwerck I, Balboa S, De Wachter M, Thompson F, Swings J, De Vos P &
13
14
15 344 Romalde JL (2008) Diversity of Vibrios associated with reared clams in Galicia (NW Spain).
16
17 345 *Syst Appl Microbiol* **31**: 215-222.
18
19 346 Beaz-Hidalgo R, Balboa S, Romalde J & Figueras MJ (2010) Diversity and pathogenecity of *Vibrio*
20
21
22 347 species in cultured bivalve molluscs. *Environm Microbiol Reports* **2**: 34-43.
23
24 348 Benson, DR, Buchholz, SE, Hanna, DG (1984) Identification of *Frankia* strains by Two-
25
26 349 Dimensional Polyacrilamide Gel Electrophoresis. *Appl Environm Microbiol* **47**: 489-494.
27
28
29 350 Borrego JJ, Castro D, Luque A, Paillard C, Maes P, García MT & Ventosa A (1996) *Vibrio tapetis*
30
31 351 sp. nov., the Causative Agent of the Brown Ring Disease Affecting Cultured Clams. *Int J Syst*
32
33 352 *Bacteriol* **46**: 480-484.
34
35
36 353 Cash P (2009) Proteomics in the study of the molecular taxonomy and epidemiology of bacterial
37
38 354 pathogens. *Electrophoresis* **30**: S133-S141.
39
40
41 355 Cash P, Argo E & Bruce KD (1995) Characterization of *Haemophilus influenzae* proteins by two-
42
43 356 dimensional gel electrophoresis. *Electrophoresis* **16**: 135-148.
44
45 357 Castro D, Santamaria JA, Luque A, Martínez-Manzanares E & Borrego JJ (1996) Antigenic
46
47
48 358 characterization of the etiological agent of the brown ring disease affecting manila clams. *Syst*
49
50 359 *Appl Microbiol* **19**: 231-239.
51
52 360 Castro D, Romalde JL, Vila J, Magariños B, Luque A & Borrego JJ (1997) Intraspecific
53
54
55 361 characterization of *Vibrio tapetis* strains by use of pulsed-field gel electrophoresis, ribotyping
56
57 362 and plasmid profiling. *Appl Environ Microbiol* **63**: 1449–1452.
58
59 363 Dopson M, Baker-Austin C & Bond PL (2004) First use of two-dimensional polyacrilamide gel
60
364 electrophoresis to determine phylogenetic relationships. *J Microbiol Methods* **58**: 297-302.

For Peer Review

365 Dumas E, Meunier B, Berdagué J-L, Chambon C, Desvaux M & Hébraud M (2008) Comparative
366 analysis of extracellular and intracellular proteomes of *Listeria monocytogenes* strains reveals a
367 correlation between protein expression and serovar. *Appl Environ Microbiol* **74**: 7399-7409
368 González-Escalona N, Martinez-Urtaza J, Romero J., Espejo RT, Jaykus L-A & DePaola A (2008)
369 Determination of molecular phylogenetics of *Vibrio parahaemolyticus* strains by Multilocus
370 Sequence Typing. *J Bacteriol* **190**: 2831-2840.
371 Gormon T & Phan-Thanh L (1995) Identification and classification of *Listeria* by two-dimensional
372 protein mapping. *Res Micriobiol* **146**: 143-154.
373 Govorun VM, Moshkovskii SA, Tikhonova OV, Goufman EI, Serebryakova MV, Momynaliev KT,
374 Lokhov PG, Khryapova EV, Kudryavtseva LV, Smirnova OV, Toropyguine IY, Maksimov BI &
375 Archakov AI (2003) Comparative analysis of proteome maps of *Helicobacter pylori* clinical
376 isolates. *Biochemistry* **68**:52-60.
377 Guindon S & Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies
378 by maximum likelihood. *Syst Biol* **52**: 696-704.
379 Kiraga J, Mackiewicz P, Mackiewicz D, Kowalczuk M, Biecek P, Polak N, Smolarczyk K, Dudek
380 MR & Cebrat S (2007) The relationships between the isoelectric point and: length of proteins,
381 taxonomy and ecology of organisms. *BMC Genomics* **8**:163.
382 Klose J (1975) Protein mapping by combined isoelectric focusing and electrophoresis in mouse
383 tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 26:
384 231-243.
385 Knight CG, KassenR, Hebestreit H & Rainey PB (2004) Global analysis of predicted proteomes:
386 Functional adaptation of physical properties. *PNAS* **101**: 8390-8395.
387 Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of
388 bacteriophage T4. *Nature* **227**: 680-685.

