INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

TAXONOMIC DESCRIPTION
Lucena *et al.*, *Int. J. Syst. Evol. Microbiol.* 2021;71:005020
DOI 10.1099/ijsem.0.005020

MICROBIOLOGY SOCIETY
OPEN ACCESS

Chryseobacterium potabilaquae sp. nov., *Chryseobacterium aquaeductus* sp. nov. and *Chryseobacterium fistulae* sp. nov., from drinking water systems

Teresa Lucena, María A. Ruvira, M. Carmen Macián, David R. Arahal, Rosa Aznar and María J. Pujalte*

Abstract

A polyphasic taxonomic study was conducted on three strains isolated from drinking water systems that had previously been deposited as *Chryseobacterium* species at the Spanish Type Culture Collection in order to complete their classification. Strains CECT 9293^T, CECT 9390^T and CECT 9393^T were isolated from sites in Barcelona, Spain, in the framework of a project aimed at generating the first MALDI-TOF database specific for bacteria present in water for human consumption. Their partial 16S rRNA sequences showed that their closest relatives among the type strains of *Chryseobacterium* exhibited 98% similarity or less, supporting their taxonomic novelty. At the same time, comparison between them revealed that strains CECT 9293^T and CECT 9393^T could perhaps be related at the species level as they shared 99.5% similarity. However, whole genome sequencing was performed and the subsequent calculation of relatedness indices, average nucleotide identity and estimated DNA–DNA hybridization, ruled out that possibility and confirmed instead that each of the strains should be considered a separate species in the genus *Chryseobacterium*. Having clarified their status, we also performed phylogenomic analyses and searched for possible environmental or non-type material sequences that could be related to any of them at the species level. In parallel, the strains were characterized phenotypically and compared to their closest relatives to determine diagnostic traits to support their formal proposal. The proposed species are *Chryseobacterium potabilaquae* sp. nov. with the type strain CECT 9293^T (=LMG 32084^T), *Chryseobacterium aquaeductus* sp. nov. with the type strain CECT 9390^T (=LMG 32085^T) and *Chryseobacterium fistulae* sp. nov. with the type strain CECT 9393^T (=LMG 32086^T).

The genus *Chryseobacterium*, formerly part of the family *Flavobacteriaceae* [1, 2], contains more than 120 species registered at LPNS (https://lpsn.dsmz.de/genus/chryseobacterium accessed on May 2021) [3]. Its taxonomy is in a status of flux and, apart from various previous emendations, the recent major changes include its assignment to the newly defined family *Weeksellaceae* [4] and the transfer of 24 species to the neighbouring genera *Epilithonimonas* (nine), *Halpernia* (three), *Kaistella* (11) and *Planobacterium* (one) [5]. *Chryseobacterium* species are widely distributed in soils, plants, animals and man-made environments and many of them are related to aquatic habitats [2]. In particular, they have been detected as part of the culturable bacterial fraction of drinking water systems in different countries [6–8].

Quality and health regulations of tap water in developed countries require microbial monitoring to guarantee the absence of pathogens, based on bacterial faecal indicators and heterotrophic bacteria enumeration. Heterotrophic plate counts are important for various reasons including monitoring bacterial regrowth and biofilm formation, the inclusion of opportunistic pathogens such as *Aeromonas*, *Klebsiella* and *Pseudomonas*, which may have consequences for public health, and the interference with the enumeration of faecal indicator micro-organisms. The characterization of microbial populations in drinking water is also important for water management, as it can reveal impaired functioning of treatments and/or distribution pipelines [8].

Author affiliations: [†]Departamento de Microbiología y Ecología and Colección Española de Cultivos Tipo (CECT), Universitat de València, Valencia, Spain.
***Correspondence:** María J. Pujalte, maria.j.pujalte@uv.es
Keywords: *Chryseobacterium*; drinking water; taxogenomics; *Weeksellaceae*.
Abbreviations: AAI, average amino acid identity; ANI, average nucleotide identity; *is*DDH, *in silico* DNA–DNA hybridization; ML, maximum-likelihood; MP, maximum-parsimony; RAST, Rapid Annotation using Subsystem Technology; TSA, tryptone soy agar; TSB, tryptone soy broth; UBCG, Up-to-date-Bacterial Core Gene; WiPCA, water ISO plate count agar.
The 16S rRNA gene sequence and draft genome accession numbers for *Chryseobacterium potabilaquae* CECT 9293^T, *Chryseobacterium aquaeductus* CECT 9390^T and *Chryseobacterium fistulae* CECT 9393^T are MN982946/CACVBR01, MN982948/CAJIMS01 and MN982950/CACVBY01, respectively.
One supplementary table and two supplementary figures are available with the online version of this article.

During a survey of drinking water bacteria populations carried out in the framework of a project aimed at generating the first MALDI-TOF database specific for bacteria present in water for human consumption, several bacteria were isolated from tap water and MALDI-TOF was used to resolve the identity of the isolates obtained. The most abundant genera in the drinking water system were *Acidovorax* and *Blastomonas*, in terms of both the total number and their presence across samples. Other important genera were *Acinetobacter*, *Aquabacterium*, *Bosea*, *Brevundimonas*, *Caulobacter*, *Citrobacter*, *Polaromonas*, *Pseudomonas*, *Rhizobium* and *Variovorax* [8]. Three isolates that were not identified through a combination of MALDI-TOF and 16S rRNA gene partial sequencing were further investigated and were found to be related to *Chryseobacterium*, although they do not correspond to any described species. This study was undertaken to resolve the taxonomic position of the three strains through a characterization comprising phenotypic, phylogenetic and genomic aspects of their biology.

ISOLATION AND MAINTENANCE

The three strains of *Chryseobacterium* characterized in this study, CECT 9293^T, CECT 9390^T and CECT 9393^T, were obtained from the Colección Española de Cultivos Tipo (CECT). Strain CECT 9293^T had been isolated from a drinking water treatment plant in Sant Joan Despí (Barcelona, Spain; GPS: 41° 21′ 9.5″ N 2° 3′ 00.5″ E) on 20 August 2013; while strains CECT 9390^T and CECT 9393^T had been isolated from a drinking water distribution network in Castelldefels (Barcelona, Spain; GPS: 41° 16′ 9.4″ N 1° 59′ 41.8″ E) on 15 May 2017. All three strains were obtained from membrane filters on water ISO plate count agar (WiPCA), incubated at 22 °C for 3 days, streaked on the same medium until pure cultures were obtained, submitted to MALDI-TOF MS bacterial identification procedure and reported as non-identified. After deposition in CECT, they were further stored in lyophilized format for a long-term maintenance, using *meso*-inositol (5%) as cryoprotectant. To ensure they would be available on a second culture collection in a different country, they were deposited at LMG (Laboratorium voor Microbiologie, Universiteit Gent, Belgium) where they are kept as LMG 32084^T, LMG 32085^T and LMG 32086^T, respectively.

