Phylogenomic Analyses and Reclassification of Species within the Genus *Tsukamurella*: Insights to Species Definition in the Post-genomic Era

Jade L. L. Teng[1,2,3,4†], Ying Tang[1†], Yi Huang[1], Feng-Biao Guo[5], Wen Wei[6], Jonathan H. K. Chen[1], Samson S. Y. Wong[1,2,3,4], Susanna K. P. Lau[1,2,3,4,7]* and Patrick C. Y. Woo[1,2,3,4,7]*

[1] Department of Microbiology, The University of Hong Kong, Hong Kong, China, [2] Research Centre of Infection and Immunology, The University of Hong Kong, Hong Kong, China, [3] State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong, China, [4] Carol Yu Centre for Infection, The University of Hong Kong, Hong Kong, China, [5] Centre of Bioinformatics, Key Laboratory for NeuroInformation of the Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China, [6] School of Life Sciences, Chongqing University, Chongqing, China, [7] Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The University of Hong Kong, Hong Kong, China

Owing to the highly similar phenotypic profiles, protein spectra and 16S rRNA gene sequences observed between three pairs of *Tsukamurella* species (*Tsukamurella pulmonis/Tsukamurella spongiae*, *Tsukamurella tyrosinosolvens/Tsukamurella carboxydivorans*, and *Tsukamurella pseudospumae/Tsukamurella sunchonensis*), we hypothesize that and the six *Tsukamurella* species may have been misclassified and that there may only be three *Tsukamurella* species. In this study, we characterized the type strains of these six *Tsukamurella* species by tradition DNA–DNA hybridization (DDH) and "digital DDH" after genome sequencing to determine their exact taxonomic positions. Traditional DDH showed 81.2 ± 0.6% to 99.7 ± 1.0% DNA–DNA relatedness between the two *Tsukamurella* species in each of the three pairs, which was above the threshold for same species designation. "Digital DDH" based on Genome-To-Genome Distance Calculator and Average Nucleotide Identity for the three pairs also showed similarity results in the range of 82.3–92.9 and 98.1–99.1%, respectively, in line with results of traditional DDH. Based on these evidence and according to Rules 23a and 42 of the *Bacteriological Code*, we propose that *T. spongiae* Olson et al. 2007, should be reclassified as a later heterotypic synonym of *T. pulmonis* Yassin et al. 1996, *T. carboxydivorans* Park et al. 2009, as a later heterotypic synonym of *T. tyrosinosolvens* Yassin et al. 1997, and *T. sunchonensis* Seong et al. 2008 as a later heterotypic synonym of *T. pseudospumae* Nam et al. 2004. With the advancement of genome sequencing technologies, classification of bacterial species can be readily achieved by "digital DDH" than traditional DDH.

Keywords: phylogenomic, reclassification, *Tsukamurella*, species, genomics

INTRODUCTION

Traditionally, classification of bacteria is performed on the basis of their phenotypic characteristics, such as the appearance of the bacterium under light microscope after Gram staining, growth requirements, biochemical tests, and chemotaxonomic characteristics. However, the results of individual biochemical tests may vary among different strains of the same bacterial species, leading to variations in their biochemical profiles. Therefore, it is often difficult to determine whether a strain belongs to a new species or if it is just a variant of an existing species. In the 1970s, the discovery of conserved small ribosomal RNA (rRNA) gene sequences has marked the beginning of a new era for the study of evolution and classification of living organisms (Fox et al., 1977; Gupta et al., 1983). These rRNA gene sequences are, in general, highly conserved within living organisms of the same genus and species, but different between organisms of different genera and species. With the subsequent invention of PCR and automated DNA sequencing technology in the 1990s, 16S rRNA gene sequences have been widely used for phylogenetic, bacterial classification and identification. Numerous bacterial genera and species have been re-classified and renamed, and many novel bacterial genera and species have been discovered (Woo et al., 2008). However, careful analysis of 16S rRNA data is needed to avoid misinterpretation of species classification, especially in cases where taxonomic claims are proposed on the basis of the 16S rRNA data alone. For example, two strains that share 96% 16S rRNA gene sequence similarity are determined to be members of different species; but in some cases, two strains that share 98% 16S rRNA gene similarity may or may not be members of the same species (Stackebrandt and Goebel, 1994; Palys et al., 1997). As a result, it is difficult to use a universal cut-off value for classifying all bacteria to the species level. Other data sets should always be used in support of the claims.

In view of these problems, DNA–DNA hybridization (DDH) has remained the gold standard for classification in bacterial taxonomy. It represents a universally applicable technique that offers genome-wide comparisons between any bacteria, and a universal DDH value of 70% is proposed to be the criterion for bacterial species delineation (Wayne, 1988). However, this traditional method is time-consuming and labor-intensive, and it is difficult to compare DDH results between different laboratories objectively. Given these drawbacks, bacterial taxonomists have been actively searching for alternative methods to replace the traditional DDH for bacterial classification and species assignment (Goris et al., 2007). In the era of genomics, massive amounts of high-quality bacterial genome sequences can be easily obtained by high-throughput sequencing approach. As a result, different attempts have been made to replace the traditional wet-lab DDH with *in silico* genome-to-genome comparison. Among the various "digital DDH" methods studied, the Genome-To-Genome Distance Calculator (GGDC) and Average Nucleotide Identity (ANI) calculator were shown to yield good correlation with wet-lab DDH values (Goris et al., 2007; Auch et al., 2010). GGDC operates on the same scale as the wet-lab DDH method and use the value of 70% as a cut-off for species delineation (Auch et al., 2010). ANI, on the other hand, estimates the average

nucleotide identity between two genomic datasets, and an ANI value in the range of 94–96% has been suggested as the substitute for a wet-lab DDH value of 70% (Konstantinidis and Tiedje, 2005; Goris et al., 2007; Richter and Rosselló-Móra, 2009).

The genus *Tsukamurella* was first proposed in 1988 by Collins and colleagues using both chemotaxonomic data and 16S rRNA gene sequence analysis (Collins et al., 1988), although taxonomic history of the genus can be dated back to 1941, when the first strain of this group of bacteria was isolated from mycetomes and ovaries of bed bugs (*Cimex lectularius*), and was initially named *Corynebacterium paurometabolum* (Steinhaus, 1941). Members of *Tsukamurella* are Gram-positive, aerobic, catalase-positive, and partially acid-fast, with morphology similar to the related genera of the order Actinomycetales, including *Mycobacterium*, *Nocardia*, *Corynebacterium*, *Gordonia*, and *Rhodococcus*. Chemotaxonomic characteristics, such as fatty acid, mycolic acid, menaquinone, and polar lipid composition, have been shown to be particularly useful to study the relatedness of actinomycete and coryneform systematics (Collins and Jones, 1982; Collins et al., 1982). At the time of writing, there are 13 species to be included in the genus *Tsukamurella* according to the current state of the taxonomy, with only seven species associated with human infections. The most common *Tsukamurella* infections in human are related to the presence of indwelling devices (Larkin et al., 1999; Shaer and Gadegbeku, 2001; Schwartz et al., 2002; Almehmi et al., 2004; Bouza et al., 2009; Liu et al., 2011). The disease spectra of *Tsukamurella* further extends to infections related to ophthalmologic diseases (Woo et al., 2003, 2009), and recently, we have reported the discovery of two novel *Tsukamurella* species, *Tsukamurella hongkongensis* and *Tsukamurella sinensis*, from patients with keratitis and conjunctivitis (Teng et al., 2015). *Tsukamurella* is also found in different environmental sources as well as in animals (Maeda et al., 2010; Jiang et al., 2013). However, as nucleotide differences between the 16S rRNA gene sequences of different *Tsukamurella* species are less than 1% (Woo et al., 2003), defining a *Tsukamurella* strain as a novel species relies heavily on DDH results. However, it is important to note that DDH results may vary depending on the identity of the DNA used and any experimental problems cannot be checked in a different laboratory using the same samples, hence it is possible that *Tsukamurella* strains may be wrongly classified on the basis of DDH data alone.

During the process of characterization of *T. hongkongensis* and *T. sinensis* with other type strains of known *Tsukamurella* species, we observed that some of the known *Tsukamurella* species shared identical 16S rRNA gene sequence and highly similar phenotypic characteristics using the same testing method (Teng et al., 2015). Our results showed that almost identical phenotypic profiles were observed between *Tsukamurella pulmonis* CCUG 35732[T] (Yassin et al., 1996) and *Tsukamurella spongiae* DSM 44990[T] (Olson et al., 2007), between *Tsukamurella tyrosinosolvens* CCUG 38499[T] (Yassin et al., 1997) and *Tsukamurella carboxydivorans* JCM 15482[T] (Park et al., 2009), and between *Tsukamurella pseudospumae* JCM 13375[T] (Nam et al., 2004) and *Tsukamurella sunchonensis* JCM 15929[T] (Seong et al., 2003; Euzéby, 2008). However,

phenotypic characteristics of these six *Tsukamurella* species obtained from our study were not comparable to those obtained from the original publications because different culture medium, reading interval and incubation conditions of the phenotypic tests have been applied in different studies (Yassin et al., 1996, 1997; Seong et al., 2003; Nam et al., 2004; Olson et al., 2007; Park et al., 2009). We further characterized these strains by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS), in which the hierarchical cluster analysis dendrogram also revealed the same clustering patterns among these three pairs of *Tsukamurella* type strains. Therefore, we hypothesized that these six *Tsukamurella* species may have been misclassified and some of them may belong to the same *Tsukamurella* species. In fact, when some of these "novel" *Tsukamurella* species were described, only a marginal DDH result of around 62% was observed between *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T] (Park et al., 2009). In this study, we re-characterized these six *Tsukamurella* species by performing the traditional DDH again. In addition, we sequenced their genomes, which represented the first genomes to be sequenced for *T. pulmonis*, *T. spongiae*, *T. tyrosinosolvens*, *T. carboxydivorans*, *T. pseudospumae*, and *T. sunchonensis*, and performed "digital DDH" studies to determine their exact taxonomic positions using a phylogenomics approach.

