METAGENOMIC CHARACTERIZATION OF DISEASE SUPPRESSIVE SOILS

MAITREYEE SARMA

DEPARTMENT OF BIOTECHNOLOGY
COLLEGE OF AGRICULTURE, DHARWAD
UNIVERSITY OF AGRICULTURAL SCIENCES,
DHARWAD – 580 005

JUNE, 2015

METAGENOMIC CHARACTERIZATION OF DISEASE SUPPRESSIVE SOILS

Thesis submitted to the
University of Agricultural Sciences, Dharwad
In partial fulfillment of the requirements for the
Degree of

MASTER OF SCIENCE (AGRICULTURE)

in

MOLECULAR BIOLOGY AND BIOTECHNOLOGY

By

MAITREYEE SARMA

DEPARTMENT OF BIOTECHNOLOGY
COLLEGE OF AGRICULTURE, DHARWAD
UNIVERSITY OF AGRICULTURAL SCIENCES,
DHARWAD – 580 005

JUNE, 2015

DEPARTMENT OF BIOTECHNOLOGY
COLLEGE OF AGRICULTURE, DHARWAD
UNIVERSITY OF AGRICULTURAL SCIENCES, DHARWAD – 580 005

CERTIFICATE

This is to certify that the thesis entitled "**METAGENOMIC CHARACTERIZATION OF DISEASE SUPPRESSIVE SOILS**" submitted by **Miss MAITREYEE SARMA,** for the degree of **MASTER OF SCIENCE (AGRICULTURE)** in **MOLECULAR BIOLOGY AND BIOTECHNOLOGY,** to the University of Agricultural Sciences, Dharwad is a record of research work carried out by her during the period of her study in this University, under my guidance and supervision and the thesis has not previously formed the basis for the award of any degree, diploma, associateship, fellowship or other similar titles.

DHARWAD
JUNE, 2015

(H.M. VAMADEVAIAH)
CHAIRMAN

Approved by :

Chairman : _____
(H. M. VAMADEVAIAH)

Members : **1.** _____
(SUMANGALA BHAT)

2. _____
(S. T. NAIK)

3. _____
(GEETA GOUDAR)

Affectionately Dedicated
to My Beloved
Parents

ACKNOWLEDGEMENT

The successful completion of my thesis is largely due to timely support, help and suggestions I have received from several persons. At the very outset, I take this opportunity to express my gratitude to them for their contributions throughout this project.

First and foremost, I would like to thank **Dr. H. M. Vamadevaiah** (Principal Scientist and Head, ARS, Hebballi Farm, UAS, Dharwad), the chairman of my advisory committee for his direction and guidance, towards the conduction of this project. I would also like to thank the other committee members **Dr. Sumangala Bhat** (Professor Department of Agricultural Biotechnology, UAS, Dharwad), **Dr. S. T. Naik** (Professor Department of Plant Pathology, UAS, Dharwad) and **Dr. Geeta Goudar** (Asst. Professor, Department of Microbiology, College of Vijayapur, UAS, Dharwad) for their guidance and specific suggestions that has helped me to continually improve my research work.

I must express my gratitude to **Dr. P.U. Krishnaraj** (Professor and Head, Department of Microbiology, College of Vijayapur, UAS, Dharwad) the scientific supervisor for my project, who has guided me with his valuable insights, and helped me immensely to complete the thesis in a timely manner. I sincerely thank the Head of the department, **Dr. I.S. Kategeri,** and other teachers of the department (**Dr. Prashanthi S.K., Dr. Ramesh Bhat, Dr. Narayan Moger**) , who has taught me the theoretical basis of my subject.

I thank my parents and my siblings wholeheartedly for their steadfast moral support through thick and thin, and for being the pillars of my strength throughout this difficult endeavour. I thank God Almighty, without whose blessings this work wouldn't have been completed.

Staff and coworkers in my laboratory (MB6, IABT), my department seniors and other students in the department have all been immensely instrumental in helping me carry out my research work and achieve real results.

I also thank my hostel-mates in B3 quarter, friends and well-wishers in Dharwad and elsewhere, for their firm support, love and care throughout the period of my work in the university.

Last but not the least, I would like to thank Mr. Siddhu from Society Xerox Centre for undertaking the formidable task of formatting my thesis.

DHARWAD
JUNE, 2015 **(Maitreyee Sarma)**

C O N T E N T S

Contd…..

LIST OF ABBREVIATIONS

Bp- Base pair

PCR- Plymerase Chain Reaction

DGGE- Denaturing Gradient Gradient Gel Electrophoresis

EDTA- Ethyl diamine tetra acetic acid

EtBr- Ethidium Bromide

g- Gram

h- Hour

HCl Hydrochloric acid

LB- Luria bertani

M Molar

min- Minute

mM- Millimolar

N Normal

Na Sodium

PCR Polymerase chain reaction

pH Hydrogen ion concentration

r DNA- Ribosomal Deoxyribose Nucleic Acid

r RNA- Ribosomal Ribose Nucleic Acid

rpm Rotations per minute

SDS Sodium dodecyl sulphate

TAE Tris acetate EDTA buffer

UV Ultraviolet

% Percent

µg Microgram

µl Microlitre

LIST OF TABLES

LIST OF PLATES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Microorganisms are the most abundant and diverse group of organisms inhabiting the earth. Their diversity makes them adaptable to almost every ecosystem present on this planet starting from polar ice caps to the intertidal hot springs. Even though small, their genomes carry huge potential to allow them to survive, to accumulate mutations aiding better adaptability than any other species present on the planet. This abundance of microorganisms also raises a question regarding their utility. The answer to this question was first realized only towards the end of the 17^{th} century. And ever since then, this enormous potential stored within the genomes of diverse microorganisms have been exploited by mankind for their use in various sectors like food (Probiotics) industry, textile industry, medicine, agriculture, etc.

The use of these microorganisms in the agricultural sector mostly involves the maintenance of soil fertility and proper disease management strategies to get good quality and high crop productivity. The use of best agronomic practices from the selection of high yielding cultivar to the use of the beneficial microorganisms is considered as some of the important variables to be taken under consideration to achieve the same. Traditionally, these beneficial microorganisms were identified by culturing them on general and selective enrichment artificial media containing essential nutrients alongwith agar required for their proper growth and development. These culture based studies helped in identifying many beneficial microorganisms which were found to be involved in nutrient recycling, managing phytopathogens and bio-remediating xenobiotics present in the soil.

A global recent estimate suggested that the soil contains around 4-5 X 10^{30} microbial cells (excluding viruses) (Singh, *et al.*, 2009). With the advent of ribosomal RNA (rRNA) sequencing technique, it proved the phenomenon of the 'Great plate anomaly' (not all bacteria can grow on culture medium) and suggested that only a minority of the microorganisms i.e. 0.1- 1% are readily culturable from most habitats. This fact had fueled the interest of the researchers to develop techniques to explore the rest 99 - 99.99% of the microorganisms which lie hidden till date. Therefore, the understanding of such hidden microbial community structure will not only allow one to study the great reservoir of genes

and their functions from previously inaccessible microbes, but also may open up exciting new possibilities in the form of discoveries (like enzymes, etc.) for the development of novel products, for the biotechnological and pharmaceutical industry.

Metagenomics is one such emerging field of environmental microbiology, which studies the community structure of microorganisms by utilizing the tools and techniques of molecular biology. The term was first coined by Jo Handelsman and others at the University of Wisconsin, Department of Plant Pathology in 1982 (Handelsman, 2004). This term was derived from the statistical concept of *meta*-analysis (the process of statistically combining seperate analyses) and genomics. Metagenomics was defined as the genomic analysis of an assemblage of microorganisms, which provides a window to the riches of the uncultured world. This technique mostly concerns the extraction, cloning and analysis of the entire genetic complement of a habitat with a goal to understand the global microbial ecology more clearly on one side and on the other side it has been driven by the increasing biotechnological demands for new enzymes and biomolecules.

The recent advancement in the DNA based molecular biology techniques like PCR, DGGE, Primer walking, Pyrosequencing alongwith the use of computer aided bioinformatics software have made the analysis of the data generated out of the metagenome to be analyzed in a proper fashion. As a result, some of the novel chitinase genes have been discovered and reported recently from soil samples (Hjort, *et al*., 2013). Even though, metagenomics requires extra time and effort to find something new it is the only answer till date to know about the uncultured world and to reveal some of the mysteries of the present day world.

One such mystery is the so called ‘Disease Suppressive soils’ in agriculture. These disease or pathogen suppressive soils have been defined as “soils in which the pathogen does not establish or persist, establishes but causes little or no damage, or establishes and causes disease for a little while but thereafter the disease is less important, although the pathogen may persist in the soil” (Borneman and Becker, 2007). This definition takes into consideration the presence of all the 3 components of the normal disease triangle, i.e. virulent pathogen, susceptible host and a conducive environment but still the disease is

either suppressed or its incidence is quite low. The traditional microbiological studies have indicated that such soils may commonly harbour antibiotic resistance genes against a broad array of antibiotics which have been found to be associated with the suppression of plant pathogens. These antibiotic resistance genes or gene clusters are collectively referred to as the 'antibiotic resistome'. The knowledge and the identification of the potential community of such microorganisms harboring such antibiotic resistome will alter the present disease management strategy with the use of pesticides; reducing which will in turn control the soil and water pollution. The soil microbiota are also known to contain a large pool of genes that encode enzymes involved in either biosynthetic or biodegradation processes, including the degradation of xenobiotics. There is a vast amount of information held within the genomes of such uncultured microorganisms, and metagenomics is one of the key emerging techniques which can be used to access and investigate this potential.

Therefore, such culture independent studies like metagenomics of disease suppressive soils need to be conducted, to overcome the limitations posed by the traditional microbiological approaches. This study will not only be helpful in understanding the microbial community composition and structure of the disease suppressive soils but may also indicate how diverse are such soils from the normal non-suppressive soils and may even provide a key to the control of soil borne plant pathogens.

Taking these perspectives under consideration, the following objectives were set for the research investigation-

1. Standardization of protocol for isolation of high molecular weight microbial DNA from disease suppressive soils.

2. Analysis of the microbial diversity present in the soil (disease suppressive) samples.

3. Screening for the presence of agriculturally important genes in the soil samples.

REVIEW OF LITERATURE

Soil is a natural resource utilized by the human society since ancient times to produce food through agriculture for their sustenance. Ever since then, human civilization is dependent on the soil for food production. Therefore, to meet the demand of the ever increasing population there was a continous use of soil amendments in the form of chemical fertilizer and pesticides (for management of pests and diseases) to boost agricultural crop productivity. This in turn had reduced the soil health status in terms of fertility due to the persistence of these chemicals in the soil. So, the main aim of present day agriculture is to sustain the present day crop production by maintaining the soil heath.

Soil basically comprises of two main elements- abiotic (sand, clay, silt) and biotic (microorganisms). Soil biotic element plays a crucial role in determining the health of the soil. It is also reported to harbor largest microbial diversity per unit mass or volume (Torsvik *et al*., 1990). These microflora present in the soil are known to affect the fertility status of the soil by making essential nutrients available to the crop plants (N_2 fixers, Phosphate solubilizers, etc.) and in some cases they are known to suppress many soil borne plant diseases (disease suppressive soils). Therefore, any changes in the soil affect the structure and composition of the microbial community present in it. For the reasons mentioned herein the present day researchers are interested in unraveling the potential of these microbial communities which may serve as environmentally safe way to sustain the agriculture crop production without hampering the soil heath status.

2.1 Disease Suppressive Soils

Disease Suppressive soils are known to mankind for over 100 years and the mechanism by which disease suppression is brought about has been a subject of study for the last few decades (Mazzola, 2002). Soils suppressive to soil borne plant diseases have been defined as those in which disease development is minimal even in the presence of a virulent pathogen and a susceptible host. Different evidences clearly points out as to how disease suppression is brought about directly or indirectly by the activity of a specific or a group of microorganisms in relation to the abiotic components of the soil. When suppressiveness has a biological origin, the identification of the causal organism becomes crucial for realizing its potential (Borneman and Becker, 2007). The knowledge of such microbes can be then used to devise effective and sustainable plant

management strategies through application of these organisms or by manipulating the soil environment by using different agronomic practices to alter their population densities.

These suppressive soils have been reported for a number of plant diseases including those caused by the cyst nematode Heterodera spp. (Kerry, 1988; Westphal and Becker, 1999), the bacterium *Streptomyces scabies* (Menzies, 1959), the fungi *Fusarium oxysporum* (Stotzky and Martin 1963; Scher and Baker, 1980), *Gaeumannomyces graminis* var. *tritici* (Cook and Rovira, 1976), *Phytophthora cinnamoni* (Broadbent and Baker, 1974), *Plasmodiophora brassicae* (Murakami *et al*., 2000), *Pythium* spp. (Hancock, 1977) and *Rhizoctonia solani* (Henis *et al*., 1978, 1979) (Mazzola, 2002). Among them also the soils suppressive to Fusarium wilts caused by *Fusarium oxysporum* and take all of wheat caused by *G. graminis* var. *tritici* have been studied extensively.

The initial studies on these soils showed the presence of a specific group of microorganisms which enhances the suppressiveness characteristics of the soil. The presence of large population of saprophytic non-pathogenic *Fusarium* spp. in the suppressive soils infected with Fusarium wilts is a classic example of such soils. Also, the activity of these groups of microorganisms in conferring disease suppressiveness was proved by introducing a small inoculum of the disease suppressive soil in the heat treated same soil (Rouxel *et al*., 1979). Whereas, Kloepper *et al*., (1980) reported the presence of a bacterial fluorescent Pseudomonad community in the same soil alongwith the non-pathogenic *Fusarium* spp. Therefore, disease suppression appeared to function via direct antagonism through competition for iron via production of siderophores (Raaijmakers *et al*.,1995), direct inhibition of saprophytic fungal growth (Duijff *et al*., 1999) and by induced systemic resistance (Lemanceau *et al*., 1993; Duijff *et al*., 1998; Mazzola, 2002).

In case of Take-all Decline disease of wheat, the reduction in disease severity and induction of soil suppressiveness in response to continous wheat monoculture is attributed to an increased populations and activity of certain antibiotic producing fluorescent Pseudomonads or parasitism by *Trichoderma* spp. (Cook and Rovira, 1976, Simon and Sivasithamparam, 1989; Raaijmakers *et al*., 1997, 1999; Mazzola, 2002). Production of antibiotics including phenazine-1-carboxylic acid, 2,4

diacetylphloroglucinol (2,4-DAPG), pyoluteorin and pyrrolnitrin play an important role in the biological control of soil borne pathogens by certain strains of fluorescent *Pseudomonas* spp. that produce these antibiotics (Thomashow and Weller 1988; Keel *et al*., 1992; Kraus and Loper, 1995). Disease suppression was eliminated by pasteurization and 2,4-DAPG producing pseudomonads were not recovered from the rhizosphere of wheat grown in this soil. Introduction of a 2,4-DAPG- producing strain into a conducive soil provided suppression of Take-all decline soil providing first biochemical evidence to support the role of these bacteria in disease suppression (Raaijmakers *et al*., 1999; Mazzola, 2002).

2.2 Metagenomics

The difficulties associated with the in situ identification, isolation and cultivation of microorganisms from soil have provided the impetus to develop molecular biological methods for the analysis of soil-borne microbial communities (Kowalchuk, *et al*, 2006). In addition, the vast complexity of soil systems impedes steps to understand the structure and functioning of communities in this habitat. Indeed, even with the most elaborate and exhaustive molecular methods, the vast diversity of organisms inhabiting soil (Torsvik *et al*. 1990) preclude our ability to characterize all the organisms in that particular habitat completely.

The traditional culture dependent techniques are known to culture microorganisms whose nutritional and physiological requirements for growth and development can be met only under stable laboratory conditions. A vast majority of enzymes and anti-microbial products used in different sectors of health and industry have been isolated from microorganisms that can be cultured *in vitro* under laboratory conditions with the use of variety of supplements to promote their growth and development in artificial media. According to the recent literature, it has been shown that only 0.1-1% of the microbes known till date, can be cultured readily on an artificial medium leaving the rest uncultured (Torsvik, *et al*, 1990). Therefore, it is evident that the microbial world needs to be more explored to find the hidden treasures in the form of novel metabolites, bioactive compounds from different habitats including soil, sediments, marine, terrestrial, sub-surfaces and other niches.

This realization that most microorganisms cannot be grown on pure culture led to the discovery of a new era of culture independent genomic analysis of the unculturable world. This culture independent genomic analyses of microbial communities called as 'Metagenomics' is used to address the challenge of studying microbes that are yet hidden from the curious eyes of the microbiologists (Handelsman, 2004). The aim of this new field of research is to elucidate the genomes of uncultured microbes whose potential role in the ecosystem is still unknown and tries to characterize it based on signatures of molecular biology. The vast amount of information that is held within the uncultured genomes have given the molecular biologists to look for novel molecules for enzyme and drug discovery for its potential use in health and industry including biotechnology.

2.3 Metagenomics of Disease Suppressive soils

With the aim to access and examine the phytopathogen suppressive soils, 7 European Laboratories ran a project named METACONTROL from 2002 (van Elsas, 2008). The aim of the project was to unravel the reservoirs of genetic loci involved in antibiosis (such as those involved in the production of polyketides and chitinases) present in the microbiota of suppressive soils. During the project, metagenomic libraries were constructed for 4 suppressive soil samples and a control soil sample. Both molecular and functional screening of metagenomic library constructed for the 4 disease suppressive soil samples was done. Functional screening was done using a dual culture assay which yielded only one positive clone which is involved in polyketide biosynthesis (PKS1). Whereas, molecular screening done using PCR amplification methods (with degenerate primer pairs targeting the polyketide biosynthesis operon, PKS1) followed by hybridization using probes specific to the gene of interest yielded the detection of 39 unique PKS sequences which was confirmed by endosequencing. The positive clones were then transferred to the Streptomyces host for screening of the antagonistic activity of the clones against *Bacillus subtilis* 1A72, *Staphylococcus aureus* 21, *Enterococcus faecalis* 40, *Neurospora crassa* HK, etc. The clones exhibited 56%, 13%, 4%, <1% antimicrobial activity against the above 4 pathogens respectively.

2.4 Isolation or extraction of high molecular weight microbial community DNA from soil

The metagenomic approach at first relies on the extraction of high molecular weight unsheared microbial DNA which is representative of the original microbial population. When isolating DNA from the environment there are 3 major problems that need to be taken into consideration which include-

(i). The DNA should be extracted from as broad as a range of microorganisms as possible so that it is representative of the original microbial population.

(ii). Shearing of DNA has to be avoided, during the extraction procedure because high molecular weight DNA is required for suitable community analysis as smaller DNA fragments may lead to formation of chimeric products.

(iii).The DNA must be free from contaminating substances which interfere with downstream DNA processing such as restriction and ligation. (Schmeisser *et al*., 2007).

Therefore, the success of any metagenomics based study is highly dependent on the soil DNA extraction method used. The high degree of cell release from soil particles or in situ cell lysis efficiency, good quality and quantity of pure DNA obtained are considered as some of the important parameters that need to be taken under consideration while developing any soil DNA extraction protocol.

Soil in particular is considered as a difficult environment to work with due to its physicochemical properties that lead to an uneven distribution of microorganisms inside the soil matrix (Lombard *et al*., 2011). Another major drawback is the presence of complex substances like humic acids and fulvic acids commonly found in soil which is known to inhibit the activity of enzymes used in molecular biology for restriction digestion, polymerase enzymes for PCR, interferes DNA hybridizations used for detection purposes, etc. (Tsai and Olson, 1992). The phenolic groups in humic acids denature biological molecules by binding to amides or are oxidized to form a quinone which covalently bonds to DNA (Young *et al*., 1993). Different methods have been tried by many researchers to get pure quality DNA free from humic acids without affecting the final yield of DNA. But, not a single protocol can be considered as universal due to the different extraction biases faced in each protocol.

