# Chapter 1

# Introduction

*The true laboratory is the mind, where behind illusions we uncover the laws of truth.*

Sir J.C.Bose

Soil is the source of innumerable life. It provides habitat for beneficial microbes and there exists a synergistic relationship between the microbial communities. Soil is essential for the conservation of biodiversity and it is vital for the existence of life on this planet. Soil can be considered as an ecosystem itself, as they support a variety of life and all accompanying complex interactions among organisms (Curtis & Sloan 2005; Ponge, 2015). The major organisms in soil are prokaryotes and can form the largest part of the soil biomass (Sikroski 2015; Daniel 2005; Hassink et al. 1993). It has been found that one gram of soil comprises approximately $10^{10}$ bacterial cells and estimated species diversity of between $4X10^3$ to $5X10^4$ species

(Raynaud 2014; Roesch et al. 2007; Torsvik et al. 1990). Fertile soils are the outcome of team work of microorganisms, which directly affect the biological fertility of that soil. In addition to fertility, soil microorganisms also play essential roles in the recycling of nutrients that are essentially central to life on the planet. Soil microbial communities play a vital role in maintaining soil quality, function and the ecosystem sustainability (Sapkota 2020; Jacoby et al. 2017; Waldrop et al. 2000; Nannipieri et al. 2003). Many previous studies revealed that soil microbial communities were affected by various factors, including plant species, soil types and agricultural practices (Zhang et al. 2021; Xue et al. 2018; Garbeva et al. 2004; Jangid et al. 2008; Berg and Smalla 2009). However, the age-old practices of conventional agriculture may not be helpful in promoting the healthy populations of microbiome, limiting production yields and intimidating sustainability. Scientific studies are exploring new and exciting prospects for the restoration and promotion of healthy microbial populations in the soil. A definite understanding of the microbiota of an ecosystem helps in sustainable agriculture, ecological restoration, reclamation of land, pathogen resistance, nutrient acquisition, and stress tolerance of the host plant (Liu et al. 2019).

Actinobacteria is one of the dominant phyla present in the soil microbial community and it is poorly understood. Actinobacteria are widely distributed in various biotopes such as soil, water, permafrost, mammals, arthropods, plants etc. (Sen et al. 2014). Furthermore, various lifestyles are encountered among Actinobacteria, and the phylum includes pathogens (e.g. *Mycobacterium* sp. *Nocardia* sp. *Tropheryma* sp. *Corynebacterium* sp. and *Propionibacterium* sp.), soil inhabitants (*Streptomyces* sp.), plant commensals (*Leifsonia* sp.), nitrogen-fixing symbionts (*Frankia sp.*), and gastrointestinal tract (GIT) inhabitants (*Bifidobacterium*) (Ventura et al. 2007).

Actinobacteria form the integral part of microbial community which is involved in antibiotic production, secondary metabolite production, bioremediation, nitrogen fixation, etc. (Girao et al. 2019). They are in constant interaction with other organisms as symbiotic or pathogenic microbes altering the immediate environment. Bioactive compounds produced by microbes exhibit pharmacological or toxicological effects in other organisms. Actinobacteria have been considered as one of the significant groups of microorganisms as they represent a broad range of valuable and prominent sources of pharmaceutically active metabolites. Actinobacteria mostly inhabit in soil, freshwater, and marine habitats playing an important role in decomposition of organic materials, such as cellulose and chitin, thereby playing a vital part in biogeochemical cycles, replenishing the supply of nutrients in the soil, and helps in humus formation (Bhatti et al. 2017). Its diversity and distribution studies will help us to reveal about microbial ecology and its applications in agriculture and pharma sector.

Human activities are creating new compounds and the microbial community is constantly evolving to cope up with this stress. The mechanism of action of microbial community *in vitro* and *in vivo* conditions may deduce the remedy for degradation of environment as well as the drug resistance in organisms. The metabolic path ways involved in xenobiotic degradation is not yet studied in a comprehensive manner (Mishra et al. 2021, Maurice et al. 2013). The detailed study of xenobiotic activity of pathogenic and non-pathogenic Actinobacteria in different niches may help us to understand the intricate mechanism of microbial community with the environment. It may open up new avenues of understanding that could have many applications in fields diverse as agriculture, biotechnology, ecosystem monitoring. In nature, Actinobacteria play a pivotal role in the cycling of organic compounds and have also been associated with soil organic matter production, owing to their black pigments called melanins, which are related in some respects to soil humic acid (Anandan et al. 2016; Schafer et al. 2010).

