Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum*

Eric Duchaud¹, Mekki Boussaha¹, Valentin Loux², Jean-François Bernardet¹, Christian Michel¹, Brigitte Kerouault¹, Stanislas Mondot¹, Pierre Nicolas², Robert Bossy², Christophe Caron², Philippe Bessires², Jean-Franois Gibrat², Stphane Claverol³, Fabien Dumetz⁴, Michel Le Hnaff⁴ & Abdenour Benmansour¹

We report here the complete genome sequence of the virulent strain JIP02/86 (ATCC 49511) of *Flavobacterium psychrophilum*, a widely distributed pathogen of wild and cultured salmonid fish. The genome consists of a 2,861,988-base pair (bp) circular chromosome with 2,432 predicted protein-coding genes. Among these predicted proteins, stress response mediators, gliding motility proteins, adhesins and many putative secreted proteases are probably involved in colonization, invasion and destruction of the host tissues. The genome sequence provides the basis for explaining the relationships of the pathogen to the host and opens new perspectives for the development of more efficient disease control strategies. It also allows for a better understanding of the physiology and evolution of a significant representative of the family Flavobacteriaceae, whose members are associated with an interesting diversity of lifestyles and habitats.

Flavobacterium psychrophilum, a member of the family Flavobacteriaceae within the phylum Bacteroidetes¹, is currently one of the most devastating fish pathogens. Whereas most members of the family are free-living organisms, others are serious pathogens of humans and various animals¹. *F. psychrophilum* is a gram-negative bacterium affecting various species of salmon and trout reared in freshwater². The epithet *psychrophilum* refers to the water temperature at which outbreaks of the disease occur (3–15 °C). The optimum generation time of 2 h is achieved at 15 °C (ref. 3). Because some strains grow weakly up to 25 °C, *F. psychrophilum* should be considered a psychrotrophic organism.

F. psychrophilum has occasionally been detected in samples of water, sediments and biofilms from rivers receiving outlet water from infected fish farms. *In vitro* studies have demonstrated that cells are able to survive for several months in stream water despite a rapid growth arrest⁴. The presence of *F. psychrophilum* within salmonid eggs and the resulting vertical transmission from broodfish to progeny suggest that this obligate fish pathogen has spread through the international trade of live fish and fish eggs⁵. Whereas in adult fish extensive necrotic lesions result in a clinical form called “cold-water disease,” young fish suffer severe mortality associated with hemorrhagic septicemia, a condition referred to as “rainbow trout fry syndrome.” Infections by *F. psychrophilum* are responsible for considerable economic losses in all major areas of salmonid aquaculture worldwide. Because a specific vaccine is currently unavailable, antibiotics represent the only recourse; however, because their use has been criticized⁶, satisfactory control methods are lacking. Moreover,

the molecular pathogenesis of *F. psychrophilum* infections is not yet understood, and only scant data on virulence factors are available. To obtain insight into the disease mechanisms of *F. psychrophilum*, we determined and analyzed the complete genome sequence of the virulent strain JIP02/86 (ref. 7).

RESULTS

General genome features

The genome of *F. psychrophilum* JIP02/86 consists of a circular chromosome of 2,861,988 bp (Fig. 1 and Table 1) and the pCP1 cryptic plasmid⁸. The origin of replication has been identified (Supplementary Fig. 1 online). The relatively high number of rRNA and tRNA genes, six and 49 respectively, is in good agreement with the rather rapid growth of the bacterium^{9,10}. The chromosome is predicted to contain 2,432 protein-coding genes. The rather small genome size of *F. psychrophilum* as compared with those of environmental members of the family is probably related to its restricted ecological niche.

Toxins

Since early studies *F. psychrophilum* was shown to be actively proteolytic¹¹, and proteases were suspected to be among its most important virulence factors¹², leading to rapid and massive tissue destruction (Fig. 2). Accordingly, the *F. psychrophilum* genome encodes 13 putative secreted proteases probably involved in virulence and/or in the destruction of host tissues (Table 2). The two metalloproteases Fpp1 (ref. 13) and Fpp2 (ref. 14) hydrolyze a broad range of

¹INRA, Unit Virologie et Immunologie Molculaires UR892, F-78350 Jouy-en-Josas, France. ²INRA, Unit Mathmatique, Informatique et Gnome UR1077, F-78350 Jouy-en-Josas, France. ³Universit Victor Segalen Bordeaux 2, Plateforme Gnomique Fonctionnelle Bordeaux, 146 Rue Lo Saignat, F-33076 Bordeaux Cedex, France. ⁴Laboratoire de Microbiologie et de Biochimie Appliques, Ecole Nationale d'Ingnieurs des Travaux Agricoles de Bordeaux, 1 Cours du Gnral de Gaulle, CS 40201, 33175 Gradignan Cedex, France. Correspondence should be addressed to E.D. (eric.duchaud@jouy.inra.fr).

