METAGENOMIC APPROACH TO ASSESS DIVERSITY OF BACTERIAL COMMUNITY IN SALINE *Pokkali* HABITATS OF KERALA

BY

SARVESHWAR SAH

(2009-11-105)

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR - 680 656

KERALA, INDIA

2012

Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala

By

Sarveshwar Sah

THESIS

Submitted in partial fulfilment of the

requirement for the degree of

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

Centre for Plant Biotechnology and Molecular Biology

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR - 680 656

KERALA, INDIA

2012

DECLARATION

I, hereby declare that this thesis entitled **"Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala"** is a bonafide record of research work done by me during the course of research and that it has not been previously formed the basis for the award to me of any degree, diploma, fellowship or other similar title, of any other University or Society.

Vellanikkara

Sarveshwar Sah
(2009-11-105)

CERTIFICATE

Certified that this thesis entitled **"Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala"** is a bonafide record of research work done independently by **Mr. Sarveshwar Sah** under my guidance and supervision and that it has not formed the basis for the award of any degree, diploma, fellowship or associateship to him.

Vellanikkara

Dr. D. Girija
(Major Advisor, Advisory Committee)
Professor and Head,
Department of Agricultural Microbiology
College of Horticulture
Vellanikkara, Thrissur

CERTIFICATE

We, the undersigned members of the advisory committee of **Mr. Sarveshwar Sah** a candidate for the degree of **Master of Science in Agriculture**, with major field in **Plant Biotechnology**, agree that the thesis entitled **"Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala"** may be submitted by **Mr. Sarveshwar Sah** in partial fulfillment of the requirement for the degree.

Dr. D. Girija
(Major Advisor, Advisory Committee)
Professor and Head,
Department of Agricultural Microbiology
College of Horticulture
Vellanikkara, Thrissur

Dr. P. A. Nazeem
(Member, Advisory committee)
Professor and Head
Centre for Plant Biotechnology and
Molecular Biology
College of Horticulture
Vellanikkara, Thrissur.

Dr. R. Keshavachandran
(Member, Advisory committee)
Professor
Centre for Plant Biotechnology and
Molecular Biology
College of Horticulture
Vellanikkara, Thrissur.

Dr. P. Suresh Kumar
(Member, Advisory Committee)
Professor, (Department of Soil Science and Agrl. Chemistry)
and Head of Radio Tracer Laboratory
College of Horticulture
Vellanikkara, Thrissur.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF APPENDICES

ACKNOWLEDGEMENT

In connection with this venture, I sincerely place on record my profound sense of gratitude and indebtedness to numerous people who have helped me in this project. And so comes the time to look back on the path traversed during the endeavour and to remember the faces behind the action with a sense of gratitude. Nothing of significance can be accomplished without the acts of assistance, words of encouragement and gestures of helpfulness from others.

First and foremost I bow my head to the ***God*** *who blessed me with the power and courage to complete this endeavour successfully in spite of the most difficult times face by me during the period of my study.*

With deep respect I express my whole hearted gratitude and unforgettable indebtedness to ***Dr. D. Girija,*** *Professor and Head, Department of Agricultural Microbiology and the chairperson of my advisory committee for her inimitable guidance, valuable suggestions, constructive criticism, unceasing interest, constant encouragement and above all the extreme patience, understanding and whole hearted co-operation rendered throughout the course of the study. I consider myself being fortunate in having the privilege of being guided by her.*

I express my sincere gratitude to ***Dr. P.A. Nazeem,*** *Professor and Head, CPBMB and member of my advisory committee for her valuable suggestions and guidance rendered me for the completion of the research programme and preparation of the thesis.*

I am grateful to ***Dr. R. Keshavachandran,*** *Professor, CPBMB and member of my advisory committee for his constructive suggestions and ever willing help rendered throughout this venture.*

I am deeply obliged to ***Dr. P. Suresh Kumar,*** *Professor, (Department of Soil Science and Agrl. Chemistry) Head of Radio Tracer Laboratory and member of my advisory committee for his unfailing support and enthusiasm, relevant suggestions and whole hearted cooperation throughout the period of investigation.*

It is my pleasant privilege to oblige ***Dr. M. R. Shylaja*** *Professor, CPBMB and member of my advisory committee, for his ardent interest, valuable suggestions, and critical scrutiny of manuscript and ever willing help which has helped a lot for refinement of this work.*

I also avail this opportunity to pay my sincere obligations and heartfelt thanks to ***Dr. P.A. Valsala****,* ***Dr. R. Sujath and Dr. P.S. Abida*** *of CPBMB for their encouragement and kind help offered at different stages of the study.*

I have no words to express my deep sense of gratitude to all my ***senior friends Sunil, Gajanan, Prathmesh, Arya, Alex, Rajesh, Sudha, Sivaji, Kiran, Suja*** *and* ***Rahul*** *for their valuable assistance and guidance extended throughout the investigation.*

Words cannot really express the true friendship that I relished from my batch mates ***Ashok, Rakesh, Achintya, Manjunath, Adarsh, Kannan, Elavarsen, Danish, Bishnu, Adnan, Pradeep, Liju, Lubina, Somya, Sahna, Keerti, Sneha, Anu, Lijitha, Sreeja, Chandana and Brinda*** *for the heartfelt help, timely suggestions and back-up which gave me enough mental strength to get through all mind-numbing circumstances.*

I have no words to express my deep sense of gratitude to ***my loving juniors Amar, Ashok, Pawan, Narender, Rohit, Pramod, Meera, Renju, Reshmi, Monika, Priya and Suraj*** *for their moral support and encouragement.*

I extend myriad of gratitude to my friends and research workers of CPBMB and Dept. of Agricultural Microbiology ***Deepa, Shidhi, Firoz, Lini, Glishma, Harish, Sunil, surya, Shylaja, Simi, Vidya, Anna, Geni, Ramya, Irin, Vinodh, Nancy, Anju*** *and* ***Niteen*** *for their kind co-operation and guidance during entire period of my study.*

I wish to express my sincere thanks to all the non-teaching staff members and labourers of CPBMB for their whole hearted cooperation and timely assistance.

I am deeply indebted to my ***Grandmother, Parents, Brothers*** *and my* ***Relatives*** *without their moral support, blessings and affection this would not have been a success. I, once again express my heartful thanks to all those who helped me in completing this venture.*

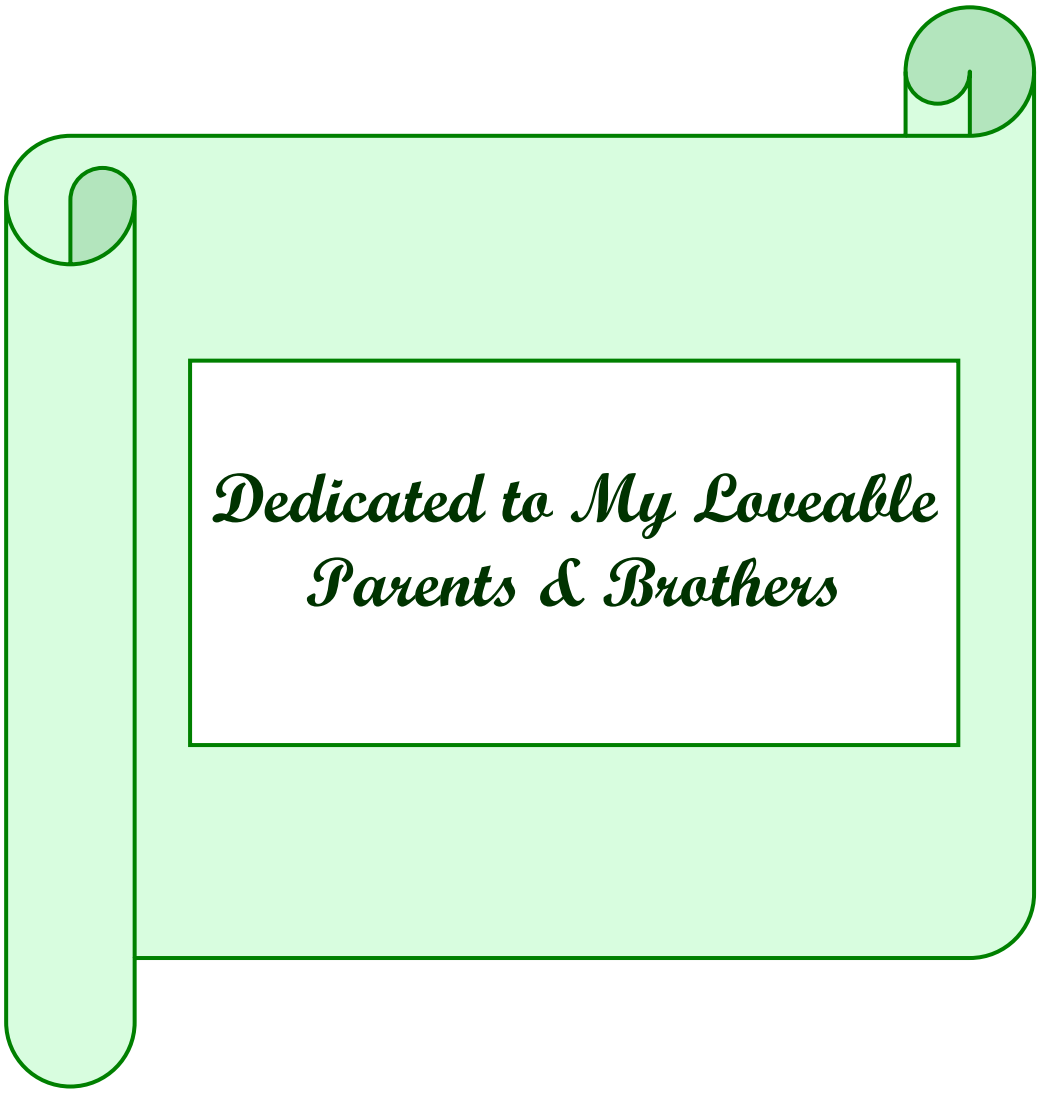
The award of fellowship by JNU is greatfully acknowledged.

Sarveshwar Sah

ABBREVIATION

A	Adenine
bp	Base pair
BLAST	Basic local alignment search tool
C	Cytosine
^{o}C	degree Celsius
cm	Centimeter
CPBMB	Centre for Plant Biotechnology and Molecular Biology
DMFO	Dimethyl formamide
DNA	Deoxyribo Nucleic Acid
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetra acetic acid
E- PCR	Exclusive PCR
G	Guanine
g	Gram
ICP	Insecticidal crystal protein
IPTG	Isopropyl thio galactoside
IPM	Integrated pet management
KAU	Kerala Agricultural University
kb	kilo base
kDa	kilo Dalton
LB	Luria Bertani
LC	Lethal concentration
M	Mole
MDa	Mega Dalton
mg	Milligram
min	minute
ml	Millilitre
mM	Millimole
μg	Microgram

µl	Microlitre
µM	Micromole
NCB1	National Centre for Biotechnology Information
ng	Nanogram
nm	Nanometer
OD	Optical Density
PCR	Polymerase chain reaction
PMSF	Phenyl methyl sulfonyl flouride
pH	Hydrogen ion concentration
%	Percentage
RNA	Ribo Nucleic Acid
RNase	Ribonuclease
rpm	Rotations per minute
SDS	Sodium dodecyl sulphate
sec	Second
SET	Sucrose, EDTA, Tris HCl
T	Thymine
TAE	Tris Acetate EDTA
TE	Tris EDTA
U	Unit
UV	Ultra violet
V	Volts
vip	Vegetative insecticidal protein
v/v	Volume by volume
w/v	Weight by volume
X-gal	5- bromo 4 chloro 3- indolyl β-D galactosidase


Dedicated to My Loveable
Parents & Brothers

Introduction

1. INTRODUCTION

The biosphere is dominated by microorganisms that have much practical significance in medicine, engineering and agriculture (Sloan *et al*., 2006). Due to their significance, genetic and biological diversity of microorganisms is an important area of scientific research (Ghazanfar and Azim, 2009). Biotechnology has become an important tool in the study of genetic and biological diversity of microbial communities. It is accepted that biotechnology has a continuous demand for novel genes, enzymes and compounds. Studies have demonstrated that natural diversity is the best supplier for these novel molecules. This can be explained by the vast richness of soil and other microbial niches (Schmeisser *et al.*, 2007). The global microbial diversity presents an enormous, largely untapped genetic and biological pool that could be exploited for the recovery of novel genes, biomolecules for metabolic pathways and various valuable products (Cowan, 2000).

Despite the obvious importance of microbes, very little is known of their diversity, for example, how many species are present in the environment and what each individual species does or its ecological function (Singh *et al*., 2008). Until recently, there were no appropriate techniques available to answer these important questions due to the limitations encountered in the culturing of microbes. Traditional methods of culturing microorganisms limit analysis to those that grow under laboratory conditions (Hugenholtz *et al*., 1998; Rondon *et al*., 2000). However, it is widely accepted that up to 99 per cent of the microbes in the environment cannot be readily cultivated (Hanada, 2003; Rappe and Giovannoni, 2003; Kamagata and Tamaki, 2005; Sekiguchi, 2006). Thus, most microbes have not been described and assessed for biotechnology.

To overcome the difficulties and limitations associated with cultivation techniques, different DNA-based molecular methods have been developed for characterizing microbial species and assemblages and these have significantly influenced our understanding of microbial diversity and ecology. In general, methods based on 16S rRNA gene analysis provide extensive information about the taxa and species present in an environment. The ability to recover and analyze 16S rRNA genes directly from environmental DNA provides a means to investigate microbial populations without the need to culture them (Ward *et al*., 1990; Amann *et al*., 1995; Hugenholtz *et al*., 1998; Dojka *et al*., 2000). The 16S rRNA gene

analysis has been used to study diverse bacterial and archaeal communities in extreme environments with a wide range of salinity, pH and temperature. These environments are interesting in terms of microbial diversity, to identify novel microorganisms and to understand the functioning of an ecosystem (Lizama *et al*., 2001). A few alkaline saline soils have been investigated, but only sediments of saline and alkaline lakes have been studied, e.g., Mono Lake in California (Humayoun *et al.*, 2003), saline meromictic Lake Kaiike in Japan (Koizumi *et al*., 2004), soda lakes in Inner Mongolia, China (Ma *et al*., 2004), saline and alkaline lakes of Wadi el Natrun in the Libyan Desert, Egypt, (Grant *et al*., 2004), athalassohaline lakes of the Atacama Desert, Chile (Demergasso *et al.*, 2004) and lakes in the Antarctic (Tindall, 2004).

However, these data usually provide only little information about the functional role of the different microbes within the community and the genetic information they contain of microbial niches (Streit and Schmitz, 2004). To overcome these difficulties, a new technology has been introduced since the late 1990s, known as 'Metagenomics'. Metagenomics is an emerging field in which the power of genomic analysis (the analysis of the entire DNA in an organism) is applied to entire communities of microbes, bypassing the need to isolate and culture individual microbial species. Metagenomics is a rapidly growing field of research that aims at studying uncultured organisms to understand the true diversity of microbes, their functions, cooperation and evolution, in environments such as soil, water, ancient remains of animals, or the digestive system of animals and humans (Ghazanfar and Azim, 2009).

Pokkali lands, known after the renowned salt tolerant land race of rice "*Pokkali*" are acclaimed for the unique way of reclamation and management of soil salinity and also for the integrated farming system involving rice - fish/prawn. It includes coastal paddy fields of Ernakulum, coastal villages of Thrissur, Alleppey and Kottayam districts. These areas are confluent with the Vembanad Lake through canals and subjected to tidal influence. *Pokkali* soils are tidal wet lands of Kerala, characterised by the accumulation of salts by tidal action over an underlying acidic soil.

The tides that occur twice a day play an important role on the fertility and productivity of this agro ecosystem. Tidal water contains high concentration of Na^+, K^+, Ca^{2+}, and Mg^+ that are important for the physiological processes in plant cells (Kramer, 1984).

A variety of microorganisms could be involved in the natural decomposition of organic matter in this area and many of them would be saline tolerant. However no systematic studies have been conducted till now on the microflora of *Pokkali* soils. Analysis of the microbial community in the *Pokkali* soils would help us to understand the mechanisms involved in tolerance to salinity.

Based on the above information, the present study entitled "Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala" was taken up with the objective to analyze the bacterial community, both culturable and non-culturable, in the saline *Pokkali* soils of Kerala through 16S rDNA sequencing approach.

Review of literature

2. REVIEW OF LITERATURE

The term “Metagenomics” was coined by Handelsman in the University of Wisconsin, Department of Plant Pathology in 1982 and it is a new field of research that allows the study of genomes recovered from environmental samples (Handelsman, 2004).

Metagenomics can be defined as "the application of modern genomics techniques to the study of communities of microbial organisms directly in their natural environments, bypassing the need for isolation and laboratory cultivation of individual species" (Chen and Pachter, 2005). It is the culture independent analysis of a mixture of microbial genomes. The Metagenomics methodology has been developed as an effective tool for the discovery of new natural products and microbial functions (He *et al*., 2007).

It can be described as the application of the genomics suit of technologies to uncultivated microorganisms. Initially; noncultured microflora and ancient DNA investigations were the prime targets of Metagenomics studies (Gabor *et al*., 2007; Singh *et al*., 2008). However, nowadays the technology is applied in the study of an array of microbial diversities like deep sea aquatic microflora, soil microbes and gastrointestinal ecosystems of humans and animals (Lu *et al*., 2007). Studies have revealed that only 0.001- 0.1 per cent of the total microbes in sea water, 0.25 per cent in freshwater, 0.25 per cent in sediments and only 0.3 per cent of soil microorganisms could be cultivable *in vitro* (Amann *et al.,* 1995). More than 99 per cent of the microorganisms present in certain environments cannot be cultured by standard techniques (Hanada, 2003; Rappe and Giovannoni, 2003).

The current Metagenomics studies have largely progressed due to the construction of efficient gene cloning vectors like Bacterial Artificial Chromosomes (BACs) or cosmids, (Xu, 2006; Babcock *et al*., 2007) which allow cloning and expression of larger and complex DNA segments or genes and the development of methods for generation and analysis of the data (Singh *et al*., 2008).

Although genomics has classically focused on pure, easy to obtain samples, such as microbes that grow readily in culture or large animals and plants, these organisms represent only a fraction of the living or once-living organisms of interest. Many species are difficult to study in isolation because they fail to grow in laboratory culture, depend on other organisms for critical processes, or have become extinct. Methods that are based on DNA sequencing

circumvent these obstacles, as DNA can be isolated directly from living or dead cells in various contexts. Such methods have led to the emergence of a new field, which is referred to as Metagenomics (Susannah and Edward, 2005).

2.1 Overview of saline soil problems and significance of halophilic bacterial study

Saline soil causes a serious problem of reduced agricultural production in over 100 countries especially in China, India, Pakistan, the United States, and Thailand.

2.1.1 Effects on plant growth

Saline soil causes both poor growth and poor yield of plants. Soluble ions in saline soil can affect plants by reducing the osmotic potential. Thus, salinity contributes to forces which reduce photosynthesis and transpiration by preventing water from entering plant roots (Waisel, 1972). In addition, salinity may reduce plant growth by the direct chemical effect of salt disrupting the nutritional and metabolic process in the plant and the indirect effect of salt altering the structure, permeability, and aeration of the soil. The effect of salinity on plant growth is affected by climate, soil conditions, agronomic practices, irrigation management, crop types and varieties, growth stage, and salt compositions (Spark, 1995). Salinity does not usually affect crop yield until the electrical conductivity exceeds a certain value for each crop. This is known as the threshold salinity level or the threshold EC value and it differs for various crops (Bresler *et al.*, 1982; Maas, 1990). In addition to the reduction in crop yield, salinity may also result in poor traffic ability, delayed seeding, and limited choice of crops.

2.1.2 Effects on soil structure properties

Soil salinity has critical effects on the permeability and infiltration of soil. When soil has high sodium ion concentration and low electrical conductivity then soil permeability, hydraulic conductivity, and the infiltration rate are decreased due to swelling and dispersion of clays and slaking of aggregates. Swelling causes the soil pores to become narrow, and slaking leads to reduction of the number of macrospores' through which water and solutes can flow, resulting in plugging of pores by the dispersed clay (Hansen *et al.*, 1999).

2.1.3 Significance of halophilic bacteria

Microorganisms in hypersaline environments need to balance their cytoplasm with the osmotic pressure exerted by the external medium. An osmotic balance can be achieved by the accumulation of compatible solutes. The solutes can be accumulated at high concentrations without interfering with cellular processes (Brown, 1976). However, when the turgor becomes too high, microorganisms need to excrete compatible solutes from their cytoplasm. Halophilic eubacteria accumulate compatible solutes by either *de novo* synthesis or uptake from the medium. Compatible solutes can be classified into 2 groups: inorganic ions (intracellular concentrations of inorganic cations: K^+, $Mg^{2+,}$ and Na^+) and organic solutes (betaines, ectoines, and glycine). The accumulation of inorganic ions, salt-in cytoplasm strategy, is found in the *Archaea* of the family *Halobacteriaceae* and the *Bacteria* of the order *Haloanaerobiales*. The accumulation of specific organic solutes, organic solutes strategy, is found in all other species of halophilic bacteria (Da Costa *et al.*, 1998). However, it may be a combined function of inorganic ions and organic solutes, since the inorganic ions are insufficient to provide osmotic balance with the external medium.

Halophilic bacteria provide a high potential for biotechnological applications for at least two reasons: (1) their activities in natural environments with regard to their participation in biogeochemical processes of C, N, S, and P formation and dissolution of carbonates, the immobilization of phosphate, and the production of growth factors and nutrients (Rodriguez-Valera, 1993), and (2) their nutritional requirements are simple. The majority can use a large range of compounds as their sole carbon and energy source. Most of them can grow at high salt concentrations, minimizing the risk of contamination. Moreover, several genetic tools developed for the nonhalophilic bacteria can be applied to the halophiles, and hence their genetic manipulation seems feasible (Ventosa *et al*., 1998).

2.2 *Pokkali* saline habitat of Kerala

Pokkali is a unique system of rice cultivation, taken up in the marshy, saline and acidic paddy fields of Ernakulum, coastal villages of Thrissur, Alleppey, and Kottayam districts. These areas are confluent with the Vembanad Lake through canals and subjected to tidal influence (Shylaraj and Sasidharan, 2005).

Rice cultivation is done usually during southwest monsoon- June to October when the salinity in the field is very low. After harvest of paddy when salinity builds up in the field from November onwards, the fields are utilized for traditional shrimp farming. The soil is inherently fertile with 3-4 per cent organic carbon content. The stubbles of the rice crop left in the field helps prawn cultivation. Taking advantage of the migratory habit of the shrimp larvae and juveniles, traditional method of shrimp culture popularly known as prawn filtration (chemmeenkettu) is prevalent in more than 4,500 hectares of low lying coastal backish water fields adjoining the Vembanad Lake in Kerala state (Menon, 1954; Muthu, 1978). *Pokkali* soil is generally sandy loam with low pH, high electrical conductivity and rich in available nutrients (George, 2011).

The high salt content coupled with very low pH and high concentration of iron and aluminium is responsible for the hazards encountered for rice production and several worker recommended reclamation measures. Several organic and inorganic amendments like pyrite, gypsum, sawdust, composed coir pith etc. have been attempted by Clarson *et al*. (1984).

A variety of microorganisms could be involved in the natural decomposition of organic matter in the area and many of them would be saline tolerant. However no studies have been conducted till now on the microflora of *Pokkali* soils. Analysis of the microbial community in the *Pokkali* soils would help us to understand the mechanisms involved in tolerance to salinity.

2.3 METAGENOMICS

Culture-independent genomic analysis of all the micro-organisms in a particular environmental niche (Handelsman *et al.,* 1998), evolved as an effort to discover more about the microbial diversity of natural environments such as soil, marine water and the gastrointestinal tracts of vertebrates and invertebrates. Metagenomics is a new and increasingly sophisticated field which in its simplest terms is concerned with the direct isolation of DNA from a defined habitat, followed by cloning (in a surrogate host such as *Escherichia coli*) of the complete genomes of the entire microbial population in that habitat (Langer *et al*., 2006).

2.3.1 Environmental Metagenomics

The first extensive large-scale environmental sequencing project was carried out by the J. Craig Venter Institute in 2004 in which they sequenced fragments of DNA derived from the entire microbial population of the nutrient-limited Sargasso Sea, an intensively studied region of the Atlantic Ocean close to Bermuda (Venter *et al*., 2004). A shot-gun sequencing approach yielded over 1.6 billion base pairs of DNA and led to the discovery of 1.2 million new genes. A total of 7, 94,061 of these genes were assigned to a conserved hypothetical protein group, the functions of which are unknown. The acid mine drainage microbiota was found to contain three bacterial and three archaeal species (Schoss and Handelsman, 2005). The three dominant bacterial genera included *Leptospirillum, Sulfobacillus, Acidomicrobium* and the dominant archaeal species was Ferraplasma acidomicrobium (Handelsman, 2004). The simplicity of the community structure allowed (Tyson *et al.*, 2004) to sequence almost the entire microflora with a high degree of coverage. It was noted that the G + C content of the genomes of the dominant taxonomic groups differed to a large extent and thus provided a means to source each of the clones (Bond *et al*., 2000).

2.3.2 Soil metagenomics

Tringle *et al.* (2005) analysis the soil Metagenomics library. Soil borne microorganisms are one of the earth's greatest source of biodiversity (Curtis *et al*., 2002), with estimates ranging between 3,000 and 11,000 microbial genomes per gram of soil (Schmeisser *et al*., 2007). Moreover, nearly 140 mega bases of sequence taken from Minnesota farmland soil contained <1per cent of sequences with any overlaps and formed no contigs (assemblies of overlapping individual clones), demonstrating that complete sequencing of highly diverse environments is virtually impossible with current technologies (Kowalchuk *et al*., 2007).

However, because of the huge diversity of soil and its history as a source of commercially important molecules in agriculture, chemical, industrial and pharmaceutical industries, it remains the most common target for studies of functional Metagenomics (MacNeil *et al*., 2001; Courtois *et al*., 2003; Daniel, 2005).

2.3.3 Marine water metagenomics

The marine environment has recently been pursued as a target of Metagenomics studies (Li and Qin, 2005). It is the largest contiguous ecosystem on earth, occupying 71 per cent of the earth's surface with an average depth of 4 km (Karl, 2007). A fosmid library was constructed from the 3000 m deep Mediterranean plankton and analysed by phylogenetic analysis of 16S rRNA genes and fosmid end sequencing. Sequence analysis revealed a high similarity with genomes from Rhizobiales within the Alphaproteobacteria, *Cenarchaem symbiosum*, Planctomycetes, Acidobacteria, Chloroflexi and Gammaproteobacteria (Martin-Cuadrado *et al*., 2007).

2.4 GUT METAGENOMICS

2.4.1 Insect gut metagenomics

In 2000 the complete genome of the fruit-fly (*Drosophila melanogaster)* was sequenced, pioneering an era of insect genomics (Adams *et al*., 2000; Myers *et al*., 2000). Further genomic research on insect maps; including the completion of over 30 additional whole genome sequences and more than 2 million ESTs, has effectively classified insects as the most diverse group of organisms on earth (Behura, 2006). Several Metagenomics studies have been carried out on the hindgut and midgut of the wood-feeding 'higher' termite (*Reticulitermes virginicus)* and the gypsy moth (*Lymantrisexia dispar*) respectively. Termites are widely known as economically important wood-degrading organisms (Sugimoto *et al*., 2000) with essential environmental roles in the turnover of carbon as well as serving as prospective sources of biochemical catalysts which can be used in the conversion of wood to biofuels (Warnecke *et al*., 2007). Significant data have recently emerged which suggests that the symbiotic bacteria resident in the hindgut of the termite play a functional role in the hydrolysis of both cellulose and xylan (Tokuda and Watanabe, 2007).

In order to gain a greater appreciation of the diversity of biological mechanisms governing lignocellulose degradation, Metagenomic analysis of the microbiota of the hindgut paunch of a wood-feeding 'higher' Nasutitermes (which do not contain hindgut flagellate protozoa) species was carried out in search of a large, complex set of bacterial genes commonly employed for cellulose and xylan hydrolysis (Warnecke *et al.,* 2007).

2.4.2 Human gut metagenomics

Microbial communities occupy all surfaces of the human body with a combined microbial cell population approximately 10 times that of human cells (Kurokawa *et al*., 2007). The distal colon has been identified as the most densely populated natural bacterial ecosystem (Marchesi and Shanahan, 2007; Frank and Pace, 2008). The human gastrointestinal microbiota is essential; bestowing metabolic functions that are otherwise absent in the host, such as improved strategies of energy harvest from ingested foods, synthesis of essential vitamins and the degradation of complex plant polysaccharides (Gill *et al*., 2006; Kurokawa *et al*., 2007). Indeed it is not uncommon for imbalances to occur in this intestinal microbial community structure with the potential for causing diseases such as Crohn's disease inflammatory bowel disease, allergy, obesity and cancer (Manichanh *et al*., 2006; Kurokawa *et al*., 2007).

13,335 16S rRNA sequences were produced from mucosal biopsy samples collected from the proximal to the distal colons of three healthy adults and one faecal sample from each person. The results yielded the largest database of 16S rRNA sequences (11,831 bacterial sequences, 1,524 archael sequences) from a single study of any environment with the identification of 395 bacterial phylotypes and one archael phylotype (Eckburg *et al*., 2005). The PCR amplified 16S rDNA products were cloned and sequenced bi-directionally, revealing that the vast majority of the 395 bacterial phylotypes discovered are members of the Bacteriodetes (48 per cent) and the Firmicutes (51 per cent); with the remaining phylotypes being represented by Proteobacteria, Verrucomicrobia, Fusbacteria, Cyanobacteia, Spirochetes.

2.5 METAGENOMIC APPROACHES

Metagenomics is a means of systematically investigating, classifying and manipulating the entire genetic material isolated from environmental samples (Zeyaullah *et al.*, 2009). The process involves isolating DNA from environmental sources and cloning it into vectors that replicate in cultured organisms. Metagenomics process relies on the efficiency of the four main steps, i.e.

1. Soil sampling and nucleic acids extraction

2.16S rRNA gene amplification

3. Library construction.

4. Analysis of Metagenomics sequences

2.5.1 Soil Sampling and nucleic acids extraction

In the Metagenomics process the samples could be analysed from any environment, soil or habitat including the Geographical Indication ecosystem (Ghazanfar and Azim, 2009). Specifically, soil microbial communities are composed of a mixture of archaea, bacteria and protists displaying a diversity of cell wall characteristics and varying in their susceptibility to lysis (Kauffmann *et al.*, 2004). Thus, some special techniques are required for their extraction. Although, various kits are commercially available for DNA isolation from environmental samples, many laboratories have developed their own methods with the aim of optimising extraction and reducing bias caused by unequal lysis of different members of the soil microbial community (Frostegard *et al.*, 1999; Krsek and Wellington, 1999; Miller *et al*., 1999).

There are two types of extraction techniques: (1) direct, in situ, extraction where the cells are lysed in the soil sample and then the DNA is recovered; and (2) indirect extraction techniques, where the cells are removed from the soil and then lysed for DNA recovery (Schmeisser *et al*., 2007). Soil is a particularly complex matrix containing many substances, such as humic acids, which can be co-extracted during DNA isolation. Removal of humic acids is essential before the DNA can be processed further. For this purpose, a range of DNA purification techniques have been developed. Sephadex G-200 spin columns have proven to be one of the best ways to remove contaminants from soil DNA (Miller *et al.*, 1999). Recently, a pulse field electrophoresis procedure using a two-phase agarose gel, with one phase containing polyvinylpyrrolidone (PVPP), was developed for removal of humics (Quaiser *et al*., 2002).

Siddhapura *et al.* (2010) studied DNA extraction and its quality assessment for PCR applications from saline soils of Coastal Gujarat and Sambhar Soda Lake, Rajasthan in India is described in a comparative manner. The mechanical and soft lysis methods were simple and efficient for rapid isolation of PCR amplifiable total genomic DNA. The results were

significant as only few extreme environments, particularly saline habitats are explored for their Metagenomics potential.

2.5.2 16S rRNA Gene Amplification

Winker and Woese, (1991) used the 16S rRNA gene sequences to study bacterial phylogeny and taxonomy has been by far the most common housekeeping genetic marker used for a number of reasons. These reasons include (i) its presence in almost all bacteria, often existing as a multigene family, or operons; (ii) the function of the 16S rRNA gene over time has not changed, suggesting that random sequence changes are a more accurate measure of time (evolution); and (iii) the 16S rRNA gene (1,500 bp) is large enough for informatics purposes.

Baker *et al.* (2003) demonstrated one basic approach is to identify microbes in a complex community by exploiting universal and conserved targets, such as rRNA genes. By amplifying selected target regions within 16S rRNA genes, microbes (specifically bacteria and archaea) can be identified by the effective combination of conserved primer-binding sites and intervening variable sequences that facilitate genus and species identification. The 16S rRNA gene in bacteria consists of conserved sequences interspersed with variable sequences that include 9 hypervariable i.e. v1 to v9 regions. The lengths of these hypervariable regions range from approximately 50 bases to 100 bases, and the sequences differ with respect to variation and in their corresponding utility for universal microbial identification.

Chakravorty *et al*. (2007) studied different hypervariable regions and demonstrated different efficacies with respect to species calls in different genera, and the V2 and V3 regions were most effective for universal genus identification.

Schmalenberger *et al*. (2001) studied parallel analysis of 3 different hypervariable regions of 16S rDNA sequence (V2–V3, V4–V5, and V6–V8 regions) was effective in determining the composition of bacterial consortia in maize rhizospheres.

2.5.3 Construction of a Metagenomic library

DNA isolation and purification is followed by the construction of DNA libraries in suitable cloning vectors and host strains. The classical approach includes the construction of small insert libraries (<10 kb) in a standard sequencing vector and in *Escherichia coli* as a host strain (Henne *et al*., 1999). However, small insert libraries do not allow detection of large gene clusters or operons. To circumvent this limitation researchers have been employing large insert libraries, such as cosmid DNA libraries (mostly in pWE15 vector of Stratagene) with insert sizes ranging from 25-35 kb (Entcheva *et al*., 2001) and/or Bacterial Artificial Chromosome (BAC) libraries with insert up to 200 kb (Beja *et al*., 2000; Rondon *et al*., 2000). Additionally, the construction of fosmid with inserts of 40 kb of foreign DNA has been reported (Beja et al., 2000). *E. coli* is still the preferred host for the cloning and expression of any metagenome-derived genes and only very recently have other hosts such as *Streptomyces lividans* been employed to identify genes involved in the biosynthesis of novel antibiotics (Courtois *et al*., 2003). Metagenomic libraries are also being developed in other Gram-negative hosts by several laboratories, and these will become available soon.

2.5.4 Analysis of Metagenomic libraries

Two methods are used for the analysis of genetic material of metagenomic library.

2.5.4.1 Sequence-based Metagenomics

Sequence-based approaches, involve screening clones for the highly conserved 16S rRNA genes for identification purposes and then sequencing the entire clone to identify other genes of interest, or large-scale sequencing of the complete metagenome to search for phylogenetic anchors in the reconstructed genomes (Riesenfeld *et al*., 2004).

Sequence-based Metagenomic studies provide information on the distribution of functions in a community, linkage of traits, genomic organization and horizontal gene transfer. Approaches typically involve either sequencing of random clones to accumulate vast stores of sequence information or identification of clones based on methods that detect a particular sequence. With both of these approaches, phylogenetic markers are sought on the clone of interest to link cloned sequences with the probable origin of the DNA.

2.5.4.2 Function-based analysis

Functional Metagenomic involves screening Metagenomic libraries for a particular phenotype, e.g. salt tolerance, antibiotic production or enzyme activity, and then identifying the phylogenetic origin of the cloned DNA. Function-based analysis enables identification of new enzymes, antibiotics or other reagents in libraries from diverse environments. Approaches include: (i) heterologous expression, in which clones that express the desired function are identified. An important limitation to heterologous expression is that the domesticated host bacterium must be able to express (transcribe and translate) the genes for the products to be detected; (ii) Selections, in which the clone expressing the desired function grows and others do not. Selections provide the most powerful approach to finding rare clones. Examples of selectable characteristics include antibiotic resistance and metal resistance.

A 'functional-anchor approach' involves identifying all of the clones that express a certain function and sequencing them completely to determine the diversity of genomic environments from which that function originates.

2.6 Metagenomic study of bacterial diversity in soil

Martín-Cuadrado *et al.* (2007) constructed a metagenomic fosmid library from 3,000m deep Mediterranean plankton, which is much warmer (~14°C) than waters of similar depth in open oceans (~2°C). It was analyzed the library both by phylogenetic screening based on 16S rRNA gene amplification from clone pools and by sequencing both insert extremities of 5,000 fosmids. Genome recruitment strategies showed that the majority of high scoring pairs corresponded to genomes from Rhizobiales within the Alphaproteobacteria, *Cenarchaeum symbiosum*, Planctomycetes, Acidobacteria, Chloroflexi and Gammaproteobacteria.

Hu *et al.* (2010) Preliminary characterization of the microbial phylogeny and metabolic potential of a deep-sea sediment sample from the Qiongdongnan Basin, South China Sea, was carried out using a Metagenomics library approach. An effective and rapid method of DNA isolation, purification, and library construction was used resulting in

approximately 200,000 clones with an average insert size of about 36 kb. End sequencing of 600 individual clones from the fosmid library generated 1,051 sequences with an average sequence length of 619 bp. Phylogenetic ascription indicated that this library was dominated by Bacteria, predominately Proteobacteria, though Planctomycetes were also relatively abundant. Sulfate-reducing and anaerobic ammoniumoxidizing bacteria, which play important roles in the cycling of sedimentary nutrients, were abundant in the library.

Litchfield and Gillevet, (2002) studied the microbial diversity of the salterns in Shark Bay, Australia, and Eilat, Israel, based on the Amplicon Length Heterogeneity (ALH) procedure with the help of 16S rDNA primers labelled with fluorescent dyes.

Cytryn *et al.* (2000) assessed diversity of *Archaea* in the Solar Lake, Sinai, Egypt, using the direct DNA extraction from the water samples. The archaeal primers, 21F (5'-TTCCGGTTGATCCYGCCGGA-3') and 958R (5'-YCCGGCGTTGAMTCCAATT-3') were employed for the amplification of 16S rDNA fragments. The PCR products were cloned and sequenced. The diversity of *Archaea* in the Solar Lake was assessed by phylogenetic analysis of detected sequences using the Accounting Research Bulletins (ARB) program package.

Antón *et al.* (2000) assessed the halophilic bacterial community in the saltern ponds at Alicante, Spain, using the DGGE analysis. Bacterial cells were collected from the water samples by centrifugation. The 16S rDNA were amplified and subsequently separated in a polyacrylamide gel containing a linear gradient of DNA denaturants. The selected DGGE bands were excised from the gel, purified, reamplified, and partially sequenced. The sequences were analyzed by BLAST at the National Center for Biotechnology Information web page and aligned using the Accounting Research Bulletins (ARB) program package.

2.7 Application of soil Metagenomics

The development and application of Metagenomics has enabled access to the uncultivated soil microbial community, availing a rich source of novel and useful biomolecules.

2.7.1 Antibiotics and pharmaceuticals

Soil Metagenomic has the potential to substantially impact on antibiotic production. Two previous studies reported the successful screening of soil Metagenomic libraries for indirubin (MacNeil *et al*., 2001; Lim *et al.,* 2005) while a range of novel antibiotics have been detected in Metagenomic libraries (Gillespie *et al.,* 2002; Brady *et al.,* 2004). A clone found in a soil Metagenomic library produces deoxyviolacein and the broad spectrum antibiotic violacein (Brady *et al.,* 2001). During the screening of seven different soil Metagenomics libraries, it was reported that these libraries exposed 11 clones producing longchain N-acyltyrosine antibiotics, and analysis of their synthesis indicated that ten of them were novel (Brady *et al.,* 2004). Metagenomics libraries have also been used for isolating natural antibiotic resistance genes. Riesenfeld *et al.* (2004) identified nine aminoglycoside and one tetracycline antibiotics resistance genes from soil.

2.7.2 Oxidoreductases/dehydrogenases

A Metagenomics study searching for the diversity of bacteria in the environment capable of utilising 4-hydroxybutyrate found five clones displaying novel 4-hydroxybutyrate dehydrogenase activity (Henne *et al*., 1999). Alcohol oxidoreductases capable of oxidising short chain polyols are useful biocatalysts in industrial production of chiral hydroxy esters, hydroxy acids, amino acids and alcohols (Knietsch *et al*., 2003). In a Metagenomic study without enrichment, a total of 24 positive clones were obtained and tested for their substrate specificity. To improve the detection frequency, an enrichment was performed using glycerol or 1, 2-propanediol, after which a further 24 positive clones were detected (Knietsch *et al.*, 2003).

2.7.3 Amidases

In a study involving general screening of a soil Metagenomic library for biocatalysts, one amidase-positive clone was detected (Voget *et al*., 2003). Amidases are used in the biosynthesis of 9-lactam antibiotics. A separate study targetting amidases of the soil

Metagenomic using enrichment detected seven amidase-positive clones, one of which encoded a novel penicillin acylase (Gabor *et al.*, 2007; Gabor and Janssen, 2004).

2.7.4 Vitamin biosynthesis

Soil Metagenomics has been applied to the search for novel genes encoding the synthesis of vitamins such as biotin (Entcheva *et al*., 2001). Seven cosmids were detected in Metagenomics libraries obtained after avidin enrichment of environmental samples and the highest levels of biotin production in this study were detected in a cosmid obtained from forest soil (Entcheva *et al*., 2001).

2.7.5 Polysaccharide degrading/modifying enzymes/ amylolytic genes

Amylases have been the focus of many Metagenomics studies with several reports available on the isolation of novel amylolytic enzymes from Metagenomics DNA libraries (Richardson *et al*., 2002; Voget *et al*., 2003; Yun *et al*., 2004; Ferrer *et al*., 2005). Celluloses have numerous applications and biotechnological potential for various industries including chemicals, fuel, food, brewery and wine, animal feed, textile and laundry, pulp and paper and agriculture (Bhat, 2000; Sun and Cheng, 2002; Wong and Saddler, 1992). Functional screening of a soil Metagenomics library for cellulases revealed a total of eight cellulolytic clones, one of which was purified and characterized (Voget *et al*., 2006). Agarases, the enzymes that can liquify agar, have been identified during the screening a soil Metagenomics library, in which a total of 4 agarolytic clones containing 12 agarase genes were identified (Voget *et al*., 2003).

