INVESTIGATIONS ON MOLECULAR DIVERSITY IN ALDER COMPATIBLE FRANKIAE

BY

RAJANI VARGHESE

THESIS SUBMITTED
IN FULFILMENT OF THE DEGREE OF
DOCTOR OF PHILOSOPHY IN BOTANY

NORTH-EASTERN HILL UNIVERSITY
SHILLONG-793 022, INDIA
2000

DS
589.92048732
VAR;1

To my Father

THE NORTH-EASTERN HILL UNIVERSITY

FEBRUARY, 2000

DECLARATION

I, Rajani Varghese, hereby declare that the subject matter of this thesis entitled "Investigations on molecular diversity in Alder compatible frankiae" is the record of work done by me, that the contents of this thesis did not form basis of the award of any previous degree to me or to the best of my knowledge to any body else, and that the thesis has not been submitted by me for any research degree in any other University/Institute.

This is being submitted to the North-Eastern Hill University for the award of the degree of Doctor of Philosophy in Botany.

1/03/2000

(Rajani Varghese)

1/31/2000

(Head of The Department)

Head
Department of Botany
School of Life Sciences
N. E. H. U., Shillong-14

(Supervisor)

ACKNOWLEDGMENTS

Foremost, I would like to express deep sense of gratitude to Dr. A. K. Misra, Reader, Department of Botany, N.E.H.U, Shillong for the valuable guidance and the support given to carry out this work.

I sincerely thank Prof. H. N. Pandey (Head), Prof. P. Tandon (former Head) and Prof. R. R. Mishra (former Head) for their kind permission to use laboratory facilities needed to carry out the research work.

I am deeply indebted to Prof. D. Ammerman and Dr. G. Steinbruck, University of Tuebingen, Germany for providing the laboratory facility to carry out a part of my work.

I extend my sincere thanks to the faculty members and the non-teaching staff of the Botany department for their help and cooperation.

I thank Dr G Ganesh and Dr. Arnab Sen for their help and useful comments. I also thank sincerely my labmates Dr Gopa Sarma, Susamma Verghese and Vineeta Chauhan for their continuos help and cooperation.

I most appreciate Dr. Barik, Dr. Kusum Arunachalam, and Sumana Chowdhary Darrysuk N. Dkhar. Mary Hmar, Asenla, Hiranjit, Sanghmitra, Sapna and especially Babu John for the help given at various stages of research in various ways

I thank also C. Marak (lab. Assistant) for not only helping me in collection of nodule samples but also for other small helps given during the study.

Special thanks are to my hostel friends Narola Jamir, Ainla Jamir and Narola Temsu for their caring attitude, encouragement and help. They were indeed my friends in need.

I am forever indebted to my parents, sister and other family members for their constant encouragement and support .

My deepest gratitude goes to my husband Dr. N.K. Jacob for his care, invaluable help, constant encouragement and support throughout the work period. My greatest debts are also to my daughters whose cooperation was a great help as well as inspiration to continue the present study.

Last but not the least, I express my sincere thanks to each and every one who rendered every possible help during my research work.

Rajani 1/03/2000

(Rajani Varghese)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

CHAPTER 1

INTRODUCTION

For several centuries it has been a common knowledge that the soils from which nitrogen has been removed by the cereal crops could be revitalized by growing leguminous plants. Yet it was during the beginning of the nineteenth century that adequate explanation started coming in about certain microorganisms that have the ability to fix atmospheric di-nitrogen to a more utilisable reduced form (NH_4^+). The ability to assimilate elementary nitrogen biologically by a comparatively few species of microorganisms is of great biological interest, specially since it is an important input to natural as well as agricultural eco-systems.

Biological nitrogen fixation in nature is brought about by two types of microorganisms - non-symbiotic or free living and symbiotic. Non-symbiotic microorganisms (e.g. certain cyanobacteria like *Rivularia*, *Scytonema* or free living bacteria like *Klebsiella*), fix nitrogen independent of a host, while, symbiotic microorganisms like *Rhizobium* and *Frankia* make use of a host plant for fixing atmospheric di-nitrogen (*Frankia* can fix nitrogen *ex planta* also). While *Rhizobium* symbiosis fixes 24-584 kg N $ha^{-1}y^{-1}$ of atmospheric nitrogen, the quantum of nitrogen fixation by actinorhizal plants is of the order of 2-362 kg N ha^{-1} y^{-1} (Shantharam and Mattoo, 1997).

The microsymbiont within the root nodules of actinorhizal plants was first described in the second half of the nineteenth century. Microorganisms that are able to induce nitrogen fixing root nodules on non-leguminous actinorhizal plants have been classified in the genus *Frankia* (Tjepkema and Torrey, 1979). The first successful isolation of *Frankia* CpI1 from *Comptonia peregrina* was done by

Callaham *et al.* (1978). Recently Normand *et al.* (1996) have classified *Frankia* as the only genus of the family *Frankiaceae*.

Members of the genus *Frankia* can be distinguished from other actinomycetes on the basis of their characteristic morphology, ability to nodulate plants and to fix nitrogen (although some strains of *Frankia* do not have the ability to nodulate or fix nitrogen). It is very typical in having high G+C content (about 68-72%; An *et al.*, 1983). Being a pleomorphic actinomycete, both *in vivo* and *in vitro* (Callaham *et al.*, 1978), *Frankia* exhibits three morphological structures. These are hyphae which are septate, sporangia containing non-motile spores (Lechevalier and Lechevalier, 1979) and pedunculate thick walled specialized structures called as 'vesicles' (Lalonde and Calvert, 1979). The enzyme nitrogenase is located in these symbiotic vesicles. Therefore, they are considered as the sites for nitrogenase activity *in vivo* and *in vitro* (Newcomb and Wood, 1987). However, *Casuarina* infective *Frankia* do not form vesicles inside the nodules. The nitrogenase is protected by the thickening of the host cell walls.

Frankia grows primarily by extension and branching of filamentous septate hyphae (~0.5-1.5 μm in diameter). Wider diameter hyphae differentiate into sporangia from which the spherical shaped asexual spores are released. Frankiae are able to use few carbon sources for energy and growth (Benson and Hanna, 1983; Hafeez *et al.*, 1984; Lechevalier *et al.*, 1983). Although, the exact nutritional requirement of *Frankia* is not fully understood, most isolates are able to use propionate, acetate, succinate, fumarate, malate and tween 80. They do not use hexoses, pentoses and disaccharide. Nitrogen utilization is restricted to ammonium, nitrite and aspartate (Shipton and Burggraf, 1982). The optimum temperature for growth is strain dependent ranging from 25°C-36°C (Tisa *et al.*, 1983).

Frankia has a varied host spectrum being in symbiosis with 25 genera of dicotyledonous plants belonging to 8 families (Baker and Schwintzer, 1990). Among actinorhizal plants, the two genera *Alnus* and *Casuarina* exhibit highest nitrogen fixing potential (Dommergues, 1996). The genus *Alnus* is one of the most extensively studied genus among actinorhizal plants. About 47 species of *Alnus* are known (Swensen and Mullin, 1997). In India, only two species of *Alnus* (*A. nepalensis* and *A. nitida*) are found distributed naturally throughout the temperate Himalaya. They are confined to the higher elevations of Meghalaya, Arunachal Pradesh and Himachal Pradesh. Although not naturally occurring, trees of *A. nepalensis* (fig 1.1) are also found in some locations of Nagaland and Tamilnadu. *A. nitida* trees (fig 1.2) are confined to higher altitudes of Himachal Pradesh (1200-1500 m above mean sea level), some time descending in to the plains along the river banks. They are morphologically quite distinct from *A. nepalensis* being larger in size, having lighter tree bark and bearing larger female cones. *Alnus glutinosa*, which was selected as an out species, in comparison is more or less like *A. nepalensis* or *A. nitida* in morphology except that its leaf blades are irregularly dentate (fig 1.3).

Alnus, together with its microsymbiont *Frankia*, is widely recognized as a good nitrogen fixer (Simonet *et al., 1991*; Guofan and Tingxiu, 1987; Domenach *et al.*, 1989). It is thought to be responsible for a high level of soil nitrogen accretion world wide (Tarrant and Trappe, 1971). In addition to this, *Alnus* plays an important role in forestry. It is cultivated in general for fuel woods or as timber or to increase soil nitrogen stability. Also, the release of compounds such as phenols, fatty acids, amino acids into soil indirectly affect the free living nitrogen fixing organisms stimulating their growth. In India, these trees are maintained in Himalayan uplands, used to colonize wasteland and to reclaim land to improve wild life habitat. Their economic value could be further explored and exploited by

Figure 1.1

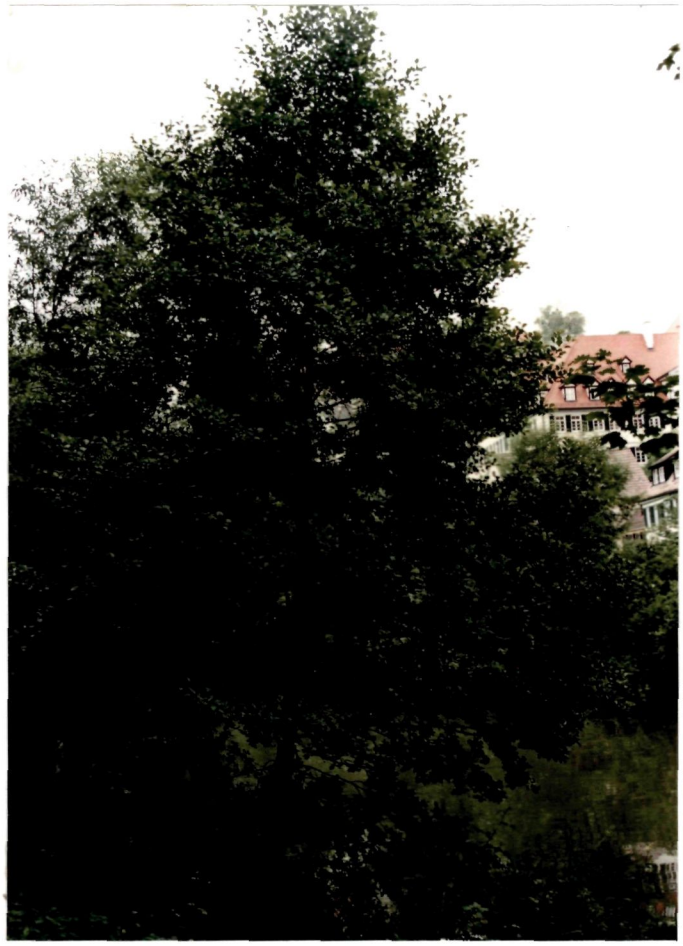
Fig 1.1 : *Alnus nepalensis* trees as found growing in the Upper Shillong region of Meghalaya.

Fig 1.2 : *Alnus nitida* tree growing in Kulu, Himachal Pradesh

Figure 1.2

Figure 1.3

Fig 1.3 : *Alnus glutinosa* tree growing near the Neckar river bank of Tuebingen, Germany.



improving their nitrogen fixing efficiency. The nitrogen fixing potential can be significantly improved further genetically by selecting superior strains of *Frankia* and the corresponding host genotype for an enhanced symbiotic performance. In order to genetically improve any genus, initially its genetic diversity has to be screened. Following which selection of most competitive and infective strains could be brought about. This is necessary since, the ability to survive in soil even in absence of a host allows greater flexibility and diversity among frankiae strains (Swensen and Mullin., 1997).

Alnus is thought to have originated in Indo-China region (Furlow, 1979). Therefore, it is expected that its microsymbiont too would have evolved with it. If it is true, there should be higher diversity for *Frankia* in this region. This assumption is supported by the study conducted by Ganesh *et al.* (1994) who found that the isolate AnpUS4 generated from *Alnus nepalensis* was genetically different from the reference strains studied. But no detailed study is available yet on the diversity in alder compatible *Frankia* from different geographical locations in India. Therefore, in the present study conserved as well as variable regions of two very important structural and functional gene units (ribosomal and nitrogen fixing genes) from *Alnus* nodulating microsymbiont, collected from various geographical locations in India, were investigated.

The ribosomal DNA is well characterised ubiquitous molecular yardstick for estimating evolutionary relationship of organisms (Ochman and Wilson, 1987; Woese, 1987). In bacteria, the usual order of structure of *rrn* operon is 16S-spacer-23S-spacer-5S with few exceptions (fig 1.4). The ribosomal RNA genes are evolutionarily homologous and functionally equivalent in all organisms. Their sequence changes are slow allowing the estimation of evolutionary relationships between even distantly related organisms. The internal transcribed spacers (ITS) in prokaryotic rRNA genes are situated between 16S and 23S (fig 1.4) and between

23S and 5S genes. These are varied in length and are assumed to have been subjected to less selection pressure. Therefore, they should have accumulated more random mutations than the coding regions. Thus, it forms an ideal region for the discriminative studies and can be PCR amplified using specific primers corresponding to the flanking conserved sites.

One of the most important functions in *Frankia* is nitrogen fixation. The rate at which nitrogen is fixed by *Frankia* varies considerably depending upon plant species it is associated with (Torrey, 1978). The presence of nitrogen fixing genes in *Frankia* makes it economically important, thereby generating interest in the genetic study of the symbiosis. The nitrogenase enzyme involved in the fixation of the molecular nitrogen is a useful tool for characterising nitrogen fixing microorganisms. Nitrogenase is an enzyme complex consisting of two components - the molybdenum-iron protein called dinitrogenase, which is an $\alpha_2\beta_2$ tetramer encoded by *nif* D and *nif* K respectively and the iron containing protein called dinitrogenase reductase, which is a homodimer encoded by *nif* H (Hirsch, 1995). In addition to the structural genes (*nif* H, *nif* D and *nif* K; fig 1.5), biological nitrogen fixation requires a host of additional genes. In *Frankia*, a 4.5 kb fragment containing eight *nif* genes has been characterised (Harriott *et al* ., 1995) and till date more than twenty additional *nif* or *nif* associated genes have been identified in other diazotrophic bacteria (Oh *et al.*, 1997). In all cases studied it is found clustered in bacterial genome or plasmids (Merrick, 1993). Also, as in almost all nitrogen fixing organisms, *nif* H-D-K genes seem to be contiguous in *Frankia* (Haselkorn, 1986; Oh *et al.*, 1997). Large sequence homology has been established between *nif* H-D-K genes from *Klebsiella pneumoniae* (Cannon *et al.*, 1979) and those of other nitrogen fixing species (Ruvkun and Ausubel, 1980) including *Frankia.*

The nitrogenase structural genes (*nif* H, *nif* D and *nif* K) are known to carry both conserved (Ruvkun *et al.*, 1980) and variable regions present in low and

Figure 1.4

Fig 1.4 : Diagrammatic representation of organization of rrn operon in prokaryotes. The underlined segment marks the analysed region of the genes.

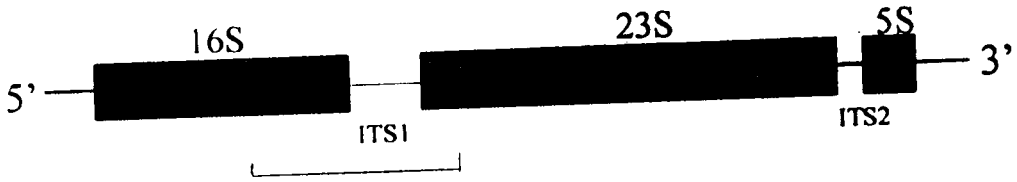
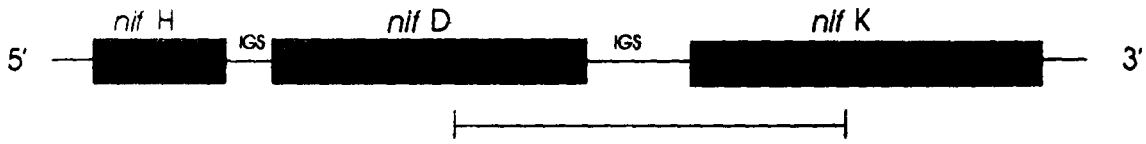

16S
23S
5S
5'
3'
ITS1
ITS2

Figure 1.5

Fig 1.5 : Schematic representation of the organization nitrogenase structural genes (*nif* H, *nif* D, *nif* K) genes. The line marked below represents the analysed region.


nif H
IGS
nif D
IGS
nif K
5'
3'

high sequence variability (Simonet *et al.*, 1990; Jamann *et al.*, 1993). Genes coding for nitrogenase and 16S rRNA are considered to have has been characterised (Harriott *et al* ., 1995) and till date more than twenty additional *nif* or *nif* associated genes have been identified in other diazotrophic bacteria (Oh *et al.*, 1997). In all cases studied it is found clustered in bacterial genome or plasmids (Merrick, 1993). Also, as in almost all nitrogen fixing organisms, *nif* H-D-K genes seem to be contiguous in *Frankia* (Haselkorn, 1986; Oh *et al.*, 1997). Large sequence homology has been established between *nif* H-D-K genes from *Klebsiella pneumoniae* (Cannon *et al.*, 1979) and those of other nitrogen fixing species (Ruvkun and Ausubel, 1980) including *Frankia.*

The nitrogenase structural genes (*nif* H, *nif* D and *nif* K) are known to carry both conserved (Ruvkun *et al.*, 1980) and variable regions present in low and high sequence variability (Simonet *et al.*, 1990; Jamann *et al.*, 1993). Genes coding for nitrogenase and 16S rRNA are considered to have evolved in parallel (Simonet *et al.*, 1991). Therefore, evolutionary studies on *Frankia* genome concentrate mostly on these two regions.

1.1 Objectives

Alnus is a largely studied genus from an ecological point of view. Very few investigations have been conducted at molecular level, especially in the Indian alder populations. Studies on *Frankia* in general, have shown existence of tremendous diversity (An *et al.*, 1985a; Mirza *et al.*, 1994; Navarro *et al.*, 1992; Rouvier *et al.*, 1996; Ritchie *et al.*, 1999). A superior and efficient host-microbe relationship could certainly affect positively the nitrogen fixing capacity of the *Frankia* strain. In order to accomplish this, selection of most infective, effective and competitive *Frankia* strain is necessary. For this, the diversity existing within the species has to be investigated which will indirectly enhance the prospect of genetically improving the

genus. Therefore, in this study, the existence of the diversity within *Frankia* nodulating *Alnus* species in India was investigated by molecular characterization of the microsymbiont directly inside the nodule. Since hidden diversity within the genus is important for the evolution and speciation of *Frankia* (Cournoyer *et al.*, 1993), the study conducted was also expected to put some light on the relationship existing among and between the species. The following approaches were tried for the purpose-

i). Alder compatible *Frankia* germplasm from different parts of India was collected. For comparative study, *Alnus glutinosa* nodules from Tuebingen, Germany, were also collected as an outgroup.

ii). DNA from field collected nodules and the reference *Frankia* strain $ACN1^{AG}$ were isolated.

iii). Polymerase Chain Reaction based amplifications of partial 16S rRNA gene, 16S-23S rDNA internal transcribed spacer (ITS) region, partial *nif* D and *nif* K genes and *nif* D-K intergenic spacer (IGS) regions, were done.

iv). Amplified Fragment Length Polymorphism was studied.

v). PCR/Restriction Fragment Length Polymorphism of 16S-23S rDNA ITS and *nif* D-K IGS regions were investigated.

vi). Partial 16S rRNA gene, 16S-23S rRNA ITS, partial *nif* D and *nif* K genes and *nif* D-K IGS regions were cloned and sequenced.

vii). Computer analyses of these sequences were done using the *Casuarina* compatible strain ORS020606 sequence retrieved from the GenBank as reference.

viii). Amino acid sequence of the partial *nif* D and *nif* K genes were developed and analysed using computers.

ix). Phylogenetic relationships among the sequences under study and those retrieved from the GenBank were worked out.

CHAPTER 2

REVIEW OF LITERATURE

CHAPTER 2

REVIEW OF LITERATURE

Biological nitrogen fixation is performed by phylogenetically diverse prokaryotic organisms. The actinomycetes, *Frankia* fixes considerable amount of nitrogen in symbioses with non-leguminous angiospermic plants. It has potential to be placed next to *Rhizobium* (Torrey and Racette, 1989). But, unlike *Rhizobium* which specifically infects the leguminous plants, *Frankia* is more versatile infecting phylogenetically diverse angiosperms (actinorhizal plants) belonging to eight families and 24 genera having least evolutionary relationship with each other. This allows a more flexible host-microbe relationship which suggests a more generalized nodulation process than the highly specific one which is observed in *Rhizobium*-legume symbiosis.

2.1 The microsymbiont *Frankia*

The first *Frankia* strain CPI1 was isolated from *Comptonia peregrina*, (Callaham *et al.*, 1978). Since then, hundreds of strains have been isolated from various actinorhizal species (Lechevalier, 1984, 1994). All of these filamentous bacteria have been assigned tentatively to the genus *Frankia* of the family *Frankiaceae* (Baker *et al.*, 1981; Lechevalier, 1994). Taxonomically, the members of the genus *Frankia* are distinguished from other actinomycetes on the basis of their host specificity, morphology (hyphae, vesicle and sporangia), biochemistry (type III cell wall and type I phospholipids) and physiology (Lechevalier, 1984; 1994).

Frankia is a very slow growing organism and its isolation in pure culture is very tedious. Therefore, studies conducted on this organism are very limited. Due to the problems encountered in isolation and cultivation of *Frankia*, the progress in

this field has been slow. With the availability of techniques for isolation and amplification of DNA directly from the nodules Misra *et al.* (1991) suggested characterization of microsymbiont without their isolation in pure culture. However, a single nodule is known to host more than one *Frankia* strain (Gautheir *et al.*, 1981; Diem *et al.*, 1983; Zhang *et al.*, 1984)) and therefore, all inferences drawn through this approach must be seen in this light. To overcome this problem, several attempts have been made to obtain genetically pure cultures of *Frankia* , resulting in generation of single spore cultures through plating (Prin *et al.*, 1991; Lumini and Bosco. 1996; Sarma *et al.*, 1997), and through entrapment of spores in calcium alginate beads (Borthakur *et al.*, 1996; Sarma *et al.*, 1997).

2.2 Actinorhizal Symbioses and host specific responses

The plants that are nodulated by *Frankia* are known as actinorhizal plants. Together, they participate in a mutually beneficial symbiotic relationship. So far, more than 200 dicotyledonous plants belonging to 8 families and 25 genera have been identified. Among these, all except *Datisca* are woody shrubs or trees. Actinorhizal plants are known to grow in marginal soils and representatives of these groups of plants can be found in most climatic zones.

Studies carried out in general show that host factors play a key role in the initiation of *Frankia*- actinorhizal symbiosis. Many *Frankia* strains studied are known to be specific towards one host. There are following four presently agreed upon host infectivity groups of *Frankia* strains (Baker, 1987) :

i). strains infective on *Alnus* species and *Myrica*,

ii). strains infective on *Casuarina* species and *Myrica* species,

iii). strains infective on Elaeagnaceae and *Myrica* and,

iv). strains strictly infective on *Elaeagnaceae*

Combination of phenotypic as well as genotypic studies conducted by Bosco *et al.* (1992) showed that strains from *Elaeagnus* species are able to cross the inoculation barriers and infect *Alnus* species. This study conducted showed the existence of broad host range in *Frankia* species.

2.3 Molecular characterization of *Frankia*

The cultured *Frankia* have relatively uniform morphology that, although actinomycetous in character, differs significantly from that of the other actinomycetes. Earlier, the uniform morphology and DNA base composition of *Frankia* gave little indication of the diversity that existed within the genus (An *et al.*, 1983). There are few morphological differences among isolates other than pigmentation and hyphal size. The high G+C content is a characteristic feature of actinomycetes. However, its relative uniformity among *Frankia* strains makes it ineffective in differentiating strains (Benson and Silvester, 1993). Therefore, for evaluating *Frankia* strain diversity earlier, methods like efficiency testing (Normand and Lalonde, 1982), serological analysis (Baker *et al.*, 1981), analysis of total protein patterns (Benson and Hanna., 1983; Benson *et al.*, 1984) and SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis; Benson and Hanna. 1983; Gardes and Lalonde, 1987) were employed. It is only when more and more strains became available for study that they were subjected to nodulation tests (Baker, 1987), isozyme analysis (Gardes *et al.*, 1987) carbon utilization studies (Shipton and Burggraf, 1982; Bloom *et al.*, 1989), restriction fragment length polymorphism (RFLP) analysis (An *et al.*, 1985b; Dobritsa, 1985; Normand *et al.*, 1988), DNA hybridisation studies (An *et al.*, 1985a; Bloom *et al.*, 1989; Fernandez *et al.*, 1989) and 16S rDNA analysis (Hahn *et al.*, 1989; Harry *et al.*, 1991; Mirza *et al.*, 1992; Nazaret *et al.*, 1991). These studies revealed tremendous diversity among *Frankia* strains. This heterogeneity among natural strains is for several

characters including some characters that are related to symbiosis (Normand and Lalonde, 1986). Such characters can be exploited for developing improved strains.

With the development of the new technique of PCR (polymerase chain reaction, Mullis *et al.*, 1986; Mullis and Faloona, 1987), amplified DNA fragments could be used as substrate for restriction pattern analysis or sequencing. Additional information inherent in the polymorphic character of the amplified product is accessible by several means. As studies on *Frankia* advanced, possibility of amplifying and sequencing of DNA directly from nodules without the need to have pure cultures, opened up possibility of DNA analysis of even those strains that have not been in culture so far (Misra *et al.*, 1991). The PCR amplified product can be digested with appropriate restriction enzyme, and the resulting fragment can be resolved elctrophoretically. The presence of a unique restriction site resulting into a unique restriction pattern can be indication of a particular species. The PCR/RFLP technique has been used to discriminate closely related *Frankia* strains (Navarro *et al.*, 1992; Jamann *et al.*, 1993; Rouvier *et al.*, 1996). The intergenic spacer regions are considered the most variable regions. However, for RFLP studies they have to be sufficiently large.

2.4 Ribosomal RNA as a taxonomic tool

The use of ribosomal RNA as a taxonomic tool has been well demonstrated in bacteria, where phylogenetic relationships, established on the basis of cellular metabolism studies, have been completely redefined by 16S rRNA sequence analyses (Woese, 1987; Woese and Fox. 1977).

16S rDNA sequences have been largely used to determine relationships among organisms. In addition to the highly conserved areas that have been used to study the relationships among distant taxa, the 16S rDNA sequence contains more variable regions that have been useful in the differentiation of genera and species

(Jensen *et al.*, 1993). Hahn *et al.* (1989) used the variable region present in the 16S rRNA gene to distinguish between the *Frankia* strains Ag45/Mut15 (*nif*$^+$) and AgB1.9 and AgW1.1 (*nif*$^-$). Similarly, 16S rDNA is also used for studies involving identification (Benson *et al.*, 1984; Gardes *et al.*, 1987; Harry et al., 1991: Jensen *et al.*, 1993), characterization (Simonet *et al.*, 1991; Nick *et al.*, 1992; Normand *et al.*, 1996) and differentiation (Bosco *et al.*, 1992; Mirza *et al.*, 1994; Murry *et al.*, 1997; Clawson *et al.*, 1998) of *Frankia* nodulating different actinorhizal host species. However, for such studies, availability of specific probes is considered crucial which increases the specificity and reliability of the study conducted (Nazaret *et al.*, 1991).

2.5 Use of *nif* genes for characterization studies

The nitrogenase (*nif*) region is quiet conserved among bacteria having a narrow distribution in the actinomycetes, being restricted to *Frankia*. The enzyme nitrogenase, involved in the fixation of molecular nitrogen can also be used as an additional tool for characterization studies (Jamann *et al.*, 1993). In *Klebsiella*, about 20 *nif* or *nif* associated genes are known. In *Frankia* however, only eight *nif* genes along with the three ORF's are known as characterized in a 4.5 kb fragment (Harriot *et al.*, 1995) This included, *nif* X, orf 3, orf 1, *nif* W, *nif* Z, *nif* B, orf 2, *nif* V, *nif* H, *nif* D and *nif* K. The nitrogenase structural genes (*nif* H, *nif* D, *nif* K) are crucial for nitrogen fixation and the organization of the *nif* structural genes in *Frankia* is (*nif* H-D-K) as found in most nitrogen fixing species. Like rDNA, both low and high sequence variability regions are present which can also be used in conjunction with rDNA for species detection (Simonet *et al.*, 1990) characterization and also to deduce relationship between organisms.

2.6 Use of 16S-23S ITS and *nif* H-D-K IGS regions for characterization studies

The most direct and rapid method to characterize organisms is to visualize the polymorphic character of internal *rrn* spacer regions. The organization of ribosomal RNA encoding (*rrn*) genes has been studied in *Frankia* strain ORS020606. It is arranged in the usual order of 16S-23S-5S as observed in bacteria (Normand *et al.*, 1992b). The study conducted by Normand *et al.* (1992b) showed presence of two *rrn* clusters in the case of *Casuarina* strain ORS020606 which though couldn't be differentiated through RFLP studies, showed a lot of variation in the non-coding region. It is this region which forms a possible target for characterization studies. Studies conducted in general shows variation in high proportion in the first intergenic spacer region between 16S-23S rRNA. Similar to the internally transcribed spacer (ITS), the intergenic spacer (IGS) regions are also assumed to accumulate more mutations than the gene flanking them. Therefore, it could be used as a parameter either alone or in combination with each other for discriminating closely related organisms at interspecific as well as intraspecific level. This could be done by carrying out PCR amplification of the spacer regions by using primers from highly conserved flanking sequences. The length and sequence polymorphism present in the PCR product can then be used in the recognition of genera and species. Rouvier *et al.* (1996) used the ITS region between 16S-23S rRNA and the IGS region between *nif* D-K to assess diversity of *Casuarina* and *Allocasuarina* microsymbiont at strain level. For this, she directly amplified the DNA isolated from nodules followed by characterizing them through restriction analysis. Five different groups could be recognized based on PCR/RFLP patterns. Similarly, genetic diversity of the *Ceanothus*-infective *Frankia* strains was investigated by PCR/RFLP of 16S-23S rRNA ITS region (Ritchie *et al.*, 1999).

