Progress in Microbes and Molecular Biology

HH PUBLISHER

Genome Report

Whole-Genome Sequence of *Chelatococcus daeguensis* Strain M38T9, Isolated from Ulu Slim Hot Spring in Malaysia

Yi Xian Goh[1], Kok Gan Chan[2,3,4], Kar Wai Hong[5]*

Article History

Received: 13 July 2022;

Received in Revised Form: 10 August 2022;

Accepted: 17 August 2022;

Available Online: 23 August 2022

[1]Department of Computational Biology, High Impact Research Building, University of Malaya, Kuala Lumpur, 50603, Malaysia; elizabel8683@gmail.com (YXG)

[2]Department of Biotechnology, Faculty of Applied Sciences, UCSI University Kuala Lumpur, 56000, Kuala Lumpur, Malaysia

[3]International Genome Centre, Jiangsu University, Zhenjiang, China.

[4]Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603, Kuala Lumpur, Malaysia; kokgan@um.edu.my (KGC)

[5]Novel Bacteria and Drug Discovery Research Group (NBDD), Microbiome and Bioresource Research Strength (MBRS), Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, Jalan Lagoon Selatan, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

*Corresponding author: Kar-Wai Hong; Novel Bacteria and Drug Discovery Research Group (NBDD), Microbiome and Bioresource Research Strength (MBRS), Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, Jalan Lagoon Selatan, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia; hong.karwai@monash.edu (KWH)

Abstract: *Chelatococcus daeguensis* strain M38T9 is a thermotolerant bacterium isolated from a hot-spring in Malaysia. The draft genome of *C. daeguensis* strain M38T9 consists of 4,218,658 bp assembled into 50 scaffolds. The GC content of the genome is 67.91 %, and the sequencing coverage of 184×. There are 4,046 predicted genes, 3,962 protein-coding genes, and 53 RNA-coding genes (tRNA: 45, rRNA: 4). The draft genome has been deposited at DDBJ/ENA/GenBank under the BioProject accession number PRJNA668056. The raw reads were deposited in the Sequence Read Archive (SRA) under accession number SRR13805582. Here we report the draft genome of this strain to expand our understanding of the genomic information available on the genus *Chelatococcus*.

Keywords: *Chelatococcus daeguensis*; nitrogen metabolism; denitrification; dissimilatory nitrate reduction; genome

1. Introduction

The genus *Chelatococcus* falls within the class of *Alphaproteobacteria* and was first reported as obligately aerobic, Gram-negative bacteria by Auling and colleagues in the year 1993[1]. According to the List of Prokaryotic names with Standing in Nomenclature (LPSN)[2], this genus comprises six species at the time of writing, namely *C. asaccharovorans*, *C. caeni*, *C. composti*, *C. daeguensis*, *C. reniformis*, and *C. sambhunathii* [1, 3-7]. Most notably, there has been a growing interest in *C. daeguensis* in the recent years for its capability in utilizing a variety of carbon sources and its capability in performing aerobic denitrification at high temperatures, as well as its capability in biodegrading crude oil, coal and toxic metals[5, 8-11]. To provide insights into the genomic basis for these mechanisms, we hereby present the draft genome of *C. daeguensis* M38T9, isolated from Ulu Slim hot spring, Malaysia (3.8986 N 101.4847 E, 110°C, pH 7).

2. Data Description

The genomic DNA from M38T9 was extracted using a Masterpure™ DNA purification kit (Epicentre, Illumina Inc., Madison, WI, USA) upon growing the cells at 37°C on Luria-Bertani (LB) agar[12]. The quality and quantity of DNA were quantified using NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA) and a Qubit version 2.0 fluorometer (Life Technologies, Carlsbad, CA, USA), respectively[13]. The sequencing library was constructed using the Nextera DNA library kit, followed by sequencing performed using Illumina HiSeq 2500 platform[14]. The quality of these raw reads was checked using FastQC (version 0.11.9)[15], and the low-quality reads were trimmed using Trimmomatic (version 0.39)[16] with the default settings for paired-end reads. Subsequently, the quality-filtered reads were *de novo* assembled by Velvet (version 1.2.10)[17]. The quality of assembly was analyzed using QUAST (version 5.0.2)[18] and BUSCO (version 3)[19]. The assembled gene was annotated using NCBI Prokaryotic Genome Annotation Pipeline (PGAP)[20] as well as MicroScope[21]. Default parameters were used for all software unless otherwise specified.

The identity of strain M38T9 was first determined using Microflex LT (Bruker Daltonics, Bremen, Germany), and the strain M38T9 was identified as Chelatococcus species[22]. The 16S ribosomal RNA (rRNA) gene of strain M38T9 was compared with the EzBiocloud Database[23]. Interestingly, the 16S rRNA gene of strain M38T9 showed the highest sequence identity with the 16S rRNA genes of *C. daeguensis* and *C. sambhunathii*, both at 99.72% of identity. However, the estimated average nucleotide identity (ANI) value determined by autoMLST[24] suggested the strain M38T9 is likely to be *C. daeguensis* as the estimated ANI for strain M38T9 by referring to *C. danguensis* strain M3 is 100% (*p*-value=0.000).