Received 12 September 2006; accepted 29 May 2007; published online 24 June 2007; doi:10.1038/nbt1313

Figure 1 Circular representation of the *Flavobacterium psychrophilum* genome. The outer scale is in kilobases. Circles 1 and 2 (from outside to inside), genes transcribed clockwise (in red) and counterclockwise (in blue). Circle 3, rRNA genes (in green) and tRNA genes (in purple). Circle 4, insertion sequence elements (in pink). Circle 5, GC skew (window size 5,000 bp).

substrates, including basic elements of muscular tissues. Sequences of the first N-terminal amino acids of purified Fpp1 and Fpp2 (SSTGQLKTMRLAQS(C/W)NGQYA and TVYNIPV, respectively¹⁴) allowed the unambiguous identification of the corresponding genes. Pathogenic fungi secrete endopeptidases belonging to the fungalsin family of metalloproteases that cleave extracellular matrix proteins and are involved in the breakdown of proteinaceous structural barriers during host tissue colonization¹⁵. A similar mechanism may apply for *F. psychrophilum*, in that we identified two paralogous genes *FP0280* and *FP0281*, tandemly arranged, predicted to encode metalloproteases of the fungalsin family.

Because many bacterial pathogens produce extracellular collagenolytic enzymes allowing the colonization of connective tissues of their vertebrate hosts, the secretion of extracellular collagenase was considered of importance to *F. psychrophilum* pathogenicity¹⁶. Surprisingly, the JIP02/86 genome contains a collagenase-encoding gene that is disrupted by an insertion sequence of the IS256 family. To further examine this unexpected finding, the integrity of the collagenase-encoding gene was assessed by polymerase chain reaction (PCR) and sequencing in 23 additional *F. psychrophilum* strains from different geographic areas and fish species. These experiments revealed that 10 of the 24 strains had an insertion sequence 256 (IS256)-disrupted gene (Supplementary Fig. 2 online). Strains containing the disrupted gene were all isolated from rainbow trout, indicating that the collagenase gene is not essential to virulence in this species and suggesting a clonal dissemination of strains with an IS256-disrupted collagenase-encoding gene in rainbow trout farms worldwide.

Bacterial hemolysins are cytolytic toxins that are important virulence factors. In another fish pathogen, *Vibrio anguillarum*, toxin VAH5 is able to lyse rainbow trout erythrocytes and a VAH5 mutant showed attenuated virulence¹⁷. In *F. psychrophilum*, *FP0063* encodes a protein 53% similar to VAH5 that may have a similar role in pathogenicity and cooperate with secreted proteases for tissue destruction.

Table 1 General features of the *Flavobacterium psychrophilum* JIP02/86 genome

Genome size	2,861,988 bp
G+C content	32.54%
Coding sequences	2,432
Pseudogenes	20
Coding density	84.5%
Average gene length	1,003 bp
Ribosomal RNAs (16S-23S-5S)	Six copies
Transfer RNAs	49 copies
Number of CDS with a predicted function	1,318
Number of CDS similar to proteins of unknown function	701
Number of CDS without similarities	401
Number of insertion sequences	74

CDS, coding sequences.

The thiol-activated cytolytic family of pore-forming toxins (TACs) has been implicated in the pathogenicity of several gram-positive bacteria¹⁸. The cytolytic activity is conferred by the C-terminal region (domain 4) of the protein. TACs are also able to trigger several signaling pathways in a variety of host cell types through domains 1–3 of the protein, as shown for listeriolysin O (LLO) and pneumolysin¹⁹. Signals triggered by LLO strongly influence the course of infection, and LLO lacking domain 4 is fully capable of inducing a cytokine response, acting as a pleiotropic pseudocytokine/chemokine. In *F. psychrophilum*, the *FP0097* gene encodes a protein that is similar to domains 1–3 of TACs. Hence, *FP0097* may function as a bacterial modulin responsible for damage to host tissue by virtue of its cytokine-inducing ability.

Motility, adhesion and biofilm formation

Cells of *F. psychrophilum* move over surfaces by gliding motility, an active movement that does not involve pili or flagella. Indeed, no genes encoding flagella, pili or any motility organelles were found in the *F. psychrophilum* genome, but 13 of the 15 *gld* genes involved in the gliding motility of *F. johnsoniae*²⁰ were identified. The missing *gldE* and *gldO* genes are the only two *F. johnsoniae* genes whose involvement in gliding motility was not unequivocally demonstrated, suggesting that *GldE* and *GldO* have an accessory function.

Adhesion mechanisms are greatly diversified among bacteria and are of particular importance in pathogenicity. Strong adhesive

Figure 2 Deep ulcerative lesion of the caudal peduncle provoked by *F. psychrophilum* in a rainbow trout (*Oncorhynchus mykiss*). Skin and muscle tissues have been destroyed, exposing the spinal cord. In the course of 'cold-water disease' of adult salmonid fish, such necrotic lesions typically occur around the adipose ('peduncle disease') or dorsal ('saddleback disease') fins, on the gills and on the flank. Scale is in centimeters.