The intergenic spacer region between *nif* H-D and *nif* D-K has also been used widely for characterizing *Frankia* strains. However, due to its larger size *nif* D-K IGS region is more preferred for characterization studies since, it is assumed to have accumulated more mutations. Information generated through PCR/RFLP analysis of *nif* D-K IGS region was used by Jamann *et al.* (1993) to characterize *Frankia* strains.

In addition to this, characterization of strains of microbes such as *E. coli* and *Nitrobacter* spp. has also been done using PCR/RFLP of ITS region (Garc'a-Mart'nez *et al.*,1996; Navarro *et al.*, 1992).

Although, the use of intergenic spacer regions for discriminating studies seems to be very fruitful, only limited amount of work has been conducted so far using this particular region. Because, the necessity of sequence information needed for synthesis of primer flanking the IGS region becomes always a limiting factor for such studies.

2.7 Phylogenetic studies based on nucleic acid comparison

Both rRNA and *nif* genes are used separately or in conjunction with each other to deduce relationship between organisms. The phylogenetic relationships have been worked out based on nucleic acid comparisons and the trees drawn generally through the phylogenetic analysis reflects the host infectivity groups for *Frankia* strains and nodule endosymbiont. For example, Ganesh *et al.* (1994) used partial 16S rRNA gene sequence for deducing the relationship of an Indian isolate with the world collections. Analysis of 16S rRNA together with *nif* H-D sequence for a number of *Frankia* isolates from the major host infectivity groups showed that the phylogenetic relationship between these isolates matches the mode of host plant infection process (Cournoyer *et al.*, 1993). Nazaret *et al.* (1991) found that 16S rDNA sequence variation among eight genomic species of *Frankia* separates them

into two major groups: one group infective on *Elaeagnus* and the second group infective on either *Alnus* or *Casuarina.* Although, Nick *et al.* (1992) also observed similar groupings but here in addition, the *Coriaria* endophytes joined the *Frankia alni* and the *Casuarina*-infective strains. Normand *et al.* (1996), based on comparative 16S rDNA sequence analysis, divided the genus *Frankia* into four main subdivisions which included:

i) a large group comprising *Frankia alni* and related organism to which *Casuarina*-infective strains, a *Myrica nagi* microsymbiont and other effective *Alnus*-infective strains are related;
ii) unisolated microsymbiont of *Dryas, Coriaria* and *Datisca* species;
iii) *Elaeagnus* infective strains and
iv) atypical strains, a group which includes the *Alnus* infective non-nitrogen fixing strains.

The hyper variable region (V1-V3) of ribosomal RNA was analysed by Normand and co-workers (personal communication) to determine the relationship of microsymbiont in *Dryas drummondii* nodules with other *Frankia* strains. They found that, although it was closely related to the microsymbiont of *Coriaria,* it was genetically distinct from other *Frankia* strains characterized by them. Similar observations were made by Bosco *et al.* (1996) where restriction analysis of PCR amplified *nif* D-K IGS region showed that *Dryas drummondii* was distantly related to the *Alnus-Casuarina* and the *Elaeagnus* infectivity groups.

Another gene which is used to deduce similar relationship is the chloroplast gene *rbcL.* More recently, analysis of this gene for a broad range of plants indicated that all symbiotic root nodulating higher plants belong to one single large group (Swensen and Mullin, 1997). Although, this study based on comparative *rbcL* sequence analysis has provided a new view of relationships among plants, only a

complete analysis of the frankiae endosymbiont including the estimated times of divergence among both host plants and *Frankia*, coupled with the geographical distribution of the host plants as they evolve, will give a clear picture of the evolution of this important symbiosis.

CHAPTER 3

MATERIALS AND METHODS

CHAPTER 3

MATERIALS AND METHODS

3.1 Collection of nodules

Nodules were collected from roots of *Alnus* trees growing in different geographic locations within India (table 3.1). From each location several trees were selected and nodule sampling was done in several replicates. As an out species, nodules were also collected from Tuebingen, Germany.

Nodules were collected from each randomly selected *Alnus* tree by tracing out the roots and rootlets to expose the nodule clusters. Usually, nodule clusters were found between 5-10 cm below the soil surface (sometimes even below 10 cm). It was observed that usually the nodule growth was maximal a month or two after monsoon. The detailed steps followed were -

1. The area around the selected tree was cleaned by removing weed plants growing over it. Hand spade used for digging was sterilised using ethyl alcohol prior to use each time.
2. The top soil was removed to expose roots and rootlets. The nodule bearing rootlets were traced to the tree of their origin prior to collection.
3. The nodules were collected in fresh plastic bags or into vials with 90% ethyl alcohol directly. In each case it was labeled with details like site of collection, date of collection, replicate number and tree species, etc.
4. Once the nodule samples were brought to the laboratory, they were thoroughly cleaned under running tap water to remove the adhering soil particles.
5. After separating the nodule lobes from each other, they were further washed with a mild detergent followed with several washes in distilled water.

Table 3.1: Collection sites of *Alnus* nodule samples from different geographical Locations.

Geographical location (Country/Province)	Site	Species
India/		
/Arunachal Pradesh	Hapoli	*Alnus nepalensis*
	Zero	*Alnus nepalensis*
/Meghalaya	Upper Shillong	*Alnus nepalensis*
	Nonkrem Hills	*Alnus nepalensis*
	Mawlai	*Alnus nepalensis*
	Cherrapunji	*Alnus nepalensis*
/Nagaland	Kohima	*Alnus nepalensis*
	Mokokchung	*Alnus nepalensis*
/Tamilnadu	Ooty	*Alnus nepalensis*
/Himachal Pradesh	Kulu	*Alnus nitida*
Germany	Tuebingen	*Alnus glutinosa*

6. Surface sterilisation of the nodules was done using 1-5 % H_2O_2 solution (5-10 minutes) followed by several washes with sterile distilled water.
7. The cleaned nodules were stored in 75% ethyl alcohol at –20°C until used.

3.2 Isolation of DNA

3.2.1 Isolation of DNA from nodules

DNA from nodules was isolated using the methods described by Simonet *et al.* (1990) or Rouvier *et al.* (1996) with minor modifications.

3.2.1.1 Isolation of DNA from nodules using the method of Simonet *et al.* (1990) :

1. Pre-washed and mildly sterilised nodules were re-sterilised using 1% H_2O_2 (v/v) for 10 minutes followed by thorough wash in several changes of sterile double distilled water.
2. The epidermis of each nodule was peeled off in PBS/PVP (Appendix- III. 2). Individual nodule lobes were taken in separate micro centrifuge tubes and crushed in 200 µl of extraction buffer (Appendix- II.1a).
3. Ten µl of 50 mg/ ml Achromopeptidase was added to it followed by 10 µl of 10 mg/ml Lysozyme. It was incubated at 37°C for 60 minutes.
4. Lysis was completed by adding 10 µl of 10% (w/v) Sodium dodecyl sulphate (SDS) to the cell suspension. It was further incubated at 80°C for 10 minutes.
5. Cell debris were removed by centrifugation at 7000 xg for 1 minute.
6. The supernatant was extracted with equal volume of phenol/ chloroform (1:1; v:v) at 13000 xg at 4°C for 30 minutes.
7. The DNA in the aqueous phase was precipitated using cold absolute ethanol. It was kept 1 hour at –20°C followed by centrifugation at 13000 xg at 4°C. Some times it was kept overnight at –20°C for better recovery of DNA. The pellet

formed was washed with 70% alcohol, briefly vacuum dried and was resolubilised in 10 µl of Tris buffer (10mM; pH-8.0).

3.2.1.2 Isolation of DNA from nodules using the method of Rouvier *et al.* (1996)

1. One lobe was taken from each nodule cluster and its epidermis was peeled off carefully in PBS/PVP as above.
2. Individual lobes were taken separately in 1.5 ml Eppendorf tubes and was crushed in 300 µl of extraction buffer (Appendix- II.1b) using sterile plastic pestle.
3. The homogenate was incubated at 65°C for 1 hour and centrifuged for 10 minutes at 7000 xg to remove plant debris including phenolics and polysaccharides trapped by PVP.
4. The supernatant was first extracted with an equal volume of phenol/chloroform (1:1; v:v) and centrifuged at 13000 xg at 4°C for 30 minutes, Then it was further extracted with chloroform/isoamylalcohol (24 :1, v:v) and centrifuged at 13000 rpm for 30 minutes, so that impurities including proteins, etc. were completely removed.
5. DNA from the aqueous phase was precipitated by addition of 2 volumes of cold ethanol, kept for overnight at –20°C, and centrifuged at 13000 xg for 30 minutes at 4°C.
6. The pelleted DNA was washed with cold ethyl alcohol (70%), vacuum dried and then resolubilised in 10 µl of Tris buffer (pH-8.0).

3.2.2 Isolation of genomic DNA from *Frankia* culture

The total genomic DNA was isolated from the *Frankia* culture ($ACN1^{AG}$) using the modified procedure described by Simonet *et al.* (1985).

1. The cells were washed with TE buffer (Appendix-II.2a) to remove traces of DPM (Appendix-I.1), the medium in which the cells were maintained.
2. The washed cells were incubated in 1 ml of TES (Appendix-II.2b) with 5 mg of Achromopeptidase for 60 minutes at 37 °C.
3. About 10 µl of 10% SDS was added and mixed by gently inverting the tube repeatedly. Lysis was allowed to continue for 15 minutes at 70°C.
4. To the lysate, equal volume of buffer equilibrated phenol was added and centrifuged at 7000 xg. The aqueous phase was collected and to this an equal volume of chloroform was added and centrifuged for half-an-hour at room temperature.
5. The aqueous phase was collected, two volumes of isopropyl alcohol were added and tube kept overnight at room temperature.
6. It was centrifuged at 13000 xg at room temperature for 30 minutes. The pellet was briefly vacuum dried and then suspended in 10µl Tris buffer (pH-8.0).

3.3 Quantification of Deoxyribonucleic acid

Quantification of DNA was done either by UV fluorescence on an ethidium bromide stained agarose gel or by spectrophotometry.

3.3.1 Quantification by direct observation

DNA was directly quantified by visual observation on ethidium bromide stained agarose gel on a transilluminator. The intensity of fluorescence is directly proportional to the total mass of DNA (Sambrook *et al.*, 1989). The quantity of the DNA sample was estimated visually by comparing the fluorescence yield by reference marker of known amount.

3.3.2 Quantification by spectrophotometric method

An optical density of one (O.D=1) at 260 nm is considered to correspond approximately to 50 µg/ ml of double stranded DNA and approximately 20 µg/ ml for single stranded oligonucleotides (Sambrook *et al.*, 1989). The purity of DNA preparations was checked using the spectrophotometric ratio A_{260}/A_{280} and by obtaining the optical density value between 1.8 and 2.0.

3.4 Gel electrophoresis

3.4.1 Agarose gel electrophoresis

Genomic DNA as well as PCR amplified DNA fragments were detected using agarose gel electrophoresis. Depending upon the size of the fragments to be separated, the concentration of the agarose gel was adjusted (0.8-1.5% w/ v). The samples were mixed with loading buffer (Sambrook *et al.*, 1989; Appendix-II.6a) and were electrophoresed along with molecular size markers (*Eco* RI / *Hind* III digested DNA or *Hinf* I digested pBR322 DNA). In cases where separation of small DNA fragments was required, 4% (w/ v) Nusieve agarose gels were used. Low melting point agarose gels were used whenever it was intended to elute the DNA by agarose digestion with the enzyme gelase. Electrophoresis was carried out in TAE buffer (occasionally in TBE buffer) (Appendix- II.4a; 4b) and run at 5–10 v/ cm. The gel was stained in ethidium bromide (0.5 µg/ ml) and visualised under UV- fluorescence (302 nm). The cutting of the DNA bands of interest for elution from preparative gels was done on a long wave length UV- transilluminator so as to avoid excessive breaks in DNA. The gels were photographed with Polaroid MP-4 land camera or by computer imaging using the software Bio-PRINT.

3.4.2 Polyacrylamide Gel Electrophoresis (PAGE)

Polyacrylamide gel electrophoresis was used for both detecting the restriction digested samples as well for sequencing. In order to detect the smaller

fragments generated through digestion of PCR products, the digested samples were run on polyacrylamide gel (8 %). It gave the size and the quantity accurately. Similar to agarose gel, depending upon the size of the fragment of interest to be separated the concentration of the gel was adjusted.(6-8%; appendix III.4). Electrophoresis was carried out in TBE buffer for 2 hours at 5V/cm and the gel was visualized by silver staining.

Sequencing was accomplished in 6 % denaturing polyacrylamide gel. Before casting the gel, the glass plates were cleaned thoroughly using liquid detergent and running tap water followed with double distilled water. The notched glass plates was treated with Gel Slick^{TM} followed with rinsing with double distilled water. After it is dried, the glass plates were further rinsed with ethanol and dried. denaturing polyacrylamide gel with a thickness of 0.4 mm was cast in the plates using spacers and combs on macromould gel casting unit (LKB).

The polymerised gels were fixed to the apparatus and pre-ran for 20 minutes at 20 mA current using 1X TBE buffer system. After inserting the comb, the wells were cleaned and 2 µl each of heat denatured samples were loaded and electrophoresed at 30 mA and 2500 V for 90 minutes (short run) or 4 hours (long run). The temperature of the gel was further maintained at 50°C by regulating the current and voltage. After electrophoresis gel were fixed for 20 minutes in 10 % acetic acid and were transferred to a supporting sheet of filter paper (Whatmann 3 mm) and vacuum dried for 40-45 minutes at 60°C. The dried gels were exposed to X-ray film (Cronex® 4) for overnight or longer and the films were developed and fixed by immersing it in respective solutions for 5 minutes. The sequences were read by placing the films on an illumination box and the data entered in a computer.

3.5 Polymerase Chain Reaction (PCR)

The DNA fragments were amplified using Polymerase Chain Reaction (PCR) which is an *in vitro* method for the enzymatic synthesis of specific DNA sequences. It makes use of two oligonucleotide primers that hybridise to opposite strands and synthesize the regions of interest in the target DNA. A repetitive series of cycle involving template denaturation, primer annealing and the extension of the annealed primers by the DNA polymerase (*Taq* polymerase*, isolated from thermostable bacterium *Thermus aquaticus*) results in the exponential accumulation of specific fragments (Mullis *et al.*, 1986; Mullis and Faloona, 1987).

PCR amplifications were performed for three DNA regions :

i). part of rRNA gene coding for 16S rRNA using *Frankia* specific primers.

ii). rRNA ITS region, that is, 3' end of 16S rDNA, the internally transcribed spacer region and 5' part of 23S rDNA region, using specific primers.

iii) *nif* IGS region, that is, 3' end of *nif* D, intergenic spacer region and 5' part of *nif* K gene using specific primers.

The annealing temperatures for the primers (table 3.2) were roughly calculated using the following formula,

Annealing Temperature = [{number of (A+T)x2}+ number of (G+C) x 4}]-5

Amplification was carried out in a total volume of 25 µl per tube. Each reaction mixture contained 3 µl of each primers (5 pM), 2.5 µl of 10X PCR buffer (Appendix-II.3a), 2.5 µl of each dNTP (1.25mM), 0.3 µl of *Taq* polymerase * (5U/ml), and 1 µl of template DNA. The use of hot start prior to amplification seemed to improve the reaction condition.

**Taq* polymerase – Cetus; Bangalore Genei; BOEHRINGER;Mannheim

Table 3.2 : Oligonucleotide primers used for PCR amplification

Primer	Sequence	Reference
FGPD 807	5' CACTGCTACCGGTCGATGAA 3'	Jamann *et al.*, 1993
FGPK 441'	5' CCGGGCGAAGTGGCT 3'	Nalin *et al.*, 1995
FGPS 989ac	5' GGGGTCCGTAAGGGTC 3'	Bosco *et al.*, 1992
FGPS 1509'	5'AAGGAGGGGATCCAGCCGCA 3'	Navarro *et al.*, 1992
FGPL 132'	5' CCGGGTTTCCCCATTCGG 3'	Ponsonnet & Nesme, 1994

Therefore, usually, prior to addition of *Taq* polymerase, the PCR mix was heated for 5 minutes at 95°C to denature the template. Amplification reaction was carried out for 35 cycles. Each cycle comprised of 1 minute of denaturation at 94°C, 1 minute of annealing at appropriate temperature (53°C for part of 16S rDNA region; 49°C for rRNA ITS region; 55°C for *nif* IGS region) and 1 minute of elongation at 72°C. It was finally followed by 10 minutes of extension at 72°C. The annealing temperature was further modified and adjusted for each primer pair with repeated experiments.

In addition to *Taq* DNA polymerase, occasionally, Expand^{TM} High Fidelity PCR system mix from BOEHRINGER, Mannheim (mixture of *Taq* and *Pwo* DNA polymerase) was also used for amplification. Since, it offers higher yield, fidelity and specificity, it was specially used for samples to be sequenced. In these cases however, 0.5 U Expand^{TM} High Fidelity PCR system was used per reaction and the reaction was set in 25µl of 1X HF buffer (Appendix II.3b).

3.6 Elution of DNA from Agarose gel

3.6.1 Elution of DNA using DEAE cellulose strip

After electrophoretic separation in a horizontal slab apparatus, a strip of DEAE cellulose membrane was placed in an incision just ahead of the band of interest. Electrophoresis was continued until binding was complete as confirmed by ethidium bromide stained fluorescence over long wave length transilluminator. The residual agarose from the DEAE strip was removed by thorough shaking in a tube containing NET buffer (Appendix- II.2c). The strip was then placed in a 1.5 ml microfuge tube and sufficient high salt NET buffer (Appendix- II.2d) was added to cover most of membrane followed by centrifugation for 5 seconds. It was further incubated at 55-68 °C for 45 minutes with occasional shaking so that bound DNA was released into the buffer. The buffer was then removed and extracted with 3

volumes of water saturated n-butanol so as to remove the residual ethidium bromide.

The DNA was precipitated with about 2.5 volumes of ethyl alcohol. The pellet formed was washed with 70% ethyl alcohol, vacuum dried briefly and finally dissolved in 10 μl of ultra pure water.

3.6.2 Elution of DNA using gelase

The DNA band of interest was cut out from the low melting point agarose gel using a razor blade and was incubated in gelase buffer (100 mg gel with 2 μl of 50X gelase buffer) at 65 °C until the agarose melted completely. It was then cooled down to 45°C and allowed to stand for 5 minutes followed by incubation in gelase enzyme (added at final concentration of 0.3 units per 100 mg of agarose) for 1 hour at 45°C. DNA was precipitated using 1 volume of ammonium acetate and 4 volume of ethyl alcohol at room temperature followed by centrifugation at 12,000 rpm for 30 minutes. The pellet formed was washed with 70% ethyl alcohol, vacuum dried and finally dissolved in 10 μl of ultra pure water or Tris buffer (10 mM, pH-8.0).

3.6.3 Electroelution of DNA

Followed by electrophoretic separation of DNA, small slots were made just ahead of the band of interest. Electrophoresis was continued with minimum buffer at 10 V/cm for approximately one minute. Buffer from the slit was removed in a microfuge tube and was observed under UV light . The process was repeated until all traces of DNA fragment were removed from gel. The DNA was the precipitated in 2.5 volumes of cold ethanol and 1/10 volume of 4M lithium chloride and recovered by centrifugation as described in section 3.2.1.1.

3.7 Purification of PCR product using QIA quick PCR purification kit

The amplified PCR products were purified using QIAquick PCR purification kit from Qiagen GmbH. It ensured efficient recovery of DNA by its adsorbance to the silica membrane in presence of high salt. Also, the impurities were efficiently washed away.

1. To 20 µl of PCR amplified product, about 100 µl of buffer PB (supplied with the kit) was added and mixed.

2. A QIAquick column was placed in a 2 µl collection tube and to it the above mix was added, and centrifuged at 10,000 xg for 1 minute.

3. The flow through was discarded and the QIAquick column was placed back into the same tube.

4. To wash the bound DNA, about 750 µl of buffer PE was added and centrifuged for 1 minute at 10,000 xg.

5. To elute DNA, the QIAquick column was placed in a fresh 1.5 ml micro centrifuge tube and to it about 30 µl of elution buffer (Appendix-II.2e) was added. It was allowed to stand for 1 minute followed by centrifugation at 13,000 rpm for 1 minute, so that, the purified DNA product got collected at the bottom of the microfuge tube.

3.8 PCR/AFLP (Amplified fragment length polymorphism)

The amplified PCR products of partial 16S rRNA and 16S-23S ITS region from different samples were run collectively together on 2% Nusieve agarose gel at 5V/cm for 1-2 hours and was photographed immediately. Similarly, the *nif* IGS PCR products from different samples were run similarly on 2% Nusieve agarose gel separately at 5V/cm for 2 hours. These were observed on a UV-transilluminator as described in section 3.4.1.

3.9 Restriction analysis of PCR products: PCR/RFLP(Restriction fragment length profile)

The amplified 16S-23S ITS as well as *nif* D-K IGS products were digested using appropriate restriction endonuclease. Four microlitres (100 ng) of the gel purified PCR products were digested in a 10 µl final volume for 2 hours with respective enzymes (*Taq* I, *Rsa* I, *Ava* I, *Hinf* I, *Alu* I and *Xma* I) as available in the laboratory. The buffers and temperatures were used as specified by the manufacturers for each endonuclease. While, 16S-23S ITS rRNA was digested with four restriction enzymes: *Taq* I, *Rsa* I, *Ava* I and *Hinf* I, *nif* D-K IGS was digested only with the restriction enzymes *Xma* I and *Alu* I.

For restriction digestion of the 16S-23S rRNA ITS amplified product, a master mix was prepared for each enzyme type separately. The master mix was prepared by addition of 8 µl of 10x buffer (buffer B for *Taq* I, buffer L for *Rsa* I, buffer 4 for *Ava* I, buffer 1 for *Xma* I, buffer A for *Alu* I, buffer H for *Hinf* I ; Appendix-II.5), 3µl of restriction enzyme (*Taq* I, *Rsa* I, *Ava* I, *Alu* I, *Xma* I and Hinf I) 4 units each and 31 µl of water to make the volume of 42 µl. From this, aliquots of 6 µl each were taken separately in different tubes and to each, 4 µl of the PCR purified DNA (from each sample) was added separately and labeled. Similarly, master mix was prepared for *nif* D-K IGS region by adding 4 µl of 10X restriction enzyme buffer (specific to each enzyme), 2µl of restriction enzyme followed by 6µl of water. From this mix, aliquots of 6 µl were taken as above in separate tubes and to each 4 µl of purified sample DNA was added. It was then incubated at the respective incubation temperatures (65°C for *Taq* I; 37°C for *Rsa* I, Hinf I, and *Alu* I; 25°C for *Ava* I, and *Xma* I) in a water bath for two hours. The reaction was stopped by increasing the temperature. The restriction digested fragments were separated using horizontal electrophoresis in TAE buffer with 4 % (w/v) Nusieve agarose gel or agarose gel containing 1µg/ml ethidium bromide. Gels were run at 2-

3 v/cm for 3 hours and immediately photographed. In addition, 8% acrylamide gel was also used followed by silver staining in order to separate the smallest fragments.

3.10 Silver staining

To detect the smaller or weaker fragments generated by restriction analysis silver staining method was employed. It proved to be a sensitive method detecting even the smallest fragments of less than 50 bp in size. The bands generated were of high resolution and sharp. For this, the following procedure was used-

1. The restriction digested products were first run in 8% acrylamide gel (Appendix- III.4) under denaturing condition.
2. Following the electrophoresis, the gel stuck on to the glass plate (treated with binding silane) was fixed in 10% acetic acid for 20 minutes after which, it was washed twice using distilled water for 5 minutes.
3. For silver staining reaction, the gel was then immersed into the silver staining solution [1% (w/v) silver nitrate solution-25 ml; 37% (w/v) Formaldehyde-0.25 ml and water added to make up the volume to 250 ml] for 20 minutes so that the silver particles got bound to the DNA. This was followed by washing in distilled water for about 10 seconds.
4. Following the above, the gel was immersed into the developing solution [sodium carbonate-6.25 g; 37 % formaldehyde (w/v) - 0.25 ml and water to make the volume 250 ml] for 5 minutes.
5. The gel was then immersed in stop solution (10 mM EDTA-$Na_2.2H_2O$ in distilled water) for 5 minutes.
6. For preserving the gel onto the glass plates, it was air dried overnight after immersing it in a solution containing 75 ml of ethanol and 11.5 ml of glycerol (v/v) for 1 hour.

3.11 Southern Hybridization

The transfer of DNA from agarose gel to nylon membrane was accomplished by the capillary transfer method at neutral pH using 20X SSC (Southern, 1975). The *Eco* RI/*Hind* III digested total genomic DNA was separated on 1.5 % agarose gel by electrophoresis in TAE buffer. After visualizing the ethidium bromide stained DNA over a transilluminator, it was depurinated by controlled acid treatment (250 mM HCl for 5 minutes). Further, the DNA was denatured by submerging in the denaturation solution (NaOH-0.5N, NaCl-1.5M) and neutralised by submerging in the neutralisation solution (Tris-HCl-0.5M, pH-7.5; NaCl-3M) twice for 15 minutes each followed by overnight blotting to positively charged nylon membrane using 20X SSC (NaCl-3M; Sodium citrate-0.3M; pH-7.0). The DNA was fixed to the membrane by vacuum baking for 2 hours at 80°C or by UV cross linking ($0.375 J/cm^2$).

3.11.1 Random primed labeling of the PCR products

Labeling of DNA for using as probes for hybridization was carried out by random primed labeling with DIG-11-dUTP followed by hybridisation and detection using DIG' system provided by BOEHRINGER Mannheim. The DIG system uses Digoxigenin (a steroid hapten isolated from *Digitalis purpurea*) to label DNA and the DIG labeled probes were hybridised to the membrane bound DNA on the Southern-blot. Hybridised probes are immuno-detected with an alkaline-phosphate conjugated anti-digoxigenin antibody and then visualised with chemiluminescent substrate CSPD®.

1. The PCR products were labeled with Digoxigenin-11-dUTP using Klenow fragment. For this, about 200 ng of DNA was first dissolved in 15 µl of water then, heat denatured by keeping it in boiling water bath for 10 minutes and was immediately transferred to cold ethanol bath..

2. To the above, about 2 µl of 10X hexanucleotide mixture (Random hexanucleotides-62.5A_{260} units/ml; Tris-HCl-500mM, pH-7.2; MgCl2-100mM; DTE-1mM; BSA-2mg/ml) followed by 2 µl of 10X labeling mixture (dATP, dCTP, dGTP-1mM each; dTTP-0.65mM; DIG-11-dUTP-0.35mM; pH-7.5) and 1 µl of Klenow fragment was added and incubated overnight at 37°C.
3. The reaction was terminated by adding 2 µl of 200 mM EDTA (pH-8.0) and the DNA was precipitated using 2.5 µl of LiCl (4M) and 70 µl of chilled ethanol for 30 minutes at –80°C.
4. The sample was then centrifuged at 13,000 x g for 20 minutes at 4°C and the pellet was washed with 50 µl of cold 70% ethanol, vacuum dried and dissolved in 50 µl of TE buffer and was stored at –20°C.

3.11.2 Hybridization

Hybridisation of the labeled probe with genomic DNA on nylon membrane was performed in a rotating tube set in a hybridisation oven. The hybridisation temperature was adjusted through manipulation by experiments or after calculating the melting temperature (T_m) using the formula:

T_m = 81.5 (°C)–16.6 log M (NaCl)+0.41 (%GC)– 550/number of base pair (% formamide) (Sambrook *et al.*, 1989)

1. Non-specific nucleic acid binding sites on the membrane were blocked prior to hybridisation by treating with standard pre hybridisation buffer (5X SSC; Blocking reagent-1%, (w/v); N-laurylsarcosine-0.1%; SDS-0.02%) for 2 hours at 63°C (20ml/100cm^2).
2. The DNA probes were heat denatured in boiling water bath for 10 minutes and hybridisation was done overnight at 63°C.

3. After hybridisation, the filters were removed and washed twice in 2X wash solution (2X SSC/0.1% SDS) at room temperature for 15 minutes each to get rid of the unbound probes.
4. The unhybridised probes were subsequently removed by two rounds of stringent washing with 0.5X wash solution (0.5X SSC/0.1% SDS) at hybridisation temperature for 15 minutes each.

3.11.3 Detection of the DIG labeled nucleic acids

The detection of the labeled probes that hybridised with the genomic DNA on the membrane was accomplished by chemiluminescent reaction using DIG Luminescent Detection Kit (BOEHRINGER; Mannheim)

1. After the post hybridisation washes, the membrane was equilibrated in buffer 1 (Maleic acid-100mM; NaCl-150mM; pH-7.5) and was agitated gently in buffer 2 (1% w/v blocking reagent in buffer 1) in a plastic bag for an hour to minimize the unspecific binding of the antibody.
2. The filter was then exposed to antibodies conjugated with alkaline phosphatase (diluted in buffer 2 at the ratio of 1: 10,000 i.e. 75 mU/ml) for 30 minutes in plastic bag.
3. Following above, the membrane was washed twice in buffer 1 for 15 minutes each and was equilibrated in buffer 3 (Tris.HCl-100mM, pH 9.5; NaCl-100mM; $MgCl_2$-50mM), which is the detection buffer for 2 minute.
4. Diluted CSPD® solution in the ratio 1:100 in buffer 3 was applied to the membrane while allowing excess CSPD® to drain off.
5. The membrane was then sealed in transparent plastic bag and incubated at 37°C (10 minutes) for achieving optimal alkaline phosphatase activity. The chemiluminescent signal was detected by exposing the membrane to Kodak X-Omat film for 2-3hours.

3.12 Cloning

3.12.1 Ligation

Following purification of the PCR products, they were prepared for ligation by utilizing the 3'-5' exonuclease activity of Klenow fragment of DNA polymerase I to remove single base (adenine) 3' overhangs on the fragments. The PCR fragments were simultaneously phosphorylated by T_4 polynucleotide kinase (Clark, 1988). After the phenol/chloroform extraction and Microspin column purification, the PCR products were ligated to blunt the dephosphorylated pUC18 vector at the *Sma* I site using rapid ligation buffer from PHARMACIA. The ligation of the PCR products into the vector pUC18 were carried out using SureClone ligation kit (PHARMACIA, Freiburg).

