Hydrogenophaga crocea sp. nov. associated with cyanobacterial mat isolated from farmland mud

Vipool Thorat1 · Bhavesh Tiwarekar1 · Kiran Kirdat1,2 · Saurabhee Chavan3 · Shital Markad3 · Shivaji Sathe2 · Tushar Lodha1 · Amit Yadav1

Received: 14 December 2021 / Revised: 18 March 2022 / Accepted: 20 March 2022 / Published online: 18 April 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
A catalase and oxidase-positive strain BA0156T was isolated from a cyanobacterial mat collected from the farmland mud cultivated with sugarcane from Ahmednagar, India. The 16S rRNA gene of strain BA0156T showed the highest percent sequence similarity with Hydrogenophaga borbori LMG 30805T (98.5%), followed by H. flava DSM 619T (98.3%) and H. intermedia DSM 5680T (98.2%). The strain BA0156T contained the major fatty acids, C16:0 (25.1%) and C17:0 cyclo (3.9%), whereas phosphatidylethanolamine and diphosphatidylglycerol were the major polar lipids. The OrthoANI and dDDH values between strain BA0156T and its closest relative H. borbori LMG 30805T were 84.6% and 28.3%, respectively. The DNA G+C content of strain BA0156T was 69.4 mol %. Furthermore, the biochemical and physiological features of strain BA0156T showed a distinct pattern from their closest phylogenetic neighbours. The phenotypic, genotypic and chemotaxonomic characteristics indicated that the strain BA0156T represents a new species for which the name Hydrogenophaga crocea (type strain BA0156T = MCC 3062T = KCTC 72452T = JCM 34507T) is proposed.

Keywords Hydrogenophaga · Hydrogenase · Hydrogen utilization

Introduction
The genus Hydrogenophaga (Phylum: Pseudomonadota, Class: Betaproteobacteria; Order: Burkholderiales, Family: Comamonadaceae) consists of 16 species with valid names and three with not valid names (Parte et al. 2020). Members of the Hydrogenophaga genus have been isolated...
from ecological niches, such as wastewater, hot springs, solvent degrading consortia, soil and others (Kämpfer et al. 2005; Yoon et al. 2008; Lin et al. 2017; Yang et al. 2017; Banerjee et al. 2021). This genus was proposed by Willems et al. (1989). Initially, members of Hydrogenophaga were described as chemolithoautotrophs, which utilises CO₂ as a carbon source and oxidizing H₂ as an energy source. This character was responsible for differentiating them from other genera of the family Comamonadaceae. However, chemolithoautotrophy was no longer the distinguishing feature separating Hydrogenophaga from other genera within the Comamonadaceae family after the description of H. atypica, H. intermedia and H. caeni were published (Contzen et al. 2000; Kämpfer et al. 2005; Chung et al. 2007). In this report, we describe new Hydrogenophaga species isolated from cyanobacterial mat collected from the mud of the field cultivated with sugarcane. The strain BA0156T was one of the bacterial isolates found during the process of obtaining an axenic culture of cyanobacterium related to genus Leptolyngbya.

Materials and methods

Bacterial isolation and culture conditions

Strain BA0156T was isolated from a cyanobacterial mat collected from a field cultivated with sugarcane at Ahmednagar district of Maharashtra state, India. The collected soil sample was initially processed to obtain cyanobacterial isolates using BG 11 (M1541; Himedia, India) media. During the purification of cyanobacterial culture, bacterial colonies, including strain BA0156T observed on a solid medium were picked and sub-cultured on Luria agar (LA) medium (M557; Himedia, India) at 30 °C for 48 h.

