GENERATION OF MOLECULAR SIGNATURES *FOR ALNUS NEPALENSIS* GENOTYPES WITH HIGH NITROGENASE ACTIVITY IN SYMBIOTIC ASSOCIATION WITH *FRANKIA*

ABSTRACT

By
VINEETA CHAUHAN

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

NORTH EASTERN HILL UNIVERSITY
SHILLONG-793 022, INDIA
2000

ABSTRACT

Kawashima *et al.* (1997) estimated that by the middle of the twenty first century the world fertilizer demand would increase three folds. This would in turn accelerate environmental deterioration further. Environmentalists are, therefore, searching for ways in which biological nitrogen fixation can be exploited to make agricultural practices more sustainable and ecologically sound.

Biological nitrogen fixation in nature is carried out by two groups of prokaryotic organisms, the free-living bacteria like *Klebsiella*, *Clostridium*, cyanobacteria, etc. Others live in intimate symbiotic association with plants or with other organisms like the Protozoans. Examples of this symbiotic association are the *Rhizobium*-legume symbiosis and the *Frankia*-actinorhizal symbiosis.

The actinorhizal symbiosis derives from an association between the filamentous, gram positive bacteria *Frankia* and an array of host woody dicotyledonous plants belonging to twenty five genera spread over eight families (Lechevalier, 1994). The eight actinorhizal families are Betulaceae, Casuarinaceae, Coriariaceae, Datiscaceae, Elaeagnaceae, Myricaceae, Rhamnaceae and Rosaceae. As against the super-specialized host range of *Rhizobium*, confined to the legumes, *Frankia* is a more versatile and dynamic organism having a wide host range.

The fixation of nitrogen resulting from the *Frankia*-actinorhizal symbiosis is central to the dynamics of several ecosystems, many of the host plants being main components of early successional communities established on poor nitrogen depleted sites. These pioneer colonizers have played a major role in increasing the fertility of nitrogen depleted soils generated due to disturbances such as flooding, landslides, fire, glacial activity, mining, volcanic eruption, etc. (Schwintzer and Tjepkema, 1990; Lumini *et al.*, 1994). Apart from

this the actinorhizal plants have tremendous economic potential as timber and fuel and in forestry and amenity planting.

Actinorhizal trees are widely distributed. They are primarily found in the polar regions and normally grow in the temperate zones. In India these plants are found at high altitudes in Himachal Pradesh, Jammu and Kashmir, Arunachal Pradesh, Sikkim, West Bengal hills and Meghalaya. Some are also found in coastal regions and plains. The genera found in India are *Alnus*, *Casuarina*, *Myrica*, *Elaeagnus*, *Coriaria* and *Hippophae*.

Alnus, commonly known as alder, belongs to the family Betulaceae comprising of six genera (*Alnus*, *Betula*, *Carpinus*, *Corylus*, *Ostrya* and *Ostrypsis*) (Lawrence, 1967). Of the six genera only alder can symbiotically associate with the actinomycete *Frankia* and thus form root nodules where atmospheric nitrogen is fixed. Forty seven species of *Alnus* have been identified so far (Swensen and Mullin, 1997). Only two species of alder, *Alnus nepalensis* and *Alnus nitida*, are found in India. *Alnus nepalensis* is found in the temperate Himalayas at an altitude of 3000-9000 ft., from Chamba to Mishmi Hills in Upper Assam. Khasi Hills, Ava, on the Khaken Hills and Kurz (Hooker J.D., 1825). *Alnus nitida* is found in the western temperate Himalayas at an altitude of 5000-9000 ft., from Kashmir to Kunawar, sometimes descending into the plains along the river banks (Hooker, 1825).

Both trees of *Alnus* found in India are deciduous, have alternately arranged leaves, are monoecious and are anemophilous. However, the two trees can be morphologically differentiated. *Alnus nepalensis* has glabrous branchlets and the fruiting spikes are in large erect panicles. *Alnus nitida* is a larger tree attaining about 100 ft. with a trunk of about 15ft. girth. They have pubescent branchlets, the tree bark is lighter, fruiting spikes are solitary or in short erect racemes and they bear larger female cones (Hooker, 1825).

Although the *Frankia*-actinorhizal symbiosis compares well with the *Rhizobium*-legume symbiosis in the quantum of nitrogen fixed (Shantaram and Mattoo, 1997), has a number of ecological and economic benefits and a wide host range, studies on this symbiosis have not made comparable strides. They actually took off after the first successful isolation of *Frankia* strain CpI1 from *Comptonia peregrina* by Callaham and coworkers in 1978 (Callaham *et al.*, 1978). Subsequently many more strains have been isolated. Despite this the progress in this field of research has been very slow. This may be attributed to the slow growth rate of *Frankia* and the lack of suitable and efficient techniques for its isolation in pure culture. However, research on *Frankia*-actinorhizal symbiosis has largely benefited from the advances in Molecular Biology and made considerable progress ever since.

Like any other symbiotic association, in the *Frankia* actinorhizal association too, the host has a prominent effect on the efficiency of symbiosis (Simon *et al.*, 1985; Prat, 1989; Verghese and Misra, 2000). However, so far, more emphasis has been laid on studies related to the microsymbiont and the host remains an enigma. Apparently the differences in plant genotype with respect to promiscuity towards compatible *Frankia* result in efficiency differences. This has also been borne out by the work of Sougoufara *et al.* (1992), who, when experimenting with a combination of *Casuarina* plant clones and *Frankia* strains, found that a particular clone was always a better nitrogen fixer irrespective of the *Frankia* strain used. Among the *Frankia* strains too, one particular strain was always a more efficient nitrogen fixer. Thus, they have predicted that nitrogen-fixing efficiency can be improved by screening for the best host and then the *Frankia* strain associated with it.

To identify the best *Alnus nepalensis* clone X *Frankia* strain combination the diversity of alder compatible *Frankia* strains have been studied in our lab. Definite molecular signatures have also been identified for a few strains (Sarma, 1999). Before the best partnership can be identified, a similar study to identify the best host is also necessary.

Alder, thought to have originated in the Indo-China region (Furlow, 1979), is an open pollinated tree and has a lot of genetic heterogeneity in forest populations. Conventional screening methods for superior hosts would be time consuming and difficult. Therefore, the main objectives of the present study were:

1. Screening trees with reference to higher nitrogenase activity in nodulated condition.
2. Fingerprinting of plants.
3. Identifying molecular markers for early screening of nursery plants.

To achieve these objectives, the following approach was followed:

1. The following two study sites were selected-
    a. Upper Shillong, and
    b. NEHU Permanent Campus, Mawlai, Shillong.

2. Fifty trees were demarcated at each of these sites.

3. One hundred fifty nodules were collected from each of the trees and ARA was performed to assess their nitrogenase activity.

4. Young leaves were collected from each tree and total genomic DNA was isolated.

5. Polymerase Chain Reaction based amplification of chloroplast 16S rRNA gene and the *trn T* and *trn L* region was done. Amplifications of nuclear 18S-28S rRNA *rrn* region, including the ITS1 and ITS2, were also done.

6. Amplified fragment length profiles were studied.

7 Sequencing of amplified nuclear 18S-28S rRNA region of a few samples was done.

8. These sequences were aligned with related sequences retrieved from the GenBank. Phylogenetic positioning of these three samples was done with respect to retrieved related sequences.

9 The aligned sequence information was utilized to design a genus specific probe as a first step towards developing PCR primers for superior genotypes.

10 Restriction digestions with appropriate enzymes were done and the RFLPs were screened.

11 To look for any correlation, the amplification and restriction fragment length profiles were compared with the nitrogenase activity as reflected by ARA.

The results obtained in the present study can be summarized as below:

1. Variability was observed in the proximal part of the chloroplast 16S rRNA gene. AFLP analysis of this region could categorize *Alnus nepalensis* into three groups. RFLP analysis of the same region also revealed three groups. However, the three groups due to RFLP were independent of those due to AFLP. This is to say that all the trees belonging to one group of AFLP were not necessarily members of the same RFLP group.
2. AFLP and RFLP analysis of the middle part of the 16S rRNA gene did not reveal any diversity. Therefore, it is highly conserved region and was not suitable for detecting intra-specific variability.
3. A lot of genetic variability was observed in the 18S-28S ITS region. AFLP analysis could divide the alder trees into nine different groups (A-I). On the basis of RFLP analysis also the alder trees could be divided into nine different PCR-RFLP (Sc1-Sc9) groups. However, like the proximal part of the 16S rRNA gene, the groups identified were independent of each other.
4. Nucleotide sequencing was done for three representative samples of *Alnus nepalensis*. When aligned sequences were subjected to the phylogenetic studies, *Alnus nepalensis* was found closest to the European species *Alnus glutinosa*.

5. The genus specific probe in the distal region of the 18S rRNA gene was designed on the basis of the aligned nucleotide sequences of related genera. This primer amplified the DNA of *Alnus nepalensis* but did not amplify at the specified conditions the DNA isolated from *Elaeagnus* , *Myrica* , *Betula* , *Carpinus* and mixed fungal cultures.
6. RFLP analysis of the 18S-28S *rrn* operon revealed that the molecular marker (Sc3) that could be used to weed out genotypes supporting lower nitrogen fixing rates.
7. Taking all the PCR-AFL/RFL profiles together, two groups could be identified as likely molecular markers for *Alnus nepalensis* genotypes supporting high nitrogenase activity.
8. Further analysis of a bigger fragment of DNA is needed for developing definite positive markers. This can be accomplished by designing an appropriate primer which in conjunction with the primer designed in the present study, would give a larger amplicon, facilitating analysis of a larger region.

REFERENCES

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.

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

Hooker J.D. (1825). The Flora of British India. M/S Bishen Singh and Mahendra Pal Singh, Dehradun. 5: 600.

Kawashima H., M.J. Bazim and J.M. Lynch. (1997). A modeling study of world protein supply and nitrogen fertilizer demand in the 21st century. Environmental Conservation: 24 (1): 50-56.

Lawrence G.H.M. (1967). Taxonomy of vascular plants. Oxford and IBH Publ. Co., New Delhi.

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

Lumini. E., M. Bosco, G. Puppi, R. Isopi, M. Frattegiani, E. Buresti and F. Favilli. (1994).Field performance of *Alnus cordata* Loisel (Italian Alder) inoculated with *Frankia* and VA-Mycorrhizal strains in mine-spoil afforestation plots. Soil Biol. Biochem. 26: 659-661.

Prat D. (1989). Effect of some pure and mixed *Frankia* strains on seedling growth in different *Alnus* species. Plant and Soil. 113: 31-38.

Sarma G. (1999). Use of molecular characteristics in *Frankia* for strain identity. Ph.D. thesis, NEHU, Shillong, India.

Schwintzer C.R. and J.D. Tjepkema. (1990). The Biology of *Frankia* and Actinorhizal plants. Academic Press, California, USA

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.

Simon. L., A. Stein and M. Lalonde. (1985). Performance of *in vitro* propagated *Alnus glutinosa* (L) Gaertn: Clones inoculated with *Frankia*. Plant and Soil. 105: 33-40.

Sougoufara B., L. Maggia, E. Duhoux and Y.R. Dommergues. (1992). Nodulation and N_2 fixation in nine *Casuarina* clone-*Frankia* strain combinations. Acta. Oecologica. 13 (4): 497-503.

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 symbiosis. Physiol. Plantarum 99: 565-573.

Verghese S.K. and A.K. Misra. (2000). PCR-RFLP based screening of *Frankia* in alder nodules having different levels of nitrogenase activity. Symbiosis: 28: 337-350.

GENERATION OF MOLECULAR SIGNATURES *FOR ALNUS NEPALENSIS* GENOTYPES WITH HIGH NITROGENASE ACTIVITY IN SYMBIOTIC ASSOCIATION WITH *FRANKIA*

By
VINEETA CHAUHAN

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

NORTH EASTERN HILL UNIVERSITY
SHILLONG-793 022, INDIA
2000

Dedicated to the memory of my father

(L) Prof. Y.S. Chauhan
(1942-1999)

The North Eastern-Hill University

December, 2000

DECLARATION

I, Vineeta Chauhan, hereby declare that the subject matter of this thesis entitled "Generation of Molecular signatures for *Alnus nepalensis* genotypes with high nitrogenase activity in symbiotic association with *Frankia*" 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 anybody 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.

Vineeta
(Vineeta Chauhan)

11/12/00
(Head of The Department)

(Supervisor)

ACKNOWLEDGEMENTS

A lot of patience, hard work, perseverance and help (by way of support and encouragement) goes into the completion and compilation of a doctoral work into the form of a thesis. My work would not be complete without acknowledging them who have made this thesis attainable.

First and foremost, my profound indebtedness goes to my supervisor Dr. A.K. Misra, Department of Botany, NEHU, Shillong, for his indefatigable supervision, consistent guidance, incessant encouragement, unwavering enthusiasm and the most needed critical comments. He has indeed proven to be an incomparable friend, mentor and guide.

I am grateful to Prof. H.N. Pandey (Head), and Prof. R.R. Misra (Former Head) of the Department of Botany, for making available to me all the facilities and amenities necessary for befitting execution of my research work.

I wish to take this opportunity to thank all my teachers especially Dr. N.K. Chrungoo, for their unconditional help and useful suggestions. I duly thank Prof. P. Tandon for allowing me access to the facilities in his laboratory.

I am deeply indebted to Dr. Philippe Normand, University of Claude-Bernard, France, for allowing me access to his unpublished data and useful suggestions whenever needed.

I extend my thankfulness to my seniors Dr. Arnab Sen, Dr. Gopa Sarma, Dr. Rajani Jacob, and Dr. Susamma Verghese for their kind help and support.

I take this opportunity to thank all my friends especially Ms. Sanghamitra Purkayastha, Ms. Sangeeta Bharali, Ms. Swapna Prabhu, Mr. Hiranjit Choudhary, Dr. Babu John, Mr. Chitta Ranjan Deb and Ms. Lalrinkimi for their invaluable help and co-operation at various stages of my work.

I would like to thank all the members of the non-teaching staff of the Department, especially Mr. Civil Marak for every help rendered during the course of my study.

My abounding thanks are especially due to all my family members especially my mother, brother, sister and Ajay for their patience, forbearance, consistent encouragement and unstinted support.

I humbly express my over-brimming gratitude to my father (L) Prof. Y.S. Chauhan, who has been an abiding source of inspiration for me from the beginning of my academic pursuits. His absence today has created a void but his perennial blessings shall always be my intellectual and spiritual strength.

CONTENTS

LIST OF FIGURES

INTRODUCTION

CHAPTER 1

INTRODUCTION

It is well known that the nitrogen cycle depends on the ability of certain groups of prokaryotes to reduce 'fix' atmospheric nitrogen into a form that is accessible to other living organisms. It is also well known that the intensive modern agriculture is dependent on the input of large quantities of industrial nitrogen fertilizers. It is estimated that if the world's arable land remains constant, the demand for nitrogen fertilizer will increase and reach 220 Tg y^{-1} by the middle of the twenty first century (Kawashima *et al.*, 1997). This is approximately three times more fertilizer than is currently used. Since we are already facing sub-soil water pollution, an increased use of nitrogen fertilizers would accelerate environmental deterioration further. The world would suffer from the conflict between the supply of sufficient protein to ever increasing human population and greater nitrogen pollution of the environment. Environmentalists are, therefore, searching for ways in which biological nitrogen fixation can be exploited to make agricultural practices more sustainable and ecologically sound. In principle, the potential exists to satisfy the fixed nitrogen needs of nearly all crop plants by the proper use of nitrogen-fixing bacteria or the nitrogenase genes that they contain. The use of biological nitrogen fixation can help meet the needs of the present without compromising the needs of the future. Biological nitrogen fixation can also be gainfully employed for regenerating degraded forests, recovering nitrogen depleted soils, mine spoils and jhum fallows, etc.

Biologically the fixation of dinitrogen is carried out by two groups of prokaryotic organisms, the free-living bacteria like *Klebsiella, Clostridium*, cyanobacteria, etc. Others live in intimate symbiotic associations with plants or with other organisms like the Protozoans. The more acclaimed *Rhizobium*-legume symbiosis and the *Frankia*-actinorhizal plants symbioses are examples of symbiotic associations of nitrogen fixing bacteria with plants. *Rhizobium* and *Frankia* possess the nitrogenase enzyme complex that breaks the nitrogen triple bonds, thus making the almost inert form of nitrogen usable. It has been reported that *Frankia* contributes 2-362 Kg N ha $^{-1}$y $^{-1}$ and the contribution of *Rhizobium*-legume symbiosis is about 24-584 Kg N ha $^{-1}$y $^{-1}$ (Shantharam and Mattoo, 1997). Thus, actinorhizal plants rival legumes in the amount of nitrogen that they fix on a global basis, yet knowledge of their biology and uses is for the most part very recent and scarce.

The microsymbiont that is able to form nodules on the roots of non-leguminous dicotyledonous plants has been classified in the genus *Frankia* of the order Actinomycetales (Tjepkema and Torrey, 1979). So far about 25 genera of actinorhizal plants belonging to eight families have been found to house *Frankia* as a microsymbiont (Lechevalier, 1994). The eight actinorhizal families are the Betulaceae, Casuarinaceae, Coriariaceae, Datiscaceae, Elaeagnaceae, Myricaceae, Rhamnaceae and Rosaceae.

Actinorhizal trees are widely distributed. They are found in every continent except the Antarctica. They are primarily found in the temperate zone. Only some species of Casuarinaceae and Myricaceae can be considered truly tropical. Actinorhizal plants are abundantly found nearer the polar regions such as Scandinavia, Canada and New Zealand. In India these plants are found at high altitudes in Himachal Pradesh, Jammu and Kashmir, Arunachal Pradesh, Sikkim, West Bengal hills and Meghalaya. Some are also found in the coastal regions and plains of India. The genera found in India are *Alnus*, *Casuarina*, *Myrica*, *Elaeagnus*, *Coriaria* and *Hippophae*.

Alnus, commonly known as alder, belongs to the family Betulaceae comprising of six genera (*Betula, Alnus, Carpinus, Corylus, Ostrya and Ostrypsis*) (Lawrence, 1967). Of these six genera only *Alnus* forms symbiotic association with *Frankia*. Two species of alder, *Alnus nepalensis* and *Alnus nitida* are found in India. *Alnus nepalensis* is found in the temperate Himalayas at an altitude of 3000-9000 ft., from Chamba to Mishmi Hills in

Upper Assam, Khasi Hills, Ava, on the Khaken Hills and Kurz (Hooker, 1825). *Alnus nitida* is found in the western temperate Himalayas at an altitude of 5000-9000 ft., from Kashmir to Kunawar, sometimes descending into the plains along the river banks (Hooker, 1825).

