

# Comparative Genomic Analysis Indicates that Niche Adaptation of Terrestrial *Flavobacteria* Is Strongly Linked to Plant Glycan Metabolism

Max Kolton<sup>1,3</sup>, Noa Sela<sup>2</sup>, Yigal Elad<sup>2</sup>, Eddie Cytryn<sup>1\*</sup>

**1** Institute of Soil, Water and Environmental Sciences, the Volcani Center, Agricultural Research Organization, Bet Dagan, Israel, **2** Department of Plant Pathology and Weed Research, the Volcani Center, Agricultural Research Organization, Bet Dagan, Israel, **3** Institute of Plant Sciences and Genetics in Agriculture, The Robert H. Smith Faculty of Agriculture, Food and Environment, the Hebrew University of Jerusalem, Rehovot, Israel

## Abstract

*Flavobacteria* are important members of aquatic and terrestrial bacterial communities, displaying extreme variations in lifestyle, geographical distribution and genome size. They are ubiquitous in soil, but are often strongly enriched in the rhizosphere and phyllosphere of plants. In this study, we compared the genome of a root-associated *Flavobacterium* that we recently isolated, physiologically characterized and sequenced, to 14 additional *Flavobacterium* genomes, in order to pinpoint characteristics associated with its high abundance in the rhizosphere. Interestingly, flavobacterial genomes vary in size by approximately two-fold, with terrestrial isolates having predominantly larger genomes than those from aquatic environments. Comparative functional gene analysis revealed that terrestrial and aquatic *Flavobacteria* generally segregated into two distinct clades. Members of the aquatic clade had a higher ratio of peptide and protein utilization genes, whereas members of the terrestrial clade were characterized by a significantly higher abundance and diversity of genes involved in metabolism of carbohydrates such as xylose, arabinose and pectin. Interestingly, genes encoding glycoside hydrolase (GH) families GH78 and GH106, responsible for rhamnogalacturonan utilization (exclusively associated with terrestrial plant hemicelluloses), were only present in terrestrial clade genomes, suggesting adaptation of the terrestrial strains to plant-related carbohydrate metabolism. The Peptidase/GH ratio of aquatic clade *Flavobacteria* was significantly higher than that of terrestrial strains ( $1.7 \pm 0.7$  and  $9.7 \pm 4.7$ , respectively), supporting the concept that this relation can be used to infer *Flavobacterium* lifestyles. Collectively, our research suggests that terrestrial *Flavobacteria* are highly adapted to plant carbohydrate metabolism, which appears to be a key to their profusion in plant environments.

**Citation:** Kolton M, Sela N, Elad Y, Cytryn E (2013) Comparative Genomic Analysis Indicates that Niche Adaptation of Terrestrial *Flavobacteria* Is Strongly Linked to Plant Glycan Metabolism. PLoS ONE 8(9): e76704. doi:10.1371/journal.pone.0076704

**Editor:** D. Ashley Robinson, University of Mississippi Medical Center, United States of America

**Received:** June 6, 2013; **Accepted:** August 26, 2013; **Published:** September 26, 2013

**Copyright:** © 2013 Kolton et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This research was supported by The Chief Scientist of the Ministry of Agriculture and Rural Development of Israel, project Number 301-0693-10. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

\* E-mail: Eddie@volcani.agri.gov.il

## Introduction

The Gram negative genus *Flavobacterium* (phylum *Bacteroidetes*) encompasses over 100 different strains that are characterized by exceptional environmental niche heterogeneity ranging from Arctic lakes to thermal springs [1]. They are important members of the bacterial community in both aquatic and terrestrial environments, where their relative abundance can in some cases reach up to 20% [2–8]. Several *Flavobacterium* species are the causative agent of severe plant diseases [9,10], while others are associated with plant protection and growth promotion [11,12] and with bioremediation in soils and marine sediments [13,14].

A myriad of recent studies strongly suggest that flavobacterial abundance is significantly enhanced in the

rhizosphere and phyllosphere of terrestrial plants [2,3,7]. For example, Bulgarelli et al., (2012) showed that the roots of *A. thaliana*, grown in different natural soils under controlled environmental conditions, are preferentially colonized by *Flavobacteria*, where the relative abundance on roots (5 to 15% of total defined families) was between 20 to 100-fold higher than in bulk and rhizosphere soil. This was also supported by a recent study by Bodenhausen et al., (2013) that determined that *Flavobacterium* were among the most dominant genera on both roots and leaves (4 to 10%) of wild *A. thaliana*.

Adaptation of *Flavobacterium* to such a wide array of heterogeneous environments is undoubtedly associated with genetic plasticity. Thus, in order to identify functional mechanisms associated with the enhanced rhizosphere and

**Table 1.** *Flavobacterium* spp. analyzed in this study.

Strain*	Genome size (Mbp)	GC content (%)	GenBank accession number	Isolation site	References
<i>Flavobacterium johnsoniae</i> UW101 (F)	6.1	34	NC009441	Soil; UK	[63]
<i>Flavobacterium</i> sp. F52 (D)	5.34	34	NZAKZQ00000000	Rhizosphere of pepper; Israel	[15]
<i>Flavobacterium</i> sp. URHB0058 (D)	5.26	34	AUEU00000000	Forest soil; North America	DOE Joint Genome Institute
<i>Flavobacterium</i> sp. WG21 (D)	5.19	36	NZMYW00000000	Wintergreen Lake; USA	University of Notre Dame
<i>Flavobacterium</i> sp. CF136 (D)	5.12	34	NZAKJZ00000000	Tree rhizosphere; USA	[64]
<i>Flavobacterium rivuli</i> DSM 21788 (D)	4.49	40	NZKB899988	Hard water stream; Germany	DOE Joint Genome Institute
<i>Flavobacterium</i> sp. B17 (D)	4.17	37	BACY00000000	Rice shoot; Japan	NCBI
<i>Flavobacterium soli</i> DSM 19725 (D)	4.0	36	AUGO00000000	Soil samples; South Korea	DOE Joint Genome Institute
<i>Flavobacterium</i> sp. ACAM 123 (D)	3.96	35	NZAJXL01000285	Burton Lake; Antarctic lake	NCBI
<i>Flavobacterium frigidum</i> PS1 (D)	3.93	34	NZAHKF00000000	Sea ice, diatom layer; Antarctic lake	[65]
<i>Flavobacterium branchiophilum</i> FL-15 (F)	3.56	33	NC016001	Diseased sheafish; Hungary	[66]
<i>Flavobacterium columnare</i> ATCC 49512 (F)	3.16	31	NC016510	Skin lesion of trout fry; France	[67]
<i>Flavobacterium antarcticum</i> DSM 19726 (D)	3.08	35	NR042998	Soil sample; Antarctica	DOE Joint Genome Institute
<i>Flavobacterium indicum</i> GPTSA100-9 (F)	2.99	31.5	NC017025	Warm spring water; India	[68]
<i>Flavobacterium psychrophilum</i> JIP02/86 (F)	2.86	33	NC009613	Kidney of rainbow trout; France	[69]

(\*) F – full completed genome; D – draft genome.

doi: 10.1371/journal.pone.0076704.t001

phyllosphere competence of *Flavobacterium*, we conducted comprehensive comparative genomic analysis of a model root-associated *Flavobacterium* strain that we recently sequenced [15] and fourteen additional *Flavobacterium* genomes isolated from plant roots, phyllosphere, soil, marine and freshwater ecosystems and fish tissue. Although some genes were ubiquitous to all flavobacteria, a myriad of functional genes were unique to terrestrial environments. These included a diverse group of genes encoding plant polysaccharide degrading enzymes, which appear to be central to plant-associated flavobacterial metabolism.

## Materials and Methods

### *Flavobacteria* isolation and characterization

Mature bell pepper (*Capsicum annuum* L. cv. Maccabi) plants were grown in potting mixtures in an experimental pest- and disease-free greenhouse as previously described [6,16,17]. Plant roots were washed briefly to remove lightly adhering soil and homogenized in sterile saline (0.85% NaCl), and homogenate 10-fold serial dilutions were plated on Casitone-Yeast Extract medium (CYE) [18]. Isolated *Flavobacterium* spp. were tested for gliding motility, presence of flexirubin, hydrolytic activities: protease, lipase, chitinase, amylase, cellulase,  $\beta$ -galactosidase, nannanase, xylanase, urease, and expression of plant growth promotion and protection features: ACC-deaminase,  $\text{NH}_3$  and hydrogen cyanide (HCN) production, using previously described standard procedures [19-24]. One of the isolates, *Flavobacterium* sp.

