Genome Features and Secondary Metabolites Biosynthetic Potential of the Class Ktedonobacteria

Yu Zheng¹, Ayana Saitou², Chiung-Mei Wang¹, Atsushi Toyoda³, Yohei Minakuchi³, Yuji Sekiguchi⁴, Kenji Ueda⁵, Hideaki Takano⁶, Yasuteru Sakai¹, Keietsu Abe¹, Akira Yokota¹ and Shuhei Yabe¹,6*  

¹ Graduate School of Agricultural Science, Tohoku University, Sendai, Japan, ² Faculty of Agriculture, Tohoku University, Sendai, Japan, ³ Comparative Genomics Laboratory, National Institute of Genetics, Mishima, Japan, ⁴ Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan, ⁵ Life Science Research Center, College of Bioresource Sciences, Nihon University, Fujisawa, Japan, ⁶ Hazaka Plant Research Center, Kennan Eisei Kogyo Co., Ltd., Miyagi, Japan

The prevalence of antibiotic resistance and the decrease in novel antibiotic discovery in recent years necessitates the identification of potentially novel microbial resources to produce natural products. Ktedonobacteria, a class of deeply branched bacterial lineage in the ancient phylum Chloroflexi, are ubiquitous in terrestrial environments and characterized by their large genome size and complex life cycle. These characteristics indicate Ktedonobacteria as a potential active producer of bioactive compounds. In this study, we observed the existence of a putative “megaplasmid,” multiple copies of ribosomal RNA operons, and high ratio of hypothetical proteins with unknown functions in the class Ktedonobacteria. Furthermore, a total of 104 antiSMASH-predicted putative biosynthetic gene clusters (BGCs) for secondary metabolites with high novelty and diversity were identified in nine Ktedonobacteria genomes. Our investigation of domain composition and organization of the non-ribosomal peptide synthetase and polyketide synthase BGCs further supports the concept that class Ktedonobacteria may produce compounds structurally different from known natural products. Furthermore, screening of bioactive compounds from representative Ktedonobacteria strains resulted in the identification of broad antimicrobial activities against both Gram-positive and Gram-negative tested bacterial strains. Based on these findings, we propose the ancient, ubiquitous, and spore-forming Ktedonobacteria as a versatile and promising microbial resource for natural product discovery.

Keywords: class Ktedonobacteria, whole genome sequencing, phylogenetic analysis, genome comparisons, NRPS, PKS, antimicrobial screening

INTRODUCTION

The discovery and clinical application of antimicrobial drugs have greatly improved our well-being with regard to microbial infection over the past several decades (Demain and Sanchez, 2009). However, misuse and overuse of antibiotics has led to the accelerating development of antibiotic resistance in pathogens, which has become a major threat to modern health care and the natural environment (Pal et al., 2016; Jiang et al., 2017). Traditionally, Actinobacteria, particularly...
the genus *Streptomyces*, has served as the major producer of antimicrobial natural products, contributing 62% of the microbe-derived antibiotics from the 1950s to 1970s (Bérczy, 2012). In contrast, the ratio has decreased to 25% in the 2000s to 2010s, with the silencing of gene clusters encoding biosynthetic enzymes required for natural products under laboratory conditions further hampering the discovery of novel antibiotics via the traditional bacterial fermentation method (Rutledge and Challis, 2015). Given that, isolation and screening of rare actinomycetes and marine-derived actinomycetes have become popular. Furthermore, the application of *in silico* biosynthetic predictions from genome mining data and novel developed heterologous expression technologies confirm the phylum *Actinobacteria* as a continuing source of novel antibiotics (Bérczy, 2012; Genilloud, 2017). Nevertheless, it is of high risk to focus the discovery of novel natural products on only a single phylum as this source could become exhausted. Thus, the exploration of new microbial resources comparable to actinomycetes is extremely important.

Numerous bioactive natural products are derived from bacterial secondary metabolites, with the enzymes required for biosynthesis of these natural products being usually encoded on certain bacterial genome loci termed biosynthetic gene clusters (BGCs). Traditionally, BGCs include non-ribosomal peptide synthase (NRPS), polyketide synthase (PKS), and ribosomally synthesized and post-translationally modified peptide (RiPP) family clusters. The BGCs of NRPS and PKS encode modular enzymes that comprise functional synthesis blocks known as modules. Therefore, NRPSs and PKSs constitute a large class of structurally diverse and biomedically useful bacterial secondary metabolites and have accordingly gained considerable attention in recent years (Doroghazi and Metcalf, 2013; Bitók et al., 2017; Vila-Farres et al., 2017). Typical NRPSs contain three essential domains: an AMP-binding (A) domain responsible for the selection and activation of monomers, a peptidyl carrier protein (PCP) domain, also termed the thiolation (T) domain, required for attachment and transfer of the adenylylated monomer, and a condensation (C) domain for catalyzing amide bond formation and chain elongation (Challis and Naismith, 2004). PKS pathways share some similarities with fatty acid synthase (Dutta et al., 2014) and can be further divided into three types: type I PKSs that comprise large successive modular enzymes containing multi-functional domains; type II PKSs constituting single or bifunctional enzymes; and type III PKSs that utilize co-enzyme A-tethered substrates rather than the acyl carrier protein (ACP) domain to attach and transfer monomers such as type I and type II PKSs (Shen, 2003). Similar to NRPSs, a typical type I PKS also comprises three essential domains: an acyltransferase (AT) domain, which transfers a malonyl group from malonyl-CoA onto the 40-phosphopantetheine prosthetic group of an ACP domain, and a ketosynthase (KS) domain responsible for elongation of the polyketide chain (Masschelein et al., 2017). Moreover, similarities shared by NRPSs and PKSs in their biosynthesis pathway lead to hybrid non-ribosomal peptide-polyketides, which further enlarge the diversity of bacterial natural products (Miller et al., 2002; Liu et al., 2004; Shou et al., 2016). RiPPs also represent another type of antimicrobial peptide and are being considered as promising alternative antimicrobial agents owing to their widespread bactericidal sensitivity (Lázár et al., 2018). RiPPs are divided into different subfamilies according to their biosynthetic machinery and structural features, such as lantipeptides, lasso peptides, thiopeptides, and bacteriocins (Ortega and van der Donk, 2016). However, despite the differences in classification, almost all RiPPs contain an N-terminal leader peptide and a C-terminal core peptide (Ortega and van der Donk, 2016). Additionally, unlike NRPSs and PKSs, RiPPs always follow a simple biosynthesis pathway wherein the leader peptide guides the post-translational modification of the precursor peptide and is removed by one or more peptidases in the final step of core peptide maturation (Letzel et al., 2014; Ortega and van der Donk, 2016).

The class *Ktedonobacteria* was first reported in 2006 and currently contains two orders, three families, four genera, and seven formally proposed species including both mesophilic (Cavaletti et al., 2006; Yabe et al., 2017b) and thermophilic strains (Yabe et al., 2010, 2011). Owing to the morphological similarities, *Ktedonobacter racemifer* SOSP1-21T, the first proposed *Ktedonobacteria* species, was initially mistakenly identified as an actinomycete (Cavaletti et al., 2006). However, phylogenetic analysis based on 16S rRNA gene sequences placed the class *Ktedonobacteria* in the phylum *Chloroflexi* (Cavaletti et al., 2006; Yabe et al., 2010). Similar to actinomycetes, isolates and environmental DNA belonging to this class were detected to be ubiquitous in various terrestrial environments from common soil (Cavaletti et al., 2006; Yabe et al., 2017b), to extreme environments, such as an acid vapor-formed spring (Jiang et al., 2016), dark oligotrophic volcanic ice cave ecosystems of Mt. Erebus in Antarctica (Tebo et al., 2015), naturally occurring CO2 gas vents (de Miera et al., 2014b), recently deglaciated high-altitude soils of the Himalaya (Stres et al., 2013), and a lava cave in a volcanic trench (Northup et al., 2011). In addition, comparable to fungi, most *Streptomyces* species inhabit the soil environment and live as saprophytic bacteria (Flärhd and Buttner, 2009). In contrast, unlike actinomycetes, the class *Ktedonobacteria* appears to preferentially predominate in oligotrophic and extreme environments, suggesting these bacteria may possess different metabolic pathways from actinomycetes.

Notably, as shown in Figure 1, some isolated *Ktedonobacteria* strains exhibit complex morphologies. The genus *Thermosporothrix* forms substrate mycelia and subsequently produces grape-like smooth exospores by budding on the aerial mycelia (Yabe et al., 2010), whereas the surfaces of exospores formed by the genus *Thermogemmatispora* are wrinkled (Yabe et al., 2011). In contrast, strain *K. racemifer* SOSP1-21T forms spherical spores with a dried plum shape on aerial mycelia (Cavaletti et al., 2006), whereas *Dictyobacter aurantiacus* S27T produces globose or subglobose terminal sporangia from the vegetative mycelia via short stalk cells (Yabe et al., 2019). Spore formation represents one of the numerous methods developed by bacteria to survive in conditions of nutrient depletion or harsh environment (Errington, 2003; Flärhd and Buttner, 2009). It has been suggested that bacteria with complex morphological differentiation may possess an active secondary metabolism (Flärhd and Buttner, 2009; Manivasagan et al., 2014).
Previously, metabolites of new acyloins and thiazoles have been identified in the fermentation broth of Thermosporothrix hazakensis SK20-1T (Park et al., 2014, 2015), whereas the extremophilic Thermogemmatispora strain T81 produced a four (methyl) lanthionine-bridged new lanthipeptide that showed antimicrobial activity against the human pathogen Staphylococcus aureus (Xu et al., 2018). Furthermore, the released genome draft data revealed that the genome sizes of strain Thermogemmatispora onikobensis ONI-1T (Komaki et al., 2016) and Thermogemmatispora carboxidivorans PM5T were 5.56 and 5.61 Mb, respectively, whereas strain K. racemifer SOSP1-21T supported a large genome of 13.66 Mb (Chang et al., 2011), which is the second largest bacterial genome identified to date following that of Sorangium cellulosum So0157-2 (14.78 Mb) (Han et al., 2013).