16S rRNA GENE PHYLOGENY

A partial 16S rRNA sequence was obtained by PCR amplification and Sanger sequencing [9] and subsequently compared by BLAST searches that placed all three strains as unnamed *Chryseobacterium* species.

Complete sequences of 16S rRNA genes were later obtained from the sequenced genomes and, after confirming they matched (100% coincidence) the respective partial sequences, they were compared through BLAST [10] and EzBioCloud [11] tools with their closer neighbours. Maximum-likelihood (ML; Fig. 1) and maximum-parsimony (MP) trees were inferred from 16S rRNA gene sequences under the GTR+CAT model by the GGDC web server [12] available at http://ggdc.dsmz.de/ using the DSMZ phylogenomics pipeline [13] adapted to single genes.

Similarities between the 16S rRNA gene sequence of strain CECT 9390^T and the type strains of its closest neighbours were 98.0% to *Chryseobacterium glaciei*, 97.8% to *Chryseobacterium polytrichastri* and 97.5% to *Chryseobacterium gambrini*. Strain CECT 9293^T shares a 96.9% sequence identity to *Chryseobacterium daecheongense*, 96.9% to *Chryseobacterium hispalense* and a 96.7% to *Chryseobacterium candidae*. Strain CECT 9393^T, on the other hand, shows 97.2% identity to *C. hispalense*, 97.2% to *C. candidae* and 97.1% to *C. daecheongense*. As can be inferred, the 16S rRNA sequence from strains CECT 9293^T and CECT 9393^T are highly similar, namely 99.5% identical, since they relate to the same organisms. Similarity values lower than 98.5% for this gene are generally considered as indicative of a separate species status, while above the threshold more resolutive methods are needed [12, 14]. Thus, while the taxonomic novelty of strain CECT 9390^T can be presumed by these results, in the case of strains CECT 9293^T and CECT 9393^T it needs to be elucidated if they represent one novel species together or one novel species each.

The position of the three strains within the *Chryseobacterium* phylogenetic tree, as defined on the basis of the 16S rRNA gene, is shown in Fig. 1. While strain CECT 9390^T merges with *C. glaciei*, the pair CECT 9293^T and CECT 9393^T forms a separate branch, without a particular link to other species. All three strains are located on the core of the genus *Chryseobacterium*, as recently defined by Nicholson *et al.* [5].

GENOME ANALYSIS AND PHYLOGENY

Genomic DNA was isolated using Jena Bioscience (Diffractia) following the standard protocol recommended by the manufacturer. The integrity of the extracted DNA was checked by visualization in a 2.0% (w/v) agarose gel electrophoresis. Its purity and quantity were checked by measuring the absorbance at 260 and 280 nm with a spectrophotometer Nanodrop 2000c (Thermo Scientific) and calculating the ratio A260/A280. Genome sequencing of the three strains was achieved at Central Support Service for Experimental Research (SCSIE) of the University of Valencia (Valencia, Spain). In the case of strains CECT 9293^T and CECT 9393^T, this was done using Illumina Miseq technology with 2×250 paired-end reads. The reads were analysed for quality control using FastQC, a tool developed by Babraham Bioinformatics to check raw sequencing data. After filtering, the remaining reads were assembled using SPAdes 3.9.0 software [15]. A plot, coverage versus length of the contigs, was performed to help in the choice of the parameters for contig filtering. After the filtration of contigs (500 bp and 10–50×kmer coverage), evaluation of the final assembly against a reference genome was done with the software Quast version 4.3 [16]. In the case of strain CECT 9390^T, genome sequencing was achieved using Sequel PacBio RS II technology (SMRT Link version 7.0) and