Morphological, physiological and biochemical characteristics

The light microscope (Model BX53; Olympus, USA) was used to determine the Gram character of strain BA0156T by using Gram staining kit (K001-KT; Himedia, India). The cell motility was analyzed by the hanging drop method under a light microscope (Model BX53; Olympus, USA) and motility test medium (M930S; Himedia, India). Oxidase and catalase activity were investigated using oxidase discs (DD018; Himedia, India) and bubble production in 3% (v/v) H2O2, respectively. The size and shape of the cells were determined by scanning electron microscopy (Zeiss, EVO 18, Version 6.02). Optimising pH, salinity and temperature (range 5 to 45 °C) was carried out in Luria broth (LB) for 5 days. The pH optimisation was determined at pH 5–12 unit with increments of 1 pH unit with different buffer systems, such as acetate (4–5 unit), phosphate (6 to 8 unit) and glycine–sodium hydroxide (9–12 unit). Salt tolerance for growth was examined at 0–4% NaCl with an increment of 0.5% at 30 °C in Luria broth (LB) for 5 days, and the growth was observed by a ‘Bioscreen C Microbiology’ reader. Biochemical properties such as utilization, assimilation of different carbon sources and enzyme activity against different substrates were examined by analytical profile index (API 20NE, API ZYM) tests and Biolog GEN III microplate assay according to the manufacturer’s instructions. H. borbori LMG 30805T and H. intermedia DSM 5680T were used as reference strains for polyphasic characterization.

Chemotaxonomic characterisation

The TSBA6 library on Sherlock Microbial Identification System (Version 6.1; MIDI) was used to determine total cell fatty acids of 48 h grown cultures, grown on tryptic soy agar at pH 7.3 ± 0.2 (M1968; Himedia) (Sasser 2001). To determine the polar lipids, cell mass was harvested from cultures in the logarithmic phase. Methanol/ chloroform/0.3% sodium chloride (2:1:0.8, by vol.) were used for extraction of polar lipids (Bligh and Dyer 1959) in addition to the modifications of Card (1973). Two-dimension chromatography was used for separation using chloroform–methanol–water (65:25:4 by vol.) in the first dimension and chloroform–acetic acid–methanol–water (40:7.5:6:2, by vol.) in the second dimension on silica gel TLC (Kieselgel 60 F254; Merck) (Minnikin et al. 1984). The dried plates were sprayed with 5% ethanolic phosphomolybdic acid to detect total lipids. These plates were further characterized by spraying ninhydrin (specifically for amino groups), molybdenum blue (specifically for phosphates), Dragendorff (for quaternary nitrogen) or α-naphthol (specifically for sugars) (Card 1973).  

Phylogenetic and genotypic analysis

The genomic DNA was extracted from the grown colonies (Ausubel et al. 1994). The 16S rRNA gene PCR was carried out with universal 16S rRNA gene primers (Baker et al. 2003). The EzBioCloud database (Yoon et al. 2017) was used to identify the closest phylogenetic relatives. The 16S rRNA gene phylogenetic trees viz. Neighbour Joining (NJ), Maximum Parsimony (MP) and Maximum Likelihood (ML) were generated using MEGA 7 (Kumar et al. 2016).

The genomic DNA of the strain BA0156T was sequenced on the Illumina MiSeq and Oxford Nanopore Technology (ONT) platform to obtain its whole genome sequence. The Illumina sequenced reads were analyzed for the quality assessment using FastQC v0.10.1 (Brown et al. 2017). The ONT sequencing data were base-called with quality filtering (> Q7) using Guppy v3.5.4. All QC passed Illumina reads (> Q30), and ONT reads were used to generate a hybrid
Table 1 Differential phenotypic characteristic between strain BA0156<sup>T</sup> and its closest phylogenetic neighbours

