Draft Genome Sequence of Virulent Strain AUSTRAL-005 of *Piscirickettsia salmonis*, the Etiological Agent of Piscirickettsiosis

Alejandro J. Yañez,[a,b,d] Cristian Molina,[b] Ronie E. Haro,[a] Patricio Sanchez,[a] Adolfo Isla,[a] Julio Mendoza,[e] Marcelo Rojas-Herrera,[c] Annette Trombert,[c] Andrea X. Silva,[b] Juan G. Cárcamo,[a,d] Jaime Figueroa,[a,d] Victor Polanco,[c] Patricio Manque,[c] Vinicius Maracaja-Coutinho,[c] Víctor H. Olavarría[a]

Instituto de Bioquímica y Microbiología, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile[a]; AUSTRAL-omics, Universidad Austral de Chile, Valdivia, Chile[b]; Centro de Genómica y Bioinformática, Facultad de Ciencias, Universidad Mayor, Santiago, Chile[c]; Interdisciplinary Center for Aquaculture Research (INCAR), Concepción, Chile[d]; Cermaq Chile, Puerto Montt, Chile[e]

We report here the draft genome sequence of a lethal pathogen of farmed salmonids, *Piscirickettsia salmonis* strain AUSTRAL-005. This virulent strain was isolated in 2008 from *Oncorhynchus mykiss* farms, and multiple genes involved in pathogenicity, environmental adaptation, and metabolic pathways were identified.

*P*iscirickettsia salmonis* is a member of the class *Gammaproteobacteria* and belongs to the family *Piscirickettsiaceae* (1, 2). *P. salmonis* is a coccoid, nonmotile, aerobic, and intracellular bacterial pathogen isolated from infected farmed salmonids in the south of Chile (3). This bacterium is the etiological agent of piscirickettsiosis or septicemia pisirickettsial of salmonids (SPS) and produces a systemic infection of several organs, such as the kidney, liver, spleen, intestine, brain, ovary, and gills, which leads to cell vacuolation and apoptosis, causing high mortality (1). *P. salmonis* is widely distributed (4, 5), being one the main pathogens responsible for significant economic losses of salmonid aquaculture worldwide. The control of SPS has been highly inefficient due to the low efficacy of vaccination (6) and the emergence of antibiotic-resistant isolates (7). These problems open the possibilities of genome sequence helping to understand the molecular mechanisms of pathogenicity and resistance.

The AUSTRAL-005 resistant strain was isolated from *Oncorhynchus mykiss* in AUSTRAL-SRS medium (7, 8). The draft genome sequence was obtained with a shotgun strategy using 454 GS Junior, Illumina MiSeq, and one Ion Torrent sequencing technology run. A total of 2,286,585 single reads and 318,027 paired-end reads, with an average length of 456 nucleotides (76× coverage), were *de novo* assembled using the GS *de novo* Assembler. The Mix software (9) was applied to improve the assembly using a draft assembly of 418 contigs to obtain extended contigs. The Mix software uses two draft assemblies to reduce contig fragmentation through overlapping of its extremes and producing extended ones. A total of 29 scaffolds were constructed, with an N_{50} of 29,089 bp; the largest assembled scaffold was 110,992 bp, with a mean length of 25,720 bp. The draft genome size was 3,529,595 bp, with a G+C content of 38.39%. We performed a clustering process with 95% identity using CD-HIT (10) to the predicted the open reading frames (ORFs) obtained with Glimmer3 (11). Next, annotation was performed using BLASTx (12),

followed by the Blast2Go tool (13). The annotation resulted in a total of 2,118 protein-coding gene predictions, with 2,035 well-annotated genes and 432 genes that encode hypothetical proteins, and 83 are unknown genes. A total of 52 tRNAs and 3 rRNAs were identified using tRNAscan-SE (14) and RNAmmer (15), respectively.

The annotation process resulted in the identification of genes associated with virulence factors, environmental adaptation, and metabolic pathways, and multiple insertion sequences were also identified. The genome analysis showed two putative toxin-antitoxin systems (TA), *higAB* and *txe/yoeB*, and four putative genes encoding proteases (*clp*, *lon*, *M22*, and *M48*). The genome sequence analysis also revealed protein secretion system types I, II, and IV. Moreover, six putative genes encoding heat shock proteins (HSP) were identified (Hsp33, GroE, Hsp90, DnaJ, Hsp70, and Hsp20). Regarding iron metabolism and its transport, we report four putative genes encoding siderophore-related proteins, one *hemH* gene, two *tonB* genes, and one *fur* gene. Surprisingly, the genome of this immobile bacterium revealed genes encoding components of flagella. This genome sequence represents a biotechnological opportunity to develop new therapies to counteract SPS.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. AZYQ00000000. The version described in this paper is the first version, AZYQ01000000.

