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 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 N50 and L50 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: JAFDUY01000000), 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 nitrogen oxide (N$_2$O), and nitrous oxide reductase NosZ which reduces N$_2$O 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