Table 2 Putative secreted protease precursors of *Flavobacterium psychrophilum* JIP02/86 involved in the virulence and/or in the destruction of host tissues

Locus tag	Gene name	Family	Product	Most likely signal peptidase cleavage site
FP0081		M50	Probable M50 family membrane-associated zinc metalloprotease precursor	29–30
FP0082		M1	Probable M1 family metalloprotease precursor	21–22
FP0086	pepO	M13	M13 family metalloprotease precursor	23–24
FP0231	fpp1	Unassigned	Psychrophilic metalloprotease Fpp1 precursor	20–21
FP0232	fpp2	M43	Psychrophilic metalloprotease Fpp2 precursor	24–25
FP280		M36	Probable M36 fungalsin family metalloprotease precursor	20–21
FP0281		M36	Probable M36 fungalsin family metalloprotease precursor	20–21
FP1024		M43	Probable M43 cytophagyalysin family metalloprotease precursor	21–22
FP1619		M43	Probable M43 cytophagyalysin family metalloprotease precursor	18–19
FP1763		S8	Probable S8 subtilisin family serine endopeptidase precursor	16–17
FP1777 + FP1776		M43	Truncated collagenase precursor (insertion sequence disrupted)	22–23
FP2364		M48	Probable M48 family membrane-associated zinc metalloprotease precursor	22–23
FP2396		S8	Probable S8 subtilisin family serine endopeptidase precursor	23–24

properties of *F. psychrophilum* cells to fish body surface, gills and eggs have been reported²¹. We identified 27 genes probably involved in bacterial adhesion (**Supplementary Table 1** online). Fifteen of them are tandemly organized and encode proteins with leucine-rich repeats similar to BspA and LrrA of the periodontopathogenic bacteria *Bacteroides forsythus* and *Treponema denticola*, respectively. These immunogenic cell surface proteins bind strongly to extracellular matrix components and function in the attachment to human oral tissue^{22,23}. Other proteins probably involved in adhesion are: (i) FP2413, which contains five fibronectin type III domains involved in cell surface binding and a CUB domain found in extracellular proteins of spermadhesin and in eukaryotic peptidase families; and (ii) FP1830, FP0016 and FP0595, which contain fibronectin type III domains.

Biofilm formation is of importance in several pathogenic bacteria, especially those living in water, conferring a selective advantage by increasing their ability to persist under adverse environmental conditions. Little is known regarding the biofilm formation capabilities of *F. psychrophilum* cells. A *tlpB* mutant was deficient in virulence and gliding motility but displayed enhanced ability to form biofilm, suggesting that gliding motility and biofilm formation are antagonistic properties²⁴. Bacterial biofilms generally consist of cells entwined in a protective matrix of extracellular polysaccharides. The *F. psychrophilum* genome contains many determinants for the biosynthesis, export, modification and polymerization of exopolysaccharides, most of them located in a 70-kilobase (kb) region (1,425,425 bp to 1,496,190 bp). Four proteins similar to alginate O-acetyltransferases of *Pseudomonas aeruginosa* probably function in biofilm formation. In *P. aeruginosa*, acetylation of the extracellular polysaccharide alginate by alginate O-acetyltransferases is important for the architecture of biofilms and increases the ability of alginate to act as a defense barrier²⁵.

Secretion systems

Secretion systems are of great significance for virulence by addressing toxins to the bacterial surface. ABC-type transport systems, the Sec-dependent transport system, the components of the main terminal branch of the general secretory pathway and the Sec-independent (TAT) transport system were all identified in the *F. psychrophilum* genome. *F. psychrophilum* is devoid of the type III and IV secretion systems usually used by gram-negative pathogens but encodes the PorT and PorR proteins. In the periodontopathogenic bacterium

Porphyromonas gingivalis, the membrane-associated protein PorT is essential for the transport of major virulence factors, such as proteinases and adhesins, from the periplasm to the cell surface²⁶, and PorR is involved in the biosynthesis of cell surface polysaccharides that probably act as anchors for these factors²⁷. Hence, this mechanism may also apply to *F. psychrophilum* and other pathogenic members of the phylum.

Cell surface proteins

Cell surface proteins are important targets of the host immune system. We identified the host-recognized antigens P18 (ref. 28), P60 (ref. 29) and FspA (ref. 30), as well as two proteins similar to those that induce protective immunity against the poultry pathogen *Ornithobacterium rhinotracheale*³¹. We also identified 35 proteins containing a signal peptide and a conserved C-terminal region of ~70 amino acids (**Supplementary Table 2** online). Most of these proteins also belong to the aforementioned toxins and adhesins. Iterative searches of databases identified 170 proteins that contain this C-terminal conserved region exclusively in members of the phylum Bacteroidetes. Hence, members of this phylum contain the C-terminal motif [YNPX_{21–23}(N/D)X₂GX_{18–27}GXY] that probably addresses and/or anchors proteins to be exposed at the bacterial surface, some of which may be of importance for virulence.

Using purified fractionized cell extract, gel electrophoresis and mass spectrometry, we confirmed the location and validated the expression of several of these predicted outer membrane proteins (**Supplementary Table 3** online). Among them, several are probably important for bacteria-host interactions.

