

Complete genome sequence of *Weeksella virosa* type strain (9751^T)

Elke Lang¹, Hazuki Teshima^{2,3}, Susan Lucas², Alla Lapidus², Nancy Hammon², Shweta Deshpande², Matt Nolan², Jan-Fang Cheng², Sam Pitluck², Konstantinos Liolios², Ioanna Pagani², Natalia Mikhailova², Natalia Ivanova², Konstantinos Mavromatis², Amrita Pati², Roxane Tapia^{2,3}, Cliff Han^{2,3}, Lynne Goodwin^{2,3}, Amy Chen⁴, Krishna Palaniappan⁴, Miriam Land^{2,5}, Loren Hauser^{2,5}, Yun-Juan Chang^{2,5}, Cynthia D. Jeffries^{2,5}, Evelyne-Marie Brambilla¹, Markus Kopitz¹, Manfred Rohde⁶, Markus Göker¹, Brian J. Tindall¹, John C. Detter^{2,3}, Tanja Woyke², James Bristow², Jonathan A. Eisen^{2,7}, Victor Markowitz⁴, Philip Hugenholtz^{2,8}, Hans-Peter Klenk¹, and Nikos C. Kyrpides^{2*}

¹ DSMZ - German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany

² DOE Joint Genome Institute, Walnut Creek, California, USA

³ Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico USA

⁴ Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA

⁵ Lawrence Livermore National Laboratory, Livermore, California, USA

⁶ HZI – Helmholtz Centre for Infection Research, Braunschweig, Germany

⁷ University of California Davis Genome Center, Davis, California, USA

⁸ Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia

*Corresponding author: Nikos C. Kyrpides

Keywords: strictly aerobic, slimy, Gram-negative, lyses proteins, inhabitant of mucosa, *Flavobacteriaceae*, GEBA

Weeksella virosa Holmes *et al.* 1987 is the sole member and type species of the genus *Weeksella* which belongs to the family *Flavobacteriaceae* of the phylum *Bacteroidetes*. Twenty-nine isolates, collected from clinical specimens provided the basis for the taxon description. While the species seems to be a saprophyte of the mucous membranes of healthy man and warm-blooded animals a causal relationship with disease has been reported in a few instances. Except for the ability to produce indole and to hydrolyze Tween and proteins such as casein and gelatin, this aerobic, non-motile, non-pigmented bacterial species is metabolically inert in most traditional biochemical tests. The 2,272,954 bp long genome with its 2,105 protein-coding and 76 RNA genes consists of one circular chromosome and is a part of the *Genomic Encyclopedia of Bacteria and Archaea* project.

Introduction

Strain 9751^T (= DSM 16922 = NCTC 11634 = JCM 21250) is the type strain of *Weeksella virosa*, which is the sole member and type species of the genus *Weeksella* [1,2]. The generic name was given in honor of O.B. Weeks for his contributions to the taxonomy of the genus *Flavobacterium*. The species epithet is derived from the Latin word 'virosa' meaning 'slimy', referring to the colony appearance of the species [1]. *W. virosa* strain 9751^T was isolated from a clinical specimen of urine and described by Holmes and coworkers in

1986 [1]. These authors collected 29 strains from clinical samples, mostly obtained in the USA but also in other nations and continents, as the basis for their species description. Most isolates came from genitourinary tract samples, predominantly from women. Since then, strains of *W. virosa* were detected by cultural methods in the oral cavity [3], the genitourinary tract [4,5] of man, in clinical specimens of pigs [6], the urine of a cow with bladder carcinoma [7], and in the midgut of a Brazilian dipteran [8]. *Weeksella*-like strains were

also found in food samples [9] and environmental samples, however, the latter matched the genus description but could not be assigned to the species *virosa* thus suggesting that they are representatives of additional species within the genus that are not yet described [10,11]. About 2% of healthy women carry *W. virosa* on their vaginal mucosa [4]. Two cases of peritonitis have been described as being caused by *W. virosa* [12,13]. Here we present a summary classification and a set of features for *W. virosa* 9751^T, together with the description of the complete genomic sequencing and annotation.

Classification and features

A representative genomic 16S rRNA sequence of strain 9751^T was compared using NCBI BLAST under default settings (e.g., considering only the high-scoring segment pairs (HSPs) from the best 250 hits) with the most recent release of the GreenGenes database [14] and the relative frequencies, weighted by BLAST scores, of taxa and keywords (reduced to their stem [15]) were determined. The five most frequent genera were *Chryseobacterium* (45.9%), *Wautersiella* (23.2%), *Myroides* (7.9%), *Elizabethkingia* (7.3%) and *Empedobacter* (7.0%) (134 hits in total). Regarding the single hit to sequences from members of the species, the average identity within HSPs was 99.2%, whereas the average coverage by HSPs was 96.8%. Among all other species, the one yielding the highest score was *Empedobacter brevis*, which corresponded to an identity of 92.0% and an HSP coverage of 97.0%. Accordingly, *E. brevis* groups as the sister genus in trees