MATERIALS AND METHODS

Bacterial Strains

The six strains included in this study were type strains and obtained from four culture collections, with *T. pulmonis* CCUG 35732[T] and *T. tyrosinosolvens* CCUG 38499[T] obtained from Culture Collection, University of Gothenburg, Sweden (CCUG), *T. spongiae* DSM 44990[T] from Leibniz Institute DSMZ—German Collection of Microorganisms and Cell Cultures, Germany (DSMZ), and *T. sunchonensis* JCM 15929[T], *T. pseudospumae* JCM 13375[T], and *T. carboxydivorans* JCM 15482[T] from Japan Collection of Microorganisms (JCM), Japan.

Phenotypic Characterizations

The six strains were phenotypically characterized in detail. Bacteria were streaked on Columbia agar with 5% defibrinated sheep blood and incubated at 25°C under aerobic condition for 2 days to observe the color of the colonies. Growth response to different temperatures (25, 37, and 42°C) were determined by spreading bacterial cells of 0.5 McFarland on Columbia agar with 5% defibrinated sheep blood agar followed by incubation under aerobic condition for up to 7 days. Hydrolysis of tyrosine and xanthine were tested as described previously (Conville and Witebsky, 2007). Biochemical data were obtained using the API 50 CH systems (bioMérieux, France) at 25°C according to manufacturers' instructions. Reading intervals and durations of the tests followed the protocol suggested by Kattar et al. (2001). Cellular fatty acids were extracted directly from lyophilized cells (approximately 30 mg of original wet weight) grown on Columbia agar with 5% defibrinated sheep blood at 25°C for 2 days and analyzed by gas chromatography as described by Microbial ID,

Inc. (Sasser, 1990) with peak-naming performed using MIDI, Inc. Sherlock Rapid Libraries and Methods.

16S rRNA Gene Sequencing

The six strains were subjected to 16S rRNA gene sequencing according to previously published protocol using primers LPW27807 (5′-TGGCTCAGGACGAACGCT-3′) and LPW27808 (5′-GAGGTGATCCAGCCGCA-3′; Teng et al., 2015). The sequences of PCR products were compared to known gene sequences in GenBank by multiple sequence alignment using Bioedit 7.1.9 (Hall, 1999). Phylogenetic tree was constructed by maximum likelihood method using the model GTR + I + G in MEGA 6.0 (Tamura et al., 2013).

MALDI-TOF MS Analysis

MALDI-TOF MS of all *Tsukamurella* strains was performed as previously described (Verroken et al., 2010; Lau et al., 2014; Teng et al., 2014) with slight modifications. Briefly, bacteria were first grown aerobically on horse blood agar at 37°C for 2 days. Two to three colonies were scraped in 500 µl of distilled water and boiled for 30 min. After centrifugation at 13,000 rpm for 2 min, 300 µl of water followed by 900 µl of 100% ethanol was added and centrifuged at 13,000 rpm for 2 min twice. Supernatant was removed and pellet was air dried thoroughly. Fifty microliters of 70% formic acid was then added and mixed well with the pellet at room temperature for 15 min. Finally the solution was mixed with 50 µl of acetonitrile and centrifuged at 13,000 rpm for 2 min. One microliter of supernatant was loaded to a 96-well-spot polished steel target plate (Bruker Daltonik, Bremen, Germany) and air dried. One microliter of matrix solution (α-cyano-4-hydroxycinnamic acid; Sigma Aldrich, St Louis, USA) was applied and allowed to air dry. Within the range of *m/z* 2,000–20,000 Da, spectra were obtained with an accelerating voltage of 20 kV and collected in linear mode using the MicroFlex LT (Bruker Daltonik). Resulting data was analyzed with MALDI Biotyper 3.1 and a score of ≥1.7 and ≥2.0, as determined by the Reference Library V.3.1.2.0. Bruker Biotyper, was considered to be confident for identification on the genus and species level, respectively. Main spectral pattern (MSP) created from representative spectra of each strain were selected for hierarchical cluster analysis using the Bruker Biotyper software with default parameters (Ketterlinus et al., 2005).

Wet-Lab DNA–DNA Hybridization

DDH studies were performed between *T. pulmonis* CCUG 35732[T] and *T. spongiae* DSM 44990[T], *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T] and *T. pseudospumae* JCM 13375[T] and *T. sunchonensis* JCM 15929[T], respectively. Preparation of genomic DNA was performed using QIAGEN Genomic tip 100-G (Qiagen, Hilden, Germany). DNA probe was prepared using reagent DIG-High Prime DNA labeling (Roche Diagnostics, Basel, Switzerland) according to Seong et al. (2003), while DNA–DNA dot blot hybridization was carried out using ULTRAhyb ultrasensitive hybridization buffer (Ambion, New York, USA), DIG wash and block buffers, anti-DIG-AP and CDP-Star (Roche Diagnostics) as described previously (Woo et al., 2014; Teng et al., 2015). DDH was performed at 42°C in

triplicate for each of the probe used. To quantify the relative levels of DNA–DNA relatedness, developed X-ray films were scanned and dot intensities were analyzed using the software ImageJ 1.46r (Abràmoff et al., 2004). Self-hybridization values were considered to be 100%, representing the maximal achievable signal. Values obtained with other test strains were compared with this standard (Woo et al., 2014; Teng et al., 2015). Reciprocal hybridizations were also performed.

Genome Sequencing, Assembly, and Annotation of the Six *Tsukamurella* Species

The draft genome sequences of the six *Tsukamurella* species were determined by high-throughput sequencing with Illumina Hi-Seq 2500. Genomic DNA of each strain was extracted from overnight cultures (37°C) grown on blood agars using the genomic DNA purification kit (Qiagen) as described previously (Woo et al., 2009; Tse et al., 2010). Extracted DNA was then sequenced by 151-bp paired-end reads with mean library size of 350 bp. Sequencing errors were corrected by k-mer frequency spectrum analysis using SOAPec[1]. *De novo* assembly was performed using SOAPdenovo2[2]. Prediction of protein coding regions and automatic functional annotation was performed using RAST (Rapid Annotations using Subsystem Technology) server version 2.0[3] (Delcher et al., 2007; Aziz et al., 2008). Genomic islands were predicted by Zisland Explorer method[4] (Wei and Guo, 2011; Wei et al., 2016).

Phylogenomic Characterizations

The phylogenetic tree based on entire genome sequences was constructed by neighbor-joining method using GGDC distance (formula 2) and *Dietzia cinnamea* (NZ_AEKG01000298.1) as the root. The tree was saved in Newick format and subsequently processed by MEGA (Tamura et al., 2013). Intergenomic distance between these six *Tsukamurella* species was calculated using GGDC 2.1[5] and ANI[6] (Goris et al., 2007; Auch et al., 2010). The web service can be used for calculating intergenomic distances, with 70% (GGDC) and 94–96% (ANI) as the substitute for a wet-lab DDH value of 70% for same species delineation.

RESULTS AND DISCUSSION

Phenotypic Characteristics

The biochemical and physiological properties of *T. pulmonis* CCUG 35732[T], *T. spongiae* DSM 44990[T], *T. tyrosinosolvens* CCUG 38499[T], *T. carboxydivorans* JCM 15482[T], *T. pseudospumae* JCM 13375[T], and *T. sunchonensis* JCM 15929[T] and are summarized in **Table 1**. Of the six *Tsukamurella*

[1] http://soap.genomics.org.cn/about.html
[2] http://soap.genomics.org.cn/
[3] http://rast.nmpdr.org/rast.cgi
[4] http://cefg.uestc.edu.cn/Zisland_Explorer/
[5] http://ggdc.dsmz.de/distcalc2.php
[6] http://enve-omics.ce.gatech.edu/ani/

species, all were aerobic, non-sporulating, Gram-positive bacilli, and grew on sheep blood agar as white or orange to red colonies of <5 mm in diameter after 48 h of incubation at 25°C in ambient air. All of them grew well at 25 and 37°C but none of them grew at 42°C. *T. tyrosinosolvens* CCUG 38499[T], *T. carboxydivorans* JCM 15482[T], *T. pseudospumae* JCM 13375[T], and *T. sunchonensis* JCM 15929[T] could hydrolyse tyrosine but only *T. carboxydivorans* JCM 15482[T] could hydrolyse xanthine. Sugar assimilation results using API 50 CH systems (bioMérieux, France) showed almost identical phenotypic profiles between *T. pulmonis* CCUG 35732[T] and *T. spongiae* DSM 44990[T] (48 of 49 sugar assimilation results were identical), between *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T] (48 of 49 sugar assimilation results were identical), and between *T. pseudospumae* JCM 13375[T] and *T. sunchonensis* JCM 15929[T] (45 of 49 sugar assimilation results were identical). The fatty acid profiles of the six *Tsukamurella* species are summarized in **Table 2**. The major fatty acids compositions of strains *T. pulmonis* CCUG 35732[T], *T. spongiae* DSM 44990[T], *T. tyrosinosolvens* CCUG 38499[T], *T. carboxydivorans* JCM 15482[T], *T. pseudospumae* JCM 13375[T], and *T. sunchonensis* JCM 15929[T] included $C_{16:0}$, $C_{18:1\omega9c}$, and 10-methyl $C_{18:0}$, which is typical of members of *Tsukamurella* (**Table 2**; Goodfellow and Kumar, 2012).

