Genomics and Simulated Laboratory Studies Reveal Thermococcus sp. 101C5 as A Novel Hyperthermophilic Archaeon Possessing Specialized Metabolic Arsenal for Enhanced Oil Recovery from High-Temperature Oil Reservoirs (101 °C)

Neelam G. Kapse
Agharkar Research Institute
Vasundhara Paliwal
Agharkar Research Institute
Sumit Singh Dagar
Agharkar Research Institute
Dolly Pal Rana
Institute of Reservoir Studies
Prashant K Dhakephalkar (pkdhakephalkar@aripune.org)
Agharkar Research Institute

Research Article

Keywords: Genome, Hyperthermophile, Microbial Enhanced Oil Recovery, Simulated reservoir conditions, Thermococcus

Posted Date: June 25th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-572353/v1

Abstract

Laboratory evaluation of hyperthermophiles with the potential for Enhanced Oil Recovery (EOR) is often hampered by the difficulties in replicating the *in situ* growth conditions in the lab. In the present investigation, genome analysis was used to gain insights into the metabolic potential of a hyperthermophile to mobilize the residual oil from depleting high-temperature oil reservoir. Here, we report the 1.9 Mb draft genome sequence of hyperthermophilic anaerobic archaeon, *Thermococcus* sp. 101C5 with a GC content of 44%, isolated from a high temperature oil reservoir of Gujarat, India. 101C5 possessed the genetic arsenal required for adaptation to harsh oil reservoir conditions, such as various heat shock proteins for thermo-adaptation, Trk potassium uptake system proteins for osmo-adaptation, and superoxide reductases against oxidative stress. MEOR potential of the strain was established by the presence of genes encoding enzymes involved in desired metabolite production like hydrogen, acetate, exopolysaccharide, bio-emulsifier, etc., which was further experimentally confirmed and validated. Also, the presence of crude oil degradative genes highlighted the ability of the strain to mobilize heavy residual oil, which was confirmed under simulated conditions in sand-pack studies. The obtained results demonstrated additional oil recoveries of 42.1% and 56.5% at 96°C and 101°C, respectively, by strain 101C5, illustrating its potential for application in high-temperature oil reservoirs. To our best knowledge, this is the first report of genome analysis of any microbe assessed for its suitability for MEOR from the high-temperature oil reservoir.

Introduction

Energy demand is globally increasing day by day, which is sufficed mainly by fossil fuels. Despite the recent technological advances in the production of biofuels (bioethanol, biodiesel, biohydrogen, etc.), crude oil has no viable alternative as of the date (Naik et al. 2010). Most biofuel-based technologies have either safety issues or application issues, whereas some lack cost-competitiveness compared to fossil fuels (De Almeida et al. 2016; Elraies & Tan 2012). Thus, crude oil continues to remain the major fuel source of our energy sector.

Conventional technologies of crude oil extraction can usually extract up to 40% – 45% of oil, leaving up to two-third of oil trapped in the reservoir (Sen 2008), which highlights the need for a process that can harness every single drop. Microbial Enhanced Oil Recovery (MEOR) is one such promising process that can extract the maximum possible oil from a reservoir (Lazar et al. 2007). The efficiency of MEOR process depends upon the production of metabolites such as acids, gases, polymers, solvents, or surfactants by the exogenous or indigenous microbes (Almeida et al. 2004). These metabolites are known to significantly alter the physical properties of reservoirs, including permeability, porosity, and wettability along with the oil properties such as viscosity (Brown 2010) and facilitate oil recovery by emulsifying the oil, reducing the viscosity of the oil, pressurizing the reservoir environment and dissolving the residual oil (Sen 2008).

The oil reservoirs are characterized by harsh environments having extreme temperatures (10 to 124°C), salinity (100 to 300 g/L), and pH (5 to 8) (Varjani & Gnansounou 2017; Rathi et al. 2015; Safdel et al. 2017). These physicochemical conditions can significantly influence microbial survival, growth, and metabolite production in oil reservoirs (Li et al. 2013; Li et al. 2010). Very few studies have reported the presence of active microbes at higher temperatures of 80–90°C (Bachmann et al. 2014; Grassia et al. 1996; Arora et al. 2014). Therefore, microbial candidates suitable for MEOR applications should be anaerobic, thermophilic, barophilic, and halophilic (Maudgalya et al. 2007). Several microbial genera like *Clostridium, Bacillus, Pseudomonas, Arcobacter, Enterobacter, Acinetobacter*, etc., that have been successfully used for MEOR applications have remained restricted to low-temperature oil reservoirs (Zhang et al. 2012). However, the MEOR process has certain limitations pertaining to high-temperature oil reservoirs, especially in countries like India, where many oil reservoirs have temperatures above 90 °C. Notably, these studies have been hampered by the fact that microbes tolerating high temperatures and pressures are difficult to cultivate.

Most MEOR experiments have exhibited excellent results in lab-scale studies (Costa et al. 2012; Al-Sulaimani et al. 2012), and some field trials have shown comparable oil recovery and better economic efficiency than chemical EOR (Al-Wahaibi et al. 2014; Cao et al. 2013; Gao & Zekri 2011; Youssef et al. 2013). The effectiveness of MEOR has remained unpredictable despite these successes. Field applications of MEOR are restricted mainly due to the limited understanding of microbial survival and metabolism within the oil reservoir (Cao et al. 2013).

In recent years, various metagenomics studies of oil reservoirs have illustrated the phylogenetic profiles of the microbial communities native to oil reservoirs (Liu et al. 2018; Hu et al. 2016; Kim et al. 2018). However, genes responsible for the production of metabolites required for MEOR were largely unmapped. Hence, there is a paucity of information evaluating the suitability of the resident microbial flora for MEOR application. Genome analysis can identify microbial strains possessing the potential to facilitate MEOR. Such strains are expected to possess genes encoding proteins involved in attributing the environmental stress resistance to the host or metabolites involved in mobilizing the residual oil, enhancing its recovery. Genes encoding properties like tolerance to thermo-, osmo- and oxidative stress present in reservoir and production of desired metabolites can be used as biomarkers. With the aid of these markers, potential MEOR strains can be identified rapidly, and an efficient MEOR process can be developed for implementation in high-temperature depleted reservoirs with a fair degree of success. With this aim, the present study was undertaken wherein hyperthermophiles were isolated from a high-temperature oil reservoir. The whole-genome sequencing of the most potent strain was performed, which provided genomic insights into its ability to survive under harsh reservoir conditions and produce metabolites desired for MEOR. The oil recovery potential of this strain was further validated under laboratory simulated reservoir conditions using sand pack columns.

