16S rRNA based identification and phylogenetic analysis of a novel dextran producing *Pediococcus pentosaceus* isolated from north-east Indian microbial diversity

Seema Patel and Arun Goyal
Department of Biotechnology
Indian Institute of Technology Guwahati
Guwahati- 781 039, Assam, India

*For correspondence - arungoyl@iitg.ernet.in*

**Abstract**

Looking at the high application potential of dextrans in food, pharmaceutical and tissue engineering fields, a dextran producing strain of lactic acid bacterium was isolated from the sugarcane field soil sample from Assam, falling under north-east Indian microbial diversity. The full-length 16S rRNA gene of the isolate was amplified by Polymerase Chain Reaction (PCR) and sequenced by Sanger’s method. The sequence was aligned with the reference sequences in the GenBank of National Centre for Biotechnological Information (NCBI) and Ribosomal Database Project (RDP) using BLAST programme. Distance matrix based phylogenetic analysis resulted in well-resolved trees with nodes supported with high bootstrap pseudoreplicate scores. The lactic acid bacterium isolate was identified as *Pediococcus pentosaceus* (GenBank Accession Number EU569832). This identification revealed the dextran production attribute of *Pediococcus* genus of lactic acid bacteria for the first time ever, heralding further rigorous investigation on this aspect. This investigation also unravelled the abundance of industrially valuable microbial flora in the soil of north-east India.

**Key words:** *Pediococcus pentosaceus*, dextran, 16S rRNA gene, phylogenetic analysis.

**Introduction**

Dextrans (C_{6}H_{10}O_{5})_{n} are a class α-(1>6) linked glucans having branchings of α-(1>2), α-(1>3), α-(1>4) linkages (1). Dextrans have enormous industrial applications as food additives, chromatography column matrices, blood plasma substitutes, treatment of anaemia etc. (2). Recently, dextrans have found use in nanoconstructs as potential vectors for anti-cancer agents (3). *Lactobacillus, Leuconostoc, Streptococcus* and *Weisella* genera of lactic acid bacteria synthesize dextrans (4). There was not any concrete report on dextran production by *Pediococcus* genus, however, Smitinont et al. (1999) had emphasized on dextran synthesizing ability of this genus (5). Patel et al. (2010) reported the dextran production ability of *Pediococcus pentosaceus* for the first time ever (6). *Pediococci* are a heterogeneous group of homofermentative lactic acid bacteria (7). Currently ten species of *Pediococcus* are recognized, including *Pediococcus damnosus, P. parvulus, P. inopinatus, P. cellicola, P. ethanolidurans, P. clauseni, P. stilesii, P. acidilactici, P. pentosaceus* and *P. dextrinicus* (8, 9). The species of this genus are cosmopolitan in distribution as they have been isolated from soil (10), plants (11), wines (12), cheese (13), wheat kernels (14), meat (15) and sausages (16). Members of the *Pediococcus* genus have great
economic importance in the fermented food industry (17). *Pediococcus* genus has been known for its use as starter cultures in fermentation processes of milk, meat, vegetable products and sausages (17, 18). *P. acidilactici* has GRAS (Generally Recognized As Safe) status and is used as probiotic culture and nutritional enhancer in silage (18, 19). In addition to their contribution to fermented foods, several *Pediococcus* strains produce bacteriocins that inhibit the growth of major Gram-positive foodborne pathogens, as well as other food spoilage bacteria (20). Bacteriocinogenic strains of *Pediococcus* genus has been found active against lactic acid bacteria, *Escherichia coli*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Listeria innocua*, *Listeria ivanovii* and *Listeria monocytogenes* (21).

Till date, there have been several studies on pediocin production from *Pediococcus* genus (22). However, only sparse literature on exopolysaccharide production aspect of *Pediococcus* strains has been documented. Bacterial exopolysaccharides have attracted huge attention in recent years due to their widespread industrial potentials and they warrant studies in *Pediococcus* genus like other exopolysaccharide producing lactic acid bacteria. In *P. damnosus*, a plasmid associated with glucan-production has been reported in wine (23). Ropy *Pediococcus damnosus* (strain 2.6) was used for production of exopolysaccharide (EPS) in a semidefined medium (24). In some Argentinian wines, the ropiness inducing lactic acid bacteria include *Pediococcus pentosaceus* (25).

