

# Lawrence Berkeley National Laboratory

## LBL Publications

### Title

High-quality draft genome sequence of *Flavobacterium suncheonense* GH29-5T (DSM 17707T) isolated from greenhouse soil in South Korea, and emended description of *Flavobacterium suncheonense* GH29-5T

### Permalink

<https://escholarship.org/uc/item/7986g7rp>

### Journal

Environmental Microbiome, 11(1)

### ISSN

2524-6372

### Authors

Tashkandy, Nisreen

Sabban, Sari

Fakieh, Mohammad

et al.

### Publication Date

2016

### DOI

10.1186/s40793-016-0159-5

Peer reviewed

EXTENDED GENOME REPORT

Open Access



# High-quality draft genome sequence of *Flavobacterium suncheonense* GH29-5<sup>T</sup> (DSM 17707<sup>T</sup>) isolated from greenhouse soil in South Korea, and emended description of *Flavobacterium suncheonense* GH29-5<sup>T</sup>

Nisreen Tashkandy<sup>1\*</sup>, Sari Sabban<sup>1</sup>, Mohammad Fakieh<sup>1</sup>, Jan P. Meier-Kolthoff<sup>2</sup>, Sixing Huang<sup>2</sup>, Brian J. Tindall<sup>2</sup>, Manfred Rohde<sup>3</sup>, Mohammed N. Baeshen<sup>1,4</sup>, Nabih A. Baeshen<sup>1,4</sup>, Alla Lapidus<sup>5</sup>, Alex Copeland<sup>6</sup>, Manoj Pillay<sup>7</sup>, T. B. K. Reddy<sup>6</sup>, Marcel Huntemann<sup>6</sup>, Amrita Pati<sup>6</sup>, Natalia Ivanova<sup>6</sup>, Victor Markowitz<sup>7</sup>, Tanja Woyke<sup>6</sup>, Markus Göker<sup>2</sup>, Hans-Peter Klenk<sup>8</sup>, Nikos C. Kyrpides<sup>1,6</sup> and Richard L. Hahnke<sup>2</sup>

## Abstract

*Flavobacterium suncheonense* is a member of the family *Flavobacteriaceae* in the phylum *Bacteroidetes*. Strain GH29-5<sup>T</sup> (DSM 17707<sup>T</sup>) was isolated from greenhouse soil in Suncheon, South Korea. *F. suncheonense* GH29-5<sup>T</sup> is part of the *Genomic Encyclopedia of Bacteria and Archaea* project. The 2,880,663 bp long draft genome consists of 54 scaffolds with 2739 protein-coding genes and 82 RNA genes. The genome of strain GH29-5<sup>T</sup> has 117 genes encoding peptidases but a small number of genes encoding carbohydrate active enzymes (51 CAZymes). Metallo and serine peptidases were found most frequently. Among CAZymes, eight glycoside hydrolase families, nine glycosyl transferase families, two carbohydrate binding module families and four carbohydrate esterase families were identified. Surprisingly, polysaccharides utilization loci (PULs) were not found in strain GH29-5<sup>T</sup>. Based on the coherent physiological and genomic characteristics we suggest that *F. suncheonense* GH29-5<sup>T</sup> feeds rather on proteins than saccharides and lipids.

**Keywords:** Aerobic, Gliding motility, Greenhouse soil, *Flavobacteriaceae*, *Bacteroidetes*, GEBA, KMG-1, Tree of Life, GGDC, Carbohydrate active enzyme, Polysaccharide utilization loci

## Introduction

*Flavobacteria/Cytophagia* have been frequently observed in aquatic and soil habitats [1–3] and play a major role in polysaccharide decomposition [2, 4, 5]. Type strains of the genus *Flavobacterium* have been isolated from many different habitats such as fresh water, sea ice and soil, and some *Flavobacterium* strains are pathogenic to humans and animals [2, 6]. Strain GH29-5<sup>T</sup> (= DSM 17707<sup>T</sup> = CIP

109901<sup>T</sup> = KACC 11423<sup>T</sup>) is the type strain of *Flavobacterium suncheonense* [2, 7], which belongs to *Flavobacteriaceae* [8]. *F. suncheonense* GH29-5<sup>T</sup> was isolated from greenhouse soil in Korea [10]. *Flavobacterium johnsoniae* UW101<sup>T</sup>, a well studied model organism, was as well isolated from soil [11, 12] and harbors a considerable number of CAZymes and PULs [13]. Thus, an investigation of the genome of strain GH29-5<sup>T</sup> will give further insights into the variety of CAZymes and the polysaccharide decomposition potential of this microorganism.

Here we present the set of carbohydrate active enzymes, polysaccharide utilization loci and peptidases of

\* Correspondence: ntashkandy@yahoo.com

<sup>1</sup>Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia

Full list of author information is available at the end of the article

*F. suncheonense* GH29-5<sup>T</sup>, together with a set of phenotypic features and the description and annotation of the high-quality draft genome sequence from a culture of DSM 17707<sup>T</sup>.

## Organism information

### Classification and features

The sequence of the single 16S rRNA gene copy in the genome is identical with the previously published 16S rRNA gene sequence (DQ222428). Figure 1 shows the phylogenetic neighborhood of *F. suncheonense* GH29-5<sup>T</sup> inferred from a tree of 16S rRNA gene sequence, as previously described [14]. The next related type species are *F. cauense* R2A-7<sup>T</sup> (EU521691), *F. enshiense* DK69<sup>T</sup> (JN790956), *F. limnosediminis* JC2902<sup>T</sup> (JQ928688) and *F. saliperosum* S13<sup>T</sup> (DQ021903) with less than 95.9 % 16S rRNA gene identity. The 16S rRNA gene sequence of strain GH29-5<sup>T</sup> has an identity of only 93.9 % with *F. aquatile* DSM 1132<sup>T</sup> (AM230485).

