Transcriptome architecture and regulation at environmental transitions in flavobacteria: the case of an important fish pathogen

Cyprien Guérin[1], Bo-Hyung Lee[2], Benjamin Fradet[2], Erwin van Dijk[3], Bogdan Mirauta[4], Claude Thermes[3], Jean-François Bernardet[2], Francis Repoila[5], Eric Duchaud[2], Pierre Nicolas[1,*], Tatiana Rochat[2,*]

[1] Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France.

[2] Université Paris-Saclay, INRAE, UVSQ, VIM, 78350, Jouy-en-Josas, France

[3] Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.

[4] Sorbonne Université, CNRS, IBPS, Laboratoire de Biologie Computationnelle et Quantitative (LCQB), 75005, Paris, France

[5] Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, 78350, Jouy-en-Josas, France.

* To whom correspondence should be addressed. E-mail: tatiana.rochat@inrae.fr, pierre.nicolas@inrae.fr

Link to the *F. psychrophilum* expression browser for interactive exploration of the dataset:

https://fpeb.migale.inrae.fr

login: fpeb_review

pwd: Flavo_Review!

Competing interests statement:

This work was financially supported by the Agence Nationale de la Recherche (grant ANR-17-CE20-0020-01 FlavoPatho).

ABSTRACT

The family *Flavobacteriaceae* (phylum *Bacteroidetes*) is a major component of soil, marine and freshwater ecosystems. In this understudied family, *Flavobacterium psychrophilum* is a freshwater pathogen that infects salmonid fish worldwide, with critical environmental and economic impact. Here, we report an extensive transcriptome analysis conducted to unravel genes and pathways associated to the phenotypic traits and lifestyle of this bacterium. This established the genome map of transcription start sites and transcribed regions, and predicted alternative sigma factor regulons and regulatory RNAs. Transcriptome profiling across 32 biological conditions covered pathogen life in freshwater, exposed to fish components and stresses mimicking host environment. The results link genes to environmental conditions and provide insights into gene regulation, highlighting similarities with better known bacteria and original characteristics linked to the phylogenetic position and the ecological niche of this aquatic pathogen. In particular, osmolarity appears as a signal for transition between free-living and within-host programs and expression patterns of secreted proteins shed light on probable virulence factors. Further investigations showed that a newly discovered sRNA widely conserved in the genus, Rfp18, is required for precise expression of proteases. A website allows interactive exploration of the detailed results which open many directions for future studies on flavobacteria.

INTRODUCTION

Aquaculture is the fastest growing food-producing sector and now accounts for half of the human fish consumption in the world (1). A primary concern of this expending sector is the control of infectious diseases, such as those caused by bacteria of the family *Flavobacteriaceae* (phylum *Bacteroidetes*). Flavobacteria are most frequently isolated from environmental sources such as soil, sediments and water, and represent an important component of their ecosystems by recycling complex organic matter. Several species are aquatic pathogens (2), one of the most widely studied being *Flavobacterium psychrophilum*, a Gram-negative aerobic yellow-pigmented bacterium displaying gliding motility and growing between 4°C and 23°C (3).

F. psychrophilum primarily affects salmonids in freshwater environments. The conditions, known as rainbow trout fry syndrome and bacterial cold-water disease, are major sanitary issues for the fish farming industry worldwide. Control strategies rely exclusively on antibiotics and outbreaks have an impact on the environment and animal welfare (4, 5). Infected fish present signs of tissue erosion, skin ulcerations, necrotic lesions and splenic hypertrophy. The bacterium is mainly found in skin lesions, dermal ulcers extending deeply into muscular tissues, and in lymphoid organs (6). In rainbow trout fry, the disease often occurs as a septicemic form and mortality reaches 70% (4).

Genome analysis has generated relevant insights into *F. psychrophilum* epidemiology and evolution (7-11). A number of infection-relevant phenotypic traits have been reported. Bacterial cells are highly proteolytic (12, 13), cytotoxic for erythrocytes (14) and macrophages (15), attach to mucus (16), survive and potentially multiply inside phagocytes (17). The Type IX secretion system (T9SS) is required for pathogenicity in rainbow trout (18). Secreted proteases and iron acquisition systems are proposed to contribute to virulence (19-21). Outside the host, *F. psychrophilum* survives long periods in freshwater while maintaining its virulence (22, 23). Nevertheless, most genes associated with these phenotypic traits remain unknown and substantial efforts are needed to unravel the molecular factors involved in the pathogen life cycle.

F. psychrophilum must adapt to diverse environmental conditions. Outside the host, water provides a natural dissemination medium for an aquatic pathogen provided that it can withstand a low nutrient environment before invading a host. Transitions between these environments imply timely and

coordinated changes in gene expression. During the last decade, advances in omics technologies opened new routes to understand bacterial adaptation. Primary transcriptome mapping and condition-dependent transcriptome profiling proved to be particularly effective in providing genome-scale information allowing functional annotation of genomes, experimental discovery of regulatory elements such as promoters, transcription factors binding sites, *cis*- and *trans*-acting RNA elements (24-28).

Transcriptomic studies are still scarce in the family *Flavobacteriaceae* (29-33). As members of the phylum *Bacteroidetes*, their transcription machinery and translation process substantially differs from most other bacteria: an unusual primary sigma factor binds atypical promoter sequences (34-36) and initiation relies on sequence properties differing from the Shine-Dalgarno (29, 37). These unique expression signals result in limitations such as the inefficacy of genetics tools developed for other bacteria. By integrating expression data reflecting a large variety of living conditions one can observe how a bacterium reshapes its transcriptional program and understand some characteristics of transcriptional networks without the use of reverse genetics (38-40). Here, we apply this strategy to address missing molecular knowledge on *F. psychrophilum*.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions

This study used the wild-type strain *F. psychrophilum* OSU THCO2-90 isolated from the kidney of a Coho salmon in Oregon in 1990 which is a model strain for molecular genetics (10, 41). Two distinct sets of growth conditions were used for RNA-Seq and condition-dependent expression profiling (***SI Appendix* M1** and **Table S1)**. Strains, plasmids and oligonucleotides are listed in ***SI Appendix* T1** and **T2**. The transcriptional reporter plasmid pCP*Gm*^{r}-P$_{less}$-mCh and derivatives carrying *remF* promoter fragments were constructed using a pCP23-derivative vector carrying a gentamycin resistance marker as backbone (21) and the promoter activity was monitored using whole-cell fluorescence (***SI Appendix* M2**). The *rfp18* deletion mutant was constructed using a pYT313-derivative plasmid (42) as described in ***SI Appendix* M3**. Proteolytic activity was quantified on azocasein substrate as previously described (18).

RNA extraction, libraries preparation and RNA-sequencing

Total RNA extractions were performed using the hot phenol method as previously described (43). DNase-treated RNA extracts were used to prepare an equimolar 18-condition RNA pool that served for 5'-end and global RNA-Seq libraries preparation (***SI Appendix* M4**). The sequencing was performed on Illumina HiSeq platform (single-end, 50 bp). Reads were aligned to the complete genome sequence (10) as described in ***SI Appendix* M5**.

Determination of putative TSSs and classification of newly defined TRs

Identification of putative TSSs was based on the exact starting positions of uniquely mapped reads of the 5'-end RNA-Seq library (***SI Appendix* M6**). Genomic segments transcribed outside genome annotation and RFAM predictions (44) were delineated based on expression-level reconstructed from global RNA-Seq reads (45) as well as predicted intrinsic terminators and high-confidence TSSs (***SI Appendix* M7**). After expert curation, a total of 1511 TRs were classified into RNA categories according to their transcriptional context (***SI Appendix* D1)**.

Condition-dependent transcriptomics

Experiments were performed in duplicate based on independent cultures. A set of 64 RNA extracts were used. Design of SurePrint G3 Custom GE 8x60K microarrays (Agilent technologies), strand-specific cDNA synthesis, hybridization procedures and data processing are described in ***SI Appendix* M8**.

Computational analysis of promoter sequences and newly defined TRs

Promoter sequences and promoter activities across samples were analyzed together to identify sigma factor motifs using the TreeMM algorithm (40) with some modifications (***SI Appendix* M9**). Subsets of new TRs were analyzed for phylogenetic conservation, RNA secondary folding and mRNA-sRNA interactions (***SI Appendix* M10**).

Real Time qPCR gene expression analysis

qPCR was performed on CFX system following manufacturer's instruction (BIO-RAD) and expression level was normalized by geometric mean of 2 reference genes (***SI Appendix* M11)**.

Online data display

The website https://fpeb.migale.inrae.fr embeds Jbrowse (version 1.12.3) (46) and a SequenceServer (version 1.0.11) to allow blast searches (47). Browsing is possible along the chromosome, down to the level of read coverage and hybridization signal of individual probes, and across the expression space

based on correlation between genes. The interface allows online extraction of specific lists of features (new RNAs, TSSs, gene clusters, DEGs), export of figures, access to genomic coordinates and nucleotide sequence for all features.

RESULTS AND DISCUSSION

Combining experimental and *in silico* strategies to unravel transcriptome architecture

Experimental and computational methodologies were combined to reconstruct the transcriptional landscape of *F. psychrophilum* OSU THCO2-90 (**Figure 1**). Transcription start sites (TSSs) and transcribed regions (TRs) outside CDSs were identified by 5'-end and global RNA-Seq (**Tables S2 and S3**). Transcriptional responses were analyzed across 32 biological conditions representative of *F. psychrophilum* living environments (***SI Appendix* T3**). Genes were partitioned into clusters according to the hierarchical clustering tree of their expression profiles (**Figure 1**). Results presented here can be explored in details at https://fpeb.migale.inrae.fr.

Principal Component Analysis of the transcriptomes revealed highly coordinated changes in gene expression with growth phase (PC1) and between key environments: fish plasma (PC2), growth on blood (PC3) and freshwater (PC4; **Figure 1**, ***SI Appendix* D2**). Analysis revealed that 57.4% of the CDSs are highly expressed (in the upper quartile of expression level) in at least one sample of the dataset. Only 5.8% (136) are highly expressed in all samples including typical housekeeping genes encoding ribosomal proteins and carbon metabolism enzymes, but also those encoding the TonB-ExbBD system, several outer membrane proteins, the T9SS core components and gliding motility proteins (**Table S4**). Only 4.4% of the CDSs, mostly of unknown function, show low expression in all samples. High congruence between biological replicates allowed to identify differentially expressed genes (DEGs) between conditions: 86% CDSs were found differentially expressed in at least one comparison (***SI Appendix* T4, Table S5**). Taken together, these numbers indicate good coverage of the expression space.