1. The blunting/ kinasing reaction was set in a microfuge tube by adding the following components : 1-16µl (500ng) of PCR product; 1 µl (10 U) of Klenow fragment; 2 µl of 10X blunting/kinasing buffer; 1µl (9.3 U) of polynucleotide kinase and sterile double distilled water to top the volume to total 20 µl.
2. The above reaction mix was gently mixed followed by brief centrifugation and incubation at 37°C for 30 minutes
3. Further, it was extracted with 20µl of phenol/chloroform at 12,000 rpm for 1 minute.
4. The PCR products were recovered by collecting the upper aqueous layer and purifying it on Microspin column (Sephacryl S-200).
5. The PCR products were then ligated into pUC18 vector at *Sma*/BAP site.
6. The ligation reaction was set by adding the following components into a microfuge tube : Column effluent – 5µl; 50ng of dephosphorylated vector – 1µl; 2X ligation buffer- 10 µl; DTT solution – 1 µl; T_4 DNA ligase – 1 µl and sterile distilled water to make the total volume of 20 µl.

7. After a brief centrifugation, the mix thus prepared was incubated at 16°C for 1–2 hours. Following which, it was used for transformation using competent *E. coli* cells.

3.12.2 Preparation of competent cells (Chung *et al.*, 1989)

The competent bacterial cells were prepared for transformation using protocol described by Chung *et al* .(1989). Cells prepared by this method were competent enough to achieve transformation at the rate of 10^7-10^8 colonies/µg of the supercoiled DNA.

1. Liquid culture of *E. coli* bacterial cell (strain XL1-Blue*) was prepared. For which, a single cell colony grown on LB agar plates (Appendix- I.3) was picked and put in 2 ml of LB medium (Appendix- I.2) and incubated for 16-20 hours at 37°C.
2. About 500 µl of the liquid culture was inoculated into 100 ml of LB medium (Appendix-I.2) and was incubated for 3 hours in an incubator shaker (225 cycles/minute) at 37°C until the OD_{600} reached about 0.5.
3. The optimally grown culture (10^6-10^7cells per ml) thus prepared was transferred to 30 ml pre-cooled cronex glass tubes and the cells were recovered by centrifugation at 4,000 rpm for 10 minutes at 4 °C.
4. The pelleted cells were resuspended in 10 ml (1/10 vol.) of ice cold TSS buffer (Appendix-II.6b) by gentle shaking.
5. The competent cell were shock frozen in an ethanol/dry ice bath and stored at -80°C until used for transformation with recombinant plasmids.

---

*XL1-Blue *E. coli* strain, supE44hsdR17 recA1 endA1 gyr A46 thi rel A1 lac$^-$F[pro AB+ lacq lacZDM15Tn10 (tet^1)] (Bullock *et al.*, 1987)

3.12.3 Transformation

Following mixing of the plasmid DNA with the competent cells, it was given heat shock, so as to allow efficient entry of DNA into cells. The cells were then grown in non-selective medium to allow synthesis of plasmid. Cells were then grown in non-selective medium to allow synthesis of plasmid encoded antibiotic resistance proteins. Then, it was plated on antibiotic containing medium to allow identification of plasmid containing colony encoded antibiotic resistance proteins. Then, it was plated on antibiotic containing medium to allow identification of plasmid containing colony.

1. To about 5-10 µl of the ligated sample in an ice cold test tube, 200 µl of the competent *E. coli* cells were added using chilled blue tips without generating air bubbles.
2. After 30 minutes of incubation on ice, the tubes were subjected to heat shock (42°C, 90 seconds) and were rapidly chilled on ice bath for 2 minutes.
3. It was followed by addition of 800µl of LB medium with glucose (20 mM) and incubated at 37°C for 45 minutes with shaking (225 rpm) to allow bacteria to recover and to express the antibiotic resistance marker encoded by the plasmid.
4. 100-200 µl of the transformed cells were plated on LB-agar plates containing Ampicilline (2.5 mg), 40 µl X-gal (2% dimethyl formamide) and 10 µl IPTG (100mM in water).
5. After the liquid was adsorbed, the inverted plates were incubated overnight at 37 °C for colour expression (Sambrook *et al.*, 1989).
6. The white colonies carrying the recombinant plasmids were separated out from the blue colonies (lac^+ colonies that result from a complementation between the amino terminal fragment of ß-galactosidase and the complete but defective poly-peptide in the bacterial genome) and selected for minipreparation of plasmid.
7. The insertions were verified by restriction analysis of the plasmids.

3.12.4 Isolation of plasmid from bacteria

Plasmids were isolated and purified from bacteria using either of the two methods described below.

3.12.4.1 QIA prep Spin Plasmid preparation Kit**

It is based on the alkaline lysis of bacterial cell followed by adsorption of the DNA onto silica in presence of high salt (Vogelstein and Gillespie, 1979).

1. About 3 ml per tube of overnight grown bacterial (*E. coli*) cultures were concentrated by centrifugation at 12,000 rpm for 30 seconds.
2. The pelleted cells were resuspended in 250µl of buffer P1 containing RNAase.
3. The suspended dispersed cells were further lysed by adding 250 µl of buffer P2 (NaOH/SDS) and were incubated for 3 minutes at room temperature. The SDS present in buffer P2 denatures the cellular proteins and the alkaline conditions cause the chromosomal and plasmid DNA to become denatured. Also, the incubation allows maximum release of plasmid DNA with minimum exposure to denaturing conditions.
4. To the lysate, about 350 µl of chilled buffer N3 (Guanidine hydrochloride buffer) was added and mixed immediately by inverting gently so that, the lysate got neutralised and the high salt binding condition was adjusted.
5. It was incubated for 5 minutes on ice, followed by centrifugation at 12,000 rpm for 10 minutes.
6. The supernatant was added to a QIA prep Spin column kept in 2 ml micro centrifuge tube followed by brief centrifugation.
7. The flow through was discarded, and the column (with DNA) was washed first with 500µl of PB buffer followed by final wash with 750 µl of PE buffer.

**QIAprep spin plasmid kit, DIAGEN, Gmbh Hilden, Germany

8. The residual alcohol (in PE buffer) was removed completely by a second centrifugation at 12,000 rpm for 1 minute.
9. The purified plasmids were eluted in 100 µl of 10 mM Tris HCL (pH-8.5) and the elute was collected in a fresh eppendorf tube by centrifugation at 12,000 rpm for 30 seconds.

3.12.4.2 "Speed Prep" method
(Ausubel *et al.*, 1989; Goode and Feinstein, 1992)

1. About 1.5ml of overnight grown bacterial (*E. coli*) cultures were transferred from liquid culture generated from single colonies and was centrifuged at 7000 rpm for 10 minutes.
2. The pelleted cells were resuspended and lysed in 100µl of Speed prep solution (Appendix-.III.3).
3. After complete lysis, 100 µl of phenol/chloroform was added and was vortexed for 10 seconds followed by centrifugation at 13,000 for 2 minutes.
4. To the upper clear portion, about 200 µl of ice cold ethanol was added and mixed by inverting and was incubated on ice for 30 minutes followed by centrifugation at 12,000 rpm for 5 minutes.
5. The pelleted plasmid DNA was washed with 70 % ethanol, followed with brief centrifugation and vacuum drying and the pellet was dissolved in 20 µl of TE buffer (10mM Tris HCl; pH-8.0).
6. RNA in the samples was removed by incubating with 1/20 volume of RNAase solution (from stock solution of 5 mg/ml) for 30 minutes at 37°C.

3.12.5 Restriction digestion of Plasmid DNA

Purified plasmids were checked for inserts by digestion with appropriate restriction endonucleases followed by gel electrophoresis. For this, approximately 1µg of the plasmid DNA (in 10 µl of 10mM Tris HCl, pH-8.0) was incubated with 1 unit each of *Eco* RI and *Hind* III in its respective buffer 'B' for 1-2 hours at 37°C

in a water bath. The reaction was then arrested by heat denaturing enzymes for 5 minutes at 65 °C.

3.13 DNA sequencing

Covalently closed circular DNA was sequenced by the dideoxy chain termination technique (Sanger *et al.*, 1977) using (^{35}S)-d ATP and T7 sequencing kit (PHARMACIA). The sequencing reaction was completed in the following four reaction steps:

1. Denaturation: The concentration of the template was adjusted so that 100 µl contained 5 µg DNA. In order to denature the double stranded template, 32 µl (1.5-2 µg) sample was incubated with 8 µl of 2M NaOH for 10 minute at room temperature. Then 7 µl of 3 M Sodium acetate (pH-4.8), 4 µl water and 120 µl ethanol were added, mixed gently and kept at –80 °C for 30 minutes. The precipitated DNA was recovered by centrifugation at 12,000 rpm (4°C), washed with 70% ethanol, briefly vacuum dried and dissolved in 10 µl of distilled water.

2. Primer annealing: The denatured plasmid DNA (1.5 –2 µg dissolved in 10 µl water) was mixed with 2 µl annealing buffer and 2 µl sequencing primer (2.5-5 pM), vortexed gently, centrifuged briefly and then incubated 5 minutes at 65°C following which, it was transferred immediately to 37°C for 10 minutes and was then kept for 5 minutes at room temperature.

3. Labeling reaction: To the above annealing reaction (14µl), 3 µl of labeling mix (dATP, dTTP, dGTP, dCTP; supplied with the kit), 1 µl of (a 35S) d ATP (37 Tbq /mM) and 2 µl diluted (1:4, in enzyme dilution buffer) T_7 DNA polymerase (3.2 U) was pipette, mixed by gentle pipetting, centrifuged briefly and then incubated 5 minutes at room temperature.

4. Termination reaction: 4.5 µl of labeling reaction was transferred into each of the four prewarmed sequencing mixes (2.5 µl of A-, C-, G-,and T mix short), mixed by gentle pipetting and incubated at 37 °C for 5 minutes). Subsequently, 5 µl of stop solution was added and the probes were heat denatured (80 °C, 2 minutes) and stored at –20 °C until loaded on the gel.

3.14 Computer analysis of data

All the details of the amplified DNA products like the band size, band position, number and size of bands generated through restriction digestion, sequence information of each nodule microsymbiont etc. were entered into the computer files for further analysis. The nucleotide sequence of each sample used for study were entered into the computer files of the program DNASIS (version-7.0). The sequences obtained from this study was aligned with the sequences retrieved from the GenBank database (http://ncbi.nl.nih.gov) resulted using BLAST search (Altschul *et al., 1990*). The Programme ALIGN (version 2.0) was used for local sequence alignment and the program CLUSTAL W (1.74) was used for multiple sequence alignment (Thompson *et al.*, 1994). So that, the evolutionary homologous regions in the rDNA and *nif* D-K portions could be juxtaposed for inferring phylogenetic trees needed for structural similarity or dissimilarity calculation. To construct the phylogenetic tree, the phylogenetic inference package PHYLIP for windows (version 3.5c, Felsenstein, 1995) was used. Similarly, the *nif* D-K sequences were also analysed at phylogenetic level using sequences retrieved from GenBank selected after BLAST search using one of the sample sequence. To analyze further, the amino acid sequence of partial *nif* D and *nif* K region was also studied.

CHAPTER 4

RESULTS AND DISCUSSION

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Collection of nodules

Three *Alnus* species, *A. nepalensis*, *A. nitida* and *A. glutinosa*, growing in different geographical locations were investigated in this study. They were all found growing between the elevations of 900-2500m above mean sea level. All the trees investigated for this study were found nodulated. The root nodules or nodule clusters (fig 4.1.1) were mostly located between 5-10 cm below the soil surface except for a few cases where they were 10-20 cm below the soil surface. On the basis of colour, the nodules could be categorized into active (pale yellow to yellowish brown) and inactive (dark brown to brownish black) nodules (Sharma and Ambasht, 1986). With the exception of the nodules collected from Upper Shillong, all the nodules collected from different locations were found to be uniform in morphology.

Little variation was observed on the basis of nodule morphology. However, nodules/ nodule clusters analysed from *Alnus nepalensis* were different with respect to their sizes compared to those of *A. nitida* or *A. glutinosa*, irrespective of the season in which they were collected. Also, unusual hair growth on nodules collected from Upper Shillong, Meghalaya, was observed occasionally which was similar to those found in case of *Myrica* nodules (fig 4.1.2).

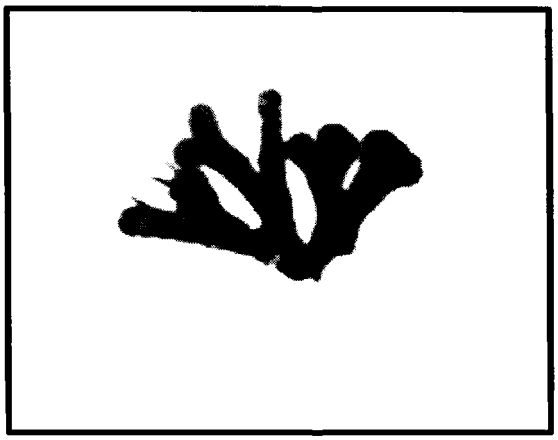
As far as the host plant involved in the symbiosis is concerned, there was morphological variation at inter-specific level. This included tree size, leaf shape, leaf colour, colour of tree bark and size of fruiting bodies, etc. Of the three species, *Alnus nitida* was the largest tree species studied having lighter tree bark, larger fruiting bodies and large globose leaves. *Alnus glutinosa*, although

Fig 4.1.1 : Root Nodules of *Alnus nepalensis*.

a) a lateral root exposed to show the cluster formation of root nodules on it.
b) distribution of nodule clusters in the roots and rootlets in soil.
c) young nodules marked by colouration.
d) a mature nodule cluster.
e) a portion of rootlet showing nodule clusters both old and decaying along with formation of new ones.

Figure 4.1.2

Fig 4.1.2 : Unusual hair growth in *Alnus* nodules collected from Upper Shillong.



morphologically similar to *A. nepalensis* in all other respects, had irregularly serrated leaf blades.

4.2 DNA extraction and gel electrophoresis

The total genomic DNA was extracted using the modified method of Rouvier *et al.* (1996) as described in section 3.2.1.2. It gave good reproducible results. While, all other nodule samples consistently gave only one 21 kb DNA band on electrophoreses (fig 4.2), Nonkrem Hill and Tuebingen samples however, occasionally gave additional bands. Although the quantity of DNA extracted per nodule sample was very small, it was nevertheless sufficient for PCR amplification. The DNA was purified prior to amplification by electroelution. The A260/A280 values were usually between 1.0-2.0 indicating low levels of contaminant proteins.

4.3 DNA amplification

4.3.1 PCR amplification of partial 16S rRNA

PCR amplification was carried out for a part of 16S rRNA using a combination of *Frankia* specific and universal primers FGPS 989ac (Bosco *et al.*, 1992) and FGPS 1509' (Navarro *et al.*, 1992) (table 3.2). The amplified product consistently obtained in all cases was a single band of ~500 bp size (fig 4.3). Each nodule sample collected from different locations was assigned a name with respect to the species and site of geographical location [Ooty-Anob, Himachal Pradesh-Anhpk, Arunachal Pradesh-Anahr, Nokhrem Hills-Anmnh, Upper Shillong-Anmus, Tuebingen-Aggtun]. From each site more than 50 nodules were analysed. In addition to the nodule samples, PCR amplification and sequence analysis of rRNA partial sequence and 16S-23S rRNA ITS region for a *Frankia* strain ACN 1^{AG} (Universite' Claude-Bernard, Lyon, France) was also carried out for comparison as a reference control.

Figure 4.2

Fig 4.2 : Gel electrophoresis of genomic DNA isolated from *Alnus* nodules collected from different geographical locations: lane 1 - Anarh; lane 2b- Aggtun; lane 3b- Anhpk; lane 4 - Anmus; lane 5 - Anob; lane 6 - Anmnh; lane 7 - ACN 1^{AG} and lane 8 - λ DNA digested with *Eco* RI/*Hin*d III double digest as molecular base pair marker (in kbp).

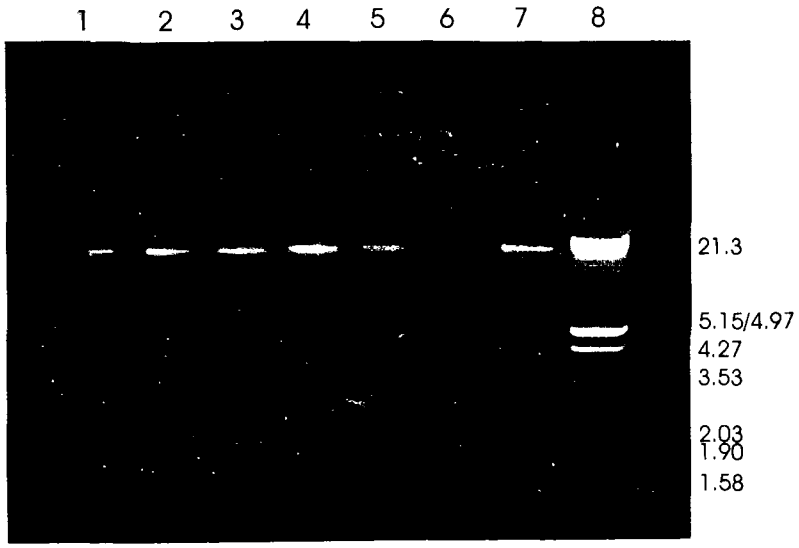
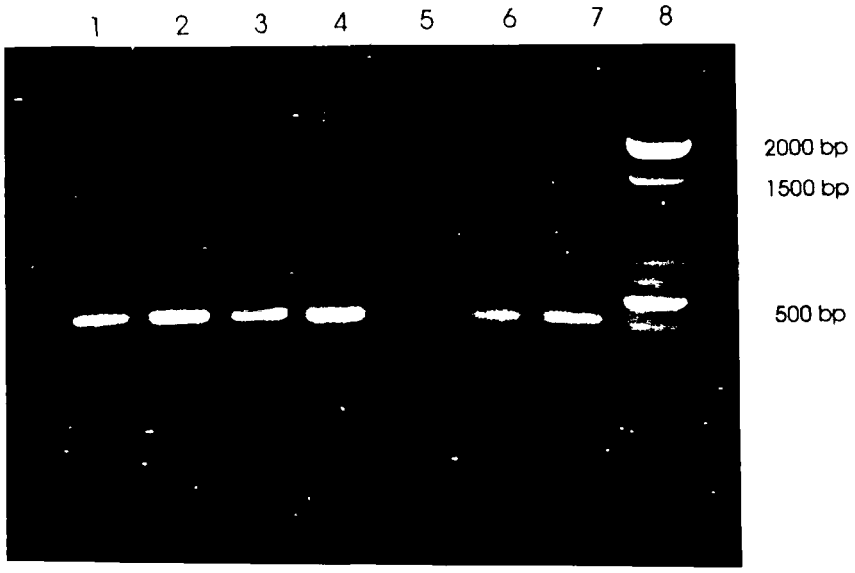

1
2
3
4
5
6
7
8
21.3
5.15/4.97
4.27
3.53
2.03
1.90
1.58

Fig 4.3 : Agarose gel electrophoresis of PCR amplified partial 16S rRNA gene isolated from nodule samples collected from seven different study sites.

Lane 1 - Anarh; lane 2 - Aggtun; lane 3 - Anhpk; lane 4 - Anmus; lane 5 - Anob; lane 6 - Anmnh; lane 7 - $ACN1^{AG}$ and lane 8 - 100 bp molecualr ladder.


1
2
3
4
5
6
7
8
2000 bp
1500 bp
500 bp

4.3.2 PCR amplification of 16S-23S rRNA ITS region

In order to characterize and differentiate between the alder compatible frankiae from the root nodules collected from different locations, the most variable internal transcribed spacer region (ITS) between 16S–23S rRNA was analysed. For PCR amplification, region specific primers were selected from literature and used (table 3.2). Primer FGPS 989ac (Bosco *et al.*, 1992) in 16S rRNA gene was selected as the forward primer and a universal primer FGPL 132' (Ponsonnet and Nesme, 1994) in the 23S rRNA gene was selected as the reverse primer. Primer FGPS 989ac is alder-casuarina *Frankia* group specific, therefore, only DNA templates of *Frankia* origin can get amplified. Although, the actual size of the internal transcribed spacer region (ITS) between 16S-23S rRNA was about 420 bp, the size of the total amplified product was about 1.1 kbp (fig 4.4). The sizes of the amplified products were more or less conserved among the *Frankia* microsymbionts studied.

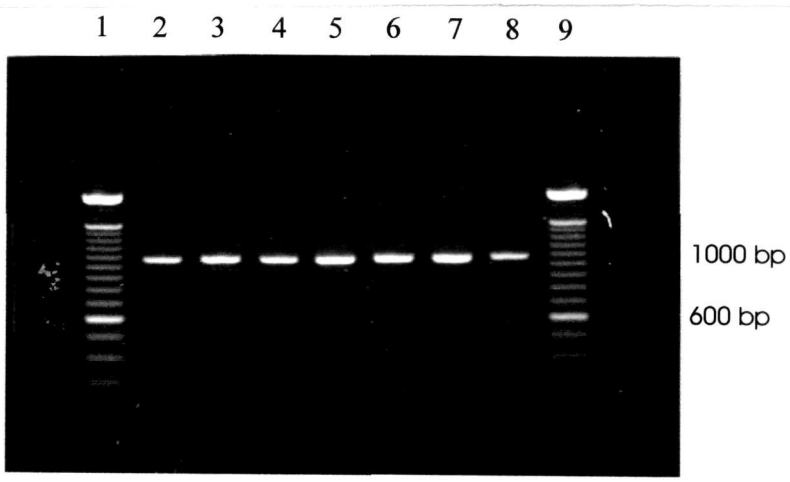
4.3.3 PCR amplification of *nif* D-K IGS region

Nitrogen fixing (*nif*) genes, though functionally conserved, do exhibit some divergence. Due to their absence in host plant, the *nif* genes are suitable as an additional tool for characterization of frankiae. Further, the *nif* D-K intergenic spacer (IGS) region is larger than *nif* H-D IGS (Normand *et al.*, 1992a). Therefore, it was expected to retain more variation during the process of evolution. It was for this reason that *nif* D-K IGS was chosen for the present investigations.

For amplifying *nif* D-K IGS region, primers FGPD 807 (Jamann *et al.*, 1993) and FGPK 441' (Nalin *et al.*, 1995) were selected. Primer FGPD 807 was derived from the conserved sites of *nif* D region which corresponded to the coordinate 807 of *nif* D (Jamann *et al.*, 1993) of the 4530-4549 bp region of *Frankia alni* strain ArI3 (Specq and Normand, 1996; Accession no.-L41344).

Fig 4.4 : Electrophoretic pattern of PCR amplified 16S-23S rDNA ITS region. Samples from seven different geographical locations were analysed.

Lane 1 and 9 - 100 bp molecualr size marker; lane 2 - Anarh; lane 3 - Aggtun; lane 4 - Anhpk; lane 5 - Anmus; lane 6 -Anob; lane 7 - Anmnh and lane 8 - $ACN1^{AG}$.


1 2 3 4 5 6 7 8 9
1000 bp
600 bp

Primer FGPK 441' corresponded to the coordinate 331 of *nif* K conserved sites in the *nif* K region between 5769 to 5754 of the strain Arl3 (Nalin *et al.*, 1995).

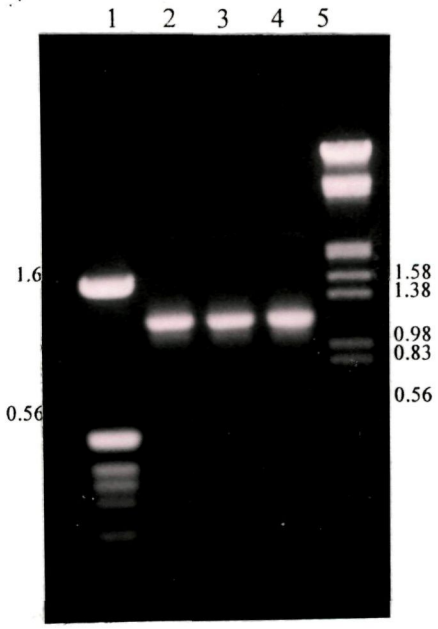
Although nodules were collected from different locations, positive results were obtained only for the samples collected from Nonkrem Hills (Meghalaya), Hapoli (Arunachal Pradesh) and Tuebingen (Germany). Among these, the one collected from Tuebingen was taken as out group for comparison. The PCR amplified products were of ~1200 bp size. The gel electrophoresis of the amplified products showed almost uniform sized fragments (fig 4.5). In case of Upper Shillong samples, infrequently faint bands, signaling amplification of the required specific region, occurred. This showed presence of *nif* genes needed for the nitrogen fixation in the sample. But, re-amplification of the same often resulted in smears instead of bands of required size. This obstructed further analysis of this region for the sample.

Analysis of partial region of 16S rRNA, 23S rRNA and the internal transcribed spacer region between 16S-23S rRNA, revealed presence of considerable variation for the *Alnus* microsymbiont studied. Similar observations was noted for partial region of *nif* D, *nif* K and *nif* D-K IGS as well.

Minor fragment length variation was present in PCR amplified products for 16S-23S rRNA ITS and *nif* D-K IGS regions. But, it was not detected on electrophoretic migration pattern, since it was limited only to few base pairs. However, the clear picture of the diversity existing with these microsymbionts were obtained only through restriction digestion analysis and sequence analysis of these amplified products.

Fig 4.5 : Agarose gel electrophoresis of *nif* D-K IGS region amplified from *Alnus* nodule microsymbionts.

Lane 1 - *Hinf* digested pBR322 as a molecualr marker (kbp); lane 2 - Anarh; lane 3 - Aggtun; lane 4 - Anmnh; lane 5 -λ DNA *Eco* RI/*Hin*d III double digest (in kbp).


1 2 3 4 5
1.6
0.56
1.58
1.38
0.98
0.83
0.56

4.4 Southern hybridisation

Southern hybridisation was carried out using amplified PCR products for 16S-23S rRNA ITS region as the probe over total genomic DNA digested with *Eco* R I/*Hind* III restriction enzymes. Against the expected one hybridisation signal, two hybridisation signals were observed (Fig 4.6). This could be due to either or both of the following reasons:

i) partial digestion of the frankial genome by *Eco* RI/ *Hind* III restriction enzyme may have resulted in two products of varying sizes.

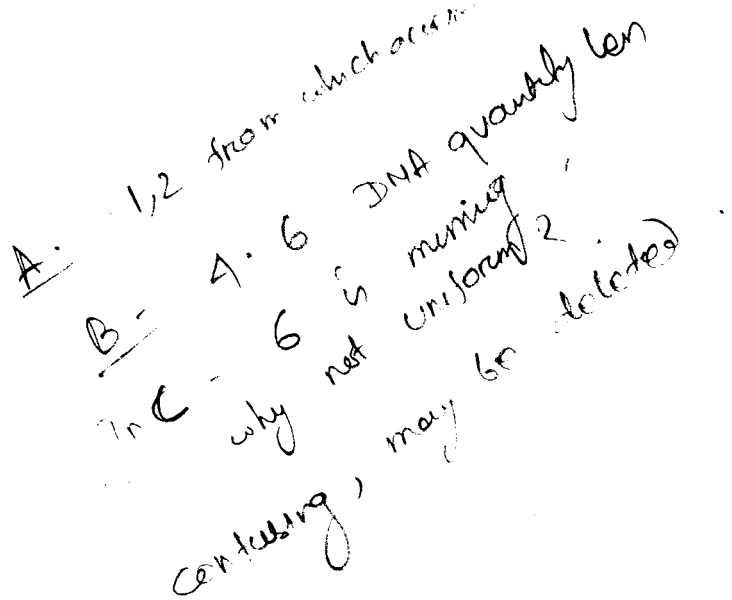
ii) *Frankia* genome contains two *rrn* operons (Normand *et al.*, 1992b). Any difference between the two might lead to two signals obtained (fig 4.6).

4.5 Restriction analysis of PCR amplified products

PCR based genomic fingerprinting method is a good method for analyzing genotypic diversity present within nodule microsymbiont. The strains of the microsymbiont within nodules can significantly differ in their nitrogen fixing efficiency or in their ability to occupy nodules in competition with other closely related strains (Benson and Hanna, 1983). So, the rapid and versatile method based on restriction endonuclease site difference in the PCR amplified 16S - 23S rRNA ITS and *nif* D-K IGS product was used as a tool to analyze the diversity existing in the nodule microsymbiont.

To perform restriction digestion PCR amplified and gel purified (Qiagen kit) 16S-23S rDNA ITS region was digested using restriction enzymes *Taq* I, *Rsa* I, *Ava* I and *Hinf* I. All these four restriction enzymes used yielded varying restriction patterns. The restricted fragments were well separated in 4% Nusieve agarose gel electrophoresis (fig 4.7). However, fragments shorter than 100 bps were not well separated in these gels.

Figure 4.6 : Hybridisation Dig-11 UTP labelled 16S-23S rDNA ITS fragment with Aggtun total genomic DNA.
Lane 1 and 2 (fig A) contains gel electrophoresis of PCR amplified 16S-23S rDNA ITS and *nif* D-K IGS region; lane 3 (fig A, B & C) contains λ DNA *Eco* RI/*Hin*d III double digest as molecular size marker and lane 4, 5 and 6 (fig B & C) contains total genomic DNA digested with *Eco* RI/*Hin*d III of ACN1^{AG}, Aggtun and Anmus; lane 5 (fig C) shows two hybridisation signal of >2kb size.



