DECIPHERING THE MICROBIAL DIVERSITY OF TATTAPANI HOT WATER SPRING USING METAGENOMIC APPROACH

MAHAJAN MAHESH MOHANRAO¹, DUSHYANT PRATAP SINGH², KUMAR KANIKA³, ETIKA GOYAL⁴ & AMIT K. SINGH⁵

¹, ⁴, ⁵Research Scholar, NRCPB, New Delhi, India
²Research Scholar, JJT University, Rajasthan, India
³Senior Scientist, NRCPB, New Delhi, India

ABSTRACT

Heat stress due to changing environmental conditions is a problem as it negatively affects plant growth and development leading to reduction in crop yield. Thermal springs are major source of potentially significant but untapped microbial diversity encoding novel genes related to high temperature tolerance. The genes involved in the heat tolerance mechanism of thermophiles can serve as a valuable resource to improve high temperature tolerance in crop plants. Metagenomics bypasses the need for isolation or cultivation of microorganisms as it deals with direct isolation of nucleic acids from environmental samples. The present study was designed to analyse the profile of microbial community inhabiting hot water spring Tattapani, Himachal Pradesh, India. The spring water temperature at the discharge point was recorded as 65°C and pH 7.5. Two sets of universal 16S rRNA gene specific primers namely U789F-U1053R and MGUF-MGUR were used to amplify partial 16S rRNA gene (from V5-V6 and V3-V6 regions respectively) from metagenomic DNA. Sequencing and bioinformatic analysis revealed that bacterial phylum Proteobacteria was most abundant, constituting 36% and 40% of the total bacteria identified using V5-V6 and V3-V6 regions, respectively. Other phyla Thermodesulfobacteria (9.3%), Firmicutes (7.5%), Deinococcus-thermus (7%), Bacteriodates (2.5%), Aquificae (2%) were assigned to sequences amplified from V5-V6 region while phyla Firmicutes (18%), Bacteriodates (6%), Chloroflexi (2%) were assigned to sequences amplified from V3-V6 region. A significant number of sequences (30.5% and 20.5%) from region V3-V6 and V5-V6 respectively did not show homology with existing database and were recorded as unclassified Bacteria. The information generated in this study will help us in understanding the microbial ecology of hot springs.

KEYWORDS: Microbial Diversity, Metagenomics, Hot Water Spring

INTRODUCTION

Natural environments are shelter for large diversity of microorganisms; most of these microorganisms are unexplored and belong to untapped genetic resource (Nichols, 2002). Extreme environments such as hot water springs, arctic regions, highly saline or alkaline regions, deep oceans etc. are a major source of potentially significant but untapped microbial diversity. Study of microbial diversity in such extreme environments is necessary because, different physico-chemical conditions, biogeography and geological history, have different microbial phenotypes (Narayan, 2008; Liu et al., 2009). Of the total microorganisms present in environment only 0.1 to 1% can be cultured by conventional laboratory methods (Amann et al., 1995). Thus remaining ~99% of microbes are unculturable. To overcome this problem culture independent approach (metagenomics) is used that...
enable us to investigate microbial diversity from diverse habitats. Metagenomic approach based on 16S rRNA gene analysis through its amplification and sequencing provides information about microbial community present in an environment.

The partial 16S rRNA gene sequence based metagenomic approach has been extensively used to study microbiota present in hot water springs. These includes, Sungai klah hot spring, Malaysia (chan et al., 2015), Tuwa hot spring, India (Mangrola et al., 2015) hot spring in Tibetan plateau, China (Huang et al., 2011), hot spring in Tengchong County in Yunnan Province, China (Briggs et al., 2013; Song et al., 2009, 2010; Jiang et al., 2010), different geothermal hot springs in Tunisia (Sayeh et al., 2010) Long Valley Caldera near Mammoth Lakes, CA, USA (Vick et al., 2010), hot springs in Seltun and Hveradalir geothermal areas, Iceland (Aguilera et al., 2010; Reigstad et al., 2009; Marteinsson et al., 2001), acidic hot spring, Kawah Hujan B, Indonesia (Aditiawati et al., 2009), hot springs located at Yellowstone National Park, USA (Meyer-Dombard et al., 2005; Hall et al., 2008; Mitchell 2009), Soldhar and Ringigad hot springs, Uttaranchal, India (Kumar et al., 2004) and Kamchatka hot spring, Russia (Bonch-Osmolovskaya et al., 1999; Reigstad et al., 2009). These studies revealed that temperature, pH and biogeography are the important factors affecting the microbial diversity of the hot water springs (Huang et al., 2011).