Characterization of promoters and alternative sigma factor regulons

153 The 1884 genomic positions identified as putative primary 5'-ends by 5'-end RNA-Seq (***SI Appendix* D3**)
154 were further analyzed to establish a list of high-confidence TSSs and to identify high-level transcriptional
155 regulation by sigma factors, the transient subunits of bacterial RNA polymerase responsible for
156 recognition of promoter sequences. Besides the primary sigma factor, hereafter named σ^A, the *F.*
157 *psychrophilum* genome contains 8 alternative sigma factors: 1 σ^{54} and 7 extracytoplasmic function (σ^{ECF})
158 sigma factors, some of them induced in specific conditions such as plasma exposure and high osmotic
159 pressure (***SI Appendix* D4**). The dataset was analyzed for *de novo* identification of sigma factor binding
160 sites by combining information from DNA sequences, 5'-end positions and condition-dependent
161 expression profiles (40). Sites were predicted for 1194 (63.4%) of the initial list of putative TSSs and their
162 genomic contexts suggest a good sensitivity of *in silico* detection of promoters and good specificity of
163 experimental mapping of TSSs (***SI Appendix* D3**). These high-confidence TSSs precede 890 CDSs (38%
164 of the CDSs annotated in the genome) which gives a lower bound on the number of distinct mRNA
165 transcription units, each often encompassing several adjacent codirectional CDSs (polycistronic mRNAs).
166 The length of 5' untranslated regions (5' UTRs), computed as the distance between a TSS and the
167 predicted start codon, was examined for these 890 CDSs. The average length is 24 bp, half of them being
168 between 16 to 65 bp. Leaderless mRNAs might have been more frequent given the absence of Shine-
169 Dalgarno sequences, but they account for only 5.5% of these mRNAs, which is consistent with a previous
170 observation in *Flavobacterium johnsoniae* (29).
171 The *de novo* prediction algorithm associated TSSs to 6 distinct sigma factor binding site motifs, numbered
172 SM1-6 according to their number of occurrences (**Figure 2A**). The 3 most abundant (SM1-3) consist of
173 variations around the TAnnTTTG consensus of the -7 box recognized by σ^A. Subtle differences between
174 these σ^A motifs may have a role in the regulation of promoter activity (***SI Appendix* D5**). SM4-6 differ
175 from the σ^A consensus and collectively represent 7% of the high-confidence TSSs, most likely under the
176 control of alternative sigma factors.
177 Motif SM4 displays the typical -24/-12 elements of σ^{54} controlled promoters (48), indicating that
178 downstream genes are part of the σ^{54} regulon of *F. psychrophilum*. Transcription level downstream SM4
179 promoters showed a strong induction during growth on blood agar and, to some extent, under high

180 osmotic pressure, plasma exposure and into freshwater (**Figure 2A, *SI Appendix* D5**). The transcribed
181 genes mainly belong to two co-expression clusters that encode functions related to quality control of
182 proteins and envelope stress response, gliding motility, pyrimidine and isoprenoid biosynthesis, as well as
183 several exported proteins of unknown function (**Table 1A**). Across bacterial species, σ^{54} factors have been
184 reported to control pathways as diverse as nitrogen assimilation, flagellar biosynthesis or carbon uptake
185 but a common theme is to control processes related to physical interaction with the environment (49). This
186 is consistent with our findings in *F. psychrophilum*. Transcriptional activation by σ^{54} strictly depends on
187 enhancer-binding activators. Since three proteins containing a σ^{54} interaction domain (PF00158) were
188 predicted in the genome, co-expression clusters within SM4 promoters might correspond to distinct
189 activators sensing different environmental signals.
190 Motif SM5 is characterized by a TAnnTTGY box at the same position (-12 to -5 bp) than the σ^{A}
191 consensus TAnnTTTG. A striking similarity is the presence of highly conserved elements TA and TTG,
192 but the distance between them is 1 bp shorter in SM5. Accordingly, guanidine at -6 and pyrimidine at -5
193 are specific of SM5 promoters. Transcribed genes are related to LPS biosynthesis, amino-acids scavenging
194 and gliding motility (**Table 1B, *SI Appendix* D5**).
195 Motif SM6 contains a conserved CGT box in the -10 region (**Figure 2A**) that strongly suggests
196 recognition by a σ^{ECF}-type sigma factor (50). Transcription level downstream SM6 promoters is
197 characterized by a strong induction under low nutrients conditions such as freshwater or growth on very
198 low-nutrient agar. Transcribed genes encode several components of the T9SS (**Table 1C, *SI Appendix***
199 **D5**), which is reminiscent of the control of T9SS genes by a σ^{ECF} reported in *Porphyromonas gingivalis,* a
200 non-motile *Bacteroidetes* (51, 52). Two transmembrane proteins with conserved "Band-7" domain are
201 also transcribed from SM6 promoters. In prokaryotes, this family contains scaffold proteins called
202 flotillins that are associated with functional membrane microdomains. Though their function is not fully
203 understood, flotillins can promote protein complexes assembly (53). The presence of one anti-sigma/σ^{ECF}
204 factor system among the SM6 associated genes suggests that this system regulates SM6 promoters.
205 Nevertheless, the biological conditions of the expression of several σ^{ECF} factors overlap (***SI Appendix* D4**),
206 which also suggests functional redundancy and partially overlapping regulons, as observed in other

bacteria (54). SM6 could thus correspond to promoters recognized by several sigma factors, as observed for computationally inferred σ^{ECF} binding sites in *Bacillus subtilis* (40). Furthermore, the conditions considered in this study might not have allowed full activation of all σ^{ECF}-type factors. Consequently, the relative contribution of the predicted 7 σ^{ECF}-type factors in the control of SM6 promoters remains to be determined, and other σ^{ECF}-controlled TSSs may remain to be discovered.

Experimental validation of the inferred regulatory motifs was performed on a case of particular interest: the *remFG-sprCDBF* operon. This operon is conserved in several *Flavobacterium* species and encodes the cell surface adhesin SprB known in *F. johnsoniae* to mediate bacterial cell attachment and propulsion over surfaces, as well as SprF which is required for secretion of SprB by the T9SS (55, 56). Two promoters carrying the SM4 (σ^{54}) and SM5 sigma factor binding sites were predicted upstream the operon. By constructing a reporter plasmid for *F. psychrophilum* (***SI Appendix* D6**) and promoter mutagenesis, we confirmed the contribution of these two regulatory elements to *remFG-sprCDBF* transcription and the importance of highly conserved nucleotides of the sigma factor motifs identified *in silico* (**Figure 2B**).

The non-coding RNA repertoire of *F. psychrophilum*

A repertoire of 1511 regions transcribed outside annotated CDSs or ubiquitous RNAs were classified according to their transcriptional context and encompasses regulatory RNA candidates as well as signatures of pervasive transcription. The condition-dependent transcriptome dataset confirmed expression of at least 87% of the TRs (**Table 2; *SI Appendix* D1**).

Antisense RNAs

Overall, 287 TRs overlapping the antisense strand of 281 CDSs were detected. This represents 12% of the total number of CDSs and very diverse biological functions (**Table S6**). Antisense RNAs (asRNAs) originate mostly from imperfect termination of transcription and from transcription initiation at non-canonical locations, as exemplified by the high proportion of asRNAs in 3', 3'PT, and indep TR categories (**Table 2, *SI Appendix* D7**). Such antisense transcription patterns have been reported in a wide range of bacterial species and was referred to as pervasive transcription. While the general role of bacterial asRNAs is still an open question, many cases of regulatory functions involving mechanisms as diverse as

transcriptional interference, modulation of mRNA stability, and translation inhibition have been documented (reviewed in (57)). In *Bacteroides fragilis*, asRNAs were reported for 15 polysaccharides utilization loci (PULs), and some were shown to negatively modulate the expression of their cognate PUL (58). The list of asRNAs reported here might be a starting point for similar functional studies in flavobacteria.

5' cis-regulatory RNAs

5' *cis*-regulatory RNAs usually adopt complex secondary structures that are essential to sense a particular signal (*e.g.* small molecule, temperature, ribosome or protein binding). We confirm the expression of 5 predictions made by scanning the genome for known *cis*-regulatory RNA families using RFAM (44). For each of these 5' regulatory RNAs, there is a clear coherence between the predicted sensed signal (*e.g.* cobalamin, thiamine) and the functions and expression profiles of the regulated genes (***SI Appendix* D8**). In particular, a cobalamin riboswitch (control by cobalamin availability), was identified upstream of an unknown TonB-dependent transporter (TBDT, THC0290_1776) located in the vicinity of the *cob* genes encoding the adenosyl cobalamin biosynthesis proteins. Interestingly, enzymes catalyzing the first steps of the pathway are missing in *F. psychrophilum* and supply of this cofactor likely relies on the scavenging of cobamide precursors or vitamin B12 itself. Upregulation of this TBDT gene in fish plasma suggests retrieval of this nutrient from the host.

To go beyond confirmation of known 5' *cis*-regulatory RNAs, secondary RNA structures were examined. This resulted in a list of 64 structured 5' TRs that most likely play a regulatory role (**Table S7, *SI Appendix* D9**). Strong clues for a leader peptide attenuation mechanism by which translation of a short peptide regulates transcriptional elongation (59) were found upstream operons encoding amino-acids biosynthesis pathways. Structured 5'UTRs were identified for mRNAs related to a large variety of functions such as aminoacyl-tRNA-synthetases, carbon metabolism enzymes and peptidases. Several 5' TRs coincided with hits of Flavo-1, a computationally inferred RNA motif widespread in *Flavobacteriaceae* (60). Nevertheless, this does not give real clues on the function of this structural element since Flavo-1 hits seemed indistinctly located in sense or antisense of the TRs and many were not in TRs. A number of structured 5' elements identified for the first time in this study were conserved outside the species. Some are present in other genera of the family *Flavobacteriaceae* (*e.g.* the 5' elements

upstream of *pheA*, *hisG*, *ftsY*, *acsA*, *alaS*, or *rimP*). Without hits in the RFAM database, they probably represent original *cis*-regulatory elements whose study may be of great interest.

Small regulatory RNAs

A total of 4 known RNAs (small signal recognition particle RNA, RNase P, transfer-messenger RNA and 6S RNA) and 85 newly described RNAs (indep TRs), named Rfp1-89, were detected (**Table S8**). Indep TRs are typically short (median length, 176 bp) and without coding potential; 40 are transcribed in antisense of CDSs, with a potential *cis*-regulatory effect (**Table 2**). Outside asRNAs whose conservation cannot be assessed independently of cognate CDSs, indep TR sequences tend to be conserved across the species *F. psychrophilum* and many may have homologs in other species of the genus (**Figure 3**). Not surprisingly, known RNAs are among the most highly conserved. A fraction of these 45 non-antisense indep TRs probably act as *bona fide* regulatory sRNAs. sRNAs are key regulators that control numerous cellular processes by fine-tuning gene expression and usually act via base-pairing with several mRNAs, resulting in modulation of stability, structure and/or translation efficiency (61). In the absence of dedicated molecular experimental studies, the discovery of their regulatory networks mainly relies on *in silico* sRNA-mRNA pairing predictions (62).