Together, these described characteristics encouraged us to consider the class Ktedonobacteria as a Gram-positive, novel microbial resource for bioactive natural compounds discovery. However, the phylogenetic position and genome features of this class remained unclear. Moreover, although the large genome size of the Ktedonobacteria species suggest that they might encode multiple secondary metabolism gene clusters (Omura et al., 2001; Li and Walsh, 2010), few descriptions of these gene clusters or their novel natural product biosynthetic potential have been published. In this study, we therefore performed genome sequencing for another six Ktedonobacteria strains. The total nine studied Ktedonobacteria strains against a variety of bacterial strains was also assessed.

### MATERIALS AND METHODS

#### Strain Cultural Conditions

Origins of the strains utilized in this study are described in Section “Strain Information and Whole Genome Sequencing.” Ts. hazakensis SK20-1T and Ts. hazakensis COM3 were maintained in pure culture at 50°C on ISP3 agar plates (Shirling and Gottlieb, 1966), whereas the other seven strains were maintained on 1/10 R2A gellan gum (Kanto Chemical Co., Inc., Tokyo, Japan) plates (Yabe et al., 2017b). The optimal growth temperature for K. racemifer SOSP1-21T, D. aurantiacus S27T, Ktedonobacterales bacterium Uno3, and Ktedonobacterales bacterium Uno16 was 30°C, whereas Tg. onikobensis ONI-1T, Thermogemmatispora sp. A3-2, and Tg. carboxidivorans PM5T were cultured at 60°C. For genomic DNA extraction, strains Ts. hazakensis SK20-1T and Ts. hazakensis COM3 were cultured at 50°C in ISP1 liquid medium (Shirling and Gottlieb, 1966). Strain Thermogemmatispora sp. A3-2 was cultured at 65°C, whereas strains D. aurantiacus S27T, Ktedonobacterales bacterium Uno3, and Ktedonobacterales bacterium Uno16 were cultured at 30°C in 1/10 R2A liquid medium (Yabe et al., 2017b).

The candidate testing microorganisms used for antimicrobial assays included Bacillus subtilis NBRC 3134, Pseudomonas aeruginosa NBRC 13275, Escherichia coli NBRC 3972, Salmonella enterica NBRC 100797, Micrococcus luteus NBRC 13867, Geobacillus stearothermophilus NBRC 13737, Staphylococcus...
aureus NBRC 13276, S. aureus NTCT8325 (MSSA) and S. aureus Mu50 (VISA) (Kuroda et al., 2001) were provided by Juntendo University. All strains were maintained in nutrient agar plates.

Genomic DNA Extraction and Whole Genome Sequencing
To extract genomic DNA from the Ktedonobacteria strains, a modified method following the protocol of the Puregene Yeast/Bact. Kit B (Qiagen, Hilden, Germany) was used. The bacterial cells were collected by centrifugation of the culture broth and an appropriate volume of purified achromopeptidase (Fujifilm Wako Pure Chemical Corporation, Osaka, Japan) and lysozyme (MP Biomedicals, LLC, Illkirch, France) were added to the tubes to break down the cell walls. Also, 10 µL of 20 mg/mL proteinase K was added and incubated at 55°C for 1 h, then the tubes were incubated at 37°C overnight gently in a shaker to digest the residual proteins. Finally, the dried DNA pellets were dissolved in 100 µL DNA hydration solution for genome sequencing or stored at −20°C for later use.

Genome sequencing of strain Ts. hazakensis SK20-1T was conducted by TaKaRa Bio Inc., Japan on a PacBio RS II platform with long sequencing and assembled de novo using SMRT Analysis v2.2.0. The genome sequences of strain Ts. hazakensis COM3 and Thermogemmatispora sp. A3-2 were obtained from paired-end sequencing with HiSeq 2500 and long sequencing with PacBio RS II/Sequel. The genome sequencing was conducted by the National Institute of Genetics in Mishima, Shizuoka, Japan. The reads were assembled de novo using HGAP3/4, and gaps between contigs were closed with their fosmid clones. Genome data of strains K. racemifer SOSP1-21T, D. aurantiacus S27T, Ktedonobacterales bacterium Uno3, Ktedonobacterales bacterium Uno16, Tg. onikobensis ONI-1T, and Tg. carboxidivorans PM5T were recovered from NCBI and DDBJ databases for analysis. Additionally, the qualities (genome completeness and contamination) of the nine studied Ktedonobacteria genomes were estimated by CheckM, a set of tools with collocated sets of genes that are ubiquitous and single-copy within a phylogenetic lineage (Parks et al., 2015). As the class Ktedonobacteria is a relatively novel taxonomy of bacteria, both the lineage-specific workflow, taxonomic-specific workflow, and 83 single copy custom marker genes (Soo et al., 2014) were used in the CheckM estimation.

Phylogenetic Analysis Based on 16S rRNA Genes and Single-Copy Marker Genes
For phylogenetic analysis, Chloroflexi-related 16S rRNA gene sequences were obtained from the All-species Living Tree Project (LTP, release s132) (Yarza et al., 2010). The 16S rRNA gene sequence of T. hazakensis COM3 was extracted from its genome sequence (determined in this study) using CommunityM1 and aligned via the ARB software package (Ludwig et al., 2004)

1https://smrt-analysis.readthedocs.io/en/latest/
2https://www.pacb.com/training/hgап4-de-novo-assembly-application/
3https://github.com/dparks1134/CommunityM

using the LTP database as a reference; the alignment was manually corrected using the ARB EDIT tool of the ARB package. Sequences (> 1,300 nt) representing various family-level taxa in the phylum were selected in ARB and their alignments exported with lane mask filtering. Neighbor joining trees based on LogDet distance were computed using PAUP*4 (Swofford, 2003) with 100 bootstrap resamplings. Maximum likelihood trees were built using RAxML v8.2 (GTR and Gamma models + I) with rapid 100 times bootstrapping (Stamatakis, 2006). The trees were visualized and inspected in ARB.

Phylogenetic relationships among members of the phylum Chloroflexi, including members of the class Ktedonobacteria, were assessed using the genome sequences determined in the present study along with other finished and draft genomes representing the major families in the phylum as described in Sun et al. (2016). Briefly, 38 (Darling et al., 2014) or 83 (Soo et al., 2014) single-copy marker gene products were extracted using hidden Markov model searches as described previously (Sekiguchi et al., 2015). Tree topologies were tested for robustness using the maximum likelihood methods from FastTree v2 (with default parameters, JTT model, CAT approximation) (Price et al., 2010) and RAxML v8.2 (JTT and Gamma models with rapid 100 times bootstrapping).

Genome Annotation and Comparisons
The nine Ktedonobacteria genomes and related genomes were primarily annotated using DFAST (DDBJ Fast Annotation and Submission Tool) (Tanizawa et al., 2018) to study their general features. The DFAST-annotated 16S rRNA gene sequences were extracted and blasted using NCBI Nucleotide-Nucleotide Blast4 program to compare their similarities. Meanwhile, the Kyoto Encyclopedia of Genes and Genomes (KEGG) PATHWAY database was used for functional characterization of the nine genomes (Moriya et al., 2007).

CGView Server (Grant and Stothard, 2008) and the Geneious prime 2019.0.4 software were used to generate the genome plot. To compare the genomic rearrangement among Ktedonobacteria species, the contigs were concatenated using the Geneious software. Then the concatenated genomic sequences were aligned with the ProgressiveMauve algorithm and visualized using the MAUVE plugin within Geneious suite (Darling et al., 2004). BPGA (Bacterial Pan Genome Analysis tool), an ultra-fast pan-genome analysis pipeline (Chaudhari et al., 2016), was used to compare genomes between different strains in the class Ktedonobacteria and with other classes.

Metabolites Analysis and Identification of Putative BGCs
The nine Ktedonobacteria genomes were submitted to antiSMASH version 4.2.0 (antibiotics and Secondary Metabolites Analysis Shell) (Blin et al., 2017) and the integrated ClusterFinder algorithm, an hidden Markov model based probabilistic algorithm to detect BGC-like regions in genomes of unknown types (Cimermancic et al., 2014), to identify both the characterized and unknown secondary metabolite
biosynthesis gene clusters in the nine Ktedonobacteria strains. Additionally, both the KnownClusterBlast and ClusterBlast modules were selected to identify similar clusters in sequenced genomes by genome comparisons. Domain functions and genetic similarities with known BGCs in these gene clusters were further predicted and annotated using Protein-Protein BLAST and Pfam analyses (Finn et al., 2016). The AMP-binding domain amino acid substrate specificities were predicted with antiSMASH referencing three prediction algorithms, NRPSPredictor2 (Röttig et al., 2011), Stachelhaus code (Stachelhaus et al., 1999), and SANDPUMA ensemble (Chevrette et al., 2017). Only when two or all three algorithms provided consistent substrate specificity predictions was a consensus prediction called. Otherwise, the prediction results were assigned an “X” to represent an unknown amino acid residue, or all of the dubious amino acid residues were listed (Chu et al., 2016).

**Phylogenetic Analysis of the PKS KS Domain, NRPS C Domain, and Lantipeptide Modification Genes**

All of the PKS KS domains and NRPS C domains identified in putative BGCs from Ktedonobacteria genomes were extracted and submitted to NCBI Protein-Protein Blast (see text footnote 4). The top three reference sequences, excluding the Ktedonobacteria sequences, were selected and preserved in the FASTA file. Sequences of the Ktedonobacteria PKS KS domains, NRPS C domains, and the corresponding reference sequences were then submitted to Natural Product Domain Seeker (NaPDoS) (Ziemert et al., 2012) for functional classification. For phylogenetic tree construction of the KS and C domains, the sequences were trimmed into equal length by MEGA7.0a, and the tree was constructed following the method described in Section “Phylogenetic Analysis Based on 16S rRNA Genes and Single-Copy Marker Genes” for 16S rRNA phylogenetic analysis. Additionally, the phylogenetic trees were displayed and annotated by iTOL v3, an online tool for the display, annotation, and management of phylogenetic trees (Letunic and Bork, 2016).

