Complete Genome Sequence of *Flavobacterium indicum* GPSTA100-9T, Isolated from Warm Spring Water

Paul Barbier,a Armel Houel,a Valentin Loux,b Julie Poulin,a Jean-François Bernardet,a Marie Touchon,c,d and Eric Duchauda

Unité de Virologie et Immunologie Moléculaires, INRA, Domaine de Vilvert, Jouy-en-Josas, France; bUnité de Mathématique, Informatique et Génome, INRA, Domaine de Vilvert, Jouy-en-Josas, France; cInstitut Pasteur, Microbial Evolutionary Genomics, Département Génomiques et Génétique, Paris, France; dCNRS, Paris, France; and Genoscope, Evry, France

We report here the complete annotated genome sequence of *Flavobacterium indicum* CIP 109464T ( = GPTSA100-9T), isolated from warm spring water in Assam, India. The genome sequence of *F. indicum* revealed a number of interesting features and genes in relation to its environmental lifestyle.

Members of the family *Flavobacteriaceae* occur in a variety of temperate and polar habitats in terrestrial, freshwater, and marine environments. Besides this interesting diversity of lifestyle, they have a very significant role in the degradation/turnover of the organic matter in these ecosystems (1, 2).

We determined the whole genome sequence of the type strain of *Flavobacterium indicum* (CIP 109464T) (11) in order to perform comparative genomic studies between this environmental species and other *Flavobacterium* species whose genome sequences have been published: two fish-pathogenic species that severely impact aquaculture worldwide (1) (*F. psychrophilum* JIP02/86 [4] and *F. branchiophilum* FL15 [13]) and another environmental species, *F. johnsoniae* UW101T (7), a model organism for characterizing gliding motility (8) and biopolymer utilization in oligotrophic environments (10).

The genome of *F. indicum* was sequenced using a combination of Sanger (ABI3730, Applied Biosystems; performed on a genomic DNA library with an average fragment length of 10 kb cloned in pCNS) and 454 (GS-FLX, Roche) sequencing with 2.6-fold and 17.4-fold coverage, respectively. The 454 reads were assembled in 145 contigs using Newbler. These contigs and the Sanger reads were assembled in four scaffolds (39 contigs) using Phrap. Scaffolds were ordered using an optical map (OpGen Technologies) (6), and gaps were closed using primer walking on gap-spanning clones or by PCR sequencing. Genome annotation, including manual validation, was performed using the AGMIAL annotation platform (3).

The complete genome of *F. indicum* consists of a circular chromosome of 2,993,089 bp with an overall G+C content of 31.8%. The genome is predicted to carry 2,671 protein-coding genes, 55 tRNA genes, and four rRNA operons.

Genome comparison with the available genome sequences of other *Flavobacterium* species confirms a loss of synteny within the genus (*F. johnsoniae* (7), *F. psychrophilum* (13), *F. branchiophilum* (4), *F. indicum* (11)).

In relation to its environmental lifestyle, the *F. indicum* genome is predicted to encode 38 adhesins, likely used for binding on different surfaces, six glycoside hydrolase precursors, various endo- and exopeptidases, and one polysaccharide utilization system (PUL) (9), confirming its ability to degrade some macromolecules such as gelatin, casein, and starch (11). However, the *F. indicum* genome, which is 2-fold smaller than the 6-Mbp-long genome of *F. johnsoniae* (7), lacks the 1.93-Mbp region enriched for genes involved in polysaccharide utilization of the latter genome. This finding corroborates the weak biopolymer-degrading ability of *F. indicum* (1) likely related to a more restricted ecological niche. The gliding motility machinery (12) is probably not functional, as the gldA gene, encoding the gliding motor, is frame-shifted and the gldE gene, involved in gliding, is absent, verifying the original description of *F. indicum* as a nongliding organism (11). In contrast with *F. psychrophilum* and *F. johnsoniae*, which contain flexirubin-type pigments, the yellowish-orange color of *F. indicum* (11) is attributable only to the presence of carotenoid biosynthesis genes.

**Nucleotide sequence accession number.** The annotated complete genome sequence of *F. indicum* CIP 109464T reported in this paper is available in GenBank under the accession number HE774682.

**ACKNOWLEDGMENTS**

This work was supported in part by grant 07-GMGE from the Agence Nationale de la Recherche of France. P.B. is a Université Evry Val d’Essonne Ph.D. fellow.

We are grateful to the INRA MIGALE bioinformatics platform (http://migale.jouy.inra.fr) for providing computational resources. We acknowledge C. Bizet (Collection de l’Institut Pasteur, Paris) for generously supplying the type strain of *F. indicum*.

**REFERENCES**


Received 22 March 2012 Accepted 27 March 2012

Address correspondence to Eric Duchaud, eric.duchaud@jouy.inra.fr.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.
doi:10.1128/JB.00420-12