A strong secondary structure was predicted for 18 (40%) of the 45 Indep TRs not listed as antisense, which suggests functionality (***SI Appendix* D10**). These include Rfp11, Rfp13, Rfp15 and Rfp18 which are conserved in genomes of other species (**Figure 3**). Condition-dependent expression profiles revealed that most of the sRNAs are expressed under specific environmental conditions, a trend well known from studies in other bacteria.

Predictions of mRNA-sRNA interactions were examined to identify putative targets (**Table S9**). Despite well-known limitations (*e.g.* high number of false positives), this approach is efficient to select candidates for functional characterization, particularly when several targets with related functions are predicted for a same sRNA (63). Among the 13 putative metalloprotease-encoding mRNAs identified in strain JIP02/86 (7), 4 are predicted as possible targets of Rfp18, a sRNA conserved outside the species *F. psychrophilum*, which likely folds into a strong secondary RNA structure (***SI Appendix* D11**). Another predicted target encodes a putative secreted adhesin (THC0290_2338) (18). The pairing region of Rfp18 was identical for all these predicted targets (third stem-loop). Regulation of proteases by sRNAs have already been reported

290 in other pathogenic bacteria, such as the collagenase ColA in *Clostridium perfringens*, the cysteine
291 protease SpeB in *Streptococcus pyogenes* or the Vsm protease in *Vibrio* species (64-67). *F. psychrophilum*
292 produces several degradative enzymes (mainly metalloproteases) that allow bacterial cells to digest
293 collagen, fibrinogen, elastin and fish muscle tissue, a trait proposed to participate to virulence and to
294 promote tissue erosion in infected fish (7, 12, 13, 41). Expression control of these putative virulence
295 determinants remains to be elucidated and most have unknown substrates.
296 A *rfp18* deletion mutant was constructed and the expression level of several predicted mRNA targets was
297 compared in wild-type and Δ*rfp18* strains by RT-qPCR. The results show that mRNA levels of two
298 metalloproteases, Fpp1 and THC0290_0300, are significantly reduced in Δ*rfp18* during stationary phase
299 (**Figure 4**). Fpp1 was reported as transcribed together with the Fpp2 metalloprotease (20), which was
300 confirmed here by identification of a single TSS. Regulation by Rfp18 did not affect Fpp2 mRNA level,
301 and Rfp18 pairing, which is predicted in the 107-bp intergenic region of the *fpp2-fpp1* operon, appears
302 thus to uncouple expression of the two genes. We confirmed expression of the homolog of Rfp18 in
303 *Flavobacterium columnare*. Furthermore, predicted targets in this other serious fish pathogen are
304 homologs to *F. psychrophilum* metalloproteases and adhesin THC0920_2338 (***SI Appendix* D11**).
305 Altogether, these results indicate that Rfp18 is required for the precise expression control of several
306 metalloproteases and evolutionary conservation underlines the importance of this regulatory mechanism.
307
308 **Transcriptional changes in response to environmental transitions**
309 Several stages can be considered in the *F. psychrophilum* life cycle: freshwater-living bacteria first attach
310 to the host surface and damage tissues for successful invasion, then they multiply, enter the bloodstream
311 and colonize organs until host's death (**Figure 5**). The shedding by infected fish can result in high
312 numbers of bacteria into tank water, a phenomenon proposed to contribute to the disease transmission (68).
313 Environmental conditions greatly vary between outside and within the host from standpoints as diverse as
314 osmotic pressure, exposure to host's defenses, shifts in nutrient sources and concentrations. Osmotic
315 pressure in body fluids is much higher than in freshwater or on fish surface. On the skin surface, the
316 bacterium has to adapt in order to resist the host defense components present in the mucus barrier such as

proteases, lysozyme, antimicrobial peptides, complement, lectins or immunoglobulins (69). Specific *in vitro* conditions were used to mimic outside, surface and inside host environmental niches: bacteria were incubated into freshwater in the presence of live fish or not, exposed to fish mucus or fish plasma (***SI Appendix* T3**). Additional conditions were analyzed to establish more direct functional links between genes and specific stimuli (*e.g.* peroxide stress, iron deprivation, hypoxia). Within-host osmotic conditions were mimicked by 0.75% NaCl supplementation in TYES broth, a concentration supporting growth but close to the maximum that *F. psychrophilum* can tolerate (12). DEGs were analyzed to predict functions and metabolic pathways involved in adaptation to these environmental conditions (**Table S5**, the ***fpeb*** website). As co-expressed genes (clusters) tend to have similar functions or to be part of the same biological pathway, we also formulate functional hypotheses for genes of unknown function. Results are reported in the following paragraphs and detailed in ***SI Appendix* D12-D15**.

Transcriptional adaptation of freshwater-living bacteria

Transcriptome profiles of bacteria maintained in freshwater differed markedly from laboratory conditions, with patterns testifying of the nutritional deprivation of *F. psychrophilum* outside the host (***SI Appendix* D12**). Dissolved O_2 was maintained near to saturation into tank water (10.7 mg L^{-1}, 10.5°C) to meet respiration needs of rainbow trout and transcriptional responses related to redox homeostasis indicate bacterial adaptation to oxidative stress. Expression of iron acquisition systems suggest availability of ferrous iron in freshwater. Fish skin contains large amounts of collagen and gelatin, whose hydrolysis-released peptides likely constitute an important source of nutrients for the bacterium during host invasion. Consistently, several peptidases were among the freshwater-induced genes. Among up-regulated genes of cluster B549, the collagenase, a SusCD family outer-membrane uptake system (THC0290_2100 and _2101) with homologs in the whole family *Flavobacteriaceae* and a peptidyl-dipeptidase Dcp2, could ensure the hydrolysis of collagen and the import of extracellular oligopeptides providing amino-acids for growth, as shown in *P. gingivalis* (70). Their up-regulation is independent of the presence of fish, suggesting that freshwater-living bacteria are transcriptionally pre-adapted to host encountering. This adaptation could be driven by nutrient starvation as expression of those genes also increases when cells enter into stationary phase.

Responses to fish components

In contrast to the collagenase, Fpp metalloproteases are induced in freshwater-living cells only when fish are present (***SI Appendix* D13**). Co-expression of cyanophycin synthetase CphA with *fpp* and THC0290_0300 metalloproteases suggests that peptides released by these proteases may partly be used for biosynthesis of cyanophycin, a branched non-ribosomal peptide composed of L-arginine and L-aspartate, commonly found in cyanobacteria and serving as a cytoplasmic reservoir for carbon, nitrogen and energy (7, 71). Other genes up-regulated specifically in the presence of fish include *ybcL* homolog, which encodes a protein inhibiting neutrophil migration in uropathogenic *E. coli* strains (72).

Genes related to fatty acids (FA) β-oxidation pathway are highly up-regulated in the presence of fish compounds such as mucus and plasma, suggesting that FA breakdown serves as an energy source. These observations indicate that FA could be scavenged from the host during the infectious process as observed in other bacterial pathogens (73). This was quite unexpected as *F. psychrophilum* was believed to solely use proteinaceous compounds but consistent with reported lipolytic activity (12).

Transcriptional changes also attest to responses against harmful conditions. Genes highly up-regulated under fish mucus exposure encode oleate hydratases that confer bacterial resistance to antimicrobial FA in some pathogenic bacteria (74, 75), multidrug efflux pumps that are used to extrude host antimicrobial peptides and FA, and exopolysaccharides biosynthesis proteins. Similar global responses involving LPS modifications and efflux pumps were reported in Gram-negative pathogenic bacteria submitted to antimicrobial compounds (76). The overlap observed between transcriptional responses under skin mucus and plasma exposure likely reflects the response to both mucosal and systemic innate immunity.

Life inside the host

Transcriptional analysis of stress responses, such as those induced by reactive oxygen species, hypoxia, or sequestration of essential metals, is an efficient way to discover genes required for virulence (***SI Appendix* D14**).

Peroxide stress response likely plays an important role in the resistance against bacterial killing during oxidative burst. Hydrogen peroxide exposure leads to a typical oxidative stress response in *F. psychrophilum*, characterized by the overexpression of antioxidative enzymes and components of [Fe-S] clusters assembly machinery. Uncharacterized transcription factors and conserved proteins of unknown function were also up-regulated suggesting their involvement in oxidative stress responses.

373 Inflammatory hypoxia can occur in infected tissues due to the activity of the numerous phagocytes
374 recruited *in situ* (77). As other pathogenic bacteria colonizing anoxic tissues (78), *F. psychrophilum*
375 exposed to oxygen limitation adapt by a strong up-regulation of the *cbb3*-type cytochrome oxidase (*cbb3*-
376 Cox). Also up-regulated, enzymes of the heme biosynthesis pathway could provide the porphyrins
377 required for assembly of newly synthetized *cbb3*-Cox complexes.
378 *F. psychrophilum* response to high osmotic pressure is characterized by the induction of osmoregulation
379 system, gliding motility genes, the T9SS C-terminal signal peptidase PorU, as well as thiol-specific
380 antioxidative enzymes. Growth on blood triggers transcriptional increase of those high osmolarity-induced
381 genes, suggesting that osmotic pressure serves as a signal for activation of the oxidative stress response.
382 Such coordinated regulation may anticipate the oxidative burst when the bacterium enters body fluids.
383 Bacterial pathogens have evolved to perceive iron scarcity as a marker of the host's internal environment
384 and have developed mechanisms to evade this nutritional immunity (79). *F. psychrophilum* response under
385 metal deprivation is characterized by the up-regulation of several TBDTs, which could be involved in iron
386 acquisition, and other uncharacterized genes expressed in iron-deficient conditions in other bacteria. FA
387 detoxification and hypoxia-induced genes are also part of the response to iron scarcity. Co-induction of
388 these genes by several stimuli reveals a common response to the multiple stresses faced during host
389 colonization.
390 Iron scarcity, hypoxia, osmotic and peroxide stress responses described above are all part of the global
391 transcriptional adaptation of *F. psychrophilum* cells exposed to fish plasma. Other genes induced by
392 plasma exposure encode efflux pumps, TBDTs that may play a role in blood-derived nutrients acquisition,
393 uncharacterized transcription factors, and several enzymes involved in O-antigen biosynthesis whose
394 modulation may participate to the resistance to killing by the host's complement.
395 A comparative proteomic analysis of *F. psychrophilum* identified 20 proteins modulated *in vivo* in
396 rainbow trout (80). Half of them corresponded to DEGs under our *in vitro* conditions (***SI Appendix* D14**),
397 which validates the strategy of mimicking within-host environments to provide a functional context to
398 unknown genes at the genome scale.
399 *Expression of putative virulence factors across environmental conditions*