The 16S rRNA gene sequence of *F. suncheonense* GH29-5<sup>T</sup> was compared with the Greengenes database [15]. Considering the best 100 hits, 99 sequences belonged to *Flavobacterium* and one sequence to *Cytophaga* sp. (X85210). Among the most frequent keywords within the labels of environmental samples were 40.4 % marine habitats (such as marine sediment, deep sea, seawater, whale fall, diatom/phytoplankton bloom, Sargasso Sea, sponge, sea urchin, bacterioplankton), 12.3 % soil habitats (such as rhizosphere, grassland, compost), 11.6 % freshwater habitats (such as lake, riverine sediment, groundwater), 8.9 % cold environments (such as Antarctic/Arctic seawater, lake ice or sediment), but also 2.7 % wastewater habitats. Interestingly, environmental 16S rRNA gene sequences with 99 % sequence identity with *F. suncheonense* GH29-5<sup>T</sup> were clones from wetland of France (KC432449) [16] and an enrichment culture of heterotrophic soil bacteria from the Netherlands (JQ855723), and with 98 % sequence identity to a soil isolate from Taiwan (DQ239767).

As described for *Flavobacterium* [17], *F. suncheonense* GH29-5<sup>T</sup> stains are Gram-negative (Table 1). The colonies are convex, round and yellow, but flexirubin-type pigments are absent and gliding motility was not observed [10]. The strain is positive for the catalase and oxidase tests [10], as are most members of the genus *Flavobacterium* [6]. Cells divide by binary fission, possess appendages and occur either as single rod shaped cells, with 0.3 μm in width and 1.5–2.5 μm in length, or as filaments (Fig. 2).

*F. suncheonense* GH29-5<sup>T</sup> grows between 15 °C and 37 °C, pH 6 and 8 and in media with up to 1 % NaCl [10], with optimal growth at pH 7.0 and without NaCl [7]. Strain GH29-5<sup>T</sup> decomposes gelatin and casein, but not starch, carboxymethyl cellulose, agar, alginate, pectin, chitin, aesculin and DNA [10]. Strain GH29-5<sup>T</sup>

produces H<sub>2</sub>S and neither reduces nitrate nor produces indole or ferments glucose [10]. Moreover, strain GH29-5<sup>T</sup> does not utilize arabinose, mannose, N-acetyl-D-glucosamine, maltose, gluconate, caprate, adipate, malate, citrate and phenylacetate [19]. Strain GH29-5<sup>T</sup> possesses alkaline phosphatase, esterase C4, esterase lipase C8, leucine arylamidase, valine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase and N-acetyl-β-glucosaminidase, but has no lipase C14, cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, α-mannosidase, α-fucosidase and urease activity [10].

### Chemotaxonomic data

The major cellular fatty acids are iso-C<sub>15:0</sub> (29.9 %), iso-C<sub>17:0</sub> 3-OH (17.7 %), iso-C<sub>15:1</sub> G (12.0 %) and iso-C<sub>15:0</sub> 3-OH (11.1 %) and MK-6 is the sole quinone [10], as common in *Flavobacterium* [6]. Besides phosphatidylethanolamine, several unidentified lipids, aminolipids and aminophospholipids were observed in strain GH29-5<sup>T</sup> [7]. The DNA G + C content was reported to be 39.0 mol % [10].

## Genome sequencing information

### Genome project history

This strain was selected for sequencing on the basis of its phylogenetic position [20, 21], and is part of Genomic Encyclopedia of Type Strains, Phase I: the one thousand microbial genomes (KMG) project [22], a follow-up of the Genomic Encyclopedia of *Bacteria* and *Archaea* (GEBA) pilot project [23], which aims at sequencing key reference microbial genomes and generating a large genomic basis for the discovery of genes encoding novel enzymes [24]. KMG-I is the part of the “Genomic Encyclopedia of *Bacteria* and *Archaea*: sequencing a myriad of type strains initiative” [25] and a Genomic Standards Consortium project [26]. The genome project is deposited in the Genomes OnLine Database [27] and the permanent draft genome sequence is deposited in GenBank. Sequencing, finishing and annotation were performed by the DOE-JGI using state-of-the-art sequencing technology [28]. A summary of the project information is shown in Table 2.

### Growth conditions and genomic DNA preparation

A culture of GH29-5<sup>T</sup> (DSM 17707) was grown aerobically in DSMZ medium 830 (R2A Medium) [29] at 28 °C. Genomic DNA was isolated using a Jetflex Genomic DNA Purification Kit (GENOMED 600100) following the standard protocol provided by the manufacturer. DNA is available from the DSMZ through the DNA Bank Network [30].

### Genome sequencing and assembly

The draft genome of strain GH29-5<sup>T</sup> was generated using the Illumina technology [31]. An Illumina Std. shotgun



Fig. 1 (See legend on next page.)

(See figure on previous page.)

**Fig. 1** Phylogenetic tree of the genus *Flavobacterium* and its most closely related genus *Capnocytophaga*. Modified from Hahnke et al. [68]. In short: the tree was inferred from 1254 aligned characters of the 16S rRNA gene sequence under the maximum likelihood (ML) criterion. The branches are scaled in terms of the expected number of substitutions per site. Numbers adjacent to the branches are support values from 1000 ML bootstrap replicates (*left*) and from 1000 maximum-parsimony bootstrap replicates (*right*) if larger than 60 %

library was constructed and sequenced using the Illumina HiSeq 2000 platform which generated 9,392,462 reads totaling 1408.9 Mbp (Table 3). All general aspects of library construction and sequencing performed at the DOE-JGI can be found at [32]. All raw sequence data were passed through DUK, a filtering program developed at DOE-JGI, which removes known Illumina sequencing and library preparation artifacts (Mingkun L, Copeland A,

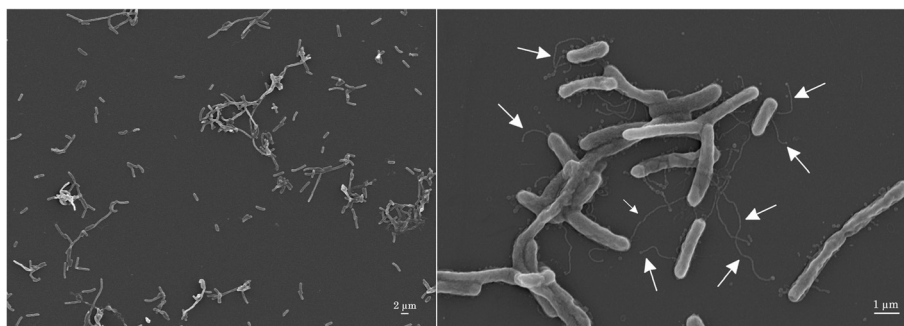
Han J: DUK. *unpublished* 2011). The following steps were performed for assembly: (1) filtered reads were assembled using Velvet [33], (2) 1–3 Kbp simulated paired-end reads were created from Velvet contigs using wgsim [34], (3) Sequence reads were assembled with simulated read pairs using Allpaths-LG [35]. Parameters for assembly steps were: 1) Velvet (“velveth 63 -shortPaired” and “velvetg -very\_clean yes -exportFiltered yes -min\_contig\_lgth 500

