Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Metagenomic analysis of Hot Springs of Surajkund and the exploration of thermostable enzymes from endogenous isolates

Snehi Soy
Birla Institute of Technology, Mesra
Usha Lakra
Birla Institute of Technology, Mesra
Pragya Prakash
Birla Institute of Technology, Mesra
Prashanth Suravajhala
Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham
Vinod Kumar Nigam
Birla Institute of Technology, Mesra
S. R. Sharma (srsharma@bitmesra.ac.in)
Birla Institute of Technology, Mesra

Article

Keywords: Thermophiles, Hot springs, metagenomics, amylase, lignocellulosic

Posted Date: March 10th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2589746/v1

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

Version of Record: A version of this preprint was published at Scientific Reports on August 30th, 2023.
See the published version at https://doi.org/10.1038/s41598-023-41515-5.

Abstract

Hot springs serve as a hub for microbial diversity, the microflora of the hot springs can be a source of numerous biologically important molecules. Extremophile bacteria were isolated from different hot springs of India to investigate their potential in industrial applications. The analysis involved both culture dependent and culture independent methods. The metagenomic analysis of the microflora was carried out using next-gen sequencing. Also, thermophilic isolates were culture and primarily screened for amylase, xylanase and cellulase enzyme production. The production for amylase was optimized and partial purification was carried out. Partially purified amylase was tested and characterized using; UPLC, DLS-ZP and TGA. The retention time for the amylase was observed to be around 0.5 minutes along with a stable thermo gravimetric profile around 400^{o}C confirming its stability at higher temperatures. One of the thermophilic isolate was able to degrade lignocellulosic waste exhibiting the potential of isolates from the kund.

Introduction

Extremophiles are the microbes that survive in diverse extreme conditions such as low and high temperature, pH, salinity, and radiation. Among extremophiles, thermophiles are heat- loving bacteria that survive at elevated temperatures. They are inhabitants of various niches such as hot springs, geothermal deposits and marine steam vents **[1, 2]**. Hot springs are unique natural environments for thermophilic microorganisms. Due to their scientific and biological significance, these thermal habitats and thermophiles have sparked interest in recent decades. In Jharkhand, there are various hot-water ecosystems, which were formed millions of years ago by water leaching from radioactive mines **[3]**. Surajkund is the hottest hot spring in the Hazaribagh area, with temperatures ranging from 85 to 90°C. This kund is well-known for its religious, medicinal, and aesthetic significance. The water from Suraj kund has a therapeutic effect due to its high sulfur content. Hot springs are a rich source of precious micro and macro elements, various compounds, and diverse microbial communities. The microbial community of the thermal springs is associated with the physicochemical parameters such as pH, warmth, micro and macro elements, oxido-reduction potential, and other chemical profiles **[4]**. Metagenomic exploration can reveal the molecular ecological traits as well as the huge and diversified gene pool of this natural ecosystem. However, hardly any studies have been done to investigate Surajkund's taxonomic diversity and metabolic features using whole-metagenome-based in-depth molecular profiling. The microbiota of this thermal habitat is underexplored at the genomic level. In previous studies, the surajkund sediment sample have shown the presence of novel bacterium JHK30^{T} identified as *Tepidiphilus thermophilus* and JS1^{T} identified as *Anoxybacillus* **[5, 6]**. Also, isolation of thermophilic bacteria at 60˚C and 70˚C with amylolytic, xylanolytic, and cellulolytic activity were reported from the same site. They were identified as *Geobacillus icigianus* and *Anoxybacillus gonensis* **[7]**. Till date, only novel bacteria or microbes with thermozyme secreting ability have been reported from this site. Yet Surajkund has not been targeted for the comprehensive metagenomic investigation. Once discovered, it may not only provide valuable information on ecological niches, but also allow for taxonomic binning and the evaluation of metabolic

and functional potential in the habitat. As well as evaluating the molecular basis of environmental sites with distinct features and their interaction. It also enables a better understanding of how thermophilic microorganisms that survive in severe natural environments have inherent traits such as high thermostability, resistance to organic solvents, and better self life.

Therefore, in this study, we examined the entire metagenomic sequences from six different kunds namely Suraj kund, Ram kund, Sita kund, Laxman kund, Brahma kund and Mixed kund of Jharkhand. Though these kunds are situated in close proximity, they show different temperatures and pH. The goal of this study is to look into diverse microflora, decipher metabolic system and profiling the function of bioresources in different kunds, which might result in metagenomic wealth for discovery of some novel microbes as well as microbes having industrial potential (bioactive compounds, enzymes, pigments, etc).

Experimental Methods

2.1 Sample collection, characterization and analysis of environmental factors

The sediment samples (water and mud) were collected from the six geographically distinct hot springs in the Surajkund dham of Jharkhand, India, viz. Surajkund hot spring, Ramkund hot spring, Laxmankund hot spring, Sitakund hot spring, Brahmakund hot spring and Mixed hot spring during November 2019 (Table 1).

The sample names were as follows Surajkund, Ramkund, Laxmankund, Sitakund, Brahmakund, and Mixed kund corresponding to six hot springs. The samples were collected in sterile thermo flasks bottles and transported to the laboratory with temperature control. Upon arrival at the laboratory, the samples were processed for bacterial isolations, whereas sample aliquots were tored at their respective kund temperature for tagenmic analysis. The pH and temperature of the kunds were recorded on the site of sampling and are tabulated **(Table 1)**.

2.2 DNA extraction for metagenomics study

DNA extraction was carried out by the method described by Gothwal et al. (2007) **[8]** with slight modification. 0.5 g of sample was processed in 1 mL TENS buffer for DNA extraction. The samples were vortexed to mix thoroughly. 900 mg of glass beads (mixture consisting of 0.1 mm, 0.5 mm and 1.0 mm diameter glass beads) were added to the sample vials. The samples were then subjected to hot SDS lysis by incubating the vials in a water bath (70°C) for 30 min with intermittent vortexing after every 5 min. The samples were then homogenized at maximum speed for 5 min. Post homogenization, the samples were subjected to liquid N_2 freeze thaw cycles thrice. The samples were then centrifuged at 10,000 rpm for 10 min at room temperature and the supernatant was collected. The pellet was washed with 750 µL TENS buffer, centrifuged and supernatant was collected. The supernatant was pooled and extracted with equal volume of PCI solution. The DNA precipitation was carried out by adding 0.1 volume of 5.0 M NaCl and 2 volumes of chilled ethanol. The pellet was washed with 70% ethanol and air dried. The dried pellet was

dissolved in sterile MQ water and stored at -80°C **[9]**. The extracted DNA was then sent for metagenomics study and the analysis was performed at Centyle Biotech Pvt. Ltd, University of Delhi, New Delhi.

2.2.1. Culture independent studies for metagenomic analysis

After the total community DNA was extracted from the sediments of six hot springs, the genomic DNA submitted for analysis was checked on 1% (w/v) Agarose gel using ETBR staining protocol. 2.0 µL of DNA samples were loaded in each well. DNA sample quantification was performed using Qubit HS DNA quantitation kit on Qubit fluorometer. When the quality control (QC) assessed was certified OK, the samples were processed for library construction. The sequencing library was prepared by random fragmentation of the DNA samples, followed by 5' and 3' adaptor ligation. Alternatively, 'tagmentation' combines fragmentation and ligation reactions into a single step that greatly increases the efficiency of the library preparation process. Adaptor ligated fragments were then PCR amplified and gel purified. The 16S rDNA V3-V4 region was amplified from the extracted DNA. PCR amplification was confirmed by agarose gel electrophoresis and the amplified products were processed for further library preparation. A second round of amplification was performed to add our proprietary index sequences to the amplified products. The indexing PCR assigns barcodes to individual samples. After the second PCR, amplified products were quantified using a Qubit fluorometer and then pooled together in an equimolar concentration. The pooled products represented the 16S Metagenomics library, which was subjected to NGS. The concentration of the Metagenomics library was estimated using an in-house qPCR. The Metagenomics library was sequenced on an Illumina sequencer in 250*2 paired end mode using Illumina's sequencing by-synthesis chemistry. For cluster generation, the library was loaded into a flow cell where fragments are captured on a lawn of surface-bound oligo complementary to the library adapters. Each fragment was then amplified into distinct clonal clusters through bridge amplification. When cluster generation was completed, the templates were ready for sequencing. Sequencing data was converted into raw data for analysis. The NGS data were generated in the form of standard fastq files. Data analysis was performed using analysis tools available on Illumina's cloud computing platform, Basespace. The quality of the data was ascertained using the FastQCtool. Accordingly, the files were quality and adapter trimmed. The Illumina 16S Metagenomics tool was used to analyze the trimmed file. The 16S Metagenomics tool performs taxonomic classification of 16S rRNA targeted amplicon reads using an Illumina-curated version of the Green Genes taxonomic database. The tool provides interactive visualizations and raw classification output, both for individual samples and as an aggregate analysis **[10, 11]**.