Actinorhizal plants have tremendous economic potential as timber and fuel wood and in forestry, biomass production, land reclamation and amenity planting. Continued urbanization is resulting in rapid deforestation that has a number of hazardous consequences, such as a rise in the carbon-dioxide concentration, resulting in an increase in global temperature. Hence, the utilization of fast growing, nitrogen fixing trees for reforestation becomes critical. These non-leguminous nitrogen fixers are known to be colonizers of nitrogen-poor, open sites and are, therefore, very useful pioneer colonizers on nitrogen depleted soils generated due to disturbances such as flooding, landslides, fire, glacial activity, volcanic eruption, mining, etc. Mixed stands of *Alnus cordata* and *Elaeagnus* and broad-leaved trees have been successfully used since 1980 for re-vegetation of more than 200 ha of mine spoils in the lignite banks in Santa Barbara, Italy (Lumini *et al.*, 1994). *Alnus glutinosa* has been extensively used for reclamation of mine spoils in Britain (Schwintzer and Tjepkema, 1990). *Alnus* releases certain organic compounds such as phenols, fatty acids and amino acids in soil. These substances may stimulate the growth of free-living nitrogen-fixing organisms. These secretions may also reduce soil pathogens, as exemplified by *Poria weirii*, a long-lived and destructive parasite of conifer roots in North America and Japan (Sprent and Sprent, 1990). *Alnus nepalensis* regenerates naturally on the landslide-affected freshly exposed and degraded sites. It is grown in forestry, agroforestry, shifting cultivation in north-eastern India, and as nurse tree in *Cinchona* plantations. It is being widely used recently for reclamation of degraded habitats in eastern and central Himalayas and in areas where there are perennial landslides to check erosion and land-slips (Sharma *et al.*, 1998). Department of Forest, Government of Meghalaya, also has used alder plantings in Cherrapunjee area for the reclamation of mine spoils. Another use of the actinorhizal symbiosis is in studies investigating the basic biology of nitrogen fixing symbiosis as there are a number of developmental and genetic parallels between *Frankia* and *Rhizobium*.

Although the *Frankia*-actinorhizal symbiosis compares well with the *Rhizobium*-legume symbiosis in the quantum of nitrogen fixed, has a number of ecological and economic benefits and a wide host range, studies on this symbiosis have not made comparable strides. They actually took off after the first successful isolation of *Frankia* strain CpI1 from *Comptonia peregrina* by Callaham and coworkers in 1978 (Callaham *et al.*, 1978). Subsequently many more strains have been isolated. Despite this the progress in this field of research has been very slow. This may be attributed to the slow growth rate of *Frankia* and the lack of suitable and efficient techniques for its isolation in pure culture. Initial studies on this association were based on morphology and host infection processes. Later, different typing methods like serotyping, chemotyping with sugars or fatty acids and electrophoretic typing methods including SDS-PAGE analysis and isoenzyme analysis, proved to be useful techniques for *Frankia* strain characterization (Jamann *et al.*, 1993). Within the past few years, molecular genetic approaches have become increasingly important to studies in behavioral ecology and population biology. Polymerase Chain Reaction (PCR) based techniques of DNA fingerprinting have revolutionized approaches to our understanding of *Frankia*- actinorhizal symbiosis.

PCR, an *in vitro* method for the enzymatic synthesis of specific DNA sequences, allows a million fold amplification of the DNA fragment of interest (Mullis *et al.*, 1986). The target region to be amplified is fixed using two oligonucleotide primers that hybridize to opposite strands and flank the region of interest in the target DNA. Initially Klenow fragment of *E. coli* DNA polymerase I was used for the amplification of DNA. With the isolation of thermostable *Taq* DNA polymerase from *Thermus aquaticus,* PCR became a simple and robust reaction that could be automated by a thermal cycling device.

PCR based techniques have enormous potential in research for detection of naturally occurring DNA polymorphism. This polymorphism can be detected by scoring band presence versus absence in banding patterns that are generated by either restriction enzyme digestion or DNA amplification procedures, or both. Amplified Fragment Length Polymorphism (AFLP), which is based on the selective amplification of genomic restriction fragments by PCR, in essence combines the power of RFLP analysis with the flexibility of PCR-based technology (Janssen *et al.*, 1996). Restriction fragment length profiles obtained by the digestion of genomic DNA and PCR amplified products have helped characterize a number of organisms (Jamann *et al.*, 1993; Laguerre *et al.*, 1996; Verghese and Misra, 2000). Often microsatellite loci are amplified using microsatellite primers that are two-three base pairs long repetitive sequences. However, all these methods for the detection of polymorphism in the DNA require relevant sequence information for the designing of appropriate primers or even the selection of a suitable restriction endonuclease. Genetic polymorphism can also be detected by the analysis of variable number of tandem repeats (VNTR) (Boerwinkle *et al.*, 1989). Use of random amplified polymorphic DNA (RAPD) markers, detected by PCR amplification of small inverted repeats throughout the genome, adds a new technology of DNA fingerprinting to the molecular analysis of relatedness between genotypes. RAPD analysis is conceptually simple. Nanogram amounts of total genomic DNA is subjected to PCR using short (10-mer) oligonucleotides of random sequence. It differs from standard PCR conditions in that only a single random oligonucleotide primer is required and no prior knowledge of the genome to be analyzed is required (Hadrys *et al.*, 1992). DNA Amplification fingerprinting (DAF) and Arbitrary Primed Polymerase Chain Reaction (AP-PCR) are modifications of RAPD approach. DAF uses short random primers of 5-8 bp to amplify short stretches of DNA from a target genome (Caetano-Anolles and Gresshoff, 1994; Caetano-Anolles and Gresshoff, 1996). Mini-hairpin primers with short arbitrary cores and primers complementary to simple sequence repeats present in microsatellites have been used to generate arbitrary signatures from amplification profiles (ASAP) (Caetano-Anolles and Gresshoff, 1994; Caetano-Anolles and Gresshoff, 1996). AP-PCR uses slightly longer primers such as universal M13 and amplification products are radiolabelled (Welsh and McClelland, 1990). The amplicons produced are best resolved by polyacrylamide gel electrophoresis and silver staining (Bassam *et al.*, 1991). Sheffield *et al.* (1989) showed that 40-45 bp GC-clamps can be attached to amplified cloned and genomic DNA fragments by using PCR, thereby allowing the separation of single-base mutations by denaturing gradient gel electrophoresis (DGGE). Asymmetric PCR, Alu-PCR, Inverse-PCR, RT-PCR, Anchored-PCR and many more PCR based techniques have proved to be valuable tools for the detection of polymorphism in DNA. PCR has even been used to study organisms that are difficult to isolate in pure culture (Misra *et al.*, 1991) or where the isolation of sufficient amount of DNA is difficult. The amplified products can finally be sequenced to confirm the differences. These sequences can then be used to infer phylogenetic relationships. Consequently, there are a myriad of applications of PCR in recombinant DNA technology research and its utility continues to increase everyday.

Research on *Frankia*-actinorhizal symbiosis has largely benefited from these advances in Molecular biology. PCR-RFLP based techniques have revealed tremendous diversity among *Frankia* strains (Nazaret *et al.*, 1991; Varghese, 2000; Verghese and Misra, 2000). Molecular systematic studies of various chloroplast and nuclear genes have in a large way helped resolve the phylogenetic relationships among the taxonomically diverse actinorhizal plants (Savard *et al.*, 1993; Maggia and Bousquet, 1994; Swensen and Mullin, 1997).

Like any other symbiotic association, in the *Frankia* actinorhizal association too, the host has a prominent effect on the efficiency of symbiosis (Simon *et al.*, 1985; Prat, 1989; Verghese and Misra, 2000). However, so far, more emphasis has been laid on studies related to the microsymbiont and the host remains an enigma. Apparently the differences in plant genotype with respect to promiscuity towards compatible *Frankia* result in efficiency differences. This has also been borne out by the work of Sougoufara *et al.* (1992), who, when experimenting with a combination of *Casuarina* plant clones and *Frankia* strains, found that a particular clone was always a better nitrogen fixer irrespective of the *Frankia* strain used. Among the *Frankia* strains too, one particular strain was always a more efficient nitrogen fixer. Thus, they have predicted that nitrogen-fixing efficiency can be improved by screening for the best host and then the *Frankia* strain associated with it.

To identify the best *Alnus nepalensis* clone X *Frankia* strain combination the diversity of alder compatible *Frankia* strains have been studied in our lab. Definite molecular signatures have also been identified for a few strains (Sarma, 1999). Before the best partnership can be identified, a similar study to identify the best host is also necessary.

Alder, thought to have originated in the Indo-China region (Furlow, 1979), is an open pollinated tree and has a lot of genetic heterogeneity in forest populations. Conventional screening methods for superior hosts would be time consuming and difficult. The objectives of the present study are presented below.

Objectives

1. Screening trees with reference to higher nitrogenase activity in nodulated condition.
2. Fingerprinting of plants.
3. Identifying molecular markers for early screening of nursery plants.

To achieve these objectives, the following approach was adopted:

1. Two study sites were selected-
    a. Upper Shillong, and
    b. NEHU Permanent Campus, Mawlai, Shillong.

2. Fifty trees were demarcated at each of these sites.

3. One hundred fifty nodules were collected from each of the trees and ARA was performed to assess their nitrogenase activity.

4. Young leaves were collected from each tree and total genomic DNA was isolated.

5. Polymerase Chain Reaction based amplification of chloroplast 16S rRNA gene and the *trn T* and *trn L* region was done. Amplifications of nuclear 18S-28S rRNA *rrn* region, including the ITS1 and ITS2, were also done.

6. Amplified fragment length profiles were studied.

7. Sequencing of amplified nuclear 18S-28S rRNA region of a few samples was done.

8 These sequences were aligned with related sequences retrieved from the GenBank. Phylogenetic positioning of these three samples was done with respect to retrieved related sequences.

9. The aligned sequence information was utilized to design a genus specific probe as a first step towards developing PCR primers for superior genotypes.

10. Restriction digestion with appropriate enzymes were done and the RFLP's were screened.

11. To look for any correlation, the amplification and restriction fragment length profiles were compared with the nitrogenase activity as reflected by ARA.

REVIEW
OF
LITERATURE

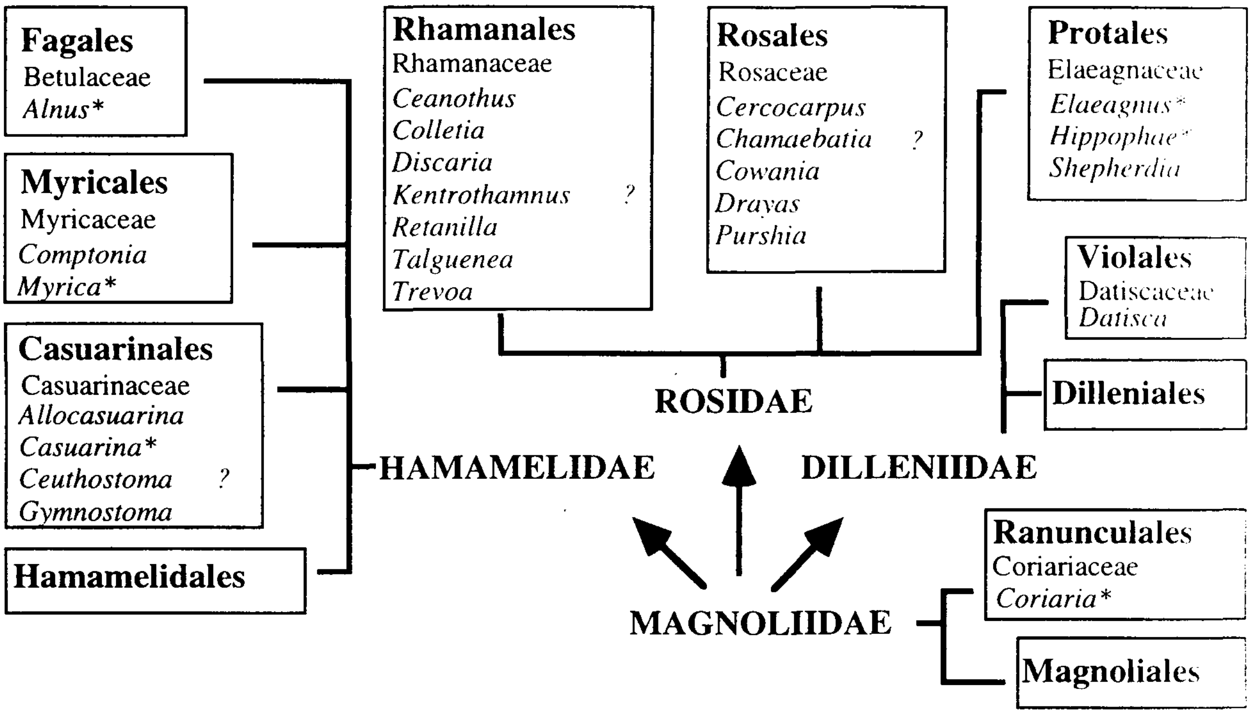
CHAPTER 2

REVIEW OF LITERATURE

The actinorhizal symbiosis may be viewed as a dynamic biological system in which both the host plant and the actinomycete *Frankia* interact in such a way that a selective advantage is conferred to the two partners. An appropriate approach to a study of this symbiosis is to look at the system as one entity and to consider individually the host and the microsymbiont, with their multiple interactions, so that a better understanding of the coadaptive significance of the symbiosis may be gained. Considerable efforts have been devoted during the last twenty years to attain a better knowledge of the biology and genetics of the microsymbiont. Efforts directed towards the genetics of the host are still in the preliminary stage and have a long way to go.

2.1 THE ACTINORHIZAL TREES : THE HOSTS

Actinorhizal trees are perennial dicots and all are woody shrubs or trees except *Datisca*, which has herbaceous shoots (Tjepkema *et al.*, 1986).They symbiotically associate with the actinomycete *Frankia* and form root nodules where atmospheric nitrogen is fixed. These plants are dispersed among eight dicotyledonous families and about twenty five genera. The eight actinorhizal families are Betulaceae, Casuarinaceae, Coriariaceae, Datiscaceae, Elaeagnaceae, Myricaceae, Rhamnaceae and Rosaceae (Fig. 2.1). Of the six genera belonging to Betulaceae (*Betula, Alnus, Carpinus, Corylus, Ostrya* and *Ostrypsis*), only *Alnus* forms symbiotic association with *Frankia.* About forty seven species of alder have been identified so far (Swensen and Mullin, 1997). Of these only two species, *Alnus nepalensis* and *Alnus nitida* are found in India. Both these trees are deciduous, have alternately arranged leaves, are monoecious and are anemophilous. However, these two



Taxonomic position of actinorhizal genera in the Magnoliopsida (dicotyledons) according to Cronquist (1988). Subclasses are shown in bold uppercase, orders in bold lowercase, families in normal lowercase and genera in italics. The genera found in India are marked with *.

Fig 2.1

trees can be distinguished from each other on the basis of their morphology. *Alnus nepalensis* has glabrous branchlets and the fruiting spikes are in large erect panicles. *Alnus nitida* is a larger tree attaining about 100 ft. height with a trunk of about 15ft. girth. They have pubescent branchlets, the tree bark is lighter, fruiting spikes are solitary or in short erect racemes and they bear larger female cones (Hooker, 1825).

Alnus nepalensis, found in the Khasi hills of Meghalaya, Sikkim, Darjeeling, Arunachal Pradesh and Nagaland, is a pioneer species in freshly exposed landslide soils. It grows on sandy eroded soils, denuded habitats, rocky slopes, landslide affected slopes, steep stream sides and in natural forests. It has been considered a useful species in social forestry and agroforestry. Shifting agriculture or "jhum" is a major economic activity in north-east India. Management of fallow species in jhum system, between two cultivations, has become an important issue for sustaining this traditional practice. *Alnus nepalensis* has been an important fallow species in the jhum system that is traditionally valued and conserved by jhum farmers (Ramakrishnan, 1994). The economic value of these trees could be further exploited by improving their nitrogen fixing efficiency. This calls for screening the genetic diversity of the microsymbiont *Frankia* and the host and then selecting superior strains of *Frankia* and the corresponding host genotype.

2.2 THE MICROSYMBIONT: *FRANKIA*

The genus *Frankia*, so named by Brunchrost in 1886 (Lechevalier and Lechevalier, 1989) had been classified into the family Frankiaceae of the order Actinomycetales together with the closely related genus *Geodermatophilus* (Hahn *et al.*, 1989). Later Normand *et al.* (1996), on the basis of 16S rDNA analysis, separated the two genera *Frankia* and *Geodermatophilus* into two separate families Frankiaceae and Geodermatophilaceae.

Frankia is a slow growing actinomycete with a doubling time of about five days (Bosco *et al.*, 1992). The members of the genus *Frankia* can be clearly distinguished from other bacterial genera on the basis of their host specificity, their morphology, cell biochemistry, physiology, DNA base composition and molecular characteristics. They possess the following characteristics (Bosco *et al.*, 1992, Normand *et al.*, 1996):

1. Gram-positive bacteria.
2. Capacity to fix nitrogen.
3. Host specificity.
4. Capacity to nodulate actinorhizal plants.
5. Presence of hyphae, specialized thick-walled organelles called diazovesicles, and multilocular sporangia containing non-motile spores.
6. A type III cell wall containing meso-diaminopimelic acid, glutamic acid, alanine, glucosamine and muramic acid.
7. Type PI phospholipids that include phosphatidylinositol, phosphatidylinositol mannosides and diphosphatidylglycerol.
8. The presence of sugar 2-O-methyl-D-mannose.
9. G+C content ranging from 66-72%.