**Table 1** Classification and general features of *F. suncheonense* GH29-5<sup>T</sup> in accordance with the MIGS recommendations [59], as developed by [60], List of Prokaryotic names with Standing in Nomenclature [61] and the Names for Life database [62]

MIGS ID	Property	Term	Evidence code
	Current classification	Domain: <i>Bacteria</i>	TAS [12]
		Phylum: <i>Bacteroidetes</i>	TAS [63, 64]
		Class: ' <i>Flavobacteriia</i> '	TAS [65, 66]
		Order: <i>Flavobacteriales</i>	TAS [9, 67]
		Family: <i>Flavobacteriaceae</i>	TAS [8, 9]
		Genus: <i>Flavobacterium</i>	TAS [6, 68]
		Species: <i>Flavobacterium suncheonense</i>	TAS [10]
		Type strain: GH29-5 <sup>T</sup>	TAS [10]
	Gram-stain	Negative	TAS [10]
	Cell shape	rod-shaped	TAS [10]
	Motility	Nonmotile	TAS [10]
	Sporulation	non-spore forming	NAS
	Temperature range	mesophilic (15–37 °C)	TAS [10]
	Optimum temperature	16–24 °C	TAS [10]
	pH range; Optimum	6–8,	TAS [10]
	Carbon source	Carbohydrates, peptides	TAS [10]
	Energy source	chemoheterotroph	TAS [10]
MIGS-6	Habitat	greenhouse soil	TAS [10]
MIGS-	Salinity	0–1 % NaCl, 0 % NaCl	TAS [10]
MIGS-22	Oxygen requirement	aerobe	TAS [10]
MIGS-15	Biotic relationship	free-living	TAS [10]
MIGS-14	Pathogenicity	unknown	TAS [69]
	Biosafety level	1	TAS [69]
MIGS-4	Geographic location	Suncheon City, South Korea	TAS [10]
MIGS-5	Sample collection	2005	TAS [10]
MIGS-	Latitude	34.954	TAS [10]
MIGS-4.2	Longitude	127.483	TAS [10]
MIGS-4.4	Altitude	not reported	TAS [10]

Evidence codes are from the Gene Ontology project [18]

Evidence codes - *IDA* inferred from direct assay (first time in publication); *TAS* traceable author statement (i.e., a direct report exists in the literature); *NAS* non-traceable author statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence)



**Fig. 2** Scanning-electron micrograph of *F. suncheonense* GH29-5<sup>T</sup> (DSM 17707<sup>T</sup>) showing appendages 50–80 nm in diameter and 0.5–8 μm in length (arrows)

-scaffolding no -cov\_cutoff 10”), (2) wgsim (“wgsim -e 0–1 100–2 100 -r 0 -R 0 -X 0”) (3) Allpaths–LG (“PrepareAllpathsInputs: PHRED\_64 = 1 PLOIDY = 1 FRAG\_COVERAGE = 125 JUMP\_COVERAGE = 25 LONG\_JUMP\_COV = 50” and “RunAllpathsLG THREADS = 8 RUN = std shredpairs TARGETS = standard VAPI\_WARN\_ONLY = True OVERWRITE = True”). The final draft assembly contained 57 contigs in 54 scaffolds. The total size of the genome is 2.9 Mbp and the final assembly is based on 331.3 Mbp of data, which provides a 114.2x average coverage of the genome.

#### Genome annotation

Genes were identified using Prodigal [36] as part of the DOE-JGI genome annotation pipeline [37], followed by manual curation using the DOE-JGI GenePRIMP pipeline [38]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information non-

redundant database, UniProt, TIGR-Fam, Pfam, PRIAM, KEGG, COG, and InterPro database. These data sources were combined to assert a product description for each predicted protein. Additional gene prediction analysis and functional annotation was performed within the IMG-ER platform [39].

#### Genome properties

The assembly of the draft genome sequence consists of 54 scaffolds amounting to 2,880,663 bp. The G + C content is 40.5 % (Table 3) which is 1.5 % higher than previously reported by Kim et al. [10] and thus shows a difference that surpasses the maximal range among strains belonging to the same species [40]. Of the 2821 genes predicted, 2739 were protein-coding genes, and 82 RNAs. The majority of the protein-coding genes (69.2 %) were assigned a putative function while the remaining ones were annotated as hypothetical proteins. The distribution of genes into COG functional categories is presented in Table 4.

**Table 2** Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Level 2: High-Quality Draft
MIGS-28	Libraries used	Illumina Std shotgun library
MIGS 29	Sequencing platforms	Illumina, Illumina HiSeq 2000, Illumina HiSeq 2500
MIGS 31.2	Fold coverage	115.3x
MIGS 30	Assemblers	Velvet v. 1.1.04; ALLPATHS v. r41043
MIGS 32	Gene calling method	Prodigal, GenePRIMP, IMG-ER
	Locus Tag	G498
	Genbank ID	AUCZ00000000
	GenBank Date of Release	12-DEC-2013
	GOLD ID	Gp0013510
	BIOPROJECT	PRJNA185581
MIGS 13	Source Material Identifier	DSM 17707
	Project relevance	Tree of Life, GEBA-KMG

**Table 3** Genome statistics

Attribute	Value	% of Total
Genome size (bp)	2,880,663	100.0
DNA coding (bp)	2,622,751	91.1
DNA G + C (bp)	1,165,575	40.5
DNA scaffolds	54	
Total genes	2821	100.0
Protein coding genes	2739	97.1
RNA genes	82	2.9
Pseudo genes	0	0.0
Genes in internal clusters	125	4.43
Genes with function prediction	1916	67.92
Genes assigned to COGs	1439	51.01
Genes with Pfam domains	2020	71.61
Genes with signal peptides	348	12.34
Genes with transmembrane helices	631	22.37
CRISPR repeats	0	

