Chapter 2
Review of Literature

2.1. Biological Nitrogen Fixation

Biological nitrogen fixation is a process in which the atmospheric nitrogen is converted to ammonia by the enzyme nitrogenase (Postgate 1998). The process was discovered by German agronomist Hermann Hellriegel (Hellriegel and Wilfarth 1888) and Dutch microbiologist Martinus Beijerinck (Beijerinck 1901) in 1991.

$N_2 + 8\ H^+ + 8\ e^- \rightarrow 2\ NH_3 + H_2$

The process is performed in nature by microorganisms that fix nitrogen known as diazotrophs. Diazotrophs are bacteria and archaea which may be free living and symbiotic. The symbiotic diazotrophs are Rhizobia, Frankia and cyanobacteria which fixes nitrogen symbiotically by partnering with a host plant. The plant provides energy in the form of carbon synthesized during photosynthesis to nitrogen fixing microorganisms while in exchange the microbes supply fixed nitrogen to the host plants.

Biological nitrogen fixation, particularly those carried out by prokaryotes is ecologically important as an input of fixed nitrogen in many terrestrial and aquatic ecosystems. It is the second most important process influencing primary productivity after photosynthesis. Annually approximately 2.5 x 10^{11} kg NH_3 is fixed from the atmosphere by legumes and bacteria (Cheng 2008). Currently, to equal the effects of 1 tonne of nitrogen biologically fixed by legume crops approximately 2 tonnes of industrially fixed nitrogen is needed as fertilizer for crop production. Therefore, biologically fixed nitrogen influences the global nitrogen cycle substantially less than industrially fixed nitrogen. Hence, a transition may be required towards agricultural diversification using legume-based crop rotations, which provide a valuable means to reduce the intensity of ammonical fertilization with the

input of less reactive organic N as the depletion of soil organic nitrogen by synthetic nitrogenous fertilizers are disputed (Mulvaney *et al*. 2009).

2.1.1. Legume rhizobia mutualism

Nitrogen fixation can only occur when the plants are in the symbiotic state and the rhizobia invades the root or stem cortex (Cooper, 2004). The two initially independent, free-living partners progresses to the symbiotic state by reciprocal signal generation and perception, which has been described as 'molecular dialogue' (Denarie *et al*. 1993). Plant lectins serves as receptors for bacterial exopolysaccharides (EPS) mediating specificity in the *Rhizobium*–legume symbiosis (Rudiger and Gabius, 2001). In legume *Rhizobium* interactions, bacterial chemotaxis towards plant root exudates is a crucial event. The roots secretes root exudates containing chemical attractants such as flavonoids and betains (Krishnan and Pueppke, 1993). Host legumes are thought to be discriminated from non hosts partly on the basis of the specific flavonoids that they release (Parniske and Downie, 2003). Their main role in the initiation of a rhizobial symbiosis is an interaction with the constitutively expressed *nodD* gene product(s) of the microsymbiont to form a protein-phenolic complex a transcriptional regulator of other rhizobial nodulation (*nod*) genes that are responsible for synthesis of reciprocal signals to the plant root. The nod genes encode approximately 25 proteins required for bacterial synthesis and export of nod factor. Nod factor is a lipo oligosaccharide signal consisting of a chitin backbone, four to five N-acetyl-glucosamine units in length, with a lipid attached to non reducing end and host specific modification on the back bone (Gage, 2004). These host specific modifications include the addition of sulphuryl, methyl, carbamoyl, acetyl, fucosyl, arabinosyl and other groups to different positions on the backbone, as well as differences in the structure of the acyl chain. These variations define much of the species specificity that is observed in the symbiosis (Perret *et al*. 2000). Rhizobial Nod factors induce in their legume hosts the expression of many plant genes like nodulins which helps in developmental processes like the infection, nodule development, and nodule function phases of symbiotic interaction. Receipt of the nod factor in the legume root tip triggers root hairs to curl around the rhizobia, which then forms an infection thread that

transports the rhizobia through the legume root. Upon reaching their destination within a root cell, rhizobia are encased in a plant derived membrane known as a symbiosome. Within the symbiosome, rhizobia ultimately differentiates into a specialized cell type called bacteroids. The cytosol of bacteroids is the site of synthesis of nitrogenase, the enzyme responsible for the reduction of atmospheric nitrogen to ammonium (Somasegaran and Hoben, 1994). Both *nif* and *fix* genes are required for nitrogen fixation after infection of a host. Rhizobia continue to differentiate inside the nodule and synthesize proteins required for nitrogen fixation and for the maintenance of the mutualistic partnership (Gage, 2004).

2.1.2. Legumes in nitrogen fixation

Legumes belonging to the family Fabaceae (Leguminosae) rank third amongst flowering plants in number of species they contain and play unique role in plant kingdom. It has been divided taxonomically into three sub-families, the Papilionoideae, Mimosoideae and Caesalpinioideae comprising 35 tribes, around 1000 genera and over 19,000 species (Lewis *et al*. 2005; Sprent, 2009).

Legumes are grown agriculturally primarily for their food grain seed as these plants are a significant source of protein, dietary fiber, carbohydrates and micronutrients, including folate, thiamin, manganese, magnesium, iron and potassium. Grain legumes also known as pulses include beans, lentils, lupins, peas, and peanuts are cultivated for their seeds which are used for human and animal consumption or for the production of oils for industrial uses. Grain legumes are rich source of dietary protein (Duranti and Gius, 1997), especially for the largely vegetarian population of sub-tropics.

Symbiotic nitrogen fixation in legumes contributes approximately 30% to the amount 63 x 10^6 to 175 x 10^6 tonnes of nitrogen fixed per year through BNF. However the most efficient nitrogen fixing system is the leguminous nitrogen fixation. The nitrogen fixed by agriculturally important legumes annually are around some 40 to 60 million metric tons (Mt) of which another 3 to 5 million Mt are fixed by legumes in natural ecosystems (Smil, 1999). N_2 fixation by legume pastures and crops provides 65% of the N currently utilized in agricultural production (Reeve *et al*. 2015).

2.2.French bean (*Phaseolus vulgaris* L.)

