Comparative Genomic Analysis of Two Chilean *Renibacterium salmoninarum* Isolates and the Type Strain ATCC 33209\textsuperscript{T}

Jorn Bethke\textsuperscript{1,2}, Alejandro J. Yáñez\textsuperscript{2,3}, and Ruben Avendaño-Herrera\textsuperscript{1,2,4,*}

\textsuperscript{1}Laboratorio de Patología de Organismos Acuáticos y Biotecnología Acuícola, Facultad de Ciencias de la Vida, Universidad Andrés Bello, Viña del Mar, Chile
\textsuperscript{2}Interdisciplinary Center for Aquaculture Research (INCAR), Viña del Mar, Chile
\textsuperscript{3}Instituto de Bioquímica y Microbiología, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile
\textsuperscript{4}Centro de Investigación Marina Quintay (CIMARQ), Universidad Andrés Bello, Quintay, Chile

*Corresponding author: E-mail: reavendano@yahoo.com; ravendano@unab.cl.

Accepted: June 29, 2018

Data deposition: This project has been deposited in DDBJ/ENA/GenBank under Accession PRJNA418717. The FASTA sequences of the genomes for Chilean isolates H-2 and DJ2R have been deposited in DDBJ/ENA/GenBank under Accessions CP024979 and CP024980, respectively. The *R. salmoninarum* ATCC 33209\textsuperscript{T} genome, used for comparative genomics, is deposited in GenBank under Accession NC_010168.1.

Abstract

*Renibacterium salmoninarum*, a slow-growing facultative intracellular pathogen belonging to the high C+G content Actinobacteria phylum, is the causative agent of bacterial kidney disease, a progressive granulomatous infection affecting salmonids worldwide. This Gram-positive bacterium has existed in the Chilean salmonid industry for >30 years, but little or no information is available regarding the virulence mechanisms and genomic characteristics of Chilean isolates. In this study, the genomes of two Chilean isolates (H-2 and DJ2R) were sequenced, and a search was conducted for genes and proteins involved in virulence and pathogenicity, and we compare with the type strain ATCC 33209\textsuperscript{T} genome. The genome sizes of H-2 and DJ2R are 3,155,332 bp and 3,155,228 bp, respectively. They genomes presented six ribosomal RNA, 46 transcription RNA, and 25 noncodingRNA, and both had the same 56.27\% G+C content described for the type strain ATCC 33209\textsuperscript{T}. A total of 3,522 and 3,527 coding sequences were found for H-2 and DJ2R, respectively. Meanwhile, the ATCC 33209\textsuperscript{T} type strain had 3,519 coding sequences. The in silico genome analysis revealed a genes related to tricarboxylic acid cycle, glycolysis, iron transport and others metabolic pathway. Also, the data indicated that *R. salmoninarum* may have a variety of possible virulence-factor and antibiotic-resistance strategies. Interestingly, many of genes had high identities with *Mycobacterium* species, a known pathogenic Actinobacteria bacterium. In summary, this study provides the first insights into and initial steps towards understanding the molecular basis of antibiotic resistance, virulence mechanisms and host/environment adaptation in two Chilean *R. salmoninarum* isolates that contain proteins of which were similar to those of *Mycobacterium*. Furthermore, important information is presented that could facilitate the development of preventive and treatment measures against *R. salmoninarum* in Chile and worldwide.

Key words: BKD, genome, pathogenicity, bacterial kidney disease, virulence factors.

