

RESEARCH ARTICLE

Open Access

Analysis of the genome content of *Lactococcus garvieae* by genomic interspecies microarray hybridization

Mónica Aguado-Urda¹, Guillermo H López-Campos², José F Fernández-Garayzábal^{1,3}, Fernando Martín-Sánchez², Alicia Gibello¹, Lucas Domínguez³, María M Blanco^{1*}

Abstract

Background: *Lactococcus garvieae* is a bacterial pathogen that affects different animal species in addition to humans. Despite the widespread distribution and emerging clinical significance of *L. garvieae* in both veterinary and human medicine, there is almost a complete lack of knowledge about the genetic content of this microorganism. In the present study, the genomic content of *L. garvieae* CECT 4531 was analysed using bioinformatics tools and microarray-based comparative genomic hybridization (CGH) experiments. *Lactococcus lactis* subsp. *lactis* IL1403 and *Streptococcus pneumoniae* TIGR4 were used as reference microorganisms.

Results: The combination and integration of *in silico* analyses and *in vitro* CGH experiments, performed in comparison with the reference microorganisms, allowed establishment of an inter-species hybridization framework with a detection threshold based on a sequence similarity of $\geq 70\%$. With this threshold value, 267 genes were identified as having an analogue in *L. garvieae*, most of which ($n = 258$) have been documented for the first time in this pathogen. Most of the genes are related to ribosomal, sugar metabolism or energy conversion systems. Some of the identified genes, such as *als* and *mycA*, could be involved in the pathogenesis of *L. garvieae* infections.

Conclusions: In this study, we identified 267 genes that were potentially present in *L. garvieae* CECT 4531. Some of the identified genes could be involved in the pathogenesis of *L. garvieae* infections. These results provide the first insight into the genome content of *L. garvieae*.

Background

Lactococcus garvieae is one of the most important bacterial pathogens that affect different farmed fish species in many countries, although its major impact is on the trout farm industry [1,2]. In addition to farmed fish, this microorganism has also been isolated from a wide range of wild fish species, from both fresh and marine water, as well as from giant fresh water prawns [3] and from wild marine mammals [4]. The host range of *L. garvieae* is not limited to aquatic species. This agent has also been identified in cows and water buffalos with subclinical mastitis [5,6] and from cat and dog tonsils [7]. In humans it has been isolated from the urinary tract,

blood, and skin and from patients with pneumonia, endocarditis or septicaemia [8-11]. Recently, intestinal disorders in humans have been associated with the consumption of raw fish contaminated with this pathogen [12], which suggests that *L. garvieae* could be considered as a potentially zoonotic bacterium [3,12]. Despite the widespread distribution and emerging clinical significance of *L. garvieae* in both veterinary and human medicine, there is almost a complete lack of knowledge about the genetic content of this microorganism.

In the last few years, research in microbial genetics has changed fundamentally, from an approach involving the characterization of individual genes to a global analysis of microbial genomes. The availability of complete genome sequences has enabled the development of high-throughput nucleic acid hybridization technologies including macro- and microarrays. Microarrays have the

* Correspondence: mmbianco@vet.ucm.es

¹Departamento de Sanidad Animal, Facultad de Veterinaria, Universidad Complutense, 28040 Madrid, Spain

capacity to monitor the genome content of bacterial strains or species very rapidly. Although whole-genome sequencing is definitely a powerful method for genetics, it is still expensive and time consuming. As an alternative, comparative genomic hybridization (CGH) experiments based on microarrays have been used to facilitate comparisons of unsequenced bacterial genomes. Array-based CGH using genome-wide DNA microarrays is used commonly to determine the genomic content of bacterial strains [13,14], but also for inter-species comparisons [14-16]. In this case, microarrays of closely related microorganisms that have been fully sequenced must be available. The primary advantage of this microarray approach is that it allows the identification of a large number of genes that are potentially present in an organism without the need for sequencing genomes. The disadvantage of this approach is that it indicates only the genes that are common between the fully sequenced relative and the strain of interest; genes unique to the strain of interest remain unknown [15,17]. In the present work the genetic content of *L. garvieae* CECT 4531 was studied by a combination of *in silico* analysis and *in vitro* microarray CGH experiments, using open reading frame (ORF) microarrays of two bacteria closely related to *L. garvieae*, namely *Lactococcus lactis* subsp. *lactis* IL1403 and *Streptococcus pneumoniae* TIGR4 [18,19].