F52, selected based on its enhanced biocontrol capacity, was sequenced and used for further analyses as described below.

### Analyzed strains

Comparative analysis was performed on 15 publically-available flavobacterial genomes from both aquatic and terrestrial environments including *Flavobacterium* sp. F52. Their names and GenBank accession numbers (<http://www.ncbi.nlm.nih.gov/Genbank>) are presented in Table 1. Additionally, the *Bacteroidetes* species *Porphyromonas gingivalis* W83 (NC002950.2) was used as an out-group genome for phylogenetic and metabolic potential analyses.

### Phylogenetic analyses

Phylogenetic relationships among all sequenced *Flavobacterium* species were investigated based on 16S rRNA gene sequences and concatenated multilocus sequence analysis (MLSA) of 10 highly conserved housekeeping genes: *serS*, *aroE*, *atpD*, *dnaE*, *guaA*, *gyrB*, *mutL*, *pyrC*, *recA*, and *rpoB*, as previously described [25,26]. Additionally, we applied an e-value-based clustering algorithm from reciprocal all-by-all BLASTP analysis of all coding sequences of the investigated genomes using the Hal automated pipeline [27]. Phylogenetic trees illustrating the relationships of 227 common concatenated core genes from the 15 flavobacterial genomes and the outlier *P. gingivalis* W83 were constructed with the Phylogenetic Maximum Likelihood (PhyML) method, using the PhyML 3.0 algorithm with 100 bootstrap replicates and the LG matrix of amino-acid substitution model (<http://www.atgc-montpellier.fr/phyml/>; [28,29]). Sequences of the 16S rRNA genes and

concatenated housekeeping genes were aligned with MUSCLE [30]. To assess the robustness of the phylogenetic tree topology, bootstrap consensus trees were inferred from 1000 replicates. Maximum likelihood phylogenetic trees were generated and/or visualized by using MEGA 5.1 software with default parameters [31].

### Flavobacterium core and niche-specific genomes

Orthologs were defined by identifying unique pairwise reciprocal best hits, with at least 60% similarity in amino acid sequence and less than 30% difference in protein length. To define the flavobacterial core genome, we used the 15 available genomes described above. For niche-specific genome identification we used the completed genomes: *F. branchiophilum* FL-15; *F. psychrophilum* JIP02/86; *F. columnare* ATCC 49512 and *F. indicum* GPTSA100-9 as representatives of the aquatic clade; and the completed genome of *F. johnsoniae* UW101 with two draft genomes *Flavobacterium* sp. F52 and *Flavobacterium* sp. CF136 as representatives of the terrestrial clade. To determine the functional roles of the core and niche-specific genes, open reading frames (ORFs) were assigned to gene ontology (GO) groups using Blast2GO software, and tested for enrichment using Fisher's exact test [32]. Only statistically different GO's were included in future analyses. These GO's were manually grouped to metabolic pathways and presented as a ratio between inspected set proteins to the reference set. Thus a ratio above or below one indicates over or under representation, respectively of specific group in an inspected core genome.

### Whole-genome functional gene comparison

The fifteen analyzed *Flavobacterium* genomes were annotated and distributed to SEED categories and subsystems using the RAST server (<http://RAST.nmpdr.org> [33]). The numbers of annotated and distributed genes across SEED categories and subsystems were normalized to the total number of genes in an individual genome. A Bray-Curtis distance matrix encompassing the relative abundances of each defined SEED categories or subsystem in the fifteen genomes was used for cluster analysis in order to assess the functional gene similarity of the flavobacterial strains. Statistically significant differences in allocation of genome resources to SEED categories and subsystems between bacterial clades was indicated by Student T-Test ( $P < 0.05$ ) using the SPSS 18 software package (SPSS inc., Chicago, Illinois). Bray-Curtis distance matrix calculation and cluster analysis with 1000 bootstrap replicates was performed by using the PAleo-ecology STatistics freeware (PAST) package, version 2.03 (<http://folk.uio.no/ohammer/past/> [34]), and the representation of the generated clusters was performed with MEGA 5.1 software [31].

### Prediction of carbohydrate and peptide utilization

Sequence similarity based annotation was performed to identify genes with functional motifs encoding carbohydrate-binding and metabolic enzymes. Predictions were automatically determined using the CAZymes Analysis Toolkit ([\[mothra.ornl.gov/cgi-bin/cat.cgi\]\(http://mothra.ornl.gov/cgi-bin/cat.cgi\)\) applying the association rule-learning algorithm \[35\] and the dbCAN prediction web server \(<http://csbl.bmb.uga.edu/dbCAN/> \[36\]\). Sequences with similar predictions based on both methods were selected and named according to CAZy classifications \[37\], and the abundance and diversity of specific carbohydrate-associated enzymes in the analyzed genomes was determined. Peptidase \(protease\) encoding genes in the flavobacterial genomes were predicted according to the MEROPS classification \(<http://merops.sanger.ac.uk/cgi-bin/blast> \[38\]\). Shannon diversity indices for CAZy classes were calculated with PAST version 2.03 \[34\] and statistical significant differences between bacterial clades was indicated by Student T-Test \( \$P < 0.05\$ \) using the SPSS 18 software package \(SPSS inc., Chicago, Illinois\).](http://</a></p>
</div>
<div data-bbox=)

## Results

### Isolation and characterization of root-associated flavobacteria

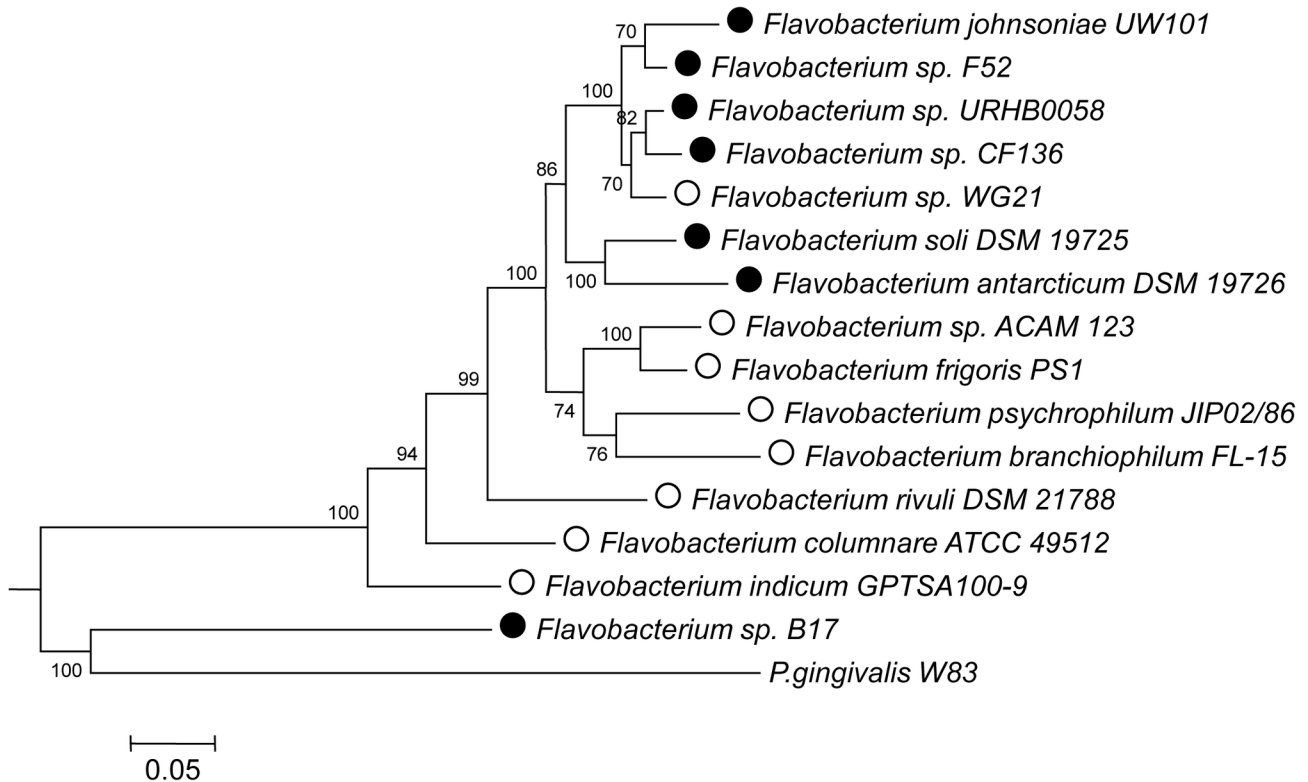
Twenty five flavobacterial strains were isolated from bell pepper roots as described above. Physiological analyses revealed that almost all of these isolates can successfully propagate at a wide range of temperatures (10-40°C) and pH values (4-9) (Table S1). Moreover characterization of enzymatic hydrolytic activities of this isolates revealed their high potential to utilize plant related polysaccharides such as mannan, xylan and cellulose (Table S1). In addition, all of the isolated root-associated flavobacteria produced ammonia and cyanide which are known players in plant-bacterial interaction (Table S1). Integrative analyses of these traits led us to select *Flavobacterium* sp. F52 as a model strain for studying *Flavobacterium*-root interactions.