Though large number of studies has been carried out on hot springs habitats, there is a lack of information on microbial communities inhabiting hot water springs in India. In this backdrop, present study was planned to study the microbial diversity inhabiting the hot water spring at Tatapani, Himachal Pradesh, India. The Tatapani hot water spring is located on the bank of river Satluj. Metagenomic DNA from the hot spring water was isolated and partial 16S rRNA gene was amplified using primers designed from V3-6 and V5-6 hypervariable regions of prokaryotic 16S rRNA genes. PCR amplified partial 16S rRNA gene fragments were cloned and sequenced. Different bioinformatic tools were used to investigate the microbial community structure of Tatapani hot water spring.

MATERIALS AND METHODS

Study Area and Sampling

The Tatapani hot water spring is located at an altitude of 655 meters above sea level, on the bank of river Satluj, Mandi district, Himachal Pradesh, India, having geographical coordinates 31° 14’ 56”N, 77° 5’ 10’’E (Figure 1a and 1b). The location of hot spring is geographically very unique as the temperature of river water was recorded near about 5°C-6°C while the temperature of hot water of the reservoir was ~65°C. Hot water samples for microbial diversity analysis were collected from the hot springs oozing out from different places on the banks of river. Water samples were pooled and used for metagenomic DNA extraction. White and light yellow coloured salt and sulphur deposits were observed all around the surrounding stones of the hot spring along with some algae which formed mats along hot springs. Different physico-chemical water quality parameters, namely; temperature, pH, electrical conductivity (EC) and total dissolved solid (TDS) were measured on-site using the relevant field meters.

Metagenomic DNA Extraction

Pooled hot spring water sample were filtered using Millipore filter membrane having pore size 0.22µm through vacuum filtration system at collection site. The microbial mass was concentrated on the surface of the filter papers which were further used to isolate the metagenomic DNA. The filter papers were sliced in to smaller pieces, placed in aseptic centrifuge tubes and processed for extraction of metagenomic DNA using RapidWater™ DNA Isolation (MO BIO Laboratories, Inc.) as per the manufacturer’s protocol. DNA extraction was performed multiple times, extracted DNA was
pooled to get concentrated DNA. The quality and quantity of DNA was evaluated by using agarose gel electrophoresis and Nanodrop™ 1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

**Targeted 16S rRNA Clone Library Preparation**

The PCR was performed using GOTaq® green master mix (Promega, USA) for amplification of partial 16S rRNA gene using diluted metagenomic DNA as template in Eppendorf master cycler. The oligonucleotide primer sets used in this study are shown in Table 1. Thermal cycler conditions used for primer set MGUF/MGUR were as follows: 4 min. of initial denaturation at 94°C; followed by 35 amplification cycles of 1 min. at 94°C, 45 sec. at 60.7°C, 1 min. at 72°C; with a final 10 min of prolonged extension at 72°C. The conditions for primer set U789F/U1053R were similar to the above conditions except for annealing and extension which were 58.6°C for 45 sec., 72°C for 40 sec. respectively. Amplified fragments were visualized on 1.2% agarose gel. The amplicons were gel purified using the column based MinElute® Gel Extraction Kit (Qiagen, Germany) as per the manufacturers protocol.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Expected Size of Amplicons</th>
<th>V Region</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGUF</td>
<td>5’TGCCAGCAGCGCGCGTA 3’</td>
<td>850bp</td>
<td>V3-V6</td>
<td>Wang, Y. and Qian, P. (2009)</td>
</tr>
<tr>
<td>MGUR</td>
<td>5’GACGGGGGCGGTGTTACAA 3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U789F</td>
<td>5’TAGATACCCSSGTAGTTC 3’</td>
<td>270bp</td>
<td>V5-V6</td>
<td>Wang, Y. and Qian, P. (2009)</td>
</tr>
<tr>
<td>U1053R</td>
<td>5’CTGACGRCRGCCATGC 3’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Cloning and Transformation of Partial 16S rRNA Gene Fragment**