Methods

Bacterial strains, culture conditions and isolation of genomic DNA

Lactococcus lactis subsp. *lactis* IL1403 (kindly provided by M.P. Gaya, INIA, Madrid, Spain) and *Streptococcus pneumoniae* TIGR4 (purchased from the American Type Culture Collection) were used as the reference sequenced microorganisms. The test strain of *Lactococcus garvieae* used for the experiments was CECT 4531 (purchased from the Spanish Type Culture Collection). The *L. lactis* subsp. *lactis* IL1403 and *L. garvieae* CECT 4531 were grown statically at 28°C in BHI broth (bio-Mérieux, Marcy l'Etoile, France). The *S. pneumoniae* TIGR4 was grown statically at 37°C in Todd Hewitt broth (Oxoid, Basingstoke, Hampshire, England). Cells were grown until the late-exponential phase of growth (OD₆₀₀~1.5-2) and harvested for isolation and purification of genomic DNA using the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) according to the manufacturer's specifications. The DNA concentrations were determined spectrophotometrically.

DNA labelling

Aliquots (1-2 µg) of genomic DNA from the three strains were labelled fluorescently with Cy3-dUTP or Cy5-dUTP (Perkin-Elmer, Foster City, CA, USA),

depending on whether the strain was used as a test or reference microorganism in the CGH experiments, respectively. Each DNA aliquot was fragmented by sonication to obtain fragments from 400 to 1000 bp. Fragmented DNA was mixed with 5 µL 10× NEBlot labelling buffer containing random sequence octamer oligonucleotides (New England Biolabs, Ipswich, MA, USA) and water to a final volume of 43.5 µL. This mixture was denatured by heating at 95°C for 5 min and then cooled for 5 min at 4°C. After this denaturing step, the remaining components of the labelling reaction were added: 5 µL of 10 × dNTP labelling mix (1.2 mM each dATP, dGTP and dCTP in 10 mM Tris pH 8.0, 1 mM EDTA) (New England Biolabs, Ipswich, MA, USA), 1.5 µL of 1 mM Cy3-dUTP or Cy5-dUTP and 1.5 µL of 10 U/µL Klenow fragment (Fermentas Life Sciences, Glen Burnie, MD, USA). The labelling reactions were incubated overnight at 37°C and then stopped by adding 2.5 µL of 0.5 M EDTA. Labelled DNA was purified from unincorporated label using a Qiaquick PCR Cleanup kit (Qiagen, Hilden, Germany) and dried under vacuum. The final DNA concentration and quality, as well as the labelling quality, were determined using a NanoDrop (NanoDrop Technologies, Wilmington, DE, USA).

Array-based comparative genome hybridization (CGH)

The *L. lactis* subsp. *lactis* IL1403 and *S. pneumoniae* TIGR4 microarrays used for the CGH analysis were purchased from Eurogentec (Serain, Belgium). The *L. lactis* microarray contains 4608 spots: 2126 duplicated ORFs, 32 negative controls and 324 empty spots. The *S. pneumoniae* microarray contains 4608 spots: 2087 duplicated ORFs, 224 negative controls and 210 empty spots.

The CGH experiments were performed by means of competitive hybridizations using DNA of *L. lactis* subsp. *lactis* IL1403 or *S. pneumoniae* TIGR4, depending on the array, as positive controls. The DNAs to be hybridized on the same array were labelled with Cy3-dUTP and Cy5-dUTP, respectively. For each microarray hybridization reaction, aliquots (1-2 µg) of labelled genomic DNAs of the reference (labelled with Cy3) and test (labelled with Cy5) strains, were mixed in 45 µL EGT hybridization solution (Eurogentec, Serain, Belgium) and denatured at 65°C for 2 min. The hybridization mixture was then loaded onto a microarray slide, covered with a coverslip and incubated at 38°C overnight. Following hybridization, the slides were washed in 2 × SSC, 0.5% SDS for 5 min followed by a second wash step in 1 × SSC, 0.25% SDS for 5 min. Finally, slides were rinsed in 0.2 × SSC and dried by centrifugation.

The results presented herein represent a compilation of sixteen separate CGH experiments: *L. lactis* subsp. *lactis* IL1403 arrays (reference microorganism) were hybridized with *S. pneumoniae* TIGR4 (test microorganism) (n = 2);

S. pneumoniae TIGR4 arrays (reference microorganism) were hybridized with *L. lactis* subsp. *lactis* IL1403 (test microorganism) (n = 2); *L. lactis* subsp. *lactis* IL1403 arrays (reference microorganism) were hybridized with *L. garvieae* CECT 4531 (test microorganism) (n = 8); *S. pneumoniae* TIGR4 arrays (reference microorganism) were hybridized with *L. garvieae* CECT 4531 (test microorganism) (n = 4). The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus [20] and are accessible through GEO Series accession number GSE19005. <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE19005>.