The isolation of *Frankia* proved to be difficult for many years, until the first reproducible isolation of *Frankia* strain CPI1 from *Comptonia peregrina* (Callaham *et al.*, 1978). Since then hundreds of isolates have been obtained in pure culture from several plant species. However, single nodules are known to host more than one *Frankia* strain (Diem *et al.*, 1983, Clawson *et al.*, 1998). Therefore, cultures obtained from them may not be pure cultures but actually may be mixed. To overcome this problem, several attempts have been made to obtain genetically pure cultures of *Frankia*. Prin *et al.* (1991) and Lumini and Bosco (1996) have made purified *Frankia* strains by generating single spore cultures through plating. Sarma *et al.* (1998) have obtained genetically pure cultures through entrapment of spores in calcium alginate beads

2.3 HOST-MICROSYMBIONT INTERACTIONS:

The interaction between the host plant and the microsymbiont starts with signal exchange and recognition of the symbiotic partners. From this step until nodule formation and functioning, many genes from both partners participate in the process. Some of them are nodule-specific plant genes called the " actinorhizin genes" (Tremblay *et al.*, 1986), or more appropriately named as " actinorhizal nodulin genes " (Guan *et al.*, 1998). Poly(A) RNA isolation from *Alnus* has made it possible to construct cDNA libraries (Mullin *et al.*, 1993). By differential screening of nodule cDNA libraries, several cDNA clones representing actinorhizal nodulin genes have been isolated. The plant origin of these cDNAs has been confirmed by DNA gel blot hybridization, also showing that all of them are encoded by small gene families. Simonet *et al.* (1990) have suggested an enzyme-mediated entry of *Frankia* into the plant cell wall rather than a physically forced penetration. Thus, they have considered genes that code for enzymes involved in penetration as "nod" genes because they permit entry into the host. The genes like the *nif* genes, allow *Frankia* to live within the host cell and the "fix" genes enable it to fix nitrogen. However, the *nod* genes of *Frankia* differ from those of *Rhizobia*.

2.4 THE INFECTION PROCESS:

The infection process begins in the rhizosphere of the host tree, where, *Frankia* colonizes. *Frankia* can infect its host in two different ways, root hair infection and the intercellular penetration (Cournoyer *et al.*, 1993). The infection process is dictated by the host plant species (Guan *et al.*, 1998).

a) **Root Hair Infection:** Responding to unknown signals due to perturbations caused by rhizosphere *Frankia* , the root hairs get deformed. *Frankia* hyphae get trapped in the deformed root hair and start branching and initiate digestion of the primary cell wall of the root hair. The host plant in turn deposits cell wall material around the site of digestion. Continued digestion by the bacteria and triggered cell wall building by the host plant leads to the formation of a tubular ingrowth termed as encapsulation. Mitotic activity in the root cortex and cell expansion result in the formation of prenodule, whose cells are infected by *Frankia* hyphae. Subsequently, one or several nodule lobe primordia are initiated in the root pericycle. *Frankia* hyphae grow from the root cortex into the developing nodule lobe primordium through cortical cells and then proliferate.

b) Intercellular Infection: In certain cases the *Frankia* hyphae penetrate the root epidermis between the epidermal cells and colonize the root cortex intercellularly (Miller and Baker, 1985). No prenodule is formed, but the colonization of the cortex directly leads to the induction of nodule primordium in the root pericycle.

In most cases when the infected cells are filled with hyphae, *Frankia* begins to differentiate vesicles where nitrogenase is formed and nitrogen fixation occurs.

2.5 MORPHOLOGY OF THE ROOT NODULES:

Actinorhizal root nodules are coralloid structures composed of several modified lateral roots (lobes). The nodule lobes contain a central vascular system and infected cells in the expanded cortex. Since the nodule lobe meristem develops only in one direction, root caps are lacking. Nodule roots generally show no root hair with the exception of *Myrica* and occasionally as an ambiguity in *Alnus nepalensis* (Varghese, 2000). They have large air spaces which, serve as oxygen diffusion pathways to aerate the nodule tissue. Due to the presence of the apical meristem, actinorhizal nodule lobes have an indeterminate growth pattern, and their tissues are of graded age. Starting from the apical meristem four different zones have been defined in a mature nodule lobe (Guan *et al.*, 1998).

i. The meristematic zone or 'zone 1' consists of small dividing cells.

ii. The infection zone or 'zone 2' contains enlarging cortical cells some of which are infected with *Frankia* and hence, are larger than uninfected cells.

iii. The fixation zone or 'zone 3' characterized by formation of vesicles by the actinomycete, where *nif* genes are transcribed, nitrogenase enzyme is formed, and nitrogen fixation occurs.

iv. The senescence zone or 'zone 4', where the infected cells degenerate and degradation of host cytoplasm and *Frankia* hyphae occurs.

The appearance of *Frankia* vesicles in the infected alder root cortical cells is an indication of the onset of nitrogen fixation. Within the nodule tissue, the extent of nodule formation and the shape, septation and spatial organization of the vesicles, are controlled by the host plant. Symbiotic vesicles can be spherical, club-shaped, elliptical or filamentous. Another characteristic feature of *Frankia* is the production of large number of spores within sporangia produced as terminal appendages of the vegetative mycelium. At times the sporangia may be intercalary in position. The spores are non-motile and show variable rates of germination. However, production of spores *in planta* is not a feature common to all strains. Therefore, depending upon the presence or absence of spore formation, there are two kinds of strains, the Sp^+ and Sp^- strains. It has been reported that the Sp^- strains can sporulate in pure culture (Normand and Lalonde, 1982).

2.6 HOST SPECIFICITY:

Frankia strains have a wide host range and they exhibit definite host-specific responses. These host-specific responses have been used as a criterion for the classification of *Frankia* strains. Baker (1987), using pure cultures in cross-inoculation tests, has grouped *Frankia* strains into the following four host-infectivity groups:

i. Strains that infect *Alnus* and *Myrica* species.

ii. Strains that infect *Casuarina* and *Myrica* species.

iii. Strains that infect *Elaeagnus* and *Myrica* species.

iv. Strains that infect only *Elaeagnus* species.

Lalonde *et al.* (1988) using a more complex approach, relying on diverse phenotypic characteristics, described two species, *Frankia alni* and *Frankia elaeagni*. Later Margheri *et al.* (1989) and Bosco *et al.* (1992) reported that a few *Frankia* strains isolated from *Elaeagnus* species were able to cross boundaries between *Alnus* and *Elaeagnus* host specificity groups. Comparing the complete 16S rDNA sequences, and constructing the phylogenetic tree, Normand *et al.* (1996) proved that there was greater diversity among *Alnus*-infective strains than among strains infecting other hosts. They grouped *Frankia* into the following four groups:

i. A large group comprising *Frankia alni* and related organisms, including *Alnus rugosa* Sp^+ microsymbionts, to which *Casuarina* infective strains, a *Myrica nagi* microsymbiont and other *Alnus*-infective strains are related.
ii. Unisolated microsymbionts of *Dryas*, *Coriaria* and *Datisca* species.
iii. *Elaeagnus*-infective strains.
iv. "Atypical" strains, a group including *Alnus*-infective non-nitrogen fixing strains like Cn7 and Dc2 isolated from *Coriaria* and *Datisca* species respectively.

However, the rules for predicting host specificity have proven something of a Gordian knot and the practical concerns that must be addressed include the effect of plant cultivation methods on apparent competence, the incidence of multiple strains within a nodule and the effect of intra- and interspecific plant differences that may determine susceptibility. Studies have shown that different plant species within a genus or different ecotypes or clones of the same host plant can respond differently to nodulation by a given *Frankia* isolate (Benson and Silvester,1993).

2.7 MECHANISM OF NITROGEN FIXATION:

Nitrogen fixation by *Frankia* can be represented by the following equation:

$$N_2 + 8H^+ + 8e^- + 16ATP = 2NH_3 + H_2 + 16ADP + 16Pi$$

In this two moles of ammonia are produced from one mole of nitrogen gas, at the expense of 16 moles of ATP and a supply of electrons and protons (hydrogen ions). This reaction is performed using the nitrogenase enzyme. This enzyme consists of two proteins, an iron protein and a molybdenum-iron protein. The former is a dimeric protein with a 63 kDa molecular mass. The MoFe-protein consists of two subunits of 55 kDa and two subunits of 60 kDa molecular mass. The reaction occurs while the N_2 is bound to the nitrogenase enzyme complex. The iron protein is first reduced by electrons donated by ferredoxin, generated by photosynthesis. The reduced iron protein binds ATP and reduces the molybdenum-iron protein, which donates electrons to N_2, producing HN=NH. In two further cycles of this process, each requiring electrons donated by ferredoxin, HN=NH is reduced to $H_2N\text{-}NH_2$ which in turn is reduced to $2NH_3$.

2.8 CONSERVED AND ASSOCIATED VARIABLE REGIONS AND THEIR STUDY:

PCR-based techniques, which have revolutionized every field of research involve the amplification of DNA sequences. Since PCR needs DNA primers, it is more useful for the study of conserved genes. These genes are ubiquitously present in all organisms and are vital for cellular function. The sequence of these genes is more or less conserved throughout the living kingdom. Any sequence change in these genes represents an evolutionary transition. Such genes have regions of large sequence conservation and regions of large variability. The conserved regions, generally the coding regions, indicate close relationships among the various genera whereas analysis of the non-coding regions could extend the utility of the molecule at lower taxonomic levels. The variable zones tend to accumulate mutations more than the coding sequences. They accumulate insertions/ deletions at a rate at least equal to that of nucleotide substitutions. The highly conserved regions can serve as primer binding sites for *in vitro* amplification by PCR. The more conserved regions are also useful, serving as targets for universal probes that react with all living organisms, or for discriminating between broad phylogenetic groups. The more variable sequence regions are appropriate targets for genus-, species- and sometimes even subspecies-specific hybridization probes. Since these oligonucleotide probes form stable hybrids only with the complementary sequence, probe binding is highly specific. The most popularly used genes for the purpose today are the cytochrome c , 16S rRNA, *rbc*L , the *nif* genes in case of nitrogen fixing organisms and nuclear 18S-28S *rrn* operon.

2.8.1 The Chloroplast 16S rRNA Gene:

The 16S rRNA genes show a high degree of functional constancy and have a large size of about 1,500bp, consisting of many loosely held independent domains each of which may undergo sequence change independent of the other. The conserved nature of the 16S rRNA gene is exhibited by the fact that the chloroplast 16S rRNA gene is more or less like the bacterial 16S rRNA gene, except for a few differences like the following (Normand *et al.*, 1996):

i. A 'T' is present in most chloroplast sequences at coordinate 30, while a 'C' is present in all bacteria including Cyanobacteria.
ii. *Nru* I restriction site is present in most chloroplast 16S rRNAs but is absent in bacterial rRNAs.
iii. *Bcl* I site is present in non-chloroplast 16S rDNA.

The conserved and variable regions in the 16S rRNA gene have been used for the construction of many specific probes for the characterization of strains. Primers FGPS849 and 1176', designed by Nazaret *et al.* (1991), allow the amplification of a 325bp fragment of a number of gram-positive bacteria. These are universal prokaryotic primers and are located in a highly conserved region, the site where the mRNA pairs with the small sub-unit of the ribosome. Simonet *et al.* (1991) selected a highly variable region of the 16S rDNA and designed primers FGPS958 and 1093'. These primers were fairly specific to

Frankia and did not amplify other bacterial strains. However, it showed positive amplification with *Geodermatophilus*. In order to distinguish between *Frankia* and *Geodermatophilus* they designed primers FGPS1493 and FGPL2051' to amplify the ITS region separating the 16S and 23S rDNA. The amplicon produced as a result of amplification using these two primers could differentiate between the genera on the basis of the band size. Further work led to the designing of *Frankia* specific primers (Bosco *et al.*, 1992). Located in the helix 31 of domain III of the 16S rRNA gene, primer FGPS989 could differentiate between *Frankia* and *Geodermatophilus*. By examining the sequences in the same region Bosco *et al.* (1992) discovered mismatches in the sequences in *Alnus* and *Elaeagnus* host compatibility groups, thus leading to the designing of primers FGPS989ac and FGPS989e. The former for *Alnus* and *Casuarina* host compatibility groups and the latter for *Elaeagnus* host compatibility group. Normand *et al.* (1996) have designed primers like FGPS5, containing a *Hind* III site, FGPS485, FGPS505', FGPS910', and a number of others, for the amplification of *Frankia* DNA. However, these primers could not exclude the possibility of amplification caused due to contamination of chloroplast 16S rDNA. Therefore, these four primers have been used in the present study for the amplification of chloroplast DNA. On the basis of the differences between the 16S rRNA gene of the chloroplast and bacteria, Normand *et al.* (1996) designed primer FGPS4-330, containing a *Bcl* 1 restriction site. This enabled effective amplification of non-chloroplast 16S rDNA preferentially from plant tissues. Tohdoh and Sugiura (1982) designed primers N-16cpA and C-16cpK for the amplification of the entire 16S rRNA gene of tobacco which is about 1.4kb in length. Primers like N-16cpB, C-16cpB, N-16cpD, C-16cpD, N-16cpF,etc were designed for sequencing the 16S rRNA gene. Even though these primers were designed for tobacco they are expected to work for most species of plants (Bult *et al.*, 1992).

2.8.2 The *rbc*L Gene:

The chloroplast gene *rbc*L, ubiquitously present in all plants and encoding the large sub-unit of the carbon fixing enzyme, ribulose-1,5-biphosphate carboxylase (RUBISCO) in photosynthesis has received special interest in phylogenetic studies. It serves as a good molecular chronometer because of its large size (>1,400bp) and high functional constancy (Soltis *et al.*, 1990). Frascaria *et al.* (1993) aligning published ORF sequences of dicots, designed seven primers for sequencing. These primers have been numbered as 334^{+}, 691^{+}, 1144^{+}, 216^{-}, 537^{-}, 979^{-}, 1303^{-}, according to their location on either of the strands. Comparing the sequences of Hamamelideae, they designed two primers, upstream and downstream the coding region respectively. Upstream the primer 34^{+} was complementary to the 3'-5' strand starting at position 34. Downstream the primer 1517^{-} was complementary to the 5'-3' strand starting at position 1517. A number of primers have been designed by Zurawski *et al.* (1984). The *rbc*L gene being a highly conserved gene these primers can be used for amplification of a wide range of plant species.

The *rbc*L ORF sequences have been determined for three species of Casuarinaceae and two species of Myricaceae by Maggia and Bousquet (1994) and have been reported to be 1428 bp long. The degree of homology was greater than 98% within families of the 'higher' Hamamelideae. Between the families, the DNA homology was 97%, whether the families belonged to the same order or to different orders. The synonymous rate was 10-20 times larger than the non-synonymous rate, except for Myricaceae. These reports and many others (Savard *et al.*, 1993; Ludovic and Taberlet, 1994) prove that the *rbc*L gene is too conserved to clarify relationships between closely related genera. Its utility is limited to interordinal or interfamilial level. It has also been reported by Bousquet *et al.* (1992b) that the *rbc*L gene evolved more rapidly in annual plants than the perennial dicots and monocots, thus suggesting an extreme diversity of substitution rate for this gene, at least among angiosperms, and caution is needed in utilizing it as a molecular clock.

2.8.3 The Non-coding Regions of the Chloroplast DNA:

Analysis of the non-coding regions of the chloroplast DNA could extend the utility of the molecule at lower taxonomic levels as they have higher rate of accumulation of mutations. In most cases sequence divergence in the non-coding sequences are higher than *rbc*L, even when only substitutions are considered (Gielly and Taberlet, 1994). Taberlet *et al.* (1991), in order to design universal primers for such non-coding regions, studied the entire cpDNA sequence for tobacco, *Marchantia* and rice. They selected the region between the *trnT* (UGU) and *trnF* (GAA) genes in the large single copy regions because of the following reasons:

1. The succession of conserved *trn* genes and several hundred base pairs of non-coding regions.
2. The absence of gene rearrangements among tobacco, rice and Marchantia.
3. The higher rate of molecular evolution of single-copy regions.

They designed primers to match more conserved sequences of the tRNA genes, avoiding palindromic sequences. They also took care to avoid sequences which would lead to hybridization of the 3' end of the primers with the same tRNA genes in the nucleus and mitochondria. Keeping these facts in mind they designed the primers B48557, A49291, B49317, A49855, B49873, A50272. Using these primers in various combinations they could amplify the intergenic spacer between *trnT* (UGU) and the *trnL* (UAA) exon, the *trnL* (UAA) intron and another intergenic spacer between the *trnL* (UAA) exon and *trnF* (GAA). These primers amplify cpDNA over a wide taxonomic range including bryophytes, pteridophytes, gymnosperms and angiosperms. The sequence of the intergenic spacers could be very useful for evolutionary studies of related species and probably of populations of the same species. Intraspecific variation of the cpDNA occurs in *Gentiana burseri* where a three nucleotide substitution separates the two subspecies *burseri* and *villarsii* (Gielly and Taberlet, 1994). They have also reported that the *trnL* intron evolves between 1.93 times faster and the intergenic spacer evolves 11.72 times faster than the *rbc*L. The size of the *trnL* intron varies from 350bp (genus *Avena*) to 600bp (genus *Euphorbia*). However, the use of these non-coding regions is limited to intrafamilial or intraspecific level. Due to the accumulation of more number of mutations, it is difficult to align the sequences of distantly related taxa.

It is not easy to establish a rule for the choice of a particular region of the chloroplast genome for resolving plant phylogenies. The more conserved genes like the *rbc*L are more appropriate for molecular analysis at the interfamilial level and the non-coding regions which exhibit a lot of variability are useful for designing molecular markers at lower taxonomic levels.

2.8.4 The ITS Sequences of the Nuclear Ribosomal DNA:

The nuclear ribosomal DNA has proved to be a powerful phylogenetic tool because of its ubiquitous presence in all organisms and the presence of repeat units in high copy number (Hamby and Zimmer, 1991). These repeat units evolve at different rates, therefore, they can be used to examine lineages with different levels of divergence. The nuclear rDNA units, separated by intergenic spacers, consist of the 18S, 5.8S and 28S coding regions. The 5.8S coding region, flanked by internal transcribed spacers, ITS1 and ITS2, is located between the 18S and 28S coding regions. In order to address questions of generic relationships in Winteraceae, Suh *et al.* (1993) used the primers 18NRBLr and 26A designed by Hamby *et al.* (1988). Internal sequencing primers C58S and N58S, the complement of C58S, were designed by Suh *et al.*. (1993). The primer C58S begins at position 96 of the coding strand of the 5.8S gene. The amplicons produced using the above mentioned primers were sequenced. The rate of substitutions per year for ITS1 has been reported to be 3.2-5.2 x 10^{-10} and 3.6-5.7 x 10^{-10} for ITS2. A similar study was conducted for *Panax* using the amplification primers C26A and N-nc18S10 and sequencing primers N18L18, nc18S10, C26A, N5.8S and ITS2 AND ITS4 (Wen and Zimmer, 1996).

