Comparative genomics and transcriptomics of *Flavobacterium columnare* isolates from genomovars I and II

Mark L. Lawrence, Attila Karsi, Hasan C. Tekedar, Michelle M. Banes & Derek Gibbs

College of Veterinary Medicine, Mississippi State University, Mississippi State, Mississippi USA

Correspondence: lawrence@cvm.msstate.edu

Abstract

Flavobacterium columnare is considered ubiquitous in the warmwater environment, and it is the etiological agent of columnaris disease, which affects cultured, wild, and ornamental fish [1–3]. It is a yellow-pigmented, motile, Gram-negative rod in the family *Flavobacteriaceae*, one of the families within the *Bacteroidetes* group [4]. *F. columnare* strains are divided into three genomovars that demonstrate varying virulence for different fish species [5–9]. *F. columnare* strain ATCC 49512 (CIP 103533 [TG 44/87]), isolated in 1987 from a skin lesion of a brown trout fry in France [10], belongs to genomovar I [11] and is avirulent in channel catfish [7]. Recently we reported the complete genome sequence of strain ATCC 49512 (CP003222.1) [12]. The final genome is 3,162,432 bp with 31.5% GC. It has 2856 predicted ORFs, 74 tRNAs, and 15 rRNAs. The average gene length is 991 bp, and it is 86.6% coding.

For comparative purposes, we have obtained a draft genome sequence of *F. columnare* strain 94-081, which was isolated in 1994 from a diseased channel catfish (*Ictalurus punctatus*) from a commercial production pond. Strain 94-081 is in genomovar II and is highly virulent in channel catfish using an experimental bath immersion. The draft genome has 3,148,440 bp with 2751 predicted ORFs and 61 RNAs. Metabolic comparison of ATCC 49512 and 94-081 revealed that ATCC 49512 has 20 unique protein functions relative to 94-081. Strain 94-081 has 38 unique protein functions, including seven unique cysteine biosynthesis proteins, seven unique branched chain amino acid biosynthesis proteins, and two proteins in heme/siroheme biosynthesis.

Comparison of ATCC 49512 with *Flavobacterium psychrophilum* JIP02/86 (NC_009613.1) revealed several metabolic differences. ATCC 49512 has 138 unique protein functions compared to *F. psychrophilum*. For example, *F. columnare* is capable of denitrification; ATCC 49512 annotation indicated that it is capable of nitrous oxide reduction, nitric oxide reduction, nitrite reduction, nitrate reduction, and nitrite/nitrate transport. Interestingly, *F. columnare* has a complete urea cycle, but it does not encode a urease; *F. psychrophilum* does not encode a urea cycle. *F. columnare* has unique amino acid metabolism functions in threonine degradation and proline uptake, and it encodes a unique aminopeptidase, zinc metalloproteinase, and endopeptidase. *F. columnare* also has 6 unique oxidative stress response proteins, a heat shock protein, an osmotic stress protein, and a periplasmic stress protein. Finally, *F. columnare* has unique single carbon metabolism features, including cobalamin biosynthesis and nickel/cobalt transport. *F. psychrophilum* has 119 unique protein functions compared to *F. columnare* ATCC 49512. These include unique amino acid metabolism functions such as histamine biosynthesis. *F. psychrophilum* is capable of riboflavin biosynthesis, and it has five unique stress response proteins. It also has unique electron acceptors/donors for respiration.

Using the 94-081 RAST annotation [13], orthology analysis was conducted between 94-081, ATCC 49512, *F. psychrophilum* JIP02/86, and *F. johnsoniae* UW101 (NC_009441.1). InParanoid orthology analysis [14] identified 2260 orthologous groups between *F. columnare* genomovars I and II (ATCC 49512 and 94-081). By contrast, 1723 orthologous groups were detected between *F. columnare* 94-081 and *F. psychrophilum*, and 1844 orthologous groups were detected between *F. columnare* 94-081 and *F. johnsoniae*.