For Peer Review

389 Lester PJ & Hubbard SJ (2002) Comparative bioinformatic analysis of complete proteomes and
390 protein parameters for cross-species identification in proteomics. *Proteomics* **2**: 1392-1405.

391 Maiden MCJ (2006) Multilocus Sequence Typing of Bacteria. *Annu. Rev. Microbiol* 60: 561-588.

392 Neuhoff V, Arold N, Taube D & Ehrhardt W (1988) Improved staining of proteins in
393 polyacrylamide gels including isoelectric focusing gels with clear background at nanogram
394 sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* **9**:255-62.

395 Norbeck AD, Callister SjJ, Monroe ME, Jaitly N, Elias DA, Lipton MS & Smith RD (2006)
396 Proteomic approaches to bacterial differentiation. *J Microbiol Methods* 67: 473-486.

397 Novoa B, Luque A, Castro D, Borrego JJ & Figueras A (1998) Characterization and infectivity of
398 four bacterial strains isolated from brown ring disease-affected clams. *J Invertebr Pathol* **71**: 34-
399 41.

400 O´Farrell PH (1975) High resolution two-kimensional electrophoresis of proteins. *J Biol Chem* **250**:
401 4007-4021.

402 Pascual J, Macián MC, Arahal DR, Garay E & Pujalte MJ (2010) Multilocus sequence analysis of
403 the central clade of the genus Vibrio by using the 16S rRNA, recA, pyrH, rpoD, gyrB, rctB and
404 toxR genes. *Int J Syst Evol Microbiol.* **60**:154-65.

405 Phillips CI & Bogyo M (2005) Proteomics meets microbiology: technical advances in the global
406 mapping of protein expression and function. *Cell Microbiol* **7**: 1061-1076.

407 Posada D (2008) jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol* **25**: 1253-1256.

408 Reid HI, Duncan HL, Laidler LA, Hunter D & Birkbeck TH (2003) Isolation of *Vibrio tapetis* from
409 cultivated Atlantic halibut (*Hipoglossus hipoglossus* L.). *Aquaculture* **221**: 65-74.

410 Rodwell, AW & Rodwell (1978) Relationships between strains of *Mycoplasma mycoides* subspp.
411 *mycoides* and *capri* studied by Two-dimensional Gel Electrophoresis of cell proteins. *J Gen*
412 *Microbiol* **109**: 259-263.

ScholarOne Support 1-434/964-4100

413 Rodríguez JM, López-Romalde S, Beaz R, Alonso, MC, Castro D & Romalde JL (2006) Molecular
414 fingerprinting of *Vibrio tapetis* strains using three PCR-based methods: ERIC-PCR, REP-PCR,
415 and RAPD. *Dis. Aquat. Org*. **69**:175-183.
416 Romalde JL, Castro D, Magariños B, Lopez-Cortes L & Borrego JJ (2002) Comparison of
417 ribotyping, randomly amplified polymorphic DNA and pulsed-field gel electrophoresis for
418 molecular typing of Vibrio tapetis. *Syst Appl Microbiol* **25**: 544–550.
419 Sawabe T, Kita-Tsukamoto K & Thompon FL (2007) Inferring the evolutionary history of vibrios
420 by means of Multilocus Sequence Analysis. *J Bacteriol* **21**: 7932-7936.
421 Sneath PHA & Sokal RR (1973) *Numerical taxonomy*. Freeman, San Francisco.
422 Stackerbrant E & Ebers J (2006) Taxonomic parameters revisited: tarnished gold standards.
423 *Microbiology today*. **33**: 152-155.
424 Stackebrandt E, Frederiksen W, Garrity GM, *et al.* (2002) Report of the ad hoc committee for the
425 re-evaluation of the species definition in Bacteriology. *Int J Syst Evol Microbiol* **52**: 1043-1047.
426 Thompson CC, Thompson FL, Vandemeulebroecke K, Hoste B, Dawyndt, P & Swings J (2004).
427 Use of recA as an alternative phylogenetic marker in the family *Vibrionaceae*. *Int J Syst Evol*
428 *Microbiol* **54**: 919-929.
429 Thompson CC, Thompson F L, Vicente AC & Swings J (2007). Phylogenetic analysis of vibrios
430 and related species by means of atpA gene sequences. *Int J Syst Evol Microbiol* **57**: 2480-2484.
431 Thompson CC, Thompson FL & Vicente ACP (2008) Identification of *Vibrio cholera* and *Vibrio*
432 *mimicus* by multilocus sequence analysis (MLSA) *Int J Syst Evol Microbiol* **58**: 617-621.
433 Thompson CC, Vicente ACP, Souza R C, Vasconcelos, ATR, Vesth T, Alves Jr N, Ussery DW,
434 Iida T & Thompson FL (2009) Genomic taxonomy of vibrios. *BMC Evolutionary Biology.*
435 **9**:258.

