

Noncontiguous finished genome sequence and description of *Weeksella massiliensis* sp. nov.

S. A. Sankar¹, C. I. Lo^{1,2}, B. Fall³, B. Sambe-Ba³,
O. Mediannikov^{1,2}, I. Diallo³, N. Labas¹, N. Faye⁴, B. Wade³,
D. Raoult^{1,2,5}, P.-E. Fournier¹ and F. Fenollar^{1,2}

1)Aix-Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine, Marseille, France, 2)Campus International UCAD-IRD, 3)Hôpital Principal, 4)Université Cheikh Anta Diop de Dakar, Laboratoire de Parasitologie générale, Dakar, Senegal and 5)Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

Strain FF8^T (= CSUR P860 = DSM 28259) was isolated in Dakar, Senegal, from the urine of a 65-year-old man with acute cystitis. This strain shows a similarity of sequence of 16S rRNA of 98.38% with *Weeksella virosa*, and its GenBank accession numbers are HG931340 and CCMH00000000. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis had a poor score, ranging from 1.32 to 1.56, that did not allow identification of the bacterium. Using a polyphasic study made of phenotypic and genomic analyses, strain FF8^T was a Gram-negative, aerobic rod and a member of the family *Flavobacteriaceae*. The sequenced genome is 2562781 bp with one chromosome but no plasmid. It exhibits a G + C content of 35.9% and contains 2390 protein-coding and 56 RNA genes, including a complete rRNA operon. On the basis of these data, we propose the creation of *Weeksella massiliensis* sp. nov.

New Microbes and New Infections © 2015 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

Keywords: Culturomics, cystitis, genome, human, taxonogenomics, urine, *Weeksella massiliensis*

Original Submission: 29 May 2015; **Revised Submission:**

21 September 2015; **Accepted:** 23 September 2015

Article published online:

Corresponding author: F. Fenollar, Aix-Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine, 27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

E-mail: florence.fenollar@univ-amu.fr

The first two authors contributed equally to this article, and both should be considered first author.

Introduction

The genus *Weeksella* (Holmes *et al.*, 1986) was first described in 1986 [1]. To date, this genus includes one species, *Weeksella virosa*, which has been isolated from human clinical specimens [1,2].

The current classification of prokaryotes relies on a polyphasic strategy combining phenotypic and genotypic characteristics [3,4]. These include 16S rRNA sequence similarity, G + C content and DNA-DNA hybridization (DDH). However, these tools have significant drawbacks, notably that the recommended threshold values do not apply to all species or genera [5,6].

Thanks to the progress made in sequencing technologies and their lowering costs, almost 40 000 bacterial genome sequences are currently available, covering many phyla [7]. Recently we proposed to integrate phenotypic characteristics, notably the MALDI-TOF spectrum, and genomic analysis and comparison in the taxonomic description of bacterial species [5,8,9]. We named this strategy taxonogenomics [5].

Strain FF8^T (= CSUR P860 = DSM 28259) was isolated from the urine of a 65-year-old man treated at the Hôpital Principal de Dakar, Senegal. That is a Gram-negative bacterium, aerobic, indole negative, nonmotile and rod shaped. This bacterium was cultivated as part of the MALDI-TOF implementation at Hôpital Principal de Dakar aiming to improve the routine laboratory identification of microorganisms [10].

Here we present a summary classification and a set of features for *Weeksella massiliensis* sp. nov., together with the description of the complete genome sequencing and annotation. These characteristics support the circumscription of the species *Weeksella massiliensis*.

Organism Information

Classification and features

In July 2013, a urine sample was collected from a 65-year-old Senegalese man with acute cystitis. From this clinical sample strain FF8 (Table 1) was isolated by cultivation on 5% sheep's blood-enriched Columbia agar (bioMérieux, Marcy l'Etoile, France). When analysed by MALDI-TOF, no identification was obtained because the strain displayed low scores.

TABLE 1. Classification and general features of *Weeksella massiliensis* strain FF8^T [19]

MIGS ID	Property	Term	Evidence code ^a
	Classification	Domain: <i>Bacteria</i> Phylum: <i>Bacteroidetes</i> Class: <i>Flavobacteriia</i> Order: <i>Flavobacteriales</i> Family: <i>Flavobacteriaceae</i> Genus: <i>Weeksella</i> Species: <i>Weeksella massiliensis</i>	TAS [34] TAS [11,12] TAS [35,36] TAS [35–37] TAS [38] TAS [1] IDA
		(Type) strain: FF8 ^T	IDA
	Gram stain	Negative	IDA
	Cell shape	Rod	IDA
	Motility	Not motile	IDA
	Sporulation	Non-spore forming	NAS
	Temperature range	Mesophile	IDA
	Optimum temperature	37°C	IDA
	Optimum pH range	7.2–7.4; 7.3	
	Carbon source	Unknown	
MIGS-6	Habitat	Human	IDA
MIGS-6.3	Salinity	Unknown	
MIGS-22	Oxygen requirement	Aerobic	TAS
MIGS-15	Biotic relationship	Free living	TAS
MIGS-14	Pathogenicity	Unknown	
MIGS-4	Geographic location	Dakar	TAS
MIGS-5	Sample collection	November 28, 2013	TAS
MIGS-4.1	Latitude	14.6937000	TAS
MIGS-4.1	Longitude	–17.4440600	TAS
MIGS-4.4	Altitude	12 m above sea level	TAS

MIGS, minimum information about a genome sequence.

^aEvidence codes are as follows: IDA, inferred from direct assay; TAS, traceable author statement (i.e., direct report exists in the literature); NAS, nontraceable author statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species or on anecdotal evidence). These evidence codes are from <http://www.geneontology.org/GO.evidence.shtml> of the Gene Ontology project [39]. If the evidence is IDA, then the property was directly observed for a live isolate by one of the authors or an expert mentioned in the acknowledgements.