Traditionally, natural isolates have been identified by phenotypic methods, which are time consuming, potentially inaccurate and unreliable. The identification of microorganisms historically has relied on phenotypic methods that are often time-consuming, because of the inherent mutability of biochemical characteristics and subject to interpretive bias (26). Recent years have witnessed an explosion in the development and application of molecular tools for exploring the microbial diversity and identifying the isolates, owing to their superiority over the conventional approaches. Many of the modern molecular tools are based on 16S ribosomal RNA sequences, complete or partial genome determination and monitoring by specific fluorescent probes coupled to flow cytometry (27). Availability of large databases or the online resource offers powerful platform for rapid in silico specificity profiling. Classification and phylogenetic studies of microorganisms based on these high throughput molecular approaches are increasingly applied to strains of lactic acid bacteria (LAB) (27). Sequence analysis of amplified 16S RNA is recognized as a reliable technique for genus and species identification in different bacterial genera including *Lactobacillus*, *Weisella*, *Actinomyces*, *Bifidobacterium*, *Propionibacterium* and *Streptococcus* (28).

In the quest of a novel dextran, a bacterium, SPA (named after the author and the place of origin) was isolated from the soil sample of microbial diversity rich north-east India. Phenotypic characterization had already established the isolate SPA as a member of lactic acid bacteria family (10). In this work, the 16S rRNA gene sequencing based identification of the isolate was conducted to assign it the generic and species name, also to trace its taxonomic position by phylogenetic analysis.

**Materials and Methods**

**Microorganism and maintenance conditions**

The lactic acid bacterium isolate SPA screened from the sugarcane field soil of Assam was grown in enzyme production medium devised by Tsuchiya *et al.* (29) at 25°C. The isolate culture was propagated as stab in modified MRS medium
(30) at 28°C and stored at 4°C. For long term preservation, the isolate was kept in 20% (v/v) glycerol at –80°C (10).

**Extraction of genomic DNA**: The bacterial cell pellet was lysed using a solution containing guanidium thiocyanate (a chaotropic agent) and SDS (a detergent), to extract DNA (31). This lysis solution was used to disrupt the cell, remove proteins, polysaccharides and partial hydrolysis of RNA. The genomic DNA was extracted by GeNeiTm Genomic DNA Extraction kit (Bangalore Genei Pvt. Ltd.). DNA was then precipitated using alcohol and washed with 70% alcohol to remove contaminants. DNA pellet was solubilized in an appropriate buffer (sodium acetate at pH 5.5 added to a final concentration of 0.3 M in isopropanol) at higher temperature (50-55°C), to increase the solubility of genomic DNA. The extracted DNA was purified using Ultra pure prep kit (KT83B, Bangalore Genei, India). Genomic DNA was purified in a gravity flow column using elution buffer, followed by treatment with lysozyme, protease K and RNAse A supplied with the kit. DNA concentrations were determined in duplicate using a spectrophotometer.

**PCR amplification**: Genomic DNA of the isolate SPA was used for amplification of 16S rRNA gene. The universal 16S rDNA primers, forward primer BG395F (sequence hidden) and reverse primer BG396R (sequence hidden) were used for the polymerase chain reaction (PCR). The PCR amplification was carried out in a reaction mixture containing ~10ng genomic DNA as template, 1µl dNTP mix (2.5 mM each), 100 ng/µl each of forward and reverse primer, 1X Taq DNA polymerase assay buffer (10X), 3U Taq DNA polymerase enzyme (Bangalore Genei Ltd., Bangalore, India) and distilled water enough to make up the volume to 50 µl reaction mixture. Amplification conditions were: 5 min initial denaturation at 94°C, 30s denaturation at 94°C, 30s of primer annealing at 54°C, 1 min elongation at 72°C for 35 cycles and a final extension of 10 min at 72°C. The reactions were carried out in a Thermal Cycler (Applied Biosystems, model ABI 2720).