2.7.6 Lipolytic genes

Metagenomics has identified a number of novel genes encoding lipolytic enzymes such as esterases and lipases. Esterase EstCE1 was derived from a soil metagenome (Elend *et al*., 2006), and this enzyme displays remarkable characteristics that cannot be related to the original environment from which they were derived. The high level of stability of this enzyme together with its unique substrate specificities make it highly useful for biotechnological applications. Environmental DNA libraries prepared from three different soil samples were screened for genes conferring lipolytic activity on *E. coli* clones. Screening

on triolein agar revealed 1 positive clone out of 7,30,000 clones, and screening on tributyrin agar revealed 3 positive clones out of 2,86,000 *E. coli* clones. Substrate specificity analysis revealed that one recombinant strain harbored a lipase and the other three contained esterases. The genes responsible for the lipolytic activity were identified and characterized (Henne *et al*., 2000). Further screening identified genes conferring Na (Li)/H antiporter activity on the antiporter-deficient Escherichia coli strain KNabc (Majernik *et al*., 2001). This powerful selection facilitated the discovery of two novel antiporter proteins in a library of 14, 80,000 clones containing DNA isolated from soil.

Table 1: Details of recently identified lipolytic and metagenome-derived biocatalysts and the origin of metagenome samples

Target gene / natural product	**Source**	**Vector used for library construction**	**Strain**	**Authors**
Lipase, Esterase	Meadow soil	pBluescript SK	*E. coli* DH5	Henne *et al*., 2000
Antibacterial, Hemolytic activities, Lipase, amylase, nuclease	North American soil	pBeloBAC11	*E.coli* DH10B	Rondon *et al.*, 2000
4-hydroxybutyrate dehydrogenase	Soil	pBluescript SK	*E.coli* DH5 Alpha	Henne *et al.*, 1999
H+ antiporters	Soil	pBluescript SK	*E.coli* KNabc	Majernik *et al*., 2001
Unique lipolytic activity	Forest topsoil	Fosmid	*E. coli*	Lee *et al*., 2004
Novel biocatalysts	unplanted field soil	Cosmid	*E. coli*	Voget *et al.*, 2003
Alcohol oxidoreductase meadow	sugar beet field, cropland soil	pBluescript SK	*E. coli* DH5 Alpha	Knietsch *et al*., 2003
Carbonyls	formation soil	n.r	*E. coli*	Knietsch *et al.*, 2003

Coenzyme B(12)-dependent glycerol and diol dehydratases	Soil	pBluescript SK	*E. coli* DH5 Alpha	Knietsch *et al*., 2003
Esterase	Soil	pCCIFOS	*E. coli* EP1300-TI	Gabor *et al*., 2007
Lipase	Soil	pEpiFOS-5	*E. coli* EP1-100	Li and Qin, 2005
Amidase	Soil and enrichment	pZero-2	*E. coli* TOP10	Gabor *et al*., 2007
Polyketide syntase	Soil	Cosmid	*E. coli*	Gillespie *et al*., 2002
Aminoglycoside and tetracycline antibiotic resistance	Soil	pJN105 pCF430	DH10B DH5 α	Riesenfeld *et al*., 2004
B-deoxyviolacein and broad spectrum antibiotic Violacein	Soil	Cosmid	*E. coli*	Brady *et al*., 2004
Amylase	Soil	pUC19	*E. coli* DH5 alpha	Yun *et al*., 2004

2.8 Characterization of culturable bacterial isolates from saline soil

Echigo *et al*. (2005) demonstrated that many halophilic bacteria that are able to grow in the presence of 20 per cent NaCl inhabit in saline environments such as Dead Sea, the Great Salt Lake soils, yards, fields and roadways in an area surrounding Tokyo, Japan. Analyses of partial 16S rRNA gene sequences of 176 isolates suggested that they were halophiles belonging to genera of the family *Bacillaceae*, *Bacillus* (11 isolates), *Filobacillus* (19 isolates), *Gracilibacillus* (6 isolates), *Halobacillus* (102 isolates), *Lentibacillus* (1 isolate), *Paraliobacillus* (5 isolates) and *Virgibacillus* (17 isolates). Sequences of 15 isolates showed similarities less than 92 per cent, suggesting that they may represent novel taxa within the family *Bacillaceae*.

Halophilic microorganisms are adapted to conditions of high salinity and require a certain concentration of NaCl for their optimum growth (Kushner and Kamekura, 2006).

They have been isolated from various saline environments such as salt lakes (eg. the Dead Sea, the Great Salt Lake), salterns, solar salts and subsurface salt formation.

Extremely halophilic microorganisms require high concentration of NaCl for their growth, with optimum concentrations of 2.5-5.2 M (15-30 per cent). *Haloarcula vallismortis* and *Haloterrigena turkmenica* for example, have been isolated from salt pool of Death Valley, California, and saline soil of Turkmenia, respectively (Gonzalez *et al*., 1978; Zvyagintseva and Tarasov, 1987).

Moderate halophiles are defined as those that grow optimally in media containing 0.5–2.5 M (3–15 per cent) NaCl, such as *Halomonas maura* isolated from a saltern in Morocco, and *Marinococcus halophilus* isolated from sea sands (Bouchotroch *et al*., 2001 and Hao *et al*., 1984). Halotolerant microorganisms possess the ability to grow in media without added NaCl and also in the presence of high concentrations of NaCl. For example, *Halobacillus salinus* isolated from a salt lake in Korea is able to grow without added salt and in media containing up to 23 per cent NaCl (Yoon *et al*., 2003).

Do halophiles inhabit non-saline environments such as garden soil, yards and field? *Bacillus clarkii*, *B. agaradhaerens* and *B. pseudofirmus* are examples of halotolerant bacteria isolated from soil samples that were shown to be tolerant up to 16 per cent or 17 per cent NaCl (Nielsen *et al.*, 1995). It has, however, been tacitly believed that habitats of halophiles able to grow in media containing higher concentrations, let's say 20 per cent (3.4 M), are restricted to saline environments.

Onishi *et al*. (1980) no reports have been published on the isolation of microorganisms able to grow at 20 per cent or higher NaCl concentrations from ordinary, non-saline soil samples. In 1980 Onishi *et al*. surveyed extensively the occurrence of halophilic bacteria in more or less saline samples collected in Japan. They adopted enrichment culture in a medium containing 4 M (23.4 per cent) NaCl, a customary concentration for the cultivation of *Halobacterium* spp. They isolated 168 strains finally, but no enrichment was obtained from one third of 287 samples of sea sands and seaweeds collected on seashore. They did not include ordinary garden soil samples. It should be pointed out that a non-pigmented haloarchaeon strain 172P1 designated later as *Natrialba asiatica*.

(Kamekura and Dyall-Smith, 1995) was isolated during their survey from dry beach sands with granular salts attached.

Akhtar *et al*. (2008) demonstrated that culturable bacterial biodiversity and industrial importance of the isolates indigenous to Khewra salt mine, Pakistan was assessed. PCR Amplification of 16S rDNA of isolates was carried out by using universal primers FD1 and rP1. These gene sequences were compared with other gene sequences in the GenBank databases to find the closely related sequences. Most of the isolates belonged to different species of genus *Bacillus*, sharing 92-99 per cent 16S rDNA identity with the respective type strain. Other isolates had close similarities with *Escherichia coli*, *Staphylococcus arlettae* and *Staphylococcus gallinarum* with 97, 98 and 99 per cent 16S rDNA similarity respectively. The abilities of isolates to produce industrial enzymes (amylase, carboxymethylcellulase, xylanase, cellulase and protease) were checked. All isolates were tested against starch, carboxymethylcellulose (CMC), xylane, cellulose, and casein degradation in plate assays. BPT-5, 11, 18, 19 and 25 indicated the production of copious amounts of carbohydrates and protein degrading enzymes. Based on this study it can be concluded that Khewra salt mine is populated with diverse bacterial groups, which are potential source of industrial enzymes for commercial applications.

Kiel and Gaylarde, (2007) demonstrated that growth on medium with no added salt was generally confluent and uncountable. Samples were not diluted as this would have influenced the total diversity detected. Many of the bacteria growing on these plates have been shown to be of the genus *Bacillus*, although Gram-negative species were also found (Kiel and Gaylarde, 2007). On salt-containing plates, countable colonies, both bacterial and fungal, were produced from all buildings, confirming the presence of halotolerant microorganisms. Salt-tolerant bacteria have been detected by both culture and molecular biological techniques on European monuments (Saiz-Jimenez and Laiz, 2000; Heyrman and Swings, 2001). Pin *et al*. (2001) reported the presence on a 14th century Austrian castle of a small number of halophilic bacterial types able to grow on up to 20 per cent salt, optimally on 5–10 per cent. They were identified by DNA analysis as new members of the genus *Halobacillus.*

Tiquia *et al*. (2007) demonstrated that the use of sodium chloride to melt highway and road snow is believed to have a significant effect on the groundwater ecosystem of the rivers

where the salt from the roads drain. As the river composition changes, the bacterial population also changes to favour those bacteria that are more suited to the higher salt concentrations. In this experiment, it was surveyed the cultivable salt-loving organisms (halophilic) on three sites that encompass the Rouge River. 16S rDNA sequences were analyzed and compared with sequences from Genbank. Results indicated that the SSU rRNA sequences of the bacterial isolates were similar to six major genera, *Bacillus*, *Staphylococcus*, *Halobacillus*, *Paenabacillus*, *Halomonas*, and *Clostridium.* Half of the isolates sequenced were similar to *Bacillus* spp. The Active Pharmaceutical Ingredient (API) assay showed that the majority of the isolates were positive for the enzymes tryptophane deaminase, gelatinase and β-galactosidase. The primary enzyme found in all isolates was arginine dihydrolase, which might be an indicator of the presence of such enzyme in halophilic and halotolerant bacteria present in the Rouge River.

Sorokin *et al.* (2008) demonstrated that the isolate had clostridia-like motile cells and formed ellipsoid endospores. It was able to fix dinitrogen gas growing on nitrogen-free alkaline medium. Strain MS 6T was a strictly fermentative bacterium without a respiratory chain, although it had a high catalase activity and tolerated aerobic conditions. It was an obligate alkaliphile with a pH range for growth between 7.5 and 10.6 (optimum at 9.0–9.5). Growth and nitrogen fixation at pH 10 were possible at a total salt content of up to 1.2 M Na+ (optimum at 0.2–0.3 M). 16S rRNA gene sequencing identified strain MS 6T as a member of the genus *Bacillus*.

Lawson *et al.* (2008) showed that the taxonomic position of a novel halophilic endospore-forming bacterium previously isolated from a desert iguana was investigated by 16S rRNA gene sequencing. Comparative sequence analyses showed the unidentified bacterium to be phylogenetically loosely associated with some other spore-forming (*Bacillus pantothenticus*, *Sporosarcina halophila*) and non-spore-forming (*Marinococcus albus*) halotolerant bacteria. Based on the phenotypic and phylogenetic distinctiveness of the unidentified bacterium, it is proposed that it is classified in the genus *Bacillus* as a new species, *Bacillus dipsosauri.*

Egamberdieva and Kucharova, (2009) found that high salinity of soils in arid and semi-arid regions results in desertification and decreased crop yield. One possibility to circumvent this problem was to use root colonising salt tolerant bacterial inoculants which

can alleviate salt stress in plants. Five wheat root tip coloniser bacteria were selected from the rhizosphere of wheat grown in saline soil and were identified by the 16S rRNA gene sequence as *Pseudomonas putida*, *Pseudomonas extremorientalis*, *Pseudomonas chlororaphis* and *Pseudomonas aurantiaca.* The isolates tolerated salt of 5 per cent NaCl and produced indole acetic acid under saline conditions. Four isolates proved to be very efficient in promoting a significant increase in the shoot, root and dry matter of wheat and were able to survive in saline soil. Four of the isolated strains appeared to be better competitive colonisers than reference strains and probably outcompeted with indigenous microorganisms of the rhizosphere. These results are promising for the application of selected environmentally save microbes in saline agricultural soils.

Materials and methods

2. MATERIALS AND METHODS

The study entitled "Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala" was carried out at the Centre for Plant Biotechnology and Molecular Biology and the Department of Agricultural Microbiology, College of Horticulture, Vellanikkara during the period 2009-2011. The materials used and the methodologies adopted in this study have been furnished in this chapter.

3.1 MATERIALS

3.1.1 Chemicals, Glassware and Plastic ware

The chemicals used for the study were of good quality (AR grade) from various agencies like MERCK, SRL and HIMEDIA. Molecular Biology Grade enzymes and buffers were supplied by Bangalore Genie Ltd. All plastic wares used were obtained from Axygen and Tarson India Ltd.

3.1.2 Equipment Items

The equipment items available at Centre for Plant Biotechnology and Molecular Biology were used for the present study. Incubation of cultures was done in Incubator Shaker, DK-S1010 (Dai Ki Scientific Co., Korea). Centrifugation was done in KUBOTA centrifuge (Kubota, Japan). A compound binocular microscope (CETI) was used for viewing the slides of bacterial cultures. DNA amplification was carried out in Eppendorf Master Cycler, Gradient (Eppendorf, Germany). Vision works LS software was used to visualize the gel and UVP GelDoc-IT^{TM} imaging system (USA) was used for imaging the gel.

3.2 Environmental soil sampling and storage

The soil samples were collected from Pizhala, Vytilla, Cochin South town, Kerala after the rice crop when the field was submerged under saline water and prawn was being grown. Four soil samples were collected from top 6 cm from the field in sterile plastic bags (HiDispo bag-12, Himedia Pvt. Ltd., Mumbai), during the month of January 2010. These were composited into one sample, transported to the laboratory and stored at 4°C. This

composite soil sample was further used for the analysis of electrochemical properties and for metagenomic DNA extraction.

3.3 Estimation of EC, TDS and pH

10g of air dried composite soil sample was suspended in 25ml of freshly prepared distilled water for 30min. EC was measured in conductivity meter (in dsm-1). Total dissolved salt was calculated following the method of Rayment and Higginson (1992).

Total dissolved salt (TDS) = EC × 640× 2.5

The clear supernatant of 1:2.5 soil water suspension was prepared. Measurement of pH was recorded potentiometrically by using pH meter (Jackson, 1958).

3.4 DNA extraction from *Pokkali* soil

Direct method of DNA extraction by soft lysis method suggested by Siddhapura *et al.* (2010) was used. Chemical composition of reagents is given in Appendix II.

Procedure

1. Soil sample (1 g) in duplicate was suspended in 10 ml of extraction buffer and incubated at 37^0C for 10–12 h under shaking at 150 rpm.
2. Sample was re-extracted in 1ml of extraction buffer and supernatant collected by low speed centrifugation (5000 rpm) for 10 min.
3. 4 ml of lysis buffer was added and incubated at 65^0C for 2 h with vigorous shaking at every 15 min.
4. Sample was centrifuged at 10,000 rpm for 10 min at 4^0C. The upper aqueous phase was extracted with equal volume of phenol: chloroform: isoamylalcohol (25:24:1) at 10,000 rpm for 20 min at 4^0C.
5. Upper aqueous phase was again extracted with equal volume of chloroform: isoamylalcohol (24:1) at 10,000 rpm for 10 min at 4^0C.
6. DNA preparation was further treated by adding 1/10 volume of 7.5M potassium acetate and subsequently precipitated by adding 2 volumes of chilled ethanol.
7. DNA precipitate was collected by centrifugation at 10,000 rpm for 10 min, air dried and suspended in 50 µl sterile distilled water.

3.5 Agarose Gel Electrophoresis

3.5.1 Preparation of Agarose gel

Agarose gel electrophoresis was performed based on the method described by Sambrook *et al*. (1989) to check the quality of DNA and also to separate the amplified products.

Materials used for agarose gel electrophoresis were:

1. Agarose (Bangalore Genei, Low EEO)
2. 50X TAE buffer (pH 8.0)
3. Electrophoresis unit (Biorad, USA), power pack, casting tray and comb
4. 6X Loading/ Tracking dye (Bangalore Genei, Bangalore)
5. Ethidium bromide solution (stock 10mg/ml; working concentration 0.5µg/ml (5µl/100ml gel))
6. Gel documentation and analysis system (UVP GelDoc-IT^{TM} imaging system)

(Chemical composition of the buffer and dyes is given in Appendix IV)

3.5.2 Electrophoresis

1. 1X TAE buffer was prepared from 50 X TAE stock solutions.
2. Agarose (1.0 per cent (w/v) for genomic DNA and 0.8 per cent (w/v) for PCR) was weighed and added to 1X TAE. It was boiled till the agarose dissolved completely and then cooled to lukewarm temperature.
3. Ethidium bromide was added to a final concentration of 0.5µg/ml as an intercalating dye of DNA and mixed well.
4. The open ends of the gel casting tray were sealed with a cellophane tape and placed on a perfectly horizontal leveled platform.
5. The comb was placed properly and molten agarose was poured into the tray, allowing it to solidify.
6. After the gel was completely set (15 to 20 minutes at room temperature), the comb and cellophane tape were carefully removed.

7. The gel was placed in the electrophoresis tank with the wells near the cathode and submerged in 1X TAE to a depth of 1cm.
8. A piece of cellophane tape was pressed on a solid surface and 1µl 6X loading dye was dispensed in small quantity on the tape. Five µl of DNA sample (in the case of PCR products, 10.0 to15.0µl) was mixed well with the dye by pipetting in and out for 2 to 3 times. Then the mixture was loaded in the wells, with the help of micropipette. λDNA/*Eco*R1+ *Hind*III double digest (Bangalore Genei) was used as the molecular weight marker.
9. The cathode and anode were connected to the power pack and the gel was run at a constant voltage of 60 volts.
10. The power was turned off when the tracking dye reached at about 3cm from the anode end.

3.5.3 Gel Documentation

The DNA bands separated by electrophoresis were viewed and photographed using Vision works LS software and UVP GelDoc-IT^{TM} imaging system.

3.5.4 Assessing the Quality and Quantity of DNA

The purity of DNA was further analyzed by using NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA). Before taking sample reading, the instrument was set to zero by taking 1µl autoclaved water as blank. The absorbance of nucleic acid sample was measured at a wavelength of 260 nm and 280 nm. 260/280 ratio was used to assess the purity of nucleic acids. A ratio of 1.8 to 2 indicated that the DNA preparation was pure and free from protein.

3.6 PCR Reaction

3.6.1 Amplification of 16S rRNA gene

The total DNA extracted from high saline soil of *Pokkali* field isolates was amplified by PCR using gene specific primers in an Eppendorf Master Cycler, Gradient (Eppendorf, Germany).

3.6.2 Composition of the reaction mixture for PCR

1. Genomic DNA (50ng)	-	1.0µl
2. 10X Taq assay buffer	-	2.5µl
3. d NTP mix (10mM)	-	1.0µl
4. Forward primer (10pM)	-	1.0µl
5. Reverse primer (10pM)	-	1.0µl
6. Taq DNA polymerase (0.3 U)	-	2.0µl
7. Autoclaved distilled water	-	16.5µl
		25.0µl

The reaction was set in a 200µl microfuge tube chilled over ice flakes. A momentary spin was given to mix completely all reagents and set in thermal cycler for amplification under suitable programme. A negative control was maintained without any template DNA.

3.6.2.1 Primer description: DNA preparations from the saline soil were used as template to amplify the region encoding 16S rRNA gene.

The details of primers (Siddhapura *et al.,* 2010) used for 16S rDNA amplification from soil metagenomic DNA are given in Table 2.

Table 2. Details of primers used for 16S rDNA amplification

Sl.No	Nucleotide sequences	No. of nucleotides	Annealing temperature (^{0}C)	Expected amplicon size (kb)
1.	Forward (5 -AGA GTT TGA TCC TGG CTC AG- 3)	20	55	1.5
2.	Reverse (5 -ACG GCT ACC TTG TTA CGA CTT- 3)	21		

3.6.3 Thermal cycler programme

Steps followed in PCR, temperature profile and time duration of each step for amplification of genes are given in table 3. The PCR product was checked on 0.9 per cent agarose gel and documented.

Table 3.Temperature profile for amplification of gene by PCR

Sl. No	Step	Temperature	Time (min)
1.	Initial denaturation	95°C	3.00
2.	Denaturation	94°C	1.30
3.	Annealing	55°C	0.40
4.	Primer extension	72°C	1.30
5.	Step 2 to 4	34 cycles	
6.	Final extension	72°C	10.00

3.7 Gel Elution of PCR amplified fragments

Products obtained in PCR reactions were loaded separately on 0.8 per cent (w/v) agarose gel and the desired band in each case was eluted by using AxyPrep DNA Gel Extraction Kit (Axygen Biosciences Pvt. Ltd). Following are the procedure as per the manufacturer's guidelines.

1. The DNA fragment of interest was excised from the gel using a sterile, sharp scalpel while avoiding much exposure to UV on a transilluminator.
2. Gel slice was weighed in a 1.5 ml microcentrifuge tube.
3. 3X gel volume of gel solubilization buffer (DEA buffer) was added.

4. The gel was resuspended in gel solubilization buffer by vortexing. The mix was heated at 75^{o}C until the gel was completely dissolved. Intermittent vortexing was given every 2 to 3 min. to enhance gel solubilization.
5. 0.5X gel solubilization volume of binding buffer (DEB buffer) was added and mixed properly.
6. A spin column was placed in a 2 ml collection tube. The solubilized gel slice was transferred into the spin column that was assembled in the 2 ml collection tube and centrifuged at 12,000 rpm for 1 min.
7. The filtrate was discarded. 500µl of wash buffer (W1) was added to the spin column and centrifuged at 12,000 rpm for 30 sec.
8. The filtrate was discarded and 700µl of desalting buffer (W2) was added and centrifuged at 12,000 rpm for 30 sec.
9. A second wash was given by adding 700µl of desalting buffer (W2), followed by centrifugation at 12,000 rpm for 30 sec to ensure the complete removal of salt. The filtrate was discarded and spin column was again placed in collection tube.
10. Column was again centrifuged for 1 min. at 12,000 rpm to remove any residual buffer.
11. Spin column was transferred to a fresh 1.5ml microcentrifuge tube. The elutent was prewarmed at 65^{o}C to improve the elution efficiency.
12. To elute the DNA, 7µl of elutent was added to the centre of the spin column. It was allowed to stand for 1 min. at room temperature. Then centrifuged at 12,000 rpm for 1 min.
13. Eluted DNA fragments were checked on 0.7 percent agarose gel and stored at -20^{o}C for further cloning works.

3.8 Transformation

Cloning and transformation were carried out with the vector pGEMT and DNA directly isolated from *Pokkali* soils by using the method of Mandel and Higa (1970).

3.8.1 Preparation of competent cells

Competent cells for plasmid transformation were prepared using Genei Competent cell preparation Kit (B) from Bangalore Genei, following the manufacturer's guideline. Reagent used for competent cell preparation is given in Appendix IV.

Day 1

1. *Escherichia coli* JM 109 was streaked on LB agar plate from stab and incubated at 37°C for 16 to 18 hours.

Day 2

1. 100ml SOC broth in 1000 ml conical flask was inoculated with 10 to 12 moderately sized colonies from SOC plates.
2. Overnight incubation was given at 37°C in a shaker at 160 rpm. When the OD_{600} reached 0.3 (3 to 3.5 h only), growth was arrested by chilling. The flask was chilled for 20 minutes.
3. The entire culture was transferred into a 50ml centrifuge tube and centrifuged at 3500 rpm for 15 minutes at 4°C.
4. The supernatant was discarded. Keeping the tubes on ice, resuspended the bacterial pellet very gently in 33.3ml ice cold solution A.
5. The tubes were kept on ice for 20 minutes and then centrifuged at 3500 rpm for 15 minutes at 4°C.
6. The supernatant was discarded and pellet chilled on ice. The pellet was resuspended in 5 to 6ml of ice cold solution A.
7. The suspension was kept on ice for 10 minutes and aliquots of 100µl were dispensed in chilled 1.5ml Eppendorf tubes.
8. The tubes were frozen on ice for few minutes before storing at -70°C. The competency of the cells prepared was confirmed by transformation using pUC18. The cells were plated on a plate containing LBA + 50mg/l ampicillin.

3.8.2 Screening of competent cells

The competent cells prepared were screened to check their transformation efficiency, by transforming them using a plasmid (pUC18) containing ampicillin resistance marker and lacZ gene. The following procedure was used:

1. Prepared 50 ml LB media and 50 ml LB broth.
2. The competent cells from -70°C were thawed on ice.

3. pUC18 is diluted to 1:10 dilution. Added 1µl of diluted pUC18 to thawed competent cells.
4. The contents were mixed gently and kept on ice for 40 min.
5. Meanwhile, the water bath was set to 42^{o}C.
6. The tube was immediately taken from ice and a heat shock at 42^{o}C was given exactly for 90 sec. Without shaking, the tube was placed back in ice for 5 min.
7. Added 250µl of LB broth to vial under sterile conditions and was inverted twice to mix the contents.
8. The tube was incubated at 37^{o}C for 1 hour with shaking.
9. The transformed cells were plated on LB agar/ampicillin (50mg/l) overlaid with IPTG (6µl) and X gal (60µl).
10. The plates were incubated overnight at 37^{o}C.

3.8.3 Cloning

The eluted specific 16s rDNA fragment was cloned in pGEMT vector system supplied by Promega, USA.

A. *Ligation*

The appropriate amount of eluted product (insert) required for ligation was calculated by estimating the quantity of eluted DNA using NanoDrop® Spectrophotometer. For this, 0.25µl of eluted band was diluted to 5µl with sterile water. The amount of PCR product required was calculated by the following relationship.

$$\text{ng of insert} = \frac{\text{ng of vector x kb size of insert}}{\text{kb size of vector}} \text{ X Insert: vector molar ratio}$$

Ligation procedure was followed as per the manufacturer's guidelines.

Procedure

1. The pGEMT vector was briefly centrifuged to collect contents at the bottom of the tubes.
2. The following ligation reaction was set up as described in Table 4

Table 4. Preparation of Ligation Mix

Components	volume
2X Ligation buffer	5.0µl
pGEMT vector (50ng)	1.0µl
Eluted product	3.0µl
T_4 DNA Ligase (3units/µl)	1.0µl
Total volume	**10.0µl**

3. The reaction was mixed by pipetting and incubated for 1 h at room temperature. Then it was kept at 4^{o}C overnight.

B. *Transformation of ligated product*

Reagents

1. Ampicillin – 5mg/ml in water
2. IPTG – 200mg/ml in water
3. X-gal – 10mg/ml in DMSO

The procedure adopted for plasmid DNA transformation was as follows:

The ligated PCR product was added to 100µl of thawed competent cells and kept on ice for 40 min. Heat shock was given at 42^{o}C for 90 seconds in a water bath and immediately placed back in ice for 5 min. LB medium (250µl) was added to the cells and incubated at 37^{o}C for 1 hr on a shaker at 160 rpm. The aliquots of transformed cells was plated on LB agar/Ampicillin (5 mg/ml) / IPTG (6µl) / X-gal (60µl) plates and incubated overnight at 37^{o}C.

C. *Transformation efficiency*

Transformation efficiency was calculated using the formula

$$\text{Transformation efficiency} = \frac{\text{colonies on plate}}{\text{Mass of DNA plated in µg}}$$

D. Confirmation of presence of insert

The putative transformants which appeared as white colonies were picked from the plate and streaked on grided LB agar plates containing ampicillin and Xgal/ IPTG. The plate was incubated overnight at 37ºC and blue colonies were also streaked as a control. This was done to check whether the white colonies retained their colour and also get more colonies of each transformant. Then it was stored at 4ºC for further use.

Presence of insert in the plasmid was confirmed by performing colony PCR with T7 and SP6 primers specific for pGEMT vector as detailed below:

1. A loopful of colony was diluted with 20µl of sterile water.
2. Cells were lysed with the help PCR machine by keeping it at 95 ^{0}C for 2 min.
3. A spin was provided to sediment the cell debris.
4. 1 µl of supernatant was taken from lysed cells as a plasmid DNA into new PCR tube.
5. A master mix was prepared and the reagents were in the sequential order as shown in the table 2.
6. The contents were mixed well and briefly centrifuged.
7. The PCR programme as given in table 5, except that the annealing temperature was 41^{0}C was run immediately.
8. The PCR products were analyzed on 1 per cent agarose gel.

Table 5. Composition of PCR reaction mix

Components	**Vol. per reaction (µl)**
Genomic DNA (50ng)	1.0
10X Taq buffer	2.5
dNTP Mix (10mM)	1.0
T7 Primer (27.5nM)	1.0
SP6 Primer(31.3nM)	1.0
Taq DNA polymerase (0.3 units)	1.0
Sterile H_2O	16.5
Total volume	**25.0**

3.8.4 Preparation of stabs

LB agar medium was melted, cooled to 42^{o}C and Ampicillin was added. This medium was poured into storage vials aseptically in a laminar air flow chamber. After solidification, single colony was inoculated into the medium using sterile bacterial loops. The vials were incubated at 37^{o}C overnight and further stored at 4^{o}C. A total of 57 stabs were prepared out of which 3 were blue colonies and rests were white colonies and 30 cloned stabs were sent for sequencing.

3.9 Sequencing of 16 S rRNA gene specific DNA fragment

Thirty stabs were sent for sequencing to DNA sequencing facility of Bangalore GeNei. The primers used for sequencing were T7, SP6 (Fig. 1).

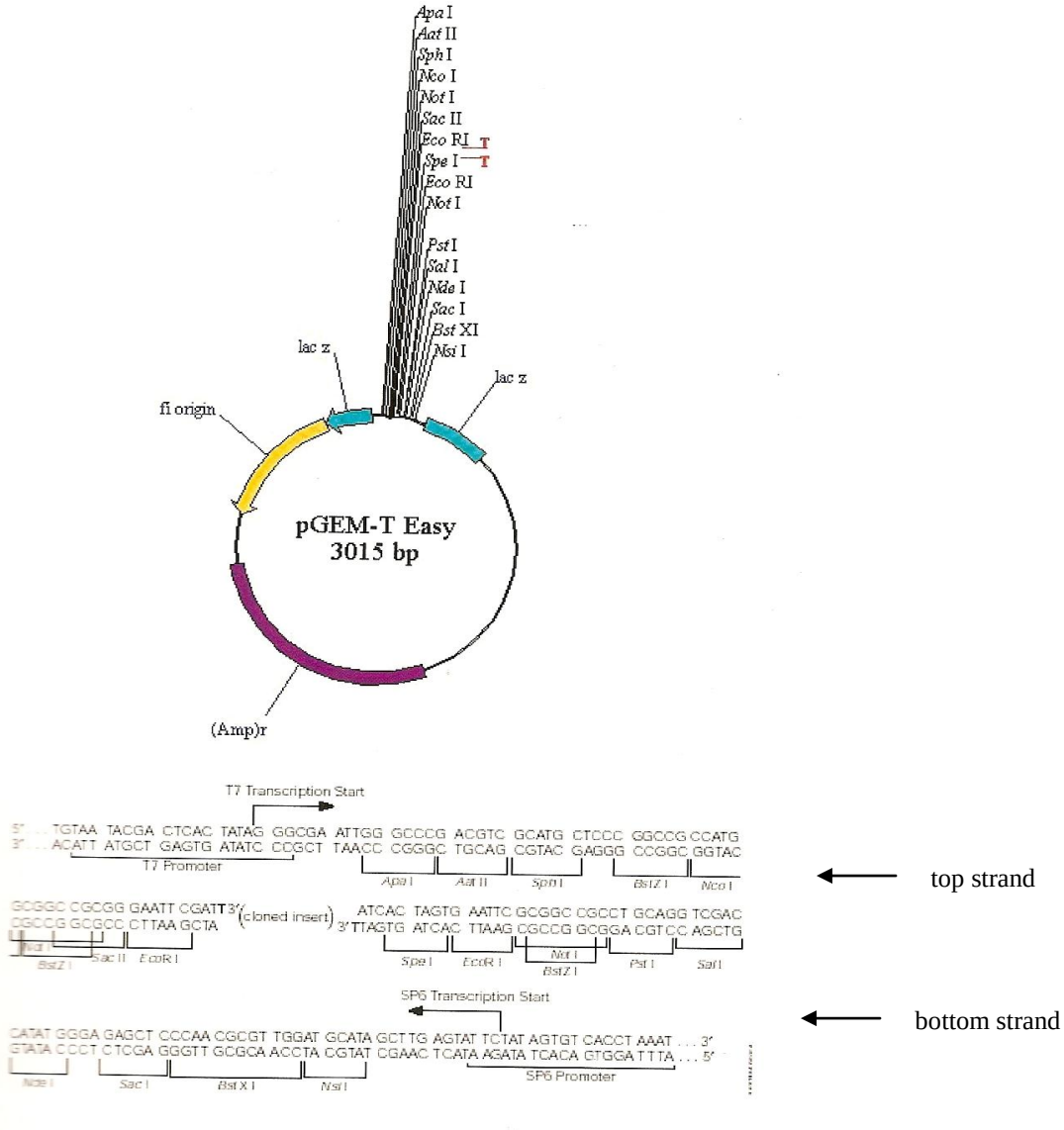


Fig. 1 Details of pGEM-T Easy Vector (Promega) used for cloning PCR product.
The *lac*Z region, promoter and multiple cloning sites are shown in the figure. The top strand of the sequence shown correspond to the RNA synthesized by T7 RNA polymerase and the bottom strand corresponds to the RNA synthesized by SP6 RNA polymerase.

3.10 Analysis of sequences

The DNA sequences obtained from Bangalore GeNei were analyzed using various online bioinformatics tools.

1. Vector sequence from the forward (T7U) and reverse sequence (Sp6) data obtained was removed using VecScreen (www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html) available at the NCBI website and Bioedit software.
2. The URL link http://wwwgeneinfinity.org/sms/sms/ was opened and used for converting SP6 sequence into reverse complements.
3. The URL link http://wwwbioinfo.nhri.org.tw/gui/ was used as online tool (i.e. Emboss explorer), merger tool was selected and forward and reverse sequences were pasted for merging.
4. The merged sequence was used for studying homology with different bacterial species using software named Blastn and URL link (http://www.ncbi.nm.nih.gov/blast/).
5. MEGA software was used for phylogenetic analysis and Tree construction. Dendrogram was constructed by Neighbor-Joining Tree method with thousand boot strapping.

3.11 Conventional Method for Characterization of Culturable Bacteria

3.11.1 Media

Soil extract agar was used for isolation of bacteria. 1000 g of garden soil was sieved and mixed with 1000 ml of tap water thoroughly. It was autoclaved for two hours at 15psi. After cooling down to room temperature, the extract was filtered using a muslin cloth to get soil extract. This was used to prepare soil extract agar. Composition is given in Appendix 1.

The ingredients were suspended in 1000 ml of distilled water. Sterilization of media was done by autoclaving at 15psi pressure and a temperature of 121°C for 15 minutes. The medium was mixed well and poured into sterile Petri plates.

The culturable bacteria and clones in metagenomic library were maintained on Luria Bertani (LB) agar (composition given in Appendix I). Medium was melted and cooled to 42^0C, added ampicillin ($50mgl^{-1}$). It was poured in Petri plates aseptically. Colonies were streaked on solidified LB plates using flame sterile loop and incubated at 37^0C overnight. The single colonies developed on the plate were incubated in vials containing the same media, incubated overnight and stored as stabs at 4^0C.

3.12 Microbiological Analysis of Saline Soil Samples

Halophilic bacteria in saline *Pokkali* soil were enumerated and isolated using the Luria agar medium containing 0, 5, 10, 15 and 20 per cent NaCl (Kushner, 1993). One g of soil was taken in test tube containing 9.0 ml of sterile water to make the 1:10 dilutions. This was shaken vigorously for two minutes. One millilitre of soil solution was transferred to the test tube containing 9.0 mL of water to make the 1:10 dilution. Furthermore serial dilutions of $1:10^2$, $1:10^3$, $1:10^4$, $1:10^5$, and $1:10^6$ were prepared. The dilutions of $1:10^2$, $1:10^3$, and $1:10^4$ were used for the enumeration of slightly (0 and 5 per cent NaCl), moderately (10 per cent NaCl) and extremely halophilic (15 and 20 per cent NaCl) bacteria using the standard pour plate method. Each selected dilution of each soil sample was analyzed in duplicate. The plates were inverted and incubated for 3 to 7 days at 37^oC. Results were recorded as colony forming units (CFU) per gram dry soil. Different colonies grown on media were selected and purified for further investigation.

3.12.1 Isolation of total DNA from bacteria

The total DNA was isolated from pure culture of three predominant bacterial isolates (BS1, BS2, BS3) following the protocol of Sambrook and Russel (2001).

(Reagents and their chemical composition for DNA isolation are given in Appendix III).

1. 25ml of overnight grown culture in LB broth was centrifuged at 10,000 rpm for 10 minutes at 4°C.
2. The pellet was resuspended in 10mM Tris HCl and 100mM NaCl

3.Centrifugation was carried out at 10,000 rpm for 10 minutes at 4°C.

4. The pellet was resuspended in 2.5ml TE and 500µl of lyzozyme from a stock of 50mg/ml and incubated at 37°C for 20 minutes.
5. 25µl of RNase A was added from a stock of 10mg/ml.
6. Incubated at room temperature for 10 minutes.
7. 2.5ml of SDS was added and incubated at 50°C for 45 minutes.
8. 50µl of proteinase K was added from a stock of 20mg/ml and incubated at 50 to 55°C for 10 minutes.
9. An equal volume of phenol was added, mixed gently and centrifuged at 10,000 rpm for 10 minutes at 4°C.
10. The aqueous phase was transferred to a fresh tube and an equal volume of phenol: chloroform (1:1) ratio was added. The mixture was mixed gently.
11. Centrifugation was carried out at 10,000 rpm for 10 minutes at 4°C and the aqueous phase was pipetted into a fresh tube.
12. An equal volume of chloroform: isoamyl alcohol (24:1) was added and mixed gently.
13. Centrifugation was carried out at 10,000 rpm for 10 minutes at 4°C and the aqueous phase was separated.
14. $1/10^{th}$ volume of 3M sodium acetate was added and kept in ice for 20 minutes.
15. An equal volume of isopropanol was added and incubated for 5 to 10 minutes.
16. Centrifugation was carried out at 10,000 rpm for 15 minutes at 4°C and the supernatant was decanted.
17. The pellet was washed in 70 per cent alcohol, by giving a centrifugation at 10,000 rpm for 3 minutes. The supernatant was discarded.
18. The pellet was dried and dissolved in 25 to 30µl of TE buffer.

DNA was detected on 0.8 per cent of agarose gel quantity and quality checked by spectrophotometry as mentioned in section 3.5.

3.13 PCR Reaction

3.13.1 Amplification of 16S rRNA gene of Cultured Bacterial DNA

The total DNA extracted from three predominant bacterial isolates (BS1, BS2, and BS3) was amplified by PCR gene specific primers in Eppendorf Master Cycler, Gradient (Eppendorf, Germany). Composition of the reaction mixture for PCR, Primer description,

Temperature profile for amplification of gene by PCR, Temperature profile for amplification of gene by PCR were same as mentioned in section 3.6.

3.13.2 Gel Elution of PCR amplified fragments

Products obtained in PCR reactions were loaded separately on 0.8 per cent (w/v) agarose gel and the desired band in each case was eluted by using AxyPrep DNA Gel Extraction Kit (Axygen Biosciences Pvt. Ltd). Procedure was same as mentioned in section 3.7.

3.14 Sequencing of 16 S rRNA gene specific DNA fragment

The eluted PCR products corresponding to 16S rDNA amplified from the bacteria named as BS1, BS2 and BS3 were sequenced at the DNA sequencing facility of Bangalore GeNei.

3.15 Analysis of sequences

The DNA sequences obtained from Bangalore GeNei were analyzed using various online bioinformatics tools.

1. The URL link http://wwwgeneinfinity.org/sms/sms/ was opened and used for converting sequence into reverse complements.
2. The URL link http://wwwbioinfo.nhri.org.tw/gui/ was used as online tool (i.e. Emboss explorer), merger tool was selected and forward and reverse sequences were pasted for merging.
3. The merged sequence was used for studying homology with different bacterial species using software named Blastn programme (http://www.ncbi.nm.nih.gov/blast/). It was used to find out the homology of nucleotide sequences (Altschul *et al*., 1997) with the NCBI database.

Results

4. RESULTS

The results of the study on "Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala" undertaken during the period 2009-2011 at the Centre for Plant Biotechnology and Molecular Biology and Department of Agricultural Microbiology, College of Horticulture, Kerala Agricultural University, Vellanikkara are presented in this chapter under different subheadings. The research work included the following aspects:-

4.1 Electrochemical properties of soil

4.2 Isolation, purification and quantification of DNA

4.3 16S rRNA gene amplification

4.4 Molecular cloning

4.5 *In silico* analysis with bioinformatics tools

4.6 Phylogenetic analysis

4.7 Conventional methods for characterization of culturable bacteria

4.1 Electrochemical properties of soil

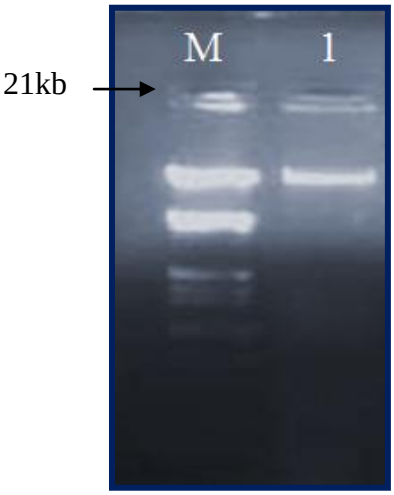
The electrical conductivity (EC), total dissolved salts (TDS) and pH of the composite soil sample under study are given in Table 6.