16S rRNA Gene Sequencing

Sequencing of the 16S rRNA gene of the six *Tsukamurella* species showed that there were 5–56 (0.3–3.8%) base differences between their 16S rRNA gene sequences and those of the type strains of other currently recognized *Tsukamurella* species. Pairwise sequence similarity calculations showed identical 16S rRNA gene sequences between *T. pulmonis* CCUG 35732[T] and *T. spongiae* DSM 44990[T], between *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T], and between *T. pseudospumae* JCM 13375[T] and *T. sunchonensis* JCM 15929[T] (**Figure 1**).

MALDI-TOF MS Analysis

Consistent with the results of phenotypic characterization, dendrogram generated from hierarchical cluster analysis of MALDI-TOF MS showed that the spectra of these six *Tsukamurella* species and the representatives of the only two *Tsukamurella* species, *Tsukamurella inchonensis* and *Tsukamurella paurometabola*, available in the Bruker reference library revealed the same clustering pattern (**Figure 2**). Specifically, same cluster was observed between *T. pulmonis* CCUG 35732[T] and *T. spongiae* DSM 44990[T], between *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T], and between *T. pseudospumae* JCM 13375[T] and *T. sunchonensis* JCM 15929[T] (**Figure 2**).

Wet-lab DNA–DNA Hybridization

Using the DDH, we found that DNA–DNA relatedness between *T. pulmonis* CCUG 35732[T] and *T. spongiae* DSM 44990[T], *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T], and *T. pseudospumae* JCM 13375[T] and *T. sunchonensis* JCM 15929[T] were demonstrated to be >70%, respectively. *T. pulmonis* CCUG 35732[T] and *T. spongiae* DSM 44990[T]

TABLE 1 | Phenotypic properties of *T. pulmonis* CCUG 35732[T], *T. spongiae* DSM 44990[T], *T. tyrosinosolvens* CCUG 38499[T], *T. carboxydivorans* JCM 15482[T], *T. pseudospumae* JCM 13375[T], and *T. sunchonensis* JCM 15929[T].

Characteristics	*T. pulmonis* CCUG 35732[T]	*T. spongiae* DSM 44990[T]	*T. tyrosinosolvens* CCUG 38499[T]	*T. carboxydivorans* JCM 15482[T]	*T. pseudospumae* JCM 13375[T]	*T. sunchonensis* JCM 15929[T]
Colony color	White	Orange, red	White	White	Orange, red	Orange, red
Hydrolysis of						
Xanthine	−	−	−	+	−	−
Tyrosine	−	−	+	+	+	+
Growth at						
25°C	+	+	+	+	+	+
37°C	+	+	+	+	+	+
42°C	−	−	−	−	−	−
Assimilation of *						
Glycerol	+	+	+	+	+	+
D-Arabinose	+	+	+	+	−	−
D-Ribose	−	−	+	+	+	+
D-Galactose	+	+	+	+	+	+
D-Glucose	+	+	+	+	+	+
D-Fructose	+	+	+	+	+	+
D-Mannose	−	−	+	+	+	+
myo-Inositol	−	−	+	+	−	+
D-Mannitol	+	+	+	+	−	+
D-Sorbitol	+	+	+	+	−	+
methyl-D-Glucopyranoside	−	−	+	+	+	+
N-acetyl glucosamine	+	+	+	+	+	+
Arbutin	−	−	+	+	−	−
Esculin ferric citrate	+	+	+	+	+	+
D-Salicin	−	−	+	−	−	−
D-Cellobiose	−	+	−	−	−	−
D-Maltose	−	−	+	+	+	+
D-Saccharose	+	+	+	+	+	+
D-Trehalose	+	+	+	+	+	+
D-Melezitose	−	−	+	+	+	+
Xylitol	−	−	+	+	−	−
D-Turanose	+	+	+	+	+	+
L-Fucose	+	+	+	+	−	−
D-Arabitol	+	+	+	+	−	+
Potassium gluconate	+	+	+	+	+	+
2-Ketogluconate	−	−	+	+	−	+

Negative results in all six isolates were not shown in this table. All results are obtained from this study.

possessed 91.1 ± 3.9% (mean ± SD, $n = 3$, *T. pulmonis* CCUG 35732[T] as probe) or 84.2 ± 3.2% DNA–DNA relatedness (mean ± SD, $n = 3$, *T. spongiae* DSM 44990[T] as probe). *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T] possessed 90.5 ± 1.3% (mean ± SD, $n = 3$, *T. tyrosinosolvens* CCUG 38499[T] as probe) or 81.2 ± 0.6% DNA–DNA relatedness (mean ± SD, $n = 3$, *T. carboxydivorans* JCM 15482[T] as probe). *T. pseudospumae* JCM 13375[T] and *T. sunchonensis* JCM 15929[T] possessed 95.1 ± 4.9% DNA–DNA relatedness (mean ± SD, $n = 3$, *T. pseudospumae* JCM 13375[T] as probe) or 99.7 ± 1.0% (mean ± SD, $n = 3$, *T. sunchonensis* JCM 15929[T] as probe). According to the criterion for species delineation using DDH, the present results (>70% DDH) suggested that these six *Tsukamurella* species should be reclassified as three distinct species. It is worth noting though, that the present DDH values are different from those obtained in the previous studies, in which the DNA–DNA relatedness between *T. pulmonis* CCUG 35732[T] and *T. spongiae* DSM 44990[T] was 48.0 ± 1.3% (mean ± SD, $n = 3$, *T. spongiae* DSM 44990[T] as probe; Olson et al., 2007) and that of *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T] was 62.7 ± 3.4% (mean ± SD, $n = 3$, *T. carboxydivorans* JCM 15482[T] as probe; Park et al., 2009). To further verify our results, we sequenced the genomes of these six *Tsukamurella* species for accurate species classification.
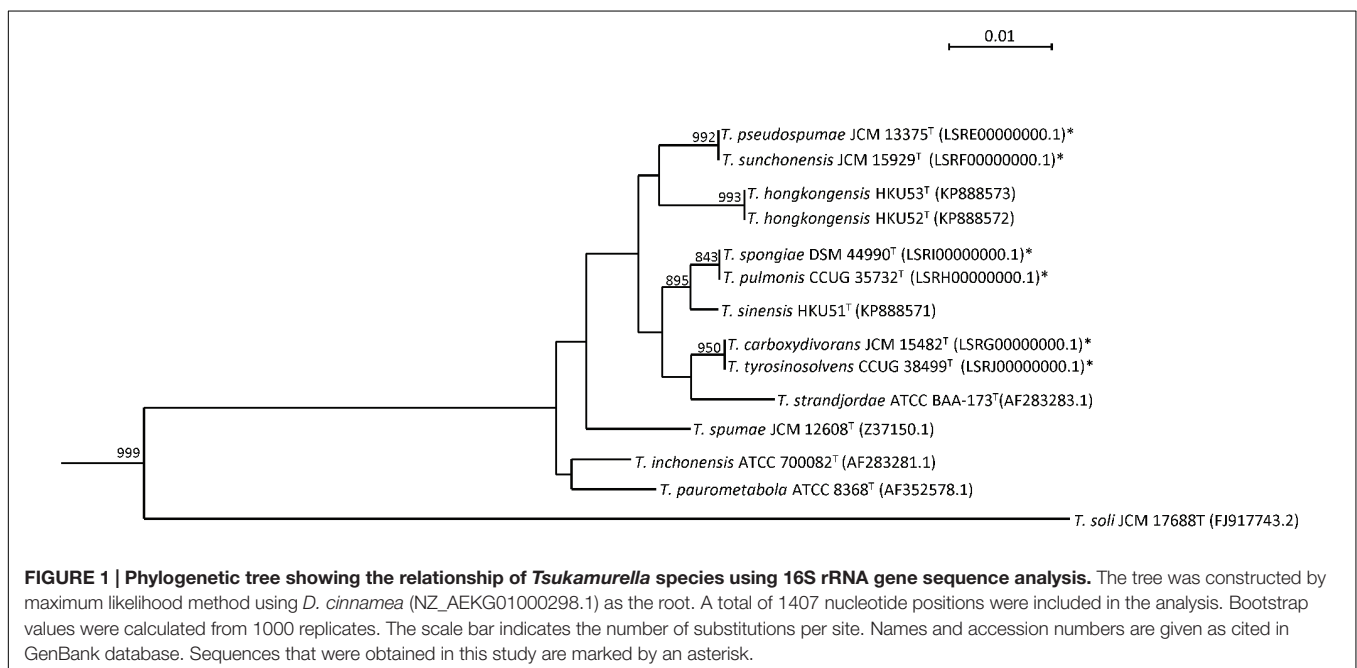
General Features of the Genomes

We sequenced the first draft genomes of the type strains of the six *Tsukamurella* species using high-throughput sequencing to investigate their genetic relatedness and to confirm their taxonomic positions. Sequencing generated 2.12 million

TABLE 2 | Fatty acid compositions (%) of *T. pulmonis* CCUG 35732^T, *T. spongiae* DSM 44990^T, *T. tyrosinosolvens* CCUG 38499^T, *T. carboxydivorans* JCM 15482^T, *T. pseudospumae* JCM 13375^T, and *T. sunchonensis* JCM 15929^T.

