Chapter 1
Introduction

Nitrogen is an essential critical limiting element used as a nutrient by plants and animals for growth and reproduction to survive. It is required to biosynthesize basic building blocks of living organisms like chlorophyll, amino acids, ATP and nucleic acids. It makes up four-fifths of the atmosphere but is metabolically unavailable directly to higher plants or animals. Although element nitrogen is present abundant in the atmosphere in the form of N_2, plants cannot use this gaseous N_2 directly until it is further converted into a more readily accessible form like ammonium (NH_4^+) or nitrate (NO_3^-) ions in a process known as nitrogen fixation. Both biological and non-biological processes are associated with the transformation of nitrogen into different chemical forms. Non-biological fixation occurs by industrial synthesis, UV radiation and lightening (Ladha and Reddy, 1995). Inorder to meet the demands of the ever increasing world population the artificially prepared nitrogen fertilizers are applied in agriculture for enhancing crop production. For the production of wheat, rice and maize alone 42 million tons of nitrogenous fertilizers are used annually on a global scale (Saikia and Jain, 2007). Approximately 25% of the nitrogen from nitrogenous fertilizers is lost by leakage, volatilization, denitrification and other factors resulting in high economic loss accompanied by severe environmental pollution. The ecosystem gets damaged owing to high levels of nitrate

ammonium ions in cultivated soils resulting in plant toxicity (Saikia and Jain, 2007). Thus, the ill effects of chemical fertilizers and the rising costs for farmers especially in third world countries which has been a subject of global concern, it is necessary to explore the benefit of biological nitrogen fixation (BNF) based methodologies for cost effectiveness, maximizing output and eliminate the ill effects of environmental pollution.

Biological nitrogen fixation (BNF) is the mechanism of reduction of atmospheric dinitrogen to a metabolically active form catalyzed by nitrogenase enzymes exclusively by microorganisms (Bohlool *et al*. 1992). It is a key process of the nitrogen cycle and is of a particular interest to ecologists because nitrogen availability can affect the rate of key ecosystem processes, including primary production and decomposition (Sur *et al*. 2010). Biological nitrogen fixation is directly proportional to agricultural sustainability (Bohlool *et al*. 1992). Annually approximately 8 x 10^{10} kg NH_3 are manufactured by ammonia industry and 1 x 10^{10} kg NH_3 are contributed by lightning worldwide per year. However, annually approximately 2.5 x 10^{11} kg NH_3 is fixed from the atmosphere by biological nitrogen fixation (BNF) which is comparatively higher (Cheng, 2008). Hence, it is necessary to explore the possibilities for improving biological nitrogen fixation and its use by farmers on a global scale.

Biological nitrogen fixation (BNF) is performed in nature by microorganisms that fix nitrogen known as diazotrophs. Diazotrophs are bacteria and archaea which may be free living and symbiotic. Free living nitrogen fixing microorganisms comprise of *Azotobacter*, *Clostridium, Chlorobium,* some methanogenic archaeal members like *Methanococcus*, *Methanosarcina*, *Methanothermobacter* etc. that fix nitrogen independent of other organisms. The free living diazotrophs require a chemical energy source if they are non photosynthetic, whereas the photosynthetic diazotrophs, such as the cyanobacteria utilize light energy (Leigh, 2002). The symbiotic diazotrophs are Rhizobia, Frankia and cyanobacteria which fixes nitrogen symbiotically by partnering with a host plant.