Metabolism

Genome analysis allowed the first metabolic reconstruction of this bacterium (**Fig. 3**). An important aspect of the lifestyle of pathogens is the acquisition of nutrients from their host. *F. psychrophilum* is unable to use carbohydrates as sources of carbon and energy⁷. Indeed, sugar kinase and phosphotransferase systems usually used by bacteria for specific carbohydrate uptake are missing, but genes encoding the enzymes of the glycolytic pathway (except glucose kinase) and the pentose phosphate pathway were identified. Therefore, these two pathways are most likely involved in the generation of precursor metabolites, such as the nucleotide-sugar precursors for exopolysaccharides biosynthesis.

Figure 3 Overview of metabolism and transport in *F. psychrophilum*. Transporters are grouped by substrate specificity: ions (green), nutrients (red), drug/metabolite efflux (black), secretion systems (purple), polysaccharide export (blue) and energy production (yellow). Arrows indicate the direction of transport.

The breakdown of host proteins by the secreted proteases results in a mixture of amino acids and oligopeptides that probably constitute the main source of carbon, nitrogen and energy. Five importers probably involved in the uptake of host proteins degradation products were identified. The *F. psychrophilum* genome is rich in peptidases that probably degrade imported peptides to amino acids that are processed by extensive amino acid catabolic pathways. Their importance is highlighted by the presence of two genes, *FP1110* and *FP1111*, that encode cyanophycinase and cyanophycin synthetase, respectively. Cyanophycin is a branched, non-ribosomally synthesized polypeptide originally described in cyanobacteria that occurs as insoluble inclusions in the cytoplasm and serves as a storage compound for carbon, nitrogen and energy³². One might predict that *F. psychrophilum* stores amino acids through this process during feast conditions that are mobilized for survival in nutrient-limited conditions.

Degradation of lipids to fatty acids is probably achieved by a putative phospholipase and by three enzymes of the esterase-lipase-thioesterase family. Beta-oxidation of fatty acids is performed by a long-chain fatty acid-CoA ligase, three fatty acid dehydrogenases, a crotonase and three thiolases.

Hence, most of the degradation products of host proteins and lipids by *F. psychrophilum* are available as tricarboxylic acid cycle precursors. Indeed, all enzymes of the tricarboxylic acid cycle are present in the *F. psychrophilum* genome, and at least two enzymes that interconvert C3 and C4 compounds were identified.

Fe acquisition, of utmost importance to bacterial pathogens, is probably performed by five TonB-dependent outer membrane Fe receptors. Moreover, *FP0089* and *FP0090* encode proteins highly similar to FeoA and FeoB that constitute a Fe²⁺ uptake system. In *Campylobacter jejuni*, *feoB* mutants were significantly affected in their ability to colonize the gastrointestinal tract during both commensal and infectious relationships. Similarly, *P. gingivalis* lost its virulence in a mouse model of infection when one of its two *feoB* paralogous genes was inactivated³³. Therefore, this iron acquisition system may also function in the pathogenicity of *F. psychrophilum*.

In agreement with its strictly aerobic metabolism, the *F. psychrophilum* genome encodes an extensive aerobic respiratory chain. All described members of the family Flavobacteriaceae contain menaquinone 6 as their only or major respiratory quinone², and the corresponding genes (*menABCDEF* and *ubiE*) have been identified. Among the 24 cytochrome oxidase family proteins identified, a probable cytochrome *cbb3*-type oxidase complex (*cbb3*) and the proteins necessary for its proper assembly were found. This complex with high O₂ affinity is usually expressed in response to microaerobic conditions and allows human pathogens to colonize anoxic tissues³⁴. Moreover, *FP1152*, located upstream of the *cbb3* gene cluster, encodes a Fnr family transcriptional regulator probably involved in the direct regulation of *cbb3* in response to O₂ as reported in other bacteria³⁵. In addition, a H-NOX family protein probably signaling shifts to low-O₂ environments³⁶ has been identified.

Psychrotrophic character and stress response

Because *F. psychrophilum* faces low temperature in its natural habitat, it has to successfully overcome the resulting physical constraints, including slower pace of biochemical reactions, increased viscosity and increased gas solubility³⁷. The *F. psychrophilum* genome encodes proteins probably involved in the regulation of membrane fluidity, in the maintenance of protein synthesis and in the production of psychrophilic enzymes and proteins with antioxidant properties.

Modification of membrane lipid composition with temperature changes is a widely distributed process for regulating membrane fluidity in bacteria³⁷. Incorporation of unsaturated fatty acids in lipid membranes is a consistent adaptation of many cold-experiencing prokaryotes³⁸. In coherence with the description of several unsaturated fatty acids in the membrane of *F. psychrophilum*³⁹, we identified *FP0413* encoding a protein similar to the $\Delta 6$ acyl-lipid desaturase of *Synechocystis* sp.⁴⁰ and two paralogous genes *FP0051* and *FP1155*, encoding proteins related to $\Delta 9$ stearoyl-[acyl-carrier-protein] desaturases of higher plants⁴¹. In addition, a cluster of genes (*criIBZY*) probably involved in the biosynthesis of carotenoids may also participate in the modulation of membrane fluidity in response to temperature shift as reported in other psychrotrophic bacteria⁴².