400 The T9SS is known as essential for *F. psychrophilum* virulence in rainbow trout, however most proteins
401 predicted as secreted are uncharacterized and those contributing to virulence are not identified (18, 21).
402 Several T9SS secreted protein encoding genes are up-regulated under within-host mimicking conditions
403 while other are up-regulated in freshwater and might have a role in host invasion (***SI Appendix* D15**).
404 Among the putative secreted adhesins, 17 tandem-arranged Leucine-rich repeat (LRR) proteins are
405 induced under fish plasma exposure. In other species, LRR proteins were shown to mediate host-pathogen
406 interactions allowing adhesion to surface receptors of host immune cells (81). Many secreted proteins are
407 predicted peptidases and some are suspected to play a role in adaptation to specific fish hosts (7, 82). Here,
408 half of secreted proteases are found modulated by many biological conditions, and some appear more
409 dedicated to outside-host or within-host conditions. Dedicated experiments confirmed variations of
410 exoproteolytic activity consistent with the transcriptional changes (***SI Appendix* D15**). Altogether, these
411 results document how *F. psychrophilum* adapts its pool of secreted degradative enzymes. Expression of
412 this large variety of degradative enzymes and their refined regulation at environmental transitions
413 highlight sophisticated adaptations to a pathogenic lifestyle.
414

415 **CONCLUSION**

416 Due to their diverse ecological niches and their important contribution to water ecosystems, there is a
417 great interest in flavobacteria. Their biology revealed original features shared with other members of the
418 phylum *Bacteroidetes* such as their gene expression signals, protein secretion machinery, mode of
419 locomotion by gliding or original outer membrane systems dedicated to nutrients acquisition. We describe
420 here the first RNA landscape of a flavobacteria and the molecular changes taking place when a pathogen
421 of this family adapts to diverse natural environments, including conditions mimicking the host. The results
422 highlight similarities with other, better known, bacteria and original characteristics linked to the position
423 in the phylum *Bacteroidetes* and the ecological niche of an aquatic pathogen. The body of information
424 collected provides many directions for future dedicated studies and a comprehensive database is made
425 available to promote its exploitation.
426

DATA ACCESS

Expression data sets are deposited in GEO (accession numbers GSE163842, GSE164189 and GSE164190). The *F. psychrophilum* codon-optimized mCherry sequence is deposited in GenBank (accession number MW401799).

ACKNOWLEDGEMENTS

This work was financially supported by the Agence Nationale de la Recherche (grant ANR-17-CE20-0020-01 FlavoPatho). We thank Marco Moroldo, Jérôme Lecardonnel and the INRAE GABI-CRB Bridge platform for microarray services, Brigitte Kerouault and Victor Folcher for excellent technical assistance. We thank David Halpern and Philippe Bouloc for helpful discussions and warm support, and Claire Morvan for critical reading of the manuscript. This work has benefited from the facilities and expertise of the high throughput-sequencing platform of IMAGIF (http://www.i2bc.paris-saclay.fr/), the INRAE MIGALE bioinformatics facility (doi:10.15454/1.5572390655343293E12), and the MicroScope annotation platform for computational resources. The authors are grateful to the staff of the INRAE IERP fish facilities. Animal experiments and sampling were performed in accordance with the European Directive 2010/2063/UE, and were approved by the ethics committee of the INRAE Center in Jouy-en-Josas, France (authorization number 2015100215242446).

AUTHOR CONTRIBUTIONS

Author contributions following the CRediT taxonomy (https://casrai.org/credit/) are as follows: Conceptualization: PN, TR; Data curation: CG, ED, PN, TR; Formal analysis: CG, BL, BM, PN, TR; Funding acquisition: TR; Investigation: CG, BL, BF, EVD, JFB, TR; Project administration: TR; Supervision: CT, FR, ED, PN, TR; Software: CG, BM, PN; Visualization: CG, PN, TR; Writing – original draft: PN, TR; Writing – review & editing: CG, BL, EVD, JFB, FR, ED, PN, TR.

CONFLICT OF INTEREST

The authors declare no competing interests.

REFERENCES

1. FAO, editor The State of World Fisheries and Aquaculture2018 2018; Rome: Food and Agriculture Organization (FAO).
2. Bernardet JF. *Flavobacteriaceae*. Bergey's Manual of Systematics of Archaea and Bacteria2015.
3. Nematollahi A, Decostere A, Pasmans F, Haesebrouck F. *Flavobacterium psychrophilum* infections in salmonid fish. Journal of fish diseases. 2003;26(10):563-74.
4. Barnes ME, Brown ML. A Review of *Flavobacterium Psychrophilum* Biology, Clinical Signs, and Bacterial Cold Water Disease Prevention and Treatment. The Open Fish Science Journal. 2011;4:1-9.
5. Bayliss SC, Verner-Jeffreys DW, Bartie KL, Aanensen DM, Sheppard SK, Adams A, et al. The Promise of Whole Genome Pathogen Sequencing for the Molecular Epidemiology of Emerging Aquaculture Pathogens. Frontiers in microbiology. 2017;8:121.
6. Evensen O, Lorenzen E. An immunohistochemical study of *Flexibacter psychrophilus* infection in experimentally and naturally infected rainbow trout (*Oncorhynchus mykiss*) fry. Diseases of aquatic organisms. 1996;25(1):53-61.
7. Duchaud E, Boussaha M, Loux V, Bernardet JF, Michel C, Kerouault B, et al. Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum*. Nature biotechnology. 2007;25(7):763-9.
8. Wiens GD, LaPatra SE, Welch TJ, Rexroad C, 3rd, Call DR, Cain KD, et al. Complete Genome Sequence of *Flavobacterium psychrophilum* Strain CSF259-93, Used To Select Rainbow Trout for Increased Genetic Resistance against Bacterial Cold Water Disease. Genome announcements. 2014;2(5).
9. Castillo D, Christiansen RH, Dalsgaard I, Madsen L, Middelboe M. Bacteriophage resistance mechanisms in the fish pathogen *Flavobacterium psychrophilum*: linking genomic mutations to changes in bacterial virulence factors. Applied and environmental microbiology. 2015;81(3):1157-67.
10. Rochat T, Barbier P, Nicolas P, Loux V, Perez-Pascual D, Guijarro JA, et al. Complete Genome Sequence of *Flavobacterium psychrophilum* Strain OSU THCO2-90, Used for Functional Genetic Analysis. Genome announcements. 2017;5(8).
11. Duchaud E, Rochat T, Habib C, Barbier P, Loux V, Guerin C, et al. Genomic Diversity and Evolution of the Fish Pathogen *Flavobacterium psychrophilum*. Frontiers in microbiology. 2018;9:138.
12. Bernardet JF, Kerouault B. Phenotypic and genomic studies of "*Cytophaga psychrophila*" isolated from diseased rainbow trout (*Oncorhynchus mykiss*) in France. Applied and environmental microbiology. 1989;55(7):1796-800.
13. Otis EJ. Lesions of cold-water disease in steelhead trout (*Salmo gairdneri*): the role of *Cytophaga psychrophila* extracellular products. [MSc thesis]: University of Rhode Island, Kingston; 1984.
14. Hogfors-Ronnholm E, Wiklund T. Hemolytic activity in *Flavobacterium psychrophilum* is a contact-dependent, two-step mechanism and differently expressed in smooth and rough phenotypes. Microbial pathogenesis. 2010;49(6):369-75.
15. Hogfors-Ronnholm E, Wiklund T. In vitro opsonin-independent interactions between cells of smooth and rough phenotypes of *Flavobacterium psychrophilum* and rainbow trout (*Oncorhynchus mykiss*) head kidney macrophages. Microbial pathogenesis. 2012;53(5-6):214-8.
16. Papadopoulou A, Dalsgaard I, Linden A, Wiklund T. In vivo adherence of *Flavobacterium psychrophilum* to mucosal external surfaces of rainbow trout (*Oncorhynchus mykiss*) fry. Journal of fish diseases. 2017;40(10):1309-20.
17. Decostere A, D'Haese E, Lammens M, Nelis H, Haesebrouck F. In vivo study of phagocytosis, intracellular survival and multiplication of *Flavobacterium psychrophilum* in rainbow trout, *Oncorhynchus mykiss* (Walbaum), spleen phagocytes. Journal of fish diseases. 2001;24:481-7.
18. Barbier P, Rochat T, Mohammed HH, Wiens GD, Bernardet JF, Halpern D, et al. The type IX secretion system is required for virulence of the fish pathogen *Flavobacterium psychrophilum*. Applied and environmental microbiology. 2020.
19. Alvarez B, Alvarez J, Menendez A, Guijarro JA. A mutant in one of two *exbD* loci of a TonB system in *Flavobacterium psychrophilum* shows attenuated virulence and confers protection against cold water disease. Microbiology (Reading, England). 2008;154(Pt 4):1144-51.
20. Pérez-Pascual D, Gomez E, Alvarez B, Mendez J, Reimundo P, Navais R, et al. Comparative analysis and mutation effects of *fpp2-fpp1* tandem genes encoding proteolytic extracellular enzymes of *Flavobacterium psychrophilum*. Microbiology (Reading, England). 2011;157:1196-204.
21. Pérez-Pascual D, Rochat T, Kerouault B, Gomez E, Neulat-Ripoll F, Henry C, et al. More Than Gliding: Involvement of GldD and GldG in the Virulence of *Flavobacterium psychrophilum*. Frontiers in microbiology. 2017;8:2168.
22. Madetoja J, Nystedt S, Wiklund T. Survival and virulence of *Flavobacterium psychrophilum* in water microcosms. FEMS microbiology ecology. 2003;43(2):217-23.
23. Vatsos IN, Thompson KD, Adams A. Starvation of *Flavobacterium psychrophilum* in broth, stream water and distilled water. Diseases of aquatic organisms. 2003;56(2):115-26.
24. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. Nature. 2010;464(7286):250-5.
25. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, et al. The *Listeria* transcriptional landscape from saprophytism to virulence. Nature. 2009;459(7249):950-6.
26. Innocenti N, Golumbeanu M, Fouquier d'Herouel A, Lacoux C, Bonnin RA, Kennedy SP, et al. Whole-genome mapping of 5' RNA ends in bacteria by tagged sequencing: a comprehensive view in *Enterococcus faecalis*. RNA. 2015;21(5):1018-30.
27. Kroger C, MacKenzie KD, Alshabib EY, Kirzinger MWB, Suchan DM, Chao TC, et al. The primary transcriptome, small RNAs and regulation of antimicrobial resistance in *Acinetobacter baumannii* ATCC 17978. Nucleic acids research. 2018;46(18):9684-98.
28. Heidrich N, Bauriedl S, Barquist L, Li L, Schoen C, Vogel J. The primary transcriptome of *Neisseria meningitidis* and its interaction with the RNA chaperone Hfq. Nucleic acids research. 2017;45(10):6147-67.