Therefore, for any soil DNA extraction, the method should be selected by taking into account its subsequent metagenomic exploration. If an exhaustive representation of the microbiota of the sampled soil is required, understanding the inherent extraction biases will help in defining the optimal DNA extraction method (Lombard *et al*., 2011). There are mainly two main approaches for extracting nucleic acids from soil samples. The first approach initially separates the cell fraction from soil followed by the lysis of bacterial fraction and the purification of nucleic acids. In the second approach, the extraction of nucleic acids from soil samples is done after in situ cell lysis which is followed by DNA purification.

2.1.1 Indirect DNA extraction from soil

The first paper on the extraction of DNA from soil samples was published by Torsvik *et al*., in 1980, obtaining nucleic acids from environmental matrices that are suitable for molecular analysis. This step indirectly opened the doors for the development of the concept of metagenomics. The pioneering works of Torsvik *et al*. (1990) and Holben *et al*., (1988) showed that the microbial cells are recovered before the cells are lysed and the DNA is extracted and purified. This indirect DNA extraction approach might preferably be used when problematic environmental matrices are to be analyzed or when cloning of large DNA fragments from soil or sediment where high proportion of DNA is of bacterial origin. The recovery of the bacterial fraction from the soil or sediments commonly involves repeated homogenization and the differential centrifugation steps as suggested by Faegri *et al*. (1977). However, protocols differ considerably with respect to the solutions used to break up soil colloids and dislodge surface attached soils that adhere to surfaces by various bonding mechanisms such as polymers, electrostatic forces and water bridging and with different strengths. Homogenization is usually achieved by shaking suspensions with gravel or blending steps. Although a complete dislodgement of the cells seems to be impossible, it is important that cells that are bound to the surface with different degrees of strength are released with similar efficiency. A clear advantage of the indirect approach is that the nucleic acids recovered are less contaminated with the co-extracted humic acids and DNA of non-bacterial origin.

2.1.2. Direct DNA extraction from soil

The second approach developed by Ogram *et al*., (1987) was based on direct or in situ lysis of microbial cells in the presence of the environmental matrix (e.g., soils, sediments, or plant material), followed by separation of nucleic acids from matrix components and the cell debris, which is by far the most frequently used approach (Trevors *et al*., 1992). The advantage of the direct nucleic acid extraction approach is that it is less time consuming as more samples can be processed in a shorter time and a higher DNA yield is achieved (Steffan *et al*., 1988). However, directly extracted DNA often contains considerable proportion of DNA that might originate from non-bacterial sources or free DNA as a less biased recovery of cells that are easily dislodged during cell separation procedure is avoided during this approach (Miller, *et al*., 1999). But, a major drawback of direct lysis methods is that the lysis treatments result in the co-extraction of the organic soil components, particularly humic acids.

2.1.2.1 Cell lysis

The efficient disruption of the bacterial and fungal cell walls is crucial for the recovery of the representative DNA that reflects the genomes of microbes present in an environmental sample and their relative abundance (More *et al*., 1994; Miller *et al*., 1999). Cell lysis can be achieved by mechanical cell disruption and by enzymatic or chemical disintegration of cell walls, and a combination of these methods are often used. To extract DNA, microorganisms are suspended in an alkaline buffer and are lysed either mechanically (like bead beating, sonication, freeze-thawing) or chemically (using chemicals such as SDS, phenol, various detergents and enzymes like lysozyme, proteinase K) or a combination of above treatments.

Physical treatments are known to destroy soil structure thereby providing a greater access to the whole bacterial community including the bacteria present deep within the soil microaggregates (Robe *et al*., 2003). Also, the use of dried, grounded and sieved soil samples is known to greatly improve the lysis efficiency (Frostegard, *et al.,* 1999). In general, these physical methods have shown efficient disruption of cells but they often result in significant DNA shearing.

Chemical lysis alone or in combination with physical methods is used extensively. The most common chemical is the detergent sodium dodecyl sulfate (SDS) which dissolves the hydrophobic material of the cell membranes. Detergents have often been used in combination with heat treatment and with chelating agents such as EDTA, Chelex 100 and diverse Tris and Sodium phosphate buffers. Uses of each of the buffers alone or in combination for DNA extractions have pros and cons of their own. So, the choice of the buffer is a compromise between the expected DNA quantity and the required DNA purity.

Addition of some Cetyltrimethyl-ammonium bromide (CTAB) and polyvinylpolypyrrolidone (PVPP) to the DNA extraction buffers have known to partially remove humic acids. Another denaturing agent guanidine isothiocyanate has also been used for mRNA extraction from seeded soils. Also, the use of enzymes like lysozyme (for cell membrane lysis), proteinase K (for removal of contaminating proteins), achromopeptidase (to improve the lysis of recalcitrant Gram positive *Frankia*) in combination with chemicals have been used to improve the DNA lysis and extraction.

The use of multivalent cations has been a standard method for removing suspended organic solids during the purification of drinking water; a process commonly referred to as chemical flocculation. In theory, these chemistries could remove organic inhibitors via flocculation during sample lysis and homogenization, potentially eliminating or reducing the need for additional treatments and the associated sample loss. Braid *et al.,* (2003) investigated chemical flocculation using multivalent cations as a potential method for eliminating soil-based inhibitors during the extraction process. The solutions of aluminium ammonium sulphate ($AlNH_4(S0_4)_2$), calcium chloride ($CaCl_2$), ferric chloride ($FeCl_3$) and magnesium chloride ($MgCl_2$) during extraction showed significant reduction in the co-purification of inhibitors with minimal loss of DNA yield.

2.1.2.2 DNA extraction and purification

Different methods for separating and purifying nucleic acids from soil components have been investigated by many researchers. In most studies, following cell lysis, a first purification step of DNA is achieved by organic solvent extraction using either phenol or chloroform, followed by ethanol, isopropanol or polyethylene glycol assisted precipitation. The crude extracts of DNA binding on to the hydroxyapatite columns were successful in extracting DNA from lysed soil and sediment samples

(Ogram *et al*., 1987). Cesium chloride density gradient centrifugation was shown to allow further enzymatic restriction of purified DNA. Steffan *et al*., (1988) noted that even extensive purification with CsCl gradient centrifugation and hydroxyapatite chromatography resulted in DNA loss and did not remove all humic acids.

However, the most recent extraction and purification procedures have avoided laborious techniques by using programmes based on one or combination of steps including size exclusion chromatography (gel filtration) with Sephadex spin-columns (Tsai and Olson, 1992; Flemming *et al*., 1994) and commercial ion exchange chromatography columns (Tebbe and Vahjen, 1993; More *et al*., 1994) and agarose gel electrophoresis (Knaebel and Crawford, 1995). Water soluble polyvinylpyrrolidone (PVP) was incorporated in agarose gels to prevent the migration of humic materials with crude DNA (Young *et al*.,1993). PVP or PVPP removes those humic acids from phenolic groups from crude DNA extracts via hydrogen bonding and the formation of PVP-phenolic complex.

The use of other commercial purification products were also reported to give good results such as Wizard DNA clean-up system (Promega) (Henne *et al*., 1999) and CentriconTM 50 and Microcon TM100 concentrators (Amicon) (Zhou *et al*., 1996), ElutipTM D column from Schleicher and Schull (Degrange and Burdin, 1995; Frostegard *et al*., 1999) silica based DNA binding SpinBind columns from FMC Bioproducts (Miller *et al*., 1999) and Tip-100 and Tip-500 columns from Qiagen (Hurt *et al*, 2001; Tebbe and Vahjen, 1993) respectively.

2.1.3 PCR amplification of extracted soil DNA

The ability to amplify specific target sequences from complex nucleic acid mixtures with the aid of PCR has revolutionized microbial ecology. PCR technologies, with the aid of appropriate primers, have made it possible for researchers to investigate rapidly genes of interest in the environment without the need to culture the organisms that harbour these genes (Kowalchuk *et al*., 2006).

The targeting of either RNA or DNA in microbial community analyses has important implications not only in the interpretation of results, but also in terms of methodological constraints. RNA is less stable than DNA, both in soil systems as well as during laboratory manipulation. In general, it is assumed that target DNA levels

provide information concerning gene numbers in a sample, while RNA levels yield information relevant to gene activity. rRNA levels have been used as an indicator of general cell activity (Poulsen *et al*., 1993; Wagner *et al*., 1994), but the correlation between rRNA content and cell activity breaks down at low growth rates, which are typical of soil systems. Indeed, many soil-borne organisms have developed a strategy of maintaining high ribosome levels in times of dormancy so that they can become immediately active when environmental conditions dictate. Thus, rRNA-based studies provide a different perspective than rDNA-based ones (Duineveld *et al*., 2001).

The information present in nucleic acids of microorganisms can be used to investigate and compare diversity at different organization levels, ranging from variability within species to diversity of communities (Johnsen *et al*., 2001). Among nucleic acid-based approaches in metagenomics, most methods currently used rely in one way or another on the polymerase chain reaction (PCR) (Albuquerque *et al*., 2009). PCR conditions needs to be optimized before going for analysis of new microbial community and when new primers are used. Most commonly optimized variables are primer annealing temperature, concentration of polymerase enzyme, cofactor Mg ions and template concentration (Boleda *et al.,* 1996). Nucleic acid quantification is one of the most important steps before performing PCR so as to ensure to obtain amplification after PCR. Dilution of template can be effective in reducing concentrations of contaminating substances which inhibits PCR mostly in case of soil DNA. Modified methods like touchdown PCR (Muyzer *et al*., 1993) and nested PCR (Nakatsu *et al*., 2000) can be used to increase specificity and efficiency of PCR.

2.2 Analysis of Microbial diversity in the disease suppressive soils

Profiling methods provide a more cursory inspection of the community, and allow simultaneous examination of multiple community members, which makes them well suited for the analysis of multiple environmental samples. The most commonly used community profiling methods include denaturing gradient gel electrophoresis (DGGE), TGGE (temperature gradient gel electrophoresis) and TTGE (temporal temperature gradient gel electrophoresis), terminal restriction fragment length polymorphism (T-RFLP) and single-strand conformational polymorphism (SSCP), but there are numerous other possibilities (Johnsen *et al*., 2001; Stephen and Kowalchuk, 2002). In all these methods, species-specific traits are used to separate individual populations from

complex mixtures of PCR products to provide a community fingerprint. These fingerprints can be compared across samples, and the development and application of statistical analysis methods had greatly increased the robustness of such comparisons (Fromin *et al*., 2002).

Ribosomal RNA genes, especially the small subunit (SSU) ribosomal RNA genes, have become the targets of choice for microbial ecologists involved in the *in situ* characterization of soil-borne microbial communities. A number of properties of this marker have made it highly useful in such studies: (1) Its ubiquitous presence in all known organisms, and its lack of horizontal gene transfer between organisms (Sneath, 1993); (2) conserved secondary structure, facilitating sequence alignment (Van de Peer *et al*., 1999); (3) presence of both highly conserved and variable regions, whose sequences are shared between broad and very narrow taxonomic groupings, respectively; (4) the high copy number of SSU rRNAs in the cell, facilitating detection; (5) appropriate size to contain large amounts of phylogenetic information; and (6) the exponentially expanding database of SSU rRNA gene sequences available for comparison.

2.2.1 DGGE

Denaturing gradient gel electrophoresis (DGGE) is technique used to separate DNA fragments (PCR products) of the same length in polyacrylamide gels of increasing concentration of denaturant for separation of bands based on their sequence composition (Fischer and Lerman, 1979, 1983; Myers *et al*., 1987). The medical research community first used DGGE for the identification of gene mutations (Borresen *et al*., 1988; Hovig *et al*., 1991).

Separation is based on the decreased electrophoretic mobility of a partially melted double stranded DNA molecule in polyacrylamide gels containing a linear gradient of DNA denaturants (a mixture of urea and formamide) or a linear temperature gradient. The melting of DNA fragments proceeds in discrete so called *melting domains:* stretches of base pairs with an identical melting temperature. Once a domain with the lowest melting temperature reaches its melting temperature (*Tm*) at a particular position in the denaturing or gradient gel, a transition of a helical to a partially melted molecule occurs, and migration of the molecule will practically halt. Sequence variation within

such domains causes the melting temperatures to differ, and molecules with different sequences will stop migrating at different positions in the gel.

By using DGGE, 50% of the sequence variants can be detected in DNA fragments up to 500 bp (Myers *et al*., 1985). This percentage can be increased to nearly 100% by the attachment of a GC-rich sequence, a so called GC-clamp, to one side of the DNA fragment (Myers *et al*., 1985; Sheffield *et al*., 1989). A sequence of guanines (G) and cytosines (C) is added to the 5' end of one of the PCR primers, co-amplified and thus introduced into the amplified DNA fragments (Sheffield *et al*., 1989; Sheffield *et al*., 1992). The GC-rich sequence acts as a high melting domain preventing the two DNA strands from complete dissociation into single strands. The length of the GC-clamp can vary between 30 and 50 nucleotides (Muyzer *et al*., 1997).

DGGE of PCR amplified 16S rDNA fragments were first used to profile community complexity of a microbial mat and bacterial biofilms (Muyzer *et al*., 1993). For this purpose bacterial genomic DNA was extracted from natural samples, and segments of the 16S rRNA genes were amplified in the polymerase chain reaction (Saiki *et al*., 1988). This resulted in a mixture of PCR products obtained from the different bacteria present in the sample. The individual PCR products were subsequently separated by DGGE. The result was a pattern of bands, for which the number of bands corresponded to the number of predominant members in the microbial communities. Various studies have been conducted where DGGE has been utilized for studying soil community analysis (Duineveld *et al*., 1998). PCR primers often targeting to specific population of microbes were chosen due to intense complexity of soil communities to known to be the major inhabitants of soil, such as actinobacteria (Heuer *et al*., 1997) or limited their analysis to specific soil environments (Duineveld *et al*., 1998). Muyzer *et al*., (1995) used DGGE analysis of PCR amplified rDNA fragments to provide information on the genetic diversity of microbial communities found around hydrothermal vents. DGGE analysis of 16S rDNA fragments has also been used to study the presence and activity of sulfate reducing bacteria in a stratified water column of Mariager Fjord in Denmark (Teske *et al*., 1996).

2.2.1.1 Statistical Analysis of DGGE Profiles

DGGE fingerprinting can be coupled with statistical analyses for the calculation of different bio-diversity indices like Shannon-Weaver Index, Sorenson's Similarity Index, Fo index etc. to compare the bacterial community present in each of the samples. Each gel represents a detailed description of a given microbial community in relation to the studied environment. The statistical tools used in this study and its interpretation is independent of both the technique and the settings of the technique itself providing an ecological and predictive value to the analyses of the structure and the diversity of the microbial community in a given environment.

(a).Sorenson's Similarity Index (Cs):

The relatedness of microbial community composition was studied by Nakatsu, *et al*., using the Sorenson's Pairwise Similarity Coefficient, Cs (Sorenson, 1948) in 6 different soil samples collected from USA and Norway. They considered that two bands in a gel are common if they migrate to the same distance. The total number of bands was determined first for the samples being compared. Then each sample was scored based on the presence and absence of each band in its profile in a pairwise manner to determine the similarity coefficient. A value of 1.0 indicated that all bands shared are common and 0.0 indicates that no bands are in common. Sorenson's index of similarity,

$$Cs = 2j/ (a+b)$$

was used to make pairwise calculations of band sharing between samples (Sorenson, 1948). In that equation, a is the number of bands in sample A, b is the number of bands in Sample B, and j is the number of bands common to A and B (Nakatsu, *et al*., 2000).

(b). Range weighted Richness (Rr):

Range weighted Richness is a parameter that can be used to mathematically define the carrying capacity of an environment (Marzorati, *et al*., 2008). This is possible since PCR-DGGE is a technique which subjects the double stranded PCR products to an increasing denaturing condition in a polyacrylamide gel. The amplicons partially melt into discrete regions called 'melting domains' creating molecules of reduced mobility (Muyzer *et al*., 1999). Several authors showed that the per cent GC content can be used to study the diversity and shift in a microbial community, even if there is no

taxonomical correlation (Torsvik *et al*., 1990; Ovreas *et al*., 1998; Nusslein and Tiedje, 1999). Therefore, the GC-content of complex microbial communities seems to be influenced by the environment itself (Foerstner *et al*., 2005).

The analysis of the microbial fingerprint of DGGE can therefore be used to correlate with the characteristics of the environment analyzed. If, an environment is very habitable, it can host a lot of different microorganisms with different genetic variability, hence a wider gradient would be needed to describe the total microbial diversity and vice versa. Thus, the broader the carrying capacity of an environment is (the number of individuals that the resources of an environment can support), the higher the probability it can host a high number of bands with a wide GC-variability (both in terms of percentage and in terms of positioning of the GC-stretches within a 16SrRNA gene). Based on DGGE, Rr< 10 can be attributed to environments particularly adverse or restricted to colonization characterized by low range weighted richness. Values intermediate between 10-30 can be correlated with a medium range weighted richness, while Rr>30 is typical of a very habitable environments (with a broad carrying capacity) characterized by high microbial diversity and with a high range weighted richness. This value was characterized according to the following formula:

$$\text{Rr} = N^2 \times Dg$$

where, N represents the total number of bands in the pattern and Dg the denaturing gradient occurring between the first and the last band of the pattern analyzed (Marzorati, *et al*., 2008).

©. Shannon- Weaver diversity index (H')-

The Shannon- Weaver index is a quantitative index which computes band intensity with species abundance, while the estimation of species richness(S) provides qualitative data relating to the total number of bands observed to the total species number and species richness(S) and correspond to the total number of unique bio-marker sequences within the ecosystem. The generated DGGE banding pattern is an "image" of the whole bacterial community, where each individual discrete band refers to a unique "sequence type" or phylotype or operational taxonomic unit (OTU) that corresponds to a discrete bacterial population. The total number of bands (S) can be determined and used for comparing communities (e.g. Duarte *et al*., 2009; Mille-

Lindblom *et al*., 2006; Nikolcheva *et al*., 2003; Solé *et al*., 2008). Diversity comparisons was done taking into account the relative intensity of each band (Pi), assuming that primers had the same extension efficiency during PCR. Shannon's diversity index (H') can be easily calculated to describe possible changes in the dominance among DGGE OTUs using the following equation:

$$(H') = -\Sigma(Pi)\ln Pi$$

where, Pi is the relative intensity of OTU, i and S is the total number of OTUs in the profile.

(d). Functional Organization (Fo) index-

Functional organization is the result of the action of the microorganisms that are most fitting to the ongoing environmental-microbiological interactions. For this reason they tend to become dominant within the structure of the microbial community. To understand the relationship between functionality and community structure, Fernandez *et al*., (2000) observed the behaviour of different microbial communities under perturbed conditions. It was suggested that the conservation of given functionality is ensured by the flexibility of a microbial community with minority community members that may become dominant in a short period following a significant perturbation.

2.2.2 Sequencing

Traditionally, microbiologists have tried to characterize the bacteria present in the disease suppressive soils by culture based techniques. But with the advent of recent rRNA amplicon sequencing technique to study the microbiota of the soil, it has increased the knowledge of the researchers about the community composition of such soils. This technique exploits the different hypervariable regions of the 16SrRNA gene to identify different sequences present in the sample. The sequences are then compared with the reference database to obtain the operational taxonomic units (OTU's).

In case of Take all decline (TAD), disease suppression is attributed to the microbial component of the soil i.e. the role of antagonistic fluorescent Pseudomonads have been studied so far. But, very little is known about the community structure of the suppressive soil samples of TAD. Therefore, Sanguin, *et al*., (2009) did a study on the rhizosphere bacterial communities which may be associated with the Take-all Decline

disease that occur during wheat monocropping. They isolated DNA from the rhizosphere soil samples collected at 3 different stages i.e. plots grown wheat for 1 year (low level of take all disease), 5yr (high level of disease) or 10yr (low level of disease suppressiveness reached). The changes in the bacterial community were assessed with a 16SrRNA-based microarray with 657 probes followed by cloning and sequencing. The outbreak stage (5 yrs.) was mainly characterized by the prevalence of *Proteobacteria*, notably *Pseudomonas* (*Gammaproteobacteria*), *Nitrospira* (*Betaproteobacteria*), *Rhizobacteriaceae*, *Sphingomonadaceae*, *Phyllobacteriaceae* (*Alphaproteobacteria*), as well as *Bacteroidetes* and *Verrucomicrobia*. By contrast, suppressiveness (10yrs.) correlated with the prevalence of a broader range of taxa, which belonged mainly to *Acidobacteria*, *Planctomycetes*, *Nitrospira*, *Chloroflexi*, *Alphaproteobacteria* (notably *Azospirillum*) and *Firmicutes* (notably *Thermoanaerobacter*).