Downstream Metagenomic analysis

We employed AMBRICATE for verifying the contiguous sequences and AMPHORA2 workflow **[12]** for ascertaining phylogenetic analysis of metagenomics shotgun sequencing data and genomic data. As the tool uses bacterial and archaeal protein coding marker genes, it may be helpful in estimating unique taxonomic composition of bacterial and archaeal communities from metagenomics shotgun sequencing data.

2.3. Culture dependent studies

2.3.1. Isolation of thermotolerant and thermophilic bacteria

The soil, mud, and water samples of Surajkund and Mixed Kund were processed for the enumeration and isolation of culturable moderate thermophiles (50°C-60°C) and thermophiles (70°C). The study was conducted by standard serial dilution and spread plate methodology as described by Kumar et al. (2014) **[13]**. Two different isolation media i.e. Nutrient Agar and Tryptone Soybean Agar/ Soyabean Casein Digest Agar were employed for the isolation of thermophiles at three different high temperatures viz 50°C, 60°C, and 70°C. To avoid desiccation and breaking of agar plates, the plates were covered in sterile autoclavable bags and incubated at respective high temperatures. After 24 h incubation, the agar plates were observed for different morphotypes and phenotypic characteristics.

The suitable colonies were picked on the basis of morphological difference and were continuously streaked on the same isolation media to obtain a pure culture. The obtained pure colonies were stored at 4°C and were revived repeatedly to retain their reproducibility. The bacterial colonies were named as BITSNS, which refers to the initials of the institute's name and the name of researchers carrying out the research.

2.3.2. Primary screening of amylase, xylanase and, cellulase producing thermophilic bacterial isolates

Primary screening was carried out using plate assay method. Respective screening medium was prepared and autoclaved at 121°C and 15 psi for 15 min. The molten medium was poured in petri plates under aseptic conditions and pure bacterial colonies were streaked and incubated at their respective temperature (50°C, 60°C, and 70°C) for 24-48 h. Plates were observed for zones of hydrolysis.

Xylanolytic activity: Xylan Congo Red Agar Plate Assay was used for the identification of xylanase producing thermophilic isolates according to the methodology described **[14]**. The slight modification in the methodology was incorporation of congo red dye (0.01% w/v) in the medium. It eliminated the destaining step with 1.0 M NaCl solution and thus prevented leaching out of isolates from agar surface. The purified isolates were inoculated on xylanase screening media plates and were incubated at their respective temperature for 24-48 h for the xylanase secretion. The formation of a clear zone of xylan hydrolysis indicated xylanase activity and the xylanolytic index was determined for the positive isolates as per the methodology **[15]**.

Amylolytic activity: Starch Agar Plate Assay **[16]** was used for the screening of amylase secreting isolates. The starch agar plates were spot inoculated with the pure isolates and incubated for 24-72 h for amylase secretion. After sufficient growth on plates, the plate was flooded with Gram's Iodine and incubated for 10 min. The surplus Gram's Iodine solution was then discarded and halo zone formation was observed which indicated starch hydrolysis and the enzymatic index (EI) for amylase activity for positive isolates were further calculated.

Cellulolytic activity: The screening of cellulase producing bacteria was conducted as described by Kasana et al. (2008) **[17]**. The isolates were spot inoculated on Carboxymethylcellulose Sodium Salt (CMC) agar plate medium and incubated for 24-48 h. After suitable growth, cellulase enzyme expression was detected by the addition of 1 mL of Gram's Iodine solution to the CMC plate. The plate was incubated at room temperature for 30 minutes and excess iodine was discarded. The isolates forming clear halo zones on CMC agar plates were selected as cellulase positive strains and the cellulolytic index was henceforth determined.

2.4. Partial purification and molecular weight determination of amylase from Geobacillus icigianus (BITSNS038)

Thermophilic bacteria isolated at 70°C, identified as *Geobacillus icigianus* showed maximum production of amylase under SmF condition. Partial purification of thermostable amylase was carried out from amylase producing strain BITSNS038 and was selected for further research. The precipitation of total protein was carried out using ammonium sulfate precipitation (80-95% saturation) method. Ammonium sulfate precipitation involved the slow addition of fine ground ammonium sulfate to the supernatant (crude amylase). The solution was continuously stirred slowly for 1-2 h at 4°C. After the amount of required ammonium sulfate addition was done, the solution was kept overnight at 4°C. The solution thus obtained was centrifuged at 10000 rpm for 10 min, and the protein pellet obtained was dissolved in sodium phosphate buffer (pH 7.0). The resultant solution mixture was dialyzed against 10 kDa membrane **[18]**.

2.4.1. Ultrafiltration

The crude and concentrated amylase sample was filtered using 0.2 µm micropore syringe filter. It was further subjected to concentration using a centrifugation-based method. The membrane was rinsed by passing an ample amount of deionized water prior to ultrafiltration. Partial purification of amylase was carried out using buffer exchange with a membrane of 10 kDa (Ultra-filtration). The calibration of flow rate of retentate and permeate was also optimized using de-ionized water. The concentration was performed using a centrifugal concentrator of 10 kDa. The centrifugal concentrators were cleaned using 0.5 N NaOH and stored in 50% ethanol. At every step of the concentration enzyme assay and protein estimation was performed for the estimation of amylase using Nelson-Somogyi method and Bradford method respectively.

2.5. Characterization of partially purified amylase

Characterization of partially purified amylase was performed using Ultra Pressure Liquid Chromatography (UPLC), DynamicLight Scattering (DLS), Zeta Potential (ZP) Thermogravimetric Analysis (TGA).

2.5.1. UPLC analysis of partially purified amylase

UPLC analysis of partially purified amylase along with the commercial amylase standard was done. The standards and sample were prepared by dissolving in de-ionized water with 0.1% TFA. The analysis was done at the 214 nm wavelength of the PDA detector **[19]**.

2.5.2. DLS-ZP analysis of partially purified amylase

DLS-ZP analysis was performed in order to study the dispersion of partially purified amylase in suspension. The suspension was prepared by separately mixing partially purified enzymes in de-ionized water (5 times dilution). The analysis was performed according to modified procedures from Jachimska and co-workers **[20]**.

ZP analysis is used to determine the charge differences of the suspended particles in liquid. A value higher than -15mV signifies formation of colloidal particles i.e., accumulation of dissolved particles. In case of amylase, partially purified sample analysis was performed with respect to standard sample using ZP analyzer from Malvern Instruments, UK **[21, 22]**.

2.5.3. TGA analysis of partially purified amylase

TGA of amylase was performed in order to study the thermal and oxidative stability of partially purified protein with respect to standard. The analysis was performed using the TGA setup with a nitrogen analyzer **[23]**. The maximum temperature used for the analysis was 900°C. Temporal variations in the enthalpy during the phase shifts were studied.

2.6. Application of G. icigianus on different lignocellulosic agricultural wastes for the production of amylase

The bacteria were grown in amylase production medium supplemented with different lignocellulosic biomass (rice husk, rice straw and wheat bran) as carbon sources without any addition of starch. The composition of the amylase production medium was as follows: Yeast extract 3.0 g/L, Tryptone 3.0 g/L, $MgSO_4.7H_2O$ 0.2 g/L, K_2HPO_4 1.0 g/L, NaCl 1.0 g/L, pH 7.0 and incubated at optimal conditions (70°C, 150 rpm, 24 h) for amylase production. Later, the concentrations (5-12 g/L) of these agricultural substrates were also optimized for amylase production.

Results And Discussion

3.1. Primary screening of amylolytic, xylanolytic and cellulolytic thermophilic bacteria

The enzyme secreting ability of 41 isolates were checked via primary screening approach on agar plate consisting of specific substrate inducer i.e starch for amylolytic activity, xylan for xylanolytic and CMC for cellulase secreting ability. Out of 41 isolates, 23 potential enzyme secreting isolates were obtained i.e., 4 isolates of 50°C, 15 isolates of 60°C and 4 isolates of 70°C having amylolytic, xylanolytic and cellulolytic enzymes secreting ability and the plate assay results are shown in the figure (Fig 1and Fig 2).