Data acquisition and analysis

The microarray was scanned after hybridization using a Scanarray HT microarray scanner (Perkin-Elmer). The signal intensity of the two fluoros was determined using ImaGene software (BioDiscovery, El Segundo, CA, USA). Microarray data were analysed using ImaGene software, Microsoft Excel and an in-house designed and built Microsoft Access database [21]. Gene calling was based on a signal-to-noise ratio (SNR) >3 for each spot. After the CGH experiments, a gene was considered to show a positive result when it was present in at least three of the four CGH assays. In the case of the *L. garvieae* CECT 4531 hybridizations with the *L. lactis* subsp. *lactis* IL1403 arrays, it was necessary to perform a larger number of assays (n = 8), owing to the poor quality of one of the batches of arrays used. Thus, the criterion chosen to determine a positive result in this case was when the gene was present in at least five of the eight CGH assays.

In silico sequence analysis

Sequence analyses were carried out to assess the performance of the inter-species CGH protocol. Using the BLAT [22] and BLAST [23] programs, the sequences of the *L. lactis* microarray probes were aligned with the *S. pneumoniae* genome sequence, and vice-versa. The BLAT search parameters were 90%, 80% and 70% sequence identity (BLAT90, BLAT80 and BLAT70) and a 100 bp minimum alignment length (owing to the fact that the length of the array probe was between 100 and 400 bp). Available *L. garvieae* sequences of the nine previously identified genes that were positive in the CGH were aligned with the *L. lactis* subsp. *lactis* IL1403 or *S. pneumoniae* TIGR4 genomes and with the sequences of the immobilized probes of these genes in the corresponding microarray using BLAST [23] and BLAST 2 sequences [24] programs.

Results

Inter-species comparison framework

In silico analyses were performed to compare the sequences of the immobilized probes in the microarray

of each reference organism with the sequences of their complete genomes available in GenBank (*L. lactis* subsp. *lactis* IL1403: NC_002662 and *S. pneumoniae* TIGR4: NC_003028). The BLAT alignment of the *L. lactis* IL1403 probes on the *S. pneumoniae* TIGR4 genome allowed the identification of 1 ORF with BLAT90, 65 ORFs with BLAT80 and 159 ORFs with BLAT70. Moreover, the BLAT alignment of the probes represented on the *S. pneumoniae* microarray on the *L. lactis* genome demonstrated 1 ORF, 63 ORFs and 165 ORFs for BLAT90, BLAT80 and BLAT70, respectively.

The CGH experiments based on swapping off the microarrays between *S. pneumoniae* and *L. lactis* identified 65 common ORFs. To evaluate the accuracy of the microarray CGH experiments, we compared these results with those of the *in silico* analysis. Out of the 65 genes, 47 (72%) showed similarities greater than 80%, 16 genes (25%) exhibited a similarity between 70% and 80%, and only 2 genes (3%) showed a similarity slightly lower than 70% (66–68%) (Table 1). In summary, 97% of the genes detected by CGH showed similarities greater than 70% at the nucleotide level.

After combined analysis of the results obtained *in silico* and *in vitro*, we established, under the hybridization conditions used in this study, a detection threshold based on a sequence similarity of $\geq 70\%$ for alignments longer than 100 bp. This was established as the reference framework for the inter-species CGH assays.

In vitro microarray CGH experiments with *L. garvieae* CECT 4531 vs reference microorganisms *L. lactis* subsp. *lactis* IL1403 and *S. pneumoniae* TIGR4, and in silico analysis of available sequences from *L. garvieae*

The microarray CGH experiments identified 267 genes in *L. garvieae* that had analogues in *L. lactis* and/or *S. pneumoniae* (Additional file 1). Of these, 111 genes (41.6%) were identified only with the *L. lactis* microarray, 70 genes (26.2%) only with the microarray of *S. pneumoniae*, and 86 genes (32.2%) were identified with both microarrays. These genes belong to diverse functional groups (Table 2). Most of the genes (96.6%) have been documented for the first time in *L. garvieae*. Only nine genes (four present in both reference microorganisms: *atpD*/SP1508, *pfk*/SP0896, *tig*/SP0400, *tuf*/SP1489; three present in *L. lactis*: *als*, *ddl*, *galK*; two present in *S. pneumoniae*: SP0766, SP1219) out of the 267 genes detected have been either identified or sequenced before in diverse strains of *L. garvieae* (Tables 3 and 4). *In silico* analysis of these previously sequenced genes (n = 9) of *L. garvieae* were performed to assess the efficacy of the methodology. Alignments of these available sequences with the genomes of the corresponding reference microorganism and their respective array probes showed nucleotide identities ranging

Table 1 *In silico* analysis of the common genes detected by CGH in the reference microorganisms