The importance of the control of RNA folding and degradation at low temperature is underlined by the presence of six ATP-dependent RNA helicases, three of them belonging to the DEAD/DEAH box RNA helicase family. In the Antarctic archaeon *Methanococcoides burtonii*, these enzymes are overexpressed at low temperature and are involved in the destabilization of secondary structures of nucleic acids⁴³. Differential regulation of gene expression may function in the bacterial adaptation to temperature shifts. In *F. psychrophilum*, the cold shock nucleic acid-binding protein *FP0276* and the chaperones *FP0864*, *FP1509* and *FP1985* might participate in these adaptive responses.

The most important selective pressure related to the exposure to low temperature is the exponential decrease in the rate of chemical reactions. Little is known about the psychrotrophic properties of *F. psychrophilum* enzymes: only the metalloproteases *Fpp1* and *Fpp2* were biochemically characterized and reported to display a psychrophilic and thermolabile behavior^{13,14}. The psychrotolerant *Flavobacterium frigidimarum* produces seven NAD(P)⁺-dependent dehydrogenases⁴⁴; four of them were psychrophilic and thermolabile, whereas the other three were unexpectedly thermophilic and thermostable. Only the genes orthologous to the *F. frigidimarum* genes encoding psychrophilic and thermolabile dehydrogenases are present in the *F. psychrophilum* genome. Hence, some of the enzymes probably involved in the virulence and in essential metabolic pathways of *F. psychrophilum* seem to be optimally active at low temperature.

At low temperature, the solubility of gases, including the noxious and highly reactive molecule dioxygen, increases rapidly. The large repertoire of proteins having antioxidant capacities reported in *Colwellia psychrerythraea* and *Desulfotalea psychrophila*³⁷ seemed essential for the survival in cold environments. Moreover, because one of the most efficient defense mechanisms against bacterial pathogens is the production of reactive oxygen species (ROS) by host macrophages⁴⁵, the resistance of *F. psychrophilum* to ROS killing is necessary to establish infection⁴⁶. Accordingly, a large panel of genes encoding enzymes probably counteracting the deleterious effects of ROS were identified (Supplementary Table 4 online): three superoxide dismutases, a bifunctional catalase-peroxidase, a thiol peroxidase and four proteins of the peroxy-redoxin family. In the strictly anaerobic bacterium *Bacteroides fragilis*, five genes (*batA*, *batB*, *batC*, *batD*, *batE*) are involved in the tolerance to O₂ (ref. 47). Surprisingly,

orthologous genes were identified in the genome of the strict aerobe *F. psychrophilum*, suggesting their involvement in the protection of bacterial cells against ROS. Among the OxyR regulons present in *Escherichia coli*, the *suf* operon is specially devoted to the synthesis of [Fe-S] clusters when Fe or S metabolism is disrupted by Fe starvation or oxidative stress⁴⁸. Moreover, this system is necessary for the virulence of the plant pathogen *Erwinia chrysanthemi*⁴⁹. *F. psychrophilum* probably uses similar strategies to counteract the deleterious effects of oxidative stress and Fe limitation, in that proteins highly similar to SufABCDSE of *E. coli* have been identified. Therefore, sophisticated stress response mechanisms and the associated regulatory repertoire reflect the ability of *F. psychrophilum* to cope with various stressful conditions, in particular the oxidative stress resulting from the synergy of ROS production by hosts and their enhanced diffusion in cold environment.

DISCUSSION

The complete genome sequence of a fish pathogen has not been published previously. Sequence analysis of *F. psychrophilum* has revealed a combination of strategies that probably confer upon it the ability to colonize and degrade fish tissues and to exploit proteinaceous compounds for growth. Versatile respiratory and stress response systems may allow *F. psychrophilum* to deal with microaerobic conditions and oxidative stress. Moreover, the ability to form biofilms and to store cyanophycin could explain the long survival of *F. psychrophilum* outside the host. These features, together with its strong adhesive properties, probably contribute to its dissemination in salmonid farms worldwide. The availability of different genetic tools⁸ and transposon mutant libraries²⁴ will allow large-scale functional analysis to assess the precise role of the predicted proteins identified. Thus, the *F. psychrophilum* genome sequence should facilitate the understanding of flavobacterial virulence mechanisms in fish and provide a basis for the development of better control strategies. Ongoing genome sequencing of other flavobacteria will provide new insights into pathogenicity and niche adaptation through comparative genomics.

METHODS

Genome sequencing strategy. The genomic DNA of *F. psychrophilum* JIP02/86 was sequenced by Qiagen GmbH using a conventional whole-shotgun strategy. Briefly, mechanically sheared 1.5-kb, 3-kb and 5-kb DNA fragments were isolated, inserted into pUC19 and cloned. Double-ended plasmid sequencing reactions were carried out using PE BigDye Terminator chemistry, and sequencing ladders were resolved on a PE 3700 automated DNA sequencer. The 33,900 useful readings were assembled into 56 contigs, providing a 9.3 X genome coverage. Gaps were closed by primer walking on gap-spanning clones, and remaining physical gaps were closed by multiplex PCR. Assessment of final assembly quality was done by long-range PCR and by ensuring that the physical map deduced from the genome sequence was identical to that obtained experimentally by pulsed-field gel electrophoresis (data not shown).