Materials And Methods

Sampling details

The produced water sample was collected from the wellhead of Nandasan 9, an oil reservoir situated in Ahmedabad (Gujarat), India. *In situ* temperature of the reservoir was 112 ℃. The sample was stored at 4˚C until processed for the enrichment of hyperthermophilic bacteria/ archaea.

Culturing and characterization of hyperthermophile

For enrichment of hyperthermophilic microbes, 10% (v/v) of the sample was inoculated into sterilized serum bottles containing molasses medium (20 ml) flushed with N_2 and incubated at 101 ℃, pH 7.0 until growth was observed. Molasses medium contained molasses 7.0%, yeast extract 0.5%, NaCl 5% supplemented with 4% SS-30 trace element solution (K_2HPO_4 0.1 %, KH_2PO_4 0.1%, $NaHCO_3$ 1%, $CaCl_2$ anhydrous 0.02%, NaCl 0.2%, $MgSO_4$ 0.02%). Isolation of pure cultures from the enriched sample was performed by Hungate's roll tube method. Following isolation, a single culture designated as 101C5 was selected for further studies based on growth and desired metabolite production. Morphological studies of the isolated strain were performed with a scanning electron microscope (ZEISS, EVO MA-15, Germany). For molecular identification, Genomic DNA of the strain 101C5 was isolated using the Nucleospin Soil DNA kit (Macherey-Nagel) using manufacturers protocol, and 16S rRNA gene sequencing was performed using primer set 86F (5'-GCT CAG TAA CAC GTG G-3') and 1340R (5'-CGG TGT GTG CAA GGA G-3') (Wright et al. 2004). Twenty µL reaction mixture contained 3 µL of genomic DNA, 1 U of Taq DNA polymerase (Sigma, India), 1X Taq buffer, 1.5 mM MgCl2, 0.2 mM of dNTP mix (Sigma) and 0.4 µM each primer. The amplification program consisted of 1 cycle at 94 ℃ for 5 min, 39 cycles at 94 ℃ for 1 min, 56 ℃ for 1 min, 72 ℃ for 1 min, 30 sec and 1 cycle of final extension at 72 ℃ for 10 min, in a thermal cycler (GeneAmp PCR system 9700, Applied Biosciences, USA).

Genome sequencing and annotation

The genome of 101C5 was sequenced on the Illumina NextSeq500 platform. *De novo* genome assembly was performed using Spades assembler version 3.13.0. Whole-genome comparison of strain 101C5 was done with publically available genomes of closes phylogenetic affiliates using online tools such as Digital DNA-DNA hybridization (DDH) (Auch et al. 2010) and Average Nucleotide Identity (ANI) (Yoon et al. 2017). Functional annotation of the genome was performed by Rapid Annotation using Subsystem Technology (RAST) (Aziz et al. 2008) and PATRIC (Wattam et al. 2017). Further, the genome was mined for genetic determinants involved in its adaptability to harsh reservoir conditions and its oil recovery potential.

Metabolite profiling of strain 101C5

The strain 101C5 was screened for the production of different metabolites (gas, acids, solvents, and VFA) desirable for the MEOR application after 14 days of incubation at 96 ℃ and 101 ℃. Gas analysis for hydrogen, carbon dioxide, and nitrogen was performed using Gas chromatography (PerkinElmer) equipped with thermal conductivity detector (TCD), Porapak Q column (SS, 1/800 _ 80 mesh) with argon as carrier gas (flow rate: 40 ml/min). Bio-emulsifier production was analyzed in terms of the emulsification index using Kerosene (Satpute et al. 2010). Acidified cell-free supernatant of the grown

culture was used for VFA and HPLC analysis. Volatile Fatty Acids (VFAs) were determined on Gas Chromatograph (Chemito-8510, India) equipped with Flame Ionization Detector, column 10% FFAP + 2% H_3PO_4, injector temperature 180 °C, oven temperature 140 °C, detector temperature 190 °C, carrier gas N_2 (40 ml min^{-1}). Organic acids and solvents were analyzed using the HPLC system (Dionex, USA) using Aminex HPX-87H, Biorad column, equipped with a refractive index (RI) detector. The working conditions were: 5 mM H_2SO_4 as a mobile phase with a flow rate of 0.7 ml/min at 40 °C. Exopolysaccharide production was determined by precipitation by isopropanol (Pawar et al. 2013), followed by gravimetric analysis. A control set of un-inoculated medium bottles was kept at similar experimental conditions.

Evaluation of MEOR potential of strain 101C5 under simulated reservoir conditions

Evaluation of reservoir properties affecting oil recovery will permit the prudent application of microorganisms for enhanced oil recovery. Sand-pack columns were designed to simulate the oil reservoir and used to evaluate the potential of 101C5 in enhanced oil recovery. Briefly, vertically oriented metal columns of 60 ml volume were uniformly packed with acid-washed dry sand. The column was then flooded with brine (2 % NaCl) followed by heavy crude oil (collected from oil reservoir located in Gujarat) under pressure till brine was washed out. The sand-packed column's void volume was determined as $V_{\text{crude oil}}$ in $-V_{\text{crude oil}}$ out. Primary recovery of crude oil was done using water, and the amount of crude oil displaced was measured. The sand pack was subsequently flooded with nutrient medium to ensure maximum possible displacement of oil till saturation of residual oil was reached. The column was inoculated with 30 ml inoculums (10^7 cells/ml) in the test and sterile distilled water in the control column and incubated at 96 °C and 101 °C for 14 days. Oil recovery was made after 14 days, and the columns were flooded with water to remove the excess oil until no more oil mixture was observed in the effluent. The amount of oil recovered was measured, and % oil recovery was calculated as (oil recovered after 14 days/ oil in place) *100.

Results And Discussion

Isolation and identification of 101C5

A hyperthermophilic strain, designated as 101C5, was isolated from a high-temperature oil reservoir at 101 °C by enrichment, followed by roll tube isolation technique. Strain 101C5 was selected for further studies based on its ability to grow at 101 °C and produce metabolites desirable for MEOR. Morphological characterization revealed 101C5 as Gram-negative cocci (0.3 µ to 0.5 µ in diameter), which occurred in clusters (Fig. S1). The strain 101C5 shared the highest 16S rRNA gene sequence homology of 100% with phylogenetic affiliate *Thermococcus litoralis* DSM 5473 strain NS-C. The genome of strain 101C5 was sequenced, and the sequence homology of the extracted housekeeping genes was determined. Relatively low sequence homologies with closest phylogenetic affiliate DSM 5473 (as identified by 16S rRNA gene sequence) for each of the housekeeping genes (90.8% for rpoB, 88.95% for gly A- serine methyltransferase and 83.31% for pfkB- carbohydrate kinase) identified strain 101C5 to be taxonomically different from DSM 5473. Further, the entire 101C5 genome was compared with other

publically available genomes of *Thermococcus* (*Thermococcus litoralis* DSM 5473 and *Thermococcus* sp. PK) using Digital DNA-DNA-hybridization (DDH) and Average Nucleotide Identity (ANI) tool. The strain 101C5 shared only 33.6% and 67.9% DDH homology with DSM 5473 and PK, respectively, which was less than the threshold of 70%. Similarly, lower ANI homologies of 87.7% and 96.5% were shared with DSM 5473 and PK, respectively, which was less than the threshold of 95% (Chun et al. 2018). Thus, strain 101C5 was identified as a putative novel member of genus *Thermococcus* based on lesser threshold homologies in DDH and ANI.