French bean synonyms common bean, bush bean, kidney bean, snap bean, haricot bean and navy bean (*P. vulgaris* L.) is a short duration non-traditional important grain legume grown worldwide both for dry beans and as green beans. Two centers were identified for the origin of *P. vulgaris* (the common bean). The small-seeded cultivars from Mesoamerica and the other in the large-seeded cultivars from Andes of South America (Gepts *et al*. 1986; Gepts and Debouck, 1991). However, over the period of domestication and dispersal, the plant was considered as a major legume food crop, grown worldwide in a broad range of environments and cropping systems and used both as a pulse and as a green vegetable. In both the developed and the developing world, *P. vulgaris* is consumed in many different forms (Rachie and Roberts 1974). Developing countries are the major producers of dry beans, while developed countries produce more green beans (Rachie and Roberts, 1974).

According to Food and Agricultural Organization (FAO) of United Nations, India is the world leader in production of dry seed of French bean followed by Brazil and Myanmar. In India, it is grown on an area of about 1 lakh ha mainly in the states of Maharashtra (60,000 ha), Jammu and Kashmir (10,000 ha), Himachal Pradesh and Uttar Pradesh Hills, Nilgiri (Tamil Nadu) and Palni (Kerala) hills, Chickmagalur (Karnataka) and Darjeeling hills (West Bengal) (Ahlawat 2008). French bean is grown during winter season in the plains of West Bengal. In the hills, it is intercropped between maize and soybean. French bean is cultivated as a crop plant in the traditional farming system in a small scale in the homesteads of eastern hills (Mukherjee, 2014).

It is placed in the tribe Phaseoleae, subfamily Papilionoideae in the family Fabaceae (Leguminosae) (Debouck, 1999; Mabberley, 1997). However, among the different species of *Phaseolus* like *P. acutifolious* A. Gray (Tepary Bean), *P. coccineus* L. (Scarlet Runner Bean), *P. lunatus* L. (Lima, Butter or Madagascar Bean) and *P. polyanthus* Greenman (Year-long Bean), *P. vulgaris* L. is the most widely distributed and has the broadest range of genetic variation (Singh, 1999). Cultivated forms are herbaceous

annuals, which are determinate or indeterminate in growth habit. It is an annual herbaceous plant with erect stem and trifoliate leaves. The plant has terminal raceme and subtending axillary bisexual flowers coloured pink, white or yellow (Begum *et al*. 2003).

P. vulgaris was the first legume in which the *Rhizobium* symbiosis was identified (Taylor *et al*. 1983), although the conflicting evidence of symbiosis still remains. *P. vulgaris* is a promiscuous legume that forms nodules with a variety of rhizobial partners (Michiels *et al*. 1998). Many different *Rhizobium* species have been recognized worldwide as symbiotic partners of common bean, including *R. etli*, *R. phaseoli* (Ramirez-Bahena *et al*. 2008), *R. tropici*, *R. gallicum*, *R. leucaenae*, *R. lusitanum*, *R. pisi*, *R. freirei* and *R. giardinii* (Van Berkum *et al*. 1996; Herrera-Cervera *et al*. 1999; Silva *et al*. 2003; Dall' Agnol *et al*. 2013). Various studies have revealed the diversity of *Phaseolus* rhizobia at the species, intraspecies and population levels and the population of rhizobia present in the nodules of *Phaseolus* sp. have been found to be highly differentiated among sites (Amarger, 2001). However, it is often considered inferior in N_2 fixation among grain legumes (Ali and Lal, 1992; Hardarson, 1993) as it lacks nodulation due to the absence of a nodulation (NOD) gene regulator (Kushwaha, 1994). However, its promising potential to fix nitrogen has been shown in several studies (Figueiredo *et al*. 2008; Ahemad and Kibret, 2014). Hence, the nitrogen requirement of French bean is different from other pulse crops and the application of nitrogen through fertilizers is imperative for exploiting its yield potential.

2.3. Rhizobia

Rhizobia are physiologically heterogeneous and genetically diverse group of bacteria that are nevertheless, classified together by variety of ability to nodulate members of Leguminosae (Somasegaran and Hoben, 1994). Based on the data derived from the use of traditional tests such as biochemical and physiological tests, the rhizobia are broadly classified into two groups, specially, on the basis of growth rate. These two groups are (i) Fast growing and (ii) Slow growing (Jordan and Allen, 1974)**.** One is the fast grower producing acid on yeast-mannitol-agar which includes *R. phaseoli, R. trifolii, R. leguminosarum, R. meliloti* and the

other slow growing strains producing alkali on yeast-mannitol-agar includes *R. lupine, Bradyrhizobium japonicum* and rhizobia of cowpea, miscellany (Tilak 1991). The genetic diversity and taxonomy as well as their plant bacteria molecular interactions of these microorganisms have been extensively studied over the last twenty years because of their ecological and economical importance.

The most up-to-date taxonomy of rhizobia based on 16SrDNA sequence belongs to three main distinct phylogenetic subclasses: α, β and γ-*Proteobacteria.* It consist of 14 genera of nodule-forming diazotrophic bacteria, eleven of which are under the order Rhizobiales in subclass α-proteobacteria (*Aminobacter*, *Azorhizobium, Bradyrhizobium, Devosia, Ensifer, Mesorhizobium, Methylobacterium, Microvirga, Ochrobactrum, Phyllobacterium* and *Rhizobium*) and two 2 genera belonging to the order of Burkholderiales in subclass β-Proteobacteria (*Burkholderia and Cupriavidus*) and one genus belonging to the order Pseudomonales in the subclass *γ*-Proteobacteria (*Pseudomonas*) (Berrada and Fikri-Benbrahim, 2014).

2.4. Rhizobium

In 1888, Beijerinck isolated a root nodule bacteria and named it *Bacillus radicicola* (Beijerinck, 1888) which was later named by Frank as *Rhizobium leguminosarum* (Frank, 1889).

Rhizobium is a type genus of the family Rhizobiaceae (Conn, 1938) belonging to the order Rhizobiales in the class α-proteobacteria (Mousavi *et al*. 2015) which was established in 1889 by Frank (Fred et al. 1932). The genera *Rhizobium*, contains about 44 recognized species including some latest novel species like *R. sphaerophysae*, *R. pusense*, *R. vallis* and *R. herbae* (Qin *et al*. 2012).

It can best be described as rod shaped, heterogeneous group of gram negative, aerobic, heterotrophic, non-spore forming microbe (Jordan and Allen, 1974; Hirsch *et al*. 1992; Moschetti *et al.* 2005). It usually contains granules of poly-β-hydroxybutyrate which are refractile by phase contrast microscopy. It produces an acidic reaction in mineral-salts medium containing mannitol or other carbohydrates. These bacteria do not produce 3-ketoglycosides (Bernaerts

and De Ley, 1963). All the strains of *Rhizobium* exhibit host range affinities. *Rhizobium* is among the best known beneficial plant-associated bacteria because of the importance of the nitrogen fixation occurring during *Rhizobium*-legume symbiosis (Hung and Annapurna, 2004).