The parameter for calculating the enzyme activity of isolates was the measurement of the enzymatic index (EI) by the given equation (1) **[15]**.

The enzymatic index was also calculated using formula $$EI = \frac{\text{diameter of hydrolysis zone}}{\text{diameter of colony}} \quad (1)$$

The isolates with hydrolysis zone 1.0 cm are considered significant **[24]**. The greater the halo zone, the higher is the EI value. Thus, higher value of EI obtained in primary screening results reflects the amylolytic, xylanolytic and cellulolytic potential of thermophiles.

3.2. Metagenomic Sequencing of the isolated DNA

More than 1.9 million high quality reads were obtained while the metagenomic sequencing of the isolated DNA with more than 0.3 million reads and more than 55% GC content on an average (Fig. 3).

16S amplicon library was constructed for NGS analysis. The analysis was carried out at different levels of Phylum, Class, Order, Family and Genus. The genuses of the bacterial diversity were confirmed using 16S rRNA technique, Brahmakund consisted of ten genera of bacteria, mainly dominated by *Traponema* with 19%, presence of *Traponema* indicates that the microbial population in the Brahmakund is dominated by the bacterial species which prove to be infectious to the human population (Fig.4a). While the Lakshman kund was well dominated by Lactobacillus (78%) (Fig.4b). The genuses *Rhizobium, Methylobacterium* share the population with 1% of the contribution each. The temperature difference of the two kunds were more than 10^{o}C. The Brahma kund temperature of 55°C indicates the favourable environment for the prevailing of infectious microbes. The 16S rRNA sequencing of Ram and Sita Kund was also performed, while Ram kund sequencing revealed that there is no specific bacterial population that is dominating the kund but it indicated the presence of radiation resistant bacteria *Rubrobacter* with 13% of the population (Fig. 4c). The genomic analysis of Sita kund also revealed the significant presence of *Rubrobacter* genus (Fig.4d).

However, the population of the infectious bacteria were found the most in the samples collected from Surajkund (Fig.4e) and samples collected from mixed kund (Fig.4f). While Surajkund samples contained *Sphingomons, Prevotella* and *Finegoldia*, the mixed kund contained *Gemmata, Sorangium* and *Fusobacterium*. *Sphingomonas* is known to be a very opportunistic bacteria and is well found in soil, water even in hospital wastes **[25]**.

Prevotella is prevalent in oral, gut and vaginal microbiota **[26]**, while *Finegoldia* is known for serious skin infections **[27]**. In case of Gemmata bacteria it is proven that they are naturally antibiotic resistant and also cause skin diseases, like *Sphingomonas* they are also known as opportunistic pathogens and are also found in hospital wastes **[25]**. *Sarangium* is not directly infectious to humans, but is known to produce EPS which cause formation of biofilms, which can ultimately cause infection to humans **[28]**. *Fusobacterium* is known to cause acute otitis in children under two years of age. If untreated it can lead

to mastoiditis and more severe symptoms of bacteremia, osteomyelitis, and potentially Lemierre syndrome and septic shock **[29]**. Apart from these finding some population of the disease-causing bacteria was found in all the kunds, the reason behind this can be the human interference that has increased with time causing the pollution of the soil and water with different kinds of wastes. However, the kunds with high temperature initially give an impression that they may not prove favourable for the presence of infectious agents growing at and around in the 37^{o}C. However, the findings in the present report suggest that the change in the micro-environment due to the pollutants and the presence of the bacteria for a very long time has made them thermotolerant and further studies may lead to identification new species or strains that are thermotolerant or thermophilic in nature. The taxonomic analysis in detail confirmed the presence of the bacterial population at phylum, order, family, class and genus levels. At smaller levels, Laxman kund, Suraj kund and Sita kund showed similar profiles at all the levels. The core microbiome analysis was also performed which represented the relative abundance of the most prevalent bacterial species (Fig.5d).

While the taxonomy plot analysis confirmed the abundance of *Lactobicillus* in the microbial diversity; the most abundant population was confirmed as; Phylum: *Firmicutes*; Class: *Bacilli*; Order: *Lactobacillales*; Family: *Bacillaceae*; Genus: *Lactobacillus* (Fig.5a- 5c).

Reports suggested presence of *Protebacteria, Actinobacteria* and *Firmicutes* in Yumthang hot spring of North Sikkim with 54.3%, 32.2% and 6.3% respectively, which is an indicator of presence of gram-negative bacteria in hot springs. The results also suggest that hot springs may harbour disease causing bacteria as most of the infectious bacteria are of gram-negative nature. *Fermecutes*have also been found abundantly in the samples analysed from the kunds of Jharkhand, India that has been reported in the present work. The dominance in the kunds of Jharkhand has been of *Lactobacillus* genus. The work from **[30]** also reports the abundance of Thermophilic actinobacteria which hold industrial importance as a source of various enzymes such as pullulanases, amylase, DNA polymerases. The present work also reports the presence of industrially important thermostable enzymes producers such as xylanases, amylases and cellulases. Proteobacteria and Firmicutes were reported in a major proportion of the microbial population by Samarasinghe et al. (2021) **[31]**. They reported the metagenomic analysis of Sri-Lankan Geothermal spring with culture dependent and independent approaches. The relative abundance was found to be maximum for *Enterobacteriales* order. They had reported the presence of microbiota in surface water and in-depth water samples, which is lacking in the present report. While the most population of microbes of at surface and in-depth water samples were different, *Klebsiella* sp. and *Pannonibacter* sp. were found to be common **[31]**. Anoxygenic photosynthesis performing bacteria were reported to be found in the Rhodobacteraceae family. The isolate was reported from a terrestrial hot spring in Japan. The bacteria were found to be chymotrypsin, alpha and beta-galactosidase, trypsin and other industrially important enzymes. As reported in other reports the isolate was found to be gram negative and of *Alphaproteo-bacteria* class **[32].** The maximum abundance was found for *Lactobacillus* and the minimum was found for *Cellulomonas* as confirmed with the 16S data as well. However, in the case of prevalence the same was not true and the most opportunistic bacteria *Sphingomonas* was found

to be the most prevalent, while the minimum was found to be *Nitrospira* (Fig. 5d). By combining the contigs, we ran a simple matrix of gene presence/absence using Galaxy Version 1.0.1 and distinct clades (Fig 5e). On a further note when a search was made using the AMPHORA2 workflow, no distinct genes were found to be significant implying that the kund harbors unique alpha diversity.

3.3. Characterization of partially purified amylase

The amylase was partially purified having molecular weight between 45-63 kDa **[33]**. Partially purified amylase characterization was done using UPLC, DLS-ZP, and TGA.

3.3.1. UPLC analysis of partially purified amylase

UPLC analysis for the presence of amylase enzymes in crude and partially purified fraction (Fig.S1a and S1b) was performed. A calibration curve of 12.5 µg/mL to 100 µg/mL was plotted and the concentration of amylase in the precipitated and subsequently filtered ultrafiltration fraction was found to be more than 1 mg/mL. The peak of amylase was observed to be at 0.55 min. The presence of amylase was confirmed in partially purified samples by the retention time in the standard samples and the partially purified amylase enzyme. The analysis was performed only to confirm the presence of amylase enzymes.

3.3.2. DLS-ZP analysis of partially purified amylase

DLS analysis of the partially purified amylase was performed in comparison with the standard drug. It is evident from the figure (Fig. 6a and Fig. 6b).

The Z average was found to be less than 200 nm in the standard sample which is far less than the average particle size of partially purified amylase of around 4000 nm. However, in comparison to the crude amylase the standard indicated higher intensity of light scattering which can be attributed to the fact that purified enzyme has reduced amount of amylase leading to the observation that standard amylase has better solubility. However, less than 100 nm of particle size was reported for immobilized and free amylase by Razi Ahmad in 2013 **[34]**. Even smaller size of amylase segregated particles was reported (4 nm in size) by Seyed and co-workers **[35]**. Although the agglomerated particles had size greater than 300 nm. The particle size of less than 100 nm was also reported by Jongchan and researchers **[36]**. The reduced size could be achieved in case of completely purified alpha-amylase.

The increase in size suggests the formation of aggregates which may be due to the partially pure nature of the enzyme. In case of zeta a potential of -12.1 mV was observed, however in comparison with the standard it was on the lower side i.e. -21 mV. However, there was non-uniformity in the counts as observed in the peaks of standard (Fig. S2a) and partially purified sample (Fig. S2b) indicating the nature of the enzyme to be partially purified. Weber et al. determined their DLS and ZP. The particle size was determined to be around 300-400 nm when the HSA (Human Serum Albumin) was stabilized using glutaraldehyde. The zeta potential measurements of the same were found to be between −17mV to −25mV **[37]** .