The most widely distributed genus of Actinobacteria, *Streptomyces* has been the focus of research by biologists because of the commercial applicability of the substances produced. Among the biological control agents derived from different microbes, Actinobacteria especially *Streptomyces* sp. are one of the most important microbial resources which can provide potential new bioactive compounds for use as insect-control agent (Kaur et al. 2013; Gadelhak et al. 2005). So, the main focus of this study is mainly limited to *Streptomyces* genera of Actinobacteria from selected biotopes of Indian subcontinent.

An area with uniform environmental conditions providing a living place for a specific assortment of plants and animals. It has its origin from the Greek word combining "bios" (meaning life) and "topos" (meaning place). The German scientist Friedrich Dahl introduced the term biotope in 1908 This term was a complement of the term biocoenosis (attributed to Karl Möbius in 1877) meaning the assemblage of animals and plants living together in a specific habitat.

Furthermore, the term biotope is not limited to physical conditions that surround a community of organisms but it also comprises the relative biota. Therefore, biotope is a topographic unit characterized by similar environmental (physical) conditions and a specific assembly of plant and animal species. In this study, I have chosen different biotopes for the isolation and characterization of Actinobacteria mainly, *Streptomyces.* Diverse biotopes are selected in order to isolate new strains of the genus. The biotopes that are selected for this study are the soils from tea garden, mangroves, rubber plantation, paddy fields, vegetable garden, and natural forests.

The world's second most popular drink is tea that has been consumed every day by billions of people. Besides legendary Darjeeling tea, Assam and Dooars tea are the most popular global brands. It is unique for its strength, taste and aroma. Tea is the backbone of Darjeeling and Dooars economy which supports the majority of the population in North Bengal. It is estimated that around 28, 000 small tea growers in Terai and Dooars of North Bengal who create employment for 1.7 lakh local people of this sub-

Himalayan region (Chowdhury et al. 2016). However, the small-scale tea sector faces a number of problems such as lack of capital, improper knowledge about the agro-techniques of tea cultivation, inadequate input availability and problem of marketing (Saikia, 2019, Karmakar & Banerjee, 2005). Apart from these, tea industry as a whole is facing multiple problems in term of its quality and production. Though the large-scale producers can overcome the losses but this is having adverse impacts on small tea estate owners (Rajbangshi & Nambiar, 2019, Kumar et al. 2008). Bioremediation and the bio-augmentation are the techniques for improvement of the capacity of a contaminated soil or water to remove pollution by the introduction of specific capable strains of microorganisms. In this context, it is vital to identify and characterize the microbial consortia mainly consist of Actinobacteria from native as well as pristine biotopes with promising qualities for bio augmentation. This will help to improve the production and quality of tea and increase its demand in international market. Another aspect is the utilization of PGPR activity of Actinobacteria strains which will reduce the input of chemical fertilizers (Aldesuquy et. al. 1998; Borah & Thakur, 2020). The bio -augmentation property of Actinobacteria (Sharma et al. 2020) is highly promising to the declining tea industry of North Bengal. Identification and characterization of key metabolites from microbiome will be highly promising to improve the health status of tea garden workers which will have wider social, economic and ecological implications.

Another important biotope, which is under studied is mangrove forests. Mangrove forests are large ecosystems prevalent in tropics and subtropics. They make up over a quarter of the total coastline in the World. Mangroves form the transient ecosystem which exhibits the edge effect and shows wide species richness. They provide habitat for different flora and fauna and also a treasure trove for microbial biome. Since it is an ecotone habit, it is rich in biodiversity. There are some criteria to be fulfilled by an ecosystem to be considered it as mangrove (Tomlinson 2016). They are: complete fidelity to the mangrove environment, plays a major role in the community, ability to form pure stands, morphological adaptation to the habitat. Such mangrove forests are estimated to cover an area of about 150,000 sq.km.in world wide (Forest survey of India, 2017). The total area of mangroves in India is estimated to be 6,740 sq. km (Kumar et al. 2008) which is drastically decreasing due to anthropogenic activities. It is a link between land and sea which serves as breeding and nursing grounds for aquatic organisms. This ecosystem is an active geographical zone which is rich in organic sediments. These sediments are the vast resources for microbial production and metabolism. Due to the presence of rich source of nutrients mangroves are called the homeland of microbes. The mangrove ecosystem is saline and highly rich in organic matter and remain mostly untapped source for screening and isolation of potential bioactive secondary metabolites (Law et al. 2019; Newman & Cragg 2007).