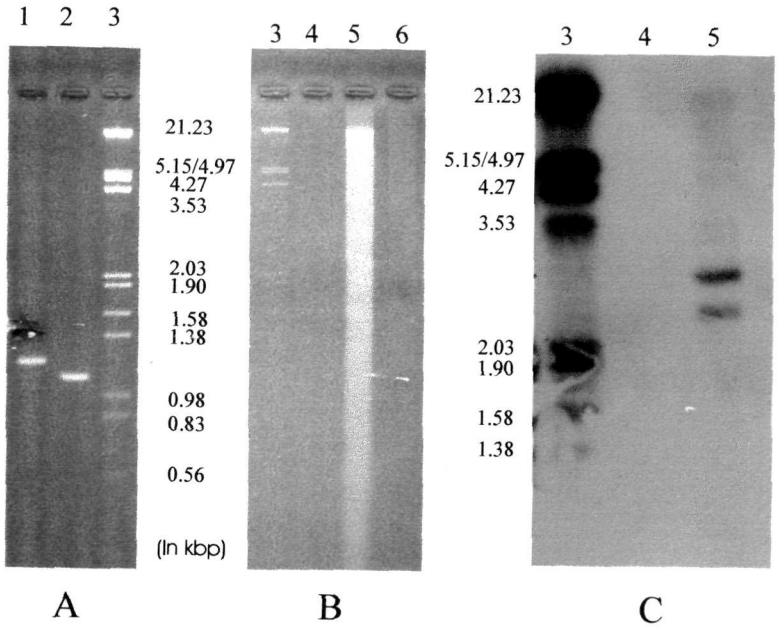
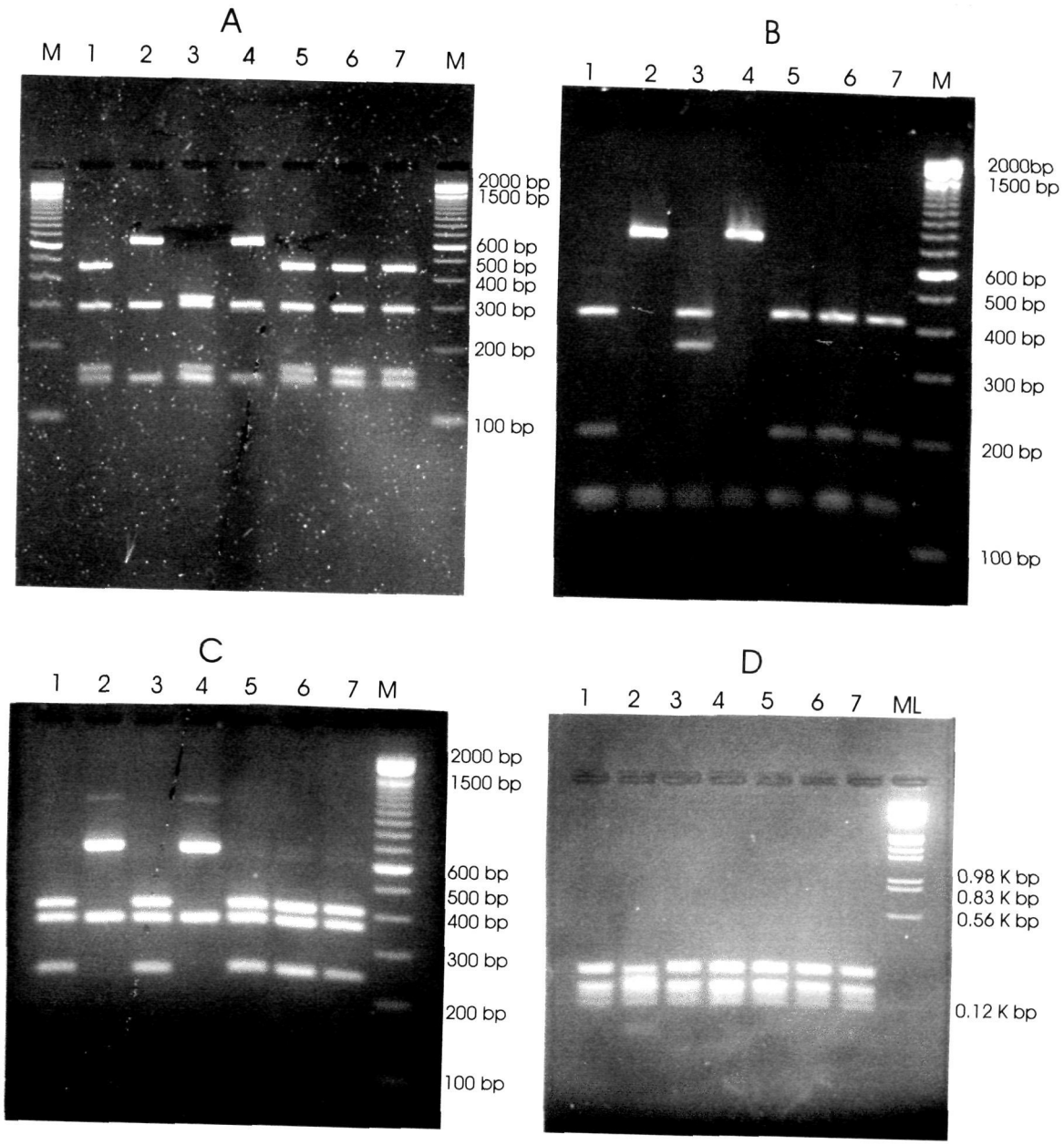

1 2 3
21.23
5.15/4.97
4.27
3.53
2.03
1.90
1.58
1.38
0.98
0.83
0.56
(In kbp)
3 4 5 6
3 4 5
21.23
5.15/4.97
4.27
3.53
2.03
1.90
1.58
1.38
A
B
C

Fig 4.7 : Restriction digestion pattern of 16S-23S rDNA ITS region digested with A) *Taq* I B) *Hinf* I C) *Ava* I and D) *Rsa* I.

Lane 1 - Anarh; lane 2 - Aggtun; lane 3 - Anhpk; lane 4 - Anmus; lane 5 - Anob; lane 6 - Anmnh and lane 7 - $ACN1^{AG}$. The molecular size marker are marked as 'M' (100 bp ladder) and 'ML' (λ DNA *Eco* RI/*Hin*d III double digest).

5, 6, 7 looks similar


A
M 1 2 3 4 5 6 7 M
2000 bp
1500 bp
600 bp
500 bp
400 bp
300 bp
200 bp
100 bp
B
1 2 3 4 5 6 7 M
2000bp
1500 bp
600 bp
500 bp
400 bp
300 bp
200 bp
100 bp
C
1 2 3 4 5 6 7 M
2000 bp
1500 bp
600 bp
500 bp
400 bp
300 bp
200 bp
100 bp
D
1 2 3 4 5 6 7 ML
0.98 K bp
0.83 K bp
0.56 K bp
0.12 K bp

In order to detect even the smaller fragments generated during the restriction digestion, the silver staining method was employed. For this, the restriction digested products for 16S-23S rRNA using the restriction enzyme *Rsa* I (Fig 4.8) and *nif* D-K IGS region using restriction enzymes *Alu* I and *Xma* I (fig 4.9) were run on 8 % acrylamide gel followed by silver staining. This enabled detection of even ~50 bp fragments. However, probably due to high resolution and sensitivity of the silver staining and/or the non-specific digestion by the enzymes, some additional bands were observed. For example, the lower most common band (~90 bp) which did not fit with the predicted pattern. Nevertheless, these nonspecific, unexplainable restriction bands did not seem to overlap the general expected pattern.

The nodules where the amplification for the regions stated succeeded, were only used for restriction digestion. Computer simulated restriction digestion of 16S-23S *rrn* ITS region with the four enzymes used (*Taq* I, *Hinf* I, *Rsa* I and *Ava* I) resolved five PCR/RFLP groups or types based on their restriction fragment size variations (table 4.1.1). However, on visual observation in gel, only two to three types could be deduced from their banding patterns. Based on the results obtained it was observed that the four selected restriction enzymes could discriminate between the strains of microsymbiont analysed. Restriction pattern type 1 comprised of the *Alnus nepalensis* nodule microsymbiont samples Anarh, Anob and the *Frankia* isolate ACN1^{AG}. Type 2 included the microsymbiont Aggtun, derived from the host *Alnus glutinosa*. Type 3 was formed by *Alnus nitida* nodule microsymbiont Anhpk. The nodule sample Anmus, isolated from *Alnus nepalensis* fell into restriction pattern type 4. This sample was distinct from all other samples collected from India. However, the close relationship of Anmus with Aggtun was particularly noted based on the PCR/RFLP pattern and the restriction fragment size variation (table 4.1.1). The restriction fragment size variation was calculated for each enzyme from the discriminative restriction sites present on the PCR amplified fragment.

Fig 4.8 : Restriction pattern of PCR amplified *Alnus* nodule rDNA (partial 16S, 23S and 16S - 23S ITS region) digested with *Rsa* I and visualized by silver staining. The digested samples were Anarh (lane 2), Aggtun (lane 3), Anhpk (lane 4), Anmus (lane 5), Anob (lane 6), Anmnh (lane 7) and $ACN1^{AG}$ (lane 8). Lane 1 contained *Pst* I digested $ACN1^{AG}$ sample and lane 9 contained a 100 bp ladder .

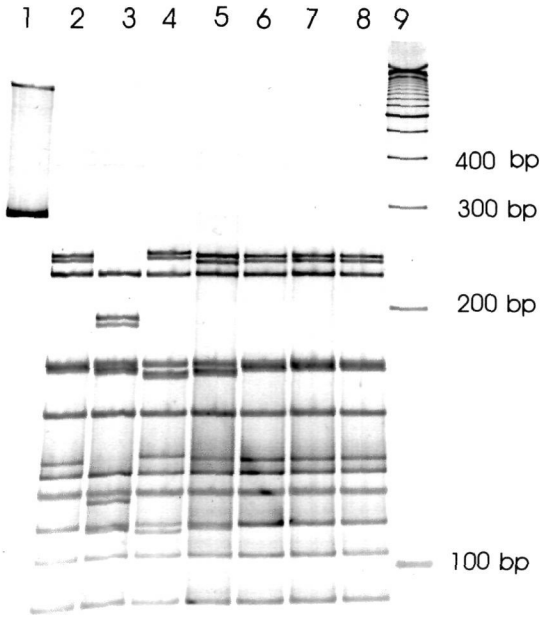

1 2 3 4 5 6 7 8 9
400 bp
300 bp
200 bp
100 bp

Fig 4.7 : Restriction digestion pattern of 16S-23S rDNA ITS region digested with A) *Taq* I B) *Hinf* I C) *Ava* I and D) *Rsa* I.

Lane 1 - Anarh; lane 2 - Aggtun; lane 3 - Anhpk; lane 4 - Anmus; lane 5 - Anob; lane 6 - Anmnh and lane 7 - $ACN1^{AG}$. The molecular size marker are marked as 'M' (100 bp ladder) and 'ML' (λ DNA *Eco* RI/*Hind* III double digest).

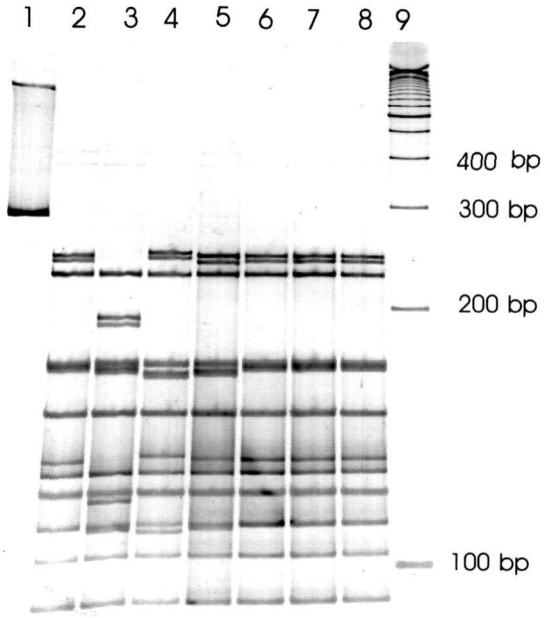

1 2 3 4 5 6 7 8 9
400 bp
300 bp
200 bp
100 bp

Fig 4.9 : Restriction digestion pattern of *nif* D-K IGS region resolved in 8 % denaturating polyacrylamide gel stained with silver nitrate. Two different enzymes were used - *Xma* I and *Alu* I for the digestion of nodule samples Anarh and Aggtun.

I. *Xma* I digested samples : lane 1- Aggtun; lane 2 - Anarh. II. *Alu* I digested samples : lane 3- Aggtun; lane 4 - Anarh; lane 5 - 00 bp molecualr size marker.

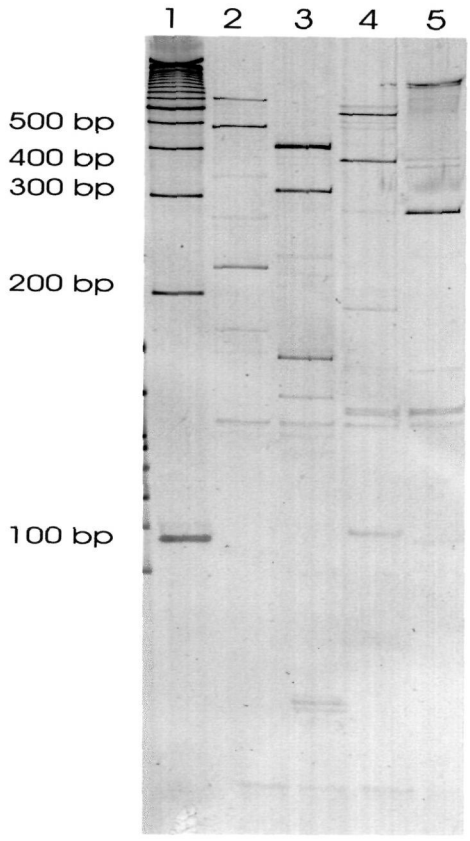

1
2
3
4
5
500 bp
400 bp
300 bp
200 bp
100 bp

Table 4.1.1: Computer predicted restriction fragment sizes (bp) for 16S-23S rRNA ITS region.

Nodule Micro-symbionts	Restriction enzyme/restriction fragment size (bp)			
	Ava I	*Hinf* I	*Rsa* I	*Taq* I
Anarh	362, 244, 405, 53	124, 132, 59, 418, 131, 198	226, 146, 119, 166, 238, 167	300, 146, 161, 458
Aggtun	244, 118, 656, 53	124, 118, 132, 159, 755	226/245, 128, 119, 176, 193, 140, 167	245, 145, 146, 626
Anhpk	362, 242, 407, 53	124, 132, 59, 424, 334	226, 146, 119, 163, 231,167	300, 146, 626,159, 142, 317
Anmus	245, 117, 648, 50	124, 118, 133, 159, 755	226/246,127, 119, 163, 239, 167	246, 145, 146, 619
Anob	362, 244, 405, 53	124, 132, 59, 418, 131, 198	226, 146, 119, 166, 238,167	300, 146, 161, 458
Anmnh	362, 245, 405, 50	124, 132, 59, 419, 131, 198	226, 146, 119, 167, 238, 166	300, 146, 161, 457
ACN1^{AG}	362, 244, 405, 53	124, 132, 59, 418, 131, 198	226, 146, 119, 166, 238, 166	300, 146, 161, 458

Restriction pattern type 5 included the nodule sample Anmnh isolated from the *Alnus nepalensis*. Both Anmus and Anmnh nodule samples were collected from Meghalaya, but from different sites.

From the restriction analysis results, it was clear that the nodule microsymbiont Anmus, from Upper Shillong, was very close to the nodule microsymbiont Aggtun from Tuebingen. The *nif* D-K IGS restriction digestion of the Aggtun and Anarh samples using the restriction enzymes *Alu* I and *Xma*I clearly differentiated them giving two distinct restriction patterns (table 4.1.2).

Thus, the restriction analysis clearly revealed presence of diversity. In addition to the four restriction enzymes used above, sequence information of the 16S - 23S rRNA ITS and *nif* D-K IGS region allowed analysis of the restriction patterns generated by several other restriction enzymes using the computer program DNASIS.

The discriminative power of PCR-RFLP method has been already assessed for a number of organisms (Navarro *et al.*, 1992; Laguerre *et al.*, 1996; Ritchie and Myrold, 1999). A distinct restriction pattern is strain specific for *Frankia* species (Dobritsa, 1985) and restriction patterns obtained can differentiate morphologically similar organism (Bloom *et al.*, 1989) or provide evidence for variation existing within a *Frankia* isolate (An *et al.*, 1985b). This is based on the presence or absence of the fragments resulting from changes in the specific restriction enzyme recognition sites in the amplified fragment.In the study conducted, the four restriction enzymes used generated five different patterns for the 16S-23S rRNA ITS region.

As expected, the restriction pattern obtained overall, did follow a geographic pattern being collected from six diverse locations. Among the samples collected from *Alnus nepalensis*, the samples Anob (Ooty), Anmnh (Nonkrem Hills) and

Table 4.1.2 Computer predicted restriction fragment sizes for *nif* D-K IGS region

Nodule Microsymbionts	Restriction enzymes/Restriction fragment size (bp)	
	Alu I	*Xma* I
Anarh	406, 308, 67, 301	1043, 141
Aggtun	704, 471	561, 81, 372, 141

Anarh (Arunachal Pradesh), although distributed at different locations, showed very close relationship to each other. Although, minor variations were noticed in the size of the products (which were confirmed by computer analysis of sequences subsequently) the electrophoretic migration pattern did not show any variation among these samples. Furthermore, ACN1^{AG}, the *Frankia* strain isolated from *Alnus crispa* did show a similar restriction pattern. However, in spite of the fact that, being from the same host (*Alnus nepalensis*), little similarity was observed between the sample Anmus (Upper Shillong) with other samples collected from India (Anob, Anarh Anmnh).

The samples Anhpk (Himachal Pradesh) collected from *Alnus nitida* showed distinct banding pattern on digestion with *Taq* I and *Hinf* I. *Rsa* I digestion although did not give any distinct pattern on agarose gel electrophoresis the variation could be resolved on a polyacrylamide gel.

Among the nodules collected from *Alnus nepalensis* the Upper Shillong sample stood distinct in its pattern. Interestingly, this pattern obtained was similar to those obtained for the sample Aggtun. For example, the restriction fragment size (table 4.1.1) generated by *Hinf* I was nearly the same for both Anmus and Aggtun even though their restriction sites were not exactly similar. This was also true with respect to *Taq* I and *Ava* I. However, *Rsa* I gave an overall uniform pattern with respect to other samples collected from India. However, the Aggtun and Anmus gave a distinct pattern on digestion with *Rsa* I.

Similar polymorphism were also obtained for *nif* D-K IGS region. In this case, two distinct restriction patterns were obtained for the two different *Alnus* species analysed. PCR amplified Anarh and Aggtun *nif* D-K IGS region on digestion with restriction enzyme *Alu* I and *Xma* I also resolved two distinct patterns (fig 4.9).

This was expected as they both belonged to two different species of *Alnus* located at two different geographical locations.

In general, based on the data obtained from the restriction fragment size variation and the restriction pattern obtained for each enzyme, the close relationship of Anmus sample with Aggtun sample was confirmed. Although, these two samples were collected separately from two different *Alnus* species growing in two different continents, the relatedness revealed on restriction analysis was very obvious.

In addition to the six enzymes used, restriction analysis was also carried out with a number of other enzymes through computer program DNASIS. It too revealed presence of heterogeneity in the samples analysed.

4.6 Ribosomal DNA sequence analysis

The usual order of structure of rDNA in prokaryote is 16S–spacer–23S–5S, with few exceptions. The primary structure is known to contain stretches of conserved and variable regions, which are being exploited for studies related to characterization, diversity, phylogeny, etc. The ITS regions in the prokaryotic rRNA gene are situated between 16S–23S (ITS 1) and between 23S–5S (ITS 2). The ITS regions are generally more variable than the conserved sequences of rRNA gene (Jensenn *et al.*, 1993). Following PCR amplification reaction, sequence analysis were carried out for the partial 16S rRNA as well as for the 16S-23S rRNA ITS region, so that species specific polymorphism could be evaluated. Ideally the DNA samples used for PCR-RFLP studies described in section 4.5 should also have been used in the sequencing experiments, but since the DNA amplified earlier was fully used in the PCR-RFLP experiments, it was not available for sequencing experiments. In some cases, where the DNA could be saved, it was used up during the sequencing experiments without giving any reliable information. Therefore,

fresh nodule samples had to be used for these experiments. However, the place of collection of the samples was the same as in the earlier experiments.

4.6.1 Cloning and sequencing of partial 16S rRNA gene

In order to carry out sequencing, the PCR amplified rDNA fragments were ligated into vector pUC18 at *Sma* I site and were cloned in *E. coli*. Positive clones were selected for minipreparation and the recombinant plasmids were isolated from overnight cultures. The Expand Polymerase® was used for amplification of DNA to be used for sequencing experiments. The frequency of errors committed during copying of the template by this enzyme system is very low. Therefore, only one colony per sample was picked up for further processing. The insertions of desired fragments were checked by digestion using restriction endonucleases *Eco* RI and *Hin*d III or *Pst* I that cleave the polylinker site of pUC18 (fig 4.10; 4.11). Sequencing of the clones were carried out using the dideoxy chain termination method (Sanger *et al.*, 1977) and the cloned plasmids were sequenced using universal M13 forward and M13 reverse primers (table 4.2). Since, the size of the amplified products was approximately 520 bp, the universal primers (M13 forward and M13 reverse) were found sufficient to determine the complete sequences of both the strands of the inserts. Each DNA sequences obtained were entered into the DNASIS computer files for further analysis.

Partial 16S rDNA sequence obtained from the nodule DNA collected from different locations as well as from the *Frankia* strain ACN1^{AG} were aligned with the corresponding sequences of *Frankia* strain ORS020606 (corresponded to nucleotide sequences between 989–1509 of 16S rRNA; Accession number-M58598) (Normand *et al.*, 1992b) using the ALIGN program (version 2.0). The forward primer FGPS 989ac corresponded to the nucleotide positions 989-1005 of

Figure 4.10

Fig 4.10 : Restriction digestion of cloned plasmid pUC18 with partial 16S rRNA gene insert using *Eco* RI/*Hin*d III.
Lane 1 - Anarh; lane 2 - Aggtun; lane 3 - Anhpk; lane 4 - Anmus; lane 5 - Anob; lane 6 - Anmnh; lane 7 - ACN1^{AG} and lane 8 - 100 bp molecular base pair ladder.

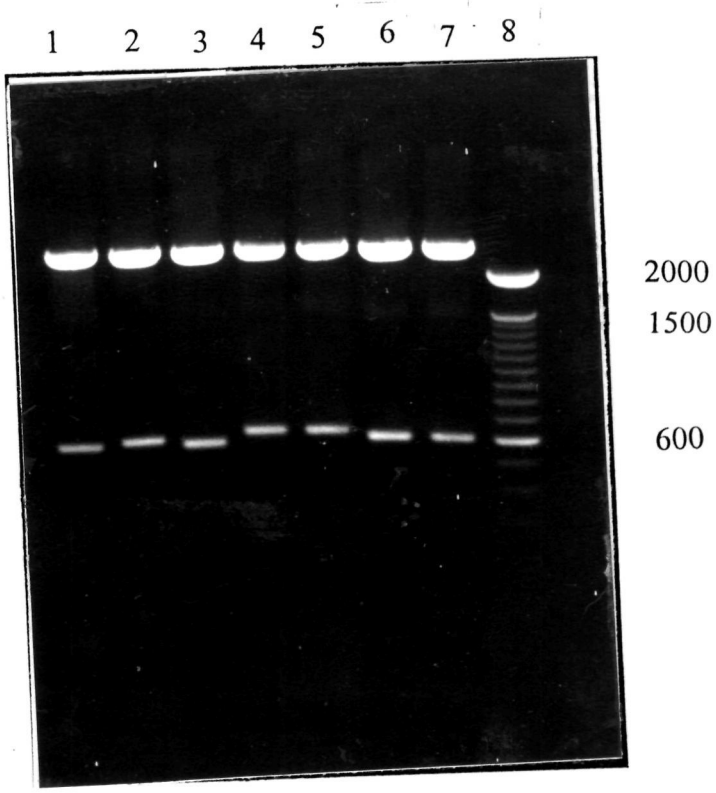

1
2
3
4
5
6
7
8
2000
1500
600

Fig 4.11 : Restriction digestion of cloned plasmid pUC18 with 16S-23S rRNA ITS region as insert using *Eco* RI/*Hin*d III.

Lane 1 and 9 - 100 bp molecular size marker;
lane 2 - Anarh; lane 3 - Aggtun; lane 4 - Anhpk;
lane 5 - Anob; lane 6 - Anmus; lane 7 - Anmnh;
lane 8 - ACN1^{AG}.

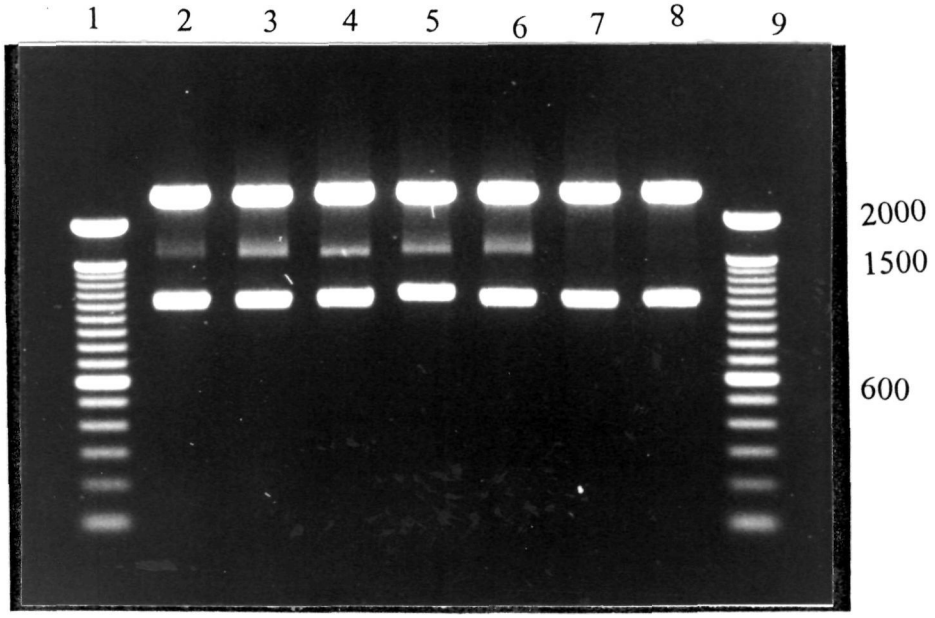

1
2
3
4
5
6
7
8
9
2000
1500
600

Table 4.2: Oligonucleotide primers used for sequencing

Primers	Sequence
M13 F	5'-GATGGATATCTTGGTTC-3'
M13R	5'-CAGGAAACAGCTATGAC-3'
FGPS 989ac	5'-GGGGTCCGTAAGGGTC-3'
RFRITF3	5'-CCGTCACGTCACGAAAGTCG-3'
RFRITR2	5'-GCTCGCGTCCACTGTGCAG-3'
RNIFF2	5'-GCATCCCTTCCGGCAGATGCA-3'

16S rRNA gene of the *Frankia* strain ORS020606. The reverse primer FGPL 1509' corresponded to the positions 1490-1509 of 16S rRNA gene of the *Frankia* strain ORS020606. The sequence alignment revealed high sequence conservation with minor substitutional variations at a few positions (fig 4.12). At position 1102 corresponding to 16S rRNA gene of ORS020606, base substitution from 'C' to 'A' (transversion; fig 4.13.2) occurred in all the samples except those collected from Upper Shillong and Tuebingen, where base 'A' was substituted by the base 'T'. The partial sequences obtained from these two locations were identical to each other. Similarly, at positions 1103, substitution of base 'G' to 'A' occurred which was common in all the cases. At position 1228, in ACN 1^{AG} strain, base 'T' was substituted by base 'C' which was not common to any other sample studied. At positions 1229 and 1233 of Aggtun and Anmus, the bases 'G' and 'T' were substituted by bases 'A' and 'C' respectively. At positions 1241, samples collected from Himachal Pradesh showed substitution of 'A' to 'G', which was unique to it. In case of samples collected from Nonkrem Hills, two bases were found substituted which were unique to it. At positions 1255 of Anmnh, base 'A' was substituted by base 'G' and at base 1283, 'C' was substituted by 'T'. Also, at position 1425 corresponding to 16S rRNA of ORS020606 strain, all the samples, excepting Anmus and Aggtun, showed base substitutions from 'G' to 'A' (fig 4.13.1). In general, substitutions were more of transition than the transversion type.

4.6.2 16S–23S rDNA ITS sequence analysis

Analysis of the gel purified PCR 16S-23S rDNA ITS products revealed minor size variations which were not significant being restricted only to a few base pairs. Therefore, they were not detected by gel electrophoresis. The minor size variations were due to either deletion or insertion of bases in the ITS region.