CGH positive gene ID ^a	BLAT80 ^b	BLAT70 ^{b, c}	% Identity in BLAST ^d
<i>atpD</i> /SP1508	++		66%
<i>dnaG</i> /SP1072	-	-	
<i>dnaJ</i> /SP0519	+		
<i>dnaK</i> /SP0517	++		
<i>enoA</i> /SP1128	++		
<i>fbaA</i> /SP0605	++		
<i>ftsZ</i> /SP1666	++		
<i>fusA</i> /SP0273	++		
<i>gapB</i> /SP2012	++		
<i>gltX</i> /SP2069	-	++	
<i>ldh</i> /SP1220	++		
<i>lepA</i> /SP1200	-	++	
<i>leuS</i> /SP0254	++		
<i>lysS</i> /SP0713	-	++	
<i>msmk</i> /SP1580	++		
<i>murA1</i> /SP1081	-	+	
<i>pepC</i> /SP0281	+		
<i>pfk</i> /SP0896	-	+	
<i>pgiA</i> /SP2070	+		
<i>pgk</i> /SP0499	++		
<i>pmg</i> /SP1655	++		
<i>ptnAB</i> /SP0284	+		
<i>ptnD</i> /SP0282	-	+	
<i>ptsI</i> /SP1176	-	++	
<i>purA</i> /SP0019	-	-	76%
<i>pyk</i> /SP0897	++		
<i>rplA</i> /SP0631	++		
<i>rplB</i> /SP0212	++		
<i>rplC</i> /SP0209	++		
<i>rplF</i> /SP0225	+		
<i>rplJ</i> /SP1355	+		
<i>rplK</i> /SP0630	++		
<i>rplL</i> /SP1354	++		
<i>rplM</i> /SP0294	++		
<i>rplN</i> /SP0219	++		
<i>rplO</i> /SP0229	++		
<i>rplQ</i> /SP0237	+		
<i>rplS</i> /SP1293	++		
<i>rplV</i> /SP0214	+		
<i>rpmI</i> /SP0960	+		
<i>rpmJ</i> /SP0233	++		
<i>rpoB</i> /SP1961	-	+	
<i>rpoC</i> /SP1960	+	++	
<i>rpoD</i> /SP1073	++		
<i>rpsA</i> /SP0862	-	+	
<i>rpsB</i> /SP2215	-	++	
<i>rpsC</i> /SP0215	++		
<i>rpsD</i> /SP0085	++		
<i>rpsE</i> /SP0227	++		
<i>rpsG</i> /SP0272	++		

Table 1: *In silico* analysis of the common genes detected by CGH in the reference microorganisms (Continued)

<i>rpsJ</i> /SP0208	++		
<i>rpsK</i> /SP0235	++		
<i>rpsL</i> /SP0271	++		
<i>rpsM</i> /SP0234	+		
<i>rpsP</i> /SP0775	+		
<i>rpsS</i> /SP0213	+		
<i>topA</i> /SP1263	-	+	
<i>tuf</i> /SP1489	++		
<i>typA</i> /SP0681	++		
<i>upp</i> /SP0745	-	-	77%
<i>yehH</i> /SP2097	+		
<i>yciA</i> /SP2096	-	-	68%
<i>yjiF</i> /SP1565	-	+	
<i>yjkl</i> /SP2209	-	+	
<i>yyaL</i> /SP0004	++		

^a Reference organisms gene ID: *Lactococcus lactis* subsp. *lactis* IL1403/*Streptococcus pneumoniae* TIGR4

^b ++ Genes detected in both alignments, *L. lactis* subsp. *lactis* IL1403 array probes vs *S. pneumoniae* TIGR4 genome, and *S. pneumoniae* TIGR4 array probes vs *L. lactis* subsp. *lactis* IL1403 genome; + positive in one of the two cases.

^c Only the results for the negative genes in BLAT80 are shown.

^d Only the results for the negative genes in both BLAT80 and BLAT70 are shown.

between 70% and 86% (Tables 3 and 4). Most of the available sequences (80%) showed similarities greater than 75%.

Discussion

In the present study, commercial microarrays of *L. lactis* subsp. *lactis* IL1403 and *S. pneumoniae* TIGR4 were used to determine the presence of homologous genes in *L. garvieae*. Both *L. lactis* and *S. pneumoniae* were chosen as reference organisms because they are closely related to *L. garvieae* [18,19] and their genomes have been fully sequenced. Although these CGH experiments cannot detect and identify genes that are likely to exist only in the target microorganism, this approach reveals genes that are common to both the reference and the target organisms, allowing the identification of a large number of genes potentially present in an organism without the need for sequencing genomes [17,25].