To elucidate the phylogenetic relationship and compare the relative amounts of genetic diversity between the genera *Alnus* and *Betula* of the family Betulaceae, the *rbc*L, 18S rDNA and ITS1 and ITS2 were analyzed (Savard *et al.*, 1993). Of the 1,400 positions of the 18S compared by them, 18 substitutions were observed between *Alnus glutinosa* and *Betula papyrifera* thus showing a homology of 98.4%. Amplification of the ITS1, 5.8S and ITS2 gave a band of 586bp, including 128bp for 5.8S, 221bp for ITS1 and 237bp for ITS2, and the average homology between *Alnus* and *Betula* was reported to be 86%. This is close to a 10-fold difference in terms of number of substitutions per site as compared to that of the 18S rDNA. Similarly a 20-fold difference in substitutions between the ITS and the *rbc*L gene has been reported.

Thus, the ITS sequences appear to be more informative and suitable than the 18S rRNA gene and the *rbc*L gene. However, they serve as better molecular chronometers at lower taxonomic levels because due to the number of differences in this region it is difficult to get proper alignment of the sequences of distantly related taxa. The other advantage of the accumulated variability in the ITS or IGS regions is that this could be exploited for developing specific molecular markers.

2.9 PHYLOGENETIC ANALYSIS OF THE HOST TREES:

2.9.1 Molecular Phylogeny of Actinorhizal Trees:

Morphological characteristics, paleobotanical evidence and DNA sequence information has been used to infer the phylogenetic relationships and evolutionary trends among the various actinorhizal genera. The most commonly used DNA sequences are that of the chloroplast gene *rbc*L and the nuclear 18S-28S rDNA region including ITS1 and ITS2. Chase *et al.* (1993) analyzed the *rbc*L gene of 499 flowering plants, including representatives from three subfamilies of legumes and eight other families where nitrogen fixing symbiosis is found (Betulaceae, Casuarinaceae, Coriariaceae, Datiscaceae, Myricaceae, Rhamnaceae, Rosaceae and Ulmaceae). On the basis of the results obtained, they placed the families with nodular nitrogen-fixing symbiosis together as part of a single clade consisting of a portion of Cronquist's dicotyledonous subclass Rosidae as well as members of the subclass Hamamelidae. This clade contained both rhizobial and actinorhizal symbiosis. However, their results were only preliminary because, due to the large size of the *rbc*L data set and the computational difficulties associated with their analysis, a complete analysis was not possible. Later Soltis *et al.* (1995), focusing entirely on higher plants engaged in root nodule symbiosis, confirmed the occurrence of all nitrogen fixing symbiosis along with several families lacking this symbiosis, as part of a single clade termed the 'nitrogen fixing clade'. This clade consists of six lineages, of which two do not contain any members known to host nitrogen-fixing bacteria. Since the ten families involved in symbiotic nitrogen fixation are only distantly related, morphologically based classification schemes proposed by taxonomists suggest a polyphyletic origin of the nitrogen fixing capacity. However, *rbc*L analysis suggests a monophyletic origin of the single large clade. All these families share a common ancestor not shared by other families of angiosperm. Those members of this clade which lack the ability to fix nitrogen could have lost the ability to form such associations during the course of evolution. Alternatively, the ancestor of the nitrogen fixing clade may have evolved the genetic components that would ultimately permit the evolution of symbiotic nodular nitrogen fixation. Following the establishment of these conditions, the necessary genetic background was present to allow parallel, recurrent evolution of symbiotic nitrogen fixation in the subsequent diversification of this clade (Soltis *et al.*, 1995).

Swensen (1996) comprehensively analyzed the *rbc*L gene and also the morphological and anatomical features of actinorhizal hosts and their nonactinorhizal relatives. They placed the actinorhizal taxa in a single large clade that included the *Rhizobium*-nodulated *Parasponia*, but excluded the legumes. This large clade was further divided into four subgroups (Clade I-IV), containing actinorhizal genera interspersed with nonactionrhizal ones. This phylogenetic reconstruction depicts widely separated groups of actinorhizal taxa, thus suggestive of multiple origin of nodulation. According to their results of *rbc*L phylogeny, if each actinorhizal lineage represents a separate origin, eight origins are possible. Alternatively, if each of the four subgroups represents a separate origin, then four origins are likely. This hypothesis of multiple origin, four times during the course of angiosperm evolution, was further confirmed by anatomical and morphological data, which shows a distinct difference among clades I-IV, suggesting that they are a result of different symbiotic origins.

2.9.2 Molecular Phylogeny of the Actinorhizal Hamamelidae:

The chloroplast gene, *rbc*L based phylogenetic analysis, using parsimony and distance-matrix approaches, of the higher Hamamelidae, including the families Betulaceae, Myricaceae and Casuarinaceae, was done by Maggia and Bousquet (1994). They report monophyletic origin of all the families. According to their findings the Myricaceae appeared to derive first, before the Betulaceae and the Casuarinaceae. In the Casuarinaceae, the genus *Gymnostoma* derived before the genera *Casuarina* and *Allocasuarina*, which were found closely related. They have also suggested a strong relationship between the evolutionary history of the host plants and their promiscuity toward *Frankia*. In their analysis, taxa appearing more recently were susceptible to a narrower spectrum of *Frankia* strains. The results also suggest that the ancestor of this group of plants was highly promiscuous, and that evolution has proceeded toward narrower promiscuity and greater specialization, perhaps culminating with the most specialized taxa being completely resistant to infection.

2.9.3 Molecular Phylogeny of the Betulaceae:

Bousquet *et al.* (1992a) compared the morphological and molecular phylogenies of the ancient family of woody dicots, the Betulaceae (birch family). The phylogeny was estimated from parsimony analysis of morphological characters of the genera *Alnus*, *Betula*, *Carpinus*, *Corylus*, *Ostrya* and *Ostryopsis* and from parsimony and distance matrix analysis of the sequences of the chloroplast gene *rbc*L encoding the large subunit of the carbon fixing enzyme ribulose-1,5-biphosphate carboxylase. The topologies obtained using these two parameters showed complete congruence. They have divided the family Betulaceae into two clades, Betuleae, containing *Alnus* and *Betula* and Coryleae, containing the other members of the family. Within the Coryleae clade, *Corylus* was the sister taxon to *Carpinus* - *Ostrya* - *Ostryopsis* subclade. This relationship was supported by both morphological and DNA sequence data, with the difference that since *Ostryopsis* was not included in the rbcL study, so *Corylus* showed to be the sister taxon to *Carpinus* and *Ostrya*. Relative-rate tests indicated that the Coryleae clade had a faster rate of *rbc*L evolution than the Betuleae clade. Heterogeneity of rates of morphological evolution also paralleled those for *rbc*L. Chen *et al.* (1999) have suggested a monophyletic origin of the family Betulaceae, with Casuarinaceae as its sister group. They have assessed the phylogeny of the Betulaceae on the basis of the *rbc*L, ITS and morphological data. From their studies also the occurrence of two sister clades, corresponding to the subfamilies Betuloideae and Coryloideae, are evident in the Betulaceae. The analysis of the three data sets combined reinforce the Coryloideae clade, and the sister group relationship between *Carpinus* and *Ostrya*, with a bootstrap value of 100%. The *Ostryopsis* – *Carpinus* - *Ostrya* clade receives support with a bootstrap value of 82%.

2.9.4 Molecular Phylogeny of Alder:

Dr Philippe Normand and his group at the University of Claude-Bernard, France, have recently done phylogenetic analysis of the actinorhizal genus *Alnus* (personal communication). They have used the nuclear ribosomal ITS region for inferring phylogenetic relationships. Their data too showed that the genus *Alnus* formed a coherent group close to *Betula*. They have positioned *Alnus nepalensis* in the subgenus

Gymnothyrsus, away from *Alnus nitida* and *Alnus maritima*. They have found that *Alnus formosana* was close to *Alnus nitida*, with which it shared fewer morphological characteristics, unlike *Alnus japonica* with which it shared more morphological similarity, yet was distantly located. According to their observations *Alnus glutinosa* appears to be the designated type of the genus since it was located in a dense cluster and showed the lowest average genetic distance with the rest of the genus for both morphological characters and evolutionary rate. Normand and his group have confirmed Japan as the center of origin of this genus because it has the highest number of species and has representatives in all deep branching lineages.

MATERIALS
AND
METHODS

CHAPTER 3

MATERIALS AND METHODS

3.1 COLLECTION SITES:

Two sites at Shillong, Meghalaya, were selected for the present study.

1. Upper Shillong
2. NEHU Permanent Campus, Mawlai.

Upper Shillong located at 25° 32' 52" North (Latitude) and 91° 53' 10.5" East (Longitude) has an altitude of approximately 6000 ft. At Upper Shillong a natural forest stand, consisting mainly of *Alnus nepalensis* and *Pinus kesiya* trees was studied. Other trees that were naturally present were *Elaeagnus* and *Myrica*. A number of other herbs and shrubs also formed a part of the natural vegetation at this site (Fig. 3.1a,b).

At the NEHU Permanent Campus, Mawlai, located at 25° 62' 14.9" North (Latitude) and 91° 89' 7.8" East (Longitude) and having an altitude of approximately 4500 ft., *Alnus nepalensis* trees were found growing naturally along with *Pinus kesiya* trees. Other naturally occurring flora included a number of herbs and shrubs (Fig. 3.1c).

3.2 COLLECTION OF LEAVES:

Fifty trees were randomly demarcated at each of these two sites. Each of the tree selected for the study was numbered and labeled by nailing on the tree trunk the number written on an aluminum sheet. The girth of each tree was measured at one-meter height. Young healthy leaves were collected from each tree. Care was taken to avoid any infected leaves. These leaves were kept in small polythene packets and labeled with the appropriate number. Leaves from ten trees were collected in a day and brought to the lab. The DNA was isolated immediately, or the leaves were stored overnight in the fridge and the DNA isolated the next day.

3.3 COLLECTION OF NODULES:

Alder nodules are mostly found in clusters on lateral roots, at a depth of 5-10cm below the soil surface. It has been reported by earlier researchers that maximum nodule growth occurs a month or two after the monsoons (Varghese, 2000). Keeping these observations in mind, the nodules were collected in the months of September-October. The weeds around each of the labeled *Alnus* tree, were cleared and then the soil around the tree was dug so as to expose the roots. The roots were traced to the tree to ascertain their origin. Clusters of nodules were collected from these trees and put into appropriately labeled polythene packets and brought to the laboratory. Care was taken to select only fresh light-brown colored nodules (Fig. 3.1d).

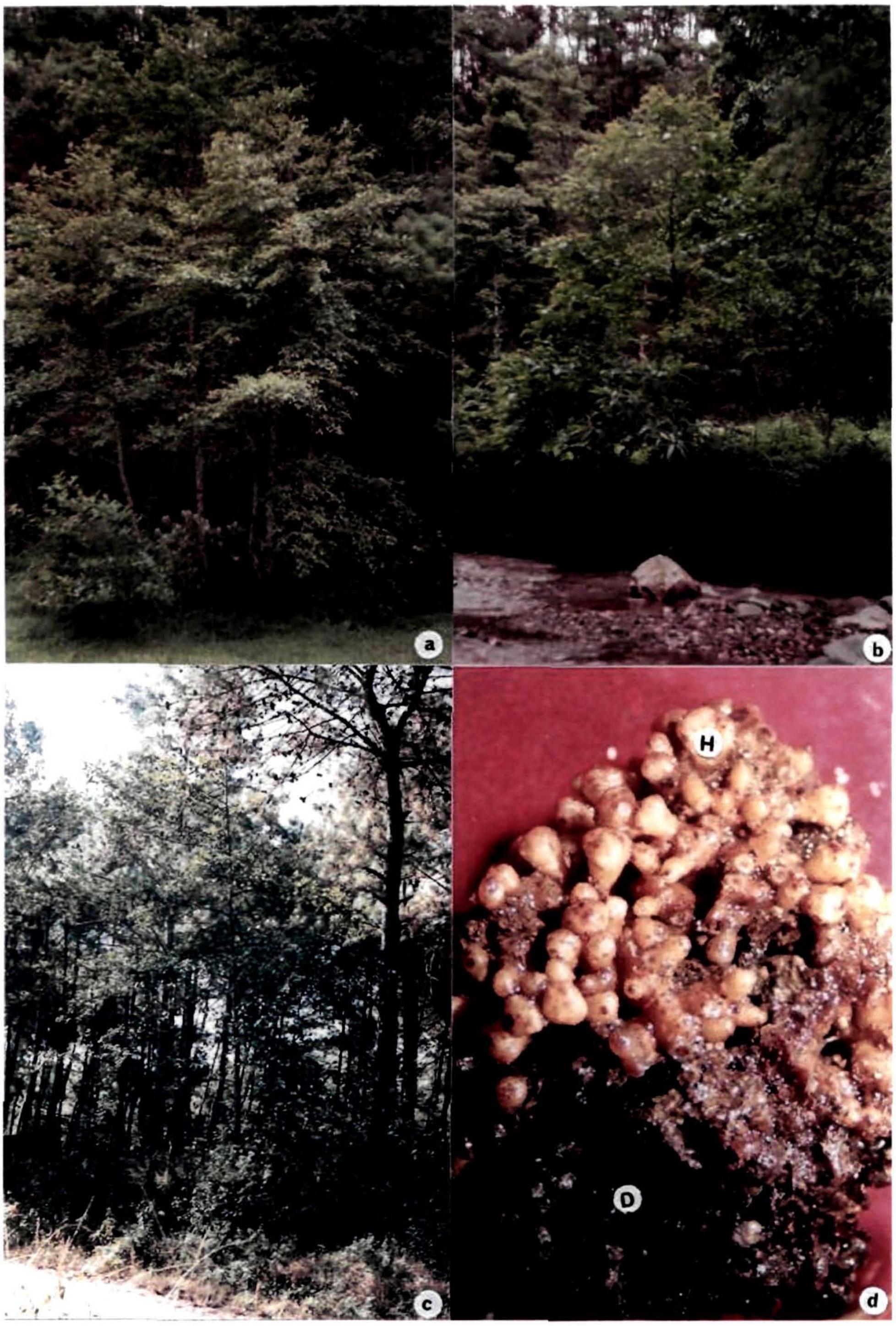
3.4 STERILIZATION OF THE NODULES:

As soon as the nodules were brought to the lab they were thoroughly cleaned.

1. The nodule clusters were washed under running tap water to remove the mud.
2. They were then transferred to a beaker containing detergent and water. Using a brush, the nodules were thoroughly washed to remove all tightly adhering soil particles.
3. They were then rinsed three-four times using double distilled water.
4. The nodules were then surface sterilized for 2-3 minutes, using 30% Hydrogen Peroxide (H_2O_2).
5. This was followed by repeated rinsing using sterile double distilled water.

The nodules, so sterilized, were then used to assess the nitrogenase activity, using the acetylene reduction assay.

Fig.3.1a.	The forest stand of *Alnus nepalensis* at Upper Shillong showing the trees growing away from water.
Fig.3.1b	The forest stand of *Alnus nepalensis* at Upper Shillong showing the trees growing near the stream.
Fig.3.1c	*Alnus nepalensis* trees growing at the NEHU Permanent Campus, Mawlai, Shillong.
Fig.3.1d	Root nodules of *Alnus nepalensis*. 'H' marks the healthy nodules and 'D' marks the degenerating nodules of the previous season.


a
b
c
H
D
d

3.5 ACETYLENE REDUCTION ASSAY:

Nitrogenase activity was measured using the Acetylene Reduction Assay (ARA) method as described by Stewart *et al.* (1968). The nitrogenase enzyme complex apart from reducing nitrogen, can also reduce acetylene (C_2H_2) gas to ethylene (C_2H_4). This property has been utilized in the Acetylene Reduction Assay where the quantity of ethylene produced or acetylene reduced is a reflection of the efficiency of the nitrogenase enzyme. The quantity of acetylene reduced, or the quantity of ethylene produced can be measured using a Gas Chromatograph equipped with H_2 flame-ionization detectors (FID). The principle behind the operation of the FID gas chromatograph is that as the gases pass through the column in the carrier gas stream, they separate into their different components, depending upon their boiling points.

The nitrogenase activity was estimated on the same day the nodules were collected.

1. The nodules after surface sterilization were dried on a filter paper. Fifty nodules from each tree were taken in a sterile stoppered vial and properly sealed with parafilm so as to make it air tight.
2. Three replicates were taken for each tree i.e., hundred fifty nodules were studied for each tree. Care was taken to avoid degenerating nodules.
3. 1 ml of air in the vial was replaced with 1ml of acetylene. In order to check the purity of the acetylene gas a control containing only acetylene gas without any nodules was also kept.
4. The vials were incubated for three hours at 28 ± 2°C .

The quantity of ethylene produced was then measured using the Gas Chromatograph (HP4890D) fitted with FID.

1. Using an air-tight syringe 1ml of the gas mixture from the vial was injected at the injection port of the gas chromatograph.

2. The injection port, oven and detector temperatures were 120°C , 90°C and 175°C respectively.
3. The carrier gas was nitrogen and the flow rates of H_2, air and N_2 were 50, 120 and 10 ml/min respectively.
4. In order to confirm the retention time 1ml of standard ethylene solution was injected into the gas chromatograph and the ethylene peak in the standard and the samples compared. Under these conditions of assay the retention time for ethylene was approximately 1.45 minutes and that of acetylene was 2.5 minutes.
5. After recording the area of ethylene gas of each of the replicates of each sample, the nodules from each vial were weighed using a Sartorius balance and the respective fresh weights recorded.

Carrying out the ARA studies for all hundred trees on a single day was not feasible. Therefore, in order to minimize the effect of changing environmental factors all ARA experiments were conducted continuously over a period of two months. The collection of each set of nodules and their ARA were performed on the same day and the conditions from the time of collection of nodules to their nitrogenase activity assay were kept constant for all the samples, thus making the comparative inference valid.

The nitrogenase activity was estimated in terms of nmoles of ethylene produced per gram fresh weight per hour using the following formula:

$$\text{N.A.} = \frac{\text{nmole } C_2H_4\text{/unit area in the standard X area of } C_2H_4 \text{ in the sample X vol. of vial}}{\text{Fresh weight X Incubation time}}$$

Where, N.A. is nmoles of ethylene produced/g fresh weight/hour.