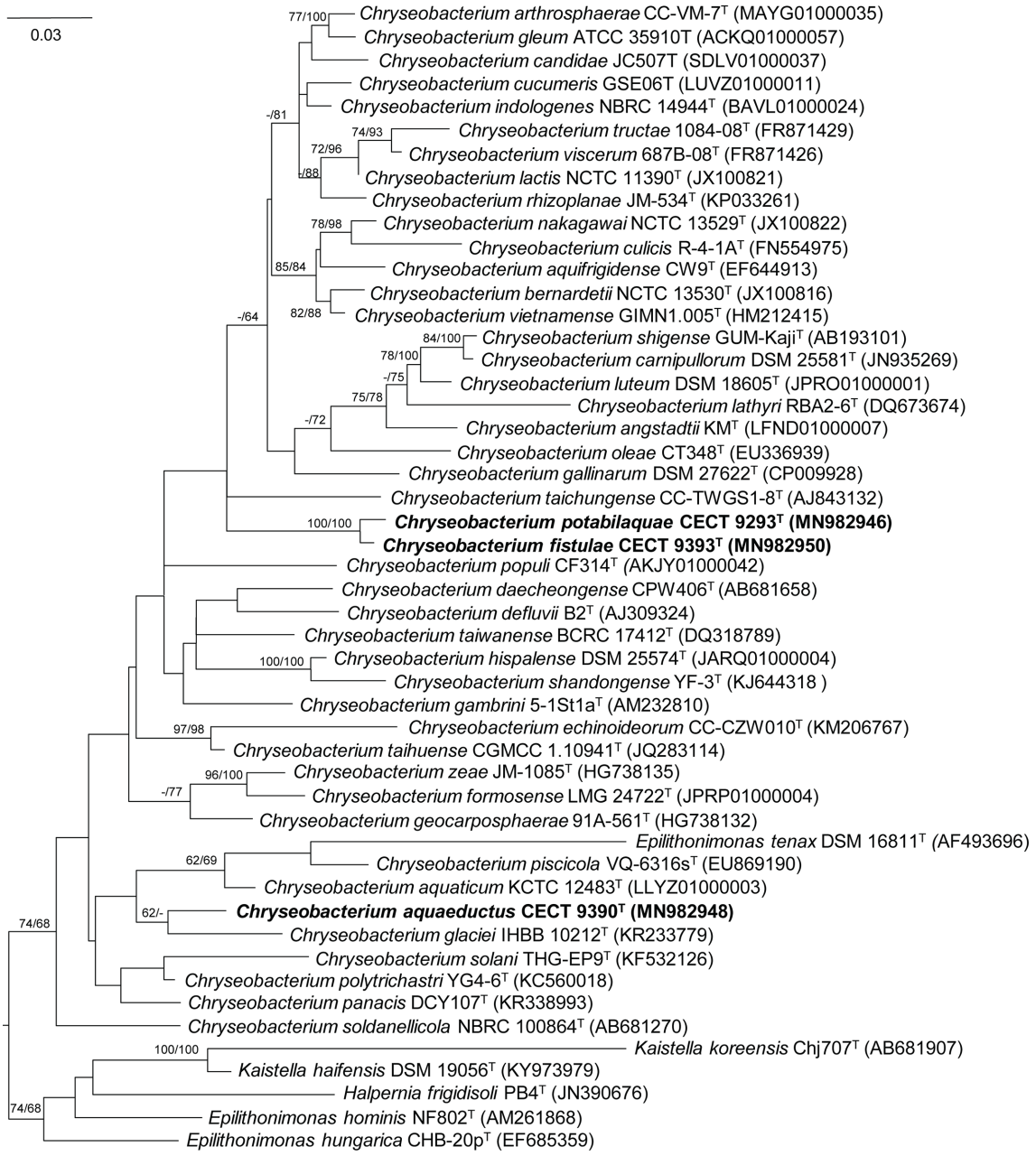

Fig. 1. ML tree inferred from 16S rRNA gene sequences under the ML analysis under the GTR+CAT model and rooted by midpoint-rooting. The branches are scaled in terms of the expected number of substitutions per site. Phylogenies were inferred by the GGDC web server [12] available at http://ggdc.dsmz.de/ using the DSMZ phylogenomics pipeline [13] adapted to single genes. The numbers above the branches are support values when larger than 60% from ML (left) and MP (right) bootstrapping.

assembled with the Hierarchical Genome Assembly Process (HGAP4) *de novo* assembly analysis application.

The bioinformatic tool CheckM version 1.0.7 [17] was used to assess the genome quality prior to annotation using Prokka version 1.12 [18] and RAST version 2.0 [19]. The process of quality assessment of reads, read-processing, assembly and annotation with Prokka was carried out in Linux OS, other tools were accessed online. The minimal standards for the quality of genome sequences and how they can be applied for taxonomic purposes [14] have been observed in this study.

Similarity between genomes was established using *in silico* DNA–DNA hybridization (*is*DDH) with the Genome-to-Genome Distance Calculator (GGDC 2.1) [12], average nucleotide identity (ANI) [20] with JSpecies software (http://jspecies.ribohost.com/jspeciesws/), and average amino acid identity (AAI) [21] with AAI matrix tools (http://

enve-omics.ce.gatech.edu/g-matrix/). Phylogenomic analysis was performed with UBCG (Up-to-date Bacterial Core Gene) [22]. This software tool is available for download at EzBioCloud [11] and employs a set of 92 bacterial core genes that are single-copy and commonly present in all bacterial genomes.

The main characteristics of the genomes are presented in Table 1. The draft genome of strain CECT 9293^T has an estimated size of 4.35 Mbp. It is composed of 149 contigs with an N50 value of 84805 nucleotides and a final assembly coverage of 283×. CheckM results of contamination and completeness were 0.98 and 99.9%, respectively. The assembly contains 3598 protein-coding sequences and 53 RNA genes. The G+C content is 32.6mol%. Only one rRNA operon is detected and its 16S rRNA gene sequences is complete and 100% coincident with the partial sequence previously amplified by Sanger technology. The genome of strain CECT 9390^T has an estimated size of 3.49 Mbp. It is composed of two contigs with an N50 value of 3462687 nucleotides and final assembly coverage of 2087×. It should be noted that strain CECT 9390^T was sequenced through PacBio technology, thus its genome is closed and could be resolved as containing one chromosome of 3.46 Mb and a 27.8 Kb independent contig, assumed to be a plasmid. CheckM results of contamination and completeness were 0.49 and 100%, respectively. The assembly contains 3203 protein-coding sequences and 67 RNA genes. Four rRNA operons are detected and its 16S rRNA gene sequence is complete and 100% coincident with the partial sequence previously amplified by Sanger technology. The G+C molar content is 34.5mol%. Finally, the draft genome of strain CECT 9393^T has an estimated size of 4.04 Mbp. It is composed of 149 contigs with an N50 value of 35975 nucleotides and final assembly coverage of 126×. CheckM results of contamination and completeness were 0.49 and 100%, respectively. The assembly contains 3433 protein-coding sequences and 53 RNA genes. The G+C molar content is 32.7 mol%. Only one rRNA operon is detected and its 16S rRNA gene sequences is complete and 100% coincident with the partial sequence previously amplified by Sanger technology.

Table 2 shows the values obtained for the three overall genomic relatedness indexes that were explored, namely ANIb, *is*DDH and AAI, relating the genomes of the three new strains with those of the type strains of neighbouring, validly named species and references used for comparison. The first noticeable finding is that the three strains show ANIb and *is*DDH values qualifying each of them as different species. The overall phenotypic resemblance and the high 16S rRNA gene sequence similarity between strains CECT 9293^T and CECT 9393^T combined with an ANIb value of 92% and 51.5% *is*DDH, indicate they clearly pertain to close, but different, yet unnamed species. All other values on Table 2 are lower than 86% for ANIb and 31% for *is*DDH. AAI values follow the same tendency, with maximum figures of 94% for the pair CECT 9293^T and CECT 9393^T, followed by a 91% between *C. daecheongense* and *C. defluvii* type strain genomes. It is noticeable that all intrageneric AAI values relating *Chyseobacterium* species are ≥74%, while values among them and species of the neighbouring genera *Epilithonimonas*, *Halpernia* and *Kaistella* lay below 70% meaning that they are well demarcated.

Phylogenomic relationships were explored using UBCG-based trees, which are based on 92 single-copy, universally distributed bacterial genes. Trees generated by using nucleotide and aminoacidic sequences were generated, including 63 genomes of *Chryseobacterium* type strains plus genomes of *Epilithonimonas, Halpernia*, *Kaistella*, *Weeksella* and *Flavobacterium* type strains. The results are depicted in Fig. 2 (aminoacid-based UBCG tree) and Fig. S1 (available in the online version of this article; nucleotide-based UBCG tree). Both trees show the position of strain CECT 9390^T paired with the type strains of *C. aquaticus* and *C. piscicola*, while strains CECT 9293^T and CECT 9393^T, paired between them, are closest to *C. daecheongense* and *C. defluvii*, all of them in the core of the genus *Chryseobacterium*. The genus boundaries are well defined from genera recently detached from it, coincident with the findings of Nicholson *et al.* [5]. It is interesting to highlight that relationships seen on the 16S rRNA gene-based tree do not coincide with the ones revealed through the phylogenomic UBCG tree: for example, the closest neighbour of CECT 9390^T is *C. glaciei* with 16S rRNA data (highest similarity, closest neighbour), but *C. aquaticum* from genome data. Aside from the important difference in information content between both datasets, quality issues with 16S rRNA sequences might also account for the results. For example, the 16S rRNA of *C. glaciei* IHBB 10212^T (KR233779) was not entirely identical to the sequences in its genome assembly (GCF_001648155), sharing only 99.7% with its seven copies.