Introduction

*Renibacterium salmoninarum*, a slow-growing facultative intracellular pathogen belonging to the high C+G content Actinobacteria phylum, is the causative agent of bacterial kidney disease, which is a chronic, progressive, and granulomatous infection threatening farmed and wild salmonids worldwide (Wiens 2011). This Gram-positive microorganism is one of few that can be transmitted both vertically, via intraovum infection (Evelyn, Ketcheson, et al. 1986; Evelyn, Prosperi-Porta, et al. 1986), and horizontally, via a shared water supply (Evelyn et al. 1981; Austin and Rayment 1985; Balfry et al. 1996). The virulence of *R. salmoninarum* has been linked to a synergistic interaction between toxins and enzymes present in extracellular products (Bruno and Munro 1982; Shieh 1988), as well as with iron acquisition mechanisms (Grayson et al. 1995; Bethke et al. 2016). Nevertheless,
the most thoroughly characterized virulence factor is a 57-kDa protein (p57) present on the bacterial cell surface and that can be secreted into growth media and host environments (Wiens 2011). Renibacterium salmoninarum can be internalized into macrophages and nonphagocytic cells, such as in the cell lines CHSE-214 and RTG-2 (McIntosh et al. 1996; González et al. 1999), where bacteria produce putative virulence factors and replicate. These transmission strategies, together with an ability to survive phagocytosis and replicate within macrophages, complicate the prophylaxis and prevention of R. salmoninarum infection (Bandin et al. 1993; Gutenberger et al. 1997; Sudheesh et al. 2007). In Chile, R. salmoninarum was initially isolated from chum salmon (Oncorhynchus keta) reared in seawater cages (Sanders and Barros 1986), but this bacterium has since been reported in Atlantic salmon (Salmo salar) and coho salmon (Oncorhynchus kisutch) farmed across distinct geographical regions of the country. While this pathogen has been present in Chile for >30 years, little or no information is available regarding virulence mechanisms and genetic characteristics.

Therefore, the general aim of this study was to gain initial genomic insights into Chilean R. salmoninarum isolates, with the specific goal of obtaining better understandings on the virulence and pathogenicity activities of proteins other than p57. This new information could be vital for the development of novel management and treatment tools against bacterial kidney disease in both the Chilean and worldwide salmonid industries.

Materials and Methods

Genomic DNA and Sequencing

Two previously characterized Chilean isolates (H-2 and DJ2R), obtained from cage-cultured Atlantic salmon with clinical signs of bacterial kidney disease in southern Chile, were used (Bethke et al. 2016, 2017). The bacteria were routinely cultured in KDM-2 agar for 15–20 days at 15°C. For sequencing, genomic DNA of the two isolates was extracted using the InstaGene Purification Matrix (Bio-Rad) according to manufacturer instructions. The DJ2R genome was sequenced using an Illumina MiSeq platform with 2 × 250 paired-end reads by the AUSTRAL-omics Institute, hosted by the Universidad Austral de Chile (Valdivia, Chile). Using the same technology and parameters, H-2 genomic DNA was sequenced by the Central Support Service for Experimental Research (SCSIE, Spanish acronym) at the University of Valencia (Valencia, Spain).

Genome Assembly, Annotation, and Analysis

Prior to genome de novo assembly, Illumina reads were analysed for quality using FASTQC (Brabaham Bioinformatics). The reads were paired and quality filtered using PEAR (Zhang et al. 2014). Later, paired reads were assembled to the R. salmoninarum ATCC 33209T reference genome (accession number NC_010168.1) using Genious v10.2.3 (Kearse et al. 2012), and assembly quality was checked by QUAST (Gurevich et al. 2013). Subsequent rapid annotation was performed using the Prokka v1.12 prokaryotic genome software (Victorian Bioinformatics Consortium) (Seemann 2014). Annotation was performed using a genus database generated from Renibacterium salmoninarum NC_010168.1 proteins and the –usegenus option. The genome maps were constructed using BRIG (BLAST Ring ImageGenerator), a prokaryote genome comparison software (Alikhan et al. 2011). A search was conducted for possible virulence and antibiotic resistance factors using crb-blast against the Virulence Factors Database (VFD) (Chen et al. 2016) and Comprehensive Antibiotic Resistance Database (CARD) (Jia et al. 2017), which respectively store information for amino acids and nucleotides. Since no pathogenic bacteria are closely related to R. salmoninarum, a 65% identity was used to filter crb-blast results obtained against the amino acid and nucleotide databases.