Among the symbiotic nitrogen fixers, the rhizobia are of special significance because of their ability to make association with a large number of

genera of leguminous plants. Rhizobia are soil bacteria which live in a mutualistic symbiotic relationship with legumes, a relationship that has existed and co-evolved for tens of millions of years. During the symbiotic process, the rhizobia enters the root hairs, multiply there and form nodules and reduces atmospheric nitrogen into a form directly assimilated by plants (ammonium) while in return the plant provides energy in the form of carbon synthesized during photosynthesis. Both *nif* and *fix* genes are required for nitrogen fixation after infection of a host. Rhizobia continue to differentiate inside the nodule and expresses the enzyme nitrogenase required for nitrogen fixation and fix nitrogen for the maintenance of the mutualistic partnership (Gage, 2004). Amongst the different nitrogen fixing endosymbiotic interactions, the most intensively studied is that established between legume plants and nitrogen-fixing endosymbiotic bacteria of the genera *Rhizobium*, *Sinorhizobium*(*Ensifer*), *Mesorhizobium*, *Bradyrhizobium* and *Azorhizobium*, collectively termed rhizobia (Weidner *et al*. 2003).

Grain legumes also known as pulses are widely recognized as an important source of food and feed proteins (Christou, 1997) and have become very important in human nutrition and as a feed for domestic animals (Cummings *et al*. 2001). Grain legumes are rich source of dietary protein (Duranti and Gius, 1997), especially for the largely vegetarian population of sub-tropics. Major part of N_2 fixed by legumes is harvested as grains, while the soil and the succeeding crops also get benefitted by N in the form of root and shoot residues. The N_2 fixation efficiency of legumes varies, and depends on the host genotype, rhizobial efficiency, soil conditions, and climatic factors. In general, it has been assumed that agriculturally important legumes fix 40 to 60 million metric tons (Mt) of N_2 annually, with another 3 to 5 million Mt fixed by legumes in natural ecosystems (Smil, 1999). This is a remarkable efficiency when compared with the small quantities of nitrogenase enzyme involved. Reported quantum of nitrogen fixation ranged from 126 to 319 kg N ha^{-1} in groundnut, 33 to 643 kg N ha^{-1} in soybean, 77 to 92 kg N ha^{-1} in pigeonpea, 25 to 100 kg N ha^{-1} in cowpea, 71 to 74 kg N ha^{-1} in green gram and 125 to143 kg N ha^{-1} in black gram (Peoples and Craswell, 1992; Gopalakrishnan *et al*. 2015). Thus, the

symbiosis between legumes and rhizobia has been widely used to improve agricultural productivity. This is particularly relevant in developing countries where agriculture is prone to nitrogen losses and legumes can represent an alternative source of protein for human and animal consume (Peoples and Craswell, 1992; Kaneko *et al*. 2002; Peoples *et al*. 2002). Grain legumes include beans, lupins, peas and peanuts (Kurlovich and Repyev, 1995).

The common bean or French bean (*Phaseolus vulgaris* L.) is an important leguminous short duration vegetable crop, highly proteinaceous in nature. Among major food legumes, *P. vulgaris* is third in importance, has broadest genetic base and is the major cultivated representative of the genus (Christou, 1997). The large amount of protein, minerals and antioxidant compounds (Xu and Chang, 2008) make this crop an excellent model food legume (Broughton *et al*. 2003). It is grown and consumed in almost every part of the world (Singh, 1999). It was introduced from the Americas into India approximately 400 years ago. It is commonly cultivated in northern part of India. However, it is also grown in other parts of India like Maharashtra, Nilgiri (Tamil Nadu) and Palni (Kerala) hills, Chickmagalur (Karnataka) and Darjeeling hills (West Bengal) (Ahlawat, 2008) and Sikkim is no exception. A wide variability of French bean is found in the various parts of the region. Climbing or pole type in French bean is popular among the tribals since it is used for mix cropping with maize, the stem of which acts as the support for the bean (Asati and Yadav, 2004).