constructed from 16S rRNA gene sequences [16-19] whereas *Bergeyella zoohelcum* (family *Flavobacteriaceae*) is looked at as the phenotypic sister taxon which may be confused with *W. virosa* when applying traditional morphological and physiological tests [11]. The highest-scoring environmental sequence was GQ383925 ('sewage water isolate XJ109 *Flavobacteriaceae* str. XJ109'), which showed an identity of 92.5% and a HSP coverage of 98.5%. The five most frequent keywords within the labels of environmental samples which yielded hits were 'skin' (11.2%), 'human' (3.7%), 'fossa' (3.4%), 'microbiom, tempor, topograph' (2.8%) and 'forearm' (2.1%) (116 hits in total). The two most frequent keywords within the labels of environmental samples which yielded hits of a higher score than the highest scoring species were 'sewag, water' (10.1%) and 'aerosol, air, citi, level, microarra, texa, two, urban' (10.0%) (2 hits in total). From these sequence-based observations and the published culture-dependent detections (see above) it can be concluded that *W. virosa* occurs mainly as an inhabitant of mucosa of animals and man.

Figure 1 shows the phylogenetic neighborhood of strain 9751^T in a 16S rRNA based tree. The sequences of the five 16S rRNA gene copies in the genome differ from each other by up to three nucleotides, and differ by up to two nucleotides from the previously published 16S rRNA sequence (M93152), which contains eleven ambiguous base calls, and differ by up to 314 nucleotides from another previously published 16S rRNA sequence (AF133539) which obviously represents a strain of the genus *Bacteroides*.

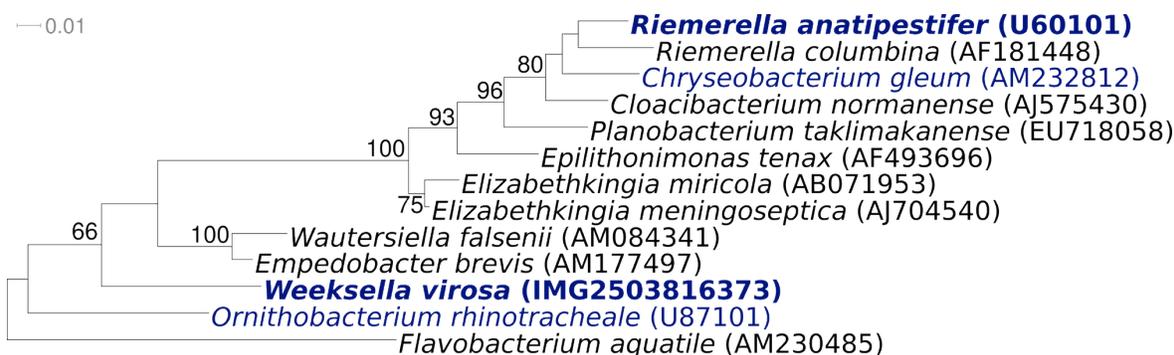


Figure 1. Phylogenetic tree highlighting the position of *W. virosa* relative to a selection of closely related other type strains within the family *Flavobacteriaceae*. The tree was inferred from 1,339 aligned characters [20,21] of the 16S rRNA gene sequence under the maximum likelihood criterion [22] and rooted in accordance with the type strain of the family. The branches are scaled in terms of the expected number of substitutions per site. Numbers above branches are support values from 700 bootstrap replicates [13] if larger than 60%. Lineages with type strain genome sequencing projects registered in GOLD [23] are shown in blue, published genomes in bold. The genome of *Riemerella anatipestifer* will be described in this issue of SIGS [24].

The cells of *W. virosa* are rod-shaped (0.6×2.0 - $3.0 \mu\text{m}$) with parallel sides and rounded ends (Figure 2). According to Holmes [1], the cells stain Gram-negative, are nonmotile and non spore-forming (Table 1). No intracellular inclusions such as polyhydroxybutyrate were detected. Colonies on nutrient agar appear circular and low convex with entire edges, smooth, shining and mucoid, reaching 2mm in diameter after 24h [1]. While the colonies are described as being non-pigmented on nutrient agar by most authors [1,37], the production of a nondiffusible yellow pigment was reported during growth on blood agar [5]. Most strains produce a diffusible dark brown pigment on tyrosine-containing agar [1]. The strains grow at 42°C but not at 5°C [1]. *W. virosa* is a strictly aerobic chemoorganotroph and is not able to reduce nitrate, nitrite or selenite and does not acidify glucose or other sugars under standard conditions [1,5]. However, under test conditions developed for fastidious organisms such as *Neisseria* (API NH strips, Biomérieux), acid is produced from glucose [IDA] as is observed for the phylogenetic neighbors *E. brevis* and *Wautersiella falseonii* [16,17]. Cytochrome oxidase, catalase and phosphatase are present [1]. Tolerance to NaCl and pH ranges have not been reported. *W. virosa* grows on McConkey agar, indicating tolerance to bile salts [1]. On the other hand, KCN (75 mg l^{-1}), cetrимide, colistin [17], and polymyxin [5] are not tolerated and inhibit growth. The species is able to