The draft genome of *C. daeguensis* strain M38T9 consists of 4,218,658 bp assembled into 50 scaffolds, with N_{50} and L_{50} of 311,252 bp and 6 bp, respectively. The GC content of the genome is 67.91 %, and the sequencing coverage of 184×. There are 4,046 predicted genes, 3,962 protein-coding genes, and 53 RNA-coding genes (tRNA: 45, rRNA: 4). The

draft genome has been deposited at DDBJ/ENA/GenBank under the BioProject accession number PRJNA668056. The raw reads were deposited in the Sequence Read Archive (SRA) under accession number SRR13805582. The version described in this paper is the first version.

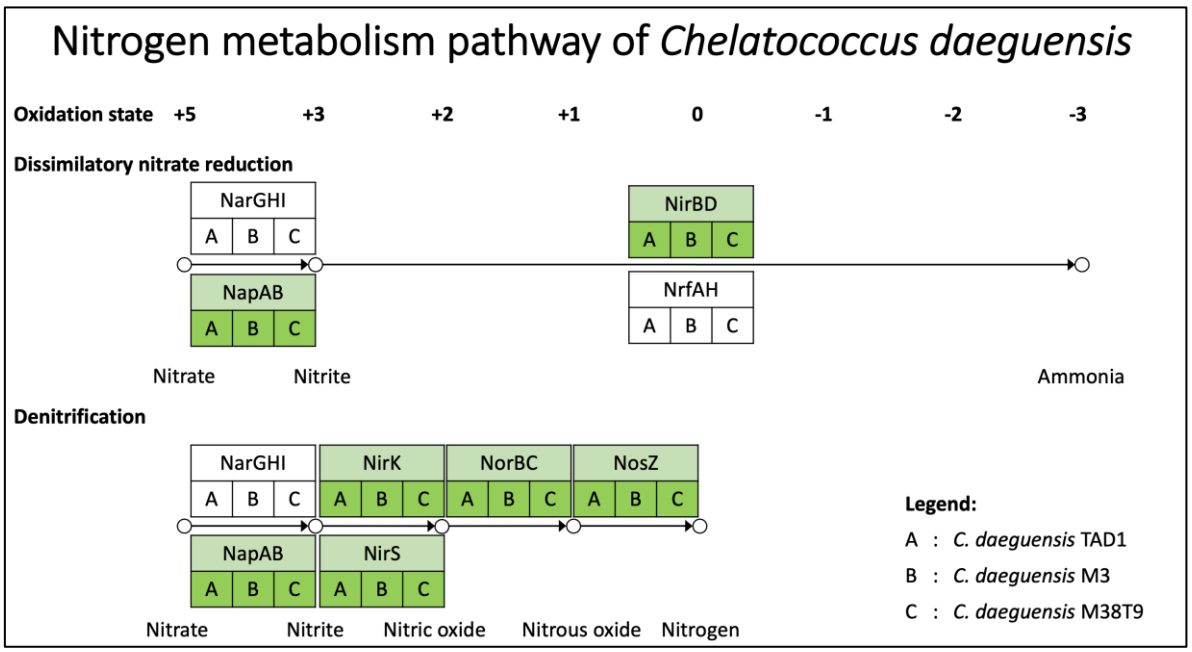

Figure 1. Complete pathway of dissimilatory nitrate reduction and denitrification in *C. daeguensis* strain TAD1, M3 and M38T9.

By comparing the metabolic pathway of three *C. daeguensis* strains, namely TAD1 (NCBI accession number: CP018095.1), M3 (NCBI accession number: LQQT01000000), and M38T9 (NCBI accession number: JAFDUY010000000), a complete metabolic pathway of denitrification has been identified (Figure 1). The denitrification metabolic pathway comprises of periplasmic nitrate reductase complex NapAB which reduces nitrate (NO_3^-) to nitrite (NO_2^-), copper-containing nitrite reductase NirK which further reduces NO_2^- to nitric oxide (NO), nitric oxide reductase complex NorBC which reduces NO to nitrous oxide (N_2O), and nitrous oxide reductase NosZ which reduces N_2O to nitrogen (N_2). Similarly, via comparative genomics, the complete pathway of dissimilatory nitrate reduction was identified in all three strains. The dissimilatory nitrate reduction comprises periplasmic nitrate reductase complex NapAB, which reduces nitrate (NO_3^-) to nitrite (NO_2^-), followed by nitrite reductase, which reduces nitrite (NO_2^-) to ammonia (NH_3). This genome report provided vital insights into the genomic basis for these mechanisms.

Author Contributions: YXG and KWH conducted the experiments and analyzed the data. KWH and KGC provided vital guidance, technical support, and proofreading for the work. All authors approved the final draft.