Strain FF8 exhibited a 98.38% 16S rRNA sequence similarity with *Weeksella virosa* strain DSM 16922^T (GenBank accession number NR_074495), the phylogenetically closest bacterial species with standing in nomenclature. These values were lower than the 98.7% 16S rRNA gene sequence threshold recommended by Meier-Kolthoff *et al.*, 2013, to delineate a new species within phylum *Bacteroidetes* [11,12] without carrying out DDH [13]. A phylogenetic tree based on the 16S rRNA sequence highlights the position of *Weeksella massiliensis* strain FF8^T among the family *Flavobacteriaceae* (Fig. 1). Different growth temperatures (25, 30, 37, 45 and 56°C) were tested. Growth was obtained between 25 and 37°C, with an optimal growth at 37°C. Growth of the strain was tested also under anaerobic and microaerophilic conditions using GENbag anaer and GENbag microaer systems, respectively (bioMérieux), and under aerobic conditions, with or without 5% CO₂. Thus the optimal growth was observed under aerobic and microaerophilic conditions. No growth was observed under anaerobic conditions. The colonies were opaque, light yellow in color with a smooth surface, not haemolytic on 5% sheep's blood-enriched Columbia agar (bioMérieux) and approximately 2 mm in diameter. A motility test was negative. Cells are

Gram-negative, non-spore-forming rods of regular shape with rounded ends (Fig. 2) and have mean diameter of 0.3 µm (range, 0.2–0.5 µm) and a mean length of 1.5 µm (range, 0.8–2.1 µm) (Fig. 3). *Weeksella massiliensis* strain FF8^T does not grow on MacConkey agar [14].

Chemotaxonomic information

This bacterium possesses catalase and oxidase. Using an API ZYM strip (bioMérieux), positive reactions were observed for alkaline phosphatase, esterase, esterase–lipase, leucine arylamidase, acid phosphatase and naphthol-AS-BI-phosphohydrolase. Negative reactions were noted for α-chymotrypsin, cystine arylamidase, valine arylamidase, trypsin, α-glucosidase, β-glucosidase, α-galactosidase, β-galactosidase, β-glucuronidase, α-mannosidase, α-fucosidase and N-acetyl-β-glucosaminidase. Strain FF8^T is susceptible to ceftriaxone, amoxicillin/clavulanic acid, penicillin, imipenem, gentamicin and doxycycline but resistant to nitrofurantoin, vancomycin, trimethoprim/sulfamethoxazole and metronidazole. The minimum inhibitory concentrations (MICs) for some antibiotics tested by *Weeksella massiliensis* strain FF8^T sp. nov. are listed in Table 2. A comparison of phenotypic characteristics with *W. virosa*, *Bergeyella zoohelcum* [15] and *Moheibacter sediminis* [16] is presented in Table 3.

Extended features descriptions

MALDI-TOF protein analysis was performed with a Microflex LT (Bruker Daltonics, Leipzig, Germany), as previously reported [17,18]. The scores previously established by Bruker allowing validating (or not) the identification of species compared to the database of the instrument were applied. Briefly, a score of ≥ 2.000 with a species with a validly published name provided allows the identification at the species level; a score of ≥ 1.700 and < 2.000 allows the identification at the genus level; and a score of < 1.700 does not allow any identification. We performed 12 distinct deposits from 12 isolated colonies of strain FF8^T. Two microliters of matrix solution (saturated solution of α-cyano-4-hydroxycinnamic acid) in 50% acetonitrile and 2.5% trifluoroacetic acid were distributed on each smear and submitted at air drying for 5 minutes. Then the spectra from the 12 different colonies were imported into the MALDI BioTyper software (version 2.0, Bruker) and analysed by standard pattern matching (with default parameter settings) against the main spectra of 6252 bacteria. Scores ranging from 1.32 to 1.56 were obtained for strain FF8^T, suggesting that it was not a member of any known species. The reference mass spectrum from strain FF8^T was incremented in our database (Fig. 4). The gel view highlighted spectrum differences with other *Flavobacteriaceae* species (Fig. 5).