ECOLOGY

A search for potential additional strains or uncultivated representatives of these three new species was performed with BLAST using complete 16S rRNA gene sequences derived from genomes: CECT 9390^T sequence had one hit with 99.3% similarity to strain E-052908, isolated from surface of an historic monument in Scotland [23], all other were less than 99%. The EzBioCloud search rendered no hit over 99%. The CECT 9293^T and CECT 9393^T sequences had no hits higher than 99.0% (except to each other) when submitted to BLAST, but when the EzBiocloud search was applied, they gave hits of 99.6–99.8% to the genome sequence of strain '*Chryseobacterium nematophagum*' Jub275, which is able to kill the nematode *Caenorhabditis elegans* [24]. Further comparison of strains CECT 9293^T and CECT 9393^T with the two strains studied by Page *et al.* [24], i.e. Jub129 and Jub275 (GCA_003710065.1 and GCA_003709475.1), was performed at the genomic level. The highest *is*DDH value found related strains CECT 9293^T and Jub275 by 63.4%, followed by 60.8% between CECT 9393^T and Jub129. The corresponding ANIb values were 94.8 and 93.7%, respectively. In all cases, the figures were

Table 1. Characteristics of the genomes of the drinking water *Chryseobacterium* strains analysed in this study (in bold) with their close type strains and some other reference organisms

Strain	Size (Mb)	G+C (mol%)	Coverage (×)	*N*50 (Kpb)	Contigs	Protein	rRNA	tRNA	Accession numbers*
***C. aquaeductus* CECT 9390^T**	3.49	34.5	2087	3462687	2†	3203	12	55	GCF_905175375 (CAJIMS01)
***C. potabilaquae* CECT 9293^T**	4.35	32.6	283	84805	149	3598	3	50	GCA_902728265 (CACVBR01)
***C. fistulae* CECT 9393^T**	4.04	32.7	126	35975	233	3433	3	50	GCA_902729325 (CACVBY01)
C. aquaticum KCTC 12483^T	3.81	33.9	125	465295	21	3387	6	64	GCA_001420285 (LLYZ01)
C. artocarpi UTM3^T	4.94	34.8	184	718111	51	4432	8	79	GCF_001684975 (MAYH01)
C. daecheongense DSM 15235^T	3.83	36.2	274	892685	9	3490	3	60	GCA_004365465 (SOQW01)
C. defluvii DSM 14219^T	3.71	36.6	284	1111720	8	3421	3	65	GCA_003634775 (RBXB01)
C. gleum ATCC 35910^T	5.57	36.8	31.57	3504888	7	4994	5	41	GCA_000143785 (ACKQ02)
C. oncorhynchi 701B08^T	4.79	35.1	82.82	735093	47	4331	8	76	GCF_002899895 (PPEI02)
C. piscicola DSM 21068^T	3.45	33.8	364	311216	28	3000	3	49	GCA_900156685 (FTOJ01)
C. viscerum 687B08^T	5.69	36.2	61.09	543073	60	4997	3	73	GCF_002899945 (PPEG02)
E. hispanica KCTC 22104^T	3.78	34.6	35.9	89519	149	3434	10	46	GCF_003385395 (QNUG01)

*Assembly accession shown first with WGS project in parentheses.
†Chromosome (3.46 Mb) and plasmid (27.8 Kb).

Table 2. Overall genome relatedness indexes between strains CECT 9390^T, CECT 9293^T and CECT 9393^T (in bold) to nearest *Chryseobacterium* species and to the type species of the nearest genus *Epilithonimonas*

(a) ANIb* and *is*DDH (blue); (b) AAI (green, intergeneric values).

(a)	1	2	3	4	5	6	7	8	9	10	11	12
1. *C. aquaeductus* CECT 9390^T		20.3	20.3	26.1	21.5	20.9	20.8	21,3	21.2	23.5	21.6	22.0
2. *C. potabilaquae* CECT 9293^T	75.9		51.5	20.9	21.7	21.5	21.1	21,7	21.5	20.5	21.4	21.0
3. *C. fistulae* CECT 9393^T	76.0	92.1		21.6	21.8	21.8	21.3	21,9	21.6	21.0	21.7	27.1
4. *C. aquaticum* KCTC 12483^T	82.3	76.3	76.7		21.4	21.4	21.4	21,7	21.6	22.9	22.9	23.7
5. *C. artocarpi* UTM3^T	76.7	77.0	77.0	76.8		22.6	22.0	25,3	27.0	21.0	25.5	20.9
6. *C. daecheongense* DSM 15235^T	76.8	77.4	77.5	77.2	78.1		30.8	23,2	22.4	20.4	22.3	20.8
7. *C. defluvii* DSM 14219^T	76.7	77.2	77.2	77.1	78.0	85.7		22,8	22.1	20.8	22.3	20.5
8. *C. gleum* ATCC 35910^T	76.2	76.5	76.9	76.3	80.8	78.4	78.1		25,3	21,1	30,0	22,7
9. *C. oncorhynchi* 701B08^T	76.5	77.0	77.1	76.8	82.9	78.1	78.1	81.2		20.8	26.3	21.0
10. *C. piscicola* DSM 21068^T	78.6	75.1	75.3	78.5	75.5	75.5	75.6	75.2	75.5		21.4	21.7
11. *C. viscerum* 687B08^T	76.4	76.3	76.6	77.3	81.1	77.9	77.9	84.2	82.0	75.5		23.3
12. *E. hispanica* KCTC 22104^T	73.4	72.2	73.2	74.1	72.3	72.4	72.7	72.9	72.4	73.3	72.7	
(b)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1. *C. aquaeductus* CECT 9390^T												
2. *C. aquaticum* KCTC 12483^T	85											
3. *C. piscicola* DSM 21068^T	79	79										
4. *C. potabilaquae* CECT 9293^T	76	77	75									
5. *C. fistulae* CECT 9393^T	77	77	75	94								
6. *C. defluvii* DSM 14219^T	78	78	76	80	80							
7. *C. daecheongense* DSM 15235^T	78	78	76	80	80	91						
8. *C. oncorhynchi* 701B08^T	77	76	75	78	79	80	80					
9. *C. artocarpi* UTM3^T	76	76	74	78	79	80	80	87				
10. *C. viscerum* 687B08^T	76	77	74	78	78	79	79	85	84			
11. *C. gleum* ATCC 35910^T	76	77	75	79	80	80	81	85	84	88		
12. *E. hispanica* KCTC 22104^T	68	69	68	68	69	68	68	67	67	67	69	

*ANIb figures are median values among reciprocal ANIb.
†Intergeneric AAI values between *Chryseobacterium* species, four *Epilithonimonas* species and three *Kaistella* species are always ≤70% (not shown).

lower than the thresholds for species differentiation (70% for DDH and 95–96% for ANI). Strains Jub129 and Jub275 were isolated from nematodes (*Caenorhabditis briggsae*) found on rotten fruits (apple, fig), an environment not resembling at all the one of strains CECT 9293^T and CECT 9393^T. Interestingly, despite the assumption of Page *et al.* that strains Jub129 and Jub275 belong to the same species, our ANI and *is*DDH data for these two strains reveal that they represent two different genomic species (55.5% *is*DDH and 93.4% ANIb). In summary, we have not found strains or uncultured clones putatively corresponding to the same species in 16S rRNA gene databases and we are thus unable to define the distribution of the species that seem to be rare and not abundant.

