Journal of Biotechnology 212 (2015) 30–31

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Genome Announcement

Complete genome sequence of *Piscirickettsia salmonis* LF-89 (ATCC VR-1361) a major pathogen of farmed salmonid fish☆

Rodrigo Pulgar[a,c,1], Dante Travisany[b,c,1], Alejandro Zuñiga[a,c], Alejandro Maass[b,c,d], Verónica Cambiazo[a,c,*]

[a] *Laboratorio de Bioinformática y Expresión Génica, INTA-Universidad de Chile, El Líbano 5524, Santiago, Chile*
[b] *Laboratorio de Bioinformática y Matemática del Genoma, Centro de Modelamiento Matemático, FCFM-Universidad de Chile, Beauchef 851, Santiago, Chile*
[c] *Fondap Center for Genome Regulation (CGR), Avenida Blanco Encalada 2085, Santiago, Chile*
[d] *Departamento de Ingeniería Matemática, FCFM-Universidad de Chile, Beauchef 851, Santiago, Chile.*

ARTICLE INFO

Article history:
Received 19 July 2015
Accepted 21 July 2015
Available online 26 July 2015

Keywords:
Piscirickettsia salmonis
Salmonid fish pathogen
Salmonid rickettsial septicemia
Salmon farming

ABSTRACT

Piscirickettsia salmonis, the causative agent of salmonid rickettsial septicemia (SRS), is a significant threat to the healthy and sustainable production of salmonid farming industry. This Gram-negative bacterium, originally isolated from a coho salmon in Southern Chile, produces a systemic infection characterized by colonization of several fish organs. *P. salmonis* is able to infect, survive, and replicate inside salmonid macrophages however little is known about its mechanisms of pathogenesis. Here, we present the whole genome sequence and annotation of the *P. salmonis* reference strain LF-89 (ATCC VR-1361). The genome contains one circular chromosome of 3,184,851 bp and three plasmids, pPSLF89-1 (180,124 bp), pPSLF89-2 (33,516 bp) and pPSLF89-3 (51,573 bp). A total of 2850 protein-coding genes, 56 tRNAs and six copies of 5S–16S–23S rRNA.

Piscirickettsia salmonis is the etiological agent of salmonid rickettsial septicaemia (SRS), a severe disease that causes major economic losses to the Chilean salmon industry (Fryer et al., 1990). This Gram-negative, intracellular bacterium was first isolated in coho salmon (*Oncorhynchus kisutch*) but now it is the cause of serious losses in all farmed salmonid fish species (Rozas and Enriquez, 2014). *P. salmonis* covers a wide geographic range and outbreaks of SRS have been reported among farmed salmonid in Canada, Norway and Ireland, however, mortalities have not been as high as those recorded in Chile. *P. salmonis* has the ability to infect, survive, replicate, and propagate in salmonid monocytes/macrophages (McCarthy et al., 2008). In fish, *P. salmonis* produces a systemic infection characterized by the colonization of several organs including kidney, liver, spleen, intestine, brain, ovary and gills (Rozas and Enriquez, 2014).

Despite the fact that *P. salmonis* has been recognized as a major fish pathogen for over 20 years, the mechanisms of *P. salmonis* pathogenesis, and immune evasion are poorly understood. Although, several prophylactic approaches have been tested, the high frequency of new epizootic events indicates that alternative strategies are required to prevent SRS. The limited effectiveness of vaccine treatments has favored therapies based primarily on the use of high doses of antibiotics, which have generated antibiotic resistance and strong criticisms to the industry (Henriquez et al., 2015). Hence, the knowledge of genome sequence and structure may be essential to develop effective *P. salmonis* vaccines. The availability of *P. salmonis* genome will provide a catalog of all the potential antigens from which to select vaccines candidates that are likely to be more effective against this pathogen. This approach known as “reverse vaccinology” has been successfully applied to identify vaccine candidates against human pathogens (Serruto et al., 2004). Moreover, the availability of *P. salmonis* genome sequence will increase our understanding of its physiology at the system level, and genome-scale metabolic and regulatory models are expected to provide a valuable framework to guide the discovery of novel drug targets (Kim et al., 2012).

Nowadays, there are eight *P. salmonis* sequencing projects available at NCBI, four of them correspond to the LF-89 strain and four to environmental isolates; however they are all draft genomes with contig numbers between 227 and 534. Here we present the first

☆ *Nucleotide sequence accession number*: The complete chromosome sequence and three plasmid sequences have been deposited in GenBank under the accession number from CP011849–CP011852 (BioProject PRJNA199690). The strain is available from the American Type Culture Collection under the accession number ATCC VR-1361.