Funding: This work was supported by the University of Malaya via PPP Grant (PG085-2015B) awarded to KWH, and High Impact Research Grants (UM-MOHE HIR Grant UM.C/625/1/HIR/MOHE/CHAN/14/1, no. H-50001-A000027 and A-000001-50001) which are awarded to Kok-Gan Chan.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Auling G, Busse HJ, Egli T, et al. description of the Gram-negative, obligately aerobic, nitrilotriacetate (NTA)-utilizing bacteria as *Chelatobacter heintzii*, gen. nov., sp. nov., and *Chelatococcus asaccharovorans*, gen. nov., sp. nov. Syst Appl Microbiol 1993; 16(1): 104-112.

2. Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, et al. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. Nucleic Acids Res 2021; 50(D1): D801-D807.

3. Jin L, Ko SR, Lee HG, et al. *Chelatococcus caeni* sp. nov., isolated from a biofilm reactor sludge sample. Int J Syst Evol Microbiol 2015; 65(Pt_3): 885-889.

4. Zhang Z, Zhao J, Yu C, et al. *Chelatococcus composti* sp. nov., isolated from penicillin fermentation fungi residue with pig manure co-compost. Int J Syst Evol Microbiol 2017; 67(3): 565-569.

5. Yoon JH, Kang SJ, Im WT, et al. *Chelatococcus daeguensis* sp. nov., isolated from wastewater of a textile dye works, and emended description of the genus *Chelatococcus*. Int J Syst Evol Microbiol 2008; 58(9): 2224-2228.

6. Gu Z, Liu Y, Wang N, et al. *Chelatococcus reniformis* sp. nov., isolated from a glacier. Int J Syst Evol Microbiol 2016; 66(11): 4525-4529.

7. Panday D and Das SK. *Chelatococcus sambhunathii* sp. nov., a moderately thermophilic alphaproteobacterium isolated from hot spring sediment. Int J Syst Evol Microbiol 2010; 60(4): 861-865.

8. Ke CY, Lu GM, Wei YL, et al. Biodegradation of crude oil by *Chelatococcus daeguensis* HB-4 and its potential for microbial enhanced oil recovery (MEOR) in heavy oil reservoirs. Bioresour Technol 2019; 287: 121442.

9. Liang W, Huang S, Liu J, et al. Removal of nitric oxide in a biotrickling filter under thermophilic condition using *Chelatococcus daeguensis*. J Air Waste Manag Assoc 2012; 62(5): 509-516.

10. Li H, Huang S, and Zhang Y. Cr(VI) removal from aqueous solution by thermophilic denitrifying bacterium *Chelatococcus daeguensis* TAD1 in the presence of single and multiple heavy metals. J Microbiol 2016; 54(9): 602-610.

11. Yang Y, Huang S, Zhang Y, et al. Nitrogen removal by *Chelatococcus daeguensis* TAD1 and its denitrification gene identification. Appl Biochem Biotechnol 2014; 172(2): 829-839.

12. Ser HL, Ab Mutalib NS, Yin WF, et al. Genome sequence of *Streptomyces* antioxidans MUSC 164T isolated from mangrove forest. Prog Microbes Mol Biol 2018; 1(1).

13. Torres M., Hong KW, Chong TM, et al. Genomic analyses of two *Alteromonas stellipolaris* strains reveal traits with potential biotechnological applications. Sci Rep 2019; 9(1): 1215.

14. Letchumanan V, Tan WS, Yin WF, et al. Genome sequence of *Vibrio* sp. OULL4 isolated from shellfish. Prog Microbes Mol Biol 2020; 3(1).

15. Brown J, Pirrung M, and McCue LA. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinform 2017; 33(19): 3137-3139.

16. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinform 2014; 30(15): 2114-2120.

17. Zerbino DR, and Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008; 18(5): 821-829.

18. Gurevich A, Saveliev V, Vyahhi N, et al. QUAST: quality assessment tool for genome assemblies. Bioinform 2013; 29(8): 1072-1075.

19. Simão FA, Waterhouse RM, Ioannidis P, et al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinform 2015; 31(19): 3210-3212.

20. Tatusova T, DiCuccio M, Badretdin A, et al. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 2016; 44(14): 6614-6624.

21. Vallenet D, Engelen S, Mornico D, et al. MicroScope: a platform for microbial genome annotation and comparative genomics. Database 2009; 2009.

22. Hong KW, Hani AA, Murni CNA, et al. Comparative genomic and phylogenetic analysis of a toxigenic clinical isolate of *Corynebacterium diphtheriae* strain B-D-16-78 from Malaysia. Infect, Genet Evol 2017; 54: 263-270.

23. Yoon SH, Ha SM, Kwon S, et al. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 2017; **67**(5):1613-1617.

24. Alanjary M, Steinke K, and Ziemert N. AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. Nucleic Acids Res 2019; 47(W1): W276-W282.