29. Baez WD, Roy B, McNutt ZA, Shatoff EA, Chen S, Bundschuh R, et al. Global analysis of protein synthesis in *Flavobacterium johnsoniae* reveals the use of Kozak-like sequences in diverse bacteria. Nucleic acids research. 2019;47(20):10477-88.
30. Ficko-Blean E, Prechoux A, Thomas F, Rochat T, Larocque R, Zhu Y, et al. Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. Nature communications. 2017;8(1):1685.
31. Gomez-Consarnau L, Gonzalez JM, Riedel T, Jaenicke S, Wagner-Dobler I, Sanudo-Wilhelmy SA, et al. Proteorhodopsin light-enhanced growth linked to vitamin-B1 acquisition in marine Flavobacteria. The ISME journal. 2016;10(5):1102-12.
32. Tang K, Lin Y, Han Y, Jiao N. Characterization of Potential Polysaccharide Utilization Systems in the Marine *Bacteroidetes Gramella Flava* JLT2011 Using a Multi-Omics Approach. Frontiers in microbiology. 2017;8:220.
33. Tekedar HC, Karsi A, Reddy JS, Nho SW, Kalindamar S, Lawrence ML. Comparative Genomics and Transcriptional Analysis of *Flavobacterium columnare* Strain ATCC 49512. Frontiers in microbiology. 2017;8:588.
34. Bayley DP, Rocha ER, Smith CJ. Analysis of *cepA* and other *Bacteroides fragilis* genes reveals a unique promoter structure. FEMS microbiology letters. 2000;193(1):149-54.
35. Vingadassalom D, Kolb A, Mayer C, Rybkine T, Collatz E, Podglajen I. An unusual primary sigma factor in the *Bacteroidetes* phylum. Mol Microbiol. 2005;56(4):888-902.
36. Chen S, Bagdasarian M, Kaufman MG, Walker ED. Characterization of strong promoters from an environmental *Flavobacterium hibernum* strain by using a green fluorescent protein-based reporter system. Applied and environmental microbiology. 2007;73(4):1089-100.
37. Wegmann U, Horn N, Carding SR. Defining the bacteroides ribosomal binding site. Applied and environmental microbiology. 2013;79(6):1980-9.
38. Kroger C, Colgan A, Srikumar S, Handler K, Sivasankaran SK, Hammarlof DL, et al. An infection-relevant transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium. Cell host & microbe. 2013;14(6):683-95.
39. Mader U, Nicolas P, Depke M, Pane-Farre J, Debarbouille M, van der Kooi-Pol MM, et al. *Staphylococcus aureus* Transcriptome Architecture: From Laboratory to Infection-Mimicking Conditions. PLoS genetics. 2016;12(4):e1005962.
40. Nicolas P, Mader U, Dervyn E, Rochat T, Leduc A, Pigeonneau N, et al. Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. Science. 2012;335(6072):1103-6.
41. Bertolini JM, Wakabayashi H, Watral VG, Whipple MJ, Rohovec JS. Electrophoretic detection of proteases from selected strains of *Flexibacter psychrophilus* and assesment of their variability. Journal of aquatic animal health. 1994;6:224-33.
42. Zhu Y, Thomas F, Larocque R, Li N, Duffieux D, Cladiere L, et al. Genetic analyses unravel the crucial role of a horizontally acquired alginate lyase for brown algal biomass degradation by *Zobellia galactanivorans*. Environ Microbiol. 2017;19(6):2164-81.
43. Rochat T, Pérez-Pascual D, Nilsen H, Carpentier M, Bridel S, Bernardet JF, et al. Identification of a novel elastin-degrading enzyme from the fish pathogen *Flavobacterium psychrophilum*. Applied and environmental microbiology. 2019.
44. Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, et al. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic acids research. 2018;46(D1):D335-D42.
45. Mirauta B, Nicolas P, Richard H. Parseq: reconstruction of microbial transcription landscape from RNA-Seq read counts using state-space models. Bioinformatics. 2014;30(10):1409-16.
46. Buels R, Yao E, Diesh CM, Hayes RD, Munoz-Torres M, Helt G, et al. JBrowse: a dynamic web platform for genome visualization and analysis. Genome biology. 2016;17:66.
47. Priyam A, Woodcroft BJ, Rai V, Moghul I, Munagala A, Ter F, et al. Sequenceserver: A Modern Graphical User Interface for Custom BLAST Databases. Mol Biol Evol. 2019;36(12):2922-4.
48. Barrios H, Valderrama B, Morett E. Compilation and analysis of sigma(54)-dependent promoter sequences. Nucleic acids research. 1999;27(22):4305-13.
49. Francke C, Groot Kormelink T, Hagemeijer Y, Overmars L, Sluijter V, Moezelaar R, et al. Comparative analyses imply that the enigmatic Sigma factor 54 is a central controller of the bacterial exterior. BMC Genomics. 2011;12:385.
50. Staron A, Sofia HJ, Dietrich S, Ulrich LE, Liesegang H, Mascher T. The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. Mol Microbiol. 2009;74(3):557-81.
51. Kadowaki T, Yukitake H, Naito M, Sato K, Kikuchi Y, Kondo Y, et al. A two-component system regulates gene expression of the type IX secretion component proteins via an ECF sigma factor. Sci Rep. 2016;6:23288.
52. Sato K, Naito M, Yukitake H, Hirakawa H, Shoji M, McBride MJ, et al. A protein secretion system linked to bacteroidete gliding motility and pathogenesis. Proc Natl Acad Sci U S A. 2010;107(1):276-81.
53. Lopez D, Koch G. Exploring functional membrane microdomains in bacteria: an overview. Curr Opin Microbiol. 2017;36:76-84.
54. Helmann JD. The extracytoplasmic function (ECF) sigma factors. Adv Microb Physiol. 2002;46:47-110.
55. Kulkarni SS, Johnston JJ, Zhu Y, Hying ZT, McBride MJ. The Carboxy-Terminal Region of *Flavobacterium johnsoniae* SprB Facilitates Its Secretion by the Type IX Secretion System and Propulsion by the Gliding Motility Machinery. Journal of bacteriology. 2019;201(19).
56. Rhodes RG, Nelson SS, Pochiraju S, McBride MJ. *Flavobacterium johnsoniae sprB* is part of an operon spanning the additional gliding motility genes *sprC*, *sprD*, and *sprF*. Journal of bacteriology. 2011;193(3):599-610.
57. Georg J, Hess WR. Widespread Antisense Transcription in Prokaryotes. Microbiol Spectr. 2018;6(4).
58. Cao Y, Forstner KU, Vogel J, Smith CJ. *cis*-Encoded Small RNAs, a Conserved Mechanism for Repression of Polysaccharide Utilization in *Bacteroides*. Journal of bacteriology. 2016;198(18):2410-8.
59. Gollnick P, Babitzke P. Transcription attenuation. Biochim Biophys Acta. 2002;1577(2):240-50.
60. Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, et al. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. Genome biology. 2010;11(3):R31.

61. Jorgensen MG, Pettersen JS, Kallipolitis BH. sRNA-mediated control in bacteria: An increasing diversity of regulatory mechanisms. Biochim Biophys Acta Gene Regul Mech. 2020;1863(5):194504.
62. Georg J, Lalaouna D, Hou S, Lott SC, Caldelari I, Marzi S, et al. The power of cooperation: Experimental and computational approaches in the functional characterization of bacterial sRNAs. Mol Microbiol. 2020;113(3):603-12.
63. Pain A, Ott A, Amine H, Rochat T, Bouloc P, Gautheret D. An assessment of bacterial small RNA target prediction programs. RNA Biol. 2015;12(5):509-13.
64. Kovacikova G, Skorupski K. Regulation of virulence gene expression in *Vibrio cholerae* by quorum sensing: HapR functions at the *aphA* promoter. Mol Microbiol. 2002;46(4):1135-47.
65. Mangold M, Siller M, Roppenser B, Vlaminckx BJ, Penfound TA, Klein R, et al. Synthesis of group A streptococcal virulence factors is controlled by a regulatory RNA molecule. Mol Microbiol. 2004;53(5):1515-27.
66. Nguyen AN, Disconzi E, Charriere GM, Destoumieux-Garzon D, Bouloc P, Le Roux F, et al. *csrB* Gene Duplication Drives the Evolution of Redundant Regulatory Pathways Controlling Expression of the Major Toxic Secreted Metalloproteases in *Vibrio tasmaniensis* LGP32. mSphere. 2018;3(6).
67. Obana N, Shirahama Y, Abe K, Nakamura K. Stabilization of *Clostridium perfringens* collagenase mRNA by VR-RNA-dependent cleavage in 5' leader sequence. Mol Microbiol. 2010;77(6):1416-28.
68. Madetoja J, Nyman P, Wiklund T. *Flavobacterium psychrophilum*, invasion into and shedding by rainbow trout *Oncorhynchus mykiss*. Diseases of aquatic organisms. 2000;43(1):27-38.
69. Ángeles Esteban M. An Overview of the Immunological Defenses in Fish Skin. ISRN Immunology. 2012;2012.
70. Madej M, White JBR, Nowakowska Z, Rawson S, Scavenius C, Enghild JJ, et al. Structural and functional insights into oligopeptide acquisition by the RagAB transporter from *Porphyromonas gingivalis*. Nat Microbiol. 2020.
71. Ziegler K, Diener A, Herpin C, Richter R, Deutzmann R, Lockau W. Molecular characterization of cyanophycin synthetase, the enzyme catalyzing the biosynthesis of the cyanobacterial reserve material multi-L-arginyl-poly-L-aspartate (cyanophycin). Eur J Biochem. 1998;254(1):154-9.
72. Lau ME, Loughman JA, Hunstad DA. YbcL of uropathogenic *Escherichia coli* suppresses transepithelial neutrophil migration. Infection and immunity. 2012;80(12):4123-32.
73. Fozo EM, Rucks EA. The Making and Taking of Lipids: The Role of Bacterial Lipid Synthesis and the Harnessing of Host Lipids in Bacterial Pathogenesis. Adv Microb Physiol. 2016;69:51-155.
74. Subramanian C, Frank MW, Batte JL, Whaley SG, Rock CO. Oleate hydratase from *Staphylococcus aureus* protects against palmitoleic acid, the major antimicrobial fatty acid produced by mammalian skin. The Journal of biological chemistry. 2019;294(23):9285-94.
75. Volkov A, Liavonchanka A, Kamneva O, Fiedler T, Goebel C, Kreikemeyer B, et al. Myosin cross-reactive antigen of *Streptococcus pyogenes* M49 encodes a fatty acid double bond hydratase that plays a role in oleic acid detoxification and bacterial virulence. The Journal of biological chemistry. 2010;285(14):10353-61.
76. Band VI, Weiss DS. Mechanisms of Antimicrobial Peptide Resistance in Gram-Negative Bacteria. Antibiotics (Basel). 2015;4(1):18-41.
77. Campbell EL, Bruyninckx WJ, Kelly CJ, Glover LE, McNamee EN, Bowers BE, et al. Transmigrating neutrophils shape the mucosal microenvironment through localized oxygen depletion to influence resolution of inflammation. Immunity. 2014;40(1):66-77.
78. Pitcher RS, Watmough NJ. The bacterial cytochrome *cbb3* oxidases. Biochim Biophys Acta. 2004;1655(1-3):388-99.
79. Skaar EP. The battle for iron between bacterial pathogens and their vertebrate hosts. PLoS pathogens. 2010;6(8):e1000949.
80. LaFrentz BR, LaPatra SE, Call DR, Wiens GD, Cain KD. Proteomic analysis of *Flavobacterium psychrophilum* cultured and in iron-limited media. Diseases of aquatic organisms. 2009;87(3):171-82.
81. Loimaranta V, Hytonen J, Pulliainen AT, Sharma A, Tenovuo J, Stromberg N, et al. Leucine-rich repeats of bacterial surface proteins serve as common pattern recognition motifs of human scavenger receptor gp340. The Journal of biological chemistry. 2009;284(28):18614-23.
82. Nakayama H, Tanaka K, Teramura N, Hattori S. Expression of collagenase in *Flavobacterium psychrophilum* isolated from cold-water disease-affected ayu (*Plecoglossus altivelis*). Biosci Biotechnol Biochem. 2016;80(1):135-44.