Table 6. Electrochemical properties of *Pokkali* soil (composite sample)

Electrical conductivity (dSm^{-1})	Total dissolved salts ($mg\ kg^{-1}$)	pH
4.03	6448	3.79

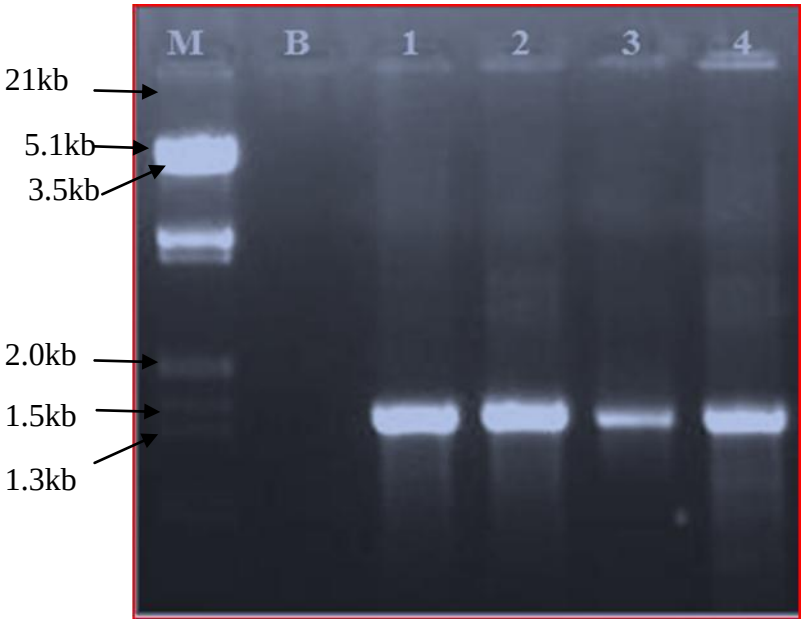
4.2 Isolation and Quantification of Metagenomic DNA

Metagenomic DNA was extracted from composite air dried soil sample. DNA isolation was carried out with soft lysis method using lysis buffer. DNA sample was run on 0.8 per cent agarose gel to check the purity of DNA and contamination by protein and RNA if any (Plate 1A). Since RNA and protein will



M- Marker (λ DNA/*Eco*R1+*Hind*III)
1- Metagenomic DNA sample

A. Metagenomic DNA from *Pokkali* soil.



M- Marker (λ DNA/*Eco*R1+*Hind*III)
B- Blank, & 1, 2, 3, 4- 16S rRNA amplified DNA fragment
B. 16S rDNA amplicons from metagenomic DNA

Plate 1. Amplification of 16S rDNA from metagenomic DNA extracted from *Pokkali* soil

Affect the PCR amplification DNA, should be free from these contaminants. No RNA contamination was observed during the study. Isolated DNA showed a single intact band of high intensity with no shearing. The quantity of DNA (isolated by soft lysis method) as assessed by spectrophotometery using NanoDrop, was 403.2ng/µl. The isolated DNA, after appropriate dilutions was used as template for 16S rDNA amplification.

4.3 16S rDNA gene amplification

PCR amplification of 16S rRNA genes was carried out with the specific primers as suggested by Siddhapura *et al*. (2010) using 100ng of the template DNA. The PCR products were electrophoresed in 0.8 per cent agarose gel. A single band of the expected size (1.5kb) was obtained (Plate1B).

4.4 Molecular cloning

4.4.1 Transformation

4.4.1.1 Preparation and screening of competent cells

The competent cells prepared in section 3.8.2 were checked for competence by transforming the plasmid (pUC18) having ampicillin resistance. A large number of blue colonies were obtained which indicated a high degree of transformation efficiency. Thus the competent cells prepared, were found ideal for cloning amplicons.

4.4.1.2 Ligation

The eluted PCR product (insert) required for ligation was calculated using the equation described in section 3.8.3. Thus, 30ng of insert was used per 50ng of vector for ligation reaction using pGEMT vector, considering the fact that the eluted DNA sample had 0.6ng DNA μl^{-1}.

A. Transformation of ligated product

The ligated product was used to transform competent cells using the heat shock method and was incubated at 37^{o}C. Large numbers of blue and white colonies were obtained after overnight incubation when the transformed cells were cultured in LB/ampicillin media, in the three plates; overlaid with X gal and IPTG (Plate 2.A).

B. Screening of the transformed colonies

The LB ampicillin agar plates containing the transformed colonies were screened for recombinant plasmid. For maintenance of clones, white colonies were picked from plate and were grown in LB ampicillin media separately. Two blue colonies were also inoculated in order to set the negative control (Plate 2.B).

C. Number of recombinants obtained per µg DNA

$$\text{Transformation efficiency} = \frac{\text{colonies on plate}}{\text{Mass of DNA plated in µg}}$$

Colonies on a single plate = 500

Mass of DNA plated = 0.09 µg (3 µl of eluted product having a concentration of 0.03µg/ µl) x Fraction of volume spread on plate

= 0.09 x 0.061 = 0.00549µg

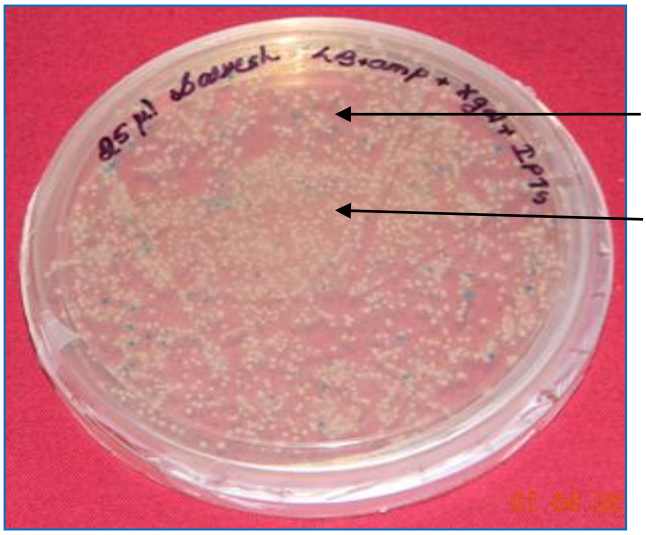
$$\text{Fraction of volume spread} = \frac{\text{volume suspension spread}}{\text{total volume suspension}}$$

$$= \frac{25\mu l}{410\mu l} = 0.061$$

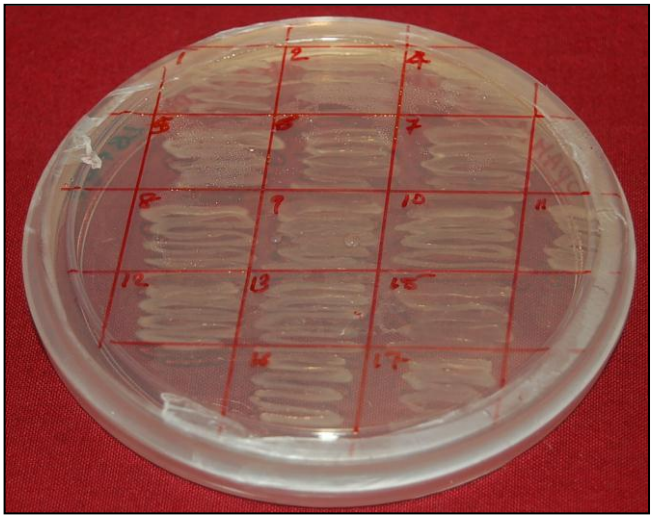
$$\text{Transformation efficiency} = \frac{500 \text{ cfu}}{0.00549\mu g}$$

$= 9.1 \times 10^4$ cfu /0.00549µg DNA

$=1.6 \times 10^7$ cfu/µg DNA



A. Blue white screening of transformed *E.coli* cells



B. Maintenance of clones
1 to 17 Recombinant clones

Plate 2. Blue white screening of recombinant clones

The transformation efficiency of competent cell was calculated as 1.6×10^7 cfu/µg DNA. The cells were competent and showed good transformation efficiency. The colonies showed luxuriant growth, with no other contamination in LBA ampicillin plates overlaid with x-gal and IPTG.

G. Confirmation of the insert by colony PCR

Confirmation of insert presence was carried out by performing colony PCR amplification of recombinant plasmid DNA using T7 and SP6 Primers. The PCR reaction was prepared as given in 3.8.3.C. The PCR products when checked on 1 per cent agarose gel showed amplicons of 1.5kb size whereas in blue colony sample showed below 0.5 kb size band (Plate. 3).

After confirmation of insert, stabs were prepared and sent for automated sequencing. The details of sequence analysis of 30 clones are given in Table 7 to Table 36.

4.5 *In silico* analysis with Bioinformatics tools

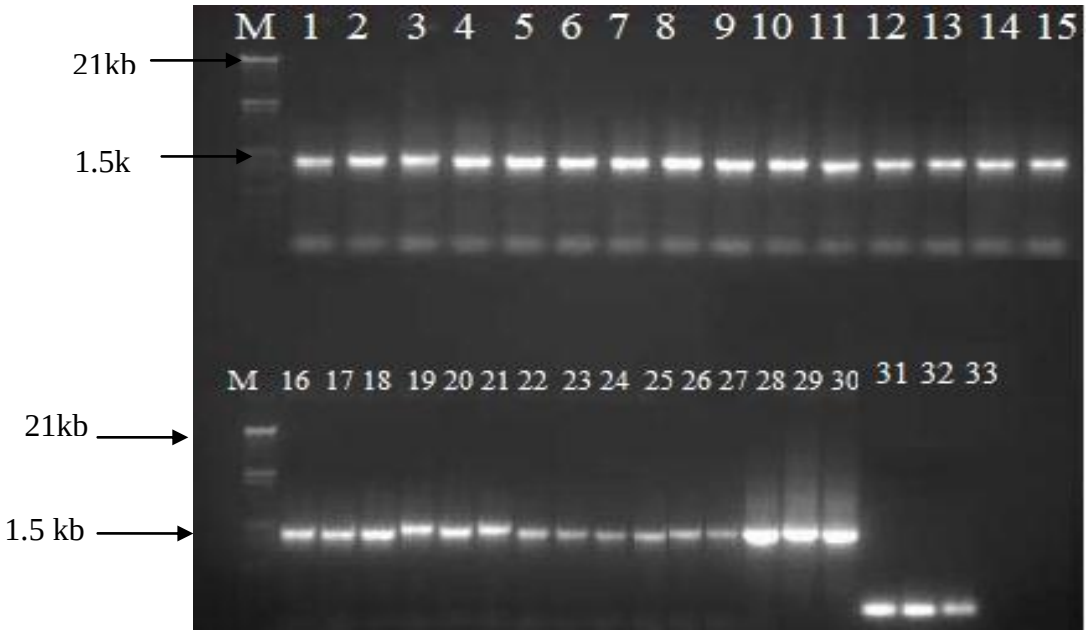
CLONE- 1

The sequence data obtained for the clone 1 with forward primer (T7U) was 707 bp in size and vector screening indicated vector sequence from 1 to 62 bp. The details of results obtained are presented in Fig. 2A. The sequence data obtained with reverse primer (SP6) was 548 bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1182 bp) indicated 92 percent query coverage with 93 percent homology with uncultured Ectothiorhodospiraceae. Details of Blastn result obtained are presented in Table 7.

CLONE- 2

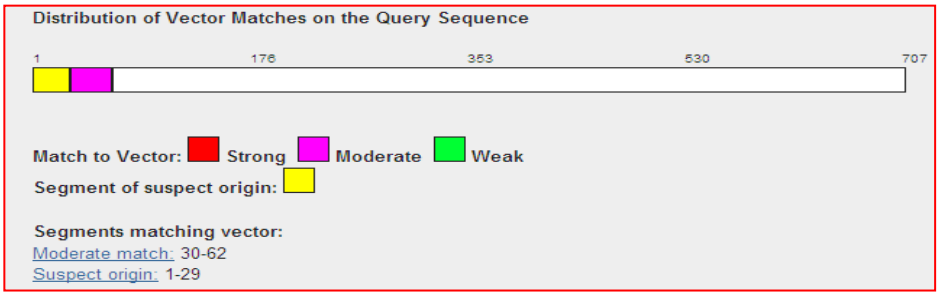
The sequence data obtained for the clone 2 with forward primer (T7U) was 695 bp in size and vector screening indicated vector sequence from 1 to 64



M- Marker (λ DNA/*Eco*RI+*Hind*III)

Lane 1-30: Recombinant clones
Lane 31-33: Non- recombinant clones (blue colonies)

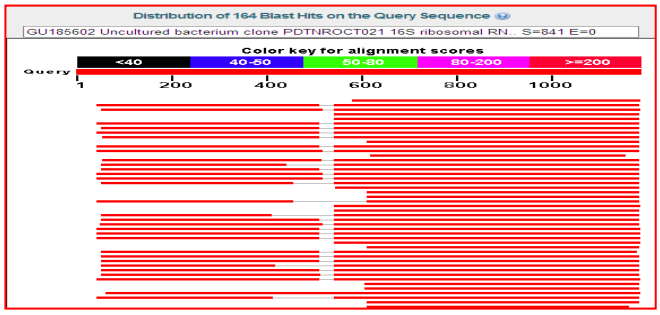
Plate 3. Confirmation of insert in recombinant clones



A. VecScreen output

Merged sequence (1182 bp)

5'GACCTTACTCTTTGCATCTTCAGAATTGTTCAAATGCTTTTGCTTTGGAACTGAGTACAGGTGTGCATGGCTTCGTCAGTTGTGTCGTGAGATGTTG
GGTTAAGTCCCGTAACGAACGCAAGCCTTGTCCATAGTGCCAGCGAGTAATGTTGGGAACTTATGGAGACTGCCCTGTGACCAACCGGAGGAAGGT
GGGGATGACGTCAATTCATCATGGCCCTTACGAGTAGGGTACACACGTGCTACAATGGAAAGTACAGAGGGCTGTAAAGCGCGAGCCGGAGCAAA
TCCCATAAAACTTTTCGTAGTCCGGATTGGAGTTGCAACTCGACTCCATGAAGTCGGAATCGTAGTAATCGCGGATCAGAATGCCGCGGTGAATAC
GTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGCTGCACCAGAAGCCGGTAGTCTAACCGCAAGGAGGACGCCGTCCACGGTGT
GGTCAAGACTGGGGGAAGTGTAACGAGGTAGCCGTAATCACGAGAGCGGCCGCTGCAGGTCGATCCTGGCTCAGATTGAACGCTGGCGGCTGCTTA
ACCTGCAGTCAACGGAAAGGGGTTCGGCCGCTGACGAGCGGCGGACGGGTGAGTAATACATAGGAATCTGCCCTGTAGTGGGGGACAACTTAGGG
AAACTTAAGCTAATACCACATACGACCTACGGGTGAAAGACTTCGGTCGCTATAGGATGAGCCTATGTCTGATTAGCTAGTTGGTAGGGTAAAGGC
CTACCAAGGCAACGATCAGTAGCTGGTCTGAGAGGACGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGG
AATATTGGACAATGGGCGAAAGCCTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTGCGGGTTGTGAAGCACTTTCGGTAGGGAAGATAATGA
CGGTACCTACAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCG
CGCGTAGCGGTTTGTTAAGTCGGGTGTGAAATCCCTGGGCTCACCTAGAATTGCATTCGATACTGGCAGCTAGAGTGTGGTAGAGGGAAGTGAATT
CCAGGTGTAGCGGTGAAATGCGTAGATATCTGG 3'

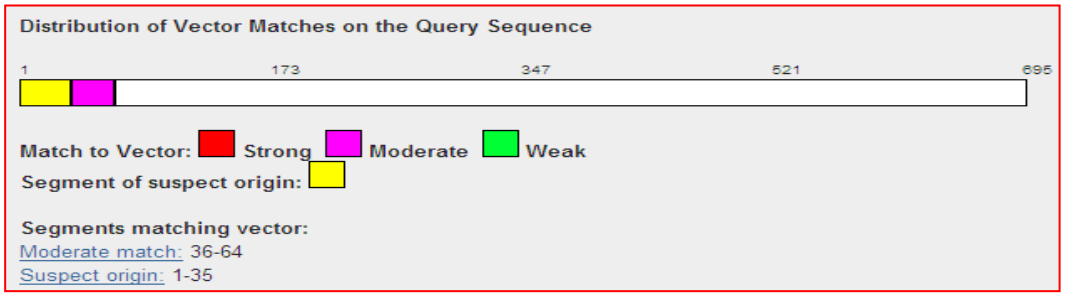


B. Blastn output

Fig. 2. Sequence analysis of clone 1

Table 7. Blastn result of clone 1

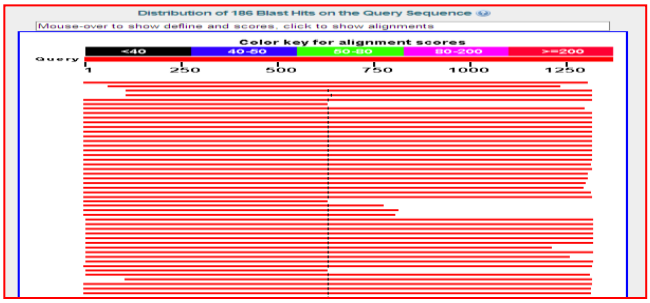
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
FN687123.1	Uncultured Ectothiorhodospiraceae	931	92	0.0	93
EF999373.1	Uncultured bacterium	917	93	0.0	92
FJ793190.1	Uncultured bacterium	889	93	0.0	91
FR851506.1	Uncultured bacterium	883	53	0.0	91
FR851685.1	Uncultured bacterium	878	53	0.0	91



A. VecScreen output

Merged sequence (1362 bp)

5'ACGGCTACCTTGTTACGACTTCACCCCAGTCATGAATCCCACCGTGGTAAGCGCCCTCCTTACGGTTAGGCTACCTACTTCTGGTGAAACCCACTC
CCATGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCGACTTCATGCACTCGA
GTTGCAGAGTGCAATCCGGACTACGATCGGCTTTCTGGGATTGGCTCCACCTCGCGGCTTGGCAACCCTCTGTACCGACCATTGTATTACGTGTGAA
GCCCTACCCATAAGGGCCATGAGGACTTGACGTCATCCCCACCTTCCTCCGGTTTATCACCGGCAGTCCCATTAAAGTGCCCAACTAAATGATGGCA
ATTAATGGCAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCAGCACCTGTGTCCAGGTTCCCTTT
CGGGCACCAACCCATCTCTGGAAAGTTCCTGGCATGTCAAGGGTAGGTAAGTTTTCGCGTTGCATCGATTATCCACATATCCACCGCTTGTGCGGTC
CCGTCATTCCTTTGAGTTTTAACCTTGCGGCCGTACTCCCCAGGCGGTCACTGTTACACGGGGAATTCTACCCCCCCTCTGCCGCACTCTAGCTTGCA
GTCACAAATGCAGTTCCCAGGTTAAGCCCGGGGATTCACACC\TGTCTTACAAAACCGCCTGCGCACGCTTTACGCCCAGTAATTCCGATTAACGCTC
GCACCCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGGTGCTTCTTCTGACGGTACCGTCATTAATCCAGGGTATTAACCTAGACCGTTTCTT
CCCATCTGAAAGAGCTTTACAACCCGAAGGCCTTCTTCACTCACGCGGCATGGCTGGATCAGGCTTGCGCCCATTGTCCAAAATTCCCCACTGCTGC
CTCCCGTAGGCTGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCGGGCCGTCCTCTCAGACCCGCTACAGATCGTCGCCTTGGTTGGCCTTTACCCCACC
AACTAGCTAATCTGATATCGGCCGCTCCGAAAACGTGAGGTCCTAAGATCCCCCACTTTCCCCCTCAGGGCATATCCGGTATTAGCGTAACTTTCGC
TACGTTATCCCCCATTTCCGGGTACGTTCCGATATATTACTCACCCGTTCGCCACTCGCCACCGGGTGCAAGCACCCGTGCTGCCGTTCGACTTGCAT
GTGTAAAGCATGCCGCCAGCGTTCAATCTGAGCCAGGATCAAACTCTAATCACTAGGGCGGCCGCCTGCACATTGACCATATGTGGGCTCCCAACG
CCG 3'



B. Blastn output

Fig. 3. Sequence analysis of clone 2

Table 8. Blastn result of clone 2

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
Z93975.1	Unidentified bacterium	1356	99	0.0	86
AY092075.1	*Stenotrophomonas maltophilia*	1116	85	0.0	84
GU183581.1	Uncultured bacterium	1057	88	0.0	94
GU183582.1	Uncultured bacterium	1051	88	0.0	94
GU724735.1	Uncultured bacterium	1033	96	0.0	96

bp. The details of results obtained are presented in Fig. 3A. The sequence data obtained with reverse primer (SP6) was 736 bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1362 bp) indicated 85 percent of the query coverage and 84 percent homology with *Stenotrophomonas maltophilia*. Details of Blastn result obtained are presented in Table 8.

CLONE – 3

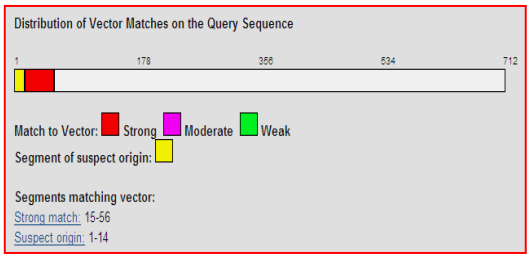
The sequence data obtained for the clone 3 with forward primer (T7U) was 712 bp in size and vector screening indicated vector sequence from 1 to 56 bp. The details of results obtained are presented in Fig. 4A. The sequence data obtained with reverse primer (SP6) was of 768 bp in size and vector screening indicated that vector sequence was 1 to 60 bp. The details of results obtained are presented in Fig. 4B.

Blastn results of the merged sequence (1299 bp) indicated that 92 percent query coverage and 98 percent homology with uncultured bacterium. Details of Blastn result obtained are presented in Table 9.

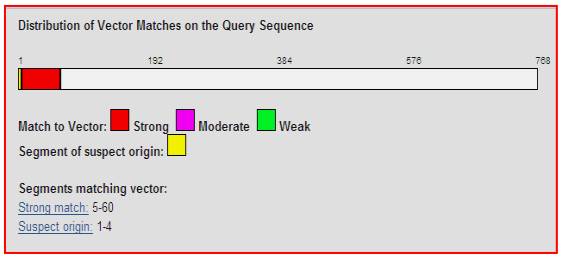
CLONE- 4

The sequence data obtained for the clone 4 with forward primer (T7U) was 694 bp in size and vector screening indicated that there was no vector sequence. The sequence data obtained with reverse primer (SP6) was 734 bp in size and vector screening indicated that vector sequence was 1 to 55 bp found. The details of results obtained are presented in Fig. 5A.

Blastn results of the merged sequence indicated (1361 bp) that 89 percent query coverage and 79 percent maximum identity with uncultured *Acaryochloris* sp. Details of Blastn result obtained are presented in Table 10.



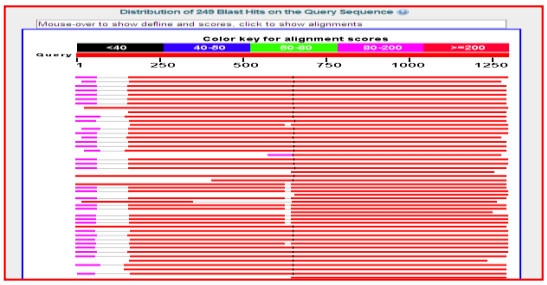
A. VecScreen output



B. VecScreen output

Merged sequence (1299 bp)

5'GTTTGATCCTGGCTCAGAACGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGGAGTTAATCGGCAGCACCAAGACTGACAACAGACAG
GTCATATAAGGCGTATACGACCTATAGGACTTATGGTGTTGCTTTCAGTTTTTGGTGCTGTTGGTTAACTTAGTGGCGGACGGGTGAGTAACACGTA
GGTAACCAACCCTAAAGAGGGGGACAACCCAGGGAAACTTGGGCTAATACCGCATAAGACCACGAACTGAAAGGTTAGTGGTAAAAGGTAGCAAT
ACGCTTTAGGATGGGCCTGCGGCCTATCAGGTAGTTGGCGGGGTAAAGGCCCACCAAGCCGAAGACGGGTAGCCGGACTGAGAGGTTGAACGGCC
ACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGCGCAATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGG
AGGAAGAAGGTCTTCGGATTGTAAACTCCTTTTGTAGGGGAAGATAATGACGGTACCCTATGAATAAGCCACGGCTAACTCTGTGCCAGCAGCCGC
GGTAAGACAGAGGTGGCAAGCGTTGTTCGGAATTACTGGGCTTAAAGGGCGCGTAGTGGTGGCGCAAGTCCGGGGTGGAAGACCATAAGTACCCC
GCCTGGGAGTACGGCCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGAAG
AACTTAGCTGGGCTTGACATGCAGGTGGTAGTGAACCGAAAGGGGAACGACCCGTAGCAATACGGGAGCTTGCACAGGTGCTGCATGGCTGTCGTC
AGCTCGTGCCGTGAGGTGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATCCTCTGTTACCTCGCAAGAGGGTCTCTGAGGAAACTGCCGGCGAT
GAGCTGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTATGTCCAGGGCTACACACGTGCTACAATGGGCGATACAGAGGGTAGCAAGA
CCGCGAGGTGGAGCAAATCCCTAAAGTCGTTCGTAGTTGGGATCGGAGTCTGCAACTCGACTCCGTGAACCTGGAATCGCTAGTAATCGCGGATCA
GAACGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAAAGCCTGTTGTACCCGAAGCAGGTGAGCTAACCCGCAAGG
GATGCAGCCTTCTAAGGTATGGCCGGTAATTGGGGTGAAGTCGTAACAAGGTAGCCG 3'

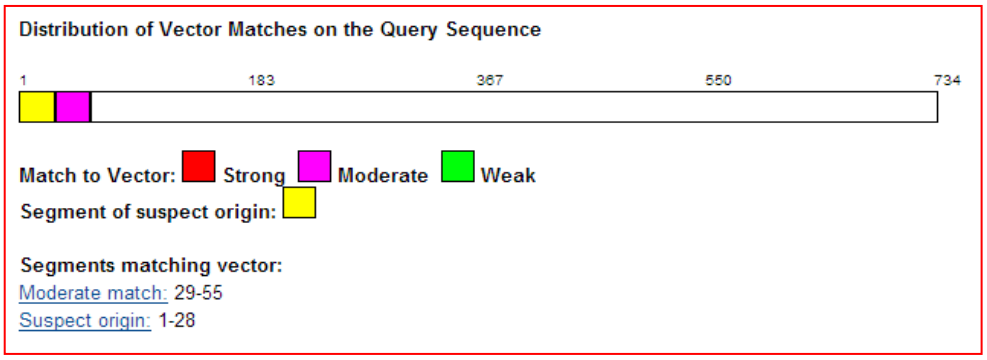


C. Blastn output

Fig. 4. Sequence analysis of clone 3

Table 9. Blastn result of clone 3

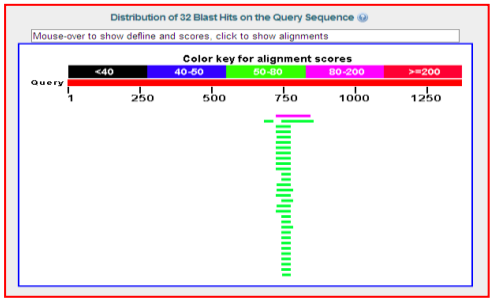
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
HO697847.1	Uncultured bacterium	1055	92	0.0	98
GO402791.1	Uncultured bacterium	1051	89	0.0	100
DO404771.1	Uncultured bacterium	1048	92	0.0	100
DO404725.1	Uncultured bacterium	1048	92	0.0	100
DO138961.1	Uncultured bacterium	1048	92	0.0	96



A. Vecscreen output

Merged sequence (1361bp)

5'TCTTTTCTTATTGTTATCCTAAAACAATGAGAATAACCCTTTGGGACTTTAAAATTTTTAAAGGACCTGCAAAAAAGTTAAGGGGGGTGTTATTGG
GCCCAAAACCGGGCGGGGGAGCGTTTTTTTTACTTGAATGAAGAAAAAAGATCCCAACTTTACGTATACCGCGGCCGAAGGTATCGGAATTGCCCT
GGTTTCTTCGGGCAGTGAGTTACAGGATCCACAGGATTATTCTAATCTGGTTTTGTGAGTTTGAAAGAAAATTCCAAAAAGGAGCCCATCCTCTCTC
CATAGGCACCAGTGGATAAGGTCGCGAACTAGATGGAGAATTTCGGAAAGGGGACGGAGGCTGGGAGGGAGGGCGGTGTTTCATCTCCACCGTTG
CGGGTAGGCCTACACGCGTCATTCAAGGGAATATACAGGGGGGGGTTAAAGCGCGGGAGCGAGCAAATCCCATATGGGCTCTTTTAGTAAGGGGA
GGTGTTAACATCTCCCACCCACGACGTCAGGGCATATATTGTAGGAGATAAGAATCGCTGAGTGATTTCCCTCCCGCGCGTAACACACGACATCTCT
CACTCCAAGGGAGCAGTCTCCACCGGAAGCAAGCAGTCTAACTGCAAGGAGGCCGCCATCCAAGGTCATTCAATGAGTGGGGAGAAGTCGTAACA
ACGATGGATGGGCCGACGTCGCTGCTCCCGGCCGCCTGGCCTCGGGATTAAGTTTGATCCTGCTCTATTGAACGCTGGCGGCCTGCTTAACACATGC
CAGTCTCACGGTAGCTAGGCCTTCTACTTCCTGACAAGTGGCTCACGGGTGATTACGGGTGGGAAGCTGCCCTGTAGTGGGGGATTCCCCGCAGCAT
GCTGATTAATACCTACTACCACTTCCGACTGAATGACCTCGGTCTGTATAGGATGATCCTAAGTCTGATTAGCTTTCTGGTATTGTATCCGCCTCCCG
GCTTGACGATCCTCTGTAGCTACCATAGGATTATGTGCGACACCGTAACCGAGACAGGGCCCAGAATCCGACGTGATCTCCCACTTCCTCAGTATTT
ATAGCCAGGGCGAAAGCCTGAGTGCCCAATGAATTGTGGGTGATCAATGGCTGGGGTTGCGCTCGTTTTTCGACTAGGCCATACATGACGGACACT
AGCTGACGACACCCGCTAGCACCGTGTGTGCAGCCTCCCTATCAGGCACGGCCCATCTCTGGACAGAATTGGTGGTGTCAAGGCTAGGGGAAGGCG
TTTTGCGTAAGTCGCGAGTGATCCCCCTGATCTCCCCCTTGTAATTGCTCCCATCCTGCCTTTCAGTATTGAGCTAGAGGGAGTGTCCTTCCTGGGGC
ACTTCAAACT3'



B. Blastn output

Fig. 5. Sequence analysis of clone 4

Table 10. Blastn result of clone 4

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E- Value	Max. Identity (%)
FM955439.1	Uncultured bacterium	97.1	92	3e-16	81
CO917911.1	Uncultured *Acaryochloris* sp.	79.0	89	1e- 09	79
IF534246.1	Uncultured gamma proteobacterium	79.1	92	5e- 09	91
IF829459.1	Uncultured bacterium	71.3	92	2e- 08	91
HO727374.1	Uncultured gamma proteobacterium	67.6	92	2e- 07	86

CLONE- 5

The sequence data obtained for the clone 5 with forward primer (T7U) was 741bp in size and vector screening indicated vector sequence from 1 to 69bp. The details of results obtained are presented in Fig. 6A. The sequence data obtained with reverse primer (SP6) was 693bp in size and vector screening indicated that vector sequence was 1 to 39 bp found. The details of results obtained are presented in Fig. 6B.

Blastn results of the merged sequence (1310 bp) indicated that 89 percent query coverage and 93 percent homology with uncultured bacterium sp. Details of Blastn result obtained are presented in Table 11.

CLONE- 6

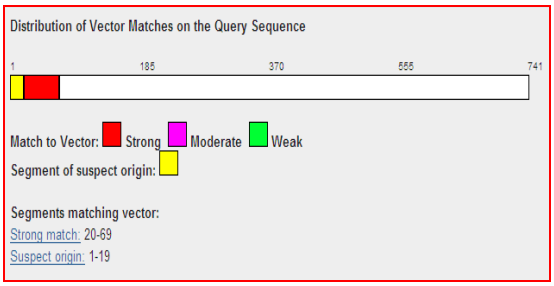
The sequence data obtained for the clone 6 with forward primer (T7U) was 694 bp in size and vector screening indicated that there was no vector sequence. The sequence data obtained with reverse primer (SP6) was 608 bp in size and vector screening indicated that there was no vector sequence found.

Blastn results of the merged sequence (1357 bp) indicated that 92 percent query coverage and 95 percent homology with *Azospira* sp. Details of Blastn result obtained are presented in Table 12.

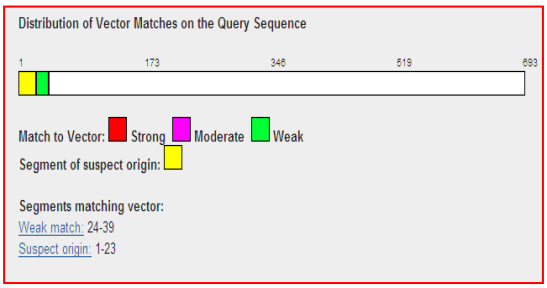
CLONE-7

The sequence data obtained for the clone 7 with forward primer (T7U) was 774bp in size and vector screening indicated vector sequence from 1 to 54bp. The details of results obtained are presented in Fig. 8A. The sequence data obtained with reverse primer (SP6) was 441 bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1158 bp) indicated that 92 percent query coverage and 95 percent homology with uncultured uncultured gama proteobacterium. Details of Blastn result obtained are presented in Table 13.



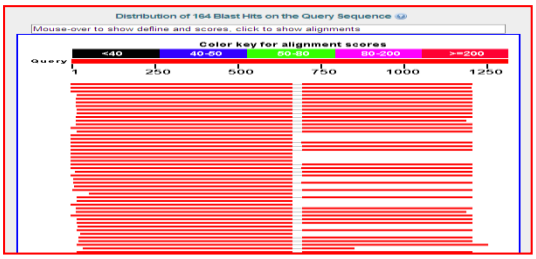
A. VecScreen output



B. VecScreen output

Merged sequence (1310 bp)

5'AGTTTGATCCTGGCTCAGGATGAACGCTAGCGGCAGGCCTAATACATGCAAGACGAACGGTAACATTTGTAGCAATACGGATGCTAGAGTGTCGT
AAGGGTGCGTAACACGTATGCAATCTGCCCTGTACAGGAGGATAGCTCCCCGAAAGGGGAATTAATACTCCGTACTATATTTAGGAGGCATCTTTTT
AAATATTAAAGCTGAGGCGGTACAGGATGAGCATGCGCATGATTAGCTAGTTGGCGGGGTAACGGCCCACCAAGGCGATGATCATTAGGGGAACT
AAGAAACGTTGATCCCCCACACTGGCACTGAGATACGGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATATTGGACAATGGGTGCGAGCCTGA
TCCAGCCATGCCGCGTGCAGGAAGACGGCCTTCTGGGTTGTGAACTGCTTTTGCCAGGGGATAAAAAGACCATGCGTGGTTAATTGAAGGTACCTG
GTGAATAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGTGGCAAGCGTTGTCCGGATTTATTGGGTTTAAAGGGTGCGTAGTGGT
CATGTAAGTCAGTGGTGAAATACGGCAGCTTAACTGTCGAGGTGCATTGATACTATGTGACTTGAGTACAGACGAGGTAGCGGAATTGACGATACC
ACGGCCCAGCGTTTTAGCGAAAACGGTGAAACTCAAAAGGAAATGGACGGGGGTCCGCACAAGCGGGTGGAGCATGTGGTTTTAATTCGATGATA
CGCGAGGAACCTTACCTGGGCTAGAATGCCCTTTGACAGATTCAGAGATGGGTTTTTTCGCAAGGACAAAGAGCAAGGTGCTGCATGGCTGTCGTC
AGCTCGTGCCGTGAGGTGTTGGCTTAAGTCCCGCAACGAGCGCAACCCCTATCCTTAGTTGCCAGCATGCAATGGTGGGGACTCTAAGGAGCCTGC
CTGCGCAAGCAGAGAGGAAGGAGGGGATGACGTCAAGTCATCATGGCCCTTACGCCCAGGGCTACACACGTACTACAATGGCGTATTCAAAGTGTT
GCCGGCCAGCGATGGTGAGCCAATCACAAAAAGTACGTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTTGGGATCGCTAGTAATCG
CGTATCAGAACGATGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACATCACGAAAGTGGGTTGTCTTAGAAGTCGTGATCCCAACC
GCAAGGAGGGTAGCGCCCAAGGTATGACTCATGATGGGGGAAGTGTAACGATAGGTAGCCGTAAT3'



C. Blastn output

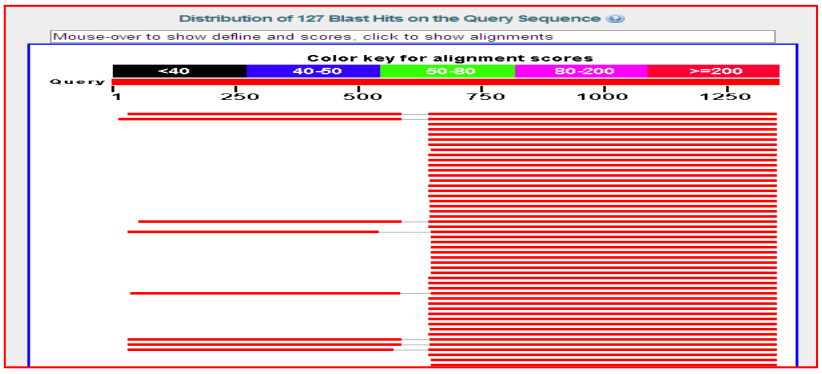
Fig. 6. Sequence analysis of clone 5

Table 11. Blastn result of clone 5

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E- Value	Max. Identity (%)
AB46240.1	Uncultured bacterium	968	89	0.0	93
HO389907.1	Uncultured bacterium	941	89	0.0	94
EO015104.1	Uncultured bacterium	941	89	0.0	93
AF314419.1	Uncultured bacterium	924	88	0.0	92
EO104332.1	Uncultured bacterium	920	88	0.0	92

Merged sequence (1357bp)

5'CCCCCTTTCCCCTTACCTTGCATCCAATCTCATTCCAGGTAACCCGGGGATTCCCCTGTTACAAACCGCCGGGCCCTGTAGCCCAGAATCCGATAA
CGTTGCACCTAGTATTCCGCGGTGTGGCAGTAGTTAGCCGGGTTTTTTGGCGGTACCGTCATTAATCAGGGTATAACCTAGACCGTTCTTCCATGAA
AGAGCTTTCAACCCGAAGGCTTTTCATCACGCGGCATGGTGGATCAGGCTGCGCCCATGTCCAAAATTCCCCATGTGCGTCCCGTAGGAGTTGGGCC
GGTTCAGTCCCAGTGTGGCGGGCCGTCCTTCAGACCCGTACAGATGTCCCTGGTAGGCCGTTACCCCACCAACTAGTAATGGGATATCGGCCGTCCG
AAAACGTGAGGTCTAAGATCCTCCACTTTCCCCCTCAGGGCATATGCGGTATTAGCGTAACTTTCGTACGTTATCCCCCATTTCCGGGTACGTTCCGA
TATATTACTCACCCGTTCGCCACTCGCCGCCGGGTGCAAGCACCCGTGCTGCCGTTCGACTGCAGAGTAAAGCAGCCGCCAGCGTTCAATTGAGCCA
GGATCAAACTTATCATGGGGAAATGGGCGACGTCGATGCTCCGGCCGCATCGGCCGCAGGGATATACGGCTACCTTGTTACCGACTTCACCCCAGT
CATGAATCCCACCGTGGTAAGCGCCCTCCTTACGGTTAGGCTACCTACTTCTGGTGAAACCCACTCCCATGGTGTGACGGGCGGTGTGTACAAGACC
CGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCGACTTCATGCACTCGAGTTGCAGAGTGCAATCCGGACTACGATCGGCT
TTCTGGGATTGGCTCCACCTCGCGGCTTGGCAACCCTCTGTACCGACCATTGTATTACGTGTGAAGCCCTACCCATAAGGGCCATGAGGACTTGACG
TCATCCCCACCTTCCTCCGGTTTATCACCGGCAGTCCCATTAAAGTGCCCAACTAAATGATGGCAATTAATGGCAAGGGTTGCGCTCGTTGCGGGAC
TTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCAGCACCTGTGTCCAGGTTCCCTTTCGGGCACCAACCCATCTCTGGAAAGTTCCTGG
CATGTCAGGGTAGGTAAGGTTTTTCGCGTTGCATCGAATAATCCACATAATCCACCGCTTGTGCGGGTCCCCGTCAATCCTTTGAGTTTACCTTGCGG
CCGTACTCCCAGGCGGTCAACTTCACGCGTTAGCTGCGGTACTAAAGAAGTCTCCTTCCCACACCTAGTTGACATCGTTTAGGCGTGACTACCAAG3'

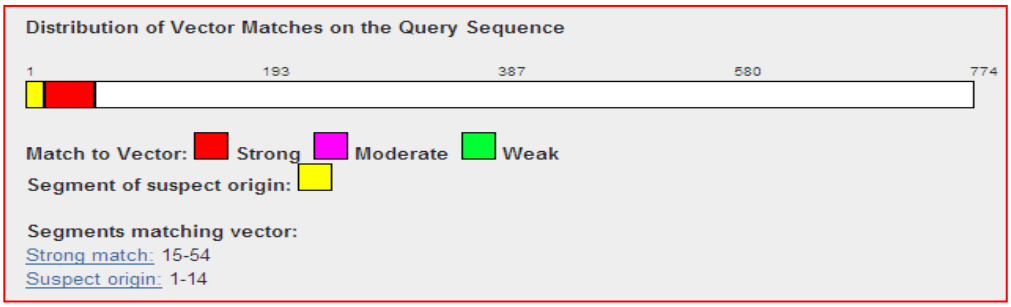


A. Blastn output

Fig. 7. Sequence analysis of clone 6

Table 12. Blastn result of clone 6

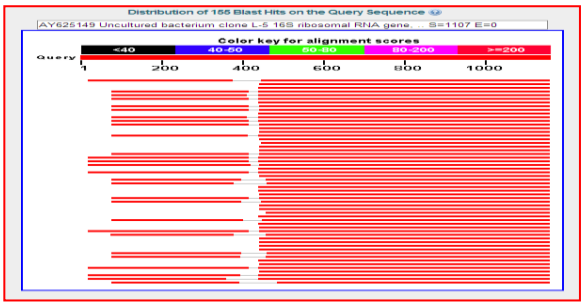
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EU802012.1	Uncultured bacterium	1142	93	0.0	95
HQ190296.1	Uncultured bacterium	1134	94	0.0	96
GU294119.1	*Azospira* sp.	1131	92	0.0	95
OC663512.1	*Azospira oryzae*	1131	88	0.0	95
AM084037.1	*Azospira* sp.	1131	88	0.0	95



A. VecScreen output

Merged sequence (1158bp)

5'AGGTTCCGGATTTCCCGGAATTTTTTTCCGACCTAAAGGTTACACCGGAGCCCTTTTCCCCGGGCAGTGGGTTCAGGTGCGCCCATTTCCATATTC
CCCAGGGCCTCCGTAGAGTTTGGGCGTGTTCAGTCCCAGTGGGGTGATCATCCTTCAGACCAGTACGGATCGTCGCCTGGTAGGCCATTACCCCACC
AACCAGGCTATCCGACGTAGGTTCCTCCAAGAGCGTGAGGCCTTGCGGTCCCCCATTCGTCCAAAGAGATCATGCGGTATTAATCCGGGTTTCCCCG
GGCTATCCCCCACTCAAGGGCAGATCCTACGTATTACTCACCCGTCCGCCACTGGGGTATGCTACCCCCGTCCGACTTGCATGAGTTAGCGTCAGCC
GCCAGCGTCAATTGAGCCAGGATCAAACTTAATCATAGGGAGGCCGCGTCAAGGCTACCTTGTTACGACTTCACCCCAGTCATTGACCACACCGTG
GACGGCGCTCCCCTTGCGGTTAAGCGACCGGCTTCTGGTGCAGCCAACTCCCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCAC
CGCCGCAGATGCTGATCGGCGATTACTAGCGATTCCGACTTCATGGAGTCGAGTTGCAGACTCCAATCCGGACTACGACCAGCTTTGTGGGATTGGC
TCCCCCTCGCGGGTTGGCGACCCTCTGTACTGGCCATTGTAGTACGTGTGTAGCCCTGGTCATAAGGGCCATGATGACTTGACGTCATCCCCACCTTC
CTCCGGTTTGTCACCGGCGGTCTCCCTAGAGTGCCCAACCCAATGCTGGCAACTAGGGACAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCT
CACGACACGAGCTGACGACAGCCATGCAGCACCTGTGTTCCGGCTCCCGAAGGCACGTCCGCATCTCTGCAGACTTCCGGACATGTCAAGACCAGG
TAAGGTTCTTCGCGTTGCATCGAATTAAACCACATACTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCACCTTGCGGCCGTACTCCCC
AGGCGGAGAACTTATCGCGTTAGCTACGACACTGATTCGCTAGGCAAACAACGTCAAGTTCTCATCGTTTACGGCGTGACTACCAGGGTATCA**3'**



B. Blastn output

Fig. 8. Sequence analysis of clone 7

Table 13. Blastn result of clone 7

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AJ313020.1	gamma proteobacteria	1164	92	0.0	95
DO323224.1	Uncultured gamma proteobacteria	1157	91	0.0	96
AJO13242.1	*Nevskia ramosa*	1134	62	0.0	95
GO388788.1	Uncultured bacterium	1129	91	0.0	95
AY326622.1	Uncultured bacterium	1129	90	0.0	95

CLONE-8

The sequence data obtained for the clone 8 with forward primer (T7U)was 348 bp in size vector screening indicated that there was no vector sequence.