	T. pulmonis CCUG 35732^T	*T. spongiae* DSM 44990^T	*T. tyrosinosolvens* CCUG 38499^T	*T. carboxydivorans* JCM 15482^T	*T. pseudospumae* JCM 13375^T	*T. sunchonensis* JCM 15929^T
$C_{14:0}$	4.0	3.4	4.1	3.9	4.1	2.5
$C_{16:1\omega9c}$	2.5	1.1	2.6	2.0	2.4	1.6
$C_{16:0}$	22.6	22.1	23.0	25.1	26.4	25.4
$C_{17:1\omega8c}$	–	–	1.2	1.2	–	–
$C_{17:0}$	–	–	1.0	1.5	–	–
$C_{18:1\omega9c}$	34.9	36.4	32.8	33.5	36.3	27.4
$C_{18:0}$	3.1	2.7	1.5	1.4	1.0	1.5
10-methyl $C_{18:0}$	10.7	10.0	13.4	10.8	10.9	21.0
$C_{20:1\omega9c}$	3.3	1.1	1.8	2.0	1.6	1.2
SF1	15.2	19.5	16.4	15.2	15.6	16.2

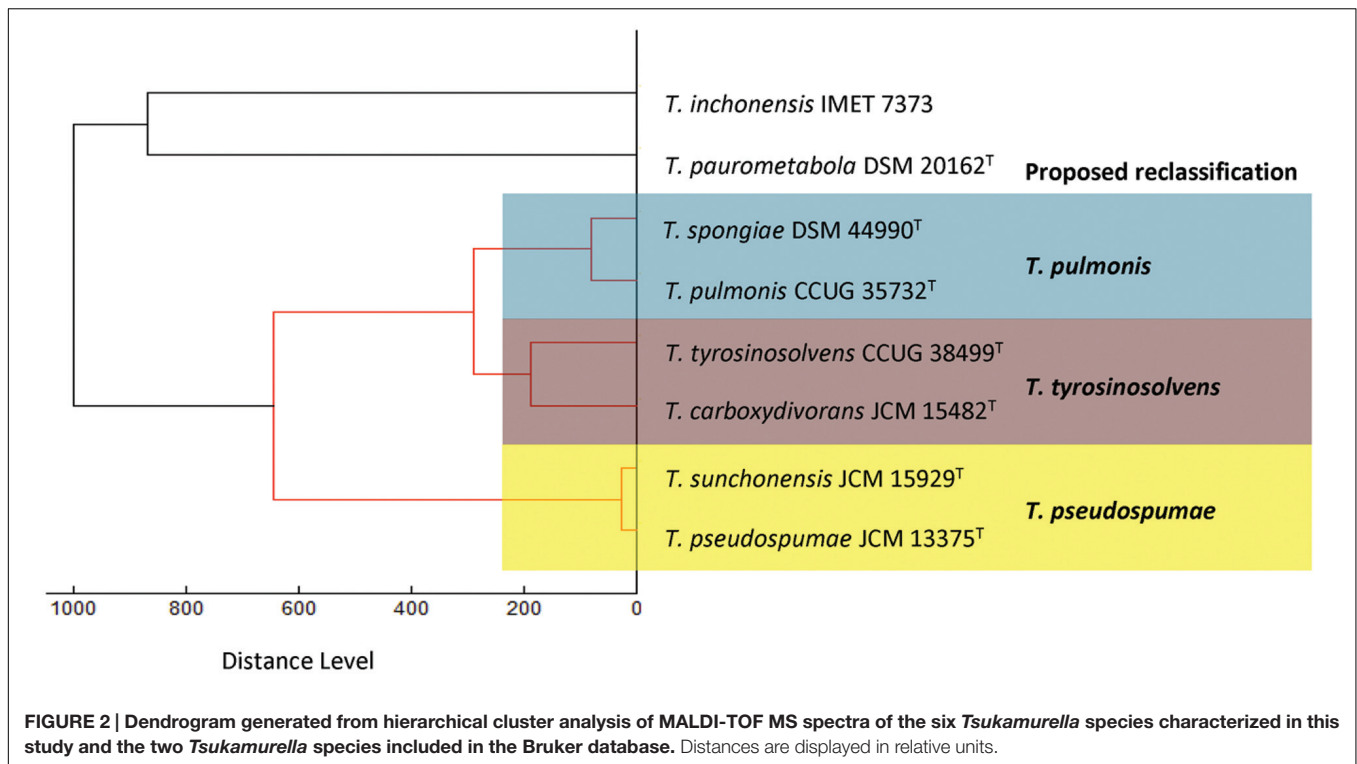SF1 (summed feature 1): C16:1ω6c and/or C16:1ω7c and/or iso-C15:0 2-OH; All results were obtained from this study.
–, Less than 1% or not detected.

FIGURE 1 | Phylogenetic tree showing the relationship of *Tsukamurella* species using 16S rRNA gene sequence analysis. The tree was constructed by maximum likelihood method using *D. cinnamea* (NZ_AEKG01000298.1) as the root. A total of 1407 nucleotide positions were included in the analysis. Bootstrap values were calculated from 1000 replicates. The scale bar indicates the number of substitutions per site. Names and accession numbers are given as cited in GenBank database. Sequences that were obtained in this study are marked by an asterisk.

TABLE 3 | Results of draft genome assembly of the six *Tsukamurella* species.

Genome assembly data	*T. pulmonis* CCUG 35732^T	*T. spongiae* DSM 44990^T	*T. tyrosinosolvens* CCUG 38499^T	*T. carboxydivorans* JCM 15482^T	*T. pseudospumae* JCM 13375^T	*T. sunchonensis* JCM 15929^T
Genome size (Mb)	4.74	4.60	5.23	4.95	4.95	5.21
G+C content (%)	70.4	71.1	71.2	71.2	70.6	70.5
No. of contigs	62	51	73	29	70	133
No. of predicted protein-coding genes	4,538	4,412	5,191	4,709	4,823	5,050

paired-end reads per strain (estimated 80× coverage). After *de novo* assembly, the six draft genomes ranged from 4.6 to 5.2 Mb in length (G+C content 70.4–71.2%) distributed in 18–74 large contigs (>500 bp; GenBank accession numbers LSRE01000000–LSRJ01000000). All contigs generated for each strain were submitted to the RAST version 2.0 annotation server respectively, resulting in 4,412–5,191 protein-coding sequences (CDSs; **Table 3**). To facilitate the subsequent genomic analysis, genome sequences of *T. paurometabola* DSM 20162^T (GenBank accession number NC_014158.1) and "*T. algeriensis*" 1534 (GenBank accession number NZ_HE997626.1), the only two available genome sequences in GenBank, were also