2.4.1. Taxonomy of Rhizobium

Rhizobium is the type genus and encompasses the largest number of species into the family. However, the original genus *Rhizobium* has undergone several subsequent changes in recent years giving rise to many other taxa. Nodulation tests on host plants to determine cross-inoculation groups (Fred *et al*. 1932), traditional phenotypic methods including analysis of colony morphology in culture medium; intrinsic antibiotic resistance; biochemical, metabolic and nutritional characteristics; bacteriophage susceptibility; and serological reactions (Graham, 1963; Moffett and Colwell, 1968; Vincent and Humphrey, 1970) were applied to identify and characterize these bacteria. In the1984 edition of Bergey's Manual of Systematic Bacteriology the genus *Rhizobium* along with *Bradyrhizobium, Agrobacterium*, and *Phyllobacterium* were the four genera included in the family Rhizobiaceae and the separation of these genera was predominantly based on the ability to stimulate the production of root or leaf nodules in host plant species (Jordan, 1984). Finally, simple characteristics based on DNA molecular analysis were added to the numerical taxonomic studies (Somasegaran and Hoben, 1994; Zakhia and De Lajudie, 2001). However, the recent taxonomy of Rhizobiaceae, as well as of any other bacterial groups is mostly supported by phylogenetic analyses based on 16SrDNA sequences. A revision and dismemberment of the genus *Rhizobium* and its relatives of the class α-proteobacteria was led by their phylogenetic studies. *Rhizobium, Agrobacterium, and Allorhizobium* of family Rhizobiaceae formed a more closely related group whereas *Sinorhizobium* (Rhizobiaceae), *Bradyrhizobium* (Bradyrhizobiaceae), *Mesorhizobium* (Phyllobacteriaceae), and *Azorhizobium* (Hyphomicrobiaceae) formed separate groups (Willems, 2006). Recently more sophisticated approaches such as matrix-assisted laser desorption/ ionization-time of flight (MALDI-TOF) mass spectrometry, and phylogeny based on gene sequencing

has replaced the older techniques. 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. Also the Multilocus Sequence analysis (MLSA) of housekeeping genes exhibiting sufficient conservation and distribution in bacterial genomes has become more favored and acceptable method to establish taxonomy of prokaryotes (Martens *et al*. 2008).

The family Rhizobiaceae were found to be composed of at least nine genera: *Rhizobium* (type genus), *Agrobacterium, Carbophilus, Chelatobacter, Ensifer, Sinorhizobium, Allorhizobium, Kaistia, and Shinella* according to workers like Euzeby (1997), Kuykendall (2005) and Parte (2014). However the genera *Agrobacterium* and *Allorhizobium* are now included in the *Rhizobium* genus (Young *et al.* 2001). According to Young (2010) the family consisted of seven genera viz. *Rhizobium, Agrobacterium, Ensifer*, *Sinorhizobium*, *Carbophilus*, *Kaistia*, and *Shinella*. *Sinorhizobium* is currently named as *Ensifer* (Young 2010). However, during the last decade, species have been arranged into five genera, *Rhizobium*, *Ensifer*, *Carbophilus*, *Kaistia*, and *Shinella* (An et al. 2006; Lin *et al.* 2008). Phylogenomic analyses showed two major superclades within the family Rhizobiaceae that corresponded to the *Rhizobium/Agrobacterium* and *Shinella/Ensifer* groups (Ormeno-Orrillo *et al*. 2015). The four highly supported clades within the *Rhizobium/ Agrobacterium* group corresponded to distinct genera are *Rhizobium, Allorhizobium, Neorhizobium* (Mousavi *et al*. 2014) and *Agrobacterium* (Costechareyre *et al*. 2010). In the Shinella/Ensifer group, the sole sequenced strain of *Shinella* formed a clade of its own, and a second clade could be equated to the genus *Ensife*r (Ormeno-Orrillo *et al*. 2015). Over the last 10 years there has been new challenges for Rhizobiaceae researchers to analyze and classify both genus and species levels as the diversity of Rhizobiaceae is vast and a deluge of genomic data are being generated by different genome and environmental sequencing project.

2.4.2. Studies on Rhizobium

2.4.2.1. Biochemical studies

Biochemical tests are still one of the main criterions utilized to characterize

the different isolates of rhizobia including different species of *Rhizobium.* Till date various types of microbiological tests has been performed on *Rhizobium* like the nodulation test, colony morphology, acid alkaline production, ketolactase production, catalase production, oxidase production, urease production, liquefaction of gelatin, indole production, reduction of methylene-blue and methyl-red, production of levan, Vogespraskauer reaction and fermentation of carbohydrates, nitrate reduction, starch hydrolysis etc (Bernaerts and De Ley, 1963; Graham and Parker, 1964; Vincent, 1970; Jordan, 1984; Kennedy, 1994; Aneja, 1996; Rodriguez-Navarro *et al*. 2000; Kucuk, *et al*. 2006; Singh, *et al*. 2008; Bedi and Naglot, 2011; Mekasha *et al.* 2015; Rai and Sen, 2015).

Yeast extract mannitol (YEM) was found to be the selective medium for the better growth of *Rhizobium* (Vincent, 1970). The morphology of the *Rhizobium* colonies appeared circular, convex, semi translucent, raised and mucilaginous .The consistency of the colonies varied depending upon the source of *Rhizobium* sp (Vincent, 1970). Most strains lacked the ability to absorb Congo red (0.0025%) dye from a yeast extract mannitol mineral salts medium whereas contaminant colonies like that of *Agrobacteria* are often a deep red. At pH 11.0 *Rhizobium* could not grow in Hofer's alkaline medium, a characteristics which was considered suitable for differentiating rhizobia from *Agrobacteria* (Allen and Allen, 1958). Majority of Rhizobial strains showed poor or no growth on glucose peptone agar medium, whereas Agrobacteria grew well (Klezkowska *et al*. 1968). The optimum pH and temperature for the growth of the *Rhizobium* isolates was found to be at pH 6-7 and 25-30ºC respectively (Singh *et al*. 2008; Bedi and Naglot, 2011; Bhatt *et al*. 2013). However *R. meliloti* can grow even at 42.5ºC (Jordan, 1984). There have been several studies and the isolates have shown disparity in tolerance to different concentration of salt (Fentahun *et al.* 2013). The production of amylase and cellulase by *Rhizobium* has been reported by workers like (Singh *et al*. 2008).

