Genome Insight and Description of Carotenoids-producing Croceicoccus Gelatinilyticus Sp. Nov., Isolated From the Tidal Flat Sediment

Tao Pei  
Guangdong Institute of Microbiology

Yang Liu  
Guangdong Institute of Microbiology

Juan Du  
Guangdong Institute of Microbiology

Kun peng Huang  
Guangdong Evergreen Feed Industry CO.,LTD

Ming rong Deng  
Guangdong Institute of Microbiology

Honghui Zhu (zhuhh@gdim.cn)  
Guangdong Institute of Microbiology  https://orcid.org/0000-0001-6854-8417

Keywords: Croceicoccus gelatinilyticus, polyphasic taxonomy, genotypes and phenotypes, genomic analysis

DOI: https://doi.org/10.21203/rs.3.rs-687094/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

A novel Gram-staining-negative and short-rod-shaped bacterial strain designated as 1NDH52<sup>T</sup> was isolated from tidal flat sediments and characterized by using a polyphasic taxonomic approach. The predominant cellular fatty acids of strain 1NDH52<sup>T</sup> were summed feature 8 (C<sub>18:1</sub>ω7c and/or C<sub>18:1</sub>ω6c) and C<sub>14:0</sub>2-OH; the major polar lipids were diphosphatidylglycerol, phosphatidylcholine, phosphatidylethanolamine, phosphatidylglycerol and sphingoglycolipid; the major respiratory quinones were Q-10 and Q-9. Phylogenetic analysis based on 16S rRNA gene sequences showed that strain 1NDH52<sup>T</sup> belonged to the genus Croceicoccus with high similarities to the close type strains Croceicoccus pelagius Ery9<sup>T</sup>, Croceicoccus sediminis S2-4-2<sup>T</sup> and Croceicoccus bisphenolivorans H4<sup>T</sup>. Phylogenetic analysis indicated that strain 1NDH52<sup>T</sup> formed an independent branch distinct from the known type strains of this genus. Digital DNA-DNA hybridization (dDDH) and average nucleotide identity (ANI) values between strain 1NDH52<sup>T</sup> and the three type strains above were well below thresholds of 70% DDH and 95-96% ANI for species definition, implying that strain 1NDH52<sup>T</sup> should represent a novel genospecies. The genomic DNA G+C content was 62.6%. The carotenoids production of the novel strain was determined by the detection of the pigment absorption spectrum and the identification of the complete biosynthetic gene cluster in its genome. Based on the phenotypic and genotypic characteristics, strain 1NDH52<sup>T</sup> is concluded to represent a novel species of the genus Croceicoccus, for which the name Croceicoccus gelatinitrophicus sp. nov., is proposed. The type strain of the species is 1NDH52<sup>T</sup> (= GDMCC 1.2381<sup>T</sup> = KCTC 82668<sup>T</sup>). The description of the genus Croceicoccus has also been emended.

Introduction

The genus Croceicoccus, a member of the family Erythrobacteraceae of the class Alphaproteobacteria, was proposed and initially contained only a species, Croceicoccus marinus (Xu et al. 2009). According to the latest List of Prokaryotic Names with Standing in Nomenclature (https://lpsn.dsmz.de/genus/croceicoccus) (Parte 2018), there are currently seven species with validly published names, including Croceicoccus marinus (Xu et al. 2009), Croceicoccus naphthovorans (Huang et al. 2015), Croceicoccus pelagius (Wu et al. 2016), Croceicoccus mobilis (Wu et al. 2016), Croceicoccus ponticola (Park et al. 2019), and Croceicoccus bisphenolivorans (Li et al. 2021). Taxonomically, the genus Croceicoccus is characterized as Gram-staining-negative, aerobic, yellow-pigmented, and coccoid or rod-shaped bacteria (Xu et al. 2009). The cellular predominant fatty acids of this genus bacteria were summed feature 8 (C<sub>18:1</sub>ω7c and/or C<sub>18:1</sub>ω6c) and the major respiratory quinone was ubiquinone 10 (Q-10). Croceicoccus have been frequently isolated from diverse marine environments, such as surface seawater from the Atlantic Ocean (Wu et al. 2016) and South Korea (Lee et al. 2018), the coastal sediment from Quanzhou Bay, Fujian (Huang et al. 2020) and the deep-sea sediment from the East Pacific polymetallic nodule region (Xu et al. 2009), the biofilm on a boat shell in the harbor of Zhoushan island, Zhejiang Province, P. R. China (Huang et al. 2015), showing their wide adaptability.

Materials And Methods

Isolation and culture of Croceicoccus strains

To investigate the culturable bacterial diversity of coastal areas, a sediment sample was obtained from the tidal flat (26°44'02.7″ N, 119°48'32.0″ E) in Ningde, Fujian province, P. R. China. Marine agar 2216 (MA; BD Difco) medium was used to isolate bacteria. The sample was serially diluted with sterilized water and spread on MA medium. After incubation at 28 °C for a week, colonies were picked out based on their characteristics and then streaked on the new MA medium. A bacterial strain designated as 1NDH52 was obtained from the Marine Culture Collection of China (MCCC). In this study, the morphological, physiological and chemotaxonomic tests of strain 1NDH52<sup>T</sup> and the three strains were performed MA medium or in marine broth 2216 (MB; BD Difco) under the same conditions unless otherwise specified.