3.6 ISOLATION OF DNA FROM LEAVES:

The Bangalore Genei Plant DNA extraction kit, the use of liquid nitrogen, the use of GR medium for isolation of chloroplast DNA, etc., were tried for the isolation of DNA from alder leaves. However, best results were obtained when Rouvier's protocol for isolation of DNA from nodules was used with a few modifications (Rouvier *et al.*, 1996). Care was taken to avoid any exogenous contamination of DNA during the course of extraction of DNA. All solutions, and plastic ware were autoclaved and sterile gloves were used for handling DNA. The protocol used is briefed as below:

1. Young, tender, uninfected leaves were cleaned with a brush to remove all adhering particles from the surface.
2. The leaves were then washed using sterile distilled water.
3. They were then surface sterilized using 30% H_2O_2 for 2-3 minutes, followed by repeated rinsing using sterile distilled water.
4. Two leaves were crushed in a mortar in 1ml of warm extraction buffer (1M Tris HCl, 5M NaCl, 0.5M EDTA, 2% [wt./vol.] CTAB and 3% [wt./vol.] polyvinyl pyrrollidone).
5. The macerate was filtered through cotton into a 1.5ml sterile microcentrifuge tube.
6. This was then centrifuged for five minutes at 8,000 xg at room temperature.
7. The supernatant was discarded and the pellet resuspended in 300μl of the extraction buffer.
8. 10μl of 20% Sodium dodecyl sulphate (SDS) was added and gently mixed.
9. The homogenate was incubated in a water bath at 65°C for one hour.
10. It was then centrifuged at 12,000 xg for fifteen minutes.
11. The supernatant was transferred to a fresh tube and extracted with equal volume of chloroform: isoamyl alcohol solution (24:1/ Vol : Vol).

12. After centrifugation at 12,000 xg for fifteen minutes, the aqueous phase was transferred to a fresh tube
13. 30μl of 3M Sodium acetate (pH 5.2) and 1.2 ml of ice cold ethanol were added to precipitate the DNA.
14. The DNA was pelleted by centrifugation at 13,000 xg for thirty minutes at 4°C.
15. The pellet was washed, 2-3 times, with ice cold 70% alcohol and vacuum dried.
16. DNA was dissolved in 10μl of ultra pure water and stored at –20°C for further use.
17. The purity of the DNA was checked by running on a 0.8% (wt./vol.) agarose gel.
18. To confirm that the DNA isolated did not have any protein contamination, it was digested with the restriction endonuclease *Eco*R 1 and then run on a 0.8% (wt./vol.) agarose gel.

3.7 QUANTIFICATION OF THE DNA:

DNA was directly quantified by visual observation after electrophoresis. When stained with ethidium bromide and observed on a transilluminator the intensity of fluorescence of the DNA, is directly proportional to the total mass of DNA (Sambrook *et al.*, 1989). During electrophoresis of the DNA known quantity of a molecular weight marker was added to one of the lanes in the gel. Using the Multi Analyst software ®(Version 1.1, Build 34) the mass of DNA present in each of the bands of the marker was calculated and then the approximate quantity of DNA present in the sample was estimated by visual comparison.

3.8 AMPLIFICATION OF DNA:

Different regions of the total genomic DNA isolated from the leaves of alder were subjected to amplification using the Polymerase Chain Reaction. This is an *in vitro* method of generation of multiple copies of the DNA fragment of interest thus facilitating a detailed and better analysis of the same. PCR involves selective amplification of the region of interest, flanked by a primer on either side. The millions of copies generated are a result of repetitive cycles involving denaturation of the double stranded DNA template, primer annealing and elongation by thermostable *Taq* Polymerase. Each amplification was run for thirty five cycles and each cycle comprised of 30 seconds of denaturation at 94°C, 30 seconds of annealing at the appropriate temperature and 30 seconds of elongation at 72°C. This was followed by an extension time of 3.00 minutes at 72°C at the end of the run.

PCR amplification was performed for the following regions:

Chloroplast DNA:

1. Entire chloroplast 16S rRNA gene (Fig.3.2).
2. Initial part of the chloroplast 16S rRNA gene.
3. Middle part of the chloroplast 16S rRNA gene.
4. *trn*T and *trn*L region of the chloroplast DNA (Fig.3.3).
5. *rbc*L gene

Nuclear DNA:

1. 18S-28S *rrn* operon (Fig.3.4).

The primers used for amplification of the above regions are listed in table 3.1.

PCR amplifications were performed using a Perkin Elmer Thermal cycler (GeneAmp PCR2400, Perkin Elmer, USA). The mix for amplification was prepared in strictly aseptic conditions under a U.V. hood. Care was taken to autoclave the plastic ware used and gloves were used while making the mix. Prior to preparing the mix the hood was cleaned using 70% alcohol and then given a 20 minutes exposure to U.V. light. In thin walled PCR tubes a 25μl mix was prepared for amplification. The reaction mix contained 2.5μl of each of the primer pair from a 0.5mM stock, 2.5μl of 10X PCR buffer, 2.5μl of each dNTP (1.25mM), 0.3μl of *Taq* Polymerase (5 units/μl.) and the rest of the volume was made up with ultra pure water. Each amplification run contained a negative control. 1μl of DNA was

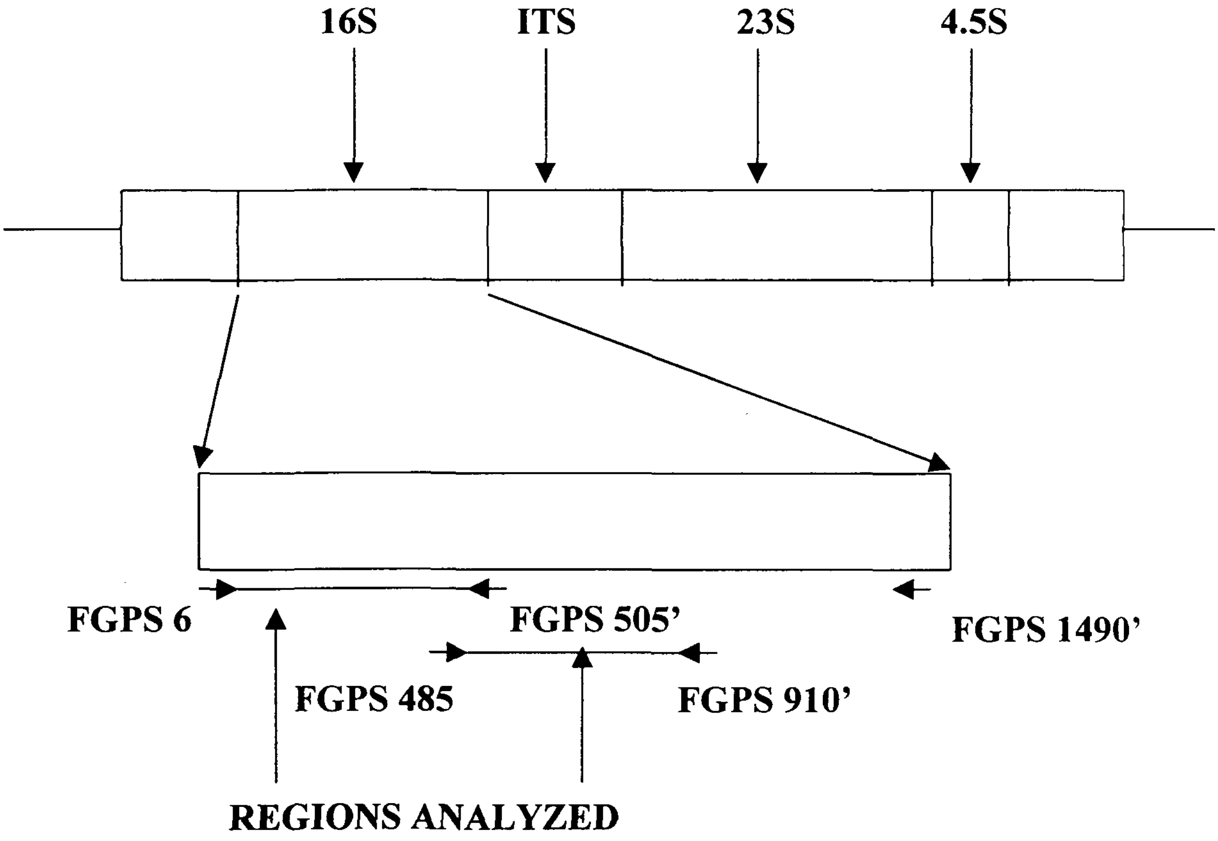


Fig.3.2 MAP OF THE *rrn* REGION OF THE CHLOROPLAST DNA HIGLIGHTING THE 16S rRNA GENE

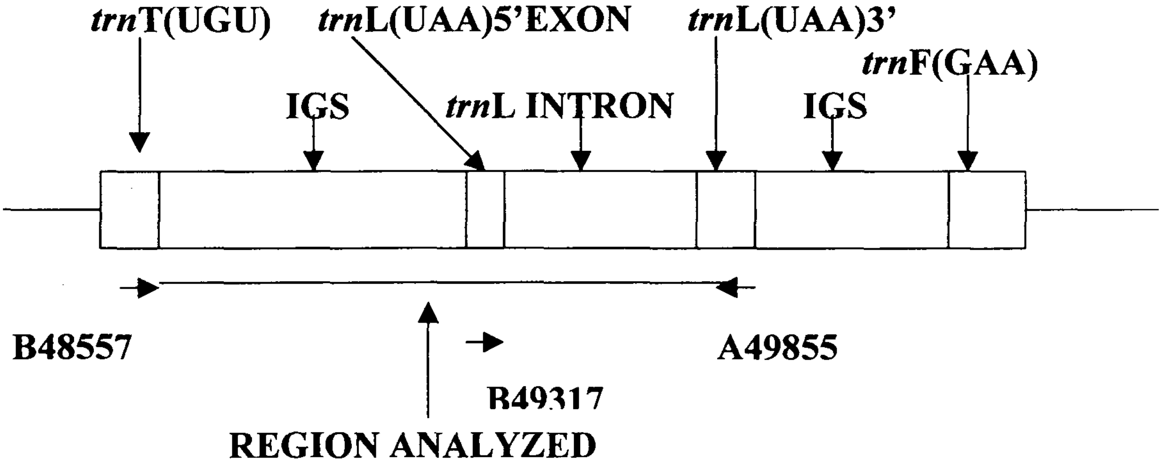


Fig.3.3 MAP SHOWING THE CONSERVED REGIONS OF THE CHLOROPLAST *trn* GENES (*trn*T, *trn*L EXON AND *trn*F) AND THE VARIABLE REGIONS (THE INTERGENIC SPACERS AND THE *trn*L INTRON).

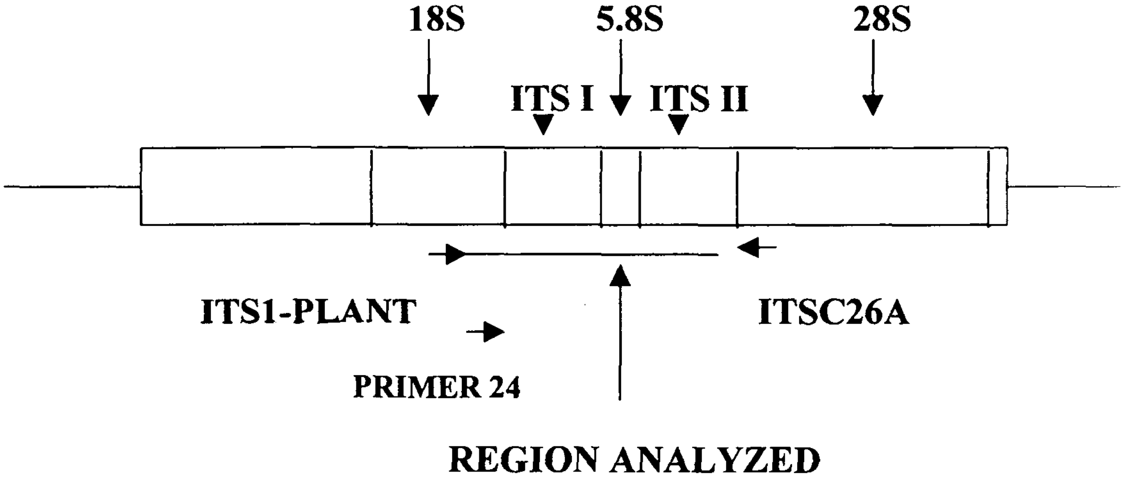


Fig. 3.4 MAP SHOWING THE 18S-28S *rrn* OPERON

TABLE 3.1 LIST OF PRIMERS USED FOR PCR AMPLIFICATION

16S rRNA gene:

LAB NO.	CODE	SEQUENCE	REFERENCE
7	FGPS 1490'	5'AAGGAGGGGATCCAGCCGCA3'	Normand *et al.*, 1996
8	FGPS 6	5'TGGAAAGCTTGATCCCTGGCT3'	Normand *et al.*, 1996
16	FGPS 505'	5'GTATTACCGCGGCTGCTG3'	Normand *et al.*, 1996
14	FGPS 485	5'CAGCAGCCGCGGTAA3'	Normand *et al.*, 1996
15	FGPS 910'	5' AGCCTTGCGGCCGTACTCCC3'	Normand *et al.*, 1996

The *trnT* - *trnL* region of the Chloroplast DNA:

LAB NO.	CODE	SEQUENCE	REFERENCE
19	B48557	5'CATTACAAATGCGATGCTCT3'	Taberlet *et al.*, 1991
20	A49317	5'CGAAATCGGTAGACGCTACG3'	Taberlet *et al.*, 1991
21	A49855	5'GGGGATAGAGGGACTTGAAC3'	Taberlet *et al.*, 1991

18S-28S Internally Transcribed Spacer of the Nuclear DNA:

LAB NO.	CODE	SEQUENCE	REFERENCE
22	ITS1-PLANT	5'CGCGAGAAGTCCACTG3'	Normand (personal communication)
23	ITSC26A	5'GTTTCTTTTCCTCCGCT3'	Wen and Zimmer, 1996
24	-	5'CCGCGAACCTGTCACAACAA3'	Present study

PRIMERS USED FOR AMPLIFICATION OF THE *rbc*L GENE

LAB NO.	CODE	SEQUENCE	REFERENCE
12	-34'	5'AAGTTCTTAATTCATGAGTCGTAG3'	Frascaria *et al.*,1993
13	1517'	5'TTAGTAAAAGATTGGGCCGAGTTT3'	Frascaria *et al.*,1993

MICROSATTELITE PRIMERS USED

LAB NO.	CODE	SEQUENCE	REFERENCE
17	-	5'AGCAGCAGCAGCAGC3'	Present study
18	-	5'GCGCGCGCGCGCGCGCGC3'	-do-

added to each of the tubes except for the negative control in which instead of DNA 1μl of ultra pure water was added.

The annealing temperature for the different primer pairs varied. The annealing temperature was calculated using the following formula:

Annealing Temperature = [4 (G+C) + 2 (A+T)] -5

In cases where the annealing temperature of the two primers differed, amplification was done at the lower annealing temperature. However, this was not a rule in all cases and the annealing temperature was modified, i.e., decreased or increased, depending upon the results obtained. In cases where a lot of non-specific amplification occurred, the stringency conditions were increased, by increasing the annealing temperature. Lower annealing temperatures were tried when it was difficult to get amplification.

3.8.1 Primer stock solutions:

The primers used for amplification were synthesized at M/S Bangalore Genei, India. In lyophilized condition, desalted and often PAGE purified primers were procured. These were then dissolved separately in ultra pure water to make 1M stock solutions respectively. A part of it was further diluted to make a 25mM stocks. Finally a 0.5mM working solution was made. The stock solutions were stored in a deep freezer at -20°C for future use.

3.8.2 Deoxyribonucleotide solutions:

Deoxyribonucleotides were either acquired from Boehringer Mannheim, Germany or Bangalore Genei, India. Powdered form of each of the four dNTPs (dGTP, dCTP, dATP and dTTP) were obtained from Boehringer Mannheim. A 5mM stock solution was made using ultra pure water and the pH adjusted to 8 using sterile NaOH solution. From the stock solutions, a common working solution was prepared that contained all the four dNTPs at the final concentration of 1.25mM each. dNTPs were also acquired from Bangalore Genei, but these were in solution. Individual 10mM stock solution of all the four dNTPs were obtained and working solution of adequate concentration were made by mixing equal quantities of each. All dNTPs were stored at -20°C.

3.8.3 *Taq* DNA Polymerase and Assay Buffer:

Taq DNA Polymerase, which is a 94kD thermostable enzyme, isolated from the bacterium *Thermus aquaticus*, was used for amplification of DNA. This enzyme lacks 3' to 5' exonuclease activity (proof reading) but has inherent 5' to 3' exonuclease activity. This enzyme and the buffer [10mM Tris-HCl (pH 9), 1.5mM $MgCl_2$, 50mM KCl and 0.01% gelatin] were procured from Bangalore Genei. The enzyme was distributed into aliquots of 10µl so that the entire stock was not thawed each time a mix was prepared. The stock was stored at -20°C.

3.9 AGAROSE GEL ELECTROPHORESIS:

The genomic DNA, PCR products and restriction digestion products were subjected to agarose gel electrophoresis. Depending on the size of DNA to be visualized, the strength of the gel was decided. The molecular weight of the DNA is inversely proportional to the gel percentage. When a big fragment was to be visualized a low percent gel (wt/vol.) was used and vice versa in case of a smaller band. The agarose gel was made in 1X TBE (Appendix: 1.1) and the buffer in the electrophoresis tank was also of 1X concentration. For running genomic DNA, a 0.8% gel was used. 2µl of the genomic DNA was mixed with loading buffer (Appendix 1.3) and loaded onto the gel. For testing PCR products normally 5µl of the product was loaded into the well of the gel. The concentration of the gel depended on the expected band size. In case of restriction digestion the entire 20µl mix was loaded onto the gel. Normally a high percentage gel (3-4%) was used to separate the smaller bands. During electrophoresis a molecular weight marker (λ DNA/ *Eco*R1 + *Hind* III double digest or pBR322 DNA/ *Hinf*1 digest) was always added to one of the wells. Electrophoresis was carried out at 5 volts/cm of the gel. After electrophoresis the gel was stained with ethidium bromide (0.5µg/ml) and visualized and photographed on the Gel Doc 1000 (BioRad).