Orthology analysis revealed several regions that are unique to the genomovar II 94-081 *F. columnare* genome relative to genomovar II (ATCC 49512), *F. psychrophilum*, and *F. johnsoniae*. One of these regions contains 19 hypothetical proteins and two bacteriophage proteins, and another contains 48 hypothetical proteins, one phage protein, and a peptidoglycan amidohydrolase. A third region unique to genomovar II has six hypothetical proteins, and a fourth unique region has twelve hypothetical proteins and VirG protein (function unknown). In summary, the genomic regions unique to genomovar II *F. columnare* largely contain proteins of unknown function and have evidence of origins from horizontal gene transfer. Orthology analysis also revealed regions of high conservation between *F. columnare* genomovars I and II that are not present in *F. psychrophilum* and *F. johnsoniae*. Not surprisingly, these regions include one encoding urea cycle proteins, one encoding nitrous oxide reductase proteins, and one encoding cobalamin synthesis and Ni/Co transport proteins. These *F. columnare*-unique proteins were also revealed by metabolic function comparison.

Searches for genomic islands were conducted in *F. columnare*, *F. psychrophilum*, and *F. johnsoniae* using IslandViewer [15]. Five genomic islands were identified in *F. columnare*; comparison of the identified genomic islands from *F. columnare* and *F. johnsoniae* by Mauve showed that one *F. columnare* island has 5% identity with a *F. johnsoniae* island. No genomic islands were identified in *F. psychrophilum*. Therefore, *F. columnare* islands are unique to this species. *F. columnare* genomic island 1 encodes four phage proteins, nine hypothetical proteins, and asparaginase; genomic island 2 encodes seven hypothetical proteins and one membrane protein; island 3 encodes an integrase, three transposons, a transcriptional activator, and 4 hypothetical proteins; island 4 encodes an integrase, tyrosine recombinase, a DNA binding protein, and three hypothetical proteins; and island 5 encodes an integrase, a DNA binding protein, a DNA repair protein, a helicase, an ATPase, and nine hypothetical proteins. Four of the five *F. columnare* genomic islands encode a protein suggesting their origins from horizontal gene transfer, and many of the encoded proteins have poorly characterized function.

Our laboratory has been developing a real-time PCR assay for quantification of *F. columnare* in fish tissues and in the environment. The primers and probe for this assay target a 203 bp region of the chondroitin AC lyase gene (GenBank AY912281) [16]. Forward primer and probe sequences were obtained from Panangala et al. [17], and the reverse primer was designed using Primer 3 (http://frodo.wi.mit.edu/primer3/). Specificity of the assay was tested against *F. johnsoniae*, *Pseudomonas aeruginosa*, *Aeromonas salmonicida*, *Edwardsiella tarda*, and *E. ictaluri*. The assay was shown to have linear correlation with *F. columnare* genomic DNA quantity over six orders of magnitude ranging from 50 to 0.0005 ng of genomic DNA. There were no significant differences between ATCC 49512 and 94-081, suggesting limited copy number and sequence variation of the target gene between the two different genomovars. The assay was found to be highly repeatable and reproducible, with a coefficient of variation between runs of $\leq 3.0\%$, indicating an acceptable level of precision for both isolates. The assay was tested in three independent runs against eight other isolates:

94-060, Matt, 94-078, and C03-133K (genomovar II); and 155-94, 90-509, C91-20, and 143-94 (genomovar I). Strain 155-94 was chosen as standard.

In summary, orthology (sequence-based) analysis and metabolic (function-based) analysis showed that genomovar I and II *F. columnare* isolates are more closely related to each other than *F. psychrophilum* and *F. johnsoniae*. Furthermore, many of the proteins unique to genomovar II have unknown function, and many show evidence of horizontal gene transfer. Finally, a real-time PCR assay was developed that can be used to consistently quantify bacterial numbers and DNA quantity from genomovar I and II *F. columnare* isolates; it is ready for fish tissue and water analyses.