3.3.3. Thermogravimetric analysis

TGA analysis of concentrated and partially purified amylase was performed in comparison with the commercial amylase standard. A temperature range of 15°C to 565°C was kept for the standard amylase. A gradual loss was observed in the initial phase of 15-125°C with a loss of just 3.65%. However, a rapid weight loss was observed in the next two cycles of 125-350°C and 350-565°C, leading to a weight loss of 99.72% (Fig.S3a).

In case of a concentrated sample, a temperature range of 30-700°C, indicating a better stability of the concentrated enzyme. However, it followed the same pattern as in case of standard amylase, in terms of weight loss. 97% weight loss was observed indicating reduction in moisture content of the sample and confirming the concentration (Fig. S3b). The partially purified sample followed the similar pattern of weight loss when compared with the concentrated sample. However, a higher weight loss of 99.3% indicated the presence of moisture in comparison to the contaminating proteins that could have been present in the concentrated amylase (Fig. S3c). Sochava and co-workers **[38]**reported in a DSC-TGA analysis of BSA (Bovine Serum Albumin), the maximum loss of protein occurred at a very reduced temperature of under 100°C compared to the present work. However, in the case of *Dunaliella tertiolecta* the temperature range was broader i.e., 170 to 900°C which indicated better stability of proteins extracted from microalgae **[39]**. Ricci and co-workers reported the DSC and TGA analysis of proteins purified from different legumes. The data was reported in comparison of different purity grades. It indicated that partially purified (medium purity) enzymes are more stable at higher temperatures compared to the completely purified enzymes indicative of the need of micro-environments for better stability, however this was not true for the protein isolated from all the sources. Total protein isolated from beans had nearly no effect on the temperature stability while better stability in medium purified samples was found to be in case of protein isolated from lentils. In case of peas the stability was found to be more in 68% purified sample compared to 77% purified sample **[40].**

3.3.4. Application of G. icigianus on different lignocellulosic agricultural wastes for the production of amylase

The use of low-cost agriculture wastes as substrates for the production of industrial enzymes is an economical and significant way to reduce cost of the overall process. *G. icigianus* (strain BITSNS038) was able to utilize a variety of inexpensive substrates such as wheat bran, rice husk, and rice straw for production of amylase (Fig. S4).

Interestingly, all lignocellulosic substrates supported amylase production by *G. icigianus*. The amylase activity was maximum for rice husk (77.5 U/mL at 10 g/L) followed by wheat bran (55.27 U/mL at 10g/L). Minimum activity was observed for rice straw with 1.52 U/mL (result not shown). The low activity compared to starch is due to presence of cellulose, hemicellulose, lignin and silica etc present in the lignocellulosic substrates. *Geobacillus* sp. have versatile catabolic activity, rapid growth rates, and are known for the degradation of hemicelluloses and starch **[41]**. These intriguing characteristics make it a potential candidate in second-generation (lignocellulosic) biorefiniries for biofuel production. Agricultural

wastes have a significant amount of lignocellulosic carbohydrate fraction that can be hydrolyzed and fermented for the bio-fuel production **[42]**. The maximum amylase production from rice husk and wheat bran is attributed to their rich profile which consists of high carbon, nitrogen, cellulose and hemicelluloses. Rice straw on the other hand has some challenges such as high inorganic composition and C/N ratio that affects its biodegrability **[43]**.

Conclusions

Hot springs are a unique niche supporting proliferation of diverse thermophilic microflora that are sources of thermostable enzymes of industrial importance. Our previous studies have proved that *Anoxybacillus gonensis* and *Geobacillus icigianus* have thermostable enzyme secretion ability and they were isolated from the hot spring of Surajkund. In particular, partially purified amylase from *G. icigianus* has been characterized and has further reflected on thermophilic aspects of the amylase. Additionally, hydrolytic enzymes secreting capacity of 50°C and 60°C isolates have exhibited several isolates that will possibly help in finding other significant strains. Metagenomics study to explore the diversity of microflora inhabiting Surajkund as well as the surrounding kunds having variable temperature reflected on one important aspect that anthropogenic activities have significantly polluted pristine water bodies. The prevalence of *Treponema, Prevotella, Finogoldia, Sphingomonas* raise the concerns of these common pathogens becoming tolerant to high temperature but at the same time presence of *Anoxybacillus* genus, and radiation resistant bacteria *Rubrobacter* reveals the potential of the hot springs as reservoir of industrially important microbes.

Declarations

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data supporting the findings of this study are available in NCBI- GenBank with BioProject ID PRJNA940280.

Acknowledgement

S Soy is supported by the DBT-JRF fellowship sponsored by the Department of Biotechnology, Government of India. The Grant No. is (DBT/2017/IBIT-R/771). U Lakra is financially supported by NFST (National Fellowship for Schedule Tribe for Higher Education) funded by the ministry of tribal affairs.

The authors are also sincerely thankful for the technical assistance of Central Instrumentation Facility, Birla Institute of Technology, Mesra and Centyle Biotech, Uttarakhand for the metagenomic analysis.

CRediT authorship contribution statement

SS: Conceptualization, Methodology, Investigation, Writing-original draft; **UL:** Investigation, Writing-introduction; **PP:** Conceptualization, Writing-original draft, review and editing; **PS:** Analysis, review and editing; **VKN:** Project administration, Supervision; **SRS:** Project administration, Supervision.

References

1. Rothschild, L. J. & Mancinelli, R. L. Life in extreme environments. *Nature*. **409**, 1092-1101; https://doi.org/10.1038/35059215(2001).
2. Pathak, A. P. & Rathod, M. G. Cultivable bacterial diversity of terrestrial thermal spring of Unkeshwar, India. *J. Biochem. Technol* .**5**, 814-818. (2014).
3. De Mandal, S. & Bhatt, P. (Eds.) Recent advancements in microbial diversity. Academic press. (2020).
4. Siering, P. L., Clarke, J. M. & Wilson, M. S. Geochemical and biological diversity of acidic, hot springs in Lassen Volcanic National Park. *Geomicrobiol. J.***23**, 129-141; https://doi.org/10.1080/01490450500533916(2006).
5. Podda, A., Lepcha, R.T. & Das, S.K. Taxonomic study of the genus *Tepidiphilus*: transfer of *Petrobacter succinatimandens* to the genus *Tepidiphilus* as *Tepidiphilus succinatimandens* comb. nov., emended description of the genus *Tepidiphilus* and description of *Tepidiphilus thermophilus* sp. nov., isolated from a terrestrial hot spring. *Int. J. Syst. Evol. Microbiol.***64**, 228-235; doi: 10.1099/ijs.0.056424-0 (2014).
6. Deep, K., Poddar, A. & Das, S. K. *Anoxybacillus suryakundensis* sp. nov, a moderately thermophilic, alkalitolerant bacterium isolated from hot spring at Jharkhand, India. *PloS one*. **8**, e85493; https://doi.org/10.1371/journal.pone.0085493(2013).
7. Soy, S., Nigam, V.K. & Sharma S.R. Cellulolytic, amylolytic and xylanolytic potential of thermophilic isolates of Surajkund hot spring. *J. Biosci.***44**, 1-12. https://doi.org/10.1007/s12038-019-9938-7 (2019).
8. Gothwal, R.K., Nigam, V.K., Mohan, M.K., Sasmal, D. & Ghosh, P. Extraction of bulk DNA from Thar Desert soils for optimization of PCR-DGGE based microbial community analysis. *Electron. J. Biotechnol.***10**, 400-408; http://dx.doi.org/10.4067/S0717-34582007000300007 (2007).
9. Paduano, S. *et al.* Microbial biodiversity of thermal water and mud in an Italian spa by metagenomics: A pilot study. *Water Sci Technol: Water Supply*. **18**, 1456-1465; https://doi.org/10.2166/ws.2017.209 (2018).
10. Modi, A., Vai, S., Caramelli, D. & Lari, M. The Illumina sequencing protocol and the NovaSeq 6000 system. In Bacterial Pangenomics. Humana, New York, NY, pp.15-42. doi: 10.1007/978-1-0716-1099-2_2 (2020).