Gel purified PCR products were ligated with pGEMT easy vector (Promega, USA). The reaction was incubated at 4°C overnight as recommended in manufactures protocol. Electrocompetent DH5α E. coli cells were transformed with ligated product using standard electroporation protocol (Sambrook and Russell, 2001). The transformed cells were picked using blue/white screening. The transformants were confirmed for the presence of the insert using colony PCR. Positive clones were used for isolation of plasmid DNA using R.E.A.L® prep 96 Plasmid Isolation Kit (Qiagen, Germany)) as per the manufacturer’s protocol.

**in silico Analysis of 16S rRNA Gene Sequences**

The partial 16S rRNA clones were sequenced and data were classified using the RDP Naive Bayesian rRNA Classifier Version 2.10, October 2014, RDP training set 14, based on nomenclatural taxonomy and Bergey’s Manual with an 80% confidence threshold on the RDP-database. DECIPHER, a search based approach was used to identify the presence of possible chimeric sequences from all 16S rRNA gene sequences (Wright et al. 2012). Bioinformatics analysis of the sequences was carried out to identify homology match. After excluding redundant and false-positive sequences, nucleic acid homology searches were performed against nucleotide databases at the National Center for Biotechnology Information (NCBI) using the Basic Local Alignment Search Tool (BLASTN) http://www.ncbi.nlm.nih.gov/BLAST to provide gene annotation. Phylogenetic trees of selected sequences were constructed from distance matrices by the neighbor-joining method (Naruya et al., 1987), using MEGA6 tool (Figure 3a and 3b). The node reproducibility for tree topology was estimated by bootstrap analysis, which included 1000 replicate data sets.

**Nucleotide Sequence Accession Numbers**

The nucleotide sequences from V3-V6 region and V5-V6 region have been deposited in the GenBank (NCBI)
RESULTS AND DISCUSSIONS

Physico-Chemical Analysis of Spring Water Sample

The physico-chemical parameters of the spring water such as temperature and pH are important factors that determine microbial diversity in hot springs (Hou et al., 2013; Cole et al., 2013). High temperature of water exerts stress on microbial species; therefore only tolerant species can survive (Abou-Shanab et al., 2007). In present study water sampling was done at different spots of hot springs present on the river bed and samples were pooled to obtain final representative water sample. The spring water was clear in appearance with smell of sulphur. The pH of water was 7.5 which showed that it was slightly alkaline in nature. The Physico-chemical parameters of water are as summarized Table 2.

Table 2: Physico-Chemical Parameters of Tattapani Hot Water Spring

<table>
<thead>
<tr>
<th>Physico-Chemical Parameters</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>65°C</td>
</tr>
<tr>
<td>pH</td>
<td>7.5</td>
</tr>
<tr>
<td>EC</td>
<td>17.3 00 dS/m</td>
</tr>
<tr>
<td>TDS</td>
<td>8400 mg/L</td>
</tr>
<tr>
<td>Odour</td>
<td>Rotten egg (qualitative)</td>
</tr>
<tr>
<td>Colour</td>
<td>Normal (Hazen scale unit)</td>
</tr>
<tr>
<td>Appearance</td>
<td>Clear (qualitative)</td>
</tr>
<tr>
<td>Taste</td>
<td>Salty (qualitative)</td>
</tr>
<tr>
<td>Sulphate</td>
<td>112 mg/L</td>
</tr>
</tbody>
</table>

Extraction of Metagenome and Cloning of Partial 16s rRNA Gene

Number of protocols has been standardized for extraction of metagenomic DNA (Zhou et al., 1996; Sharma et al., 2007; Aditiawati et al., 2009) and there are various kits available for extraction and purification of DNA from environmental water samples. Some of them are easy to use and reproducible. However, none of these methods are universally applicable to all water type. Each water type requires optimization of DNA extraction protocol. In the present study metagenomic DNA isolation was done by Rapid™ DNA Isolation (Mo Bio Laboratories’, Inc.) (Figure 2a). The isolated DNA was of good quality with no visible shearing and its concentration was calculated as ~90 ng/µl. The purity ratio was measured as 1.87 at wavelength 260/280 using Nanodrop™ 1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Purified metagenomic DNA was used for PCR amplification of partial 16S rRNA gene from two variable regions (Figure 2b). Amplified 16S rRNA gene fragments were cloned in pGEM®-T easy cloning vector. After blue white screen around 105 positive clones were picked having 850bp 16S rRNA gene fragment and 161 clones were picked having 270bp 16S rRNA gene fragment.