Phenotypic and biochemical characteristics

The morphological characteristics of colonies and cells of strain 1NDH52<sup>T</sup> were examined by the stereomicroscope (SZX10, Olympus) and the transmission electron microscope (H7650, Hitachi) by using cultures on the MA medium for 48 h according to the instruments’ manuals, respectively. Motility was tested in MB broth supplemented with 0.5% agar. Growth of strain 1NDH52<sup>T</sup> and the three reference type strains was tested at 28 °C on the R2A agar, nutrient agar (NA), Luria-Bertani agar (LB) and tryptic soy agar (TSA) media for four weeks. Anaerobic growth of strain 1NDH52<sup>T</sup> was tested in an anaerobic pouch (MGC, Mitsubishi) on MA medium at 28 °C for four weeks. Gram staining was performed by using a Gram-stain kit (HKM) according to the manufacturer’s instructions. Catalase and oxidase activities were assessed through bubble production in 10% (v/v) aqueous hydrogen peroxide solution and oxidation of 1% N,N,N',N'-tetramethyl-1,4-phenylenediamine, respectively. Growth was evaluated at various temperatures (4, 10, 15, 20, 28, 32, 37, 40, 43, 45 and 50 °C) on MA medium for four weeks. The growth pH range and optimum pH were determined in the MB broth adjusted to pH 4.0-11.0 (at increments of 1 pH units) with citrate/phosphate (pH 4.0-7.0), Tris/HCl (pH 8.0-9.0) or sodium carbonate/sodium bicarbonate (pH 10.0-11.0) buffers. Tolerance to NaCl was performed in artificially modified MB broth supplemented with NaCl concentrations of 0, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10% (w/v). Substrate hydrolysis was tested by using the MB medium as the basal medium supplemented with 1% soluble starch, 1% skimmed milk, 1% carboxymethyl cellulose and 1% Tween 20, 40, 60 and 80, respectively. The hydrolysis of these substrates was determined according to the previously described methods (Liu et al. 2021a). Additional enzyme activities and biochemical properties were examined by using API ZYM and API 20NE strips following the manufacturer’s instructions.

Chemotaxonomic characterization
For the analysis of cellular fatty acids, cells of strain 1NDH52T and the three reference type strains were collected from the third quadrant on MA medium at 28 °C. The harvested cells were saponified, methylated and extracted according to the protocol of the Sherlock Microbial Identification System (MIDI). The prepared fatty acids were analyzed by gas chromatography (model 7890B; Agilent) using the Microbial Identification software package with the Sherlock MIDI 6.1 system and the Sherlock Aerobic Bacterial Database (TSBA 6.1). Polar lipids were extracted by using the chloroform/methanol system (Minnikin et al. 1984) and separated by two-dimensional thin-layer chromatography (Silica gel 60 F254, Merck). The plates dotted with the sample were subjected to two-dimensional development, with the first solvent of chloroform-methanol-water (65:25:4, v/v) followed by the second solvent of chloroform-methanol-acetic acid-water (85:12:15:4, v/v). The extracted lipids were identified by spraying with 10% ethanolic molybdophosphoric acid, ninhydrin, molybdenum blue, α-naphthol-sulphuric acid and Dragendorff’s reagent. The respiratory quinone was extracted and determined by using the high-performance liquid chromatography (Agilent 1200, ODS 250×4.6 mm×5 μm, flowing phase, methanol-isopropanol, 2:1; 1 mL-min−1) according to the method described by Collins et al. (1977).

Sequencing and phylogenetic analysis of the 16S rRNA gene

Genomic DNA of strain 1NDH52T was extracted from fresh cells by using the HiPure Bacterial DNA Kit (Magen Biotech Co., Ltd.) following the manufacturer’s instructions. The 16S rRNA gene was amplified using universal bacterial primers 27F and 1492R (Weisburg et al. 1991) with the PCR MasterMix (G-clone (Beijing) Biotech Co., Ltd.) in a 50 μL PCR system. PCR products were sequenced by Suzhou Geneviz Biotechnology Co., Ltd. (Suzhou, P. R. China). The sequences were assembled by using the software DNAMAN version 8 (Lynnsoft Biosoft, www.lynnsoft.com). The 16S rRNA gene sequences of related taxa were obtained from the EzBioCloud (www.ezbiocloud.net) (Yoon et al. 2017) and NCBI databases. Multiple sequence alignment was performed by using the software MAFFT version 7.037 under the L-INS-i interactive refinement (Katoh and Standley 2014). Pairwise identities of 16S rRNA gene sequences were calculated by using DNAMAN. Phylogenetic tree based on 16S rRNA gene sequences was reconstructed by using the software IQ-TREE version 2.1.2 (Minh et al. 2020) based on the maximum likelihood (ML) method (Felsenstein 1981) under the K2P + I + G4 nucleotide substitution model, which was selected by the ModelFinder (Kalyaanamoorthy et al. 2017). Support for the inferred ML tree was inferred by the ultrafast bootstrapping with 10,000 replicates (Hoang et al. 2018). The visualization and annotation of the resulting phylogenetic tree were performed by using the software MEGA version X (Kumar et al. 2018). The type strain Altericroceibacterium indicum MSSRF26T was used as an outgroup in all phylogenetic analyses.