FIGURE 2 | Dendrogram generated from hierarchical cluster analysis of MALDI-TOF MS spectra of the six *Tsukamurella* species characterized in this study and the two *Tsukamurella* species included in the Bruker database. Distances are displayed in relative units.
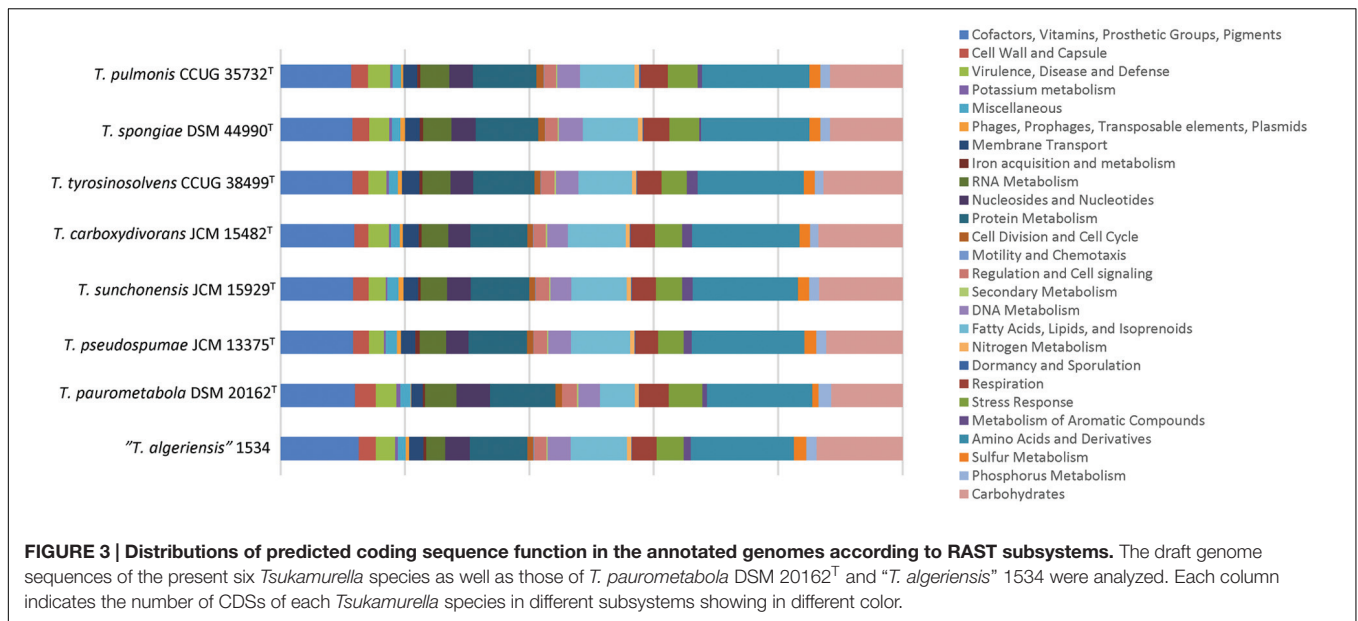
submitted to RAST for annotation. Each CDS in annotated genomes was grouped into different RAST subsystems based on the predicted functional role. Among the 4,412–5,191 CDS in these six draft genomes, only 1,629–1,762 CDSs can be categorized into RAST systems, representing 34–38% of total CDSs. Overall, the majority of CDSs were classified into subsystems of amino acid and derivatives (444–490 CDSs, 16.9–18.1%), carbohydrates (297–380 CDSs, 11.7–13.5%), cofactors, vitamins, prosthetic groups, pigments (294–335 CDSs, 11.4–11.9%) and protein metabolism (255–271 CDSs, 9.2–10.2%). The remaining 2,783–3,452 (62–66%) CDSs could not be classified into any subsystems, in which 1,450–1,720 (31.6–34.6%) CDSs were only annotated as hypothetical proteins. When we compared the distribution of CDSs in each subsystem of these six *Tsukamurella* genomes with those of *T. paurometabola* and "*T. algeriensis*" genomes, all of them have a similar percentage of their genome dedicated to each subsystem (**Figure 3**).

Phylogenomic Analyses of the Six *Tsukamurella* Species

With the availability of these six *Tsukamurella* genome sequences, as well as the genome sequences of *T. paurometabola* DSM 20162[T] and "*T. algeriensis*" 1534 from GenBank, we used two independent "digital DDH" methods, GGDC and ANI, to estimate the overall similarity between the genomes of two strains, which in turn determined the genome-based species delineation among these six *Tsukamurella* species. Consistent with the results obtained by traditional DDH, *in silico* genome-to-genome comparison unambiguously showed that for *T. pulmonis*

CCUG 35732[T] and *T. spongiae* DSM 44990[T], *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T], and *T. pseudospumae* JCM 13375[T] and *T. sunchonensis* JCM 15929[T], GGDC and ANI values for more than 70 and 94–96%, respectively were observed (**Table 4**). As these calculated values are above the threshold for same species designation, it was suggested that the current six *Tsukamurella* species should be reclassified as only three distinct *Tsukamurella* species.
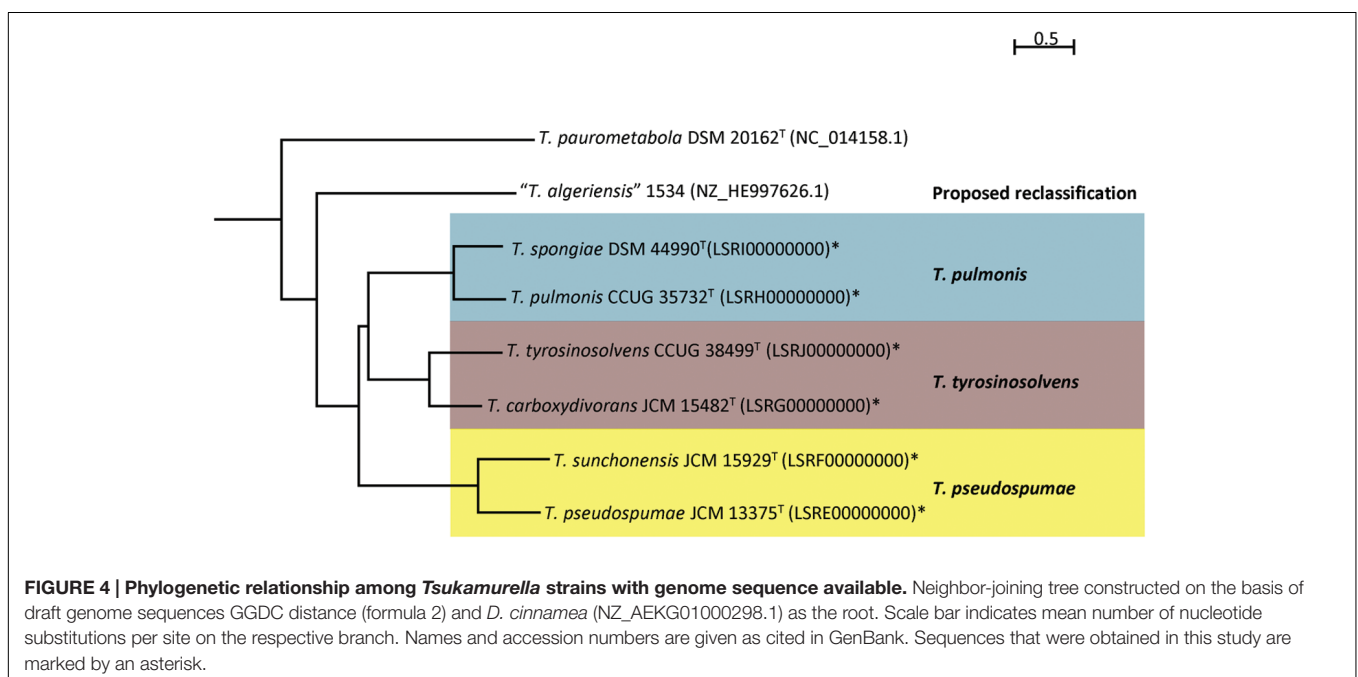
The proposed reclassification is also supported by the phylogenomic analyses (**Figure 4**) and whole genome sequence comparison (**Figure 5**) using the draft genome sequences of the six *Tsukamurella* species. The phylogenetic tree constructed using the draft genomes sequences (**Figure 4**) was concordant to the clustering pattern as observed in phenotypic (**Table 1**) and MALDI-TOF MS analyses (**Figure 2**). As for genome sequence comparison, high protein sequence identities (>90% as indicated by blue color) were observed in each orthologous gene between the genomes of *T. pulmonis* CCUG 35732[T] and *T. spongiae* DSM 44990[T] (**Figure 5A**), between the genomes of *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T] (**Figure 5B**), and between the genomes of *T. pseudospumae* JCM 13375[T] and *T. sunchonensis* JCM 15929[T] (**Figure 5C**). Interestingly, although relatively high sequence similarity was observed between the genomes of *T. tyrosinosolvens* CCUG 38499[T] and *T. carboxydivorans* JCM 15482[T], some regions were found to be uniquely present in the type strain of *T. tyrosinosolvens* only. These unique regions are likely to be genomic islands as predicted by Zisland Explorer method (**Figure 5B**). RAST analysis showed that most of the ORFs within these putative genomic islands were annotated as hypothetic proteins, phage