Iron uptake is a known essential factor in the pathogenicity of several bacteria. Therefore, a search was conducted for genes possibly involved in iron regulation; this search used the IdeR/DtxR binding sites of Corynebacterium diphtheria (Yellaboina et al. 2004) and Mycobacterium tuberculosis (Prakash et al. 2005), two known pathogenic bacteria from the Actinobacteria phylum. To evaluate overall relatedness between isolates and the type strain genomes, an average nucleotide identity analysis was performed using EZBioCloud OrthoANLu (Yoon et al. 2017). Further analyses and a detection of single nucleotide polymorphisms (SNPs) were performed by aligning genomes using the Mauve v2.3.1 plugin in the Geneious v10.2.3 software (Kearse et al. 2012). Additional genes of interest were located using key words such as toxin, drug, lipoprotein, secretion, iron, siderophore, heme/hemin, resistance, histidine kinase, and shock. Further examination of the resulting genes was conducted using Artemis v16.0.0 (Rutherford et al. 2000).

Genome Deposit in Public Databases

This project has been deposited in DDBJ/ENA/GenBank under Accession PRJNA418717. The FASTA sequences of the genomes for Chilean isolates H-2 and DJ2R have been deposited in DDBJ/ENA/GenBank under Accessions CP029236 and CP029237, respectively. The R. salmoninarum ATCC 33209T genome, used for comparative genomics, is deposited in GenBank under Accession NC_010168.1.

Results

General Genomic Features

The H-2 and DJ2R genomes were assembled into 3,155,332 bp and 3,155,228 bp, respectively. The two
Chilean isolate genomes presented six ribosomal RNA (rRNA), 46 transcription RNA, and 25 noncoding RNA, and both had the same 56.27% G + C content described for the type strain ATCC 33209<sup>T</sup>. A total of 3,522 and 3,527 coding sequences were found for H-2 and DJ2R, respectively. Meanwhile, the ATCC 33209<sup>T</sup> type strain had 3,519 coding sequences. All these results were similar to and in concordance with the sequence of the ATCC 33209<sup>T</sup> type strain, as reported by Wiens et al. (2008) (fig. 1). In this regard, nucleotide identities, on average, showed high similarities, with 99.99% similarity found between the type strain ATCC 33209<sup>T</sup> and the two Chilean isolates. Nevertheless, coverage percentages varied. The H-2 and DJ2R isolates, respectively, showed 96.28% and 98.35% coverage with the type strain. Between the two isolates, however, there was 94.69% coverage, suggesting differences between genomes.

**Antibiotic Resistance**

Regarding antibiotic resistance, crb-blast results against the amino acid database showed ten proteins with identities between 65.7% and 80.2%. Some of these putative resistance proteins corresponded to ABC transporters of *Paenibacillus* sp. LC231, including tetracycline tetA (80.2%), tetA (65.7%), and tetB (74.3%). Also related to resistance was a mutation in the rpsL 30S ribosomal protein S12 (82.9%) of LC231, including tetracycline tetA (80.2%), tetA (65.7%), tetB (74.3%), and tetracycline tetA (80.2%), tetA (65.7%), and tetB (74.3%).

**Iron Acquisition**

Iron acquisition mechanisms are an essential virulence factor in many pathogenic bacteria. These bacteria need such mechanisms to cope with host nutritional immunity (Cassat and Skaar 2013). In Gram-positive bacteria with a high CG content, such as Actinobacteria, iron-acquisition related genes are regulated by IdeR/DtxR. Therefore, a search was conducted for IdeR binding sites. A total of ten IdeR/DtxR binding sites were found (supplementary fig. 1, Supplementary Material online), and, as expected, iron acquisition-related genes were present next to these binding sites. Some of the detected genes encoded proteins related to iron transport, siderophores interaction, siderophores exportation, haeme acquisition, and bacterioferritin, among others. Other genes with no downstream IdeR-binding site, but with involvement in iron acquisition, were also found.