The lantipeptide modification genes and precursor peptides were identified via a comprehensive analysis of antiSMASH 4.2.0 and BAGEL 3.0, a web-based comprehensive mining suite to identify and characterize RiPPs in microbial genomes (van Heel et al., 2013). Reference sequences of modification genes were obtained from the NCBI database with the same method as for KS domains and C domains. The phylogenetic trees were also constructed, displayed, and annotated following the aforementioned methods. For the analysis of precursor peptides, the extracted sequences were aligned with MEGA7.0 and the figures indicating cleavage sites were visualized using WebLogo 3 (Crooks et al., 2004).

**Antimicrobial Screening of the Class Ktedonobacteria**

Strains *Ts. hazakensis* COM3 and *Thermogemmatispora* sp. A3-2 were selected as representatives of the genus *Thermosporothrix* and *Thermogemmatispora*, respectively. Strain *Ts. hazakensis* COM3 was cultured in basic medium (peptone 2 g/L, yeast extract 1 g/L, MgSO\(_4\) 1 g/L, NaCl 1 g/L, and pH 7.0) at 50°C for 7 days, whereas strain *Thermogemmatispora* bacterium A3-2 was cultured in R2A medium at 60°C for 7 days. Strains *K. racemifer* SOSP1-21T, *Dictyobacter aurantiacus* S27T, *Ktedonobacterales* Uno3, and *Ktedonobacterales* Uno16 were cultured in Bennett’s medium (yeast extract 1 g/L, beef extract 1 g/L, N-Z amine type A 2 g/L, maltose 10 g/L, and pH 7.0) at 30°C for 14 days. Additionally, 2% Diaion® HP-20 polyaromatic adsorbent resin was added to each medium. Following 2 L culture, the cell pellets and HP-20 polyaromatic adsorbent resin were collected and extracted with the appropriate volume of acetone. Extractions from the above six strains were dried via evaporator and then redissolved in 5 mL of 80% methanol (MeOH) for antimicrobial activity screening versus a series of candidate testing microorganisms.

The antimicrobial activity screening of the class Ktedonobacteria was performed using the standard paper disc method. Paper discs (Toyobo Co., Ltd., Osaka, Japan; 6 mm) were saturated with 10 µL bacterial extracts from the above *Ktedonobacterales* strains and dried at room temperature. Then, the paper discs were placed on the surface of agar plates inoculated with each testing candidate microorganism and incubated for 24 h at the optimal temperature. Results of antimicrobial activity were determined by measuring diameters of zones of inhibition formed around each paper disc by the bacterial extracts. All antimicrobial tests were performed in triplicate.

**RESULTS**

**Strain Information and Whole Genome Sequencing**

As summarized in Table 1, the nine described *Ktedonobacteria* isolates were all terrestrially derived, including five thermophilic strains and four mesophilic strains. *Ts. hazakensis* COM3 was isolated from the same compost with *Ts. hazakensis* SK20-1T (Yabe et al., 2010) and was identified as a homologous strain to *Ts. hazakensis* SK20-1T by 16S rRNA gene sequence. However, *Ts. hazakensis* COM3 was distinguished from *Ts. hazakensis* SK20-1T through the formation of red pigment in culture medium (unpublished), and thus was also studied for genome analysis in this study. Strain *Thermogemmatispora* sp. A3-2 was isolated from the same geothermal soils as *Tg. onikobensis* ONI-1T and was identified as a novel species belonging to the genus *Thermogemmatispora*. *Thermogemmatispora* sp. A3-2 is preserved in our laboratory and in the Biological Resource Center, NITE (NBRC) and China General Microbiological Culture Collection (CGMCC). Moreover, *Thermogemmatispora* sp. A3-2 grew at 40–78°C, the highest growth temperature in the class *Ktedonobacteria* identified to date. *Ktedonobacterales* bacterium Uno3 and *Ktedonobacterales* bacterium Uno16 were isolated from soil-like microbe masses termed “Tengu-no-mugimeshi” at a volcanic region located in Gunma, Japan

---

https://www.megasoftware.net/
(Okada, 1937). Optimal growth temperature of the two isolates was 30°C.

Prior to this study, the draft genome sequences of K. racemifer SOSP1-21T (Chang et al., 2011), Tg. onikobensis ONI-1T (Komaki et al., 2016), and Tg. carboxidivorans PM5T have been publically released. In addition, genome sequencing of D. aurantiacus S27T (GenBank accession: BIFQ01000001 and BIFQ01000002), Ktedonobacterales bacterium Uno3 (GenBank accession: BIFR01000001 and BIFR01000002), and Ktedonobacterales bacterium Uno16 (GenBank accession: BIFT01000001, BIFT01000002, BIFT01000003, and BIFT01000004) was performed by our group in another study (Wang et al., 2019) and the genome data were deposited in IMG or DDBJ/EMBL/GenBank databases. In the present study, we conducted genome sequencing for three strains including Ts. hazakensis SK20-1T (GenBank accession: BIFX01000001, BIFX01000002, and BIFX01000003), Ts. hazakensis COM3 (GenBank accession: AP019376), and Thermogemmatispora sp. A3-2 (GenBank accession: AP019377). Sequencing platforms, sequence coverage, assembly methods, and accession numbers of the nine Ktedonobacteria genomes are summarized in Supplementary Table S1.

As shown in Figure 2, draft genome of Ktedonobacterales bacterium Uno3 possesses two circular contigs (5.30 Mb and 2.40 Mb, respectively) whereas the complete genome of Thermogemmatispora sp. A3-2 comprises a single circular chromosome. Ts. hazakensis COM3 possesses one putative linear chromosome, whereas D. aurantiacus S27T contains two linear contigs (6.13 Mb and 2.75 Mb, respectively). Ktedonobacterales bacterium Uno16 comprises two linear contigs (5.58 Mb and 3.14 Mb) and two circular plasmids of 199,150 and 43,380 bp in size, respectively. Initially, the second circular contig of Ktedonobacterales bacterium Uno3 (2.40 Mb) and the second linear contig of strains D. aurantiacus S27T (2.75 Mb) and Ktedonobacterales bacterium Uno16 (3.14 Mb) were thought to be part of the chromosomes because they were quite large in size compared with normal bacterial plasmids. However, the above “chromosomes” were determined to be incomplete due to the absence of most bacterial house-keeping genes in the subsequent CheckM assessment (Supplementary Table S2), suggesting that they may be “megaplasmids” (Thomas and Summers, 2008; Wu et al., 2009).

### General Genome Features

As given in Table 2 and Supplementary Table S1, genome sizes of the nine Ktedonobacteria strains ranged from 5.54 to 13.66 Mb, which were larger than the other classes in the phylum Chloroflexi. The average genome size of the genus Thermosporothrix and Thermogemmatispora was 7.48 and 5.57 Mb, respectively. Excluding strain K. racemifer SOSP1-21T, the average genome size of the three mesophilic strains including strains D. aurantiacus S27T, Ktedonobacterales bacterium Uno3, and Ktedonobacterales bacterium Uno16 was 8.51 Mb. However, the nine Ktedonobacteria strains exhibited a much lower GC content (an average of 55.1%) than that of actinomycete strains, especially the genus Streptomyces (Chater and Chandra, 2006). Additionally, the class Ktedonobacteria encoded multiple copies of ribosomal RNA operons (8~28 copies per genome) and 16S rRNA genes (3~9 copies per genome) on their genomes. The data of Tg. onikobensis ONI-1T was not counted due to the poor genome quality. In addition, a maximum of 1.49% 16S rRNA gene variation was observed in strain D. aurantiacus S27T (Supplementary Table S3). Moreover, 2364 (strain Tg. onikobensis ONI-1T) to 8227 (strain K. racemifer SOSP1-21T) of the detected coding sequences (CDSs) were termed as hypothetical proteins with unknown functions by DFAST annotation (Tanizawa et al., 2018), representing an average of 57.0% of the total CDSs in the nine Ktedonobacteria strains.

To investigate the absence/presence of genes responsible for autotrophic metabolites, the nine studied Ktedonobacteria

| Order | Strain | Isolation source | Growth temperature (optimal) | Morphology | References |
|-------|--------|-----------------|-------------------------------|------------|------------|
| Ktedonobacterales | Dictyobacter aurantiacus S27T | Waterlogged rice paddy field | 20~37°C (25~30°C) | Sporangia on aerial mycelia | Yabe et al., 2017b |
| Ktedonobacterales bacterium Uno3 | Soil-like microbe masses | 11~37°C (20~30°C) | Exospores on aerial mycelia | Wang et al., 2019 |
| Ktedonobacterales bacterium Uno16 | Soil-like microbe masses | 20~37°C (20~30°C) | Exospores on aerial mycelia | Wang et al., 2019 |
| Ktedonobacter racemifer SOSP1-21T | Soil | 17~40°C (28~33°C) | Exospores on aerial mycelia | Cavaletti et al., 2006 |
| Thermosporothrix hazakensis SK20-1T | Ripe compost for livestock treatment | 31~58°C (50°C) | Exospores on aerial mycelia | Yabe et al., 2010 |
| Thermosporothrix hazakensis COM3 | Ripe compost for livestock treatment | 31~58°C (50°C) | Exospores on aerial mycelia | This study |
| Thermogemmatisporales | Thermogemmatispora onikobensis ONI-1T | Geothermal soils | 50~74°C (60~65°C) | Exospores on aerial mycelia | Yabe et al., 2011 |
| Thermogemmatispora carboxidivorans PM5T | Geothermally heated biofilm | 40~65°C (65°C) | Exospores on aerial mycelia | King and King, 2014 |
| Thermogemmatispora sp. A3-2 | Geothermal soils | 40~78°C (60~65°C) | Exospores on aerial mycelia | This study |
FIGURE 2 | Continued
genomes were submitted to the KEGG PATHWAY database for functional characterization. Additionally, the presence of a type I carbon monoxide dehydrogenases (cox) gene in the genome of Tg. carboxidivorans PM5<sup>T</sup> and Thermogemmatispora sp. A3-2, April 2019 | Volume 10 | Article 893. Notably, this is the first report of the formation of a monophyly between Ktedonobacteria and Actinobacteria. The genome-based phylogenies determined via a comprehensive analysis of 16S rRNA gene sequences and universally conserved protein sequences. Initially, a phylogenetic association between the class Ktedonobacteria and the phylum Chloroflexi was suspected owing to the low bootstrap values in the 16S rRNA gene sequence phylogenetic tree and the dissimilarities in morphological, physiological, and chemotaxonomic data between the two (Cavaletti et al., 2006). However, our phylogenetic analyses based on both 16S rRNA gene sequences (Figure 3A) and 38 and 83 single-copy marker genes (Figure 3B and Supplementary Figure S1) clearly inferred that the members of the class Ktedonobacteria belong to the phylum Chloroflexi. The genome-based phylogenies determined in this study added evidence supporting the idea that the members of the class Ktedonobacteria evolved from a common descendent of the phylum Chloroflexi, although they largely shared some important phenotypic traits with those of members of the phylum Actinobacteria. Notably, this is the first report of the formation of a monophyly between Ktedonobacteria and
Dehalococcoidetes, another class in the phylum Chloroflexi and which is characterized with small genomes and obligate niche adaptation to reductive dehalogenation as the sole catabolic metabolism (Kaster et al., 2014).