PHENOTYPIC CHARACTERIZATION

In order to perform a comparative analysis of the phenotypic profile of the drinking water strains, *Chryseobacterium aquaticum* CECT 7302^T and *C. piscicola* CECT 7357^T were characterized in parallel. Four more strains, *C. artocarpi* CECT 8497^T, *C. oncorhynchi* CECT 7794^T, *C. viscerum* CECT 7793^T and *Epilithonimonas hispanica* CECT 7129^T, were also tested as control but their results are not included

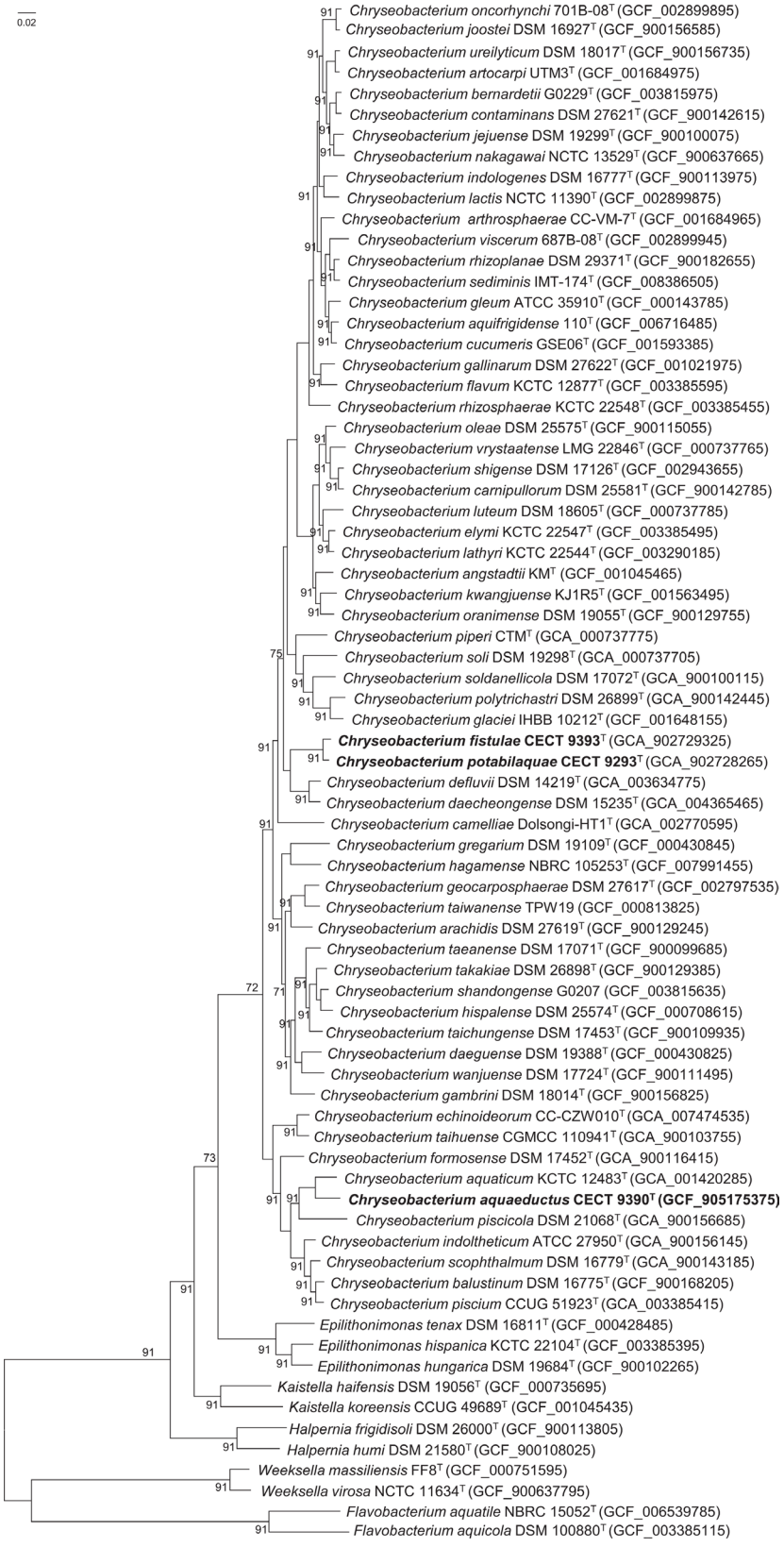

Fig. 2. Phylogenetic tree generated with the UBCG [22] by using amino acid sequences. The numbers at the nodes indicate the gene support index (maximal value is 92). Genome accession numbers are indicated in parentheses. Bar, 0.05 substitutions per position.

or discussed as they are not close relatives of any of the novel species. Procedures for testing phenotyping traits followed various published recommendations [1, 2, 25]. Unless otherwise indicated, the strains were cultured on Reasoner's 2A (R2A) medium (Difco) at 26 °C.

Cell morphology was determined on wet mounts prepared from 48 h cultures on R2A agar, by using phase contrast microscopy in a Leica DMRB fluorescence microscope (Fig. S2). Colony morphology and pigmentation were recorded from 48 h R2A and TSA (tryptone soy agar; Difco) plate cultures. Flexirubin-type pigment production was tested as described by Bernardet *et al.* [1]. Temperature (4, 15, 26, 30, 37 and 40 °C) and pH (pH 4–10.5, at 0.5 pH unit intervals) ranges for growth were determined on R2A broth (0.5 g yeast extract, 0.5 g protease peptone, 0.5 g casein hydrolysate, 0.5 g glucose, 0.5 g soluble starch, 0.3 g sodium pyruvate, 0.3 g K_2HPO_4, 0.05 g Mg $SO_4 \cdot 7H_2O$, 12.0 g agar powder) incubated for 48 h (up to 7 days at 15 and 4 °C). Optimal values were taken from the fastest grown tubes. Ability to grow in tryptone soy broth (TSB; Difco) plus 3% NaCl and salinity range for growth (0–1.0, 1.5–2.0, 2.5–3.0, 4.0–5.0 and 6.0% sea salts in 5 g l^{-1} tryptone and 1 g l^{-1} yeast extract) were determined after 3 days incubation. Growth on cetrimide agar (Pronadisa), marine agar 2216 (MA; Difco) and MacConkey agar (Pronadisa) were tested after 2 days incubation at 26 °C. Extracellular hydrolytic activities on casein and starch were determined on TSA supplemented with 10% (v/v) casein suspension or 0.2% (w/v) soluble starch, respectively. Tween 80 agar [26] and DNAse agar (Oxoid) were used to detect esterase and DNase activities. Activity on starch was revealed after lugol addition, and 1 M HCl was used to show DNase activity. Oxidase test was performed with Oxoid oxidase discs and catalase was tested with 10 vol% H_2O_2. API 20E, API 20NE, API 50CHE and API ZYM strips were inoculated following the manufacturer's instructions. Oxidation/fermentation medium of Hugh and Leifson (Difco) was used to further test the aerobic acidification of the following carbohydrates: D-glucose, D-fructose, trehalose, D-xylose, maltose, cellobiose, sucrose, D-glycerol and D-mannitol, by incubating the tubes at 26 °C for 6 days.

