Marine Genomics 31 (2017) 53–61

Contents lists available at ScienceDirect

Marine Genomics

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

Comparative genome analysis of two *Streptococcus phocae* subspecies provides novel insights into pathogenicity

J. Bethke [a,b], R. Avendaño-Herrera [a,b,c,*]

[a] Laboratorio de Patología de Organismos Acuáticos y Biotecnología Acuícola, Facultad de Ciencias Biológicas, Universidad Andrés Bello, Viña del Mar, Chile
[b] Interdisciplinary Center for Aquaculture Research (INCAR), Concepción, Chile
[c] Centro de Investigación Marinas Quintay (CIMARQ), Quintay, Chile

ARTICLE INFO

Article history:
Received 17 June 2016
Received in revised form 28 September 2016
Accepted 29 September 2016
Available online 5 October 2016

Keywords:
Streptococcus phocae subsp.
Pathogenicity
Atlantic salmon
Seal pathogen

ABSTRACT

Streptococcus phocae is a beta-hemolytic, Gram-positive bacterium that was first isolated in Norway from clinical specimens of harbor seal (*Phoca vitulina*) affected by pneumonia or respiratory infection, and in 2005, this bacterium was identified from disease outbreaks at an Atlantic salmon farm. A recent comparative polyphasic study reclassified *Streptococcus phocae* as subsp. *phocae* and subsp. *salmonis*, and there are currently two *S. phocae* NCBI sequencing projects for the type strains ATCC 51973^T and C-4^T. The present study compared these genome sequences to determine shared properties between the pathogenic mammalian and fish *S. phocae* subspecies. Both subspecies presented genomic islands, prophages, CRISPRs, and multiple gene activator and RofA regulator regions that could play key roles in the pathogenesis of streptococcal species. Likewise, proteins possibly influencing immune system evasion and virulence strategies were identified in both genomes, including Streptokinases, Streptolysin S, IgG endopeptidase, Fibronectin binding proteins, Daunorubicin, and Penicillin resistance proteins. Comparative differences in phage, non-phage, and genomic island sequences may form the genetic basis for the virulence, pathogenicity, and ability of *S. phocae* subsp. *salmonis* to infect and cause disease in Atlantic salmon, in contrast to *S. phocae* subsp. *phocae*. This comparative genomic study between two *S. phocae* subsp. provides novel insights into virulence factors and pathogenicity, offering important information that will facilitate the development of preventive and treatment measures against this pathogen.

1. Introduction

Streptococcus phocae, a beta-hemolytic bacterium of the pyogenic streptococcal group (Köhler, 2007), was first isolated in Norway from clinical specimens of harbor seal (*Phoca vitulina*) affected by pneumonia or respiratory infection (Skaar et al., 1994). This pathogen was further identified in other pinnipeds, including Cape fur seal (*Arctocephalus pusillus pusillus*) (Henton et al., 1999), ringed seal (*Phoca hispida*) (Raverty et al., 2004), grey seal (*Halichoerus grypus*) (Vossen et al., 2004), Caspian seal (*Phoca caspica*) (Kuiken et al., 2006), California sea lions (*Zalophus californianus*) (Johnson et al., 2006), spotted seal (*Phoca largha*) (Hueffer et al., 2011), and Steller sea lions (*Eumetopias jubatus*) (Lee et al., 2016), as well as in other marine mammals, including southern sea otters (*Enhydra lutris nereis*) (Imai et al., 2009). Furthermore, *S. phocae* has been isolated in several countries of northwestern Europe, North America, and Africa, in addition to being found in the Caspian Sea and South Korea. Since 1999, *S. phocae* has also been isolated from diseased Atlantic salmon (*Salmo salar*), causing serious economic losses in the salmon industry in Chile during the summer months (Romalde et al., 2008).

A recent comparative polyphasic study analyzed strains originating from different hosts (seal and salmon) to clarify the taxonomic position (Avendaño-Herrera et al., 2014a). This study reclassified Atlantic salmon *S. phocae* as *Streptococcus phocae* subsp. *salmonis* and proposed C-4^T as the type strain. This subspecies level status is shared only with *Streptococcus phocae* subsp. *phocae* isolated from seals, the type strain of which is ATCC 51973^T (Skaar et al., 1994). These two subspecies differ in biochemical and physiological traits, such as in hemolytic capacities, temperature growth range, and host specificity (Avendaño-Herrera et al., 2014a).