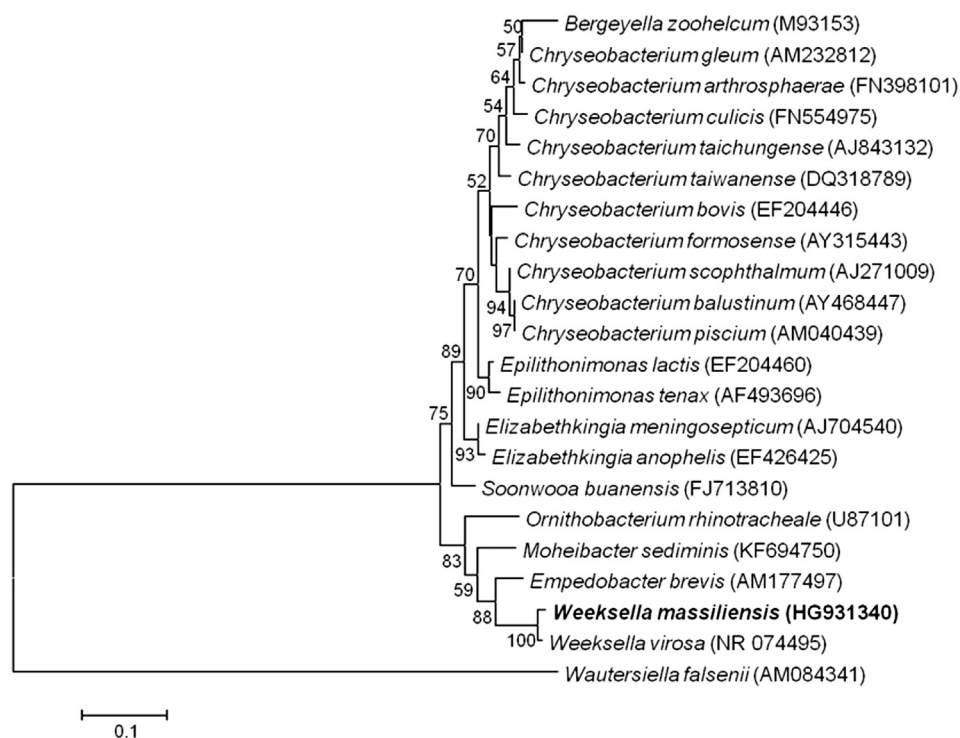


FIG. 1. Phylogenetic tree showing position of *Weeksella massiliensis* strain FF8^T relative to other type strains within family *Flavobacteriaceae*. Strains and their corresponding 16S rRNA GenBank accession numbers are: *Bergeyella zoohelcum* strain ATCC 43767, M93153 (GA: [AGYA000000000](#)); *Chryseobacterium gleum* strain ATCC 35910, M58772 (GA: [ACKQ000000000](#)); *Chryseobacterium arthrosphaerae* strain CC-VM-7, FN398101; *Chryseobacterium culicis* strain R4-1A, FN554975; *Chryseobacterium taichungense* strain CC-TWGS1-8^T, AJ843132; *Chryseobacterium taiwanense* strain BCRC 17412, DQ318789 (GA: [JWTA000000000](#)); *Chryseobacterium bovis* strain H9, EF204446; *Chryseobacterium formosense* strain CC-H3-2, AY315443 (GA: [JRP000000000](#)); *Chryseobacterium scopthalmum* strain LMG 13028, AJ271009; *Chryseobacterium balustinum* strain LMG 8329, AY468447; *Chryseobacterium piscium* strain LMG 23089^T, AM040439; *Epilithonimonas lactis* strain HI, EF204460 (GA: [JPLY000000000](#)); *Epilithonimonas tenax* strain DSM 16811, AF493696 (GA: [AUAA000000000](#)); *Elizabethkingia meningoseptica* strain ATCC 13253^T, AJ704540 (GA: [BARD000000000](#)); *Elizabethkingia anophelis* strain Ag1, EF426426 (GA: [AHHG000000000](#)); *Soonwooa buanensis* strain HM0024^T, FJ713810; *Ornithobacterium rhinotracheale* strain DSM 15997, NRI02940 (GA: [CP003283](#)); *Moheibacter sediminis* strain M0116, KF694750; *Empedobacter brevis* strain NBRC 14943, NRI12974 (GA: [ARNT000000000](#)); *W. massiliensis* strain FF8^T, HG931340 (GA: [CCMH000000000](#)); *Weeksella virosa* strain DSM 16922^T, NR074495 (GA: [CP002455](#)); and *Empedobacter falsenii* strain NF 993^T, AM084341. Sequences were aligned using MUSCLE [42], and phylogenetic tree inferred by maximum likelihood method with Kimura 2-parameter model from MEGA 6 software [43]. Numbers at nodes are percentages of bootstrap values obtained by repeating analysis 1000 times to generate majority consensus tree. *E. falsenii* was used as outgroup. Scale bar = rate of substitution per site of 0.1.

Genome Sequencing Information

Genome project history

The organism was selected for sequencing on the basis of its phylogenetic position, 16S rRNA similarity and phenotypic differences with other members of the family *Flavobacteriaceae*. It was the second genome within *Weeksella* genus and the first genome of *W. massiliensis* sp. nov. The GenBank accession number is [CCMH000000000](#) and consists of 54 contigs. Table 4 shows the project information and its association with minimum information about a genome sequence (MIGS, version 2.0) compliance [19]; associated MIGS records are also summarized in Supplemental Table S1.

Growth conditions and genomic DNA preparation

Weeksella massiliensis strain FF8^T (= CSUR P860 = DSM 28259) was grown aerobically on 5% sheep's blood-enriched Columbia agar (bioMérieux) at 37°C. Bacteria grown on four petri dishes were resuspended in 5 × 100 µL of Tris-EDTA buffer; 150 µL of this suspension was diluted in 350 µL Tris-EDTA buffer 10×, 25 µL proteinase K and 50 µL sodium dodecyl sulfate for lysis treatment. This preparation was incubated overnight at 56°C. Extracted DNA was then purified using 3 successive phenol-chloroform extractions (Thermo Fisher Scientific, Waltham, MA, USA) and ethanol precipitations at -20°C overnight. After centrifugation, the DNA was suspended in 65 µL elution buffer. The genomic DNA

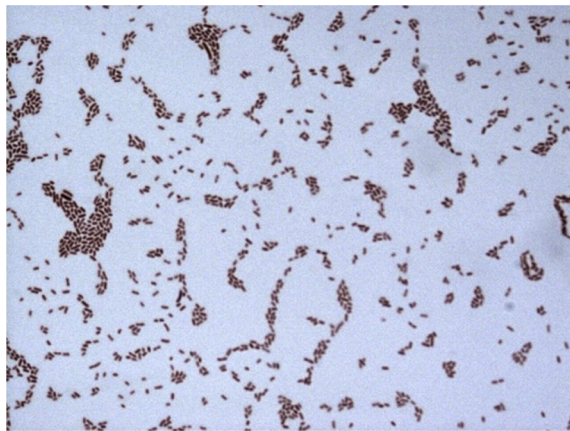


FIG. 2. Gram staining of *Weeksella massiliensis* strain FF8^T.

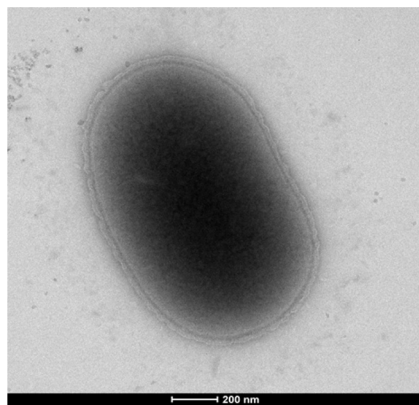


FIG. 3. Transmission electron microscopy of *Weeksella massiliensis* strain FF8^T. Cells were observed with Tecnai G20 transmission electron microscope operated at 200 keV. Scale bar = 200 nm.

concentration was measured at 53.7 ng/μL using the Qubit assay with the high sensitivity kit (Life Technologies, Carlsbad, CA, USA).