3.10 AMPLIFIED FRAGMENT LENGTH PROFILES (AFLP):

The PCR products obtained for different samples showed differences under identical conditions of amplification. In order to study these variations, the amplicons of different samples were run at low voltage (5volts/cm), for four hours on high percentage agarose gels (1.5% for 18S-28S ITS, 2% for 16S rRNA gene and *trn*T and *trn*L region) containing ethidium bromide. After electrophoresis the gels were visualized in the Gel Doc 1000 and captured and saved in a floppy. The different profiles were analyzed using the Multi Analyst software®.

3.11 PURIFICATION OF AMPLIFIED DNA:

The amplified DNA was purified using a modified protocol of Byrnes *et al.* (1995).

1. 20µl of the amplified product was electrophoresed on a 0.8% agarose gel prepared in 1X TBE buffer containing ethidium bromide.
2. The separated DNA fragments were visualized using a U.V transilluminator and the band of interest was excised out using a clean scalpel blade.
3. The excess agarose was removed and the agarose slice containing the band of interest was also cut into smaller slices which were then transferred to a 0.5ml micro-centrifuge tube.
4. 30-40µl of cold saturated phenol (pH. 8) was added to the tube and left overnight at 4°C.
5. The tube was then kept at 60°C for 20-30 minutes till the agarose was completely dissolved.

6. It was then centrifuged at 12,000 xg for ten minutes at room temperature.
7. The aqueous phase was transferred to a fresh microcentrifuge tube and equal volume of chloroform : isoamyl alcohol (24:1) was added.
8. The tube was then centrifuged at 12,000 xg for five minutes.
9. The aqueous phase was then transferred to a fresh tube and the DNA precipitated with two volumes of ice cold ethanol. The tube was kept overnight at 4°C for precipitation.
10. It was then centrifuged at 12,000 xg for five minutes at 4°C.
11. The pellet was then washed with 70% alcohol and vacuum dried.
12. After drying the pellet was dissolved in 4-5µl of ultra pure water.
13. To confirm that the elution was proper 1µl of the elluent was again electrophoresed.

3.12 NUCLEOTIDE SEQUENCING:

The 18S-28S Internally Transcribed Spacer, on amplification using the primer pair ITS1-PLANT/ITSC26A, showed multiple bands including the expected 750 base pair band. To confirm the origin of this band three samples were selected. The band of interest was excised out and purified using the protocol described earlier. These samples were then packed in gel ice and sent to M/S Bangalore Genei, India for sequencing. The sequencing was done using an automated DNA sequencer that uses fluorescent label dye terminators or fluorescent label primers. The ABI's Ampli *Taq* FS dye terminator cycle sequencing chemistry, based on Sanger's dideoxy chain termination method (Sanger *et al.*, 1977) was used.

3.13 COMPUTER ANALYSIS OF THE NUCLEOTIDE SEQUENCES:

Using the site http://ncbi.nlm.nih.gov, individual sequences were fed into the GenBank for Blast Analysis. The sequences that were retrieved from the GenBank and our sequences were then aligned using the multiple sequence alignment program CLUSTAL W (1.75) (Thompson *et al.*, 1994). The inference package PHYLIP for Windows (version 3.5, Felsenstein, 1993), was used to construct phylogenetic trees using the Parsimony method. Neighbor Joining trees were constructed using the program PHYLODENDRON (version 0.8d, beta by D.G. Gilbert, January 1999).

3.14 DESIGNING GENUS SPECIFIC PROBE:

The aligned sequences of the 18S-28S ITS region, including the distal part of the 18S rRNA gene, the 5.8S rRNA and the initial part of the 28S rRNA, of alder (*Alnus nepalensis, Alnus incana, Alnus crispa, Alnus glutinosa, Alnus matsumurae*), *Betula*, *Carpinus* and *Corylus* were compared. A region showing similarities between the different species of alder and differences between the different genera was selected to design a genus specific probe for *Alnus*. This primer was located at 94-110bp downstream of the sequence of *Alnus nepalensis* when sequenced using the primer ITS1-PLANT This region was conserved between the different species of *Alnus* but showed inter generic differences. This primer (primer 24) was then used for amplification of the 18S-28S ITS region of *Alnus nepalensis,* in conjunction with primer ITSC26A. To confirm that it was a genus specific probe, DNA was isolated from *Betula*, *Carpinus*, *Myrica* and *Elaeagnus* and tested for amplification.

3.15 RESTRICTION DIGESTION OF PCR PRODUCTS:

The amplicons produced for the different regions were digested using different restriction endonucleases. The enzymes to be used were decided using the Mac Vector software®, wherein a known sequence was subjected to digestion with different restriction endonucleases and the enzyme showing maximum variability was selected. Different enzymes were used for different regions. A list of the enzymes used, their sites and reaction conditions are given in the table 3.2. The restriction digestion was done in

Table:3.2 LIST OF RESTRICTION ENZYMES USED

Proximal part of the 16S rRNA gene:

ENZYME	SOURCE	TARGET SEQUENCE	BUFFER	INCUBATION TEMPERATURE
*Msp*1	*Moraxella* species	C↓CGG* GGC↑C	Buffer C*	37°C

Middle Part of the 16S rRNA gene:

ENZYME	SOURCE	TARGET SEQUENCE	BUFFER	INCUBATION TEMPERATURE
*Msp*1	*Moraxella* species	C↓CGG GGC↑C	Buffer C*	37°C
*Mbo*1	*Moraxella bovis*	↓GATC CTAG↑	Buffer B*	37°C
*Alu*1	*Arthrobacter luteus*	AG↓CT TC↑GA	Buffer L*	37°C

18S-28S ^{rrn} operon:

ENZYME	SOURCE	TARGET SEQUENCE	BUFFER	INCUBATION TEMPERATURE
*Scr*F1	*Streptococcus cremoris*	CC↓NGG GGN↑CC	Buffer B**	37°C

---

*- Bangalore Genei; **- Boehringer Mannheim

microcentrifuge tubes in which a 20μl mix was prepared. 5-8μl of the amplicon was digested with 2 units of enzyme. 2μl (1X) of the appropriate buffer (10X) was added and the remaining volume was made up with ultra pure water. The tube was sealed with parafilm and incubated for three hours at the optimal temperature suggested by the manufacturers. To stop the digestion, 2μl of the gel loading buffer (Bromophenol Blue) was added. The samples were then electrophoresed on a 3% agarose gel at 40 volts for eight hours. The gel was stained with ethidium bromide and visualized using the Gel Doc 1000. The profiles were analyzed using the Multi-Analyst software®.

RESULTS
AND
DISCUSSION

CHAPTER 4

RESULTS AND DISCUSSION

4.1 COLLECTION OF MATERIAL

4.1.1 Collection sites:

Alnus nepalensis trees at both Upper Shillong and NEHU Permanent Campus Mawlai, were found growing along with *Pinus kesiya* trees. At Upper Shillong some of the alder trees were found growing adjacent to a stream. Other trees were growing away from the water. The soil at both the sub-sites was moist. This could be attributed to the high rainfall in this region. However, for obvious reasons, the soil adjacent to the stream had relatively higher moisture. The pH of the soil at both the sites was measured. At Upper Shillong the soil pH ranged from 4.47 ± 0.02 to 4.82 ± 0.01 in areas away from water. Near the water a pH of approximately 5.57 ± 0.04 was recorded. At the NEHU Permanent Campus, the soil pH was about 5.21 ± 0.04.

4.1.2 Collection of leaves:

At each site 50 *Alnus nepalensis* trees were randomly selected for study. At Upper Shillong most of the trees had a girth of approximately 17-25 inches at 1 meter height from ground level. At the NEHU Permanent Campus, the girth of the trees showed a wider range. Majority of the trees had a girth of 7-15 inches while others were of approximately 17-25 inches girth. It was not possible to get all hundred trees having exactly the same girth. The leaves showed luxuriant growth for major part of the year except for the winter season (December-February), when there was maximum leaf fall. Young light green coloured leaves were selected for molecular analysis. Bigger leaves at times showed fungal infection and so were avoided. Moreover, the concentration of phenolics in the mature leaves was higher than the young ones and therefore getting a clean preparation of DNA from mature leaves was difficult.

4.1.3 Collection of nodules:

Nodules were collected from each tree demarcated for the study. During the dry season it was difficult to find nodules. Therefore, as suggested by earlier workers, nodules were collected in September-October after the monsoon. Each tree analyzed was found nodulated and healthy. The nodules were found in clusters on lateral roots, at a depth of 5-10cm below the soil surface. Near the water at times the roots were found exposed and healthy nodule clusters were found on these roots. The fresh nodules were light brown in colour while the degenerating ones, of the previous season, were dark brown in colour. While selecting nodules for assessing the nitrogenase activity only fresh light-brown coloured nodules were taken.

4.2 ACETYLENE REDUCTION ASSAY:

Acetylene reduction assay (Stewart *et al.*, 1968), was carried out for fresh light brown coloured nodules. The degenerating nodules were eliminated because they contained a lot of dead, woody tissue. To get best results the assay for nitrogenase activity was done on the day the nodules were collected. Fifty nodules were taken at a time for assessing the nitrogenase activity. Three replicates were taken for each tree, i.e., hundred fifty nodules from each tree were assessed and the ARA values averaged out. Differences were seen in the ARA values of different nodules of a tree. This could be due to the occurrence of different *Frankia* strains in different nodules. In fact, earlier workers have reported the occurrence of more than one strain of *Frankia* in a single nodule. The average nitrogenase activity for different trees have been listed in the tables 4.1 and 4.2 and have been graphically represented in fig. 4.1 and 4.2. Different trees showed a range of ARA values. A few trees at both the sites showed very high nitrogenase activity. It was observed that

Table 4.1-4.2

Table 4.1 Average nitrogenase activities of the trees at the NEHU Permanent Campus, Mawlai, Shillong. Values of hundred and fifty nodules were averaged out.

Table 4.2 Average nitrogenase activities of the trees at Upper Shillong. Values of hundred and fifty nodules were averaged out.

TABLE 4.1 NEHU PERMANENT CAMPUS, MAWLAI, SHILLONG

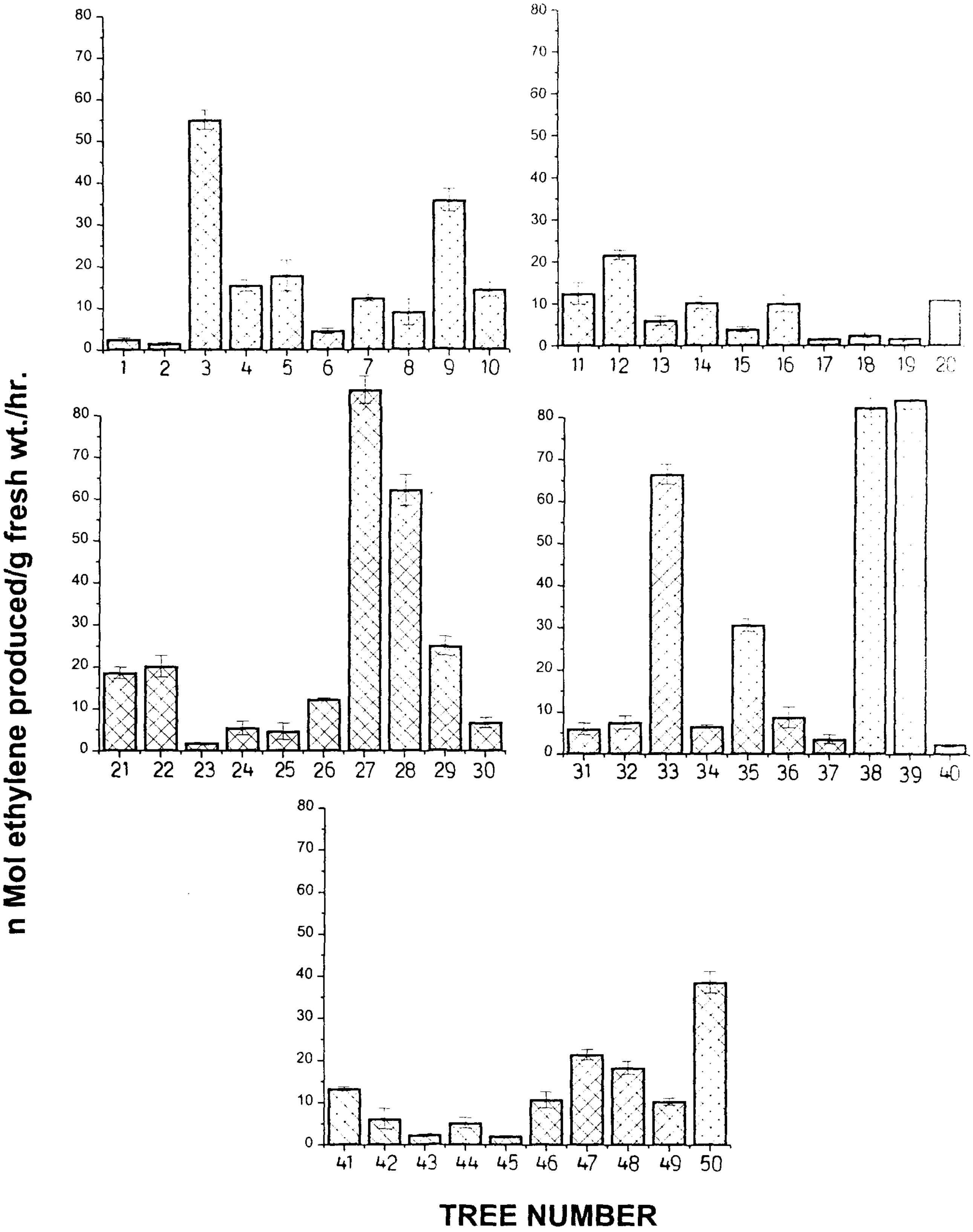
Sl.No.	SAMPLE	AVERAGE NITROGENASE ACTIVITY (nMC_2H_4 produced/g fresh wt./hr.)
1.	P.C:1	2.53±0.40
2.	P.C:2	1.58±0.29
3.	P.C:3	55.02±2.34
4.	P.C:4	15.45±1.35
5.	P.C:5	17.79±3.68
6.	P.C:6	4.58±0.63
7.	P.C:7	12.35±0.73
8.	P.C:8	8.99±3.09
9.	P.C:9	35.63±2.67
10.	P.C:10	14.32±1.71
11.	P.C:11	12.48±2.58
12.	P.C:12	21.59±1.11
13.	P.C:13	5.98±1.12
14.	P.C:14	10.27±1.45
15.	P.C:15	3.96±0.61
16.	P.C:16	10.03±1.95
17.	P.C:17	1.61±0.17
18.	P.C:18	2.43±0.51
19.	P.C:19	1.72±0.81
20.	P.C:20	10.87±2.02
21.	P.C:21	18.59±1.36
22.	P.C:22	20.16±2.56
23.	P.C:23	1.83±0.13
24.	P.C:24	5.46±1.66
25.	P.C:25	4.65±1.96
26.	P.C:26	12.31±0.31
27.	P.C:27	86.00±3.24
28.	P.C:28	62.09±3.72
29.	P.C:29	25.00±2.26
30.	P.C:30	6.74±1.21
31.	P.C:31	5.96±1.40
32.	P.C:32	7.52±1.59
33.	P.C:33	66.51±2.41
34.	P.C:34	6.60±0.40
35.	P.C:35	30.63±1.50
36.	P.C:36	8.75±2.48
37.	P.C:37	3.54±1.08
38.	P.C:38	82.28±2.22
39.	P.C:39	84.15±2.21
40.	P.C:40	2.41±0.56
41.	P.C:41	13.31±0.47
42.	P.C:42	6.25±2.40
43.	P.C:43	2.44±0.20
44.	P.C:44	5.28±1.26
45.	P.C:45	2.11±0.21
46.	P.C:46	10.71±1.91
47.	P.C:47	21.50±1.24
48.	P.C:48	18.33±1.52
49.	P.C:49	10.31±0.85
50.	P.C:50	38.63±2.57

TABLE 4.2 UPPER SHILLONG

Sl.No.	SAMPLE	AVERAGE NITROGENASE ACTIVITY (nMC_2H_4 produced/g fresh wt./hr.)
1.	U.S:1	3.845±0.72
2.	U.S:2	0.342±0.07
3.	U.S:3	11.78±1.79
4.	U.S:4	4.06±1.40
5.	U.S:5	2.56±0.61
6.	U.S:6	3.89±1.92
7.	U.S:7	0.652±0.22
8.	U.S:8	6.03±1.36
9.	U.S:9	1.98±0.51
10.	U.S:10	6.29±1.89
11.	U.S:11	8.35±2.28
12.	U.S:12	15.51±0.21
13.	U.S:13	2.14±0.91
14.	U.S:14	1.20±0.27
15.	U.S:15	38.84±1.06
16.	U.S:16	2.16±0.72
17.	U.S:17	2.97±0.09
18.	U.S:18	2.16±0.88
19.	U.S:19	2.04±0.72
20.	U.S:20	1.31±0.40
21.	U.S:21	6.36±2.59
22.	U.S:22	10.13±1.03
23.	U.S:23	6.95±1.93
24.	U.S:24	0.460±0.06
25.	U.S:25	9.22±0.02
26.	U.S:26	1.23±0.51
27.	U.S:27	7.76±0.97
28.	U.S:28	2.58±0.86
29.	U.S:29	12.08±1.25
30.	U.S:30	9.81±1.15
31.	U.S:31	5.69±1.40
32.	U.S:32	4.04±1.19
33.	U.S:33	84.67±2.19
34.	U.S:34	70.92±4.47
35.	U.S:35	4.27±2.04
36.	U.S:36	5.62±0.44
37.	U.S:37	13.01±2.50
38.	U.S:38	19.16±1.48
39.	U.S:39	20.04±1.50
40.	U.S:40	12.16±0.85
41.	U.S:41	9.41±0.58
42.	U.S:42	16.65±0.36
43.	U.S:43	2.74±0.65
44.	U.S:44	6.76±0.94
45.	U.S:45	5.90±1.43
46.	U.S:46	3.12±1.09
47.	U.S:47	8.45±0.82
48.	U.S:48	9.57±1.54
49.	U.S:49	39.70±2.55
50.	U.S:50	11.43±2.05

GRAPHICAL REPRESENTATION OF THE AVERAGE ARA VALUES OF THE FIFTY TREES AT THE NEHU PERMANENT CAMPUS, MAWLAI.