Blastn result (348 bp) indicated that 75 percent query coverage and 75 percent homology with uncultured bacterium. Details of Blastn result obtained are presented in Table 14.

CLONE - 9

The sequence data obtained for the clone 9 with forward primer (T7U) was 780 bp in size and vector screening indicated vector sequence from 1 to 58 bp. The details of results obtained are presented in Fig. 10A.

Blastn result (722 bp) indicated that 96 percent query coverage and 95 percent homology with *Thiobacillus prosperus*. Details of Blastn result obtained are presented in Table 15.

CLONE – 10

The sequence data obtained for the clone 10 with forward primer (T7U) was 769 bp in size and vector screening indicated vector sequence from 1 to 30 bp. The details of results obtained are presented in Fig. 11A. The sequence data obtained with reverse primer (SP6) was 490 bp in size and vector screening indicated that there was no vector sequence.

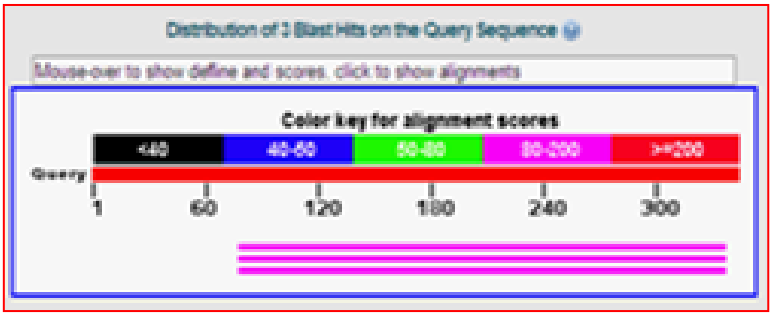
Blastn results of the merged sequence (1226 bp) indicated that 97 percent query coverage and 94 percent homology with Desulfobacteraceae family. Details of Blastn result obtained are presented in Table 16.

CLONE – 11

The sequence data obtained for the clone 11 with forward primer (T7U) was of 622 bp in size and vector screening indicated vector sequence from 1 to 50 bp. The details of results obtained are presented in Fig. 12A. The sequence data obtained with reverse primer (SP6) was 556 bp in size and vector screening

Merged sequence (342bp)

5’AACAATCCACGGCGGGGATTACGGTACTTGTTTCCACTTCTTCACATCAAAAAAATACCTTAATATCCTGTCAAAATTAAACGTGGTTAGCAAGCG
ACTTCTGGAACTCGATCTCATGGTGTGACAGGCGGTGTGTCTTGCCGGGAATGATCCCCGCTGTGATCCGCATTATTCAAAAATTCCGGGTTCTCGT
ATACTAGAACCTAATGCATTCTGAAAATGAAAGCGACTTTTTGGGATTACTCCCTTCCAGAATTGCCACCCTTTGTGCCGCCCATTGTACACATGAG
AAACCTGGACCTAAAGCAGAATACTTGAAGCATCCCCAACTTCCTCAAGTTA3’

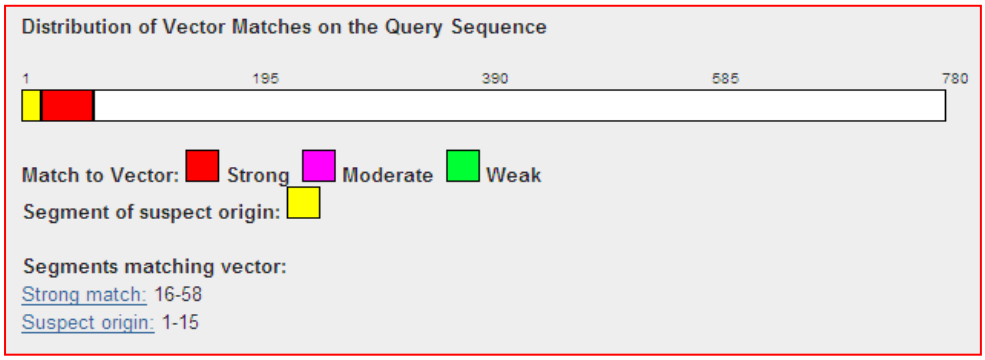


A. Blastn output

Fig. 9. Sequence analysis of clone 8

Table 14. Blastn result of clone 8

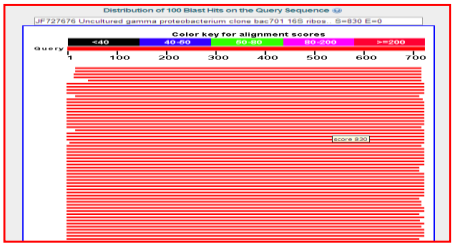
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
HM069044.1	Uncultured bacterium	121	75	4e-24	76
HM243997.1	Uncultured bacterium	115	75	2e-22	75
HM243306.1	Uncultured bacterium	115	75	2e-22	75



A. VacScreen output

Merged sequence (722bp)

5'GTTTGATCCTGGCTCAGATTGAACGCTGGCGGCATGCTTAACACATGCAAGTCGAACGGTAACAGGCCCTTCGGGGTGCTGACGAGTGGCGGACG
GGTGAGTAACGCGTGGGAATCTGCCTAGTAGTGGGGGATAGCCCGGAGAAATCCGGATTAATACCGCATACGCCCTACGGGGGAAAGCAGGGGAT
CTTCGGACCTTGCGCTATTGGATGAGCCCGCGTCTGATTAGCTTGTTGGTAGGGTAATGGCCTACCAAGGCAACGATCAGTAGCTGGTCTGAGAGGA
TGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGACAATGGGCGCAAGCCTGATCCAGCCAT
GCCGCGTGAGTGAAGAAGGCCTTCGGGTTGTAAAGCTCTTTCAGATGGGAAGAAACGGTCTAGGTTAATACCCTGGATTAATGACGGTACCGTCAG
AAGAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCTAGCGTTATCCGGATTCACTGGGCGTAAAGCGCGTGCAGGCGGTT
TCGTAAGTCGGACGTGAAAGCTCCTGCTTACTGGGAGAGGCCGTTCGATACTGCGGGACTTGAGGTCGGGAGAGGGACGTGAATTCCCGGTGTAGT
GGTGGAATGCGTAGATATCGGGGAGGAACACCAGTGGCGAAGCGGCGTCCTG**3'**

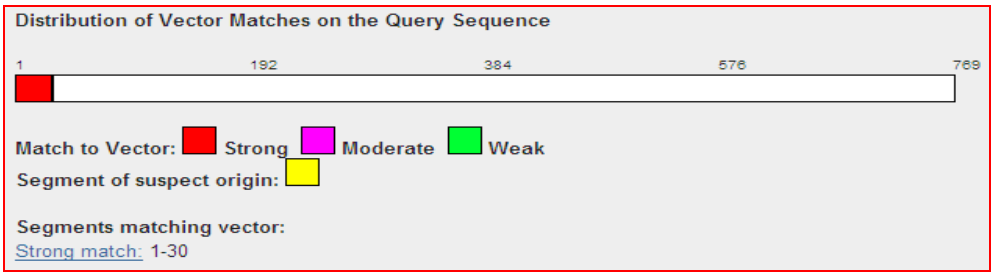


B. Blastn output

Fig. 10. Sequence analysis of clone 9

Table 15. Blastn result of clone 9

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EU653291.1	*Thiobacillus prosperus*	917	96	0.0	95
EU553290.1	*Thiobacillus prosperus*	911	96	0.0	90
HO678144.1	Uncultured bacterium	904	97	0.0	90
HO678143.1	Uncultured bacterium	898	93	0.0	89
AV034139.1	*Thiobacillus prosperus*	869	100	0.0	90



A. VecScreen output

Merged sequence (1226 bp)

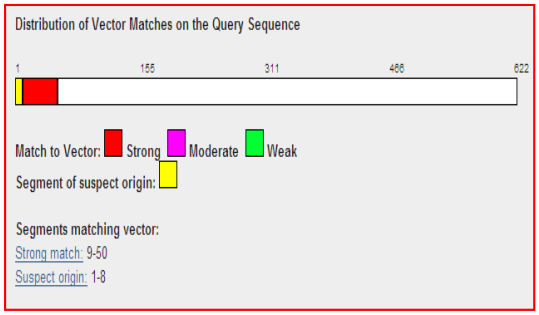
5'CGGCTACCTTGTTACGACTTCACCCCATCATCAGCCATACCTTAGGTGCCTGCCCCCCCTTGCGGGTTGGCCCGGCAACTTCTGGTACAACCAACT
CTCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCAACTTCATGAAGTCG
AGTTGCAGACTCCAATCTGAACTGAGACCGGCTTTGTGGGATTGGCTCCACCTTACGGCATTGCGACCCTCTGTACCGGCCATTGTAGTACGTGTGT
AGCCCTGGGCATAAAGGCCATGAGGACTTGACGTCATCCCCACCTTCCTCCACGTTAACCGTGGCAGTATCTCCAGAGTCCCCAACCGAATTGCTGG
TAACTGGAGACAGGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCAGCACCTGTCTTGGGGCTCCC
CGAAAGGCACTTCCTACTTTCATAAGAATTCCCCAGATGTCAAGCCCAGGTAAGGTTCTTCGCGTTGCGTCAAATAAAACACATACTCCACCGCTTG
TGCGGGCCCCCGTCAATTCCTTTGAGTTTTAATCTTGCGACCGTACTCCCCAGGCGGGATCATTAATGCGTTAACTGCGGGACTGCCGGGGTCAACA
CCCGCAAAACCTATTGATCACCGTTTACTGGGGGGGACAACCAAGGGTATCCAAATCTGGTTGATATATAGCGTGTTCCCCGGACAGAGTTTTGGCT
GAGGCCTCATCATTCCGCGGCGTGGTGGTCAGGGTACCCCATTCTGCAAAATCTCATGGGCCTGCCGTAGGAGTGTGGACCGTGTTCCAGTTCCAGG
GGGCGGCCATCTTCAGACCAGTACCATGTGGCAATGGTAGGCCCTCCTACCAACCAGTAAAAGGACGCGGGTCATCTCCAAACAACAGCTTCATGA
AGCGGCTATCTGGATCCATTCAACCGAAACCAAATGGATGTTATCCGGTATTAGCACCCCTTTGAAGTGTATCCCAAATTCGGAGGTAGATTACCCA
CGCATATCACCCGTGCGCCACTTATGCCCAAAGCAAGCTTCAGTTTTTTTCGTTTGACTGGGCGGAGAGAAGAAAAAGCACGCCGCCAGCGTCATTT
GACCACGATATCTTTATCTCTCTCTCTATATCACTGGACCCCTCTGTTTTTTTTTTTTTTTGGAGC3'

B. Blastn output

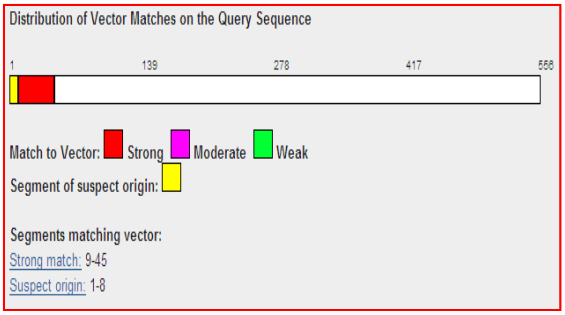
Fig. 11. Sequence analysis of clone 10

Table 16. Blastn result of clone 10

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AJ582707.1	Uncultured Desulfobacteraceae bacterium	1134	97	0.0	94
AY799892.1	Uncultured bacterium	1081	84	0.0	92
AY799901.1	Uncultured bacterium	1059	84	0.0	94
AY868687.1	Uncultured bacterium	1053	58	0.0	82
GU291339.1	Uncultured bacterium	1040	59	0.0	92



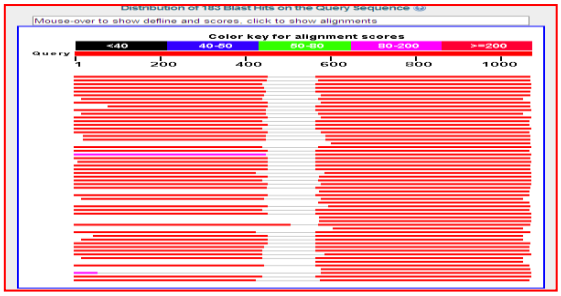
A. Vecscreen output



B. Vecscreen output

Merged sequence (1073bp)

5'GTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAATGCATGCAAGTCGAACGGAGAAAGTAGCAATATTTTCTCAGTGGCGAACGGGTG
AGTAACGCGTTGGTGACCTGCCCCAAAGCGGGGGATAACAGTTCGAAAGGATTGCTAATACCCCATGTTGTCATCGGGATTAGATGCCTGATGACT
AAAGGAGTAATTCGCTTTGGGAGGGGCCTGCGTCCCATCAGCTAGTTGGTAAGGTAATGGCTTACCAAGGCTTCGACGGGTAGGGGACCTGAGAGG
GTGACCCCCCACAAATGGAACTGAAACACGGTCCATACACCTACGGGTGGGCAGCAGCAGGGGAATATGGGTATCTGCGAAAGCGGGGAACAAGC
AACGCCGGCGTGTGCGATGAAGCCTTCGGGTCGTAAGGACTTTTCCAGGGATGGAGGAAGGCAGAACCCTCGAAAAATTCGGGTTAAAATGCCGA
GCGCGTAAAACTAGAAGCGAGGTTCCGAATTTTGGGGAAAACGGGCGCCGGTCTCACATAGGAGAGGAAACTTCGTCTATGGGGGAGAGTTGCCA
TCAGGTAGTAGGGATCCGAAAGGTGACCGCCCTTCGGGGAGCCTGCACAGGTGTTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAA
GTCCGTAACGAGCGCAACCCTCGCTGTGTGTTATACGTGTCACACGGGACTGCCGGTATTAAGCCGGAGGAAGGTGGGGACGAAGTCAAGTCAGCA
TGACCTATATGTCCTGGGATACACACACGCTACAATGGCCAGTACAACGGGTAGATAAGCCGCGAGGCGGAGCCAATCCTTAAAGCTGGTCTCAGT
TCAGATTGCAGGCTGCAACCCGCCTGCATGAAGACGGAGTTGCTAGTAACCGCAGGTCAGCTATACTGCGGTGAATACGTTCCCGGGCCTTGTACA
CACCGCCCGTCACGTCATGGGAGCTGGCAACACCTGAAGCCGGTATCCTAACCCGCAAGGGAGGGAGCCGTCGAGGGTGGGGTTGGTGACTGGGA
CGAAGTCGTAACAAGGTAGCCG3'



C. Blastn output

Fig. 12. Sequence analysis of clone 11

Table 17. Blastn result of clone 11

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AJ009487.1	Uncultured bacterium	760	89	0.0	93
AM490732.1	Uncultured bacterium	759	88	0.0	93
HQ697693.1	Uncultured bacterium	754	88	0.0	93
FM955656.1	Uncultured bacterium	742	88	0.0	93
AJ249100.1	Uncultured bacterium	742	88	0.0	94

indicated vector sequence from 1 to 45 bp. The details of results obtained are presented in Fig. 12B.

Blastn results of the merged sequence (1073 bp) indicated that 89 percent query coverage and 93 percent homology with uncultured bacterium. Details of Blastn result obtained are presented in Table 17.

CLONE – 12

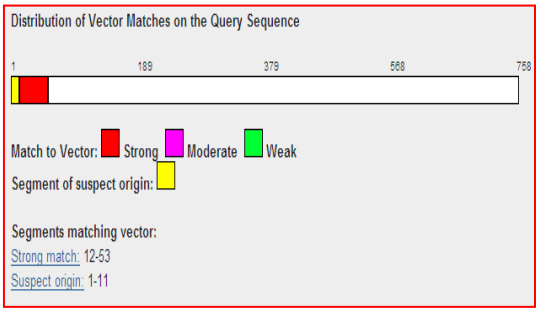
The sequence data obtained for the clone 12 with forward primer (T7U) was 758 bp in size and vector screening indicated vector sequence from 1 to 53bp.The details of results obtained are presented in Fig. 13A. The sequence data obtained with reverse primer (SP6) was 584 bp in size and vector screening indicated vector sequence from 1 to 70 bp. The details of results obtained are presented in Fig. 13B.

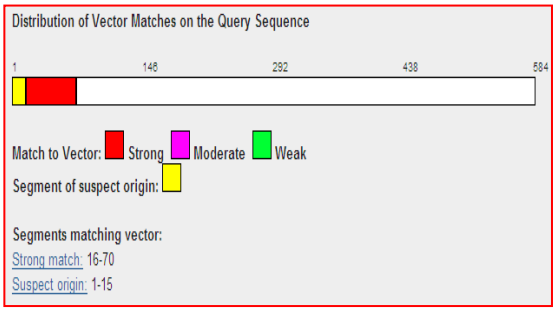
Blastn results of the merged sequence (1205 bp) indicated that 90 percent query coverage and 94 percent homology with Acidobacteria bacterium. Details of Blastn result obtained are presented in Table 18.

CLONE-13

The sequence data obtained for the clone 13 with forward primer (T7U) was 675 bp in size and vector screening indicated vector sequence from 1 to 89 bp. The details of results obtained are presented in Fig. 14A. The sequence data obtained with reverse primer (SP6) was 752 bp in size and vector screening indicated that no vector sequence found.

Blastn results of the merged sequence (1299 bp) indicated that 97 percent query coverage and 95 percent homology with uncultured Desulfobacteraceae. Details of Blastn result obtained are presented in Table 19.

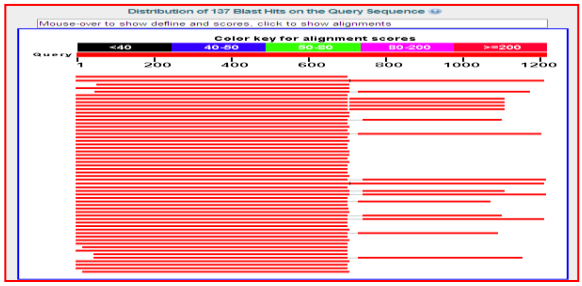




A. VecScreen output B. **VecScreen output**

Merged sequence (1205bp)

5'GTTTGATCCTGGCTCAGAGCGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGCGCGGCCGTAGCAATACGGCTGGCGCGTGGCGAACG
GGTGAGTAACACGTGGGCAACCTTCCCCCGGAAGTGGAATAACCCGCCGAAAGGCGGGCTAATACCACATAACATCCCGATGGCACAAGTCGTTG
GGATCAAAGCCGCAAGGCGTCTGGGGAGGGGCCCGCGGCCGATCAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCTGCGACCGGTAGCTGGCCT
GAGAGGGTGGCCGGCCACACTGGAACTGCGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGCAATGGGCGAAAGCCTGACG
CAGCGACGCCGCGTGGGGGATGAAGCCTTTCGGGGTGTAAACCCCTTTTAGCCCGGACGAATCCCTGGCAACAGGGTCGACGGTACGGGCAGAAA
AAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAAAACGTAGGGGGCCAGCGTTGCTCGGAATGACTGGGCGTAAAGGGTCTGTAGGCGGTGCG
GCAAGTCGGAAGTGAAATCCTGCGGCTTAACTGCAGGGCTGCTTCCGAAACTGCCGCGCTAGAGTGCGGGAGAGGGGAGTGGAATTGCCGGTGTA
GCGGTGAAATGCGTAGATATCGGCAGGAACACCGGGGCCATCCTGGAAATCAGGATTTTTGCAAGGACGCCGGTAGAGGTGTGCATGGCTGTGTCA
GTGTCCCGTGAGGTGTTGGGTTAAGTCCCCGCAACGAGCGCAACCCCTGTCCCCTAGCTGCCAGCGGTTCGGCCGGGCACTTAGGGGGACTGCTCC
GGATAACGGAGAGGAAGGTGGGGATGAGGTCAAGTCAGCATGGCCTTTATGTCCGGGGCTACACACGTGTTACAATGCGCACCACAAAGCGTTGC
CAACCCGCGAGGGGGATCCAATCGCATAAAAGCGCGCTCAGTTCGGATTGCAGTCTGCAACTCGACTGCATGAAGCCGGAATCGCTAGTAATGGCG
TATCAGAACGACGCCGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACATCACGAAAGGGGGCTGTACTAGAAGTCGCCAGGCCAACCCG
CAAGGGAGGCAGGCGCCCAAGGTGTGGTCCCTGATTGGGGTGAAGTCGTAACAAGGTAGCCG3'

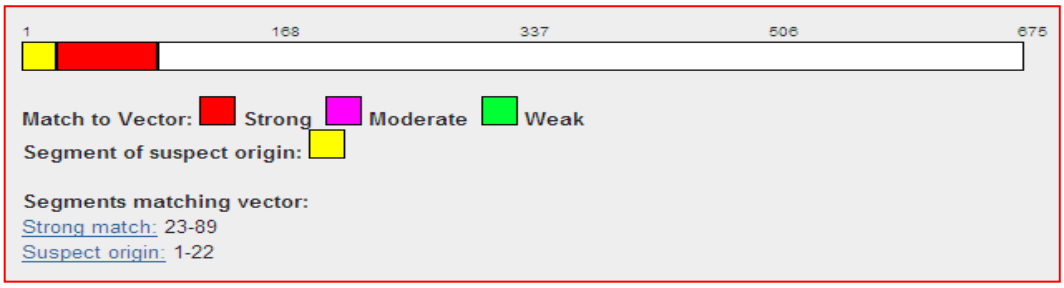


C. Blastn output

Fig. 13. Sequence analysis of clone 12

Table 18. Blastn result of clone 12

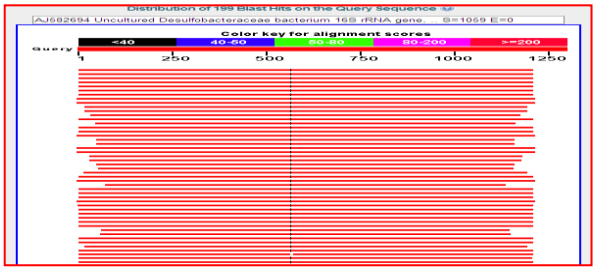
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EF197038.1	Uncultured bacterium	1068	57	0.0	94
AB254784.1	Uncultured bacterium	1035	99	0.0	93
AF523901.1	Uncultured bacterium	1027	93	0.0	95
DO791386.1	Uncultured bacterium	983	94	0.0	91
AM749736.1	Uncultured Acidobacteria bacterium	898	90	0.0	94



A. VecScreen output

Merged sequence (1299bp)

5'CGGCTACCTTGTTACGACTTCACCCCAATCATCAACCGTACCTTAGGTACCTGCCCCCCCTTGCGGGGTCAGCCCGGCAACTTCTGGTACAGCCAA
CTTTCGTGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGTGGCATGCTGATCCACGATTACTAGCGATTCCAACTTCACGGAGTC
GAGTTGCAGACTCCGATCCGAACTGAGAGCGGCTTTTTGGGATTAGCTCCTCCTCGCGGACTAGCAACCCTTTGTACCGCCCATTGTAGCACGTGTG
TAGCCCTGGATATAAAGACCATGAGGACTTGACGTCGTCCCCACCTTCCTCCCCGTTAACCGAGGCAGTATCTCTAGAGTGCCTAACTGAATAAGGG
CAACTAGAGAGAGGGGTTGCGCTCGTTGCGGGACTTAACCCAAATCTCACGACACGAGCTGACGACAGCCATGCAGCACCTGTCACTCTGTCCCTG
CAAGCAGGGAAGACCCCGTTTCCGGGGCTGTCAAAGGATGTCAAATCCAGGTAAGGTTCTTCGCGTTGCGTCGAATTGAACCATTAAGTGCCGCCT
GGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATGACGGGGCCCGCACAAGCGGTGGAGCATGTGGTTCAATTCGACGCAACGCGAAGAACCTT
ACCTGGATTTGACATCCTTTGACAGCCCCGGAAACGGGGCTCTTCCCTGCTTGCAGGGACAGAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCTCTCTAGTTGCCCTTATTCAGTTAGGCACTCTAGAGATACTGCCTCGGTTAACG
GGGAGGAAGGTGGGGACGACGTCTAGTCCTCATGGTCTTTATATCCAGGGCTACACACGTGCTACAATGGGCGGTACAAAGGGCTGCTAGTCCGCG
AGGAGGAGCTAATCCCAAAAAGCCGCTCTCAGTTCGGATCGGGTCTGCAACTCGACTCCGTGAAGTTGGAATCGCTAGTAATCGTGGATCAGCATG
CCACGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAAAGTTGGCTGTACCAGAAGTTGCCGGGCTGACCCCGCAAGGGGGG
GCAGGTACATAAGGTACGGTTGATGATTGGGGTGAAGTCGTAACAAGGTATTCCGTAATCACTAGGACGGCCGCCTGCAGGTTTGTATCCTATATG
GGACGATCTCCCAACGCGTGGAGCCGGCCCGCGTTTTTTTCCG3'



B. Blastn output

Fig. 14. Sequence analysis of clone 13

Table 19. Blastn result of clone 13

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
HM244005.1	uncultured Desulfobacteraceae	1116	97	0.0	95
HM243765.1	Uncultured bacterium	1116	92	0.0	98
FJ716434.1	Uncultured bacterium	1110	92	0.0	98
FJ716431.1	Uncultured bacterium	1105	92	0.0	98
HM244002.1	Uncultured bacterium	1099	92	0.0	98

CLONE - 14

The sequence data obtained for the clone 14 with forward primer (T7U) was 736 bp in size and vector screening indicated vector sequence from 1 to 61bp. The details of results obtained are presented in Fig. 15A. The sequence data obtained with reverse primer (SP6) was 669 bp in size and vector screening indicated vector sequence from 1 to 74 bp. The details of results obtained are presented in Fig. 15B.

Blastn results of the merged sequence (1269 bp) indicated that 98 percent query coverage and 98 percent homology with Chloroflexi bacterium. Details of Blastn result obtained are presented in Table 20.

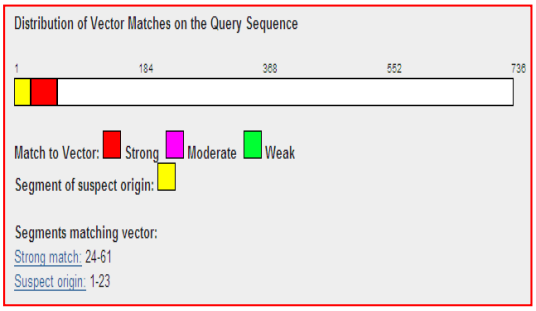
CLONE - 15

The sequence data obtained for the clone 15 with forward primer (T7U) was 626 bp in size and vector screening indicated vector sequence from 1 to 36bp. The details of results obtained are presented in Fig. 16A. The sequence data obtained with reverse primer (SP6) was 641bp in size and vector screening indicated vector sequence from 1 to 43 bp. The details of results obtained are presented in Fig. 16B.

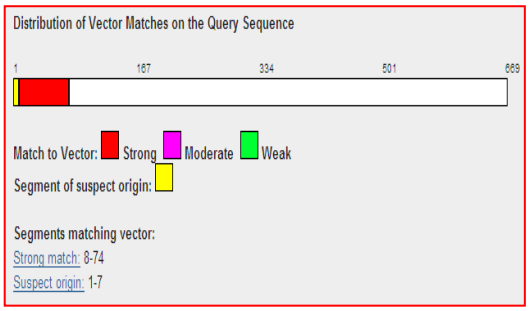
Blastn results of the merged sequence (1181 bp) indicated that 59 percent query coverage and 97 percent homology with uncultured bacterium. Details of Blastn result obtained are presented in Table 21.

CLONE-16

The sequence data obtained for the clone 16 with forward primer (T7U) was 746 bp in size and vector screening indicated vector sequence from 1 to 57bp. The details of results obtained are presented in Fig. 17A. The sequence data obtained with reverse primer (SP6) was 707 bp in size and vector screening



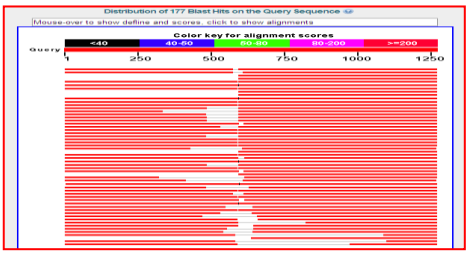
A. VecScreen output



B. VecScreen output

Merged sequence (1269bp)

5'CCTCAGGTAAAATCAAAGGAATTGCGGGGCCCGCACAAGCAGCGGAGCTTTGGTTTAATTGGGCAACCGAGGGAACCTTTCCTGGGCTTGACATG
CACGTAGTAGTGAAGCGAAAGCGGATCGATCCTTCGGGAAGCGTGCACAGGTGTGCATGGCTGTCGTCAGCTCGTGCCGTGAGGTGTTCGGTTAAG
TCCGGTAACGAGCGCAACCCTCGCTGTTAGTTACATGTGTCTAACGGGACTGCCGGTTTAATGCCGGAGGAAGGTGGGGATGACGTCAAGTCAGCA
TGGCCTCTATGCCTAGGGCTACACACACGCTACAATGGCCGGTACAATGGGTTGCCAAGCCGCGAGGCGGAGCTAATCCTCCAAAGCCGGTCTCAG
TTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGTCGGAGTTGCTAGTAACCGCAGGTCAGCCATACTGCGGTGAATGCGTTCCCGGGCCTTGTACA
CACCGCCCGTCACGTCATGGGAGCTGGTAACACCTGAAGCCGGTGGGCTAACCGAAAGGGGGCAGCCGTCGAAGGTGGGACTGGTGACTGGGACG
AAGTCGTAACAAGGTAGCCGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGGGATCCCTTGGATTCGTCT
GAGGGTGAGAGTGGCGAACGGGTGAGTAATACGTAGGTGACCTGCCCCAAAGTAGGGAATAACCACTGGAAACGGTGGCTAATACTCTATGTGCT
CGGTGGTTCGGTCTGCCGAGTAAAGTACCTCTTTGAGGTGGCGCTTTGGGAGGGGCCTGCGGCCCATCAGGTAGTTGGTAGGGTAATGGCCTACCA
AGCCGATGACGGGTAGGGGGCGTGAGAGCGTGACCCCCCACACTGGCACTGAGATACGGGCCAGACTCCTACGGGAGGCAGCAGTGAGGAATATT
GCCCAATGGACGCAAGTCTGAGGCAGCAACGCCGCGTGGGCGATGAAGGCCTTCGGGTTGTAAAGCCCTTTTCTGGGGGACGAGGAAGGACGGTA
CCTCAGGAATAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGATTCACTGGGCGTAAAGCGCGTGCA
GGCGGTTTCGTAAGTCGGACGTGAAAGCTCCTGGCTTACTGGGAGAGGTCGTTCGATACTGCGGGGCTTGAGGTCGGGAGAGGGACGTGGAATTCC
CGGTGTAGTGGTGGAATGCGTAGAT3'

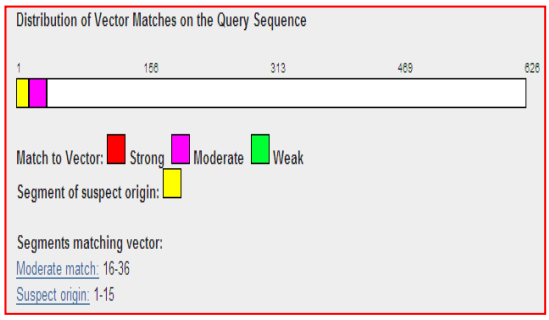


C. Blastn output

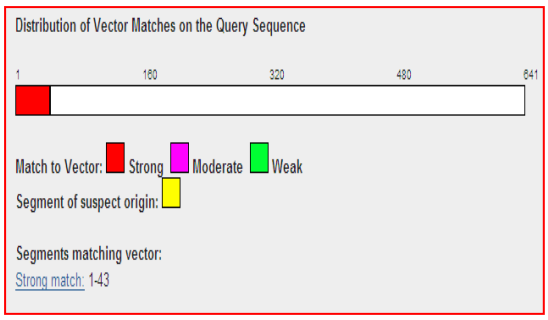
Fig. 15. Sequence analysis of clone 14

Table 20. Blastn result of clone 14

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
CO463738.1	Uncultured Chloroflexi bacterium	1077	98	0.0	98
FN863178.1	Uncultured bacterium	1059	96	0.0	95
EU376188.1	Uncultured bacterium	1037	83	0.0	94
GU682761.1	Uncultured bacterium	1030	83	0.0	94
EU488039.1	Uncultured bacterium	1044	99	0.0	94



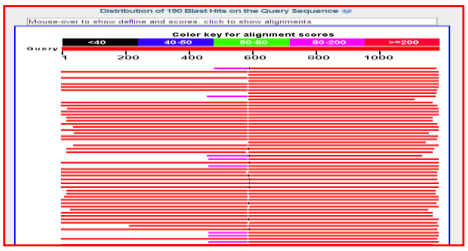
A. VecScreen output



B. VecScreen output

Merged sequence

5'GTTTGATCCTGGCTCACATTGAACGCTGGCGGCATGCTTAACACATGCAAGTCGAACGGTAACAGGGGCTTCGGCCCGCTGACGAGCGGCGGACG
GGTGAGTAATACATAGGAATCTGCCCTGTAGTGGGGGACAACTTAGGGAAACTTAAGCAATACACTACACCTACGGGTGAAAGACTTCGGTCGCTA
TAGGATGAGCCTATGTCTGATTAGCTAGTTGGTAGGGTAAAGGCCTACCAAGGCAACGATCAGTAGCTGGTCTGAGAGGACGATCAGCCACACTGG
GACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCAATGCCGCGTGTGTGAAGA
AGGCCTGCGGGTTGTGAAGCACTTTCGGTAGGGAAGATAATGACGGTACCTACAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAAT
ACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGCGGTTTGTTAATCGGGTGTGAATCCCTGGGCTCACCTAGGAATTGC
ATTCATACTGGAGGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTGGATGCAACGGAAGAACCTTACTACTCTTGACATCCT
CAGAACTTGTCAGAAATGACTTGGTGCCTTCGGGAACTGAGTGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCC
CGTAACGAGCGCAACCCTTGTCCATAGTTGCCAGCGAGTAATGTCGGGAACTCTATGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGA
CGTCAAGTCATCATGGCCCTTACGAGTAGGGCTACACACGTGCTACAATGGAAAGTACAGAGGGCTGCTAAAGCGCGAGCCGGAGCAAATCCCAT
AAAACTTTTCGTAGTCCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGTAATCGCGGATCAGAATGCCGCGGTGAATACGTTCC
CGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGCTGACCAGAAGCCGGTAGTCTAACCGCAAGGAGGACGCCGTCCACGGTGTGGTCAA
TGACTGGGGTGAAGTCGTAACAAGGTAGCCG**3'**

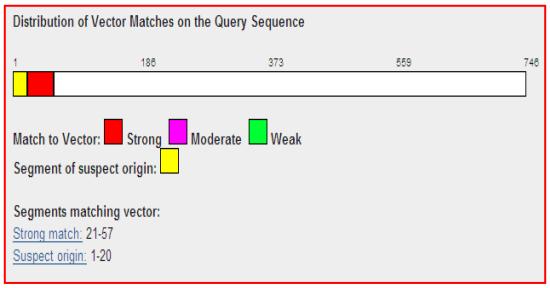


C. Blastn output

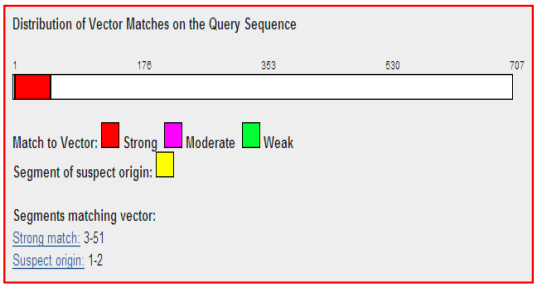
Fig. 16. Sequence analysis of clone 15

Table 21. Blastn result of clone 15

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
FI154783.1	Uncultured bacterium	1007	59	0.0	97
EU700176.1	Uncultured bacterium	924	99	0.0	94
EU700160.1	Uncultured bacterium	924	50	0.0	94
EU735702.1	Uncultured bacterium	908	99	0.0	94
JF9735702.1	Uncultured bacterium	896	90	0.0	93



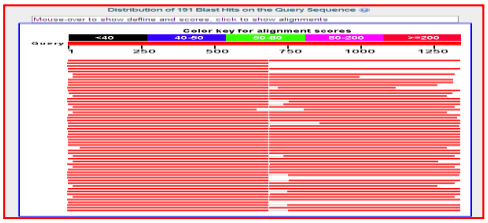
A. VecScreen output



B. VecScreen output

Merged sequence (1336bp)

5'CGGCTACCTTGTTACGACTTCGTCCCAGTCACCGACCCCACCCTCGACGGCTCCCTCCCTTACGGGTTAGGATACCGGCTTCAGGTGTTGCCAGCT
CCCATGACGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCAGTATGGCTGACCTGCGGTTACTAGCAACTCCGTCTTCATGCAGGC
GGGTTGCAGCCTGCAATCTGAACTGAGGACAGCTTTGAGGATTGGCTCCACCTCGCGGTTTCGCGACCTACTGTACTGCCCATTGTAGCGTGTGTGT
AGCCCAGGACATAAAGGCCATGCTGACTTGACGTCATCCCCACCTTCCTCCGGCTTAAGACCGGCGGTCCCGCGTGACACATATAACACACGGCGA
GGGTTGCGCTCGTTAGCGGACTTAACCGAACATCTCACGACACGAGCTGACGACAGCCATGCAGCACCTGTGCAAGCTCCCCGAAGGGTCGGTCTG
CTTTCGCTTCCCTACCACTTGCATGTCAAGCCCTGGTAAGGTTCTTCGTGTAACATCGAATTAAACCACACGCTCCGCTGCTTGTGCGGGCCCCCGTC
AATTCCTTTGAGTTTTAACCTTGCGGCCGTAGTCCCCAGGCGGTAGACTTAACGCGTTAGCTCCGGCACAGATGGATTTGAACCCACCTACGCCAAG
TCTACATCGGGTCCTCCGGATTTTAGCATTCACCCTCCCCGGAATCCATTCCCTCTCACACTCCAACCCGGTAGTATGACGACCTTTCCAGTAACCAG
GAGCTTCACATCAACTTCGAGCGCTAGCTCGCGCTTACGCCAGTAAATCCGGTAACGCTCGCTCTACGTTTACCGCGGCTGCTGCACGTAGTTAGCC
GAGGCTTATTCTCAGGGTACTGTCCTTCCTCATCCCCTGCAAAAGCGCTTTACAACCCGAAGGCCTTCATCGCGCACGCGGCGTTGCTCTTTCAGGCT
TTCGCCCATTACCAATATTCCCTACTGCTGCCACCCGTAGGTGTCTGGCCCGTGTTTCAGTGCCAGTGTGGGGGGCCACCCTCTCAGGTCTCCCTACC
CGTCATCGCCTTGGTAGGCCGTTACCCTACCAACTAGCTGATGGGACGCAGGCCCCTCCCGAAGCTAATATGGGATCCTTTACTTATCCTTCTCAAA
CTGGAATAAGCACATGGGGTATTAGCAGCCCTTTCGAGCTGTTATCCCCCGCTTCGGGGCAGGTCACCAACGCGTTACTCACCCGTTCGCCACTAAG
AACATTGCTGCTCTTCGTTCGACTTGCATGTATTAGGCACGCCGCCAGCGTTCATCCTGAGCCAGGATCAAACTC**3'**



C. Blastn output

Fig. 17. Sequence analysis of clone 16

Table 22. Blastn result of clone 16

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AB109437.1	*Levilinea saccharolytica*	1063	97	0.0	97
AJ009456.1	Uncultured bacterium	1079	96	0.0	94
HQ602616.1	Uncultured bacterium	1072	50	0.0	94
HQ900950.1	Uncultured bacterium	1054	99	0.0	94
AJ009500.1	Uncultured bacterium	1042	90	0.0	93

indicated vector sequence from 1 to 51bp. The details of results obtained are presented in Fig. 17B.

Blastn results of the merged sequence (1336 bp) indicated that 97 percent query coverage and 97 percent homology with *Levilinea saccharolytica*. Details of Blastn result obtained are presented in Table 22.