The ability of *Rhizobium* to metabolize broad range of sugars, organic acids and aromatic compounds is well documented (Stowers, 1985). The studies in carbon nutrition and

metabolism in free-living cells has provided a baseline for comparison among strains. For *R. leguminosarum* bv. *trifolii*, the carbon utilization patterns have been used for phenotypic comparison between plasmid-cured strains (Baldani *et al*. 1992). The BIOLOG substrate utilization patterns have been used to confirm phenotypic similarities between electrophoretic types given by multilocus enzyme electrophoresis (MLEE) (Leung *et al*. 1994). Numerical taxonomy of *Rhizobium* strains from legumes was also done by workers to find the phylogenetic relatedness with other strains. Novikova *et al*. (1994) studied the phenotypic properties (growth characteristics, utilization of carbon and nitrogen sources, and intrinsic antibiotic resistance) of 53 *Rhizobium* strains isolated from root nodules of the temperate–zone legumes *Astragalus* sp, *Oxytropis campanulata*, *Hedysarum alpinum*, *Ononis arvensis*, *Glycorrhizza* sp and *Coronilla varia*. The strains were grouped into 14 clusters on the basis of their metabolic properties. A similar numerical taxonomic analysis based on physiological and biochemical tests was done by the Unweighted pair group method with average (UPGMA) clustering method by Maatallah *et al*. (2002).

Though the genus *Rhizobium* can be differentiated with other morphologically or physiologically similar genera as well as within its own species with certain differential characteristics (Jordan, 1984) the biochemical and physiological attributes are not found to be suitable for the identification of *Rhizobium* (Tilak 1991).

2.4.2.2. Antibiotic resistance profile

Antibiotic resistance is a plasmid borne character in bacteria (Cole and Elkan, 1979). The bacteria are known to possess resistance or susceptibility towards known antibiotic substances, which need the mechanism of self-protection (Gray and Fitch, 1983; Prasuna, 2014). There are 3 known determinants of bacterial permeability to an antibiotic: hydrophobicity, electrical charge and amount of the antibiotic (Deora and Singhal, 2010; Bhattacharya *et al*. 2013). The intrinsic antibiotic resistance (IAR) method can be used as a fingerprint to identify strains. This methodology uses the natural resistance of strains to given levels of various antibiotics for identification (Beck *et al*. 1993) and

the detection of antibiotic markers is considered more practical than other methods because the methodology is simple, reliable and non-expensive (Mueller *et al*. 1988; Brockman and Bezdicek, 1989; Abaidoo *et al.* 2002). Specific antibiotics help in identifying certain markers to distinguish bacterial isolates from each other (Dakora, 1984; Prasuna, 2014). *Rhizobium* has been found to be more sensitive to tetracycline, penicillin G, vancomycin and streptomycin (Jordan, 1984). Resistance to streptomycin has been one of the most frequently used markers for *Rhizobium* (Borges *et al*. 1990; Amer, 2008).

Disk diffusion test based on the presence or absence of a zone of inhibition without regard to the size of the zone has been peformed by several workers to determine the IAR of the strains (Cole and Elkan, 1979; Gupta *et al*. 1983; Hungria *et al.* 2001; Alexandre *et al*. 2006; Singh *et al.* 2008; Rashid *et al*. 2012; Bhattacharya *et al*. 2013; Hewedy *et al.* 2014; Ahmed and Abdelmageed, 2015). However, disk diffusion tests based solely on the zone of inhibition are not acceptable for determining antimicrobial susceptibility. Reliable results can only be obtained with disk diffusion tests that use the principle of standardized methodology and zone diameter measurements correlated with minimal inhibitory concentrations (MICs) with strains known to be susceptible or resistant to various antimicrobial agents. The most thoroughly described disk diffusion method is the current standardized method for which interpretive standards have been developed and supported by laboratory and clinical data which is based on the method originally described by (Bauer *et al*. 1966). The CLSI (Clinical and Laboratory Standards Institute) subcommittee on Antimicrobial Susceptibility testing has recommended the current standardized method.

2.4.2.3. Heavy metal resistance

The key pollutants causing serious illness to plants, ecosystem and humans by their non-degradable nature are the heavy metals and the most important process of our biosphere, the symbiotic nitrogen fixation is no exception (Stan *et al*. 2011). In order to preserve the natural soil properties and microbial biomass phytoremediation is suggested to be practised for the reclamation and removal of heavy metals (Gopalakrishnan *et al*. 2015).

The microorganisms used for the photremediation process are *Bacillus* sp, *Pseudomonas* sp, *Azotobacter* sp, *Enterobacter* sp, and *Rhizobium* sp. (Ma *et al*. 2011; Gopalakrishnan *et al*. 2015). Due to the importance of legumes in animal and human consumption and their use in maintaining soil fertility, some attention has been given to the effects that heavy metals exert on *Rhizobium* isolates (Pereira *et al*. 2006; Stan *et al*. 2011; Vasilica *et al*. 2011). The increase in the concentration of Cu, Zn and Pb either sole or in combination led to the reduction in the bacterial counts of *Rhizobium* sp. (Stan *et al*. 2011; Gopalakrishnan *et al*. 2015). Effect of heavy metals on the growth, abundance, morphology and physiology of various strains have been well documented (Gopalakrishnan *et al*. 2015). A variation in the protein profile due to the change in physiology of *R. leguminosarum* bv. *viciae* isolated from the heavy metal polluted site served as a marker for stress response analysis (Pereira *et al.* 2006). A variation in the number of plasmids in tolerant strains was observed by different workers relating to the heavy metal resistance to the plasmids (Pereira *et al*. 2006).

2.4.2.4. Rhizobium as PGPR

Rhizobacteria that benefit plant growth and development are called 'PGPR' first defined by Kloepper and Schroth (Kloepper and Schroth, 1978). Plant growth promoting rhizobacteria (PGPR) promotes 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). The legume associated rhizobia and the woody plant associated *Frankia* sp. are known as intracellular PGPR (iPGPR) as these bacteria lives inside cells in specialized root structures (Gopalakrishnan *et al*. 2015). The most studied and longest exploited PGPR are the rhizobia (including the *Allorhizobium, Azorhizobium, Bradyrhizobium, Mesorhizobium*, *Rhizobium*, and *Sinorhizobium*) for their ability to fix N_2 in their legume hosts (Vessey, 2003).