ORF prediction and annotation. The prediction of coding sequences was generated using the self-training gene detection software SHOW based on Hidden Markov Models (<http://genome.jouy.inra.fr/ssb/SHOW/>). The ribosome-binding sites and transcriptional terminators were detected using the SHOW and Petrin softwares, respectively, whereas tRNA and rRNA were detected using the tRNA-scan and rRNA-scan softwares, respectively. Genome annotation was performed using the AGMIAL annotation platform⁵⁰. Because of a lack of a reference set of annotation for genomes in the family Flavobacteriaceae, web-based softwares and databases were used to manually curate all predicted genes. ISs were classified using the IS finder database (<http://www-is.biotoul.fr/>), ABC transporters using the ABCISSE database (<http://www.pasteur.fr/recherche/unites/pmtg/abc/>) and proteases using the

MEROPS database (<http://merops.sanger.ac.uk/>). Genome comparisons were performed using the web interface MaGe, which allows graphic visualization enhanced by a synchronized representation of synteny groups (<http://www.genoscope.cns.fr/agg/mage/>). Orthologous genes between two genomes were defined as gene couples satisfying the bidirectional best-hit criterion or a BLASTP alignment threshold—that is, a minimum of 35% sequence identity on 80% of the length of the smallest protein. These orthologous genes were subsequently used to search for conserved gene clusters, for example, synteny groups among several bacterial genomes. All possible kinds of chromosomal rearrangements were allowed (inversion, insertion/deletion). A gap parameter, representing the maximum number of consecutive genes that were not involved in a synteny group, was set to five genes.

Nucleotide sequence accession number. The genomic sequence reported in this article has been deposited in the EMBL database under the accession number AM398681. The annotated genome and the relevant AGMIAL database (genome navigator, Blast server and metabolic comparisons) are available at (<http://migale.jouy.inra.fr/psychrophilum>).

Cell surface protein identification. Isolation of the outer membrane fraction, sodium dodecyl sulfate–polyacrylamide gel electrophoresis, protein elution and protein identification by liquid chromatography–tandem mass spectrometry were performed following published procedures²⁸ detailed in **Supplementary Table 3**.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

This work received financial support from the Institut National de la Recherche Agronomique. We wish to thank Arnim Wiezer and Holger Wedler for genome sequencing; Claudine Médigue and Stéphane Cruveiller for their contribution to genome comparison; Stéphane Chaillou and Vincent Fromion for their contribution with the reconstruction of metabolic pathways; Pierre Boudinot, Carmen Buchrieser and Eduardo Rocha for helpful discussions and for critical reading of the manuscript.

AUTHOR CONTRIBUTIONS

E.D. wrote the manuscript, analyzed the data and coordinated annotation strategy and the analysis of genome information; M.B., C.M. and S.M. contributed to data analysis; J.-F.B. co-wrote the paper and determined the genome size together with B.K.; V.L. contributed to the maintenance and improvement of the AGMIAL web interface; P.N., R.B., C.C., P.B. and J.-F.G. contributed to annotation and data analysis; S.C., F.D. and M.L.H. performed the proteomic analysis; A.B. initiated the project and contributed to the writing of the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturebiotechnology/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