In experiments that involve inter-species comparison it is necessary to establish a framework that allows accurate comparison and interpretation of the results. Thus, the first efforts were focused on establishing that framework by the combination and integration of *in silico* analyses and *in vitro* microarray CGH experiments to compare the reference organisms *L. lactis* subsp. *lactis* IL1403 and *S. pneumoniae* TIGR4. Signal intensity has been used to assess the level of similarity between two genes in inter-species CGH experiments [15]. However,

Table 2 Functional groups of genes identified in *L. garvieae* CECT 4531 according to the COG database

Functional Group	Homologous in <i>L. lactis</i> subsp. <i>lactis</i> IL1403	Homologous in <i>S. pneumoniae</i> TIGR4
Amino acid transport and metabolism	14	10
Carbohydrate transport and metabolism	24	15
Cell cycle control, cell division, chromosome partitioning	4	2
Cell wall/membrane/envelope biogenesis	5	4
Coenzyme transport and metabolism	1	1
DNA replication, recombination and repair	8	12
Energy production and conversion	11	6
Inorganic ion transport and metabolism	4	5
Intracellular trafficking, secretion and vesicular transport	4	2
Lipid transport and metabolism	2	0
Nucleotide transport and metabolism	15	11
Phage capsid proteins	1	0
Post translational modification, protein turnover, chaperones	8	8
Signal transduction mechanisms	2	3
Transcription	7	6
Translation, ribosomal structure and biogenesis	64	60
Unknown function	23	11
Total	197	156

Table 3 *In silico* analysis of the available sequences of the genes detected in *L. garvieae* by CGH

Gene ID	GenBank accession number of <i>L. garvieae</i> sequence	<i>L. garvieae</i> strain	Similarity with <i>L. lactis</i> subsp. <i>lactis</i> IL1403 gene (%)	Similarity with array probe (%)
<i>als</i>	EF450031	UNIUD074	77	76
<i>atpD</i>	AX111128	from patent W00123604	86	86
<i>ddl</i>	AF170808	<i>E. serolicida</i>	72	75
<i>galk</i>	EU153555	DSM 20684	78	79
<i>pfk</i>	AB024532	SA8201	85	84
<i>tig</i>	AB024531	SA8201	82	-
<i>tuf</i>	AX109994	from patent W00123604	80	77

Results for the *L. lactis* subsp. *lactis* IL1403 array based-CGH

this approach may be influenced, and therefore biased, by different factors, such as regional sample labelling effects, probe accessibility or local hybridization issues [13]. For these reasons, in the present study signal intensity was not considered for determining whether a gene was positive or not in the inter-species CGH experiments.

These analyses revealed that nearly all the genes common to *L. lactis* and *S. pneumoniae* that were detected by swap microarray CGH experiments (97%) exhibited a sequence similarity of at least 70% (Table 1). Only two genes (*dnaG* and *yciA*) detected in the microarray CGH experiments showed a sequence similarity slightly lower than 70% (66 and 68%, respectively; Table 1). Variability in the factors that influence the CGH signals, such as systematic errors (e.g. dye effects), copy number variation, and sequence divergence between the analysed

samples [13], may explain these results. The comparison of the results of both analyses, *in silico* and *in vitro*, for the reference microorganisms (Table 1) allowed us to establish that, under our experimental conditions, it was possible to detect and identify inter-species hybridization with a detection threshold based on a sequence similarity of $\geq 70\%$.

Therefore, our threshold value of sequence similarity $\geq 70\%$ was set up directly from the comparison of the results of the *in silico* and *in vitro* analyses of the present study. This threshold value was used subsequently to interpret the results of the microarray-based CGH experiments comparing *L. garvieae* and the reference microorganisms. Less stringent hybridization conditions would probably have allowed the identification of a larger number of genes, but this would have also resulted in lower specificity. Given that the final aim of the

Table 4 *In silico* analysis of the available sequences of the genes detected in *L. garvieae* by CGH

Gene ID	GenBank accession number of <i>L. garvieae</i> sequence	<i>L. garvieae</i> strain	Similarity with <i>S. pneumoniae</i> TIGR4 gene (%)	Similarity with array probe (%)
SP1508	AX111128	from patent WO0123604	82	82
SP0896	AB024532	SA8201	80	79
SP0766	AM490328	JIP 31-90 (2)	71	79
	AJ387925	CIP 102507 T	70	70
	AJ387923	<i>E. serolicida</i> ATCC49156	70	70
SP04000	AB024531	SA8201	74	-
SP1489	AX109994	from patent WO0123604	80	79
SP1219	AB364641	20-92	84	86
	AB364640	Lc.1236	84	85
	AB364639	Lc. 925	85	85
	AB364638	Lc. 881	84	84
	AB364637	Lc. 337	85	85
	AB364633	LMG9472	85	85
	AB364632	ATCC43921	84	84
	AB364627	G50202	84	84
	AB364626	KGLA5224	84	84
	AB364625	EH5803	83	83
AB364624	KG9408	84	84	

Results for the *S. pneumoniae* TIGR4 array based-CGH

experiment was the identification of genes potentially present in *L. garvieae*, it was preferred to maintain stringent hybridization conditions, therefore increasing the specificity and the reliability of the results. Hence, the genes detected in the CGH experiments should have an analogue in *L. garvieae* with a nucleotide sequence identity greater than 70% with the respective gene in the reference organism.