The potential phosphate(P) solubilizers within the rhizobia group are *R. leguminosarum*, *R. meliloti*, *Mesorhizobium mediterraneum*, and *B. japonicum* (Vessey, 2003; Egamberdiyeva *et al*. 2004; Rodrigues

et al. 2006; Afzal and Bano, 2008; Gopalakrishnan *et al*. 2015). The Phosphate Solubilzing Microorganisms (PSM) are able to convert inorganic and organic soil phosphorus into the bioavailable form through various mechanisms of solubilization and mineralization facilitating uptake by plant roots (Sharma *et al*. 2013) in the soluble forms [the monobasic ($H_2PO_4^-$) and the dibasic (HPO_4^{2-}) ions] in the rhizosphere (Richardson, 2001). Production of 2-ketogluconic acid with a phosphate-solubilizing ability has been identified in *R. leguminosarum* and *R. meliloti* (Halder *et al*. 1990; Gopalakrishnan *et al*. 2015). The solubilization of phosphate by a strain of *R. leguminosarum* bv. *trifolii* ARPV02 isolated from *P. vulgaris* has been found (Abril *et al*. 2007) and *Rhizobium* species nodulating *Crotalaria* species (Sridevi *et al*. 2007) have found to improve plant P nutrition.

Rhizobial species such as *R. meliloti*, *R. tropici*, *R. leguminosarum* bv. *viciae*, *R. leguminosarum* bv. *trifolii*, *R. leguminosarum* bv. *phaseoli*, *S. meliloti* and *Bradyrhizobium* spp. are known to produce siderophores (Chabot *et al.* 1996; Antoun *et al*. 1998; Arora *et al*. 2001; Gopalakrishnan *et al*. 2015). Siderophores are low molecular weight iron binding molecules secreted by several microorganisms which sequesters the Fe^{3+} iron from the environment making it available in a soluble form (Fe^{2+}) accessible for the plant (Bashan and de Bashan, 2005). Most siderophores belong to the hydroxymate and catechol classes (Derylo *et al.* 1994). Siderophores also has the ability to relieve the plants from heavy metal stress facilitating the iron uptake (Neubauer *et al.* 2000; Gopalakrishnan *et al*. 2015). Siderophores produced by PGPB has a higher affinity for the iron available in the soil in comparison to the fungal pathogens protecting the plants from various fungal diseases (Bashan and de Bashan, 2005). *R. leguminosarum* ATCC 14479, a symbiont of red clover *Trifolium pratense*, was found to produce a trihydroxamate siderophore (Wright and William, 2010).

The IAA producers are *Azorhizobium caulinodans*, *Bradyrhizobium japonicum*, *B. elkanii*, *Mesorhizobium loti*, *Rhizobium japonicum*, *R. leguminosarum*, *R. lupine*, *R. meliloti*, *R. phaseoli*, *R. trifolii* and *Sinorhizobium* spp. (Antoun *et al*. 1998; Biswas *et al*. 2000; Boiero *et al*.

2007; Afzal and Bano 2008; Weyens *et al.* 2009; Chi *et al.* 2010; Gopalakrishnan *et al*. 2015). Indole-3-acetic acid (IAA) is a phytohormone involved in root initiation, cell division and cell enlargement (Salisbury, 1994; Vessey, 2003) accelerating the plant growth and development. *Rhizobium* strains secrete IAA which plays an important role in the formation and development of root nodules and shows positive influence on plant growth (Nutman 1977; Kumari *et al*. 2009). IAA production in *Rhizobium* takes place *via* indole-3-pyruvic acid and indole-3-acetic aldehyde pathway (Gopalakrishnan *et al*. 2015). *Rhizobium* isolated from *P. mungo* produced a large amount of IAA (142 µg/mL) from L-tryptophan supplemented basal medium (Ghosh *et al.* 2008) and also from *Indigofera* (Kumari et al. 2009).

Rhizobium along with other bacterial genera like *Pseudomonas*, *Xanthomonas*, *Erwinia*, *Arthrobacter*, *Bacillus*, *Acinetobacter* and *Aeromonas* are found to be hydrocyanic acid (HCN) producers (Bhawsar, 2014). The cyanogenic bacteria have been found to be effective in biocontrol of weeds, pathogenic fungi and nematodes. For the chemical insecticides, weedicides and fungicides used in agriculture polluting the environment, the use of these cyanogens are a superlative replacement (Kremer and Souissi, 2001). *Rhizobium* sp isolated from pea and lentil has been found to be cyanogen producers (Ahemad and Khan, 2010; Ahemad and Khan, 2012).

Rhizobial strains producing ACC deaminase are *R. leguminosarum* bv. *viciae*, *R. hedysari*, *R. japonicum*, *R. gallicum*, *B. japonicum*, *B. elkani*, *M. loti* and *S. meliloti* (Gopalakrishnan *et al*. 2015). ACC(1-Aminocyclopropane-1-carboxylate)deaminase is an enzyme responsible for the cleavage of the plant ethylene precursor, ACC into ammonia and α-ketobutyrate (Honma and Shimomura, 1978). The production of enzyme lowers the plant ethylene level which when present in high concentration inhibits the growth of plants (Honma and Shimomura, 1978; Glick *et al*. 2007). *Rhizobia* which possesses the enzyme normally helps in the growth of the root and shoots providing relief from stresses, such as heavy metals, pathogens, drought, radiation, salinity etc.

The rhizobial strains having biological control agent property are *R. leguminosarum* bv. *trifolii*, *R. leguminosarum* bv. *viciae*, *R. meliloti,*

R. trifolii, *S. meliloti* and *B. japonicum* (Ozkoc and Deliveli, 2001; Shaukat and Siddiqui, 2003; Bardin *et al*. 2004; Chandra *et al*. 2007; Gopalakrishnan *et al.* 2015). *R. leguminosarum* bv. *phaseoli* was found to inhibit the growth of *Fusarium solani* leading to a significant reduction in the root rot disease (Buonassisi *et al*. 1986; Antoun *et al*. 1998). Similarly *B. japonicum*, *R. meliloti* and *R. leguminosarum* were found to control the fungal pathogens like *Macrophomina phaseolina*, *Rhizoctonia solani* and *Fusarium solani* infecting the plants like okra and sunflower.