**Table 4** Number of genes associated with the general COG functional categories

Code	Value	% age	Description
J	178	11.5	Translation, ribosomal structure and biogenesis
A	–	–	RNA processing and modification
K	83	5.3	Transcription
L	76	4.9	Replication, recombination and repair
B	1	0.1	Chromatin structure and dynamics
D	24	1.5	Cell cycle control, cell division, chromosome partitioning
Y	–	–	Nuclear structure
V	44	2.8	Defense mechanisms
T	53	3.4	Signal transduction mechanisms
M	165	10.6	Cell wall/membrane/envelope biogenesis
N	10	0.6	Cell motility
Z	–	–	Cytoskeleton
W	–	–	Extracellular structures
U	15	1.0	Intracellular trafficking, and secretion
O	93	6.0	Posttranslational modification, protein turnover, chaperones
C	84	5.4	Energy production and conversion
G	51	3.1	Carbohydrate transport and metabolism
E	109	7.1	Amino acid transport and metabolism
F	62	4.0	Nucleotide transport and metabolism
H	99	6.4	Coenzyme transport and metabolism
I	77	5.0	Lipid transport and metabolism
P	74	4.8	Inorganic ion transport and metabolism
Q	29	1.9	Secondary metabolites biosynthesis, transport and catabolism
R	131	8.4	General function prediction only
S	83	5.3	Function unknown
–	1382	49.0	Not in COGs

## Insights from the genome sequence

### Comparative genomics

We conducted a comparative genomics analysis of *F. suncheonense* (AUCZ00000000) with a selection of closely related (according to 16S rRNA gene sequence similarities) *Flavobacterium* type strains, i.e., *F. enshiense* (AVCS00000000), *F. cauense* (AVBI00000000), *F. saliperosum* (AVFO00000000) and *F. columnare* (CP003222) and the type species *F. aquatile* (JRH00000000). The genome sizes of the five type strains were 3.1 Mbp on average with the biggest difference of 0.5 Mbp between the genomes of *F. suncheonense* and *F. saliperosum*, on the one hand, and *F. enshiense*, on the other hand. Genome sizes were 3.1 Mbp (*F. cauense*), 3.2 Mbp (*F. columnare*), 3.4 Mbp (*F. enshiense*), 2.9 Mbp (*F. suncheonense*) and 2.9 Mbp (*F. saliperosum*). However, since these genomes have not yet been sequenced completely, their sizes might slightly change in the future.

An estimate of the overall similarity between *F. suncheonense* and the five reference strains was conducted using the Genome-to-Genome Distance Calculator (GGDC 2.0) [41, 42]. It reports model-based DDH estimates (digital DDH or dDDH) along with their confidence intervals [42], which allow for genome-based species delineation and genome-based subspecies delineation. The recommended distance formula 2 is robust against the use of incomplete genome sequences and is thus especially suited for this dataset.

The result of this comparison is shown in Table 5 and yields dDDH of below 22 % throughout, which confirms the expected status of distinct species. Furthermore, the G + C content was calculated from the genome sequences of the above strains and their pairwise differences were assessed with respect to *F. suncheonense*. Differences were 2.4 % (*F. cauense*), 2.8 % (*F. enshiense*), 1 % (*F. saliperosum*), 9.1 % (*F. columnare*) and 8.3 % (*F. aquatile*). These differences confirm the status of distinct species, because, if computed from genome sequences, these differences can only vary up to 1 % within species [40].

**Table 5** Pairwise comparison using the GGDC (Genome-to-Genome Distance Calculator) of *F. suncheonense* with a selection of currently available *Flavobacterium* genomes, *F. enshiense* (AVCS00000000), *F. cauense* (AVBI00000000), *F. saliperosum* (AVFO00000000) and *F. columnare* (CP003222), plus the type species *F. aquatile* (JRH00000000)

<i>F. suncheonense</i> versus	% dDDH	% C.I. dDDH	HSP length/total length [%]	Identities HSP/length [%]	Identities/total length [%]
<i>F. aquatile</i>	18.7	2.6	4	76	3
<i>F. cauense</i>	21.2	3.0	45	79	36
<i>F. columnare</i>	20.9	2.6	4	79	3
<i>F. enshiense</i>	20.2	2.9	29	78	23
<i>F. saliperosum</i>	21.0	3.0	41	79	33

Digital DDH values (dDDH) and the respective confidence intervals (C.I.) are specified for GGDC's recommended formula 2. The columns "HSP length / total length [%]", "Identities / HSP length [%]" and "Identities / total length [%]" list similarities as calculated from the intergenomic distances, which were also reported by the GGDC (Formulae 1–3)

### Gliding motility

McBride and Zhu [43] described the diversity of genes involved in gliding motility among members of phylum *Bacteroidetes*. The machinery for gliding motility is composed of adhesin-like proteins, the type IX secretion system, and additional proteins [43]. Even though strain GH29-5<sup>T</sup> was never observed to glide [10], all necessary genes for gliding motility were identified in its genome (Table 6).

### Carbohydrate active enzymes and peptidases

Cottrell and Kirchman [44] showed that members of the *Cytophaga-Flavobacteria* group preferentially consume polysaccharides and proteins rather than amino acids.

**Table 6** Gliding motility-related genes in strain GH29-5<sup>T</sup> compared to genes in *Flavobacterium* strains studied by McBride and Zhu [43]

locus tag prefix	<i>F. suncheonense</i> GH29-5 <sup>T</sup>	<i>F. rivuli</i> DSM 21788 <sup>T</sup>	<i>F. johnsoniae</i> ATCC 17061 <sup>T</sup>
Gliding motility	-	-	+
Adhesin-like			
<i>remA</i>	00716	-	0808
<i>remB</i>	01803	-	1657
<i>sprB</i>	+ <sup>b</sup>	-	0979
ATP-binding cassette transporter			
<i>gldA</i>	02505	05270	1516
<i>gldF</i>	02374	00760	2722
<i>gldG</i>	02375	00765	2721
Additional proteins			
<i>gldB</i> <sup>a</sup>	00808	13390	1793
<i>gldC</i>	00807	13385	1794
<i>gldD</i> <sup>a</sup>	01936	18865	1540
<i>gldE</i>	00405	18860	1539
<i>gldH</i> <sup>a</sup>	02655	10515	0890
<i>gldJ</i> <sup>a</sup>	00438	11845	1557
Peptidoprollyl isomerase (' <i>Flavobacteriia</i> ', protein folding)			
<i>gldI</i>	01009	08180	2369
Type IX secretion system (secretion of RemA/RemB)			
<i>gldK</i> <sup>a</sup>	00758	18605	1853
<i>gldL</i> <sup>a</sup>	00757	18600	1854
<i>gldM</i> <sup>a</sup>	00756	18595	1855
<i>gldN</i> <sup>a</sup>	00755	18590	1856
<i>sprA</i> <sup>a</sup>	01807	06065	1653
<i>sprE</i> <sup>a</sup>	02154	19150	1051
<i>sprT</i> <sup>a</sup>	02545	05475	1466