The genus *Thermococcus* comprises strictly anaerobic and hyperthermophilic archaea belonging to the order Thermococcales within the phylum of euryarchaeota. The euryarchaeal rrder Thermococcales consists of three genera *Pyrococcus*, *Palaeococcus* and *Thermococcus*. Members of the genus *Thermococcus* include 19 different species and are common inhabitants of extreme niches such as hot springs, hydrothermal vents, deep subsurface oil reservoirs, etc., with growth temperature ranging from 45 °C to 90 °C (Adams 1993; Summit and Baross, 1998). Majority of *Thermococcus* sp. are organoheterotrophs fermenting organic compounds into CO_2 and H_2 as end products (Teske et al. 2019), while few are capable of lithotrophic CO-dependent hydrogenogenic growth such as *Thermococcus* strain AM4 and *Thermococcus onnurineus* (Svetlitchnyi et al. 2001). Though the presence of *Thermococcus* sp. has been reported from high-temperature oil reservoirs (Zahner et al. 2010; Orphan et al. 2000), only two reports have documented its role in the MEOR application (Takahata et al. 2001; Lal et al. 2009), thus warranting further study in this area.

General genome features and comparative genomics of 101C5 with related species

The sequencing of the 101C5 genome yielded a genome size of 1,926,104 bp with 119 contigs and GC content of 44%. A circular map illustrating the graphical display of the distribution of genome annotations was generated (Fig. S2) (Wattam et al. 2017), wherein, outermost circle 1 represented the contigs, circle 2 represents CDS on the forward strand, circle 3 represents CDS on the reverse strand, circle 4 represents RNA genes, circle 5 represents CDS with homology to known antimicrobial resistance genes, circle 6 represents GC content, and innermost circle 7 represents GC skew. The colors of the CDS on the forward and reverse strand indicate the subsystem that protein-coding genes belong to. RAST functional annotation predicted 146 subsystems having 2304 total genes, with 2258 protein-coding genes and 46 RNA encoding genes. Majority of the 101C5 proteins showed similarity to those of *Thermococcus* sp. PK. The phylogenomic tree constructed using PATRIC indicated that 101C5 was placed at the same node with PK, revealing a close taxonomic relationship of 101C5 with *Thermococcus* sp. PK (Fig. 1). However, 101C5 was placed distantly from DSM5473, indicating the taxonomic difference between both species.

Genome reports of six different species of *Thermococcus* namely, *T. sibricus, T. barophilus, T. kodakaraensis*, *T. gammotolerans, T. onnurineus* and *T. litoralis* are publically available and have provided insights into their genetic makeup and metabolic machinery. Except for *T. sibricus*, none of the reported species of *Thermococcus* originated from oil reservoirs. *T. sibricus* was isolated from a high-temperature oil reservoir located in Western Siberia, growing anaerobically at 78 ℃. The genome size of

T. sibricus was 1.84 Mb with 2061 protein coding genes, providing metabolic insights into substrate utilization and energy generation. The *T. sibricus* genome established its indigenous origin from oil reservoir based on the presence of enzymes involved in metabolizing organic matter (cellulose, agarose, laminarin, and triglycerides) from oceanic sediments and alkanes from crude oil (Mardanov et al. 2009). However, this study did not provide any information regarding the adaptive starategy employed by *T. sibricus* to survive in harsh reservoir conditions. *T. barophilus* MP a hyperthermophilic, piezophilic archaeon capable of growing at 95°C and 40 MPa was isolated from chimney samples. The MP genome consisted of a circular chromosome of size 2.01 Mb and a circular plasmid of size 54,159 bp, harboring the hydrogenase complexes and genes involved in carboxydotrophic pathway (Vannier et al. 2011). *T. kodakaraensis* KOD1 was isolated from Kodakara Island, Japan and possessed a genome size of 2.08 Mb with 2306 coding sequences. KOD1 genome reported several proteins involved in nucleotide metabolism, stress response, etc (Fukui et al. 2005). *T. gammotolerans,* a radioresistant archaea isolated from hydrothermal chimneys had a genome size of 2.045 Mb with 2157 coding sequences. Genome analysis of *T. gammotolerans* provided insights into catabolism, hydrogenase complexes, detoxification systems and DNA repair (Zivanovic et al. 2009). *T. onnurineus* NA1 isolated from a deep-sea hydrothermal vent possessed a genome size of 1.84 Mb with 1976 coding sequences. Genome analysis of NA1 reported the carboxydotrophic pathway in combination with heterotrophy as well as provided the first evidence of lithotrophic feature of *Thermococcus* (Lee et al. 2008). *T. litoralis* NS-C isolated from a hot spring Italy had a genome size of 2.3 Mb with 2724 coding sequences. The NS-C genome provided crucial information regarding the mobile genetic elements inteins (Gardner et al. 2012).

101C5 genome was compared with the publically available *Thermococcus* genomes of PK, EP1, DSM 5473 and oil based on general genome features using RAST (Table S1). The genome size of *Thermococcus* sp. used for comparison in the present study ranged from 1.82 – 2.22 Mb, with DSM 5473 having the largest genome size among the genomes used. Interestingly, the genomes sued in the study had similar GC contents (39.3% to ~44%). A noticeable difference in the number of genes between 101C5 and DSM 5473 was observed mainly in carbohydrate, DNA, protein, phosphorus, and amino acids metabolism, membrane transport and stress response. RAST analysis revealed a total of 31, 26, 12 and 31 unique genes respectively, in 101C5 genome compared with other genomes, namely, PK, EP1, DSM 5473 and oil.

BRIG tool (Alikhan et al. 2011) was used for comparing multiple genomes of *Thermococcus* sp. available in the NCBI database with 101C5 genome. The circular map (Fig. S3) representing the BLASTn results of each query genome (*Thermococcus litoralis* DSM 5473, *Thermococcus* sp. PK, *Thermococcus* sp. EP1, and *Thermococcus* sp. oil) against the reference strain 101C5 was generated. BRIG image displayed lighter color intensities and pronounced gaps in each of the query genomes, highlighting the difference between strain 101C5 and the other *Thermococcus* genomes used and further confirmed the novelty of strain 101C5.