Fig 4.12 : Nucleotide sequence of 16S-23S rRNA ITS region from *Alnus* nodule microsymbionts collected from different geographical locations aligned with the corresponding sequences of the *Casuarina* strain ORS020606. The junction between 16S rRNA end and beginning of 16S-23S rRNA ITS region is marked by an asterisk. Similarly, the junction between the ITS end and beginning of 23S rRNA is marked by a dot above the sequence. The highlighted nucleotide region (bold) between 1503-1508 is the complementary binding site for the shine Dalgarno sequence.

```
ORS020606       ( 989)  GGGGTCCGTAAGGGTCCTGCACAGGTGGTGCATGGCTGTCGTCAGCTCGT
ACN1AG          (   1)  ................................................
Anarh           (   1)  ................................................
Anhpk           (   1)  ................................................
Anmnh           (   1)  ................................................
Anob            (   1)  ................................................
Aggtun          (   1)  ................................................
Anmus           (   1)  ................................................

ORS020606       (1039)  GTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTCGTCCTAT
ACN1AG          (  51)  ................................................
Anarh           (  51)  ................................................
Anhpk           (  51)  ................................................
Anmnh           (  51)  ................................................
Anob            (  51)  ................................................
Aggtun          (  51)  ................................................
Anmus           (  51)  ................................................

ORS020606       (1089)  GTTGCCAGCGAGTCGTGTCGGGGACTCATAGGAGACTGCCGGGGTCAACT
ACN1AG          ( 101)  ............AA....................................
Anarh           ( 101)  ............AA....................................
Anhpk           ( 101)  ............AA....................................
Anmnh           ( 101)  ............AA....................................
Anob            ( 101)  ............AA....................................
Aggtun          ( 101)  ............TA....................................
Anmus           ( 101)  ............TA....................................

ORS020606       (1139)  CGGAGGAAGGTGGGGATGACGTCAAGTCATCATGCCCCTTACGTCCTGGG
ACN1AG          ( 151)  ................................................
Anarh           ( 151)  ................................................
Anhpk           ( 151)  ................................................
Anmnh           ( 151)  ................................................
Anob            ( 151)  ................................................
Aggtun          ( 151)  ................................................
Anmus           ( 151)  ................................................

ORS020606       (1189)  CTGCACACATGCTACAATGGCCGGTACAAAGGGCTGCGATGCCGTGAGGT
ACN1AG          ( 201)  ....................................C...........
Anarh           ( 201)  ................................................
Anhpk           ( 201)  ................................................
Anmnh           ( 201)  ................................................
Anob            ( 201)  ................................................
Aggtun          ( 201)  ....................................A...C.....
Anmus           ( 201)  ....................................A...C.....

ORS020606       (1239)  GGAGCGAATCCCAAAAAGCCGGTCTCAGTTCGGATCGGGGTCTGCAACTC
ACN1AG          ( 251)  ................................................
Anarh           ( 251)  ................................................
Anhpk           ( 251)  ..G.............................................
Anmnh           ( 251)  ................G.......................T.....
Anob            ( 251)  ................................................
Aggtun          ( 251)  ................................................
Anmus           ( 251)  ................................................
```

(contd...)

```
ORS020606       (1289)  GACCCCGTGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAATGCTGCGG
ACN1AG          ( 301)  .....................................................
Anarh           ( 301)  .....................................................
Anhpk           ( 301)  .....................................................
Anmnh           ( 301)  .....................................................
Anob            ( 301)  .....................................................
Aggtun          ( 301)  .....................................................
Anmus           ( 301)  .....................................................

ORS020606       (1339)  TGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGAAAGTC
ACN1AG          ( 351)  .....................................................
Anarh           ( 351)  .....................................................
Anhpk           ( 351)  .....................................................
Anmnh           ( 351)  .....................................................
Anob            ( 351)  .....................................................
Aggtun          ( 351)  .....................................................
Anmus           ( 351)  .....................................................

ORS020606       (1389)  GGTAACACCCGAAGCCGGTGGCCTAACCCTTGTGGGGGGAGCCGTCGAAG
ACN1AG          ( 401)  ..................................A..................
Anarh           ( 401)  ..................................A..................
Anhpk           ( 401)  ..................................A..................
Anmnh           ( 401)  ..................................A..................
Anob            ( 401)  ..................................A..................
Aggtun          ( 401)  .....................................................
Anmus           ( 401)  .....................................................

ORS020606       (1439)  GTGGGACCGGCGATTGGGACGAAGTCGTAACAAGGTAGCCGTACCGGAAG
ACN1AG          ( 451)  .....................................................
Anarh           ( 451)  .....................................................
Anhpk           ( 451)  .....................................................
Anmnh           ( 451)  .....................................................
Anob            ( 451)  .....................................................
Aggtun          ( 451)  .....................................................
Anmus           ( 451)  .....................................................

                                          *
ORS020606       (1489)  GTGCGGCTGGATCACCTCCTTTCTAAGGAGCGTCTGGCTGGTCTGTCC-T
ACN1AG          ( 501)  ............C....................................C.
Anarh           ( 501)  ............C....................................C.
Anhpk           ( 501)  ............C..................................C....-C
Anmnh           ( 501)  ............C....................................C.
Anob            ( 501)  ............C....................................C.
Aggtun          ( 501)  ............C.....................................-.
Anmus           ( 501)  ............C.....................................-.

ORS020606       (1538)  GTAGTGGGGTGGGCTGGTGCAGAGCCAGGGCCGGCTGTGGATGCCGGTCT
ACN1AG          ( 551)  ..T.G..........C...G.....A....A....T.G.............
Anarh           ( 551)  ..T.G..........C...G.....A....A....T.G.............
Anhpk           ( 550)  ..T.-..........C...G.........A....T.G.............
Anmnh           ( 551)  ..T.G..........C...G.....A....A....T.G.............
Anob            ( 551)  ..T.G..........C...G.....A....A....T.G.............
Aggtun          ( 550)  ..T.-..........T...G.........A....C.................
Anmus           ( 550)  ..T.-..........T...G.........A....C.................
```

(contd...)

```
ORS020606       (1588)   GGTTGCTCGTGGGTG--GAACGCTGACGATGGGGTTTC-GGCTGGTT--G
ACN1AG          ( 601)   .........A.....TG....A....T....T.-----.-T..C....CT.
Anarh           ( 601)   .........A.....TG....A....T....T.-----.-T..C....CT.
Anhpk           ( 599)   .........A.....T-.G...........T..C.G.T.....-------.
Anmnh           ( 601)   .........A.....TG....A....T....T.-----.-T..C....CT.
Anob            ( 601)   .........A.....TG....A....T....T.-----.-T..C....CT.
Aggtun          ( 599)   ........A.....--............A...C.G.T..T...G.--C
Anmus           ( 599)   ........A.....--............A.---.G.-T.........-.

ORS020606       (1633)   TCCGGGG----TCTAGTACTCCTCTGTCCTTATCTTCGGGTGGGGTGG--
ACN1AG          ( 647)   ..T..CT----.........................G.CT...T.------
Anarh           ( 647)   ..T..CT----.........................G.CT...T.------
Anhpk           ( 644)   ..T..T.----C.....................------C.T.T.G.T--
Anmnh           ( 647)   ..T..CT----.........................G.CT...T.------
Anob            ( 647)   ..T..CT----.........................G.CT...T.------
Aggtun          ( 645)   .G.T...CG--.....---.....GG..G..C.TTC.G.......CT
Anmus           ( 641)   ..T.CT.GATG...........-G...GGC.TC.......G...T.----

ORS020606       (1677)   G-GGGTGG-AACGGTCCGGGTGGGTGGCTGG-GGT---CCGTGGCACGCT
ACN1AG          ( 677)   .-.....-.....---..T.......GGC.-...GGT.......G....
Anarh           ( 677)   .-.....-.....---..T.......GGC.-...GGT.......G....
Anhpk           ( 673)   .-.....-.....GG.T..GT..G.TTG.T-.....---.............
Anmnh           ( 677)   .-.....-.....---..T.......GGC.-...GGT.......G....
Anob            ( 677)   .-.....-.....---..T.......GGC.-...GGT.......G....
Aggtun          ( 690)   .-T..GT.-G....G.TT......G......T.....---T...........
Anmus           ( 688)   .C......C.....GT.T.............T.....---G...........

ORS020606       (1721)   -GTTGGGTCCTGAGGGAGT---------------GAGGCTC-TCTCGG--
ACN1AG          ( 721)   -.....................------------C....CGG.G.G.T--
Anarh           ( 721)   -.....................------------C....CGG.G.G.T--
Anhpk           ( 717)   -..................CAGGCCGGTGTGGTC......T-...T.ATC
Anmnh           ( 721)   -.....................------------C....CGG.G.G.T--
Anob            ( 721)   -.....................------------C....CGG.G.G.T--
Aggtun          ( 735)   A.....................-----------------.....T-C....T--
Anmus           ( 735)   -.....................-----------------.....T-C....T--

ORS020606       (1752)   -----CTGGTGGGTT-CC--TGTGCTG-----TGGTCCTG-GTGGGGATC
ACN1AG          ( 753)   ----.........-..-T.-.-..A.C-----C.C.T.....-.....TGC.
Anarh           ( 753)   ----.........-..-T.-.-..A.C-----C.C.T.....-.....TGC.
Anhpk           ( 765)   TGCTT..TA.....G-..-----C.........CC.....T.........---
Anmnh           ( 753)   ----.........-..-T.-.-..A.C-----C.C.T.....-.....TGC.
Anob            ( 753)   ----.........-..-T.-.-..A.C-----C.C.T.....-.....TGC.
Aggtun          ( 767)   ----.......T..GA..--CC.T...CATGA.CC.............G.T
Anmus           ( 766)   ----.......T.GC-..CT.C.....-----ATC....T-CG....G.T

ORS020606       (1788)    ----------GTGGTGG--G-GGG----TCTGCT------GGTGGAACG
ACN1AG          ( 785)    CCCTGTCCTCT....A..--.T.....-----GG..T.CGCGAG.....TG..
Anarh           ( 785)    CCTGTCCTCT-....A..--.T.....-----GG..T.CGCGAG.....TG..
Anhpk           ( 801)    ----------.....-.-.TGTTC.G..AG------.....TG..
Anmnh           ( 785)    CCTGTCCTCT-....A..--.T.....----GG..T.CGCGAG.....TG..
Anob            ( 785)    CCTGTCCTCT-....A..--.T.....-----GG..T.CGCGAG.....TG..
Aggtun          ( 809)    ----------.......TG.-.....-----.....-----...-..G..
Anmus           ( 804)    ----------.....-.--.-...----.....-----...-.....
```

(contd...)

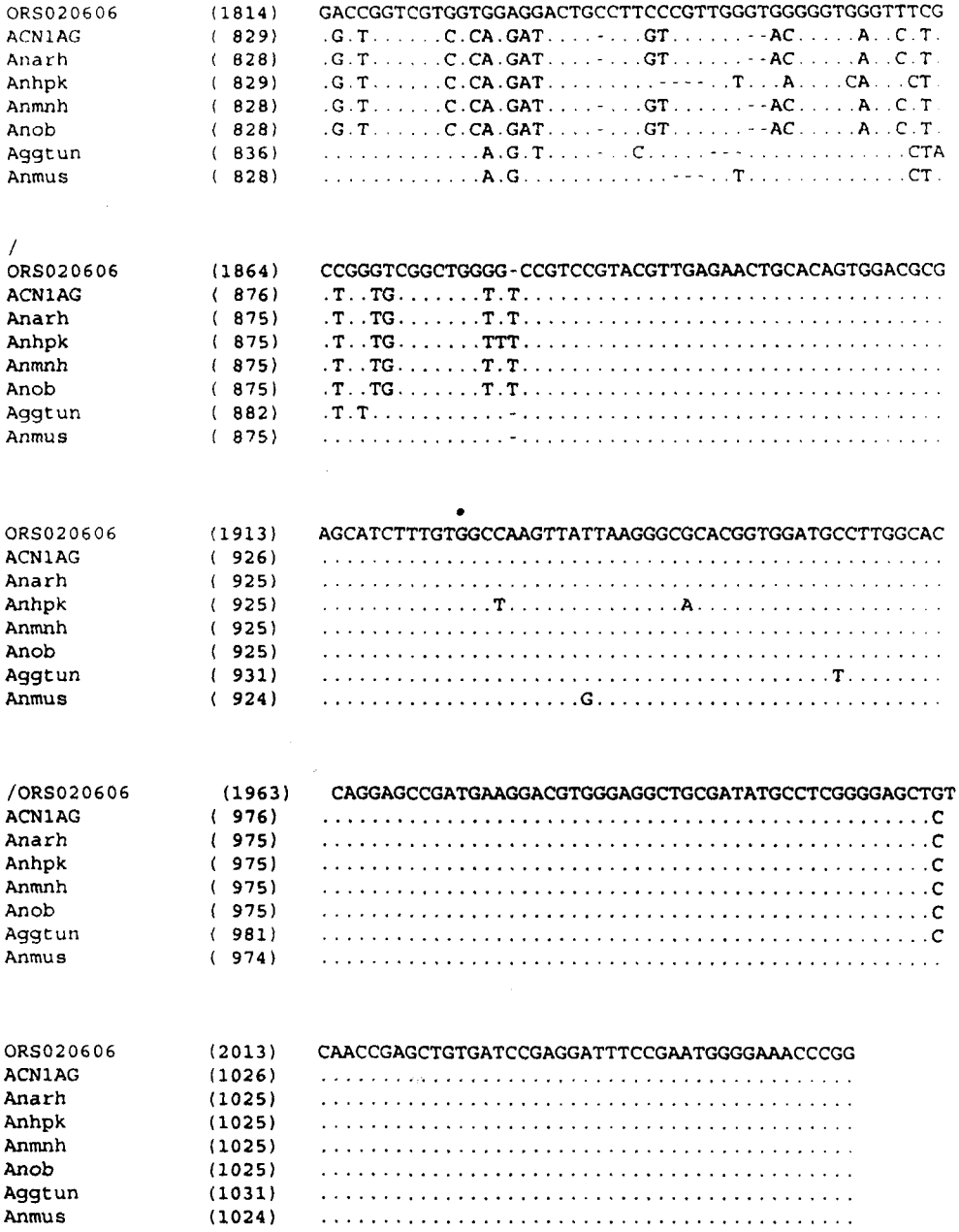

ORS020606 (1814) GACCGGTCGTGGTGGAGGACTGCCTTCCCGTTGGGTGGGGGTGGGTTTCG
ACN1AG (829) .G.T.....C.CA.GAT.....GT.....-AC.....A..C.T.
Anarh (828) .G.T.....C.CA.GAT.....GT.....-AC.....A..C.T.
Anhpk (829) .G.T.....C.CA.GAT.....T...A....CA...CT.
Anmnh (828) .G.T.....C.CA.GAT.....GT.....-AC.....A..C.T.
Anob (828) .G.T.....C.CA.GAT.....GT.....-AC.....A..C.T.
Aggtun (836) .....A.G.T.....C.....T.....CTA
Anmus (828) .....A.G.....T.....CT.
/
ORS020606 (1864) CCGGGTCGGCTGGGG-CCGTCCGTACGTTGAGAACTGCACAGTGGACGCG
ACN1AG (876) .T..TG.....T.T.....
Anarh (875) .T..TG.....T.T.....
Anhpk (875) .T..TG.....TTT.....
Anmnh (875) .T..TG.....T.T.....
Anob (875) .T..TG.....T.T.....
Aggtun (882) .T.T.....-.....
Anmus (875) .....-.....
ORS020606 (1913) AGCATCTTTGTGGCCAAGTTATTAAGGGCGCACGGTGGATGCCTTGGCAC
ACN1AG (926) .....
Anarh (925) .....
Anhpk (925) .....T.....A.....
Anmnh (925) .....
Anob (925) .....
Aggtun (931) .....T.....
Anmus (924) .....G.....
/ORS020606 (1963) CAGGAGCCGATGAAGGACGTGGGAGGCTGCGATATGCCTCGGGGAGCTGT
ACN1AG (976) C
Anarh (975) C
Anhpk (975) C
Anmnh (975) C
Anob (975) C
Aggtun (981) C
Anmus (974) .....
ORS020606 (2013) CAACCGAGCTGTGATCCGAGGATTTCCGAATGGGGAAACCCGG
ACN1AG (1026) .....
Anarh (1025) .....
Anhpk (1025) .....
Anmnh (1025) .....
Anob (1025) .....
Aggtun (1031) .....
Anmus (1024) .....

Fig 4.13.1 : **Autoradiogram showing a portion of 16S rDNA gene in the 3' region between 1418-1456 bp corresponding to the reference *Frankia* species ORS020606. At nucleotide base position 1425, the base 'A' is substituted by 'G' in the samples Anmus and Aggtun.**

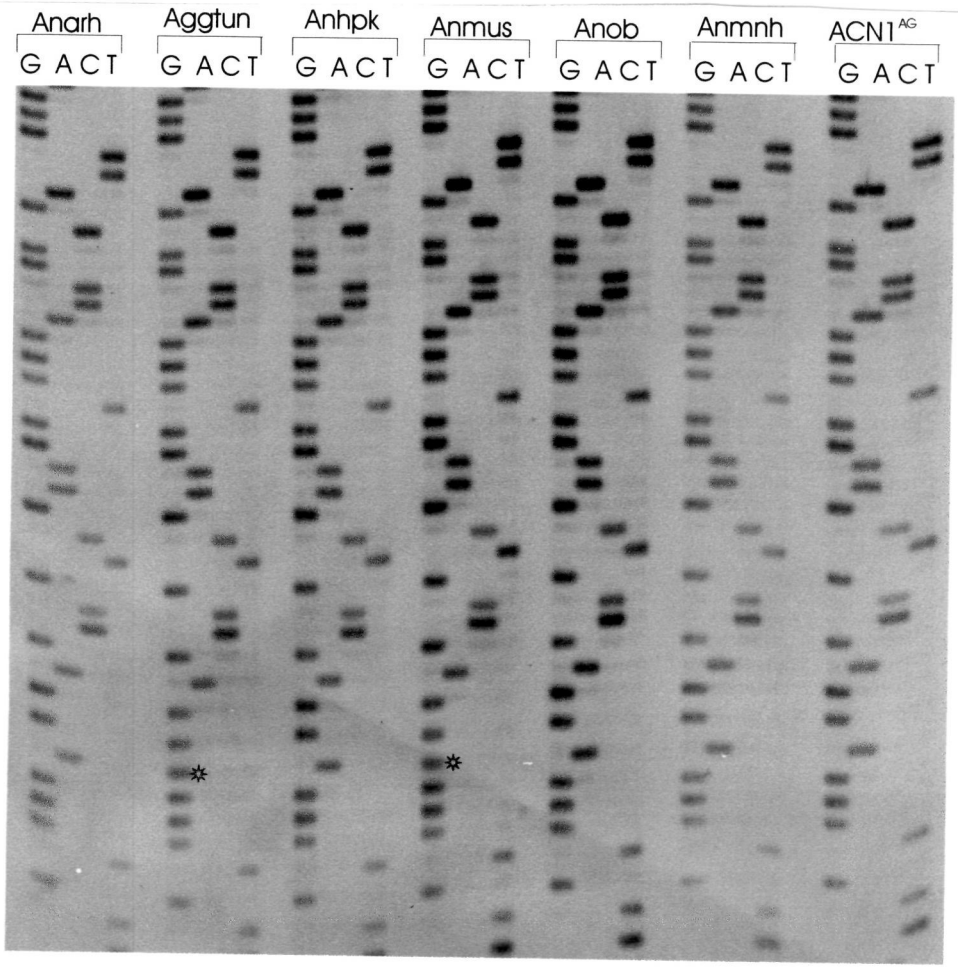
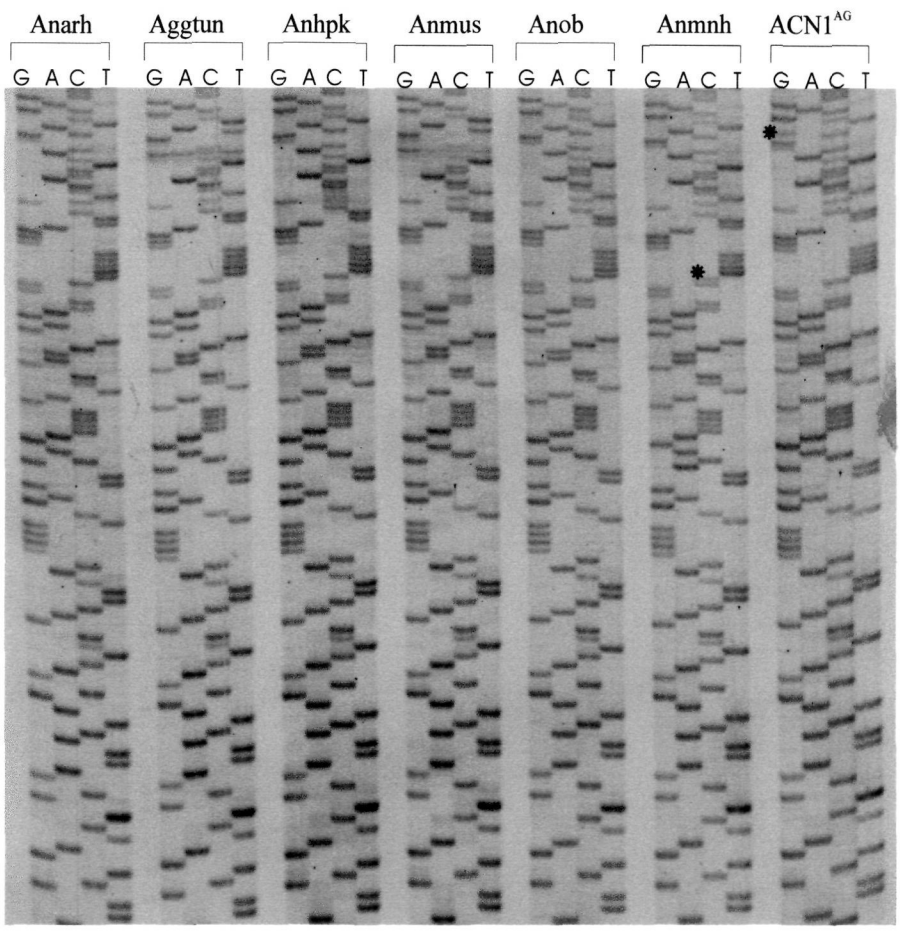

Anarh
G A C T
Aggtun
G A C T
Anhpk
G A C T
Anmus
G A C T
Anob
G A C T
Anmnh
G A C T
ACN1 AG
G A C T

Fig 4.13.2 : Autoradiogram of a sequencing gel showing the conserved nature of 16S rDNA gene. The sequences shown correspond to the nucleotide base sequences between 1239-1372 of the reference *Frankia* strain ArI3 rDNA. The asterisk marks the base substitution incase of Anmnh and ACN1^{AG} samples.


Anarh
Aggtun
Anhpk
Anmus
Anob
Anmnh
ACN1AG
G A C T
G A C T
G A C T
G A C T
G A C T
G A C T
G A C T

Cloning and sequence analysis of the PCR products revealed that the sequence variations among the samples were considerably more in the non-coding internal transcribed spacer region than in the coding region of the 16S and the 23S rRNA gene (fig 4.14; 4.15). Also, analysis of the 520 base pairs from the 3' region of the 16S rRNA gene from seven frankiae samples belonging to 3 different *Alnus* species showed no insertion or deletion leading to size variation. But, in the case of the internal transcribed spacer region, size variation was observed .

Sample collected from Upper Shillong, Meghalaya, had the shortest (413 bp) and that collected from Tuebingen, Germany, had the longest (420 bp) ITS. The Upper Shillong sample was very distinct from the rest of the samples collected from India in being very close to Tuebingen sample. Further, the nucleotide sequences between 1650-1661 (fig 4.12) corresponding to the reference strain ORS020606 were present in Anmus and Aggtun samples, but were found absent between 1652-1666 in Anhpk and between 1650-1661 in ACN1^{AG}, Anmnh, Anarh and Anob. The sample collected from Himachal Pradesh was also unique in having the ITS region which was different with respect to others. For example, the regions between 1739-1740 and 1751-1752 incorporated 15 bp and 7 bp fragments respectively, which were absent in others. Similarly, a 16 bp fragment was found inserted between positions 2310-2315 in Anarh, ACN1^{AG}, Anmnh and Anob. This insertion was, however, seen absent in the samples Aggtun and Anmus. Variation for a base or two in the partial gene sequence of 23S rRNA adjacent to the ITS region was also detected. In this case, at base positions 4 and 19 of 23S rRNA of Anhpk sample the bases 'C' and 'G' were substituted by 'T' and 'A' respectively. Similarly, at position 11, the base 'T' was substituted by 'G' for the Upper Shillong sample Anmus. Also, at position 31 of 23S rRNA, the base 'C' was substituted by the base 'T' for the samples (Aggtun) collected from Tuebingen. Sequence analysis of the ITS region revealed that, both Anmus and Aggtun samples

Figure 4.14 : Autoradiogram of a sequencing gel showing variations in the 16S-23S rDNA ITS region. Sequencing was carried out using cloned PCR products of nodule microsymbionts collected from seven different locations. The astersik marks some of the variations.

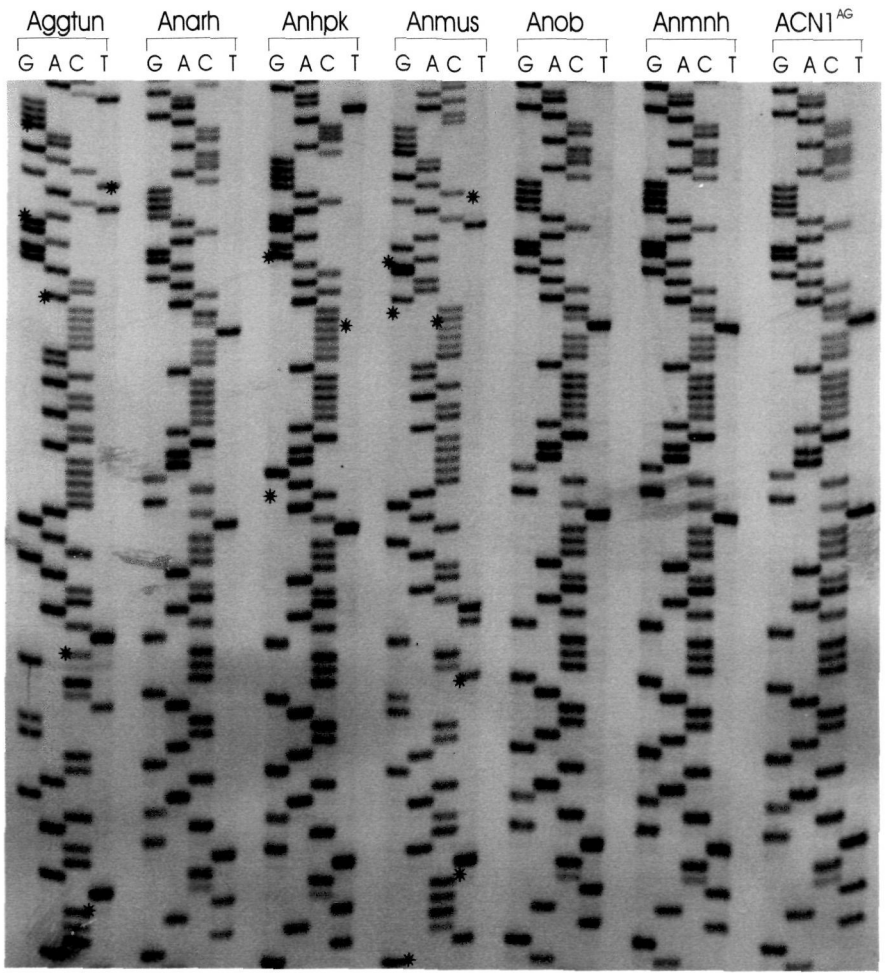
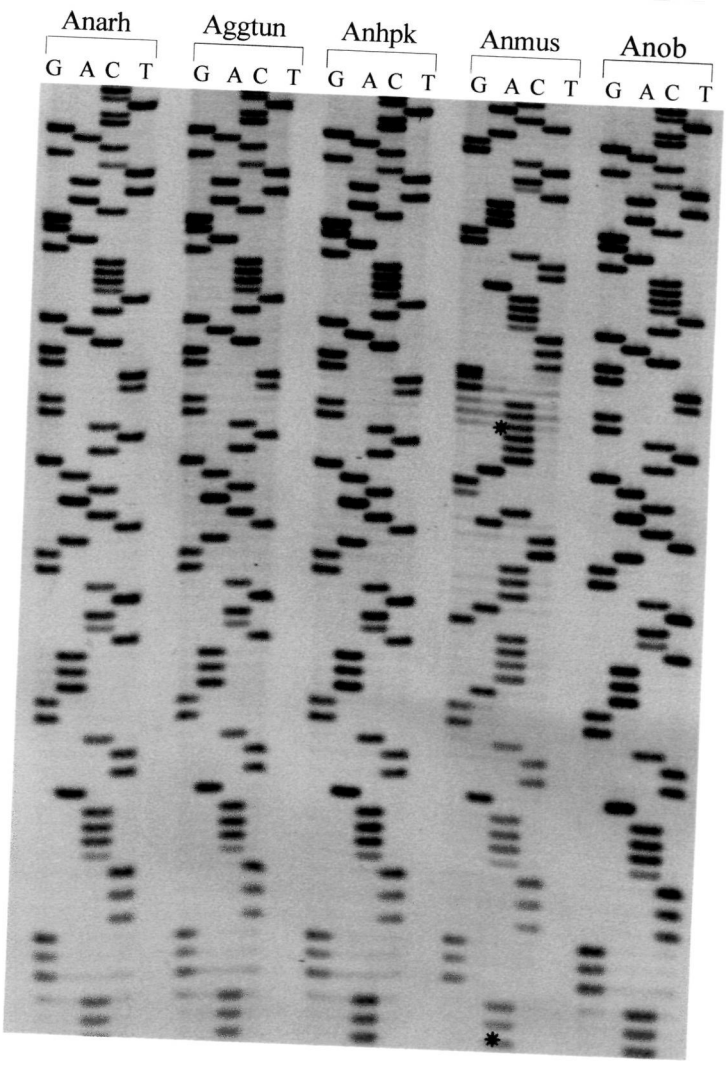

Aggtun
Anarh
Anhpk
Anmus
Anob
Anmnh
ACN1AG
G A C T
G A C T
G A C T
G A C T
G A C T
G A C T
G A C T

Fig 4.15 : Autoradiogram of a sequencing gel showing showing dimer formation. The sequences shown correspond to 5' region of the 23S rDNA. The asterisk marks the sequence forming primer dimer.


Anarh
G A C T
Aggtun
G A C T
Anhpk
G A C T
Anmus
G A C T
Anob
G A C T

differed significantly from each other not only with respect to size differences but also with respect to sequence variation.

Overall, the partial sequences of 16S, 23S and the 16S-23S rRNA ITS together showed significant differences among the nodule microsymbiont studied, revealing six genotypes. However, when 16S-23S rRNA ITS alone was taken into consideration, only four genotypes resulted which showed that this region is also to some extent conserved among closely related species.

4.7 Sequence analysis of *nif* D-K intergenic spacer region

The PCR amplified products were cloned and sequenced using an automatic cycle sequencer. The cloned recombinant plasmids were checked for the desired fragments prior to sequencing with *Eco* RI/*Hind* III restriction endonucleases (fig 4.16). Universal primers M13 forward, M13 reverse and specific primer RNIFF2 were used for accomplishing sequencing. The resulting nucleotide sequence obtained as electropherogram (fig 4.17) was edited manually. Sequence analysis of the various samples revealed size variation which exceeded not more than a few base pairs. Sequencing of partial *nif* D and partial *nif* K gene was also carried out by radioactive method mentioned earlier.