| Characteristics                              | BA0156<sup>T</sup> | LMG 30805<sup>T</sup> | DSM 5680<sup>T</sup> |
|----------------------------------------------|--------------------|------------------------|------------------------|
| Gram staining                                | Negative           | Negative               | Negative               |
| Spore formation                              | No                 | No                     | No                     |
| Motility                                     | Yes                | Yes                    | Yes                    |
| Salinity range for growth (%)                | 0.5–3              | 0.5–4                  | 0.5–4                  |
| PH range for growth                          | 6–11               | 6–8                    | 6–8                    |
| Maximum growth temperature (°C)              | 40                 | 40                     | 40                     |
| **API ZYM test**                              |                    |                        |                        |
| Alkaline phosphatase                         | +                   | −                      | +                      |
| Lipase                                       | +                   | −                      | +                      |
| Valine arylamidase                           | +                   | −                      | +                      |
| Cysteine arylamidase                         | +                   | −                      | +                      |
| Acid phosphatases                            | +                   | −                      | +                      |
| **API 20NE test**                             |                    |                        |                        |
| Assimilation of                              |                    |                        |                        |
| Gluconate                                    | +                   | +                      | −                      |
| Adipate                                      | +                   | −                      | −                      |
| Malate                                       | +                   | −                      | −                      |
| Citrate                                      | +                   | −                      | −                      |
| **Biolog GEN III**                            |                    |                        |                        |
| Quinic acid                                  | +                   | −                      | −                      |
| Dextrin                                      | −                   | −                      | +                      |
| d-turanose                                   | −                   | −                      | +                      |
| β-methyl-d-glucoside                         | −                   | −                      | +                      |
| N-acetyl-d-glucosamine                       | −                   | −                      | +                      |
| N-acetyl-d-mannosamin                        | −                   | −                      | +                      |
| d-fructose                                   | −                   | −                      | +                      |
| d-galactose                                  | −                   | −                      | +                      |
| l-rhamnose                                   | −                   | −                      | +                      |
| d-sorbitol                                   | −                   | −                      | +                      |
| d-mannitol                                   | −                   | −                      | +                      |
| d-arabinol                                   | −                   | −                      | +                      |
| Myo-inositol                                 | −                   | −                      | +                      |
| Glycyl-l-proline                             | −                   | −                      | +                      |
| l-alanine                                    | −                   | −                      | +                      |
| l-aspartic acid                              | −                   | −                      | +                      |
| α-keto-glutaric acid                         | −                   | −                      | +                      |
| l-malic acid                                 | −                   | −                      | +                      |
| Tween 40                                     | −                   | −                      | +                      |
| Y-amino-butyric acid                         | −                   | −                      | +                      |
| Formic acid                                  | −                   | −                      | +                      |
| l-galactonic acid lactone                    | −                   | −                      | +                      |
| 3-methyl glucose                             | −                   | +                      | +                      |
| d-fucose                                     | −                   | +                      | +                      |
| l-Fucose                                     | −                   | +                      | +                      |
| Glycerol                                    | −                   | +                      | +                      |
| Acetoacetic acid                             | −                   | +                      | +                      |
| Propionic acid                               | −                   | +                      | +                      |
| Acetic acid                                  | −                   | +                      | +                      |
| β-hydroxy-d, l-butyric acid                  | −                   | +                      | +                      |
| l-glutamic acid                              | −                   | +                      | +                      |
| d-gelacturonic acid                          | +                   | −                      | +                      |
assembly in Unicycler v0.4.8 (Wick et al. 2017). Assembly quality and completeness analysis were carried out using QUAST v5.0.2 (Gurevich et al. 2013) and CheckM v1.1.3 (Parks et al. 2015). The annotation was performed by NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Zhao et al. 2012).

The eggNOG-mapper v2 (Huerta-Cepas et al. 2019; Cantalapiedra et al. 2021) was used for functional annotation (Clusters of Orthologous Groups (COG)) of the PGAP annotated proteins. The genome was analyzed for the presence of genes for hydrogen oxidation and the methylerythritol phosphate pathway (MEP pathway). The genes for hydrogen oxidation were searched using HMMER 3.1b2 (Eddy 2011) by utilising the HMM profiles described by Gan et al. (2017). The representative sequences for a different group of hydrogenases were downloaded from HydDB (Søndergaard et al. 2016). The genomes of strain BA0156T and closely related type strains were blast searched against these sequences. The presence of genes for the MEP pathway was done using BlastKOALA (Kanehisa et al. 2016). The genome was also analyzed using antiSMASH 6.0.1 (Blin et al. 2021) to predict secondary metabolite biosynthesis gene clusters. Comparison of orthologous gene clusters among the genomes under study was made using OrthoVenn2 (Xu et al. 2019).