Meanwhile, an obvious formation of a monophyly between strain Ktedonobacterales bacterium Uno16 and D. aurantiacus S27T was observed in both 16S rRNA and genomic phylogenetic trees (Figure 3). The 16S rRNA genetic similarity and average nucleotide identity (ANI) between the two strains were 95.60% and 74.26%, respectively, suggesting strain nucleotide identity (ANI) between the two strains were very dissimilar (Figure 4).

TABLE 3 | Comparisons of contig1 and contig2 in strains D. aurantiacus S27T, Ktedonobacterales bacterium Uno16, and Ktedonobacterales bacterium Uno3.

| Genes | D. aurantiacus S27T | Ktedonobacterales bacterium Uno16 | Ktedonobacterales bacterium Uno3 |
|-------|---------------------|----------------------------------|----------------------------------|
| contig1 | contig2 | contig1 | contig2 | contig1 | contig2 |
| Essential genes for growth | | | | | |
| Housekeeping genes*1 | 104 | 11 | 104 | 3 | 104 | 4 |
| Translation*2 | 77 | 5 | 77 | 3 | 74 | 3 |
| Replication and repair*2 | 66 | 8 | 66 | 9 | 67 | 3 |
| Citrate cycle (TCA cycle)*2 | 23 | 1 | 21 | 1 | 20 | 3 |
| Oxidative phosphorylation*2 | 42 | 4 | 44 | 8 | 41 | 3 |
| Adaptation to environment | | | | | |
| Transporter related genes*3 | 223 | 178 | 213 | 185 | 179 | 109 |
| Cytochrome P450 related genes*3 | 15 | 17 | 10 | 18 | 6 | 13 |
| Regulator genes*3 | 193 | 128 | 175 | 159 | 120 | 82 |

*1 The result of CheckM using dataset of domain bacteria. *2 The result of KEGG pathway functional characterization. *3 The result of DFAST annotation.

Dehalococcoidetes, another class in the phylum Chloroflexi and which is characterized with small genomes and obligate niche adaptation to reductive dehalogenation as the sole catabolic metabolism (Kaster et al., 2014).

Meanwhile, an obvious formation of a monophyly between strain Ktedonobacterales bacterium Uno16 and D. aurantiacus S27T was observed in both 16S rRNA and genomic phylogenetic trees (Figure 3). The 16S rRNA genetic similarity and average nucleotide identity (ANI) between the two strains were 95.60% and 74.26%, respectively, suggesting strain nucleotide identity (ANI) between the two strains were very dissimilar (Figure 4).

The pan-genome concept was firstly coined by Tettelin et al. (2005) in to catalog the entire genomic genes of a given phylgetic clade, including core genes (detected in all strains), accessory genes (detected in two or more strains), and unique genes (strain-specific genes), and was also used to estimate the genetic diversity of a studied group. In this study, we performed the pan-genome analysis to compare the nine strains within class Ktedonobacteria, and with the class Dehalococcoidetes, which formed a monophyly with the class Ktedonobacteria on the 38 and 83 single-copy marker genes based phylogenetic trees (Figure 3B and Supplementary Figure S1). As summarized in Supplementary Table SSA, 81.00~83.73% of the total genes were shared between strains Ts. hazakensis SK20-1T and Ts. hazakensis COM3, while the numbers were 77.50~77.80% in the genus Thermogemmatispora. By contrast, only 55.22~54.64% of genes were shared between D. aurantiacus S27T and Ktedonobacterales bacterium Uno16. Moreover, the percentages of core genes changed to 57.45~63.26% (Supplementary Table SSB) and 32.48~36.59% (Supplementary Table SSC), when contig1 (chromosome) and contig2 (“megaplasmid”) were aligned separately. Additionally, we found that Ktedonobacterales bacterium Uno3 shared the highest core gene ratio (38.84%) with the genus Dictyobacter among the class Ktedonobacteria. As for K. racemifer SOSP1-21T, the highest core gene ratio was found to be with the genus Thermosporothrix, at 21.06% of its total genes. The close relationships between K. racemifer SOSP1-21T with the genus Thermosporothrix, and Ktedonobacterales bacterium Uno3 with the genus Dictyobacter were in accordance with the result of the above phylogenetic analyses based on 83 single-copy marker genes (Figure 3B).

To compare the genomic rearrangement between D. aurantiacus S27T and Ktedonobacterales bacterium Uno16, which were proposed to comprise putative "megaplasmids," homologous regions in genomes were aligned between the two strains. As described above, strain Ktedonobacterales bacterium Uno16 was considered a novel species within the genus Dictyobacter by phylogenetic analyses based on both 16S rRNA gene sequences and universally conserved protein sequences. However, although genomic rearrangement in contig1 of strain Ktedonobacterales bacterium Uno16 was observed to be highly homologous to the contig1 of D. aurantiacus S27T, contig2 of the two strains were very dissimilar (Figure 4).
**A** phylogeny based on 16S rRNA gene sequence

- **Dicytbacter aurantiacus** S-27, LC219088
- Ktedonobacteraceae bacterium Uni18, LC278467
- Ktedonobacter sp. Uni19, LC278455
- Thermosporothrix hazakensis, COM4, this study
- Thermosporothrix rakukenkasa, FA, AB999923
- Thermogemmatispora sp. A3-2, LC235485
- Thermogemmatispora thermoaerophilica PM5, KP191206
- Thermogemmatispora orchiomoris, OR1-1, AB451912
- Chloroflexus aggregans, DSM 9488, CP013377
- Chloroflexus aggregans, He-2, KP193061
- Chloroflexus aurantiacus J-10-1, D83885
- Roseiflexus castenholzi, DSM 19841, CP000904
- Kahlottia papyrodurans, KJ63, JB48544
- Herpetosiphon aurantiacus, CP000875
- Herpetosiphon pyrrocorum, ATCC23276, AF030593
- Thermotogaceae phyl. phyla. KJ, KF193530
- Thermotoga phyl. phyla. WM50, EJ794997
- Thermoclostridium caudatum, K5, KF105829
- Thermomicrobium mauxi, ATCC 27502, M34115
- Sphaeroidobacter thermophilus, DSM 20764, AJ420142
- Leptotypea tertitarsis, VMK-2, AB109638
- Flexibacter submaris, MQ-QFX1, AB158677
- Aquicrocum halophilus, GOM-1, AB426072
- Leviceichromatium vallismortense, KMB-1, AB100439
- Moraxella gullothi, GMC-1, AB179431
- Rhodosphaera novella, RDO-1, AB138680
- Piscichromatium novella, RTM-1, AB138680
- Aquificales thermophiles, UN-1, AB264613
- Thermoproteales, SAW-1, AB665773
- Dehalogenimonas alleni gen. sp. IP-3, JQ994286
- Dehalogenimonas formidicivorans, NS-14, KP271189
- Dehalogenimonas lykomorphogenum, BL-DC-5, EU794919
- Dehalococcoides mccartyi, 165, CP000627
- Caldimicrobium aphelit, STL-6-01, AB108767
- Caldimicrobium aphelit, D2-15-10-12, HM148993
- K. aphelit, D2-15-10-12, HM148993
- Arthrobacter marina, AI5, AB781687
- Thermofexus hugenotii, JAD, KF265101

**B** based on 83 single-copy marker gene products

- Chloroflexus aurantiacus J-10-1, D83885
- Chloroflexus aggregans, DSM 9488, CP013377
- Chloroflexus aggregans, He-2, KP193061
- Redfield, Redfield phyl. phyla, DDBJ/EMBL/GenBank databases, 2016
- Redfield phyl. phyla
- Roseiflexus castenholzi, DSM 19841, CP000904
- Kahlottia papyrodurans, KJ63, JB48544
- Herpetosiphon aurantiacus, CP000875
- Herpetosiphon pyrrocorum, ATCC23276, AF030593
- Kahlottia papyrodurans, KJ63, JB48544
- Helicrobacter thermophilus, DSM 20764, AJ420142
- Leptotypea tertitarsis, VMK-2, AB109638
- Flexibacter submaris, MQ-QFX1, AB158677
- Aquicrocum halophilus, GOM-1, AB426072
- Leviceichromatium vallismortense, KMB-1, AB100439
- Moraxella gullothi, GMC-1, AB179431
- Rhodosphaera novella, RDO-1, AB138680
- Piscichromatium novella, RTM-1, AB138680
- Aquificales thermophiles, UN-1, AB264613
- Thermoproteales, SAW-1, AB665773
- Dehalogenimonas alleni gen. sp. IP-3, JQ994286
- Dehalogenimonas formidicivorans, NS-14, KP271189
- Dehalogenimonas lykomorphogenum, BL-DC-5, EU794919
- Dehalococcoides mccartyi, 165, CP000627
- Caldimicrobium aphelit, STL-6-01, AB108767
- Caldimicrobium aphelit, D2-15-10-12, HM148993
- K. aphelit, D2-15-10-12, HM148993
- Arthrobacter marina, AI5, AB781687
- Thermofexus hugenotii, JAD, KF265101