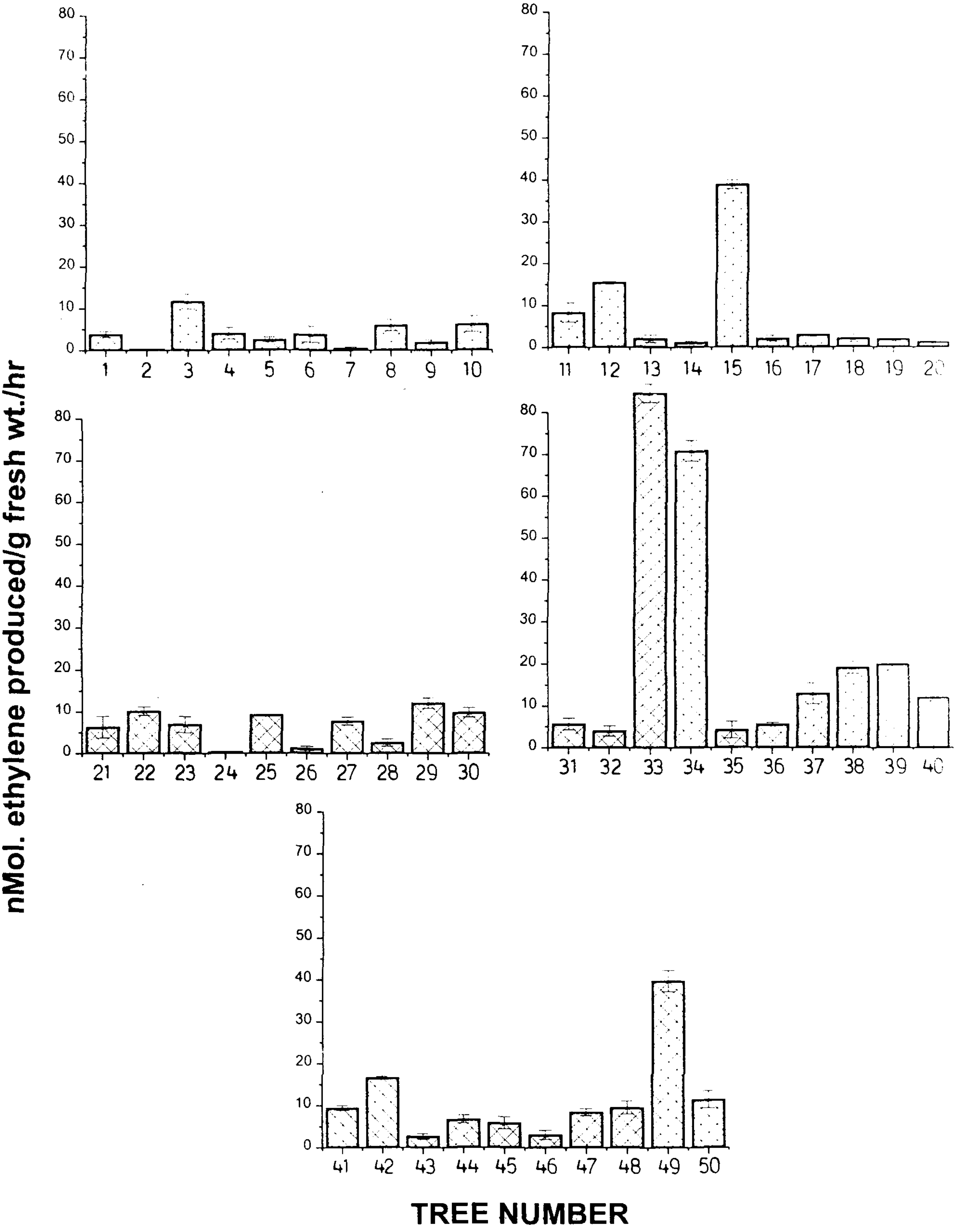
Nitrogenase activity was measured using the method of Stewart *et al.*, 1968. The ARA values of one hundred and fifty nodules was averaged out and graphically represented.



NEHU PERMANENT CAMPUS

Fig.4.2 GRAPHICAL REPRESENTATION OF THE AVERAGE ARA VALUES OF THE FIFTY TREES AT UPPER SHILLONG.

Nitrogenase activity was measured using the method of Stewart *et al.*, 1968. The ARA values of one hundred and fifty nodules was averaged out and graphically represented.



UPPER SHILLONG

more number of trees at the NEHU Permanent Campus belonged to the high nitrogen fixing group.

A comparison of the nitrogenase activity estimates of nodules collected from trees growing near the stream, with the ones collected from trees growing away from the stream at Upper Shillong, was also done (Table 4.3). The mean ARA values for the two sub-sites were subjected to t-test. The two means were significantly different and the range of the ARA values was 0.342 - 84.67 for the trees growing near the stream and 0.460 - 39.70 for the trees growing away from water. Availability of sufficient water would have a general invigorating affect on the metabolic activity of the tree. This is reflected in the superior nitrogenase activity shown by the trees growing near the stream.

4.3 ISOLATION OF DNA:

As a first step for molecular analysis a number of protocols were tried for the isolation of total genomic DNA from young leaves of *Alnus nepalensis*. Best results were obtained when the protocol used by Rouvier *et al.* (1996) for isolation of DNA from nodules was used with some modifications. Freshly collected leaves gave a better yield of DNA. Young leaves gave a cleaner preparation of DNA. It was difficult to get rid of the pigmentation if the leaves were mature and dry. Moreover, the mature and dry leaves accumulated too much phenolics which would interfere with the amplification if not eliminated completely. A 3% concentration of PVPP (Polyvinyl Pyrrolidone) was added to the extraction buffer to get rid of the phenolics.

Once the DNA was isolated it was electrophoresed on a 0.8% agarose gel to confirm the presence of DNA. On electrophoresis an approximately 20kb band was obtained. At times a distinct DNA band was not visible but on amplification with appropriate primers, it got amplified. This was because nanogram amount of DNA is sufficient for amplification. To check for protein contamination of the DNA, the DNA was digested with the restriction

TABLE: 4.3 COMPARISON OF THE ARA VALUES OF THE TREES GROWING NEAR AND AWAY FROM WATER AT UPPER SHILLONG.

TREES NEAR WATER		TREES AWAY FROM WATER	
TREE NO.	ARA VALUE	TREE NO.	ARA VALUE
U.S:1	3.84	U.S:16	2.16
U.S:2	0.342	U.S:17	2.97
U.S:3	11.78	U.S:18	2.16
U.S:4	4.06	U.S:19	2.04
U.S:5	2.56	U.S:20	1.31
U.S:6	3.89	U.S:21	6.36
U.S:7	0.652	U.S:22	10.13
U.S:8	6.03	U.S:23	6.95
U.S:9	1.98	U.S:24	0.460
U.S:10	6.29	U.S:25	9.22
U.S:11	8.35	U.S:26	1.23
U.S:12	15.51	U.S:27	7.76
U.S:13	2.14	U.S:28	2.58
U.S:14	1.20	U.S:29	12.08
U.S:15	38.84	U.S:30	9.81
U.S:31	5.69	U.S:41	9.41
U.S:32	4.04	U.S:42	16.65
U.S:33	84.67	U.S:43	2.74
U.S:34	70.92	U.S:44	6.76
U.S:35	4.27	U.S:45	5.90
U.S:35	5.62	U.S:46	3.12
U.S:37	13.01	U.S:47	8.45
U.S:38	19.16	U.S:48	9.57
U.S:39	20.04	U.S:49	39.70
U.S:40	12.16	U.S:50	11.43

Mean m_1	13.905	Mean m_2	7.638
S_1^2	443.885	S_2^2	61.913

$$t = \frac{m_1 - m_2}{S}\sqrt{\frac{n}{2}}$$

=46.736** at 48 degrees of freedom.

enzyme *Eco*R 1 and then run on a 0.8% agarose gel. When the gel was visualized under UV light in the transilluminator, a smear was observed which confirmed the absence of proteins. This DNA was then utilized for amplification studies.

4.4 AMPLIFICATION OF DNA:

4.4.1 Amplification of the 16S rRNA gene:

The chloroplast 16S rRNA gene has been widely used for the characterization of plants because of its large size, conserved nature and the presence of independent domains that can undergo sequence change independently. In the present study, we amplified the entire 16S rRNA gene using the primer FGPS 1490' in conjunction with primer FGPS 6. These two primers and other primers used in this study were designed by Normand *et al.* (1996) and were used for the amplification of the 16S rRNA gene of *Frankia*. However, these primers could also amplify the chloroplast 16S rRNA gene. The amplicon produced was run on a 0.8% agarose gel. When visualized using the Gel Doc 1000, after staining with ethidium bromide, the expected approximately 1.5kb band, was observed (Fig. 4.3). It is known that the chloroplast and bacterial 16S rRNA gene are more or less similar and have the same size. Therefore, it was not possible to differentiate on the basis of the band size. Unlike the bacterial 16S rRNA gene, the chloroplast 16S rRNA gene has a *Nru*1 site in the distal region. We therefore, digested our amplicon with this enzyme. Since, the amplicon got digested with the enzyme *Nru*1 (Fig. 4.4) it was confirmed that the amplicon was of plant origin. However, it was not possible to get amplification of the entire gene, for all hundred samples. It could have been due to fragmentation of the DNA. Hence, two different primer pairs were used to amplify the initial part and middle part of the 16S rRNA gene.

Fig 4.3-4.8

AMPLIFICATION OF THE CHLOROPLAST 16S rRNA GENE:

Fig.4.3: Amplification of the entire chloroplast 16S rRNA gene with the primers FGPS 6/FGPS 1490'.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1 & 5: Amplification positive.
Lanes 2, 3, 4, & 6: Amplification negative.
Lane 7: Negative control.

Fig.4.4: Restriction digestion of the amplicon of primer FGPS 6/FGPS 1490' with the restriction enzyme *Nru* 1:

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1: Undigested DNA
Lanes 2 & 3: Digested DNA.

Fig.4.5 & 4.6: Amplification of the proximal part of the chloroplast 16S rRNA gene with the primers FGPS 6/FGPS 505'.

Fig.4.5: Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest)
Lanes 1, 2, 3 & 5: Amplification positive
Lane 4: Amplification negative.
Lane 6: Negative control.

Fig.4.6: Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest)
Lanes 1-7: Amplification positive.

Fig.4.7 & 4.8: Amplification of the middle part of the chloroplast 16S rRNA gene with the primers FGPS 485/FGPS 910':

Fig.4.7: Amplification at 45°C.

LaneM: Molecular weight marker (Lambda DNA/*EcoR*I + *Hind*III Double Digest).
Lanes 1-7: Amplification positive.

Fig.4.8: Amplification at 50°C.

LaneM: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-7: Amplification positive.

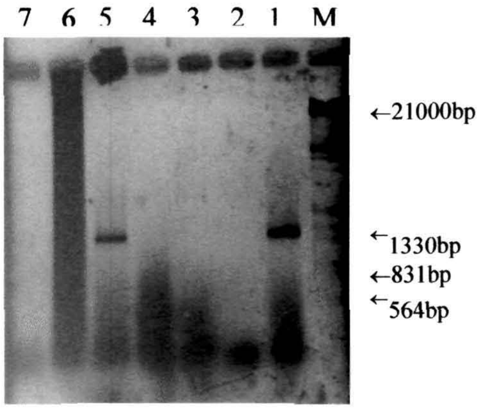


Fig.4.3

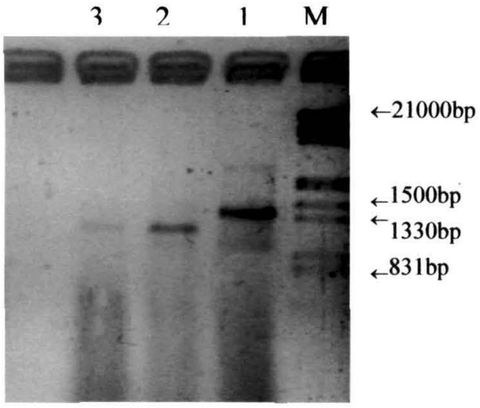


Fig.4.4

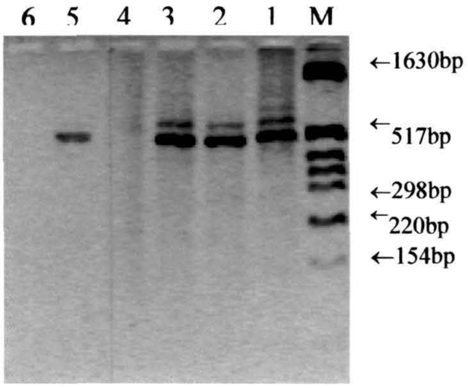


Fig.5

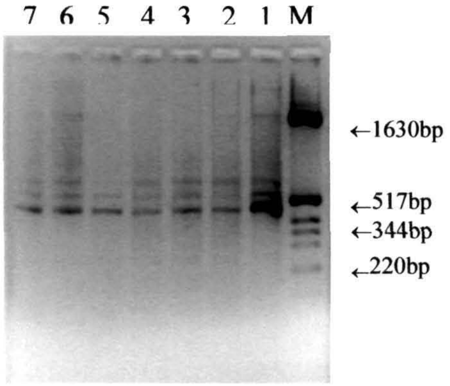


Fig.4.6

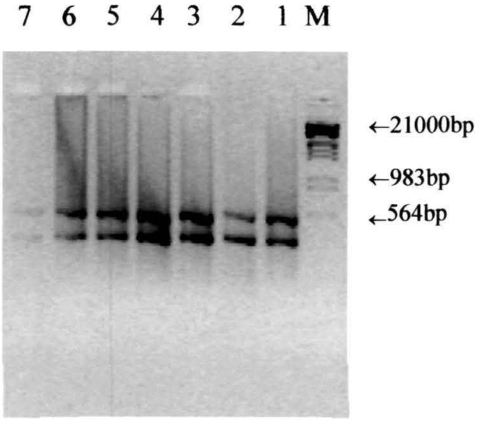


Fig.4.7

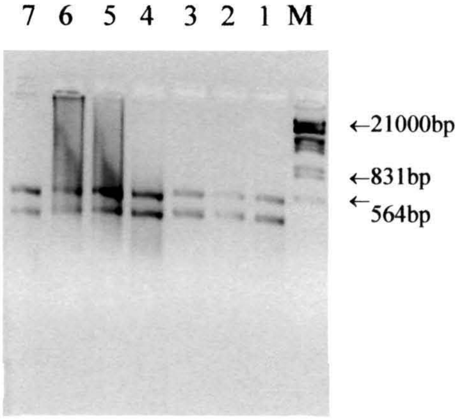


Fig.4.8

Using a specific primer why more than one bands are seen?

To be explained / explain okay.

4.4.2 Amplification of the proximal part of the chloroplast 16S rRNA gene:

The proximal part of the 16S rRNA gene was amplified using the primer pair FGPS 6 and FGPS 505'. Primer FGPS 6 is located at coordinate 6 and primer FGPS 505' at position 505' on the 16S rRNA gene of *Frankia*. It has been reported that these primers also amplify chloroplast 16S rRNA gene. *Alnus nepalensis* samples were amplified using these primers and then electrophoresed on a 2% agarose gel since the expected band was approximately 500bp. When the gel was observed using the Gel Doc 1000, differences were seen in the amplification profiles of different samples (Fig. 4.5 & 4.6).

4.4.3 Amplification of the middle part of the chloroplast 16S rRNA gene:

Primer pair FGPS 485 and FGPS 910' was utilized for amplifying the middle part of the chloroplast 16S rRNA gene. The primers FGPS 485 and FGPS 910', located at 485bp and 910bp respectively on the *Frankia* 16S rRNA gene, were designed and used by Normand *et al.* (1996). Like primer FGPS 6 and FGPS 505', these primers could also amplify chloroplast 16S rRNA gene. Our samples were amplified using these primers at an annealing temperature of 45°C. Two bands of approximately 400bp and 550bp were visualised after electrophoresis on a 2% agarose gel (Fig. 4.7 & 4.8). To eliminate one of the bands, stringency conditions were increased by raising the annealing temperature to 50°C and then to 55°C. Increasing stringency did not eliminate either of the bands and even at 55°C both the bands persisted. Comparative sequence analysis revealed that primer FGPS 485 completely matched with the alder sequence but primer FGPS 910' could anneal only if a few mismatches were allowed. At low stringency conditions both primers could bind to alternate annealing sites. These mismatch sites varied for different species (Table 4.4-4.7).

ALTERNATE ANNEALING SITES OF PRIMER 14 & 15 FOUND ON THE CHLOROPLAST 16S rRNA GENE OF *ALNUS GLUTINOSA*.

TABLE 4.4 PRIMER14 (FGPS 485)

Strict Position	Bases Matched	Strand	Subseq. Found
403	15/15	+	CAGCAGCCGCGGTAA
106	10/15	-	CAGCctaCGgGGTAT
744	10/15	+	TAGCtaaCGCGtTAA
950	10/15	+	CAGCtcgtGCcGTAA

TABLE 4.5 PRIMER 15 (FGPS 910')

Strict Position	Bases Matched	Strand	Subseq. Found
790	16/20	-	AttCTTGCGaaCGTACTCCC
989	13/20	-	AtCgTTGCGGgatTACcCCa
316	12/20	-	ctCCacGCGGCatTgCTCCg

ALTERNATE ANNEALING SITES OF PRIMER 14 & 15 FOUND ON THE CHLOROPLAST 16S rRNA GENE OF *ALNUS INCANA* .