2.3 Screening for the presence of agriculturally important genes in the disease suppressive soils

Disease suppressive soils are considered to be the reservoirs of genes or genetic loci which might have a role in controlling phytopathogens. Therefore, there is a considerable commercial and research interest in identifying those potential genetic loci for use in developing bioactive compounds that might contribute to disease suppression. Hjort, K. *et al*., (2010) screened for the presence of chitinase genes in phytopathogen suppressive soils of clubroot disease caused by *Plasmodiophora*. The screening for the presence of chitinase genes was done in 3 different ways- (a).from a metagenomic library constructed from microbial cells isolated from soil, (b).from directly extracted DNA and (c).from bacterial isolates with antifungal and chitinase activities. The amplified chitinase genes in the samples were then digested with 3 different restriction enzymes to generate the chitinase T-RFLP. The pooled PCR products were then cloned and sequenced. The most abundant chitinase found in the soil DNA and the metagenomic library corresponded to the TRF^{103} which showed closest identity to *Streptomyces mutomycini* and/or *Streptomycis clavifer*.

.

MATERIALS AND METHODS

The metagenomic work for characterization of the disease suppressive soils was done to analyze, whether the kinds of bacteria present in such soils have any relevance to suppressiveness based on the metagenomic analysis of the soil DNA. The following stated materials and methods were employed in the course of research work, to obtain the results for the set objectives.

3.1 Soil Sampling:

The soil samples for research work were collected from the farm of Jawahar Lal Nehru Krishi Vishwavidyalaya (JNKVV), Jabalpur, Madhya Pradesh. The soil samples collected were black clay soil with a high humic acid content. The soil samples were collected from the root zone of the crops plants Soybean, Okra and Rice which were showing suppressiveness to nematode infestation of Root Knot Nematode (*Meloidogyne incognita*) in case of Okra, Reniform nematode (*Rotylenchulus reniformis*) in case of Soybean and Sheath Rot disease (*Rhizoctonia bataticola*) in case of Rice. Two replications of each soil sample were taken alongwith one control soil sample from a nearby fallow plot. Based on these, the replications were mixed and samples were taken for further analysis.

Table 1: Samples taken for study

Sl. No.	Sample Name	Crop
1.	S	Soybean
2.	O	Okra
3.	R	Rice
4.	C	Control

3.2 SOIL MICROBIAL COMMUNITY DNA EXTRACTION AND PURIFICATION

Before DNA extraction, the collected soil samples were crushed and sieved through a 2 mm sieve, so as to remove root and leaf residues, to avoid amplification from other genomes. Direct method of DNA extraction was used for soil metagenome DNA extraction, where combination of different available methods was used which is described below. The composition of soil DNA extraction buffer was a modification of the buffer composition as given by Miller *et al*., (1999). For removal of humic acid

inhibitors, chemical flocculation method was adopted, which was done using $CaCl_2$ as a chemical flocculant (Braid *et al*., 2003).

3.2.1 DNA Extraction

For soil microbial community DNA extraction, two hundred fifty mg of each soil sample, was weighed and transferred in a 2 ml microcentrifuge tube. To the samples 1000 µl of soil DNA extraction buffer (100mM Tris-Cl (pH-9), 100mM EDTA (pH-9) and 1.5M NaCl) was added; along with this 120 µl of $CaCl_2$ (120mM) and 200 µl of 20% sodium dodecyl sulphate was added and the mixture was mixed properly by vortexing for 30 sec-1min. The samples were then vortexed at 1400 rpm at 80°C for 90 min. The tubes were then centrifuged at 13200 rpm for 15 min at room temperature. 1000 µl of the clear supernatant was then transferred to a fresh 2 ml microcentrifuge tube.

3.2.2 DNA Purification

Purification of DNA was done using organic solvents. To 1000 µl of collected supernatant, equal volume of chloroform: isoamyl alcohol (24:1) was added, to remove protein and other impurities. The content was thoroughly mixed and centrifuged at 12200 rpm for 20-30 mins at room temperature. The clear supernatant was transferred to a fresh 1.5 ml microcentrifuge tube. To this supernatant, equal volume of ice-chilled isopropanol was added along with one tenth the volume of 3M Sodium Acetate for precipitation of DNA. This was kept for overnight incubation at -20°C followed by centrifugation at 13200 rpm for 15 min at 4°C. The pellet thus obtained was washed with 70% ethanol and air-dried to remove the traces of ethanol. The dried DNA samples were dissolved in 30 µl of $T_{10}E_1$. The concentration of extracted DNA was checked using NanoDrop ND 1000 Spectrophotometer, using $T_{10}E_1$ as blank. The isolated DNA was checked in an agarose gel of 0.8% strength.

3.3. PCR Amplification of the isolated DNA using 16S rDNA primers

The protocols used in this work were based on the amplification of the 16S rRNA gene by PCR, directly from soil samples (Nakatsu, *et al*., 2000). Polymerase chain reaction was done with diluted purified DNA in 1:5 ratios of DNA and water. 16S rDNA primer pairs targeting partial variable domains of 16S rRNA gene fragment of Bacteria. The reaction mixture was prepared for final volume of 10 µl which contained 0.25pmol of each forward and reverse primers, 0.1mM of dNTP's, 1X Taq buffer A containing

1.5mM $MgCl_2$ and 1 unit of Taq DNA polymerase (GeNei, India). PCR was performed in automated thermal cycler (Eppendorf master cycler, Germany). After completion of PCR, amplified products were stored at 4°C and were checked on 1% agarose gel by agarose gel electrophoresis.

Table: 2 List of Primer Pairs used

Primer Pairs	16Sr DNA target region	Primer sequence	Reference
PRBA 338	Bacteria V3 region (338-358)	5'AC TCC TAC GGG AGG CAG CAG 3'	Lane (1991)
PRUN 518	Universal V3 region (534-518)	5'ATT ACC GCG GCT GCT GG 3'	Muyzer *et al.*(1993)
PRBA 968	Bacteria V6 region (968-983)	5'AA CGC GAA GAA CCT TAC 3'	Nu¨ bel *et al.*(1996)
PRBA 1406	Bacteria V9 region (1406-1392)	5'ACG GGC GGT GTG TAC 3'	Lane *et al.* (1988)
E783	Eubacteria region (783-797)	5' CAGGATTAGATACCC 3'	Wang and Qian(2009)
E926	Eubacteria region (909-926)	5'ACTCAAAKGAATTGACGG 3'	Wang and Qian (2009)

Table 3: PCR Programme for 3 different Primer Pairs

Temperature	**PRBA 338 & PRUN 518**	**PRBA 968 & PRBA 1406**	**E783 & E926**
Initial Denaturation	95^0 C for 5 mins.	95^0 C for 5 mins	95^0 C for 5 mins
Denaturation	94^0 C for 45 secs.	94^0 C for 45 secs.	94^0 C for 50 secs.
Primer Annealing	57.4^0 C for 30 secs.	55^0 C for 45 secs.	48.3^0 C for 30 secs.
Extension	72^0 C for 50 secs.	72^0 C for 1 min, 30 secs.	72^0 C for 50 secs.
Final Extension	72^0 C for 5 mins	72^0 C for 10 mins.	72^0 C for 7 mins
Number of Cycles	30	35	40

3.4 Diversity Analysis

3.4.1 Diversity Analysis Using DGGE

DGGE analysis was done to check for the bacterial diversity present in the disease suppressive soil samples collected from the rhizosphere of 3 different crop plants. For DGGE, PCR was done using same template DNA of all the 10 soil samples. PCR conditions and the amplification program details were same as above. The same set of primer pairs was used for DGGE analysis targeting the V-3 and V-6 region of the 16S rDNA of bacteria and variable region of Eubacteria but with a 40 base GC- rich clamp (CGC CCG CCG CGC GCG GCG GGC GGG GCG GGG GCA CGG GGG G) attached at the 5'end of the forward primer of PRBA 338F-GC -PRUN 518R, PRBA 968F-GC-PRBA1406R and E 783F-GC- E926R. The PCR product of expected size (200 bp, 450 bp and 200bp respectively) was subjected to DGGE analysis by following protocol of Muyzer *et al.*, (1995). Acrylamide gel gradient of 12% was used for proper migration of bands in the gel. The polyacrylamide gels were prepared with denaturing gradient from 40-70% using a gradient maker, where 100% denaturant contains 7M urea and 40% (v/v) formamide.

DGGE was performed with Ingeny Phor U-2 system (Leiden, The Netherlands). Sample volume of 20 µl of PCR product was loaded onto 12% polyacrylamide gels in 1X tris-acetate EDTA buffer (TAE buffer, pH-8). The electrophoresis was run for 18 h at 150 V at 60°C buffer temperature. After complete run, gels were removed from the unit and transferred to the OHP sheets without causing any breakage or damage to the gel. Gel was stained by silver staining protocol as given by Gustavo and Peter (1994). After staining, gel was dried well and was analyzed using SynGene Gene Tools.

3.4.2 Statistical Analysis of the DGGE Fingerprints

DGGE fingerprinting can be coupled with statistical analyses for the calculation of different bio-diversity indices like Shannon-Weaver Index, Sorenson's Similarity Index, Fo index and Moving window analysis to compare the bacterial community present in each of the samples. Each gel represents a detailed description of a given microbial community in relation to the studied environment. The statistical tools used in this study and its interpretation is independent of both the technique and the settings of the technique itself providing an ecological and predictive value to the analyses of the structure and the diversity of the microbial community in a given environment.

(a).**Sorenson's Similarity Index**

Relatedness of microbial communities was determined using similarity coefficients of bands common to two samples. The working definition was that two bands are common if they migrated the same distance on the gel. Each sample was scored based on the presence or absence of each band in its profile when compared to the profile of each of the other samples. Sorenson's index of similarity,

$$Cs = 2j/ (a+b)$$

was used to make pairwise calculations of band sharing between samples (Sorenson, 1948). In that equation, a is the number of bands in sample A, b is the number of bands in Sample B, and j is the number of bands common to A and B (Nakatsu, *et al*., 2000).

(b*).* **Range weighted Richness**

A range-weighted richness (Rr) index can be mathematically expressed by defining the total number of bands observed multiplied by the percentage of denaturing gradient needed to generate the total diversity of the sample. This value describes a carrying capacity of an environment containing wide species GC variability (both in terms of percentage and in terms of positioning of the GC stretches within the 16S rDNA gene). This value was characterized according to the following formula:

$$\text{Rr} = N^2 \times Dg$$

where, N represents the total number of bands in the pattern and Dg the denaturing gradient occurring between the first and the last band of the pattern analyzed.

©. **Shannon- Weaver diversity index**

Bands were scored in the GeneSys Tools (SynGene Synoptics Ltd), by giving lowest score to the least intense band in the gel. On the basis of this scored data, Shannon's diversity index (H') can be easily calculated to describe possible changes in the dominance among DGGE OTUs using the following equation:

$(H') = -\Sigma(Pi)\ln Pi$

where, Pi is the relative intensity of OTU, i and S is the total number of OTUs in the profile (Shannon and Weaver, 1971).

(d). **Functional Organization (Fo) index**

In order to graphically represent the structure of the bacterial community (species distribution), the Pareto-Lorenz evenness curve (Lorenz, 1905) was constructed based on the DGGE profile. For each DGGE lane in the profile, the respective bands are ranked from high to low, based on their intensities. The cumulative normalized number of bands is used as the X-axis and their respective cumulative normalized intensities represent the Y-axis. The F_o index is the horizontal y-axis projection on the intercept with the vertical 20% x-axis line, i.e. the combined relative abundance of 20% of the OTUs. The more the PL- curve diverges from the 45^0 diagonal (the theoretical perfect evenness line), the less evenness can be observed in the structure of the studied community.

3.4.3 Diversity Analysis using sequence based approach

The bacterial diversity present in the disease suppressive soil samples was also analyzed using the Targeted Amplicon Sequencing Technology on Illmina MiSeq platform was done at Xcelris Labs, Ahmedabad, Gujarat.

The disease suppressive soil samples for the 3 different crops and a control sample was taken for amplicon sequencing. The workflow for sequencing is given below.

3.4.3.1 Qualitative and quantitative analysis of gDNA

Quality of gDNA was checked on 1% agarose gel (loaded 5 µl) for the single intact band. The gel was run at 110 V for 30 mins. 1 µl of each sample was loaded in Nanodrop 8000 for determining A260/280 ratio. The DNA was quantified using Qubit dsDNA BR Assay kit (Thermo Fisher Scientific Inc.). 1 µl of each sample was used for determining concentration using Qubit® 2.0 Fluorometer.

3.4.3.2 Preparation of libraries for 2 x 300 bp Run Chemistry

The amplicon libraries were prepared using Nextera XT Index Kit (Illumina inc.) as per the 16S Metagenomic Sequencing Library preparation protocol (Part # 15044223 Rev. B). Primers for the amplification of the V3-V4 hyper-variable region (Table-1) of 16S rDNA gene of Eubacteria and Archaea were designed in Xcelris NGS Bioinformatics Lab.

Table 4: Primer Pair Used for Sequencing

Sl No.	Oligo Name	Oligo Sequence	Length of Primer	Product Size (Approx.)
1.	V3-Forward	CCTACGGGNGGCWGCAG	17	460bp
	V-4 Reverse	GACTACHVGGGTATCTAATCC	21	

The amplicons with the Illumina adaptors were amplified by using i5 and i7 primers that add multiplexing index sequences as well as common adapters required for cluster generation (P5 and P7) as per the standard Illumina protocol. The amplicon libraries were purified by 1X AMpureXP beads and checked on Agilent High Sensitivity (HS) chip on Bioanalyzer 2100 and quantified on fluorometer by Qubit dsDNA HS Assay kit (Life Technologies).

3.4.3.3 Cluster Generation and Sequencing

After obtaining the Qubit concentration for the library and the mean peak size from Bioanalyser profile, library was loaded onto MiSeq at appropriate concentration (10-20pM) for cluster generation and sequencing. Paired-End sequencing allows the template fragments to be sequenced in both the forward and reverse directions on MiSeq. The kit reagents were used in binding of samples to complementary adapter oligos on paired-end flow cell. The adapters were designed to allow selective cleavage of the forward strands after re-synthesis of the reverse strand during sequencing. The copied reverse strand was then used to sequence from the opposite end of the fragment.

3.4.3.4 Sequence analysis

Raw sequences were analyzed using MG-RAST server. The raw data was submitted in Fasta, Fastq format. The input processing steps consisted of quality filtering, demultiplexing (split on barcodes), length filtering, dereplication, and removal of model organism sequences.

The framework then employs RDP's (Ribosomal Database Project) naïve Bayesian classifier for rapid assignment of sequences to the taxonomic groups at the phylum, class, order, family and genus levels and the framework proceeds to perform unsupervised preliminary analyses on the samples acquired from the RDP classifier results. These analyses include:

- Calculations of total and percent abundance of bacteria in every sample,
- Bar chart representation of the number of sequences acquired for each sample
- Bar chart representation of Shannon and Simpson's diversity indices
- Pie chart representations of samples based on their bacterial compositions at each taxonomic level ranging from phylum to genus
- Rarefaction curves to illustrate the degree of diversity covered by each sample
- Hierarchical clustering dendogram that illustrate how samples clustered based on their bacterial composition at different taxonomic levels
- Lowest common ancestors

The diversity of the soil microbial communities can be checked out by the presence of unique or abundance of particular phylum or genus in a particular disease suppressive soil samples.

3.5 Screening for the presence of agriculturally important genes using PCR

The disease suppressive soil metagenome was screened for the presence of agriculturally important genes which might be responsible for conferring the disease suppressive nature to that particular soil. The metagenome soil DNA was used as a template for such metagenomic exploration. The DNA from the pooled soil samples of the 3 crop plants and the control soil were screened for the presence of the agriculturally important genes like *chiA*, *cry12*, *cry13*, and *cry14*, *gdh*, pks,. The genes were selected based on the fact that these genes might have a role in conferring disease suppressiveness characteristic of the soil.

Polymerase chain reaction was done with diluted purified DNA in 1:5 ratios of DNA and water. The reaction mixture was prepared for final volume of 10 µl which contained 0.25pmol of each forward and reverse primers, 0.1mM of dNTP's, 1X Taq buffer A containing 1.5mM $MgCl_2$ and 1 unit of Taq DNA polymerase (GeNei, India). PCR was performed in automated thermal cycler (Eppendorf master cycler, Germany). After completion of PCR, amplified products were stored at 4°C and were checked on 1% agarose gel by agarose gel electrophoresis.

Table 5: List of Primer Pairs used

Primer Pairs	Primer sequence	Expected Amplicon Size (bp)	Reference
Cry 12	FP: 5' CTCCCCCAACATTCCATCC3' RP:5'AATTACTTACACGTGCCATACCTG3'	363bp	Ejiofar and Johnson (2002)
Cry 13	FP:5'CTTTGATTATTTAGGTTTAGTTCAA3' RP: 5' TTGTAGTACAGGCTTGTGATTC 3'	313bp	Bravo *et al.*(1998)
Cry 14	FP:5'ATAATGCGCGACGACCTACTGTTGT3' RP:5'TGCCGTTATCGCCGTTATT3'	456bp	Ejiofar and Johnson (2002)
ChiA	F2: 5'CGTGGACATCGACTGGGARTWYCC3' R2:5' CCCAGGCGCCGTAGARRTCRTARSWCA3'	270-300bp	Hobel, *et al.* (2005)
Gdh	FP:5'CTGCCAGTDAACGAYGGYCGYCGYCTG3' RP:CACAACCAGAGGCTGGTGGTGGTGTA3'	1400bp	Pujol and Kado, (1999)
Pks1	K1F: 5'TSAAGTCSAACATCGGBCA3' M6R: 5'CGCAGGTTSCSGTACCAGTA3'	1200-1400bp	Ayuso-Sacido and Genilloud (2005)

Table 6: PCR Programme for 3 different Cry Primer Pairs

Temperature	Cry 12F & Cry12R	Cry13F & Cry 13R	Cry14F &Cry14R
Initial Denaturation	94^0C for 5mins	94^0C for 5mins	94^0C for 8 mins
Denaturation	94^0Cfor 50secs	94^0C for 1 min	94^0C for 1 min
Primer Annealing	45^0C for 45 secs	52^0C for 1 min	50^0C for 1 min
Extension	72^0C for 30secs	72^0C for 2mins	72^0C for 2mins
Final Extension	72^0C for 5 mins	72^0C for 20mins	72^0C for 20mins
Number of Cycles	30	40	35

Table 7: PCR Programme for ChiA, Gdh and Pks Primer Pairs

Temperature	**ChiA F2 & ChiA R2**	**GdhF & GdhR**	**Pks1F & Pks1R**
Initial Denaturation	95^0C for 5mins	95^0C for 5mins	94^0C for 5mins
Denaturation	94^0C for 30 secs	95^0C for 2 mins	94^0C for 1 min
Primer Annealing	55^0C for 30 secs	55^0C for 1 min	60^0C for 1 min
Extension	72^0C for 1 min 30 secs	72^0C for 2mins	72^0C for 2mins
Final Extension	72^0C for 10 mins	72^0C for 10 mins	72^0C for 10 mins
Number of Cycles	32	35	35

4. EXPERIMENTAL RESULTS

The aim of the present research was to study the community diversity of the metagenome of three disease suppressive soils in comparison with the control soil sample using the culture independent methods. The research so conducted focused mostly on the isolation of high molecular weight microbial DNA from the disease suppressive soil samples followed by analyzing the microbial diversity using the PCR-DGGE and sequence based approach and the interpretation of results. The metagenome was also screened for the presence of agriculturally important genes which might help in deciphering the genetic potential of the disease suppressive soil samples.