In fish, *S. phocae* subsp. *salmonis* produces exophthalmia with the accumulation of purulent and hemorrhagic fluid around the eyes, ventral petechial hemorrhages, skin abscesses, and, in some cases, muscle liquefaction with the formation of deep ulcerative areas (Romalde et al., 2008). In marine mammals, the clinical signs caused by *S. phocae* subsp. *phocae* vary by host and can include damage associated with phocine or canine distemper, conditions characterized by bronchointerstitial pneumonia, lymphocytic necrosis and depletion in the lymphoid organs, the presence of typical intracytoplasmic inclusion bodies

* Corresponding author at: Universidad Andrés Bello, Quillota 980, Piso 4, Torre C, Viña del Mar, Chile.
E-mail addresses: ravendano@unab.cl, reavendano@yahoo.com (R. Avendaño-Herrera).

http://dx.doi.org/10.1016/j.margen.2016.09.005
1874-7787/

in multiple epithelia (Vossen et al., 2004; Kuiken et al., 2006), and, as more recently discovered, gross enlargement of the uterus and associated lymph nodes (Hueffer et al., 2011).

Isolates of *S. phocae* from Atlantic salmon have been biochemically, antigenically, and genetically characterized as a homogeneous taxon (Romalde et al., 2008; Valdés et al., 2009) that is very distinct from seal isolates, including ATCC 51973^T (Skaar et al., 1994; Vossen et al., 2004). Despite this knowledge, scarce information exists regarding the pathogenic and virulence mechanisms of this bacterium. González-Contreras et al. (2011) suggest that surface hydrophobicity and binding capacity to fish cells, as well as the effects of live *S. phocae* cells, play important roles in pathogenicity during the early stages of the infective process, stages in which the pathogen must overcome the antibacterial activity of fish serum. Moreover, *S. phocae* survival in mucus could be a relevant *in vivo* factor facilitating host colonization and invasion. Cortez-San Martin et al. (2012) found that *S. phocae* subsp. *salmonis* can internalize in different fish and mammalian cell lines, although it is incapable of invading cell nuclei. Interestingly, Salazar et al. (2016) suggest that the non-specific humoral and cellular barriers of Atlantic salmon are immunologically insufficient against *S. phocae* subsp. *salmonis*, thereby facilitating successful infection.

There are currently two *S. phocae* NCBI sequencing projects, corresponding to the type strains ATCC 51973^T (Avendaño-Herrera and Poblete-Morales, 2015) and C-4^T (Avendaño-Herrera et al., 2014b). The present study compared these genome sequences to determine shared properties between the pathogenic mammalian and fish *S. phocae* subspecies. The aim of this genomic assessment was to provide novel understandings on *S. phocae* pathogenesis, particularly in relation to relevant markers for diagnosing pathogen development and epidemiology/transmission dynamics. Ultimately, genome comparison data could contribute to the development of novel preventive approaches against this pathogen.

2. Materials and methods

2.1. Sequencing and genomic comparison

Previously sequenced genomes for *S. phocae* subsp. *salmonis* (Accession no. JSAP00000000; Avendaño-Herrera et al., 2014a, 2014b) and *S. phocae* subsp. *phocae* (Accession no. LHQM00000000; Avendaño-Herrera and Poblete-Morales, 2015) were used for comparative analyses. Both *S. phocae* subsp. genome sequences were uploaded to the RAST (Aziz et al., 2008) and SEED servers (Overbeek et al., 2014). The whole genomes were aligned with the Mauve software (Darling et al., 2004). Circular plots of DNA sequences were generated using the BLAST Ring Image Generator (BRIG) (Alikhan et al., 2011), and linear plots were generated with the Easyfig 2.2.2 Genome Comparison Visualizer using BLASTN (Sullivan et al., 2011). Pathogenicity islands were detected and analyzed using Island Viewer (Dhillon et al., 2015). Clustered regularly interspaced short palindromic repeat (CRISPR) elements were detected using two web-based programs, CRISPI (Rousseau et al., 2009) and the CRISPR Finder (Grissa et al., 2007). To detect, identify, and graphically display putative prophage sequences, the PHAge Search Tool was used (Zhou et al., 2011). Genes of interest were located in Mauve or Artemis v.16.0.0 (Rutherford et al., 2000) using key words such as toxin, drug, iron, siderophore, surface, and resistance. To evaluate overall relatedness between both *S. phocae* subspecies, an average nucleotide identity (ANI) analysis was performed using OrthoANI (Lee et al., 2015).

