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 \(^{b,*}\)

\(^a\) 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

**A R T I C L E  I N F O**

Article history:
Received 5 December 2017
Accepted 4 April 2018
Available online 12 April 2018

Keywords:
- Bacterial genomes
- Coding potential
- Comparative analysis
- Draft genome
- *Piscirickettsia salmonis*
- Reference transcriptome
- Refined annotation
- Salmon Rickettsial Syndrome
- Salmonids

**A B S T R A C T**

**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.


© 2018 Pontificia Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. All rights reserved. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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).

\(^\dagger\) 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 GenJet 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.
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, 2520, 2642, 2511 and 2520 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 lux 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

This work was supported by InnovaChile–CORFO (FCR–CSB 09CIEI-6951) grant to D.E.F., S.H.M. and J.H.V.; FIE V014 to S.H.M. and F.A.G.; FONDECYT INICIACION 11101434 to J.H.V. and FONDECYT 1170342 to S.H.M.; CONICYT National PhD Scholarship and CONICYT-PAI 7813110018 to A.D.M.

References