**Discussion**

The present study provides the first genomic insights for two *R. salmoninarum* isolates from fish farmed in Chile. The genomes of the two Chilean isolates were de novo assembled and compared to the type strain ATCC 33209<sup>T</sup>, showing high sequence similarity. This similarity was expected due to the genetic homogeneity described for this fish pathogen (Grayson et al. 1999, 2000; Rhodes et al. 2000; Alexander et al. 2001). Nevertheless, some differences in coverage percentages were found by ANIu analysis, as expected. Multiple proteins that could be related to antibiotic resistance were found, indicating that *R. salmoninarum* could have different strategies for coping with antibiotics. On-going research will soon allow us to clarify this hypothesis.

The detected resistance-related mechanisms included the tetracycline resistance proteins tetA and tetB (supplementary table 1, Supplementary Material online). A response regulator (mtrA) showing 71.8% identity with *Mycobacterium* spp. was found. Notably, a capsular polysaccharide synthesis (capE) protein was found that belonged to the non-related *Staphylococcus haemolyticus* JCSC1435 bacterium (66.8% identity; supplementary table 1, Supplementary Material online). A number of other known virulence factors related to resistance, secretion systems, toxins, lipoproteins, histidine kinase sensors, and shock were found (supplementary table 2, Supplementary Material online). Additionally, 274 SNP events were found in the Chilean isolates, as compared to the type strain ATCC 33209<sup>T</sup> genome (supplementary table 3, Supplementary Material online). Some SNPs were located in genes coding for virulence factors, such as p57, which is the most characterized virulence factor of *R. salmoninarum*. A SNP was present in the Chilean H-2 isolate, specifically where the mtrA gene is located (i.e., between nucleotides 945,101 and 946,777). This included a change from thymine to cytosine in the gene sequence at nucleotide 677. This modification would involve a valine<sup>226</sup> to alanine modification in the amino acidic sequence.
associated with the lipoprotein LpqB, which is located directly downstream MtrAB and directly affects the phosphorylation of MrtA. Prior research supports that LpqB mutants with transposon insertion are multidrug sensitive (Nguyen et al. 2010). This particular lipoprotein was also found directly downstream the MtrAB two-component system in \textit{R. salmoninarum}, indicating that LpqB might play a role in drug resistance in this bacterium. High identity percentages were also found for 16S (88.7%) and 23S (83.9%) rRNA-mutated genes, a finding that suggests some mutation(s) in these genes may confer antibiotic resistance, such as in \textit{Mycobacterium} species.

The detected proteins that may play important roles as virulence factors evidenced high identity percentages with...
Mycobacterium spp., a related pathogenic genus. The protein with the highest identity (84%) was the GroEL chaperone (Hsp60) of M. tuberculosis H37Rv. This immune-dominant antigen plays essential roles in virulence, disease development, and survival under stressful conditions (Goulhen et al. 1998; Neckers and Tatu 2008). The location of the GroEL-like protein seems to depend on the organism and has been reported in compartments other than the cytoplasm (Scopio et al. 1994; Esaguay and Aguas 1997; Frisk et al. 1998) and in association with the cell surface (Gillis et al. 1985; Engraber and Loos 1992; Phadnis et al. 1996). Indeed, M. tuberculosis Hsp60 has various “moonlighting” functions and can act as secreted signalling molecules, modulators of host immunity, surface-located bacterial ligands, and bacterial cell-wall components, leaving the cytosol to function extracellularly or at the surface of the bacterial cell wall (Henderson 2017). Interestingly, GroEL also actively participates in cell adherence for Clostridium difficile (Hennequin et al. 2001) and Legionella pneumophila, playing a key role in supporting intracellular behaviour through host–cell contact and internalization (Fernandez et al. 1996; Garduño, Faulkner, et al. 1998; Garduño, Garduño, et al. 1998). Therefore, the detected R. salmoninarum GroEL protein is most likely involved in different processes, as other bacterial pathogens. The intracellular behaviour of GroEL could have key functions in infection and disease development (McIntosh et al. 1996; González et al. 1999). Nevertheless, further studies are needed to identify the location and function of GroEL as a virulence factor in R. salmoninarum.