11. Saurabh, S., Prasad, D., Masi, A. & Vidyarthi A. S. Next generation sequencing and transcriptome analysis for identification of ARF and Aux/IAA in pointed gourd (*Trichosanthes dioica* Roxb.), a non-model plant. *Sci. Hortic.***301**, 111152; https://doi.org/10.1016/j.scienta.2022.111152 (2022).
12. Kumar, M., Yadav, A. N., Tiwari, R., Prasanna, R. & Saxena, A.K. Deciphering the diversity of culturable thermotolerant bacteria from Manikaran hot springs. *Ann .Microbiol*. **64**,741-751; https://doi.org/10.1007/s13213-013-0709-7 (2014).
13. Kerepesi, C., Bánky, D. & Grolmusz, V. AmphoraNet: the webserver implementation of the AMPHORA2 metagenomic workflow suite. *Gene.***533**, 538-40; doi: 10.1016/j.gene.2013.10.015 (2014).
14. Nagar, S., Mittal, A. & Gupta, V. K. A cost effective method for screening and isolation of xylan degrading bacteria using agro waste material. *Asian J. Biol. Sci.***5**, 384-394. doi: 10.3923/ajbs.2012.384.394 (2012).
15. Florencio, C., Couri, S. & Farinas, C.S. Correlation between agar plate screening and solid-state fermentation for the prediction of cellulase production by *Trichoderma* strains. *Enzyme Res.* 2012. Article ID 793708. https://doi.org/10.1155/2012/793708 (2012).
16. Vaikundamoorthy, R., Rajendran, R., Selvaraju, A., Moorthy, K. & Perumal, S. Development of thermostable amylase enzyme from *Bacillus cereus* for potential antibiofilm activity. *Bioorg. Chem.* **77**, 494-506. doi: 10.1016/j.bioorg.2018.02.014. (2018).
17. Kasana, R. C., Salwan, R., Dhar, H., Dutt, S. & Gulati, A. A rapid and easy method for the detection of microbial cellulases on agar plates using Gram's iodine. *Curr. Microbiol* . **57**, 503-507; doi: 10.1007/s00284-008-9276-8 (2008).
18. Anand, J., Ramamoorthy, K., Muthukumar, G. & Nagaraj, S. Production and partial purification α-amylase producing *Streptomyces* sp. SNAJSM6 isolated from *Sargassum myriosystum.J. Agardh Indian Journal of Geo Marine Sciences*. **48**, 1245 – 1251; http://nopr.niscair.res.in/handle/123456789/49705 (2019).
19. Das, K., Doley, R. & Mukherjee, A. K. Purification and biochemical characterization of a thermostable, alkaliphilic, extracellular α-amylase from *Bacillus subtilis* DM-03, a strain isolated from the traditional fermented food of India. *Biotechnol. Appl. Biochem.* **40**, 291-298; doi: 10.1042/BA20040034 (2004).
20. Jachimska, B., Wasilewska, M. & Adamczyk, Z. Characterization of globular protein solutions by dynamic light scattering, electrophoretic mobility, and viscosity measurements. *Langmuir*, **24**, 6866-6872; https://doi.org/10.1021/la800548p (2008).
21. Teodor, E., Litescu, S.C., Lazar, V. & Somoghi, R. Hydrogel-magnetic nanoparticles with immobilized L-asparaginase for biomedical applications. *J .Mater .Sci.: Mater. Med.* **20**, 1307-14; doi: 10.1007/s10856-008-3684-y (2009).
22. Ashrafi, H. *et al.* Nanostructure L-asparaginase-fatty acid bioconjugate: synthesis, preformulation study and biological assessment. *Int. J. Biol .Macromol.* **62**,180-7; doi: 10.1016/j.ijbiomac.2013.08.028 (2013).
23. Zhang, H., Takenaka, M. & Isobe, S. DSC and electrophoretic studies on soymilk protein denaturation. *J Therm Anal Calorim.***75**, 719-26; https://doi.org/10.1023/B:JTAN.0000027168.18317.78 (2004).

24. Gaur, R. & Tiwari, S. Isolation, production, purification and characterization of an organic-solvent-thermostable alkalophilic cellulase from *Bacillus vallism*. *ortis* RG-07. *BMC Biotechnol.***15**, 1-12; https://doi.org/10.1186/s12896-015-0129-9 (2015).
25. Ryan, M. P. & Adley, C.C. *Sphingomonas paucimobilis*: a persistent Gram-negative nosocomial infectious organism. *J. Hosp. Infect.***75**, 153-157; doi: 10.1007/s00284-021-02608-4 (2010).
26. Iljazovic, A. *et al.* Perturbation of the gut microbiome by *Prevotella* spp. enhances host susceptibility to mucosal inflammation. *Mucosal immunol*. **14**, 113-124; https://doi.org/10.1038/s41385-020-0296-4 (2021).
27. Neumann, A., Björck, L. & Frick, I.M. *Finegoldia magna*, an anaerobic Gram-positive bacterium of the normal human microbiota, induces inflammation by activating neutrophils. *Front .Microbiol.***65**, https://doi: 10.3389/fmicb.2020.000 (2020).
28. Nkosi, N. Characterisation of Extracellular Polymeric Substances (EPS) from Bacteria Isolated from Acid Mine Decant. University of Johannesburg (South Africa). (Masters Dessertation Thesi**s)** https://hdl.handle.net/102000/0002 (Accessed: 22 August 2017). (2020).
29. Thurnheer, T., Karygianni, L., Flury, M. & Belibasakis, G.N. *Fusobacterium* species and subspecies differentially affect the composition and architecture of supra-and subgingival biofilms models. *Front Microbiol* **10**, 1716; doi: 10.3389/fmicb.2019.01716 (2019)
30. Panda, A.K. *et al*. Bacterial diversity analysis of Yumthang hot spring, North Sikkim, India by Illumina sequencing. *Big Data Anal***2**, 1-7; https://doi.org/10.1186/s41044-017-0022-8 (2017).
31. Samarasinghe, S.N., Wanigatunge, R.P. & Magana-Arachchi, D.N. Bacterial Diversity in a Sri Lankan Geothermal Spring Assessed by Culture-Dependent and Culture-Independent Approaches. *Curr. Microbiol*. **78**, 3439-3452; doi: 10.1007/s00284-021-02608-4 (2021).
32. Muramatsu, S. *et al*. *Neotabrizicola shimadae* gen. nov., sp. nov., an aerobic anoxygenic phototrophic bacterium harbouring photosynthetic genes in the family *Rhodobacteraceae*, isolated from a terrestrial hot spring. *Antonie van Leeuwenhoek*. **115**, 731-740; doi: 10.1007/s10482-022-01728-665 (2022).
33. Soy, S., Nigam, V.K., & Sharma, S.R. Enhanced production and biochemical characterization of a thermostable amylase from thermophilic bacterium *Geobacillus icigianus* BITSNS038. *J. Taibah Univ. Sci.***15**, 730-745; https://doi.org/10.1080/16583655.2021.2002549(2021).
34. Ahmad, R., Khatoon, N. & Sardar, M. Biosynthesis, characterization and application of TiO_2 nanoparticles in biocatalysis and protein folding. *J. Proteins Proteomics*. **4**,115–121(2013).
35. Dehnavi, S.M., Pazuki, G. & Vossoughi, M. PEGylated silica-enzyme nanoconjugates: a new frontier in large scale separation of α-amylase. *Sci. Rep.***5**, 1-9; https://doi.org/10.1038/srep18221(2015).
36. Lee, J. *et al. Bacillus licheniformis* α-amylase: Structural feature in a biomimetic solution and structural changes in extrinsic conditions. *Int. J. Biol. Macromol*. **127**, 286-296; https://doi.org/10.1016/j.ijbiomac.2019.01.053 (2019).
37. Weber, C., Kreuter, J. & Langer K. Desolvation process and surface characteristics of HSA-nanoparticles. *Int. J. Pharm*. **196**, 197–200; doi: 10.1016/S0378-5173(99)00420-2. (2000).