ScholarOne Support 1-434/964-4100

For Peer Review

436 Thompson FL, Gevers D, Thompson CC, Dawyndt P, Naser S, Hoste B, Munn CB & Swings J
437 (2005). Phylogeny and molecular identification of Vibrios on the basis of Multilocus Sequence
438 Analysis. *Appl Environ Microb* **71**: 5107-5115.
439 Thompson FL, Austin B & Swings J (2006) *The biology of Vibrios*. ASM Press. Washington DC.
440 Zhang W, Li F & Nie L (2010) Integrating multiple ´omics´ analysis for microbial biology:
441 application and methodologies. *Microbiology* **156**: 287-301.

For Peer Review

FIGURE LEGENDS

Figure 1: Reference mapping protein of whole cells of *V. tapetis* separated by 2-DE using a 24 cm, linear pH 4-7 IPG strip in the first dimension. (A): Manila clam isolate CECT4600^T, (B): GR0202RD isolated from carpet shell clam and (C) fish isolate HH6087. Proteins were separated on a linear pH 4-7 gradient in the first dimension and visualized using Coommassie G-250 staining. Scatter plots demonstrating the reproducibility of the gels are shown at the upper right of pair of gels.

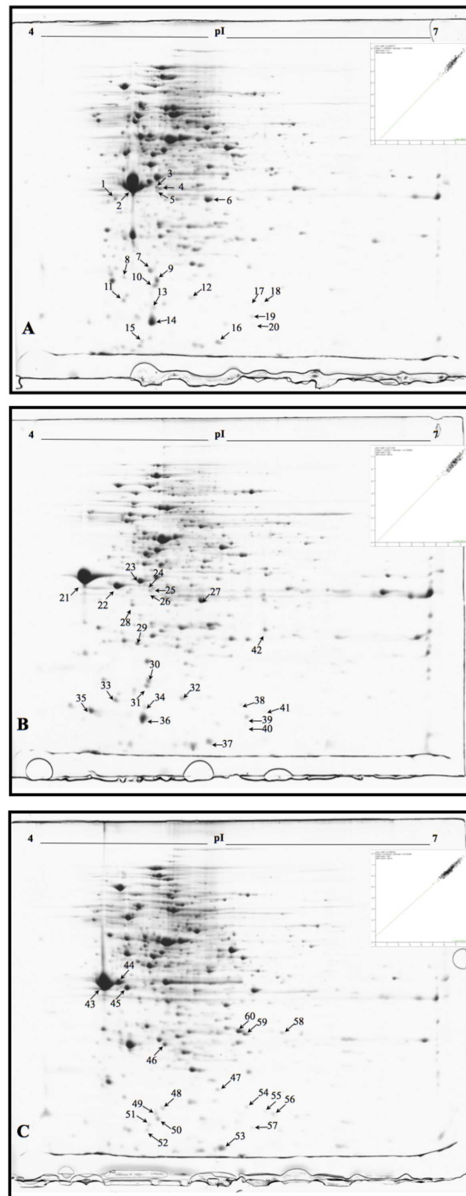
Figure 2: Conserved and variable areas among strains in the gels: A) great similarity in the high molecular weight region; B) hipervariable protein pattern in the middle molecular weight region; and C) variable pattern in low molecular weight region showing more similarity between CECT4600^T and GR0202RD strains.