Taxonomic Profile Analysis Using V5-6 and V3-6 Region of 16S rRNA

To investigate taxonomic profile present in environmental samples, 16S rRNA gene analysis with its highly conserved and variable sequence provides a major target region for studying microbial diversity (Muyzer et al., 1993; Becker et al., 2004; Wang and Qian, 2009). In recent metagenomic studies, the number of phylotypes in the same number of 16S rRNA gene sequences varied substantially for samples from different environments and geographical sites (Sogin, 2006; Huber, 2007; Roesch, 2007). Primer usage is undoubtedly one of the most critical limiting factors affecting 16S rRNA gene analysis (Armougom and Raoult, 2009). The amplification efficacy of the different 16S rRNA gene primers used can lead to under or over estimation of some microbial phyla. In various studies V3 and V6 have been most
extensively examined regions (Huse et al., 2008; Wang et al., 2009). However different primers targeting different regions of 16S rRNA gene were used for amplification in different studies. In present study, two sets of primers targeting V3-V6 region were used for 16S rRNA gene amplification (Wang et al., 2009). To study microbial diversity of Kawah Hujan B hot spring, Indonesia, primers were designed from variable region V7-V8 (Aditiawati et al., 2009), for Malaysian hot spring amplicons of V3-V4 region were amplified (Chan et al., 2015), regions V1-3 and V4-7 were amplified for bacterial diversity analysis of Siloam hot water spring (Tekere et al., 2011). Different sets of primers lead to difference in the capture depth of the microbes in different environmental samples because of varying amplification efficiencies and coverage rates of the primers. The bacteria in rare biospheres can never be identified if the primers used are not applicable to them.

Previous studies have shown that diverse microbial body structures can inhabit hot springs with a wide temperature range from ambient to boiling. (Barnes et al., 1994; Pace 1997; Bonch-Osmolovskaya et al., 1999; Marteinsson et al., 2001; Kumar et al., 2004; Meyer-Dombard et al., 2005; Lau et al., 2006, 2009; Hall et al., 2008; Aditiawati et al., 2009; Mitchell 2009; Reigstad et al., 2009; Song et al., 2009, 2010; Aguilera et al., 2010; Jiang et al., 2010; Sayeh et al., 2010; Vick et al., 2010). In silico analysis of obtained results revealed that majority of the total partial 16S rRNA gene fragments belong to domain bacteria, where as only one sequence belonged to domain Archaea whose phylum identified was Woesearchaeota. The bacterial community structure detected in present study is given in figure 4. An evident phylum Proteobacteria was most abundant, constituting 36% and 40% of the total bacteria identified using V5-6 and V3-6 regions, respectively. Phylum Proteobacteria comprised mainly of hydrogenophilus with pseudomonas, sphingomonas, rhodovibrio, nitrospira, serpens, klebsiella and Escherichia/Shigella being present in small portion. The majority of Hydrogenophilus were represented by the order Hydrogenophilales. Hydrogenophilales generally includes anaerobic, rod-shaped, gram-negative, chemolithoautotrophic bacteria. Other phyla Thermodesulfobacteria (9.3%), Firmicutes (7.5%), Deinococci-thermus (7%), Bacteriodates (2.5%), Aquifera (2%) were assigned to V5-6 region amplified sequences while phyla Firmicutes (18%), Bacteriodates (6%), Chloroflexi (2%) were assigned to V3-6 region amplified sequences. A significant number of sequences, 30.5% and 20.5% from region V3-6 and V5-6 respectively could not match to existing database which were recorded as unclassified Bacteria. The most dominant phylotypes inhabiting Tattapani hot water spring were found to be Proteobacteria, Firmicutes, Thermodesulfobacterium and Deinococcus-Thermus, which are widely distributed in both terrestrial and aquatic environment. Apart from this other phylotypes of this spring were Bacteriodates, Chloroflexi and Aquifera. The phylotypes Proteobacteria, Firmicutes, Thermodesulfobacterium, Deinococcus-Thermus, Bacteriodates, Chloroflexi and Aquifera have been reported to inhabit Siloam hot water spring, Limpopo, South Africa (Tekere et al., 2011); Malaysian hot spring (Chan et al., 2015); hot spring on the Tebetian Plateau, China (Huang et al., 2011).