Genome sequencing and assembly

Genomic DNA of *Weeksella massiliensis* FF8^T was sequenced on a MiSeq sequencer (Illumina, San Diego, CA, USA) using the

TABLE 2. Antimicrobial susceptibility and MICs of *Weeksella massiliensis* strain FF8^T sp. nov.

Antibiotic	MIC (mg/L)	Interpretation
Penicillin	0.125	Susceptible
Ceftriaxone	0.06	Susceptible
Amoxicillin/clavulanic acid	0.16	Susceptible
Imipenem	0.15	Susceptible
Gentamycin	0.06	Susceptible
Vancomycin	3	Resistant
Trimethoprim/sulfamethoxazole	4	Resistant

MIC, minimum inhibitory concentration.

mate pair strategy. The genomic DNA (gDNA) was bar coded in order to be mixed with 11 other projects with the Nextera Mate-Pair sample prep kit (Illumina). The mate pair library was prepared with 1 μg of genomic DNA using the Nextera Mate-Pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with Mate-Pair junction adapters. The fragmentation pattern was validated on an Agilent 2100 Bio-Analyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 lab chip. The DNA fragment sizes ranged from 1 to 10 kb with an optimal size at 3.1 kb. No size selection was performed, and only 143 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal at 645 bp on a Covaris device S2 in microtubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies), and the final concentration library was measured at 6.03 nmol/L. The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing were performed in a single 27-hour run in a 2 × 151 bp. Total information of 6 Gb was obtained from a 1158K/mm² cluster density with a cluster passing quality control filters of 88.9% (21 834 000 clusters). Within this run, the index representation for *Weeksella massiliensis* was determined to 8.32%. The 1 613 495 paired reads were filtered according to the read qualities. These reads were trimmed then assembled thought CLCgenomicsWB4 software. Finally, the draft genome of *Weeksella massiliensis* consists of 17 scaffolds made of 54 contigs and generated a genome size of 2.56 Mb with G + C content of 35.9%.

Genome annotation

Open reading frames (ORFs) were predicted using Prodigal [20] with default parameters, but the predicted ORFs were excluded if they were spanning a sequencing gap region. The predicted bacterial protein sequences were searched against the GenBank database [21] and the Clusters of Orthologous Groups (COGs) database using BLASTP. The tRNAscanSE tool [22] was used to find tRNA genes, whereas ribosomal RNAs were found by using RNAmmer [23] and BLASTn against the GenBank database. Lipoprotein signal peptides and the number of transmembrane helices were predicted using SignalP [24] and TMHMM [25], respectively. ORFans were identified if their BLASTP E value was lower than 1e⁻⁰³ for alignment length greater than 80 aa. If alignment lengths were smaller than 80 aa, we used an E value of 1e⁻⁰⁵. Such parameter thresholds have already been used in previous works to define ORFans. Artemis [26] was used for data management, and DNA Plotter [27] was

TABLE 3. Differential characteristics of *Weeksellia massiliensis* strain FF8^T (data from this study), *W. virosa* strain DSM 16922^T [1], *Bergeyella zoohelcum* strain D658^T [15], *Moheibacter sediminis* strain M0116^T [16], *Elizabethkingia meningoseptica* [40] and *Chryseobacterium bovis* [41]

Character	<i>Weeksellia massiliensis</i>	<i>Weeksellia virosa</i>	<i>Bergeyella zoohelcum</i>	<i>Moheibacter sediminis</i>	<i>Elizabethkingia meningoseptica</i>	<i>Chryseobacterium bovis</i>
Cell diameter (µm)	0.2–0.5	NA	NA	0.2–0.3	0.5–1.0	0.5–0.9
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	–	–	–	–	–	–
Motility	–	–	–	–	–	–
Endospore formation	–	–	–	–	–	–
Production of						
Alkaline phosphatase	+	+	+	+	+	+
Acid phosphatase	+	+	+	+	+	+
Catalase	+	+	+	+	+	+
Oxidase	+	+	+	+	+	+
Nitrate reductase	–	–	–	–	–	–
Urease	–	–	+	–	–	–
α-Galactosidase	–	–	–	–	+	NA
β-Galactosidase	–	–	–	–	–	+
β-Glucuronidase	–	–	–	–	–	NA
α-Glucosidase	–	–	–	+	+	+
β-Glucosidase	–	–	–	+	–	NA
Esterase	+	+	NA	+	–	+
Esterase lipase	+	+	NA	+	+	+
Naphthol-AS-BI-phosphohydrolase	+	+	+	+	+	+
N-acetyl-β-glucosaminidase	–	–	–	–	NA	+
Utilization of						
5-keto-gluconate	–	NA	NA	+	NA	NA
D-Xylose	+	–	+	–	+	–
D-Fructose	–	–	–	–	+	+
D-Glucose	–	–	–	–	+	+
D-Mannose	–	–	NA	–	NA	+
Habitat	Human	Human	Parasite saprophytic	Sediment	Human	Animal

NA, data not available.