FIGURE 3 | Distributions of predicted coding sequence function in the annotated genomes according to RAST subsystems. The draft genome sequences of the present six *Tsukamurella* species as well as those of *T. paurometabola* DSM 20162[T] and "*T. algeriensis*" 1534 were analyzed. Each column indicates the number of CDSs of each *Tsukamurella* species in different subsystems showing in different color.
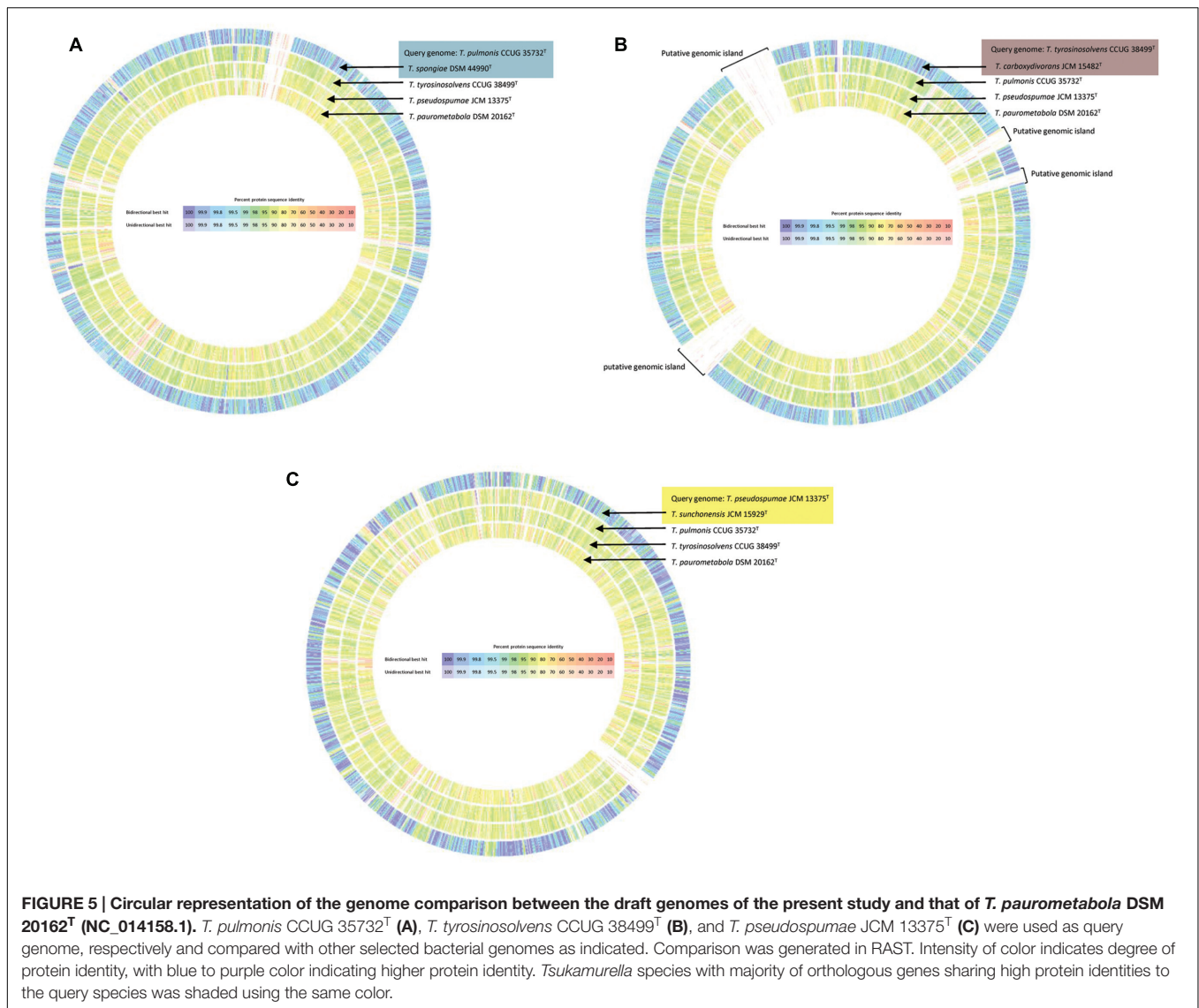
TABLE 4 | Intergenomic distance between the draft genomes of the six currently recognized *Tsukamurella* species represented by ANI (upper triangle) and GGDC (lower triangle) values (%).

Strains	*T. pulmonis* CCUG 35732[T]	*T. spongiae* DSM 44990[T]	*T. tyrosinosolvens* CCUG 38499[T]	*T. carboxydivorans* JCM 15482[T]	*T. pseudospumae* JCM 13375[T]	*T. sunchonensis* JCM 15929[T]
T. pulmonis CCUG 35732[T]	–	**99.1**	87.3	87.2	84.6	84.8
T. spongiae DSM 44990[T]	**92.9**	–	87.3	87.3	84.5	84.8
T. tyrosinosolvens CCUG 38499[T]	34.3	34.3	–	**98.8**	85.1	85.1
T. carboxydivorans JCM 15482[T]	34.0	33.9	**90.4**	–	85.0	84.8
T. pseudospumae JCM 13375[T]	28.4	28.4	29.4	29.4	–	**98.1**
T. sunchonensis JCM 15929[T]	28.8	28.8	29.5	29.5	**82.3**	–

Species with >70% GGDC and >94–96% ANI values, suggesting same species delineation, are bolded.

FIGURE 4 | Phylogenetic relationship among *Tsukamurella* strains with genome sequence available. Neighbor-joining tree constructed on the basis of draft genome sequences GGDC distance (formula 2) and *D. cinnamea* (NZ_AEKG01000298.1) as the root. Scale bar indicates mean number of nucleotide substitutions per site on the respective branch. Names and accession numbers are given as cited in GenBank. Sequences that were obtained in this study are marked by an asterisk.

FIGURE 5 | Circular representation of the genome comparison between the draft genomes of the present study and that of *T. paurometabola* DSM 20162^T (NC_014158.1). *T. pulmonis* CCUG 35732^T **(A)**, *T. tyrosinosolvens* CCUG 38499^T **(B)**, and *T. pseudospumae* JCM 13375^T **(C)** were used as query genome, respectively and compared with other selected bacterial genomes as indicated. Comparison was generated in RAST. Intensity of color indicates degree of protein identity, with blue to purple color indicating higher protein identity. *Tsukamurella* species with majority of orthologous genes sharing high protein identities to the query species was shaded using the same color.

related proteins and mobile element proteins. As the type strain of *T. tyrosinosolvens* was isolated from blood culture of a patient with cardiac pacemaker implants, whereas *T. carboxydivorans* JCM 15482^T was isolated from soil, the pathogenic role of these genomic islands warrants further investigations. Given all genomic evidence, *T. spongiae* DSM 44990^T, *T. carboxydivorans* JCM 15482^T, and *T. sunchonensis* JCM 15929^T likely represent previously misclassified isolates that actually belong to the same species of *T. pulmonis*, *T. tyrosinosolvens*, and *T. sunchonensis*, respectively. According to Rules 23a and 42 of the Bacteriological Code (Lapage et al., 1992), we propose that *T. spongiae* Olson et al. 2007 (Olson et al., 2007), *T. carboxydivorans* Park et al. 2009 (Park et al., 2009), and *T. sunchonensis* Seong et al. 2008 (Seong et al., 2003) should be reclassified as later heterotypic synonyms of *T. pulmonis* Yassin et al. 1996 (Yassin et al., 1996), *T. tyrosinosolvens* Yassin et al. 1997 (Yassin et al., 1997), and *T. pseudospumae* Nam et al. 2004 (Nam et al., 2004), respectively.

PROPOSAL FOR A RECLASSIFICATION OF SPECIES WITHIN THE GENUS *TSUKAMURELLA*

In this study, we sequenced the first draft genomes of six *Tsukamurella* species and unambiguously determined their phylogenetic positions using a phylogenomics approach. Consistent with the results of 16S rRNA gene sequencing, phenotypic and MALDI-TOF MS analyses, additional genomic evidences, including intergenomic distance comparisons using GGDC and ANI, phylogenetic analysis and whole genome comparison based on the draft genomes, support that the currently recognized six *Tsukamurella* species should be reclassified as three species only. According to Rules 23a and 42 of the *Bacteriological Code* (Lapage et al., 1992), it is proposed that *T. spongiae* Olson et al. 2007 (Olson et al., 2007) should be reclassified as a later heterotypic synonym of *T. pulmonis*

Yassin et al. 1996 (Yassin et al., 1996), *T. carboxydivorans* Park et al. 2009 (Park et al., 2009) as a later heterotypic synonym of *T. tyrosinosolvens* Yassin et al. 1997 (Yassin et al., 1997), and *T. sunchonensis* Seong et al. 2008 (Seong et al., 2003) as a later heterotypic synonym of *T. pseudospumae* Nam et al. 2004 (Nam et al., 2004), respectively. With the decreasing costs of high-throughput sequencing genomic data may be considered as one of the taxonomic criteria for species delineation of bacteria, especially in the case where two strains share highly similar phenotypic characteristics and 16S rRNA gene sequences.

Emended Description of *T. pulmonis* Yassin et al. 1996 (Yassin et al., 1996)

Tsukamurella pulmonis (pul. mo' nis. L. gen. masc. n. pulmonis, of the lung, referring to the organ from which the bacterium was isolated).

Heterotypic Synonym: *T. spongiae* Olson et al. 2007 (Olson et al., 2007)

The characteristics of *T. pulmonis* are described by Yassin et al. (1996), with the following amendments. Grows best on Columbia agar with 5% defibrinated sheep blood agar as white or orange to red and rough or easily emulsified colonies, 2 mm in diameter, with irregular spreading margins after 48 h of incubation at 25°C under aerobic conditions. As determined by API 50 CH kit, cells are able to assimilate glycerol, D-arabinose, D-galactose, D-glucose, D-fructose, D-mannitol, D-sorbitol, N-acetyl glucosamine, esculin ferric citrate, D-saccharose, D-trehalose, D-turanose, L-fucose, D-arabitol, and potassium gluconate. The major fatty acids (>10% of the total) are $C_{16:0}$, $C_{18:1\omega9c}$, 10-methyl $C_{18:0}$ and summed feature 1 ($C_{16:1\omega6c}$ and/or $C_{16:1\omega7c}$ and/or iso-$C_{15:0}$ 2-OH).