Genome sequencing, overall genome relatedness indices and phylogenomic analysis

The genome sequence of strain 1NDH52T was determined by using the Illumina Novaseq PE150 platform of Shanghai Majorbio Bio-Pharm Technology Co., Ltd. About 1 Gbp of clean data for strain 1NDH52T was generated, reaching 200-fold coverage. The high-quality reads were assembled by using the software SPAdes version 3.8.1 with default parameters (Bankevich et al. 2012). The available genomes of the three close type strains and other related type strains within the family Erythrobacteraceae were obtained from the GenBank database. All genomes used in this study were assessed by using the software CheckM version 1.1.2 (Parks et al. 2015), which inspected the existence of gene markers specific to the class Sphingomonadales (UJD3310). Overall genome relatedness indices (OGRIs) including digital DNA-DNA hybridization (ddDH) and average nucleotide identity (ANI) were estimated by using the genome-to-genome distance calculator version 2.1 online service with the recommended formula 2 (Auch et al. 2010) and the software FastANI version 1.31 (Jain et al. 2018), respectively. Gene prediction and annotation of genomes of strain 1NDH52T and the close type strains were performed by using the software Prokka version 1.13 (Seemann 2014) and the fully automated Rapid Annotation using Subsystem Technology (RAST) pipeline (Overbeek et al. 2014) with default parameters. The functional roles of annotated genes for each genome were assigned and grouped in subsystem feature categories. The phylogenomic tree was reconstructed based on an up-to-date 92 bacterial core gene sets by the software UBCG version 3.0 (Na et al. 2018) with default parameters.

Determination of the pigment and the biosynthetic genes

The genus Croceicoccus is characterized by carotenoid biosynthesis (Xu et al. 2009). Therefore, the pigment absorption spectrum was performed according to the method described by Liu et al. (2021b). Cultures in the exponential phase of growth were scraped from the MA medium plates and then the cell pellet was mixed with 5 mL acetic acid/methanol (1:1, v/v) and incubated in the dark at 28 °C for an hour. The cell suspensions were centrifuged at 8000 g for 10 min and filtrated by a 0.22 μm organic phase filter after incubation in aceton/methanol. The absorption spectrum of the supernatant was monitored with a Lambda 45 UV/VIS spectrophotometer (Perkin Elmer). The genes related to the carotenoid biosynthesis among strain 1NDH52T and related type strains were determined and compared based on the genome annotation by the RAST.

Nucleotide sequences

The GenBank/EMBL/DDBJ accession numbers of 16S rRNA gene and genome sequences of strain 1NDH52T are MZ165339 and JAHCBO000000000, respectively.

Results And Discussion

Morphological and physiological characterization

Colonies of strain 1NDH52T were yellow-colored, circular, convex and smooth with a diameter of about 2 mm after cultivation on MA medium at 28 °C for 48 h (Fig. S1a). Cells of the strain were observed to be Gram-staining-negative and rod-shaped (0.5 μm wide and 1.0-1.2 μm long, Fig. S1b). Its activities of catalase and oxidase were positive. Anaerobic growth was not observed on MA medium. As listed in Table 1, the ranges of temperature, pH and NaCl for growth of strain 1NDH52T were 10-40 °C, 7.0-10.0 and 1-8% (w/v), and the optima of them were 28 °C, 7.0-8.0 and 2.0% (w/v) NaCl. Strain 1NDH52T could not grow on TSA medium, but the two type strains C. pelagius Ery9T and C. bisphenolivorans H4T could do. Strain 1NDH52T could be also distinguished from C. sediminis...
S2-4-2T, C. pelagius Ery9T and C. bisphenolivorans H4T by the hydrolysis of Tweens 40 and 60. Differential physiological and biochemical characteristics between strain 1NDH52T and the related reference strains were shown in Table 1. In the API ZYM and API 20NE tests, strain 1NDH52T could be distinguished from the three close relatives by the utilization of some carbohydrates, lipase (C14) and cystine arylamidase. Additional phenotypic characteristics that served to differentiate strain 1NDH52T from the three reference strains are shown in Table 1 and the species description.