utilize β -hydroxybutyrate as a substrate and to hydrolyze casein, gelatin and Tween 20 [1,17]. Tryptophan is cleaved to give indole, pyruvate and ammonia when tested with Ehrlich's reagent, but the reaction can not be detected when Kovacs' reagent is used [1]. *W. virosa* is inert in most traditional biochemical tests [1], it does not utilize glucose as a substrate under standard conditions [16]. *W. virosa* does not hydrolyze starch, esculin or DNA, and is negative for gluconate oxidation, urease, phenylalanine deaminase, arginine deaminase, arginine dihydrolase, lysine or ornithine decarboxylase, β -D-galactosidase [1], alkalization of galacturonate [17]. The following enzymes are present as concluded from tests using API ZYM galleries: acid and alkaline phosphatase, lysine arylamidase, aspartate arylamidase, alanine arylamidase and methionine arylamidase [1]. In addition, the following substrates were hydrolyzed in the latter galleries by *W. virosa*: naphthol-AS-BI-phosphodiamide, bis-(para-nitrophenyl)-phosphate, glycyl-glycyl- β -naphthylamide hydrobromide, glycyl-L-phenylalanyl- β -naphthylamide, glycyl-L-prolyl- β -naphthylamide, L-leucyl-glycyl- β -naphthylamide, α -L-glutamyl- β -naphthylamide, and N-carbobenzoxy-glycyl-glycyl-L-arginine- β -naphthylamide [1]. *W. virosa* is susceptible to most β -lactams, tetracycline, chloramphenicol, nalidixic acid, erythromycin and sulfamethoxazole-trimethoprim [11]. In contrast, the species is resistant to aminoglycosides [11].

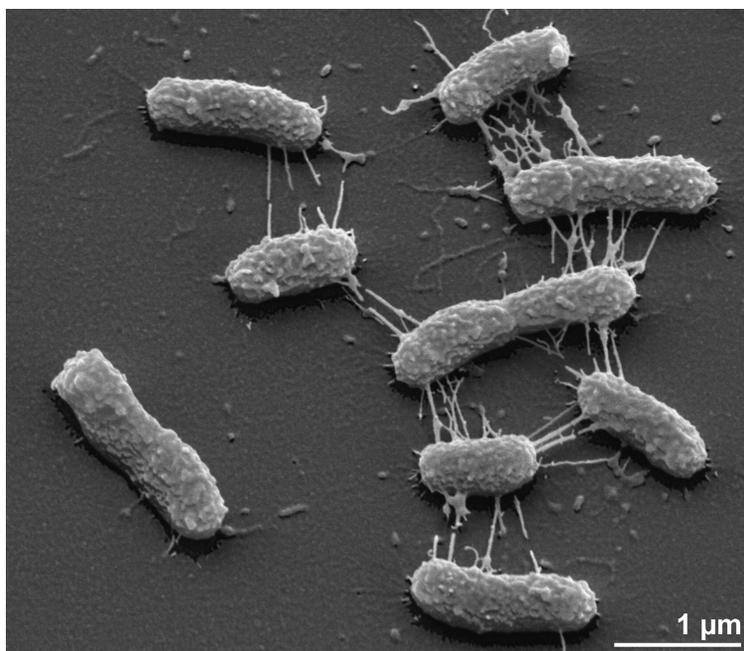


Figure 2. Scanning electron micrograph of *W. virosa* strain 9751^T

Table 1. Classification and general features of *W. virosa* 9751^T according to the MIGS recommendations [25].

MIGS ID	Property	Term	Evidence code
		Domain <i>Bacteria</i>	TAS [26]
		Phylum <i>Bacteroidetes</i>	TAS [27]
		Class ' <i>Flavobacteria</i> '	TAS [28]
	Current classification	Order ' <i>Flavobacteriales</i> '	TAS [29]
		Family <i>Flavobacteriaceae</i>	TAS [30-33]
		Genus <i>Weeksella</i>	TAS [1,34]
		Species <i>Weeksella virosa</i>	TAS [1,34]
		Type strain 9751 (= CL345/78)	TAS [1,34]
	Gram stain	negative	TAS [1]
	Cell shape	rod-shaped	TAS [1]
	Motility	non-motile	TAS [1]
	Sporulation	none	TAS [1]
	Temperature range	10°C or less up to 42°C	TAS [1]
	Optimum temperature	not reported	NAS
	Salinity	not reported	NAS
MIGS-22	Oxygen requirement	aerobic	TAS [1]
	Carbon source	proteins	TAS [1]
	Energy source	chemoorganotroph	TAS [1]
MIGS-6	Habitat	human urogenital tract	TAS [1,4]
MIGS-15	Biotic relationship	saprophyte	TAS [1,11]
MIGS-14	Pathogenicity	none	NAS
	Biosafety level	1	TAS [35]
	Isolation	human urine	TAS [1]
MIGS-4	Geographic location	North Carolina, USA	TAS [1]
MIGS-5	Sample collection time	1986 or before	TAS [1]
MIGS-4.1	Latitude	not reported	NAS
MIGS-4.2	Longitude	not reported	NAS
MIGS-4.3	Depth	not reported	NAS
MIGS-4.4	Altitude	not reported	NAS