FIGURE LEGENDS

Figure 1. Global assessment of the RNA landscape and condition-dependent transcriptome. Upper panel: Overview of the strategy that combines the advantages of RNA-Seq and microarrays: detection of low-abundance transcripts and bp-level resolution; accurate and cost-effective quantification of transcript levels without PCR amplification biases. Lower left panel: 3D representation of the 64 samples with coordinates on the PC axes (Principal Component Analysis); lower right panel: Heatmap representation of gene-centered expression profiles and pairwise comparisons of samples identifying condition-relevant DEGs.

Figure 2. Sigma factor binding sites. A. Logo representation of the 6 sigma factor binding motifs identified *in silico* and average expression downstream the corresponding TSSs across conditions. Expression levels have been normalized by applying the transformation used for quantile-normalization of CDS expression levels. **B.** Mutagenesis of sigma factor binding sites of the *remFG-sprCDBF* operon. Schematic representation of transcriptional fusions and mutagenesis of highly conserved nucleotides in SM4 and SM5 motifs. Promoter activity was measured using whole-cell fluorescence of *F. psychrophilum* strains carrying the mCherry reporter plasmid. Values represent the mean and standard deviation of three independent experiments.

Figure 3. Conservation and secondary structure of Indep TRs. Heatmap representation of conservation profile for the 45 Indep TRs which are not antisense of CDSs and 4 RFAM RNAs. Length and evidence for secondary structure (low Minimum Free Energy z-score indicates significant folding) are represented as bar-plots on the right-hand side of the heatmap.

Figure 4: Rfp18 deletion affects the mRNA level of two metalloprotease encoding genes. RNAs from wild-type and Δ*rfp18* cultures sampled in exponential (OD600 = 0.5), transition (1.2) and stationary phase (2) were used for RT-qPCR assays. Relative quantification ($2^{-\Delta\Delta Ct}$; RQ) of putative target mRNA is calculated using wild-type in exponential phase of growth as the reference condition. Values are the mean ± *s.e.m* from five biological replicates. (*) indicates significant difference using two-way ANOVA analysis (Bonferroni adjusted p-value < 0.05).

Figure 5. Graphical summary of the main functions identified to contribute to *F. psychrophilum* adaptation. *In vitro* conditions analyzed (uppercase letters) are positioned to reflect three environmental

niches: freshwater living bacteria, external surface associated bacteria and within-host bacteria. The three conditions related to fish-derived components are in colored uppercase (freshwater-living bacteria, blue; fish mucus, yellow; fish plasma, red). Arrows indicate transcriptional up-regulation.

Table 1: Promoters with predicted alternative sigma factor binding sites.

TSS position	Downstream genes (locustag)	Downstream gene products	Gene cluster
A. Selection of SM4 promoters			
867211	THC0290_0729	10 kDa chaperonin GroES	B254
	THC0290_0728	60 kDa chaperonin GroEL	
1168709	THC0290_0994	ATP-dependent endopeptidase Lon (S16 family)	B254
1654192	THC0290_1425	Chaperone protein DnaK	B254
1798337	THC0290_1544	Chaperone protein GrpE	B254
	THC0290_1545	Chaperone protein DnaJ	
2117454	THC0290_1828	Chaperone protein HtpG	B254
753817	THC0290_0618	Outer membrane chaperone Skp (OmpH)	B258
1110250	THC0290_0942	ClpB, ATPase with chaperone activity	B415
2425654	THC0290_2103	Rhomboid family intramembrane serine protease	B415
25239	THC0290_0021 to THC0290_0026	Gliding motility operon *remF-remG-sprC-sprD-sprB-sprF*	B419
356002	THC0290_0297	Protein of unknown function	B419
	THC0290_0298	1-deoxy-D-xylulose-5-phosphate synthase Dxs (isoprenoid biosynthesis)	
710561	THC0290_0583	Uncharacterized N-acetyltransferase YitI	B419
840245	THC0290_N_0510	Putative regulatory RNA Rfp36	B419
966765	THC0290_0811	Putative muramidase	B419
1042927	THC0290_0885	Probable lipoprotein precursor	B419
1655528	THC0290_1426	Protein of unknown function precursor	B419
1806311	THC0290_1552	Orotate phosphoribosyltransferase PyrE (pyrimidine metabolism)	B419
2172256	THC0290_1870	Protein of unknown function precursor	B419
2618622	THC0290_2263	Thioredoxin family protein	B419
490781	THC0290_0402	Outer membrane protein precursor, AsmA family	
552525	THC0290_0455	Protein of unknown function with C-type lysozyme inhibitor family domain	
2117336	THC0290_1827	Probable lipoprotein	
	THC0290_1826	Probable lipoprotein with OmpA-like domain	
B. Selection of SM5 promoters			
3491	THC0290_0003	LysE-type exporter protein	
25333	THC0290_0021 to THC0290_0026	Gliding motility operon *remF-remG-sprC-sprD-sprB-sprF*	
479610	THC0290_0394	Asp/Glu-specific dipeptidyl-peptidase, Dpp11	
870544	THC0290_0733	Probable sigma-54-dependent transcriptional regulator	
	THC0290_0732	Lipopolysaccharide assembly protein LptE	
999712	THC0290_0839	Protein of unknown function	
	THC0290_0838	dTDP-4-amino-4,6-dideoxygalactose transaminase	
	THC0290_0837	Oligosaccharide translocase WzxE	
1187638	THC0290_1009	Probable transmembrane protein of unknown function	
1582507	THC0290_1371	Putative glutamine cyclotransferase	
1716045	THC0290_1475	Glycosyl transferase (group 1 family)	
2046887	THC0290_1761	Probable lipoprotein precursor	
2501832	THC0290_2166	Putative enzyme with P-loop containing nucleotide triphosphate hydrolase domain	
2676614	THC0290_2321	Metallo-beta-lactamase superfamily protein	
2773735	THC0290_2400	RmlD, dTDP-4-dehydrorhamnose reductase (synthesis of dTDP-L-rhamnose)	
C. Selection of SM6 promoters			
9250	THC0290_0009	Putative inner membrane protein (Band 7 family)	
49622	THC0290_0031	Protein of unknown function YceA, rhodanese-related sulfurtransferase	
182564	THC0290_N_0107	Putative regulatory RNA Rfp11	
758025	THC0290_0621	Outer membrane protein; Homolog in *P. gingivalis* (PG0189) interacts with the PorM/PorN complex of T9SS	

882660	THC0290_0743	GldN, core component of the T9SS machinery
1229139	THC0290_1050	ATPase component of a probable ABC-type transporter
1750688	THC0290_1507	Putative inner membrane protein (Band 7 family)
1954718	THC0290_1681	TPR-domain containing protein
2057098	THC0290_1774	ECF-type sigma factor
	THC0290_1773	Anti-sigma factor
2382309	THC0290_2050	RmlB, dTDP-glucose 4,6-dehydratase (synthesis of dTDP-L-rhamnose)
2473290	THC0290_2146	PorU, C-terminal signal peptidase of the T9SS

A full list of promoters is available in **Table S2**, see also the genome browser at https://fpeb.migale.inrae.fr/tss_annotation.html.

Table 2: Summary of the 1511 regions transcribed outside annotated CDSs.

Category of TRs[a]	Total number (#)	Median length (bp)	TRs of length < 150bp			TRs of length ≥ 150 bp			Expression[d]
			#	CDS[b]	AS[c]	#	CDS[b]	AS[c]	
5’	231	80	182	0	0	49	4	4	95%
3’	586	108	333	0	21	253	14	166	79%
3’ PT	67	314	13	0	4	54	5	35	78%
Intra	542	99	374	2	0	168	10	4	91%
Indep	85	176	40	0	14	45	4	26	94%
All	1511	105	942	2	39	569	37	235	87%

[a] classification by transcriptional context: [5’] = in 5' of a CDS; [3’] = in 3' of a CDS; [3’-PT] = resulting from partial termination of transcription (*i.e.* downstream a 3’ region terminated by a predicted intrinsic terminator); [intra] = between two CDSs, part of a polycistronic mRNA; and [indep] = independent of previously annotated features.

[b] number of TRs with predicted CDS

[c] number of TRs in antisense of CDSs or new TRs

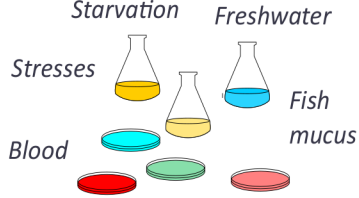
[d] percentage of TRs whose expression level is above the lower quartile of CDS expression in at least one sample.

A detailed description of the 1511 TRs and all features from genome annotation is available in **Table S3**, see also the ***fpeb*** website at https://fpeb.migale.inrae.fr/structural_annotation.html.