The CGH hybridizations using *L. lactis* subsp. *lactis* IL1403 and *S. pneumoniae* TIGR4 microarrays identified 267 analogous genes in *L. garvieae* (Additional file 1). Only 3.4% of these genes (nine out of 267) have been characterized or sequenced previously by other groups in different strains of *L. garvieae* [[18,26-29], and GenBank sequences: AX109994, AB364624, AB364625, AB364626, AB364627, AB364632, AB364633, AB364637, AB364638, AB364639, AB364640, AB364641, EU153555]. The alignments of the available sequences of these nine previously of these nine previously identified genes in *L. garvieae* with both the sequences of these genes from the reference microorganisms and those from the array probe showed nucleotide similarities greater than 70% (70-86%) between them (Tables 3 and 4). These data are consistent with the detection threshold value discussed previously. Therefore it is reasonable to assume that the other genes detected in *L. garvieae* CECT 4531 by CGH experiments will also have at least 70% sequence similarity with the respective genes in the reference microorganisms. The

positive result obtained in both CGH experiments for the *tig/SP0400* gene (Tables 3 and 4), was unexpected given the absence of similarity between the available sequence and the probes on both microarrays. This result could be explained by the fact that the available sequence for *L. garvieae* is partial, and it represents a part of the gene that does not correspond with the probe.

We classified the ORFs into clusters of orthologous genes (COGs) [30]. The 267 genes identified in *L. garvieae* CECT 4531 (Additional file 1) belong to diverse biological functional groups (Table 2). Most of the genes detected in *L. garvieae* (about 66%) were related to meaningful biological functions such as those related to ribosomal functions, sugar metabolism or energy conversion systems, which are usually represented in *Lactobacillales* [31]. The remaining genes identified included "housekeeping genes", such as *gyrB*, *sodA*, *recA*, *ileS*, *rpoD*, *dnaK* and *ddl* [19], genes of diverse functional groups and genes with unknown functions. Some of them are of interest because they could be involved in the pathogenesis of *L. garvieae* infections. For example, the gene *als*, which has been described as an important factor for host colonization by El Tor biotypes of *Vibrio cholerae* [32], has also been suggested to be one of the genes required for survival of *L. garvieae* in fish [27]. In addition, the gene *mycA*, which was detected for the first time in *L. garvieae* in the present study, encodes an antigen that cross-reacts with myosin, and members of

this family of proteins have been suggested to play an important role in the pathogenesis of streptococcal infections [33].

Sequencing of the genes identified in this work is beyond the scope of this initial study, but the data provided can be the starting point for future genetic analysis of *L. garvieae* strains from different ecological niches or adapted to different host species.

This study provides the first insight into the genome content of *L. garvieae* and suggests that CHG could be a useful approach for studying the genetic content of other Gram-positive catalase-negative cocci of human and veterinary relevance.

Conclusions

In the present work, a comparative analysis based on microarray interspecies hybridization and on the use of bioinformatic tools was used for the first time to study the genetic content of *L. garvieae* CECT 4531. It is important to remark that the integration of results from bioinformatics and microarray-based CGH requires the definition of a framework that allows an accurate comparison and interpretation of the results obtained. Once this framework was established, it was possible to identify 267 genes potentially present in *L. garvieae* CECT 4531. Some of the identified genes, such as the *als* and *mycA* genes, could be involved in the pathogenesis of *L. garvieae* infections.

In summary, these results provide the first insight into the genome content of *L. garvieae* and could be useful for future understanding of the genetics of this pathogenic microorganism.

Additional file 1: Genes potentially identified in *L. garvieae* CECT 4531 and their homologues in *L. lactis* subsp. *lactis* IL1403 and *S. pneumoniae* TIGR4.

Acknowledgements

This work was supported partially by projects AGL2005-04775 and AGL2009-12447 of the Ministerio Español de Ciencia e Innovación. M. Aguado-Urda was a recipient of a grant from Centro de Vigilancia Sanitaria Veterinaria (VISAVET), and a PhD grant from the Universidad Complutense de Madrid. The work of Dr. López-Campos and Dr. Martín-Sánchez was partially funded by the COMBIOMED Network and ONTOMINEBASE research project (Ministerio Español de Ciencia e Innovación). The authors thank M.P. Gaya for providing the *Lactococcus lactis* subsp. *lactis* IL1403 strain.