* Corresponding author at: Laboratorio de Bioinformática y Expresión Génica, INTA-Universidad de Chile, El Líbano 5524, Santiago, Chile.
E-mail address: vcambiaz@inta.uchile.cl (V. Cambiazo).

[1] These authors have contributed equally to this work.

http://dx.doi.org/10.1016/j.jbiotec.2015.07.017
0168-1656/

Table 1
Genome features of *Piscirickettsia salmonis* LF-89 (ATCC VR-1361).

Features	Chromosome	Plasmid		
		pPSLF89-1	pPSLF89-2	pPSLF89-3
Length (bp)	3,184,851	180,124	33,516	51,573
G + C content (%)	39.73	38.9	40.61	39.06
CDS	2634	138	35	43
rRNA genes	18	0	0	0
tRNA genes	56	0	0	0

complete genome sequence of this bacterium, which comprises one circular chromosome and three previously unreported plasmids. Thus, our work provides a genetic background to understand *P. salmonis* biology and pathogenesis.

Genomic DNA from *P. salmonis* was purified from exponential growth cultures (1 mL, OD_{600}: 0.5) using the DNeasy Blood & Tissue Kit for DNA (Qiagen). Genome sequencing was performed using one lane of an Illumina GAIIx and two single-molecule-real-time (SMRT) cells of the PacBio RSII. For Illumina, a shotgun library with a total of 70,531,338 paired end reads of 101 bp was generated (2,064x raw coverage). Whereas the PacBio RSII instrument produced a total of 158,140 reads with mean read length of 6432 bp, N50 size of 9350 bp and a total of 1,017,189,225 bp (294x raw coverage).

PacBio reads were processed using HGAP3 and the assembled contigs were polished using Quiver to generate 16 high quality contigs completing 3,661,513 bp. The longest contig size was 1,291,149 bp with an N50 size of 874,056 bp and a mean contig length of 228,844 bp. The PacBio reads were also assembled using PBcR of WGS-Assembler 8.2, resulting in 12 contigs with a total of 3,361,721 bp. The longest contig size was 934,543 bp with an N50 of 612,582 bp and mean contig length of 242,396 bp. Both assemblies were manually aligned and curated using MUMmer 3.23 (Kurtz et al., 2004). We use NUCmer (Delcher et al., 2002) to find overlaps between the HGAP3 and the WGS-Assembler assemblies. The resulting overlaps were manually aligned and then checked using blastx and blastn to avoid chimeric regions and assembly errors. The final assembly consisted of a circular chromosome of 3,184,851 bp and three circular plasmids, pPSLF89-1 (180,124 bp), pPSLF89-2 (33,516 bp) and pPSLF89-3 (51,573 bp). Lastly, an iterative in-house procedure of error correction was performed mapping the Illumina reads over *P. salmonis* genome using the BWA-aligner 0.7.5a-r405. Ori-Finder software was used to center the chromosome at the origin of replication. Genome average GC content was 39.68 mol%.

Putative coding sequences were identified using Glimmer 3.02, GeneMarkS, RNAmmer 1.2 and tRNAscan-SE 1.21. Functional annotation of CDSs was performed using Metanor of GenDB. Protein domains were predicted using InterPRO-scan against the Interpro database collection. Final annotation was submitted to "NCBI Microbial Genome Submission Check" and approved on June 17th, 2015. We predicted 74 RNA genes (18 rRNA and 56 tRNA) and a total of 2850 protein-coding genes, 2634 were contained in the chromosome, 138 in plasmid pPSLF89-1, 35 in pPSLF89-2 and 43 in pPSLF89-3 (see Table 1).