**Chemotaxonomic characterization**

The fatty acid compositions of strain 1NDH52T and the three close type strains are listed in Table 2. The major fatty acids (>10% of total fatty acids) of strain 1NDH52T were summed feature 8 (C_{16:1}ω7c and/or C_{8:1}ω6c; 57.4%) and C_{14:0} 2-OH (12.3%), and the minor fatty acids (>5%) was C_{16:1} (5.7%). As showed in Table 2, the fatty acids profile of strain 1NDH52T was in agreement with those of the three type strains, but strain 1NDH52T was distinguishable from the three reference strains by the proportions of some fatty acids, such as iso-C_{15:0} and C_{16:1} 2-OH. The polar lipids of strain 1NDH52T were diphospha tidglycerol (DPG), phospha tidlychol (PC), phosphatidyethanolamine (PE), phosphatidylglycerol (PG), sphingoglycolipid (SGL), an unidentified glycolipid (GL) and two unidentified lipids (L1-2) (Fig. S2). As shown in Table 1, strains 1NDH52T exhibited similar polar lipid profiles to almost all type strains of this genus with the PC, PE, PG and SGL as major components. An obvious difference was that strains 1NDH52T and C. ponticola GM-16T contained the DPG, but the others did not. The major respiratory quinones of strain 1NDH52T were Q-10 (83.8%), Q-9 (13.9%), and the minor quinone was Q-8 (2.3%) (Fig. S3), which was different from other related type strains.

**Phylogenetic analysis based on 16S rRNA gene**

The almost complete 16S rRNA gene sequence of strain 1NDH52T (1347 bp) was determined by the Sanger sequencing. Its complete 16S rRNA gene sequence (1489 bp, locus_tag="KIH45_RS18855") from its genome sequence as shown below shared a 100% similarity with that by the PCR. In this case, the complete 16S rRNA gene sequence of strain 1NDH52T was used for the following analysis. The sequence comparison by using the BLAST and EzBioCloud search results showed that strain 1NDH52T belonged to the genus Croceicoccus with similarities to C. pelagius Ery9T (98.7%), C. sediminis S2-4-2T (98.9%) and C. bisphenolivorans H4T (99.2%). The ML phylogenetic tree based on 16S rRNA gene sequences showed that strain 1NDH52T clustered together with type strain C. sediminis S2-4-2T with 72% bootstrap value (Fig. 1), indicating the close relationship with each other. Furthermore, strain 1NDH52T formed a separate branch different from all type strains of recognized species of the genus Croceicoccus. The 16S rRNA gene sequence-based analyses indicated that strain 1NDH52T may represent a member of a putative novel species of the genus Croceicoccus.

**Comparative genomics and phylogenomic analysis**

The general genome characteristics of strain 1NDH52T and the related type strains are shown in Table 3. Based on the estimation from the CheckM, all genomes were determined to be of high quality based on their higher than 98% completeness and less than 2.5% contamination (Table 3) (Parks et al. 2015). The genome sizes of strains 1NDH52T and other type strain were from 3,001,363 to 4,116,752 bp. The genome size of strains 1NDH52T was 3,790,455 bp and only lower than strain C. mobilis Ery22T with 4,116,752 bp. The genomic sequence of strain 1NDH52T contained 3800 genes, 3 rRNA genes and 50 tRNA genes. The genomic DNA G + C content of strain 1NDH52T was 62.6% and under the range from 62.1 to 65.2% for the genus Croceicoccus (Table 3). Besides, the two 16S rRNA gene sequences of strain 1NDH52T respectively from the PCR and its genome by using a local BLAST sharing 100% identity confirmed the authenticity of the final genome assembly again.

In recent years, genomic information has increasingly been used for prokaryotic species definition and classification (Chun et al. 2018). The two OGRIs including dDDH and ANI were conducted in this study. The dDDH and ANI values between strain 1NDH52T and the three close type strains C. pelagius Ery9T, C. sediminis S2-4-2T, and C. bisphenolivorans H4T were 37.4, 30.2, 22.4%, and 89.7, 87.2, 83.2%, respectively, which were well below the recognized thresholds of 70% dDDH (Wayne et al. 1987) and 95-96% ANI (Richter and Rossello-Mora 2009) for bacterial species delineation, indicating that strain 1NDH52T should represent a novel genospecies. In the phylogenetic tree based on the 92 bacterial core gene sets, strains 1NDH52T, C. sediminis S2-4-2T, C. pelagius Ery9T and C. bisphenolivorans H4T clustered together; moreover, strain 1NDH52T formed an independent branch with the support of 100% bootstrap value (Fig. 2). Therefore, the comprehensive analyses of OGRIs and genome-based phylogeny indicated that strain 1NDH52T should represent a novel genospecies of the genus Croceicoccus.