Author details

¹Departamento de Sanidad Animal, Facultad de Veterinaria, Universidad Complutense, 28040 Madrid, Spain. ²Área de Bioinformática y Salud Pública, Instituto de Salud Carlos III, 28220 Madrid, Spain. ³Centro de Vigilancia Sanitaria Veterinaria (VISAVET), Facultad de Veterinaria, Universidad Complutense, 28040 Madrid, Spain.

Authors' contributions

MAU carried out the microarray experiments and the bioinformatics analyses, and participated in the analysis of the data and drafting of the manuscript. GHLC designed the microarray experiments, and participated in

analysis of the data and drafting of the manuscript. JFFG participated in and supervised drafting of the manuscript. FMS participated in and supervised the design of the microarray experiments and the analysis of the microarray data. AG participated in the design of the study and drafting of the manuscript. LD participated in the design of the study. MB conceived the study, and participated in its design and coordination. All the authors read and approved the final manuscript.

Received: 3 December 2009 Accepted: 16 March 2010

Published: 16 March 2010

References

1. Chen SC, Liaw LL, Su HY, Ko SC, Wu CY, Chung HC, Tsai YH, Yang KL, Chen YC, Chen TH, Lin GR, Cheng SY, Lin YD, Lee JL, Lai CC, Weng YJ, Chu SY: *Lactococcus garvieae*, a cause of disease in grey mullet, *Mugil cephalus*, in Taiwan. *J Fish Dis* 2002, **25**:727-732.
2. Vela AI, Vázquez J, Gibello A, Blanco MM, Moreno MA, Liébana P, Albendea C, Alcalá B, Méndez A, Domínguez L, Fernández-Garayzábal JF: Phenotypic and genetic characterization of *Lactococcus garvieae* isolated in Spain from lactococcosis outbreaks and comparison with isolates of other countries and sources. *J Clin Microbiol* 2000, **38**:3791-3795.
3. Vendrell D, Balcázar JL, Ruiz-Zarzuola I, de Blas I, Gironés O, Múzquiz JL: *Lactococcus garvieae* in fish: a review. *Comp Immunol Microbiol Infect Dis* 2006, **29**:177-198.
4. Evans JJ, Pasnik DJ, Klesius PH, Al-Ablani S: First report of *Streptococcus agalactiae* and *Lactococcus garvieae* from a wild bottlenose dolphin (*Tursiops truncatus*). *J Wild Dis* 2006, **42**:561-569.
5. Collins MD, Farrow JAE, Phillips BA, Kandler O: *Streptococcus garvieae* sp. nov. and *Streptococcus plantarum* sp. *J Gen Microbiol* 1983, **129**:3427-3431.
6. Teixeira LM, Merquior VLC, Vianni MCV, Carvalho MGS, Fracalanza SEL, Steigerwalt AG, Brenner DJ, Facklam RR: Phenotypic and genotypic characterization of atypical *Lactococcus garvieae* strains isolated from water buffalos with subclinical mastitis and confirmation of *L. garvieae* as a senior subjective synonym of *Enterococcus seriolida*. *Int J Syst Bacteriol* 1996, **46**:664-668.
7. Pot B, Devriese LA, Ursi D, Vandamme P, Haesebrouck F, Kersters K: Phenotypic identification and differentiation of *Lactococcus* strains isolated from animals. *Syst Appl Microbiol* 1996, **19**:213-222.
8. Elliot JA, Collins MD, Pigott NE, Facklam RR: Differentiation of *Lactococcus lactis* and *Lactococcus garvieae* from humans by comparison of whole-cell protein patterns. *J Clin Microbiol* 1991, **20**:2731-2734.
9. Facklam RR, Elliot JA: Identification, classification, and clinical relevance of catalase-negative, Gram-positive cocci, excluding the streptococci and enterococci. *Clin Microbiol Rev* 1995, **8**:479-495.
10. Fefer JJ, Ratzan KR, Sharp SE, Saiz E: *Lactococcus garvieae* endocarditis: report of a case and review of the literature. *Diagn Microbiol Infect Dis* 1998, **32**:127-130.
11. Li WK, Chen YS, Wann SW, Liu YC, Tsai HT: *Lactococcus garvieae* endocarditis with initial presentation of acute cerebral infarction in a healthy immunocompetent man. *Inter Med* 2008, **47**:1143-1146.
12. Wang CY, Shie HS, Chen SC, Huang JP, Hsieh IC, Wen MS, Lin FC, Wu D: *Lactococcus garvieae* infections in humans: possible association with aquaculture outbreaks. *Int J Clin Pract* 2007, **61**:68-73.
13. van Hijum SAFT, Baerends RJS, Zomer AL, Karsens HA, Martín-Requena V, Trelles O, Kok J, Kuipers OP: Supervised Lowess normalization of comparative genome hybridization data-application to lactococcal strain comparisons. *BMC Bioinf* 2008, **9**:93.
14. Hakenbeck R, Balmelle N, Weber B, Gardes C, Keck W, de Saizieu A: Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infect Immun* 2001, **69**:2477-2486.
15. Dong Y, Glasner JD, Blattner FR, Triplett EW: Genomic interspecies microarray hybridization: rapid discovery of three thousand genes in the maize endophyte, *Klebsiella pneumoniae* 342, by microarray hybridization with *Escherichia coli* K-12 Open Reading Frames. *Appl Environ Microbiol* 2001, **67**:1911-1921.
16. Fukuya S, Mizoguchi H, Tobe T, Mori H: Extensive genomic diversity in pathogenic *Escherichia coli* and *Shigella* strains revealed by comparative genomic hybridization microarray. *J Bacteriol* 2004, **186**:3911-3921.
17. Zhang L, Reddi U, Srinivasan U, Li S, Borchardt SM, Pillai P, Mehta P, Styka AN, DeBusscher J, Marrs CF, Foxman B: Combining microarray technology and molecular epidemiology to identify genes associated