<sup>a</sup>essential gliding motility genes after McBride and Zhu [43]

<sup>b</sup>partial gene sequences, located at the beginning of AUCZ00000022 and at the end of AUCZ00000002

This phenotypic feature was attributed by Fernández-Gómez et al. [4] to higher numbers of peptidases and additionally higher numbers of glycoside hydrolases and carbohydrate-binding modules in the genomes of *Bacteroidetes* compared to other bacteria. *F. suncheonense* GH29-5<sup>T</sup> was isolated from greenhouse soil, hydrolyzes casein and gelatin, but did not utilize any of the tested saccharides [10, 19]. Therefore, we compared the predicted CDS against the CAZyme [45, 46] and dbCAN [47] database. The CAZyme annotation (Additional file 1, Table S1) was a combination of RAPSearch2 search [48, 49] and HMMER scanning [50] as described in Hahnke et al. [14]. The genome of strain GH29-5<sup>T</sup> comprised a small number of carbohydrate active enzymes (49) including 36 glycosyl transferases, nine glycoside hydrolases, four carbohydrate binding modules and six carbohydrate esterases (Table 7). Furthermore, sulfatases were suggested as important enzymes for the metabolic potential of *Bacteroidetes* to degrade sulfated algae polysaccharides such as carrageenan, agarans and fucans. Only, three sulfatases were identified in the genome of strain GH29-5<sup>T</sup> (Additional file 1, Table S2).

### Polysaccharide utilization loci

CAZymes of *Flavobacteria* that are suggested to be involved in polysaccharide decomposition are frequently observed to be organized in gene clusters. Such polysaccharides-utilization loci (PULs) consist of a TonB-dependent receptor, a SusD-like protein and carbohydrate active enzymes [51, 52]. In strain GH29-5<sup>T</sup> five TonB-dependent transporters were identified of which G498\_00119, G498\_01595, G498\_02575 were associated to siderophores and G498\_00706, G498\_00915 were associated with a SusD-like protein. The gene cluster up-stream of the TonB-dependent transporter G498\_00706 comprised five hypothetical proteins.

**Table 7** Carbohydrate active enzymes (CAZy) in the genome of strain GH29-5<sup>T</sup>

CAZy family	GH2	GH3	GH20	GH23	GH25	GH73	GH92
Counts	1	1	1	2	1	1	1
CAZy family	GH <sup>a</sup>		CBM50	CBM <sup>a</sup>			
Counts	1		3	1			
CAZy family	GT2	GT4	GT5	GT9	GT19	GT28	GT30
Counts	14	11	1	2	1	1	1
CAZy family	GT51	GT56					
Counts	4	1					
CAZy family	CE4	CE11	CE14	CE <sup>a</sup>		AA1	AA <sup>a</sup>
Counts	2	1	2	1		1	1

<sup>a</sup>genes attributed to an enzyme class, but not to a family



## Peptidases

The MEROPS annotation was carried out by searching the sequences against the MEROPS 9.10 database [53] (access date: 2014.10.16, version: pepunit.lib) as described in Hahnke et al. [14]. The genome of strain GH29-5<sup>T</sup> comprised 117 identified peptidase genes (or homologues), mostly serine peptidases (S, 50), metallo peptidases (M, 50) and cysteine peptidases (C, 14) (Table 8, Additional file 1: Tables S3 and S4). Hence, the low number of carbohydrate active enzymes and the high number of peptidases in the genome of strain GH29-5<sup>T</sup> reflects its currently known substrate range being proteins rather than saccharides.

## Conclusions

The genome of *F. suncheonense* GH29-5<sup>T</sup> contains a relatively low number of carbohydrate active enzymes in contrast to genomes of other *Flavobacteriaceae* such as *Flavobacterium branchiophilum* [54], *Flavobacterium rivuli* [14], *Formosa agariphila* [55], *Polaribacter* [4, 56], '*Gramella forsetii*' [57] and *Zobellia galactanivorans* [17]. This is surprising, since greenhouse soil might be a rich source of plant litter. McBride et al. [13] described the genome features of *Flavobacterium johnsoniae* UW101<sup>T</sup>, a bacterium that was as well isolated from soil [11, 58]. Both the genomes of *F. johnsoniae* UW101<sup>T</sup> and *F. suncheonense* GH29-5<sup>T</sup> have an almost equal number of 31 and 39 peptidases per Mbp, respectively. The genomes, however, differ remarkably in the number of CAZymes, with 47 genes per Mbp in the genome of *F. johnsoniae* UW101<sup>T</sup> and only 18 genes per Mbp in the genome of *F. suncheonense* GH29-5<sup>T</sup>. Thus, this small set of CAZymes contributes only little to a pool of

enzymes, which might be essential for a *Flavobacterium* to feed on soil components.

A systematic collection of genome sequences, such as GEBA [23] and KMG-1 [22], will provide the scientific community with the possibility for a systematic discovery of genes encoding for novel enzymes [24] and support microbial taxonomy. In addition, genome sequences also provide further taxonomically useful information such as the G + C content [40], which, as seen in this report might significantly differ from the values determined with traditional methods.

Based on the observed large difference in the DNA G + C content and the additional information on cell morphology obtained in this study, an emended description of *F. suncheonense* is proposed.