FIGURE 3 | Phylogenetic positions of the members of the class Ktedonobacteria based on comparative analysis of 16S rRNA gene sequences (A) and conserved protein sequences (B). (A) Neighbor-joining phylogenetic tree of public data (accession numbers are shown at the end of the index of each node) and the 16S rRNA gene sequence of T. hazakensis COM3, which was extracted from the genome sequence obtained in this study. Sequences from the bacterial phylum Nitrospira were used to root the tree (not shown). Reproducible nodes are marked based on bootstrap values from neighbor-joining and maximum-likelihood inferences (closed circle, >95% for both inferences; open circle, >90%; closed square, >80%; open square, >70%). Nodes that lack symbols were not reproducible among trees. Bar: 5% estimated sequence divergence. (B) Maximum-likelihood phylogenetic inference of representative genomes of the class Chloroflexi. The tree was built using RAxML based on up to 83 universally conserved proteins. Reproducible associations based on bootstrap values (closed circle, >95% for both RAxML and FastTree inferences; open circle, >90%; closed square, >80%; open square, >70%) are indicated at interior nodes. Accession numbers of genomes (IMG or DDBJ/EMBL/GenBank databases) are shown at the end of the index of each taxon. Alignments of homologous proteins from genomes of members of the phylum Deinococci were used to root the tree (not shown). Bar: 10% estimated sequence divergence.
Overview of the Putative Secondary Metabolites Biosynthetic Gene Clusters

To evaluate the biosynthetic potential of the class *Ktedonobacteria*, antiSMASH version 4.2.0 plus the integrated ClusterFinder algorithm were used to predict both characterized and unknown functioned secondary metabolite BGCs in the nine genomes. As shown in Figure 5A and Supplementary Table S6A, a total of 593 putative characterized and unknown BGCs were predicted by ClusterFinder, with 37∼117 BGCs per genome and counting for 11.75∼34.49% of the genomic sequences. To gain a deeper understanding of the *Ktedonobacteria* secondary metabolites, we focused on the putative BGCs with characterized functions predicted by antiSMASH. Initially, 104 antiSMASH-identified putative BGCs were found in the nine *Ktedonobacteria* genomes, including 14 NRPS clusters, 17 PKS clusters, 18 hybrid PKS/NRPS clusters, 27 lantipeptide clusters, and 28 other BGCs (Figure 5B and Supplementary Table S6B). However, these BGCs were distributed very unevenly in the nine strains. Strain *Ts. hazakensis* SK20-1 and *Ts. hazakensis* COM3 encoded 20 and 18 BGCs, dedicating 15.28% and 12.64% of their total genome, respectively (Supplementary Table S6B). In contrast, an average of four BGCs were identified in the genome of *Thermogemmatispora* strains, constituting only 1.78∼2.48% of the total genome on average. Strains *D. aurantiacus* S27T, *Ktedonobacterales* bacterium Uno3 encoded 13, 12, and 13 antiSMASH-identified putative BGCs, dedicating 6.46%, 7.54%, and 12.89% of their total genomes, respectively. In accordance with the result of ClusterFinder, 35.27% of the genomic sequences on contig 2 of *Ktedonobacterales* bacterium Uno3 were predicted to encode antiSMASH-identified putative BGCs. Moreover, composition of the antiSMASH-identified putative BGC types were very diverse in the nine *Ktedonobacteria* strains. The main BGC types in *Ts. hazakensis* SK20-1T and *Ts. hazakensis* COM3 were NRPS/T1PKS hybrid BGCs and lantipeptide BGCs, while *D. aurantiacus* S27T, *Ktedonobacterales* bacterium Uno3, and *K. racemifer* SOSP1-21T were more abundant in NRPS and PKS BGCs. Unlike *D. aurantiacus* S27T, *Ktedonobacterales* bacterium Uno16 encoded more Lantipeptide and NRPS/T1PKS hybrid BGCs on its genome. Overall, it was observed that BGCs responsible for peptide compounds production including NRPS, lantipeptide, lasso peptide, and thiopeptide BGCs predominate in the nine *Ktedonobacteria* genomes, while PKS BGCs encoding the polyketide compounds were the second most abundant BGC types. Herein, we analyzed composition and organization of the *Ktedonobacteria* NRPS, PKS, NRPS/PKS hybrid, and lantipeptide BGCs in the next section. In addition, the 104 antiSMASH-identified BGCs exhibited very limited similarity with known clusters in the MIBiG database (Supplementary Table S6B).

**NRPS, PKS, NRPS/PKS Hybrid, and Lantipeptide BGCs**

A total of 14 putative antiSMASH-identified NRPS BGCs were identified albeit distributed unevenly in seven strains, ranging from 20.4 kb (cluster Toni_5 in strain *Ts. onikobensis* ONI-1T) to 154.4 kb (cluster Uno3_7 in strain *Ktedonobacterales* bacterium Uno3) in size. Domain composition and organization analysis (Supplementary Table S7A) indicated that some of the NRPS BGCs such as cluster Txc_10 and cluster Dar_12 could produce peptide products with a ring moiety, by the existence of multiple epimerization (E) domain and heterocyclization (HC) domain (Chen L.Y. et al., 2016; Bloudoff et al., 2017) in the clusters. In addition, the existence of fatty acyl-AMP ligase (FAAL) domain in cluster Uno3_5 suggests that the final product of this gene cluster may constitute a lipopeptide or its derivatives. The FAAL domain was first discovered in *Mycobacterium tuberculosis* and activates fatty acids such as acyl adenylates, subsequently catalyzing their transfer onto the ACPs of PKSs or non-ribosomal peptide synthetases to produce lipidic metabolites (Hayashi et al., 2011).

As for PKS BGCs, 6 T1PKS BGCs and 11 T2PKS BGCs were identified. The 6 T1PKS BGCs spanned from 34.2 kb (cluster...
Krac_15 in strain *K. racemifer* SOSP1-21T) to 46.6 kb (cluster Uno16_10 in strain *Ktedonobacterales* bacterium Uno16) in size whereas the 11 T2PKS BGCs spanned from 41.3 kb (cluster Uno16_1 in strain *Ktedonobacterales* bacterium Uno16) to 42.5 kb (cluster Krac_11 in strain *K. racemifer* SOSP1-21T). Domain composition and organization of the identified PKS clusters are summarized in Supplementary Tables S7B,C. Notably, cluster Uno16_10 contained two additional NRPS-A domains that were
predicted to function as a long chain fatty acid CoA ligase and an acyl CoA synthetase by Protein-Protein Blast, indicating that the final products of this cluster may comprise derivatives of lipids (Hisanaga et al., 2004; Soubene and Kuypers, 2008).

Eighteen putative antiSMASH-identified BGCs were classified to be NRPS-T1PKS hybrid clusters, which were distributed mainly in strains Ts. hazakensis SK20-1T, Ts. hazakensis COM3, and Ktedonobacteriales bacterium Uno16. The 18 putative hybrid NRPS-T1PKS clusters ranged in size from 54.4 kb (cluster Uno16_4 in strain Ktedonobacteriales bacterium Uno16) to 333.5 kb (cluster Uno3_11 in strain Ktedonobacteriales bacterium Uno3). Domain composition and organization and the AMP-binding domain amino acid substrate specificities of the hybrid clusters are summarized in Supplementary Tables S7D,E. Based on genetic similarity and domain composition and organization in these gene clusters, the 18 putative hybrid NRPS-T1PKS clusters could be divided into four groups. Moreover, the hybrid NRPSs and PKSs biosynthesis pathways obviously enlarge the diversity of bacterial natural products.

After a comprehensive analysis of antiSMASH and BAGEL3 (van Heel et al., 2013), the 27 lantipeptide BGCs were further classified into four classes according to their biosynthetic machinery (Zhang et al., 2012). Class I contains 11 lantipeptide BGCs that are synthesized by two separated modification genes, LanB and LanC. In addition, cluster Tsx_13 from strain Ts. hazakensis SK20-1T comprised two LanB genes and one LanC gene. However, the two LanB genes were short in size and both lacked a lantibiotic biosynthesis dehydratase C-terminal domain, according to Pfam analysis. As the C-terminal domain is necessary for the final glutamate-elimination step in the generation of lantipeptide (Ortega et al., 2015), we propose these two LanB genes may be inactive. Class II contains 11 lantipeptide BGCs that are synthesized by a single bifunctional enzyme termed LanM. Moreover, after conducting a ClustalW alignment using MEGA7.0 and generating sequence logos with WebLogo 3 (Crooks et al., 2004), we observed F(E/D)LD (Supplementary Figure S2A) and GG (Supplementary Figure S2B) cleavage sites in class I and class II, respectively.

Based on these analyses, the antiSMASH-identified Ktedonobacteria BGCs showed very limited homology with known BGCs. The different composition and organization of tailoring genes, transport-related genes, and regulatory genes further increased the diversity of these BGCs and indicated that these clusters may encode novel natural products with novel functions.

Phylogenetic Analysis of the PKS KS Domain, NRPS C Domain, and Lantipeptide Modification Genes

Considering that class Ktedonobacteria constitutes a relatively new bacterial taxa, to date only very limited knowledge is available regarding their secondary metabolites. The domain-specific phylogenetic analysis of Ktedonobacteria-originated secondary metabolite BGCs identified in the present study may provide a better understanding of the functional and evolutionary classification of their domains. Furthermore, as C and KS domains are responsible for peptide and polyketide chain elongation in NRPS and PKS biosynthesis, respectively, the two domains represent the best candidates for domain-specific phylogenetic analysis (Rausch et al., 2007; Jenke-Kodama and Dittmann, 2009). As shown in the outlying ring of Figure 6A, the most abundant functional type among the Ktedonobacteria KS domains was assigned to hybrid KS by NaPDoS classification, owing to the high occurrence of hybrid NRPS-T1PKS clusters identified in the Ktedonobacteria genomes. Modular KS, the second most abundant KS domain functional type in the class Ktedonobacteria, was also mainly derived from hybrid NRPS-T1PKS clusters. The modular PKS is responsible for the incorporation of one building block and contains at least three domains: KS, AT, and ACP (Jenke-Kodama et al., 2005). With regard to evolutionary classification, the Ktedonobacteria-originated KS domains and reference sequences appeared to form four clades. The first clade included a sequence from strain Ktedonobacteriales bacterium Uno3, a sequence from strain D. aurantiacus S27T, and two sequences derived from the phylum Cyanobacteria.