The type strain, CCUG 35732^T (=JCM 10111^T = DSM 44142^T), was isolated from human with mycobacterial lung infection. The G+C content of the type strain is 70.4 mol%. The reference strain, DSM 44990 (=JCM 14882), is the nomenclatural type of *T. spongiae* Olson et al. 2007 (Olson et al., 2007), and was isolated from a deep-water marine sponge. The G+C content of strain DSM 44990 is 71.1 mol%. The GenBank accession numbers of whole genome sequence of the type strain and reference strain are LSRH00000000.1 and LSRI00000000.1, respectively.

Emended Description of *T. tyrosinosolvens* Yassin et al. 1997 (Yassin et al., 1997)

Tsukamurella tyrosinosolvens (ty. ro. si. no. sol' vens. Gr. masc. n. tyros, cheese; *tyrosine*, an amino acid; L. pres. part. *solvens*, dissolving; M.L. adj. *tyrosinoslovens*, tyrosine dissolving, referring to the hydrolysis of tyrosine which is characteristics of this species.

Heterotypic Synonym: *T. carboxydivorans* Park et al. 2009 (Park et al., 2009)

The characteristics of *T. tyrosinosolvens* are described by Yassin et al. (1997), with the following amendments. As determined by API 50 CH kit, cells are capable of assimilating glycerol,

D-arabinose, D-ribose, D-galactose, D-glucose, D-fructose, D-mannose, *myo*-inositol, D-mannitol, D-sorbitol, *methyl*-D-glucopyranoside, N-acetyl glucosamine, arbutin, esculin ferric citrate, D-maltose, D-saccharose, D-trehalose, D-melezitose, xylitol, D-turanose, L-fucose, D-arabitol, potassium gluconate, and 2-ketogluconate. The major fatty acids (>10% of the total) are $C_{16:0}$, $C_{18:1\omega9c}$, 10-methyl $C_{18:0}$ and summed feature 1 ($C_{16:1\omega6c}$ and/or $C_{16:1\omega7c}$ and/or iso-$C_{15:0}$ 2-OH).

The type strain, CCUG 38499^T (=JCM 10112^T = DSM 44234^T), was isolated from blood culture of a patient with cardiac pacemaker implants. The G+C content of the type strain is 71.2 mol%. The reference stain, JCM 15482 (=KCCM 52885), is the nomenclatural type of *T. carboxydivorans* Park et al. 2009 (Park et al., 2009), and was isolated from soil sample from a roadside. The G+C content of strain JCM 15482 is 71.2 mol%. The GenBank accession numbers of whole genome sequence of the type strain and reference strain are LSRJ00000000.1 and LSRG00000000.1, respectively.

Emended Description of *T. pseudospumae* Nam et al. 2004 (Nam et al., 2004)

Tsukamurella pseudospumae (pseu.do.spu'mae. Gr. adj. *pseudes* false; L. gen. n. *spumae* of foam and specific epithet of bacterial species; N. L. n. *pseudospumae* the false *spumae*, referring to the close relationship to *Tsukamurella spumae*).

Heterotypic Synonym: *T. sunchonensis* Seong et al. 2008 (Seong et al., 2003)

The characteristics of *T. pseudospumae* are described by Nam et al. (2004), with the following amendments. As determined by API 50 CH kit, cells are capable to grow on brain–heart infusion agar. Positive for hydrolysis of tyrosine. Assimilates glycerol, D-ribose, D-galactose, D-glucose, D-fructose, D-mannose, *methyl*-D-glucopyranoside, N-acetyl glucosamine, esculin ferric citrate, D-maltose, D-saccharose, D-trehalose, D-melezitose, D-turanose, and potassium gluconate. The major fatty acids (>10 % of the total) are $C_{16:0}$, $C_{18:1\omega9c}$, 10-methyl $C_{18:0}$ and summed feature 1 ($C_{16:1\omega6c}$ and/or $C_{16:1\omega7c}$ and/or iso-$C_{15:0}$ 2-OH).

The type strain, JCM 13375^T (=DSM 44118^T), was isolated from extensive foam in the aeration basin of an activated sludge process. The G+C content of the type strain is 70.6 mol%. The reference strain, JCM 15929 (=NRRL B-24668), is the nomenclatural type of *T. sunchonensis* Seong et al. 2008 (Seong et al., 2003), and was also isolated from activated sludge foam. The G+C content of strain JCM 15929 is 70.5 mol%. The GenBank accession numbers of whole genome sequence of the type strain and reference strain are LSRF00000000.1 and LSRE00000000.1, respectively.

DATA ACCESSIBILITY

This Whole Genome Shotgun project has been deposited in DDBJ/ENA/GenBank under the accession numbers LSRE00000000–LSRJ00000000. The version described in

this paper is version LSRE00000000.1–LSRJ00000000.1. The GenBank accession numbers for the draft genome sequences of *Tsukamurella pseudospumae* JCM 13375[T], *Tsukamurella sunchonensis* JCM 15929[T], *Tsukamurella carboxydivorans* JCM 15482[T], *Tsukamurella pulmonis* CCUG 35732[T], *Tsukamurella spongiae* DSM 44990[T] and *Tsukamurella tyrosinosolvens* CCUG 38499[T] are LSRE00000000.1, LSRF00000000.1, LSRG00000000.1, LSRH00000000.1, LSRI00000000.1, and LSRJ00000000.1, respectively.

AUTHOR CONTRIBUTIONS

JT conceived of the study, designed the study, carried out the molecular lab work, participated in data analysis, and drafted the manuscript; YT carried out the molecular lab work, participated in data analysis, and helped draft the manuscript. YH, F-BG, WW, and JC participated in data analysis; SW contributed reagents; SL revised the manuscript and contributed reagents; PW conceived of the study, designed the study, contributed reagents, and drafted the manuscript. All authors gave final approval for publication.

REFERENCES

Abràmoff, M. D., Magalhaes, P. J., and Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics Int.* 11, 36–42.

Almehmi, A., Pfister, A. K., McCowan, R., and Matulis, S. (2004). Implantable cardioverter-defibrillator infection caused by *Tsukamurella*. *W. V. Med. J.* 100, 185–186.

Auch, A. F., Klenk, H. P., and Göker, M. (2010). Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. *Stand. Genomic Sci.* 2, 142–148.doi: 10.4056/sigs.541628

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75.doi: 10.1186/1471-2164-9-75

Bouza, E., Pérez-Parra, A., Rosal, M., Martín-Rabadán, P., Rodríguez-Créixems, M., and Marín, M. (2009). *Tsukamurella*: a cause of catheter-related bloodstream infections. *Eur. J. Clin. Microbiol. Infect. Dis.* 28, 203–210.doi: 10.1007/s10096-008-0607-2

Collins, M. D., Goodfellow, M., and Minnikin, D. E. (1982). A survey of the structures of mycolic acids in *Corynebacterium* and related taxa. *J. Gen. Microbiol.* 128, 129–149. doi: 10.1099/00221287-128-1-129

Collins, M. D., and Jones, D. (1982). Lipid composition of *Corynebacterium paurometabolum* (Steinhaus). *FEMS Microbiol. Lett.* 13, 13–16. doi: 10.1111/j.1574-6968.1982.tb08217.x

Collins, M. D., Smida, J., Dorsch, M., and Stackebrandt, E. (1988). *Tsukamurella* gen. nov., harboring *Corynebacterium paurometabolum* and *Rhodococcus aurantiacus*. *Int. J. Syst. Bacteriol.* 38, 385–391.doi: 10.1099/00207713-38-4-385

Conville, P. S., and Witebsky, F. G. (2007). "*Nocardia, Rhodococcus, Gordonia, Actinomadura, Streptomyces*, and other aerobic actinomycetes," in *Manual of Clinical Microbiology,* 9th Edn, eds P. R. Murray, E. J. Baron, J. H. Jorgensen, M. R. Louise, and M. A. Pfaller (Washington, DC: American Society for Microbiology), 515–542.

Delcher, A. L., Bratke, K. A., Powers, E. C., and Salzberg, S. L. (2007). Identifying bacterialgenes and endosymbiont DNA with Glimmer. *Bioinformatics* 23, 673–679.doi: 10.1093/bioinformatics/btm009

Euzéby, J. (2008). Validation List 123. *Int. J. Syst. Evol. Microbiol.* 58, 1993–1994.

Fox, G. E., Magrum, L. J., Balch, W. E., Wolfe, R. S., and Woese, C. R. (1977). Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proc. Natl. Acad. Sci. U.S.A.* 74, 4537–4541.doi: 10.1073/pnas.74.10.4537

Goodfellow, M., and Kumar, Y. (2012). "Genus I. *Tsukamurella* Collins, Smida, Dorsch and Stackebrandt 1988, 387," in *Bergey's Mannual of Systematic Bacteriology*, 2nd Edn, eds M. Goodfellow, P. Kämpfer, H.-J. Busse, M. E. Trujillo, K.-I. Suzuki, W. Ludwig et al. (New York, NY: Springer), 500–509.

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91.doi: 10.1099/ijs.0.64483-0

Gupta, R., Lanter, J. M., and Woese, C. R. (1983). Sequence of the 16S ribosomal RNA from *Halobacterium volcanii*, an archaebacterium. *Science* 221, 656–659.doi: 10.1126/science.221.4611.656

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Jiang, Y., Chen, X., Han, L., Li, Q., Huang, X., Qiu, S. M., et al. (2013). Diversity of cultivable actinomycetes in 6 species of herbivore feces. *Int. J. Microbiol. Res.* 1, 76–84.