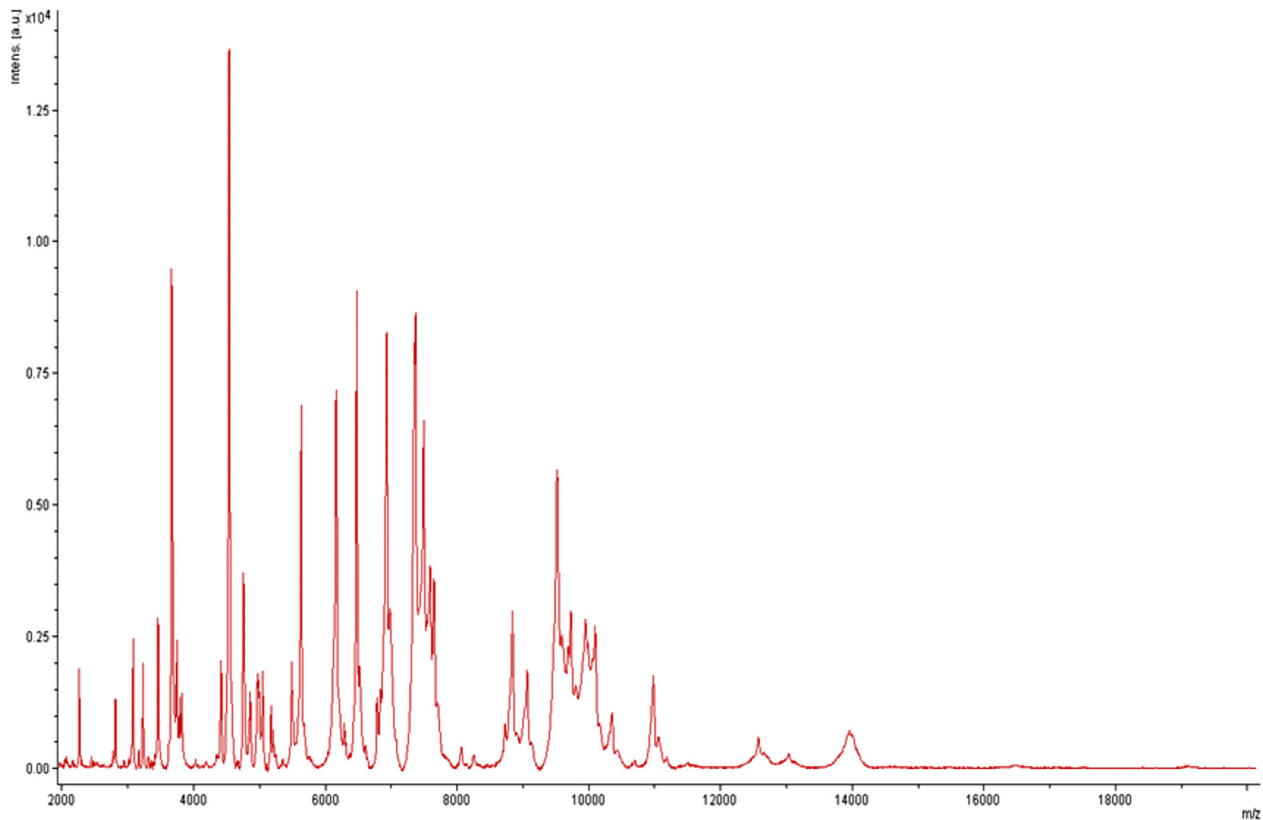


FIG. 4. Reference mass spectrum from *Weeksellia massiliensis* strain FF8^T. This reference spectrum was generated by comparison of 12 individual colonies.

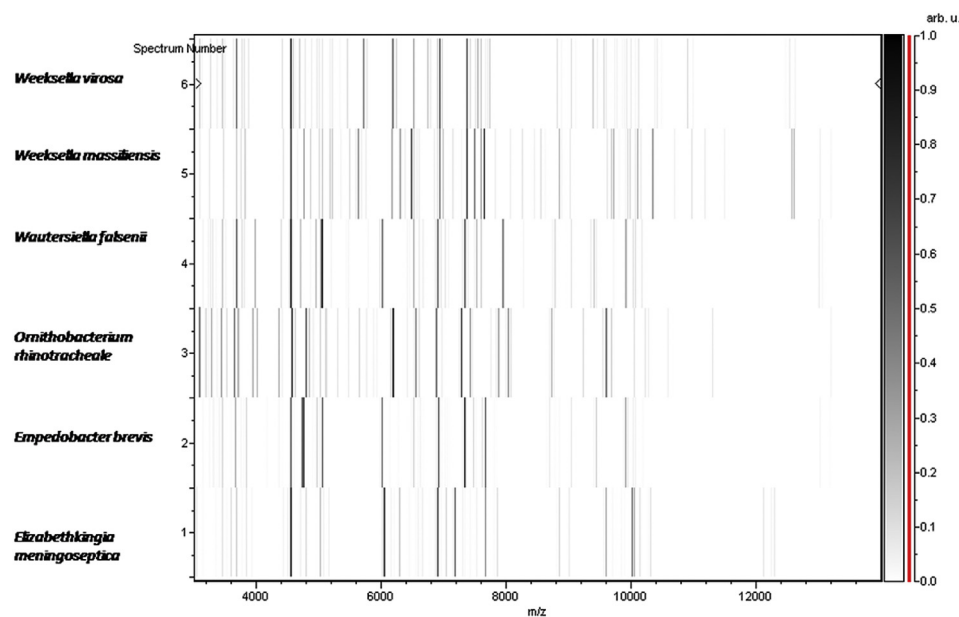


FIG. 5. Gel view comparing *Weeksella massiliensis* strain FF8^T to members of family *Flavobacteriaceae*. Gel view displays raw spectra of all loaded spectrum files arranged in pseudo-gel-like look. X-axis records *m/z* value. Left y-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity expressed by grayscale scheme code. Color bar and right y-axis indicate relation between color peak; peak intensity displayed in arbitrary units. Displayed species are indicated at left.

used for visualization of genomic features. The Mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment [28]. To estimate the mean level of nucleotide sequence similarity at the genome level, we used calculated the average genomic identity of orthologous gene sequences (AGIOS) parameter using an in-lab pipeline named Marseille Average Genomic Identity (MAGi). Briefly, this software combines the Proteinortho software [29] for detecting orthologous proteins in pairwise comparisons of genomes, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm. The script

was created to calculate the average genomic identity of orthologous gene sequences (AGIOS) between genomes by MAGi. The MAGi script created to calculate AGIOS values written in Perl and BioPerl modules. Genome-to-genome distance (GGDC) analysis was also performed using the GGDC Web server (<http://ggdc.dsmz.de/>) as previously reported [30,31].