The other three clades were more diverse in representation, including sequences from the phyla Actinobacteria, Cyanobacteria, Firmicutes, Chloroflexi, and the order Myxobacteria. Moreover, part of the Ktedonobacteria KS domains clustered with Actinobacteria, Cyanobacteria, and Firmicutes reference sequences, indicating that they may be phylogenetically related to these phyla. However, the majority of the Ktedonobacteria-originated KS domains formed independent branches. Furthermore, the highest similarity between Ktedonobacteria KS domains and reference sequences was only 71%, as observed between cluster Uno3_11 and a Brevibacillus laterosporus reference sequence. According to the NaPDoS functional classification, the most abundant types of Ktedonobacteria-originated C domains are LCL and DCL. The LCL type C domain catalyzes formation of a peptide bond between two L-amino acids whereas the DCL type links an L-amino acid to a growing peptide ending with a D-amino acid in NRPS biosynthesis (Rausch et al., 2007). In agreement with our analysis of NRPS and hybrid clusters, the C domains, which are replaced by HC domains in strain SK20-1T and strain COM3, were classified as cyclization domains by NaPDoS, which catalyze both peptide bond formation and subsequent cyclization of cysteine, serine, or threonine residues. With regard to evolutionary classification, the Ktedonobacteria C domains showed more diversity compared with the KS domains. As shown in Figure 6B, only a small percentage of C domains were related to the reference sequences from Cyanobacteria, Actinobacteria, Firmicutes, Myxobacteria, Proteobacteria, and Acidobacteria. Rather, a large proportion of Ktedonobacteria C domains formed independent branches, indicating they are distinct from those that originated from other phyla in evolutionary taxonomy.

As modification genes are responsible for the maturation of lantipeptide, the phylogeny of these modification enzymes are closely correlated with the structure of final lantibiotic natural products (Willey and van der Donk, 2007; Zhang et al., 2012). In the present study, the phylogenetic analyses were also applied to Ktedonobacteria modification genes to
determine their evolutionary relationship with other bacterial phyla or classes. As confirmed in the phylogenetic tree (Supplementary Figure S3), the majority of the Ktedonobacteria LanB and LanC modification genes formed independent clusters although they showed limited similarities with reference sequences from the phylum Actinobacteria. Compared with LanB and LanC modification genes, greater diversity was observed among LanM genes (Supplementary Figure S3). The two LanM genes from cluster Dar_6 and the LanMM cluster in strain S27 were related to LanMs from the phylum Cyanobacteria. The other Ktedonobacteria LanMs showed similarities with Actinobacteria and Myxobacteria sequences, whereas the majority formed independent branches, indicating that they may modify structurally different products from the known lantibiotic natural products.

**Antimicrobial Activities of Ktedonobacteria**

Cell pellets and HP-20 polyaromatic adsorbent resin crude extract from six representative Ktedonobacteria strains were tested for antimicrobial activities against both Gram-positive and Gram-negative bacterial strains. All six Ktedonobacteria strains demonstrated activity against at least two of the nine tested bacterial strains (Table 4). In addition, all six Ktedonobacteria strains were positive against G. stearothermophilus NBRC 13737, S. aureus NBRC 13276, and S. aureus Mu50.

Overall, the three mesophilic Ktedonobacteria strains including D. aurantiacus S27T, Ktedonobacterales bacterium Uno3, and Ktedonobacterales bacterium Uno16, exhibited a broader antibacterial spectra against both Gram-positive and Gram-negative (P. aeruginosa NBRC 13275, E. coli NBRC 3972, and S. enterica NBRC 100797) bacterial strains. In comparison, strain Ts. hazakensis COM3 and Thermogemmatispora sp. A3-2, the two thermophilic strains, only inhibited the Gram-positive bacterial strains (M. luteus NBRC 13867, G. stearothermophilus NBRC 13737, and S. aureus Mu50). Moreover, the three mesophilic Ktedonobacteria strains also exhibited stronger inhibitory capacities against each Gram-positive bacterial strains than that shown by the two thermophilic strains, as evidenced by the formation of larger inhibition clear zones (Table 4 and Supplementary Figure S4).

**DISCUSSION**

**Phylogenetic and Evolutional Analysis**

Members of the phylum Chloroflexi, especially the class Chloroflexi, are generally known as filamentous anoxygenic phototrophs, also termed filamentous green non-sulfur bacteria (Björnsson et al., 2002). However, the phylum Chloroflexi also constitutes a deeply branched bacterial lineage encompassing classes of bacteria with diverse metabolic types from strictly

**TABLE 4 |** Antimicrobial activity of the representative Ktedonobacteria strains assessed using the paper disc method.

| Strain | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------|---|---|---|---|---|---|---|---|---|
| D. aurantiacus S27T | 11.2 | 10.0 | 15.0 | 9.0 | 18.5 | 18.5 | 13.0 | 16.0 | 21.0 |
| Ktedonobacterales bacterium Uno16 | 8.4 | 10.0 | 11.5 | 9.0 | 21.5 | 21.5 | 10.0 | 15.0 | 20.0 |
| Ktedonobacterales bacterium Uno3 | – | 9.0 | 11.5 | 9.5 | 22.0 | 22.0 | 8.4 | 13.5 | 16.0 |
| K. racemifer SOSP1-21T | – | – | – | – | – | – | 7.5 | – | 7.5 |
| Ts. hazakensis COM3 | 7.0 | – | – | – | 21.0 | 21.0 | 7.5 | 9.8 | 14.5 |
| Thermogemmatispora sp. A3-2 | – | – | – | 10.0 | 10.0 | – | – | 7.8 |

1. B. subtilis NBRC 3134; 2. P. aeruginosa NBRC 13275; 3. E. coli NBRC 3972; 4. S. enterica NBRC 100797; 5. M. luteus NBRC 13867; 6. G. stearothermophilus NBRC 13737; 7. S. aureus NBRC 13276; 8. S. aureus NTCT8325 (MSSA); 9. S. aureus Mu50 (VISA).
anaerobic chlorinated hydrocarbon reducers to filamentous aerobic heterotrophs and vastly distinct genome sizes from 1.38 to 13.67 Mb (Cavaletti et al., 2006; Moe et al., 2009; Chang et al., 2011; Kaster et al., 2014). Moreover, comparative genomics and phylogenetic analyses suggest that these different classes may represent related but distinct phyla under a Chloroflexi “superphylum” (Gupta et al., 2013). Notably, maximum likelihood phylogenomic analysis based on informative amino acid sequences confirmed the class Ktedonobacteria and the SAR202 clade of dark-ocean bacterioplankton as the two deepest-branching classes related to the common ancestor of the phylum Chloroflexi (Landry et al., 2017). As the phylum Chloroflexi constitutes one of the earliest bacterial branches in the terrestrial environment (Hug et al., 2016), the class Ktedonobacteria is therefore considered to be among the earliest diverging bacterial lineages.

In the phylogenetic tree based on conserved protein sequences, class Ktedonobacteria clustered together with Dehalococcoidetes, a class of a strictly anaerobic, slow growing, and highly niche-specialized bacteria in terrestrial aquifer environments that utilize organohalide respiration as their sole source of energy (Kaster et al., 2014). However, our pan-genome analysis between class Ktedonobacteria and class Dehalococcoidetes indicated the two bacterial classes do not share a high number of conserved core genes with each other. Moreover, the two classes are clearly distinct in morphologies, genome sizes, habitats, and metabolic styles. Nonetheless, the formation of a monophyly in the phylogenetic tree based on conserved protein sequences suggest they may have evolved from a common ancestor. Herein, we hypothesize that the two bacterial classes then evolved separately to adapt to different environments and the evolutionary results reflect on their genomes. The class Dehalococcoidetes evolved a sole energy conservation mode via organohalide respiration and developed a symbiotic metabolic style, depending on co-living microbial species to acquire electron donors and cofactors (Kaster et al., 2014). As a result, species in the class Dehalococcoidetes may gradually drop unnecessary genes in their genomes, resulting in small genome sizes. On the contrary, species in the class Ktedonobacteria are very diverse in habitats, ranging from common soil to oligotrophic environments. Thus, species in the class Ktedonobacteria may tend to obtain foreign genes or foreign secondary metabolites gene clusters to adapt to these environments, resulting in relatively large genomes. Moreover, the class Ktedonobacteria developed a spore-forming morphology to survive nutrient depletion or harsh environments (Errington, 2003; Flärdh and Buttner, 2009). Of course, the currently sequenced genomes of the phylum Chloroflexi are still very limited and further studies are needed to completely characterize the evolution of class Ktedonobacteria.

Genome Features of the Class Ktedonobacteria

As described above, species in the class Ktedonobacteria are featured with low GC content and relatively large genome sizes, among which genomes of strains D. aurantiacus S27T, Ktedonobacterales bacterium Uno16, Ktedonobacterales bacterium Uno3, and K. racemifer SOSPI-21T are comparable to that of actinomycetes strains (Chen W.H. et al., 2016). As for the studies of strains in genus Thermosporothrix and genus Thermogemmatispora, their genome sizes are also quite large among thermophilic bacteria given that growth temperature and genome size in bacteria are negatively correlated and thermophilic bacteria tend to have a small genome (Sabath et al., 2013). Also, it has been reported that the Firmicutes and Gammaproteobacteria have the largest average 16S rRNA gene copy numbers (6.788 ± 2.694 and 5.141 ± 2.411, respectively) and the maximum variation (5.458 ± 5.633 and 5.298 ± 7.052, respectively) (Ibáñez et al., 2019). Thus, the average copy numbers of 16S rRNA genes (6.375 per genome, excluding strain Tg. onikobensis ONI-1T) in Ktedonobacteria genomes is quite high in bacterial species, although their variations are lower than Firmicutes and Gammaproteobacteria. Moreover, the relatively high percentage of functional unknown hypothetical proteins indicated that the class Ktedonobacteria may possess as yet unknown metabolites and produce some novel natural products through novel mechanisms (Galperin, 2001).