Based on the genome annotation of RAST, the sub-system category distribution of 1NDH52T and the seven related type strains was compared. As shown in Table S1, strain 1NDH52T has 1321 annotated genes, accounting for 34.8% of the whole genome, most of which are necessary for bacterial survival, such as DNA, RNA, protein, amino acids and derivatives, fatty acids, lipids and isoprenoids, and carbohydrates metabolism. The similar distribution patterns of key metabolic-related genes were found in other type strains. Strain 1NDH52T and the seven type strains presented some unique features related to nitrogen metabolism, gene transfer agent, protein and nucleoprotein secretion system, cAMP signaling, iron acquisition, and metabolism. Among them, strain 1NDH52T contained genes encoding nitrate reductase, indicating that it could reduce nitrate, which was also confirmed by the nitrate reduction test in the API 20NE, but type strains C. sediminis S2-4-2T and C. ponticola GM-16T did not; strain 1NDH52T contained genes encoding type IV protein secretion systems, but type strain C. pelagius Ery9T did not. Instead, the reference type strains contained genes related to the ton and tol transport systems, but strain 1NDH52T did not.

**The determination of the pigment and related genes**

The result showed that the acetone-methanol-soluble pigment was characterized by the two maximum absorption peaks at 450 and 478 nm, suggesting that strain 1NDH52T could produce carotenoid (Figure S4a). As shown in Figure S4b and Table S2, all key genes including cytB, cytI, cytY, cytG, cytE and cytZ related...
Conclusion

Strain 1NDH52\textsuperscript{T} showed consistent chemotaxonomic traits characterized by the genus *Croceicoccus*, such as summed feature 8 (C\textsubscript{18:1}\textomega 7\textc and/or C\textsubscript{18:1}\textomega 6\textc) and C\textsubscript{14:0} 2-OH as the predominant fatty acids, DPG, PC, PE, PG and SGL as major polar lipids, Q-10 as the dominant respiratory quinone. Moreover, the microorganism can be distinguished from the three close type strains by using a combination of the phylogenetic analysis based on 16S rRNA gene and genome sequences and phenotypic properties, such as the gelatin hydrolysis, the utilization of some carbohydrates and no growth on TSA medium. Therefore, strain 1NDH52\textsuperscript{T} is considered to represent a novel species of the genus *Croceicoccus*, for which the name *Croceicoccus gelatinilyticus* sp. nov. is proposed. As a consequence of the newly reported characteristics not included in the original and emended descriptions of the genus *Croceicoccus* (Xu et al. 2009; Huang et al. 2015), the description of this genus has been emended.

Description of *Croceicoccus gelatinilyticus* sp. nov.

*Croceicoccus gelatinilyticus* (ge.la.ti.ni.ly’ti.cus N.L. n. *gelatium*, gelatin; Gr. masc. adj. [λυτικός] lytkos, able to dissolve; N.L. masc. adj. *lyticus*, dissolving; N.L. masc. adj. *gelatinilyticus*, gelatin-dissolving).

Colonies of strain 1NDH52\textsuperscript{T} were yellow, circular, convex and smooth with a diameter of 2 mm after cultivation on MA medium at 28 °C for 48 h. The activities of catalase and oxidase are positive. No anaerobic growth occurs on MA medium. Growth at 10-40 °C (optimum, 28 °C), at pH 7.0-10.0 (optimum, 7.0-8.0) and in 1-8% NaCl (optimum, 2%, w/v), and on R2A, LB, MA and NA media; no growth on TSA media. Do not hydrolyze the skimmed milk, Tweens 20 and 80, starch, casein and carboxymethyl cellulose. Hydrolysis of Tweens 40 and 60. In the API ZYM tests, strain 1NDH52\textsuperscript{T} shows positive results for trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, a-galactosidase, N-acetyl-β-glucosaminidase, a-mannosidase and a-fucosidase; negative for alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), leucine arylamidase, valine arylamidase, cystine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase and β-glucosidase. In the API 20NE tests, strain 1NDH52\textsuperscript{T} shows positive results for nitrate reduction, indole production, D-glucose fermentation, arginine dihydrolase, urease, β-galactosidase, utilization of D-glucose, D-mannose and phenylacetic acid; negative results for nitrite reduction, β-glucosidase, gelatin hydrolysis, utilization of L-arabinose, D-mannitol, N-acetyl-glucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid and trisodium citrate. The predominant cellular fatty acids are summed feature 8 (C\textsubscript{18:1}\textomega 7\textc and/or C\textsubscript{18:1}\textomega 6\textc) and C\textsubscript{14:0} 2-OH. The major polar lipids are diphosphatidylglycerol, phosphatidylcholine, phosphatidylethanolamine, phosphatidylglycerol, sphingoglycolipid, glycolipid, phospholipid and lipid. Besides the major respiratory quinones are ubiquinones 10 and 9 and the minor quinone is ubiquinones 8.

The type strain, 1NDH52\textsuperscript{T} (= GDGCC 1.2381\textsuperscript{T} = KCTC 82668\textsuperscript{T}), was isolated from the tidal flat sediment in Ningde city, Fujian province, P. R. China. The genomic DNA G + C content of the type strain is 62.6%.