Genome Properties

The genome of *W. massiliensis* strain FF8^T is 2 562 781 bp long with a 35.9% G + C content (Fig. 6). Of the 2446 predicted genes, 2390 were protein-coding genes and 56 were RNAs genes. Three rRNA genes (one 16S rRNA, one 23S rRNA and one 5S rRNA) and 53 predicted tRNA genes were identified in the genome. A total of 1428 genes (58.38%) were assigned a putative function. Eighty genes were identified as ORFans (3.27%). The remaining genes were annotated as hypothetical proteins. The properties and the statistics of the genome are summarized in Table 5. The distribution of genes into COGs functional categories is presented in Table 6.

Insights From Genome Sequence

Here we compared the genome sequence of *W. massiliensis* strain FF8^T with those of *Weeksella virosa* strain DSM 16922^T

TABLE 4. Project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality draft
MIGS-28	Libraries used	Paired-end 3.1 kb library
MIGS-29	Sequencing platforms	Illumina MiSeq
MIGS-31.2	Fold coverage	30.88x
MIGS-30	Assemblers	CLCGENOMICSWB4
MIGS-32	Gene calling method	Prodigal
	Locus tag	Not indicated
	GenBank ID	CCMH00000000
	GenBank date of release	August 22, 2014
	GOLD ID	Gp0102101
	BioProject ID	PRJEB5516
MIGS-13	Source material identifier	DSM 28259
	Project relevance	MALDI-TOF implementation in Dakar

MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; MIGS, minimum information about a genome sequence.

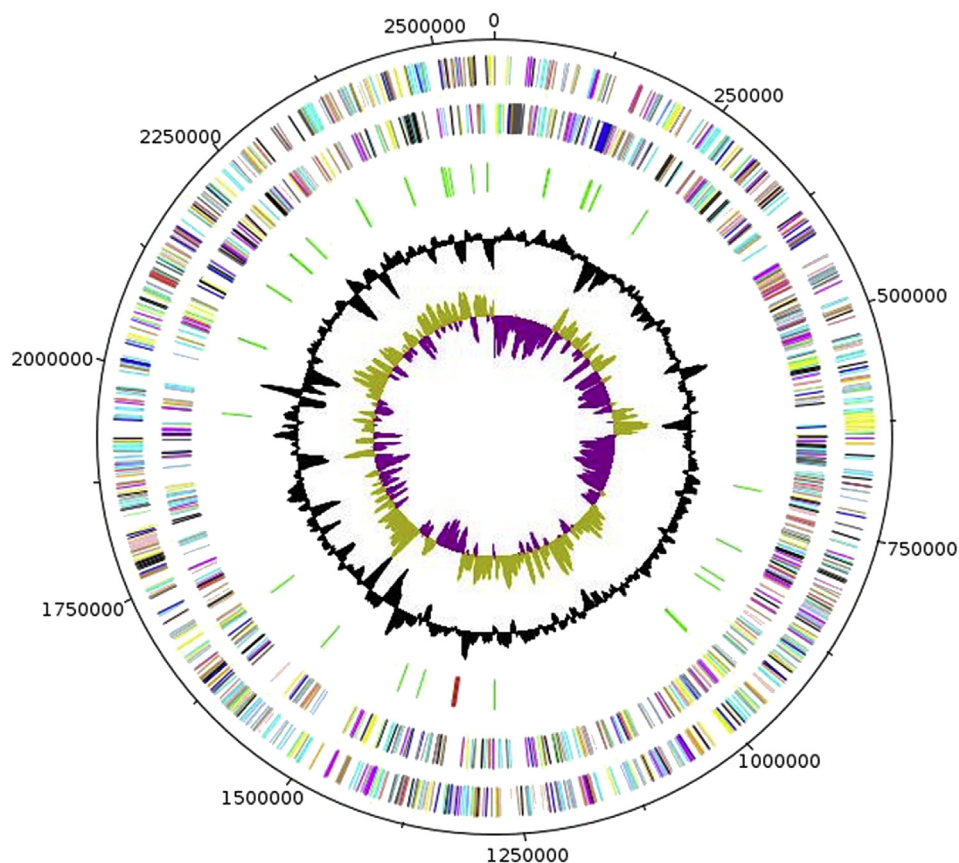


FIG. 6. Graphical circular map of *Weeksella massiliensis* strain FF8^T chromosome. From outside in, outer two circles show open reading frames oriented in forward (colored by COGs categories) and reverse (colored by COGs categories) directions, respectively. Third circle marks rRNA gene operon (red) and tRNA genes (green). Fourth circle shows G + C content plot. Innermost circle shows GC skew, with purple indicating negative values and olive positive values. COGs, Clusters of Orthologous Groups (COGs) database.

TABLE 5. Nucleotide content and gene count levels of genome

Attribute	Value	% of total ^a
Genome size (bp)	2 562 781	100
DNA coding (bp)	413 280	16.12
DNA G + C (bp)	920 038	35.9
DNA scaffolds	07	—
Total genes	2446	100
Protein coding genes	2390	97.71
RNA genes	56	2.28
Pseudo genes	ND	—
Genes in internal clusters	ND	—
Genes with function prediction	1428	58.38
Genes assigned to COGs	1567	64.06
Genes with Pfam domains	1022	41.78
Genes with signal peptides	279	11.40
Genes with transmembrane helices	528	21.58
ORFan genes	80	3.27
CRISPRs	2	—

COGs, Clusters of Orthologous Groups (COGs) database; CRISPR, clustered regularly interspaced short palindromic repeat; ND, not determined.
^aTotal is based on either size of genome (bp) or total number of protein-coding genes in annotated genome.