- Bernardet, J.-F. & Nakagawa, Y. An introduction to the family Flavobacteriaceae. in *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community*, edn. 3, rel. 3.15 (eds. Dworkin, M. et al.) (Springer-Verlag, New York, 2003).
- Bernardet, J.-F. & Bowman, J.P. The genus *Flavobacterium*. in *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community*, edn. 3, rel. 3.20 (eds. Dworkin, M. et al.) (Springer-Verlag, New York, 2006).
- Holt, R.A., Rohovec, J.S. & Fryer, J.L. Bacterial cold-water disease. in *Bacterial Diseases of Fish* (eds. Inglis, V., Roberts, R.J. & Bromage, N.R.) 3–22 (Blackwell Scientific Publications, Oxford, 1993).
- Madetoja, J., Nystedt, S. & Wiklund, T. Survival and virulence of *Flavobacterium psychrophilum* in water microcosms. *FEMS Microbiol. Ecol.* **43**, 217–223 (2003).
- Bovo, G. et al. Work Package 1 Report: Hazard Identification for Vertical Transfer of Fish Disease Agents. (European Commission Project VESO-1601, VESO, Oslo, 2005).
- Gustafson, R.H. & Bowen, R.E. Antibiotic use in animal agriculture. *J. Appl. Microbiol.* **83**, 531–541 (1997).
- Bernardet, J.-F. & Kerouault, B. Phenotypic and genomic studies of “*Cytophaga psychrophila*” isolated from diseased rainbow trout (*Oncorhynchus mykiss*) in France. *Appl. Environ. Microbiol.* **55**, 1796–1800 (1989).
- Alvarez, B., Secades, P., McBride, M.J. & Guijarro, J.A. Development of genetic techniques for the psychrotrophic fish pathogen *Flavobacterium psychrophilum*. *Appl. Environ. Microbiol.* **70**, 581–587 (2004).
- Klappenbach, J.A., Dunbar, J. & Schmidt, T. rRNA operon copy number reflects ecological strategies of bacteria. *Appl. Environ. Microbiol.* **66**, 1328–1333 (2000).
- Rocha, E.P. Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. *Genome Res.* **14**, 2279–2286 (2004).
- Pacha, R.E. & Porter, S. Characteristics of myxobacteria isolated from the surface of freshwater fish. *Appl. Microbiol.* **16**, 1901–1906 (1968).
- Bertolini, J.M., Wakabayashi, H., Watral, V.G., Whipple, M.J. & Rohovec, J.S. Electrophoretic detection of proteases from selected strains of *Flexibacter psychrophilus* and assessment of their variability. *J. Aquat. Anim. Health* **6**, 224–233 (1994).
- Secades, P., Alvarez, B. & Guijarro, J.A. Purification and characterization of a psychrophilic, calcium-induced, growth-phase-dependent metalloprotease from the fish pathogen *Flavobacterium psychrophilum*. *Appl. Environ. Microbiol.* **67**, 2436–2444 (2001).
- Secades, P., Alvarez, B. & Guijarro, J.A. Purification and properties of a new psychrophilic metalloprotease (Fpp2) in the fish pathogen *Flavobacterium psychrophilum*. *FEMS Microbiol. Lett.* **226**, 273–279 (2003).
- Brouta, F. et al. Purification and characterization of a 43.5 kDa keratinolytic metalloprotease from *Microsporium canis*. *Med. Mycol.* **39**, 269–277 (2001).
- Ostland, V.E., Byrne, P.J., Hoover, G. & Ferguson, H.W. Necrotic myositis of rainbow trout, *Oncorhynchus mykiss* (Walbaum): proteolytic characteristics of a crude extracellular preparation from *Flavobacterium psychrophilum*. *J. Fish Dis.* **23**, 329–336 (2000).
- Rodkhum, C., Hirono, I., Crosa, J.H. & Aoki, T. Four novel hemolysin genes of *Vibrio anguillarum* and their virulence to rainbow trout. *Microb. Pathog.* **39**, 109–119 (2005).
- Palmer, M. The family of thiol-activated, cholesterol-binding cytotoxins. *Toxicon* **39**, 1681–1689 (2001).
- Baba, H. et al. Induction of γ -interferon and nitric oxide by truncated pneumolysin that lacks pore-forming activity. *Infect. Immun.* **70**, 107–113 (2002).
- Nelson, S.S. & McBride, M.J. Mutations in *Flavobacterium johnsoniae* secDF result in defects in gliding motility and chitin utilization. *J. Bacteriol.* **188**, 348–351 (2006).
- Kondo, M., Kawai, K., Kurohara, K. & Oshima, S. Adherence of *Flavobacterium psychrophilum* on the body surface of the ayu *Plecoglossus altivelis*. *Microbes Infect.* **4**, 279–283 (2002).
- Sharma, A. et al. Cloning, expression, and sequencing of a cell surface antigen containing a leucine-rich repeat motif from *Bacteroides forsythus* ATCC 43037. *Infect. Immun.* **66**, 5703–5710 (1998).
- Ikegami, A., Honma, K., Sharma, A. & Kuramitsu, H.K. Multiple functions of the leucine-rich repeat protein LrrA of *Treponema denticola*. *Infect. Immun.* **72**, 4619–4627 (2004).
- Alvarez, B., Secades, P., Prieto, M., McBride, M.J. & Guijarro, J.A. A mutation in *Flavobacterium psychrophilum* *tlpB* inhibits gliding motility and induces biofilm formation. *Appl. Environ. Microbiol.* **72**, 4044–4053 (2006).
- Franklin, M.J. & Ohman, D.E. Identification of *algI* and *algJ* in the *Pseudomonas aeruginosa* alginate biosynthetic gene cluster which are required for alginate O acetylation. *J. Bacteriol.* **178**, 2186–2195 (1996).
- Sato, K. et al. Identification of a new membrane-associated protein that influences transport/maturation of gingipains and adhesins of *Porphyromonas gingivalis*. *J. Biol. Chem.* **280**, 8868–8877 (2005).
- Shoji, M. et al. Construction and characterization of a nonpigmented mutant of *Porphyromonas gingivalis*: cell surface polysaccharide as an anchorage for gingipains. *Microbiology-SGM* **148**, 1183–1191 (2002).
- Dumetz, F. et al. A protective immune response is generated in rainbow trout by an OmpH-like surface antigen (P18) of *Flavobacterium psychrophilum*. *Appl. Environ. Microbiol.* **72**, 4845–4852 (2006).
- Merle, C., Faure, D., Urdaci, M.-C. & Le Hénaff, M. Purification and characterization of a membrane glycoprotein from the fish pathogen *Flavobacterium psychrophilum*. *J. Appl. Microbiol.* **94**, 1120–1127 (2003).
- Crump, E.M., Burian, J., Allen, P.D. & Kay, W.W. Identification and expression of a host-recognized antigen, FspA, from *Flavobacterium psychrophilum*. *Microbiology-SGM* **151**, 3127–3135 (2005).
- Schuijffel, D.F., VanEmpel, P.C.M., Pennings, A.M.M.A., VanPutten, J.P.M. & Nuijten, P.J.M. Successful selection of cross-protective vaccine candidates for *Ornithobacterium rhinotracheale* infection. *Infect. Immun.* **73**, 6812–6821 (2005).
- Krehenbrink, M., Oppermann-Sanio, F.B. & Steinbüchel, A. Evaluation of non-cyanobacterial genome sequences for occurrence of genes encoding proteins homologous to cyanophycin synthetase and cloning of an active cyanophycin synthetase from *Acinetobacter* sp. strain DSM 587. *Arch. Microbiol.* **177**, 371–380 (2002).
- Dashper, S.G. et al. A novel *Porphyromonas gingivalis* FeoB plays a role in manganese accumulation. *J. Biol. Chem.* **280**, 28095–28102 (2005).
- Pitcher, R.S. & Watmough, N.J. The bacterial cytochrome cbb(3) oxidases. *Biochim. Biophys. Acta Bioenerg.* **1655**, 388–399 (2004).
- Cosseau, C. & Batut, J. Genomics of the ccoNOQP-encoded cbb(3) oxidase complex in bacteria. *Arch. Microbiol.* **181**, 89–96 (2004).
- Boon, E.M., Huang, S.H. & Marletta, M.A. A molecular basis for NO selectivity in soluble guanylate cyclase. *Nat. Chem. Biol.* **1**, 53–59 (2005).
- D'Amico, S., Collins, T., Marx, J.C., Feller, G. & Gerday, C. Psychrophilic microorganisms: challenges for life. *EMBO Rep.* **7**, 385–389 (2006).