Furthermore, to support the taxonomic position of strain BA0156T obtained using the 16S rRNA gene phylogeny, a core-genome phylogenetic analysis was carried out using BPGA v1.3.0 (Chaudhari et al. 2016) and UBCG v3.0 (Na et al. 2018) tools. The BPGA tool searched for core genes from the genomes of 10 members phylogenetically close to strain BA0156T. The BPGA pipeline generated orthologous protein clusters using the integrated USEARCH algorithm (Edgar 2010). Among these, protein sequences from 20 random orthologous gene clusters were aligned and concatenated. Finally, the phylogenetic tree was reconstructed using the BPGA concatenated sequences by the neighbour-joining method in MEGA 7. The BPGA final data set included 22,960 amino acid positions used for the tree generation. The UBCG identified the bacterial core gene set from the strain BA0156T and other closely related genomes. These genes were concatenated, filtered for aligned positions and then processed for reconstructing the phylogenetic tree. The length of the UBCG concatenated alignment was 89,943 bp containing 92 marker genes for BA0156T identified using HMMER (Potter et al. 2018) and predicted using Prodigal search (Hyatt et al. 2010).

The orthoANI and digital DNA–DNA hybridisation (dDDH) values were calculated using the orthoANI calculator (Yoon et al. 2017) and Genome-to-Genome Distance Calculator (GGDC) webserver, respectively. The dDDH values were tested using recommended formula 2 (Richter and Rosselló-Móra 2009; Auch et al. 2010; Meier-Kolthoff et al. 2014).

### Results and discussion

#### Morphological, physiological and biochemical characteristics

Colonies of strain BA0156T were yellowish-brown with an entire margin and convex elevation. Colonies of strain BA0156T were motile, short rod in shape with 0.4–0.5 μM width and 0.7–1.3 μM in length (Fig. S1) determined using scanning electron microscopy. The strain BA0156T grew optimally at 30 °C and pH 7.0 with a range of 5–40 °C and 6.0–11.0 pH unit. It grew optimally at 0.5% salinity with a 0.5–3% salinity range. (Table 1).

According to API 20NE test, unlike its closest relatives, strain BA0156T showed the ability to assimilate adipate,
malate and citrate. The strain BA0156\textsuperscript{T} and strain LMG 30805\textsuperscript{T} were found assimilating gluconate. Strain DSM 5680\textsuperscript{T} was tested negative for all API 20NE substrates. Strain BA0156\textsuperscript{T} and strain DSM 5680\textsuperscript{T} were tested positive for alkaline phosphatase, lipase, cysteine arylamidase, valine arylamidase enzyme activities, whereas strain LMG 30805\textsuperscript{T} showed negative results for these enzyme activities. According to Biolog Gen III system results, strain BA0156\textsuperscript{T} utilised the substrate quinic acid but its closest relatives LMG 30805\textsuperscript{T} and DSM 5680\textsuperscript{T} did not. Strain BA0156\textsuperscript{T} was found negative for utilisation of 3-methyl glucose, D-fucose, L-fucose, glycerol, acetoacetic acid.
propionic acid, acetic acid, β-Hydroxy-D, L-butyric acid, L-glutamic acid, whereas both the comparative strains were positive for utilisation of these substrates. According to the BIOLOG Gen III system, strain BA0156 was found resistant to fusidic acid, and both the comparative strains were found sensitive to it. Strain BA0156 showed sensitivity towards chemicals, such as sodium butyrate and 4% NaCl, whereas strains LMG 30805T and DSM 5680T showed resistance to these chemicals. Other results related to biochemical tests are shown in Tables 1 and S1.

**Chemotaxonomic characterisation**

Major fatty acids found in strain BA0156 were C16:0 (25.1%) and C17:0 cyclo (3.9%) (Table S2). The major polar lipids of strain BA0156 were phosphatidylethanolamine and diphosphatidylglycerol, while phosphatidylethanolamine and unknown phospholipid PL4 were the major polar lipids of the strain LMG 30805T. Phosphatidylglycerol and unknown lipids such as L1, L2, L3 and L4 are minor polar lipids of the BA0156T strain, whereas diphosphatidylglycerol, phosphatidylglycerol, unknown phospholipid such as PL1 to PL4, unknown aminophospholipid such as APL1 and unknown lipids such as L1 and L2 were detected as the minor polar lipids of strain LMG 30805T. Phosphatidylethanolamine and phosphatidylglycerol were the major polar lipids of strain DSM 5680T. Diphosphatidylglycerol and the unknown lipid L1 and L4 were also detected in small amounts in strain DSM 5680T (Supplementary Figure S2).