### Genomic features of Flavobacterium strains

A summary of the features of each of the 15 flavobacteria genomes used for the comparative analyses conducted in this study is provided in Table 1. The flavobacterial genomes vary in size by approximately two-fold (ranging from 2.9 to 6.1 Mbp) with the number of gene coding sequences (CDSs) ranging from 2,432 to 5,017. The genomes of *Flavobacterium* strains isolated from temperate soil and rhizosphere environments were predominantly larger than those from aquatic ones, with the exception of *F. antarcticum*, *F. rivuli* and *Flavobacterium* sp. WG21 (Table 1, Figure S1).

### Phylogenetic analysis

We estimated flavobacterial phylogenetic relationships by constructing maximum likelihood trees based on 16S rRNA genes (Figure S2), MLSA (Figure 1) and whole-genome comparisons of common proteins (Figure S3). MLSA and 16S rRNA gene-based analyses failed to show a clear-cut environmental-based distribution of flavobacteria (Figure 1; Figure S2), although MLSA-based phylogenetic analysis revealed that most of soil and plant-associated strains formed a strongly related paraphyletic sub-cluster (Figure 1).



**Figure 1. Phylogenetic relationships of MLSA genes in *Flavobacterium* strains.** Maximum Likelihood analysis based on concatenated alignments of ten highly conserved housekeeping genes: *serS*, *aroE*, *atpD*, *dnaE*, *guaA*, *gyrB*, *mutL*, *pyrC*, *recA*, and *rpoB*. Bootstrap values are shown next to the branch nodes. Open and black circles represent aquatic and terrestrial isolates, respectively.

doi: 10.1371/journal.pone.0076704.g001

### Functional characterization of flavobacterial isolates

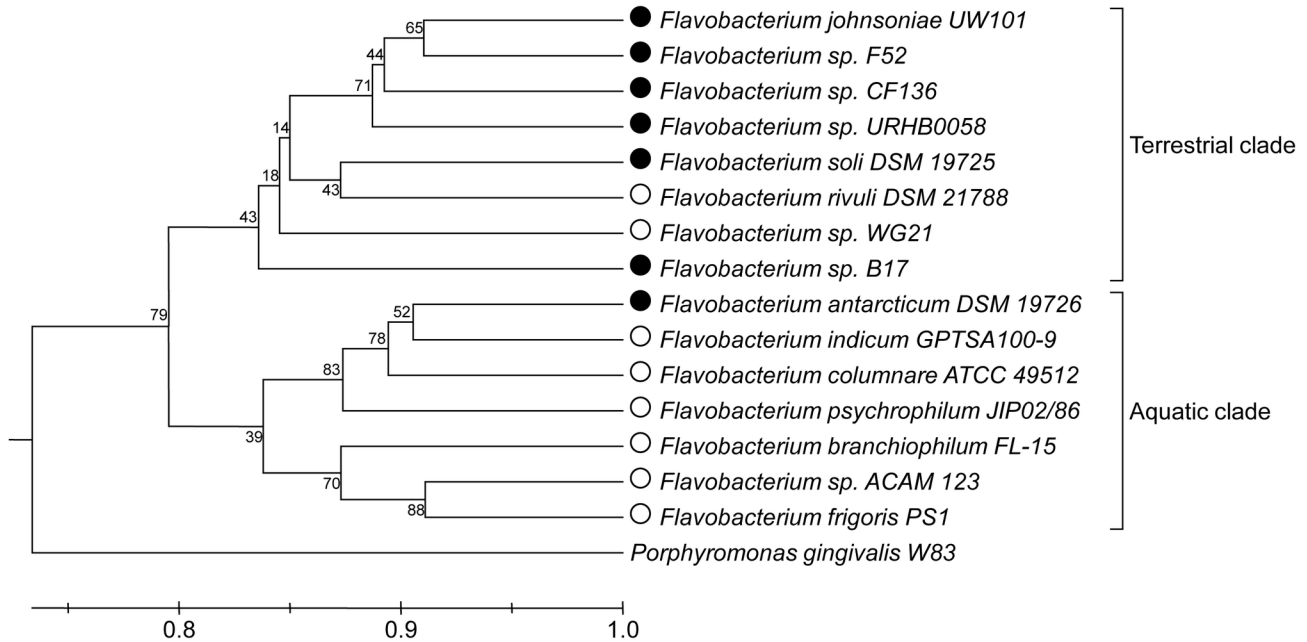
Analysis of protein distribution of *Flavobacterium* strains based on functional (SEED) subgroups generated two distinct clusters: a mostly terrestrial clade consisting of temperate soil and plant-associated flavobacteria, and a predominantly aquatic clade (Figure 2). The exceptions to this division were the aquatic strains *Flavobacterium* sp. WG21 and *F. rivuli*, who functionally clustered with terrestrial isolates; and the soil isolates *F. antarcticum* and *F. soli*, which were associated with the aquatic clade (Figure 2).

In order to identify proteins potentially associated with plant-flavobacterial interactions, we focused on SEED categories that were substantially over-represented among members of the terrestrial clade. These included genes associated with cell signaling pathways and genes encoding for carbohydrate metabolism (Figure 3A) as well as genes associated with Ton and Tol transport systems; metal and drug resistance; and denitrification. Additionally, our analysis revealed that SEED sub-systems associated with metabolism of plant related carbohydrates such as xylose, arabinose and pectin were substantially over-represented in the terrestrial clade (Figure 3B). On the other hand, the aquatic clade was characterized by higher representation of genes associated with cell wall and

capsule synthesis, protein metabolism (Figure 3A) and genes associated with phosphate transport.

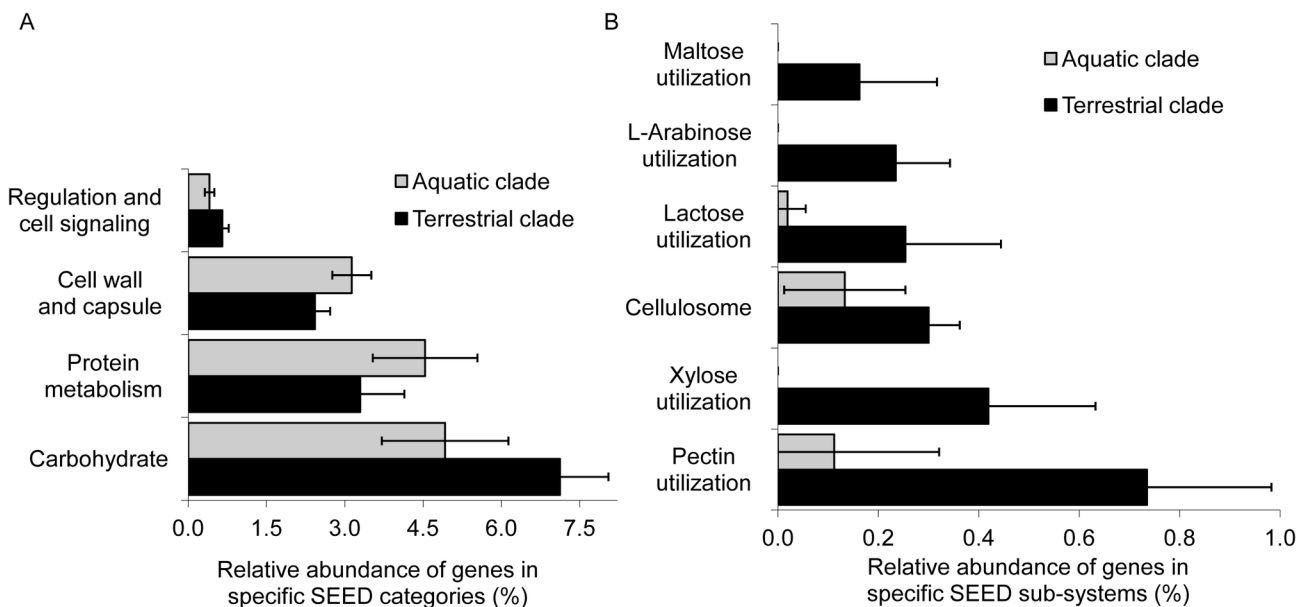
### *Flavobacterium* core and niche-specific genomes