38. Sochava, I.V., Belopolskaya, T.V. & Smirnova, O.I. DSC study of reversible and irreversible thermal denaturation of concentrated globular protein solutions. *Biophys. Chem.***22**, 323–336; https://doi.org/10.1016/0301-4622(85)80056-9 (1985).
39. Shuping, Z., Yulong, W., Mingde, Y., Chun, L. & Junmao, T. Pyrolysis characteristics and kinetics of the marine microalgae *Dunaliella tertiolecta* using thermogravimetric analyzer. *Bioresour. technol.***101**, 359-365; https://doi.org/10.1016/j.biortech.2009.08.020 (2010).
40. Ricci, L. *et al*. On the thermal behavior of protein isolated from different legumes investigated by DSC and TGA. *J. Sci. Food. and Agric*. **98**, 5368-5377; doi: 10.1002/jsfa.9078. (2018).
41. Hussein, E.I. *et al*. Exploring the microbial diversity in Jordanian hot springs by comparative metagenomic analysis. *Microbiology Open*. **6**, e00521; doi: 10.1002/mbo3.521(2017).
42. Turner. P., Mamo, G. & Karlsson, E.N. Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microb. Cell Factories*. **6**, 1-23; https://doi.org/10.1186/1475-2859-6-9 (2007).
43. Bakker, R.R.C., Elbersen, H.W., Poppens, R.P. & Lesschen, J.P. Rice straw and wheat straw-potential feedstocks for the biobased economy. NL Agency (2013).

Table 1

Table 1. Sampling study of hot springs of Surajkund

Sl No.	Name of Kunds	Temperature (°C)	pH	Sample
1.	Suraj kund	90	7.5	mud+water
2.	Laxman	60	8	mud+water
3.	Brahma	50	8	microbial mats+ mud+water
4.	Ram	80	9	mud+water
5.	Sita	35	7	mud+water
6.	Mixed	85	8	mud+water

Figures

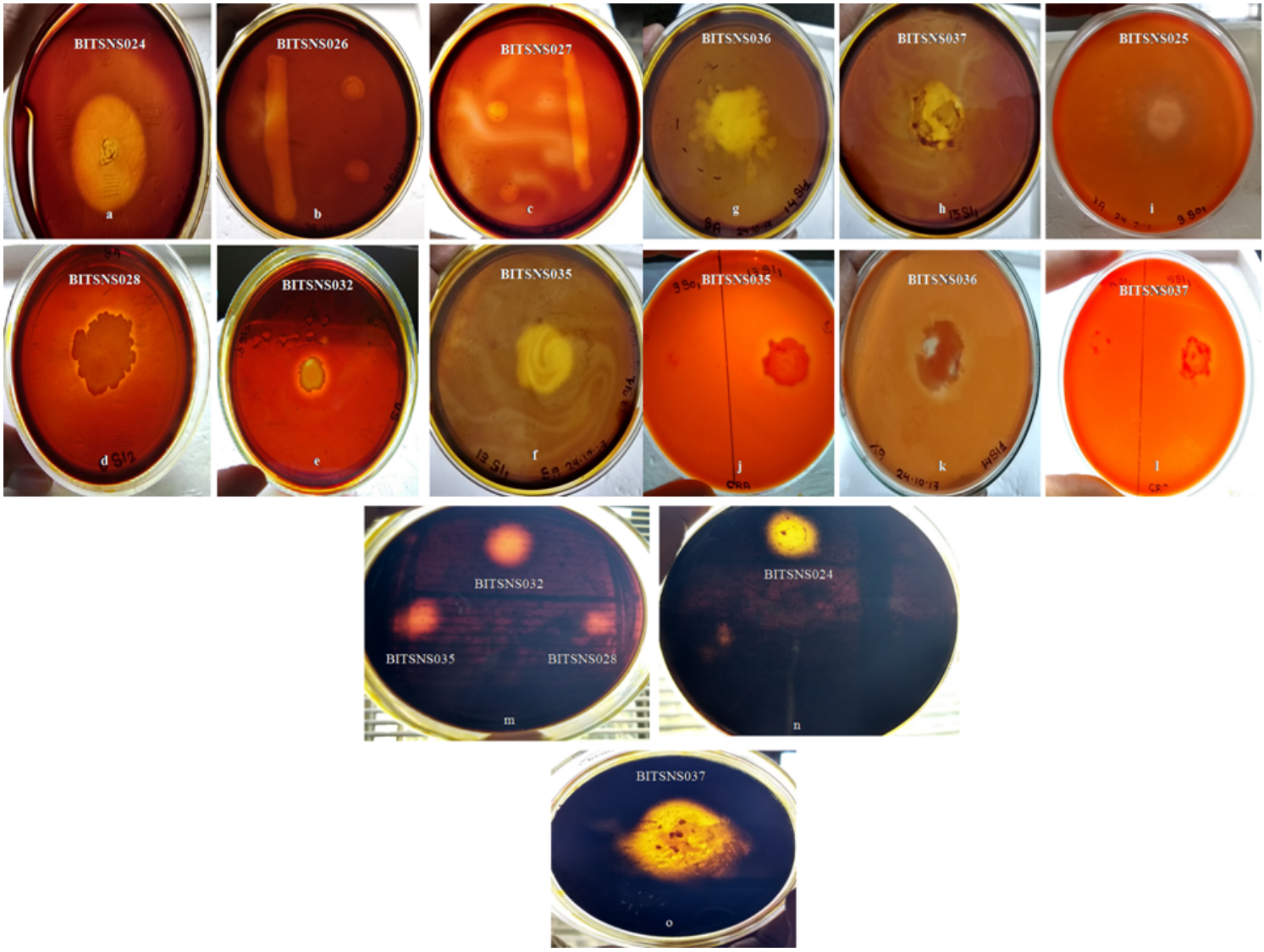

Figure 1

Hydrolysis zone as indicative for amylase (a-h), xylanase (i-l) and cellulase (m-o) activity screening of 60°C isolates

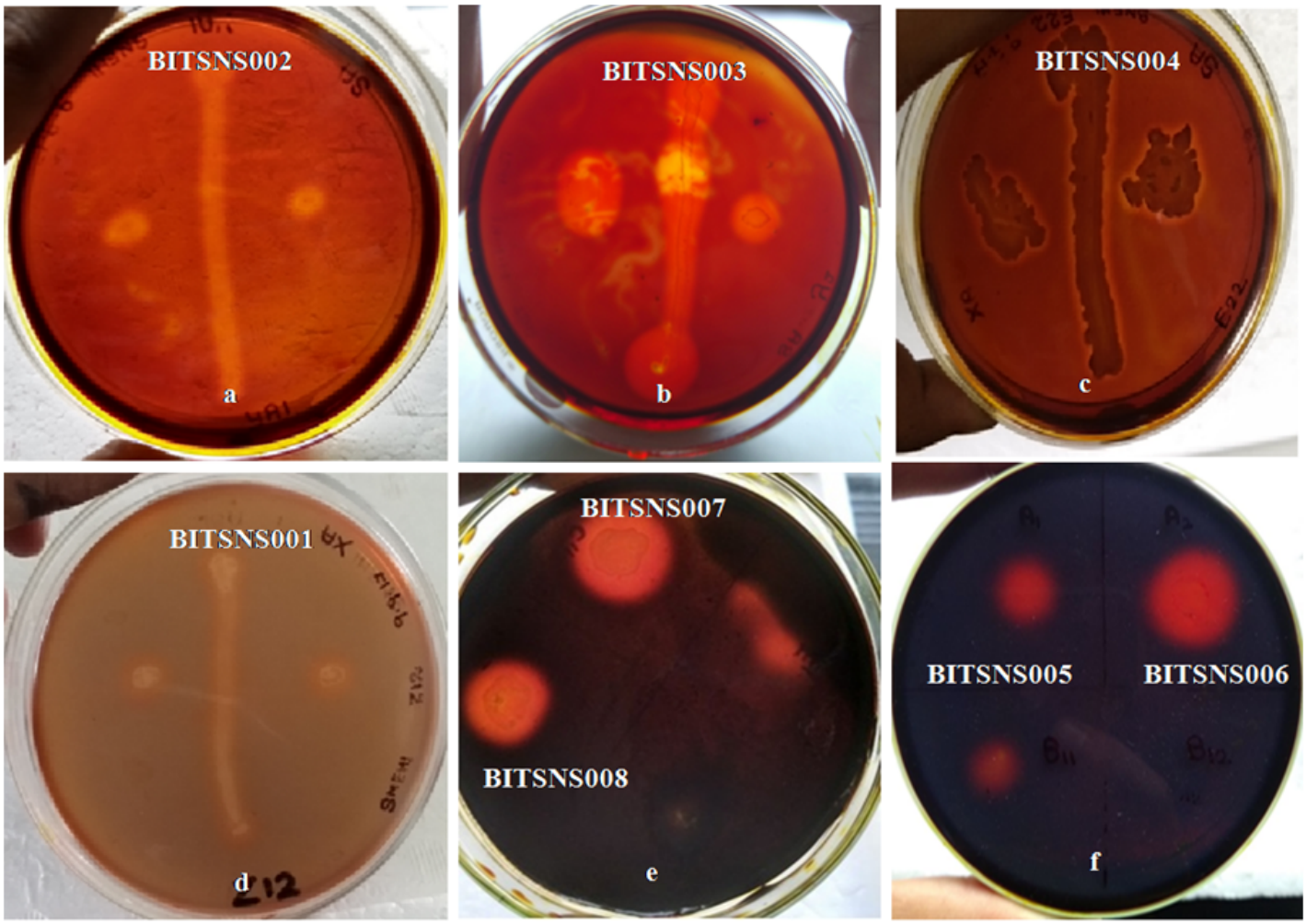

Figure 2

Amylolytic (a-c), xylanolytic (d) and cellulolytic (e-f) activity as hydrolysis zone on media plates for 50°C isolates