Emended description of the genus *Croceicoccus* Xu et al. 2009 and Huang et al. 2015

In addition to the characteristics reported by Xu et al. (2009) and Huang et al. (2015), the following properties of the genus *Croceicoccus* are observed. The polar lipids contain diphosphatidylglycerol, phosphatidylcholine, phosphatidylethanolamine, phosphatidylglycerol, sphingoglycolipid, glycolipid, phospholipid and lipid. Besides the major respiratory ubiquinone 10, some species contain ubiquinones 9 and 8. The genomic DNA G + C contents range from 62.1 to 65.2%.

Declarations

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

We are very grateful for Mr. Fan Yang at Institute of Microbiology, Guangdong Academy of Sciences (Guangzhou, China) for purchasing type strains.

Funding information

This work was jointly supported by the Key-Area Research and Development Program of Guangdong Province (2018B020205003), the GDAS’ Project of Science and Technology Development (2020GDASYL-20200103028) and the Science and Technology Plan Project of Guangdong Province (2019B030316017).

References

Auch AF, von Jan M, Klenk H-P, Goekeler M (2010) Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Standards in Genomic Sciences 2:117-134. https://doi.org/10.4056/sigs.531120

Bankevich A et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455-477. https://doi.org/10.1089/cmb.2012.0021
Wayne LG et al. (1988) Report of the Ad Hoc committee on reconciliation of approaches to bacterial systematics. Int J Syst Bacteriol 268:433-434. https://doi.org/10.1016/s0176-6724(88)80120-2

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173:697-703. https://doi.org/10.1128/jb.173.2.697-703.1991

Wu YH et al. (2016) Croceicoccus pelagius sp. nov. and Croceicoccus mobilis sp. nov., isolated from marine environments. Int J Syst Evol Microbiol 66:4506-4511. https://doi.org/10.1099/ijsem.0.001381

Xu XW et al. (2009) Croceicoccus marinus gen. nov., sp. nov., a yellow-pigmented bacterium from deep-sea sediment, and emended description of the family Erythrobacteraceae. Int J Syst Evol Microbiol 59:2247-2253. https://doi.org/10.1099/ijs.0.004267-0

Yoon S-H et al. (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. International Journal of Systematic and Evolutionary Microbiology 67:1613-1617. https://doi.org/10.1099/ijsem.0.001755

Tables

Table 1 Differential characteristics between strain 1NDH52^T and the related type strains
| Characteristics                  | 1NDH52<sup>T</sup> | C. sediminis S2-4-2<sup>T</sup> | C. pelagius Ery<sup>T</sup> | C. bisphenolivorans H<sup>T</sup> | C. mobilis Ery<sup>T</sup> | C. ponticola GM-16<sup>T</sup> | C. marinus E4A<sup>T</sup> | C. naphthovorans PQ-2<sup>T</sup> |
|---------------------------------|---------------------|---------------------------------|-----------------------------|---------------------------------|-----------------------------|-----------------------------|-----------------------------|---------------------------------|
| Cell width (μm)                 | 0.5                 | 0.6                             | 0.3-0.5                     | 0.8-0.9                         | 0.5-0.8                     | 0.4-0.8                     | 0.8                         | 0.5-0.6                          |
| Cell length (μm)                | 1.0-1.2             | 1.2                             | 1.0-1.4                     | 1.6-1.8                         | 1.0-1.6                     | 0.4-1.0                     | 1                           | 0.8-1.2                         |
| Growth for                      |                     |                                 |                             |                                 |                             |                             |                             |                                 |
| temperature range (optimum, °C) | 10-40 (28)          | 15-40 (30)                      | 15-45 (30-37)               | 25-45 (32)                      | 15-45 (30-37)               | 20-35 (30)                  | 4-42 (25)                    | 15-50 (32)                      |
| pH range (optimum)              | 7.0-10.0 (7.0-8.0)  | 6.0-8.0 (7.0)                   | 5.5-8.5 (7.0)               | 5.0-9.5 (8.0)                   | 5.5-8.5 (7.0-7.5)           | 5.5-8 (7.0)                 | 6.0-9.0 (7.0)                | 6.5-9.5 (7.0)                   |
| NaCl range (optimum, w/v)       | 1-8 (2)             | 0-10 (1)                        | 0-10 (3)                    | 0-7.0 (0.5)                     | 0-7.5 (5)                   | 0-4 (1-2)                   | 0-10 (0-1)                  | 0.5-8 (2)                       |