Rubber plantations are considered as a key ecosystem in tropical areas such as Southeast Asia, Latin America, etc. (Dechner et al. 2018). The rubber tree is an indigenous species of Amazon rainforests and utilities as the sole source of natural rubber for the industry. The site for natural rubber biosynthesis and storage is the secondary laticifers present in the inner bark of rubber tree are (Hao & Wu 2000). Rubber is widely cultivated as plantation crop in Asia, Africa. In India, rubber is mainly cultivated in Kerala.

The increase in rubber plantations led the researchers to focus on the ecological and environmental consequences of these so called, ‘artificial ecosystems’. These studies have analysed the effects of rubber plantation on soil fertility, soil organic carbon and soil microbial biomass (Diniz et al. 2020; Yang et al. 2004; Zhang et al. 2007). A soil microbial community in the rubber plantation varies according to the age of the plantation, cover vegetation of the plantation, use of agrochemicals etc. Most studies have focused on microbial biomass and activity in surface or subsurface soil of rubber plantations (Guo et al. 2015; Zhang et al. 2007). There are very few studies available for the effects of rubber cultivation on composition of microbial population in the rhizosphere and bulk soil of rubber trees (Zhou et al. 2017). So efforts have made to isolate and characterise the indigenous microflora from the rubber plantation soil.

Rice (*Oryza sativa* L.) is the most important cereal crop in the world, feeding more than 50% of the human population and is the major human food in Asia. The rice ecosystem or paddy field may be a prime habitat for microorganisms adapted to fluctuating nutritional levels and oxygen and light availability. The application of fertilizers in paddy field to increase rice yield affect the microbial community. Soil microorganisms are crucial for nutrient cycling, soil fertility, and crop productivity (Jacoby, 2017; Garbeva et al. 2004). Maintaining the health of soil microbiota is imperative for soil fertility and optimal crop yield. Microbial communities within rice fields in different habitats vary in diversity and response to environmental changes. Microbial communities in the rice rhizosphere can be significantly affected by seasonal changes (Dong et al. 2021; Schmidt & Eickhorst 2013). Soil properties and the microbial community in rice field is also affected by fertilization treatments (Ahn et al. 2016; Pascual 2000). Chemical fertilizers are likely to stimulate the growth of gram-positive bacteria in rice soils, while organic amendments increase the relative abundance of bacteria and fungi and decrease the abundance of Actinobacteria. Zhang et al. 2012). Nitrogen fertilizers increase soil biomass in organic carbon-rich rice soils (Yu et al. 2020). It has been shown that in these agroecosystems, rice exudates and nutrients from straw incorporation were shown to influence the bacterial communities' composition. It is vital to gather the information about microbiota prevailing in this unique biotope.

The forest is an important biotope as they are abundant with microflora and also act as carbon sinks. Previous studies by various have shown that soil microbial communities are affected by the changes in land use patterns and climate (Jesus et al. 2009; Tripathi et al. 2012; Lee-Cruz et al. 2013;

Rodrigues et al. 2013; Kerfahi et al. 2016). While most of the research have focused on fungi, and minimal information is available on forest soil bacteria which play a vital role in the environment. In forest soils, bacteria inhabit numerous habitats with definite properties, which vary from bulk soil, rhizosphere, litter, and deadwood habitats, where their communities are influenced by nutrient availability and biotic interactions. Bacteria interact with plant roots in symbiotic or non-symbiotic manner and exhibit commensality property with mycorrhiza fungi as symbiotic or mycorrhiza helpers. Bacteria also facilitate multiple critical steps in the nitrogen cycle, including Nitrogen fixation. However, the understanding of bacterial communities in forest soils has advanced in current years, but it is still insufficient. The precise level of the involvement of bacteria to forest ecosystem processes will be known only in the future, when the activities of all soil community members are studied synchronously. (Lladó et al. 2017). So, the forest soil is also chosen for the isolation of Actinobacteria in this study. The isolated microbes will be characterised at both biochemical and molecular level using the available protocols. The genome sequencing of few isolated Actinobacteria and its detailed bioinformatic analysis is proposed to be undertake in this research.