## Emended description of *F. suncheonense* GH29-5<sup>T</sup> Kim et al. 2006 emend. Dong et al. 2013

The description of *Flavobacterium suncheonense* is as given by Kim et al. [10] and Dong et al. [7], with the following modifications: the DNA G + C content is 40.5 mol%, and amendments: possesses appendages of 50–80 nm in diameter and 0.5–8 μm in length.

## Additional file

**Additional file 1: Table S1.** Carbohydrate active enzymes (CAZymes) in the genome of *F. suncheonense* GH29-5<sup>T</sup>. **Table S2.** Sulfatases in the genome of *F. suncheonense* GH29-5<sup>T</sup>. **Table S3.** Peptidases or homologues in the genome of *F. suncheonense* GH29-5<sup>T</sup>. **Table S4.** Simple peptidases inhibitors in the genome of *F. suncheonense* GH29-5<sup>T</sup>. (DOCX 497 kb)

**Table 8** Peptidases and simple peptidase inhibitors in the genome of strain GH29-5<sup>T</sup>

Peptidase	M01	M03	M12	M13	M14	M16	M20	M23	M24
Counts	4	1	2	2	5	2	3	6	2
Peptidase	M28	M36	M38	M41	M42	M43	M48	M50	M61
Counts	3	1	4	1	1	2	1	1	1
Peptidase	M79	M90							
Counts	1	1							
Peptidase	S01	S06	S08	S09	S12	S14	S16	S24	S26
Counts	1	1	3	16	5	1	3	1	1
Peptidase	S33	S41	S46	S49	S51	S54	S66		
Counts	6	3	2	1	1	4	1		
Peptidase	C01	C25	C26	C40	C44	C45	C56		
Counts	1	1	5	2	3	1	1		
Peptidase	N11		T02		U32	U73		A08	A28
Counts	1		1		4	1		1	1
Inhibitor	I39	I87							
Counts	4	1							

## Abbreviations

DOE: Department of Energy; EMBL: European molecular biology laboratory; GEBA: Genomic encyclopedia of *Bacteria* and *Archaea*; JGI: Joint Genome Institute; IMG-ER: Integrated microbial genomes – expert review; KMG: One thousand microbial genomes project; RDP: Ribosomal database project (East Lansing, MI, USA)

## Acknowledgments

The authors gratefully acknowledge the help of Andrea Schütze for growing cells of GH29-5<sup>T</sup> and of Evelyne-Marie Brambilla (both at DSMZ), for DNA extraction and quality control. This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231. AL was supported by the St. Petersburg State University grant (No 1.38.253.2015). R.L.H. was supported by the Bundesministerium für Ernährung und Landwirtschaft No. 22016812 (PI Brian J. Tindall). We would also like to thank the Center of Nanotechnology at King Abdulaziz University for their support.

## Authors' contributions

HPK and NCK initiated the study. RLH, SS, NT, MF, NCK and HPK designed research and project outline. SS, NT, MF, RLH, JPMK, MG, BJT, HPK and NCK drafted the manuscript. AL, JH, MP, TBKR, MH, AP, NNI, VM, TW and NCK sequenced, assembled and annotated the genome. MNB and NAB provided financial support. SH performed CAZY and MEROPS analysis. RLH investigated the CAZymes and PUL. JPMK conducted comparative genomics. JPMK and RLA performed 16S rRNA based phylogeny. MR performed electron microscopy. All authors read and approved the final manuscript.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia. <sup>2</sup>Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. <sup>3</sup>HZI – Helmholtz Centre for Infection Research, Braunschweig, Germany. <sup>4</sup>Center of Nanotechnology, King Abdulaziz University, Jeddah, Saudi Arabia. <sup>5</sup>Centre for Algorithmic Biotechnology, St. Petersburg State University, St. Petersburg, Russia. <sup>6</sup>Department of Energy Joint Genome Institute, Genome Biology Program, Walnut Creek, CA, USA. <sup>7</sup>Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. <sup>8</sup>School of Biology, Newcastle University, Newcastle upon Tyne, UK.