CLONE -17

The sequence data obtained for the clone 17 with forward primer (T7U) was 664 bp in size and vector screening indicated vector sequence from 1 to 43 bp. The details of results obtained are presented in Fig. 18A. The sequence data obtained with reverse primer (SP6) was 697 bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1309 bp) indicated that 95 percent query coverage and 93 percent homology with uncultured bacterium. Details of Blastn result obtained are presented in Table 23.

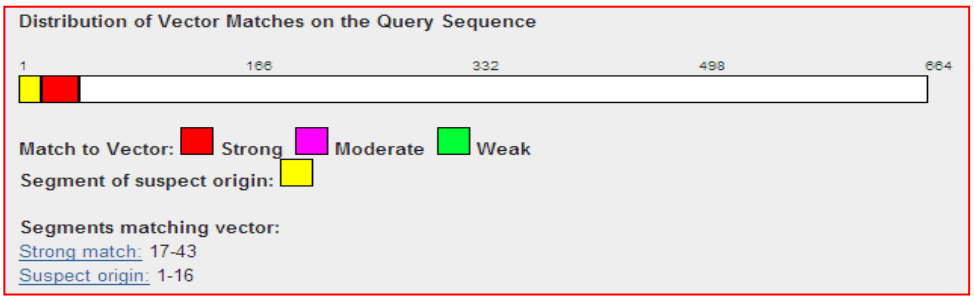
CLONE – 18

The sequence data obtained for the clone 18 with forward primer (T7U) was 666 bp in size and vector screening indicated vector sequence from 1 to 83 bp. The details of results obtained are presented in Fig. 19A. The sequence data obtained with reverse primer (SP6) was 334 bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (911 bp) indicated that 61 percent query coverage and 95 percent homology with Chloroflexi bacterium. Details of Blastn result obtained are presented in Table 24.

CLONE- 19

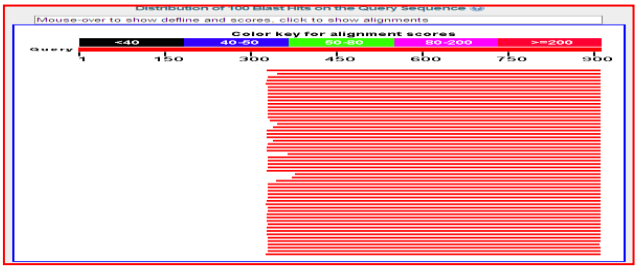
The sequence data obtained for the clone 19 with forward primer (T7U) was 666 bp in size and vector screening indicated that there was no vector



A. VecScreen output

Merged sequence (1309bp)

5'ATAGCGTATGACCCAATTTCGGAGTACGCTTTCACGTTAGTTTGCGCGGGGGGGCAGGGTAAGCTGAATTATTGGAAGGACGGCCAAGCGGAGA
GATGGCACAGGTTATTTCCGGAAGAAGTGATCATACTCTCAGCGTTGCCTCACAAGAGCATGTTATCTTGGTTTTCCCCAGTGCAGATTGCCCAGGG
CCTGCCATATATGTGCTAGAGTGCGTTCTCAGTGGAACGATCGTGCAACACTCATAGTAGATCTGTTAGGTATACAGTAACTATATGGTGACTGGGT
AAGACGGAACAGAGGCAGGCAAGGATGCAAGTTCGGAAGATCGCCAGCTTCGGGCAGGGGGGCGTACGCGATACAAGGGCAAGAACGGATGGTT
GGTCCCGAGGGGGAGGGAAGATCAAAAACCTGCTACTCACTCGTCCGGCAGTAAGTTAACCAACAGCATCAAGCCGGCAAGCTATCCAAGTGTCA
GGCTATCAAGTCGAAGGATCTAAATTAAATCCCTTAATCATCTCCCCGTCACTCCCGTCAGCTTGGGAGTACCTCAGGTTAGCGGGGTGGATAAACT
CAATGGGGCATCCATGCAAGGGCACGCAACGAGCGGGGGAAGGAGAACGATAAGGAAACCTTAATCATAGGGACGCCCTCAGTTTTTTTTTCTTAA
GTGACTCCCAACGCTTGGAGCCGCGGGATTAAAGTTTGATTGTGGCTCATATTGCCGCTGGTGCATGCTTAAGCTTGCCATCCTCCCGATAGTTAGA
CTATCGGTTTGCTTGGAGTGACTCACGTGGGGTGACGGGCAGTGTGTTGAGGCCCGTGGGGGATTCCTGCGCTTCAACAGTATGCTACGCTACGACT
GCATGAAGTCGATCTGCATAGGATGATCCTTAGTACGATTATGTGCTGATTTGTATTTGCCTACGAATGGACTTCCACTGCACGGTCATAAATAAGA
ATGTGAGAAGCCCCTGCCCCAAAACCCCGACCAAAAAATTGCCGGCCACCCCCCCCCCTCGCGGCAAATTTGTCAACAGGGTTAACCTTGAAGTCC
CCCGCCGAAACCCGCCGGGAAAAGGGGGGAAGGGGTTGGCAAACTCCTTTCGGGAAATTTTACCTAAAATTGGCGGAAAACACGGGGTTAGAGAA
AAACCCCTTGGATCTCCTGGGGACCCTCCCGTTCCGAAAAGAGCCCCCGGCGGGGGGGGGGCGTTCTTTCTCTTTTAATCTCAAGTTTAAAAAAACC
CTGGGGGACCAACGCTTTGTGGGGGCCCCCCCCCAAAAATCCCCTGGGGGTTTTAAACGG **3'**

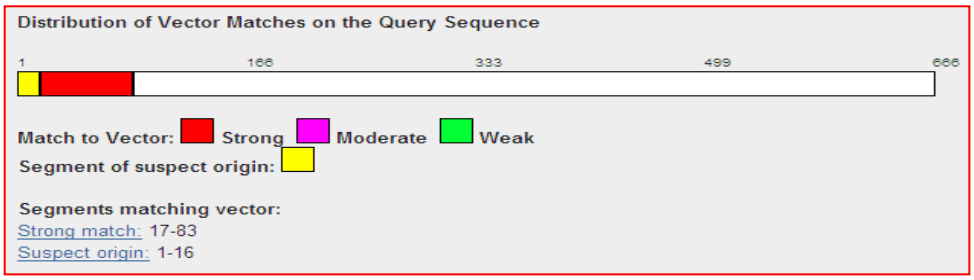


B. Blastn output

Fig. 18. Sequence analysis of clone 17

Table 23. Blastn result of clone 17

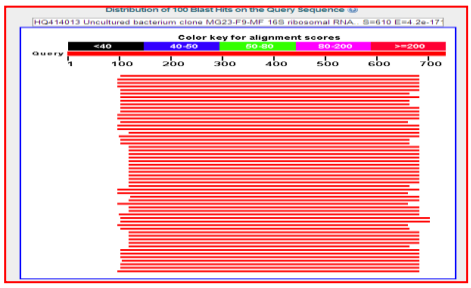
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AJ105437.1	Uncultured bacterium	1263	95	0.0	93
AJ409456.1	Uncultured bacterium	1259	96	0.0	91
HQ202616.1	Uncultured bacterium	1172	59	0.0	94
HQ200950.1	Uncultured bacterium	1154	99	0.0	91
AJ109500.1	Uncultured bacterium	1142	90	0.0	93



A. VecScreen output

Merged sequence (911bp)

5'GAAACGATACCAACGAGGCCGATCCAGGGACCAAGGGCAAGGACAAGGGCTAAAAAAGGGAGGGGAGCACTCCAAAAGCCAATCAGGTGGGA
GGCAGGGCATTCCAAAAGTGGAGGGCGGGAGCGGCAGGTCAGCGAACGGCGGGACGATAGGTGCGGGACTTGGACACACCGGCGGGGGGCTGCA
TCACTATCATCTCCTACACCGGAAGGGAGTTTTTTTTTTTTTCGATTAGGCAGTTAGAGAGGGACCTCTGAGAAAAAAAAAAAAAAATTTTTTTTTGA
AGCAGGATAACGGCACAACTTTTTTTTTTTTTTTTTTCTTTTAGGCAGCTACCTTGTTACTACTTCGTCCCACTCACCAGTCCCACCTTCCACGGCTGC
CCCCTTTCGGTTACCCCACCGGCTTCAGGTGTTACCAGCTCCCATGACGTGACGGGCGGTGTGTACAAGGCCCGGGAACGCATTCACCGCAGTATGG
CTGACCTGCGGTTACTAGCAACTCCGACTTCTTGCTTGCGAGTTGCGGCCTGCAATCCGAACTGACACCGGCTTTGGAGGATTAACTCCACCTCGCG
GCTTGGCAACCCATTGTACCGGCCATTGTATCGTGTGTGTACCACTAGGCATAGAGGCCATGCTGACTTGACGTCATCCGCACCTTCCTCCGGCATT
AAACCGGCAGTCCCGTTAGACACATGTAACTAACAGCGAGGGTTGCGCTCGTTACCGGACTTAACCGAACACCTCACGGCACGAGCTGACGACAGC
CATGCAGCACCTGTGCACGCTTCCCGAAGGATCGATCCGCTTTCGCTTCACTACTACGTGCATGTCAAGCCCAGGTAAAGTTCCTCGTGTTGCCTCG
AATTAAACACACGCTCCGCTGCTTGTGCGGACCCCCGTCAATTCC3'

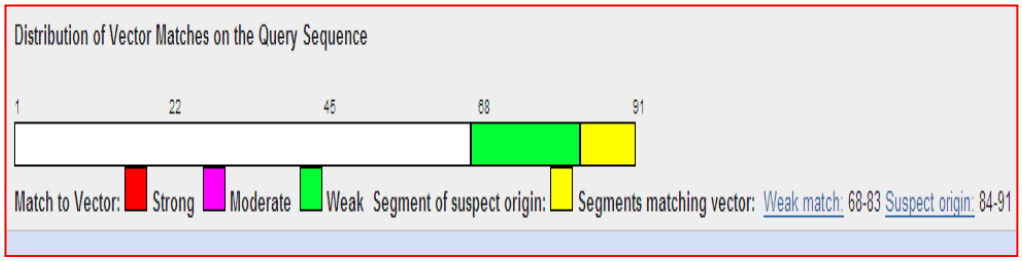


B. Blastn output

Fig. 19. Sequence analysis of clone 18

Table 24. Blastn result of clone 18

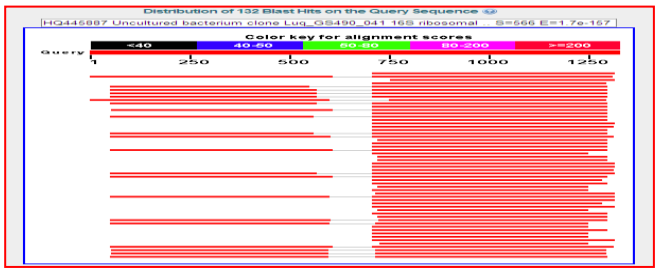
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EF688197.1	Uncultured bacterium	909	65	0.0	94
FN563178.1	Uncultured Chloroflexi bacterium	907	61	0.0	95
DO811835.1	Uncultured bacterium	893	63	0.0	94
DO394932.1	Uncultured bacterium	876	63	0.0	93
GU389465.1	Uncultured bacterium	865	63	0.0	93



A. VecScreen output

Merged sequence (729bp)

5'CTCGTGGGGAACTTTACTCGATCGATCTATTTTTTCTTCATAGAGATATCTTTTTTATCTACGTTGGAAATGGCCAAGGGCCGAAATGGCCGCGGG
ATTACGCTACCTTGTTACGACTCCCACCCCAGTCATTGACCACCCGTGGTGCACGTCCTCCCAAATGTCACACTATCCACTTCAGGTGCAGCCAACTC
CCATGGTGTGACGGGCGGTGTGTACAGGCCCGGGAACGTATTCACTGCGACTTCTGTTCGCATTACACGTTCCGACTTCTGGAGTCAGTTGCAACTC
CAATCCGGACTACACCGGTTTTCTGGGATTAGCTCCCCCTCGCGGGTTTGCACCCTCTGTACCGGCCATTGTACACGTGTGTACCCTGCCCATAAGGC
CATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCCGAGTCTCCCTAAAGTTCCCGGCCGACCCGCTGGCACTAGGGACAAGGTTGCG
CTCGTTACGGAACTTAACCCACATCTCACGACACAGCTGACACAGCCATGCACCCTGTCACTCGTCCCGAAGCACCAAGTCATCTCTGACAAGTTCC
AGATGTCGGGCAGGAGTCTTCCGTTGCTCAATTACCCATGCTCACCGCTTGGCGGCCCCCAATTCCTTTGAATTTAACTGCGCGACTCCCAGGGGAA
ATTACCGTACTGCCCCTAAAATTACCCCCAGGCAATTCCACGTTAGGGG3'



B. Blastn output

Fig. 20. Sequence analysis of clone 19

Table 25. Blastn result of clone 19

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AF424161.1	Uncultured gamma poteobacterium	676	79	0.0	87
CP001905.1	*Thioalkalivibrio* sp.	649	79	0.0	86
CP001099.1	*Thioalkalivibrio sulfidophilus*	649	79	0.0	86
FR749905.1	*Thioalkalivibrio versutus*	638	79	2e-179	86
EU709879.1	*Thioalkalivibrio* sp.	638	79	2e-179	86

sequence. The sequence data obtained with reverse primer (SP6) was 91 bp in size and vector screening indicated vector sequence was from 68 to 91bp. The details of results obtained are presented in Fig. 20A.

Blastn results of the merged sequence (729 bp) indicated that 79 percent query coverage and 86 percent homology with *Thioalklivibrio* sp. Details of Blastn result obtained are presented in Table 25.

CLONE- 20

The sequence data obtained for the clone 20 with forward primer (T7U) was 674 bp in size and vector screening indicated vector sequence from 1 to 51bp. The details of results obtained are presented in Fig. 21A. The sequence data obtained with reverse primer (SP6) was 711bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1320 bp) indicated that 79 percent query coverage and 85 percent homology with uncultured Acidobacteria sp. Details of Blastn result obtained are presented in Table 26.

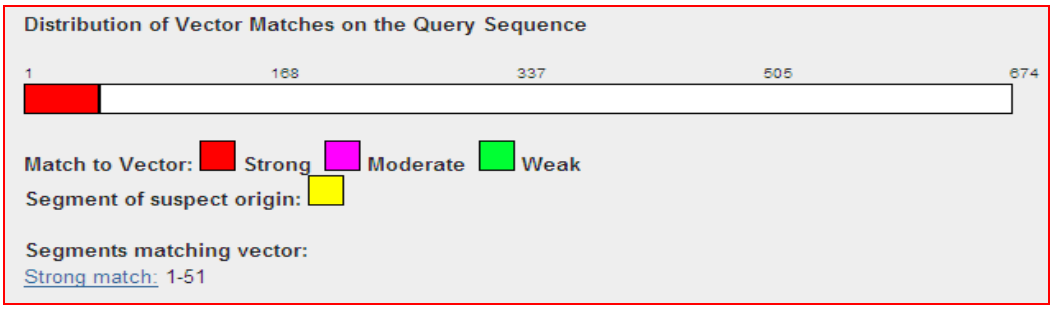
CLONE – 21

The sequence data obtained for the clone 21 with forward primer (T7U) was 931bp in size and vector screening indicated vector sequence from 1 to 52 bp. The details of results obtained are presented in Fig. 22A. The sequence data obtained with reverse primer (SP6) was 450 bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1328 bp) indicated that 79 percent query coverage and 92 percent homology with *Acidobacterium* sp. Details of Blastn result obtained are presented in Table 27.

CLONE -22

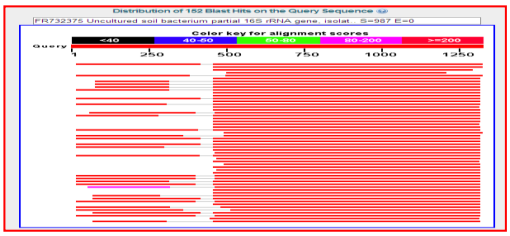
The sequence data obtained for the clone 22 with forward primer (T7U) was 972 bp in size and vector screening indicated vector sequence from 1 to 52



A. VecScreen output

Merged sequence (1320bp)

5'ATTGGAATAGGGGTGGGCCCCCTGGGGGTCCCTCCGCAATTAACCGGTAAATTTCCGCGCCTGGGGAGTACGGGCGGGGAGGTTGAAACTCAAA
GAATGACGGGGGCCCGCCCCAGCGGTGGAACAAGGTGTTTCGTTTGAGCAACGCGGAGAACCTACCCCGGCCTGATCCGGCGACGCATCTGAGAA
ATCAGGATTTTTGCAAGGGACGCCGGTAGAGGTGGCATGGCGTGTCAGTCGTCCCGTGAGGGGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTG
TCCCTAGTTGCCAGCGGTTCGGCCGGGCACTTTAGGGGGACTGCTCCGGATAACGGAGAGGAAGGTGGGGAGAGGTCAAGTCAGCATGGCCGTTAT
GTCCGGGGTACACAAATGTTACAATGCGCACCACAAAGCGTTGCCAACACACGAGGGGGAGCCAATCGCATAAAAGCGCGCTCAGTTCGGATTCTC
GTCTCCAACTGACTCAGAAGCCGGAATCGCTAGTAATGGGTATCACTACGACGCCGGAATACGTTCCCGGGCCTGTACACACCGCCCGTCACATCA
CGAAAGGGGGCTGACTAGAAGTCGCCAGGCCAATTTTTTGGGAGGCAGGGGCTTCGACGTTAGGGTCCCGATGGGGGAATGTCTCAATTTATCTCT
TCTCTTTTCTCATGACCTCAATCTTCCTCTTTGTGCCTGGCTCAGAGCGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGCGCGGCCGTAG
CAATACGGCTGGCGCGTGGCGAACGGGTGAGTAACACGTGGGCAACCTTCCCCCGGAAGTGGAATAACCCGCCGAAAGGCGGGCTAATACCACAT
AACATCCCGATGGCACAAGTCGTTGGGATCAAAGCCGCAAGGCGTCTGGGGAGGGGCCCGCGGCCGATCAGCTAGTTGGTGAGGTAACGGCTCAC
CAAGGCTGCGACCGGTAGCTGGCCTGAGAGGGTGGCCGGCCACACTGGAACTGCGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATC
TTGCGCAATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGGGGGATGAAGCCTTTCGGGGTGTAAACCCCTTTAAGCCCGGACGAATCCCTGGGAA
CAGGGTCGACGGTACGGGCAAAAAAAGCCCCGGCTAACTAGTGGCCAGCAGCCGCGGGAAAACCTAAGGGGCCCAGGTTGCTCCGAAATGACTGG
GGCTAAGGGCCTGTAGGGGGGCGCCATTCGAAGTGAATCTGGGGCTTATTTCAGGGGTGTTCCAAACTCCCCCCAAAG 3'

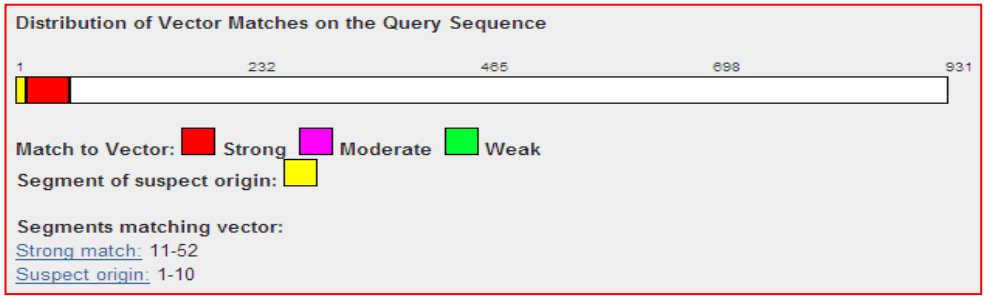


B. Blastn output

Fig. 21. Sequence analysis of clone 20

Table 26. Blastn result of clone 20

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EF197088.1	Uncultured bacterium	747	79	0.0	88
AB254784.1	Uncultured bacterium	737	91	0.0	86
AF523900.1	Uncultured bacterium	726	79	0.0	90
OO791386.1	Uncultured bacterium	680	79	6e-172	87
FJ592748.1	Uncultured Acidobacteria bacterium	614	79	2e-171	85

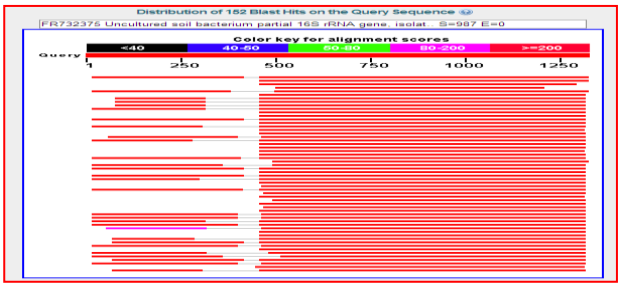


details of results obtained are presented in Fig. 25A. The sequence data obtained

A. VecScreen output

Merged sequence (1328bp)

5'TTTTTGTTTATTTTTTGCAAACGAGCGCAACCCCTGTCCCTTAGTTCCCACGGTTCGGCCGGGCACTTTAGGGGGATTGTCCGGATAACGGAGAGG
AAGGTGGGGATGAGGTCAAGTCAGCATGGCCTTTATCTCCGGGGTACACACGTGTTACAATGCGCACCACAAAGCGTTGCCAACCCGGAGGGGGA
GCCAATCGCATAAAAGCGCGCTCAGTTCGGATTGCAGTTGCAACTGATGCATGAAGCCGGAATCGTAGTAATGGCGTATCAGAACGACGCCGTGAA
TACGTTCCCGGGCCTTGTACACACCGCCCGTCACATCACGAAAGGGGGCTGTACTAGAAGTCGCCAGGCCAACCCGCAAGGGAGGCAGGGGCACA
AGGTGTGGTCCCTGATTGGGGGAAGTCGTAACAAGGTATTTCTTTTAATCATTAGGAACGGCCGCTCATTTTTTTCTTGCTCAGAGCGAACGCTGGC
GGCGTGCCTAACACATGCAAGTCGAACGCGCGGCCGTAGCAATACGGCTGGCGCGTGGCGAACGGGTGAGTAACACGTGGGCAACCTTCCCCCGG
AAGTGGAATAACCCGCCGAAAGGCGGGCTAATACCACATAACATCCCGATGGCACAAGTCGTTGGGATCAAAGCCGCAAGGCGTCTGGGGAGGGG
CCCGCGGCCGATCAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCTGCGACCGGTAGCTGGCCTGAGAGGGTGGCCGGCCACACTGGAACTGCGA
CACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGCAATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGGGGGATGAAGCCTTTC
GGGGTGTAAACCCCTTTTAGCCCGGACGAATCCCTGGCAACAGGGTCGACGGTACGGGCAGAAAAAGCCCCGGCTAACTACGTGCCAGCAGCCGC
GGTAAAACGTAGGGGGCCAGCGTTGCTCGGAATGACTGGGCGTAAAGGGTCTGTAGGCGGTGCGGCAAGTCGGAAGTGAAATCCTGCGGCTTTACT
GCAGGGCTGCTTCCGAAACTGCCGCGCTAGAGTGCGGGAGAGGGGAGTGGAATTGCCGGTGTAGCGGTGAAATGCGTAGATATCGGCAGGAACAC
CGGAGGCGAAGCGGCTCCCTGGACCGCAACTGACGCTGAGAGACGAAAGCTGGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCCAGCCCTAA
ACCATCGGACTAGCGTGGGCCCGTTGGGGGTCGTGCCGCACCTAACGCGGTTAAGTTCCGCGCCTGGGGGAGTACGGCGCGAGTTGAA3'

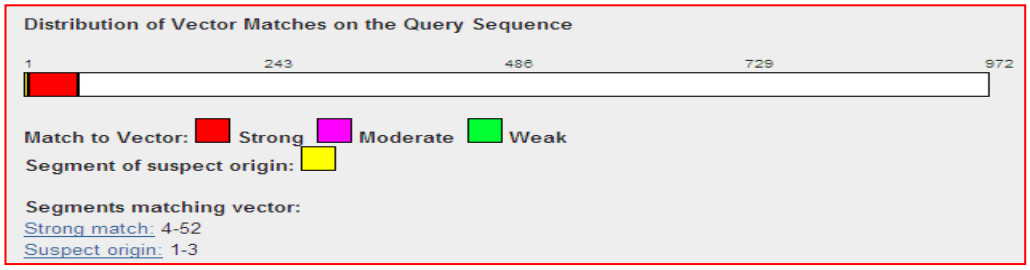


B. Blastn output

Fig. 22. Sequence analysis of clone 21

Table 27. Blastn result of clone 21

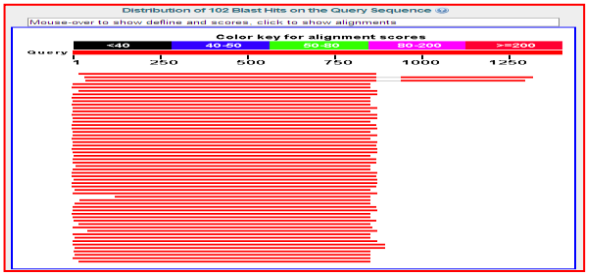
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AB254784.1	Uncultured bacterium	1258	79	0.0	92
OO291986.1	Uncultured bacterium	1194	91	0.0	91
EF197038.1	Uncultured bacterium	1168	79	0.0	91
AF523900.1	Uncultured bacterium	1127	79	0.0	95
AM749736.1	Uncultured Acidobacteria bacterium	1129	79	0.0	92



A. **VecScreen output**

Merged sequence (1394bp)

5'CGGCTACCTTGTTACGACTTCACCCCAATCATCAGCCATACCTTAGGTGCCTGCCCCCCCTTGCGGGTTGGCCCGGCAACTTCTGGTACAACCAAC
TCTCATGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCAACTTCATGAAGTCG
AGTTGCAGACTCCAATCTGAACTGAGACCGGCTTTGTGGGATTGGCTCCACCTTACGGCATTGCGACCCTCTGTACCGGCCATTGTAGTACGTGTGT
AGCCCTGGGCATAAAGGCCATGAGGACTTGACGTCATCCCCACCTTCCTCCACGTTAACCGTGGCAGTATCTCCAGAGTCCCCAACCGAATTGCTGG
TAACTGGAGACAGGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCAGCACCTGTCTTGGGGCTCCC
CGAAAGGCACTTCCTACTTTCATAAGAATTCCCCAGATGTCAAGCCCAGGTAAGGTTCTTCGCGTTGCGTCGAATTAAACCACATACTCCACCGCTT
GTGCGGGCCCCCGTCAATTCCTTTGAGTTTTAATCTTGCGACCGTACTCCCCAGGCGGATCACTTAATGCGTTAACTGCGGCACTGCAGGGGTCAAC
ACCCGCAACACCTAGTGATCACCGTTTACTGTGTGGACTACCAGGTATCTAATCCTGTTTGCTCCCCACACCTTCCCGCTCAGCGTCAGTCTTGTCAGA
AGTCGCCTTCGCCACCGGTGTTCTCTGATATCTACGCATTCACGCTACCCAGGATTCCCTCCTCTACATACTCAGCACAGTTCAATGCCTTCGGGTGG
ACCGGGATTCAACTGACTGACACCGCTACGCCTTAGCCATAATTAGGTTTCCGACAGAGTTTTGTCCGAAGGCTCATCTCACGCGGGTGTGGTCAGG
GTTGCCCCCATGCGCAAAATTCTCATGTGCCTCCGTAGGAGTTGGACCGTGTTCCAGTTCCAGGTGGTGGCCATCCTTCAGACCAGTACCCATGTCG
CCTGGTAGGCCCTTACCTACCAACCAGTAAAGGGACGCGGGTCATTCCAAACAACAGCTTTCATGAAGCGGTATCTTTGATCCATTCAACCGAAACC
AAATGGATGTTATCCGGTATTAGCACCCCTTTGAAGTGTTATCCCAAATTCGGAGGTAGATTACCCACGCATTATCACCCGTGCGCCACTTATCGCC
CAAAGCAAGCTTCAGGCTTTCTCGTTGACTGGCGAAAGAGTTAAGCACGCCGCCAGCGTTCATTTGACCAGGATCTTATTTCTTCTTTAATCAGTGG
AGACGTCCCTCTTTTTTTTTTTTTTTTAGATAGAG3'



B. Blastn output

Fig. 23. Sequence analysis of clone 22

Table 28. Blastn result of clone 22

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AJ582707.1	Uncultured Desulfobacteraceae	1352	60	0.0	95
AY799897.1	Uncultured bacterium	1299	86	0.0	94
AY799901.1	Uncultured bacterium	1277	84	0.0	94
AY889587.1	Uncultured bacterium	1271	60	0.0	93
GU291339.1	Uncultured bacterium	1253	60	0.0	92

bp. The details of results obtained are presented in Fig. 23A. The sequence data obtained with reverse primer (SP6) was 490 bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1394 bp) indicated that 60 percent query coverage and 95 percent homology with Desulfobacteraceae bacterium. Details of Blastn result obtained are presented in Table 28.

CLONE - 23

The sequence data obtained for the clone 23 with forward primer (T7U) was 990 bp in size and vector screening indicated that there was no vector sequence.

Blastn result (990 bp) indicated that 94 percent query coverage and 95 percent homology with Rhodocyclalales bacterium. Details of Blastn result obtained are presented in Table 29.

CLONE-24

The sequence data obtained for the clone 24 with forward primer (T7U) was 950 bp in size and vector screening indicated vector sequence from 1 to 51bp. The with reverse primer (SP6) was 260 bp in size and vector screening indicated that there was no vector sequence.

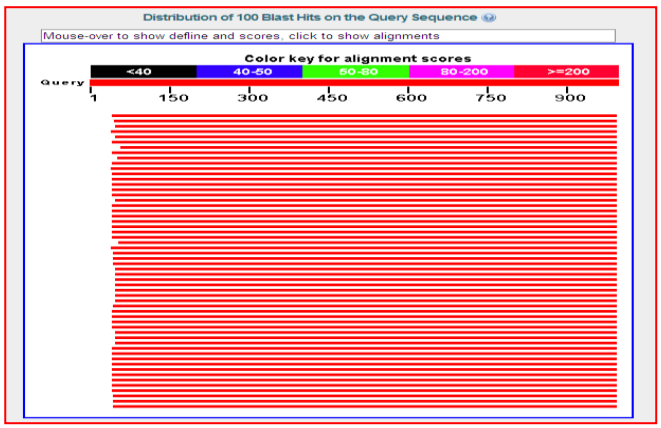
Blastn results of the merged sequence (1159 bp) indicated that 75 percent query coverage and 92 percent homology with *Levilinea saccharolytica*. Details of Blastn result obtained are presented in Table 30.

CLONE- 25

The sequence data obtained for the clone 25 with forward primer (T7U) was 958 bp in size and vector screening indicated vector sequence from 1 to 51bp. The details of results obtained are presented in Fig. 26A. The sequence data obtained with reverse primer (SP6) was 733 bp in size and vector screening indicated that there was no vector sequence.

Merged sequence (990bp)

5'CGATTGGGCGACGTCGCAAAAGGCCCGGCCGTAGGCCGCGGGATTACGGCTACCTTGTTACGACTTCACCCCAGTCATGAATCCCACCGTGGTAA
GCGCCCTCCTTACGGTTAGGCTACCTACTTCTGGTGAAACCCACTCCCATGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGG
CATGCTGATCCGCGATTACTAGCGATTCCGACTTCATGCACTCGAGTTGCAGAGTGCAATCCGGACTACGATCGGCTTTCTGGGATTGGCTCCACCT
CGCGGCTTGGCAACCCTCTGTACCGACCATTGTATTACGTGTGAAGCCCTACCCATAAGGGCCATGAGGACTTGACGTCATCCCCACCTTCCTCCGG
TTTATCACCGGCAGTCCCATTAAAGTGCCCAACTAAATGATGGCAATTAATGGCAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGAC
ACGAGCTGACGACAGCCATGCAGCACCTGTGTCCAGGTTCCCTTTCGGGCACCAACCCATCTCTGGAAAGTTCCTGGCATGTCAAGGGTAGGTAAG
GTTTTTCGCGTTGCATCGAATTAATCCACATAATCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTTGAGTTTTAACCTTGCGGCCGTACTCCCCAGG
CGGTCAACTTCACGCGTTAGCTGCGGTACTAAAGAAGTCTCCTTCCCCAACACCTAGTTGACATCGTTTAGGGCGTGGACTACCAGGGTATCTAATC
CTGTTTGCTCCCCACGCCTTCGTGCATGAGCGTCAGTGTCGACCCAGGGGGCTGCCTTCGCCATCGGTATTCCTCCACATCTCTACGCATTTCACTGC
TACACGTGGATCTACCCCCCTCTGGCGCACTCTAGCTGGCAGTCACAATGCAGTTCCAGGTTAGCCCGGGATTCACACCTGTTCTTACAAACCGCCT
GGCCACGCTTTACGCCAGTAT3'

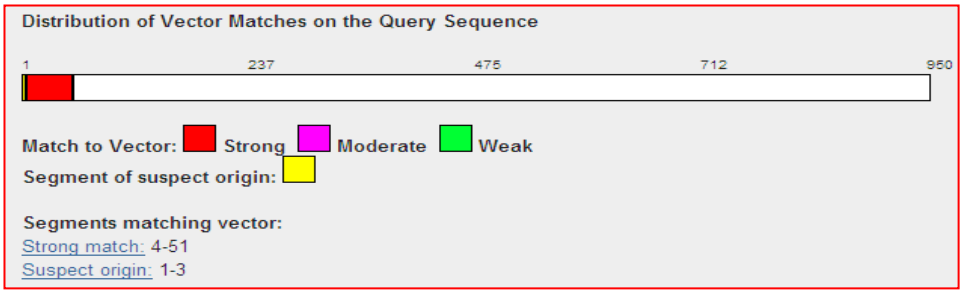


A. Blastn output

Fig. 24. Sequence analysis of clone 23

Table 29. Blastn result of clone 23

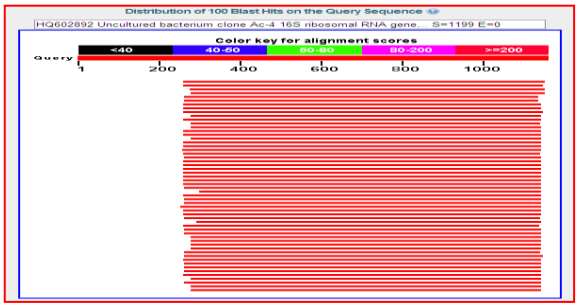
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EU802012.1	Uncultured bacterium	1535	95	0.0	95
AB186833.1	Uncultured bacterium	1515	95	0.0	95
EF562063.1	Uncultured Rhodocyclales bacterium	1502	94	0.0	95
JF502952.1	Uncultured bacterium	1498	95	0.0	95
EF562098.1	Uncultured Rhodocyclales bacterium	1496	94	0.0	95



A. VecScreen output

Merged sequence (1159bp)

5'TCATCCCGTAGCCGTTCCTCCAATAGTGATGGGAGCAGGCCCCTCCGAAGTGAATATGGGATCCTTATTATCTTTCAAAGGAATAAGCACATGGG
GTATAGCAGCCCTTTCGAGCTGTTTCCCCGTGGGGCAGGTCACCAAGCGTTATCACCCGTGCATAAGAACATTGTTTTCGTTGATTGGCGAGTAGAT
ATTAGGCAGCCCAGCGTCATCTGAGCAGGATCAATTTTTTTTTCTAGGAGCCCCTTTTTTTTTCTTTTTGGCTACCTTGTTACGACTTCGTCCCAGTCA
CCGACCCCACCCTCGACGGCTCCCTCCCTTACGGGTTAGGATACCGGCTTCAGGTGTTGCCAGCTCCCATGACGTGACGGGCGGTGTGTACAAGGCC
CGGGAACGTATTCACCGCAGTATGGCTGACCTGCGGTTACTAGCAACTCCGTCTTCATGCAGGCGGGTTGCAGCCTGCAATCTGAACTGAGGACAG
CTTTGAGGATTGGCTCCACCTCGCGGTTTCGCGACCTACTGTACTGCCCATTGTAGCGTGTGTGTAGCCCAGGACATAAAGGCCATGCTGACTTGAC
GTCATCCCCACCTTCCTCCGGCTTAAGACCGGCGGTCCCGCGTGACACATATAACACACGGCGAGGGTTGCGCTCGTTAGCGGACTTAACCGAACAT
CTCACGACACGAGCTGACGACAGCCATGCAGCACCTGTGCAAGCTCCCCGAAGGGTCGGTCTGCTTTCGCTTCCCTACCACTTGCATGTCAAGCCCT
GGTAAGGTTCTTCGTGTAACATCGAATTAAACCACACGCTCCGCTGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTTAACCTTGCGGCCGTAGTC
CCCAGGCGGTAGACTTAACGCGTTAGCTCCGGCACAGATGGATTTGAACCCACCTACGCCAAGTCTACATCGTTTACAGCTAGGACTACCGGGGTCT
CTAATCCCGTTTGCTACCCTAGCTTTCGCGTCTGACGTCAGATATGTGCCAGGAGTCGCTTTCGCACTGGTGTCTCGGATATCTACGCATTTCACACT
ACACCCGGATCACTCCTCTCACACTCTCAGCCCGGTAGTATGACGACTCTCCAGTAGGCAGGAGCTTTCACATCACTTTCGAGCGCTACTG 3'

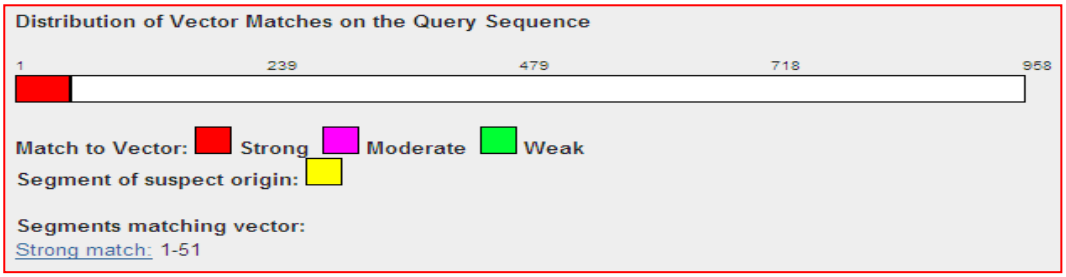


B. Blastn output

Fig. 25. Sequence analysis of clone 24

Table 30. Blastn result of clone 24

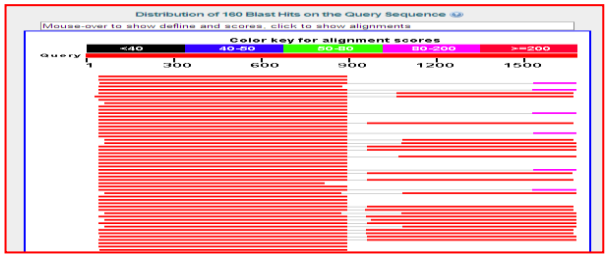
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AJ009468.1	Uncultured bacterium	1288	76	0.0	92
AJ249094.1	Uncultured bacterium	1273	76	0.0	92
AB109439.1	*Levilinea saccharolytica*	1269	75	0.0	92
AB267002.1	Uncultured bacterium	1266	75	0.0	92
HG602819.1	Uncultured bacterium	1264	75	0.0	92



A. VecScreen output

Merged sequence (1689bp)

5'CGAATTGGGCCGACGTCGCATGCTCCCGGCCGCCCTGGCCGCGGGATTAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAATGC
ATGCAAGTCGAACGGAGAAAGTAGCAATATTTTCTCAGTGGCGAACGGGTGAGTAACGCGTTGGTGACCTGCCCCAAAGCGGGGGATAACAGTTC
GAAAGGATTGCTAATACCCCATGTTGTCATCGGGATTAGATGCCTGATGACTAAAGGAGTAATTCGCTTTGGGAGGGGCCTGCGTCCCATCAGCTA
GTTGGTAAGGTAATGGCTTACCAAGGCTTCGACGGGTAGGGGACCTGAGAGGGTGACCCCCCACAATGGAACTGAAACACGGTCCATACACCTACG
GGTGGCAGCAGCAGGGAATATTGGTAATCTGCGAAAGCGGGAACCAGCAACGCCGCGTGTGCGATGAAGGCCTTCGGGTCGTAAAGCACTTTTCG
AGGGAATGAGGAAGGACAGTACCCTCGGAATAAGTCTCGGCTAACTACGTGCCAGCAGCCGCGGTAAAACGTAGGAGGCGAGCGTTATCCGGATT
TACTGGGCGTAAAGCGCGTGCAGGCGGTTCGGTAAGTTGGATGTGAAAGCTCCTGGCTCAACTGGGAGAGGTCGTTCAATACCGTCGGACTTGAGG
ATGGTAGAGGGAGGTGGAATTCCGGGTGTAGTGGTGAAATGCGTAGATATCCGGAGGAACACCAGTGGCGAAAGCGGCCTCCTGGACCATTTCTG
ACGCTCAGACGCGAAAGCTAGGGGAGCAAACGGGATTAGAGACCCGGTAGTCTAGCTGTAACGATGTAGACTTGGCGTCGGTGGTTTAATCATCGT
GCGTAGCTACGCGATAGTCTACGCTGGGACTACGCGCAGTAACTCAAGATGACGGGGCCGCCAGCAGCGGACGGTGTATCGAGCTCCGGAGACTA
CAGTAGGGGGCCAACCGGATTTGAGACCCCGGTAGTCCTAGCTGTAAACGATGTAGACTGGCGTCGGGGGTTAAATCCATCGGTCCGTAGCTAACG
GATAAGTTTACCGCTTGGGGATCCGGCGGCAAGGTTAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCAGCGGAGCGTGTGGTTTAATTGAGGT
ACAGAAGAACCTTCCCAGGGTTTGACATGCAGGTAGTAGGGATCCGAAAGGTGACCGACCCTTCGGGGAGCCTGCACAGGTGTTGCATGGCTGTGT
CAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGTAACGAGCGCAACCCTCGCTGTGTGTTATACGTGTCACACGGGACTGCCGGTATTAAGCCGGAG
GAAGGTGGGGACGAAGTCAAGTCAGCATGACCTATATGTCCTGGGCTACACACACGCTACAATGGCCAGTACAACGGGTAGTAAGCCGCGAGGCG
GAGCCAATCCTTAAAGCTGGTCTCAGTTCAGATTGCAGGCTGCAACCCGCCTGCATGAAGACGGAGTTGCTAGTAACCGCAGGTCAGCTATATGCG
GTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCATGGGAGCTGGCAACACCTGAAGCCGGTATCCTAACCCGCAAGGGAGGGAGCCG
TCGAGGGTGGGGTTGGTGACTGGGACGAAGTTGTAACAAGGTACCGTAATTACTAGGGGGAAA 3'



B. Blastn output

Fig. 26. Sequence analysis of clone 25

Table 31. Blastn result of clone 25

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EF667677.1	Uncultured bacterium	1258	50	0.0	93
EF667725.1	Uncultured bacterium	1253	50	0.0	93
HM214933.1	Uncultured bacterium	1245	59	0.0	93
EU444437.1	Uncultured bacterium	1216	49	0.0	93
AB293389.1	Uncultured Chloroflexi bacterium	1188	59	0.0	91

Blastn results of the merged sequence (1689 bp) indicated that 59 percent query coverage and 91 percent homology with Chloroflexi bacterium. Details of Blastn result obtained are presented in Table 31.