*Flavobacterium* core and aquatic and terrestrial niche-specific genomes were generated using the methodology described in the materials and methods section, with orthologous genes being characterized as sharing at least 60% similarity in amino acid sequence and less than 30% difference in protein length. The *Flavobacterium* core genome consisted of only 227 predicted proteins, representing  $7.3 \pm 1.3$  and  $5.2 \pm 0.5\%$  of all predicted proteins across aquatic and terrestrial *Flavobacterium* clades, respectively. In contrast, 798 orthologous ORFs were terrestrial niche-specific, while only 32 unique ORFs were aquatic niche-specific (Table S2). Interestingly, 56% (18/32) of the aquatic niche-specific and 39% (315/798) of the terrestrial niche-specific ORFs encoded for hypothetical proteins, whereas only 4% (9/227) of the *Flavobacterium* core genome ORFs were hypothetical (Table S2). The *Flavobacterium* core genome was primarily associated with basic housekeeping functions such as nucleic acid metabolism, protein synthesis, glycolysis related proteins and amino acid metabolism, with underrepresentation of genes encoding bacterial cell wall modification and transport (Figure S4A). Conversely, the terrestrial-specific genome was



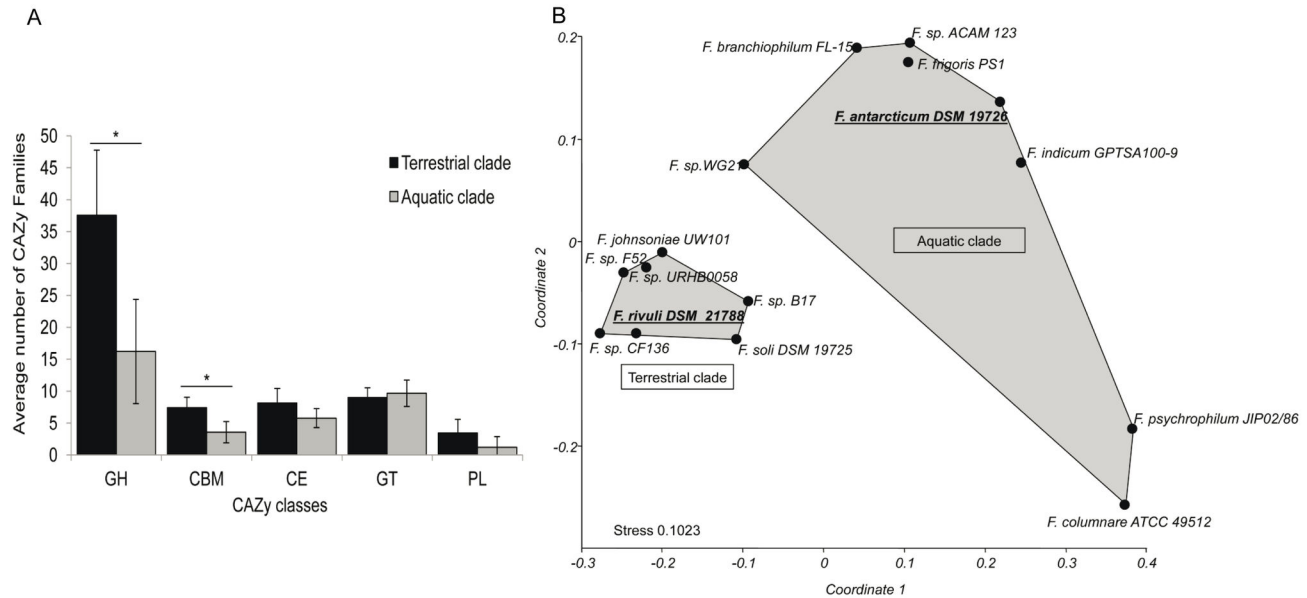
**Figure 2. Hierarchical clustering of *Flavobacterium* strains based on functional similarity.** *Flavobacterium* strains functionally clustered into terrestrial and aquatic clades. The distribution of genes across SEED subsystems was normalized to the total number of genes in a particular genome. Hierarchical clustering was performed using the Bray-Curtis distance matrix. Bootstrap values are shown next to the branch nodes. Open and black circles represent aquatic and terrestrial isolates, respectively.

doi: 10.1371/journal.pone.0076704.g002



**Figure 3. Statistically significant differences in distribution of SEED subsystems between aquatic and terrestrial clades.** (A) Major SEED categories, and (B) carbohydrate sub-systems, showing significant differences between aquatic and terrestrial clades. The distribution of genes across SEED categories and subsystems were normalized to the total number of genes in a particular genome and Student T-test ( $P < 0.05$ ) was applied to test statistical significance.

doi: 10.1371/journal.pone.0076704.g003



**Figure 4. Distribution of carbohydrate-utilizing enzymes.** (A) Distribution of CAZy enzyme classes in the terrestrial and aquatic flavobacterial clades; and (B) Distribution of *Flavobacterium* strain glycoside hydrolases visualized using nMDS of a Bray-Curtis distance matrix. GH- glycoside hydrolases; CE-carbohydrate esterase; PL-polysaccharide lyases; CBM- carbohydrate binding modules; GT- glycosyl transferases. Asterisks indicate statistically significant differences ( $P < 0.05$ ) between clades.

doi: 10.1371/journal.pone.0076704.g004

significantly enriched with genes related to hydrolysis of glycosidic bonds, nitrogen metabolism and cell signaling and communication, with underrepresentation of CDSs involved in protein biosynthesis and general cellular metabolic processes (Figure S4B).

### Carbohydrate and peptide utilization

The SEED analyses described above, stimulated more comprehensive investigation of genes involved in carbohydrate metabolism. Abundance and diversity of Carbohydrate-Active EnZyme database (CAZy)-characterized enzymes were found to be significantly higher in the terrestrial clade genomes than in the aquatic clade genomes (Figure 4A, Table S3). Specifically, the terrestrial clade genomes contained significantly higher variation and abundance of glycoside hydrolase (GH) and Carbohydrate-Binding Module (CBM) families (Figure 4A). Bray-Curtis based nMDS ordination of GH families clearly indicated environmental niche differentiation of bacteria belonging to the two clades (Figure 4B). Two exceptions were *F. antarcticum* which was associated with the aquatic clade, and the stream isolate *F. rivuli*, which was strongly affiliated with the terrestrial clade (Figure 4B). The terrestrial niche-specific genome contained unique CAZy enzymes involved in metabolism of glucans, arabinose and rhamnagalacturonan that are exclusively associated with terrestrial plant hemicellulose (Table 2, Table S3).

While the terrestrial clade strains had significantly higher relative abundances of genes encoding GH enzymes, the aquatic strains appeared to be more tailored for peptide utilization (Figure 5A), supporting previous findings by

Fernández-Gómez et al. (2013). This is strongly visualized by the peptidase/GH ratio, which averaged  $9.7 \pm 4.7$  in the aquatic clade vs.  $1.7 \pm 0.7$  in the terrestrial clade. The fish pathogen *F. psychrophilum* had the highest ratio (16.7), whereas the soil and rhizosphere isolates *F. johnsoniae* and *Flavobacterium* sp. CF136 had the lowest ratios (1.1) (Figure 5B).