References

[1] Davis H.S., 1922. A new bacterial disease of fresh water fishes. U.S. Bureau of Fisheries Bulletin, 38:261–280

[2] Ordal E.J., Rucker R.R., 1944. Pathogenic myxobacteria. Proceedings of The Society for Experimental Biology and Medicine, 56:15–18

[3] Austin B., Austin D.A., 1999. Bacterial Fish Pathogens: Disease of Farmed and Wild Fish, Edinburgh, UK: Heriot-Watt University

[4] Bernardet J.-F., Segers P., Vancanneyt M., Berthe F., Kersters K., Vandamme P., 1996. Cutting a Gordian knot: embedded classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. Nov. (basonym, *Cytophaga aquatilis* Strohl and Tait 1978) International Journal of Systematic Bacteriology, 46:128–148

[5] Arias C.R., Welker T.L., Shoemaker C.A., Abernathy J.W., Klesius P.H., 2004. Genetic fingerprinting of *Flavobacterium columnare* isolates from cultured fish. Jornal of Applied Microbiology, 97:421–428

[6] Shoemaker C.A., Olivares-Fuster O., Arias C.R., Klesius P.H., 2008. *Flavobacterium columnare* genomovar influences mortality in channel catfish (*Ictalurus punctatus*) Veterinary Microbiology, 127:353–359

[7] Soto E., Mauel M.J., Karsi A., Lawrence M.L., 2008. Genetic and virulence characterization of *Flavobacterium columnare* from channel catfish (*Ictalurus punctatus*). Journal of Applied Microbiology, 104:1302–1310

[8] Darwish A.M., Ismaiel A.A., 2005. Genetic diversity of *Flavobacterium columnare* examined by restriction fragment length polymorphism and sequencing of the 16S ribosomal RNA gene and the 16S-23S rDNA spacer. Molecular and Cellular Probes, 19:267–274

[9] Triyanto A., Wakabayashi H., 1999. Genotypic diversity of strains of *Flavobacterium columnare* from diseased fish. Fish Pathology, 34:65–71

[10] Bernardet J.-F., 1989. *Flexibacter columnaris*: first description in France and comparison with bacterial strains form other origins. Diseases of Aquatic Organisms, 6:37–44

[11] Michel C., Messiaen S., Bernardet J.-F., 2002. Muscle infections in imported neon tetra, *Paracheirodon innesi* Meyers: limited occurrence of microsporidia and predominance of severe forms of columnaris disease caused by an Asian genomovar of *Flavobacterium columnare.* Journal of Fish Diseases, 25:253–263

[12] Tekedar H.C, Karsi A., Gillaspy A.F., Dyer D.W., Benton N.R., Zaitshik J., Vamenta S., Banes M.M., Gülsoy N., Aboko-Cole M., Waldbieser G.C., Lawrence M.L., 2012. Genome sequence of the fish pathogen *Flavobacterium columnare* ATCC 49512. Journal of Bacteriology, 194:2763–2764

[13] Aziz R.K., Bartels D., Best A.A., DeJongh M., Disz T., Edwards R.A., Formsma K., Gerdes S., Glass E.M., Kubal M., Meyer F., Olsen G.J., Olson R., Osterman A.L., Overbeek R.A., McNeil L.K., Paarmann D., Paczian T., Parrello B., Pusch G.D., Reich C., Stevens R., Vassieva O., Vonstein V., Wilke A., Zagnitko O., 2008. The RAST Server: Rapid Annotations using Subsystems Technology. BMC Genomics, 9:75

[14] Ostlund G., Schmitt T., Forslund K., Köstler T., Messina D.N., Roopra S., Frings O., Sonnhammer E.L., 2010. InParanoid 7:new algorithms and tools for eukaryotic orthology analysis. Nucleic Acids Research, 38(Database issue):196–203

[15] Langille M.G., Brinkman F.S., 2009. IslandViewer: an integrated interface for computational identification and visualization of genomic islands. Bioinformatics, 25:664–665

[16] Xie H.X., Nie P., Chang M.X., Liu Y., Yao W.J., 2005. Gene cloning and functional analysis of glycosaminoglycan-degrading enzyme chondroitin AC lyase from *Flavobacterium columnare* G4. Archives of Microbiology, 184:49–55

[17] Panangala V.S., Shoemaker C.A., Van Santen V.L., Dybvig K., Klesius P.H., 2007. Multiplex-PCR for simultaneous detection of 3 bacterial fish pathogens, *Flavobacterium columnare, Edwardsiella ictaluri,* and *Aeromonas hydrophila.* Diseases of Aquatic Organisms, 74:199–208