We performed a comparative analysis between the predicted proteomes of *P. salmonis* LF-89 and four other strains with publicly available draft genome sequences and detected 1626 proteins encoded by genes present in the four strains. *In silico* analysis using the β-barrel predictor BOMP and the TMHMM program identified 30 outer membrane proteins (OMP) from the set of common proteins; these OMPs may represent good candidates for vaccine development. It is known that intracellular pathogens have evolved complex mechanisms of iron acquisition to compete successfully with the host for the metal. In a previous work we predicted a cluster of genes encoding all enzymes necessary for siderophore biosynthesis and transport (Pulgar et al., 2015). Siderophores have the potential to be modified in order to improve the delivery of antibiotics by using them as Trojan horses, a biotechnological approach that have been employed to counter bacterial antibiotic resistance (Gorska et al., 2014). In addition to the iron acquisition genes, the genome of *P. salmonis* contains other 143 putative virulence genes that were identified by blastp search against VFDB (Chen et al., 2012). Among them, we found genes encoding the core components ($n = 16$) of the Dot/Icm type IV secretion system, of which four components (*dotB*, *dotA*, *icmK* and *icmE*) were identified previously (Gomez et al., 2013). Type IV secretion system has been well described in intracellular pathogens, such as *Legionella* spp. and *Coxiella* spp., in *P. salmonis*, the 16 *dot*/*icm* genes were organized into three clusters that were interspersed in the genome.

In conclusion, the complete genome sequence of *P. salmonis* will provide new insights into virulence factors and pathogenicity of this bacterium. A detailed analysis of this genome and comparative studies between genomes of strains with distinct virulence phenotypes will help in the development of new and efficient vaccines and antimicrobial agents.

Acknowledgements

This work was supported by Fondap Grant 15090007 Center for Genome Regulation, Fondecyt 1120254 to VC, Fondecyt 3130742 to RP, Basal Grant of the Center for Mathematical Modeling UMI2807 UCHILE-CNRS and Grant CIRIC-INRIA Chile. DT is a recipient of Doctoral fellowship from CONICYT (21150895). We acknowledge the National Laboratory for High Performance Computing at the Center for Mathematical Modeling (PIA ECM-02-CONICYT) and the sequencing center OMICS-Solutions.

References

Chen, L., Xiong, Z., Sun, L., Yang, J., Jin, Q., 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res. 40, D641–D645.

Delcher, A.L., Phillippy, A., Carlton, J., Salzberg, S.L., 2002. Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res. 30, 2478–2483.

Fryer, J., Lannan, C., Garcés, L., Larenas, J., Smith, P., 1990. Isolation of a rickettsiales-like organism from diseased coho salmon (*Oncorhynchus kisutch*) in Chile. Fish Pathol. 25, 107–114.

Gomez, F.A., Tobar, J.A., Henriquez, V., Sola, M., Altamirano, C., Marshall, S.H., 2013. Evidence of the presence of a functional Dot/Icm type IV-B secretion system in the fish bacterial pathogen *Piscirickettsia salmonis*. PLoS ONE 8, e54934.

Gorska, A., Sloderbach, A., Marszall, M.P., 2014. Siderophore–drug complexes: potential medicinal applications of the 'Trojan horse' strategy. Trends Pharmacol. Sci. 35, 442–449.

Henriquez, P., Bohle, H., Bustamante, F., Bustos, P., Mancilla, M., 2015. Polymorphism in gyrA is associated to quinolones resistance in Chilean *Piscirickettsia salmonis* field isolates. J. Fish Dis. 38, 415–418.

Kim, T.Y., Sohn, S.B., Kim, Y.B., Kim, W.J., Lee, S.Y., 2012. Recent advances in reconstruction and applications of genome-scale metabolic models. Curr. Opin. Biotechnol. 23, 617–623.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., Salzberg, S.L., 2004. Versatile and open software for comparing large genomes. Genome Biol. 5, R12.

McCarthy, U.M., Bron, J.E., Brown, L., Pourahmad, F., Bricknell, I.R., Thompson, K.D., Adams, A., Ellis, A.E., 2008. Survival and replication of *Piscirickettsia salmonis* in rainbow trout head kidney macrophages. Fish Shellfish Immunol. 25, 477–484.

Pulgar, R., Travisany, D., Zúñiga, A., Domínguez, C., Maass, A., González, M., Cambiazo, V., 2015. Transcriptional response of Atlantic salmon families to *Piscirickettsia salmonis* infection highlights the relevance of the iron-deprivation defence system. BMC Genomics 16, 495.

Rozas, M., Enriquez, R., 2014. Piscirickettsiosis and *Piscirickettsia salmonis* in fish: a review. J. Fish Dis. 37, 163–188.

Serruto, D., Adu-Bobie, J., Capecchi, B., Rappuoli, R., Pizza, M., Masignani, V., 2004. Biotechnology and vaccines: application of functional genomics to *Neisseria meningitidis* and other bacterial pathogens. J. Biotechnol. 113, 15–32.