### Discussion

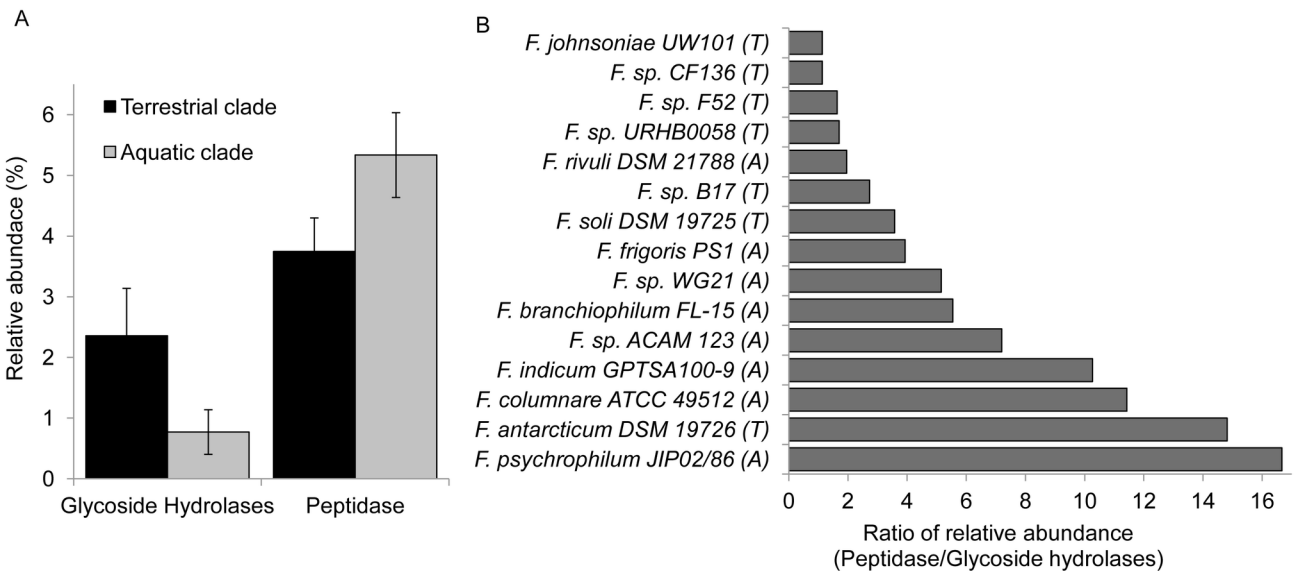
A major challenge of microbial ecology is to determine the functional capacity of dominant organisms in specific environmental niches. *Flavobacteria* are ubiquitous in a wide array of terrestrial and aquatic environments where they are believed to play a pivotal role in mineralization of poorly degradable macromolecules and thus serve as carbon flux regulators in these ecosystems [39]. A myriad of recent molecular and culture-based data has strongly indicated that flavobacteria are highly enriched in the phyllosphere and rhizoplane of an array of plants, where they may represent over 20% of the bacterial community [2-4,6-8]; suggesting that they may play a pivotal role in these environments. However, we know practically nothing about the genetic characteristics that are responsible for their capacity to thrive in these environments. In order to identify specific functional properties that are required for flavobacterial adaptation to plant interfaces, we compared the recently sequenced root-associated *Flavobacterium* sp. F52 genome [15] to 14 additional *Flavobacterium* genomes from a wide range of terrestrial and aquatic ecosystems (Table 1). Phylogenetic and functional gene-based analyses of flavobacterial genomes

**Table 2.** Terrestrial clade-unique CAZy domains.

Strain	Rhamnogalacturonan			Xylan		Arabinose			Mannose		Fucose	Pectin	Chitin			
	GH 28	GH 78	GH 106	GH 10	GH 115	CE 2	CBM 6	GH 27	GH 43	GH 51	GH 1	GH 130	GH 95	PL 1	GH 18	GH 89
<i>Flavobacterium johnsoniae</i>	7	3	2	3	1	1	4	2	10	2	1	5	2	7	4	1
<i>Flavobacterium</i> sp. F52	6	3	2	1	1	1	6	2	13	2	2	5	2	6	3	1
<i>Flavobacterium</i> sp. CF136	10	4	2	3	1	1	3	4	25	6	1	4	5	11	3	1
<i>Flavobacterium</i> sp. URHB0058	11	3	2	1	3	1	3	1	10	3	1	4	2	5	3	1
<i>Flavobacterium rivuli</i>	4	1	2	0	0	1	1	2	13	3	2	2	3	1	0	0
<i>Flavobacterium</i> sp. B17	0	0	0	1	2	0	4	1	8	1	0	1	1	0	1	1
<i>Flavobacterium soli</i>	1	0	0	0	3	1	2	1	7	0	1	2	0	1	0	0
<i>Flavobacterium</i> sp. WG21	3	0	0	0	0	0	1	0	1	0	1	0	0	2	3	0
<i>Flavobacterium antarcticum</i>	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
<i>Flavobacterium branchiophilum</i>	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Flavobacterium columnare</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Flavobacterium indicum</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Flavobacterium psychrophilum</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Flavobacterium frigoris</i> PS1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Flavobacterium</i> sp. ACAM 123	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0

GH- glycoside hydrolases; CE-carbohydrate esterase; PL-polysaccharide lyases; CBM- carbohydrate binding modules; GT- glycosyl transferases.

doi: 10.1371/journal.pone.0076704.t002



**Figure 5. Relative abundances of hydrolytic enzymes.** (A) Relative abundances of peptidases and glycoside hydrolases in terrestrial and aquatic flavobacterial clades; (B) Peptidase/GH ratios of *Flavobacterium* strains analyzed in this study; letters A and T in the parentheses indicates aquatic or terrestrial origin of strains, respectively. Functional predicted peptidases and glycoside hydrolases were normalized to the total number of genes in a particular genome.

doi: 10.1371/journal.pone.0076704.g005

generated two clusters comprised primarily of aquatic and terrestrial strains, respectively (Figure 2).

As discussed above, comparative analysis of the two clades revealed significant differences in the allocation of carbohydrate-utilizing genes (Figure 3A), and highly significant

overrepresentation of genes that degrade plant related carbohydrates such as xylose, arabinose and pectin in the terrestrial clade (Figure 3B). Furthermore, the terrestrial clade-affiliated flavobacteria harbored a wide array of unique CAZy enzymes involved in the metabolism of glucans containing



arabinose (GH27 and GH43) and rhamnogalacturonan (GH78 and GH106) that are exclusively associated with terrestrial plant hemicellulose (Table 2) [40,41]. These GH's were absent in all of the analyzed aquatic strains (exceptions discussed below), potentially indicating co-evolution of terrestrial flavobacterial plant-associated glycoylucanases with terrestrial plants. It was previously suggested that the scope of GH's in a particular microorganism and their capacity to utilize particular carbon sources, is dictated by the ecological niche that they inhabit [42]. For instance, *Bacteroidetes* from phytoplankton-rich marine environments have higher potential for degradation of sulfated algae-derived polysaccharides like fucoidan and (xylo)mannan than their relatives from the oligotrophic areas, whose metabolic potential appeared to be more geared towards peptide utilization and light harvesting as described below [43].

Our results strongly support this notion, and we postulate that GHs may be accurate genomic predictors for identification of plant-associated flavobacteria. For example, SEED-defined functional cluster analyses (Figure 2), as well as detailed analysis of the flavobacterial GH distributions (Figure 4B, Table 2, Table S3), specifically designated the stream isolate *F. rivuli* [44] as a member of the terrestrial clade, contrary to its apparent aquatic origin. This was supported by the presence of glycoside hydrolases GH78 and GH106 (Table 2, Table S3), which are responsible for rhamnogalacturonan cleavage which is incorporated mainly into cell walls of all land plants [40,41]. The notion that *F. rivuli* is in essence soil-derived is supported by a previous microbial fingerprinting-based study that analyzed stream and soil flavobacterial communities from in and around the stream from which *F. rivuli* was isolated [45]. Based on these analyses the authors inferred that the flavobacterial stream community, originated from soil washout [45]. Conversely, functional and GH clustering (Figure 2, Figure 4B) placed *F. antarcticum*, originally isolated from an Antarctic soil [46], within the aquatic clade, which is supported by the absence of GH78 and GH106 in this bacterium (Table 2, Table S3). We hypothesize that this stems from the absence of terrestrial plants in the isolated environment, emphasizing the notion that terrestrial plant glycans dictate the arsenal of GH's in the terrestrial flavobacterial clade.

Almost all *Flavobacterium* GH's are clustered together with TonB-dependent receptors and transducers in small genomic islands, suggesting an integrated regulation of adhesion and degradation of polymers [47]. There is significant evidence that extracellular chitinase genes, exclusive to the terrestrial flavobacterial genomes are the product of horizontal gene transfer, most likely from *Firmicutes* phylum (data not shown). In contrast, based on BLASTP (<http://blast.ncbi.nlm.nih.gov>) similarities, most of the soil-unique GH's are highly conserved within the *Bacteroidetes* phylum indicating that these genes were lost by the aquatic strains. Interestingly, the cluster affiliation of the soil strain *F. soli*, isolated from a recently-formed 0.2 km<sup>2</sup> volcanic island off the coast of South Korea [48], changed depending on the specific genomic element being analyzed. For instance functional gene-based clustering determined that it was affiliated with the aquatic clade (Figure

2), whereas GH-based nMDS analysis placed it strongly within the terrestrial clade (Figure 4B).