CLONE- 26

The sequence data obtained for the clone 26 with forward primer (T7U) was 873 bp in size and vector screening indicated vector sequence from 1 to 49 bp. The details of results obtained are presented in Fig. 27A.

Blastn results of the merged sequence (824 bp) indicated that 84 percent query coverage and 93 percent homology with *Azospira* sp. Details of Blastn result obtained are presented in Table 32.

CLONE-27

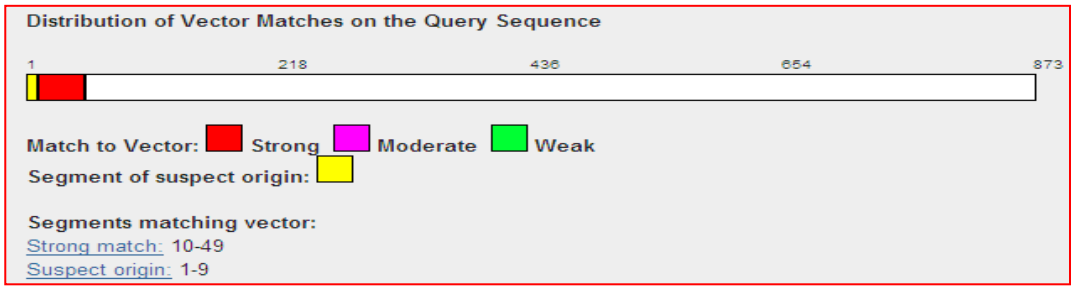
The sequence data obtained for the clone 27 with forward primer (T7U) was 953 bp in size and vector screening indicated vector sequence from 1 to 51bp. The details of results obtained are presented in Fig. 28A.

Blastn results of the merged sequence (902 bp) indicated that 89 percent query coverage and 93 percent homology with Ectothiorhodospiraceae bacterium. Details of Blastn result obtained are presented in Table 33.

CLONE - 28

The sequence data obtained for the clone 28 with forward primer (T7U) was 242 bp in size and vector screening indicated that there was no vector sequence. The sequence data obtained with reverse primer (SP6) was 185 bp in size and vector screening indicated that there was no vector sequence found.

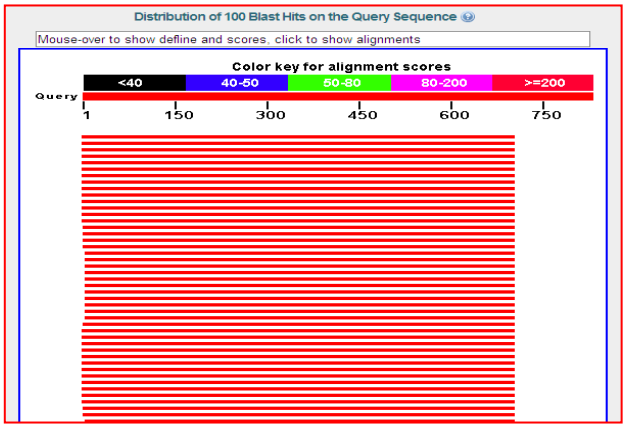
Blastn (424 bp) indicated that 25 percent query coverage and 88 percent homology with Chloroflexi bacterium. Details of Blastn result obtained are presented in Table 34.



A. VecScreen output

Merged sequence (824 bp)

5'ACGGCTACCTTGTTACGACTTCACCCCAGTCATGAATCCCACCGTGGTAAGCGCCCTCCTTACGGTTAGGCTACCTACTTCTGGTGAAACCCACTC
CCATGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCGACTTCATGCACTCGA
GTTGCAGAGTGCAATCCGGACTACGATCGGCTTTCTGGGATTGGCTCCACCTCGCGGCTTGGCAACCCTCTGTACCGACCATTGTATTACGTGTGAA
GCCCTACCCATAAGGGCCATGAGGACTTGACGTCATCCCCACCTTCCTCCGGTTTATCACCGGCAGTCCCATTAAAGTGCCCAACTAAATGATGGCA
ATTAATGGCAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCAGCACCTGTGTCCAGGTTCCCTTT
CGGGCACCAACCCATCTCTGGAAAGTTCCTGGCATGTCAAGGGTAGTAAGTTTTTCGCGTTGCATCGAATTAATCCACATAATCCACCGCTTGTGCG
GTCCCGTCAATCTTTGAGTTTACCTTGCGCCGTACTCCCCAGCGTCACTCCGCGTAGCTGCGGTACTAAAAAGTCCCTCCCCACCCCTATGACTCGTT
AGGCGTGACTACAGTTTATCTGTTGCTCCCAGCTTCTGCTGACGCATTCACCAGGGCTGCTCCCACGATCCCCACTCCACCATACGTCACGGATTCAC
CCCTTCCCCACCTACCTGCCAAGTTCGTACCGGATCCTCTAACCGCACT3'

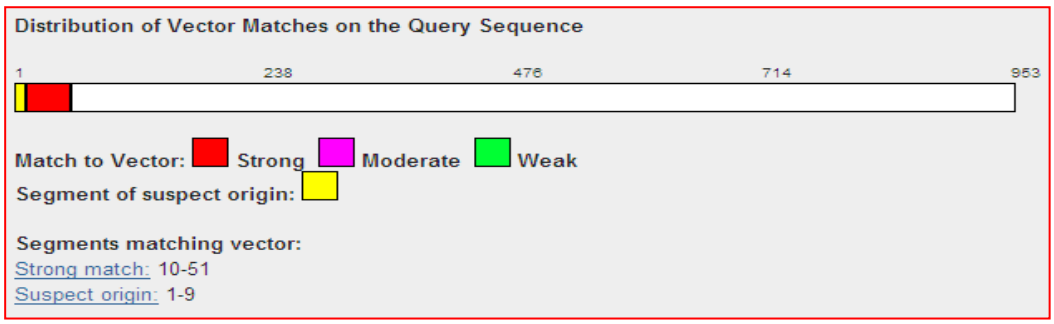


B. Blastn output

Fig. 27. Sequence analysis of clone 26

Table 32. Blastn result of clone 26

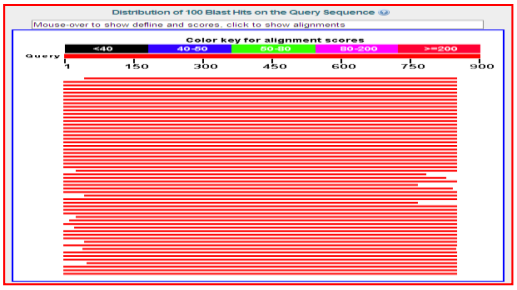
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EU002012.1	Uncultured bacterium	1044	84	0.0	93
GU724735.1	Uncultured bacterium	1038	84	0.0	93
GU724723.1	Uncultured bacterium	1038	84	0.0	93
GU294119.1	*Azospira* sp.	1030	84	0.0	93
FN436195.1	Uncultured bacterium	1036	84	0.0	91



A. VecScreen output

Merged sequence (424bp)

5'AGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCATGCTTAACACATGCAAGTCGAACGGTAACAGGGGCTTCGGCCCGCTGACGAGCGGCGGAC
GGGTGAGTAATACATAGGAATCTGCCCTGTAGTGGGGGACAACTTAGGGAAACTTAAGCTAATACCACATACGACCTACGGGTGAAAGACTTCGGT
CGCTATAGGATGAGCCTATGTCTGATTAGCTAGTTGGTAGGGTAAAGGCCTACCAAGGCAACGATCAGTAGCTGGTCTGAGAGGACGATCAGCCAC
ACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCAATGCCGCGTGTGT
GAAGAAGGCCTGCGGGTTGTAAGCACTTTCGGTAGGGAAGATAATGACGGTACCTACAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCG
GTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGTTTGTTAAGTCGGGTGTGAAATCCCTGGGCTCAACCT
AGGAATTGCATTCGATACTGGCAGGCTAGAGTGTGGGTAGAGGGAAGTGGAATTCCAGGTGTAGCGGTGAAATGCGTAATATCTGGAGGAACATC
AGTGGCGAAGGCGACTTCCTGGACAACACTGACGCTGAGCGCGAAGCGTGGGGAGCAACAGGATAAATACCCTGGTAGTCCACGCCGTAACGATG
AAACTACCGTTGGGATTTAAATCTTATGGCGCACTACGCGAAAGTCTCCGCTGGGGAGTCGGCCGCAGTTAAACTCAAGATGACGGGGCCCCAACG
TGGAAATGTGTTATCATGCACCAAACCTCTATCTGATCTCATGT**3'**



B. Blastn output

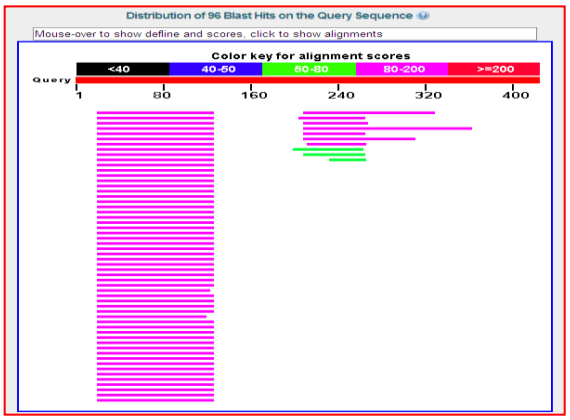
Fig. 28. Sequence analysis of clone 27

Table 33. Blastn result of clone 27

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
FN687123.1	Uncultured Ectothiorhodospiraceae bacterium	1195	89	0.0	93
EF999373.1	Uncultured bacterium	1181	94	0.0	91
FR851506.1	Uncultured bacterium	1147	94	0.0	90
FR851685.1	Uncultured bacterium	1147	94	0.0	90
IF044400.1	Uncultured gamma proteobacterium	1136	94	0.0	90

Merged sequence (424bp)

5'GCGGGAAGCGTTCCGGGCTTCACCGACGTCAGGTCATGGGAGTGGTAACACTGAAGCCGGTGGGTTAACCGAAAGGGGGCAGGGGGTAGAAGGT
GGGACTGGTGACTGGGACGAAGTGTAACAAGGTTTTTTTTAATCATTAGAGAACTCCCTTTCTTTTTTTTTTTTTTTTGTTGAAAGTCCCCGGGCCGTC
TCCTGGCGCGGGATTAAGTTTGATCCTGGCTCAGATGACGCTGGGGCGTGCCTAACACATGCAGTCGAACGGGATCCCTTGGATTCTCTGAGGGTGA
GAATGGCGAACGGGTGATAATACGTAGGTGACCTGCCCCAAAGTAGGAATAACCCTGGAAACGTGGCTAATACTCTATGTGCTCGTGGTTCGGTCT
GCCAGTAAAGACCTCTTTGAGGTGGCCTTTGGGAGGGG3'



A. Blastn output

Fig. 29. Sequence analysis of clone 28

Table 34. Blastn result of clone 28

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EU488019.1	Uncultured bacterium	139	25	1e-29	90
GU180179.1	Uncultured bacterium	134	25	6e-28	89
EF125411.1	Uncultured bacterium	134	25	6e-28	89
EF018058.1	Uncultured bacterium	130	25	8e-27	88
GO249570.1	Uncultured Chloroflexi bacterium	128	25	3e-26	88

CLONE - 29

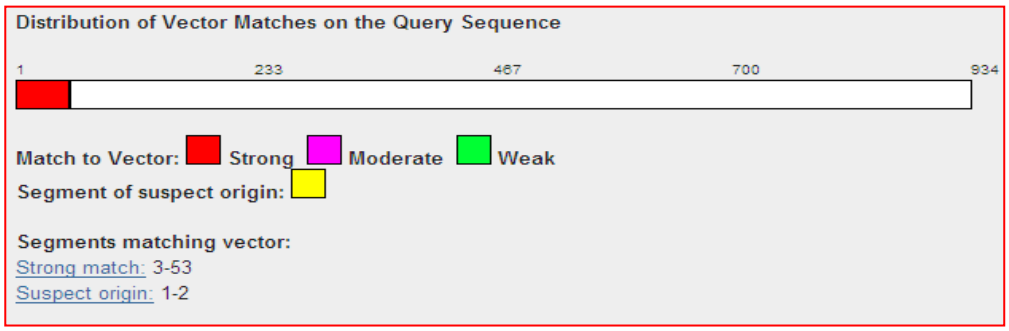
The sequence data obtained for the clone 29 with forward primer (T7U) was 934 bp in size and vector screening indicated vector sequence from 1 to 53bp. The details of results obtained are presented in Fig. 30A. The sequence data obtained with reverse primer (SP6) was 471bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1333 bp) indicated that 89 percent query coverage and 94 percent homology with Acidobacteria bacterium. Details of Blastn result obtained are presented in Table 35.

CLONE-30

The sequence data obtained for the clone 30 with forward primer (T7U) was 917 bp in size and vector screening indicated vector sequence from 1 to 59 bp. The details of results obtained are presented in Fig. 31A. The sequence data obtained with reverse primer (SP6) was 291 bp in size and vector screening indicated that there was no vector sequence.

Blastn results of the merged sequence (1131 bp) indicated that 88 percent query coverage and 91 percent homology with Acidobacteria bacterium sp. Details of Blastn result obtained are presented in Table 36.



A. VecScreen output

Merged sequence (1333bp)

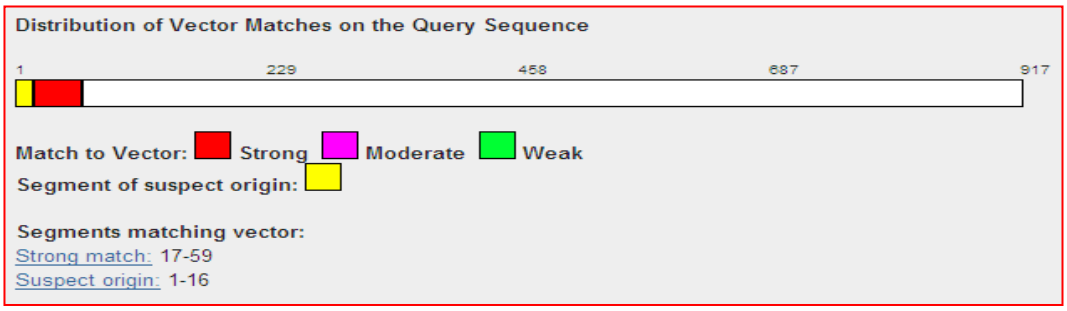
5'CGAGCGCAACCCCTGTCCTAGTTCCAGCGGTTCGGCCGGGCATTAGGGGGATGTCCGGATAACGGAGAGGAAGGTGGGGATGAGGTCAAGTCAG
CATGGCCTTTATGTCCGGGGTACACACGTGTTACAATGCGCACCACAAAGCGTTGCCAACCCGCGAGGGGGAGCCAATCGCATAAAAGCGCGCTCA
GTTCGGATTGCAGTCTGCAACTCGACTGCATGAAGCCGGAATCGCTAGTAATGGCGTATCAGAACGACGCCGTGAATACGTTCCCGGGCCTTGTAC
ACACCGCCCGTCACATCACGAAAGGGGGCTGTACTAGAAGTCGCCAGGCCAACCCGCAAGGGAGGCAGGCGCCCAAGGTGTGGTCCCTGATTGGG
GGAAGTCGTAACAAGGTATCCGTAATCACTAGAGCGGCCGCCTGCAGTTTTTTCTATATGGACTCCCAACGCTTGGAGCATGGCTCAGAGCGAACG
CTGGCGGCGTGCCTAACACATGCAAGTCGAACGCGCGGCCGTAGCAATACGGCTGGCGCGTGGCGAACGGGTGAGTAACACGTGGGCAACCTTCC
CCCGGAAGTGGAATAACCCGCCGAAAGGCGGGCTAATACCACATAACATCCCGATGGCACAAGTCGTTGGGATCAAAGCCGCAAGGCGTCTGGGG
AGGGGCCCGCGGCCGATCAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCTGCGACCGGTAGCTGGCCTGAGAGGGTGGCCGGCCACACTGGAAC
TGCGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGCAATGGGCGAAAGCCTGACGCAGCGACGCCGCGTGGGGGATGAAGC
CTTTCGGGGTGTAAACCCCTTTTAGCCCGGACGAATCCCTGGCAACAGGGTCGACGGTACGGGCAGAAAAAGCCCCGGCTAACTACGTGCCAGCAG
CCGCGGTAAAACGTAGGGGGCCAGCGTTGCTCGGAATGACTGGGCGTAAAGGGTCTGTAGGCGGTGCGGCAAGTCGGAAGTGAAATCCTGCGGCT
TAACTGCAGGGCTGCTTCCGAAACTGCCGCGCTAGAGTGCGGGAGAGGGGAGTGGAATTGCCGGTGTAGCGGTGAAATGCGTAGATATCGGCAGG
AACACCGGAGCGAAGCGCTCCCTGGACCGCAACTGACGCTGAGAGACGAAAGCTGGGGGAGCAACAGGATTAGATACCCTGGTAGTCCCAGCCCT
AAACATCGGACTAGCGTGGGCCCGTGGGGTCGTGCCGCAGCTACGCGTTAGTTCGCGCCTGGAGTACGGCGGAGTTGAAACTCAAGGAATGACGG
3'

B. Blastn output

Fig. 30. Sequence analysis of clone 29

Table 35. Blastn result of clone 29

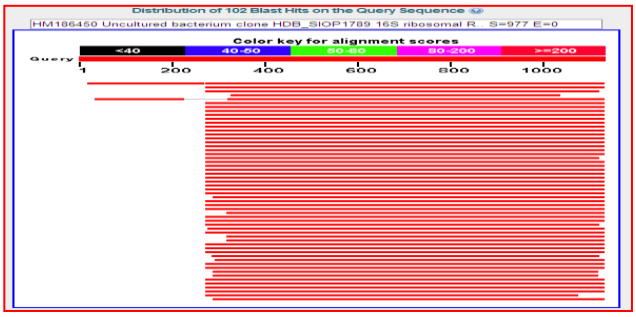
Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AB254764.1	Uncultured bacterium	1234	95	0.0	94
DQ791336.1	Uncultured bacterium	1177	85	0.0	90
EF197038.1	Uncultured bacterium	1149	68	0.0	93
AF523900.1	Uncultured bacterium	1114	62	0.0	95
AM749736.1	Uncultured Acidobacteria bacterium	1103	89	0.0	94



A. VecScreen output

Merged sequence (1131bp)

5'TATTCGGGGGTCCATACTTTGCCACAAAGGTGCCAACCGCGAGGGGGAGCCAATTGCATAAAAGCGCGTCAGTTCGGATTGCATTGCAACTGATG
CATGAAGCCGGAATCGTAGTAATGGCGTATCAGAAGACGCCGGAATACGTTCCCGGGCTTTACACACCGCCCGTCACATCACGAAAGGGGGCTGTA
TAGAAGTTCCAGGCCAACCCGCAAGGGAGGCAGGCGCAAAACAAGGTGTGGTCCTGATTGGGGGAAGTCGTAACAAGGTATCCTTTAATCATGGCT
CAGAGCGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAACGCGCGGCCGTAGCAATACGGCTGGCGCGTGGCGAACGGGTGAGTAACACGTG
GGCAACCTTCCCCCGGAAGTGGAATAACCCGCCGAAAGGCGGGCTAATACCACATAACATCCCGATGGCACAAGTCGTTGGGATCAAAGCCGCAA
GGCGTCTGGGGAGGGGCCCGCGGCCGATCAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCTGCGACCGGTAGCTGGCCTGAGAGGGTGGCCGGC
CACACTGGAACTGCGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGCAATGGGCGAAAGCCTGACGCAGCGACGCCGCGTG
GGGGATGAAGCCTTTCGGGGTGTAAACCCCTTTTAGCCCGGACGAATCCCTGGCAACAGGGTCGACGGTACGGGCAGAAAAAGCCCCGGCTAACTA
CGTGCCAGCAGCCGCGGTAAAACGTAGGGGGCCAGCGTTGCTCGGAATGACTGGGCGTAAAGGGTCTGTAGGCGGTGCGGCAAGTCGGAAGTGAA
ATCCTGCGGCTTAACTGCAGGGCTGCTTCCGAAACTGCCGCGCTAGAGTGCGGGAGAGGGGAGTGGAATTGCCGGTGTAGCGGTGAAATGCGTAGA
TATCGGCAGGAACACCGGAGGCGAAGCGGCTCCCTGGACCGCAACTGACGCTGAGAGACGAAAGCTGGGGGAGCAAACAGGATTAGATACCCTGG
TAGTCCCAGCCCTAAACCATCGGACTAGGCGTGGGCCCCGGTTGGGGGTCCGTGCCGCAGCTAACGCGTTAGTTCCGCGCCT3'



B. Blastn output

Fig. 31. Sequence analysis of clone 30

Table 36. Blastn result of clone 30

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AB254784.1	Uncultured bacterium	1267	97	0.0	93
DQ791385.1	Uncultured bacterium	1199	75	0.0	91
EF197088.1	Uncultured bacterium	1188	74	0.0	92
AF523900.1	Uncultured bacterium	1133	62	0.0	95
AN749736.1	Uncultured Acidobacteria bacterium	1366	88	0.0	91

4.6 Phylogenetic analysis

Phylogenetic analysis is the study of hereditary molecular differences, on the basis of 16S rDNA, to gain information of an organism's evolutionary relationships. The result of a molecular phylogenetic analysis is expressed in a phylogenetic tree. Molecular phylogenetic is one aspect of molecular systematic, a broader term that also includes the use of molecular data in taxonomy and biogeography.

Phylogenetic tree was constructed using the neighbour-joining method using Mega software. Similarity for most of the sequences with those of known bacteria was too low to identify the sequence as representing a particular species. The phylogenetic tree based on partial 16S rRNA gene placed 30 clones from *Pokkali* soil samples into 4 groups– Proteobacteria, Acidobacteria, Chloroflexi, Cyanobateria (Fig. 32).

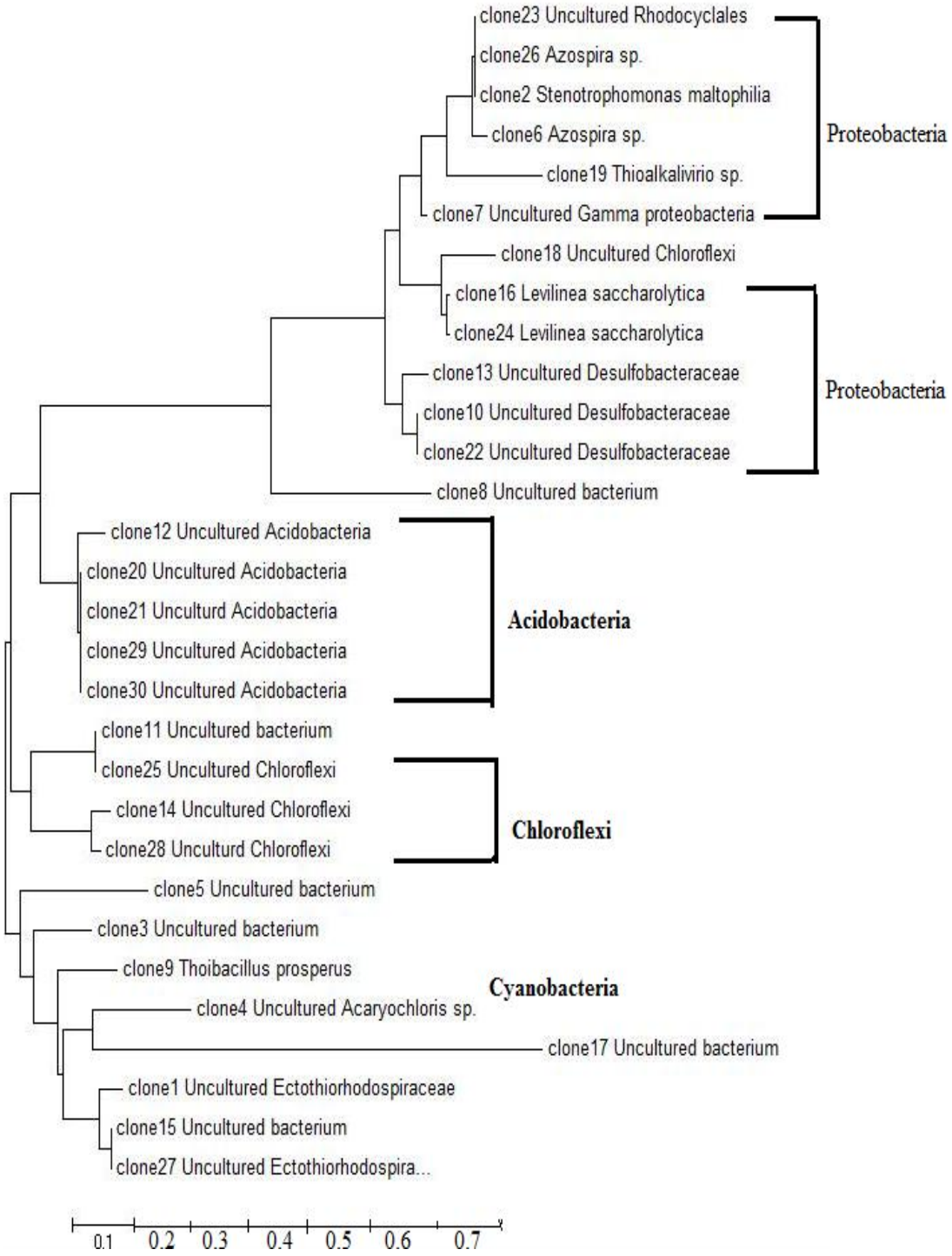


Fig. 32. Phylogenetic relationships on basis of partial 16S rDNA gene sequences of clones recovered from *Pokkali* soil sample.

Table 37. Distribution of *Pokkali* microbiome

Organism	Phylum	No. of samples
Beta-Proteobacteria	Proteobacteria	3
Gamma-Proteobacteria	Proteobacteria	6
Delta-Proteobacteria	Proteobacteria	3
Chloroflexi	Chloroflexi	6
Acidobacteria	Acidobacteria	5
Cyanobacteria	Cynobacteria	1
Unidentified bacteria		6

4.6.1 Diagrammatic representation of *Pokkali* microbiome

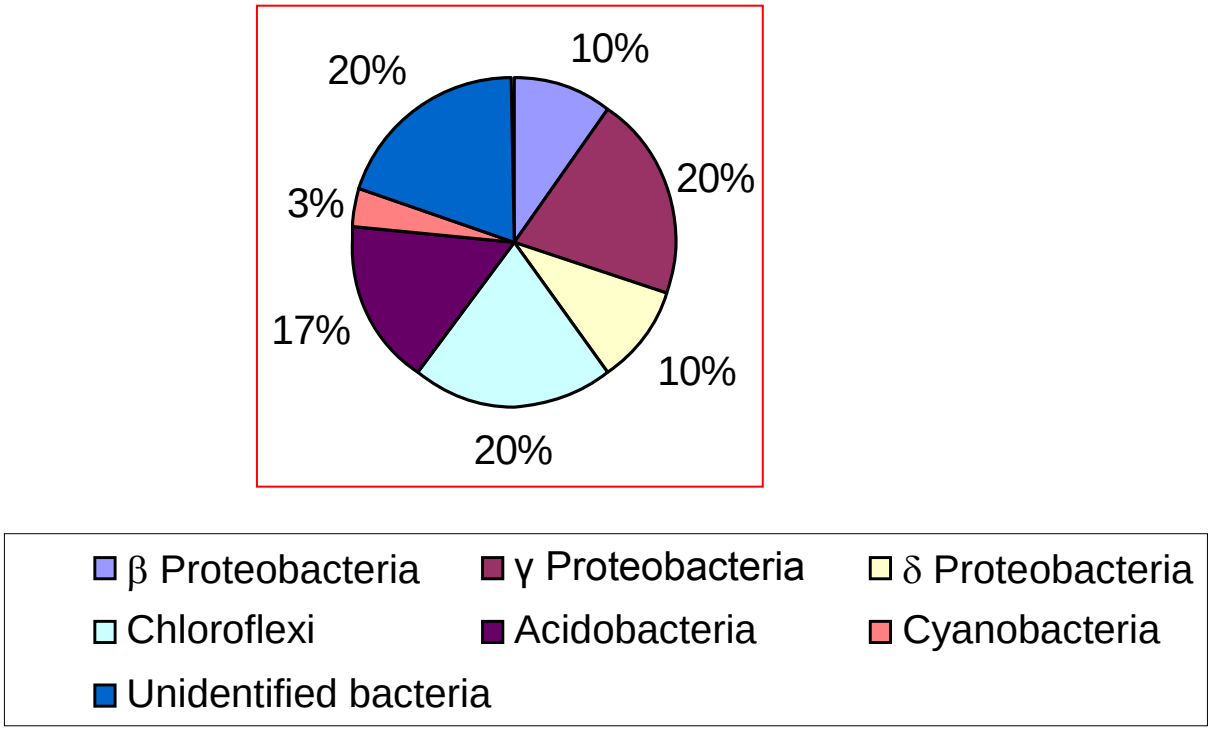


Fig. 33. Per cent distribution of *Pokkali* microbiome

4.6.1.1 Proteobacteria

Twelve clones out of 30 were grouped under Proteobacteria. Proteobacteria belonged to β-proteobacteria (10 percent), γ-proteobacteria (20 per cent)and δ-proteobacteria (10 percent).

4.6.1.2 Chloroflexi

20 per cent of total cloned sequences were found as Chloroflexi sequence types.

4.6.1.3 Acidobacteria

17 per cent of total cloned sequences were found as Acidobacteria sequence types.

4.6.1.4 Cyanobacteia

3 per cent of total cloned sequences were found as Cyanobacteria sequence types.

4.6.1.5 Uncultured bacterium

20 per cent of total cloned sequences were found as unknown uncultured bacteria sequence types.

4.7 CHARACTERIZATION OF CULTURABLE BACTERIA

The air dried composite soil sample from *Pokkali* was used for isolation and characterization of culturable bacteria.

4.7.1 Enumeration of bacteria by serial dilution plate method

The total counts of bacteria obtained on soil extract agar containing different concentrations of NaCl are presented in Table 38.

Table 38. Number of bacteria by serial dilution plate count in per gm *Pokkali* soil sample

Sl. no.	Salt concentration *per cent*	Bacteria cfu /g soil
1.	0	45×10^4
2.	5	22×10^4
3.	10	12×10^4
4.	15	2×10^3
5.	20	0

4.7.2 Colony characteristics

Of the several colonies obtained on soil extract agar, three morphotypes were found to be predominant. These were named as BS1, BS2 and BS3 and were characterized by colony characters (Table 39) and 16SrDNA typing.

Table 39. Colony characters of bacterial isolates from *Pokkali* soil

Isolate	Pigmentation	Size	Form	Margin
BS1	Yellow	Small	Circular	Convex
BS2	Slightly yellow	Large	Circular	Entire
BS3	Creamy white	Large	Irregular	Undulate

4.7.3 Culturable bacterial DNA isolation

The total DNA was isolated from pure cultures of *Pokkali* isolates followed by the protocol of Sambrook and Russel (2001). DNA samples were run on 0.8 *per cent* agarose gel to check the purity of DNA and contamination by protein and RNA if any (Plate. 4A). No RNA contamination was observed and the DNA was detected as single intact band of high intensity.

4.7.4 16S rDNA gene amplification

On PCR with specific primers, single band of expected size (1.5kb) was obtained (Plate 4B).

4.7.5 Sequence analysis of 16S rDNA from culturable bacteria

Sequence data obtained from 16S rDNA amplicons from the three selected bacterial isolates were subjected to Blastn analysis.

BS-1

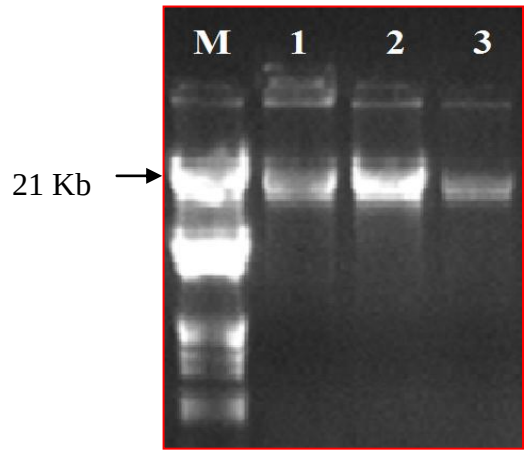
The sequence data obtained for sample BS-1 with the forward primer (T7U) was 199 bp in size and with reverse primer (SP6) was 346 bp in size.

Blastn results of the merged sequence (545 bp) indicated 83 percent query coverage with 98 percent homology with *Bacillus aryabhattai*. Details of Blastn result obtained are presented in Table 40.

BS-2

The sequence data obtained for sample BS-2 with the forward primer (T7U) was 538 bp in size and with reverse primer (SP6) was 702 bp in size.

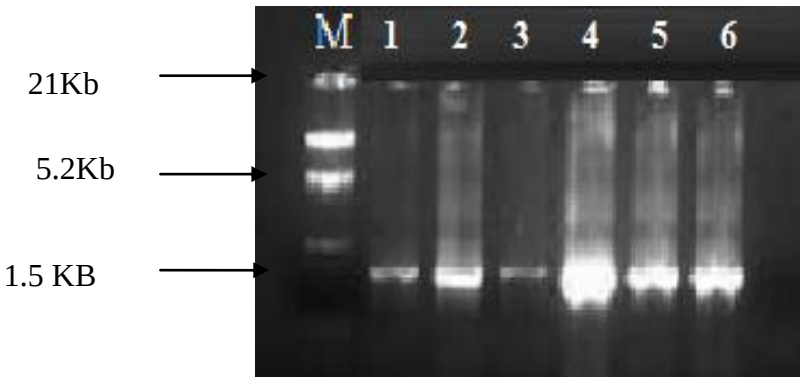
Blastn results of the merged sequence (1227 bp) indicated 99 percent query coverage with 99 percent homology with *Bacillus megaterium*. Details of Blastn result obtained are presented in Table 41.



M- Marker (λ DNA/*Eco*RI+*Hind*III)

Lane 1-3 Bacterial isolates- 1:BS1, 2: BS2, 3:BS3

A. Total DNA from culturable bacteria



M- Marker (λ DNA/*Eco*RI+*Hind*III)

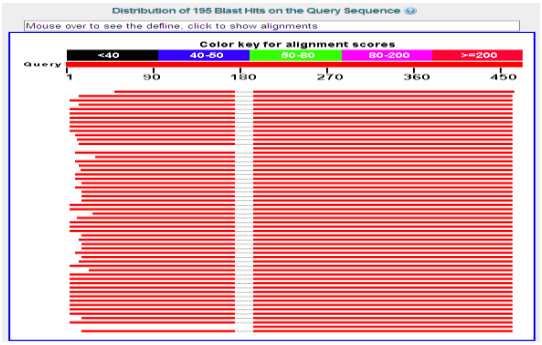
Lane 1-6 Bacterial isolates- 1-2: BS1, 3-4: BS2, 5-6: BS3

B. 16S rDNA amplicons

Plate 4. 16S rDNA amplification from culturable bacteria

Merged sequence (545 bp)

5'GCCCATAATCCATGGCAAGTCCGAGCGAACTGATTAGAAGCTTGCTTCTATGACGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCT
GTAAGACTGGGATAACTTCGGGAAACCGAAGCTAATACCGGATAGGATCTTCTCCTTCATGGAAGATGATTGAAATATGGTGCCAGGTATTTCTTTT
TTTTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGATACACACGTGCTACAATGGATGGTACAAAGGGCTGCAAGACCGCGAGGTCAA
GCCAATCCCATAAAACCATTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGT
GAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGAAGTTAACCCCGGTAAAT **3'**



A. Blastn output

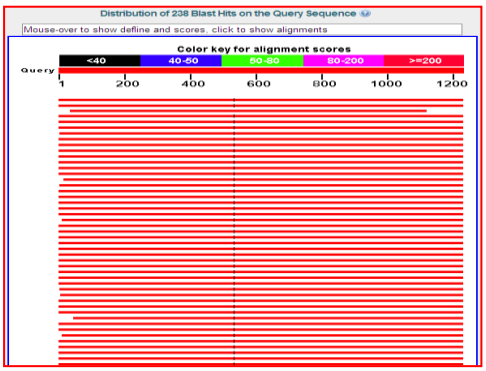
Fig. 34. Sequence analysis of sample BS1

Table 40. Blastn result of sample BS1

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
AB618947.1	*Bacillus* sp.	481	83	2e-132	98
JN135237.1	*Bacillus aryabhattai*	479	91	8e-132	98
JN178876.1	Uncultured bacterium	479	93	8e-132	98
HQ857752.1	*Bacillus aryabhattai*	479	91	8e-132	98
JF513170.1	*Bacillus subtilis*	479	93	8e-132	98

Merged sequence (1227bp)

5'CCTGCAAGTCGAGCGAACTGATTAGAAGCTTGCTTCTATGACGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCTGTAAGACTGGGA
TAACTTCGGGAAACCGAAGCTAATACCGGATAGGATCTTCTCCTTCATGGGAGATGATTGAAAGATGGTTTCGGCTATCACTTACAGATGGGCCCGC
GGTGCATTAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCAACGATGCATAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGG
CCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGCTTTCGGGTC
GTAAAACTCTGTTGTTAGGGAAGAACAAGTACAAGAGTAACTGCTTGTACCTTGACGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCA
GCCGCGGTAATACGTATGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGCCACGCCGTA
AACGATGAGTGCTAAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTCGCAAGACTGAA
ACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCTCTGA
CAACTCTAGAGATAGAGCGTTCCCCTTCGGGGGACAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCC
GCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCATTTAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTC
AAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGATGGTACAAAGGGCTGCAAGACCGCGAGGTCAAGCCAATCCCATAAAAC
CATTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGC
CTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGAGTAACCGTAAGGAGCAGCCGC3'



B. Blastn output

Fig. 35. Sequence analysis of sample BS2

Table 41. Blastn result of sample BS2

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
GU257962.1	*Bacillus* sp.	1284	99	0.0	99
GO169375.1	*Bacillus megaterium*	1284	99	0.0	99
AF548883.1	*Bacillus* sp.	1284	88	0.0	88
EU571105.1	*Bacillus megaterium*	2227	99	0.0	99
HQ857752.1	*Bacillus aryabhattai*	2244	99	0.0	99

BS-3

The sequence data obtained for sample BS-3 with the forward primer (T7U) was 438 bp in size and with reverse primer (SP6) was 689 bp in size.

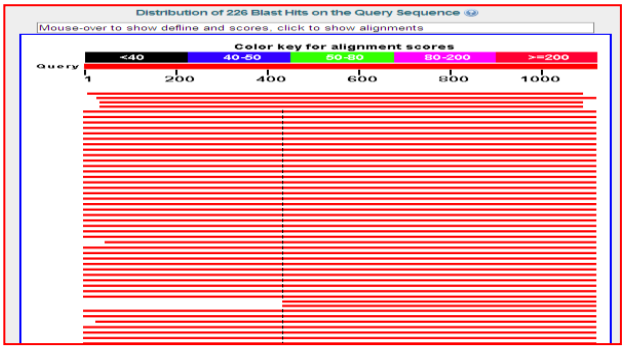
Blastn results of the merged sequence (1122 bp) indicated 99 percent query coverage with 99 percent homology with *Bacillus* sp. Details of Blastn result obtained are presented in Table 42.

The main objective of the study was to analyze the bacterial community, both culturable and non-culturable, in the saline *Pokkali* soils of Kerala through 16S rDNA sequencing approach. By using Metagenomics approach 76.7 per cent of unculturable bacteria and remaining 23.3 per cent culturable bacteria were detected. The most diverse group of bacteria Proteobacteria, Chloroflexi, Acidobacteria, Cyanobacteria and unidentified bacteria were detected.

By using conventional method only genus *Bacillus* bacterial species were detected. Metagenomic approach successfully revealed the composition and diversity of bacterial community in *Pokkali* soil.