In total 11 distinct classes were identified out of which 7 classes were belonged to region V3-6 and all 11 classes to region V5-6. All the classes from region V3-6 were detected in 16S rRNA fragments from region V5-6. The most dominant class of the phylum Proteobacteria in this hot water spring was the betaproteobacteria (78%) followed by gammaproteobacteria (19%) and alphaproteobacteria (4%). Other major classes detected for both regions were Clostridia, Sphingobacteria, Bacilli, Anaerolineae, unclassified Bacteroidetes and unclassified Firmicutes. Classes unique to region V5-6 were Thermodesulfobacteria, Deinococci, Cytophagia and Aquifera.
Out of 15 distinct identified orders, 12 orders belonged to V5-6 region and 9 orders were obtained from V3-6 region. The order *Hydrogenophilales* was found to be most abundant in both the regions. Other major orders identified for region V5-6 were *Thermodesulfobacterales, Clostridiales, Pseudomonadales, Thermales, Enterobacterales, Desulfurococcales, Pseudomonadales and Bacillales* and for region V3-6 were *Clostridiales, Thermoanaerobacterales, Sphingobacteriales and Coriobacteriales*.

A total of 17 and 14 different genera were assigned to region V5-6 and V3-6 respectively. The most abundant bacterial genus identified using both variable regions was *Hydrogenophilus*. Other detectable genera for region V5-6 were *Caldimicrobium, Thermus, Pseudomonas, Hyperthermus* etc., and for region V3-6 were *Fervidicola, Desulfovirgula, Salisaeta, Thermobrachium, Rhodothermus* etc. Although, the dominant microbial groups present in extreme environments are more emphasised, it is worth noting that the less abundant microbes found in extreme environment could be an important individuals of the unexplored community which can reuse as unculturable genetic resource (Sogin et al., 2006).

**CONCLUSIONS**

Analysis of DNA directly extracted from environmental samples allows us to study natural microbial communities without the need for cultivation. The spring water rich in evident bacterial phylum *Proteobacteria* and other bacterial phyla found in the present study were *Thermodesulfobacteria, Firmicutes, Deinococci-thermus, Bacteriodates and Aquificea*. Isolation of genes responsible for high temperature tolerance in these bacteria can be used for developing high temperature tolerance in crop plants. This study will act as a foundation for further studies on microbial community diversity of extreme habitats.

**ACKNOWLEDGEMENTS**

The authors thank to Dr. T. R. Sharma, Project Director ICAR-NRCPB, New Delhi. Special thanks to Dr. Vinay Sharma ICAR-CPRI, Shimla for their help during sample collection.

**REFERENCES**


APPENDICES

Figure 1: Location of Hot Water Spring Tattapani, Himachal Pradesh. (a)-Google Map Showing; Hot Water Location of the Spring; (b)-Actual Sites of the Sampling of Hot Water

Figure 2: Agarose Gel Electrophoresis a-(Lane 1-3) Metagenomic DNA and Lane M- Lambda DNA /EcoR1+HindIII marker. b- (Lane 1,2) Amplicon Obtained Using Primer Pair U789F-U1053R; (Lane 3,4) Amplicons Obtained Using Primer Pair MGUF-MGUR
Figure 3a: Neighbor-Joining Tree Showing the Phylogenetic Relationships of Partial 16S rRNA Gene Sequences Amplified from V3-V6 Region
Figure 3b: Neighbor-Joining Tree Showing the phylogenetic Relationships of Partial 16S rRNA Gene Sequences Amplified from V5-V6 Region
Figure 4: Bacterial Community Structure Obtained from V3-6 (a-d) and V5-6 (e-f) Region; a & e – Phyla, b & f- Classes, c & g – Order, d & h- Genera