Figure 3: Comparison of the phylogenetic reconstruction based on concatenated of 16S rRNA, *atpA*, *pyrH*, *rpoA, rpoD* and *recA* partial gene sequences obtained by the Maximum Likelihood method (GTR model) (Black) and the Maximum Parsimony tree generated on the basis of the protein patterns (Grey). Numbers on the branches show the phylogenetic distances.

For Peer Review

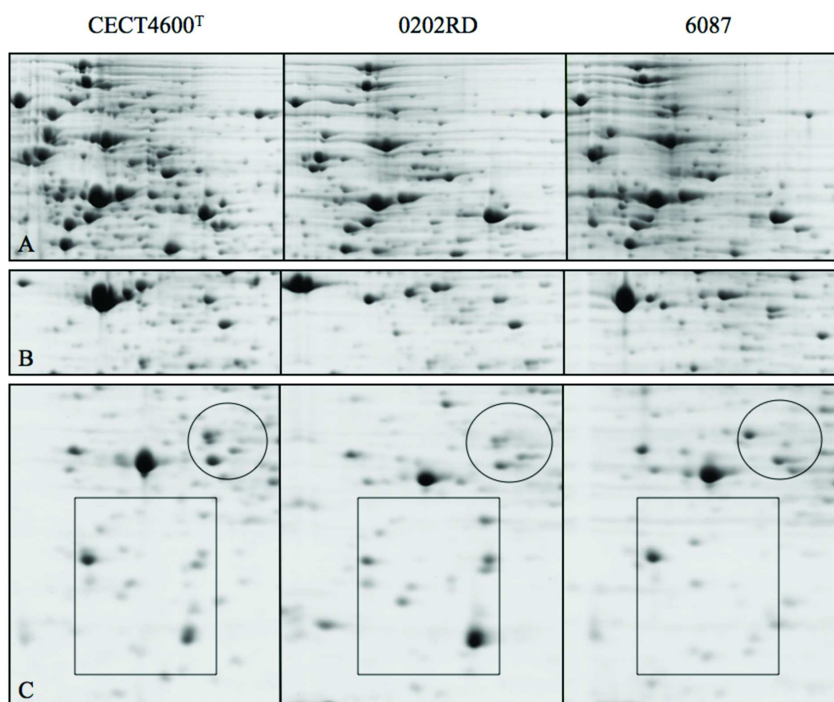
Table 1: Selected proteins of the three strains of *V. tapetis* identified by MS.

Protein	Spot	Strain
50S ribosomal protein L9	16	CECT4600 ^{T}
	37	GR0202RD
	53	HH6087
Peptidyl-prolyl cis-trans isomerase B (rotamase B)	14	CECT4600 ^{T}
	36	GR0202RD
Ribose-phosphate pyrophosphokinase	3	CECT4600 ^{T}
	24	GR0202RD
Succinyl-CoA synthase α subunit	6	CECT4600 ^{T}
	27	GR0202RD
3-hydroxydecanoyl-(acyl carrier protein) dehydratase	17	CECT4600^{T}
	38	GR0202RD
Riboflavin synthase β subunit	40	GR0202RD
	57	HH6087
Hypothetical protein LIC12719	2	CECT4600^{T}
β-lactamase	5	CECT4600^{T}
Hypothetical protein pVT1_26	18	CECT4600^{T}
6,7-dimethyl-8-ribityllumazine synthase	20	CECT4600^{T}
3-Hidroxydecanoyl-ACP dehydratase	38	GR0202RD
Transaldolase	23	GR0202RD
Peroxiredoxin-2	29	GR0202RD
Putative type VI secretion protein VasQ-1	30	GR0202RD
Hypothetical protein MED222_13960	33	GR0202RD
Hypothetical protein pVT1_26	41	GR0202RD
Phage intregrase	48	HH6087
ABC transporter related	55	HH6087
Transcriptional regulator LysR family	59	HH6087
Uridine phosphorilase	60	HH6087