Merged sequence (1122bp)

5'ATGGATTGAGAGCTTGCTCTTATGAAGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCCATAAGACTGGGATAACTCCGGGAAACC
GGGGCTAATACCGGATAACATTTTGAACTGCATGGTTCGAAATTGAAAGGCGGCTTCGGCTGTCACTTATGGATGGACCCGCGTCGCATTAGCTAGT
TGGTGAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGG
GAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGCTTTCGGGTCGTAAAACTCTGTTGTT
AGGGAAGAACAAGTGCTATTTGAGTAAGCTGGCACCTTTGACGGGACCTAACCAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA
CGCCGTAAAACGATGAGTGCTAAGTGTTAGAGGGTTCCCGCCCTTTAGTGCTGAAGTTAACGCATAAGCACTCCGCCTGGGGAGTACGGCCGCAAG
GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTTGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATC
CTCTGAAAACCCTAGAGATAGGGCTTCTCCTTCGGGAGCAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGT
CCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCATCATTAAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGAC
GTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACGGTACAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAA
AACCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCG
GGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGGGTAACCTTT3'



A. Blastn output

Fig. 36. Sequence analysis of sample BS3

Table 42. Blastn result of sample BS3

Accession No.	Descriptions	Max. Score	Query Coverage (%)	E-Value	Max. Identity (%)
EF105468.1	*Bacillus* sp.	1284	99	0.0	99
EF105471.1	*Bacillus* sp.	1284	99	0.0	99
EF105461.1	*Geobacillus* sp.	1284	88	0.0	88
EF1054601.1	*Geobacillus* sp.	2227	99	0.0	99
FJ601638.1	*Bacillus* sp.	2244	99	0.0	99

Table 43. Sequence analysis of 30 clones

Clone no.	Total sequence length		Vector sequence		Sequence after screening		Sequence length (bp)	Query coverage (%)	Maximum identity (%)	Name of Bacteria with maximum identity
	Forward (T7U)	Reverse (SP6)	Forward (T7U)	Reverse (SP6)	Forward (T7U)	Reverse (SP6)				
1.	707	548	1-62	-	645	548	1182	92	93	Uncultured Ectothiorhodospiraceae
2.	695	736	1-64	-	631	736	1362	85	84	*Stenotrophomonas maltophilia*
3.	712	768	1-56	1-60	656	708	1299	92	98	Uncultured bacterium
4.	694	734	-	1-55	694	679	1361	89	79	Uncultured *Achryochloris* sp.
5.	741	693	1-69	1-39	672	654	1310	89	93	Uncultured bacterium
6.	694	608	-	-	694	608	1357	92	95	*Azospira* sp.
7.	774	441	1-54	-	720	441	1158	92	95	Uncultured γ-proteobacterium
8.	348	-	-	-	348	-	348	75	75	Uncultured bacterium
9.	780	-	1-58	-	722	-	722	96	95	*Thiobacillus prosperus*
10.	769	490	1-30	-	739	490	1226	97	94	Uncultured Desulfobacteraceae
11.	622	556	1-50	1-45	572	511	1073	89	93	Uncultured bacterium

12.	758	584	1-53	1-70	705	514	1205	90	94	Uncultured Acidobacteria bacterium
13.	675	752	1-89	-	586	752	1299	97	95	Uncultured Desulfobacteraceae
14.	736	669	1-61	1-74	675	595	1269	98	98	Uncultured Chloroflexi bacterium
15.	626	641	1-36	1-43	590	592	1181	59	97	Uncultured bacterium
16.	746	707	1-57	1-51	689	656	1336	97	96	*Levilinea saccharolytica*
17.	664	697	1-43	-	621	697	1309	95	93	Uncultured bacterium
18.	666	334	1-83	-	583	334	911	61	95	Uncultured Chloroflexi bacterium
19.	666	91	-	68-91	666	67	729	79	86	*Thioalkalivirio* sp.
20.	674	711	1-51	-	625	711	1320	79	85	Uncultured Acidobacteria bacterium
21.	931	450	1-52	-	879	471	1328	79	92	Uncultured Acidobacteria bacterium
22.	972	490	1-52	-	920	490	1329	60	95	Uncultured Desulfobacteraceae bacterium
23.	990	-	-	-	990	-	990	94	95	Uncultured Rhodocydales bacterium
24.	950	260	1-51	-	661	266	1159	75	92	*Levilinea saccharolytica*

25.	958	733	1-51	-	907	733	1689	59	91	Uncultured Chloroflexi bacterium
26.	873	-	1-49	-	824	-	824	84	93	*Azospira* sp.
27.	953	-	1-51	-	902	-	902	89	93	Uncultured Ectothiorhodospiraceae
28.	242	185	-	-	242	185	424	25	88	Uncultured Chloroflexi bacterium
29.	934	471	1-53	-	881	471	1333	89	94	Uncultured Acidobacteria bacterium
30.	917	291	1-59	-	858	291	1131	88	91	Uncultured Acidobacteria bacterium

Table 44. Taxonomic distribution of 30 sequenced 16S rRNA gene clones

Cl. No.	**Organism**	**Phylum**	**Class**	**Order**	**Family**	**Genus**	**Species**	**Unique feature**
1.	Uncultured Ectothiorhodospiraceae bacterium	Proteobacteria	γ-Proteobacteria	Chromatiales	Ectothiorhodospiraceae	-	-	Purple sulphur bacteria, deposit S globules outside the cell
2.	*Stenotrophomonas maltophilia*	Proteobacteria	γ-Proteobacteria	Xanthomonadales	Xanthomonadaceae	*Stenotrophomonas*	*maltophilia*	Human pathogen, causes respiratory infections, found in aqueous environment
3.	Uncultured bacterium	-	-	-	-	-	-	-
4.	Uncultured *Acaryochloris* sp.	Cyanobacteria	-	Chroococcales	-	*Acaryochloris*	-	Cyanobacteria, chlorophyll d, utilizes far red light
5.	Uncultured bacterium	-	-	-	-	-	-	-
6.	*Azospira* sp.	Proteobacteria	β-Proteobacteria	Rhodocyclales	Rhodocyclaceae	*Azospira*	-	N fixing β-proteobacteria, found in deep waters
7.	Uncultured γ-proteobacterium	Proteobacteria	γ-Proteobacteria	-	-	-	-	-
8.	Uncultured bacterium	-	-	-	-	-	-	-

9.	*Thiobacillus prosperus*	Proteobacteria	γ-Proteobacteria	-	-	*Thiobacillus*	*prosperus*	Obligate chemolithoautotroph, Fe oxidising, acid and salt tolerant
10.	Uncultured Desulfobacteraceae bacterium	Proteobacteria	δ-Proteobacteria	Desulfobacterales	Desulfobacteraceae	-	-	Reduce sulphate to sulphide, strictly anaerobic, Deltaproteobactera
11.	Uncultured bacterium	-	-	-	-	-	-	-
12.	Uncultured Acidobacteria	Acidobacteria	Acidobacteria	-	-	-	-	Acidophilic, mostly unculturable
13.	Uncultured Desulfobacteraceae bacterium	Proteobacteria	δ-Proteobacteria	Desulfobacterales	Desulfobacteraceae	-	-	Reduce sulphate to sulphide, strictly anaerobic, Deltaproteobactera
14.	Uncultured Chloroflexi bacterium	Chloroflexi	Chloroflexi	-	-	-	-	Aerobic thermophiles, anoxygenic phototrophs and anaerobic halorespirers
15.	Uncultured bacterium	-	-	-	-	-	-	-

16.	Levilinea saccharolytica	Chloroflexi	Anaerolineae	Anaerolineales	Anaerolineaceae	Levilinea	saccharolytica	Filamentous, anaerobic, belongs to chloroflexi, non- photosynthetic
17.	Uncultured bacterium	-	-	-	-	-	-	-
18.	Uncultured Chloroflexi bacterium	Chloroflexi	Chloroflexi	-	-	-	-	Aerobic thermophiles, anoxygenic phototrophs and anaerobic halorespirers
19.	Thioalkalivibrio sp.	Proteobacteria	γ-Proteobacteria	Chromatiales	Ectothiorhodospiraceae	Thioalkalivibrio	-	Extremely salt tolerant, chemolithoautotrophic, oxidizes S
20.	Uncultured Acidobacterium	Acidobacteria	Acidobacteria	-	-	-	-	Acidophilic, mostly unculturable
21.	Uncultured Acidobacterium	Acidobacteria	Acidobacteria	-	-	-	-	Acidophilic, mostly unculturable
22.	Uncultured Desulfobacteraceae bacterium	Proteobacteria	δ-Proteobacteria	Desulfobacterales	Desulfobacteraceae	-	-	Reduce sulphate to sulphide, strictly anaerobic, Deltaproteobactera
23.	Uncultured Rhodocyclales bacterium	Proteobacteria	β-proteobacteria	Rhodocyclales	-	-	-	Aqeuous conditions, ologotrophic, denitrifying bacteria, β-proteobacteria, aerobic

24.	Levilinea saccharolytica	Chloroflexi	Anaerolineae	Anaerolineales	Anaerolineaceae	Levilinea	saccharolytica	Filamentous, anaerobic, belongs to chloroflexi, non- photosynthetic
25.	Uncultured Chloroflexi bacterium	Chloroflexi	Chloroflexi	-	-	-	-	Aerobic thermophiles, anoxygenic phototrophs and anaerobic halorespirers
26.	Azospira sp.	Proteobacteria	β-Proteobacteria	Rhodocyclales	Rhodocyclaceae	Azospira	-	N fixing β-proteobacteria, found in deep waters
27.	Uncultured Ectothiorhodospiraceae	Proteobacteria	γ-proteobacteria	Chromatiales	Ectothiorhodospiraceae	-	-	Purple sulphur bacteria, deposit S globules outside the cell
28.	Uncultured Chloroflexi bacterium	Chloroflexi	Chloroflexi	-	-	-	-	Aerobic thermophiles, anoxygenic phototrophs and anaerobic halorespirers
29.	Uncultured Acidobacteria	Acidobacteria	Acidobacteria	-	-	-	-	Acidophilic, mostly unculturable
30.	Uncultured Acidobacteria	Acidobacteria	Acidobacteria	-	-	-	-	Acidophilic, mostly unculturable

Discussion

5. DISCUSSION

Conventional methods for the analysis of diversity of microbial communities in different environments are based on the cultivation of microorganisms on laboratory media. However, these methods are successful in isolating only less than 1 per cent of total population of bacteria from natural habitats (Ward *et al.,* 1992). It means that more than 99 per cent of genetic information and biotechnological potential of those cannot be studied or used by conventional methods. The development of molecular biological techniques that do not depend upon culturing has proven effective to gain insight into the diversity and structure of microbial communities. In particular, the ability to recover and analyze 16S rRNA genes directly from environmental DNA provides a mean to investigate microbial populations without the need to culture them. Bacterial identification using such DNA-based tools, such as polymerase chain reaction, is increasingly popular due to their specificity and speed, compared to culture based methods (Louie *et al*., 2000; Ghosh *et al*., 2010).

The 16S rRNA gene analyses have been used to study diverse bacterial and archaeal communities in extreme environments with a wide range of salinity, pH and temperature. These environments are interesting to study microbial diversity, to identify novel microorganisms and to understand the functioning of an ecosystem (Lizama *et al*., 2001). *Pokkali* soils of Kerala are saline and acidic. Microbial floras in these soils have not been characterized systematically and it is possible that these microbes may be tolerant to acidity and salinity.

Based on the above information, the present study entitled "Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala" was taken up with the objective of analyzing the bacterial community, both culturable and non-culturable, in the saline *Pokkali* soils of Kerala through 16S rDNA sequencing approach.

5.1 *Pokkali* soils of Kerala

"*Pokkali*" are acclaimed for unique way of reclamation and management of soil salinity and also for the integrated farming system involving rice - fish/prawn. These include coastal paddy fields of Ernakulum, coastal villages of Thrissur, Alleppey, and Kottayam districts and are confluent with the Vembanad Lake through canals. A single-crop of rice is taken in the low saline phase of the production cycle (June to mid-October) on mounds. This is followed by prawn farming during the high saline phase (November to April). In this traditional method of rice cultivation, no chemical fertilizers or plant protection chemicals are applied. The *Pokkali* fields are subjected to periodic submergence due to daily tidal inflows and outflows. These soils are highly fertile owing to the presence of large quantities of organic matter arising from decomposed aquatic weeds and paddy stubbles. The high organic matter content of the soil may be associated with high microbial activity. However, no reports are available on the diversity of microflora in these soils.

5.2 Soil sampling

Soil characteristics can vary significantly from one spot to another, even in a small garden or field. Sample depth is also a factor for various nutrients, and other soil components vary during the year, so sample timing may also affect results. Mixing soil from several locations to create an "average" (or "composite") sample is a common procedure (Foster, 1988). Microbial habitat is often limited to 6inches of soil and as depth increases, microbial activity decreases (Griffiths *et al*., 1977). Microhabitats for soil micro organisms include the surfaces of the soil aggregates, and the complex pore spaces between and inside the aggregates. Some pore spaces are inaccessible for microorganisms owing to size restrictions. The metabolism and the survival of soil microorganisms are strongly influenced by the availability of water and nutrients (Hassink *et al.,* 1993).

Electrochemical properties of the composite soil sample were measured. The electrical conductivity (EC) was 4.03dsm^{-1} and the pH 3.79. The total dissolved salt (TDS) was calculated from the EC value as 6448 mg/kg. These values indicate that the soil was typically an acid saline soil. Soil electrical conductivity (EC) is a measurement that correlates with soil properties that affect crop productivity, including soil texture, cation exchange capacity (CEC), drainage conditions, organic matter level, salinity, and subsoil characteristics (Anderson-Cook *et al.*, 2002; Jackson, 1958). George (2011) reported that *Pokkali* soil greatly varies with respect to electrochemical properties, soil texture, and available nutrient status.

Soil chemical and biological changes over a period of time result in breaking down or combining compounds. These processes change once the soil is removed from its natural ecosystem (flora and fauna that penetrate the sampled area) and environment (temperature, moisture, and solar light/radiation cycles). As a result, the chemical composition analysis accuracy can be improved if the soil is analyzed soon after extraction usually within 24 hours. The chemical changes in the soil can be slowed during storage and transportation by freezing it. Longer term soil sample preservation (many months) can be accomplished by air drying.

5.3 Isolation, Purification and Quantification of DNA

The key factor for any metagenomic study is the isolation of quality environmental DNA in appreciable amount from a given environment. It is also one of the bottlenecks in metagenomic studies. The extracted DNA should be of high quality and in good yield to pursue molecular biology applications (Charles *et al.*, 1998; Kauffmann *et al.*, 2004). Over the past 10 years, several techniques have been described for DNA extraction from environmental samples, in addition to commercial kits (Purohit and Singh, 2009; Tsai and Olson, 1992; Santosa, 2001). These protocols are classified as direct and indirect methods. Indirect method involves bacterial cell extraction from the environmental sample followed

by cell lysis and DNA recovery (Holben *et al*., 1988). Direct extraction involves cell lysis within the sample matrix followed by separation of DNA from cell debris (Ogram *et al.,* 1987). In the latter method, lysis can be achieved either by soft or harsh treatments. Soft lysis is based on the disruption of the microorganism by enzymatic and chemical means whereas harsh lysis approaches involve mechanical cell disruption by bead beating, sonication, freeze-thawing and grinding (Siddhapura *et al.,* 2010).

In the present study, the DNA was isolated by the direct method (soft lysis) following the procedure of Siddhapura *et al.* (2010). Lysis buffer contained enzymes (lysozyme and proteinase K) and detergents (SDS, CTAB and N-lauroyl sarcosine). These enzymes and detergents help in complete lysis of bacterial cells and release of DNA. The denatured contaminating molecules were precipitated and removed with the help of phenol and chloroform. Both phenol and chloroform denature proteins; get solubilised in organic phase or interphase, while nucleic acids remain in aqueous phase. Chloroform is mixed with phenol to increase the efficiency of nucleic acid extractions. The increased efficiency is due to the ability of chloroform to denature proteins. It helps in removal of lipids, thus improving separation of nucleic acids into the aqueous phase. Chloroform: isoamyl alcohol improves deproteinization. DNA thus extracted in aqueous phase was precipitated with sodium acetate and ethanol. Final DNA pellet was dissolved in sterile distilled water.

Upon agarose gel electrophoresis, a single sharp band was observed (Plate 1A). Absence of smear on the gel indicated that no mechanical or chemical shearing of DNA occurred. It has been earlier reported that direct extraction method yields a better DNA yield than the indirect method. One of the most important pre-requisites for DNA manipulations is good quality DNA. If the DNA is sheared during the process of isolation, proper PCR amplification may not take place. The quantity of DNA isolated by soft lysis direct DNA method assessed by NanoDrop was 403.2ng/µl.

5.4 Amplification of 16S rRNA genes

The rRNA is the most conserved (least variable) gene in all cells. Portions of the rDNA sequence from distantly-related organisms are remarkably similar. This means that sequences from distantly related organisms can be precisely aligned, making the true differences easy to measure. For this reason, genes that encode the rRNA (rDNA) have been used extensively to determine taxonomy, phylogeny (evolutionary relationships) and to estimate rates of species divergence among bacteria. Thus the comparison of 16S rDNA sequence can show evolutionary relatedness among microorganisms. This work was pioneered by Carl Worne, who proposed the three Domain system of classification - Archaea, Bacteria, and Eucarya - based on such sequence information.

The 16S rDNA sequence has hypervariable regions, where sequences have diverged over evolutionary time. These are often flanked by strongly-conserved regions. Primers are designed to bind to conserved regions and amplify variable regions. The DNA sequence of the 16S rDNA gene has been determined for an extremely large number of species. In fact, there is no other gene that has been as well characterized in as many species. Sequences from tens of thousands of clinical and environmental isolates are available over the internet through the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov) and the Ribosomal Database Project (www.cme.msu.edu/RDP/html/index.html). These sites also provide search algorithms to compare new sequences to their database.

In structure based metagenomic studies, forward and reverse primers are designed based on conserved regions, the DNA in between is amplified by polymerase chain reaction (PCR). Based on the sequence homology of the amplified DNA which includes variable regions also (Fig. 34), the bacterium can be identified to the Genus level. Various scientists have used different primers from the 16S rRNA gene to amplify the gene. Schmalenberger *et al*. (2001) studied parallel analysis of 3 different hypervariable regions of 16S rDNA

sequence (V2–V3, V4–V5, and V6–V8 regions) was effective in determining the composition of bacterial consortia in maize rhizospheres. Chakravorty *et al*. (2007) studied different hypervariable regions and demonstrated different efficacies with respect to species calls in different genera, and the V2 and V3 regions were most effective for universal genus identification.

In the present study, primers designed by Siddhapura *et al.* (2010) from the conserved regions v1 and v9 were used for amplification by PCR.

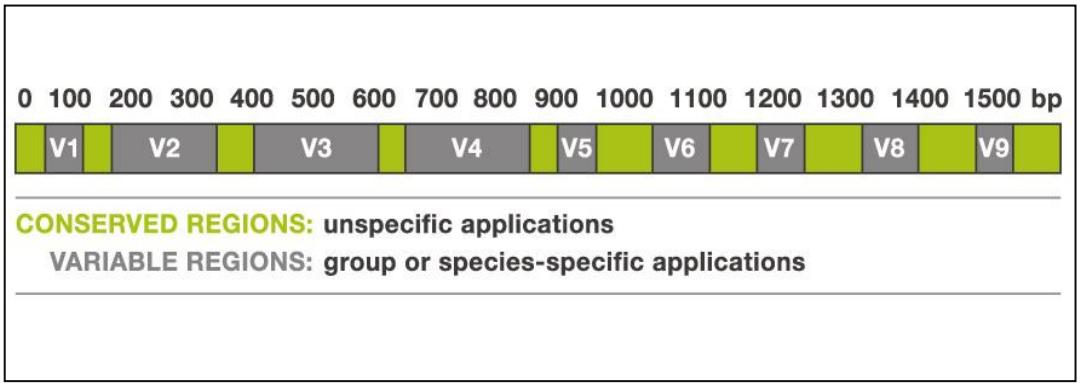


Fig. 37. Conserved and hypervariable regions in the 16S rRNA gene

PCR is a revolutionary method developed by Kary Mullis in the 1980s. PCR is based on using the ability of DNA polymerase to synthesize new strand of DNA complementary to the offered template strand. Because DNA polymerase can add a nucleotide only onto a pre-existing 3'-OH group, it needs a primer to which it can add the first nucleotide. This requirement makes it possible to delineate a specific region of template sequence that the researcher wants to amplify. At the end of the PCR reaction, the specific sequence will be accumulated in billions of copies (amplicons).

DNA template is the sample of DNA that contains the target sequence. At the beginning of the reaction, high temperature is applied to the original double-stranded DNA molecule to separate the strands from each other. DNA polymerase

is the type of enzyme that synthesizes new strands of DNA complementary to the target sequence.

Siddhapura *et al*. (2010) optimized PCR conditions for amplification of 16S rDNA from metagenomic DNA extracted from saline soils of Rajasthan and Gujrat, using the same set of primers. They reported that amplification was evident at annealing temperatures ranging from 52.4 to 56.9^{0}C, but the concentration of the product varied. In the present investigation, annealing temperature of 55^{0}C was found to give good amplification and a single, crisp band of 1500bp was obtained (Plate. 1B).

5.5 Cloning of 16S rDNA amplicons in *E. coli*

Before cloning of 16S rDNA amplicons in the plasmid vector and mobilization of the recombinant plasmid into *E. coli* cells, gel purification was carried out. Amplicons for cloning should be free of all contaminating molecules like unutilized primers, enzymes, template DNA, non-specific amplicons like primer dimer etc. present in PCR reaction mix. The amplicons were ligated in pGEMT vector, which is specifically designed for direct cloning of PCR products. It is 3kb in size and contains T7 and SP6 RNA polymerase promoters that flank a multiple cloning region within the α peptide coding region of the enzyme β-galactosidase. Thus, due to insertional inactivation of the α peptide region, the recombinants could be directly identified by the blue-white screening of indicator plates. The vector contains unique restriction sites for several restriction enzymes within the multiple cloning sites (MCS), thus facilitating easy ligation of insert and its release by digestion with restriction enzyme (Peterson *et al*., 2002).

Competent cells of *E. coli* JM109 were prepared by the $CaCl_2$ method (Mandel and Higa, 1970) and their efficiency was confirmed by transforming the cells with plasmid (pUC18) having *lacZ* gene and ampicillin resistance. *E. coli* cells alone could not grow on ampicillin containing media, as they lack the gene for ampicillin resistance. But the competent cells harbouring the plasmid could grow in the media. In the present study, large number of blue colonies were

observed on LB/ampicillin plate after overnight incubation at 37°C, confirming α-complementation of *lacZ* gene and the competence of *E. coli* (JM109) cells for transformation (Kushawa, 2008).

Bacterial transformation is the process by which bacterial cells take up naked DNA molecules. If the foreign DNA has an origin of replication recognized by the host cell DNA polymerases, the bacteria will replicate the foreign DNA along with their own DNA. When transformation is coupled with antibiotic selection techniques, bacteria can be induced to uptake certain DNA molecules, and those bacteria can be selected for that incorporation. The bacterial cell membrane is permeable to chloride ions, but is non-permeable to calcium ions. As the chloride ions enter the cell, water molecules accompany the charged particle. This influx of water causes the cells to swell and is necessary for the uptake of DNA. The exact mechanism of this uptake is unknown. It is known, however, that the calcium chloride treatment be followed by heat. When *E. coli* are subjected to 42^0C heat, a set of genes are expressed which aid the bacteria in surviving at such temperatures. This set of genes is called the heat shock genes. The heat shock step is necessary for the uptake of DNA. At temperatures above 42^0C, the bacteria's ability to uptake DNA becomes reduced, and at extreme temperatures the bacteria will die.

In the present study, the ligated product containing 16S rDNA fragment was used to transform the competent cells. Blue white screening was used to select the recombinants. The white colonies could be easily distinguished and picked up from the selection media containing 5-bromo-4-chloro-3-indolyl β-D galactoside (X-gal) and isopropyl thiogalactoside (IPTG). pGEMT vector contains polycloning sites inside β-galactosidase gene. Insertion of a new sequence would disrupt the reading frame of galactosidase encoding gene. As a result of α-complementation, the bacterial cell and vector together provided the complete protein, because one part of the gene was present in bacteria while the other in vector (Ullmann *et al*., 1967). The colonies which have not taken up the plasmid further utilized the substrate and appeared as blue colonies on chromogenic

substrate, X-gal (Horwitz *et al.*, 1964). Due to the disruption of α-complementation, all the transformed colonies harbouring the recombinant plasmid appeared as white (Plate 2.A).

Transformation efficiency is a measure of the efficiency by which cells can take up extracellular DNA and express genes encoded by it. It is influenced by the competence of cells. In the present investigation, a transformation efficiency of $1.6x10^7$ cfu/µg DNA was obtained.

Confirmation of insert in recombinant clones was carried out by colony PCR using T7 and SP6 primers (Plate 3). Colony PCR is the fastest method to screen recombinant clones for the presence of the insert DNA. Primers are used which generate a PCR product of known size. Thus, any colonies which give rise to an amplification product of the expected size are likely to contain the correct DNA sequence. Amplicons of expected size (1.5kb) were obtained from white colonies. Amplicons of approximately 100bp were obtained in blue colonies, which could be primer dimer. Several scientists have cloned the PCR amplified DNA fragments and further transformation to *E. coli* cells (Escudero *et al.*, 2000; Johnson *et al.*, 1996).

Thirty randomly picked clones were further characterized by 16S rDNA sequencing and sequence analysis.

5.6 Sequence Analysis

Sequences obtained from Bangalore GeNei were subjected to *in silico* analysis using various bioinformatics tools.

5.6.1 VecScreen

This software hosted by NCBI is useful for quick identification of segments of a nucleic acid sequence that may be of vector origin. It helps to combat the problem of vector contamination in public sequence databases. If vector sequences are not removed, it may lead to erroneous conclusions about the

biological significance of the sequence waste time and effort in analysis of contaminated sequence, delay the release of the sequence in a public database and also pollute public databases with contaminated sequence.

5.6.2 BLAST (Basic Local Alignment Search Tool)

BLAST program uses a heuristic approach (a mathematical problem solving method). BLAST first creates a dictionary of three or four letter words from the query sequence (depending on Word size) and searches for matching words in the database sequence. A substitution matrix is then constructed for assigning scores. It matches its dictionary words one after the other and looks for match score (Altschul *et al*., 1997).

5.7 Bacterial diversity of *Pokkali* soil

Of the 30 clones sequenced, 23 showed homology with uncultured bacteria and seven with culturable bacteria (Table 43 and 44). Of the unculturable bacteria, six sequences did not show homology to any of the accessions in the NCBI database, indicating that these could be novel isolates. Seventeen clones showed homology with bacterial sequences already available in NCBI database. Uncultured Acidobacteria was most predominant with five clones (Clones 12, 20, 21, 29 and 30). Four clones showed homology with uncultured Chloroflexi and three with uncultured Desulfotobacteraceae. Two clones were grouped under uncultured Ectothiorhodospiraceae and one each under uncultured *Acaryochloris*, uncultured Rhodocyclales and uncultured γ-Proteobacteria. None of the unculturable bacteria could be identified up to the species level. Only one clone was classified up to Genus (*Acaryochloris*) and five clones up to Family (Ectothiorhodospiraceae-two, Rhodocyclaceae-one and Desulfobacteraceae-three). One clone could be identified up to Order (Rhodocyclales) and ten clones up to Class (γ-Proteobacteria-one, Acidobacteria-five and Chloroflexi-four).

Among the seven clones of culturable bacteria, four could be identified up to species level (*Stenotrophomonas maltophilia*, *Thiobacillus prosperus* and

Levilinea saccharolytica) and three up to Genus (*Azospira*-two and *Thioalkylivibrio*-one). Brief description of the different groups of bacteria identified in the present study is given below:

5.7.1 Proteobacteria

This group comprised the largest number of bacteria identified from the *Pokkali* microbiota, in the present investigation. Of the 30 clones sequenced, twelve were grouped under Phylum Proteobacteria (Table 44). Of these, seven were unculturables bacteria and five culturables. Six belonged to γ-Proteobacteria, three δ-Proteobacteria and three β-Proteobacteria.

5.7.1.1 γ-Proteobacteria

Six cloned sequences belonged to this group which included three unculturable bacteria (clones 1, 7 and 27) and three culturable ones (clones 2, 9 and 19). The non-culturable bacteria included two clones (1 and 27) of Ectothiorhodospiraceae, which come under the purple sulphur bacteria. These photoautotrophic bacteria use reduced sulphur compounds as electron donors for anoxygenic photosynthesis and deposit S globules outside the cells (Frigaard and Dahl, 2009). Clone 7 was identified only up to Class, as uncultured γ-Proteobacteria.

The three culturable bacteria under γ-Proteobacteria included *Stenotrophomonas maltophilia*, *Thiobacillus prosperus* and *Thioalkalivibrio* sp. *Stenotrophomonas maltophilia* is a ubiquitous, aerobic, non-fermentative, gram-negative *bacillus* that is closely related to the *Pseudomonas* species which is found in various aquatic environments. It is an uncommon pathogen in humans which causes ailments like respiratory infections, urinary disease, and wound infections. It was first isolated in 1943 and, at that time, was named *Bacterium bookeri*. It was later classified within the genus *Pseudomonas*, then *Xanthomonas*, and then finally *Stenotrophomonas* in 1993 (Denton and Kerr, 1998; Looney, 2009). *S. maltophilia* is the only species of *Stenotrophomonas* known to infect

humans, whereas its closest genetic relatives are plant pathogens. It is frequently isolated from soil, water, animals, plant matter, and hospital equipment (Falagas, 2009).

Thiobacillus prosperus, an obligate chemolithoautotroph is Fe oxidising, acid and salt tolerant bacteria. It is a halotolerant, metal-mobilizing bacterium similar to *T. ferrooxidans.* It is able to grow aerobically on pyrite, sphslerite, chalcopyrite, and galena as well as on H_2S. It shows vigorous extraction of metal ions from sulfidic ores and is resistant to cobalt. Due to its high tolerance of salt (even 6 per cent in the case of isolate), *T. prosperus* is suitable for industrial leaching in salt containing environments.

The third clone was identified as *Thioalkalivibrio* sp. It is extremely salt tolerant, chemolithoautotrophic, sulphur oxidizer which is closely related to *Thioalkalivibro denitrificans.* It was isolated as a dominant sulfide-oxidizing species from a full-scale bioreactor removing sulfide from biogas. It is capable of growth at up to 4M of total sodium and 3.8 M of potassium (Sorokin and Kuenen, 2005).

5.7.1.2 β-Proteobacteria

Three cloned sequences (one unculturable and two culturable bacteria) were classified as β-Proteobacteria (Table 44). Clone 23 was identified as uncultured Rhodocyclales bacterium. It is an aerobic denitrifying bacteria found in oligotrophic conditions like aqueous environments. Clones 6 and 26 were identified as *Azospira* sp., a nitrogen fixing bacterium found in deep waters. The genus *Azospira* includes a single validly published species, *Azospira oryzae*, strains of which have been isolated from surface-sterilized roots of gramineae such as Kallar grass (*Leptochloa fusca*) and rice (*Oryza*), or resting stages (sclerotia) of a basidiomycete (Reinhold *et al.*, 1986). Nitrogen-fixing ability was determined by growth and acetylene reduction in semi-solid (0.2 per cent w/v,

agar), nitrogen-free SM medium (Reinhold *et al.*, 1986). A PCR-based assay using the PolF/PolR PCR primer set and a PCR protocol as described by Poly *et al.* 2001 was used to detect presence of the *nifH* gene which is responsible for N fixation in soil. The Genus *Azospira* is also a member of Rhodocyclales and belongs to Family *Rhodocyclaceae*. This family of gram-negative bacteria includes many genera previously assigned to the family *Pseudomonadaceae*. The family contains mainly aerobic or denitrifying rod-shaped bacteria, which exhibit very versatile metabolic capabilities. Many occur in waste water and play an important role in biological remediation in waste water treatment.

5.7.1.3 δ-Proteobacteria

Three cloned sequences showed maximum homology with uncultured Desulfobacteraceae which are sulphate reducing bacteria (reduce sulphate to sulphide). They are strictly anaerobic chemoorganotrophs, chemolithoheterotrophs, or chemolithoautotrophs with respiratory metabolism. Growth occurs only in the presence of sulphate, sulphite, or thiosulfate as electron acceptors. Elemental sulphur inhibits its growth. *Desulfobacter* sp. is most common in anoxic brackish or marine sediments but some types may be found in anoxic sediments or in activated sludge (Bak and Pfennig, 1987).

5.7.2 Chloroflexi

Chloroflexi is second largest group of bacteria present in *Pokkali* soil which includes four uncultured and two culturable bacteria. Clones 14, 18, 25 and 28 were identified as uncluturable Chloroflexi bacteria. These are aerobic thermophilic anoxygenic phototrophs (Sekiguchi *et al.*, 2003). The culturable bacteria (Clones 16 and 24) under the Class Chloroflexi included *Levilinea saccharolytica*. This is a filamentous anaerobic, non- photosynthetic bacteria under Chloroflexi.

5.7.3 Acidobacteria

It is third diverse group of bacteria present in *Pokkali* soil which includes five cloned sequences, among the 30 clones under investigation. Acidobacteria has a widespread occurrence in a variety of ecosystems (Hugenholtz *et al.,* 1998; Barns *et al.,* 1999) and it represents the second most abundant phylum after the Proteobacteria in several soils (Janssen, 2006). Genomic and culture traits indicate the use of carbon sources that span simple sugars to more complex substrates such as hemicellulose, cellulose, and chitin. The genomes encode low-specificity major facilitator superfamily transporters and high-affinity ATP-binding cassette transporters (ABC-transporter) for sugars, suggesting that they are best suited to low-nutrient conditions. They appear capable of nitrate and nitrite reduction but not N_2 fixation or denitrification (Dore and Karl, 1996).

5.7.4 Cyanobacteria

One cloned sequence has shown maximum homology with Uncultured *Acaryochloris sp.* is a unique cyanobacterium that is able to produce chlorophyll *d* as its primary photosynthetic pigment and thus efficiently use far-red light for photosynthesis. *Acaryochloris* species have been isolated from marine environments in association with other oxygenic phototrophs, which may have driven the niche-filling introduction of chlorophyll *d* (Miyashita, 1996).

5.7.5 Unculturable unidentified bacteria

Six of the cloned sequence have shown maximum homology with unknown uncultured bacterial species, which amounts to 20 per cent. This may include novel classes' of bacteria. Bacterial population has changed in response to the changing environment in a coastal system, such as dissolved organic carbon (Takenaka *et al*., 2007). Reports on the microflora in acid saline and acid sulphate soils of Kerala are not available. This indicates that *Pokkali* soils are good reservoirs of a diverse group of bacterial population.

5.8 Conventional method for characterization of culturable bacteria

A conventional method to assess the diversity of microbes in any environment involves pure culturing on appropriate media and characterization by cultural, morphological, biochemical and physiological tests. These techniques often fail to isolate several microorganisms in the environment. In conventional genomics also genetic material is extracted from pure cultures and then subject to molecular analysis. Therefore, a large majority of the microorganisms may not be represented in the results. Moreover, pure culturing is often laborious. In the present investigation, an attempt was also made to isolate and identity the predominant morphotypes of bacteria from *Pokkali* soil.

5.8.1 DNA isolation from bacteria

Gram-positive bacteria have a rigid cell wall that can be difficult to lyse, and the methods used for the isolation of genomic DNA from Gram-negative bacteria are generally not successful with Gram-positive bacteria. Molecular epidemiologic studies and many other screening applications may also require preparation of genomic DNA from a large number of samples. Clean DNA is generally required for any *in vitro* amplification procedures, such as PCR. Several procedures are available for preparing genomic DNA from Gram-positive bacteria, but they are time-consuming and are not suitable for application to large sample numbers. Usually, large culture volumes are required to obtain sufficient amounts of genomic DNA for various molecular manipulations. Moreover, they involve a lengthy lysis step, which can result in considerable DNA damage.

In the present study, total DNA was isolated from three bacterial isolates (BS1, BS2 and BS3) according to the procedure of Sambrook and Russel (2001). Similar procedure was used by several workers for isolating DNA from *Bacillus thuringiensis* (Ben- Dov *et al*., 1997; Beron *et al*., 2004; Rajesh *et al*., 2006).

Tris EDTA is one of the reagents used in DNA isolation. EDTA present in it could effectively chelate magnesium ions and mediate aggregation of nucleic acid. Lysozyme helps in lysis of bacterial cells. The detergent used was SDS which acts as a nuclease inhibitor and also dissolves membranes.

The protocol yielded a single discrete band on 0.8 per cent agarose gel. No RNA contamination was observed (Plate 4A). The OD_{260}/OD_{280} values of DNA ranged between 1.84 and 2.05 indicating that DNA was good without much protein contamination. Quantity of DNA was found to be highest in the isolate BS1 (1974.9 ng/µl) and the lowest BS3 (391.0 ng/µl).

5.8.2 16S rRNA amplification

DNA extracted from the three bacterial isolates was used to amplify the 16S rDNA with the same primers as in metagenomic DNA (Siddhapura *et al* 2010). Amplicons of expected size (1500bp) were obtained (Plate 4B).

5.8.3 Sequence analysis

In the present study sequence data obtained from 16S rDNA amplicons from three selected bacterial isolates were subjected to Blastn analysis.

In the isolate **BS1-** Blastn revealed that 83 per cent query coverage and 98 per cent identity with *Bacillus* sp., and also shows maximum homology with *Bacillus aryabhattai*, *Bacillus subtils* and *Bacillus megaterium* (Table 40). Echigo *et al*. (2005) made similar observations in that many halophilic bacteria in *Bacillus* sp. that are able to grow in the presence of 20 per cent NaCl inhabit in saline environments such as Dead Sea, the Great Salt Lake soils, and yards.

In the isolate **BS-2** Blastn result shows that 99 per cent sequence similarity and 99 per cent homology with *Bacillus* sp., and also shows homology with *Bacillus megaterium* and *Bacillus aryabhattai*. (Table 41). Similar reports had

been made by earlier workers Kiel and Gaylarde, (2007). Many of the bacteria growing on these plates have been shown to be of the genus *Bacillus*.

In the isolate **BS-3** Blastn result shows that 96 per cent query coverage and 90 per cent homology with *Bacillus* sp., 94 per cent query coverage and 89 per cent homology with *Geobacillus* sp, 99 per cent query coverage and 100 per cent homology with *Bacillus cerus* and 99 per cent query coverage and 100 per cent homology with *Bacillus anthracis* (Table 42). Similar kinds of observations were finding by earlier workers (Lawson *et al.*, 2008; Akhtar *et al*., 2007).

The bacteria obtained by conventional culturing on media failed to appear in the metagenomic data. This may be due to the fact that out of thousands of clones obtained in the metagenomic library, only thirty were sequenced. If at least 200 colonies could be sequenced, it would have given a more representative picture of the microbiota in *Pokkali* soil. It could also be assumed that the number of culturable bacteria in the present investigation was also insufficient to give a true representation of culturable bacteria.

In the present study, bacterial microbiota were assessed successfully from saline *Pokkali* habitats of Kerala through metagenomic approach. Total environmental DNA was isolated directly from saline soil and 16S rRNA genes were amplified. The cloned sequences which showed homology with several bacteria were grouped under Proteobacteria, Chloroflexi, Acidobacteria and Cynobacteria in which most of them were saline tolerant, acidophilic, nitrogen fixing, sulphur reducing, sulphur oxidizing, strict anaerobes and pathogenic. By using conventional method of serial dilution plate count on nutrient agar media the number of bacterial population were observed and DNA were isolated from culturable bacterial isolates. Sequence analysis showed maximum homology with *Bacillus* sp.

Further investigations are required for the function derived approach for gene prospecting (acid and salt tolerance); more number of clones may be

sequenced to get a true representation of microbiota. Emphasis must be laid on the identification of unknown uncultured species which may come under novel ones so far not reported.

Summary

6. SUMMARY

The study on "Metagenomic approach to assess diversity of bacterial community in saline *Pokkali* habitats of Kerala" was carried out at the Centre for Plant Biotechnology and Molecular Biology (CPBMB) and Department of Agricultural Microbiology, College of Horticulture, Vellanikkara during the period 2009-2011. The main objective of the study was to analyze the bacterial community, both culturable and non-culturable, in the saline *Pokkali* soils of Kerala through 16S rDNA sequencing approach. The salient findings of the study are summarized below:

1.The soil samples were collected from Pizhala, Vytilla, Cochin South town, Kerala.

2.Electrical conductivity (EC) of the soils sample was estimated to be 4.03dsm^{-1}, total dissolved salt (TDS) 6448mg/kg and pH 3.79.

3.Metagenomic DNA isolation was carried out using direct method (soft lysis), following the procedure of Siddhapura *et al.* (2010). The extract contained good concentration of DNA (403.2 ng/µl) and purity was indicated by an OD_{260}/OD_{280} value of 1.84.

4.PCR amplification of 16S rRNA genes was carried out with the specific primers as suggested by Siddhapura *et al*. (2010) using 100ng of the metagenomic DNA as template. Agarose gel electrophoresis on 0.8 per cent revealed presence of a single band of the expected size (1.5kb).

5.The 1.5kb amplicon of 16S rRNA gene was purified from agarose gel and cloned into pGEMT vector. The ligated product was transformed in competent *E.coli* cells. The transformation efficiency was found to be as 1.6×10^7 cfu/µg DNA.

6.Thirty randomly picked recombinant clones were confirmed for the presence of insert by colony PCR using T7 and SP6 primers. 1.5kb amplicons were obtained from white colonies (true recombinants) whereas only primer dimmer was present in blue colonies.

7.Sequencing of the cloned 16S rDNA fragments was carried out using universal primers T7 and SP6.

8.Forward and reverse sequencing data was subjected to VecScreen to remove the vector sequence. These were merged to get the complete sequence and then subjected to *in silico* analysis using bioinformatics tools.

9.Blasn analysis revealed taxonomic position of the cloned sequences. Four different phyla of bacteria viz. Proteobacteria, Acidobacteria, Chloroflexi, Cyanobacteria were identified. Six clones were grouped under unculturable, unidentified bacteria.

10.Conventional method was used for characterization of three predominant culturable bacteria of *Pokkali* soil.

11.Three predominant bacterial isolates (BS1, BS2, and BS3) from 0%, 5%, 10%, 15%, 20% salt extract agar media were selected after serial dilution plate count for study.

12.PCR amplification of 16S rRNA genes was carried out with the specific primers as suggested by Siddhapura *et al*. (2010) using 100ng of the template DNA.

13.The sequence obtained was subjected to nucleotide Blastn search. All the three showed homology with available organism nucleotide data base.

14.The sequence data showed homology with genus Bacillus spp.

The present investigation revealed that 76.7% of the total bacterial community in Pokkali soil was unculturable and only the remaining 23.3% is culturable. The study therefore, highlights the importance of metagenomic approach in assessing the diversity of microbial communities in specific environments.

References

REFERENCES

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W. 2000. The genome sequence of Drosophila melanogaster. *Science.* 287: 2185–2195.

Akhtar, N., Ghauri, M.A., Iqbal, A., Anwar, M. A., Akhtar, K. 2008. Biodiversity and phylogenetic analysis of culturable bacteria indigenous to khewra salt mine of pakistan and their industrial importance. *Brazilian J. Microbiol.* 39: 143-150.

Altschul, S.F., Thomas, L., Madden, A.A, Schaffer, A., Zhang, J., Zhang, z., Miller, W. and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Res.* 25: 3389-3402.

Amann, R.I., Ludwig, W. and Schleifer, K.H. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59: 143-169.

Anderson-Cook, C.M., M.M. Alley, J.K. Roygard, R. Khosla, R.B. Noble, and J.A. Doolittle. 2002. Differentiating soil types using electromagnetic conductivity and crop yield maps. *Soil Science Society of America Journal* 66: 1562-1570.

Antón, J., Rosselló-Mora, R., Rodriguez-Valera, F., and Amann, R. 2000. Extremely halophilic Bacteria in crystallizer ponds from solar salterns. *Appl. Environ. Microbiol.* 66: 3052-3057.

Babcock, D.A., Wawrik, B., Paul, J.H., Mc Guinness, L. and Kerkhof, L.J. 2007. Rapid screening of a large insert BAC library for specific 16S rRNA genes using TRFLP. *Microbiol. Methods.* 71: 156-161.

Bak, f. and Pfennig, N. 1987. Chemolithotropic growth of desulfovibrio sulfo dimutans, new species by disproportionation of inorganic sulfer compound. *Arch. Microbiol.* 147: 184-189.

Baker, G.C., Smith, J.J., Cowan, D.A. 2003. Review and reanalysis of domain-specific 16S primers. *J. Microbiol. Methods.* 55: 541-555.