- with invasive group B *Streptococcus*. *Interdiscip Perspect Infect Dis* 2008, **2008**:314762.
18. Michel C, Pelletier C, Bousaha M, Douet DG, Lautraite A, Tailliez P: **Diversity of lactic acid bacteria associated with fish and the fish farm environment, established by amplified rRNA gene restriction analysis.** *Appl Environ Microbiol* 2007, **73**:2947-2955.
 19. Coenye T, Vandamme P: **Extracting phylogenetic information from whole-genome sequencing projects: the lactic acid bacteria as a test case.** *Microbiol* 2003, **149**:3507-3517.
 20. Edgar R, Domrachev M, Lash AE: **Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.** *Nucleic Acids Res* 2002, **30**:207-210.
 21. López-Campos GH, García-Albert L, Martín-Sánchez F, García-Sáez A: **Analysis and management of HIV peptide microarray experiments.** *Methods Inf Med* 2006, **45**:158-162.
 22. Kent WJ: **BLAT-the BLAST-like alignment tool.** *Genome Res* 2002, **12**:656-664.
 23. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403-410.
 24. Tatusova TA, Maden TL: **BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences.** *FEMS Microbiol Lett* 1999, **2**:247-250.
 25. Zhang R, Zhang CT: **The impact of comparative genomics on infectious disease research.** *Microbes and Infect* 2006, **8**:1613-1622.
 26. Hirono I, Yamashita H, Park CI, Yoshida T, Aoki T: **Identification of genes in a KG - phenotype of *Lactococcus garvieae*, a fish pathogenic bacterium, whose proteins react with antiKG-rabbit serum.** *Microb Pathog* 1999, **27**:407-417.
 27. Menéndez A, Fernández L, Reimundo P, Guijarro JA: **Genes required for *Lactococcus garvieae* survival in a fish host.** *Microbiology* 2007, **153**:3286-3294.
 28. Ozawa Y, Courvalin P, Gaiimand M: **Identification of enterococci at the species level by sequencing of the genes for D-alanine: D-alanine ligases.** *Syst Appl Microbiol* 2000, **23**:230-237.
 29. Poyart C, Quesnes G, Trieu-Cuot P: **Sequencing the gene encoding manganese dependent superoxide dismutase for rapid species identification of enterococci.** *J Clin Microbiol* 2000, **38**:415-418.
 30. Tatusov RL, Galperin MY, Natale DA, Koonin EV: **The COG database: a tool for genome-scale analysis of protein functions and evolution.** *Nucleic Acids Res* 2000, **1**:33-36.
 31. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welkeri D, Huges J, Goh Y, Benson A, Baldwin K, Lee JH, Díaz-Muñiz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D: **Comparative genomics of the lactic acid bacteria.** *PNAS* 2006, **103**:15611-15616.
 32. Yoon SS, Mekalanos JJ: **2,3-butanediol synthesis and the emergence of the *Vibrio cholerae* El Tor biotype.** *Infect Immun* 2006, **74**:6547-6556.
 33. Kil KS, Cunningham MW, Barnett LA: **Cloning and sequence analysis of a gene encoding a 67-kilodalton myosin-cross-reactive antigen of *Streptococcus pyogenes* reveals its similarity with class II major histocompatibility antigens.** *Infect Immun* 1994, **62**:2440-2449.

doi:10.1186/1471-2180-10-79

Cite this article as: Aguado-Urda et al.: Analysis of the genome content of *Lactococcus garvieae* by genomic interspecies microarray hybridization. *BMC Microbiology* 2010 **10**:79.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