It is interesting to find that strains D. aurantiacus S27T, Ktedonobacterales bacterium Uno16, and Ktedonobacterales bacterium Uno3 may possess “megaplasmids.” As described in Section “Strain Information and Whole Genome Sequencing,” the draft genome of Ktedonobacterales bacterium Uno3 includes two circular contigs (Figure 2). As the presence of housekeeping genes are essential to define a bacterial chromosome (Thomas and Summers, 2008), here, it is probably more appropriate to classify the second circular contig of Uno3 to be “megaplasmid” due to the absence of most housekeeping genes (Table 3 and Supplementary Table S2). The presence of megaplasmid was also observed in Thermomicrobium roseum DSM 5159, an aerobic CO oxidizing thermophile strain in the phylum Chloroflexi and composed one chromosome (2,006,217 bp) and one megaplasmid (919,596 bp) in which few standard housekeeping genes were found (Wu et al., 2009). As for the linear contig2 of strains D. aurantiacus S27T and Ktedonobacterales bacterium Uno16, the definition of “megaplasmid” could be further evidenced with the absence of most housekeeping genes, translation genes, DNA replication and repair genes, and genes involved in TCA cycle and oxidative phosphorylation (Table 3), which are essential genes for growth. Furthermore, the high dissimilarities in sequences observed between contig2 of the two strains (Figure 4) indicated they originated differently. Meanwhile, high numbers of transporter related genes, cytochrome P450 related genes, and regulator genes were identified in contig2 of the two strains (Table 3). Commonly, these genes in bacterial genomes could contribute to their adaptation or resistance to various environmental pressures such as antibiotics, heavy metals, extreme Antarctic environments, among others (Kelly and Kelly, 2013; Dziewit et al., 2015; Romaniuk et al., 2018). Thus, it could be that strains D. aurantiacus S27T and Ktedonobacterales bacterium Uno16 acquired the foreign “megaplasmids” naturally to adapt to various environments. Additionally, we found that housekeeping genes are also absent in eight contigs of SOSPI-21T
In this study, a total of 593 ClusterFinder-identified putative BGCs were predicted in nine *Ktedonobacteria* strains among which 104 putative BGCs were also predicted by antiSMASH. As the genome size of bacteria is considered to correlate with the number of gene clusters (Donadio et al., 2007), the high number of putative BGCs predicted could be a consequence of their large genomes. Moreover, we observed that the class *Ktedonobacteria* encode numerous NRPS, PKS, NRPS/PKS hybrid, and lantipeptide gene clusters, which may assist them in fighting against their competitors and predators in their niches (Challis and Naismith, 2004; Zhang et al., 2012; Johnston et al., 2015; Ortega and van der Donk, 2016; Naughton et al., 2017). This hypothesis could be supported by the observation that the genus *Thermogemmatispora*, which exhibited the lowest number of BGCs in the class *Ktedonobacteria*, predominates in geothermal niches in which the microbial community richness is markedly decreased (de Miera et al., 2014b; Jiang et al., 2016). Owing to a lack of competitors and predators (Yabe et al., 2017a), the genus *Thermogemmatispora* may tend to reduce the number of BGCs in their genomes as secondary metabolites are not necessary but are always costly to produce. Furthermore, the existence of a homologous CAL domain in cluster Uno16_7 and the puwainaphycins BGC resulted in a 20% genetic similarity between the two clusters.

In addition, the difference of transport-related genes and regulatory genes with known BGCs further increases the diversity of known *Ktedonobacteria* BGCs. Notably, the BGC similarities revealed by antiSMASH comparison and the results of *Ktedonobacteria* ketosynthase and condensation domain phylogenetic analysis (Figure 6) suggest that a portion of the secondary metabolite BGCs in the class *Ktedonobacteria* may have been acquired from other bacterial phyla via horizontal gene transfer (HGT). HGT is common in bacteria and may provide new functions and subsequent niche adaptations through the acquisition of valuable secondary metabolite BGCs (Tooming-Klunderud et al., 2013; Ziemert et al., 2014; Jensen, 2016; Adamek et al., 2018). BGCs encoding natural products usually include both core biosynthetic and tailoring enzymes, along with regulatory genes for biosynthesis, resistance genes to the natural products they produce, and transport genes such as efflux pumps to export these natural products from their cells to the extracellular environment (Jensen, 2016). In addition, the regulatory genes required for function may facilitate integration of the HGT-acquired genes into existing genomes (Lawrence, 1999). Thus, the existence of regulatory, resistance, and transport genes in NRPS, PKS, and their hybrid gene clusters might facilitate their HGT (Ziemert et al., 2014). However, as the class *Ktedonobacteria* is among the earliest diverging bacterial lineages, arising much earlier than the phyla *Actinobacteria*, *Cyanobacteria*, *Firmicutes*, and *Proteobacteria* (Hug et al., 2016; Landry et al., 2017), we cannot exclude the possibility that the class *Ktedonobacteria* constitutes the actual HGT origin of secondary metabolite gene clusters.

The representative *Ktedonobacteria* strains produced probable bioactive compounds against both Gram-positive and
Gram-negative bacterial strains (Table 4 and Supplementary Figure S4), suggesting their potential utility for antimicrobial screening. Among the six screened Ktedonobacteria strains, D. aurantiacus S27T, Ktedonobacterales bacterium Uno3, and Uno16 were the most capable. Given that Gram-negative bacterial strains are becoming increasingly antibiotic resistant owing to their protective outer membranes and constitutively active efflux pumps, these mesophilic Ktedonobacteria strains may contribute to the development of novel antibiotics targeting Gram-negative bacteria (Miller, 2016; Domalaon et al., 2018). In turn, the antibacterial activities against S. aureus NBRC 13276, S. aureus NTCT8325, and S. aureus Mu50 might be explained by the production of novel lanthipeptides or other antimicrobial peptides by the Ktedonobacteria strains, as reported previously (Xu et al., 2018). Moreover, as the candidate testing microorganisms used for antimicrobial assays and activities assessed in the present study were finite and few isolated Ktedonobacteria species are available for study, this class likely harbors as-yet undisclosed bioactive functions.

**CONCLUSION**

In this study, we performed whole genome sequencing, comprehensive phylogenetic analysis, genome comparisons, and biosynthetic potential analysis of the class Ktedonobacteria, one of the deepest-branching classes related to the common ancestor of the earliest diverging bacterial lineages Chloroflexi. Our results further confirmed the classification of the spore-forming Ktedonobacteria to the phylum Chloroflexi at both the 16S rRNA and genomic level. The existence of a putative “megaplasmid” in some of the Ktedonobacteria strains was also observed in this study. Meanwhile, we found the class Ktedonobacteria were characterized with relatively large genome sizes, multiple copies of ribosomal RNA operons, and a high ratio of hypothetical proteins with unknown functions. The possibilities of reductive TCA cycle were also observed in the class Ktedonobacteria. Analysis of putative secondary metabolites BGCs in the nine Ktedonobacteria strains yielded a total of 593 ClusterFinder-identified putative BGCs (104 antiSMASH-identified putative BGCs), representing the first study to mine the biosynthetic potential of the class. The overall novelty and diversity of these BGCs provided convincing evidence that the class Ktedonobacteria possesses broad potential to produce new metabolite products with novel structures and mechanisms. Furthermore, bioactivity screening of the studied nine Ktedonobacteria strains revealed a wide spectrum of antibacterial activities. However, it should be noted that the current understanding of the Ktedonobacteria secondary metabolite pathways is still very limited, and their relationship with pathways of other bacterial phyla remains to be elucidated. Given the potential ability to produce novel and diverse bioactive natural products, we consider the ancient, ubiquitous, and mycelia-forming Ktedonobacteria to represent a versatile and promising microbial resource for pharmaceutical and biotechnological use.

**AUTHOR CONTRIBUTIONS**

YZ and SY conceived and designed the study. YZ, AS, and YM performed the experiments. YZ, AT, YuS, YaS, SY, C-MW, KU, HT, KA, and AY analyzed the results and data. YZ wrote the manuscript. All authors have read and approved the manuscript, and reviewed and confirmed the manuscript for publication.

**FUNDING**

This study was supported in part by MEXT/IJSPS KAKENHI (Grant Nos. 18K05406 and 16H06279), the Institute for Fermentation, Osaka (Grant No. G-2018-1-038), and the China Scholarship Council (CSC No. 201606330094).

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.00893/full#supplementary-material

**FIGURE S1** | Maximum likelihood phylogenetic inference of members of the class Ktedonobacteria among known bacterial phyla. The tree was constructed using RAxML or FastTree based on up to 38 marker genes (using taxon-outgroup configuration Config 3 described in Sekiguchi et al., 2015) and sequences were collapsed at the phylum level except for classes in the Proteobacteria. Ranks are indicated by prefix; p_ (phylum), c_ (class). Parameters used for phylogenetic inference using RAxML and FastTree are described in the “Materials and Methods” section. Numbers at the notes indicate bootstrap values (%) with 100 times determination. The scale bar represents 10% estimated sequence divergence.

**FIGURE S2** | Sequence logos of precursor peptide conserved motifs identified in the nine Ktedonobacteria strains are underlined in red. (A) The “FIE/DIL” cleavage site for Class I lantipeptide gene clusters. (B) The “GG” cleavage site for Class II lantipeptide clusters.

**FIGURE S3** | Phylogenetic analysis of the Ktedonobacteria modification genes in lantipeptide clusters. (A) Phylogenetic analysis of the Ktedonobacteria LanB genes. Amino acid sequences extracted from the Ktedonobacteria LanB and the three top hit reference sequences from Protein-Protein Blast were aligned by MEGA v. 7.0. Maximum-likelihood method was used to build the tree. The background colors represent evolutionary classification of the LanB modification genes. (B) Phylogenetic analysis of the Ktedonobacteria LanC genes. The tree was built as for (A). (C) Phylogenetic analysis of the Ktedonobacteria LanM genes. The tree was also built as for (A).