Barns, S.M., Takala, S.L., Kuske, C.R. 1999. Wide distribution and diversity of members of the bacterial kingdom *Acidobacterium* in the environment. *Appl. Environ. Microbiol.* 65**:** 1731-1737.

Behura, S.K. 2006. Molecular marker systems in insects: current trends and future avenues. *Mol. Ecol.* 15: 3087-3113.

Beja, O., Suzuki, M.T., Koonin, E.V., Aravind, L., Hadd, A., Nguyen, L.P., Villacorta, R., Amjadi, M., Garrigues, C., Jovanovich, S.B., Feldman, R.A. and DeLong, E.F. 2000. Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ. Microbiol*. 2: 516-529.

Ben-Dov, E.A., Zaritsky, E., Dahan, Z., Barak, R., Sinai, R., Manasherob, A., Khamraev, E., Troitskaya, A., Dubitsky, N.B. and Margalith, Y. 1997. Extended screening by PCR for seven *cry* group genes from field collected strains of *Bacillus thuringiensis*. *Appl. Environ. Microbiol.* 63: 4883-4890.

Beron, C.M., Curatti, L. and Salerno, G.L. 2004. New strategy for identification of Novel *cry*-type genes from *Bacillus thuringiensis* strains. *Appl. Environ. Microbiol*. 71(2): 761-765.

Bhat, M.K., 2000. Research review paper: cellulases and relate enzymes in biotechnology. *Biotechnol. Adv*. 18: 355-383.

Bond, P.L., Smriga, S.P. and Banfield, J.F. 2000. Phylogeny of microorganisms populating thick, subaerial, lithotrophic biofilm at an extreme acid mine drainage site. *Appl. Environ. Microbiol*. 66: 3842-3849.

Bouchotroch, S., Quesada, E., del Moral, A., Llamas, I., Bejar, L. and V. 2001. *Halomonas maura* sp. nov., a novel moderately halophilic, exopolysaccharide-producing bacterium. *Int. J. Syst. Evol. Microbiol*. 51: 1625-32.

Brady, S.F., Chao, C.J. and Clardy, J. 2004. Longchain N-acyltyrosine synthases from environmental DNA. *Appl. Environ. Microbiol.* 70: 6865-6870.

Brady, S.F., Choa, C.J., Handelman, J. and Clardy, J. 2001. Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA. *Org. Lett.* 3: 1981-1984.

Bresler, E., McNeal, B.L., and Carter, D.L. 1982. Saline and sodic soils. Quoted in D.L. Spark (ed.). (1995). Environmental Soil Chemistry (p. 230). California: Academic Press.

Brown, A. D. 1976. Microbial water stress. *Bacteriol. Rev.* 40: 803-846.

Chakravorty, S., Helb, D., Burday, M., Connell, N., Alland, D. 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J. Microbiol. Methods*. 69: 330-339.

Charles, C.S., Ivor, T.K., William, L.S., Rita, R.C. 1998. Capacity of Nine Thermostable DNA Polymerases to Mediate DNA Amplification in the Presence of PCR-Inhibiting Samples. *Appl. Environ. Microbiol.* 64 (10): 548-554.

Chen, K. and Pachter, L. 2005. Bioinformatics for whole genome shotgun sequencing of microbial communities. *PLoS. Comp. Biol*. 1: 2-24.

Clarson, D., Ramasawmi, P.P. and Sreeramalu, U.S. 1984. Influence of amendments on rice in a sodic soil. *Madras Agril. J.* 71(10): 681-684.

Courtois, S., Cappellano, C.M. Ball, M., Francou, F.X., Normand, P., Helynck, G., Martinez, A., Kolvek, S.J., Hopke, J., Osburne, M.S., August, P.R., Nalin, R., Guérineau, M., Jeannin, P., Simonet, P. and Pernodet, J.L. 2003. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol*. 69: 49-55.

Cowan, D.A. 2000. Microbial genomes the untapped resource. *Trends. Biotechnol*. 18: 14-6.

Curtis, T.P., Sloan, N.T. and Scannell, J.N. 2002. Estimating prokaryote diversity and its limits. *Proc. Natl. Acad. Sci.* USA. 99: 10494–10499.

Cytryn, E., Minz, D., Oremland, R.S., and Cohen, Y. 2000. Distribution and diversity of Archaea corresponding to the limnological cycle of a hypersaline stratified lake (Solar Lake, Sinai, Egypt). *Appl. Environ. Microbiol.* 66: 3269-3276.

Da Costa, M.S., Santos, H. and Galinski, E.A. 1998. An overview of the role and diversity of compatible solutes in bacteria andarchaea. *Adv. Biochem. Eng. Biotechnol*. 61: 117-53

Daniel, R. 2005. The metagenomics of soil. *Nat. Rev. Microbiol.* 3: 470–478.

Demergasso, C., Casamayor, E.O., Chong, G., Galleguillos, P., Escudero, L., Pedros-Alio, C. 2004. Distribution of prokaryotic genetic diversity in athalassohaline lakes of the Atacama Desert, Northern Chile. *FEMS. Microbiol. Ecol.* 48: 57–69.

Denton, M. and Kerr, K.G. 1998. Microbiological and clinical aspects of infection associated with Stenotrophomonas maltophilia. *Clin. Microbiol. Rev*. 11:57-65.

Dojka, M.A., Harris, J.K., and Pace N.R. 2000. Expanding the known diversity and environmental distribution of an uncultured phylogenetic division of Bacteria. *Appl. Environ*. *Microbiol.* 66: 1617–1621.

Dore, J. and Karl, D. 1996. Nitrification in the euphotic zone as a source for nitrite, nitrate, and nitrous oxide at Station ALOHA. *Limnol. Oceanogr.* 41:1619-28

Echigo, A., Hino, M., Fukushima, T., Mizuki, T., Kamekura, M., and Usami, R. 2005. Endospores of halophilic bacteria of the family *Bacillaceae* isolated from non-saline Japanese soil may be transported by Kosa event (Asian dust storm). *Appl. Environ. Microbiol*. 57: 1525-1530.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E. 2005. Diversity of the human intestinal microbial flora. *Science*. 308: 1635–1638.

Egamberdieva, D. and Kucharova, Z. 2009. Selection for root colonising bacteria stimulating
Wheat growth in saline soils. *Biol. Fertil. Soil*. 45: 563-571.

Elend, C., Schmeisser, C., Leggewie, C., Babiak, P., Carballeira, J. D., Steele, H.L., Reymond, J.L., Jaeger, K.E. and Streit, W.R. 2006. Isolation and biochemical characterization of two novel metagenome-derived esterases. *Appl. Environ. Microbiol*. 72: 3637-3645.

Entcheva, P., Liebl, W., Johann, A., Hartsch, T. and Streit, W.R. 2001. Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia. *Appl. Environ. Microbiol*. 67: 89 – 99.

Escudero, I.R., Estela, A., Porcar, M., Martinez, C., Oguiza, J.A., Escriche, B., Ferre, J. and Caballero, P. 2000. Molecular and insecticidal characterization of a *Cry11* protein toxic to insects of the families Noctuidae, Tortricidae, Plutellidae and Chrysomelidae. *Appl. Environ. Microbiol.* 72(7): 4796-4804.

Falagas, M.E., Kastoris, A.C., Vouloumanou, E.K. 2009. Attributable mortality of Stenotrophomonas maltophilia infections: a systematic review of the literature. Future Microbiol; 4:1103-1111.

Ferrer, M., Golyshina, O.V., Chernikova, T.N., Khachane, A.N., Reyes-Duarte, D., Martins Dos Santos, V.A.P., Strompl, C., Elborough, K., Jarvis, G., Neef, A., Yakimov, A. Timmis, M.M., and Golyshin, P.N. 2005. Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. *Environ. Microbial.* 7: 1996 – 2010.

Foster, R. C. (1988). Microenvironments of soil microorganisms. *Biol. Fertil. Soils.* 6: 189-203.

Frank, D.N. and Pace, N.R. 2008. Gastrointestinal microbiology enters the Metagenomic era. *Curr. Opin. Gastroenterol*. 24: 4-10.

Frigaard, N.U. and Dahl, C. 2009 *Advances in Microbial Physiology (*ed) R K Poole (Amsterdam: Elsevier). 54: 103-200.

Frostegard, A., Courtois, S., Ramisse, S., Clerc, D., Bernillon, L.G., Francoise, J., Pascale, N. and Pascal, S. 1999. Quantification of bias related to the extraction of DNA directly from soils. *Appl. Environ. Microbiol.* 65: 5409-5420.

Gabor, E.M., Vries, E.J. and Janssen, D.B. 2007. Construction, characterization, and use of small-insert gene banks of DNA isolated from soil and enrichment cultures for the recovery of novel amidases. *Environ. Microbiol.* 6, 948-958.

Gabor, E.M. and Janssen, D.B. 2004. Increasing the synthetic performance of penicillin acylase PAS2 by structure-inspired semi-random mutagenesis. *Protein Eng. Des. Sel*. 17: 571-579.

Gabor, E.M., Liebeton, K., Niehaus, F., Eck, J. and Lorenz. P. 2007. Updating the Metagenomics toolbox. *Biotechnol. J*. 2: 201-206.

George, A. 2011. Availability indices of boron in major soil group of Kerala. M.Sc. (Ag) thesis. Kerala Agricultural University, Thrissur, 137p.

Ghazanfar, S. and Azim, A. 2009. Metagenomics and its Application in Rumen Ecosystem:

Potential Biotechnological Prospects. *Pakistan J. of Nutr*. 8: 1309-1315.

Ghosh, A., Dey, N., Bera, A., Tiwari1, A., Sathyaniranjan, K.B., Chakrabarti, K., Chattopadhyay, D. 2010. Culture independent molecular analysis of bacterial communities in the mangrove sediment of Sundarban, India. *licensee BioMed Central Ltd*. 6: 1-11.

Gill, S.R., Pop, M., DeBoy, R., Eckburg, P.B., Taurnbaugh, P.J. 2006. Metagenomic analysis of the human distal gut microbiome. *Science.* 312: 1255-1359.

Gillespie, D.E., Brady, S.F., Bettermann, A.D., Cianciotto, N.P., Liles, M.R., Rondon, M.R., Clardy, J., Goodman, R.M. and Handelsman, J. 2002. Isolation of antibiotics turbomycin a and b from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* 68: 4301- 4306.

Gonzalez, C., Gutierrez, C. and Ramirez, C. 1978. *Halobacterium vallismortis* sp. nov. an amylolytic and carbohydrate-metabolizing, extremely halophilic bacterium. *Can. J. Microbiol.* 24: 710-715.

Grant, S., Sorokin, D.Y., Grant, W.D., Jones, B.E., and Heaphy, S. 2004. A phylogenetic analysis of *Wadi el Natrun* soda lake cellulose enrichment cultures and identification of cellulase genes from these cultures. *Extremophiles.* 8: 421-429.

Griffiths, R.P., Hayasaka, S.S., McNamara, T.M., and Morita, R.Y. 1977. Comparison between two methods of assaying relative microbial activity in marine environments. *Appl. Environ. Microbiol*. 34:801-805.

Hanada, S. 2003. Filamentous anoxygenic phototrophs in hot springs. *Microbes. Environ*. 18: 51-61.

Handelsman, J. 2004. Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol. Mol. Biol. Rev*. 68: 669-685.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J. and Goodman, R.M. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5: 245–249.

Hansen, J., Ruedy, R., Glascoe, J. and Sato, M. 1999. Soils texture analysis of surface temperature change. *J. Geophys. Res*. 104: 30997–31022.

Hao, M.V., Kocur, M., Komagata, K. 1984. *Marinococcus* gen nov., a new genus for motile cocci with meso-diaminopimelic acid in the cell wall and *Marinococcus albus* sp. nov. and *Marinococcus halophilus* (Novitsky and Kushner) comb. nov. *J. Gen. Appl. Microbiol*, 30: 449-459.

Hassink, J., Bouwman, L. A., Zwart, K. B. and Brussaard, L. 1993. Relationship between habitable pore space, soil biota, and mineralization rates in grassland soils. *Soil. Biol. Biochem.* 25: 47-55.

He, J.Z., Zhu, Y.G., Zheng, Y.M., Zhang, L.M. and Shen, J.P. 2007. Methodology and application of soil metagenomics. Chinese Academy of Science (Article in Chinese) 1: 212-218.

Henne, A., Daniel, R., Schmitz, R.A. and Gottschalk, G. 1999. Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4 -hydroxybutyrate. *Appl. Environ. Microbiol.* 65: 3901–3907.

` Henne, A., Schmitz, R.A., Bomeke, M., Gottschalk, G., and Daniel, R. 2000. Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Appl. Environ. Micro*biol. 66: 3113-3116.

Heyrman, J. and Swings, J. 2001. 16S rDNA sequence analysis of bacterial isolates from biodeteriorated mural paintings in the Servilia tomb (necropolis of Carmona, Seville, Spain). *Syst. Appl. Microbiol*. 24: 417-422.

Holben, W.E., Jansson, J.K., Chelm, B.K. and Tiedje, J.M. 1988. DNA probe method for the detection of specific microorganisms in the soil bacterial community. *Appl. Environ. Microbiol.* 54: 703-711.

Horwitz, J.P., Chua, J., Curby, R.J., Tomson, A.J. and Fisher, B.E. 1964. Substrates for cytochemical demonstration of enzyme activity. *J. Med. Chem*. 7: 574-576.

Hu, Y., Fu, C., Yin, Y., Cheng, G., Lei, F., Yang, X., Li, J., Ashforth, E.J., Zhang, L., and Zhu, B. 2010. Construction and Preliminary Analysis of a Deep-Sea Sediment Metagenomic Fosmid Library from Qiongdongnan Basin, South China Sea. *Mar. Biotechnol.* 12: 719–727.

Hugenholtz, P., Goebel, B.M. and Pace, N.R. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol*. 180**:** 4765–4774.

Humayoun, S.B., Bano, N., Hollibaugh, J.T. 2003. Depth distribution of microbial diversity in mono lake, a meromictic soda lake in California. *Appl. Environ. Microbiol.* 69: 1030–1042.

Jackson, M.L. 1958. *Soil chemical analysis.* Prentice Hall of India Private Ltd., New Delhi, 498p.

Janssen, P.H. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl. Environ. Microbiol.* 72: 1719–1728.

Johnson, T.M., Rishi, A.S., Nayak, P. and Sen, S.K.1996. Cloning of a *cryIIIA* endotoxin gene of Bacillus thuringiensis var. Tenebrionis and its transient expression in indica rice. *J.Bioscience.* 21(5):673-685.

Kamagata, Y. and Tamaki, H. 2005. Cultivation of uncultured fastidious microbes. *Microbes. Environ.* 20: 85-91.

Kamekura, M. and DyallSmith, M.L. 1995. Taxonomy of the family Halobacteriaceae and the description of two new genera *Halorubrobacterium* and *Natrialba*. *J. Gen. Appl. Microbiol*. 41: 333-350.

Karl, D.M. 2007. Microbial oceanography: paradigms, processes and promise. *Nat. Rev*. *Microbiol*. 5: 759–769.

Kauffmann, I.M., Schmitt, J. and Schmid, R.D. 2004. DNA isolation from soil samples for
cloning in different hosts. *Appl. Microbiol. Biotechnol*. 64: 665–670.

Kiel, G. and Gaylarde, C. 2007. Bacterial diversity in biofilms on external surfaces of historic buildings in Porto Alegre. *World J. Microbiol. Biotechnol*. 22: 293-297.

Knietsch, A., Waschkowitz, T., Bowien, S., Henne, A. and Daniel, R. 2003. Metagenomes of complex microbial consortia derived from different soils as sources for novel genes conferring formation of carbonyls from short-chain polyols on *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* 5: 46–56.

Koizumi, Y., Kojima, H., Oguri, K., Kitazato, H., Fukui, M. 2004. Vertical and temporal shifts in microbial communities in the water column and sediment of saline meromictic Lake Kaiike (Japan), as determined by a 16S rRNA-based analysis and related to physicochemical gradients. *Environ. Microbiol.* 6: 622–637.

Kowalchuk, G.A., Speksnijder, A.G.C.L., Zhang, K., Goodman, R.M. and Veen, J.A. 2007. Finding the needles in the metagenome haystack. *Microb. Ecol.* 53: 475–485.

Kramer, D. 1984. Cytological aspects of salt tolerance in higher plants. Salinity tolerance in plant. John Wiley and Sons, New York, pp. 3-17.

Krsek, M. and Wellington, E.M. 1999. Comparison of different methods for the isolation and purification of total community DNA from soil. *J. Microbiol. Methods*. 39: 287- 294.

Kurokawa, K., Itoh, T., Kuwahara, T., Oshima, K., Hoh, H. 2007. Comparative Metagenomic revealed commonly enriches gene sets in human gut microbiomes. DNA *Res*. 14: 169–181.

Kushawa, R. 2008. Developement and analysis of ESTs (Express sequence tags) in Black pepper (*Piper nigrum* L.) M.Sc. (Agri.). thesis, Kerala Agricultural Unoversity, thrissur, India p. 130.

Kushner, D. J. 1993. Growth and nutrition of halophilic bacteria, In *The Biology of Halophilic Bacteria*, pp. 87–103.

Kushner, D.J. and Kamekura, M. 2006. *Physiology of halophilic eubacteria.* In *Halophilic bacteria* Volume I. *Boca Raton, CRC. Press.* 1988: 109-140.

Langer, M., Gabor, E.M., Liebeton, K., Meurer, G., Niehaus, F., Schulze, R., Eck, J. and Lorenz, P. (2006) Metagenomics: an inexhaustible access to nature's diversity. *Biotechnol. J.* 1: 815–821.

Lawson, P.A., Deutch, C.E., and Collins, M.D. 2008. Phylogenetic characterization of a novel salt-tolerant *Bacillus* species: description of *Bacillus dipsosauri* sp. Nov. *Journal.* of *Applied. Microbiol*. 81: 109-112.

Lee, J.M., Nahm, S.H., Kim, Y.M. and Kim B.D. 2004. Characterization and molecular genetic mapping of microsatellite loci in pepper. *Theor.Appl. Genet.* 108(4): 619-627.

Li, X. and Qin, L. 2005. Metagenomics-based drug discovery and marine microbial diversity. *Trends. Biotechnol*. 23: 539-543.

Lim, H. K., Chung, E.J., Kim, J.C., Choi, G.J., Jang, K.S., Chung, Y.R., Cho, K.L. and Lee, S.W. 2005. Characterization of a forest soil metagenome clone that confers indirubin and indigo production on *Escherichia coli*. *Appl. Environ. Microbiol.* 71: 7768-7777.

Litchfield, C.D and Gillevet, P.M. 2002. Microbial diversity and complexity in hypersaline environments: a preliminary assessment. *J. Ind. Microbiol. Biotechnol.* 28: 48-55.

Lizama, C., Monteoliva, S. M., Prado, B., Ramos, C.A., Weckesser, J and Campos, V. 2001. Taxonomic study of extreme halophilic archaea isolated from the ‘‘Salar de Atacama’’. *Chile. Syst. Appl. Microbiol*. 24: 464-474.

Looney, W.J., Narita, M., Mühlemann, K. 2009. Stenotrophomonas maltophilia: an emerging opportunist human pathogen. *Lancet. Infect. Dis*. 9: 312-320.

Louie, M., Louie, L. and Simor, A.E. 2000. 'The role of DNA amplification technology in the diagnosis of infectious diseases.' *CMAJ.* 163: 301-309.

Lu, J., Domingo, J.S., and Shanks, O.C. 2007. Identification of chicken-specific fecal microbial sequences using a Metagenomic approach. *Water. Res.* 41: 3561–3574.

Ma, Y., Zhang, W., Xue, Y., Zhou, P., Ventosa, A., Grant, W. 2004. Bacterial diversity of the Inner Mongolian Baer Soda Lake as revealed by 16S rRNA gene sequence analyses of *Extremophiles. Chile. Syst. Appl. Microbiol.* 8: 45-51.

Maas, E.V. 1990. Crop salt tolerance. In K.K. Tanji (ed.). *Agricultural Salinity Assessment and management* (pp. 266-304). *New York : American Society of Civil Engineers*.

MacNeil, I.A., Tiong, C.L., Minor, C., August, R.P., Grossman, T.H., Loiacono, K.A., Lynch, B.A., Phillips, T., Narula, S., Sundaramoorthi, R., Tyler, A., Aldredge, T., Long, H., Gilman, M., Holt, D. and Osburne, M.S. 2001. Expression and isolation of antimicrobial small molecules from soil DNA libraries. *J. Mol. Microbiol. Biotechnol*. 3: 301-308.

Majernik A., Gottschalk E. and Daniel R., 2001. Screening of environmental DNA libraries for the presence of genes conferring Na+ (Li+)/H+ antiporter activity on Escherichia coli: characterization of the recovered genes and the corresponding gene products, *J. Bacteriol*. 183: 6645-6653.

Manichanh, C., Rigottier, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R. and Jarrin, C. 2006. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Science.* 5: 205-211.

Marchesi, J. and Shanahan, F. 2007. The normal intestinal microbiota. *Curr. Opin. Infect. Dis.* 20: 508–513.

Martin-Cuadrado, A.B., Lopez-Garcia, P., Alba, J.C., Moreina, D., Strittmatter, A. 2007. Metagenomics of the deep Mediterranean, a warm Bathypelagic habitat. *PLoS. ONE. J.* 2: 914-921.

Mandel, M. and Higa, A. 1970. *J. Mol. Bio*. 53: 159-163.

Menon, M.K. 1954. On the paddy field prawn fishery of Travancore-Cochin and an experiment in prawn culture. *Proc. Indo-Pacific Fish Counc*. 5: 131-135.

Miller, D. N., Bryant J.E., Madsen E.L. and Ghiorse, W.C. 1999. Evaluation and optimization of DNA extraction and purification procedure for soil and sediment samples. *Appl.Environ.Microbiol.* 65: 4715-4724.

Miyashita, H., Ikemoto, H., Kurano, N., Adachi, K., Chihara, M. and Miyachi, S. 1996. Chlorophyll *d* as a major pigment. *Nature.* 383: 402-408.

Muthu, M.S. 1978. A general review of penaeid prawn culture. Summer Institute in Breeding and rearing of marine prawns. *CMFRI Special Publication No.* 3: 25-33.

Myers, E.W., Sutton, G.G., Delcher, A.L., Dew, I.M., Fasulo, D.P., Flanigan, M.J., Kravitz, S.A., Mobarry, C.M. 2000. A whole-genome assembly of Drosophila. *Science.* 287: 2196-2204.

Nielsen, P., Fritze, D. and Priest, F.G. 1995 Phenetic diversity of alkaliphilic *Bacillus* strains: proposal for nine new species. *Microbiol*. 141: 1745-1761.

Ogram, A., Sayler, G. S. and Barkay, T. 1987. The extraction and purification of microbial DNA from sediments. *J. Microb. Methods.* 7: 57-66.

Onishi, H., Fuchi, H., Konomi, K., Hidaka, O. and Kamekura, M. 1980. Isolation and distribution of a variety of halophilic bacteria and their classification by salt-response. *Agric. Biol. Chem.* 44:1253-1258.

Peterson, D.G., Schulze, S.R., Sciara, E.B., Lee, S.A., Bowers, J.E., Nagel, A., Jiang, N., Tibbits, D.C., Wessler, S.R and Paterson, A.H. 2002. Integration of cot analysis, DNA cloning, and high - throughout sequencing facilitates genome characterization and gene discovery. *Genome Res*. 12: 795-807.

Pin, C. L., Rukstalis, J. M., Johnson, C. and Konieczny, S. F. 2001. The transcription factor Mist 1 is required to maintain exocrine pancreas cell organization and acinar cell identity. *J. Cell Bio*. 155: 19-30.

Poly, f., Monrozier, L.J. and Bally, R. 2001. Improvement in the RFLP procedure for studying the diversity of nifH genes in communities of nitrogen fixers in soil. *Res. Microbiol*. 152: 95-103.

Purohit, M.K. and Singh, S.P. 2009. Assessment of various methods for extraction of Metagenomic DNA from saline habitats of coastal Gujarat (India) to explore molecular diversity. *Lett. Appl. Microbiol*. 49(3): 338-344.

Quaiser, A., Ochsenreiter, T., Klenk, H.P., Kletzin, A., Treusch, A.H., Meurer, G., Eck, J., Sensen, C.W. and Schleper, C. 2002. First insight into the genome of an uncultivated crenarchaeote from soil. *Environ. Microboil*. 4: 603–611.

Rajesh, N.R., Alagawadi, A.R. and Krishnaraj, P.U. 2006. *Cry* diversity of *Bacillus thuringiensis* isolates of Western Ghat region and their bioefficacy against *Spodoptera litura* and *Helicoverpa armigera. Entomon.* 31(2): 89-98.

Rappe, M.S. and Giovannoni, S.J. 2003. The uncultured microbial majority. *Annu. Rev. Microbiol.* 57: 369-394.

Rayment, G.E. and Higginson, F.R. 1992, *Australian Laboratory Handbook of Soil and Water Chemical Methods*, Melbourne, *Inkata Press*. (Australian Soil and Land Survey Handbooks, vol. 3).

Reinhold, B., Hurek, T., Niemann, E.G. and Fendrik, I. 1986 Close association of *Azospirillum* and diazotrophic rods with different root zones of Kallar grass. *Appl. Environ. Microbiol.* 52: 520-526.

Richardson, T.H., Xuqiu, T., Gerhard, F., Walter, C., Mark, C., David, L., John, M., Short, J.M., Robertson, D.E. and Miller, C. 2002. A novel, high performance enzyme for starch liquefaction. Discovery and optimization of a low pH, thermostable alpha-amylase. *J. Biol. Chem*. 277: 26501-26507.

Riesenfeld, C.S., Goodman, R.M. and Handelsman, J. 2004. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ. Microbiol.* 6: 981-989.

Rodriguez-Valera, F. 1993. Introduction to saline environments. In R.H. Vreeland and L.I. Hochstein (eds.). The Biology of Halophilic Bacteria (pp. 1-12). Boca Raton: CRC Press.

Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., Loiacono, K.A., Lynch, B.A., MacNeil, I.A., Minor, C., Tiong, C.L., Gilman, M., Osburne, M.S., Clardy, J., Handelsman, J. and Goodman, R.M. 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol*. 66: 2541-2547.

Saiz-jimenez, C. and Laiz, C. 2000. Occurrence of halotolerant/ halophilic bacterial communities in deteriorated monuments. *Int. Biodeterioration and Biodegradation.* 46: 319-326.

Sambrook, J. and Russel, E.F. 2001. *Molecular cloning*: *A Laboratory Manual.* Cold Spring Harbar Laboratory Press, New York, USA. 1541p.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular cloning*: A Laboratory Manual. Second edition. Cold Spring Harbar Laboratory Press, New York, USA. 1322p.

Santosa, D.A. 2001. Rapid extraction and purification of environmental DNA for molecular cloning applications and molecular diversity studies. *Mol. Biotechnol.* 17: 59–64.

Schmalenberger, A., Schwieger, F., Tebbe, C.C. 2001. Effect of primers hybridizing to different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling. *Appl. Environ. Microbiol.* 67: 3557-63.

Schmeisser, C., Helen, S. and Wolfgang, R.S. 2007. Metagenomics, biotechnology with nonculturable microbes. *Appl. Microbiol. Biotechnol.* 75: 955-962.

Schoss, P.P. and Handelsman, J. 2005. Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol*. 6: 229-233.

Sekiguchi, Y. 2006. Yet-to-be cultural microorganisms relevant to methane fermentation processes. *Microbes. Environ.* 21: 1-15.

Sekiguchi, Y., Yamada, T., Hanada, S., Ohashi, A., Harada, H. and Kamagata., Y. 2003. *Anaerolinea thermophila* gen. nov., sp. nov. and *Caldilinea aerophila* gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain Bacteria at the subphylum level. *Int. J. Syst. Evol. Microbiol.* 53: 1843-1851.

Shylaraj, K.S. and Sasidharan, N.K. 2005. A high yielding salinity tolerant rice variety for the coastal saline ecosystems of Kerala. Rice Research Station, Kerala Agricultural University, Vyttila, Cochin. *J. of Tropical Agriculture.* 43 (1-2): 25-28.

Siddhapura, P.K., Vanparia, S., Purohit, M.K., Singh, S.P. 2010. Comparative studies on the extraction of metagenomic DNA from the saline habitats of Coastal Gujarat and Sambhar Lake, Rajasthan (India). *Int. J. Biological Macromol*. 47: 375-379.

Singh, B., Gautam, S.K., Verma, V., Kumar, M. and Singh, B. 2008. Metagenomics in animal gastrointestinal ecosystem. *Potential. biotechnol. prospects. Anaerobe*. 14: 138-140.

Sloan, W.T., Lunn, M., Woodcock, S., Head, I.M., Nee, S. and Curtis, T.P. 2006. Quantifying the role of immigration and chance in shaping prokaryote community structure. *Environ. Microbiol*. 8: 732-740.

Sorokin, I.D. and Kuenen. J. G. 2005. Haloalkaliphilic sulfur-oxidizing bacteria in soda lakes. *FEMS Microbiol. Rev*. 29: 685-702.

Sorokin, I.D., Kravchenko, I. K., Tourova, T. P., Kolganova, T. V., Boulygina, E. S., and Sorokin, D. Y. 2008. *Bacillus alkalidiazotrophicus sp. nov., a diazotrophic*, low salt-tolerant alkaliphile isolated from Mongolian soda soil. *Int. J. of Systematic and Evolutionary Microbiol*. 58: 2459-2464.

Spark, D.L. 1995. Environmental Soil Chemistry. California: Academic Press.

Streit, W.R. and Schmitz, R.A. 2004. Metagenomics – the key to the uncultured microbes. *Curr. Opin. Microbiol*. 7: 492-498.

Sugimoto, A., Bigneil, D.E. and MacDonald, J.A. 2000. Global impact of termites on the carbon cycle and atmospheric trace gases. In Termites: Evolution, Sociality, Symbiosis, Ecology. ed. Abe, T., Bignell, D.E. and Higashi, M. pp. 409–435. Dordrecht: Kluwer Academic.

Sun, Y. and Cheng, J. 2002. Hydrolysis of lignocellulosic material from ethanol production: A review. *Biores. Technol.* 83:1-11.

Susannah, G.T. and Edward, M.R. 2005. Metagenomics: DNA sequencing of environmental samples. Nature Rev. Gen. *6: 805-814.*

Takenaka, T., Tashiro., T. Ozaki., A. Takakubo., H. 2007. Planktonic bacterial population dynamics wih environmental changes in coastal area of Suruga Bay. *Microbes. Environ*. 22: 257-267.

Tindall, B.J. 2004. Prokaryotic diversity in the antarctic: The Tip of the Iceberg. *Microb. Ecol.* 47: 271–283.

Tiquia, S.M., Davis, D., Hadid, H., Kasparian, S., Ismail, M., Sahly, R., Shim, J., Singh, S., and Murray K.S. 2007. Halophilic and halotolerant bacteria from river waters and shallow groundwater along the rouge river of south-eastern Michigan. *Environ. Technol*. 28: 297-307.

Tokuda, G. and Watanabe, H. 2007. Hidden cellulases in termites: revision of an old hypothesis. *Biol. Lett*. 3: 336-339.

Tringe, S.G. and Rubin, E.M. 2004, Metagenomics: Concept, methodology, ecological inference and recent advances. *Nature. Reviews. Genet.* 6: 805-814.

Tringle, S.G., Von Mering, C., Kobuyashi, A., Salamov, A.A and Chen, K. 2005. Comparative metagenomics of microbial communities. *Science.* 308: 554-557.

Tsai, Y.L and Olson, B.H. 1992. Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. *Appl. Environ. Microbiol*. 58: 2292–2295.

Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature.* 438: 37-43.

Ullmann, A., Jacob, F. and Monod, J. 1967. Characteriztion by *in vitro* complementation of a peptide corresponding to an operator- proximal segment of the β galactosidase structural gene of *Escherichia coli. J. Mol. Biol*. 24: 339-341.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L and Rusch, D. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science.* 304: 66–74.

Ventosa, A., Nieto, J.J. and Oren, A. 1998. Biology of moderately halophilic aerobic
bacteria. *Microbiol. Mol. Biol. Rev.* 62(2): 504-544.

Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K.E. and Streit, W.R. 2003. Prospecting for novel biocatalysts in a soil metagenome. *Appl. Environ. Microbiol.* 69: 6235-6242.

Voget, S., Steele, H.L. and Streit, W.R. 2006. Characterization of a metagenome derived halo tolerant cellulase. *J. Biotechnol*. 126: 26-36.

Waisel, Y. 1972. *Biology of Halophytes*. New York: Academic Press. 395p.

Ward, D.M., Bateson, M.M., Weller, R., and Ruff-Roberts, A.L. 1992. Ribosomal RNA analysis of microorganism as the occur in *Advances in Microbial Ecology,* Plenum Press., New York, 12: 219-286.

Ward, D.M., Weller, R. and Bateson, M.M. 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature.* 345: 63-65.

Warnecke, F. and Hugenholtz, P. 2007. Building on basic metagenomics with complementary technologies. *Genome. Biol.* 8: 231-236.

Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M and Richardson, T.H. 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*. 450: 560-565.

Winker, S. and Woese, C.R. 1991. A definition of the domains Archaea, Bacteria and

Eucarya in terms of small subunit ribosomal RNA characteristics. *Syst. Appl. Microbiol*. 14: 305-10.

Wong, K.K.Y. and Saddler, J.N. 1992. Applications of hemicellulases in the food, feed and pulp and paper industries. In Coughlan PP, Hazlewood GP (eds) Hemicellulose and Hemicellulases. *Portland. Press*. London, 40: 127-143.

Xu, J. 2006. Microbial ecology in the age of genomics and metagenomics: concepts, tools and recent advances. *Mol. Ecol.* 15: 1713-1731.

Yoon, J.H., Kang, K.H. and Park, Y.H. 2003. *Halobacillus salinus* sp. nov., isolated from a salt lake on the coast of the East Sea in Korea. *Int. J. Syst. Evol. Microbiol*, 53:687-93.

Yun, J., Seowon, K., Sulhee, P., Hyunjin, Y., Myo- Jeong K., Sunggi, H. and Sangyeol, R., 2004. Characterization of a novel amylolytic enzyme encoded by a gene from a soilderived metagenomic library. *Appl. Environ. Microbiol*. 70(12): 7229-7235.

Zeyaullah, M., Kamli, M.R., Islam, B., Atif, M., Benkhayal, F.A., Nehal, M., Rizvi, M.A. and Arif, A. 2009. Metagenomics An advanced approach for non-cultivable microorganisms. *Biotechnol. and Mol. Biol. Reviews*, 4(3): 49-54.

Zvyagintseva, I.S. and Tarasov, A.L. 1987. Extreme halophilic bacteria from saline soils. *Microbiol.* 56: 839-844.

Appendices

Appendix I

Composition of different media used in the study

1. Luria Bertani (LB) broth

Tryptone	-	10 g
Yeast Extract	-	5 g
NaCl	-	5 g
pH adjusted to	-	7± 0.2
Distilled water	-	to make up to 1000ml

2. Luria Bertani agar medium

Tryptone	-	10 g
Yeast Extract	-	5 g
NaCl	-	5 g
Agar	-	20g
pH adjusted to	-	7± 0.2
Distilled water	-	make up to 1000ml

3. Nutrient agar medium

Peptone	-	5 g
Beef extract	-	3 g
NaCl	-	3 g
Agar	-	15 g
pH adjusted to	-	7
Distilled water	-	make up to 1000ml

(Two egg yolks to be added into the medium under sterile conditions, after autoclaving, when the temperature is around 40 to 50^0C).

4. Soil extract agar medium

Glucose	-	1g
Dipotassium phosphate	-	0.5g
Soil extract	-	100 ml
Agar	-	15g
Distilled water	-	make up to 1000ml

AppendixII

Reagent used for Metagenomic DNA isolation

1. Extraction buffer - 10 ml

A. 100mM Tris HCl

1M Tris HCl (pH - 8.0) - 1ml

Distilled water - 100 ml

B. 100mM EDTA - 0.372g

C. 1.5M NaCl

1M NaCl - 0.75 ml

Distilled water - 100 ml

2. Lysis buffer - 4 ml

A. 20% SDS - 0.8g

B. Lysozyme - 20mg/ ml

C. Proteinase K - 10mg/ml

D. N-lauroyl sarcosine - 10mg/ml

E. 1% CTAB - 4g

4. Phenol: chloroform: isoamylalcohol (25:24:1)

To 25 parts of phenol, 24 parts of chloroform, 1 part of isoamylalcohol was added and mixed properly. The mixture was stored in refrigerator before use.

5. Chloroform: isoamylalcohol (24:1)

To 24 parts of chloroform, 1 part of isoamylalcohol was added and mixed properly. The mixture was stored in refrigerator before use.

6. Potassium acetate 7.5M

Potassium acetate - 20.412 g

Distilled water - 50 ml

7. Chilled ethanol (70%)

To 70 parts of absolute ethanol, 30 parts of double distilled water was add

8.Sterile distilled water - 20–50 µl

Appendix III

1. Reagent used for DNA isolation

1. 10mM Tris HCl

   1M Tris HCl (pH - 8.0) - 1ml

   Distilled water - 100 ml

2. 100mM NaCl

   1M NaCl - 10 ml

   Distilled water - 100 ml

3. TE buffer

   (Tris HCl - 50mM; EDTA - 20mM)

   Tris HCl - 0.05 M (pH - 8.0) - 0.394 g

   0.02M EDTA (pH - 8.0) - 0.372 g

   Distilled water - 100 ml

4. Lysozyme stock

   Lysozyme - 50 mg

   Distilled water - 1 ml

   Stock was prepared by dissolving 50 mg lyzozyme in 1 ml water and was stored under refrigerated conditions.

5. RNase A stock

   RNase A - 10 mg

   Distilled water - 1 ml

   Stock was prepared by dissolving 10 mg RNase An in 1 ml water and was stored under refrigerated conditions at -20^{0}C.

6. 2 % SDS in TE buffer

   SDS - 2 g

   TE buffer - 100 ml

7. Proteinase K

   Proteinase K - 20 mg

   Distilled water - 1 ml

   Stock was prepared by dissolving 20 mg Proteinase K in 1 ml water and stored under refrigerated conditions at -20^{0}C.

8. Phenol: chloroform(1:1 v/v)

   To 1 part of Tris saturated phenol, 1 part of chloroform was added.

9. Chloroform: isoamylalcohol (24:1 v/v)

   To 24 parts of chloroform, 1 part of isoamylalcohol was added and mixed properly. The mixture was stored in refrigerator before use.

10. 3M Sodium acetate

Sodium acetate	-	20.412 g
Distilled water	-	50 ml

11. Chilled isopropanol

12. 70% ethylacohol

To 70 parts of absolute ethanol, 30 parts of double distilled water was added.

Appendix IV

I. Buffers and dyes used in gel electrophoresis

1. 6x Loading/ tracking dye

Bromophenol blue	-	0.25%
Xylene cyanol	-	0.25%
Glycerol	-	30%

The dye was prepared and kept in fridge at 4^0C

2. Ethidium bromide (intercalating dye)

The dye was prepared as a stock solution of 10 mg/ ml in water and was stored at room temperature in a dark bottle.

3. 50x TAE buffer (pH 8.0)

Tris base	-	242.0 g
Glacial acetic acid	-	57.1 ml
0.5M EDTA (pH 8.0)	-	100 ml
Distilled water	-	1000 ml

The solution was prepared and stored at room temperature

II. Reagents used for plasmid isolation

1. A. Solution I (SET)

Sucrose	-	25%
Tris HCl	-	50mM
EDTA	-	50mM
pH	-	8.0

B. 5mg of lysozyme per ml of SET

2. Solution II (Lysis buffer)

NaOH	-	0.2 N
SDS	-	1 %

3. Solution III

CH_3COOK	-	3M
pH	-	5.5

IV. Reagent used for competent cell preparation

Solution A

Ice - cold 100mM $CaCl_2$

Abstract

METAGENOMIC APPROACH TO ASSESS DIVERSITY OF BACTERIAL COMMUNITY IN SALINE *Pokkali* HABITATS OF KERALA

By

SARVESHWAR SAH

ABSTRACT OF THE THESIS

Submitted in partial fulfilment of the

requirement for the degree of

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

Centre for Plant Biotechnology and Molecular Biology

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR - 680 656

KERALA, INDIA

2012

ABSTRACT

The use of traditional microbiological culturing methods for the study of microbes has its limitations. It has been estimated that 99 per cent of microbes cannot be cultured easily. Over the past decade, "Metagenomics," which is culture-independent genomic analysis of microbes, has been developed to overcome the drawbacks in culture-based analysis of microbial communities. Soil is considered to be a complex environment and a major reservoir of microbial genetic diversity. Metagenomics is a powerful tool in the analysis of diversity of microbial communities in specific environments.

An attempt was made to analyse the diversity of bacterial community in the acid saline *Pokkali* soil from Vyttila, through Metagenomic approach. Environmental DNA was isolated from the soil by direct method. 16S rDNA which encodes 16S rRNA found in the 30S subunit of prokaryotic ribosome is used as the barcode for bacterial identification. 1500bp fragment of 16S rDNA was amplified using specific primers through polymerase chain reaction (PCR). The amplicons were ligated in plasmid vector pGEMT and cloned in *E. coli*, to construct a Metagenomic library. Sequence analysis of 30 randomly picked colonies revealed that 23 clones (76.7%) were of unculturable bacteria and 7 clones (23.3%) were culturable. The most predominant phylum was Proteobacteria, which included Ectothiorhodospiraceae, *Azospira* sp., *Stenotrophomonas maltophilia, Thiobacillus prosperus, Levilinea saccharolytica*, Desulfobacteraceae, *Thioalkalivibrio* sp. and Rhodocylales. Other phyla included Chloroflexi, Acidobacteria and Cynobacteria. A range of bacteria including acidophiles, extreme halophiles, denitrifier, S oxidizers, sulphate reducers, aerobes, strict anaerobes, thermophiles, mesophiles, photosynthetic bacteria and even human pathogens were obtained, of the 23 unculturable bacteria detected, six did not show homology with any bacterial sequence available in NCBI database, indicating the possibility that these could be novel. The phylogenetic

tree based on partial 16S rRNA gene placed 30 clones from *Pokkali* soil in three major cluster- Proteobacteria, Chloroflexi and Acidobacteria.

Three predominant bacteria isolated by the traditional method belonged to the genus *Bacillus*.

Metagenomic approach successfully revealed the composition and diversity of bacterial community in *Pokkali* soil. A function-derived strategy could be used for bioprospecting of gene related to acid and salt tolerance genes.