4.1.1 Soil DNA isolation

The DNA isolated from the four soil samples was of good quality with no visual shearing. The average yield of DNA obtained was in the range of 2-3µg per gram of soil sample with purity ratio of 1.50-1.56 at 260/280 measured using NanoDrop ND 1000 Spectrophotometer (Plate 1).

For enhancing the precipitation of the isolated DNA and increasing the yield, the DNA was precipitated using isopropanol and 3M Sodium Acetate with an enhanced incubation of overnight. Chemical flocculation method using $CaCl_2$, was quite effective in obtaining DNA free from humic substances.

4.2 PCR amplification of the soil DNA using universal primer pairs for bacteria and eubacteria

For amplification of the soil DNA from the four soil samples, the DNA was diluted in the ratio of 1:5 with sterile water for use as template in the PCR program. The annealing temperature for the 3 primer pairs used PRBA338-PRUN518 (V-3 region of Bacteria), PRBA968-PRBA1406 (V-6 region of Bacteria) and E783-E926 (variable region for Eubacteria) was found to be 57.4^0C, 55^0C and 49.3^0C respectively. The amplicon size of 200bp was observed for the first and third primer pairs whereas an amplicon size of 450bp was observed for the second one (Plate 2, 3& 4) in an Ethidium bromide stained 1% agarose gel.

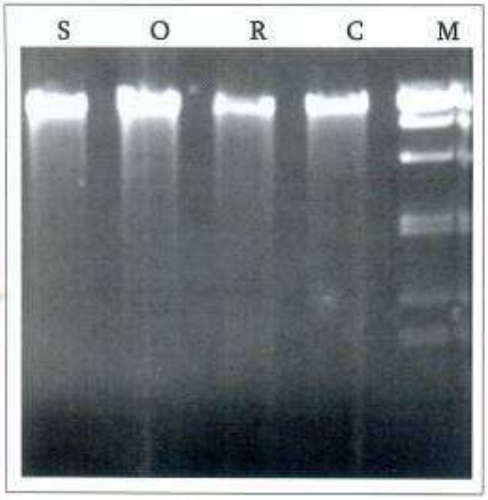


Plate 1: DNA isolated from the 3 suppressive soil samples and control soil sample.

S- Soybean soil sample, O- Okra soil sample, R- Rice soil sample and C- Control soil sample. M -Mlu single digest marker.

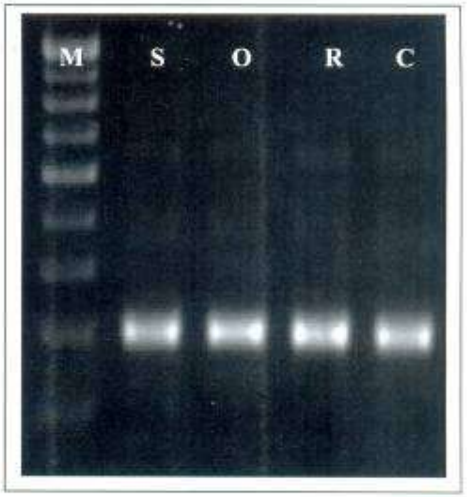


Plate 2: PCR amplification of V3 region of the 16SrDNA from soil DNA (4 samples) using the primer pairs PRBA 338F and PRUN 518R giving an amplified product size of 200bp. M- 100bp ladder.

Plate 3: PCR amplification of V6 region of the 16SrDNA from soil DNA (4 samples) using the primer pairs PRBA 968F and PRBA 1406R giving an amplified product size of 200bp. M- 500bp ladder.

Plate 4: PCR amplification of variable region of the 16SrDNA from soil DNA (4 samples) using the primer pairs E 783F and E 926R giving an amplified product of 200bp. M- 100bp ladder.

4.3. Diversity Analysis

4.3.1 Using DGGE

After confirmation of the presence of bands in agarose gel, the same template was used for amplification of the specified regions of 16SrDNA in the four soil samples. The same set of primer pairs were used but with a slight modification i.e. at the 5'end of the forward primers a 40bp rich GC-clamp was attached (e.g.- GC-PRBA338F & PRUN518R). The PCR products so obtained were loaded onto 12% polyacrylamide gels. The DGGE fingerprint for the primer pair GC-PRBA338 and PRUN518 obtained showed the presence of 24, 23, 26 and 21 OTU's (operational taxonomic units, where one OTU corresponds to single band on the DGGE gel which indicates one separate taxonomic unit) in S, O, R and C samples respectively (Plate 5). Whereas, the other set of primer pair targeting the V-6 region of bacteria showed the presence of 21, 20, 19 and 21 OTU's respectively in the S, O, R and C samples respectively (Plate 6). The 3rd primer pair targeting the variable region of Eubacteria showed the presence of 26, 25, 22 and 16 OTU's respectively (Plate 7).

For analyzing the community microbial diversity in each of the four soil samples using the three primer pairs, the bands were scored using the GeneSys Tools from SynGene Synoptics Ltd. The similarity between the samples was calculated statistically using the Sorenson's Pairwise Similarity Index which is depicted in the form of a similarity matrix for each of the three primer pairs (Table 8, 9 & 10). Relatedness of the microbial community present in between the 3 suppressive soil samples with the control soil sample was determined using this index. A value of 1.0 indicates all OTU's are shared and 0.0 indicates no OTU's are in common. The diversity, richness and functional organization for the four samples were calculated using the Shannon-Weiner Index (Plate- 5), Range Weighted Richness (Plate-6) and Pareto-Lorenz Evenness Curve (Plate-7) respectively.

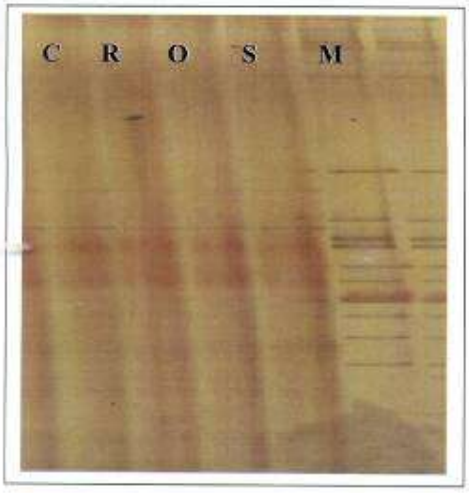


Plate 5: (Left) DGGE profile for the 3 suppressive soil samples (S, O and R) in comparison with the control soil sample (C) using the primer pair PRBA 338-GCF-PRUN518R.

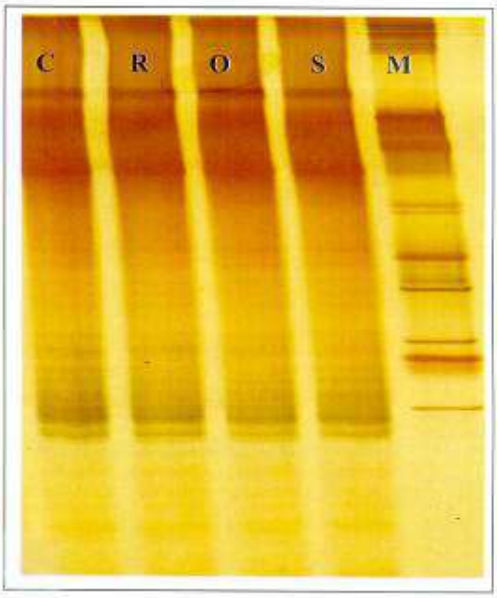


Plate 6: (Right) DGGE profile for the 3 suppressive soil samples (S, O and R) in comparison with the control soil sample (C) using the primer pair PRBA 968F-GC - PRBA1406 R.

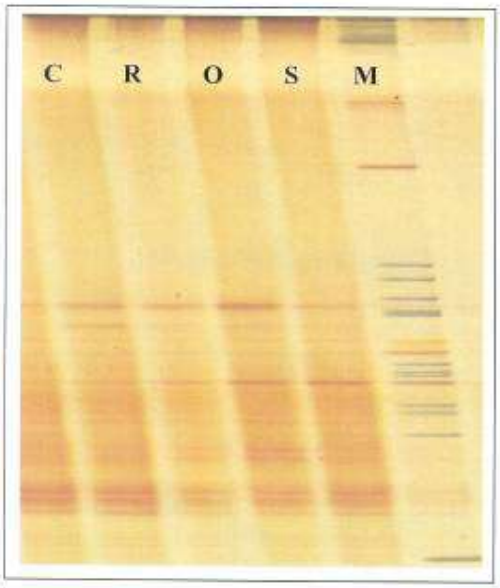


Plate 7: (Right) DGGE profile for the 3 suppressive soil samples (S, O and R) in comparison with the control soil sample (C) using the primer pair E-783F-GC – E926 R.

Table 8- Similarity matrix between control soil sample and the 3 disease suppressive soils samples for PRBA 338F-GC and PRUN 518R

Combination (Cs)	Soybean	Okra	Rice	Control
Soybean	1	0.69	0.62	0.51
Okra	0.69	1	0.61	0.53
Rice	0.62	0.61	1	0.53
Control	0.53	0.53	0.53	1

Table 9- Similarity matrix between control soil sample and the 3 disease suppressive soils samples for PRBA 968F-GC and PRBA 1406R

Combination (Cs)	Soybean	Okra	Rice	Control
Soybean	1	0.81	0.76	0.76
Okra	0.81	1	0.73	0.81
Rice	0.76	0.73	1	0.76
Control	0.76	0.81	0.76	1

Table 10- Similarity matrix between control soil sample and the 3 disease suppressive soils samples for E 783F-GC and E 926R

Combination (Cs)	Soybean	Okra	Rice	Control
Soybean	1	0.82	0.83	0.66
Okra	0.82	1	0.78	0.64
Rice	0.83	0.78	1	0.63
Control	0.66	0.64	0.63	1

4.3.1.2 Microbial diversity analysis for the primer pair PRBA 338F-GC & PRUN 518R

Similarities between the four soil samples as calculated using the Sorenson's Pairwise Similarity Index showed that the combination of Soybean and Okra soil share the highest similarity (0.69) followed by Soybean and Rice (0.62) and Okra and Rice (0.61). The latter two combinations shared almost the same similarity coefficients. Whereas, the Soybean and Control soil sample share the least similarity of 0.51. In terms of diversity, which was calculated using the Shannon-Weiner

Diversity index (Table-11) showed that the Control soil sample has a significant biological diversity when compared with the other 3 disease suppressive soil samples (p=0.01). The Shannon Diversity ranged from 1.81-2.92 in the 4 samples. The range weighted richness in the 4 soil samples ranged from 56.80 to 87.08(Table 12). The Rice soil sample showed the highest range weighted richness followed by Soybean, Okra and Control. The functional organization Index showed that the index varied from 68% in the control soil sample to 92% in the Rice soil sample (Table 13).

4.3.1.2 Microbial Diversity Analysis Using the Primer Pair PRBA 968F-GC & PRBA 1406R

Similarities between the four soil samples as calculated using the Sorenson's Pairwise Similarity Index for the primer pair targeting the V-6 region of bacteria showed that the combinations of Soybean and Okra soil sample along with Okra and Control soil sample combination share the highest similarity of 0.81. Almost all other combinations show similar coefficients of similarity. In terms of diversity, which was calculated using the Shannon-Weiner index almost all the four samples showed similar microbial diversity ranging from 2.65- 2.49 (Table 11). The okra soil sample showed significant difference in terms of microbial community composition in comparison with the control soil sample. The range weighted richness in the four soil samples ranged from 59.99 - 73.29 (Table 12). The Soybean and Control soil sample showed the highest range weighted richness followed by Okra and Rice. The functional organization Index showed that the F_o values were same for Soybean, Rice and Control soil samples (86%) except for Okra soil sample (75%) (Table 13).

4.3.1.3 Microbial Diversity Analysis Using the Primer Pair E 783F-GC & E 926R

Similarities between the four soil samples as calculated for the Eubacteria primer pair showed that the similarity coefficients of Soybean and Rice (0.83) is almost at par with the Soybean and Okra combination (0.82) (Table 11). Whereas, the Rice and Control soil sample share the least similarity coefficient of 0.63. The diversity values as calculated by the Shannon-Weiner Index showed that the microbial diversity ranged from 2.49-2.60 for all the four soil samples. Significant difference (p=0.01) was observed in the 3 soil samples in comparison with the control soil sample. The range weighted richness for the four soil samples ranged from 81.12 to 30.72 (Table 12). The highest and lowest range weighted richness

value was shown by Soybean soil sample and the Control soil sample respectively. In terms of functionality, which was calculated using Functional Organization Index showed that the F_o values were almost at par for all the four soil samples ranging from 86%- 91% (Table 13).

Table 11- Shannon Weiner Diversity (H') Index for the control soil sample and 3 disease suppressive soil samples

Primer Pair used	Soybean	Okra	Rice	Control
PRBA 338F-GC and PRUN 518R	2.77**	2.1**	1.81**	2.92
PRBA 968F-GC and PRBA 1406R	2.49^{NS}	2.65**	2.50^{NS}	2.49
E 783F-GC and E 926R	2.59**	2.60**	2.56**	2.49

**p=0.01 level of significance,*p=0.05 level of significance, ns=non-significant

Table 12- Range Weighted Richness (Rr) for the control soil sample and 3 disease suppressive soil samples

Primer Pair used	Soybean	Okra	Rice	Control
PRBA 338F-GC and PRUN 518R	74.19	68.14	87.08	56.80
PRBA 968F-GC and PRBA 1406R	73.29	66.48	59.99	73.29
E 783F-GC and E 926R	81.12	75	58.08	30.72

Table 13- Functional Organization (F_o) Index for the control soil sample and 3 disease suppressive soil samples

Primer Pair used	Soybean	Okra	Rice	Control
PRBA 338F-GC and PRUN 518R	75%	85%	92%	68%
PRBA 968F-GC and PRBA 1406R	86%	75%	86%	86%
E 783F-GC and E 926R	87%	87%	86%	91%

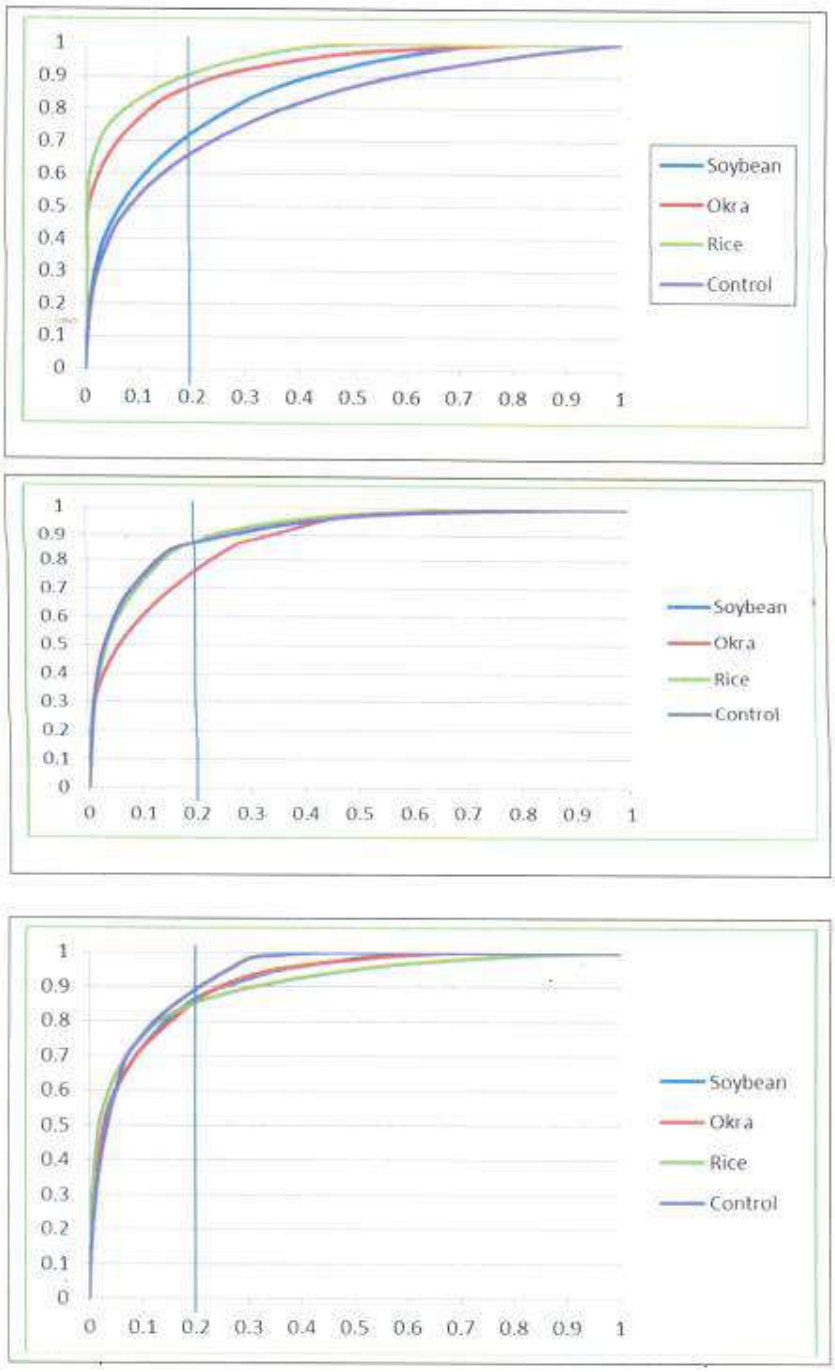


Figure 1: The Pareto-Lorenz evenness curve is the plot of the cumulative proportion of OTU abundances (y-axis) against the cumulative proportion of OTUs (x-axis) for the three primer pairs from top to bottom (PRBA 338F-GC-PRUN 518R, PRBA 968F- GC-PRBA1406R and E 783F-GC–E 926R).

4.3.2 Analysis of Microbial Diversity Using Sequence Based Approach

For better understanding of the soil microbial community structure and diversity of the four soil samples, the amplicons targeting the V-3 region of the bacterial 16SrDNA was subjected to next generation sequencing using the Targeted Amplicon Sequencing using Illumina MiSeq platform. The total DNA from the disease suppressive soil samples of Soybean, Okra, Rice and the Control soil sample yielded 354, 381, 294 and 281Mb of data assembled to 536768, 554840, 575913 and 372900 sequence reads respectively.

Phylogenetic Sequence Profile

The distribution of the domains as calculated from the data identified bacteria as the most dominant phyla followed by eukaryota (Figure 2). Phylogenetic profile was carried out using the M5RNA system databases in MG-RAST. Among the classified bacteria, the major phyla observed were Firmicutes, Actinobacteria, Proteobacteria, Planctomycetes, Acidobacteria, Bacteroidetes; however, each of their abundance was different among the four different soil samples (Figure 4). Very few sequences can be attributed to the phyla Nitrospirae, Verrucomicrobia, Chloroflexi and Gemmatimonas. The M5RNA- based phylogenetic analyses classified 87.4%, 60.6%, 58.2% and 63.7% sequences in Soybean, Okra, Rice and Control soil samples respectively.

Another parameter called alpha diversity summarizes the diversity of organisms in a sample with a single index. It helps in understanding the species richness and relative abundances of distinct species in a dataset. The alpha diversity was found to be 20.72, 58.36, 72.09 and 86.3 for the Soybean, Okra and Rice soil samples and control soil sample respectively.

The rarefaction curve showed the variation in the number of OTU's identified at a given percentage of identity as a function of the number of sequence reads obtained per sample. Rarefaction curve for the four samples is shown in figure (Figure 3). Analysis of the rarefaction curves provided the information on the sequence coverage obtained with the amplicon sequencing technique. The plateau of the curve was an indication of the species count (the number of OTU's in relation to number of reads per sample).