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). These evidence codes are from of the Gene Ontology project [36]. If the evidence code is IDA, then the property was directly observed by one of the authors or an expert mentioned in the acknowledgements.

Chemotaxonomy

The major respiratory quinone of *W. virosa* is menaquinone 6 and the major polyamine is homospermidine, as is the case for all members of the family *Flavobacteriaceae* [11,38-40]. No sphingophospholipids were detected [1]. The polar lipids of *W. virosa* have not yet been described. The major whole-cell fatty acids of *W. virosa* are *iso*-C_{15:0} (46%), *iso*-C_{15:0}2-OH (10%), *iso*-C_{17:1}ω12t (8%) and *iso*-C_{17:0}3-OH (7%) as described for CDC group IIf, the preliminary name given to these strains prior to being formally named *W. virosa* [41]. A comparison of the patterns of *W. virosa* and '*W.*

zoohecum' obtained at that time [41] with more recently published patterns of *B. zoohecum* and *E. brevis* and phylogenetic neighbors [17,19] seems to cast doubts on the comparability of these early patterns. They are the only ones listing the presence of high amounts of *iso*-C_{15:0}2-OH and *iso*-C_{17:1}ω12t, which are not listed for phylogenetically related genera later on [19]. However, *iso*-C_{15:0}2-OH and isomers of *iso*-heptadecene are included in the summed features of the Microbial Identification System applied in many recent analyses including [17,19].

Genome sequencing and annotation

Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position [42], and is part of the *Genomic Encyclopedia of Bacteria and Archaea* project [40]. The genome project is deposited in the Genomes OnLine Database [23] and the

complete genome sequence is deposited in GenBank. Sequencing, finishing and annotation were performed by the DOE Joint Genome Institute (JGI). A summary of the project information is shown in Table 2.

Table 2. Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	Tree genomic libraries: one 454 pyrosequence standard library, one 454 PE library (17 kb insert size), one Illumina library
MIGS-29	Sequencing platforms	Illumina GAii, 454 GS FLX Titanium
MIGS-31.2	Sequencing coverage	2,107.5 × Illumina; 64.3 × pyrosequence
MIGS-30	Assemblers	Newbler version 2.5-internal-10Apr08-1-threads, Velvet version 0.7.63, phrap
MIGS-32	Gene calling method	Prodigal 1.4, GenePRIMP
	INSDC ID	CP02455
	Genbank Date of Release	February 15, 2011
	GOLD ID	Gc01619
	NCBI project ID	50581
	Database: IMG-GEBA	2503754024
MIGS-13	Source material identifier	DSM 16922
	Project relevance	Tree of Life, GEBA

Growth conditions and DNA isolation

W. virosa 9751^T, DSM 16922, was grown on DSMZ medium 220 (Caso Agar) [37] at 30°C. DNA was isolated from 0.5-1 g of cell paste using MasterPure Gram-positive DNA purification kit (Epicentre MGP04100) following the standard protocol as recommended by the manufacturer, with modification st/DL for cell lysis as described in Wu *et al.* [43]. DNA is available through the DNA Bank Network [44,45].

Genome sequencing and assembly

The genome was sequenced using a combination of Illumina and 454 sequencing platforms. All general aspects of library construction and sequencing can be found at the JGI website [46]. Pyrosequencing reads were assembled using the Newbler assembler version 2.5-internal-10Apr08-1-threads (Roche). The initial Newbler assembly consisting of 27 contigs in one scaffold was converted into a phrap assembly by making fake reads from the consensus, to collect the read pairs in the 454 paired end library. Illumina GAii sequencing data (4,788 Mb) was assembled with Velvet [47] and the consensus sequences were

shredded into 1.5 kb overlapped fake reads and assembled together with the 454 data. The 454 draft assembly was based on 131.6 Mb 454 draft data and all of the 454 paired end data. Newbler parameters are -consed -a 50 -l 350 -g -m -ml 20. The Phred/Phrap/Consed software package [48] was used for sequence assembly and quality assessment in the subsequent finishing process. After the shotgun stage, reads were assembled with parallel phrap (High Performance Software, LLC). Possible mis-assemblies were corrected with gapResolution, Dupfinisher [49], or sequencing cloned bridging PCR fragments with subcloning or transposon bombing (Epicentre Biotechnologies, Madison, WI). Gaps between contigs were closed by editing in Consed, by PCR and by Bubble PCR primer walks (J.-F.Chang, unpublished). A total of 60 additional reactions were necessary to close gaps and to raise the quality of the finished sequence. Illumina reads were also used to correct potential base errors and increase consensus quality using a software Polisher developed at JGI [50]. The error rate of the completed genome sequence is less than one in 100,000. Together, the combi-

nation of the Illumina and 454 sequencing platforms provided $2,171.8 \times$ coverage of the genome. The final assembly contained 384,925 pyrosequence and 63,008,730 Illumina reads.