API ZYM

| Enzyme                          | 1NDH52<sup>T</sup> | C. sediminis S2-4-2<sup>T</sup> | C. pelagius Ery<sup>T</sup> | C. bisphenolivorans H<sup>T</sup> | C. mobilis Ery<sup>T</sup> | C. ponticola GM-16<sup>T</sup> | C. marinus E4A<sup>T</sup> | C. naphthovorans PQ-2<sup>T</sup> |
|---------------------------------|---------------------|---------------------------------|-----------------------------|---------------------------------|-----------------------------|-----------------------------|-----------------------------|---------------------------------|
| Esterase (C4)                   | +                   | +                               | +                           | +                               | -                           | +                           | w                           | w                               |
| Lipase (C14)                    | +                   | +                               | +                           | +                               | -                           | -                           | +                           | -                               |
| Cystine arylamidase             | +                   | +                               | +                           | +                               | w                           | +                           | w                           | -                               |
| Trypsin                         | -                   | -                               | -                           | -                               | +                           | -                           | +                           | +                               |
| Acid phosphatase                | +                   | +                               | +                           | +                               | -                           | +                           | +                           | -                               |
| Naphthol-AS-BI-phosphohydrodase | +                   | +                               | +                           | +                               | +                           | -                           | -                           | +                               |
| β-Glucuronidase                 | -                   | +                               | -                           | -                               | -                           | +                           | +                           | -                               |
| α-Glucosidase                   | -                   | +                               | +                           | +                               | -                           | -                           | -                           | w                               |
| β-Glucosidase                   | +                   | +                               | -                           | -                               | -                           | +                           | +                           | -                               |

API 20NE

| Test                           | 1NDH52<sup>T</sup> | C. sediminis S2-4-2<sup>T</sup> | C. pelagius Ery<sup>T</sup> | C. bisphenolivorans H<sup>T</sup> | C. mobilis Ery<sup>T</sup> | C. ponticola GM-16<sup>T</sup> | C. marinus E4A<sup>T</sup> | C. naphthovorans PQ-2<sup>T</sup> |
|--------------------------------|---------------------|---------------------------------|-----------------------------|---------------------------------|-----------------------------|-----------------------------|-----------------------------|---------------------------------|
| Reduction of nitrate to nitrite | +                   | -                               | +                           | +                               | +                           | -                           | +                           | +                               |
| β-Glucosidase                   | +                   | +                               | +                           | +                               | -                           | +                           | -                           | +                               |
| Gelatin hydrolysis              | +                   | -                               | -                           | -                               | -                           | +                           | -                           | -                               |
| D-Glucose                       | -                   | +                               | +                           | -                               | +                           | -                           | +                           | -                               |
| L-Arabinose                     | +                   | -                               | -                           | -                               | -                           | +                           | +                           | -                               |
| D-Mannose                       | -                   | -                               | -                           | -                               | -                           | +                           | -                           | +                               |
| D-Mannitol                      | +                   | -                               | -                           | -                               | +                           | -                           | -                           | +                               |
| D-Maltose                       | +                   | +                               | +                           | -                               | +                           | -                           | +                           | -                               |
| Malic acid                      | +                   | +                               | +                           | -                               | -                           | +                           | -                           | -                               |
| N-Acetyl-Glucosamine, potassium glutonate, capric acid, adipic acid, trisodium citrate | + | - | - | - | - | - | - | - |

Polar lipids

| Polar lipids                  | DPG, PC, PE, PG, PL, SGL, GL, Ls | PG, PE, PG, PL, SGL, GL, L | PG, PE, PC, SGL, PL, L | PG, PE, PC, SGL, GL, L | PG, PE, PC, SGL, GL, L | DPG, PG, PC, SGL, GL, L | PG, PC, GL, PL | PG, PE, PC, SGL, GL, L |
|--------------------------------|---------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|---------------------------------|
| Respiratory quinone(s)         | Q-10, Q-9, Q-8                 | Q-10                        | Q-10                        | Q-10                        | Q-10                        | Q-10                        | Q-10                        | Q-10                            |

In the API ZYM tests, all strains showed positive results for alkaline phosphatase, esterase lipase (C8), leucine arylamidase and valine arylamidase; negative results for α-chymotrypsin, α-galactosidase, N-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase. In the API 20NE tests, all strains showed negative results for nitrite reduction, indole production, D-glucose fermentation, arginine dihydrolase, urease, β-galactosidase and the utilization of phenylacetic acid. +, positive; w, weakly positive; –, negative; *, data were inconsistent with the previous studies (Park et al. 2019; Huang et al. 2020; Li et al. 2021).