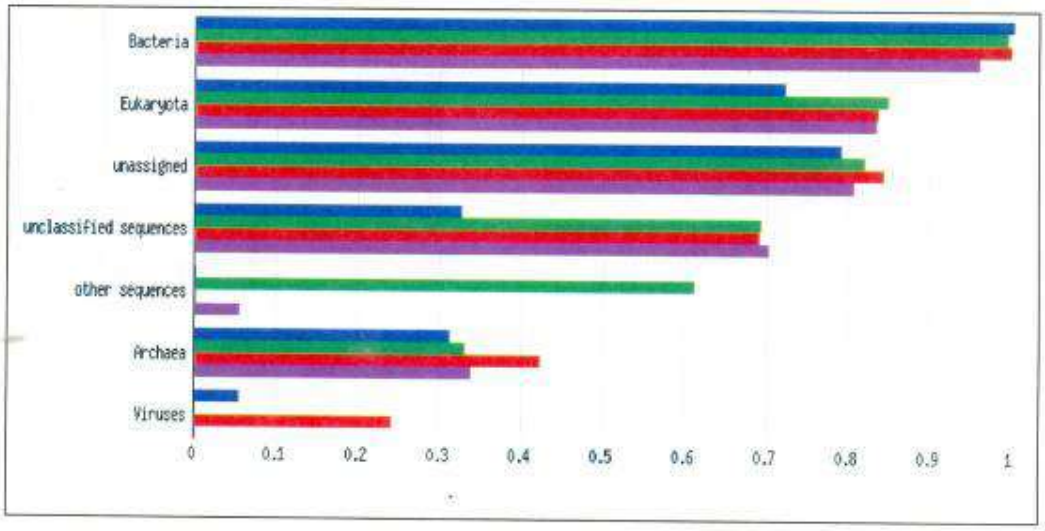


Figure 2: Domain wise distribution of sequences identified Bacteria as the most common organism among the 4 soil samples. - Soybean, -Okra, Rice and Control (The X-axis indicates 0-1 scale for 0-100%).

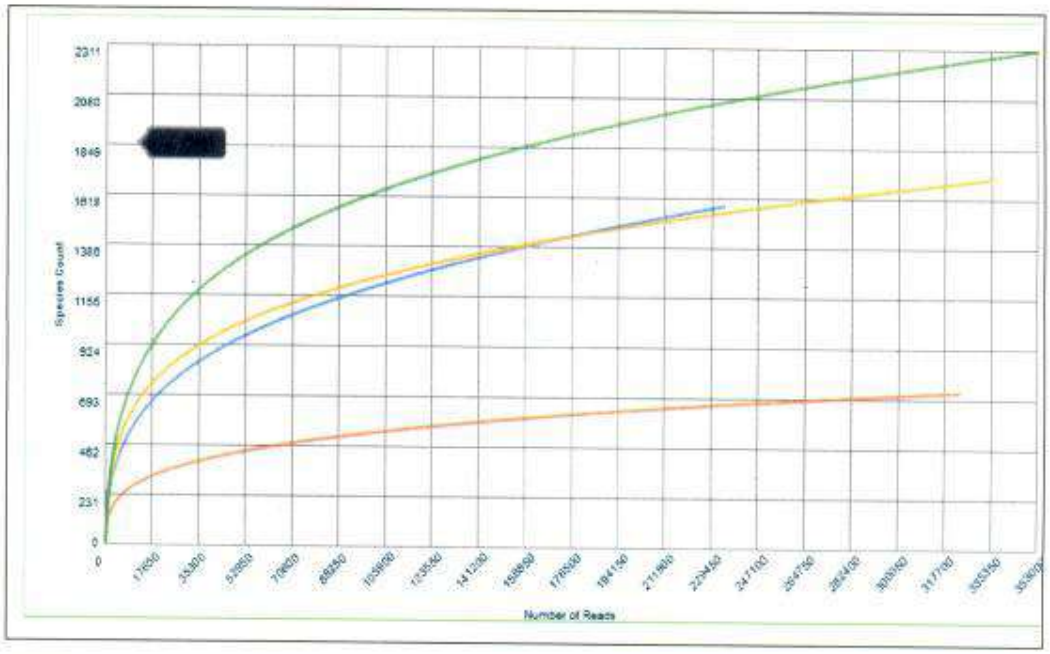


Figure 3: Rarefaction curve of the 3 disease suppressive soil samples in comparison with the control soil sample showing the variation in the no. of OTU's identified at a given percentage of identity as a function of the number of sequence reads obtained per sample.

Control - Okra - Rice - Soybean

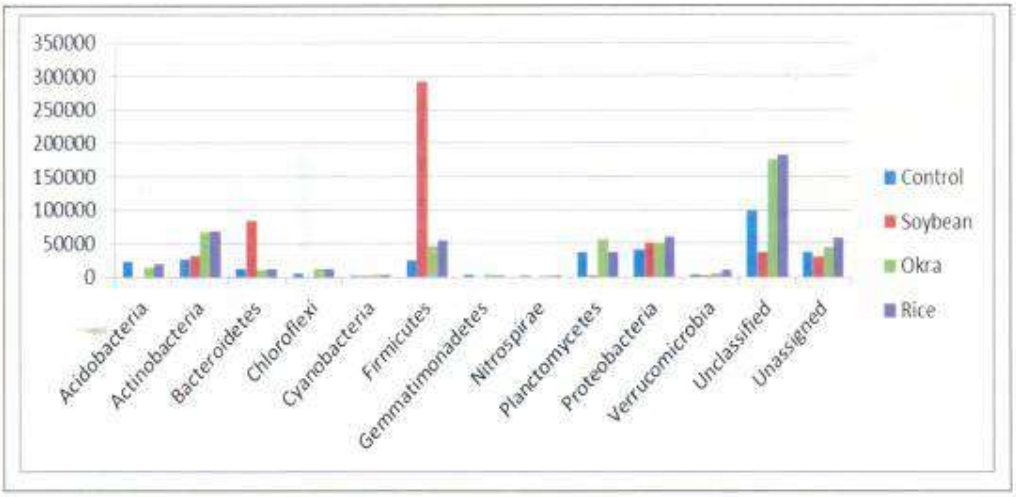


Figure 4: Comparison of the Phylum level abundance of the three disease suppressive soil samples in comparison with the control (adjacent fallow land) soil sample

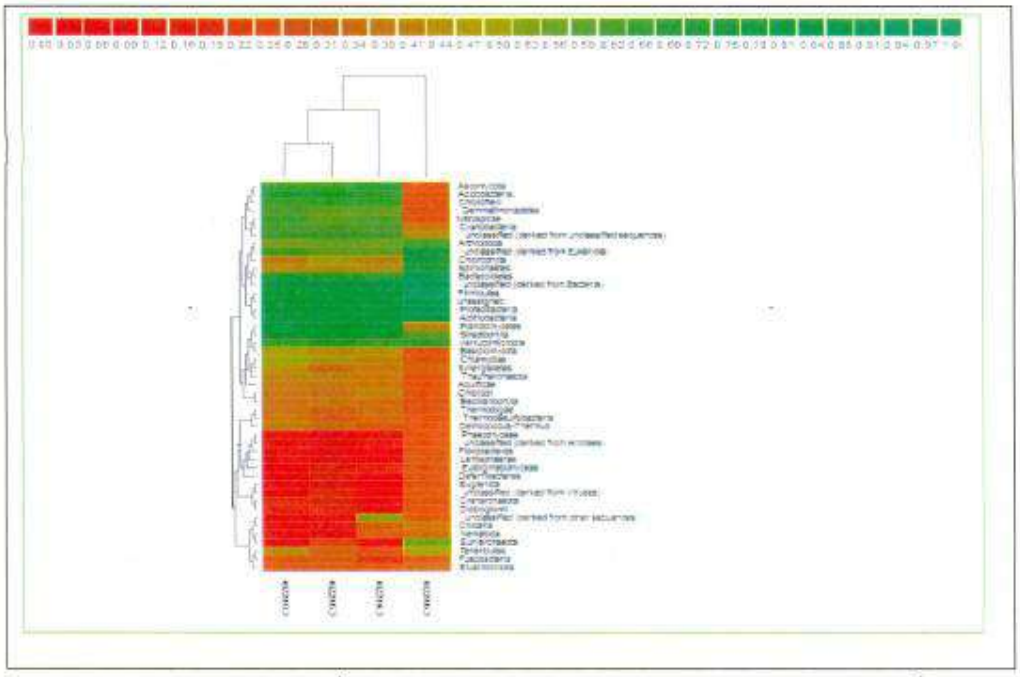
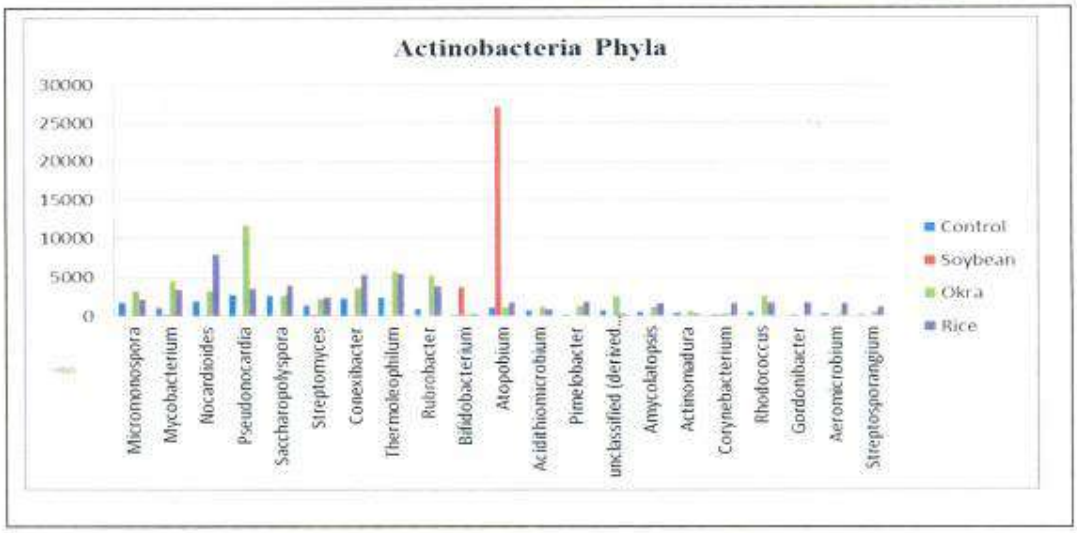
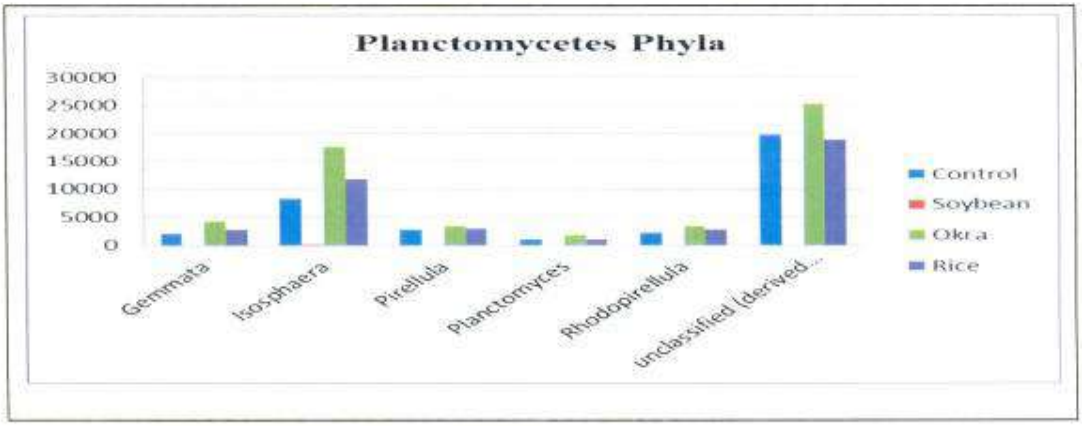


Figure 5: Heat dendogram showing similarities between the control soil sample- 4622993.3 and the 3 disease suppressive soil samples Okra- 4622994.3, Rice- 4622995.3 and Soybean- 4622996.3 at the phylum level. Shades represent effective raw scores from 0 to 1 of all identified sequences.

Figure 6: Comparison of the genera level abundance for the three disease suppressive soil samples in comparison with the control (adjacent fallow land) soil sample under each major phylum (Acidobacter, Bacteroidetes, Chloroflexi, Nitrospira, Gemmatimonadetes and Firmicutes).





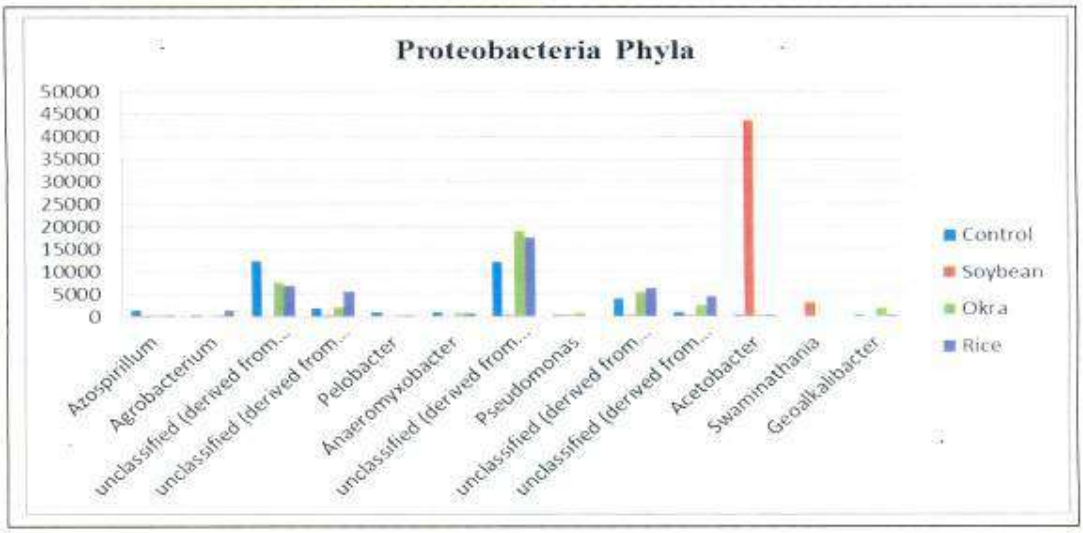


Figure 7: Comparison of the genera level abundance for the three disease suppressive soil samples in comparison with the control (adjacent fallow land) soil sample under each major phylum (Actinobacteria, Proteobacteria and Planctomycetes).

A heatmap dendogram indicates the similarity/dissimilarity between groups of vectors. Vectors in a dendogram are usually ordered with respect to their level of similarity; similar vectors are placed next to each other, while more distantly related vectors are placed farther apart. Each square in the heatmap dendogram represents the abundance level of a single category in a single sample. Figure showed the dendogram based on similarity/dissimilarity between the four (control soil sample-4622993.3 and the 3 disease suppressive soil samples Okra- 4622994.3, Rice-4622995.3 and Soybean- 4622996.3) soil samples on X-axis and different phylum in the Y-axis. Shades represent effective raw scores from 0 (indicates 0% similarity) to 1(indicates 100% similarity) of all identified sequences (Figure 5). For example, the Firmicutes in Soybean soil sample is highly abundant in comparison with the other 3 soil samples. In terms of microbial community composition, the Control and Rice soil sample share higher similarity than the Okra and Soybean soil samples.

Soybean soil sample showing disease suppressiveness to Reniform nematode **(***Rotylenchulus reniformis***)** is characterized by relative abundance Bacteria belonging to the major genera for each of the phyla are Firmicutes (*Dialister*, *Lactobac*illus), Actinobacteria (*Atopbium, Bifidobacterium*), Bacteroidetes (*Bacteroides, Prevotella*, *Alistipes, Empedobacter, Wautersiella, Porphyromonas*) and Proteobacteria (*Acetobacter, Swaminathania*) (Figure 6 & 7).

Okra soil sample showing disease suppressiveness to Root Knot nematode (*Meloidogyne incognita*) is characterized by relative abundance of Bacteria belonging to the major genera for each of the phyla are Actinobacteria (*Micromonospora, Mycobacterium, Nocardiodes, Pseudonocardia, Saccharopolyspora, Streptomyces, Conexibacter, Thermoleophilum, Rubrobacter, Atopbium, Pimelobacter, Rhodococcus, Acidithiomicrobium, Amycolatopsis*), Acidobacteria (*Candidatus Koribacter*, *Candidatus Solibacter*, *Candidatus Chloracidobacterium*), Bacteroidetes (*Flexibacter*, *Chitinophaga*, *Terrimonas*) Firmicutes (*Bacillus, Lactobacillus*), Chloroflexi (*Herpetosiphon, Ktedonobacter, Thermomicrobium, Chloroflexus, Roseiflexus, Dehalogenimonas*), Gemmatimonadetes (*Gemmatimonas*), Nitrospirae (*Nitrospira*), Verrucomicrobia (*Pedosphaera*), Planctomycetes (*Gemmata, Isosphaera, Pirellula, Planctomyces, Rhodopirellula*, unclassified derived from Planctomycetaceae) and Proteobacteria (*Geoalkalibacter,* unclassified derieved from *Alphaproteobacteria,*

Betaproteobacteria, Deltaproteobacteria, Gammaproteobacteria, Proteobacteria) (Figure 6 & 7).

Rice soil sample showing disease suppressiveness to Sheath Rot disease is characterized by relative abundance of Bacteria belonging to the major genera of each of the phylum are Acidobacteria (*Candidatus Koribacter*, *Candidatus Solibacter*, *Candidatus Chloracidobacterium*), Actinobacteria (*Micromonospora, Mycobacterium, Nocardiodes, Pseudonocardia, Saccharopolyspora, Streptomyces, Conexibacter, Thermoleophilum, Rubrobacter, Atopbium, Pimelobacter, Rhodococcus, Acidithiomicrobium*, *Corynebacterium, Gordonibacter, Aeromicrobium, Streptosporangium, Amycolatopsis*), Bacteroidetes (*Flexibacter*, *Chitinophaga*, *Terrimonas*), Chloroflexi (*Herpetosiphon, Ktedonobacter, Thermomicrobium, Chloroflexus, Roseiflexus, Dehalogenimonas*), Firmicutes (*Bacillus, Lactobacillus*), Gemmatimonadetes (*Gemmatimonas*), Nitrospirae (*Nitrospira*), Verrucomicrobia (*Pedosphaera*), Planctomycetes (*Gemmata, Isosphaera, Pirellula, Planctomyces, Rhodopirellula*, unclassified derived from Planctomycetaceae) and Proteobacteria (*Geoalkalibacter, Agrobacterium,* unclassified derieved from *Alphaproteobacteria, Betaproteobacteria, Deltaproteobacteria, Gammaproteobacteria, Proteobacteria*) (Figure 6 & 7).

Control soil sample is characterized by the relative abundance of Bacteria belonging to the major generas of each of the phyla are Acidobacteria (*Candidatus Koribacter*, *Candidatus Solibacter*, *Candidatus Chloracidobacterium*), Actinobacteria (*Micromonospora, Mycobacterium, Nocardiodes, Pseudonocardia, Saccharopolyspora, Streptomyces, Conexibacter, Thermoleophilum, Rubrobacter, Atopbium, Rhodococcus*), Bacteroidetes (*Flexibacter*, *Chitinophaga*), Chloroflexi (*Herpetosiphon, Ktedonobacter, Thermomicrobium, Chloroflexus, Roseiflexus, Dehalogenimonas*), Gemmatimonadetes (*Gemmatimonas*), Nitrospirae (*Nitrospira*), Verrucomicrobia (*Pedosphaera*), Planctomycetes (*Gemmata, Isosphaera, Pirellula, Planctomyces, Rhodopirellula*, unclassified derived from Planctomycetaceae) and Proteobacteria (unclassified derieved from *Alphaproteobacteria, Betaproteobacteria, Deltaproteobacteria, Gammaproteobacteria, Proteobacteria*) (Figure 6 & 7).