The most sensitive method to elucidate biological N_2 fixation activity is the acetylene reduction assay (ARA). The assay is based on the reduction of acetylene (C_2H_2) to ethylene (C_2H_4) by the nitrogenase enzyme providing a useful assay for the quantification of the N_2-fixation process (Dilworth, 1966). The gas ethylene formed during the reduction can be very sensitively detected by gas chromatography. The method is reported to be simple, economical and a thousand times more sensitive than the ^{15}N method (Hardy *et al*. 1968). Estimates of fixed nitrogen in the soil using the acetylene reduction method provided data closer to the natural environment than in the isotope ^{15}N method (Yoshida and Ancajas, 1970). An increase in the nitrogenase activity has been observed in the *Rhizobium* strain Rh7 among individual application and in RH3 + UTPF109 and RH6 + UTPF68, among co-applications in common bean by (Samavat *et al*. 2012). The nitrogenase activity of *R. leguminosarum* bv. *viciae* STDF-Egypt 19 was found to increase in co inoculation with the mixture of AMF (*Acaulospora laevis*, *Glomus geosporum*, *Glomus mosseae* and *Scutellospora armeniaca*) over control in faba bean (*Vicia faba*).

Inoculation of legumes with symbiotic *Rhizobium* as a biofertilizer has been practiced for almost 100 years and has had a major impact worldwide on crop yields (Bashan and de Bashan, 2005). Biofertilizer are defined as substances that contains living microorganisms which helps in the growth promotion by increasing the nutrient status of the plant. Commercial rhizobia inoculants for use on legume crops were first introduced in the 1890s (Fred *et al*. 1932; Vessey 2003). The peat based inoculants were found to be superior among the various carrier materials. However, powdered lignite or charcoal or a combination of them has been

universally adopted by Indian producers for it is easily available (Tilak, 1991). More advanced formulations like the synthetic carriers in micro and macro bead forms or powders on seed coatings are being formulated (Bashan and de Bashan, 2005). The yield of grains like pigeonpea, greengram, blackgram, cowpea, gram and lentil ranged from 2-65% over the control due to *Rhizobium* inoculation (Tilak, 1991). Inoculation of the common bean with the superior strain of *Rhizobium tropici* (H 12, H 20, PRF 81 and CIAT 899) with the input of low level of N fertilizer increased the yield, nodulation and its N_2 fixation rates (Hungria *et al*. 2003).

2.4.2.5. Molecular studies on Rhizobium

Characterization of the *Rhizobium* genome at the molecular level is the most discriminating method for assessing the variability among strains and isolates of the bacteria (Demezas *et al*. 1991; Thies *et al*. 2001). *Rhizobium* researchers have used various primers to obtain 'PCR-fingerprints' which are used to characterize rhizobial isolates at the strain level. The primers most frequently used are designed to target specific DNA fragments, e.g. 16S ribosomal RNA genes, repetitive element sequences (REP, ERIC and BOX), 16S–23S rRNA intergenic spacer regions or genes for nitrogen fixation and nodulation (de Bruijn, 1992; Thies et al. 2001).

The RAPD technique is a potential tool for the identification of the genetics and systematic of different populations. This technique use arbitrary primers to detect changes in the DNA sequence at sites in the genome which anneal by the primer and generate randomly amplified polymorphic DNA (RAPD) fragments (Harrison *et al*. 1992; Richardson *et al.* 1995). Harrison *et al.* (1992) applied RAPD primers to the DNA from isolates of *R. leguminosarum* biovar. *trifolii* and the resulting fingerprints allowed strain differentiation. Phylogenetic relationships are constructed from RAPD fingerprints by statistical analysis. Young and Cheng (1998) working on fast and slow growing rhizobia used this technique to depict the genetic relationships of six strains of rhizobia, including three strains of *R. fredii* and three strains of *B. japonicum*. They found the RAPD markers to be highly useful in the construction of genetic maps, referred to as RAPD mapping. RAPD technique is a potentially useful tool for the study

of genetics and systematic (Young and Cheng, 1998). The estimated diversity through RAPD analysis was more evident than the diversity on the basis of morphological and biochemical characters (Suman *et al*. 2001) and was found to be an efficient means for rapidly typing a large number of strains and estimation of their diversity.

However, in RAPD-PCR, because of the short primer length (generally 10 mers), and the need to generate a sufficient number of DNA fragments for reliable characterization of the template genomic DNA, minute variations in the concentration and purity of template DNA, as well as the amount of primers and Taq polymerase, may affect the efficiency with which strains can be discriminated (Berg *et al*. 1994).

rep-PCR is a genotypic fingerprinting method that generates specific strain patterns by the amplification of repetitive elements present in the genome of Gram-negative and Gram-positive bacteria probably located in distinct intergenic positions of bacteria (Versalovic *et al*. 1991; de Bruijn, 1992; Leung *et al.* 1994). Five rep-PCR methods, such as REP-PCR (primer sets Rep1R-I and Rep2-I); ERIC-PCR (primer sets ERIC1R and ERIC2); ERIC2-PCR (primer ERIC2); BOX-PCR (primer BOX A1R) and (GTG)5-PCR [primer (GTG)5] are commonly used for genotyping of different bacterial strains (Versalovic *et al.* 1994). Three main sets of repetitive DNA elements containing highly conserved palindromic inverted repeat sequences have been used for typing bacteria: the 35 to 40-pb repetitive extragenic palindromic (REP) sequences found in about 500 copies dispersed around the chromosomes of *Escherichia coli* and *Salmonella typhimurium* (Stern *et al*. 1984), the 124 to 127 bp enterobacterial repetitive intergenic consensus (ERIC) sequences present in many copies in the genomes of enterobacteria (Hulton *et al*. 1991), and the BOX elements, an invertedly repeated DNA element found in *Streptococcus pneumoniae* (Martin *et al*. 1992). This third class consists of differentially conserved subunits, namely box A, box B, and box C though only the box A-like subunit appears highly conserved across diverse bacteria (Menna *et al*. 2009). These repetitive elements, located in the intergenic regions of many bacterial genomes, are considered to be highly conserved (Martin et al. 1992) and as such are useful for elucidating

relationships within and between bacterial species including rhizobia (de Bruijn, 1992; Chen *et al*. 2000). Many authors proposed these methods for identification and phylogenetic grouping of *Rhizobium* isolates (Labes *et al*. 1996; Sikora and Redzepovic 2003; Blazinkov *et al.* 2007). A polyphasic approach, combining ERIC, REP and BOX-PCR profiles, has been used to improve the precision of genetic discrimination (Louws *et al*. 1994; De Bruijn *et al*. 1996). Blazinkov *et al.* (2007) assessed genetic diversity and characterization of indigenous *R. leguminosarum* bv. *viciae* strains isolated from different regions in Croatia with both RAPD and rep-PCR. The fingerprints obtained were specific and reproducible allowing the distinction of bacteria at the (sub) species and strain level (Blazinkov *et al*. 2007). The genotypic characterization and phylogenetic analysis of *R. leguminosarum* subsp. *ciceri* strains isolated from perennial wild chickpeas collected from high altitudes in Erzurum-Turkey was done by rep (REP, ERIC and BOX) PCR fingerprinting methods.