Soil being the most assorted and multifaceted microbial habitat on earth surpasses the quantity of inhabiting microbial communities than other environments. Some phylogenetic surveys on soil environs made evident that the number of prokaryotic species present in a specific soil sample is far more than the known cultured prokaryotes (Daniel 2005). Metagenomic analyses endow extensive information about the structure, architecture, and predicted gene function and behaviour of varied environmental assemblages. The first step in soil metagenomics is DNA isolation from the soil along with production and screening of clone libraries that aids us to explore the cultivation-independent genetic reservoir of soil microbial neighbourhood (Sarkar et al. 2021). The unmatched diversity of soils promises sustained exploration of diverse industrial, agricultural and environmental agents in the future. Each environment has its unique challenges to metagenomic-investigation requiring explicitly designed approaches considering both biotic and physicochemical factors (Kakirde et al. 2010; Nesme et al. 2016). It's a promising approach in unfolding the useful potential of the soil microbial community, which might produce greater insight into the health of a soil than taxonomy-based metrics, and also delivers Information about microbial communities and their interaction (Berg et al. 2020). It's a powerful tool for accessing almost 100% of genetic resources from environment and also aids in gene level studies of the communities. Functional metagenomic strategies helps us to explore discovery of new genes, enzymes and natural products etc. (Kennedy et al. 2011).

The successful application of soil metagenomics demands high-quality DNA extraction, purification along with cloning methods for the predicted downstream bioinformatics analyses (Kakirde et al. 2010). The average insert size of a clone library or the length of sequence read for a high-throughput sequencing approach is also crucial. An appropriate metagenomics screening strategy should be adapted to address the specific questions of interest (Kakirde et al. 2010). Hence, 16S amplicon or metabarcoding analysis based on the V3-V4 region will be undertaken for the soils collected from different biotopes to have a comprehensive account of the microbiome. Qualitative analysis of the different soils of the different biotopes by physiochemical methods is also planned under this study. This may give an insight into the interaction of microbes and their environment.

Bioinformatic tools has become indispensable in the genomics era for the analysis of voluminous data. The plethora of biological data are available to the scientific community due to the advancements in sequencing technologies (Albarano et al. 2020). Since the release of the first genomic sequence of *Haemophilus influenza* in 1995 by Fleischmann, thousands of genomes are sequenced and are available in biological databases. The completion of human genome project in 2003 was an added booster. In this scenario, Bioinformatics, a new discipline has gained popularity. It is the application of computational techniques to analyse and organise the biological data. The genome sequences which are available in databases are utilized for the study of interrelationships between various species in terms of structure and function. Codon usage study provides information about usage of codons and gene expression in genomes (Liu 2020). Codon usage patterns and preferences shows variation within and between the species (Zhou & Li 2009) The various indices employed in codon usage study are GC content, GC3 content, relative synonymous codon usage (RSCU), (Peden 1999), effective number of codons (Nc) (Wright 1990) and frequency of optimal codons (Fop) (Ikemura 1981). The basic principle is that the genetic code is degenerate and the genetic codes of different organisms are often biased towards the use of one of several codons. The codon usage bias or differences among the usage of the synonymous codons have been the important factor for the evolution of proteome diversity. The difference in the usage for synonymous codons exists widely within the genomes due to mutation, natural selection, and random drift (Shen et al. 2020). Thus, a complete understanding of the biases in codon usage can help us to explore the evolution of genes in different genomes. Codon adaptation index (CAI) is an important tool for analysing the patterns of codon usage within a gene relative to the reference set of genes, mainly, ribosomal protein genes (Sharp & Li 1987; Wu et al. 2005). CAI is used to predict the highly expressed genes in various organisms. Apart from these, comparative genomics-based analysis helps to elucidate the function and evolutionary processes that act on genomes (Lukjancenko et al. 2012). The characterisation and profiling of the core genome and pan genome of the concerned organisms is an effective method to unravel the puzzles of speciation and genomic variations. The core genome is the conserved pool of genes present in a species and the dispensable genome is the set of genes present in some strains of a species but not present in all strains of a species (Medini et al. 2005). The pan genome is the complete sets of genes within a species, including the core and the dispensable genomes (Morneau 2021). Hence in this study, steps are taken to compared and analyse the newly isolated and sequenced genomes from various biotopes with the selected genomes from public domain.