Table- 14- Distribution of phylum among the 4 soil samples

	Soybean	**Okra**	**Rice**	**Control**
Acidobacteria	0% (0)	2.93% (14405)	3.85% (20243)	7.45% (23749)
Actinobacteria	6.09% (32131)	13.73% (67514)	12.92% (67890)	8.30% (26458)
Bacteroidetes	16.10% (84986)	2.23% (10960)	2.45% (12857)	3.94% (12581)
Chloroflexi	0% (0)	2.71% (13363)	2.43% (12802)	1.69% (5381)
Cyanobacteria	0% (9)	0.72% (3523)	0.75% (3939)	0.63% (5381)
Firmicutes	55.32% (292018)	9.21% (45296)	10.44% (54859)	7.91% (25223)
Gemmatimonadetes	0% (0)	0.64% (3142)	0.43% (2244)	0.95% (3054)
Nitrospirae	0% (0)	0.45% (2228)	0.37% (1954)	0.83% (2660)
Planctomycetes	0% (5)	11.45% (56304)	7.14% (37536)	11.73% (37402)
Proteobacteria	9.67% (51032)	10.39% (51115)	11.37% (59733)	13.01% (41493)
Verrucomicrobia	0.09% (507)	1.13% (5570)	2.02%	1.10% (3525)

The data was compared by M5RNA with a maximum e-value of 0.001, a minimum identity of 85% and a minimum alignment length of 20. Number of hits is shown in brackets.

A non-parametric analysis of variance was used to statistically analyse the 16S sequence data at the phylum level using the freely available offline “R” software. The data was used to generate Bray-Curtis dissimilarity matrices and the statistical significance was tested using the non-parametric methods for testing the difference among a group of observations.

Table- 15- Non-parametric Analysis of Variance

	Df	Sum of Squares	Mean Sum of Squares	F- Model	R^2	Pr> (F)
Soil Type	1	0.13681	0.136807	2.2546	0.36047	0.01667*
Crop Type	2	0.12136	0.060680	1.0000	0.31977	0.60139
Residuals	2	0.12136	0.060687		0.31976	
Total	5	0.37952			1.00000	

**p=0.01 level of significance,*p=0.05 level of significance.

4.4 Screening for the presence of agriculturally important genes in the metagenome of the 3 disease suppressive soil samples and the control soil sample

The metagenome of the 3 disease suppressive soil samples and the control soil sample were screened for the presence of agriculturally important genes. Of, all the 6 primer pairs used for screening only the genes coding for the chitinase gene (Plate 8) and *cry14* gene (Plate 9) showed an amplified product size of 300bp and 400bp respectively.

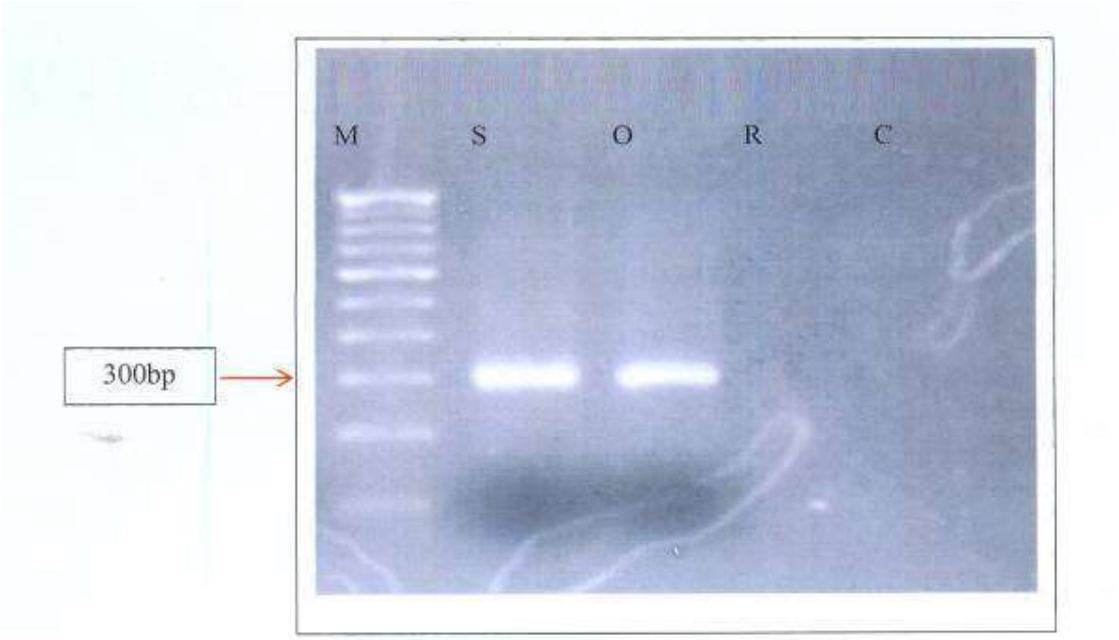


Plate 8: PCR amplification of *chia A* gene using the primer pairs ChiAF2 and ChiAR2 gives an amplified product of 300bp. M- 100bp DNA ladder.

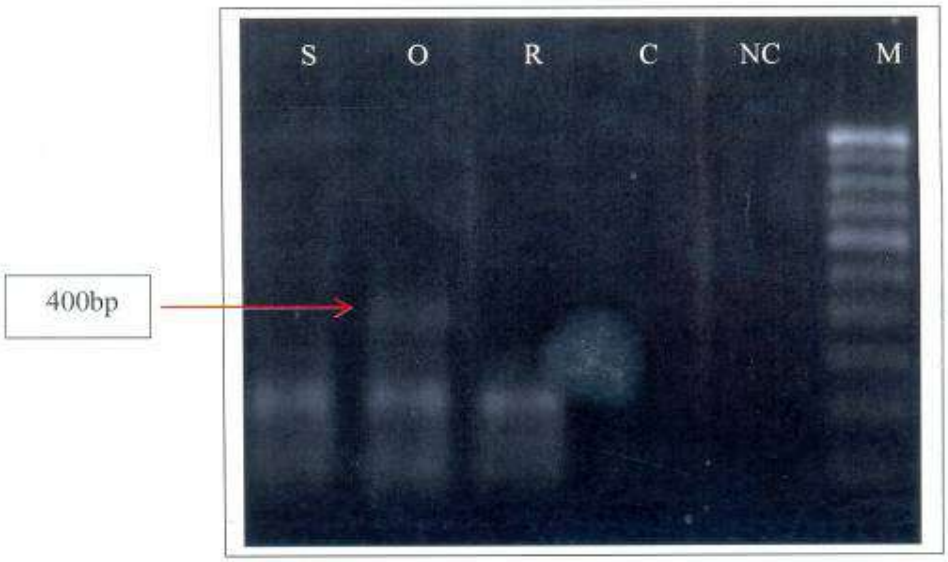


Plate 9: PCR amplification of *cry14* gene using the primer pairs Cry 14F and Cry14R gives an amplified product of 400bp.M- 100bp DNA ladder.

5. DISCUSSION

The study aimed at understanding the soil microbial community composition and structure of the disease suppressive soils in comparison with the control soil. The disease suppressive soils have been the subject of study for a last few decades as they are known to possess genes having potential role in controlling phytopathogens. The conventional tools have been used to identify microbial populations in such soils like fluorescent Pseudomonas which produces 2,4-DAPG to protect the wheat plants from the Take-all decline disease after continous monocropping (Mazzola, 2002). But, the genetic potential that is hidden in those disease suppressive soils need to be explored for identifying the potential genes that might have a role in disease suppression. Identification of such genes would lead to an effective management of phytopathogens in an eco-friendly way without harming the soil heath status.

The above study was conducted to understand the microbial community composition both in terms of species richness and diversity of three different disease suppressive soil samples in comparison with the control soil sample (nearby fallow land). Keeping the above considerations in view, different culture independent techniques like PCR-DGGE and Next generation sequencing technology of the Illumina MiSeq platform were used in studying the metagenomes of the four samples taken.

5.1 Metagenomic DNA yield and purification of isolated DNA

The metagenomic approach mostly relies on the extraction of high molecular weight unsheared microbial DNA which is representative of the original microbial population. Therefore the success of any metagenomics based study is highly dependent on the soil DNA extraction method used. The high degree of cell release from soil particles or in situ cell lysis efficiency, good quality and quantity of pure DNA obtained are considered as some of the important parameters that need to be taken under consideration while developing any soil DNA extraction protocol. Humic substances are a major component of soil organic matter. Though the chemical composition is highly complex, these compounds readily co-purify with DNA and are difficult to remove without additional, laborious and time intensive treatments to obtain DNA suitable for downstream molecular methods like that of PCR and restriction digestion (Romanowski *et al.*, 1992; Trevors and van Elsas, 1989). So as

to maximize cell lysis and in to minimize the contamination of inhibitors (humic substances), a protocol for metagenomic soil DNA extraction was standardized during the study, which was a combination of different available DNA extraction protocols. The DNA isolation and purification protocol standardized during this study resulted in obtaining pure high yield DNA, which can be employed for various molecular biological applications. The average yield obtained with this standardized protocol was in the range of 2-3 µg per 1 gram of soil and was quite less than that obtained by Braid *et al.,* (2003). The DNA yield in all the four samples was not significantly different from each other.

Purity ratio of the extracted DNA at 260/280 was 1.56-1.57, which was higher than that of metagenomic DNA extraction using CTAB (1.35) (Zhou *et al.,* 1996). The use of $CaCl_2$ as a chemical flocculant resulted in clear and transparent DNA i.e. free from humic acid contamination. Also, the overnight incubation of Isopropanol alongwith1/10^{th} volume of 3M Sodium Acetate for precipitation of DNA yielded higher quantity of DNA.

5.2 PCR amplification of the 4 soil metagenomic DNA samples

PCR amplification of the extracted soil DNA was done using three sets of primer pairs which targeted the variable regions of Bacteria and Eubacteria. The use of primer pairs PRBA 338F-PRUN518R and PRBA 968F-PRBA1406R targeting the V-3 and V-6 region of Bacteria gave an amplicon size of 200bp and 470bp respectively. The Eubacterial primer pair E 783F-E 926R targeting the variable region gave an amplicon size of 200bp.

5.3 DGGE Analysis of Disease Suppressive Soils in comparison with the Control Soil sample

Community-level molecular techniques are widely used in comparative microbial ecology to assess the diversity of microbial communities and their response to changing environments (Marzorati *et al*., 2008). Among the techniques widely used for study, DGGE is one of the most well established molecular tools in microbial ecology studies (Muyzer and Smalla, 1998; Muyzer *et al*., 1998; Boon *et al*., 2002). This fingerprinting technique is commonly used to provide, on a polyacrylamide gel, a profile representing the genetic structure and diversity of a microbial community from

a specific environment with a high versatility, reliability and reproducibility (Muyzer and Smalla, 1998).

For better interpretation of the data that gets generated out of a DGGE fingerprint, statistical tools are applied for calculation of different biodiversity indices like Shannon-Weaver Index (Gavin, *et al*, 2005), Sorenson's Similarity Index (Nakatsu, *et al*, 2000), etc.

The use of this culture independent technique based on PCR-DGGE coupled with different statistical analysis have given some significant insights into the microbial community composition (both in terms of similarity and diversity) of the four soil samples.

Sorenson's Pairwise similarity index values varied for each of the three primer pairs used in the study. The lesser similarity values were found for the primer pair PRBA 338F-PRUN 518R. All the three primer pairs used in the study indicated that the combination of Soybean and Okra soil samples shared the highest coefficient of similarity which might indicate that the soils infested with the same group of pathogens (nematode) might harbour similar microbial community composition.

The diversity indices as calculated using the Shannon-Weaver Index showed that there is a significant diversity ($p=0.01$) in the three suppressive soil samples in comparison with the control soil sample for the primer pair PRBA 338F- PRUN 518R and E 783F- E 926R. This data supports the hypothesis suggested by the culture based study, that there is a significant increase in a certain group of microorganisms in the disease suppressive soil sample which might have a role potential in suppressing the disease and in turn reducing the soil microbial diversity (Cook and Rovira, 1976; Simon and Sivasithamparam, 1989; Raajimakers *et al*., 1997, 1999).

The range weighted richness values for all the four soil samples using the three different primer pairs were found to be above 30, ranging from 30.72 to as high as 87.08. This indicates that the soil samples have a broader carrying capacity with high range weighted richness i.e. typical of a very habitable environment. As, these soil samples were collected from the root rhizosphere of three crop plants it can support higher number of individuals with the resources available in that environment (Marzorati, *et al*., 2008).

The functional organization (F_o) index as indicated by the Pareto-Lorenz Curve (Lorenz, 1905) showed that the combined relative abundances of 20% of the OTU's for the 3 disease suppressive soil samples in comparison with the control soil sample ranged from 68%- 92% for the three primer pairs. The more the PL-Curve deviates from the 45^o angle, less evenness can be observed in the structure of the studied community. Wittebolle *et al*., (2008) suggested that, the interpretation of these curves is based on the assumption that the distribution of species within a microbial community relates to the capacity to optimize and conserve functionality even under perturbed conditions. Therefore, based on this assumption it can be interpreted that, the area around 80% PL-Curve represents a specialized community in which a small amount of the species is dominant and all others are present in low numbers, with a large difference between the two groups. This community can be highly functionally organized (high Fo) but is fragile to external changes because disruption may require longer recovery times (Marzorati, *et al*., 2008).

5.4 Sequence Based Analysis of Disease Suppressive Soils in comparison with the Control Soil sample

Natural biological suppression of soil- borne diseases is a function of the activity and composition of soil microbial communities (Penton *et al*, 2014). Therefore, for better understanding of the microbial community composition, different culture independent approaches have been used till date to analyze such soils. These cultivation independent methods employ PCR-amplification, cloning and sequence analysis of 16SrRNA or other phyllogenetically informative genes have made it possible to assess the composition of microbial species in natural environments (Murat *et al*, 2011). Such a study was conducted by Sanguin *et al*., (2009) to understand the changes in rhizobacterial community composition during the Take-all decline of field-grown wheat. The study was based on development and utilization of a taxonomic 16S rRNA based microarray coupled with clonal sequencing and quantitative PCR.

Until recently, this approach has been too time consuming and expensive for routine use. With the advent of Next generation sequencing platforms, it has largely eliminated these obstacles by increasing sequencing capacity by orders of magnitude. A 454-pyrosequencing targeting the 28S LSU rRNA gene was used to

study the fungal community structure in *Rhizoctonia solani* suppressive soils in comparison with a non-suppressive soil (Penton *et al*., 2014).

Similarly, in this study 16S rDNA targeted amplicon sequencing (targeting the V-3 region) was done using the Illumina MiSeq platform for the three suppressive soil samples in comparison with the control soil sample. Phylogenetic profile was carried out using the M5RNA system databases in MG-RAST (Metagenome Rapid Annotation using Subsystem Technology). For providing information regarding taxonomic diversity, the MG-RAST server utilizes different reference databases for their analysis. The highlights of the data analyzed can be summarized as follows-

The total DNA from the Soybean, Okra, Rice soil samples and the Control soil sample yielded 354, 381, 294 and 281Mb of data respectively. The distribution of the domains as calculated from of the data identified bacteria as the most dominant phyla followed by eukaryota.

The relative abundance of the three disease suppressive soil samples in comparison with the control sample at phylum level showed Firmicutes to be the most dominant phyla followed by Proteobacteria and Actinobacteria. The alpha diversity was found to be 20.72, 58.36, 72.09 and 86.3 for the Soybean, Okra and Rice and control soil sample respectively. This data indicates, highest microbial diversity is present in the control soil sample in comparison with the three disease suppressive soil samples. The analysis of the heatmap dendogram indicated that at phylum level, the Control and Rice soil samples share highest similarity in terms of microbial community composition thereby contradicting the result obtained from the PCR-DGGE analyses(Sorenson's Pairwise similarity index) using the same primer pair. This may be due to the fact that while calculating the Sorensons's Index visual presence and absence of OTU's were taken into consideration and not the relative intensity of each OTU.

Genera level classification of 3 disease suppressive soil samples in comparison with the control soil sample based on number of reads per sample indicated the higher relative abundance of *Paenibacillus, Heliobacterium, Bacillus, Lactobacillus, Acetivibrio, Amycolatopsis*, *Isoptericola*, *Micromonospora*, *Nocardia*, *Pseudonocardia* and *Candidatus Solibacter*, *Candidatus Chloracidobacterium*,

Atopbium, *Prevotella*, *Dialister*, *Lactobacillus*, *Ktedonobacter*, *Pedospharea* and *Acetobacter*.

Higher relative abundance of Candidatus *Solibacter* belonging to the Acidobacter phyla was observed in the case of Okra and Rice soil samples. This group of microorganisms in soil is reported to be capable of nitrate and nitrite reduction. Another group of Acidobacteria i.e Candidatus *Chloracidobacterium* present in the Okra and Rice soil samples is an aerobic phototrophic *Acidobacterium* capable of chlorophyll (Chl)–based phototrophy (Bryant *et al*., 2007).

Among the Actinobacteria phyla, higher relative abundance of *Amycolatopsis*, *Isoptericola*, *Micromonospora*, *Nocardia, Heliobacterium, Atopbium*, *Corynebacterium* and *Pseudonocardia* was observed in the three disease suppressive soil samples.

The higher relative abundance of the genera *Amycolatopsis*, *Isoptericola*, *Micromonospora*, *Nocardia* and *Pseudonocardia* belonging to the Actinobacteria phyla has been observed in the Okra and Rice soil samples. These genera of bacteria have been reported to have antagonistic activity against Soybean pathogen *Xanthomonas campestris* pv. *glycine* by production of some antifungal and antibacterial compounds (Minigma *et al*., 2014**)**.

Higher relative abundance of *Corynebacterium* genus in the Rice soil sample belonging to the Actinobacteria phyla was observed. The organism is known to harbor Lantibiotic gene clusters (gene encoded antimicrobial peptides) in the lantibiotic encoding operons (Orla *et al*., 2011).

Higher prevalence of *Heliobacterium* genus (*Heliobacterium modesticaldum*) in the three disease suppressive soil samples than the control soil sample was observed. The organism is an anoxygenic phototroph which is known to fix atmospheric nitrogen in the plant thereby promoting plant growth (Matthew *et al*., 2008).

Among the Bacteroidetes phyla the genus *Prevotella* was found to be abundantly available in the Soybean soil sample. *Prevotella paludivivens* is an anaerobic, gram negative, hemicellulose decomposing bacteria that have been reported to be associated with the rice root rhizosphere. This is quite interesting as

this group of microorganism is known to decompose organic matter in the form of rice straw in the field under anoxic conditions (Atsuko *et al*., 2007).

Among the Firmicutes phyla, higher relative abundance of *Paenibacillus*, *Lactobacillus*, *Sporoanerobacter*, *Dialister* and *Bacillus* was observed. Slightly higher relative abundance of *Bacillus* genus (*Bacillus cereus*) in Okra and Rice samples in comparison with the Control soil sample is seen. *Bacillus cereus* is known to produce two antibiotics which reversibly are known to affect the growth of Damping off of Alfa-alfa (Laura *et al*., 1993).

Higher relative abundance of *Paenibacillus* genus in the disease suppressive soil samples in comparison with the control soil sample is quite interesting as this microorganism is known to suppress the activity of *Botrytis cinerea* (grey mold of strawberry) (Helbig *et al*., 2001), *Fusarium oxysporum* and *Rhizoctonia solani* damping off of sesame (Ryu *et al*., 2006).

Among the Chloroflexi phyla, higher relative abundance of *Ktedonobacter* genus (*Ktedonobacter racemifer*) was observed in the Okra and Rice soil samples in comparison with the Control soil sample. *Ktedonobacter* is reported to harbor chitinolytic enzymes when it was isolated from a metagenomic library of Chitin amended soil samples (Mariana, *et al*., 2015).

The relative abundance of *Pedosphaera* genus (*Pedosphaera parvula*) belonging to the Verrucomicrobiae phyla was observed in the Okra, Rice and the Control soil sample. This bacterium is known to produce an antibacterial protein which inhibits the growth of *Bacillus subtilis* (Mark *et al*., 2015).

Among the Proteobacteria phyla, higher relative abundance of *Acetobacter* genus (*Acetobacter indonesiensis*) was observed in the Soybean soil sample. The *Acetobacter* have been reported to be the potential biocontrol agent of *Fusarium oxysporum* f. sp. *melonis* and *Xanthomonas campestris* both under in vivo and in vitro studies (Suarez-Estrella *et al*., 2013).