**Electrophoresis**: Fifteen µl of PCR amplification product was electrophoresed on 1% (w/v) low EEO; agarose gel in 1X TBE buffer (45 mM Tris-borate, pH 8.3 and 1 mM Na 2 EDTA) at 100V for 2h. The gel was stained with ethidium bromide in a final concentration of 0.5µg/ml, visualized and photographed under UV light. A low range DNA ladder (1 Kb) (Bangalore Genei Pvt Ltd, India) was used as a molecular weight marker.

**Sequencing of the 16S rRNA gene**: The 16S rRNA gene from the isolate was amplified as mentioned above. The amplification product was eluted and purified from the gel slice using the GeneiPure™ Gel Extraction Kit (QIAGEN) and sequenced by Sanger’s method using an automated genetic analyser (Make: ABI, Model: 3100) which uses the sequencing analysis software v.5.1 with the kb basecaller (32).

**Sequence alignment of the 16S rRNA gene**: The 16S rRNA sequence of the isolate SPA was compared with above 100 homologous sequences taken from the National Centre for Biotechnological Information (NCBI) genbank and Ribosomal Database Project (RDP) (33). Multiple sequence alignment of the sequence was conducted by ClustalW algorithm. Phylogenetic analysis was performed by applying distance matrix method. The isolate was assigned accession number.

**Results and Discussion**

16S rRNA gene sequence analysis after PCR amplification was performed for identifying
the isolate SPA. Profiles obtained by PCR amplification allowed identification of the isolate at both genus and species level. The Polymerase Chain Reaction of 16S rRNA revealed a 1,491bp amplicon by 1% agarose gel electrophoresis (Fig. 1). The full length sequencing of 1491bp 16S rRNA gene was identified by Sanger’s method (Fig. 2). The 16S rRNA sequence was compared with that of reference bacteria obtained from National Centre for Biotechnological Information (NCBI) Genbank (http://www.ncbi.nlm.nih.gov) and Ribosomal Database Project (RDP). Similarity searches were carried out using the BLAST algorithms available at (http://www.ncbi.nlm.nih.gov/BLAST/). Multiple alignments comparing the sequences were performed using ClustalW (http://www.ebi.ac.uk/clustalw/) (34). The alignment was checked visually and corrected manually using the sequence editor. The homology in sequences identified by above methods is used to find out common ancestry.

After the alignment the sequences were subjected to distance matrix based on nucleotide sequence homology using the Kimura 2 parameter (35). Pairwise distance between and within the genotypes at the nucleotide levels were calculated with Kimura 2 parameters. Kimura 2 parameter is used to correct the rates of transition and transversion and to remove the base bias. Substitutions in base occurs randomly and the transitions (C to T or A to G) are more frequent than transversions (A to T and C to G). This is the method of distance corrections in a phylogenetic analysis. The last step followed for identification was the phylogenetic tree based on nearest neighbour joining method, which actually identified the homology of the organism. Trees were drawn, the distance matrix for which was generated by MEGA 3.1 software (35) (Fig. 3). The stability of branching pattern and the statistical significance of the tree topology were confirmed by bootstrapping (32). Bootstrap values indicate the number of times a node was supported in 1000 sampling replications (35). A cut-off of 97-98% similarity in 16S rRNA sequence was recommended as a criterion for demarcating species (37).

The isolate SPA and the strain Pedicoccus pentosaceus KC007 (Genbank Accession Number EU569832) clustered together (Fig. 3). The sequence similarity was calculated to be 96% (Table 1). Hence, the isolate SPA was identified as Pedicoccus pentosaceus and assigned the Genbank Accession Number EU569832. The closest homolog of the isolate was Pedicoccus sp. MMZ60A (EU157914).
Fig. 2. 16S rRNA gene sequence (1491 bp full length) of the isolate SPA.

Fig. 3. Phylogenetic Tree made in MEGA 3.1 software using Neighbour Joining method. The Phylogenetic tree demonstrates the relationship of isolate SPA to other reference Pediococcus species from NCBI and RDP.