Stress resistance mechanisms: Survival under extreme reservoir conditions

An oil reservoir is characterized as an extreme niche with extremes of temperature, pH, salinity, pressure, oxidative stress, etc. These harsh conditions can have a detrimental effect on exogenously injected microbes and can severely impair MEOR efficiency. Thus, adaptation to the environment becomes crucial for microbes to survive, which is controlled by several mechanisms at the molecular level. Genome analysis of strain 101C5 revealed the presence of diverse and abundant genes known to have a role in combating these physiological stresses (Table S2).

Temperature is one of the critical factors affecting microbial growth and metabolism in oil reservoirs. Elevated temperature affects the integrity of the cell membranes and can thus hamper bacterial growth. 101C5 genome harbored genes encoding various heat shock proteins, spermidine synthases, Clp, etc., having a role in thermo-tolerance. Spermidine synthases are physiological polyamines that are reported to be strong candidates responsible for thermophily in bacteria and archaea (Takami et al. 2004). Genes conferring resistance against oxidative stress (which might be transiently introduced during water flushing used for oil production), such as superoxide reductases, rubrerythrin, etc., were also detected in the 101C5 genome. Rubrerythrins are non-haem iron proteins having a potential role in oxidative stress protection in anaerobes (Sztukowska et al. 2002). 101C5 genome was also found to be well equipped with the genetic arsenal required for adaptation to the saline environment present in a reservoir (osmo-adaptation), such as Trk like transporters, genes involved in compatible solutes (choline and betaine) synthesis. Under suddenly imposed and prolonged osmotic stress, uptake of K^+ ions *via* Trk-like transporter could act as an essential feature in the cellular defense of strain 101C5 (Kraegeloh et al. 2005). Genes encoding betaine synthesis were also detected in the genome, which indicated its osmo-adaptive trait towards hypersaline environments. Plausible genes involved in tolerance of strain 101C5 to high pressure were also assessed. Previous reports have suggested the involvement of specific heat shock proteins in piezo tolerance (Aertsen et al. 2004; Simonato et al. 2006; Karatzas et al. 2007). The presence of GroEL in the 101C5 genome suggested its role as a protectant against high-pressure stress, which may also aid in repairing damages caused by high pressure.

Thus*, in silico* analysis of 101C5 predicted the excellent reservoir adaptive features of the strain, suggesting its strong adaptability to harsh conditions.

Prediction of functional genes in 101C5 genome to elucidate its potential role in MEOR

Genome mining revealed an arsenal of candidate genes involved in producing various metabolites like hydrogen, exopolysaccharide (EPS), organic acids, bio-emulsifier (Table S3), highlighting the potential of strain 101C5 for MEOR. 101C5 genome contains a wide range of genes encoding various membrane-bound hydrogenases, NiFe hydrogenases, hydrogenase maturation, expression, and formation proteins. These enzymes are actively involved in the generation of H+ ions and the production of bio-hydrogen, which could enhance light oil recovery by the mechanisms of reservoir re-pressurization (Van Hamme et al. 2003). Numerous NiFe hydrogenases have been identified and well characterized from *Thermococcus litoralis* (Rákhely et al. 1999). NiFe hydrogenases are multisubunit membrane bound enzymes involved in hydrogen production.

Biopolymers or exopolysaccharide selectively plug the high permeability zones and divert the displacement fluid towards the low permeability oil-rich zones in the reservoir thus, providing better sweep efficiency of oil, leading to enhanced oil recovery (Guo et al. 2015). Genes encoding Phosphomannomutase, mannose-1-phosphate guanylyltransferase, mannosyl-3-phosphoglycerate phosphatase, and mannosyl-3-phosphoglycerate synthase suggested mannose to be the significant component of EPS produced by strain 101C5. EPS is also known to act as a bio-emulsifier. Emulsification is one of the most efficient mechanisms in MEOR processes. Bio-emulsifiers work by desorbing the oil from rocks and forming a stable oil/water emulsion, consequently facilitating oil mobilization. 101C5 harbored genes encoding key enzymes involved in the Mannoprotein nature of bio-emulsifier production, indicative of crude oil emulsifying and solubilizing ability of strain 101C5. 101C5 genome harbored enzymes such as Pyruvate: Ferredoxin oxidoreductase (POR), Acetyl-CoA synthetase (ADP-forming) (ACS), along with other auxiliary enzymes involved in acetate production. Organic acids, such as acetic acid, play a vital role in oil recovery by acting on the reservoir properties, particularly by dissolving the carbonate rock, thereby increasing its porosity and permeability.

101C5 genome presented the ability to produce gases, EPS, bio-emulsifiers and organic acids simultaneously, suggesting its strong potential for MEOR, which was further validated with wet-lab evidence. Thus, genome-level analysis aided in identifying candidate genes that can act as biomarkers for selecting potential MEOR strains.

Genomic insights into crude oil degradation ability of 101C5

Genome analysis revealed the presence of a cascade of genes involved in the production of enzymes that play a significant role in the biotransformation/ bio-degradation of crude oil. Genes encoding enzymes such as dioxygenases, 4-oxalocrotonate tautomerase, acylphosphatase, etc.having a role in benzoate degradation, alcohol dehydrogenase and haloacid dehalogenase, rubredoxin involved in alkane degradation were present in 101C5 genome (Table S4). The presence of these genes indicated the biotransformation potential of 101C5, implicating its prospects in improving the flow characteristics of heavier crudes. Also, several oxidoreductases, dehydrogenases, and dioxygenases involved in the degradation of PAHs, cyclic hydrocarbons and other aromatic compounds were detected in the genome. Members of genus *Thermococcus* have been reported as efficient hydrocarbon degraders, particularly in high-temperature oil reservoirs, making them ideal candidates for MEOR (Nikolova & Gutierrez 2020). Heavy crude oil, because of its high viscosity and density, is challenging to extract, transport, and refine by conventional technologies (Shibulal et al. 2014). However, if strain 101C5 can bio-transform or partially degrade heavy oil into light oil, then this trait can be considered desirable by many.

Thus, genome analysis unraveled the ability of strain 101C5 to produce MEOR desired metabolites, which was further validated in wet-lab experiments.