Pool of total RNA extracts:
18 biological conditions

5'-end RNA-Seq

Global RNA-Seq



Individual RNA extracts:
32 biological conditions mimicking life cycle transitions *in vitro*

Transcription start sites:
1884 5'-ends

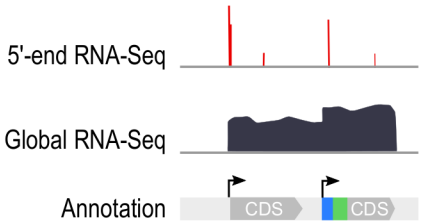
Transcription outside CDSs:
1511 regions

Design of microarray probing transcribed regions and TSSs

High-density microarray

Global RNA landscape

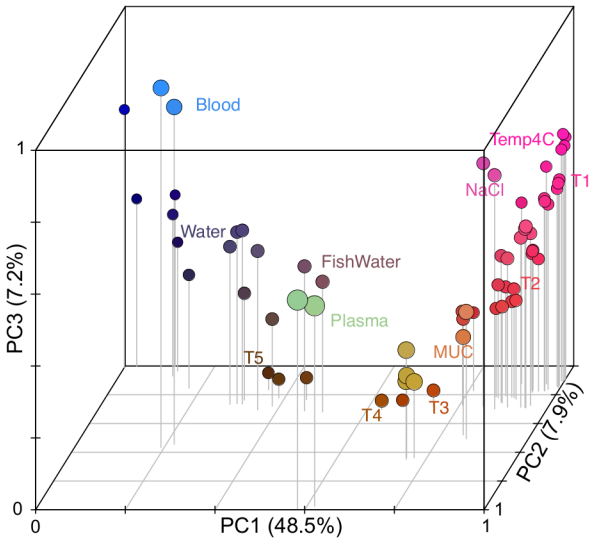
Condition-dependent transcriptome

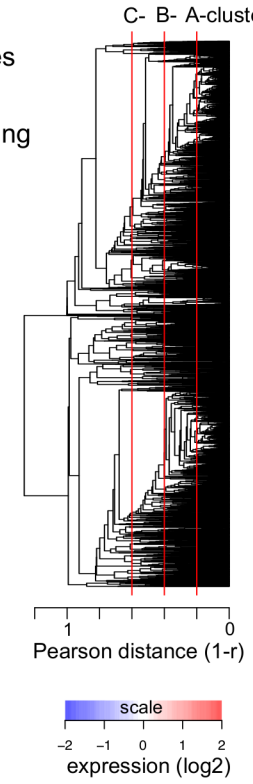


1194 promoters with sigma factor binding sites

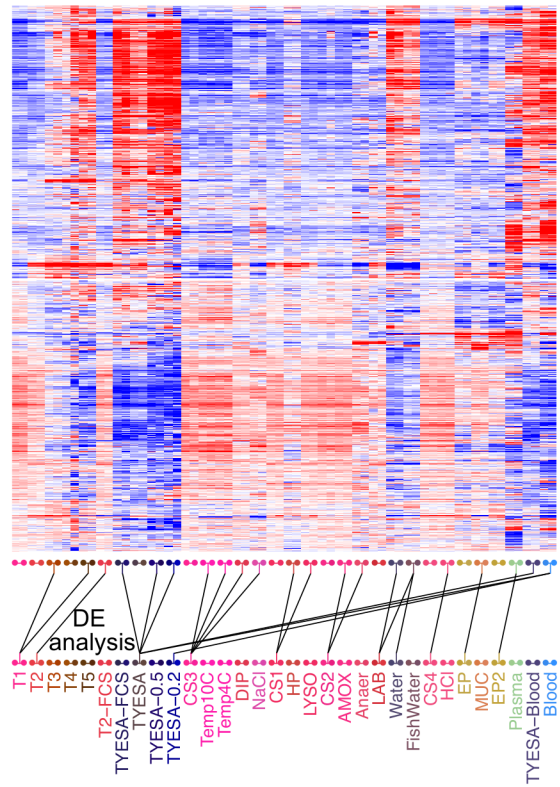
66 *cis*- and 45 *trans*-acting regulatory RNAs