The sequences were entered into DNASIS program in computer for analysis and were aligned with the known sequences retrieved from the GenBank. As a reference sequence, *Alnus* infective *Frankia alni* strain ArI3 (FRANIFV; Accession number: L41344) was selected. The aligned sample sequence available for Anarh and Aggtun, corresponded to the nucleotide sequence between 819 to 2058 bp (that is, extending from 819 bp of *nif* D to 339 bp of *nif* K) of ArI3. The presence of the triplet codon TGA at the end of the *nif* D gene marked the termination codon for the *nif* D gene and the beginning of *nif* K gene was marked by the presence of GTG triplet codon (fig 4.18). The triplet codon GTG, while acting as the initiation codon,

Fig 4.16 : **Restriction digestion of recombinant plasmid DNA with *nif* D-K IGS insert using *Eco* RI/*Hin*d III endonuclease. Lane 1 - Anarh; lane 2 - Aggtun; lane 3 - Anmnh; and lane 4- λ DNA *Eco* RI/*Hin*d III double digest (in kbp).**

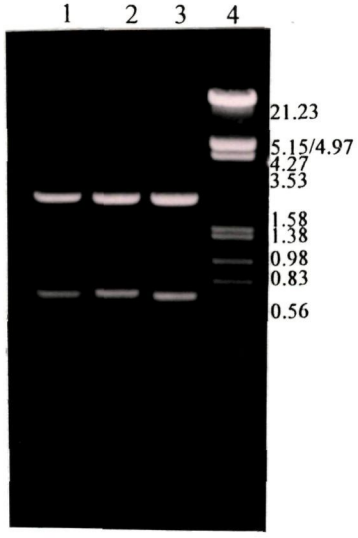

1
2
3
4
21.23
5.15/4.97
4.27
3.53
1.58
1.38
0.98
0.83
0.56

Fig 4.17 : Electropherogram showing partial sequences of *nif* D, *nif* K and *nif* D-K IGS region. The junction between the coding (*nif* D) and the non-coding (IGS) region for Anarh sample (1a) and Aggtun sample (1b) is shown revealing variations (asterisk indicate the termination codon UGA). Similarly, the junction between the IGS and *nif* K region is shown for Anarh (2a) and Aggtun (2b) samples.

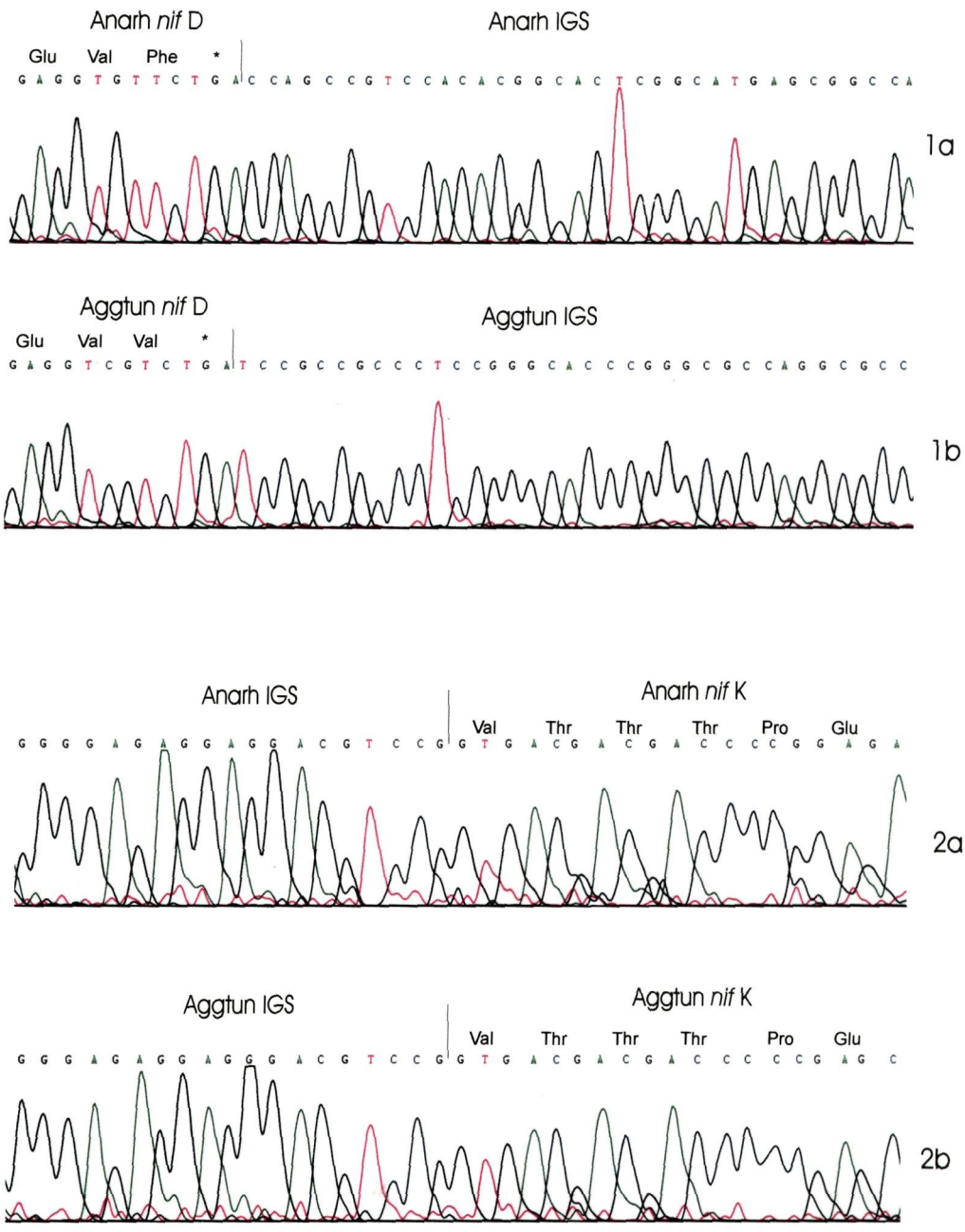
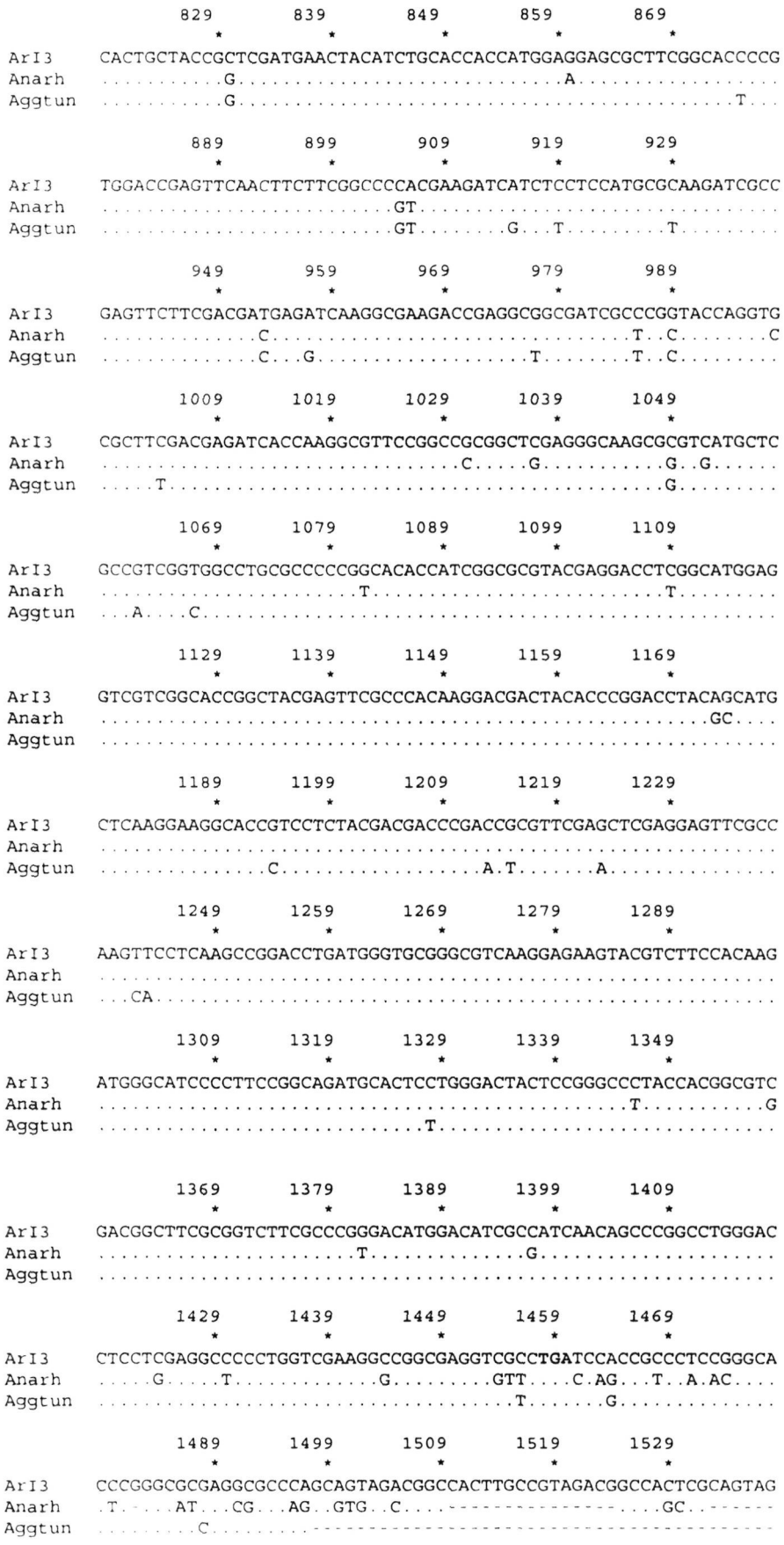
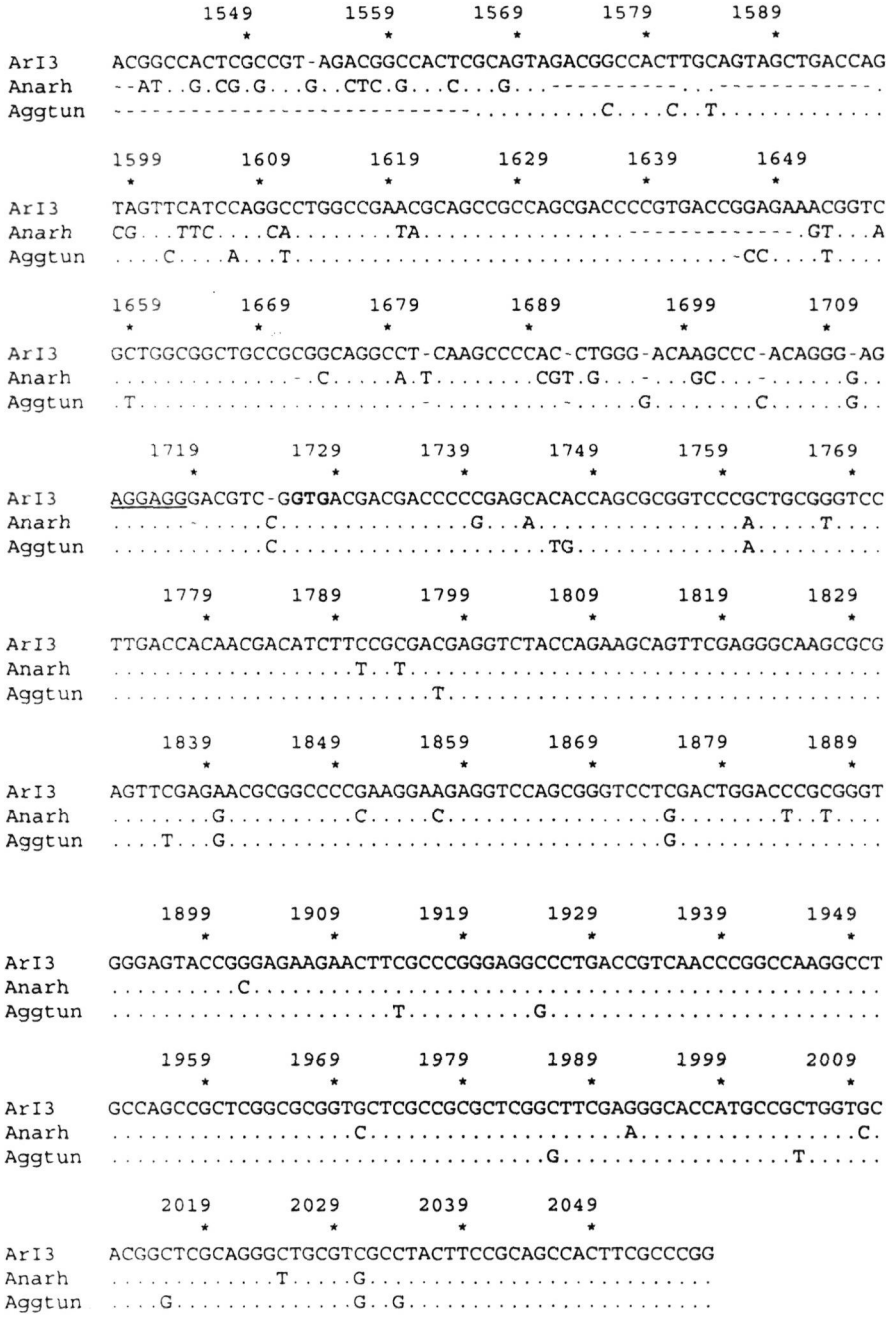

Anarh *nif* D
Anarh IGS
Glu Val Phe *
1a
Aggtun *nif* D
Aggtun IGS
Glu Val Val *
1b
Anarh IGS
Anarh *nif* K
Val Thr Thr Thr Pro Glu
2a
Aggtun IGS
Aggtun *nif* K
Val Thr Thr Thr Pro Glu
2b

Fig 4.18 : Nucleotide sequence alignment of *nif* D-K IGS region from nodule microsymbionts Anarh and Aggtun with the reference *Frankia* strain ArI3, a *Alnus rubra* isolate (accession number: L41344; specq and Normand, 1996). The codon TGA (bold) at coordinate 1460 of the reference strain ArI3 represents the end codon for the *nif* D gene. The codon GUG (bold) at coordinate 1728 of the reference strain ArI3 represents the initiation codon for the *nif* K gene. The underlined nucleotide sequences 6/7 bases upstream of the initiation codon GTG is the Shine - Dalgarno sequence (Shine and Dalgarno, 1974).


829 839 849 859 869
ArI3 CACTGCTACCGCTCGATGAACTACATCTGCACCACCATGGAGGAGCGCTTCGGCACCCCG
Anarh .....G.....A.....
Aggtun .....G.....T...
889 899 909 919 929
ArI3 TGGACCGAGTTCAACTTCTTCGGCCCCACGAAGATCATCTCCTCCATGCGCAAGATCGCC
Anarh .....GT.....
Aggtun .....GT.....G...T.....T.....
949 959 969 979 989
ArI3 GAGTTCTTCGACGATGAGATCAAGGCGAAGACCGAGGCGGCGATCGCCCGGTACCAGGTG
Anarh .....C.....T..C.....C
Aggtun .....C...G.....T.....T..C.....
1009 1019 1029 1039 1049
ArI3 CGCTTCGACGAGATCACCAAGGCGTTCCGGCCGCGGCTCGAGGGCAAGCGCGTCATGCTC
Anarh .....C.....G.....G..G.....
Aggtun .....T.....G.....
1069 1079 1089 1099 1109
ArI3 GCCGTCGGTGGCCTGCGCCCCCGGCACACCATCGGCGCGTACGAGGACCTCGGCATGGAG
Anarh .....T.....T.....
Aggtun ...A....C.....
1129 1139 1149 1159 1169
ArI3 GTCGTCGGCACCGGCTACGAGTTCGCCCACAAGGACGACTACACCCGGACCTACAGCATG
Anarh .....GC....
Aggtun .....
1189 1199 1209 1219 1229
ArI3 CTCAAGGAAGGCACCGTCCTCTACGACGACCCGACCGCGTTCGAGCTCGAGGAGTTCGCC
Anarh .....
Aggtun .....C.....A.T.....A.....
1249 1259 1269 1279 1289
ArI3 AAGTTCCTCAAGCCGGACCTGATGGGTGCGGGCGTCAAGGAGAAGTACGTCTTCCACAAG
Anarh .....
Aggtun ...CA.....
1309 1319 1329 1339 1349
ArI3 ATGGGCATCCCCTTCCGGCAGATGCACTCCTGGGACTACTCCGGGCCCTACCACGGCGTC
Anarh .....T.....G
Aggtun .....T.....
1369 1379 1389 1399 1409
ArI3 GACGGCTTCGCGGTCTTCGCCCGGGACATGGACATCGCCATCAACAGCCCGGCCTGGGAC
Anarh .....T.....G.....
Aggtun .....
1429 1439 1449 1459 1469
ArI3 CTCCTCGAGGCCCCCTGGTCGAAGGCCGGCGAGGTCGCCTGATCCACCGCCCTCCGGGCA
Anarh .....G.....T.....G.....GTT....C.AG...T..A.AC....
Aggtun .....T.....G.....
1489 1499 1509 1519 1529
ArI3 CCCGGGCGCGAGGCGCCCAGCAGTAGACGGCCACTTGCCGTAGACGGCCACTCGCAGTAG
Anarh .T.....AT...CG...AG..GTG..C.....GC.....
Aggtun .....C.....

Cont'd


1549 1559 1569 1579 1589
ArI3 ACGGCCACTCGCCGT-AGACGGCCACTCGCAGTAGACGGCCACTTGCAGTAGCTGACCAG
Anarh --AT..G.CG.G...G..CTC.G...C...G.....
Aggtun .....C....C..T.....
1599 1609 1619 1629 1639 1649
ArI3 TAGTTCATCCAGGCCTGGCCGAACGCAGCCGCCAGCGACCCCGTGACCGGAGAAACGGTC
Anarh CG...TTC....CA.....TA.....GT...A
Aggtun ....C....A...T.....-CC....T....
1659 1669 1679 1689 1699 1709
ArI3 GCTGGCGGCTGCCGCGGCAGGCCT-CAAGCCCCAC-CTGGG-ACAAGCCC-ACAGGG-AG
Anarh .....-C.....A.T.....CGT.G...-...GC...-.....G..
Aggtun .T.....-.....-.....G.....C.....G..
1719 1729 1739 1749 1759 1769
ArI3 AGGAGGGACGTC-GGTGACGACGACCCCCGAGCACACCAGCGCGGTCCCGCTGCGGGTCC
Anarh .....C.....G...A.....A.....T....
Aggtun .....C.....TG.....A.....
1779 1789 1799 1809 1819 1829
ArI3 TTGACCACAACGACATCTTCCGCGACGAGGTCTACCAGAAGCAGTTCGAGGGCAAGCGCG
Anarh .....T..T.....
Aggtun .....T.....
1839 1849 1859 1869 1879 1889
ArI3 AGTTCGAGAACGCGGCCCCGAAGGAAGAGGTCCAGCGGGTCCTCGACTGGACCCGCGGGT
Anarh .....G.....C.....C.....G.....T..T....
Aggtun ....T...G.....G.....
1899 1909 1919 1929 1939 1949
ArI3 GGGAGTACCGGGAGAAGAACTTCGCCCGGGAGGCCCTGACCGTCAACCCGGCCAAGGCCT
Anarh .....C.....
Aggtun .....T.....G.....
1959 1969 1979 1989 1999 2009
ArI3 GCCAGCCGCTCGGCGCGGTGCTCGCCGCGCTCGGCTTCGAGGGCACCATGCCGCTGGTGC
Anarh .....C.....A.....C.
Aggtun .....G.....T.....
2019 2029 2039 2049
ArI3 ACGGCTCGCAGGGCTGCGTCGCCTACTTCCGCAGCCACTTCGCCCGG
Anarh .....T.....G.....
Aggtun ....G.....G..G.....

codes for the valine tRNA. Normand *et al.* (1988) had observed a potential Shine-Dalgarno sequence about 7 bp upstream before the start codon for *nif* H gene of the *Frankia* strain ArI3. Here too, about 7 bp upstream of the GTG initiation codon, a potential Shine-Dalgarno sequence (5'-AGGAGG-3') was observed. The presence of this sequence further confirmed GTG as the initiation codon for the *nif* K gene.

From the *nif* D-K sequence alignment, considerable amount of sequence heterogeneity was found among the samples (fig 4.18). Further, it was noticed that, although conserved, the coding region incorporated significant number of variations in the form of base substitutions. The extent of variation was, however, lesser in comparison to the IGS regions which incorporated deletions and additions as well. In general, the *nif* D-K IGS size is known to vary among *Frankia* strains (Nalin *et al.*, 1995). Although, only two *nif* D-K IGS sequences were obtained, the length variation was evident, although it extended only for a few base pairs. The Anarh *nif* D-K IGS was longer in size (211 bp) by 9 bp than the Aggtun sample (202 bp). The reason for the presence of high sequence diversity may be attributed to their different geographical location and being microsymbiont of two different species: Anarh being the microsymbiont of the host *Alnus nepalensis* (Arunachal Pradesh, India) and Aggtun, the microsymbiont of the host *Alnus glutinosa* (Tuebingen, Germany). 16S rRNA gene sequence analysis had also revealed similar pattern of variation.

The nucleotide and amino acid sequences are undoubtedly more effective to resolve the variations between the organisms as it gives more characters to analyze. The stringent functional constraints on rRNA molecules render them as the slowest evolving molecules having the potential to record the entire evolutionary history of an organism (Woese, 1987; Clawson *et al.*, 1999). Sequence analysis combined with PCR-RFLP enabled to detect the heterogeneity present within the nodulating

microsymbiont confirming the polymorphism present within the genus, both at inter and intra-specific level. In general, the sequence data were in concurrence with the information obtained with the restriction analysis.

Sequence analysis of partial 16S and 23S rRNA genes (520 bases at 3' rRNA gene and 132 bases at 5' 23S rRNA gene) from the analysed samples confirmed the conserved nature of the ribosomal RNA gene. This is as expected considering the primary and the secondary structure of the rRNAs. Nevertheless, when compared with the available sequence information (*Frankia* strain ORS020606) with the seven samples from *Frankia* nodulating three *Alnus* species, the variation present was considerably higher (11 nt from 520 bases analysed from 16S rRNA and 5 nt from 132 bases from 23S rRNA genes). Considering the ability of *Frankia* to nodulate a wide variety of hosts these variations are not very unexpected. The 3' end of 16S rRNA is the most conserved part of the gene. However, at nucleotide base pair position 1425 corresponding to ORS020606, base substitution from 'G' to 'A' was observed for all the samples except Anmus and Aggtun (fig 4.13.1).

The internally transcribed spacer regions are supposedly under less functional constraints and so susceptible to more random mutations. This makes the region a better candidate to study the variations between closely related organisms. Sequence analysis had shown that there was minor size variation in the 16S-23S rRNA ITS region for the nodule samples analysed. This size variation, however, did not exceed more than a few base pairs. The size of ITS ranged from 413-420 bp which was slightly more than what was observed in the *Frankia* sp. strain ORS020606 (411bp, Normand *et al.*, 1992b).

The sequence analysis of ITS region revealed variations as expected. With one exception (Anmus), the general trend was more towards interspecific level than intraspecific level. Of the seven samples studied, the ITS sequence analysis gave

only four genotypes. Excepting Anmus, the samples analysed from *Alnus nepalensis* (Anob, Anmnh and Anarh) grouped together with *Frankia* isolate ($ACN1^{AG}$). These samples exhibited 102 nt variations which were common to each other. Surprisingly, as against the 102 common variations in the ITS region which grouped them together, 16S rRNA analysis showed variation which differentiated Anmnh (2 nt) and $ACN1^{AG}$ (1 nt) sample.

Although belonging to two different species, Anmus and Aggtun, showed 100% homology limiting for the analysed region of 16S rRNA. But, outside 16S rRNA, it revealed considerable amount of variations in the ITS region (76 nt). In addition to this, two substitutional variations was also seen in 23S rRNA. Furthermore, Anhpk sample, belonging to a separate species, showed variations for bases both at the coding and the non coding regions. Based on the nucleotide information, it stood distinct from the rest. That is, while analysis of the partial 16S rRNA and 23S rRNA showed variations for 1 or 2 nucleotides , the ITS region of this sample, in addition to the other variations such as insertion or deletion, incorporated a long stretch of nucleotides (15 bp). One thing is clear from the observations that the overall substitutional variations were more of energetically favoured transitions rather than the transversions.

Like in the case of the ribosomal RNA genes, the analysed regions of the *nif* genes also showed considerable amount of homology in the coding region. As expected, here too the non-coding intergenic spacer region showed variations like single base substitution, deletion or addition and also variations for long stretches of nucleotides. The comparison of the nucleic acid and the deduced amino acid showed varying degrees of homology in Anarh and Aggtun with respect to the *Frankia* strain ArI3. That is, the samples showed 98% and 95% homology for the analysed partial *nif* D gene and 97% each for the partial *nif* K gene. Although, samples from only two different geographical locations were analysed, it seemed that the

variations in the IGS region was more than that was observed in the ITS sequence. Unlike the ITSs, that are known to have at least some role to play in pre-rRNA processing, the IGS sequences do not have any functional constraints, which makes this region open to more random variations.

4.8 G+C comparison in the coding and the non-coding regions

The G+C nucleotide content is strongly co-related with the optimum growth temperature of prokaryotes (Galtier *et al.*, 1999). From the data analysed, it seemed to be more in case of partial sequences of *nif* D-K regions (ranging between 64-72 %; table 4.3A) than in the partial sequences of rRNA region (ranging between 52-63 %; table 4.3B). It was also noticed that the G+C content was especially high at third codon positions, an observation in line with the finding of Normand and Bousquet (1989) and Hirsch *et al.* (1995). The per cent G+C content seemed to be more in the ITS and the IGS regions than that in the corresponding coding regions (table 4.3A; 4.3B).

Actinorhizal plants are known to grow even under varied environmental stresses. Probably, the ability of *Frankia* to survive even under extreme conditions explains its high G+C content. The overall proportion of guanine and cytosine residues in the DNA base composition of *Frankia* genome is high ranging 68-72 % (An *et al.*, 1983). In fact, this unique feature of high G+C content is used as one of the tools to distinguish it from other non-actinomycete organisms. The total G+C nucleotide content of the 16S and 23S rRNA gene for the *Frankia* sp. strain ORS020606 is 60% and 57% respectively (Normand *et al.*, 1992b). For the nodule samples analysed in the study, however, the G+C content calculated for the 16S-23S rRNA ITS region (63%) seemed higher than that observed for the partial 16S (52%) and 23S (60-61%) rRNA genes. This was also the case with *nif* D-K IGS region which showed higher G+C content (70-72%) than the analysed partial *nif* D

Table-4.3 : Molecular size and per cent G+C content as observed in 16S-23S rRNA and *nif* D-K regions

A 16S-23S rRNA region

Micro-symbiont / Strain	16S- 23S rRNA region					
	16S rRNA (partial)		ITS		23S rRNA (partial)	
	Size (bp)	G+C (%)	Size (bp)	G+C (%)	Size (bp)	G+C (%)
Anarh	520	52.3	415	63.4	132	61.4
Anmus	520	52.5	413	63.9	132	61.3
Anob	520	52.5	414	63.3	132	61.4
Anmnh	520	52.8	414	63.3	132	61.7
Aggtun	520	52.5	420	63.3	132	60.6
ACN1^{AG}	520	52.3	415	63.4	132	61.7

B) *nif* D-K region

micro symbiont / strain	*nif* D-K region					
	nif D (partial)		IGS		*nif* K (partial)	
	Size(bp)	G+C (%)	Size(bp)	G+C (%)	Size(bp)	G+C(%)
Anarh	642	64	211	72.5	332	67.8
Aggtun	642	64	202	70.8	332	68.9

(64%) and *nif* K (67-68%) genes. Probably, the absence of major functional or structural constraints in the ITS or IGS regions allowed more directed mutations favouring higher G+C content as an adaptation to survive in extreme conditions.

Further, as previously observed by Normand *et al.* (1988, 1992a), the codon usage for the partial *nif* D and *nif* K gene sequences was found to be highly skewed in favour of G or C in the third position of degenerate codons. This, however is not unusual for actinomycetes and other micro-organisms with high G+C content (Bibb *et al.*, 1984). This skew resulted in 96% and 97% G+C occurrence at third position for Anarh and Aggtun samples respectively, which is greater than the overall G+C (64%) observed in the analysed *nif* D partial sequences. Similarly, in the case of partial sequences of *nif* K gene (70-72%), the skew resulted in 92% and 96% G+C occurrence in the third codon position of Anarh and Aggtun samples respectively.

Again, the high G+C content and codon usage observed for the partial *nif* D and *nif* K genes from the samples analysed are similar to those of *nif* H in *Frankia* strain ArI3 (Normand *et al.*, 1988; Normand and Bosquet, 1989). This is in concurrence with the codon usage tabulated from GenBank for other *Frankia* strains.

4.9 Partial *nif* D and *nif* K amino acid sequence analysis

The amino acid sequence information were deduced (using the program DNASIS) from the partial DNA sequences of the *nif* D and *nif* K genes. It was aligned with the known amino acid sequences of the reference *Frankia alni* strain ArI3 retrieved from the GenBank (Accession number: L41344). Amino acid sequence analysis of partial *nif* D and *nif* K gene of the samples analysed aligned with the reference sequence revealed differences at few positions (fig 4.19.1; 4.19.2). The samples Anarh and Aggtun showed an amino acid sequence homology of 98% and 95% respectively for the partial *nif* D region. The partial *nif* K amino

Figure 4.19.1
&
Figure 4.19.2

Fig 4.19.1 : Alignment of amino acid sequences for partial *nif* D gene. The sequences are aligned with corresponding amino acid sequences (23-235) of *Frankia* sp ArI3, *Alnus rubra* isolates (specq and Normand, 1996). The asterisk at the end of the sequence marks the termination codon (TGA) for *nif* D gene sequence.