The screening of the metagenomes for the three disease suppressive soils in comparison with the control soil sample showed that the bacterial diversity at different taxonomic levels. The higher relative abundance of these groups of microorganisms may have some role in reducing the phytopathogenic activity of the suppressive soil

samples. But simple relative abundance cannot implicate that the transition of the bacterial community is due to disease suppressiveness characteristic of the soil sample and not due to crop type.

Therefore, for better understanding of the bacterial distribution pattern in the three disease suppressive soil samples, the non-parametric analysis of variance (Anderson, 2001) was carried out using the relative abundance of the different phyla for the four soil samples. The statistical analyses proved that there is a strong correlation between the relative abundance of bacterial phylum (5% level of significance) and the disease suppressive soil type. Therefore, the relative abundance of each of the phylum of the three disease suppressive soil samples in comparison with the control soil sample can be strongly attributed to the disease suppressiveness characteristic of the soil and not due to the crop type.

5.5 Screening for the presence of agriculturally important genes in the Disease suppressive soil samples

The metagenomes of the 3 disease suppressive soil samples alongwith the control soil sample was screened for the presence of agriculturally important genes which might have a role in controlling phytopathogens. The metagenomes were screened for the presence of chitinase genes, polyketide synthase genes, glutamate dehydrogenase genes and different cry genes. Of them, only chitinase genes (chitinase screening primers designed by Hobel, *et al*., 2005) were found to be present in the Soybean and Okra soil samples giving an amplicon size of 300bp. Hjort, *et al*., (2010) after sequencing of the clones having the chitinase genes revealed that the chitinase gene sequences gives best database matches to 4 bacterial Firmicutes group, β- Proteobacteria, γ- Proteobacteria and Actinobacteria. Therefore, the presence of the chitinase genes in the 2 disease suppressive soil samples can be supported by the sequencing data that is generated by the targeted amplicon sequencing of the above samples.

SUMMARY AND CONCLUSION

The above study was conducted to understand the microbial community composition and structure of the three disease suppressive soil samples in comparison with the control soil sample. Different culture independent techniques were used in the study for better understanding of the microbial community composition of the four soil samples. The study concentrated mainly on understanding the soil microbial diversity in the three disease suppressive soil samples and the control soil samples with a major focus on identifying some specific group of microorganisms or genes which might have a role in suppressing the phytopathogens. The results obtained are summarized as below;

- At first the protocol was standardized for isolation of high molecular weight microbial DNA from the four soil samples.
- The purified DNA from control and three disease suppressive soils were directly used for PCR and were sufficient in getting effective PCR amplification using three distinct sets of primer pairs, which indicated that the quality of the DNA extracted was good enough for further downstream molecular biology applications.
- PCR-DGGE analysis was done for four soil samples using the two primer pairs of bacteria (PRBA 338F-GC- PRUN 518R & PRBA 968F-GC- PRBA 1406R) and one Eubacterial primer pair (E 783F-GC- E 926R). DGGE analysis concluded coverage of maximum diversity using bacterial specific primer PRBA338-PRUN518, which was higher than rest primer pairs, i.e., bacterial specific primer PRBA968-PRBA1406, eubacteria specific primer E783-E926.
- Higher microbial diversity was observed in case of the Control soil sample in comparison with the Disease suppressive soil samples. This data supports the hypothesis suggested by the culture based study, that there is a gradual increase in a certain group of microorganisms in the disease suppressive soil sample which might have a potential role in suppressing the disease and in turn reducing the soil microbial diversity.
- High range weighted richness was observed for the four soil samples indicating the presence of a very habitable environment with a broad carrying capacity.
- The Fo index indicated the presence of a specialized community in the four soil samples in which a small amount of the species is dominant and all others are

present in low numbers, with a large difference between the two groups. This community can be highly functionally organized (high Fo).

- Sequencing results show that distribution of the domains as calculated from the data of identified bacteria as the most dominant phyla followed by eukaryota.
- The relative abundance of the three disease suppressive soil samples in comparison with the control sample at phylum level showed Firmicutes to be the most dominant phyla followed by Proteobacteria and Actinobacteria.

- Genera level classification of three suppressive soil samples in comparison with the control soil sample based on number of reads per sample indicated the higher prevalence of Candidatus *Solibacter, Amycolatopsis, Isoptericola, Micromonospora, Nocardia, Heliobacterium, Atopbium, Corynebacterium, Pseudonocardia, Prevotella, Bacillus, Lactobacillus, Acetobacter* and *Ktedonobacter.*
- The relative abundance of the genera level classification of different phyla identified the presence of few genera that might have a potential role in controlling phytopathogens.
- Higher prevalence of *Paenibacillus* and *Bacillus* genera in the Suppressive soil samples in comparison with the control soil sample is quite interesting as this microorganism is reported to suppress the activity of some phytopathogens.
- The higher relative abundance of the genera *Amycolatopsis*, *Isoptericola*, *Micromonospora*, *Nocardia* and *Pseudonocardia* belonging to the Actinobacteria phyla have been reported to have antifungal and antibacterial activity against Soybean pathogen *Xanthomonas campestris* pv. *glycine*.
- Higher relative abundance of *Ktedonobacter* genus (*Ktedonobacter racemifer*) was observed in the Okra and Rice soil samples in comparison with the Control soil sample have been reported to harbor chitinolytic enzymes.
- Higher prevalence of *Heliobacterium* genera (*Heliobacterium modesticaldum*) in the three suppressive soil samples in comparison was observed. The above organism is an anoxygenic phototroph which is known to fix atmospheric nitrogen in the in the plant thereby promoting plant growth.
- Higher relative abundance of *Acetobacter* genus (*Acetobacter indonesiensis*) was observed in the Soybean soil sample have been reported to be the potential

biocontrol activity against *Fusarium oxysporum* f. sp. *melonis* and *Xanthomonas campestris*.

- The non-parametric analysis of variance carried out using the relative abundance of the different phyla for the four samples indicated that there is a strong correlation between the relative abundance of the bacterial phylum (5% level of significance) and the disease suppressive soil type.
- The PCR screening of the metagenomes for the presence of agriculturally important genes showed the presence of chitinase genes in the Soybean and Okra soil samples.

Future Line of Work

- The cultivation strategies for in vitro screening of pathogens can be searched for the generas identified in the disease suppressive soil samples like *Paenibacillus*, *Heliobacterium*.
- The chitinase genes which amplified in the soybean and okra soil samples can be cloned and can be used for screening of functional metagenomic libraries for the presence of chitinase genes in it.
- Transcriptome analysis of nucleic acid extracted directly from soil will be helpful in studying in particular those genes which are involved in the disease suppression.

REFERENCES

Albuquerque, P., Mendes, M. V., Santos, C. L., Moradas-Ferreira, P. and Tavares, F., 2009, DNA signature-based approaches for bacterial detection and identification. *Sci. Total Environ.*, 407 **:** 3641-3651.

Anderson, M.J., 2001, A new method for non-parametric multivariate analysis of variance. *Austral. Ecol.*, 26 : 32-46.

Atsuko, U., Hiroshi, A., Atsuya, S.,Daisuke, S. and Katsuji, U., 2007, *Prevotella paludivivens* sp. nov., a novel strictly anaerobic, Gram-negative, hemicellulose decomposing bacterium isolated from plant residue and rice roots in irrigated rice-field soil. *Int. J. Syst. Evol. Microbiol.*, 57 : 1803-1809.

Ayuso-Sacido A. and Genilloud, O., 2005, New PCR primers for screening of NRPS and PKS-I systems in actinomycetes: detection and distribution of these biosynthetic gene sequences in major taxonomic groups. *Microbiol. Ecol.*, 49 : 10-24.

Boleda, M. D., Briones , P., Farres, J., Tyfield, L. and Pi, R., 1996, Experimental design: A useful tool for PCR optimization. *Biotech.*, 21 : 134-140.

Boon, N., Windt, D. W., Verstraete, W. and Top, E.M., 2002, Evaluation of Nested PCR-DGGE (denaturing gradient gel electrophoresis) with group specific 16SrRNA primers for analysis of bacterial communities from different wastewater treatment plants. *FEMS Microbiol. Ecol*. 39 : 101-112.

Borneman, J. and Becker, J.O., 2007, Identifying microorganisms involved in specific pathogen suppression- in soil. *Annu. Rev. Phytopathol*. 45 : 153-172.

Borresen, A.L., Hovig, E. and Brogger, A., 1988, Detection of base mutations in genomic DNA using denaturing gradient gel-electrophoresis (DGGE) followed by transfer and hybridization with gene-specific probes. *Mutat. Res*. 202 : 77–83.

Braid, M. D., Daniels, L. M. and Kitts, C. L., 2003, Removal of PCR inhibitors from soil DNA by chemical flocculation. *J. Microbiol. Meth.,* 52 : 389-393.

Bravo, A., Sergio, S., Lorena, L., Hernesto, O., Carolina, A., Anabel, O., Miriam, O., Laura, L., Francisco, J.V., Guadalupe, P., Maria, E.N.V., Mario, S. and Rodolfo, Q., 1998, Characterization of *cry* genes in a Mexican *Bacillus thuringiensis* strain collection. *Appl. Environ. Microbiol*., 64(12) : 4965-4972.

Broadbent, P. and Baker, K.F., 1974, Behaviour of *Phytophthora cinnamoni* in soils suppressive and conducive to root rot. *Aust. J. Agric. Res*., 25 : 121-137.

Bryant, D.A., Amaya, M. G. C., Julia, A. M., Aline, G.M. C., Christian, G. K., Mary, M.B., Luke, J. T., Jessica, H., William, C. N., John, F. H. and David, M. W., 2007, *Candidatus Chloroacidobacterium thermophilum*: An aerobic phototrophic Acidobacterium. *Science*, 317 : 523.

Cook, R. J. and Rovira, A.D., 1976, The role of bacteria in the biological control of *Gaeumannomyces graminis* by suppressive soils. *Soil Biol. Biochem*. 8 : 267-273.

Degrange, V., and Bardin, R., 1995, Detection and counting of *Nitrobacter* populations in soil by PCR. *Appl. Environ. Microbiol*., 61 : 2093–2098.

Duarte, S., Pascoal, C., Garabétian, F., Cássio, F. and Charcosset, J.Y., 2009, Microbial decomposer communities are mainly structured by the trophic status in circumneutral and alkaline streams. *Appl. Environ. Microbiol*., 75 : 6211-6221.

Duijff, B.J., Pouhair, D., Olivain, C., Alabouvette, C. and Lemanceau, P., 1998, Implication of systemic induced resistance in the suppression of Fusarium wilt of tomato by *Pseudomonas flourescens* WCS417r and non-pathogenic *Fusarium oxysporum* Fo47. *Eur. J. Plant Pathol*., 104 : 903-910.

Duijff, B.J., Recorbert, G., Baker, P.A.H.M., Loper, J.E. and Lemanceau, P., 1999, Microbial antagonism at the root level is involve in the suppression of Fusarium wilt by the combination of non-pathogenic *Fusarium oxysporum* Fo47 and *Pseudomonas putida* WCS358. *Phytopathology*, 89 : 1073-1979.

Duineveld, B. M., Rosado, A. S., van Elsas, J. D. and van Veen, J. A., 1998, Analysis of the dynamics of bacterial communities in the rhizosphere of the chrysanthemum via denaturing gradient gel electrophoresis and substrate utilization patterns. *Appl. Environ. Microbiol.,* 64 : 4950-4957.

Duineveld, B.M., Kowalchuk, G.A., Keijzer, A., van Elsas, J.D. and van Veen, J.A., 2001, Analysis of bacterial communities in the rhizosphere of chrysanthemum via denaturing gradient gel electrophoresis of PCR-amplified 16S rRNA as well as DNA fragments coding for 16S rRNA. *Appl. Environ. Microbiol*., 67 : 172–178.

Ejiofar, A.O. and Johnson, T., 2002, Physiological and molecular detection of crystalliferous Bacillus thuringiensis strain from habitats in the South Central United States. *J. Indu. Microbiol. Biotechnol*., 28(5) : 284-290.

Faegri, A., Torsvik, V. L. and Goksoyr, J., 1977, Bacterial and fungal activity in soil: separation of bacteria and fungi by a rapid fractionated centrifugation technique. *Soil Biol. Biochem*., 9 : 105-112

Fernandez, A. S., Hashsham, S. A., Dollhopf, S. L., Raskin, L., Glagoleva, O., Dazzo, F. B., Hickey, R. F., Criddle, C. S. and Tiedje**,** J. M.**,** 2000, Flexible community structure correlates with stable community function in methanogenic bioreactor communities perturbed by glucose. *Appl. Environ. Microbiol*., 66 : 4058–4067.

Fischer, S. G. and Lerman, L. S., 1979, Length independent separation of DNA restriction fragments in two dimensional gel electrophoresis. *Cell*, 16 : 191–200.

Fischer, S. G. and Lerman, L.S., 1983, DNA fragments differing by single base pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. *Proc. Natl. Acad. Sci. USA*, 80 : 1579–1583.

Flemming, C. A, Leung, K., Lee, H., Trevors, J. T. and Greer, C., 1994, Survival of lux-lac marked biosurfactant-producing *Pseudomonas aeruginosa* UG2L in soil monitored by non-selective plating and polymerase chain reaction. *Appl. Environ. Microbiol.,* 60 : 1606-1613.

Foerstner, K.U., Mering, V. C., Hooper, S. D., and Bork, P., 2005, Environments shape the nucleotide composition of genomes. EMBO Rep 6 : 1208–1213.

Fromin, N., Hamelin, J., Tarnawski, S., Roesti, D., Jourdain-Miserez, K., Forestier, N., Teyssier-Cuvelle, S., Gillet, F., Aragno, M. and Rossi, P., 2002, Statistical analysis of denaturing gel electrophoresis (DGE) fingerprinting patterns. *Environ. Microbiol*., 4 : 634 643.

Frostegard, A., Courtois, S., Ramisse, V., Clerc, S., Bernillon, D., Le Gall, F., Jeanin, P., Nesme, X. and Simonet, P., 1999, Quantification of bias related to the extraction of DNA directly of soils. *Appl. Environ. Microbiol*., 65 : 5409-5420.

Gavin, P.G., Lucas, V.S., Graham, J.R., Petrie, A., Wilson, M. and Spratt, D.A., 2005, Statistical analyses of complex Denaturing Gradient Gel Electrophoresis profiles. *J. Clin. Microbiol*., 43(8) : 3971-3978.

Gustavo, C. A., Peter, M.G. and Basam, B.J., 1991, Fast and sensitive silver staining of DNA in Polyacrylamide gels. *Anal. Biochem*, 196 : 80-83.

Hancock, J.G., 1977, Factors affecting soil populations of *Pythium ultimum* in the San Joaquin Valley of California. *Hilgardia*, 45 : 107-122.

Handelsman, J., 2004, Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol. Mol. Biol. Rev*., 68(4) : 669.

Helbig, J., 2001, Biological control of *Botrytis cinerea* Pers. ex Fr. in strawberry by *Paenibacillus polymyxa* (isolate 18191). *J. Phytopathol*. 149 : 265-273.

Henis, Y., Ghaffar, A., and Baker, R., 1978, Integrated control of *Rhizoctonia solani* damping-off of radish: effect of successive plantings, PCNB and *Trichoderma harzianum* on pathogen and disease. *Phytopathology*, 68 : 900-907.

Henis, Y., Ghaffar, A., and Baker, R., 1979, Factors affecting suppressiveness to *Rhizoctonia solani* in soil. *Phytopathology*, 69 : 1164-1169.

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.

Heuer, H., Krsek, M., Baker, P., Smalla, K. and Wellington, E. M. H., 1997, Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel electrophoretic separation in denaturing gradients. *Appl. Environ. Microbiol*., 63 : 3233-3241.

Hjort, K., Haria, P., Annelie, E., Flavia, M. and Sara, S., 2013, Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics. *Appl*. *Microbiol*. *Biotechnol*., 10.

Hjort, K., Maria, B., Modupe, F.A., Janet, K.J., Kornelia, S. and Sara, S., 2010, Chitinase genes revealed and compared isolates, DNA extracts and metagenomic library from a phytopathogen-suppressive soil. *FEMS Microbiol. Ecol*., 71 : 197-207.

Hobel, C. F. V., Viggo, T. M., Gudmundur, □. H. and Jakob, K. K., 2005, Investigation of the Microbial ecology of Intertidal hot springs by using diversity analysis of 16SrRNA and Chitinase genes. *Appl*. *Environ*. *Microbiol*., 71(5) : 2771-2776.

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.

Hovig, E., Smithsorensen, B., Brogger, A. and Borresen, A. L., 1991,Constant denaturant gel-electrophoresis, a modification of denaturing gradient gel electrophoresis, in mutation detection. *Mutat. Res*. 262 : 63–71.

Hurt, R. A., Qiu, X. Y., Wu, L. Y., Roh, Y., Palumbo, A. V., Tiedje, J. M. and Zhou, J. H., 2001, Simultaneous recovery of RNA and DNA from soils and sediments. *Appl. Environ. Microbiol*., 67 : 4495-4503.

Johnsen, K., Jacobsen, C. S., Torsvik, V. and Sørensen, J., 2001, Pesticide effects on bacterial diversity in agricultural soils—a review. *Biol. Fertil. Soils,* 33 : 443–453.

Keel, C., Schnider, U., Maurhofer, M., Voisard, C., Laville, J., Burger, P., Wirthner, P., Haas, D. and Defago, G., 1992, Suppression of root diseases by *Pseudomonas flourescens* CHAO: importance of the secondary metabolite 2,4-diacetylphloroglucinol. *Mol. Plant-Microbe Interact*. 5 : 4-13.

Kerry, B. R., 1988, Fungal parasites of Cyst nematodes. *Agric. Eco-syst. Environ*. 24 : 293-305.

Kloepper, J. W., Leong, J., Teintze, M. and Schroth, M.N., 1980, Enhanced plant growth by siderophores produced by plant growth-promoting rhizobacteria. *Nature*, 286 : 885-886.

Knaebel, D.B. and Crawford, R.L., 1995, Extraction and purification of microbial DNA from petroleum-contaminated soils and detection of low numbers of toluene, octane and pesticide degraders by multiplex polymerase chain reaction and Southern analysis. *Mol. Ecol*. 4(5) : 579-591.

Kowalchuk, G. A., Barbara, D., Etienne, Y. and van Veen, J. A., 2006, Assessing bacterial and fungal community structure in soil using ribosomal RNA and other structural gene markers. Soil Biology, Nucleic Acids and Proteins in Soil P. Nannipieri, K. Smalla (Eds.) ©Springer-Verlag Berlin Heidelberg, 8 : 159-186.

Kraus, J. and Loper, J. E., 1995, Characterization of genomic region required for production of the antibiotic pyoluteorin by the biological control agent *Pseudomonas flourescens* Pf. 5. *Appl. Environ. Microbiol*., 61 : 849-854.

Lane, D. J., 1991, 16S/23S rRNA sequencing: In Nucleic Acid Techniques in Bacterial Systematics (ed. E. Stackebrandt and M. Goodfellow): Chichester: John Wiley, 184-189.

Lane, D. J., Field, K. G., Olsen, G. J. and Pace, N. R., 1988, Reverse transcriptase sequencing of ribosomal RNA for phylogenetic analysis. *Meth. Enzymol*., 167 : 138-144.

Laura, A., Silo-Suh, Benjamin, J. L., Sandra, J. R., Haiyin, H., John, C. and Handelsman, J., 1993, Biological activities of two fungistatic antibiotics produced by *Bacillus cereus* UW85. *Appl. Environ. Microbiol.,* 60(6) : 2023-2030.

Lemanceau, P., Baker, P.A.H.M, Kogel, D. W. J., Alabouvette, C. and Schippers, B., 1993, Antagonistic effect of non-pathogenic *Fusarium oxysporum* strain Fo47 and pseudobactin 358 upon pathogenic *Fusarium oxysporum* f.sp. *dianthi*. *Appl. Environ. Microbiol*., 59 : 74-82.