Metabolite profiling of hyperthermophile 101C5

The highest production of CO_2, organic acids and emulsifying activity was obtained with strain 101C5 in molasses medium (20 ml) in 14 days at 101 °C (Table 1). This validates the potential of the hyperthermophilic strain 101C5 towards microbial enhanced oil recovery. For successful implementation of the MEOR process, microbial performance in terms of sustained growth and production of desired metabolites such as CO_2, biosurfactants, solvents, acids, etc., is of utmost importance. Microbial production of carbon dioxide is a desired property for MEOR as it enhances oil recovery by increasing the reservoir pressure and reducing the viscosity of heavy oil (Al-Sulaimani et al. 2011). In the present study, 101C5 produced significant quantities of gas, with carbon dioxide (43.8 ml molasses g^{-1}) as the major component. It was interesting to note that strain 101C5 produced substantial quantities of hydrogen (33.1 ml molasses g^{-1}) and lower quantities of CO_2 (3.2 ml molasses g^{-1}) at 96 °C, however at 101 °C, a high quantity of CO_2 was produced with negligible H_2 production (Table 4). Bio-acids such as lactate and acetate, which are produced as the byproducts of anaerobic fermentation of sugars, can be helpful in carbonate reservoirs or sandstone formations cemented by carbonates. Acids are known to dissolve the carbonate rocks, resulting in improved porosity and permeability and also aids in the solubilization of oil (Adkins et al. 1992; Fardeau et al. 2004). Strain 101C5 produced 1477 mg/l of acetate, 466.4mg/l of lactate using molasses as fermentable C source and Yeast extract as N source. Bio-emulsifier production was analyzed in terms of the emulsification index. Arora et al. (2014) have reported EI of 43 % for thermophilic *Clostridium* sp. growing at 96 °C. The highest EI_{24}% of 60.8% was obtained in the present study by strain 101C5 at 101 °C. This was suggestive of bio-emulsifier activity known to form stable oil-water emulsions. This property is known to mobilize trapped oil in reservoirs.

101C5 mediated oil recovery under simulated reservoir conditions

Reservoir characteristics can drastically influence the final outcome of MEOR as they significantly influence the microbial performance and, thus, the efficiency of MEOR. Such reservoir characteristics include *in situ* temperature, pressure, the porosity of reservoir rocks, carbonaceous or otherwise nature of the formation rocks, viscosity and pour point of the oil in place, etc. Thus, it becomes imperative to evaluate the oil recovery potential of the microbe under simulated reservoir conditions. The pore size of the rock will determine the microbial penetration and extent of the growth in the pores and crevices. The efficiency of MEOR will be directly affected by lesser pore size. Oil properties can have a significant impact on the oil recovery process. The high viscosity and low flow characteristics of heavy oil make it difficult to extract oil from the reservoir. The viscosity of the oil will determine the concentration of metabolites required for dissolution/ emulsification/ dissipation of oil. Additional oil recoveries as compared to water control were observed at 96 °C and 101 °C (Table 2). With increasing temperature, increased oil recovery was observed. A maximum oil recovery of 56.5% was observed at 101 °C by strain 101C5. To our best knowledge, this is the first report of the Microbial Enhanced Oil Recovery at 101 °C.

Further, >56% recovery of the residual oil by MEOR is one of the highest ever reported efficiency of MEOR for any microorganisms at any temperature in sand pack studies. This ability of strain 101C5 to efficiently recover heavy oil may be attributed to the significant quantities of acetic acid as well as gases

(H_2 and CO_2) generated by the culture, which may have reduced the viscosity of the heavy oil, allowing it to display improved flow features and thereby, facilitating enhanced oil recovery. Organic acids are reported to be the active metabolite involved in the viscosity reduction of heavy crudes, thereby improving their mobility. Laboratory experiments in Daqing Oilfield revealed the viscosity reduction of 36 % with the aid of microbial acids (Han et al. 2001). Studies have also reported the importance of CO_2 in the viscosity reduction of crude oil by expanding it when produced by microbes in a supercritical state of high temperature and high pressure (She et al. 2019). Treatment of heavy oils with strain 101C5 can thus improve their flow characteristics, thereby increasing the additional oil recoveries.

Conclusion

Low digital DNA-DNA hybridization homology (< 70%) and ANI value (< 95%) with the closest phylogenetic affiliate identified 101C5, a hyperthermophilic, halophilic and oxygen tolerant anaerobic isolate, as a putative novel member of the genus *Thermococcus.* Genome analysis of 101C5 revealed its adaptability to extreme reservoir conditions. 101C5 genome harbored the gene repertoire that encoded the production of metabolites desired in MEOR. Strain 101C5 exhibited the excellent ability to recover residual heavy oil at 96 ℃ and 101 ℃ in simulated sand pack studies. Such data validated the potential of *Thermococcus* sp. to facilitate MEOR as identified by the genome analysis. To the best of our knowledge, this is the first report of MEOR facilitated by *Thermococcus* sp. at 101°C in simulated lab studies.

Declarations

Acknowledgments

Gratefully acknowledged: (i) Financial assistance from Institute of Reservoir Studies, ONGC, Ahmedabad and, (ii) CSIR-SRF to Ms. Neelam G. Kapse by the Government of India.

Author's contribution

NGK performed the experiments, did the analysis and wrote the manuscript. VP performed the experiments. SSD reviewed the manuscript. DPR designed the experiments. PKD designed the experiments, performed data validation and reviewed the manuscript.

Funding

The research was supported by Institute of Reservoir Studies, ONGC, Ahmedabad, India

Availability of data and materials

The draft genome sequence of strain 101C5 has been deposited at DDBJ/EMBL/GenBank under accession number WHIZ00000000.

Compliance with ethical standards

Conflicts of interest

The authors declare that there are no conflicts of interest.

Ethical statement

This article does not contain any studies with animals performed by any of the authors.