Fig 4.19.2 : Alignment of amino acid sequences for the partial *nif* K gene. The sequences are aligned with the corresponding sequences of *Frankia* sp ArI3, *Alnus rubra* isolate (Specq and Normand, 1996). The initiation codon for both the microsymbionts, Ararh and Aggtun were GUG (also act as the initiation codon) which is recognised by the initiation valine tRNA.

```
                30        40        50        60        70        80
                 *         *         *         *         *         *
ArI3     HCYRSMNYICTTMEERFGTPWTEFNFFGPTKIISSMRKIAEFFDDEIKAKTEAAIARYQ
Anarh    ..............................S..............................
Aggtun   ..............................S..VF............V.............

                90       100       110       120       130       140
                 *         *         *         *         *         *
ArI3     VRFDEITKAFRPRLEGKRVMLAVGGLRPRHTIGAYEDLGMEVVGTGYEFAHKDDYTRTYS
Anarh    ............................................................A
Aggtun   ........................I.....................................

               150       160       170       180       190       200
                 *         *         *         *         *         *
ArI3     MLKEGTVLYDDPTAFELEEFAKFLKPDLMGAGVKEKYVFHKMGIPFRQMHSWDYSGPYHG
Anarh    ............................................................
Aggtun   ......L.....NS........H.....................................

               210       220       230
                 *         *         *
ArI3     VDGFAVFARDMDIAINSPAWDLLEAPWSKAGEVA*
Anarh    ................................F*
Aggtun   ................................V*
```

Fig 4.19.1 : Amino acid sequence alignment for the partial *nif* D gene.

```
                10        20        30        40
                 *         *         *         *
ArI3     MTTTPEHTSAVPLRVLDHNDIFRDEVYQKQFEGKREFENAAPKEEVQRVL
Anarh    V......N..............................D....D.....
Aggtun   V.......A..............................D..........

                60        70        80        90
                 *         *         *         *
ArI3     DWTRGWEYREKNFAREALTVNPAKACQPLGAVLAALGFEGTMPLVHGSQG
Anarh    ..................................................
Aggtun   ................G.................................

               110
                *
ArI3     CVAYFRSHFA
Anarh    ..........
Aggtu    ..........
```

Fig 4.19.2 : Amino acid sequence alignment for the partial *nif* K gene

acid sequences of Anarh and Aggtun showed difference of four amino acids. But, when compared with the reference strain ArI3, they showed homology of 97% each.

4.10 Phylogenetic analysis

4.10.1 Analysis of the nodule microsymbiont based on the ITS Sequences

The internal transcribed spacer region between 16S-23S rRNA was analysed for the nodule microsymbiont isolated. Not much information is available for this specific region from *Frankia* species or any other related organism. The obtained sequences were aligned with the corresponding sequences of ORS020606, a *Casuarina* infective strain. Since, the ITS region stretched to a maximum of 420 nucleotide, a set of 450 characters incorporating the most variable region were taken for the comparative sequence analysis. Analysis was carried out using bothneighbour joining method (Kimura 2 parameter) and the maximum parsimony analysis (fig 4.20; 4.21). The tree formed by neighbour joining showed close relationship between Upper Shillong sample (Anmus) and Tuebingen sample (Aggtun). It was marked by a bootstrap confidence value of 93.0 as was revealed by the partial 16S rRNA analysis. Neighbour joining analysis showed that the sample Anmnh collected also from Meghalaya was very distant from Anmus sample.

Another interesting feature observed in the analyses is the relationship of ORS020606 with both Upper Shillong and Tuebingen samples. Both neighbbour joining and DNA parsimony analysis showed that the Anmus and Aggtun samples were closer to the *Casuarina* strain ORS020606 and Anhpk rather than the other microsymbiont used in the study.

Figure 4.20

Fig 4.20 : DNA parsimony consensus tree derivedfrom 450 nucleotide positions in 16S rDNA of *Alnus* nodule microsymbionts. The numbers at the fork correspond to the bootstrap values out of 100 replicates.

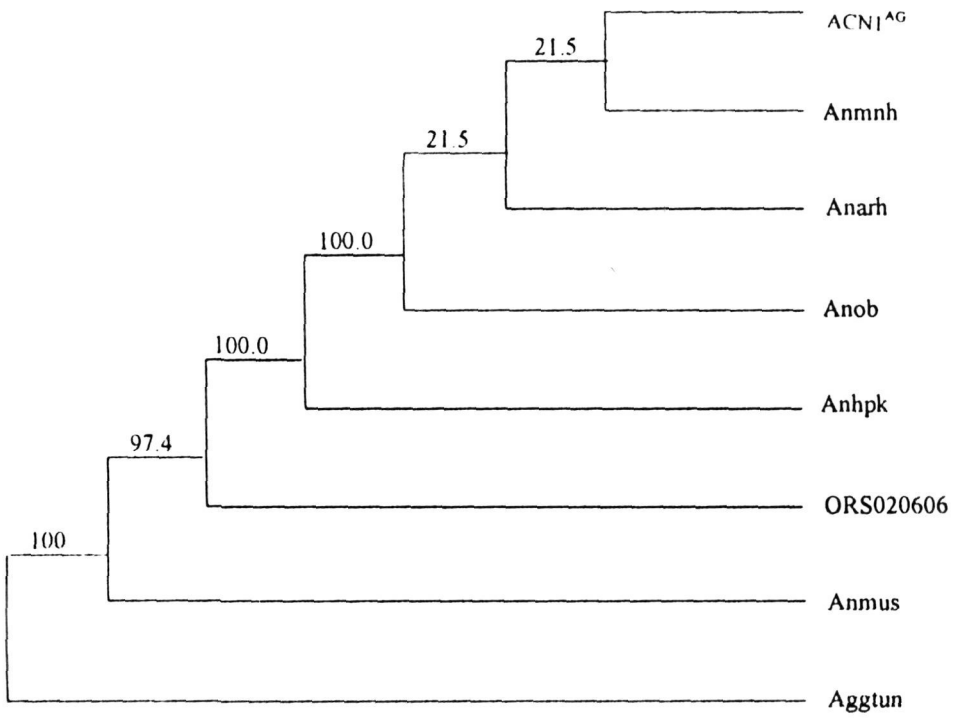
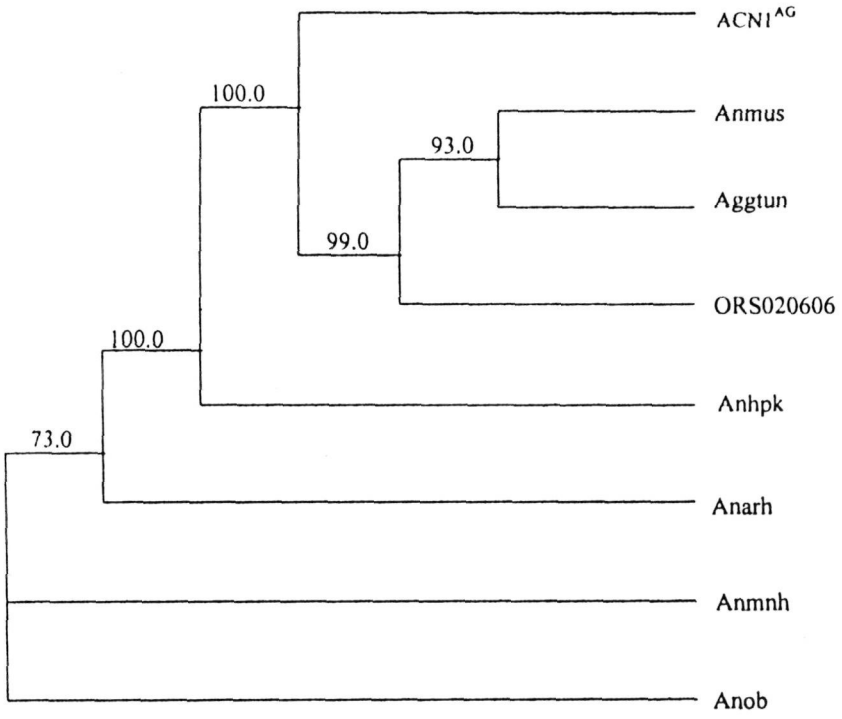

ACN1AG
21.5
Anmnh
21.5
Anarh
100.0
Anob
100.0
Anhpk
97.4
ORS020606
100
Anmus
Aggtun

Figure 4.21

Fig 4.21 : Evolutionary relationship between the *Alnus* nodule microsymbionts as revealed by the tree constructed through the neighbour joining method. The numbers at each node represent the bootstrap values out of 100 replicates


ACN1AG
Anmus
Aggtun
ORS020606
Anhpk
Anarh
Anmnh
Anob
100.0
93.0
99.0
100.0
73.0

4.10.2 Phylogenetic relationship based on partial 16S rRNA gene sequences

The comparison of rRNA sequences is a powerful tool for deducing phylogenetic relationships. In this study, sequence information of a number of *Alnus* nodule microsymbiont collected from various locations has been used to deduce the inter as well as intra specific polymorphism. For analysis, DNA sequences of a variety of related organisms selected from literature (table 4.4) was retrieved from GenBank data library (http://ncbi.nl.nih.gov) and were searched for sequence homology using multiple alignment program CLUSTAL W (version 1.74; Thompson *et al.*, 1994). From the aligned sequences, regions containing a number of insertions in one sequence and deletions in another sequence were excluded from the analysis. For phylogenetic analysis different approaches were employed. Phylogenetic analysis using maximum likelihood and distance matrix approaches were performed with programs DNADIST and DNAML from the PHYLIP package (version 3.5c; Felsenstein, 1995). Trees were constructed from the distance matrices using the PHYLIP program NEIGHBOUR JOINING and UPGMA. Similarly, maximum parsimony analysis was carried out by DNAPARS (DNA parsimony program), SEQBOOT (bootstrap) and CONSENSE (consensus tree program) programmes of PHYLIP inference package, version 3.5c (Felsenstein, 1995). In each case, the similar tree topology resulted. The reliability of the phylogenetic trees obtained were checked by 100 bootstrap replicates.

The computer analysis of the partial 16S rRNA sequence data revealed relationship between 19 microorganisms including the nodule microsymbiont and *Frankia* isolates. A set of 460 characters (nucleotide positions) having 32 variable sites of 16S rRNA 3' end gene was taken for the analysis. In general, analyses using programmes DNA maximum likelihood, neighbour joining, UPGMA and maximum parsimony analysis, showed more or less similar tree topology showing

Table-4.4 : Bacterial strains and nodule microsymbiont used for the 16S rRNA phylogenetic analysis

Bacterial Strains/ Micro-symbionts	Host	Characteristic/ Geographic Location	GenBank Accession Number	Source / Reference
Anmnh*	*Alnus nepalensis*	Nonkhrem Hills, Meghalaya, India	-	this study
Anmus*	*Alnus nepalensis*	Upper Shillong, Meghalaya, India	-	this study
Anarh*	*Alnus nepalensis*	Hapoli, Arunachal-Pradesh, India	-	this study
Anob*	*Alnus nepalensis*	Ooty, Tamilnadu, India	-	this study
Anhpk*	*Alnus nitida*	Kulu, Himachal-Pradesh, India	-	this study
Aggtun*	*Alnus glutinosa*	Tuebingen, Germany	-	this study
ACN1^{AG}	*Alnus* sp	*Alnus* infective, Université, Claude- Bernard, Lyon, France	-	this study
Rugosa	*Alnus rugosa*	Nod+, Quebec City, Canada	L40956	Normand *et al.*, 1996
Dryas	*Dryas drummondii,*	Gaspé, Canada	L40616	Normand *et al.*, 1996
Myrica	*Myrica nagi,*	Lawjynriew Meghalaya, India	L40622	Normand *et al.*,1996
DARJ16	*Alnus nepalensis*	*Alnus* infective, Darjiling, India	U60284	Misra and Normand (1997)
AVN17s	*Alnus* sp	*Alnus* infective, France	L40613	Normand *et al.*, 1996
ArgP5	*Alnus rugosa*	*Alnus* infective, Quebec City, Canada	L40612	Normand *et al.*, 1996
ACN14a	*Alnus crispa*	Tadousac, Canada	M88466	Johnson (1992)
ORS020606	*Casuarina* sp	*Casuariana* infective	M58598	Normand *et al.* (1992)
HRN27-14	*Hippophae-rhamnoides*	*Elaeagnus* infective France	L40617	Normand *et al.* (1996)
SCN10a	*Shepherdia-canadensis*	*Elaeagnus* infective Canada	L40619	Normand *et al.* (1996)
Ptl1	*Purshia tridentata*	*Purshia tridentata* isolate	L41048	Normand *et al.* (1996)
Myco-bacterium	*Mycobacterium ratisbonense*		AF055331	Reischl *et al.* (1998)

The asterisk marked samples are from nodules.

two major clusters. The resultant tree in all the cases were rooted using the outgroup species *Mycobacterium ratisbonense* (Accession number: AF055331; Reisch *et al.*, 1998). All the nodule microsymbiont were clustered together forming one group along with the *Frankia* isolates $ACN1^{AG}$ and nodule microsymbiont DARJN (Misra and Normand, GenBank Accession No. U60284) except for Aggtun and Anmus which clustered together with Myrica, Rugosa and ACN14a (fig 4.21; 4.23) (Normand *et al.*, 1996). The analysis in general, revealed close relationship between Upper Shillong (Anmus) and Tuebingen (Aggtun) samples as was confirmed by the high bootstrap confidence value (98.0). The two *Elaeagnus* infective strains SCN10a and HR27-14 were clustered together separately showing a bootstrap confidence value of 83.0 (NJ). Both, PtI1 and *Dryas* had no close neighbours in all the trees used for the analysis.

4.10.3 Phylogenetic analysis based on *nif* D-K IGS sequences

Similarly, the available *nif* D-K IGS sequences were used for BLAST search to retrieve homologous sequences from GenBank data library (table 4.5). Phylogenetic analysis of the data set was carried out similar to 16S rRNA analysis using programs DNADIST, DMAML, NEIGHBOUR, DNAPARS and CONSENSE of the PHYLIP package (version 3.5c; Felsenstein, 1995). The confidence of the obtained trees were confirmed by boot strap analysis.

Computer analysis of *nif* D-K intergenic spacer region from the sequence information obtained from the nodule microsymbiont Anarh, Aggtun and the GenBank retrieved sequences allowed construction of phylogenetic tree. In both the parsimony and neighbour joining methods, the sample Anarh and Aggtun clustered together in one group with *Alnus* infective ArI3 strain. However, the bootstrap confidence value shown in each case was not satisfactory (fig 4.24; 4.25). The phylogenetic analysis of the microsymbiont and related organisms based on the

Fig 4.22 : DNA parsimony strict consensus phylogenetic tree derived from the aligned *Frankia* 16S rDNA partial sequences. A total of 460 nucleotides positions having 32 variable sites were compared for phylogenetic inference. The numbers at the fork indicate the bootstrap values out of 100 replicates. The tree was rooted with the out group species *Mycobacterium ratisbonense.*

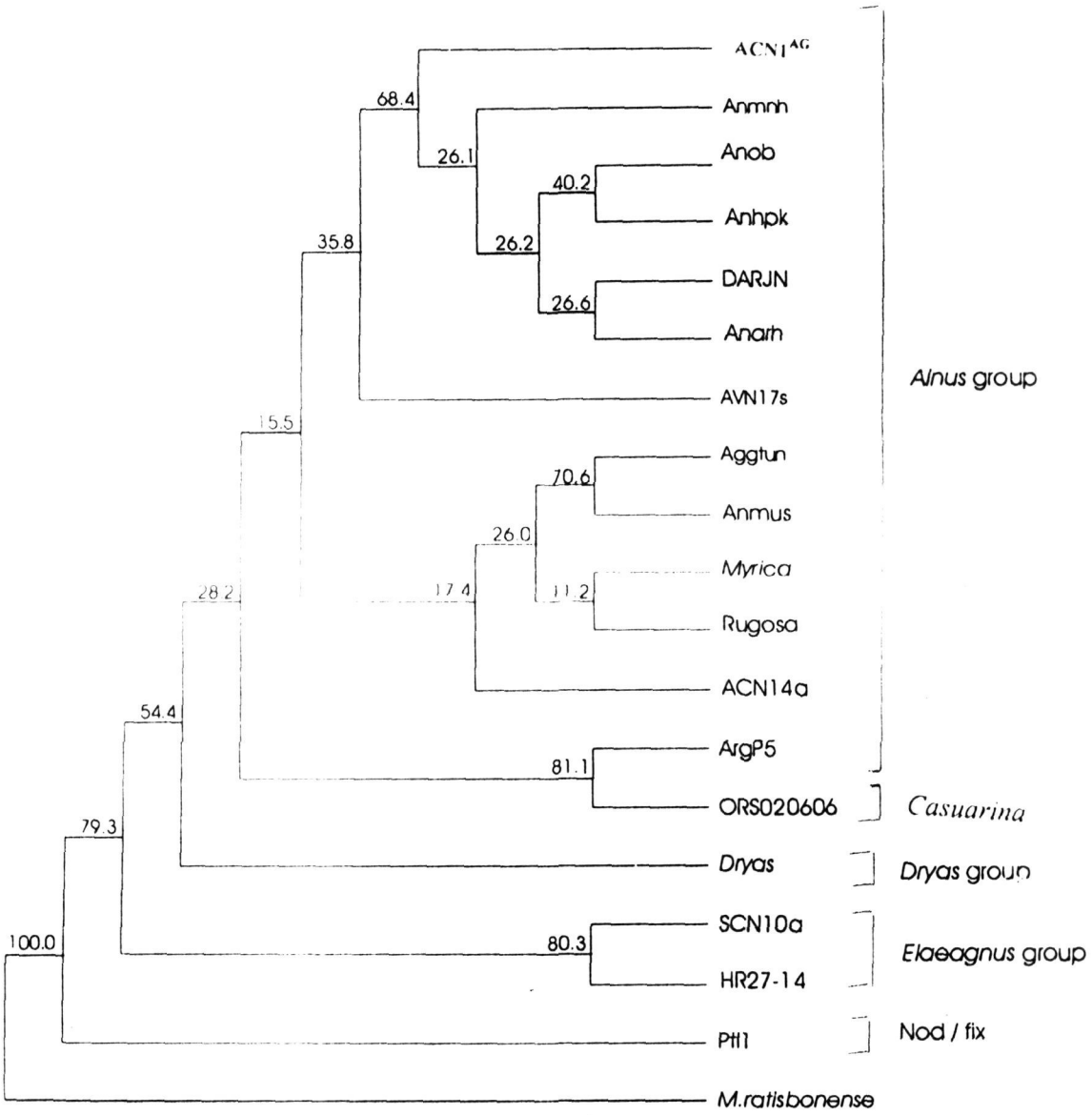

ACN1AG
Anmnh
Anob
Anhpk
DARJN
Anarh
AVN17s
Aggtun
Anmus
Myrica
Rugosa
ACN14a
ArgP5
ORS020606
Dryas
SCN10a
HR27-14
Ptl1
M.ratisbonense
68.4
26.1
40.2
26.2
35.8
26.6
15.5
70.6
26.0
28.2
17.4
11.2
54.4
81.1
79.3
100.0
80.3
Alnus group
Casuarina
Dryas group
Elaeagnus group
Nod / fix

Fig 4.23 : Neighbour joining strict consensus phylogenetic tree derived from the aligned *Frankia* 16S rRNA sequences. The numbers at the fork indicates the bootstrap values out of 100 replicates.

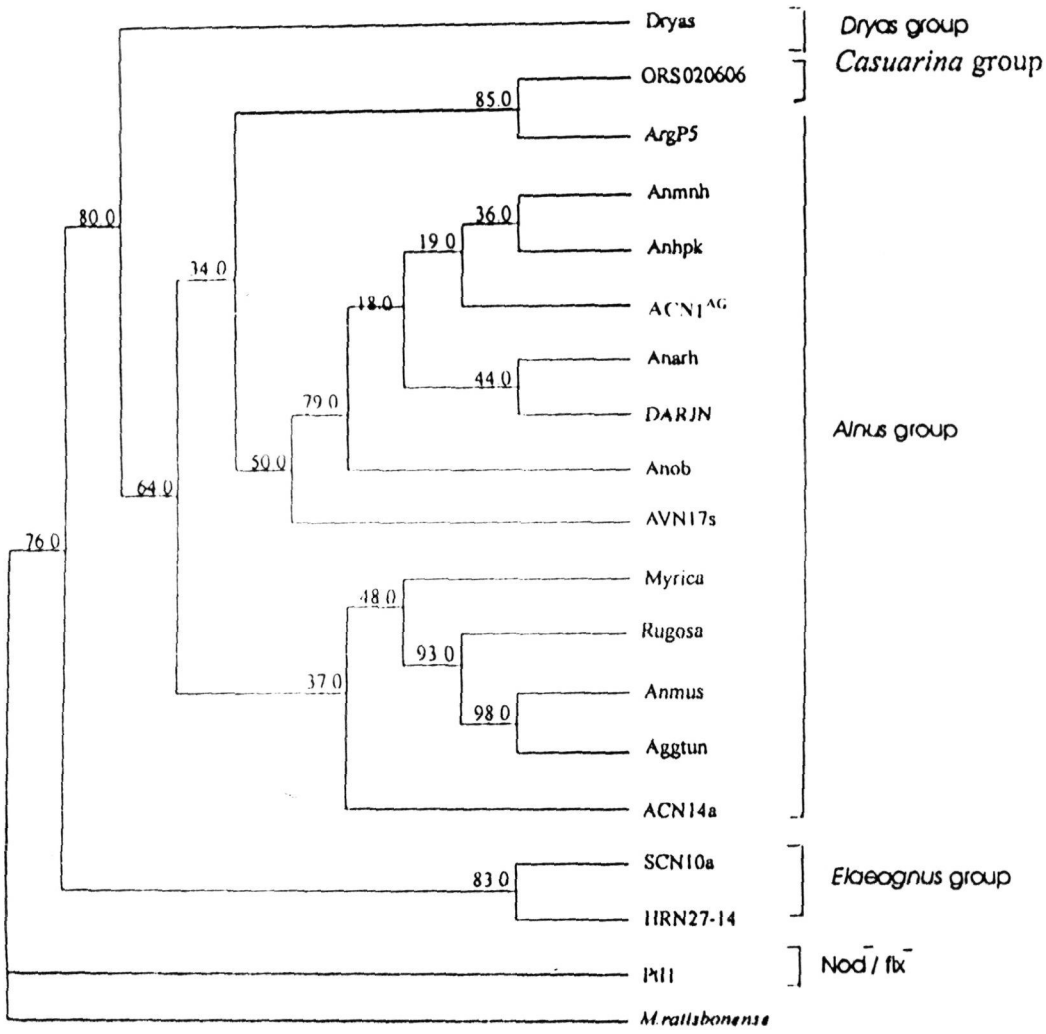

Dryas
ORS020606
ArgP5
Anmnh
Anhpk
ACN1AG
Anarh
DARJN
Anob
AVN17s
Myrica
Rugosa
Anmus
Aggtun
ACN14a
SCN10a
HRN27-14
Pkl1
M.ratisbonense
85.0
80.0
36.0
19.0
34.0
18.0
44.0
79.0
50.0
64.0
76.0
48.0
93.0
37.0
98.0
83.0
Dryas group
Casuarina group
Alnus group
Elaeagnus group
Nod⁻ / fix⁻

Table 4.5 : Bacterial strains and nodule microsymbionts used for the *nif* D-K phylogenetic analysis

Bacterial strains /Nodule microsymbionts	Host	Characteristic/ Geographic location	GenBank Accession Number	Reference/ Source
Anarh	*Alnus nepalensis*	Hapoli, Arunachal Pradesh	-	this study
Aggtun	*Alnus glutinosa*	Tuebingen, Germany	-	this study
FaC1	*Alnus viridis ssp. crispa*	*Alnus* isolate	U53363	Oh *et al.*, 1996
ArI3	*Alnus rubra*	Oregon	L41344	Specq and Normand., 1996
Kp54	*Gymnostoma poissonianum*	Kauaoua, New Caledonia	U63694	Navarro *et al.*, 1997
CN61	*Gymnostoma nodiflorum*	Canala, New Caledonia	U63693	Navarro *et al.*, 1997
MG59	*Gymnostoma glaucescens*	Me Aïu, New Caledonia	U63691	Navarro *et al.*, 1997
TC24	*Gymnostoma chamaecypari s*	Tontouta, New Caledonia	U63692	Navarro *et al.*, 1997
Eun1f	*Elaeagnus umbellata*	Illinois	L37664	Nalin *et al.*, 1995 Navarro *et al.*, 1997
EaN1Pec	*Elaeagnus angustifolia*	Ohio	U63698	Navarro *et al.*, 1997
HRN18a	*Hippophae rhamnoides*	Alps, France	U63696	Navarro *et al.*, 1997
SCN10a	*Shepherdia canadensis*	Quebec, Canada	U63695	Navarro *et al.*, 1997
AVINIFA	*Azotobacter vinelandii*	-	M11579	Brigle *et al.*, 1985

Fig 4.24 : Phylogenetic tree constructed using the parsimony method using PHYLIP DNAPARS program of the aligned *Frankia nif* D-K IGS sequences. The sequences were aligned using the Clustal W program. The numbers at the left of the branches correspond to bootstrap values out of 100 replicates. The tree was rooted using *Azotobacter vinelandii* as an outgroup.

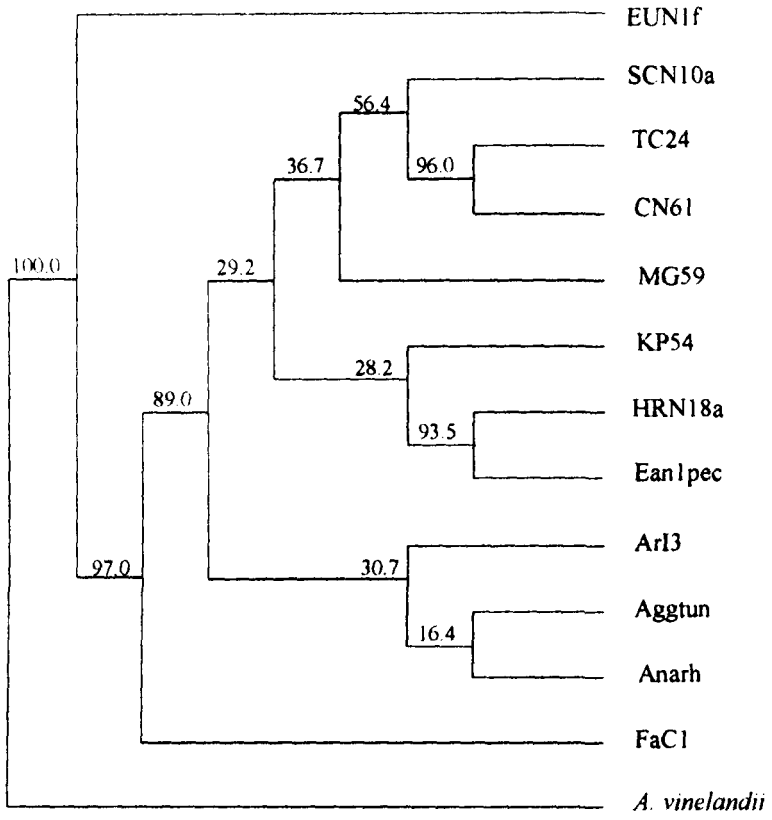
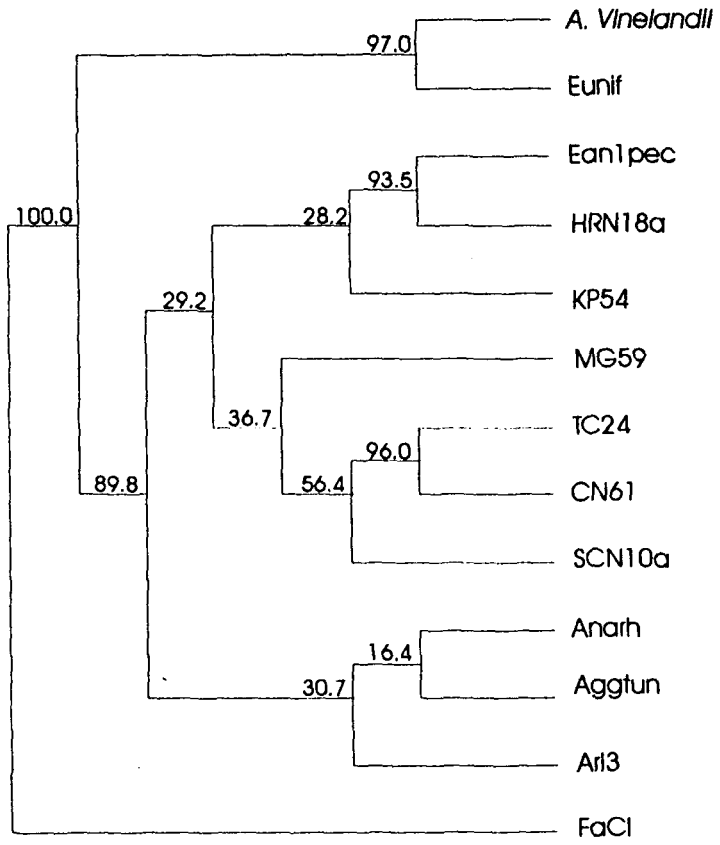

EUN1f
SCN10a
56.4
TC24
36.7
96.0
CN61
100.0
29.2
MG59
KP54
28.2
89.0
HRN18a
93.5
Ean1pec
Arl3
97.0
30.7
Aggtun
16.4
Anarh
FaC1
A. vinelandii

Figure 4.25

Fig 4.25 : Neighbour joining phylogenetic tree constructed using the aligned *Frankia nif* D-K IGS sequences. The horizontal distances between nodes of the tree represent the relative evolutionary distance and the bootstrap values.


A. Vinelandii
97.0
Eunif
Ean1pec
93.5
100.0
28.2
HRN18a
KP54
29.2
MG59
36.7
TC24
96.0
89.8
56.4
CN61
SCN10a
Anarh
16.4
30.7
Aggtun
Arl3
FaCl

partial 16S rDNA sequences gave an overall impression about their evolutionary relationship.