Genome annotation

Genes were identified using Prodigal [51] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [52]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant database, UniProt, TIGR-Fam, Pfam, PRIAM, KEGG, COG, and InterPro databases. Additional gene prediction analysis and

functional annotation was performed within the Integrated Microbial Genomes - Expert Review (IMG-ER) platform [53].

Genome properties

The genome consists of a 2,272,954 bp long chromosome with a GC content of 35.9% (Figure 3 and Table 3). Of the 2,181 genes predicted, 2,105 were protein-coding genes, and 76 RNAs; 56 pseudo-genes were also identified. The majority of the protein-coding genes (65.5%) were assigned with a putative function while the remaining ones were annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 4.

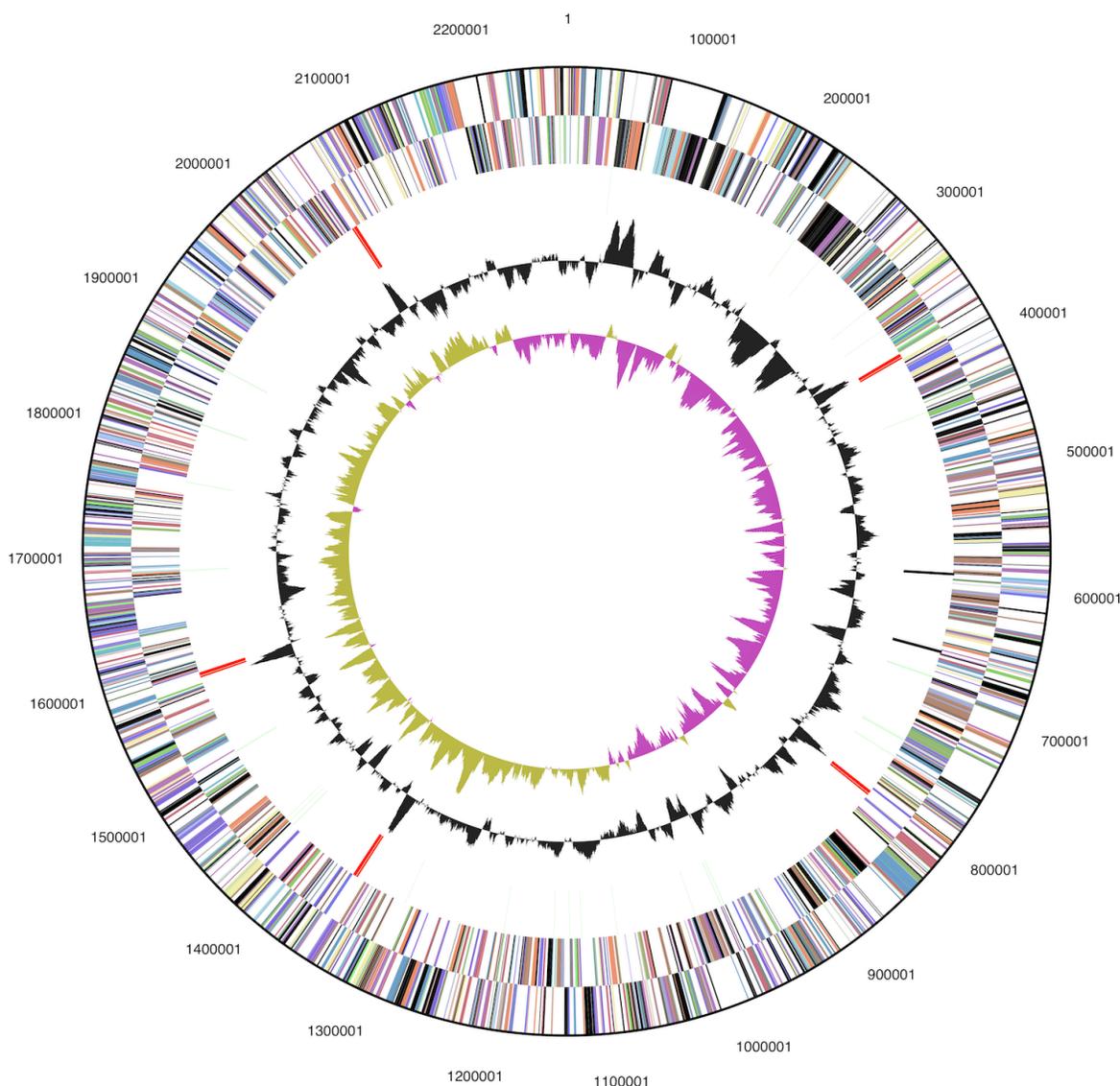


Figure 3. Graphical circular map of the chromosome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