The sigma factors SigA and SigH were also among the feasible virulence factors detected in R. salmoninarum. These factors are critical in regulating different cellular responses that sustain multiple stages of the host–pathogen interaction. For example, SigA is indispensable for Mycobacterium spp. growth, while SigH is a central regulator of the oxidative and heat stress responses induced during macrophage infection (Sachdeva et al. 2010).

Another possible virulence factor in R. salmoninarum was sodA, an iron-cofactored superoxide dismutase. These dismutases are ubiquitous metalloenzymes that catalyse the conversion of superoxide anions to hydrogen peroxide, thus playing a key role during infection. Prior research has found that superoxide dismutase-attenuated strains of Mycobacterium are less resistant to hydrogen peroxide and have a poorer survival outlook in mouse models (Edwards et al. 2001).

Also crucial for bacterial multiplication and, therefore, the pathogenicity of several bacteria is the ability to uptake iron during infection (Skaar 2010; Saha et al. 2013). In Gram-positive bacteria with high CG content, such as Actinobacteria, iron acquisition-related genes are regulated by IdeR/DtxR. Ten IdeR binding sites were found upstream of important iron acquisition-related genes in R. salmoninarum. Some of these genes may be inactivated by point mutations or frame shifts; however, IdeR binding sites were found in relation to clusters associated with siderophores transport (Rsal33209_1684 to Rsal33209_1681) (supplementary fig. 1A, Supplementary Material online), the intact coding sequence of a fepr homolog (RSal33209_3347) (supplementary fig. 1B, Supplementary Material online) (Wiens et al. 2008). Also, two IdeR binding sites related to a haemin uptake cluster were found (supplementary fig. 1C, Supplementary Material online). It has been demonstrated that R. salmoninarum can synthesizes siderophores (shown by a CAS assay) and has the ability to use a variety of iron sources, including haemin (Bethke et al. 2016). These data would indicate that R. salmoninarum may have different strategies to confront nutritional immunity in the host.

Regarding the detected SNP event (i.e., thymine to cytosine), it is worth highlighting that changes in certain virulence-related proteins may affect function during infection. Such is the case of the p57 virulence factor, where the valine226 to alanine change may impact protein function in ways similar to those reported by Wiens et al. (2002); namely, a single Ala139 to Glu substitution enhances the binding capacity of p57 to Chinook leukocytes, but further experiments are needed to corroborate this. Likewise, other SNPs could have significant functional repercussions, thereby influencing the virulence and pathogenicity of the different R. salmoninarum isolates. Some of these detected SNPs included: Arg242 to Leu in the cold shock protein; Ser273 to Arg in the LipO lipoprotein; Thr99 to Lys in the MFS transporter, and cysteine to thymine in nucleotide 391 of the type II secretion protein F, a change that generates the TAA stop codon and, consequently, a truncated protein. Therefore, the variety of proteins that could be involved in antibiotic-resistance and virulence found for R. salmoninarum may vary due to SNP events. This would mean that virulence and pathogenic competence could vary among R. salmoninarum isolates, but further experimentation is needed to corroborate this.

In conclusion, while R. salmoninarum has existed in the Chilean salmon industry for >30 years, this is the first report to provide a more detailed insight into the genomic characteristics of Chilean isolates of this pathogen. Crucial genomic information was obtained in regards to possible genes involved in virulence and pathogenicity of two Chilean R. salmoninarum isolates. This foundation of knowledge will be of aid in the future as a starting point to develop new preventive and treatment measures applicable against this pathogen worldwide.

**Supplementary Material**

Supplementary data are available at Genome Biology and Evolution online.

**Funding**

This work was supported by funding of the Comisión Nacional de Investigación Científica y Tecnológica (CONICYT, Chile) [Grant Numbers FONDAP No. 15110027 and FONDECYT No. 1150695]. J.B. also acknowledges support received by CONICYT [Doctoral Scholarship No. 21140421].
Literature Cited


Associate editor: Howard Ochman