38. Mansilla, M.C., Cybulski, L.E., Albanesi, D. & de Mendoza, D. Control of membrane lipid fluidity by molecular thermosensors. *J. Bacteriol.* **186**, 6681–6688 (2004).
39. Bernardet, J.F. *et al.* Cutting a Gordian knot: emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatidis* nom. nov. (Basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *Int. J. Syst. Bacteriol.* **46**, 128–148 (1996).
40. Reddy, A.S., Nuccio, M.L., Gross, L.M. & Thomas, T.L. Isolation of a delta 6-desaturase gene from the cyanobacterium *Synechocystis* sp. strain PCC 6803 by gain-of-function expression in *Anabaena* sp. strain PCC 7120. *Plant Mol. Biol.* **22**, 293–300 (1993).
41. Murata, N. & Wada, H. Acyl-lipid desaturases and their importance in the tolerance and acclimatization to cold of cyanobacteria. *Biochem. J.* **308**, 1–8 (1995).
42. Chintalapati, S., Kiran, M.D. & Shivaji, S. Role of membrane lipid fatty acids in cold adaptation. *Cell. Mol. Biol.* **50**, 631–642 (2004).
43. Lim, J., Thomas, T. & Cavicchioli, R. Low temperature regulated DEAD-box RNA helicase from the Antarctic archaeon, *Methanococcoides burtonii*. *J. Mol. Biol.* **297**, 553–567 (2000).
44. Oikawa, T. *et al.* Purification, characterization, and overexpression of psychrophilic and thermolabile malate dehydrogenase of a novel Antarctic psychrotolerant, *Flavobacterium frigidimarum* KUC-1. *Biosci. Biotechnol. Biochem.* **69**, 2146–2154 (2005).
45. Secombes, C.J. The nonspecific immune system: cellular defences. in *The Fish Immune System: Organism, Pathogen and Environment* (eds. Iwama, G. & Nakanishi, T.) 63–103, (Academic Press, San Diego, 1996).
46. Nematollahi, A., Pasmans, F., Haesebrouck, F. & Decostere, A. Early interactions of *Flavobacterium psychrophilum* with macrophages of rainbow trout *Oncorhynchus mykiss*. *Dis. Aquat. Org.* **64**, 23–28 (2005).
47. Tang, Y.P., Dallas, M.M. & Malamy, M.H. Characterization of the *BatI* (*Bacteroides aerotolerance*) operon in *Bacteroides fragilis*: isolation of a *B. fragilis* mutant with reduced aerotolerance and impaired growth in *in vivo* model systems. *Mol. Microbiol.* **32**, 139–149 (1999).
48. Outten, F.W., Djaman, O. & Storz, G. A *suf* operon requirement for Fe-S cluster assembly during iron starvation in *Escherichia coli*. *Mol. Microbiol.* **52**, 861–872 (2004).
49. Nachin, L., El Hassouni, M., Loiseau, L., Expert, D. & Barras, F. SoxR-dependent response to oxidative stress and virulence of *Erwinia chrysanthemi*: the key role of SufC, an orphan ABC ATPase. *Mol. Microbiol.* **39**, 960–972 (2001).
50. Bryson, K. *et al.* AGMIAL: implementing an annotation strategy for prokaryote genomes as a distributed system. *Nucleic Acids Res.* **34**, 3533–3545 (2006).