3. Results

3.1. Genomic properties

The whole *S. phocae* subsp. *phocae* genome contained 1,700,445 bp, presented a GC content of 39.47%, and included 296 RAST-classified subsystems, 1747 coding sequences, and 37 RNA sequences, of which 35 were transfer RNA sequences. Meanwhile, the *S. phocae* subsp. *salmonis* genome contained 1,602,091 bp, presented a GC content of 39.58%, and had 287 subsystems, 1657 coding sequences, and 36 RNA sequences, of which 34 were transfer RNA sequences. The *S. phocae* subsp. *phocae* and *S. phocae* subsp. *salmonis* genomes presented genomic islands covering 2% and 4.3% of the entire genome, respectively. This gross difference in size might be due to the number of prophages found in each genome, with eight in *S. phocae* subsp. *phocae* and two in *S. phocae* subsp. *salmonis*. In turn, this difference could be related to the number of CRISPR arrays found in each genome (Fig. 1). Furthermore, BLASTN comparisons of the whole genomes showed that gene organization was not highly conserved, and a high number of regions were inverted in *S. phocae* subsp. *salmonis* in relation to *S. phocae* subsp. *phocae*, (Fig. 2). Nevertheless, the OrthoANI comparison showed 97.48% relatedness between the types of both *S. phocae* subspecies.

3.2. Prophages and virulence mechanisms

Some proteins encoded by prophages may notably influence bacterial virulence (Feiner et al., 2015). Therefore, the presence of prophages was evaluated using the PHAge Search Tool (Zhou et al., 2011). For *S. phocae* subsp. *phocae*, eight prophages were found with lengths ranging from 9.4 kbp to 62.3 kbp. According to PHAST scoring methods, two of these prophages were complete while the other six were incomplete. Altogether, the identified prophages covered 13% of the whole *S. phocae* subsp. *phocae* genome sequence. In turn, only two prophages where found in the *S. phocae* subsp. *salmonis* genome, one of which was intact (45.2 kbp) while the other was incomplete (15.8 kbp). Together, these prophages covered 3.8% of the entire genome.

In the case of *S. phocae* subsp. *phocae*, the prophage proteins coding sequences that could influence bacterial virulence included YaaA, metallopeptidases, AhpC, SsrA, hyaluronoglucosaminidase, ComX, and Streptodornase D. For *S. phocae* subsp. *salmonis*, of the prophage proteins that could possibly influence virulence, only hyaluronoglucosaminidase was identified (Table 1).

3.3. Genomic islands

Genomic islands are gene clusters, probably of horizontal origin, that encode genes related to virulence, antimicrobial resistance, and pathogenicity. In *S. phocae* subsp. *phocae*, four genomic islands were found that represented 2% of the whole genome. In *S. phocae* subsp. *salmonis*, five genomic islands were identified that covered 4.3% of the entire genome. Some of the genes in these genomic islands included hemolysin III, transcription regulators, cold shock proteins, hyaluronidase, multidrug and toxic compound extrusion family transporters, and secY.

3.4. Other virulence factors

In addition to the virulence-related proteins present in the genomic islands and acquired by the prophages, other proteins were found in the genomes that could play roles in bacterial pathogenicity and virulence. Proteins such as Streptokinases, Streptolysin S, Exfoliative toxin A, IgG endopeptidase, Collagen-like proteins, Fibronectin binding proteins, the Antiphagocytic M protein, Paratox, Daunorubicin and Penicillin resistance proteins, Betalactamase type C, the two component system involved in ferric iron transportation, and the heme uptake system were found in both genomes. Additionally, *S. phocae* subsp. *phocae* presented other proteins important for virulence, including proteins involved in iron uptake through siderophore transport, phospholipase A2, two-component lantibiotic system, and the precursor for the epidermin ladder peptide serine protease EpiP.