Table 3. Genome Statistics

Attribute	Value	% of Total
Genome size (bp)	2,272,954	100.00%
DNA coding region (bp)	2,037,648	89.65%
DNA G+C content (bp)	816,484	35.92%
Number of replicons	1	
Extrachromosomal elements	0	
Total genes	2,181	100.00%
RNA genes	76	3.48%
rRNA operons	5	
Protein-coding genes	2,105	96.52%
Pseudo genes	56	2.57%
Genes with function prediction	1,429	65.52%
Genes in paralog clusters	99	4.54%
Genes assigned to COGs	1,403	64.33%
Genes assigned Pfam domains	1,534	70.33%
Genes with signal peptides	506	23.20%
Genes with transmembrane helices	451	20.68%
CRISPR repeats	1	

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

Code	value	%age	Description
J	142	9.5	Translation, ribosomal structure and biogenesis
A	0	0.0	RNA processing and modification
K	54	3.6	Transcription
L	94	6.3	Replication, recombination and repair
B	0	0.0	Chromatin structure and dynamics
D	18	1.2	Cell cycle control, cell division, chromosome partitioning
Y	0	0.0	Nuclear structure
V	29	1.9	Defense mechanisms
T	29	1.9	Signal transduction mechanisms
M	154	10.3	Cell wall/membrane/envelope biogenesis
N	4	0.3	Cell motility
Z	0	0.0	Cytoskeleton
W	0	0.0	Extracellular structures
U	25	1.7	Intracellular trafficking, secretion, and vesicular transport
O	73	4.9	Posttranslational modification, protein turnover, chaperones
C	82	5.5	Energy production and conversion
G	47	3.1	Carbohydrate transport and metabolism
E	127	8.5	Amino acid transport and metabolism
F	57	3.8	Nucleotide transport and metabolism
H	95	6.3	Coenzyme transport and metabolism
I	72	4.8	Lipid transport and metabolism
P	101	6.7	Inorganic ion transport and metabolism
Q	24	1.6	Secondary metabolites biosynthesis, transport and catabolism
R	171	11.4	General function prediction only
S	103	6.9	Function unknown
-	778	35.7	Not in COGs

Acknowledgements

This work was performed under the auspices of the US Department of Energy Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-

AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396, UT-Battelle and Oak Ridge National Laboratory under contract DE-AC05-000R22725, as well as German Research Foundation (DFG) INST 599/1-2.