References

1. Adams MW (1993) Enzymes and proteins from organisms that grow near and above 100 C. Annu Rev Microbiol 47(1):627–658
2. Adkins JP, Cornell LA, Tanner RS (1992) Microbial composition of carbonate petroleum reservoir fluids. Geomicrobiol J 10(2):87–97
3. Aertsen A, Vanoirbeek K, De Spiegeleer P, Sermon J, Hauben K, Farewell A, Nyström T, Michiels CW (2004) Heat shock protein-mediated resistance to high hydrostatic pressure in Escherichia coli. Appl Environ Microbiol 70(5):2660–2666
4. Alikhan NF, Petty NK, Zakour NLB, Beatson SA (2011) BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12(1):1–10
5. Almeida PFD, Moreira RS, Almeida RCDC, Guimaraes AK, Carvalho AS, Quintella C, Esperidia MCA, Taft CA (2004) Selection and application of microorganisms to improve oil recovery. Eng Life Sci 4(4):319–325
6. Al-Sulaimani H, Al-Wahaibi Y, Al-Bahry S, Elshafie A, Al-Bemani A, Joshi S, Ayatollahi S (2012) Residual-oil recovery through injection of biosurfactant, chemical surfactant, and mixtures of both under reservoir temperatures: induced-wettability and interfacial-tension effects. SPE Reservoir Eval Eng 15(02):210–217
7. Al-Sulaimani H, Joshi S, Al-Wahaibi Y, Al-Bahry S, Elshafie A, Al-Bemani A (2011) Microbial biotechnology for enhancing oil recovery: current developments and future prospects. Biotechnol Bioinf Bioeng 1(2):147–158
8. Al-Wahaibi Y, Joshi S, Al-Bahry S, Elshafie A, Al-Bemani A, Shibulal B (2014) Biosurfactant production by Bacillus subtilis B30 and its application in enhancing oil recovery. Colloids Surf B 114:324–333
9. Arora P, Ranade DR, Dhakephalkar PK (2014) Development of a microbial process for the recovery of petroleum oil from depleted reservoirs at 91–96 C. Bioresour Technol 165:274–278
10. Auch AF, von Jan M, Klenk HP, Göker M (2010) Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Standards in genomic sciences 2(1):117–134
11. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F (2008) The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, *9*(1), pp.1–15

12. Bachmann RT, Johnson AC, Edyvean RG (2014) Biotechnology in the petroleum industry: an overview. International Biodeterioration Biodegradation 86:225–237
13. Brown, L.R., 2010. Microbial enhanced oil recovery (MEOR). *Current opinion in Microbiology*, *13*(3), pp.316–320
14. Cao G, Liu T, Ba Y, Xu D, Wang T, Zhao F, Shu Q (2013) Microbial flooding after polymer flooding pilot test in Ng3 of Zhong1 area, Gudao oilfield. Petroleum Geology Recovery Efficiency 20(6):94–96
15. She, H., Kong, D., Li, Y., Hu, Z. and Guo, H., 2019. Recent advance of microbial enhanced oil recovery (MEOR) in China. *Geofluids*, *2019*
16. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol MicroBiol 68(1):461–466
17. Costa AS, Romão LPC, Araújo BR, Lucas SCO, Maciel STA, Wisniewski A Jr, Alexandre MDR (2012) Environmental strategies to remove volatile aromatic fractions (BTEX) from petroleum industry wastewater using biomass. Biores Technol 105:31–39
18. De Almeida DG, Da Silva S, Luna RDCF, Rufino JM, Santos RD, Banat VA, I.M. and Sarubbo LA (2016) Biosurfactants: promising molecules for petroleum biotechnology advances. *Frontiers in microbiology*, *7*, p.1718
19. Elraies KA, Tan IM (2012) The application of a new polymeric surfactant for chemical EOR. *Introduction to Enhanced Oil Recovery (EOR) Processes and Bioremediation of Oil-Contaminated Sites*
20. Fardeau ML, Salinas MB, l'Haridon S, Jeanthon C, Verhé F, Cayol JL, Patel BK, Garcia JL, Ollivier B (2004) Isolation from oil reservoirs of novel thermophilic anaerobes phylogenetically related to Thermoanaerobacter subterraneus: reassignment of T. subterraneus, Thermoanaerobacter yonseiensis, Thermoanaerobacter tengcongensis and Carboxydibrachium pacificum to Caldanaerobacter subterraneus gen. nov., sp. nov., comb. nov. as four novel subspecies. Int J Syst Evol Microbiol 54(2):467–474
21. Fukui T, Atomi H, Kanai T, Matsumi R, Fujiwara S, Imanaka T (2005) Complete genome sequence of the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1 and comparison with Pyrococcus genomes. Genome research 15(3):352–363
22. Gao CH, Zekri A (2011) Applications of microbial-enhanced oil recovery technology in the past decade. Energy Sour Part A Recover Utilization Environ Eff 33(10):972–989
23. Gardner AF, Kumar S, Perler FB (2012) Genome sequence of the model hyperthermophilic archaeon Thermococcus litoralis NS-C
24. Grassia GS, McLean KM, Glénat P, Bauld J, Sheehy AJ (1996) A systematic survey for thermophilic fermentative bacteria and archaea in high temperature petroleum reservoirs. FEMS Microbiol Ecol 21(1):47–58
25. Guo H, Li Y, Yiran Z, Wang F, Wang Y, Yu Z, Haicheng S, Yuanyuan G, Chuyi J, Xian G, 2015, August. Progress of microbial enhanced oil recovery in China. In *SPE Asia Pacific Enhanced Oil Recovery*

Conference. Society of Petroleum Engineers

26. Han P, Sun F, Shi M, 2001, January. Microbial EOR laboratory studies on the microorganisms using petroleum hydrocarbon as a sole carbon source. In *SPE Asia Pacific Improved Oil Recovery Conference*. Society of Petroleum Engineers
27. Hu P, Tom L, Singh A, Thomas BC, Baker BJ, Piceno YM, Andersen GL, Banfield JF (2016) Genome-resolved metagenomic analysis reveals roles for candidate phyla and other microbial community members in biogeochemical transformations in oil reservoirs. *MBio*, *7*(1)
28. Karatzas KA, Zervos A, Tassou CC, Mallidis CG, Humphrey TJ (2007) Piezotolerant small-colony variants with increased thermotolerance, antibiotic susceptibility, and low invasiveness in a clonal Staphylococcus aureus population. Appl Environ Microbiol 73(6):1873–1881
29. Kim DD, O'Farrell C, Toth CR, Montoya O, Gieg LM, Kwon TH, Yoon S (2018) Microbial community analyses of produced waters from high-temperature oil reservoirs reveal unexpected similarity between geographically distant oil reservoirs. Microbial biotechnology 11(4):788–796
30. Kraegeloh A, Amendt B, Kunte HJ (2005) Potassium transport in a halophilic member of the bacteria domain: identification and characterization of the K + uptake systems TrkH and TrkI from Halomonas elongata DSM 2581T. J Bacteriol 187(3):1036–1043
31. Lal B, Reddy MRV, Agnihotri A, Kumar A, Sarbhai MP, Singh N, Khurana RK, Khazanchi SK, Misra TR, Resources Institute and Institute of Reservoir Studies (2009) *Process for enhanced recovery of crude oil from oil wells using novel microbial consortium*. U.S. Patent 7,484,560
32. Lazar, I., Petrisor, I.G. and Yen, T.F., 2007. Microbial enhanced oil recovery (MEOR).*Petroleum Science and Technology*, *25*(11), pp.1353–1366
33. Lee HS, Kang SG, Bae SS, Lim JK, Cho Y, Kim YJ, Jeon JH, Cha SS, Kwon KK, Kim HT, Park CJ (2008) The complete genome sequence of Thermococcus onnurineus NA1 reveals a mixed heterotrophic and carboxydotrophic metabolism. J Bacteriol 190(22):7491–7499
34. Li H, Chen S, Mu BZ, Gu JD (2010) Molecular detection of anaerobic ammonium-oxidizing (anammox) bacteria in high-temperature petroleum reservoirs. Microbial ecology 60(4):771–783
35. Li H, Wang XL, Mu BZ, Gu JD, Liu YD, Lin KF, Lu SG, Lu Q, Li BZ, Li YY, Du XM (2013) Molecular detection, quantification and distribution of alkane-degrading bacteria in production water from low temperature oilfields. International Biodeterioration Biodegradation 76:49–57
36. Liu YF, Galzerani DD, Mbadinga SM, Zaramela LS, Gu JD, Mu BZ, Zengler K (2018) Metabolic capability and in situ activity of microorganisms in an oil reservoir. *Microbiome*, *6*(1), pp.1–12
37. Mardanov AV, Ravin NV, Svetlitchnyi VA, Beletsky AV, Miroshnichenko ML, Bonch-Osmolovskaya EA, Skryabin KG (2009) Metabolic versatility and indigenous origin of the archaeon Thermococcus sibiricus, isolated from a siberian oil reservoir, as revealed by genome analysis. Appl Environ Microbiol 75(13):4580–4588
38. Maudgalya S, Knapp RM, McInerney M, 2007, January. Microbially enhanced oil recovery technologies: a review of the past, present and future. In *Production and Operations Symposium*. Society of Petroleum Engineers