(GenBank accession number CP002455), *Chryseobacterium gleum* strain ATCC 35910^T (ACKQ000000000), *Elizabethkingia meningoseptica* ATCC 13253^T (BARD000000000) and *Empedobacter brevis* strain NBRC 14943^T (ARNT000000000). The draft genome of *W. massiliensis* has a larger size than that of *W. virosa* (2.56 Mb and 2.27 Mb, respectively) but smaller than those of *C. gleum*, *E. meningoseptica* and *E. brevis* (5.57 Mb, 3.96 Mb and 3.79 Mb, respectively). *Weeksella massiliensis* has a higher G + C content than *E. brevis* (35.89% and 32.7%, respectively) but lower than *W. virosa*, *C. gleum* and *E. meningoseptica* (35.9%, 36.8% and 36.4%, respectively). Because it has been suggested in the literature that the G + C content deviation is at most 1% within species, these data prove that this strain is a new taxon [32]. The gene content of *W. massiliensis* is larger than that of *W. virosa* (2446 and 2171, respectively) but smaller than those of *C. gleum*, *E. meningoseptica* and *E. brevis* (5369, 3423 and 3655, respectively). However, the distribution of genes into COGs categories was similar in all compared genomes. In addition, *W. massiliensis* shared 2390, 2049, 5289, 3369 and

TABLE 6. Number of genes associated with general COGs functional categories^a

Code	Value	%	Description
J	139	5.82	Translation, ribosomal structure and biogenesis
A	0	0	RNA processing and modification
K	83	3.47	Transcription
L	122	5.10	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	17	0.71	Cell cycle control, cell division, chromosome partitioning
V	43	1.80	Defense mechanisms
T	46	1.92	Signal transduction mechanisms
M	158	6.61	Cell wall/membrane biogenesis
N	4	0.17	Cell motility
U	27	1.13	Intracellular trafficking and secretion
O	86	3.60	Posttranslational modification, protein turnover, chaperones
C	106	4.44	Energy production and conversion
G	56	2.34	Carbohydrate transport and metabolism
E	148	6.19	Amino acid transport and metabolism
F	56	2.34	Nucleotide transport and metabolism
H	87	3.64	Coenzyme transport and metabolism
I	80	3.35	Lipid transport and metabolism
P	107	4.48	Inorganic ion transport and metabolism
Q	39	1.63	Secondary metabolites biosynthesis, transport and catabolism
R	211	8.83	General function prediction only
S	116	4.85	Function unknown
—	823	33.64	Not in COGs

COGs, Clusters of Orthologous Groups (COGs) database.

^aTotal is based on total number of protein-coding genes in annotated genome.

3589 orthologous genes with *W. virosa*, *C. gleum*, *E. meningoseptica* and *E. brevis*, respectively.

Because only one genome was available for the genus *Weeksella*, we used the genomes from closely related species for the calculation of AGIOS values. The AGIOS values ranged from 66.45% to 73.15% between *W. virosa* and other members of the family *Flavobacteriaceae* (Table 7). *W. massiliensis* exhibited similar values when compared to representatives of other genera, from 66.65% to 72.25%, but a higher value when compared to *W. virosa* (87.85%). In addition, digital DDH

similarities between the genomes were calculated by the GGDC Web server version 2.0, as recommended (Table 8) [30–33].

Conclusions

On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the creation of *Weeksella massiliensis* sp. nov. that contains strain FF8^T as type strain. The strain was isolated from the urine of a 65-year-old Senegalese man who sought care at the Hôpital Principal de Dakar, Senegal.

Description of *Weeksella massiliensis* sp. nov.

Weeksella massiliensis (mas.il.i.en'sis. L. gen. fem. n. *massiliensis*, of *Massilia*, the Latin name for Marseille, where strain FF8^T was cultivated). Colonies were 2 mm in diameter and are opaque and light yellow with a smooth surface on 5% sheep's blood-enriched Columbia agar. It is not haemolytic on blood agar. Cells are Gram negative and not motile, with a mean diameter of 0.3 µm (range, 0.2–0.5 µm) and a mean length of 1.5 µm (range, 0.8–2.1 µm). Cells were catalase and oxidase positive. Positive reactions were observed for alkaline phosphatase, esterase, esterase–lipase, leucine arylamidase, phosphatase acid and naphthol-AS-BI-phosphohydrolase activities. Negative reactions were noted for α-chymotrypsin, cystine arylamidase, valine arylamidase, trypsin, α-glucosidase, β-glucosidase, α-galactosidase, β-galactosidase, β-glucuronidase, α-mannosidase, α-fucosidase and N-acetyl-β-glucosaminidase activities. *Weeksella massiliensis* strain FF8^T is susceptible to

TABLE 7. Numbers of orthologous proteins shared between genomes (upper right) and AGIOS values obtained (lower left)

	<i>Weeksella massiliensis</i>	<i>Weeksella virosa</i>	<i>Chryseobacterium gleum</i>	<i>Elizabethkingia meningoseptica</i>	<i>Empedobacter brevis</i>
<i>W. massiliensis</i>	2390	1584	1421	1254	1647
<i>W. virosa</i>	87.85	2049	1297	1174	1551
<i>C. gleum</i>	66.69	66.45	5289	1523	1658
<i>E. meningoseptica</i>	66.65	66.54	73.15	3369	1523
<i>E. brevis</i>	72.25	71.80	67.39	67.42	3589

Bold indicates numbers of proteins per genome.

TABLE 8. Pairwise comparison of *W. massiliensis*, *W. virosa*, *C. gleum*, *E. meningoseptica* and *E. brevis* using GGDC

	<i>Weeksella massiliensis</i>	<i>Weeksella virosa</i>	<i>Chryseobacterium gleum</i>	<i>Elizabethkingia meningoseptica</i>	<i>Empedobacter brevis</i>
<i>W. massiliensis</i>	100%	32.9%	22.1%	19.7%	19.7%
<i>W. virosa</i>		100%	26.7%	22.6%	18.9%
<i>C. gleum</i>			100%	19.9%	24.1%
<i>E. meningoseptica</i>				100%	16.8%
<i>E. brevis</i>					100%

GGDC, genome-to-genome distance.

ceftriaxone, amoxicillin/clavulanic acid, penicillin, imipenem, gentamicin and doxycycline but resistant to nitrofurantoin, vancomycin, trimethoprim–sulfamethoxazole and metronidazole. The G + C content of the genome is 35.9%. The 16S rRNA and genome sequences are deposited in GenBank under accession numbers HG931340 and CCMH00000000, respectively. The type strain FF8^T (= CSUR P860 = DSM 28259) was isolated from the urine of a 65-year-old man with acute cystitis at Hôpital Principal de Dakar, Senegal.