In general, the genomes of the terrestrial clade were approximately 40 percent larger than the aquatic clade genomes (Table 1, Figure S1), with the exceptions discussed above behaving according to their corresponding clades for this parameter as well. Bacteria genome size can be strongly shaped by environmental complexity [49]. For instance free-living bacteria that have more heterogeneous lifestyles generally harbor large genomes, especially in complex environments with strong physico-chemical gradients [50]. Conversely, facultative pathogens generally contain smaller genomes, while, obligatory pathogens are readily characterized by extremely reduced genome size [51]. Whole genome comparison of the flavobacteria genomes coincides with this general rule, with an apparent correlation between genome size and environmental heterogeneity. The evidently free living soil isolate *F. johnsoniae* had the largest genome size followed by the rhizosphere strain *Flavobacterium* sp. F52, whereas the poplar root and rice endophytes *Flavobacterium* sp. CF136 and *Flavobacterium* sp. B17, respectively, had the smallest genomes among the terrestrial plant-associated flavobacteria analyzed in this study. The same tendency was observed within the aquatic clade. The free living strain *Flavobacterium* sp. GW21 had the largest genome, whereas the fish pathogens harbored much smaller genomes (Table 1). Temperate soils are characterized by a highly complex multi-dimensional architecture that is often associated with strong spatial and temporal gradients in a wide array of biotic and a-biotic parameters including: plant biomass water potential, oxygen, nutrients, temperature and pH [52,53]. Therefore, it is perhaps not surprising to find larger genomes with overrepresentation of sensory and signal transduction proteins in the terrestrial clade niche-specific genome (Figure 3A, Figure S4, Table S2).

*Flavobacteria*, the most abundant class of *Bacteroidetes* in marine environments [54], are assumed to specialize in the degradation of polymers and particulate organic matter (POM) [5,54-57]. Amino acids, peptides and proteins appear to be important components of POM, and these components account for up to 15% of total biologically available nitrogen in the ocean [58]. Therefore, given the general limitation of nitrogen in ocean environments, the ability to utilize peptide-incorporated POMs may be beneficial for bacteria propagation. *Flavobacteria* adaptation to oligotrophic condition appears to be primarily associated with enhanced peptide utilization and the capacity to utilize alternative energy sources such as light, which is harvested with proteorhodopsin [59]. Moreover, the synthesis and expression of proteolytic and peptidoglycan degrading enzymes are much higher than in mesotrophic or eutrophic areas [43]. This is demonstrated by the high number of peptidases found in genome sequences of the Proteorhodopsin-containing marine *Flavobacterium* Dokdonia sp. MED134 and *Polaribacter* sp. MED152 isolated from oligotrophic Mediterranean sea [60,61].

We believe that the high peptidase/GH ratio (>6) detected in the aquatic clade genomes, implies that this clade is primarily associated with protein-amino acid metabolism, in contrast to the terrestrial strains that had a much lower peptidase/GH ratio



(~1.5), who seem to depend primarily on plant related polysaccharide degradation. This is supported by our analysis (Figure 5B) and by a recently published comparative genomic studies of *Bacteroidetes*, which indicated that the peptidase/GH ratio in the genomes of aquatic *Bacteroidetes* were above six, while the soil species ratio was approximately one [43,47,62]. The exceptions to this rule are *F. antarcticum*, whose peptidase/GH ratio (15) (Figure 5B) was far higher than other soil isolates and *F. rivuli*, whose peptidase/GH ratio (1.9) was much more similar to terrestrial strains. The *F. antarcticum* aquatic-like peptidase/GH ratio is also supported by its smaller genome size (3.08 Mbp), and we hypothesize that its aquatic-like attributes are most likely due to the lack of terrestrial plant material in the Antarctic soils and from the fact that it potentially originated from the surrounding seawater. Conversely, the terrestrial-like peptidase/GH ratio of *F. rivuli*, its larger genome size (4.49 Mbp) and arsenal of genes encoding plant glycan metabolizing enzymes, strongly indicate that it is of terrestrial origin. The intermediate peptidase/GH ratio (3.6) and genome size (4.0 Mbp) of *F. soli* suggest that it is adapted to aquatic/terrestrial interfaces, or perhaps represents a transitional stage between aquatic and terrestrial lifestyles. This is potentially supported by the fact that it was isolated from a new volcanic island of the coast of Korea, which contains almost no soil and very sparse vegetation. Large-scale genome screening of flavobacterial strains from a myriad of terrestrial and aquatic environments is required to enhance our understanding of these relationships and our comprehension of how flavobacterial genomes evolved to adapt to different environments.

## Supporting Information

### Figure S1. Genome size of aquatic and terrestrial flavobacterial clades.

(TIF)

**Figure S2. Phylogenetic relationships of 16S rRNA genes in *Flavobacterium* strains.** Maximum Likelihood analysis based on concatenated alignments of 16S rRNA genes. Bootstrap values are shown next to the branch nodes. Open and black circles represent aquatic and terrestrial isolates, respectively.

(TIF)

**Figure S3. Phylogenetic relationships of core genes in *Flavobacterium* strains.** Maximum Likelihood method is based on Hal-software-generated 15 concatenated core genome polypeptides with 100 bootstrap replicates. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Open and black circles represent aquatic and terrestrial isolates, respectively.

(TIF)

**Figure S4. Over- and under-represented gene ontologies (GO's) in the *Flavobacterium* core and niche-specific genomes.** (A) *Flavobacterium* core genome and (B) terrestrial clade niche-specific genome. Values over one (dotted line)

represent GO's terms that are overrepresented, where values under one are underrepresented. Enrichment of GO's terms was tested using a one-tail Fisher's exact test, between *Flavobacterium* core and niche-specific genomes and the whole protein set of *F. johnsoniae*.

(TIF)

### Table S1. General characteristics of the pepper root-associated flavobacterial isolates.

Phenotypic and growth characteristics were measured as follow: Flexirubin: *Flavobacterium* spp. were grown on CYE-agar for 24 hours and the exposure to 50  $\mu$ l of 10 N KOH followed by exposure to 42  $\mu$ l 12 N HCl. Flexirubin-positive cells were yellow at neutral pH and red under alkaline conditions. Gliding motility: *Flavobacterium* spp. were grown on PY2 agar medium (2 g of peptone, 0.5 g yeast extract, 5 g agar and DDW up to 1 liter) at 25°C. Colony spreading was observed by binocular 24-48 hours after initial plating. Hydrogen cyanide (HCN) production: *Flavobacterium* spp. were grown on PY2 agar medium supplemented with glycine (4.4 g per liter) and cyanogenesis was revealed as previously described by 23. Results were read after 4 days of culture at 28°C. A change in filter paper colour from yellow to orange-brown indicated production of HCN (Yellow (1) - limited cyanide production; orange (2) - moderate cyanide production; light brown (3) - relatively high cyanide production; and brown (4) - high cyanide production). Amonia (NH<sub>3</sub>) production was estimated by using Nesler's reagent according to the manufacturer's instructions. Generation times of *Flavobacterium* spp. under different conditions were calculated from measured OD<sub>600nm</sub> at 25°C and 200 rpm. and indicates as follow: (5)- 60-90 min, (4)- 90-120 min, (3)- 120-200 min, (2)- 200-500 min, (1)- 500-800 min, (0)- 800-1200 and N.D for not detectible growth or generation time was longer than 1200 min. Hydrolytic activities were measured based on standard procedures.

(XLSX)

**Table S2. Flavobacterial clade-specific proteins.** Orthologs were defined by identifying unique pairwise reciprocal best hits, with at least 60% similarity in amino acid sequence and less than 30% difference in protein length. To determine the flavobacterial core genome, we used the 15 available genomes. For niche-specific core genome identification we used the completed genomes *F. branchiophilum* FL-15, *F. psychrophilum* JIP02/86; *F. columnare* ATCC 49512; *F. indicum* GPTSA100-9; *F. johnsoniae* UW101 and two draft genomes *Flavobacterium* sp. F52 and *Flavobacterium* sp. CF136.

(XLSX)

### Table S3. Summary of distribution carbohydrate utilizing enzymes.

Only sequences with similar predictions from dbCAN prediction web server and CAZymes Analysis Toolkit were selected and named according to CAZy classifications. GH-glycoside hydrolases; CE-carbohydrate esterase; PL-polysaccharide lyases; CBM- carbohydrate binding modules; GT- glycosyl transferases.

(XLS)

## Acknowledgements

We thank Stefan Green for assistance with sequence assembly, Galit Kahana and Yael Meller-Harel for technical assistance and Prof. Edouard Jurkevitch and Prof. Uri Gophna for valuable comments.