Received: 10 December 2015 Accepted: 23 May 2016

Published online: 16 June 2016

**References**

- Bodenhausen N, Horton MW, Bergelson J. Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana*. *PLoS One*. 2013;8:2.
- Kirchman DL. The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol Ecol*. 2002;39:91–100.
- Kolton M, Harel YM, Pasternak Z, Graber ER, Elad Y, Cytryn E. Impact of biochar application to soil on the root-associated bacterial community structure of fully developed greenhouse pepper plants. *Appl Environ Microbiol*. 2011;77:4924–30.
- Fernández-Gómez B, Richter M, Schüller M, Pinhassi J, Acinas SG, González JM, Pedrós-Alió C. Ecology of marine *Bacteroidetes*: a comparative genomics approach. *ISME J*. 2013;7:1026–37.
- Kolton M, Sela N, Elad Y, Cytryn E. Comparative genomic analysis indicates that niche adaptation of terrestrial *Flavobacteria* is strongly linked to plant glycan metabolism. *PLoS One*. 2013;8:1–11.
- Bernardet JF, Bowman JP. Genus I. *Flavobacterium* Bergey, Harrison, Breed, Hammer and Huntoon 1923, 97<sup>Al</sup> emend. Bernardet, Segers, Vancanneyt, Berthe, Kersters and Vandamme 1996, 139. In: *Bergey's Manual of Systematic Bacteriology, The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*, vol. 4. 2nd ed. New York: Springer; 2011. p. 112–54.
- Dong K, Xu B, Zhu F, Wang G. *Flavobacterium hauense* sp. nov., isolated from soil and emended descriptions of *Flavobacterium subsaxonicum*, *Flavobacterium beibuense* and *Flavobacterium rivuli*. *Int J Syst Evol Microbiol*. 2013;63:3237–42.
- Bernardet JF. Family I. *Flavobacteriaceae* Reichenbach 1992b, 327<sup>VP</sup> (effective publication: Reichenbach 1989b, 2013.) emend. Bernardet, Segers, Vancanneyt, Berthe, Kersters and Vandamme 1996, 145 emend. Bernardet, Nakagawa and Holmes 2002, 1057. In: *Bergey's Manual of Systematic Bacteriology, The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*, vol. 4. 2nd ed. New York: Springer; 2011. p. 106–314.
- Reichenbach H. Order 1. *Cytophagales* Leadbetter 1974, 99<sup>Al</sup>. In: Staley JT, Bryant MP, Pfennig N, Holt JG, editors. *Bergey's Manual of Systematic Bacteriology*, vol. 3. New York: Springer New York; 1989. p. 2011–3.
- Kim BY, Weon HY, Cousin S, Yoo SH, Kwon SW, Go SJ, Stackebrandt E: *Flavobacterium daejeonense* sp. nov. and *Flavobacterium suncheonense* sp. nov., isolated from greenhouse soils in Korea. *Int J Syst Evol Microbiol*. 2006;56:1645–9.
- Bernardet JF, Segers P, Vancanneyt M, Berthe F, Kersters K, Vandamme P. Cutting a gordian knot: emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov. (basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *Int J Syst Bacteriol*. 1996;46:128–48.
- Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci U S A*. 1990;87:4576–9.
- McBride MJ, Xie G, Martens EC, Lapidus A, Henrissat B, Rhodes RG, Goltsman E, Wang W, Xu J, Hunnicutt DW, Starosik AM, Hoover TR, Cheng YQ, Stein JL: Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl Environ Microbiol*. 2009;75:6864–75.
- Hahnke RL, Stackebrandt E, Meier-Kolthoff JP, Tindall BJ, Huang S, Rohde M, Lapidus A, Han J, Trong S, Haynes M, Reddy TBK, Huntemann M, Pati A, Ivanova NN, Mavromatis K, Markowitz V, Woyke T, Göker M, Kyrpides NC, Klenk H-P: High quality draft genome sequence of *Flavobacterium rivuli* type strain WB 3.3-2<sup>T</sup> (DSM 21788<sup>T</sup>), a valuable source of polysaccharide decomposing enzymes. *Stand Genomic Sci*. 2015;10:46.
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL: Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 2006;72:5069–72.
- Bouali M, Zrafi I, Bakhrouf A, Chaussonnerie S, Sghir A. Bacterial structure and spatiotemporal distribution in a horizontal subsurface flow constructed wetland. *Appl Microbiol Biotechnol*. 2014;98:3191–203.
- Thomas F, Barbeyron T, Tonon T, Génicot S, Czjzek M, Michel G. Characterization of the first alginolytic operons in a marine bacterium: From their emergence in marine *Flavobacteria* to their independent transfers to marine *Proteobacteria* and human gut *Bacteroides*. *Environ Microbiol*. 2012;14:2379–94.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25:25–9.
- Lee K, Park S-C, Yi H, Chun J. *Flavobacterium limnosediminis* sp. nov., isolated from sediment of a freshwater lake. *Int J Syst Evol Microbiol*. 2013;63:4784–9.
- Göker M, Klenk H-P. Phylogeny-driven target selection for large-scale genome-sequencing (and other) projects. *Stand Genomic Sci*. 2013;8:360–74.
- Klenk HP, Göker M. En route to a genome-based classification of *Archaea* and *Bacteria*? *Syst Appl Microbiol*. 2010;33:175–82.
- Kyrpides NC, Woyke T, Eisen JA, Garrity G, Lilburn TG, Beck BJ, et al. Genomic encyclopedia of type strains, Phase I: The one thousand microbial genomes (KMG-I) project. *Stand Genomic Sci*. 2014;9(3):628–634.
- Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, Hooper SD, Pati A, Lykidis A, Spring S, Anderson IJ, D'haeseleer P, Zemla A, Singer M, Lapidus A, Nolan M, Copeland A, Han C, Chen F, Cheng J-F, Lucas S, Kerfeld C, Lang E, Gronow S, Chain P, Bruce D, et al. A phylogeny-driven genomic encyclopaedia of *Bacteria* and *Archaea*. *Nature*. 2009;462:1056–60.
- Piao H, Froula J, Du C, Kim TW, Hawley ER, Bauer S, Wang Z, Ivanova N, Clark DS, Klenk HP, Hess M. Identification of novel biomass-degrading enzymes from genomic dark matter: Populating genomic sequence space with functional annotation. *Biotechnol Bioeng*. 2014;111:1550–65.
- Kyrpides NC, Hugenholtz P, Eisen J a, Woyke T, Göker M, Parker CT, Amann R, Beck BJ, Chain PSG, Chun J, Colwell RR, Danchin A, Dawyndt P, Dedeurwaerdere T, DeLong EF, Dettler JC, De Vos P, Donohue TJ, Dong X-Z, Ehrlich DS, Fraser C, Gibbs R, Gilbert J, Gilna P, Glöckner FO, Jansson JK, Keasling JD, Knight R, Labeda D, Lapidus A, et al. Genomic encyclopedia of bacteria and Archaea: sequencing a myriad of type strains. *PLoS Biol*. 2014;12:e1001920.
- Field D, Amaral-Zettler L, Cochrane G, Cole JR, Dawyndt P, Garrity GM, Gilbert J, Glöckner FO, Hirschman L, Karsch-Mizrachi I, Klenk HP, Knight R, Kottmann R, Kyrpides N, Meyer F, San Gil I, Sansone SA, Schriml LM, Sterk P, Tatusova T, Ussery DW, White O, Wooley J. The Genomic Standards Consortium. *PLoS Biol*. 2011;9:e1001088.
- Reddy TBK, Thomas AD, Stamatis D, Bertsch J, Isbandi M, Jansson J, Mallajosyula J, Pagani I, Lobos EA, Kyrpides NC: The Genomes OnLine Database (GOLD) v.5: a metadata management system based on a four level (meta)genome project classification. *Nucleic Acids Res*. 2014;43:D1099–106.
- Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, Clum A, Goodwin L, Woyke T, Lapidus A, Klenk HP, Cottingham RW, Kyrpides NC: The fast changing landscape of sequencing technologies and their impact on microbial genome assemblies and annotation. *PLoS One*. 2012;7:e48837.
- List of growth media used at the DSMZ. <http://www.dsmz.de/>.
- Gemeinholzer B, Dröge G, Zetzsche H, Haszprunar G, Klenk H-P, Güntsch A, Berendsohn WG, Wägele J-W. The DNA bank network: the start from a German initiative. *Biopreserv Biobank*. 2011;9:51–5.
- Bennett S. Solexa Ltd. *Pharmacogenomics*. 2004;5:433–8.
- DOE Joint Genome Institute, <http://www.jgi.doe.gov/>.
- Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 2008;18:821–9.