Reference mapping protein of whole cells of V. tapetis separated by 2-DE using a 24 cm, linear pH 4-7 IPG strip in the first dimension. (A): Manila clam isolate CECT4600T, (B): GR0202RD isolated from carpet shell clam and (C) fish isolate HH6087. Proteins were separated on a linear pH 4-7 gradient in the first dimension and visualized using Coommassie G-250 staining. Scatter plots demonstrating the reproducibility of the gels are shown at the upper right of pair of gels.
101x254mm (300 x 300 DPI)

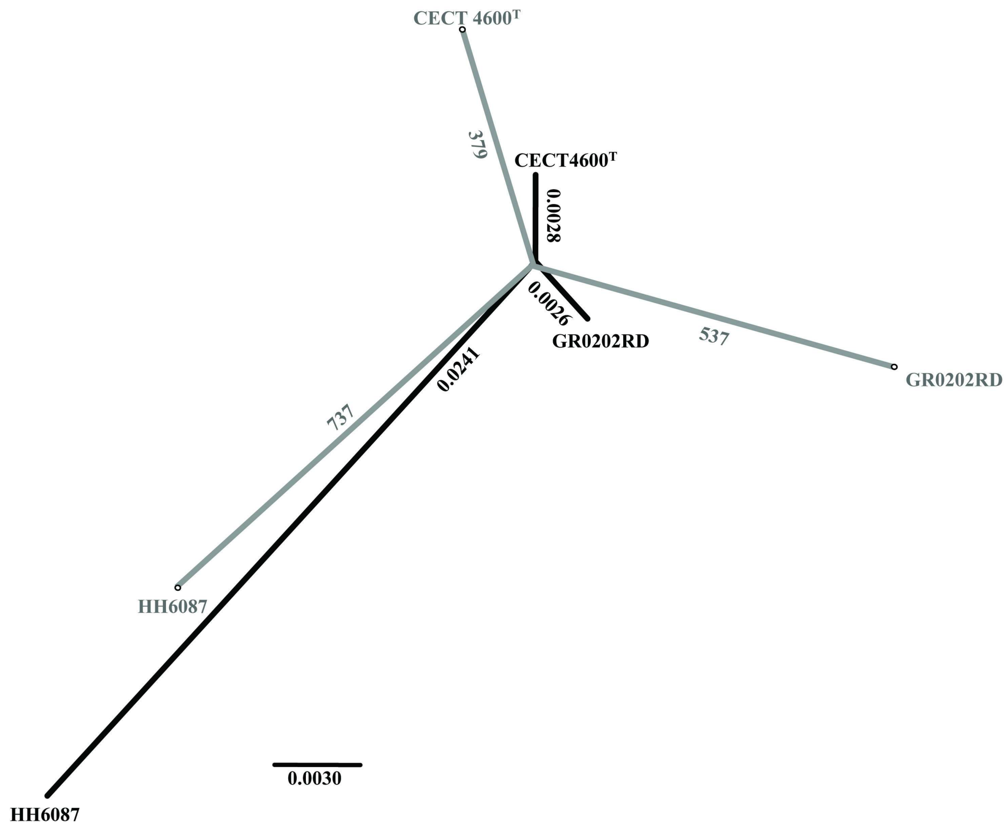
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Conserved and variable areas among strains in the gels: A) great similarity in the high molecular weight region; B) hipervariable protein pattern in the middle molecular weight region; and C) variable pattern in low molecular weight region showing more similarity between CECT4600T and GR0202RD strains.
190x143mm (300 x 300 DPI)

view

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Comparison of the phylogenetic reconstruction based on concatenated of 16S rRNA, atpA, pyrH, rpoA, rpoD and recA partial gene sequences obtained by the Maximum Likelihood method (GTR model) (Black) and the Maximum Parsimony tree generated on the basis of the protein patterns (Grey). Numbers on the branches show the phylogenetic distances.
250x207mm (300 x 300 DPI)