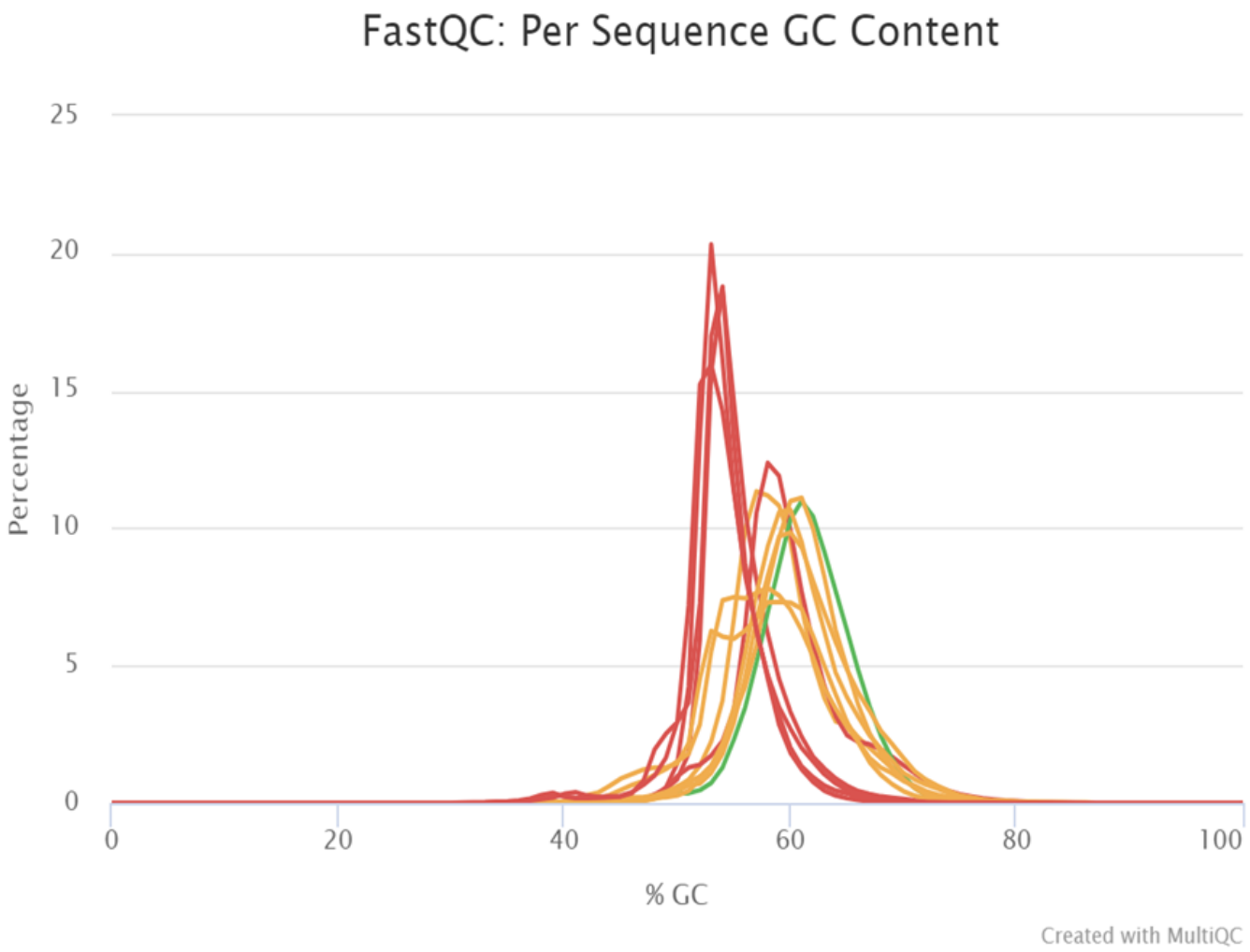

Figure 3

Comparative analysis of the GC content present in the microbiota isolated from the hot springs of Jharkhand

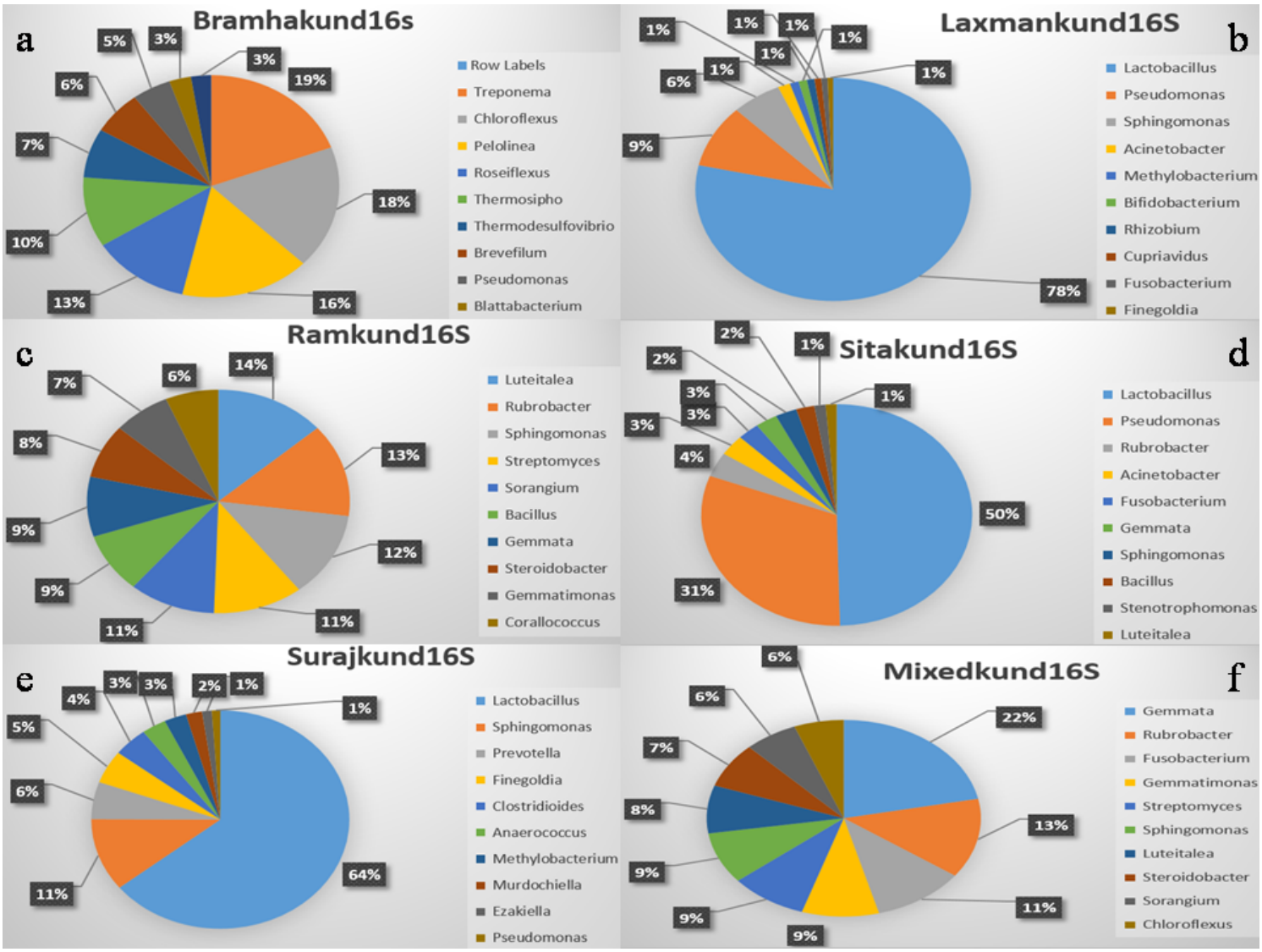

Figure 4

The representation of the 16S rRNA identification of bacterial population of (a) Brahma kund (b) Laxman kund (c) Ram kund and (d) Sita kund (e) Suraj kund (f) Mixed kund