39. Naik SN, Goud VV, Rout PK, Dalai AK (2010) Production of first and second generation biofuels: a comprehensive review. Renewable sustainable energy reviews 14(2):578–597
40. Nikolova C, Gutierrez T (2020) Use of microorganisms in the recovery of oil from recalcitrant oil reservoirs: Current state of knowledge, technological advances and future perspectives. *Frontiers in microbiology*, *10*, p.2996
41. Orphan VJ, Taylor LT, Hafenbradl D, Delong EF (2000) Culture-dependent and culture-independent characterization of microbial assemblages associated with high-temperature petroleum reservoirs. Appl Environ Microbiol 66(2):700–711
42. Pawar ST, Bhosale AA, Gawade TB, Nale TR (2013) Isolation, screening and optimization of exopolysaccharide producing bacterium from saline soil. Journal of Microbiology Biotechnology Research 3(3):24–31
43. Rákhely G, Zhou ZH, Adams MW, Kovács KL (1999) Biochemical and molecular characterization of the [NiFe] hydrogenase from the hyperthermophilic archaeon, Thermococcus litoralis. European journal of biochemistry 266(3):1158–1165
44. Rathi R, Lavania M, Sawale M, Kukreti V, Kumar S, Lal B (2015) Stimulation of an indigenous thermophillic anaerobic bacterial consortium for enhanced oil recovery. RSC Adv 5(107):88115–88124
45. Safdel M, Anbaz MA, Daryasafar A, Jamialahmadi M (2017) Microbial enhanced oil recovery, a critical review on worldwide implemented field trials in different countries. Renew Sustain Energy Rev 74:159–172
46. Satpute SK, Banpurkar AG, Dhakephalkar PK, Banat IM, Chopade BA (2010) Methods for investigating biosurfactants and bioemulsifiers: a review. Crit Rev Biotechnol 30(2):127–144
47. Sen R (2008) Biotechnology in petroleum recovery: the microbial EOR. Progress in energy combustion Science 34(6):714–724
48. Shibulal B, Al-Bahry SN, Al-Wahaibi YM, Elshafie AE, Al-Bemani AS, Joshi SJ (2014) Microbial enhanced heavy oil recovery by the aid of inhabitant spore-forming bacteria: an insight review. *The Scientific World Journal*, *2014*
49. Simonato F, Campanaro S, Lauro FM, Vezzi A, D'Angelo M, Vitulo N, Valle G, Bartlett DH (2006) Piezophilic adaptation: a genomic point of view. Journal of biotechnology 126(1):11–25
50. Summit M, Baross JA (1998) Thermophilic subseafloor microorganisms from the 1996 North Gorda Ridge eruption. Deep Sea Res Part II 45(12):2751–2766
51. Svetlitchnyi V, Peschel C, Acker G, Meyer O (2001) Two membrane-associated NiFeS-carbon monoxide dehydrogenases from the anaerobic carbon-monoxide-utilizing eubacterium Carboxydothermus hydrogenoformans. Journal of Bacteriology, 183(17), p.5134
52. Sztukowska M, Bugno M, Potempa J, Travis J, Kurtz Jr DM (2002) Role of rubrerythrin in the oxidative stress response of Porphyromonas gingivalis. Mol Microbiol 44(2):479–488
53. Takahata Y, Hoaki T, Maruyama T (2001) Starvation survivability of Thermococcus strains isolated from Japanese oil reservoirs. Archives of microbiology 176(4):264–270

54. Takami H, Takaki Y, Chee GJ, Nishi S, Shimamura S, Suzuki H, Matsui S, Uchiyama I (2004) Thermoadaptation trait revealed by the genome sequence of thermophilic Geobacillus kaustophilus. Nucleic acids research 32(21):6292–6303
55. Teske A, Edgcomb V, Rivers AR, Thompson JR, de Vera Gomez A, Molyneaux SJ, Wirsen CO (2009) A molecular and physiological survey of a diverse collection of hydrothermal vent Thermococcus and Pyrococcus isolates. Extremophiles 13(6):905–915
56. Van Hamme JD, Singh A, Ward OP (2003) Recent advances in petroleum microbiology. Microbiology molecular biology reviews 67(4):503–549
57. Vannier P, Marteinsson VT, Fridjonsson OH, Oger P, Jebbar M (2011) Complete genome sequence of the hyperthermophilic, piezophilic, heterotrophic, and carboxydotrophic archaeon Thermococcus barophilus MP. J Bacteriol 193(6):1481–1482
58. Varjani SJ, Gnansounou E (2017) Microbial dynamics in petroleum oilfields and their relationship with physiological properties of petroleum oil reservoirs. Bioresour Technol 245:1258–1265
59. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S (2017) Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic acids research 45(D1):D535–D542
60. Wright ADG, Williams AJ, Winder B, Christophersen CT, Rodgers SL, Smith KD (2004) Molecular diversity of rumen methanogens from sheep in Western Australia. Appl Environ Microbiol 70(3):1263–1270
61. Yoon SH, Ha SM, Lim J, Kwon S, Chun J (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110(10):1281–1286
62. Youssef N, Simpson DR, McInerney MJ, Duncan KE (2013) In-situ lipopeptide biosurfactant production by Bacillus strains correlates with improved oil recovery in two oil wells approaching their economic limit of production. International Biodeterioration Biodegradation 81:127–132
63. Zahner, R.L., Govreau, B.R. and Sheehy, A., 2010, January. MEOR success in southern California. In *SPE Improved Oil Recovery Symposium*. Society of Petroleum Engineers
64. Zhang F, She YH, Chai LJ, Banat IM, Zhang XT, Shu FC, Wang ZL, Yu LJ, Hou DJ (2012) Microbial diversity in long-term water-flooded oil reservoirs with different in situ temperatures in China. Sci Rep 2(1):1–10
65. Zivanovic Y, Armengaud J, Lagorce A, Leplat C, Guérin P, Dutertre M, Anthouard V, Forterre P, Wincker P, Confalonieri F (2009) Genome analysis and genome-wide proteomics of Thermococcus gammatolerans, the most radioresistant organism known amongst the Archaea. Genome biology 10(6):1–23