Acknowledgements

The authors thank the Xegen Company (<http://www.xegen.fr/>) for automating the genomic annotation process. This study was funded by the Méditerranée-Infection Foundation. We also thank J.-P. Baudoin for creating the electron microscope photos.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.nmni.2015.09.013>.

Conflict of Interest

None declared.

References

- [1] Holmes B, Steigerwalt AG, Weaver RE, Brenner DJ. *Weeksellia virosa* gen. nov., sp. nov. (formerly group IIF) found in human clinical specimens. *Syst Appl Microbiol* 1986;8:185–90.
- [2] Validation list no. 23. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. *Int J Syst Bacteriol* 1987;37:179–80.
- [3] Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* 2010;60:249–66.
- [4] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006;33:152–5.
- [5] Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of new bacterial species. *Int J Syst Evol Microbiol* 2014;64:384–91.
- [6] Sentausa E, Fournier PE. Advantages and limitations of genomics in prokaryotic taxonomy. *Clin Microbiol Infect* 2013;19:790–5.
- [7] Reddy TBK, Thomas A, Stamatis D, Bertsch J, Isbandi M, Jansson J, et al. The Genomes OnLine Database (GOLD) v.5: a metadata management system based on a four level (meta) genome project classification. *Nucl Acids Res* 2014;43:1099–106.
- [8] Lagier JC, Armougom F, Mishra AK, Nguyen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes timonensis* sp. nov. *Stand Genomic Sci* 2012;6:315–24.
- [9] Lagier JC, Karkouri KE, Mishra AK, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Enterobacter massiliensis* sp. nov. *Stand Genomic Sci* 2013;7:399–412.
- [10] Fall B, Lo CI, Samb-Ba B, Perrot N, Diawara S, Gueye MW, et al. The ongoing revolution of MALDI-TOF mass spectrometry for microbiology reaches tropical Africa. *Am J Trop Med Hyg* 2015;92:641–7.
- [11] Validation list no. 143. *Int J Syst Evol Microbiol* 2012;62:1–4.
- [12] Krieg NR, Ludwig W, Euzéby J, Whitman WB. Phylum XIV. *Bacteroidetes* phyl. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, et al., editors. *Bergey's manual of systematic bacteriology*. 2nd ed. New York: Springer; 2011. p. 4–25.
- [13] Meier-Kolthoff JP, Göker M, Spröer C, Klenk HP. When should a DDH experiment be mandatory in microbial taxonomy? *Arch Microbiol* 2013;195:413–8.
- [14] Reina J, Gil J, Salva F, Gomez J, Alomar P. Microbiological characteristics of *Weeksellia virosa* (formerly CDC group IIf) isolated from the human genitourinary tract. *J Clin Microbiol* 1990;28:2357–9.
- [15] Vandamme P, Bernardet JF, Segers P, Kersters K, Holmes B. New perspectives in the classification of the *Flavobacteriia*: description of *Chryseobacterium* gen. nov., *Bergeyella* gen. nov., and *Empedobacter* nom. rev. *Int J Syst Bacteriol* 1994;44:827–31.
- [16] Zhang RG, Tan X, Zhao XM, Deng J, Lv J. *Moheibacter sediminis* gen. nov., sp. nov., a member of the family *Flavobacteriaceae* isolated from sediment, and emended descriptions of *Empedobacter brevis*, *Empedobacter falsenii* and *Weeksellia virosa*. *Int J Syst Evol Microbiol* 2014;64:1481–7.
- [17] Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Infect Dis* 2009;49:543–51.
- [18] Seng P, Rolain JM, Fournier PE, La Scola B, Drancourt M, Raoult D. MALDI-TOF–mass spectrometry applications in clinical microbiology. *Future Microbiol* 2010;5:1733–54.
- [19] Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008;26:541–7.
- [20] 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.
- [21] Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res* 2012;40:48–53.
- [22] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997;25:955–64.
- [23] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucl Acids Res* 2007;35:3100–8.
- [24] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004;340:783–95.
- [25] Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001;305:567–80.
- [26] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinformatics* 2000;16:944–5.
- [27] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 2009;25:119–20.
- [28] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 2004;14:1394–403.

- [29] Lechner M, Findeib S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* 2011;12:124.
- [30] Auch AF, Klenk HP, Göker M. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. *Stand Genomic Sci* 2010;2:142–8.
- [31] Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
- [32] Meier-Kolthoff JP, Klenk HP, Göker M. Taxonomic use of DNA G + C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–6.
- [33] Auch AF, Von Jan M, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2010;2:117–34.
- [34] 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.
- [35] Bernardet JF. Class II. *Flavobacteriia* class. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, et al., editors. *Bergey's manual of systematic bacteriology*. 2nd ed. New York: Springer; 2011. p. 4–105.
- [36] List of new names and new combinations previously effectively, but not validly, published. *Int J Syst Evol Microbiol* 2012;62:1017–9.
- [37] Garrity GM, Holt JG. Taxonomic outline of the *Archaea* and *Bacteria*. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, et al., editors. *Bergey's manual of systematic bacteriology*. 2nd ed. New York: Springer; 2011. p. 155–66.
- [38] 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;52:1049–70.
- [39] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000;25:25–9.
- [40] Kim KK, Kim MK, Lim JH, Park HY, Lee ST. Transfer of *Chryseobacterium meningosepticum* and *Chryseobacterium miricola* to *Elizabethkingia* gen. nov. as *Elizabethkingia meningoseptica* comb. nov. and *Elizabethkingia miricola* comb. nov. *Int J Syst Evol Microbiol* 2005;55:1287–93.
- [41] Hantsis-Zacharov E, Senderovich Y, Halpern M. *Chryseobacterium bovis* sp. nov., isolated from raw cow's milk. *Int J Syst Evol Microbiol* 2008;58:1024–8.
- [42] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–7.
- [43] Tamura K, Stecher G, Peterson D, Filipksi A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 2013;30:2725–9.