## References

- Bernardet JF, Bowman JP (2011) Genus. *Flavobacterium* Bergey et al. 1923. In: W Whitman. Bergey's Manual of Systematic Bacteriology. 2 ed. Baltimore, USA: The Williams & Wilkins Co. pp. 112-154.
- Bodenhausen N, Horton MW, Bergelson J (2013) Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana*. PLOS ONE 8: e56329. doi:10.1371/journal.pone.0056329. PubMed: 23457551.
- Bulgarelli D, Rott M, Schlaeppi K, van Themaat EVL, Ahmadinejad N et al. (2012) Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 488: 91-95. doi:10.1038/nature11336. PubMed: 22859207.
- Johansen JE, Binnerup SJ (2002) Contribution of Cytophaga-like bacteria to the potential of turnover of carbon, nitrogen, and phosphorus by bacteria in the rhizosphere of barley (*Hordeum vulgare*. p. L.). Microbial Ecology 43: 298-306
- Kirchman DL (2002) The ecology of *Cytophaga-Flavobacteria* in aquatic environments. FEMS Microbiol Ecol 39: 91-100. doi:10.1111/j.1574-6941.2002.tb00910.x. PubMed: 19709188.
- Kolton M, Harel YM, Pasternak Z, Graber ER, Elad Y et al. (2011) Impact of biochar application to soil on the root-associated bacterial community structure of fully developed greenhouse pepper plants. Appl Environ Microbiol 77: 4924-4930. doi:10.1128/AEM.00148-11. PubMed: 21622786.
- Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J et al. (2012) Defining the core Arabidopsis thaliana root microbiome. Nature 488: 86-90. doi:10.1038/nature11237. PubMed: 22859206.
- Manter DK, Delgado JA, Holm DG, Stong RA (2010) Pyrosequencing reveals a highly diverse and cultivar-specific bacterial endophyte community in potato roots. Microb Ecol 60: 157-166. doi:10.1007/s00248-010-9658-x. PubMed: 20414647.
- Declercq AM, Haesebrouck F, Van den Broeck W, Bossier P, Decostere A (2013) Columnaris disease in fish: a review with emphasis on bacterium-host interactions. Vet Res 44: 27. doi: 10.1186/1297-9716-44-27. PubMed: 23617544.
- Starliper CE (2011) Bacterial coldwater disease of fishes caused by *Flavobacterium psychrophilum*. J Advanced Res 2: 97-108. doi: 10.1016/j.jare.2010.04.001.
- Sang MK, Chun SC, Kim KD (2008) Biological control of *Phytophthora* blight of pepper by antagonistic rhizobacteria selected from a sequential screening procedure. Biol Contr 46: 424-433. doi:10.1016/j.biocontrol.2008.03.017.
- Sang MK, Kim KD (2012) The volatile-producing *Flavobacterium johnsoniae* strain GSE09 shows biocontrol activity against *Phytophthora capsici* in pepper. J Appl Microbiol 113: 383-398. doi: 10.1111/j.1365-2672.2012.05330.x. PubMed: 22563881.
- Fu YY, Tang XX, Lai QL, Zhang CH, Zhong HZ et al. (2011) *Flavobacterium beibuense* sp. nov., isolated from marine sediment. Int J Syst Evol Microbiol 61: 205-209. doi:10.1099/ijs.0.018846-0. PubMed: 20190020.
- Jit S, Dadhwal M, Prakash O, Lal R (2008) *Flavobacterium lindanitolerans* sp. nov., isolated from hexachlorocyclohexane-contaminated soil. Int J Syst Evol Microbiol 58: 1665-1669. doi: 10.1099/ijs.0.65578-0. PubMed: 18599713.
- Kolton M, Green SJ, Harel YM, Sela N, Elad Y et al. (2012) Draft genome sequence of *Flavobacterium* sp strain F52, isolated from the rhizosphere of bell pepper. *Capsicum annuum* L: cv. Maccabi). Journal of Bacteriology 194: 5462-5463
- Elad Y, David DR, Harel YM, Borenshtein M, Ben Kalifa H et al. (2010) Induction of systemic resistance in plants by biochar, a soil-applied carbon sequestering agent. Phytopathology 100: 913-921. doi:10.1094/PHYTO-100-9-0913. PubMed: 20701489.
- Graber ER, Harel YM, Kolton M, Cytryn E, Silber A et al. (2010) Biochar impact on development and productivity of pepper and tomato grown in fertigated soilless media. Plant Soil 337: 481-496. doi: 10.1007/s11104-010-0544-6.

## Author Contributions

Conceived and designed the experiments: MK YE EC. Performed the experiments: MK. Analyzed the data: MK YE EC NS. Contributed reagents/materials/analysis tools: MK NS. Wrote the manuscript: MK YE EC NS.

- McBride MJ, Kempf MJ (1996) Development of techniques for the genetic manipulation of the gliding bacterium *Cytophaga johnsonae*. J Bacteriol 178: 583-590. PubMed: 8550486.
- Agarwal S, Hunnicutt DW, McBride MJ (1997) Cloning and characterization of the *Flavobacterium johnsoniae* (*Cytophaga johnsonae*) gliding motility gene, *gldA*. Proc Natl Acad Sci U S A 94: 12139-12144. doi:10.1073/pnas.94.22.12139. PubMed: 9342376.
- Berić T, Urdaci MC, Stanković S, Knežević-Vukčević J (2009) RAPD analysis of genetic diversity and qualitative assessment of hydrolytic activities in a collection of Bacillus sp. isolate. Arch Biol Sciences 61: 645-652. doi:10.2298/ABS0904645B.
- Johansen JE, Nielsen P, Binnerup SJ (2009) Identification and potential enzyme capacity of flavobacteria isolated from the rhizosphere of barley (*Hordeum vulgare*. p. L.). Canadian Journal of Microbiology 55: 234-241
- Reichenbach H, Kohl W, Bottger-Vetter A, Achenbach H (1980) Flexirubin-type pigments in *Flavobacterium*. Arch Microbiol 126: 291-293. doi:10.1007/BF00409934.
- Soltani A, Khavazi K, Asadi-Rahmani H, Omidvari M, Abaszadeh Dahaji P et al. (2010) Plant growth promoting characteristics in some *Flavobacterium* spp. isolated from soils of Iran. J Agric Sci 2: 106.
- Yu Y, Li H-R, Zeng Y-X, Chen B (2011) Bacterial diversity and bioprospecting for cold-active hydrolytic enzymes from culturable bacteria associated with sediment from Nella Fjord, Eastern Antarctica. Mar Drugs 9: 184-195. doi:10.3390/md9020184. PubMed: 21566794.
- Loper JE, Hassan KA, Mavrodi DV, Davis EW II, Lim CK et al. (2012) Comparative Genomics of Plant-Associated *Pseudomonas* spp.: Insights into Diversity and Inheritance of Traits Involved in Multitrophic Interactions. PLOS Genet 8: e1002784. PubMed: 22792073.
- Yan YL, Yang J, Dou YT, Chen M, Ping SZ et al. (2008) Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501. Proc Natl Acad Sci U S A 105: 7564-7569. doi:10.1073/pnas.0801093105. PubMed: 18495935.
- Robbertse B, Yoder RJ, Boyd A, Reeves JB, Spatafora JW (2011) Hal: an automated pipeline for phylogenetic analyses of genomic data. PLOS Curr Trees Life 4: RRN1213 PubMed: 21327165.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W et al. (2010) New algorithms and methods to estimate Maximum-Likelihood phylogenies: Assessing the performance of PhyML. Syst Biol 3. 0: 59: 307-321
- Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol 25: 1307-1320. doi:10.1093/molbev/msn067. PubMed: 18367465.
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5: 113. doi:10.1186/1471-2105-5-113. PubMed: 15318951.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739. doi:10.1093/molbev/msr121. PubMed: 21546353.
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674-3676. doi: 10.1093/bioinformatics/bti610. PubMed: 16081474.
- Aziz RK, Devoid S, Disz T, Edwards RA, Henry CS et al. (2012) SEED Servers: High-performance access to the SEED genomes, annotations, and metabolic models. PLOS ONE 7: e48053. doi:10.1371/journal.pone.0048053. PubMed: 23110173.
- Hammer O, Harper D, Ryan P (2001) PAST: paleontological statistics software package for education and data analysis. Paleontol Electronica 4: 9.
- Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC (2010) CAZymes Analysis Toolkit (CAT): Web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced

- organism using CAZy database. *Glycobiology* 20: 1574-1584. doi: 10.1093/glycob/cwq106. PubMed: 20696711.
36. Yin Y, Mao X, Yang J, Chen X, Mao F et al. (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40: W445-W451. doi:10.1093/nar/gks479. PubMed: 22645317.
  37. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V et al. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* 37: D233-D238. doi: 10.1093/nar/gkn663. PubMed: 18838391.
  38. Rawlings ND, Barrett AJ, Bateman A (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 40: D343-D350. doi:10.1093/nar/gkr987. PubMed: 22086950.
  39. Nagata T (2008) Organic matter-bacteria interactions in seawater. In: DL Kirchman. *Microbial ecology of the ocean*: Wiley. pp. 207-242.
  40. Niklas KJ (2004) The cell walls that bind the tree of life. *BioScience* 54: 831-841. doi:10.1641/0006-3568(2004)054[0831:TCWVBT]2.0.CO;2.
  41. Popper ZA, Tuohy MG (2010) Beyond the green: understanding the evolutionary puzzle of plant and algal cell walls. *Plant Physiol* 153: 373-383. doi:10.1104/pp.110.158055. PubMed: 20421458.
  42. Naumoff DG (2011) Hierarchical classification of glycoside hydrolases. *Biochemistry (Mosc)* 76: 622-635. doi:10.1134/S0006297911060022. PubMed: 21639842.
  43. Gómez-Pereira PR, Schüler M, Fuchs BM, Bennke C, Teeling H et al. (2012) Genomic content of uncultured *Bacteroidetes* from contrasting oceanic provinces in the north Atlantic Ocean. *Environ Microbiol* 14: 52-66. doi:10.1111/j.1462-2920.2011.02555.x. PubMed: 21895912.
  44. Ali Z, Cousin S, Frühling A, Brambilla E, Schumann P et al. (2009) *Flavobacterium rivuli* sp. nov., *Flavobacterium subsaxonicum* sp. nov., *Flavobacterium swingsii* sp. nov. and *Flavobacterium reichenbachii* sp. nov., isolated from a hard water rivulet. *Int J Syst Evol Microbiol* 59: 2610-2617. doi:10.1099/ijs.0.008771-0. PubMed: 19622646.
  45. Cousin S (2009) Flavobacterial community structure in a hardwater rivulet and adjacent forest soil, Harz mountain, Germany. *Curr Microbiol* 58: 409-415. doi:10.1007/s00284-008-9323-5. PubMed: 19011940.
  46. Yi H, Oh HM, Lee JH, Kim SJ, Chun J (2005) *Flavobacterium antarcticum* sp nov., a novel psychrotolerant bacterium isolated from the Antarctic. *Int J Syst Evol Microbiol* 55: 637-641. doi:10.1099/ijs.0.63423-0. PubMed: 15774636.
  47. Fernández-Gómez B, Richter M, Schüler M, Pinhassi J, Acinas SG et al. (2013) Ecology of marine *Bacteroidetes*: a comparative genomics approach. *ISME J* 7: 1026-1037. doi:10.1038/ismej.2012.169. PubMed: 23303374.
  48. Yoon J-H, Kang S-J, Oh T-K (2006) *Flavobacterium soli* sp. nov., isolated from soil. *Int J Syst Evol Microbiol* 56: 997-1000. doi:10.1099/ijs.0.64119-0. PubMed: 16627644.
  49. Eitema TJG, Andersson SGE (2009) The *a-proteobacteria*: the Darwin finches of the bacterial world. *Biol Lett* 5: 429-432. doi:10.1098/rsbl.2008.0793. PubMed: 19324639.
  50. Wang Y, Yang JK, Lee OO, Li TG, Al-Suwailm A et al. (2011) Bacterial niche-specific genome expansion is coupled with highly frequent gene disruptions in deep-sea sediments. *PLOS ONE* 6: e29149. doi:10.1371/journal.pone.0029149. PubMed: 22216192.
  51. Ochman H, Davalos LM (2006) The nature and dynamics of bacterial genomes. *Science* 311: 1730-1733. doi:10.1126/science.1119966. PubMed: 16556833.
  52. Heuer H, Smalla K (2012) Plasmids foster diversification and adaptation of bacterial populations in soil. *FEMS Microbiol Rev* 36: 1083-1104. doi:10.1111/j.1574-6976.2012.00337.x. PubMed: 22393901.
  53. Young IM, Crawford JW (2004) Interactions and self-organization in the soil-microbe complex. *Science* 304: 1634-1637. doi:10.1126/science.1097394. PubMed: 15192219.
  54. Alonso C, Warnecke F, Amann R, Pernthaler J (2007) High local and global diversity of *Flavobacteria* in marine plankton. *Environ Microbiol* 9: 1253-1266. doi:10.1111/j.1462-2920.2007.01244.x. PubMed: 17472638.
  55. DeLong EF, Franks DG, Alldredge AL (1993) Phylogenetic diversity of aggregate-attached vs free-living marine bacterial assemblages. *Limnol Oceanogr* 38: 924-934. doi:10.4319/lo.1993.38.5.0924.
  56. Pinhassi J, Sala MM, Havskum H, Peters F, Guadayol Às, et al. (2004) Changes in bacterioplankton composition under different phytoplankton regimens. *Applied and Environmental Microbiology* 70: 6753-6766.
  57. Riemann L, Steward GF, Azam F (2000) Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. *Appl Environ Microbiol* 66: 578-587. doi:10.1128/AEM.66.2.578-587.2000. PubMed: 10653721.
  58. Ogawa H, Tanoue E (2003) Dissolved organic matter in oceanic waters. *J Oceanogr* 59: 129-147. doi:10.1023/A:1025528919771.
  59. Gómez-Consarnau L, González JM, Coll-Lladó M, Gourdon P, Pascher T et al. (2007) Light stimulates growth of proteorhodopsin-containing marine *Flavobacteria*. *Nature* 445: 210-213. doi:10.1038/nature05381. PubMed: 17215843.
  60. González JM, Fernández-Gómez B, Fernández-Guerra A, Gómez-Consarnau L, Sánchez O et al. (2008) Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp MED152 (*Flavobacteria*). *Proc Natl Acad Sci U S A* 105: 8724-8729. doi: 10.1073/pnas.0712027105. PubMed: 18552178.
  61. González JM, Pinhassi J, Fernández-Gómez B, Coll-Lladó M, González-Velázquez M et al. (2011) Genomics of the proteorhodopsin-containing marine *Flavobacterium* Dokdonia sp Strain MED134. *Appl Environ Microbiol* 77: 8676-8686. doi:10.1128/AEM.06152-11. PubMed: 22003006.
  62. Riedel T, Gómez-Consarnau L, Tomasch J, Martin M, Jarek M et al. (2013) Genomics and physiology of a marine *Flavobacterium* encoding a proteorhodopsin and a xanthorhodopsin-like protein. *PLOS ONE* 8: e57487. doi:10.1371/journal.pone.0057487. PubMed: 23526944.
  63. McBride MJ, Xie G, Martens EC, Lapidus A, Henrissat B et al. (2009) Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl Environ Microbiol* 75: 6864-6875. doi:10.1128/AEM.01495-09. PubMed: 19717629.
  64. Brown SD, Utturkar SM, Klingeman DM, Johnson CM, Martin SL et al. (2012) Twenty-one Genome sequences from diverse bacteria Isolated from the rhizosphere and endosphere of *Populus deltoides*. *J Bacteriol* 194: 5991-5993. doi:10.1128/JB.01243-12. PubMed: 23045501.
  65. Raymond JA, Kim HJ (2012) Possible role of horizontal gene transfer in the colonization of sea ice by algae. *PLOS ONE* 7: e35968. PubMed: 22567121.
  66. Touchon M, Barbier P, Bernardet JF, Loux V, Vacherie B et al. (2011) Complete genome sequence of the fish pathogen *Flavobacterium branchiophilum*. *Appl Environ Microbiol* 77: 7656-7662. doi:10.1128/AEM.05625-11. PubMed: 21926215.
  67. Tekedar HC, Karsi A, Gillaspay AF, Dyer DW, Benton NR et al. (2012) Genome sequence of the fish pathogen *Flavobacterium columnare* ATCC 49512. *J Bacteriol* 194: 2763-2764. doi:10.1128/JB.00281-12. PubMed: 22535941.
  68. Barbier P, Houel A, Loux V, Poulain J, Bernardet JF et al. (2012) Complete genome sequence of *Flavobacterium indicum* GPSTA100-9T, isolated from warm spring water. *J Bacteriol* 194: 3024-3025. doi:10.1128/JB.00420-12. PubMed: 22582381.
  69. Duchaud E, Boussaha M, Loux V, Bernardet JF, Michel C et al. (2007) Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum*. *Nat Biotechnol* 25: 763-769. doi:10.1038/nbt1313. PubMed: 17592475.