Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

Short communication

Draft genomes and reference transcriptomes extend the coding potential of the fish pathogen *Piscirickettsia salmonis*

Angela D. Millar [a,b], Paz Tapia [b], Fernando A. Gómez [c], Sergio H. Marshall [c], Derie E. Fuentes [a,*], Jorge H. Valdes [b,*,1]

[a] *Aquaculture and Marine Ecosystems Division, Fraunhofer Chile Research Foundation, Santiago, Chile*
[b] *Biocomputing and Applied Genetics Division, Fraunhofer Chile Research, Santiago, Chile*
[c] *Laboratorio de Genética e Inmunología Molecular, Instituto de Biología, Facultad de Ciencias, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile*

ARTICLE INFO

ABSTRACT

Background: Draft and complete genome sequences from bacteria are key tools to understand genetic determinants involved in pathogenesis in several disease models. *Piscirickettsia salmonis* is a Gram-negative bacterium responsible for the Salmon Rickettsial Syndrome (SRS), a bacterial disease that threatens the sustainability of the Chilean salmon industry. In previous reports, complete and draft genome sequences have been generated and annotated. However, the lack of transcriptome data underestimates the genetic potential, does not provide information about transcriptional units and contributes to disseminate annotation errors. *Results:* Here we present the draft genome and transcriptome sequences of four *P. salmonis* strains. We have identified the transcriptional architecture of previously characterized virulence factors and trait-specific genes associated to cation uptake, metal efflux, antibiotic resistance, secretion systems and other virulence factors. *Conclusions:* This data has provided a refined genome annotation and also new insights on the transcriptional structures and coding potential of this fish pathogen.

1. Introduction

Piscirickettsia salmonis, a Gram-negative intracellular facultative bacterium, was first isolated from Coho salmon (*Oncorhynchus kisutch*) and currently produces systemic infection in salmonids and non-salmonid species [1]. Several drafts and complete genome sequences have been made publicly available in the last years; however, no transcriptome data associated have been generated for each genome sequence released to date.

Using next generation sequencing, we have generated draft genomes and reference transcriptomes for four *P. salmonis* strains (LF-89; EM-90; S-GIM and T-GIM). Bioinformatics and comparative genome analyses have provided an improved view of the coding potential of this species, which has been validated by reference transcriptome mapping.

* Corresponding authors.
E-mail addresses: derie.fuentes@fraunhofer.cl, (D.E. Fuentes), jorge.valdes@gmail.com. (J.H. Valdes).
[1] Current address: Centro de Genómica y Bioinformática, Facultad de Ciencias, Universidad Mayor, Santiago, Chile.
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

2. Experimental procedures

Chilean strains LF-89 (ATCC VR-1361) and EM-90 were isolated from Coho salmon in 1989 and from Atlantic salmon in 1990 respectively; T-GIM strain was isolated from Los Lagos region from rainbow trout in 2010 and the S-GIM strain was isolated from the Aysen region in 2012 from Atlantic salmon (Table 1).

The four *P. salmonis* strains were maintained on blood agar and grown on liquid culture in modified MC-8 medium [2]. DNA extraction was carried out with the AxyPrep™ Bacterial Genomic DNA Miniprep Kit (Axygen) according to the manufacturer's instructions. The genomic sequences were obtained using Illumina MiSeq technology with a 150 bp paired end library. Assembly was performed using the CISA software [3]. Gene annotation was performed using an adaptation of the Tanenbaum's protocol [4]. For reference transcriptome sequencing, total RNA was extracted with the GeneJET RNA Purification kit combined with the TURBO™ DNA-free™ kit (Thermo Fisher Scientific Inc.). The Ribo-Zero rRNA Removal Kit (Bacteria) (Illumina) was used to deplete rRNA, and the TruSeq RNA Library Preparation Kit v3 (Illumina) was used for library preparation

Table 1
General characteristics of the strains used in this study.

Strain name	ID	Year	Host	Genome accession number (NCBI)	Transcriptome accession number (NCBI)
P. salmonis LF-89	LF-89	1989	*O. kisutch*	AMFF00000000	SRX3259096
P. salmonis EM-90	EM-90	1990	*S. salar*	JRHP00000000	SRX3259647
P. salmonis T-GIM	T-GIM	2010	*O. mykiss*	JRHD00000000	SRX3259625
P. salmonis S-GIM	S-GIM	2012	*S. salar*	PDFD00000000	SRX3259646

ID: strain identifier; NCBI: National Center for Biotechnology Information (http://ncbi.nlm.nih.gov).