Fatty acid methyl esters were extracted from biomass grown for 48 h on WiPCA at 26 °C and prepared according to standard protocols as described for the MIDI Microbial Identification System [27]. Cellular fatty acid content was analysed by gas chromatography with an Agilent 6850 chromatographic unit, with the MIDI Microbial Identification System using the TSBA6 method [28] and identified using the Microbial Identification Sherlock software package.

Results of the phenotypic analysis are reported in the species descriptions and in Table 3, which displays the differential features between each strain and their phylogenetically closest neighbours. All three strains fit the defining characteristics of the genus *Chryseobacterium* according to the latest emendation by Nicholson *et al.* [5]. In some instances, traits not determined experimentally, such as major polar lipids, could be inferred from gene content; thus, the presence of a phosphatidyl serine decarboxylase gene points to the ability to synthesize phosphatidyl ethanolamine, the major polar lipid in most *Chryseobacterium* [29]. The presence of hexaprenil diphosphate synthase genes (but not others for larger isoprene chain) as part of the subsystem for 'isoprenoid for quinones' suggests MK-6 as major quinone. The search for homospermidine synthase gene, whose product is responsible for *sym*-homospermidine synthesis (the main polyamine of the genus according to [30]) gave no result. It is also noticeable the presence of several genes involved in capsular and extracellular polysaccharide synthesis in all three genomes, in agreement with the very mucous colony types that the strains display. Genes for β-lactamases, chloramphenicol acetyltransferase and RND efflux systems are present in all three strains, while genes for MATE family MDR pump occur only in strains CECT 9293^T and CECT 9393^T.

Regarding the fatty acid composition (Table S1), all three strains had $C_{15:0}$ iso as a main component followed by $C_{17:0}$ iso 3-OH, but in a rather wide range of abundance (24.0–54.8% and 10.8–17.5%, respectively). Other fatty acids added yet more differences between the strains both in presence and relative amount.

Strain 9390^T could be easily differentiated from their closest relatives, *C. aquaticum* and *C. piscicola*, by a combination of temperature growth range, DNAse and *N*-acetyl-β-glucosidase activities, ability to produce acid aerobically from trehalose and sucrose, and to grow in TSB 3% NaCl and MA. Differentiation between 9293^T and 9393^T relys only in one characteristic, among the many tested: nitrate reduction ability, but they are differentiated from both *C. defluvii* and *C. daecheongense* by at least five traits; another additional six features allow further differentiation from each of the two relatives, as shown in Table 3. Genomic information adds some other distinctive characters, as the presence of some of the genes involved in carotenoid synthesis or the presence of restriction–modification systems and CRISPR-Cas proteins.

Taken together, the results under this heading provide phenotypic support for the separate species status that was shown by the phylogenetic analysis. Therefore, the assignment of each of the three water system isolates, strains CECT 9293^T, CECT 9390^T and CECT 9393^T, to novel species is proposed with the following descriptions.

DESCRIPTION OF *CHRYSEOBACTERIUM POTABILAQUAE* SP. NOV.

Chryseobacterium potabilaquae (po.ta.bil.a'quae. L. masc. adj. *potabilis* potable; L. fem. n. *aqua* water; N.L. gen. n. *potabilaquae* of drinking water).

Cells are bacilli with slightly pointed ends, 1.5–3.5 µm×0.5–0.6 µm in size, Gram-reaction-negative and non-motile. Colonies on R2A medium are elevated, mucous and pale

Table 3. Differential characteristics between strains CECT 9390^T, CECT 9293^T, CECT 9393^T and their closest phylogenomic relatives

Strains: 1, *Chryseobacterium aquaeductus* CECT 9390^T; 2, *Chryseobacterium aquaticum* CECT 7302^T; 3, *Chryseobacterium piscicola* CECT 7357^T; 4, *Chryseobacterium potabilaquae* CECT 9293^T; 5, *Chryseobacterium fistulae* CECT 9393^T; 6, *Chryseobacterium defluvii* B2^T [2]; 7, *Chryseobacterium daecheongense* CPW 406^T [2]. +, Positive; –, negative; W, weak; ND, not determined. All data from this study unless otherwise indicated.

Characteristic	1	2	3	4	5	6	7
Growth on/at:							
TSB 3% NaCl	+	+	–	–	–	–	–
MA	+	+	–	–	–	ND	ND
Cetrimide agar	–	–	–	–	–	+	+
5 °C	+	–	–	–	–	–	–
37 °C	–	+	–	–	–	+	+
Nitrate reduction to nitrite	–	–	–	–	+	–	+
DNAse	–	+	–	–	–	+	+
Urease	–	–	–	+	+	–	–
N-Acetyl-β-glucosidase	–	+	W	–	–	ND	ND
Acid from:							
D-Glucose	–	–	–	+	+	+	–
D-Fructose	–	–	–	–	–	–	+
Trehalose	–	+	–	+	+	+	+
D-Xylose	–	–	–	–	–	–	+
Maltose	–	–	–	–	–	+	–
Cellobiose	–	–	–	–	–	+	+
Sucrose	–	+	–	–	–	–	–
D-Glycerol	–	–	–	–	–	–	+
Genomic traits:							
β-Carotene 3-hydroxylase gene	+	+	+	–	+	+	+
CRISPR-Cas systems	–	+ (Cas 1,2,9)	+ (Cas 1,2,9)	+ (Cas 1,2,9)	–	+ (Cas 1,2)	–
Restriction–modification systems	–	+	+	–	+	+	+