34. wgsim. [https://github.com/lh3/wgsim].
35. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB: High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A*. 2011;108:1513–8.
36. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010;11:119.
37. Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen IM-A, Pati A, Nielsen T, Markowitz VM, Kyrpides NC. The standard operating procedure of the DOE-JGI microbial genome annotation pipeline (MGAP v.4). *Sciences*. 2015;10:86.
38. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods*. 2010;7:455–7.
39. Markowitz VM, Mavromatis K, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: A system for microbial genome annotation expert review and curation. *Bioinformatics*. 2009;25:2271–8.
40. Meier-Kolthoff JP, Klenk H-P, Göker M. Taxonomic use of DNA G + C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol*. 2014;64(Pt 2):352–6.
41. Auch AF, Klenk H-P, Göker M. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. *Stand Genomic Sci*. 2010;2:142–8.
42. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*. 2013;14:60.
43. McBride MJ, Zhu Y. Gliding motility and Por secretion system genes are widespread among members of the phylum *Bacteroidetes*. *J Bacteriol*. 2013;195:270–8.
44. Cottrell MT, Kirchman DL. Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl Environ Microbiol*. 2000;66:1692–7.
45. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res*. 2009;37:D233–8.
46. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42:D490–5.
47. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res*. 2012;40:W445–51.
48. Ye Y, Choi J-H, Tang H. RAPSearch: a fast protein similarity search tool for short reads. *BMC Bioinformatics*. 2011;12:159.
49. Zhao Y, Tang H, Ye Y. RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics*. 2012;28:125–6.
50. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heeger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M: Pfam: the protein families database. *Nucleic Acids Res*. 2014;42:D222–30.
51. Bjursell MK, Martens EC, Gordon JI. Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *J Biol Chem*. 2006;281:36269–79.
52. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL. Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell*. 2010;141:1241–52.
53. Rawlings ND, Waller M, Barrett AJ, Bateman A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res*. 2014;42(Database issue):D503–9.
54. Touchon M, Barbier P, Bernardet J-F, Loux V, Vacherie B, Barbe V, Rocha EPC, DuChaud E. Complete genome sequence of the fish pathogen *Flavobacterium branchiophilum*. *Appl Environ Microbiol*. 2011;77:7656–62.
55. Mann AJ, Hahnke RL, Huang S, Werner J, Xing P, Barbeyron T, Huettel B, Stüber K, Reinhardt R, Harder J, Glöckner FO, Amann RI, Teeling H. The genome of the alga-associated marine flavobacterium *Formosa agariphila* KMM 3901<sup>T</sup> reveals a broad potential for degradation of algal polysaccharides. *Appl Environ Microbiol*. 2013;79:6813–22.
56. Xing P, Hahnke RL, Unfried F, Markert S, Huang S, Fuchs BM, Barbeyron T, Harder J, Schweder T, Glöckner FO, Amann RI, Teeling H. Niche separation of two *Polaribacter* strains isolated from the German Bight of the North Sea during a spring diatom bloom. *ISME J*. 2014;9:1410–22.
57. Kabisch A, Otto A, König S, Becher D, Albrecht D, Schüler M, Teeling H, Amann RI, Schweder T. Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* "*Gramella forsetii*" KT0803. *ISME J*. 2014;8:1492–502.
58. Stanier RY. Studies on nonfruiting myxobacteria I. *Cytophaga johnsonae*, n. sp., a chitin-decomposing myxobacterium. *J Bacteriol*. 1947;53:297–315.
59. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli S V, Ashburner M, Axelrod N, Baldauf S, Ballard S, Boore J, Cochrane G, Cole J, Dawyndt P, De Vos P, DePamphilis C, Edwards R, Faruque N, Feldman R, Gilbert J, Gilna P, Glöckner FO, Goldstein P, Guralnick R, Haft D, Hancock D, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol*. 2008;26:541–7.
60. Field D, Amaral-Zettler L, Cochrane G, Cole J, Dawyndt P, Garrity GM, Gilbert J, Glöckner FO, Hirschman L, Karsch-Mizrachi I, Klenk HP, Knight R, Kottmann R, Kyrpides N, Meyer F, Gil IS, Sansone SA, Schriml LM, Sterk P, Tatusova T, Ussery DW, White O, Wooley J: The Genomic Standards Consortium. *PLoS Biol*. 2011;9:8–10.
61. Euzéby JP. List of bacterial names with standing in nomenclature: a folder available on the internet. *Int J Syst Bacteriol*. 1997;47:590–2.
62. Garrity G. NamesforLife. BrowserTool takes expertise out of the database and puts it right in the browser. *Microbiol Today*. 2010;37:9.
63. Krieg NR, Ludwig W, Euzéby J, Whitman WB. Phylum XIV. *Bacteroidetes* phyl. nov. In: Krieg NR, Staley JT, Brown DR, editors. *Bergey's Manual of Systematic Bacteriology, The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*, vol. 4. 2nd ed. New York: Springer; 2010. p. 25–469.
64. Euzéby J. Validation List N° 143. *Int J Syst Evol Microbiol*. 2012;62:1–4.
65. Bernardet J-F. Class II. *Flavobacteriia* class. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig WWWB, editors. *Bergey's Manual Of Systematic Bacteriology, The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*, vol. 4. 2nd ed. New York: Springer; 2010. p. 106–314.
66. Euzéby J. Validation List N° 145. *Int J Syst Evol Microbiol*. 2012;62:1017–1019.
67. Bernardet JF. Order I. *Flavobacteriales* ord. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB, editors. *Bergey's Manual of Systematic Bacteriology, The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*, vol. 4. 2nd ed. New York: Springer; 2011. p. 105–329.
68. Holmes B, Owen RJ. Proposal that *Flavobacterium breve* be substituted as the type species of the genus in place of *Flavobacterium aquatile* and emended description of the genus *Flavobacterium*: status of the named species of *Flavobacterium*. Request for an Opinion. *Int J Syst Bacteriol*. 1979;29:416–26.
69. BAuA 2010 – 2012 update, Classification of bacteria and archaea in risk groups. [http://www.baua.de/TRBA\\_466:19](http://www.baua.de/TRBA_466:19).

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