**Phylogenetic and genotypic analysis**

The genome sequencing run of BA0156 on the Illumina MiSeq platform generated 5,273,152 reads (150 × 2 chemistry). The ONT run generated 61,220 reads. The final BA0156 assembly (GenBank accession number, CP049989) contained one scaffold of 4,561,131 bp, corresponding to 99.8% completeness and 382 × depth coverage. The DNA G + C content was 69.4 mol %. The BA0156 genome (CP049989) consists of 4273 protein-coding genes, 44 tRNA and two rRNA operons.

The 16S rRNA gene sequence of strain BA156 showed the highest similarity with H. borbori LMG 30805T (98.5%), followed by H. flava DSM 619T (98.3%) and H. intermedia DSM 5680T (98.2%). The strain BA0156 clustered in a monophyletic clade along with H. borbori LMG 30805T and H. intermedia DSM 5680T as inferred from 16S rRNA gene phylogeny (Fig. 1). This phylogenetic cladding was supported by the core-genome phylogenetic tree constructed using BPGA and UBCG with strong bootstrap values (Fig. 2). Furthermore, the orthoANI value for strain
BA0156<sup>T</sup> against *H. borbori* (LMG 30805<sup>T</sup>) was 84.6%, with *H. intermedia* (DSM 5680<sup>T</sup>) it was 84.1%, and with *H. flava* (DSM 619<sup>T</sup>) it was 79.1%. The digital DNA–DNA hybridization (dDDH) values for strain BA0156<sup>T</sup> against LMG 30805<sup>T</sup>, DSM 5680<sup>T</sup> and DSM 619<sup>T</sup> were 28.3, 27.5 and 21.8%, respectively. The orthoANI and dDDH values computed for the strain BA0156<sup>T</sup> strongly support its unique genomic structure among its closest relative species and a novel species (Table 2).

The genome sequence of BA0156<sup>T</sup> contained 351 (8.5%) genes for amino acid transport and metabolism, 339 (8.4%) genes for transcription and 337 (8.2%) genes for energy production and conversion as the abundant assigned categories (Fig. 3). Total 741 (18.0%) genes were assigned to putative functions. The distribution of gene components between BA0156<sup>T</sup> and its closely related type strain genomes describe the similar composition of COG functional categories. The OrthoVenn2 assigned the 4269 protein sequences from BA0156<sup>T</sup> in 3333 orthologous clusters with 873 singletons. The Venn diagram represents 2862 gene clusters shared by BA0156<sup>T</sup> with its closely related type strains (Supplementary Figure S3). The individual gene cluster denotes homologous genes from different organisms under study. Total, 26 gene clusters corresponding to 55 proteins were unique to BA0156<sup>T</sup>. These clusters include aliphatic sulfonate ABC transporter permease SsuC, phytoene desaturase, TerB family tellurite resistance protein, acetyl hydrolase, BCD family MFS transporter and hypothetical proteins.

Strains BA0156<sup>T</sup>, LMG 30805<sup>T</sup> and DSM 5680<sup>T</sup> which clade together in the DNA sequence-based phylogenetic analysis, showed the absence of genes for hydrogen oxidation (highlighted in bold font), however, other species of the genus showed the presence of the genes encoding proteins involved in hydrogen oxidation (Table 3). The BLAST search against the HydDB sequences showed the potential for hydrogen oxidation in a group of closely related Hydrogenophaga strains used to describe the genus (Gan et al. 2017).