Table 2
Data summary of the draft genome and transcriptome sequences generated in this study.

ID Strain	Genome length (bp)	%GC	N50	Contigs	Genes	tRNA	TU's in the assembly	Average N° of ORF in TU's
LF-89	3,029,115	39.5	15,344	406	2894	49	2541	4.7
EM-90	3,035,439	40.0	13,604	534	3293	54	3250	4.5
T-GIM	2,936,546	39.5	20,087	342	2998	53	2642	4.9
S-GIM	3,061,568	40.7	14,584	547	2508	55	2511	5.1

bp: base pairs; %GC: %GC content; tRNA: transfer RNA; TU's: transcriptional units; ORF: open reading frame.

and sequencing. Quality trimming was performed using Trimmomatic [5], reads were assembled using Trinity, and TransDecoder was used to find coding sequences [6], which were annotated using BLAST against the SwissProt and NR databases.

3. Results and discussion

Genome sizes ranged approximately from 2.9 to 3 million bp and the % G + C was similar among all the isolates (average 40%) (Table 2). We identified 2541, 3250, 2642, 2511 and 3250 transcriptional units (TU) for strains LF-89, EM-90, S-GIM and T-GIM respectively. Up to 48 ORFs were predicted to belong to the same TU which encode for ribosomal structure and biosynthesis proteins, while the average ORF per TU ranged between 4,5 and 5,1 for the four strains (Table 2). Since complete genomes are available for strains LF-89 and EM-90, in this case transcriptome data was mapped against draft genome sequences provided in this study to avoid any potential inconsistencies generated by using different bacterial cultures.

We inspected the transcriptional architecture of previously characterized virulence factors and trait-specific genes associated to cation uptake, metal efflux, antibiotic resistance, secretion systems (types I, II, III, IV and VI), peptidoglycan transport, siderophore uptake and metabolism. We also identified and refined the annotations of several transporters for organic molecules, genes potentially involved in bacterial chemotaxis as well as components of the pili and the flagellum, which have been shown to play a key role in the adhesion and colonization processes of the bacteria in the host [7].

Finally, we detected between two and five copies of *luxR* gene, but we didn't identify *luxI* gene orthologues. The LuxI/LuxR system is involved in the quorum sensing of Gram-negative bacteria as *Vibrio fischeri* [8], and *A. salmonicida* [9], controlling biofilm formation, secretion of virulence factors and bioluminescence [10]. Recently genome data have shown a significant amount of species that only encode the *luxR* gene component [11].

Our results, in addition to those previously reported genome [12,13,14,15,16,17,18,19] and transcriptome information [20], provide an extended and refined genomic annotation and have also revealed unreported aspects of transcriptional architecture and its variability among the strains with available genome data. These results have also provided new insights on the unexplored coding potential and have added more detailed information about previous findings regarding virulence and pathogenicity mechanisms of this microbial pathogen.

Conflicts of interest

The authors declare no conflict of interest.

Financial support

References

[1] Mauel MJ, Miller DL, Styer E, et al. Occurrence of Piscirickettsiosis-like syndrome in tilapia in the continental United States. J Vet Diagn Invest 2005;17(6):601–5. https://doi.org/10.1177/104063870501700616.

[2] Marshall SH, Gómez FA, Ramírez R, et al. Biofilm generation by *Piscirickettsia salmonis* under growth stress conditions: a putative in vivo survival/persistence strategy in marine environments. Res Microbiol 2012;163(8):557–66. https://doi.org/10.1016/j.resmic.2012.08.002.

[3] Lin SH, Liao YC. CISA: Contig Integrator for Sequence Assembly of Bacterial Genomes. PLoS One 2013;8(3):e60843. https://doi.org/10.1371/journal.pone.0060843.

[4] Tanenbaum DM, Goll J, Murphy S, et al. The JCVI standard operating procedure for annotating prokaryotic metagenomic shotgun sequencing data. Stand Genomic Sci 2010;2(2):229–37. https://doi.org/10.4056/sigs.651139.