The 16SrRNA gene is the most commonly used molecular marker to study the rhizobial genetic diversity in different sections of the rhizosphere (Thies *et al*. 2001). 16SrRNA or rDNA contain highly conserved regions because of their crucial structural and functional constraints, but also contain highly variable signatures. Comparison of 16SrDNA nucleotide sequences can be applied for classification of isolates at species and higher levels (Cilia *et al*. 1996) and have been used extensively to determine taxonomy, phylogeny (evolutionary relationships) and to estimate rates of species divergence among bacteria. In fact, there is no other gene that has been as well characterized in as many species. Still, highly conserved ribosomal genes and in particular the 16S and 23S rRNA genes appear to be very useful to infer phylogeny among organisms as distantly related as the different genera included in the rhizobial group (Thies *et al*. 2001).

The 16SRNA gene sequence are amplified with various universal primers and the PCR products are then sequenced. The Gene sequences are submitted to and maintained within various databases such as the Ribosome Database Project or GenBank. Sequences from tens of thousands of clinical and environmental isolates are available

over the internet through the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov) and the Ribosomal Database Project (www.cme.msu.edu/RDP/html/index.html). Continued development of databases through DNA sequencing is essential and is a prerequisite for a good primer design. Today, the accumulated 16SrRNA sequences (about 10,000) constitute the largest body of data available for inferring relationships among organisms. To differentiate between strains within a species 16SrDNA analysis may not be sufficient. However, due to the relatively higher sequence divergence in the interspacer region between the 16S and the 23SrDNA, it can be used to differentiate organisms at the intraspecific levels (Oliveira *et al*. 2000). Denaturing gradient gel electrophoresis (DGGE) combined with 16S-23SrDNA PCR segment has been used in a preliminary screening to aid recognition of sample diversity (Thies *et al.* 2001) and to infer the phylogenetic affiliation of the community members (Oliveira *et al*. 2000).

However, variations in 16SrRNA genes can also be estimated by restriction fragment length polymorphism (RFLP) analysis of 16SrDNA sequences product or intergene spacer region (IGS) such as that found between the 16S and 23SrRNA genes amplified by polymerase chain reaction (PCR) also known as Amplified ribosomal DNA restriction analysis (ARDRA). It is now a common method for the genotypic classification of bacteria including the root nodule bacteria (Safronova *et al*. 2004). ARDRA has been used extensively in rhizobial ecology studies to characterise isolates. To study the genetic diversity of rhizobia belonging to different genera PCR–RFLP analysis of the 16S–23S rDNA intergenic transcribed spacer (ITS) has been demonstrated as a useful tool (Rahmani *et al*. 2011).

Other than 16SrRNA, the Multi locus specific analysis (MLSA) of protein coding housekeeping genes like atpD (ATP synthase F1, betasubunit), recA (recombinase A), rpoB (RNA polymerase, betasubunit), atpD, glnII and recA showing conserved and variable regions are also used as molecular evolutionary markers to define the specific position of the rhizobial isolate. In comparison to the separate analysis, the combined analysis can illustrate the phylogeny

better (Cao et al. 2014). Also several symbiotic genes like *nod C* and *nif H* in combination are studied to draw plant bacteria co-evolution (Thies *et al*. 2001). Several authors have constructed phylogenies based on *nod* gene sequences and have concluded that they are not congruent with those based on SSU rRNA. For example the *nod ABC* and *D* genes of *R. trifolli* are closer to *R. viciae* and *S. meliloti.* Therefore, some authors have considered that the study of the genes involved in nitrogen fixation must be worked out separately from that of the bacteria that now carry them (Young and Haukka, 1996). However, symbiotic genes are not useful in rhizobial taxonomy as they are prone to horizontal transfer between rhizobial species and genera as these genes are frequently located on plasmids or chromosomal symbiotic islands eventually leading to complications in the process of rhizobial characterization (Young *et al*. 2006; Galardini *et al*. 2011). However, the modern *Rhizobium* taxonomy is based on the polyphasic approach whereby a range of genotypic and phenotypic data and information on groups of isolates are integrated to generate a consensus type of taxonomy for a more stable classification (Vandamme *et al.* 1996; Zakhia and De Lajudie, 2001; Vinuesa *et al*. 2005).

2.5. Whole genome sequencing of *Rhizobium*

In the early to mid 80's, sequences of individual genes for *nif*, *nod* and *fix* for rhizobia started appearing in public domains but, full scale genome sequencing started in 1997 when funds became available (Perret and Broughton, 2001). It is generally accepted that all taxonomic information about a bacterium is incorporated in the complete nucleotide sequence of its genome (Goris *et al*. 2007). In bacterial taxonomy, whole genomic DNA relatedness is considered to provide the absolute resolution. To date, 29 complete and 141 draft whole genome sequences (WGS) from members of the family Rhizobiaceae are available from the GenBank database (Ormeno-Orrillo *et al.* 2015). The breakthrough in unearthing the gene maneuvering propagation of nodule primordial opened a new dimension in functional genomes of *Rhizobium*-legume symbiosis. *Mesorhizobium loti* was the first sequence of a symbiotic bacterium followed by *Sinorhizobium meliloti* (Puhler *et al*. 2004), *Bradyrhizobium*