yellow. Grows on TSA, but growth is better and faster on R2A medium. Flexirubin-type pigments are produced, according to the KOH test. Does not grow on MacConkey agar, cetrimide agar, MA or in TSB plus 3% (w/v) NaCl. Salinity tolerance range for growth is 0–2.5%. Oxidase and catalase tests are positive. Temperature range for growth is 15–30 °C, no growth is observed at 4 or 37 °C. The pH range for growth is pH 5.0–8.5, with slight growth at pH 9.0. Aerobic, unable to ferment carbohydrates, but acid is produced (weakly) in aerobic conditions from D-glucose and trehalose. Nitrate is not reduced to nitrite or further. The following activities are negative: arginine dihydrolase, lysine decarboxylase, ornithine decarboxylase, β-galactosidase, tryptophanase (indole), phenylalanine deaminase, Voges–Proskauer and acid production from D-glucose, D-mannitol, D-sorbitol, myo-inositol, L-rhamnose, sucrose, melibiose, amygdalin and L-arabinose on API 20E; positive for urea, aesculin and gelatin hydrolysis, alkaline phosphatase, leucine and valine arylamidases, trypsin and naphthol-AS-BI-phosphohydrolase; weak positive response for acidic phosphatase. Hydrolyses casein, Tween 80 and starch (weak), but not DNA. Tyrosine clearing is negative but growth on tyrosine medium produces a diffusible brown pigmentation. No acidification is observed in API 50CH/E, except for a slight response on D-glucose, maltose and trehalose. Acid production in aerobic oxidation/fermentation test is positive for D-glucose and trehalose but negative for D-xylose, L-arabinose, D-fructose, maltose, cellobiose, sucrose, D-glycerol and D-mannitol. Major fatty acids are $C_{15:0}$ iso and $C_{17:0}$ iso 3-OH, followed by $C_{16:1}$ *ω*7*c*/$C_{16:1}$ ω6*c*, $C_{16:0}$ 10-methyl/$C_{17:1}$ iso *ω*9*c* and $C_{15:0}$ anteiso.

The type strain is CECT 9293^T (=LMG 32084^T), which was isolated from a drinking water treatment plant in Barcelona, Spain.

Genome of the type strain is 4.35 Mb in size and comprises 3598 protein-coding genes and 53 RNA genes (a single rRNA operon and 50 tRNAs). The G+C content is 32.6 mol%.

DESCRIPTION OF *CHRYSEOBACTERIUM AQUAEDUCTUS* SP. NOV.

Chryseobacterium aquaeductus (a.quae.duc'tus. L. gen. n. *aquaeductus* of a conveyance of water).

Cells are Gram-reaction-negative, short, non-motile bacilli, 0.8–2.0 µm×0.6–0.8 µm, that do not glide. Colonies on R2A medium are yellow, mucoid and shiny. The pigmentation on TSA medium is ochre. Flexirubin-type pigments are produced, according with the KOH test. Chemoorganotrophic, strictly aerobic, positive for catalase and oxidase tests. Carbohydrates are not fermented. Grows on MA after 4 days but not on MacConkey or cetrimide agar. Grows in TSB with 3% NaCl. Salinity range of growth is 0–3.0%. Temperature range for growth is 4–30 °C, no growth is observed at 37 °C or more. pH range for growth is pH 5.5–9.5. Negative for nitrate reduction to nitrite (or further), arginine dihydrolase, lysine decarboxylase, ornithine decarboxylase, urease, β-galactosidase, tryptophanase (indole), phenylalanine deaminase, Voges–Proskauer and acid production from D-glucose, D-mannitol, D-sorbitol, *myo*-inositol, L-rhamnose, sucrose, melibiose, amygdalin and L-arabinose; positive for aesculin and gelatin hydrolysis, alkaline phosphatase, leucine and valine arylamidases and trypsin; weak positive response for esterase lipase, acidic phosphatase and naphthol-AS-BI-phosphohydrolase. Hydrolyses casein, Tween-80 and starch, but not DNA. Tyrosine clearing is negative but growth on tyrosine medium produces a strong diffusible brown pigmentation. No acidification is observed in any of the 49 carbohydrates of API 50CH/E, with only a very slight response in maltose, trehalose, amygdalin and D-gentibiose. No acidification is observed in aerobic oxidation/fermentation medium with any of the following carbohydrates: D-xylose, L-arabinose, D-glucose, D-fructose, trehalose, maltose, cellobiose, sucrose, D-glycerol and D-mannitol. Major fatty acids are $C_{15:0}$ iso, $C_{17:0}$ iso 3-OH and $C_{16:1}$ *ω7c*/$C_{16:1}$ ω6*c*, followed by $C_{15:1}$ ω5*c* and $C_{15:0}$ anteiso.

The type strain is CECT 9390^T (=LMG 32085^T), which was isolated from a drinking water distribution network in Barcelona, Spain.

Genome of the type strain is 3.49 Mb in size (one chromosome of 3.46 Mb and one plasmid of 27.8 Kb) and comprises 3203 protein-coding genes plus 68 RNA genes. It contains four rRNA operons and 55 tRNA genes. The G+C content is 34.5 mol%.

DESCRIPTION OF *CHRYSEOBACTERIUM FISTULAE* SP. NOV.

Chryseobacterium fistulae (fis'tu.lae. L. gen. n. *fistulae* of a water pipe).

Cells are bacilli with slightly pointed ends, 1.5–3.5 µm×0.5–0.6 µm in size, Gram-reaction-negative and non-motile. Colonies on R2A medium are elevated, mucous and pale yellow. Grows on TSA, but growth is better and faster on R2A medium. Flexirubin type pigments are produced, according with the KOH test. Does not grow on MacConkey agar, cetrimide agar or MA or in TSB plus 3% (w/v) NaCl. Salinity tolerance range for growth is 0–2.0%. Temperature range for growth is 15–30 °C, no growth is observed at 4 or 37 °C. The pH range for growth is pH 5.0–8.5, with slight growth at pH 9.0. Oxidase and catalase tests are positive. Aerobic, unable to ferment carbohydrates, but acid is produced (weakly) in aerobic conditions from D-glucose and trehalose. Nitrate is reduced to nitrite. The following activities are negative: arginine dihydrolase, lysine decarboxylase, ornithine decarboxylase, β-galactosidase, tryptophanase (indole), phenylalanine deaminase, Voges–Proskauer and acid production from D-glucose, D-mannitol, D-sorbitol, *myo*-inositol, L-rhamnose, sucrose, melibiose, amygdalin and L-arabinose on API 20E; positive for urea, aesculin and gelatin hydrolysis, alkaline and acidic phosphatases, leucine and valine arylamidases, trypsin, and naphthol-AS-BI-phosphohydrolase. Hydrolyses casein, Tween-80 and starch (weak), but not DNA. Tyrosine clearing is negative but growth on tyrosine medium produces a diffusible brown pigmentation. No acidification is observed in API 50CH/E, except for a slight response on trehalose. Acid production in aerobic oxidation/fermentation (O/F) test is positive for D-glucose and trehalose but negative for D-xylose, L-arabinose, D-fructose, maltose, cellobiose, sucrose, D-glycerol and D-mannitol. Major fatty acids are $C_{15:0}$ iso, $C_{17:0}$ iso 3-OH, $C_{16:0}$ 10-methyl/$C_{17:1}$ iso *ω9c* and $C_{15:0}$ anteiso followed by $C_{16:1}$ *ω7c*/$C_{16:1}$ ω6*c* and $C_{15:1}$ ω5*c*.