Analysis using the antiSMASH 6.0.1 identified eight predicted secondary metabolite biosynthetic gene clusters (BGCs). One of the identified clusters was for terpene biosynthesis with a length of ~ 20 kb and shared 100% similarity to carotenoid biosynthetic gene cluster from *Rhodobacter sphaeroides* (Accession: S71770). The remaining identified clusters did not show any significant homology with other organisms. Isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) are the fundamental units in isoprenoid biosynthesis and isoprenoid-based biofuel production. These compounds are synthesised via the mevalonate pathway (MVA pathway) and the methylerythritol phosphate pathway (MEP pathway). Furthermore, the pathway analysis of the strains BA0156<sup>T</sup>, LMG 30805<sup>T</sup> and DSM 5680<sup>T</sup> showed the presence of a complete set of genes viz. dxs, dxe, ispD, ispE, ispF, ispG and ispH in the MEP pathway responsible for IPP and DMAPP synthesis.

**Description of Hydrogenophaga crocea sp. nov.**

*Hydrogenophaga crocea* (cro.ce’a. L. fem. adj. crocea pertaining to the saffron colour of the colony).

Gram-stain negative, aerobic, catalase and oxidase positive. Non-spore forming, motile short rods,
yellowish-brown colour with an entire margin and convex elevation. Cell size on solid medium (Luria agar) is 1–2 mm and 0.4 to 0.5 μM width and 0.7 to 1.3 μM in length by scanning electron microscopy. Able to grow at 5–40 °C (Optimum 30 °C), pH 6–11 (Optimum 7.0), salinity 0.5–3% (Optimum 0.5%). Assimilates adipate, malate and citrate and gluconate. Positive for alkaline phosphatase, esterase, esterase lipase, lipase, leucine arylamidase, cysteine arylamidase, valine arylamidase enzyme activities. Utilised D-glucuronic acid, quinic acid, phenylacetic acid, L-pyroglutamic acid, D-galacturonic acid, methyl pyruvate, L-lactic acid, D-Fructose-6-P04, D-gluconic acid and glucuronamide. Resistant to Fusidic acid, D-serine, 1% sodium lactate, rifamycin SV, lincomycin, guanidine-HCl, vancomycin, tetrazolium violet, tetrazolium Blue, lithium chloride, and aztreonam such as chemicals and antibiotics. C16:0 (25.1%) and C17:0 cyclo (3.9%) are the major fatty acids, whereas phosphatidylethanolamine and diphosphatidylglycerol are the major polar lipids. The DNA G+C content is 69.4 mol %.

The type strain is BA0156T (= MCC 3062T = KCTC 72452T = JCM 34507T), isolated from cyanobacterial mat collected from sugarcane field.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00203-022-02865-2.

Acknowledgements The authors sincerely thank Dr T. D. Nikam and Dr Kirti Nitnavare, Department of Botany, Savitribai Phule Pune University (Pune, India), for providing the cyanobacterial sample.

Author contributions VT, KK, SC and SM carried out the polyphasic taxonomy experiments; VT, KK, and BT did the phenotypic and genome data analysis and wrote the first draft of the manuscript. SS provided guidance and in-house facilities. TL and AY did the formal analysis, reviewed, edited and finalized the manuscript. AY conceptualized, led the investigation and provided the funding for the study. All the authors reviewed and approved the final version of the paper.

Funding Authors acknowledge the project funding and fellowships to KK and BT by the Department of Science and Technology (DST), Government of India under grant number SERB/EEQ/2016/000752; authors also acknowledge the funding by Department of Biotechnology (DBT), Government of India under grant number BT/COORD. II/01/03/2016 (NCMR) used for in-house laboratory facilities. Authors gratefully acknowledge the University Grant Commission (UGC) of the Government of India for providing CSIR–UGC NET–JRF fellowship to KK (Ref. No. 857/CSIR-UGC NET JUNE 2017).

Declarations

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

Ethical statement No human or animal subjects were recruited in this study.
Yoon J-H, Kang S-J, Ryu SH et al (2008) *Hydrogenophaga bisanensis* sp. nov., isolated from wastewater of a textile dye works. Int J Syst Evol Microbiol 58:393–397. https://doi.org/10.1099/ijs.0.65271-0

Yoon S-H, Ha S-M, Kwon S et al (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67:1613. https://doi.org/10.1099/ijsem.0.001755

Zhao Y, Wu J, Yang J et al (2012) PGAP: pan-genomes analysis pipeline. Bioinformatics 28:416–418. https://doi.org/10.1093/bioinformatics/btr655

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.