References

1. Holmes B, Steigerwaldt AG, Weaver RE, Brenner DJ. *Weeksella virosa* gen. nov., sp. nov. (formerly group If) found in human clinical specimens. *Syst Appl Microbiol* 1986; **8**:185-190.
2. Garrity G. NamesforLife. BrowserTool takes expertise out of the database and puts it right in the browser. *Microbiol Today* 2010; **37**:9.
3. Leung WK, Jin LJ, Yam WC, Samaranayake LP. Oral colonization of aerobic and facultatively anaerobic gram-negative rods and cocci in irradiated, dentate, xerostomic individuals. *Oral Microbiol Immunol* 2001; **16**:1-9. [PubMed doi:10.1034/j.1399-302x.2001.160101.x](#)
4. Mardy C, Holmes B. Incidence of vaginal *Weeksella virosa* (formerly group If). *J Clin Pathol* 1988; **41**:211-214. [PubMed doi:10.1136/jcp.41.2.211](#)
5. Reina J, Gil J, Salva F, Gomez J, Alomar P. Microbiological characteristics of *Weeksella virosa* (formerly CDC group If) isolated from the human genitourinary tract. *J Clin Microbiol* 1990; **28**:2357-2359. [PubMed](#)
6. Vela AI, García N, Latre MV, Casamayor A, Sánchez-Porro C, Briones V, Ventosa A, Domínguez L, Fernández-Garayzábal JF. *Aerococcus suis* sp. nov., isolated from clinical specimens from swine. *Int J Syst Evol Microbiol* 2007; **57**:1291-1294. [PubMed doi:10.1099/ijs.0.64537-0](#)
7. Brun R, Urraro C, Medaglia C, Russo V, Borzachiello G, Roperto F, Roperto S. *Lymphoepithelioma*-like carcinoma of the urinary bladder in a cow associated with bovine papillomavirus type-2. *J Comp Pathol* 2008; **139**:121-125. Published online August 3, 2008. [PubMed doi:10.1016/j.jcpa.2008.06.002](#)
8. Gouveia C, Asensi MD, Zahner V, Rangel EF, Oliveira SM. Study on the bacterial midgut microbiota associated to different Brazilian populations of *Lutzomyia longipalpis* (Lutz & Neiva) (Diptera: Psychodidae). *Neotrop Entomol* 2008; **37**:597-601. [PubMed doi:10.1590/S1519-566X2008000500016](#)
9. Botha WC, Jooste PJ, Hugo CJ. The incidence of *Weeksella*- and *Bergeyella*-like bacteria in the food environment. *J Appl Microbiol* 1998; **84**:349-356. [PubMed doi:10.1046/j.1365-2672.1998.00346.x](#)
10. Botha WC, Jooste PJ, Britz TJ. The taxonomic relationship of certain environmental flavobacteria to the genus *Weeksella*. *J Appl Bacteriol* 1989; **67**:551-559. [PubMed](#)
11. Hugo CJ, Bruun B, Jooste PJ. The genera *Bergeyella* and *Weeksella*. In: The Prokaryotes, vol. 7: *Proteobacteria*: Delta, Epsilon Subclass. Dworkin M, Falkow S, Rosenberg E, Schleifer K-H and Stackebrandt E (eds). 2006; 523-538.
12. Boixeda D, de Luis DA, Meseguer MA, Aller R, Martin de Argila C. Lopez Sanroman. A case of spontaneous peritonitis caused by *Weeksella virosa*. *Eur J Gastroenterol Hepatol* 1998; **10**:897-898. [PubMed doi:10.1097/00042737-199810000-00016](#)
13. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? *Lect Notes Comput Sci* 2009; **5541**:184-200. [doi:10.1007/978-3-642-02008-7_13](#)
14. 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**:5069-5072. [PubMed doi:10.1128/AEM.03006-05](#)
15. Porter MF. An algorithm for suffix stripping. *Program: electronic library and information systems* 1980; **14**:130-137.
16. Bernardet JF, Nakagawa Y. An introduction to the family *Flavobacteriaceae*. In: The Prokaryotes, vol. 7: *Proteobacteria*: Delta, Epsilon Subclass. Dworkin M, Falkow S, Rosenberg E, Schleifer K-H and Stackebrandt E (eds). 2006; 455-480.
17. Kämpfer P, Avesani V, Janssens M, Charlier J, De Baere T, Vaneechoutte M. Description of *Wautersiella falsenii* gen. nov., sp. nov., to accommodate clinical isolates phenotypically resembling members of the genera *Chryseobacterium* and *Empe-*

- dobacter*. *Int J Syst Evol Microbiol* 2006; **56**:2323-2329. [PubMed](#) [doi:10.1099/ijs.0.64393-0](#)
18. O'Sullivan LA, Rinna J, Humphreys G, Weightman AJ, Fry JC. Culturable phylogenetic diversity of the phylum 'Bacteroidetes' from river epilithon and coastal water and description of novel members of the family *Flavobacteriaceae*: *Epilithonimonas tenax* gen. nov., sp. nov. and *Persicivirga xylanidelens* gen. nov., sp. nov. *Int J Syst Evol Microbiol* 2006; **56**:169-180. [PubMed](#) [doi:10.1099/ijs.0.63941-0](#)
 19. Hugo CJ, Segers P, Hoste B, Vancanneyt M, Kersters K. *Chryseobacterium joostei* sp. nov., isolated from the dairy environment. *Int J Syst Evol Microbiol* 2003; **53**:771-777. [PubMed](#) [doi:10.1099/ijs.0.02232-0](#)
 20. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000; **17**:540-552. [PubMed](#)
 21. Lee C, Grasso C, Sharlow MF. Multiple sequence alignment using partial order graphs. *Bioinformatics* 2002; **18**:452-464. [PubMed](#) [doi:10.1093/bioinformatics/18.3.452](#)
 22. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 2008; **57**:758-771. [PubMed](#) [doi:10.1080/10635150802429642](#)
 23. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* 2010; **38**:D346-D354. [PubMed](#) [doi:10.1093/nar/gkp848](#)
 24. Han C, Goodwin L, Pitluck S, Liolios L, Pagani I, Ivanova N, Mikhailova N, Pati A, Chen A, Paliannappan K, et al. Complete genome sequence of *Riemerella anatipestifer* type strain (Bruner^T). *Stand Genomic Sci* 2011; **4**(2).
 25. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](#) [doi:10.1038/nbt1360](#)
 26. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](#) [doi:10.1073/pnas.87.12.4576](#)
 27. Garrity GM, Holt JG. The Road Map to the Manual. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, 2001, p. 119-169.
 28. Ludwig W, Euzéby J, Whitman WG. Draft taxonomic outline of the *Bacteroidetes*, *Planctomycetes*, *Chlamydiae*, *Spirochaetes*, *Fibrobacteres*, *Fusobacteria*, *Acidobacteria*, *Verrucomicrobia*, *Dictyoglomi*, and *Gemmatimonadetes*. http://www.bergeys.org/outlines/Bergeys_Vol_4_Outline.pdf. Taxonomic Outline 2008
 29. Garrity GM, Holt JG. 2001. Taxonomic outline of the *Archaea* and *Bacteria*, p. 155-166. In G. M. Garrity, D. R. Boone, and R. W. Castenholz (ed.), *Bergey's Manual of Systematic Bacteriology*, 2nd ed, vol. 1. Springer, New York.
 30. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. List No. 41. *Int J Syst Bacteriol* 1992; **42**:327-328. [doi:10.1099/00207713-42-2-327](#)
 31. Reichenbach H. Order 1. *Cytophagales* Leadbetter 1974, 99AL. In: Holt JG (ed), *Bergey's Manual of Systematic Bacteriology*, First Edition, Volume 3, The Williams and Wilkins Co., Baltimore, 1989, p. 2011-2013.
 32. 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**:128-148. [doi:10.1099/00207713-46-1-128](#)
 33. Bernardet JF, Nakagawa Y, Holmes B. Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae*, and emended description of the family. *Int J Syst Evol Microbiol* 2002; **52**:1049-1070. [PubMed](#) [doi:10.1099/ijs.0.02136-0](#)
 34. Validation list No° 23. *Int J Syst Bacteriol* 1987; **37**:179-180. [doi:10.1099/00207713-37-2-179](#)
 35. Classification of bacteria and archaea in risk groups. <http://www.baua.de> TRBA 466.
 36. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene Ontology: tool for the unification of biology. *Nat Genet* 2000; **25**:25-29. [PubMed](#) [doi:10.1038/75556](#)