<table>
<thead>
<tr>
<th>Identity</th>
<th>Alignment Results</th>
<th>Sequence Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolate SPA</td>
<td>0.98</td>
<td>Studied sample</td>
</tr>
<tr>
<td>EU157914</td>
<td>0.94</td>
<td>Pediococcus sp.strain:MMZ60A</td>
</tr>
<tr>
<td>DQ267152</td>
<td>0.94</td>
<td>Pediococcus pentosaceus</td>
</tr>
<tr>
<td>EU569832</td>
<td>0.96</td>
<td>Pediococcus pentosaceus strain:KC007</td>
</tr>
<tr>
<td>FJ844982</td>
<td>0.97</td>
<td>Pediococcus acidilactici strain:IMAU20070</td>
</tr>
<tr>
<td>EU147312</td>
<td>0.96</td>
<td>Pediococcus acidilactici strain:BFE 8384</td>
</tr>
<tr>
<td>FJ4055228</td>
<td>0.83</td>
<td>Pediococcus clausenii strain:Z15</td>
</tr>
<tr>
<td>EU331259</td>
<td>0.96</td>
<td>Pediococcus parvulus strain:Bpe301</td>
</tr>
<tr>
<td>FJ867642</td>
<td>0.91</td>
<td>Pediococcus ethanolidurans strain P2</td>
</tr>
<tr>
<td>AY956788</td>
<td>0.89</td>
<td>Pediococcus cellicola strain:Bpe260</td>
</tr>
<tr>
<td>EU331313</td>
<td>0.95</td>
<td>Pediococcus damnosus strain:Bpe260</td>
</tr>
</tbody>
</table>

16S rRNA based identification of Pediococcus pentosaceus
Other close homologs of the isolate SPA clustering separately are *Pediococcus pentosaceus* (DQ267152), *Pediococcus calssenii* ZJ5 (FJ405228), *Pediococcus parvulus* Bpe301 (EU331259), *Pediococcus ethnolidurans* P2 (FJ867642), *Pediococcus cellicola* Z-8 (AY956788), *Pediococcus damnosus* Bpe260 (EU331313). This isolate is farthest from the strain *Pediococcus acidilactici* IMAU20070 (FJ844982) and *Pediococcus acidilactici* BFE 8384 (EU147312) (Fig. 3). Distance matrix based on Nucleotide sequence homology (using Kimura-2 parameter) corroborated the above finding. Phylogenetic analysis of the complete 16S rRNA dataset resulted in trees with much greater resolution and well supported with high bootstrap pseudoreplicate score (1000 pseudoreplicates).

**Conclusion**

The isolate SPA was identified to the species level by full length sequence analysis of its 16S rRNA gene. The sequence analysis followed by alignment revealed the identity of the isolate SPA as *Pediococcus pentosaceus* (Genbank Accession Number EU569832). The genetic relationships of this isolate with its neighbours were traced and a phylogenetic tree was constructed. Based on this finding, it is concluded that direct sequence analysis of amplified 16S rRNA gene is a promising rapid and accurate method for species determination of lactic acid bacteria. This study, revealed the dextran production attribute of *Pediococcus* genus of lactic acid bacteria for the first time ever, heralding further rigorous investigation on this aspect. This discovery is expected to be a major break-through in the history of dextran production. This unique isolate SPA discussed in this work represented an emerging group of dextran producing *Pediococcus*. This investigation also unravelled the abundance of industrially valuable microbial flora in soil. Screening of the microbial biodiversity enriched North Eastern region of India needs proper attention. The meteoric rise in the usage of dextrans in food, pharmaceutical and cosmetics industries emphasizes the importance of exploration of the new strains and characterization of their traits.

**Acknowledgement**

This research work was supported by a project grant by the Department of Biotechnology, Ministry of Science and Technology, New Delhi, India, to AG. Fellowship to SP by Indian Institute of Technology Guwahati is acknowledged.

**References**


17. DellaGlio, F. and Torriani, S. (1986). DNA-DNA homology, physiological characteristics and distribution of lactic acid bacteria isolated...


16S rRNA based identification of *Pediococcus pentosaceus*