Kattar, M. M., Cookson, B. T., Carlson, L. D., Stiglich, S. K., Schwartz, M. A., Nguyen, T. T., et al. (2001). *Tsukamurella strandjordae* sp. nov., a proposed new species causing sepsis. *J. Clin. Microbiol.* 39, 1467–1476.doi: 10.1128/JCM.39.4.1467-1476.2001

Ketterlinus, R., Hsieh, S. Y., Teng, S. H., Lee, H., and Pusch, W. (2005). Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software. *Biotechniques* 38(Suppl. 6), 37–40. doi: 10.2144/05386SU07

Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2567–2572.doi: 10.1073/pnas.0409727102

Lapage, S. P., Sneath, P. H. A., Lessel, E. F., Skerman, V. B. D., Seeliger, H. P. R., and Clark, W. A. (eds) (1992). *International Code of Nomenclature of Bacteria (1990 Revision). Bacteriological Code*. Washington, DC: American Society for Microbiology.

Larkin, J. A., Lit, L., Sinnott, J., Wills, T., and Szentivanyi, A. (1999). Infection of aknee prosthesis with *Tsukamurella* species. *South. Med. J.* 92, 831–832.doi: 10.1097/00007611-199908000-00019

Lau, S. K., Tang, B. S., Teng, J. L., Chan, T. M., Curreem, S. O., Fan, R. Y., et al. (2014). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry for identification of clinically significant bacteria that are difficult to identify in clinical laboratories. *J. Clin. Pathol.* 67, 361–366.doi: 10.1136/jclinpath-2013-201818

Liu, C. Y., Lai, C. C., Lee, M. R., Lee, Y. C., Huang, Y. T., Liao, C. H., et al. (2011). Clinical characterics of infections caused by *Tsukamurella* spp. and antimicrobial susceptibilities of the isolates. *Int. J. Antimicrob. Agents* 38, 534–537.doi: 10.1016/j.ijantimicag.2011.07.018

Maeda, Y., Stanley, T., Stirling, J., Griffiths, M., Calvert, A., Stuart Elborn, J., et al. (2010). No evidence of transmission of bacteria between reptiles and a CF

patient-a case report of a young adult CF patient and reptiles. *Zoonoses Public Health* 57:e47–e53.doi: 10.1111/j.1863-2378.2009.01293.x

Nam, S. W., Kim, W., Chun, J., and Goodfellow, M. (2004). *Tsukamurella pseudospumae* sp. nov., a novel actinomycete isolated from activated sludge foam. *Int. J. Syst. Evol. Microbiol.* 54, 1209–1212.doi: 10.1099/ijs.0.02939-0

Olson, J. B., Harmody, D. K., Bej, A. K., and McCarthy, P. J. (2007). *Tsukamurella spongiae* sp. nov., a novel actinomycete isolated from a deep-water marine sponge. *Int. J. Syst. Evol. Microbiol.* 57, 1478–1481.doi: 10.1099/ijs.0.64837-0

Palys, T., Nakamura, L. K., and Cohan, F. M. (1997). Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int. J. Syst. Bacteriol.* 47, 1145–1156.doi: 10.1099/00207713-47-4-1145

Park, S. W., Kim, S. M., Park, S. T., and Kim, Y. M. (2009). *Tsukamurella carboxydivorans* sp. nov., a carbon monoxide-oxidizing actinomycete. *Int. J. Syst. Evol. Microbiol.* 59, 1541–1544. doi: 10.1099/ijs.0.005959-0

Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19126–19131.doi: 10.1073/pnas.0906412106

Sasser, M. (1990). "Identification of bacteria through fatty acid analysis," in *Methods in Phytobacteriology*, eds Z. Klement, K. Rudolph and D. C. Sands (Budapest: Akademiai Kiado), 199–204.

Schwartz, M. A., Tabet, S. R., Collier, A. C., Wallis, C. K., Carlson, L. C., Nguyen, T. T., et al. (2002). Central venous catheter-related bacteremia due to *Tsukamurella* species in the immunocompromised host: a case series and review of the literature. *Clin. Infect. Dis.* 35:e72–e77.doi: 10.1099/ijs.0.005959-0

Seong, C. N., Kim, Y. S., Baik, K. S., Goodfellow, M., et al. (2003). *Tsukamurella sunchonensis* sp. nov., a bacterium associated with foam in activated sludge. *J. Microbiol.* 41, 83–88.

Shaer, A. J., and Gadegbeku, C. A. (2001). *Tsukamurella* peritonitis associated with continuous ambulatory peritoneal dialysis. *Clin. Nephrol.* 56, 241–246.

Stackebrandt, E., and Goebel, B. M. (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.* 44, 846–849.doi: 10.1099/00207713-44-4-846

Steinhaus, E. A. (1941). A study of the bacteria associated with thirty species of insects. *J. Bacteriol.* 42, 757–790.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729.doi: 10.1093/molbev/mst197

Teng, J. L., Huang, Y., Tse, H., Chen, J. H., Tang, Y., Lau, S. K., et al. (2014). Phylogenomic and MALDI-TOF MS analysis of *Streptococcus* sinensis HKU4T reveals a distinct phylogenetic clade in the genus *Streptococcus*. *Genome Biol. Evol.* 6, 2930–2943.doi: 10.1093/gbe/evu232

Teng, J. L., Tang, Y., Wong, S. S., Ngan, A. H., Huang, Y., Tsang, C. C., et al. (2015). *Tsukamurella hongkongensis* sp. nov. and *Tsukamurella sinensis* sp. nov. isolated from patients with keratitis, catheter-related bacteraemia and conjunctivitis in Hong Kong. *Int. J. Syst. Bacteriol.* 66, 391–397.doi: 10.1099/ijsem.0.000733

Tse, H., Tsoi, H. W., Leung, S. P., Lau, S. K., Woo, P. C., and Yuen, K. Y. (2010). Complete genome sequence of *Staphylococcus lugdunensis* strain HKU09-01. *J. Bacteriol.* 192, 1471–1472.doi: 10.1128/JB.01627-09

Verroken, A., Janssens, M., Berhin, C., Bogaerts, P., Huang, T. D., Wauters, G., et al. (2010). Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of *Nocardia* species. *J. Clin. Microbiol.* 48, 4015–4021.doi: 10.1128/JCM.01234-10

Wayne, L. G. (1988). International Committee on Systematic Bacteriology: announcement of the report of the ad hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *Zentralbl. Bakteriol. Mikrobiol. Hyg. A* 268, 433–434.doi: 10.1016/S0176-6724(88)80120-2

Wei, W., Gao, F., Du, M. Z., Hua, H. L., Wang, J., and Guo, F. B. (2016). Zisland Explorer: detect genomic islands by combining homogeneity and heterogeneity properties. *Brief. Bioinform*.doi: 10.1093/bib/bbw019 [Epub ahead of print].

Wei, W., and Guo, F. B. (2011). Prediction of genomic islands in seven human pathogens using the Z-Island method. *Genet. Mol. Res.* 10, 2307–2315.doi: 10.4238/2011.October.5.1

Woo, P. C., Lau, S. K., Teng, J. L., Tse, H., and Yuen, K. Y. (2008). Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Eur. J. Clin. Microbiol. Infect. Dis.* 14, 908–934.doi: 10.1111/j.1469-0691.2008.02070.x

Woo, P. C., Lau, S. K., Tse, H., Teng, J. L., Curreem, S. O., Tsang, A. K., et al. (2009). The complete genome and proteome of *Laribacter hongkongensis* reveal potential mechanisms for adaptations to different temperatures and habitats. *PLoS. Genet.* 5:e1000416.doi: 10.1371/journal.pgen.1000416

Woo, P. C., Ngan, A. H., Lau, S. K., and Yuen, K. Y. (2003). *Tsukamurella* conjunctivitis: a novel clinical syndrome. *J. Clin. Microbiol.* 41, 3368–3371.doi: 10.1128/JCM.41.7.3368-3371.2003

Woo, P. C., Wu, A. K., Tsang, C. C., Leung, K. W., Ngan, A. H., Curreem, S. O., et al. (2014). *Streptobacillus hongkongensis* sp. nov., isolated from patients with quinsy and septic arthritis, and emended descriptions of the genus *Streptobacillus* and *Streptobacillus moniliformis*. *Int. J. Syst. Bacteriol.* 64, 3034–3039.doi: 10.1099/ijs.0.061242-0

Yassin, A. F., Rainey, F. A., Brzezinka, H., Burghardt, J., Rifai, M., Seifert, P., et al. (1996). *Tsukamurella pulmonis* sp. nov. *Int. J. Syst. Bacteriol.* 46, 429–436. doi: 10.1099/00207713-46-2-429

Yassin, A. F., Rainey, F. A., Burghardt, J., Brzezinka, H., Schmitt, S., Seifert, P., et al. (1997). *Tsukamurella tyrosinosolvens* sp. nov. *Int. J. Syst. Bacteriol.* 47, 607–614. doi: 10.1099/00207713-47-3-607

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.