[5] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170.

[6] Haas BJ, Papanicolaou A, Yassour M, et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 2013;8:1494–512. https://doi.org/10.1038/nprot.2013.084.

[7] Sudheesh PS, Al-Ghabshi A, Al-Mazrooei N, et al. Comparative pathogenomics of bacteria causing infectious diseases in fish. Int J Evol Biol 2012;457264. https://doi.org/10.1155/2012/457264.

[8] Fuqua WC, Winans SC, Greenberg EP. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. J Bacteriol 1994;176(2):269–75. https://doi.org/10.1128/jb.176.2.269-275.1994.

[9] Tsai CS, Winans SC. LuxR-type quorum-sensing regulators that are detached from common scents. Mol Microbiol 2010;77(5):1072–82. https://doi.org/10.1111/j.1365-2958.2010.07279.x.

[10] Engebrecht J, Nealson K, Silverman M. Bacterial bioluminescence: isolation and genetic analysis of functions from *Vibrio fischeri*. Cell 1983;32(3):773–81. https://doi.org/10.1016/0092-8674(83)90063-6.

[11] Hudaiberdiev S, Choudhary KS, Vera Alvarez R, et al. Census of solo LuxR genes in prokaryotic genomes. Front Cell Infect Microbiol 2015;5(20). https://doi.org/10.3389/fcimb.2015.00020.

[12] Nourdin-Galindo G, Sánchez P, Molina CF, et al. Comparative Pan-Genome Analysis of *Piscirickettsia salmonis* Reveals Genomic Divergences within Genogroups. Front Cell Infect Microbiol 2017;7(459). https://doi.org/10.3389/fcimb.2017.00459.

[13] Bohle H, Henríquez P, Grothusen H, et al. The Genome Sequence of an Oxytetracycline-Resistant Isolate of the Fish Pathogen *Piscirickettsia salmonis* Harbors a Multidrug Resistance Plasmid. Genome Announc 2017;5(5):e01571-16. https://doi.org/10.1128/genomea.01571-16.

[14] Bravo C, Martinez V. Whole-genome comparative analysis of the pathogen *Piscirickettsia salmonis*. Vet Microbiol 2016;196:36–43. https://doi.org/10.1016/j.vetmic.2016.10.015.

[15] Fuentealba P, Aros C, Latorre Y, et al. Genome-scale metabolic reconstruction for the insidious bacterium in aquaculture *Piscirickettsia salmonis*. Bioresour Technol 2017;223:105–14. https://doi.org/10.1016/j.biortech.2016.10.024.

[16] Pulgar R, Travisany D, Zuñiga A, et al. Complete genome sequence of *Piscirickettsia salmonis* LF-89 (ATCC VR-1361) a major pathogen of farmed salmonid fish. J Biotechnol 2015;212:30–1. https://doi.org/10.1016/j.jbiotec.2015.07.017.

[17] Bohle H, Henríquez P, Grothusen H, et al. Comparative Genome Analysis of Two Isolates of the Fish Pathogen *Piscirickettsia salmonis* from Different Hosts Reveals Major Differences in Virulence-Associated Secretion Systems. Genome Announc 2014;2 (6):e01219-14. https://doi.org/10.1128/genomea.01219-14.

[18] Yañez AJ, Molina C, Haro RE, et al. Draft Genome Sequence of Virulent Strain AUSTRAL-005 of *Piscirickettsia salmonis*, the Etiological Agent of Piscirickettsiosis. Genome Announc 2014;2(5):e00990-14. https://doi.org/10.1128/genomea.00990-14.

[19] Eppinger M, McNair K, Zogaj X, Dinsdale EA, Edwards RA. Klose et al. Draft Genome Sequence of the Fish Pathogen *Piscirickettsia salmonis*. Genome Announc 2013;1(6): e00926-3. https://doi.org/10.1128/genomea.00926-13.

[20] Machuca A, Martinez V. Transcriptome Analysis of the Intracellular Facultative Pathogen *Piscirickettsia salmonis*: Expression of Putative Groups of Genes Associated with Virulence and Iron Metabolism. PLoS One 2016;11(12): e0168855. https://doi.org/10.1371/journal.pone.0168855.