Lombard, N.**,** Prestat, E., van Elsas, J.D., and Simonet, P., 2011, Soil specific limitations for access and analysis of soil microbial communities by metagenomics. *FEMS Microbiol. Ecol*., 78 : 31-49.

Lorenz, M.O., 1905, Methods of measuring concentration of wealth. *J. Am. Stat. Assoc.,* 9 : 209–219.

Mariana, S. C., Francesca, B., Anna, M.K., Flavia, M. and van Elsas, J.D., 2015, A novel salt-tolerant chitobiosidase discovered by genetic screening of a metagenomic library derived from chitin-amended agricultural soil. *Environ. Biotechnol*.

Mark, M. O. M., Ruth, H., Joseph, S., Jonathan, K., Fiona, D., Julian, R. M. and Alan, D. W. D., 2015, Inhibition of the growth of *Bacillus subtilis* DSM10 by a newly discovered antibacterial protein from the soil metagenome. *Bioengineered*, 6(2) : 89-98.

Marzorati, M., Wittebole, L., Boon, N., Daniele, D. and Verstrate, W., 2008, How to get more out of molecular fingerprints: practical tools of microbial ecology. *Environ. Microbiol*., 10 : 1462-2920.

Matthew, W. S., Michael, T. M., Wesley D. S., Patricia C. C., Kate M. C., Amber, L. C., Liza, C. D., Barbara, M. H., Deborah, O. J., Lauren, E. K., Ahmet, K., Surobhi, L., Stephen, D. M., Lawrence, E. P., Heather, L. T., Zi, T. W., Jason, R., Min, C., Robert E. B. and Jeffrey, W. T., 2008, The Genome of *Heliobacterium modesticaldum*, a Phototrophic representative of the *Firmicutes* containing the simplest photosynthetic apparatus. *J. Bacteriol*., 190(13) : 4687–4696.

Mazzola, M**.,** 2002, Mechanism of natural soil suppressiveness to soil borne diseases. *Antonie van Leewenhooek*, 81 : 557-564.

Menzies, J. D., 1959, Occurrence and transfer of a biological factor in soil that suppress potato scab. *Phytopathology*, 49 : 648-652.

Mille-Lindblom, C., Fischer, H. and Tranvik, L. J., 2006, Litter-associated bacteria and fungi-a comparison of biomass and communities across lakes and plant species. *Freshwater Biol.*, 51 : 730-741.

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

Minigma, R., Pathom-aree, W., Savitr, T., Arinthip, T., Kannika, D., 2014, Isolation of rhizospheric and roots endophytic actinomycetes from *Leguminosae* plant and their activities to inhibit soybean pathogen, *Xanthomonas campestris* pv. *glycine*. *World J. Microb. Biot.*, 30(1) : 271-280.

More, M. I., Herrick, J. B., Silva, M. C., Ghiorse, W. C. and Madsen, E. L., 1994, Quantitative cell lysis of indigenous microorganisms and rapid extraction of microbial DNA from sediment. *Appl. Environ. Microbiol.*, 60(5) : 1572–1580.

More., 2011, Metagenomic analysis of forest and farm soil. Thesis.

Murakami H., Tsushima, S. and Shishido, Y., 2000, Soils suppressiveness to clubroot disease of Chinese cabbage caused by *Plasmodiophora brassicae*. *Soil Biol. Biochem.*, 32 : 1637-1642.

Murat, E. A., Ferris, M. J. and Taylor, C. M., 2011, A framework for analysis of metagenomic sequencing data. *Biocomputing*.

Muyzer, G. and Ramsing, N. B., 1995, Molecular methods to study the organization of microbial communities. *Wat. Sci. Tech.,* 32 : 1-9.

Muyzer, G. and Smalla, K., 1998, Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie Van Leeuwenhoek,* 73 : 127–141.

Muyzer, G., 1999, DGGE/TGGE a method for identifying genes from natural ecosystems. *Curr. Opin. Microbiol.*, 2 : 317-322.

Muyzer, G., Brinkhoff, T., N¨ubel, U., Santegoeds, C., Sch¨afer H and Wawer, C., 1997, Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. In: Akkermans ADL, van Elsas JD and de Bruijn FJ (Eds) Molecular Microbial Ecology Manual (3.4.47 1-2 Kluwer Academic Publishers, Dordrecht, The Netherlands.

Muyzer, G., Brinkhoff, T., Nubel, U., Santagoeds, C., Schafer, H. and Wawer, C., 1998, Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. *Mol. Microbiol. Ecol. Man.* 3.4.4 **:** 1–27.

Muyzer, G., de Waal, E. C. and Uitterlinden, A.G., 1993, Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis and polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.*, 59 : 695–700.

Myers, R. M., Maniatis, T. and Lerman, L. S., 1987, Detection and localization of single base changes by denaturing gradient gel electrophoresis. *Methods Enzymol.*, 155 : 501-527

Myers, R. M., Sheffield, V. C. and Cox, D. R., 1989, Mutation detection by PCR, GC-clamps, and denaturing gradient gel electrophoresis. In: Erlich HA (Ed) PCR Technology – Principles and Applications for DNA amplification Stockton Press, New York. 71-88.

Myers, R.M., Fischer, S.G., Lerman, L.S. and Maniatis, T., 1985, Nearly all single base substitutions in DNA fragments joined to a GC-140 clamp can be detected by denaturing gradient gel electrophoresis. *Nucleic Acids Res.*, 13 : 3131–3145

Nakatsu, C. H., Torsvik, V. and Øvreås, L., 2000, Analysis using denaturant gradient gel electrophoresis (DGGE) of 16S rDNA PCR products. *Soil Sci. Soc. AM. J.*, 46 : 1382-1388.

Nikolcheva, L.G., Cockshutt, A.M. and Bärlocher, F., 2003, Determining diversity of freshwater fungi on decomposing leaves: comparison of traditional and molecular approaches. *Appl. Environ. Microbiol*., 69 : 2548-2554.

Nu¨bel, U., Engelen, B., Felske, A., Jiri S., Alois, W., Amann, R. I., Wolfgang, L., and Horst, B., 1996, Sequence Heterogeneities of Genes Encoding 16S rRNAs in *Paenibacillus polymyxa* detected by Temperature Gradient Gel Electrophoresis. *J. of Bacteriol*., 178 (19) : 5636–5643.

Nüsslein, K. and Tiedje, J.M., 1999, Soil bacterial community shift correlated with change from forest to pasture vegetation in a tropical soil. *Appl. Environ. Microbiol.,* 65 **:** 3622–3626.

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

Orla, O. S., Maire, B., Ross, R. P., Paul, D. C. and Colin, H., 2011, Further identification of novel Lantibiotic operons using LanM-based genome mining. *Probiotics Antimicrob. Proteins*, 3(1) : 27-40.

Overas, L., Forney, L. Daae, F. L. and Torsvik, V., 1997, Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR – amplified gene fragments coding for 16Sr RNA. *Appl. Environ .Microbiol.,* 63 : 3367-3373.

Peer, V. D., Robbrecht, Y. E., de Hoog, S., Caers, A., de Rijk, P. and de Wachter, R, 1999, Database on the structure of small subunit ribosomal RNA. *Nucleic Acids Res*, 27 : 179–183.

Penton, C. R., Gupta, V. V. S. R., Tiedje, J. M., Neate, S. M., Ophel-Keller, K., Gillings, M, Harvey, P., Pham, A. and Roget, D. K., 2014, Fungal community structure in Disease Suppressive Soils Assessed by 28S LSU Gene Sequencing. *Plos One*, 9(4) : e93893.

Poulsen, L. K., Ballard, G. and Stahl, D.A., 1993, Use of ribosomal-RNA fluorescence in situ hybridization for measuring the activity of single cells in young and established biofilms. *Appl. Environ. Microbiol*., 59 : 1354–1360

Pujol, C. and Kado, C., 1999, *gdhB*, a gene encoding a second quinoprotein glucose dehydrogenase in *Pantoea citrea*, is responsible for pink disease of pineapple. *Microbiol*., 145 : 1217-1226.

Raaijmakers, J.M., van der Sluis I, Koster, M., Baker, P.A.H.M., Weisbeek, P.J. and Schippers, B., 1995, Utilization of heterologous siderophores and rhizosphere competence of fluorescent *Pseudomonas* spp. *Can. J. Microbiol*., 41 : 126-135.

Raajimakers, J. M., Weller, D.M. and Thomashow, L. S., 1997, Frequency of antibiotic producing Pseudomonas spp. in natural environments. *Appl. Environ. Microbiol*., 63 : 881-887.

Raajimakers, J.M. Bonsall, R.F. and Weller, D.M., 1999, Effect of population density of *Pseudomonas flourescens* on production of 2, 4-diacetylpholoroglucinol in the rhizosphere of wheat. *Phytopathology*, 89 : 470-475.

Robe, P., Nalin, R., Capellano, C. and Vogel, T. M. 2003, Extraction of DNA from soil. *Eur. J. Soil Biol*., 39 : 183-190.

Romanowski, G., Lorenz, M. G., Sayler, G. and Wackernagel, W., 1992, Persistence of free plasmid DNA in soil monitored by various methods, including a transformation assay. *Appl. Environ. Microbiol*., 58 : 3012–3019.

Rouxel, F., Alabouvette, C. and Louvet, J., 1979, Recherches sur la rèsist-ance des sols aux maladies. IV. Mise en èvidence du r□le des Fusarium autochtones dans la rèsistance d'un sol á la fusriose vasculaire du melon. *Ann. Phytopathol*., 11 : 199-207.

Ryu, C-M, Kim, J., Choi, O., Kim, S.H. and Park, C.S., 2006, Improvement of biological control capacity of *Paenibacillus polymyxa* E681 by seed pelleting on sesame. *Biol. Control*, 39 : 282-289.

Saiki, R. K., Gelfand, D. H., Stoffel, S. J., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A., 1988, Primer directed enzymatic amplification of DNA with thermostable DNA polymerase. *Science*, 239 : 487–491

Sangiun, H., Sarniguet, A., Gazengal, K., Mo□nne-Locoz, Y. and Grundmann, G. L., 2009, Rhizosphere bacterial communities associated with disease suppressiveness stages of take-all decline in wheat monoculture. *New Phytol.*, 184 : 694-707.

Scher, F. M. and Baker, R., 1980, Mechanism of biological control in a *Fusarium*-suppressive soil. *Phytopathology*, 70 : 412-417.

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

Shannon, C. E. and Weaver, W., 1971, The mathematical theory of communication. University of Illinois Press, Urbana.

Sheffield, V. C., Beck, J. S., Stone, E. M. and Myers, R. M., 1992, A simple and efficient method for attachment of a 40base pair, GC-rich sequence to PCR amplified DNA. *BioTechniques*, 12 : 386–387

Sheffield, V. C., Cox, D. R. and Myers, R. M., 1989, Attachment of a 40bp G+C rich sequence (GC-clamp) to genomic DNA fragments by polymerase chain reaction results in improved detection of single base changes. *Proc. Natl. Acad. Sci. USA*, 86 : 232–236.

Simon, A**.** and Sivasithamparam, K., 1989, Pathogen suppression: A case study in biological suppression of *Gaeumannomyces graminis* var*. tritici* in soil. *Soil Biol. Biochem*. 21 : 331-337.

Singh, B. K., Campbell, C.D., Sorenson, S. J. and Zhou, J., 2009, Soil genomics. *Nature Rev. Microbiol.*, 7 : 756.

Sneath, P.H.A., 1993, Evidence from *Aeromonas* for genetic crossing-over in ribosomal sequences. *Int. J. Syst. Bacteriol.*, 43 : 626–629.

Solé, M., Chatzinotas, A., Sridhar, K. R., Harms, H. and Krauss, G., 2008, Improved coverage of fungal diversity in polluted groundwaters by semi-nested PCR. *Sci. Total Environ.*, 406 : 324-330.

Sorenson, T., 1948, A method of establishing groups of equal amplitude in plant sociology based on similarity of species content. *Biol. Skr.*, 4 : 1.

Steffan, R. J., Goksøyr, J., Bej, A. K. and Atlas, R. M., 1988, Recovery of DNA from soils and sediments. *Appl. Environ. Microbiol*., 54 : 2908-2915.

Stephen, J. R. and Kowalchuk, G. A., 2002, Ribotyping methods for assessment of in situ microbial community structure. In: Bitton G (ed) Encyclopedia of environmental microbiology. *JohnWiley and Sons, New York*, 5 : 2728–2741.

Stotzky, G. and Martin, R. T., 1963, Soil mineralogy in relation to the spread of Fusarium wilt of banana in Central America. *Plant Soil*, 18 : 317-337.

Suárez-Estrella, F., Arcos-Nievas, M. A., López, M. J., Vargas-García, M. C. and Moreno, J., 2013, Biological control of plant pathogens by microorganisms isolated from agro-industrial composts. *Biol. Control*, 67 : 3, 509-515.

Tebbe, C. C. and Vahjen, W., 1993, Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and yeast. *Appl. Environ. Microbiol.,* 59(8) : 2657-2665.

Teske, A., Wawer, C., Muyzer, G. and Ramsing, N. B., 1996, Distribution of sulfate reducing bacteria in a stratified Fjord (Mariager Fjord, Denmark) as evaluated by most probable number counts and denaturing gradient gel electrophoresis of PCR amplified ribosomal DNA fragments. *Appl. Environ. Microbiol*. 62 : 1405–1415.

Thomashow, L. S**.** and Weller, D. M., 1988, Role of Phenazine antibiotic from *Pseudomonas flourescens* in biological control of *Gaeumannomyces graminis* var*. tritici. J.Bacteriol.* 170 : 3499-3508.

Torsvik, V. L., and Goksoyr, J. and Daae, F. L., 1990, High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol*., 56: 782-787.

Trevors, J. T, van Elsas, J. D., Lee, H. and van Overbeek, L.S., 1992, Use of alginate and other carriers for encapsulation of microbial cells for use in soil. *Micron Releases,* 1 : 61-69.

Trevors, J. T. and van Elsas, J. D., 1989, A review of selected methods in environmental microbial genetics. *Can. J. Microbiol.,* 35 : 895-902.

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

van Elsas, J.D., 2008, The metagenomics of Disease Suppressive Soils. j.d.van.elsas@rug.nl.

Wagner, M., Erhart, R., Manz, W., Amann, R., Lemmer, H., Wedi, D. and Schleifer, K.H., 1994, Development of an ribosomal-RNA-targeted oligonucleotide probe specific for the genus *Acinetobacter* and its application for in situ monitoring in activated-sludge. *Appl. Environ. Microbiol.*, 60 : 792–800.

Wang, Y. and Qian, P-Y., 2009, Conservative fragments in Bacterial 16SrRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *Plos One*, 4(10) : e7401.

Westphal, A. and Becker, J. O., 1999, Biological suppression and natural population decline of *Heterodera schachtii* in a California field. *Phytopathology,* 89 : 434-440.

Wittebolle, L., Vervaeren, H., Verstraete, W., and Boon, N., 2008, Quantifying community dynamics of nitrifiers in functionally stable reactors. *Appl. Environ. Microbiol.,* 74 : 286–293.

Young, C. C., Burghoff, R. L., Keim, L. G., Minak-Bernero, V., Lute, J. R., and Hinton, S. M., 1993, Polyvinylpyrrolidone-agarose gel electrophoresis purification of polymerase chain reaction-amplifiable DNA from soils. *Appl. Environ. Microbiol*., 59(6) : 1972–1974.

Zhou, J., Bruns, M. A. and Tiedje, J. M., 1996, DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.*, 62 : 316-322.

APPENDIX I

Composition of DNA isolation buffer			
Sl. No.	Name of the component	Final concentration in Mm	pH
1	Tris-HCl	100	9
2	Na-phosphate buffer	100	9
3	Sodium chloride	1500	-
4	EDTA	100	9
Note: The final volume to be made up with nanopure water			

Preparation of stock solutions				
Sl. No.	Name of Component	Concentration (in M)	pH	Quantity of component for final volume of 100 ml (in gm)
1	Tris-HCl	1	9	12.11
2	Sodium chloride	5	-	29.22
3	Sodium dodecyl sulfate	20%	-	20.00
Note: Maintain pH of Tris by using HCl and makeup final volume by using nanopure water				

Preparation of stock solutions of Na-phosphate buffer		
Name of component	Concentration (in M)	Quantity of component for final volume of 100 ml(in gm)
Na_2HPO_4 (dibasic)	1	14.96
NaH_2PO_4 (Monobasic)	1	13.79
Note: The final volume to be made up using nanopure water		

Preparation of 1M sodium phosphate buffer			
Name of the component	pH	NaH_2PO_4 (in ml)	Na_2HPO_4 (in ml)
Sodium phosphate buffer	9	4.5	94.5

APPENDIX II

Agarose gel electrophoresis

a) Loading dye composition(6X)
1 0.25% Bromo phenol blue (BPB)
2 40% sucrose
3 Water to make up the volume

b) Ethidium bromide
10 mg/ml in distilled water. Stored at 4^0C in dark bottle

c) Preparation of 1% Agarose gel (100 ml)	
1 Agarose	1 gm
2 1X TAE	100 ml
3 EtBr (10mg/ml)	4 µl

d) 50X TAE composition		
1	Tris base	242 g
2	Glacial acetic acid	57.1 ml
3	0.5 M EDTA (pH 8.0)	100 ml
Total volume made up to 1000 ml with double distilled water		

APPENDIX III

Denaturant gradient gel electrophoresis

Preparation of denaturant gradient gel			
Sl.No.	Components	Solution A 70% denaturant - 12% gel	Solution B 45% denaturant – 12% gel
1	45% polyacrylamide	17.8 ml	17.8 ml
2	50X TAE	2 ml	2 ml
3	Formamide	32 ml	12 ml
4	Urea	42 gm	12.6 gm
Note: The final volume made up to 100 ml of solution A and solution B using nanopure water			

Preparation of staking gel		
Sl.No.	Components	Volume(in µl)
1	45% polyacrylamide	1111
2	50X TAE	200
3	20% APS	50
4	TEMED	5
5	Nanopure water	8620

Preparation of silver staining solutions		
Sl.no.	Solutions used	Reagents
1	Fixer	100 ml ethanol and 5 ml glacial acetic acid
2	Wash	Deionised water
3	Impregnation solution	1.5 gm $AgNO_3$ and 1 ml formaldehyde
4	Rinse	Deionised water
5	Developer solution	15 gm NaOH and 2 ml formaldehyde
6	Stop solution	100 ml ethanol and 5 ml glacial acetic acid
Note: Make up the volume of each solution to 1000 ml by using nanopure water		

METAGENOMIC CHARACTERIZATION OF DISEASE UPPRESSIVE SOILS

MAITREYEE SARMA **2015** **DR. H. M. VAMADEVAIAH**
(MAJOR ADVISOR)

ABSTRACT

Disease suppressive soils have been an unsolved mystery from the time of its discovery. Culture based techniques have indicated that the presence of a specialized community of microflora harbouring antibiotic resistance genes. Therefore, for better understanding of the community structure and composition of these soils the metagenomic analysis of 3 disease suppressive soil samples alongwith a Control soil sample was done. The soil metagenome DNA was isolated directly from the 4 soils, followed by analyzing the bacterial diversity present in it using PCR-DGGE and Sequence based approach (NGS). The PCR-DGGE and sequencing results were further analyzed using both statistical and bioinformatic tools (MG-RAST). The metagenomes were also screened for the presence of agriculturally important genes. The results of the above analyses showed that the high quality and quantity of DNA was instrumental in reflecting the bacterial diversity present in the 4 soil samples. The DGGE results for the 3 suppressive soil samples significantly differed from the Control soil samples in terms of species diversity. Also, the Fo index values obtained from the DGGE results clearly points out to the presence of a bacterial community having high functionality. The sequence based results also supports the result obtained from the DGGE analyses i.e. higher relative abundance of a specific group of bacterial microflora which was found to be statistically correlated with the disease suppressive property of that soil. Therefore, the above study helped in approving the assumptions made from the culture based studies done on analyzing the role of bacteria in conferring the disease suppressiveness characteristic to a particular soil.