P. vulgaris L. is nodulated with various fast growing *Rhizobium* sp*.,* such as *R. leguminosarum* bv. *phaseoli* (Jordan, 1984) , *R. tropici* (Martinez-Romero *et al*. 1991) *, R. etli* (Segovia *et al*. 1993), *R. gallicum* and *R*. *giardinii* (Amarger *et al*. 1997). However, it is often considered inferior in N_2 fixation in comparison to other grain legumes (Ali and Lal 1992; Hardarson, 1993) as it lacks nodulation due to the absence of a nodulation (NOD) gene regulator. This fact is attributed to intrinsic characteristics of the host plant, particularly the nodulation promiscuity (Michiels *et al*. 1998). The extreme sensitivity to nodulation-limiting factors, such as the high rate of N fertilizer used in intensive agriculture, nutrient deficiency, high temperature and soil dryness also hinders in the N_2

fixation (Graham, 1981). The efficiency of symbiosis are also greatly affected by the genotypic variation in the beans as well as compatibility of *Rhizobium* plant cultivar. As a result of this variability the nitrogen fixing performance of soil native rhizobia or use of commercially available inocula are often limited. Hence, application of nitrogen through fertilizers is imperative for exploiting its yield potential and is being a common practice. However, the usage of high amounts of inorganic nitrogen fertilizers has detrimental environmental consequences (Yadegari and Rahmani, 2010). The use of Rhizobial inoculant is the best alternative to N_2 fertilizer. Rhizobia can be used as inoculants for enhanced N fixation and studies demonstrated their predominance in nodules for 5–15 years after initial inoculation (Lindstrom *et al*. 1990), and confirming that they are effective colonizers persisting in soil for many years in the absence of their host (Sanginga *et al*. 1994). The symbiosis between *Rhizobium* and legumes has been found to be a cheaper and usually more effective agronomic practice for providing an adequate supply of N for legume based crop compared to the nitrogen based fertilizers (Zahran, 1999). The *Rhizobium* legume symbiosis are the most extensively examined system as it is considered to be superior to other nitrogen fixing systems due to its high potential (El-Deeb and Al-Sheri, 2005) .

Rhizobium belongs to the intracellular PGPR (iPGPR) group of bacteria that promote plant growth either directly (nitrogen fixation, phosphate solubilization, iron chelation and phytohormone production) or indirectly (suppression of plant pathogenic organisms, induction of resistance in host plants against plant pathogens and abiotic stresses (Gopalakrishnan *et al*. 2015). Several investigations has been done on the host-endophyte interactions in some major field crops but only recently has common bean received attention (Hohenberg *et al*. 1982). During the last decade much research has focused on the beneficial effect of *Rhizobium* inoculation alone or simultaneous inoculation with *Rhizobium* and plant growth-promoting rhizobacteria (PGPR), so called co-inoculation, showing the potential to enhance plant growth, nodulation and nitrogen fixation of several legumes including common bean. To aid this, efficient

strains must be matched with compatible hosts and the resulting association selected for suitable environments (Rys and Bonish, 1981). So, the identification and exploration of potential rhizobia with various plant growth promoting properties will be useful for sustainable agriculture. Although only 57 % of 650 genera of leguminous plants have been studied for nodulation, the numbers of rhizobial species are increasing with the exploration of large number of legume species. The number of rhizobial species has increased considerably from 8 in the year 1980 to 53 in 2006 (Willems, 2006). Dispersion of host plants to new geographical locations might serve as a major source for these new rhizobia species. Studies on the various aspects of *Rhizobium* including the recent advancements in the taxonomic research with the aid of specific molecular tools have lead to the identification and exploration of many more rhizobial species.

Traditional phenotypic methods including analysis of colony morphology in culture medium, biochemical, metabolic and nutritional characteristics, bacteriophage susceptibility and serological reactions and nodulation tests on host plants to determine cross-inoculation have been applied to identify and characterize these bacteria groups (Graham, 1963; Moffett and Colwell, 1968; Vincent and Humphrey, 1970). Based on data derived from the use of traditional tests such as biochemical and physiological tests *R. phaseoli, R. trifolii, R. leguminosarum* and *R. meliloti* are considered to be fast growers which produces acid on yeast-mannitol-agar medium, while *R. lupine, R. japonicum* and rhizobia of cowpea, miscellany are slow growing species which produce alkali on yeast mannitol-agar medium (Fred *et al*. 1932). However, some of the rhizobia have been placed in a new bacterial genus, *Bradyrhizobium* (Jordan, 1982). On the basis of the capability of *Rhizobium* to nodulate legumes, seven cross inoculation group were recognized and thus could be used in the classification. In a cross-inoculation group, *Rhizobium* isolated from one legume members of the group would nodulate all other members of that group. For example, a strain isolated from *Medicago sativa* can nodulate *Melilotus* and *Trigonella* species belonging to the same cross-inoculation group, but cannot nodulate clovers or soybeans (Lieberman *et al*.

