Abstract

Rhizobium, a type genus of the family Rhizobiaceae is considered the best known beneficial plant-associated bacteria because of its importance in the nitrogen fixation occurring during its symbiosis with the legumes. 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 present study covered large area of North Bengal and Sikkim to characterize the French bean associated *Rhizobium* microsymbiont.

The plant infectivity test performed by the Koch's postulation and the Scanning Electron Microscopy confirmed the strains collected from different regions of North Bengal and Sikkim to be *Rhizobium*. The several biochemical studies involving biochemistry of one and half dozen of native strains of French bean microsymbiont *Rhizobium* found in the North Bengal and Sikkim region revealed considerable chemical and behavioral diversity. Based on the carbohydrate utilization, the strains clustered more or less as per their geographical location apart from a few exceptions. The physiological characterization revealed that the *Rhizobium* strains showed some amount of variations in response to the different stresses. All the strains grew in 1% NaCl but variations in growth occurred as concentration of the salt increased. However, growth of the strains decreased at 4% NaCl, although some of the strains showed some amount of growth even at such high concentration. Maximum growth of all the strains were obtained at 30ºC while the growth decreased with the increase and decrease in temperature. However, some of the strains showed a good amount of growth at 10ºC, 20ºC and 40ºC as well. The growth of the studied strains were found to be the maximum in pH 7 compared to 4 and 9. A statistically significant result based on the growth of strains at different concentration of heavy metals were obtained by 't- test' at a *p* value of less than or equal to 0.001. In this experiment cobalt (Co) was found to be

the most toxic of the heavy metal and lead (Pb), the least potent inhibitor to the *Rhizobium* growth. Each of the strains showed a wide variation in their result in response to different concentration of 27 antibiotics. The heatmap generated using gplot package of R-statistical software revealed three clusters depending upon the sensitivity against the antibiotics. *Rhizobium* strain MIR-6 was highly resistant against Norfloxacin (Nx) and Ciprofloxacin (Cf) while the strain CBR was resistant against Cefadroxil (Cq). Also, the clustering of *Rhizobium* strains based on IAR profiles did not reflect their geographical origin.

The positive *in vitro* characteristics like the solubilization of phosphates, siderophore production, IAA production, HCN production and the antagonistic activity against the phytopathogens shown by the *Rhizobium* strains DJG, RSm-3, CBR, SKM-N, RSv-1, SAM-12, BIJ, KPG-5N, RMa-13 and RGJ including the reference strain MTCC 99 labeled it as a potent strain for the *in vivo* green house study. An increase in different growth parameters of the plants like plant height, shoot fresh weight, shoot dry weight, root fresh weight and root dry weight were found compared to the uninoculated control. However, out of the eleven strains, DJG strain of Darjeeling followed by RSm-3 and KPG-5N were found to be the most promising strain for the future use as a biofertilizer as the growth parameters of the plant were found to be the maximum compared to the other strains as well as the uninoculated control.

The molecular study of *Rhizobium* was incorporated in this present study as it is considered to be the most discriminating method for assessing the variability among different strains of the bacteria. A wide range of genetic diversity was observed within the *Rhizobium* strains while performing the DNA fingerprinting techniques like RAPD, rep-PCR and 16SrDNA PCR-RFLP. A total of 133 major bands were scored in RAPD with 97.74% polymorphism. The amplification reactions with rep-PCR (REP, ERIC and BOX) primers also generated sufficient number of distinct polymorphic bands with 96.55% polymorphism for reliable strain discrimination. Each rep-PCR pattern generated a unique fingerprint and no similarity was observed in clustering of the isolates nor were the clusters found to form according to their geographical

location in any of the rep-PCR fingerprinting demonstrating the heterogeneity among the strains. However, both the markers have been found to provide a cheap, rapid and effective means to evaluate the genetic diversity among the *Rhizobium* strains under study. Hence, since both the two DNA markers are useful in their own way, both the DNA fingerprinting studies can be conducted to analyze the genetic diversity among the strains.

Attempt was also made to analyze a portion of 16SrDNA genome through PCR-RFLP. The PCR-RFLP of the conserved region of 16SrDNA performed with seven restriction enzymes *Alu* I, *Hae* III, *Hinf* I, *Hpa* II, *Mbo*1, *Taq* I and *Eco*RI revealed a polymorphism of 64.28 %. However, the value was found to be less compared to other molecular markers. Thus the PCR-RFLP of partial 16SrDNA region of the genome was not considered to be a good option for the study of *Rhizobium* diversity.

Attempt was also made to explore the partial 16SrDNA PCR products, by sequencing the conserved 16SrRNA gene of eleven *Rhizobium* strains. The size of the sequences were found to be around 500bp. The sequences were subjected to alignment with global partial 16SrDNA database from GenBank through BLAST program. The studied eleven partial *Rhizobium* 16SrDNA sequences shared a maximum similarity of 99-100% with the gene sequences of *R. etli*, *R. phaseoli*, *R. leguminosarum*, *R. tropici* and *R. paranaense* present in the database. The dendrogram constructed with RAxML software revealed the eleven *Rhizobium* strains to club with the standard *Rhizobium* strains in the GenBank.

The whole genome sequence of the newly sequenced strain RSm-3 was performed in the sequencing center of the University of NewHampshire, in collaboration with Dr. Louis Tisa of Hubbard Genome Center. The BLASTN program revealed that the complete genome sequence of the strain RSm-3 matched with *Rhizobium leguminosarum* bv *trifolii* WSM2304 with 96% identity and E-value 0.0. The whole genome sequencing of the strain RSm-3 using the Illumina HiSeq2000 platform generated 1585078 reads with a total of 58 contigs. The total size of the genome was estimated to be 7.3Mb with 60.6% GC content. The IMG annotation with the Genome ID 2684623037 revealed

the total protein coding genes to be 6564 with 53 count of RNA and 46 tRNA.

The codon usage study of the genes involved in the pathways of nitrogen fixation, siderophore formation and Indole Acetic Acid (IAA) formation of the newly sequenced strain RSm-3 along with five strains of *Rhizobium* from IMG database viz. *R. leguminosarum* bv. *viciae* 3841, *R. leguminosarum* bv. *trifolii* WSM1325, *R. etli* bv. *phaseoli* IE4803 and *R. etli* bv. *mimosae* Mim1 showed a similar synonymous codon usage pattern with more than 60% GC content. In all the cases, the effective number of codon values ranged from 21± 2 to 61 ± 0 for all five genomes suggesting that these high GC-rich genomes exhibited considerable heterogeneity in codon usage. All the studied genes (NFGs, SBGs and IPGs) clustered more or less with the ribosomal protein genes depicting their moderately high expression level. Thus, it was found from the expression pattern analysis and the RSCU data analysis, that the GC compositional constrains along with the translational selection have an influence on the codon usage as well as expression pattern on these genes in *Rhizobium*. Moreover, highly expressed genes did not posses any leading strand bias indicating the absence of replicational and transcriptional pressure. The core genes estimated to be around 3871 were found to be the house keeping genes involved in some major metabolic and signal transduction pathways.