The topology of the tree drawn through comparative analysis of partial sequences of 16S rDNA were similar to those deduced by Normand *et al.* (1996) and Clawson *et al.* (1998) showing similar groupings. The trees obtained through both neighbour joining and DNA parsimony analysis showed consistent features such as, clustering together of strains/microsymbiont from *Alnus* species (Betulaceae) with those from *Casuarina* sp. (Casuarinaceae) and *Myrica* sp. (Myricaceae). The SCN10a and HR27-14 strains from the *Elaeagnus* group formed a separate cluster of their own. This is in agreement with the inferences made by Nazaret *et al.* (1991) who showed the close relationship of *Casuarina* strains with the *Alnus* infectivity group and its distance from *Elaeagnus* infectivity group. Furthermore, nod^-/fix^- PtI1 strain was placed distantly in the tree.

The sequence information of the internal transcribed spacer in conjunction with partial 16S and 23S rRNA gene gave a general view of relationship existing between the frankial nodule microsymbiont. The microsymbiont analysed in the study joined the *Alnus* infective group. The nearest neighbour for the nodule microsymbiont Anob sample was Anhpk and for Anarh was DARJN. The distant relationship of the Anmus with other nodule samples studied from India was clearly reflected in its rDNA sequence variation pattern. In the trees drawn, it clustered together with Aggtun as close neighbour followed by Myrica, Rugosa and ACN14a.

The microsymbiont from nodules collected from Upper Shillong, Meghalaya, stood distinct from the other nodule samples analysed from the same host (*Alnus nepalensis*), although collected from diverse locations. The evolutionary separation of Anmus from other samples was reflected in the PCR-RFLP pattern which was

Divergence of genotypes in a geographical location, making it similar to a different region

16 genotype of are also known.

confirmed by rDNA sequence analysis. This is in accordance with the observation made by Ganesh *et al.* (1994) who suggested the possibility of existence of genetically different *Frankia* strains in this region. It is surprising that the microsymbiont Anmus, although distinct from rest of the samples analysed from India, showed close relationship with Aggtun sample. The phylogenetic tree derived from both partial 16S rRNA and 16S-23S rRNA ITS analysis showed close relationship existing between them (fig 4.21;4.22; 4.23). But this relationship observed in the molecular data was not clear morphologically. Whatever be the relationship between Anmus and Aggtun, it is evident that the microsymbiont analysed from Upper Shillong is genetically different from the rest of the samples analysed from India. At morphological level too, the occurrence of nodule hair occasionally made it distinct from others. So, the study gives insight into the diversity among the *Frankia* nodulating *Alnus* species in India. Also, it suggests that Anmus can be a distinct *Alnus* infective *Frankia* strain with a distinct genetic history.

It may not seem to speculative to conjecture that one of the following two scenarios may explain the proximity of Anmus sequence from India and Aggtun sequence from Germany -

(i) The enormous pool of genetic diversity in Meghalaya, India, contributed a strain to the European environment, where it diversified only marginally, while getting adapted to associating with a different alder species.

(ii) An evolved *A. glutinosa* compatible strain from Germany got a chance to migrate to India, may be along with the increased mobility of man. The strain ended up infecting *A. nepalensis* trees and diversified marginally in the process.

CHAPTER 5

CONCLUSIONS

CONCLUSIONS

1. The present investigations were able to demonstrate the presence of considerable variability among the alder compatible *Frankia* found in India. This diversity was confirmed by -
   (a) The PCR/RFLP studies.
   (b) The nucleotide sequence analysis.

2. While the nodule samples collected from various parts of India showed considerable diversity, even the samples collected from Meghalaya also had significant differences.

3. Quite unexpectedly, one of the samples (Anmus) collected from Indian alder (*Alnus nepalensis*) was found to be closer to a sample collected from *Alnus glutinosa* tree growing in Tuebingen, Germany.

4. The phylogenetic analysis of the sequence data confirmed the proximity of the Indian sample Anmus to the Tuebingen sample.

5. This suggested that the strain was highly divergent from the other *Alnus nepalensis* nodule microsymbiont present in India.

6. The study conducted, therefore, confirmed that the ITS between 16S-23S rRNA genes may be an effective marker for detecting genetic differences at the intra and inter-specific levels.

REFERENCES

REFERENCES

Altshul S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman (1990). Basiclocal alignment search tool. J. Mol. Biol. **215**: 403-410.

An C.S., W.S. Riggsby and B.C. Mullin (1983). Deoxyribonucleic acid base composition of 12 *Frankia* isolates. Can. J. Bot. **61**: 2859-2862.

An C.S., W.S. Riggsby and B.C. Mullin. (1985a). Relationships of *Frankia* isolates based on deoxyribonucleic acid homology studies. Int. J. Syst. Bacteriol. 35: 140-146.

An C.S., W.S. Riggsby and B.C. Mullin (1985b). Restriction pattern analyses of genomic DNA of *Frankia* isolates. Plant and Soil **87**: 43-48.

Ausubel F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidmann, J.A. Smith and K. Struhl (eds.) (1989). Current protocols in molecular biology. Green Publishing Associates (Massachusetts General Hospital/Harvard Medical School.

Baker D., L.P. William and J.G. Torrey (1981). Immunochemical analysis of relationship among isolated frankiae (Actinomycetales). Int. Syst. Bacteriol. **31**: 148-151.

Baker D.D. (1987). Relationships among pure cultured strains of *Frankia* based on host specificity. Physiol. Plant. **70**: 245-248.

Baker D.D. and C.R. Schwintzer (1990). Introduction. In The Biology of *Frankia* and Actinorhizal plants, C.R. Schwintzer and J.D. Tjepkema (eds.). Academic Press, California, pp. 3-11.

Baker D.D. and B.C. Mullin (1994). Diversity of *Frankia* nodule endophytes of the actinorhizal shrub *Ceanothus* as assessed by RFLP patterns from single nodule lobes. Soil Biol. Biochem. **26**: 547-552.

Benson D.R. and D. Hanna (1983). *Frankia* diversity in alder stand as estimated by sodium dodecyl sulphate polyacrylamide gel electrophoresis of whole cell proteins. Can. J. Bot. **61**: 2919-2923.

Benson D.R. and W.B. Silvester (1993). Biology of *Frankia* strains, actinomycete symbionts of Actinorhizal plants. Microbiol. Rev. **57**: 293-319.

Benson D.R., S.E. Bucholz and D.G. Hanna (1984). Identification of *Frankia* strains by 2-D-polyacrylamide gel electrophoresis. Appl. Environ Microbiol. **47**: 489-494.

Bibb M.J., P.R. Findley and M.W. Johnson (1984). The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein coding sequences. Gene **30**: 157-166.

Bloom R.A., B.C. Mullin and R.L. Tate (1989). DNA restriction patterns and DNA hybridisation studies of *Frankia* isolates from *Myrica pensylvanica* (bayberry). Appl. Environ. Microbiol. **55**: 2156-2160.

Borthakur M., A. Sen and A.K. Misra (1996). Exogenous nitrogen sources do not drastically reduce nitrogenase activity in polymer entrapped *Frankia*. Indian J. Exp. Biol. 35: 173-175.

Bosco M., M. P. Fernandez, P. Simonet, R. Materassi and P. Normand (1992). Evidence that some *Frankia* sp. strains are able to cross boundaries between *Alnus* and *Elaeagnus* host specificity groups. Appl. Environ. Microbiol. **58**: 1569-1576.

Bosco M., E. Lumini and P. Normand (1996). PCR-RFLP direct fingerprinting of uncultured Frankia microsymbiont from Dryas drummondii nodules. Annali di Microbiologia ed Enzimologia. **46**: 115-123.

Bullock W.O., J. M. Fernandez and J. M. Short (1987). XL1- Blue: a high efficiency plasmid transforming *rec* A *Escherichia coli* strain with beta-galactoside selection. BioTechniques. **5**: 376-379.

Busse M.T. and P.J. Bottomley (1989). Growth and nodulation response of Rhizobium meliloti to water stress induced by permeating and nonpermeating solutes. Appl. Environ. Microbiol. **55**: 2431-2436.

Callaham D., P. Del Tredici and J.G. Torrey (1978). Isolation and cultivation in vitro of the actinomycete causing root nodulation in *Comptonia peregrina*. Science **199**: 899-902.

Cannon F.C., G.E. Riedal and F.M. Ausubel (1979). Overlapping sequences of *Klebsiella pneumoniae nif* DNA cloned and characterized. Mol. Genet. **174**: 59-66.

Chung C.T., S.L. Niemela and R.H. Miller (1989). One step preparation of competant *Escherichia coli*: Transformation and storage of bacterial cells in the same solution. Proc. Natl. Acad. Sci. USA. **86**: 2172-2175.

Clark J.M. (1988). Novel non-templated nucleotide addition reactions catalysed by prokaryotic and eukaryotic DNA polymerases. Nucl. Acids Res. **16**: 9677-9686.

Clawson M. L., M. Caru and D.R. Benson (1998). diversity of *Frankia* strains in root nodules of plants from the families Elaeagnaceae and Rhamnaceae. Appl. Environ. Microbiol. **64**: 3539-3543.

Cournoyer B., M. Gouy and P. Normand (1993). Molecular Phylogeny of the symbiotic actinomycetes of the genus *Frankia* matches host-plant infection process. Mol. Biol. Evol. **10**: 1303-1316.

Diem H.G., D. Gautheir, Y.R. Dommergues (1983). An effective strain of *Frankia* from *Casuarina* sp. Canadian J. Botany **61**: 2815-2821.

Dobritsa S.V. (1985). Restriction analysis of the *Frankia* spp. genome. FEMS Microbiol. Lett. **29**: 123-128.

Domenach A.M., F. Kurdali and R. Bardin (1989). Estimation of dinitrogen fixation in alder forest by the method based on natural 15N abundance. Plant and Soil 118: 51-59.

Dommergues Y.R. (1996). Nitrogen fixation in actinorhizal plants: Application. Acta Botanica Gallica. **143**: 663-679.

Felsenstein J. (1995). PHYLIP (phylogeny inference package), Version 3.5c. Department of Genetics, University of Washington, Seatle.

Fernandez M.P., H. Meugnier, P.A.D. Grimont and R. Bardin (1989). Deoxyribonucleic acid relatedness among members of the genus *Frankia*. Int. J. Syst. Bacteriol. **39**: 424-429.

Furlow J.J. (1979). The systematics of the American species of *Alnus* (Betulaceae). Rhodora **81**: 1-121, 151-248.

Galtier N., N. Tourasse and M. Gouy (1999). A non hyperthermophilic common ancestor to extant life forms. Science **283**: 220-221.

García-Martínez J., A. Martínez-Murcia, A.I. Antón and F. Rodríguez-Valera (1996). Comparison of the small 16S to 23S intergenic spacer region of the rRNA operons of some *Escherichia coli* strains of the ECOR collection and *Escherichia coli* K-12. J. Bacteriol. **178**: 6374-6377.

Gardes M. and M. Lalonde (1987). Identification and subgrouping of *Frankia* strains using sodium dodecyl sulphate-polyacrylamide gel electrophoresis. Physiol. Plant. **70**: 237-244.

Gardes M., J. Bosquet and M. Lalonde (1987). Isozyme variations among 40 *Frankia* strains. Appl. Environ. Microbiol. **53**: 1596-1603.

Gauthier D., H.G. Diem, Y.R Dommergues (1981). In vitro nitrogen fixation by two actinomycetes strains isolated from Causarina nodules. Appl. Environ. Microbiol. **41**, 306-308.

Ganesh G., A.K. Misra, C. Chapelon and P. Normand (1994). Morphological and molecular characterization of *Frankia* sp. isolate from nodules of *Alnus nepalensis*. Don. Arch. Microbiol. **161**: 152-155.

Goode B.L., and S.C. Feinstein (1992). "Speedprep" purification of template for double stranded DNA sequencing. Bio Techniques **12**: 148-151.

Hafeez F., A.D.L. Akkerman and A.H. Chaudhary (1984). Morphology, Physiology and infectivity of two *Frankia* isolates An1 and An2 from root nodules of *Alnus nitida*. Plant and Soil 78: 45-59.

Hahn D., M.P. Lechevalier, A. Fischer and E. Stackebrandt (1989). Evidence for a close relationship between members of the genera *Frankia*, *Geodermatophilus* and "*Blastococcus*" and emendation of the family Frankiaceae. Appl. Environ. Microbiol. **11**: 236- 242.

Harriott O.T., T.J. Hosted and D.R. Benson (1995). Sequences of *nif* X, *nif* W, *nif* Z , *nif* B and two ORF in the *Frankia* nitrogen fixation gene cluster. Gene. **161**: 63-67.

Harry D.E., D.C. Yang and J.O. Dawson. 1991. Nucleotide sequence and diversity in the 16S ribosomal RNA from Frankia. Plant Soil. **131**: 143-146.

Haselkorn R. (1986). Organization of the genes for nitrogen fixation in photosynthetic bacteria and cyanobacteria. Annu. Rev. Microbiol. **40**: 525-547.

Hirsch A.M., H.I. Mckhann, A. Reddy, J. Liao, Y. Fang and C.R. Marshall (1995). Assessing horizontal transfer of *nif* HDK genes in eubacteria: nucleotide sequence of *nif* K from *Frankia* strain HFPCcI3. Mol. Biol. Evol. **12**: 16-27.

Jamann S., M.P. Fernandez and P. Normand (1993). Typing method for N_2 fixing bacteria based on PCR-RFLP application to the characterization of *Frankia* strains. Mol. Ecology **2**: 17-26.

Jensen M.A., J.A. Webster and N. Straus (1993). Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphisms. Appl. Environ. Microbiol. **59**: 945-952.

Laguerre G., P. Mavingui, M.R. Allard, M.P. Charnay, P. Louvier, S.I. Mazurier, L.R.Gouy and N. Amarger (1996). Typing of *Rhizobia* by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: application to *Rhizobium*

leguminosarum and its different biovars. Appl. Environ. Mcrobiol. **62**: 2029-2036.

Lalonde M. and H.E. Calvert (1979). Production of *Frankia* hyphae and spores as infective inoculant for *Alnus* species. In Symbiotic nitrogen fixation in the Management of Temperate Forests, J.C. Gordon, C.T. Wheeler and D.A Perry (eds.), Oregon State University, Corvallis, Oreg. pp. 95-110.

Lechevalier M.P. (1984). The taxonomy of genus *Frankia*. Plant and Soil **78**: 1-6.

Lechevalier M.P. (1994). Taxonomy of the genus *Frankia* (Actinomycetales). Int. Syst. Bacterio. **44**: 1-8.

Lechevalier M.P. and H.A. Lechevalier (1979). The taxonomic position of Actinomycete endophytes. In Symbiotic nitrogen fixation in the Management of the Temperate Forests. J.C. Gordon, C.T. Wheeler and D.A Perry (eds.), Oregon State University, Corvallis, Oreg. pp. 111-122.

Lechevalier M.P., D. Baker and F. Horriere (1983). Physiology, chemistry, serology, and infectivity of two *Frankia* isolates from *Alnus incana* subsp. *Rugosa*. Can. J. Bot. **61**: 2826-2833.

Lumini E. and M. Bosco (1996). PCR - restriction fragment length polymorphism identification and host range of single spore isolates of flexible *Frankia* sp. UFI132715. Appl. Environ. Micorbiol. **62**: 3026-3029.

Merrick M.J. (1993). Organization and regulation of nitrogen fixation genes in bacteria. In New horizons in Nitrogen Fixation, Proceedings of the 9th International congress on Nitrogen Fixation, R. Palacious, J. Mora and W.E. Newton (eds.), Kluwer Academic, Dordrecht, pp. 43-54.

Mirza M.S., D. Hahn and A.D.L. Akkermans (1992). Isolation and characterization of *Frankia* strains from *Coriaria nepalensis*. Syst. Appl. Microbiol. 15: 289-295.

Mirza M.S., S. Hameed and A.D.L. Akkermans (1994). Genetic diversity of *Datisca cannabina*-compatible *Frankia* strains as determined by sequence analysis of the PCR-amplified 16S rRNA gene. Appl. Environ. Microbiol. **60**: 2371-2376.

Misra A.K., E. Paget, P. Simonet and P. Normand (1991). Isolation and amplification of *Frankia* DNA from nodules of *Alnus nepalensis* and studies on overcoming foreign DNA contamination. Indian J. Expt. Biol. **29**: 946-948.

Mullis K.B. and F.A. Faloona (1987). Specific synthesis of DNA in vitro via polymerase catalysed chain reaction. Methods Enzymol. **155**: 335-350.

Mullis K.B., F.A. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold spring Harbor Symp. Quant. Biol. **51**: 263-273.

Murry M.A., A.S. Konopka, S.D. Pratt and T.L. Vandergon (1997). The use of PCR based typing methods to assess the diversity of *Frankia* nodule endophytes of the actinorhizal shrub *Ceanothus*. Physiol. Plant. **99**: 714-721.

Nalin R., A.M. Domenach and P. Normand (1995). Molecular structure of the *Frankia* spp. *nif* D-K intergenic spacer and design of *Frankia* genus compatible primer. Mol. Ecology **4**: 483-491.

Navarro E., P. Simonet, P. Normand and R. Bardin (1992). Characterization of natural populations of *Nitrobacter* spp. using PCR-RFLP analysis of the ribosomal intergenic spacer. Arch. Microbiol. **157:** 107-115.

Nazaret S., B. Cournoyer, P. Normand and P. Simonet (1991). Phylogenetic relationships among *Frankia* genome species determined by use of amplified 16S rDNA sequences. J. Bacterial. **173**: 4072-4078.

Newcomb W. and S.M. Wood (1987). Morphogenesis and fine structure of *Frankia* (Actinomycetales): the microsymbiont of nitrogen fixing root nodules. Int. Rev. Cytol. **109**: 1-88.

Nick G., E. Paget, P. Simonet, A. Moiroud and P. Normand (1992). The nodular endophytes of *Coriaria* spp. form a distinct lineage within the genus *Frankia*. Mol. Ecol. **1**: 175-181.

Normand P. and M. Lalonde (1982). Evaluation of *Frankia* strains isolated from provenances of two *Alnus* species. Can. J. Microbiol. **28**: 1133-1142.

Normand P. and M. Lalonde (1986). The genetics of actinorhizal *Frankia*: - a review. Plant and Soil **90**: 429-453.

Normand P., P. simonet. and R. Bardin (1988). Conservation of *nif* sequences in *Frankia*. Mol. Gen. Genet. **213**: 238-246.

Normand P. and J. Bousquet (1989). Phylogeny of nitrogenase sequences in *Frankia* and other nitrogen fixing microorganisms. J. Mol. Evol. **29**: 436-447.

Normand P., M. Gouy, B. Cournoyer and P. Simonet (1992a). Nucleotide sequence of *nif* D from *Frankia alni* strain ArI3: Phylogenetic inferences. Molecular Biology and Evolution **9**: 495-506.

Normand P., B. Cournoyer, P. Simonet and S. Nazaret (1992b). Analysis of ribosomal operon in the actinomycete *Frankia*. Gene **111**: 119-124.

Normand P., S. Orso, B. Cournoyer, P. Jeannin, C. Chapelon, J. Dawson, L. Evtushenko and A.K. Misra (1996). Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. Intl. J. Syst. Bacteriol. **46**: 1-9.

Ochman H. and A.C. Wilson (1987). Evolution in bacteria: evidence for a universal substitution rate in cellular genomes, J. Mol. Evol. **26**: 74-86.

Oh B.S., P.G. Twigg, J.S. Hong, B.C. Mullin and C.S. An (1997). *Nif* V is contiguous to *nif* HDK in *Frankia* strain FaC1. Physiol. Plant. **99**: 702-713.

Prin Y., L. Maggia, B. Picard, H.G. Diem and P. Goullet (1991). Electrophoretic comparison of enzymes from 22 single spore cultures obtained from *Frankia* strain ORS140102. FEMS Microbiol. Lett. **77**: 223-228.

Ponsonnet C. and X. Nesme (1994). Identification of *Agrobacterium* strains by PCR-RFLP analysis of pTi and chromosomal regions. Arch. Microbial. **161**: 300-309.

Ritchie N.J. and A.M. David (1999). Geographic distribution and genetic diversity of *Ceanothus*- infective *Frankia* strains. Appl. Environ. Microbiol. **65**: 1378-1383.

Rouvier C., Y. Prin, P. Reddell, P. Normand and P. Simonet (1996). Genetic diversity among *Frankia* strains nodulating members of the family *Casuarinaceae* in Australia revealed by PCR and restriction fragment length polymorphism analysis with crushed root nodules. Appl. Environ. Microbiol. **62**: 979-985.

Ruvkun G.B. and F.M. Ausubel (1980). Interspecies homology of nitrogenase genes. Proc. Natl. Acad. Sci. USA **74**: 5463-5467.

Sambrook J., E.F. Fritsch and T. Maniatis (1989). Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press.

Sanger F., S. Nicklen and A.R. Coulson (1977). DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. USA. **74**: 5463-5467.

Sarma G., A. Sen, R. Varghese, and A.K. Misra (1997). A novel technique for isolation of *Frankia* and generation of single spore cultures. Canadian J. Microbiol. **44**: 490-492.

Shantharam S. and A.K. Mattoo (1997). Enhancing biological nitrogen fixation : An appraisal of current and alternative technologies for N input into plants. Plant and Soil **194**: 205-216.

Sharma E., and R.S. Ambasht (1986).Root nodules age class transition, production and decomposition in an age sequence of *Alnus nepalensis* plantation stands in the Eastern Himalayas. J. Appl. Ecol.**23**: 689-701.

Shine J. and L. Dalgarno (1974). The 3'- terminal sequence of *Escherichia coli* 16S ribosomal RNA complementary to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**: 1342-1346.

Shipton W.A. and A.J.P. Burggraf (1982). A comparison of the requirement for various carbon and nitrogen sources and vitamins in some *Frankia* isolates. Plant and Soil **69**: 149-161.

Simonet P., P. Normand, A. Moiroud and M. Lalonde (1985). Restriction enzyme digestion patterns of *Frankia* plasmids. Plant and Soil **87**: 49-60.

Simonet P., P. Normand, A. Moiroud and R. Bardin (1990). Identification of *Frankia* strain in nodules by hybridization of polymerase chain reaction products with strain specific oligonucleotide probes. Arch. Microbiol. **153**: 235-240.

Simonet P., A.K. Misra, S. Nazaret, B. Cournoyer and P. Normand (1991). *Frankia* genus-specific characterization by polymerase chain reaction. Appl. Environ. Microbiol. **57**: 3278- 3286.

Southern E.M. (1975). Detection of specific sequences among DNA fragment separated by gel electrophoresis. J. Mol. Biol. **98**: 503.

Swensen S.M. and B.C. Mullin (1997). Phylogenetic relationships among actinorhizal plants. The impact of molecular systematics and implications for the evolution of actinorhizal symbioses. Physiol. Plantarum **99**: 565-573.

Tarrant R.F. and J.M. Trappe (1971). The role of *Alnus* in improving the forest environment. Plant and Soil (special volume) 335- 348.

Thompson J.D., D.G. Higgins and T.J. Gibson (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting position specific gap penalities and weight matrix choice. Nucleic Acid Res. **22**: 4673-4680.

Tisa L., M. McBride and J.C. Ensign (1983). Studies on growth and morphology of *Frankia* strains EAN1pec, Eu1c, Cp1 and $ACN1^{AG}$. Can. J. Bot. **61**: 2768-2773.

Tjepkema J.D. and J.G. Torrey (1979). Symbiotic nitrogen fixation in actinomycete-nodulated plants. Preface. Bot. Gaz. (Chicago) Suppl. **140**: i-ii.

Torrey J.G. (1978). Nitrogen fixation by actinomycete-nodulated angiosperms. Bioscience **28**: 586-592.

Torrey J.G. and C. Racette (1989). Specificity among the *Casuarinaceae* in root nodulation by *Frankia*. Plant and Soil **118**: 157-164.

Vogelstein B. and D. Gillspie (1979). Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. USA **76**: 615-619.

Woese C. R. (1987). Bacterial evolution. Microbiol. Rev. **51**: 221-271.

Woese, C.R. and G.E. Fox (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. USA **74**: 5088-5090.

Zhang Z., M.F. Lopez, J.G. Torrey (1984). A comparison of cultural characteristics and infectivity of Frankia isolates from root nodules of *Casuarina* species. Plant and Soil **78**: 79-90.

APPENDIX

APPENDIX

I. CULTURE MEDIA

1. Defined Propionate Minimal medium
(DPM, Baker and O'keafe, 1984)

KH_2PO_4	1.00 g
$MgSO_4$	0.10 g
$CaCl_2$	0.01 g
Sodium Propionate	1.20 g
Hoagland's stock*	1 00 ml
$FeSO_4$-EDTA stock**	1.00 ml
Distilled water	1000.00 ml
pH 6.8	

* Hoagland's stock (g/l) :

H_3BO_3	0.710
$ZnCl_2$	0.220
$CoCl_2$	0.025
$Na_2MoO_4.2H_2O$	0.02
$MnCl_2$	0.450
$CuSO_4$	0.020

** $FeSO_4$-EDTA Stock (g/l) :

$FeSO_4$	10.008
Na_2EDTA	13.410

2. Luria Bertani (LB) Medium

NaCl	10.00 g
Yeast extract	5.00 g
Bacto-tryptone	10.00 g
Water	1000.00 ml
pH-7.5	

3. LB Agar Plates

Agar	15.00 g
LB medium	1000.00 ml

II. BUFFERS

1. DNA Extraction Buffers

a). Extraction Buffer (Simonet *et al.*, 1990)

Tris base	50.00 mM
EDTA	20.00 mM
PVP	1.00 %
pH-8.0	

b). Extraction Buffer (Rouvier *et al.*, 1996)

Tris base	100.00 mM
EDTA	20.00 mM
NaCl	1.40 M
CTAB	2.00 % (w/v)
PVP	1.00 % (w/v)
pH-8.0	

2. Isolation buffers

a). TE Buffer (Simonet *et al.*, 1985)

Tris base	50.00 mM
EDTA	20.00 mM
pH-8.0	

b). TES (Simonet *et al.*, 1985)

Tris base	50.00 mM
EDTA	20.00 mM
pH-8.0	

c). NET Buffer

NaCl	0.15 M
EDTA	0.10 mM
Tris	20.00 mM
pH- 8.0	

d). High Salt NET Buffer

NaCl	0.10 mM
Tris	1.00 M
EDTA	20.00 mM
pH- 8.0	

e). Elution Buffer

Tris. HCl	10.00 mM
pH-8.5	

3. **PCR Buffers**

a). PCR Buffer

Tris HCl	10.00 mM
KCl	50.00 mM
$MgCl_2$	2.00 mM
Gelatin	0.01 %
pH-8.4	

b). HF Buffer

Tris HCl	2.00 mM
KCl	10.00 mM
DTT	0.10 M
EDTA	0.01 mM
Tween 20 (v/v)	0.05 %
Nonidet	0.05 %
Glycerol	5.00 %
$MgCl_2$	1.50 mM
pH-7.5	

4. **Electrophoresis Buffer**

a). TAE Buffer(1X)

Tris	0.040 M
Sodium Acetate	0.020 M
EDTA	0.002 M

pH 8.0

b). TBE Buffer

Tris	0.100 M
Boric acid	0.083 M
EDTA	0.001 M
pH 8.3	

5. **Restriction Enzyme Buffers**

a). Buffer B (Boehringer Mannheim)

Tris HCl	10.00 mM
$MgCl_2$	5.00 mM
NaCl	100.00 mM
2-Mercaptoethanol	1.00 mM
pH-8.00	

b). Buffer H (Boehringer Mannheim)

Tris HCl	50.00 mM
$MgCl_2$	10.00 mM
NaCl	100.00 mM
DTE (Dithioerythritol)	1.00 mM
pH- 8.0	

c). Buffer L (Boehringer Mannheim)

Tris HCl	10.00 mM
$MgCl_2$	10.00 mM
DTE	1.00 mM
pH- 8.0	

d). Buffer 1 (New England Biolab)

Bis- Tris Propane HCl	10.00 mM
$MgCl_2$	10.00 mM
DTT (Dithiothreitol) -	10.00 mM
pH- 7.0	

e). Buffer 2 (New England Biolab)

Tris HCl	10.00 mM
$MgCl_2$	10.00 mM
NaCl	50.00 mM
DTT	1.00 mM
pH - 7.9	

f). Buffer 4 (New England Biolab)

Tris Acetate	20.00 mM
Magnesium Acetate	10.00 mM
Potasium Acetate	50.00 mM
DTT	1.00 mM
pH - 7.9	

6. **Other buffers**

a). Type III Loading Buffer (6X; Sambrook *et al* ., 1989)

Bromophenol blue	2.50 mg
Xylene Cyanol FF	2.50 mg
Glycerol	3.00 % in water

b). Transformation and Storage (TSS) Buffer

LB medium with :	
PEG 6000	10.00 % w\v
$MgCl_2$	50.00 mM
DMSO	5.00 %
pH 6.5	

III. OTHER SOLUTIONS

1. **PBS**

NaCl	0.08 g
NaH_2PO_4	0.12 g
KH_2PO_4	0.02 g
Water	100.00 ml

2. **PBS/PVP**

PVP	3.00 g
PBS	100.00 ml

3. **Speed prep solution**

Tris HCl	50.00 mM
Triton X-100	4.00 %
LiCl -	2.50 M
EDTA	62.50 mM

4. **8 % Acrylamide gel (total volume 50 ml)**

40 % Acrylamide /2 % bisacrylamide solution (29:1)	6.25 ml
5X TBE	10.00 ml
Urea	20.00 g
Water	15.00 ml
Ammonium per sulfate (10%)	400 µl
TEMED	25 µl

5. **6 % Acrylamide gel (total volume 50 ml)**

38 % Acrlamide /2 % bis-acrylamide solution (29:1)	6.25 ml
5X TBE	10 ml
Urea	20 g
Water	15 ml
Ammonium per sulphate (10 %)	400 µl
TEMED	20 µl