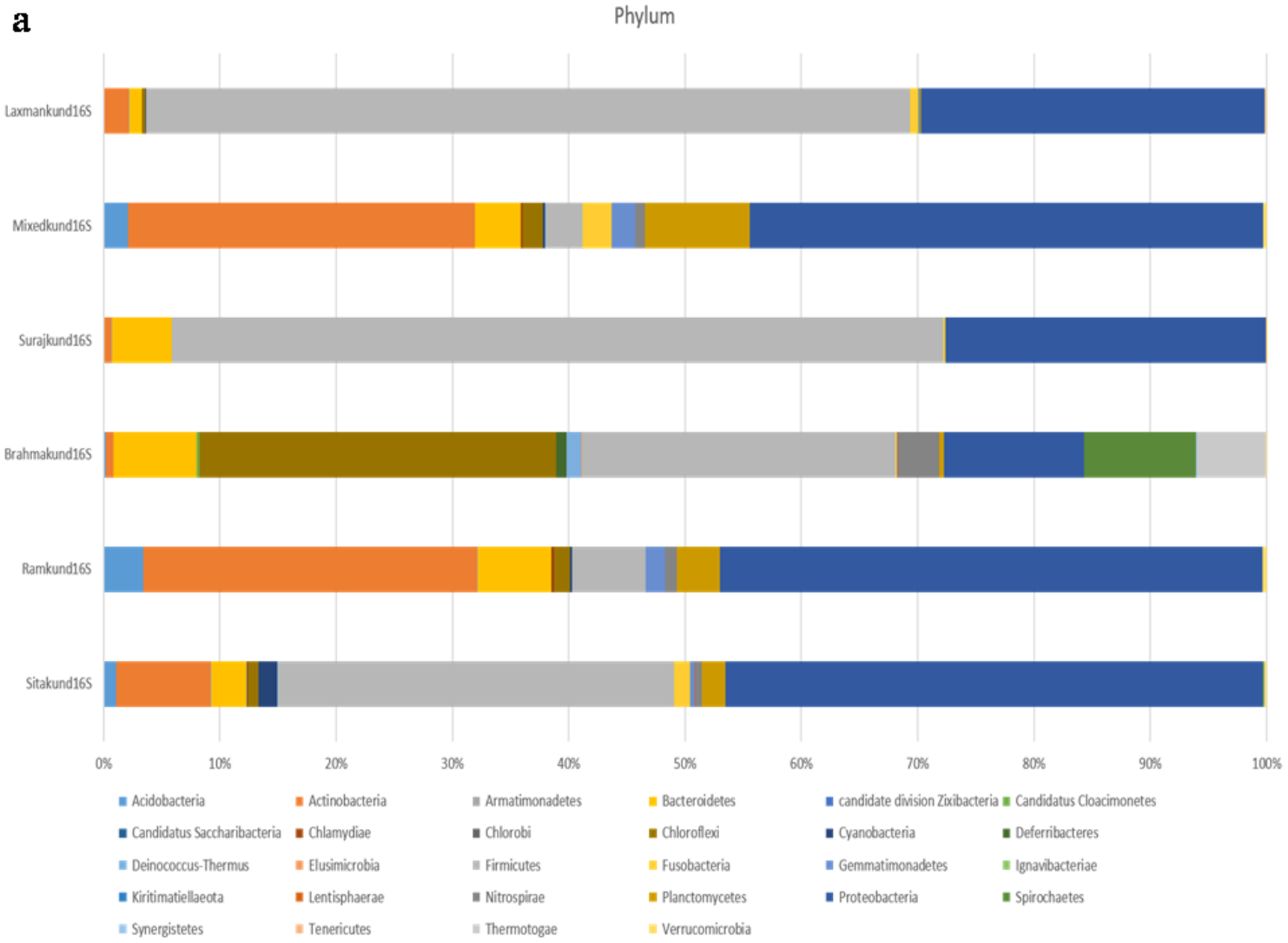

Figure 5

Taxonomic analysis of isolated microbiota from hot springs of Jharkhand and its classification on the basis of (a) Phylum (b) Class (c) Family (d) relative abundance along with the prevalence of different bacterial genus in the kunds (e) analysis of gene presence/absence.

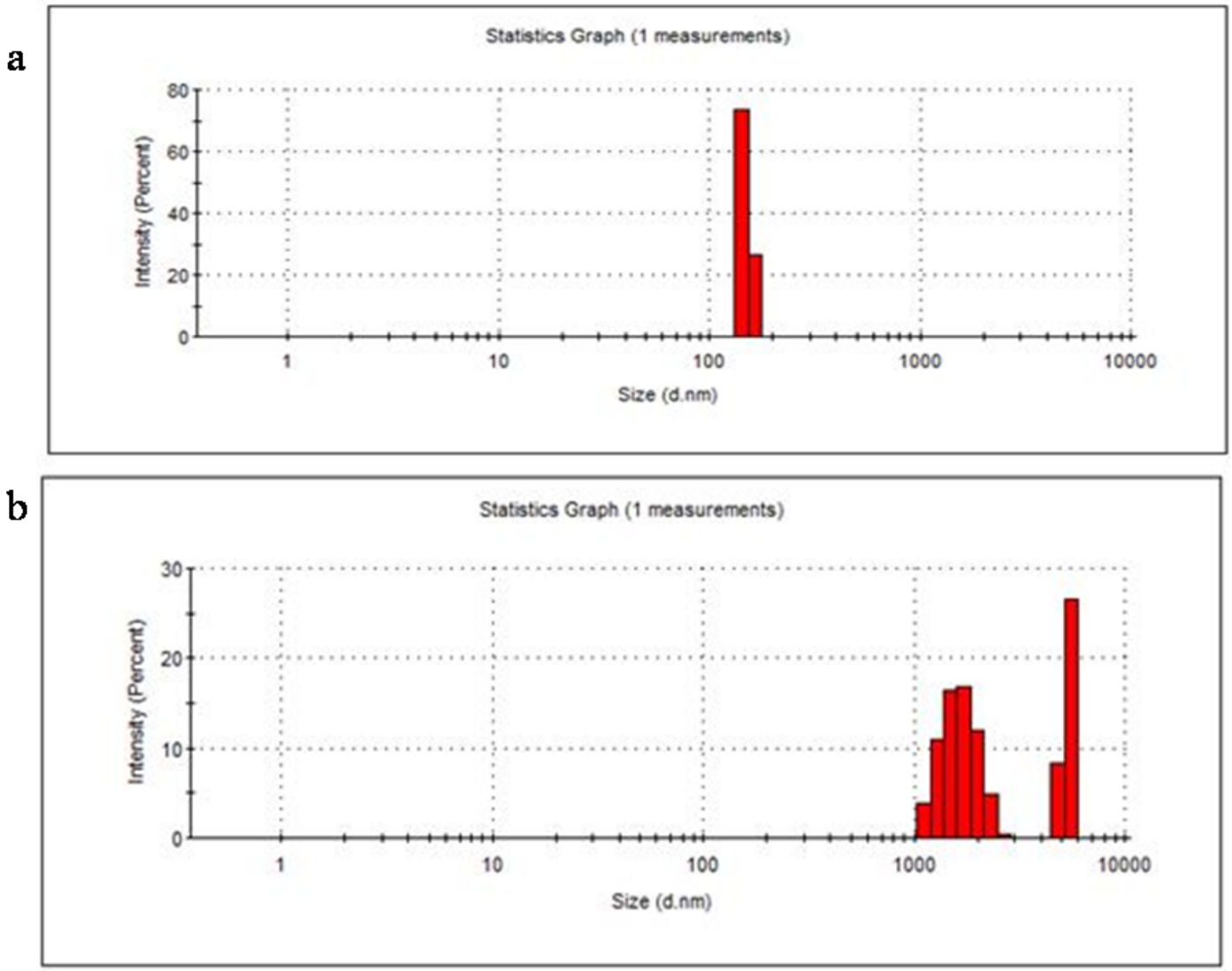

Figure 6

DLS analysis of (a) standard commercial amylase (b) partially purified amylase

Supplementary Files

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

- SupplementryfileSR.docx