ACKNOWLEDGMENTS

REFERENCES

1. **Fryer JL, Hedrick RP.** 2003. *Piscirickettsia salmonis*: a gram-negative intracellular bacterial pathogen of fish. J. Fish Dis. **26**:251–262. http://dx.doi.org/10.1046/j.1365-2761.2003.00460.x.
2. **Fryer JL, Lannan CN, Garces LH, Larenas JJ, Smith PA.** 1990. Isolation

of a *Rickettsiales*-like organism from diseased coho salmon (*Oncorhynchus kisutch*) in Chile. Fish Pathol. **25**:107–114. http://dx.doi.org/10.3147/jsfp.25.107.

3. **Fryer JL, Lannan CN, Giovannoni SJ, Wood ND.** 1992. *Piscirickettsia salmonis* gen. nov., sp. nov., the causative agent of an epizootic disease in salmonid fishes. Int. J. Syst. Bacteriol. **42**:120–126. http://dx.doi.org/10.1099/00207713-42-1-120.

4. **Cusack RR, Groman DB, Jones SRM.** 2002. Rickettsial infection in farmed Atlantic salmon in eastern Canada. Can. Vet. J. **43**:435–440.

5. **Arkush KD, McBride AM, Mendonca HL, Okihiro MS, Andree KB, Marshall S, Henriquez V, Hedrick RP.** 2005. Genetic characterization and experimental pathogenesis of *Piscirickettsia salmonis* isolated from white seabass *Atractoscion nobilis*. Dis. Aquat. Organ. **63**:139–149. http://dx.doi.org/10.3354/dao063139.

6. **Kuzyk MA, Burian J, Machander D, Dolhaine D, Cameron S, Thornton JC, Kay WW.** 2001. An efficacious recombinant subunit vaccine against the salmonid rickettsial pathogen *Piscirickettsia salmonis*. Vaccine **19**:2337–2344. http://dx.doi.org/10.1016/S0264-410X(00)00524-7.

7. **Yáñez AJ, Valenzuela K, Matzner C, Olavarría V, Figueroa J, Avendaño-Herrera R, Carcamo JG.** 2013. Broth microdilution protocol for minimum inhibitory concentration (MIC) determinations of the intracellular salmonid pathogen *Piscirickettsia salmonis* to florfenicol and oxytetracycline. J. Fish Dis. **37**:505–509. http://dx.doi.org/10.1111/jfd.12144.

8. **Yáñez AJ, Silva H, Valenzuela K, Pontigo JP, Godoy M, Troncoso J, Romero A, Figueroa J, Carcamo JG, Avendaño-Herrera R.** 2013. Two novel blood-free solid media for the culture of the salmonid pathogen *Piscirickettsia salmonis*. J. Fish Dis. **36**:587–591. http://dx.doi.org/10.1111/jfd.12034.

9. **Soueidan H, Maurier F, Groppi A, Sirand-Pugnet P, Tardy F, Citti C, Dupuy V, Nikolski M.** 2013. Finishing bacterial genome assemblies with Mix. BMC Bioinformatics **14**(Suppl 15):S16. http://dx.doi.org/10.1186/1471-2105-14-S15-S16.

10. **Li W, Jaroszewski L, Godzik A.** 2001. Clustering of highly homologous sequences to reduce the size of large protein databases. Bioinformatics **17**:282–283. http://dx.doi.org/10.1093/bioinformatics/17.3.282.

11. **Delcher AL, Bratke KA, Powers EC, Salzberg SL.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23**:673–679. http://dx.doi.org/10.1093/bioinformatics/btm009.

12. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.

13. **Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, Robles M.** 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21**:3674–3676. http://dx.doi.org/10.1093/bioinformatics/bti610.

14. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25**:955–964. http://dx.doi.org/10.1093/nar/25.5.0955.

15. **Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35**:3100–3108. http://dx.doi.org/10.1093/nar/gkm160.