Apart from these, GC MS analysis of few isolates are also proposed under this study. This will help to identify the potential metabolites present in the isolates.

Agricultural production is often mired by numerous biotic stresses. Various chemical pesticides are used to minimize crop loss, which are hazardous (Mittal & Gupta 2008). Agrochemicals such as pesticides and fertilizers are essential inputs to improve the productivity of the farmer in terms of farm produce, but their use is also associated with chemically induced injuries to health. One of the most important plantation crops in India is tea. The tea industry in India is one of the old agriculture-based industry since 18th century (Borborah & Gogoi 2007). As of 2019, India is ranked the second largest tea producer in the world with production of 1,339.70 million kgs. The estimated tea production in India stood at 73.44 million kgs in April 2021. The total tea export was US$ 755.86 million in April 2020 to March 2021 and for March 2021 it was US$ 53.35 million (https://www.ibef.org).Though, tea is grown in various parts of India, Darjeeling tea which is unique in its flavour is grown in Terai and Dooars region in West Bengal and is considered as an economical backbone in tea cultivation areas of Assam and Northern part of West Bengal. In Terai and Dooars area of North Bengal, tea production is the only major and key industries on which teagardens, workers, tea traders are dependent for their major economy. Tea industry provides employment opportunities to large number of people and is cultivated by small tea growers which is largely unorganized and is ignorant of occupational hazards including the agrochemical exposure. [Tea industry Annual report, 2010]. Large number of families live on the tea estates. The recent information gathering showed that tea farmers have been applying different pesticides to reduce pest incidence in their tea garden. It is also reported that nearly more than 20 different types of pesticides being used in tea garden. Though farmers have been using WHO approved pesticides, they are practicing cocktail preparation of two or more than two pesticides without knowing their compatibility, which is of serious concern to health and environment. These are toxic chemicals and they are to be handled with extreme care. They are not easily degradable; they remain in the soil, leach to ground water and contaminate the environment. Further more, it also enters the body of organisms, which leads to bio accumulation in the food chains and adversely impairs the human health. The problem of heavy metal toxicity affects the health of the people. It has been reported that metal toxicity in the blood samples is above the permissible level among the people living in tea gardens. (NBMCH report). Moreover, occurrence of physical and genetic abnormalities and declining reproductive health has also been reported. The reports of the health status of tea garden workers makes it evident that the soil micro flora has an effect on human health. The respiratory ailments such as tuberculosis and skin disorders are prevalent among the tea garden workers (Gayathri & Arjunan, 2019). The health survey conducted among the tea garden workers of various tea gardens of Northeastern India has established the fact that workers are suffering from various ailments such as gastrointestinal disorders, respiratory disorders, skin diseases due to the unhealthy environmental condition (Rajput et al. 2021; Ahmmed & Hossain 2016; www.tezu.ernet.in/ project reports).

The bacteria and fungi produce produces secondary metabolites which are an important source of antimicrobials and other bioactive compounds. The genome mining has emerged as an effective approach in identifying and characterizing new compounds in recent years. The ‘antibiotics and secondary metabolite analysis shell or antiSMASH’ is extensively used tool for identifying and analysing biosynthetic gene clusters (BGCs) in bacterial and fungal genome sequences (Blin et al. 2019). It also helps to compare and coordinate with the available secondary metabolites present in the databases. This has prompted us to explore the bioactive compounds of Actinobacterial populations from different biotopes of India. The secondary metabolites present in the selected genomes will be identified using antiSMASH software as well as by GC MS analysis.

Keeping these facts in mind,

**Objectives of the present study are:**

- Survey of health status of tea garden workers in Terai and Dooars region.
- Screening and Isolation of Actinobacteria in pure culture from various biotopes such as tea gardens, paddy fields, mangroves etc.
- Identification and Morphological and biochemical characterisation of isolated microbes using standard protocols.
- Molecular documentation of different Actinobacterial strains using16S rRNA primers.
- Bioassays of selected strains using different methods.
- Whole genome sequencing of few potential Actinobacteria such as Streptomyces.
- GC MS analysis of microbial cultures for their metabolites.
- Comparative analysis of different genomes through bioinformatic tools.
- Metagenomic analysis of soil microbiome from different biotopes.■