37. List of growth media used at DSMZ: http://www.dsmz.de/microorganisms/media_list.php.
38. Vandamme P, Bernardet JF, Segers P, Kersters K, Holmes B. New perspectives in the classification of the flavobacteria: description of *Chryseobacterium* gen. nov., *Bergeyella* gen. nov., and *Empedobacter* nom. rev. *Int J Syst Bacteriol* 1994; **44**:827-831. [doi:10.1099/00207713-44-4-827](https://doi.org/10.1099/00207713-44-4-827)
39. Bernardet JF, Nakagawa Y, Holmes B. Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. *Int J Syst Evol Microbiol* 2002; **52**:1049-1070. [PubMed doi:10.1099/ijs.0.02136-0](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
40. Hamana K, Nakagawa Y. Polyamine distribution profiles in the eighteen genera phylogenetically located within the *Flavobacterium-Flexibacter-Cytophaga* complex. *Microbios* 2001; **106**:7-17. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
41. Hollis DG, Daneshvar MI, Moss CW, Baker CN. Phenotypic characteristics, fatty acid composition, and isoprenoid quinone content of CDC group IIg bacteria. *J Clin Microbiol* 1995; **33**:762-764. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
42. Klenk HP, Göker M. En route to a genome-based classification of *Archaea* and *Bacteria*? *Syst Appl Microbiol* 2010; **33**:175-182. [PubMed doi:10.1016/j.syapm.2010.03.003](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
43. Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, et al. A phylogeny-driven genomic encyclopaedia of *Bacteria* and *Archaea*. *Nature* 2009; **462**:1056-1060. [PubMed doi:10.1038/nature08656](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
44. Gemeinholzer B, Dröge G, Zetzsche H, Haszprunar G, Klenk HP, Güntsch A, Berendsohn WG, Wägele JW. The DNA Bank Network: the start from a German initiative. *Biopreservation and Biobanking* (In press).
45. DNA bank Network. <http://www.dnabank-network.org>
46. DOE Joint Genome Institute. <http://www.jgi.doe.gov>
47. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008; **18**:821-829. [PubMed doi:10.1101/gr.074492.107](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
48. Phrap and Phred for Windows, MacOS, Linux, and Unix. <http://www.phrap.com>
49. Han C, Chain P. 2006. Finishing repeat regions automatically with Dupfinisher. Proceeding of the 2006 international conference on bioinformatics & computational biology. Edited by Hamid R. Arabnia & Homayoun Valafar, CSREA Press. June 26-29, 2006: 141-146.
50. Lapidus A, LaButti K, Foster B, Lowry S, Trong S, Goltsman E. POLISHER: An effective tool for using ultra short reads in microbial genome assembly and finishing. AGBT, Marco Island, FL, 2008.
51. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**:119. [PubMed doi:10.1186/1471-2105-11-119](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
52. 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**:455-457. [PubMed doi:10.1038/nmeth.1457](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
53. Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 2009; **25**:2271-2278. [PubMed doi:10.1093/bioinformatics/btp393](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)
54. Faber MD, del Busto R, Cruz C, Mezger E. Response of *Weeksella virosa* peritonitis to imipenem/cilastin. *Adv Perit Dial* 1991; **7**:133-134. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101099/ijs.0.02136-0/)