Tables

Table 1
Metabolite production by strain 101C5 under wet lab conditions

Metabolites	At 96 ℃	At 101 ℃
H_2 (ml molasses g^{-1})	33.1 ± 0	0.14 ± 0
CO_2 (ml molasses g^{-1})	3.2 ± 0	43.8 ± 2.0
Organic acid (mg/L)	525.2 ± 1.34	1943.4 ± 0
Emulsification index %	52.3 ± 0.10	60.8 ± 0
Exopolysaccharide (mg/ml)	1.45 ± 0.35	3.5 ± 0.70

Table 2
Ability of 101C5 to mediate heavy oil recovery under simulated reservoir conditions

Oil recovery parameters	96 ℃		101 ℃	
	Control	Test	Control	Test
Volume of column (ml)	60	60	60	60
Volume of sand added (ml)	30	30	30	30
Total volume of oil added (ml)	14.8	14.8	15.8	19.2
Void volume (ml)	14.4	14.4	15.2	17.6
Primary recovery (using water) (ml)	1.6	0	1.7	0.7
Oil in place after primary recovery (ml)	12.8	14.4	13.5	16.9
Culture density of inoculum	–	2.0×10^7 cells/ml	–	1.6×10^7 cells/ml
Volume of residual medium/ water in the columns (ml)	30	30	30	30
Oil recovered after 14 days of incubation (ml)	0.6	6.9	1.6	9.7
% recovery of oil	5.4 ± 1.1	42.1 ± 8.2	11.8 ± 0	56.5 ± 1.06

Figures

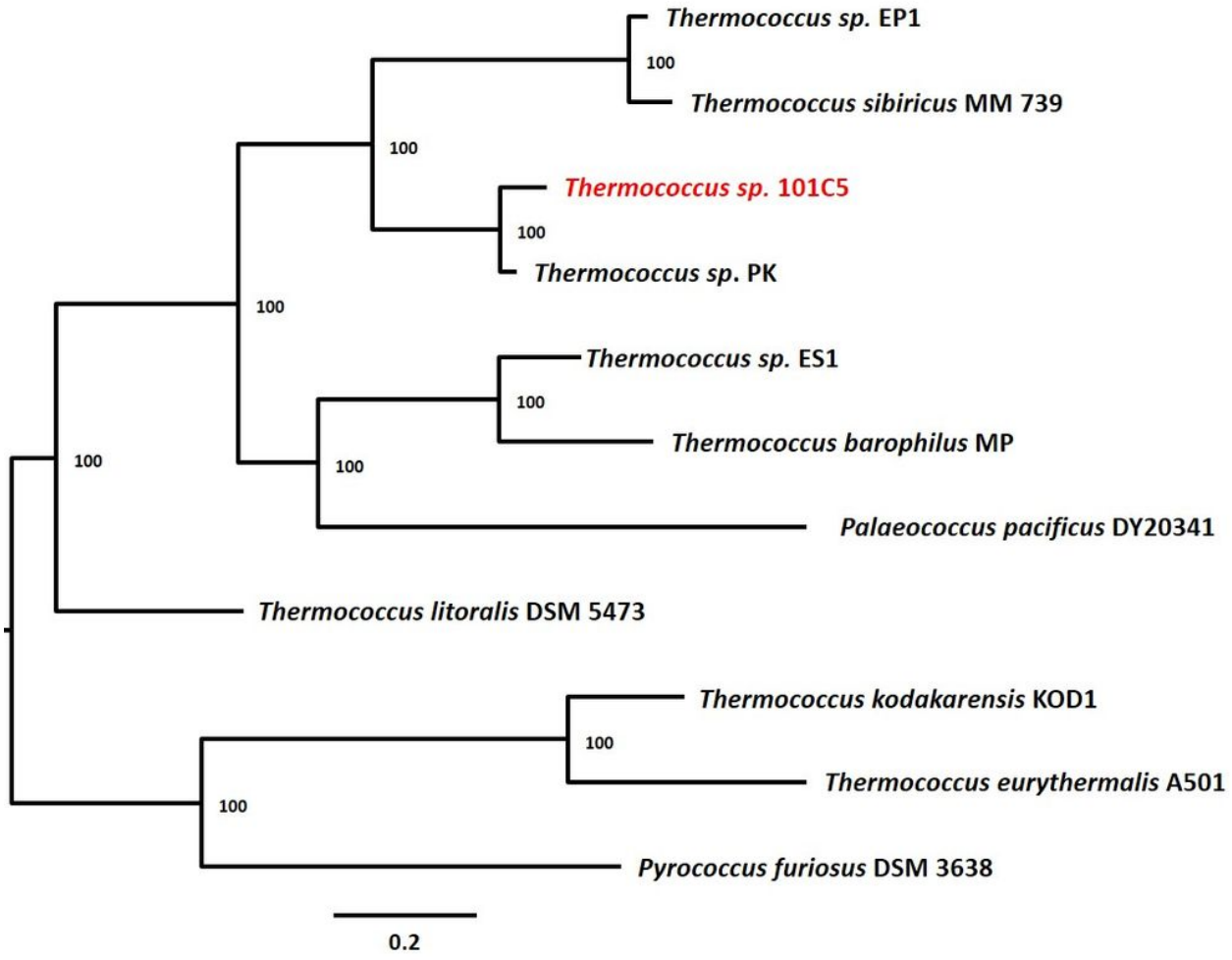

Figure 1

Phylogenomic tree showing the taxonomic relationship of strain 101C5 and closely related species

Supplementary Files

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

- Supplementarydoc.docx