**FIGURE S4** | Antimicrobial screening of six representative Ktedonobacteria strains against bacterial and fungal strains. The antimicrobial activity is indicated by formation of clear zone of S. aureus NTCT8325 (MSSA) (A). S. aureus Mu50 (VISA) (B), Escherichia coli NBRC3972 (C), Bacillus subtilis NBRC3134 (D), Staphylococcus aureus NBRC13276 (E) Geobacillus steaeromorphus NBRC15737 (F), Pseudomonas aeruginosa NBRC13275 (G), Salmonella enterica NBC100797 (H), Micrococcus kutscheri NBRC139867 (I) as inhibited by Thermoparvophy indices COM3 (1), Ktedonobacter racemifer SOSP1-21T (2), Dicyobacter aurantiacus S27T (3), Ktedonobacterales bacterium Uno16 (4), Ktedonobacterales bacterium Uno16 (5), Thermogemmatispora sp. A3-2 (6), and negative control (N).

**TABLE S1** | Detailed sequencing information and genome features of the nine studied Ktedonobacteria strains and related taxon.

**TABLE S2** | CheckM estimation of the nine studied Ktedonobacteria strains.
TABLE S3 | Similarities of multiple 16S ribosomal RNA operons within the nine studied Ktedonobacteria strains.

TABLE S4 | cox gene (A) and iKEGG pathway (B) analysis of the Ktedonobacteria strains.

TABLE S5 | Pan-genome analysis of the Ktedonobacteria genomes (A), D. aurantiacus S27* (B), and Ktedonobacterales bacterium Uno16 contig1 (C).

REFERENCES

Adamek, M., Alanjary, M., Sales-Orellis, H., Goodfellow, M., Bull, A. T., Winkler, A., et al. (2018). Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in Amycolatopsis species. BMC Genomics 19:426. doi: 10.1186/s12864-018-4809-4

Arce-Rodríguez, A., Puente-Sánchez, F., Avendaño, R., Martínez-Cruz, M., de Moor, J. M., Pieper, D. H., et al. (2019). Thermoplasmatales and sulfur-oxidizing bacteria dominate the microbial community at the surface water of a CO2-rich hydrothermal spring located in Tenorio Volcano National Park, Costa Rica. Extremophiles 23, 17–187. doi: 10.1007/s00792-018-01072-6

Bédy, J. (2012). Thoughts and facts about antibiotics: where we are now and where we are heading. J. Antibiot. (Tokyo) 65, 385–395. doi: 10.1093/ja/2012.27

Bitok, J. K., Lemetre, C., Termei, M. A., and Brady, S. F. (2017). Identification of biosynthetic gene clusters from metagenomic libraries using PPTase complementation in a Streptomycetes host. FEMS Microbiol. Lett. 364:fxn155. doi: 10.1093/femsle/fxn155

Björnsson, L., Hugenholtz, P., Tyson, G. W., and Blackall, L. L. (2002). Filamentous Streptomyces reveals the biosynthetic potential of the marine-derived strain Streptomyces marokkonensis. Photosynth. Res. 72, 4360ñ4369. doi: 10.1023/A:10218271-148-8-2309

Blin, K., Wolf, T., Chevrette, M. G., Lu, X., Schwalen, C. J., Kautsar, S. A., et al. (2014). Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. Cell 158, 412–421. doi: 10.1016/j.cell.2014.08.034

Brooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. Genome Res. 14, 1189–1190. doi: 10.1109/GR.2004.1371316

Chang, Y. J., Land, M., Hauser, L., Chertkov, O., Del Rio, T. G., Nolan, M., et al. (2014). Structural and mutational analysis of the nonribosomal peptide synthetase intermodule organization in epimerization domain containing nonribosomal peptide biosynthesis. ACS Chem. Biol. 9, 1129–1137. doi: 10.1021/cb500550x

Chevrette, M. G., Aicheler, F., Kohlbacher, O., Currie, C. R., and Medema, M. H. (2017). SANDPUMA: ensemble predictions of nonribosomal peptide chemistry

TABLE S6 | List of the putative BGCs identified by ClusterFinder (A) and antiSMASH (B) in nine Ktedonobacteria strains and their similarities to known clusters.

TABLE S7 | Domain composition and organization of the putative NRPS clusters (A), TPKS clusters (B), T2PKS clusters (C), and NRPS-TPKS clusters (D) identified in nine Ktedonobacteria strains, and AMP-binding domain amino acid substrate specificities in the NRPS-TPKS clusters (E).

reveal biosynthetic diversity across Actinobacteria. Bioinformatics 33, 3202–3210. doi: 10.1093/bioinformatics/btx400

Chu, J., Vila-Farres, X., Inoyama, D., Termei, M., Cohen, L. J., Gordon, E. A., et al. (2016). Discovery of MRSA active antibiotics using primary sequence from the human microbiome. Nat. Chem. Biol. 12, 1004–1006. doi: 10.1038/nchembio.2207

Cimermancic, P., Medema, M. H., Claesen, J., Kurtia, K., Wieland, Brown, L. C., et al. (2014). Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. Cell 158, 412–421. doi: 10.1016/j.cell.2014.08.034

Domingo, S., Monciardini, P., Schumann, P., Rohde, M., and Sosio, M. (2007). Polypeptide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomes. Nat. Prod. Rep. 24, 1073–1109. doi: 10.1039/b504505c

Domalaa, R., Iwouw, T., Zhanel, G. G., and Schweizer, F. (2018). Antibiotic hybrids: the next generation of agents and adjuvants against gram-negative pathogens! Clin. Microbiol. Rev. 32:e00077-17. doi: 10.1128/CMR.00077-17

Dwivedi, K. R., Arroyo, P., de Luis Calabuig, E., and Ansola, G. (2014a). Effects of varying CO2 flows on bacterial communities in mesocosms created from two soils. Int. J. Greenhouse Gas Control 46, 205–214. doi: 10.1016/j.ijggc.2016.01.013

Dwivedi, K. R., Arroyo, P., de Luis Calabuig, E., Falagán, J., and Ansola, G. (2014b). High-throughput sequencing of 16S RNA genes of soil bacterial communities from a naturally occurring CO2 gas vent. Int. J. Greenhouse Gas Control 29, 176–184. doi: 10.1016/j.ijggc.2014.08.014

Dutta, S., Whicher, J. R., Hansen, D. A., Hales, W. A., Chemler, J. A., Congdon, G. R., et al. (2014). Structure of a modular polyketide synthase. Nature 510, 512–517. doi: 10.1038/nature13423

Dziewit, L., Pyszak, A., Szuplewska, M., Matlakowska, R., Mięlnicki, S., Wiberg, D., et al. (2015). Diversity and role of plasmids in adaptation of bacteria inhabiting the Lubin copper mine in Poland, an environment rich in heavy metals. Front. Microbiol. 3:152. doi: 10.3389/fmicb.2015.00152

Errington, J. (2003). Regulation of endospore formation in Bacillus subtilis. Nat. Rev. Microbiol. 1, 117–126. doi: 10.1038/nrmicro750

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mitchell, T. L., and Bateman, A. (2014). Pfam: the next generation of the protein families database. Nucleic Acids Res. 42, W222–W230. doi: 10.1093/nar/gku417

Flardh, K., and Buttner, M. J. (2009). Streptomyces morphogenetics: dissecting differentiated in a filamentous bacterium. Nat. Rev. Microbiol. 7, 36–49. doi: 10.1038/nrmicro1968

Galperin, M. Y. (2001). Conserved ‘hypothetical’ proteins: new hints and new puzzles. Comp. Funct. Genomics 2, 14–18. doi: 10.1002/cfg.66

Geniloud, O. (2017). Actinomycetes: still a source of novel antibiotics. Nat. Prod. Rep. 34, 1203–1232. doi: 10.1039/c7np00026j
description of Thermogemmatisporaceae fam. nov. and Thermogemmatisporales ord. nov. within the class Ktedonobacteria. Int. J. Syst. Evol. Microbiol. 61, 903–910. doi: 10.1099/ijs.0.024877-0

Yabe, S., Sakai, Y., Abe, K., and Yokota, A. (2017a). Diversity of Ktedonobacteria with actinomycetes-like morphology in terrestrial environments. Microbes Environ. 31, 61–70. doi: 10.1026/2092-421X.16144

Yabe, S., Sakai, Y., Abe, K., Yokota, A., Také, A., Matsumoto, A., et al. (2017b). Dictyobacter aurantiacus gen. nov., sp. nov., a member of the family Ktedonobacteraceae, isolated from soil, and emended description of the genus Thermosporothrix. Int. J. Syst. Evol. Microbiol. 67, 2615–2621. doi: 10.1099/ijsem.0.001985

Yabe, S., Wang, C. M., Zheng, Y., Sakai, Y., Abe, K., and Yokota, A. (2019). Formation of sporangiospores in Dictyobacter aurantiacus (class Ktedonobacteria in phylum Chloroflexi). J. Gen. Appl. Microbiol. 65 (in press).

Yarza, P., Ludwig, W., Euzéby, J., Amann, R., Schleifer, K. H., Glöckner, F. O., et al. (2010). Update of the all-species living tree project based on 16S and 23S rRNA sequence analyses. Syst. Appl. Microbiol. 33, 291–299. doi: 10.1016/j.syapm.2010.08.001

Zhang, Q., Yu, Y., Vélasquez, J. E., and van der Donk, W. A. (2012). Evolution of lanthipeptide synthetases. Proc. Natl. Acad. Sci. U.S.A. 109, 18361–18366. doi: 10.1073/pnas.1210393109

Ziemert, N., Lechner, A., Wetz, M., Millán-Aguíñaga, N., Chavarria, K. L., and Jensen, P. R. (2014). Diversity and evolution of secondary metabolism in the marine actinomycete genus Salinispora. Proc. Natl. Acad. Sci. U.S.A. 111, E1130–E1139. doi: 10.1073/pnas1324161111

Ziemert, N., Podell, S., Penn, K., Badger, J. H., Allen, E., and Jensen, P. R. (2012). The natural product domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. PLoS One 7:e34064. doi: 10.1371/journal.pone.0034064

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.

Copyright © 2019 Zheng, Saitou, Wang, Toyoda, Minakuchi, Sekiguchi, Ueda, Takano, Sakai, Abe, Yokota and Yabe. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.