Principal component analysis

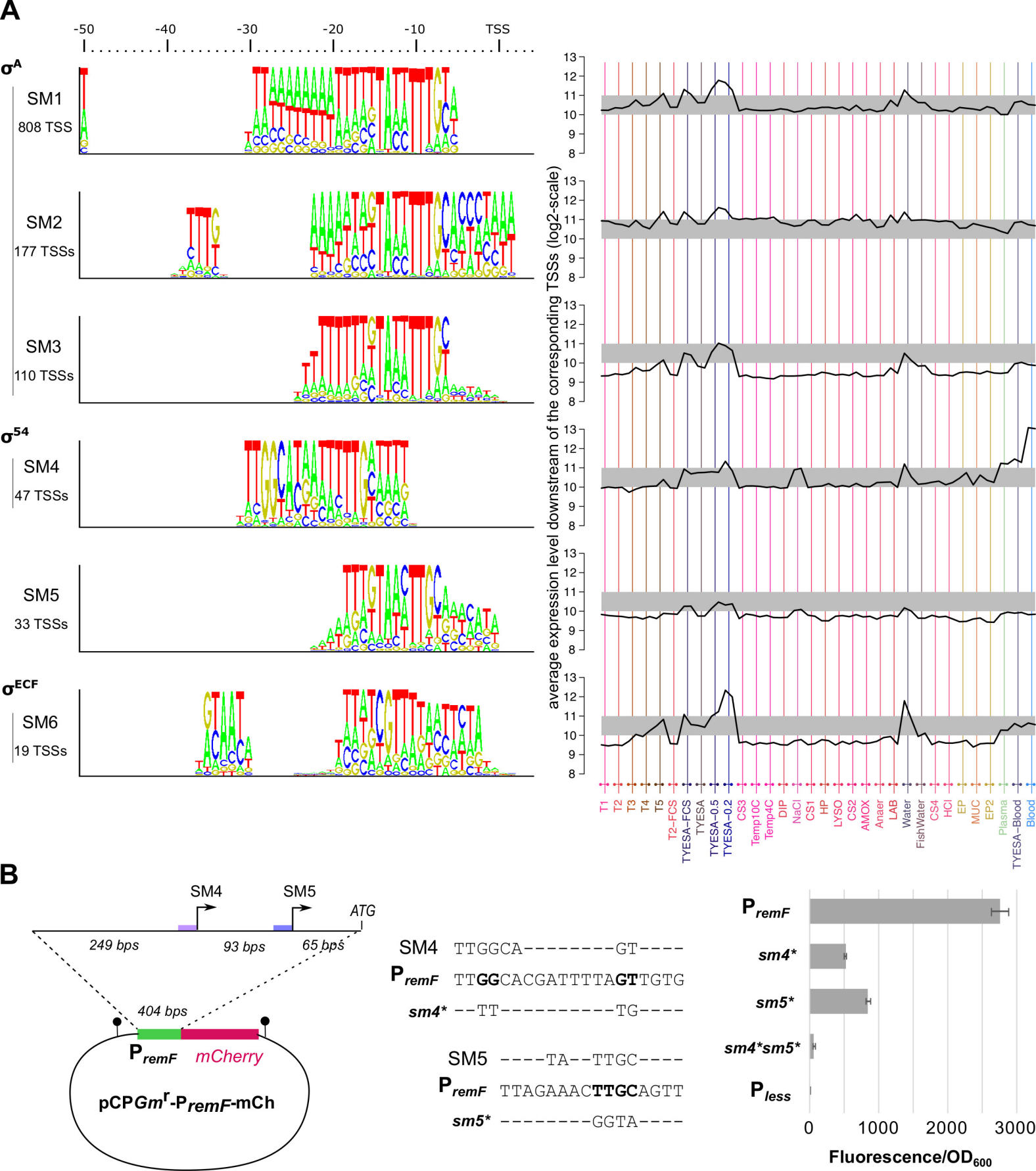


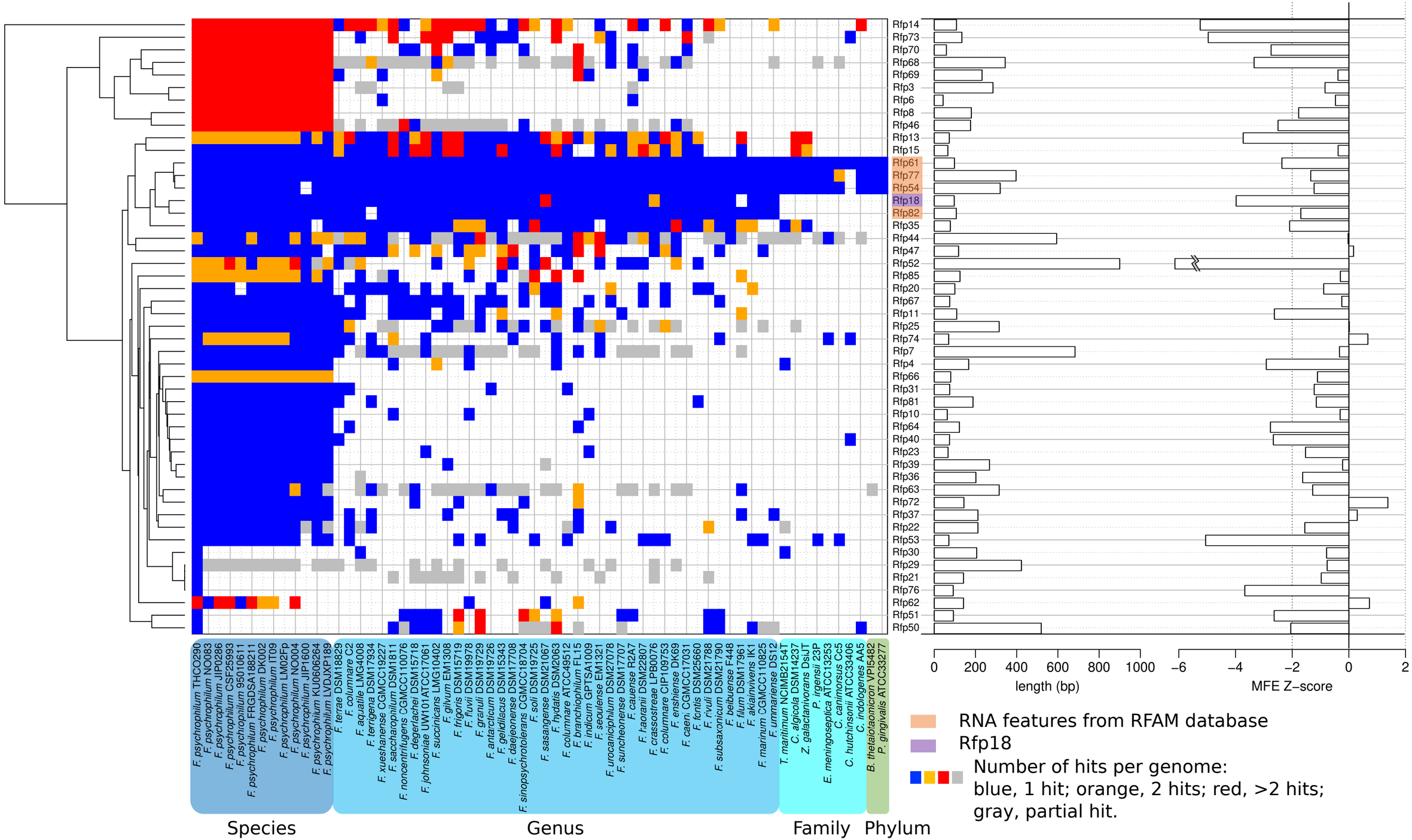


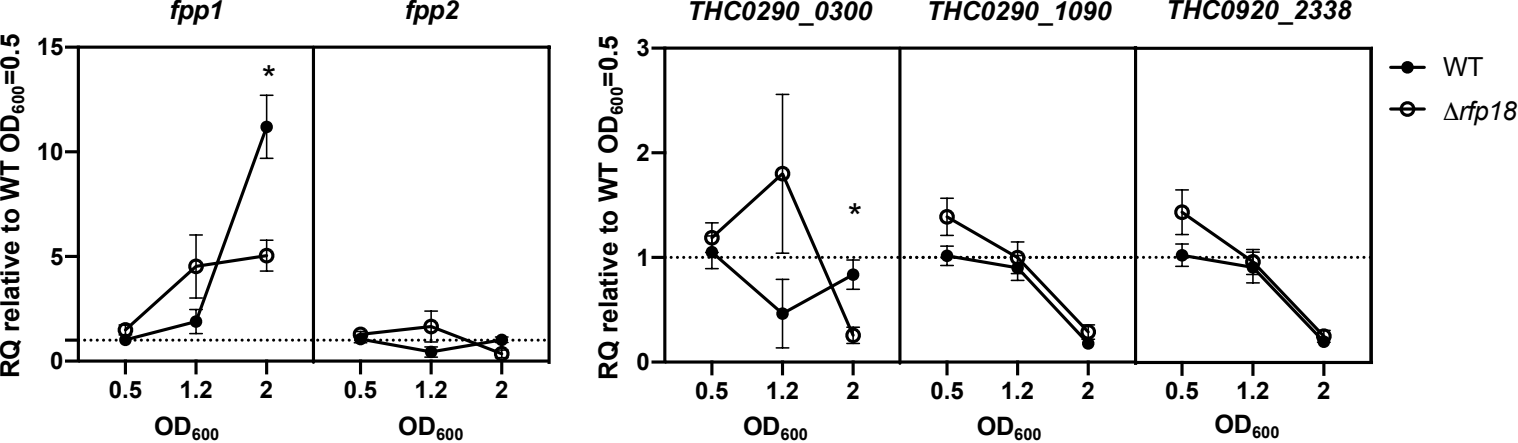
Gene expression profiling

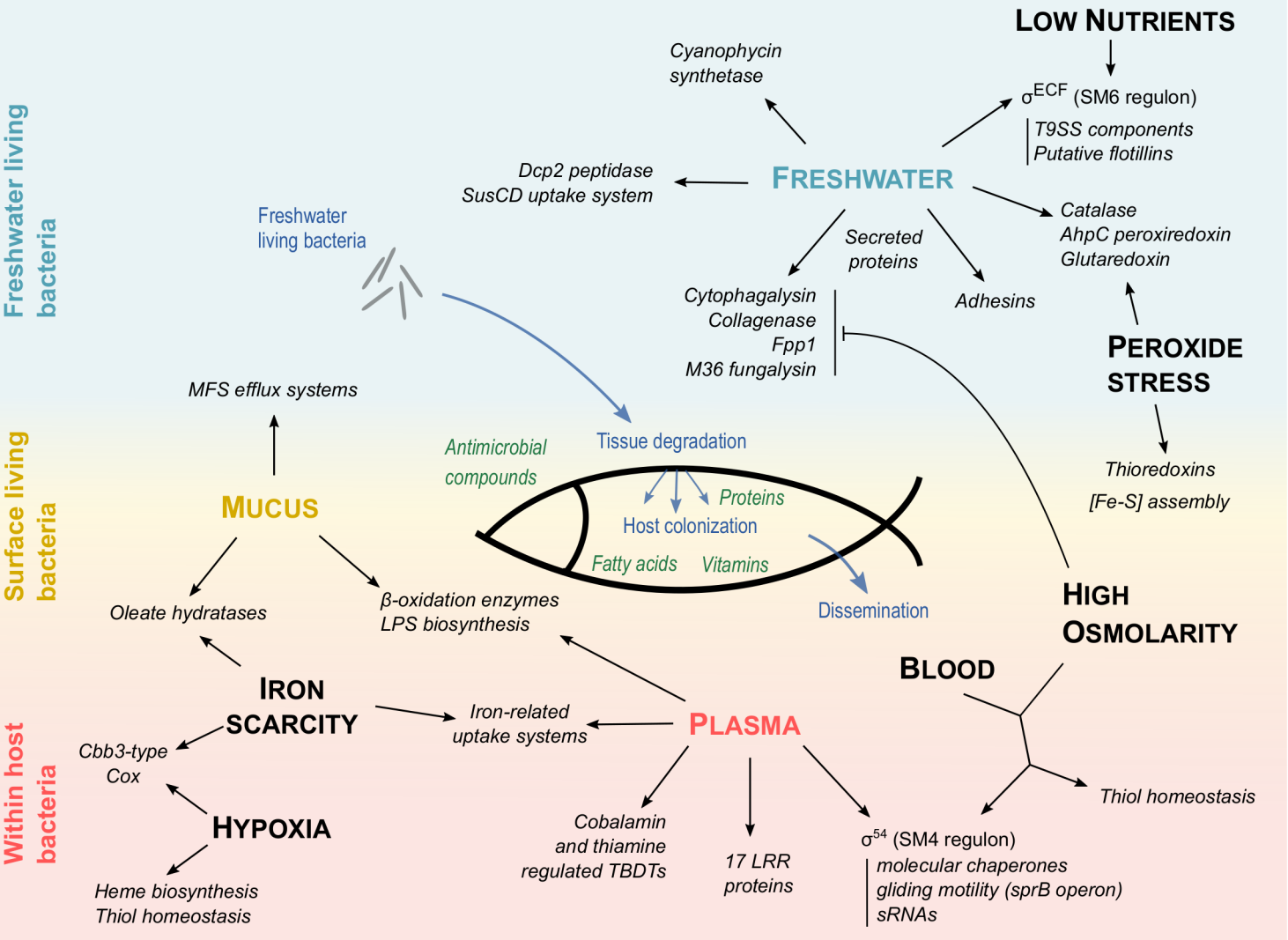


fpeb, a database allowing interactive exploration of *F. psychrophilum* expression data (https://fpeb.migale.inrae.fr)


A
-50
-40
-30
-20
-10
TSS
σA
SM1
808 TSS
SM2
177 TSSs
SM3
110 TSSs
σ54
SM4
47 TSSs
SM5
33 TSSs
σECF
SM6
19 TSSs
average expression level downstream of the corresponding TSSs (log2-scale)
T1
T2
T3
T4
T5
T2-FCS
TYESA-FCS
TYESA
TYESA-0.5
TYESA-0.2
CS3
Temp10C
Temp4C
DIP
NaCl
CS1
HP
LYSO
CS2
AMOX
Anaer
LAB
Water
FishWater
CS4
HCl
EP
MUC
EP2
Plasma
TYESA-Blood
Blood
B
SM4
SM5
ATG
249 bps
93 bps
65 bps
404 bps
PremF
mCherry
pCPGmr-PremF-mCh
SM4 TTGGCA--------GT----
PremF TTGGCACGATTTTAGTTGTG
sm4* --TT----------TG----
SM5 ----TA--TTGC----
PremF TTAGAAACTTGCAGTT
sm5* --------GGTA----
PremF
sm4*
sm5*
sm4*sm5*
Pless
0
1000
2000
3000
Fluorescence/OD600


Rfp14
Rfp73
Rfp70
Rfp68
Rfp69
Rfp3
Rfp6
Rfp8
Rfp46
Rfp13
Rfp15
Rfp61
Rfp77
Rfp54
Rfp18
Rfp82
Rfp35
Rfp44
Rfp47
Rfp52
Rfp85
Rfp20
Rfp67
Rfp11
Rfp25
Rfp74
Rfp7
Rfp4
Rfp66
Rfp31
Rfp81
Rfp10
Rfp64
Rfp40
Rfp23
Rfp39
Rfp36
Rfp63
Rfp72
Rfp37
Rfp22
Rfp53
Rfp30
Rfp29
Rfp21
Rfp76
Rfp62
Rfp51
Rfp50
0
200
400
600
800
1000
length (bp)
−6
−4
−2
0
2
MFE Z−score
F. psychrophilum THCO290
F. psychrophilum NO083
F. psychrophilum JIP0286
F. psychrophilum CSF25993
F. psychrophilum 95010611
F. psychrophilum FRGDSA188211
F. psychrophilum DK002
F. psychrophilum IT09
F. psychrophilum LM02Fp
F. psychrophilum NO004
F. psychrophilum JIP1600
F. psychrophilum KU0606264
F. psychrophilum LVDJXP189
F. terrae DSM18829
F. columnare C2
F. aquatile LMG4008
F. terrigena DSM17934
F. xueshanense CGMCC19227
F. saccharophilum DSM1811
F. noncentrifugens CGMCC110076
F. degerlachei DSM15718
F. johnsoniae UW101ATCC17061
F. succinicans LMG10402
F. gilvum EM1308
F. frigoris DSM15719
F. fluvi DSM19978
F. granuli DSM19729
F. antarcticum DSM19726
F. gelidilacus DSM15343
F. daejeonense DSM17708
F. sinopsychrotolerans CGMCC18704
F. soli DSM19725
F. sasangense DSM21067
F. hydatis DSM2063
F. columnare ATCC49512
F. branchiophilum FL15
F. indicum GPTSA1009
F. seoulense EM1321
F. urocaniciphilum DSM27078
F. suncheonense DSM17707
F. cauense R2A7
F. haoranii DSM22807
F. crassostreae LPB0076
F. columnare CIP109753
F. enshiense DK69
F. caeni CGMCC17031
F. fontis DSM25660
F. rivuli DSM21788
F. subsaxonicum DSM21790
F. beibuense F448
F. filum DSM17961
F. aklanivivens IK1
F. marinum CGMCC110825
F. ummariense DS12
T. maritimum NCIMB2154T
C. algicola DSM14237
Z. galactanivorans DsiJT
P. irgensii 23P
E. meningoseptica ATCC13253
C. canimorsus Cc5
C. hutchinsonii ATCC33406
C. indologenes AA5
B. thetaiotaomicron VPI5482
P. gingivalis ATCC33277
Species
Genus
Family
Phylum
RNA features from RFAM database
Rfp18
Number of hits per genome:
blue, 1 hit; orange, 2 hits; red, >2 hits;
gray, partial hit.


fpp1
fpp2
THC0290_0300
THC0290_1090
THC0920_2338
RQ relative to WT OD600=0.5
OD600
0.5
1.2
2
WT
Δrfp18
*


Freshwater living bacteria
Surface living bacteria
Within host bacteria
LOW NUTRIENTS
Cyanophycin synthetase
σECF (SM6 regulon)
T9SS components
Putative flotillins
Dcp2 peptidase
SusCD uptake system
FRESHWATER
Catalase
AhpC peroxiredoxin
Glutaredoxin
Freshwater living bacteria
Secreted proteins
Cytophagalysin
Collagenase
Fpp1
M36 fungalysin
Adhesins
PEROXIDE STRESS
Thioredoxins
[Fe-S] assembly
MFS efflux systems
Antimicrobial compounds
Tissue degradation
MUCUS
Proteins
Host colonization
Fatty acids
Vitamins
Dissemination
Oleate hydratases
β-oxidation enzymes
LPS biosynthesis
HIGH OSMOLARITY
BLOOD
IRON SCARCITY
Iron-related uptake systems
PLASMA
Cbb3-type Cox
Thiol homeostasis
HYPOXIA
Cobalamin and thiamine regulated TBDTs
17 LRR proteins
σ54 (SM4 regulon)
molecular chaperones
gliding motility (sprB operon)
sRNAs
Heme biosynthesis
Thiol homeostasis