Table 2 Cellular fatty acid profiles of strain NDH52<sup>T</sup> and the three type strains
| Fatty acids | 1NDH52<sup>T</sup> | *C. sediminis* S2-4-2<sup>T</sup> | *C. pelagius* Ery9<sup>T</sup> | *C. bisphenolivorans* | H4<sup>T</sup> |
|------------|-------------------|-------------------------------|-----------------------------|---------------------|---------|
| C<sub>12:0</sub> | 0.5 | 0.7 | 0.7 | 1.2 | |
| C<sub>14:0</sub> | 1.8 | TR | TR | TR | |
| C<sub>15:0</sub> | 1.0 | TR | TR | ND | |
| C<sub>16:0</sub> | 5.7 | 4.1 | 4.6 | 3.5 | |
| C<sub>17:0</sub> | TR | 0.5 | 0.6 | TR | |
| C<sub>18:0</sub> | TR | 0.6 | 0.8 | 0.6 | |
| C<sub>14:0 2-OH</sub> | 12.3 | 10.5 | 14.6 | 7.5 | |
| C<sub>15:0 2-OH</sub> | 3.1 | 2.2 | 2.2 | 1.6 | |
| C<sub>16:0 2-OH</sub> | 2.9 | 3.2 | ND | 1.7 | |
| C<sub>16:1 2-OH</sub> | 0.7 | ND | ND | TR | |
| C<sub>16:1 ω6c</sub> | 0.8 | 2.6 | 1.8 | 1.2 | |
| C<sub>17:1 ω6c</sub> | 4.5 | 6.6 | 3.7 | 4.7 | |
| C<sub>17:1 ω8c</sub> | TR | 0.6 | TR | 1.0 | |
| C<sub>18:1 2-OH</sub> | 1.0 | ND | 2.7 | 1.3 | |
| C<sub>18:1 ω6c</sub> | 1.3 | 1.2 | 2.1 | 1.4 | |
| C<sub>18:1 ω7c 11-methyl</sub> | 4.6 | ND | ND | 5.6 | |
| Summed feature 3<sup>a</sup> | 3.0 | 9.2 | 5.6 | 8.2 | |
| **Summed feature 8<sup>a</sup>** | 54.7 | 54.5 | 55.7 | 57.2 | |

Data on the fatty acids of all strains were obtained from this study. Values were percentages of total fatty acids. Less than 0.5% and/or the absence of fatty acids for all strains were not shown. The predominant cellular fatty acids for all strains (> 10%) were in bold. TR, trace amount (< 0.5%); ND, not detected.

<sup>a</sup> Summed features are fatty acids that cannot be resolved reliably from another fatty acid using the chromatographic conditions chosen. The MIDI system groups these fatty acids together as one feature with a single percentage of the total. The summed feature 3, C<sub>16:1 ω6c</sub> and/or C<sub>16:1 ω7c</sub>; summed feature 8, C<sub>18:1 ω6c</sub> and/or C<sub>18:1 ω7c</sub>.

**Table 3** The general genome characteristics of strain 1NDH52<sup>T</sup> and the related type strains

**Figures**
| Genomic features | 1NDHS2<sup>T</sup> | C. sediminis S2-4-2<sup>T</sup> | C. pelagius Ery9<sup>T</sup> | C. bisphenolivorans H4<sup>T</sup> | C. mobilis Ery22<sup>T</sup> | C. ponticola GM-16<sup>T</sup> | C. marinus E4A9<sup>T</sup> |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| DNA size (bp)    | 3,790,455        | 3,548,496        | 3,306,530        | 3,602,538        | 4,116,752        | 3,244,470        | 3,001,363        |
| Contigs numbers  | 43               | 9                | 40               | 74               | 37               | 17               | 1               |
| DNA G + C content (%) | 62.6          | 63.0            | 62.8            | 62.8            | 62.6            | 62.1            | 65.2            |
| The longest contig (bp) | 757,386       | 1,712,580       | 586,724         | 344,509         | 940,960         | 1,264,975       | 3,001,363       |
| Contigs N50 (bp) | 376,040          | 574,689         | 327,690         | 262,101         | 346,704         | 411,713         | 3,001,363       |
| Number of genes  | 3,800            | 3,306           | 3,238           | 3,533           | 3,928           | 3,126           | 2,862           |
| Number of CDSs   | 3,746            | 3,255           | 3,174           | 3,481           | 3,877           | 3,072           | 2,805           |
| Number of rRNA genes | 3              | 3               | 9               | 3               | 3               | 3               | 6               |
| Number of tRNA genes | 50             | 47              | 54              | 46              | 47              | 50              | 50              |
| Completeness (%) | 99.3            | 99.5            | 99.3            | 99.6            | 99.4            | 98.7            | 100             |
| Contamination (%) | 2.3             | 1.0             | 0.5             | 1.0             | 1.7             | 0.1             | 0               |
| Genbank accession number | JAHCBF0000000000 | SUNC00000000 | LYWY0000000000 | LRSE0000000000 | BMIP0000000000 | RXOL0000000000 | CP019602        |
Figure 1

The ML tree based on 16S rRNA gene sequences showing the phylogenetic positions of strain 1NDH52T, the close type strains and related type strains within the family Erythrobacteraceae. Type strain Altericroceibacterium indicum MSSRF26T was used as an outgroup. Bootstrap values greater than 50% were shown at branch points. Bar: 0.01 represents the number of substitutions per site.
Figure 2

The phylogenomic tree based on bacterial core gene sets showing the phylogenetic positions of strain 1NDH52T the close type strains, and related type strains within the family Erythrobacteraceae. Type strain Altericroceibacterium indicum MSSRF26T was used as an outgroup. The nodes were labeled with bootstrap values. Bar, 0.05 represents the number of substitutions per site.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GDMCC1.23811NDH52.pdf
- KCTC826681NDH52.pdf
- Supplementarymaterials.pdf