The type strain is CECT 9393^T (=LMG 32086^T), which was isolated from a drinking water distribution network at Barcelona, Spain.

The genome of the type strain is 4.04 Mb in size and comprises 3433 protein-coding genes and 53 RNA genes (one rRNA operon and 50 tRNAs). The G+C content is 32.7 mol%.

Funding information
Partial financial support was received from UV-INV-AE19-1199655 grant of the Universitat de València and AICO-2020-181 grant (Generalitat Valenciana) to MJP.

Acknowledgements
We thank María Olcina for technical support and Dr. Aharon Oren for nomenclatural advice. We thank SCSIE, Central Support Service for Experimental Research, for technical support in genome sequencing.

Author contributions
Conceptualization: T.L., M.C.M., D.R.A., R.A. and M.J.P. Data curation, formal analysis and investigation: T.L., M.A.R., D.R.A. and M.J.P. Funding acquisition: D.R.A., R.A. and M.J.P. Project administration and resources: T.L., M.A.R. and M.C.M. Supervision and writing original draft: M.J.P. All authors contributed to review and editing.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Bernardet JP, Nakagawa Y, Holmes B. Subcommittee on the taxonomy of Flavobacterium and Cytophaga-like bacteria of the International Committee on Systematics of Prokaryotes. 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.
2. Hugo C, Bernardet JF, Nicholson A, Kampfer P, *et al*. *Chryseobacterium*. Trujillo M, Dedysh S, DeVos P, Hedlund B and Kämpfer P (eds). In: *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons; 2019.
3. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol* 2020;70:5607–5612.
4. García-López M, Meier-Kolthoff JP, Tindall BJ, Gronow S, Woyke T, *et al*. Analysis of 1,000 type-strain genomes improves taxonomic classification of *Bacteroidetes*. *Front Microbiol* 2019;10:2083.
5. Nicholson AC, Gulvik CA, Whitney AM, Humrighouse BW, Bell ME, *et al*. Division of the genus *Chryseobacterium*: Observation of discontinuities in amino acid identity values, a possible consequence of major extinction events, guides transfer of nine species to the genus *Epilithonimonas*, eleven species to the genus *Kaistella*, and three species to the genus *Halpernia* gen. nov., with description of *Kaistella daneshvariae* sp. nov. and *Epilithonimonas vandammei* sp. nov. derived from clinical specimens. *Int J Syst Evol Microbiol* 2020;70:4432–4450.
6. Ultee A, Souvatzi N, Maniadi K, König H. Identification of the culturable and nonculturable bacterial population in ground water of a municipal water supply in Germany. *J Appl Microbiol* 2004;96:560–568.
7. Gallego V, García MT, Ventosa A. *Chryseobacterium hispanicum* sp. nov., isolated from the drinking water distribution system of Sevilla, Spain. *Int J Syst Evol Microbiol* 2006;56:1589–1592.
8. Sala-Comorera L, Caudet-Segarra L, Galofré B, Lucena F, Blanch AR, *et al*. Unravelling the composition of tap and mineral water microbiota: Divergences between next-generation sequencing techniques and culture-based methods. *Int J Food Microbiol* 2020;334:108850.
9. Arahal DR, Sánchez E, Macián MC, Garay E. Value of *recN* sequences for species identification and as a phylogenetic marker within the family "*Leuconostocaceae*". *Int Microbiol* 2008;11:33–39.
10. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, *et al*. BLAST: A more efficient report with usability improvements. *Nucleic Acids Res* 2013;41:W29-33.
11. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, *et al*. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.
12. 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.
13. Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, *et al*. Complete genome sequence of DSM 30083(T), the type strain (U5/41(T)) of Escherichia coli, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci* 2014;9:2.
14. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, *et al*. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461–466.
15. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, *et al*. Assembling genomes and mini-metagenomes from highly chimeric reads. Deng M, Jiang R, Sun F and Zhang X (eds). In: *Research in Computational Molecular Biology. RECOMB 2013. Lecture Notes in Computer Science*, Vol. 7821. Berlin, Heidelberg: Springer;
16. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–1075.
17. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
18. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.
19. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, *et al*. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 2014;42:D206-14.
20. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 2016;32:929–931.
21. Rodriguez-R LM, Konstantinidis KT. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints* 2016;4:e1900v1.
22. Na SI, Kim YO, Yoon SH, SM H, Baek I, *et al*. UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J Microbiol* 2018;56:281–285.
23. Suihko M-L, Alakomi H-L, Gorbushina A, Fortune I, Marquardt J, *et al*. Characterization of aerobic bacterial and fungal microbiota on surfaces of historic Scottish monuments. *Syst Appl Microbiol* 2007;30:494–508.
24. Page AP, Roberts M, Félix M-A, Pickard D, Page A, *et al*. The golden death bacillus *Chryseobacterium nematophagum* is a novel matrix digesting pathogen of nematodes. *BMC Biol* 2019;17:10.
25. Bernardet JF. Flavobacteriaceae. Trujillo M, Dedysh S, DeVos P, Hedlund B and Kämpfer P (eds). In: *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons; 2019.
26. Smibert RM, Krieg NR. Phenotypic characterization. Reddy C, Beveridge T, Breznak J, Marzluf G and Schmidt T (eds). In: *Manual of Methods for General Bacteriology*. Washington, DC: American Society for Microbiology; 2007. pp. 607–654.
27. Sasser M. *Identification of Bacteria by Gas Chromatography of Cellular Fatty Acids*. MIDI Technical Note 101. Newark: DE: MIDI Inc; 1990.
28. MIDI. *Sherlock Microbial Identification System Operating Manual, version 6.1. 2008*. Newark, DE: MIDI Inc;
29. Wu Y-F, Wu Q-L, Liu S-J. *Chryseobacterium taihuense* sp. nov., isolated from a eutrophic lake, and emended descriptions of the genus *Chryseobacterium*, *Chryseobacterium taiwanense*, *Chryseobacterium jejuense* and *Chryseobacterium indoltheticum*. *Int J Syst Evol Microbiol* 2013;63:913–919.
30. Kämpfer P, Vaneechoutte M, Lodders N, De Baere T, Avesani V, *et al*. Description of *Chryseobacterium anthropi* sp. nov. to accommodate clinical isolates biochemically similar to *Kaistella koreensis* and *Chryseobacterium haifense*, proposal to reclassify *Kaistella koreensis* as *Chryseobacterium koreense* comb. nov. and emended description of the genus *Chryseobacterium*. *Int J Syst Evol Microbiol* 2009;59:2421–2428.