1985). The use of cross-inoculation groups for taxonomic purposes has been limited because of the relatively few species examined in the assessment, and little attention recently has been given to adding new information (Lieberman *et al*. 1985). However, the cross-inoculation grouping system is not perfect since rhizobia have often been found to cross -infect between groups (Gaur *et al*. 1974). Rhizobial genome organization has been widely studied. Generally, these bacteria have one chromosome and several plasmids and/or mega plasmids that may represent 50% of the genome (Alexandre *et al*. 2006). Symbiotic genes are typically located in mega plasmids, known as pSyms (Brom *et al*. 2002) while the non symbiotic plasmids may encode locally adaptive traits that confer phenotypic advantages such as heavy metals or antibiotic resistance genes. Intrinsic antibiotic resistance (IAR) profiles has been found to be a discriminatory method to screen diversity in new isolates (Alexandre *et al*. 2006). Studies on the effects of heavy metals on growth, abundance, morphology and physiology of various strains of *R. leguminosarum* have been well documented (Castro *et al*. 1997; Lakzian *et al*. 2002; Chaudhary *et al*. 2004). It has been found that the continuous exposure to heavy metals adversely affects the genetic diversity and nodulation of the host plants (Hirsch *et al*. 1993; Paton *et al*. 1997) and variations in the expression of symbiotic genes including nod genes (Stan *et al*. 2011). It is necessary to isolate and study the native rhizobial strains from heavy metals contaminated soils to identify the potential of *Rhizobium*–legume symbiosis of particular strain for the remediation of the affected area. *R. fredii* and *R. meliloti* were found to exhibit higher metal tolerance against Tellurium (Te) and Selenium (Se) (Kinkle *et al*. 1994) .

The population genetics of *Rhizobium* species have been carried out by using serological or multilocus enzyme electrophoresis techniques (Paffetti *et al*. 1996). However, characterization of the *Rhizobium* genome at the molecular level has been found to be the most discriminating method for assessing the variability among strains of the bacteria (Demezas *et al*. 1991; Thies *et al.* 2001). The DNA-DNA homologies have indicated that there is a sharp line of demarcation between the fast growing species (*R leguminosarum*, *R*

trifolii, *R. phaseoli* and *R. meliloti*) and the slow growing ones (*R. japonicum* and *R. lupini*). The randomly amplified polymorphic DNA (RAPD) markers have been found to be a commonly used technique to study the genetic polymorphism in *Rhizobium*. The other molecular techniques that are frequently used for the study of rhizobial diversity are analysis of repetitive sequences, including repetitive extragenic palindromic (REP) and enterobacterial repetitive intergenic consensus (ERIC) sequences amplified with PCR (De Bruijn 1992; Sikora and Redzepovic, 2003). Restriction fragment length polymorphism (RFLP) of PCR amplified genes coding for 16SrRNA (ARDRA) is now a common method for the genotypic classification of bacteria including root nodule bacteria (Safronova *et al*. 2004). The PCR RFLP analysis of 16S-23S rDNA IGS sequences have also been successfully applied for the differentiation of rhizobial strains at species and intraspecies level (Laguerre *et al*. 1996; Safronova *et al*. 2004) .