(Kaneko *et al.* 2002) and *Mesorhizobium huakuii* 7653R (Wang *et al.* 2014). The genome sequencing of 6 strains belonging to the genus *Rhizobium* has been completed (Wang *et al*. 2013). Those sequenced strains of *Rhizobium* are as follows: *R. leguminosarum* bv. *trifolii* strain WSM1325(Reeve *et al*. 2010b), *R. leguminosarum* Rlv3841(Young *et al.* 2006b), *R. etli* CFN42 (Gonzalez *et al*. 2006), *R. etli* CIAT 652(Gonzalez et al. 2010), *R. leguminosarum* bv. *trifolii* WSM2304 (Reeve *et al*. 2010a), and *R. rhizogenes* K84 (Slater *et al.* 2009). From India, so far whole genome sequence of *R. lupini* HPC (L) isolated from saline desert soil, Kutch (Gujarat) has been done (Agarwal and Purohit, 2013). Under the Vavilov centers of diversity and GEBA-RNB (Genomic Encyclopedia of Bacteria and Archaea, The Root Nodulating Bacteria chapter) projects at the U.S. Department of Energy, Joint Genome Institute, 107 RNB strains isolated from diverse legume hosts in various geographic locations around the world are being sequenced using Illumina, PacBio or Roche sequencing platforms followed by the DOE-JGI genome annotation pipeline. This project will aid in the development of a phylogenetically balanced genomic representation of the microbial tree of life and allow for the large-scale discovery of novel rhizobial genes and functions (Reeve *et al*. 2015). High-quality permanent draft genome sequence of *R. leguminosarum* bv. *viciae* strain GB30, an effective microsymbiont of *Pisum sativum* growing in Poland (Mazur *et al*. 2015) has been sequenced. Phylogenetic analysis revealed that GB30 was closely related to *R. leguminosarum* bv. *trifolii* CB782 and WSM1689, both part of the GEBA-RNB project (Mazur *et al.* 2015; Reeve *et al.* 2015). The genome project was deposited in the Genome Online Database (Pagani *et al*. 2012) and the high-quality permanent draft genome sequence in IMG (Markowitz *et al*. 2013). Sequencing, finishing and annotation were performed by the JGI using state of the art sequencing technology. The completion of the genomes of strains and sequences for a number of other *Rhizobium* strains spanning different habitats and ecological niches has bolstered nitrogen fixation research. The whole genome sequence of *R. leguminosarum* biovar. *viciae* strain 3841 revealed the 7.75 Mb genome to comprise of a circular chromosome and six circular plasmids with 61% G+C

overall. All three rRNA operons and 52 tRNA genes and protein-encoding genes were found on the chromosome while most functional classes occurred on plasmids as well (Young *et al*. 2006b).

2.6..The bioinformatics approach

In the pre-genomic era although the amalgamation of the knowledge of plant physiology, biochemistry, genetics and molecular biology gave idea about the understanding of the mechanism of nitrogen fixation, the knowledge remained erratic and discrete (Benedito *et al*. 2006). However, things changed with the accessibility of complete genome sequences of symbiotic as well as non-symbiotic diazotrophs. Knowledge of the whole genome became the stepping stone in understanding the working principle of the bacterial cell (Puhler *et al.* 2004). The breakthrough in unearthing the gene maneuvering propagation of nodule primordial opened a new dimension in functional genomes of *Rhizobium*-legume symbiosis.

The ease of access in retrieving complete genome sequences of symbiotic bacteria from databases gave insights into the working principle of bacterial cell and exposed new facts pertaining to evolution, structure, interaction between plants and microbes and the diversity amongst them. With the new advancement in technology, various software for genomic and proteomic analysis have accelerated research in areas of codon usage, proteome analysis, molecular modeling, and phylogenetic analysis of whole genome, nitrogen fixing genes or biosynthetic genes of important pathway etc. With the current availability of inundate data, computational methods have become indispensable to biological investigations. Comparative genomics study with the aid of bioinformatics tools has given an insight upon the unique story of every genome with the genetic code and its usage preference being one of the most interesting aspects of biological science. In the recent years various parameters of codon usage such as GC content, GC3 content, Fop (Frequency of Optimal codons) (Ikemura, 1985), relative synonymous codon usage (RSCU) (Sharp *et al*. 1986), effective number of codons (Nc) (Wright, 1990), Codon Bias Index (CBI), Codon Adaptation Index (CAI) (Sharp and Li, 1987) etc. have been developed which proved to

be noteworthy in studies regarding codon usage patterns. It has been postulated that major trends in codon usage patterns across genomes are determined by compositional bias, mutational pressure and/or translational selection in high or low G+C containing organisms (Knight *et al*. 2001). Highly expressed genes are influenced by translational selection compared to lowly expressed ones which are influenced by mutational pressure (dos Reis *et al*. 2003). Gradually bioinformatics analysis of codon usage in *E.coli* (Peden, 2000) paved its pathway towards exploring codon usage analysis upon mammalian, bacterial, bacteriophage, viral and mitochondrial genes. The codon usage bias of the nitrogen fixing diazotrophs like rhizobia has received increasing attention (Wang *et al.* 2013). With the onset of preliminary work on codon usage of nitrogen fixing diazotrophs (Mathur and Tuli, 1991), differences in codon usage and GC content in *Bradyrhizobium* genes was reported (Ramseier and Gottfert, 1991). Translational selection was the moderate factor governing the codon bias in nitrogen fixing genes of *B. japonicum* USDA 110 (Sur *et al*. 2005). The three strains of *Frankia* genomes Cc13, ACN14a and EAN1pec revealed that codon usage was highly biased and variations were noticed among the three strains (Sen *et al*. 2008). A function of rare TTA codon was also explored in the genome of diazotrophic actinomycete *Frankia* (Sen *et al*. 2012). Considerable amount of heterogeneity was reported during synonymous codon usage analysis in *Azotobacter vinelandii* (Sur *et al*. 2008). However, little is known about the conserved features of the codon usage patterns in a typical rhizobial genus. The first report of the codon usage pattern in the genus *Rhizobium* was given by Wang *et. al* (2012). They studied the codon usage patterns of 6 completely sequenced *Rhizobium* strains and revealed that the selection pressure, gene expression level and the ENC (effective number of codons) played a role in the codon usage in *Rhizobium*. The usage of Cys (cysteine) codons was found to be a conserved feature of the genus that shaped the codon usage patterns in *Rhizobium* genomes. The comparison of codon usage between highly and lowly expressed genes showed that 20 unique preferred codons were shared among *Rhizobium* genomes, revealing another conserved feature of the genus.

Thus, the whole genome sequence of the strains could therefore provide important insight into the mechanisms required by the effective microsymbionts to adapt to a particular edaphic environment and will definitely benefit the on-going comparative and functional analyses of the plant microbe interactions required for the successful establishment of agricultural crops.