The phylogenetic analysis based on gene and genome sequencing is considered to be one of the most trustworthy methods to study the Rhizobiaceae family at the molecular level. Sequence analysis of 16SrRNA genes has highlighted the evolutionary diversity of the symbiotic bacteria and has been helpful in revealing the identification and classification of bacteria for the past two decades (Rajendhran and Gunasekaran, 2011). New bacterial isolates are identified based on the 16S sequence homology analysis with existing sequences in the databases. The species are identified based on the closest match obtained from comparative tools such as BLAST (http://www.ncbi.nlm.nih.gov) and Seq-match (http:// rdp.cme.msu.edu).

The sequencing of the whole genome of symbiotic as well as non-symbiotic diazotrophs has exposed new evidences pertaining to evolution and structure, interactions between plants and microbes and diversity amongst the diazotrophs (Sur *et al*. 2010). However, not much whole genome sequencing of *Rhizobium* has been done so far worldwide. The whole genome projects for nitrogen fixing microbes are great since they are the sources of natural nitrogen in plants and soil. These projects have resulted in the availability of tremendous amount of biological data. The whole genome

attributes of *Rhizobium*, in conjunction with the other genomes are important for on-going comparative and functional analysis of the plant microbe interactions required for the successful establishment of agricultural crops. Bioinformatics research in plant microbiology, constituting the associated plants and microorganisms is directed at acquiring the total nucleotide sequence of nitrogen fixing microorganisms and applying knowledge to successive post genomic studies. The science of bioinformatics is associated with the challenge of decoding the huge number of genomic sequences present in databases which in turn gives the idea about the proteome complement, proteins, codon usage etc.

Comparative genomics have emerged as one of the interesting areas of study. For comparing genomes of nitrogen fixing bacteria it is necessary to study their codon usage, analyze their proteomes, molecular phylogeny and protein coding gene (Sur *et al*. 2010). Protein coding genes remain distributed between leading and lagging strand of DNA. Genes transcribed on leading and lagging strand of replication have been reported to have distinct codon usage pattern in some prokaryotes like *Borrelia burgdorferi* (McInerney, 1998). Asymmetrical replication plays a vital role in codon usage variation among genes. Replicational selection governs on genomes having more genes on leading strand whereas, transcriptional selection is responsible for enrichment of highly expressed genes on leading strand. Thus, 'replicational- transcriptional' selection plays a pivotal role in determining the codon usage pattern on genes. This is actually a new paradigm of codon selection study in prokaryotes. Orthologous and paralogous relationship among genes are very important for both functional and evolutionary aspects of comparative genomics (Tatusov *et al*. 2003). Robustness of genome annotation, construction of evolutionary scenarios involving vertical inheritance, lineage-specific gene loss and horizontal gene transfer require accurate identification of orthologous genes (Tatusov *et al*. 2003). Hence, a functional classification system based on orthologous relationship between genes appears to be a natural framework for comparative genomics and evolutionary studies among organisms.

North Bengal and Sikkim being a part

of Eastern Himalayan biodiversity hotspot have a rich biodiversity of microbial flora (http://www.natureasia.com/en/india 1038/India). However, not much work has been done to study the microbial biodiversity of this region, especially no work has been done on *Rhizobium* and its diversity from these regions.

***The following objectives were undertaken for my research work**:*

- Isolation and authentication of *Rhizobium* strains.
- Morphological and biochemical characterization of the strains following standard methods.
- Physiological characterization of the *Rhizobium* strains including the heavy metal and antibiotic resistance test.
- Study of the use of local strains of *Rhizobium* as PGPR by *in vitro* and *in vivo* study.
- Molecular characterization of the *Rhizobium* strains by RAPD, rep-PCR and PCR-RFLP of 16SrDNA analysis.
- Exploration of the conserved partial 16SrRNA gene by sequencing and its molecular phylogeny analysis.
- Whole genome sequencing of the isolated *Rhizobium* strains.
- Bioinformatics study of the whole genome sequence of *Rhizobium* strains by the codon usage study of nitrogen fixation related genes, siderophore biosynthesis genes and Indole Acetic Acid producing genes.
- Comparative genomics of newly sequenced *Rhizobium* strain with other *Rhizobiun* strains available in public domain database.