TABLE 4.6 PRIMER14 (FGPS 485)

Strict Position	Bases Matched	Strand	Subseq. Found
642	15/15	+	CAGCAGCCGCGGTAA
344	10/15	-	CAGCctaCGgGGTAt
983	10/15	+	TAGCtaaCGCGtTAA
1190	10/15	+	CAGCtcgtGCcGTAA

TABLE 4.7 PRIMER15 (FGPS 910')

Strict Position	Bases Matched	Strand	Subseq. Found
1029	16/20	-	AttCTTGCGaaCGTACTCCC
1602	13/20	-	AGCCcTGCcttCGgcaTCCC
555	12/20	-	CtCCacGCGGCatTgCTCCg
1324	12/20	-	TGaCTTGacGtCaTcCTCaC
1483	12/20	+	AGCCaTaCGGCgGTgaattC

4.4.4 Amplification of the *trn*T-*trn*L region:

Non-coding regions are known to harbor greater genetic variability. In order to look for polymorphism the IGS and the intron lying between the *trn*T and *trn*L exon was studied. Primer B48557, A49317 and A49855 are located in the *trn* genes (Fig.3.3), which, are highly conserved in nature. Initial experiments were carried out using primer A49855 in conjunction with primer B48557 and A49317 respectively. Primer A49317 in conjunction with primer A49855 was used to amplify the *trn*L intron at an annealing temperature of 45°C The amplicon produced, on electrophoresis, showed a band of approximately 550bp and occasionally a bigger band of approximately 1.2kb was also present (Fig. 4.9). In some cases smaller bands were also present. To eliminate the additional bands the annealing temperature was raised to 50°C. At this temperature the bigger and smaller bands completely disappeared, but the 550bp band remained (Fig. 4.10). Primer B48557 along with primer A49855 amplified the IGS region between the *trn*T (UGU) and *trn*L (UAA) 5' exon as well as the *trn*L intron. Since the primer pair B48557/A49855 allowed analysis of a larger region, including the region amplified with the primers A49317/ A49855, we amplified all samples using this primer pair at an annealing temperature of 45°C. Amplicons produced were electrophoresed on a 0.8% agarose gel and observed and photographed using the Gel Doc 1000. For some samples a single band of approximately 1.2kb was observed while in others additional smaller bands were present (Fig. 4.11 and 4.12) and therefore, we went in for AFLP analysis (Section 4.5.3).

4.4.5 Amplification of the 18S-28S *rrn* operon:

Primer ITS1-PLANT in the distal region of the 18S rRNA gene and primer ITSC26A in the initial part of the 28S rRNA gene were used to amplify the ITS I and ITS II flanking the 5.8S rRNA gene. Prior to amplifying all samples, experiments were conducted to standardize the amplification conditions. Three samples were selected and two different dilutions (1X & 1/10X;) of DNA were used for amplification, at an annealing temperature of

Fig 4.9-4.13

AMPLIFICATION OF THE *trn*T-*trn*L REGION OF THE CHLOROPLAST DNA:

Fig.4.9: Amplification of the *trn*L intron using primers A49317/A49855 at 45°C.

LaneM: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).

Lanes 1-4: Amplification positive.

Lane 5: Negative control.

Fig.4.10: Amplification of the *trn*L intron using primers A49317/A49855 at 50°C.

LaneM: Molecular weight marker (Lambda DNA/*EcoR*I + *Hind*III Double Digest).

Lanes 1-4: Amplification positive.

Lane 5: Negative control.

Fig.4.11 & 4.12: Amplification of the *trn*T - *trn*L region using primers B48557/A49855 at 45°C.

Fig.4.11: Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).

Lanes 1, 3: Amplification positive.

Lane 2: Amplification negative.

Lane 4: Negative control.

Fig.4.12: Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).

Lanes 1, 2 & 4: Amplification positive.

Lane 3: Amplification negative.

Lane 5: Negative control.

AFLP ANALYSIS OF THE *trn*T - *trn*L REGION OF THE CHLOROPLAST DNA:

Fig.4.13: AFL profiles for the *trn*T - *trn*L region of the chloroplast DNA:

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).

Lanes1-15: Amplification profiles of different sample.

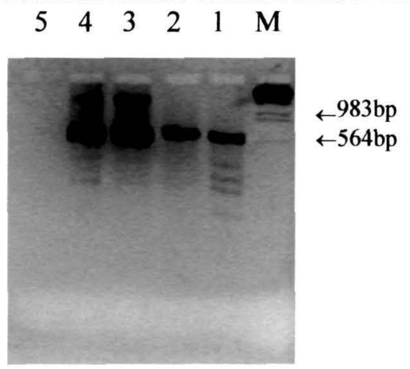


Fig 4.9

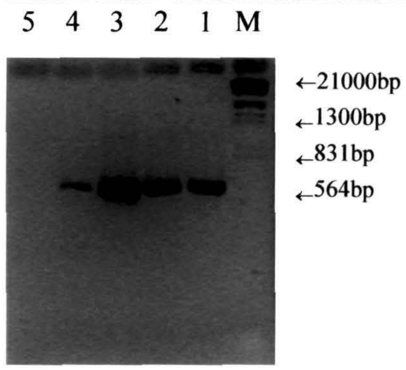


Fig.4.10

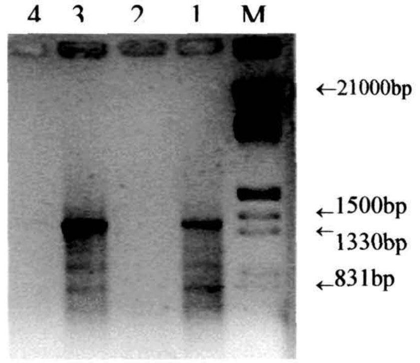


Fig.4.11

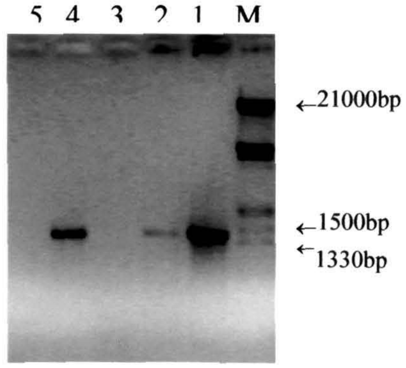


Fig.4.12

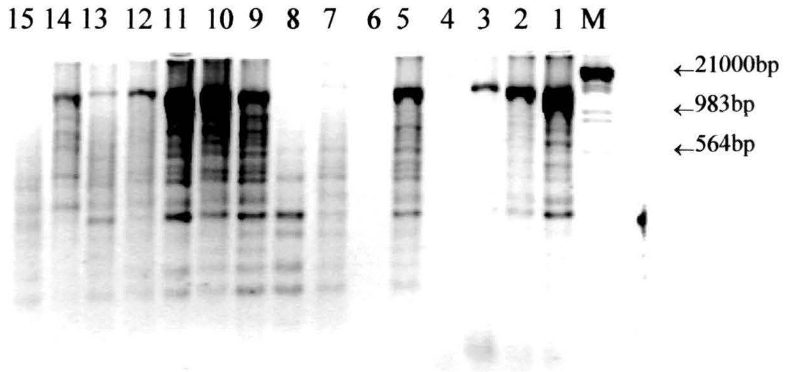


Fig.4.13

45°C. After electrophoresis it was observed that two of the three samples got amplified but, instead of a single expected band of approximately 750bp, multiple bands were present (Fig. 4.14). However, the banding pattern of the two samples differed from each other. To decrease the chances of non-specific amplification the annealing temperature was increased to 47°C and the three samples amplified. An additional dilution of 1/100X was used for the sample that did not amplify earlier. On increasing the stringency conditions, only one of the samples showed amplification but the additional bands that were present at 45°C persisted. The other sample, which had initially amplified, failed to amplify on increasing the stringency (Fig. 4.15). Since, raising the annealing temperature did not eliminate any of the extra bands amplification of all the samples was subsequently done at 45°C (Fig. 4.16 & 4.17). Sequence analysis of *Hamamelis mexicana* (Table 4.8 & 4.9) confirmed our hypothesis of the occurrence of the additional bands due to the presence of alternate annealing sites for the primers in our template DNA. However, amplification of all hundred samples showed differences in the banding patterns. These differences probably occurred because of multiple annealing sites in DNA templates obtained from different trees. To overcome the problem of multiple annealing sites, it was decided to design a more specific probe (Section 4.8).

4.4.6 Amplification of the *rbc*L gene:

The *rbc*L gene coding for the large subunit of the carbon fixing enzyme RUBISCO was amplified using the primer -34^{+} in conjunction with the primer 1517'. The amplicon produced was of approximately 1.5kb. It was difficult to get amplification for this region using these primers because the primers did not perfectly match with the alder sequence. This region was not further analyzed using other primers because literature survey revealed that this gene (Savard *et al.*, 1993, Ludovic and Taberlet, 1994) was a highly conserved gene and so it was not suitable for detecting intra-specific variations.

Fig 4.14-4.17

AMPLIFICATION OF THE 18S-28S *rrn* OPERON USING THE PRIMER PAIR ITS1-PLANT/ITSC26A:

Fig.4.14: Amplification at 45°C.

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest).
Lanes 2 & 5: Amplification positive.
Lanes 1, 3, 4, 6: Amplification negative.
Lane 7: Negative control.

Fig.4.15: Amplification at 47°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-5 &7: Amplification negative.
Lane 6: Amplification positive.
Lane 8: Negative control.

Fig.4.16 & 4.17: Amplification of different samples at 45°C

Fig.4.16: Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 4, 7, 14: Amplification positive.
Lane 1-3, 6, 8-13: Amplification negative.
Lane 15: Negative control.

Fig.4.17: Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 4-7: Amplification positive.
Lanes 1-3: Amplification negative.

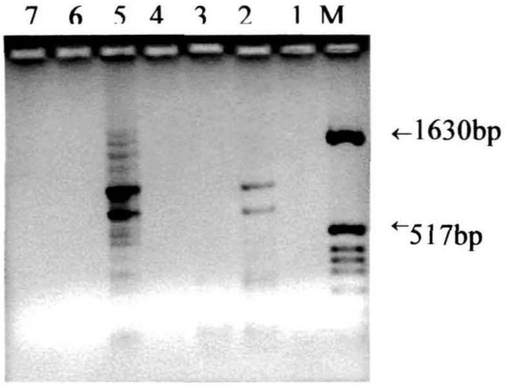


Fig.4.14

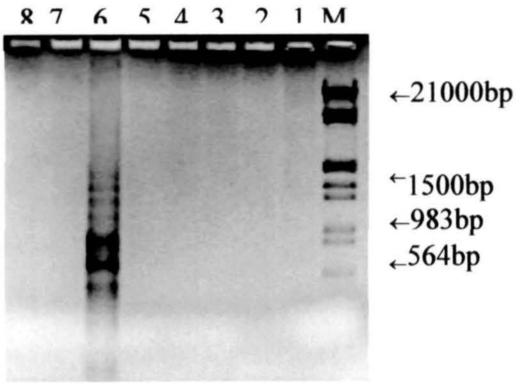


Fig.4.15

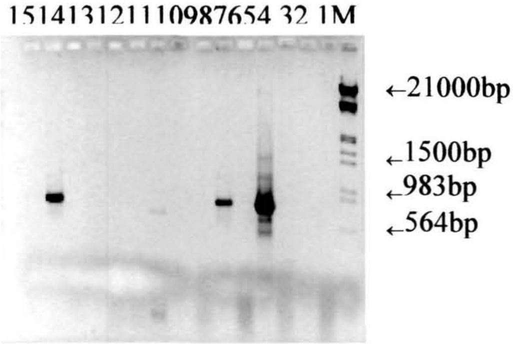


Fig.4.16

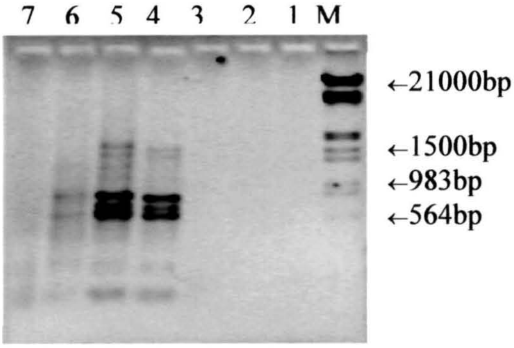


Fig.4.17

ALTERNATE ANNEALING SITES OF PRIMER 22 &23 FOUND IN THE 18S-28S ITS REGION OF *HAMAMELIS MEXICANA*.

TABLE 4.8 PRIMER22 (ITS1-PLANT)

Strict Position	Bases Matched	Strand	Subseq. Found
21	9/16	+	tGCGgaAgGatCAtTG
139	9/16	+	CcCGAtccGcCaACcG
146	9/16	+	CGCcAaccGTgCcCcG
167	9/16	-	CGCGcacgGggCACgG
194	9/16	-	CgaGcGcAcgCCggTG
221	9/16	+	CGCGtcAAGgaaACgc
299	9/16	-	AtCGAagAtgCCAtTG
336	9/16	-	GctGAGAtaTCCgtTG
459	9/16	-	AGgcAGAcGTgCcCTc
558	9/16	-	GcgGcGAAGcgCACgG
584	9/16	+	CGacAGAAcgtCACga
644	9/16	-	CaCGAcAtGaCactTG
659	9/16	-	AGCGAcAgGgCaACgc
738	9/16	-	aGCGgGtAGTCCcgcc

TABLE 4.9 PRIMER 23 (ITSC26A)

Strict Position	Bases Matched	Strand	Subseq. Found
239	11/17	-	GTTgCgTTTCCTtgaCg
634	11/17	+	GtgTCaTgTCgTgCGtT
234	10/17	-	GTTTCcTTgaCgCgttT
530	10/17	+	GgaTaTTggCCTCCcgT
618	10/17	-	GcTTaTcaaCCaCCaCT

4.4.7 Amplification using microsattelite primers:

Though a little cumbersome analysis of the entire genome is the best way to look for variability. In an attempt to do this, we designed two microsattelite primers (Primer No. 17 and 18). However, these primers did not amplify alder DNA, except for an occasional single band. A pre-requisite for designing an appropriate microsattelite primer is the sequence information of the DNA to be amplified. At the start of the work no sequence information was available for alder and so a suitable microsattelite primer could not be designed. Hence, this approach was also abandoned.

4.5 AMPLIFIED FRAGMENT LENGTH PROFILES:

should have some different term!

4.5.1 AFLP analysis of the proximal part of the 16S rRNA gene:

Since amplification of the proximal part of the 16S rRNA gene showed differences in the amplification profiles of different samples, we decided to go in for analysis of the Amplification Fragment Length Profiles. The amplicons of all hundred samples were run on 2% agarose gels at low voltage (50 volts) for four hours. After electrophoresis the gel was visualized and photographed using the Gel Doc 1000. Analysis of all hundred samples showed three different profiles (Fig. 4.18). The three profiles have been named as profile 1, 2, and 3. The approximate size of the bands present in the different profiles was estimated using the Multi-Analyst software®. Samples with profile 1 showed a single band of approximately 470bp. Profile 2 had two bands of approximately 470bp and 570bp and profile 3 showed three bands of 470bp and 570bp and 660bp (Table 4.10). To confirm the origin of the additional bands, a computer-simulated analysis of alder sequences retrieved from the GenBank was done using the Mac Vector software®. Since, *Alnus nepalensis* sequences were not available *Alnus incana* and *Alnus glutinosa* sequences were used to search for alternate annealing sites for the primers. Primer FGPS 505' completely matched with the sequence of *Alnus incana* at 659bp and *Alnus glutinosa* at 420bp. Apart from this

Fig 4.18-4.19

AFLP ANALYSIS OF THE CHLOROPLAST 16S rRNA GENE:

Fig.4.18: AFL profiles for the proximal part of the chloroplast 16S rRNA gene:

LaneM: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 2, 4 & 6: Profile 1
Lanes 3, 8, 9, 11, 12: Profile 2
Lanes 1, 5, 7, 10 & 13: Profile 3

Fig.4.19: AFL profiles for the middle part of the chloroplast 16S rRNA gene:

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest)
Lanes1-12: Amplicons of different sample showing similar profiles.

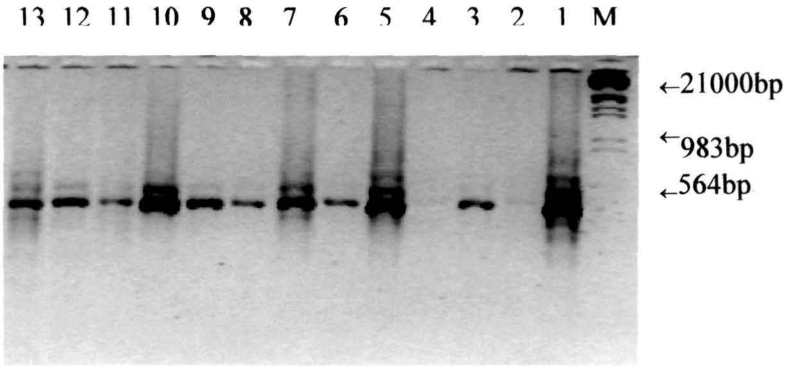


Fig.4.18

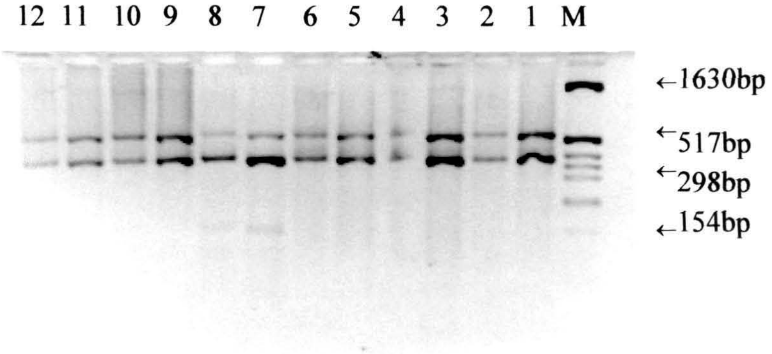


Fig.4.19

TABLE 4.10 COMPARISON OF THE BANDING PATTERNS OF THE AMPLIFICATION FRAGMENT LENGTH PROFILES OBTAIND FOR THE INITIAL PART OF THE 16S rRNA GENE USING THE PRIMERS FGPS 6/FGPS 505'

PROFILE No.	APPROXIMATE BAND SIZE (BASE PAIRS) 660	570	470
1.	-	-	+
2.	-	+	+
3.	+	+	+

+ = Band presence
- = Band absence

Approximate band sizes were found out using the Multi-Analyst software®.

ALTERNATE ANNEALING SITES OF PRIMER 8 & 16 FOUND ON THE CHLOROPLAST 16S rRNA GENE OF *ALNUS INCANA*

TABLE 4.11 PRIMER 8 (FGPS6)

Strict Position	Bases Matched	Strand	Subseq. Found
40	13/21	+	TGGAtgGCTatATttCTGGg
180	13/21	+	TGGAgAGtTcGATCCtgGct
1757	13/21	+	TGGAAgtCTTctTtCgTttC

TABLE 4.12 PRIMER 16 (FGPS 505')

Strict Position	Bases Matched	Strand	Subseq. Found
659	18/18	-	GTATTACCGCGGCTGCTG
327	12/18	+	CTAaTACCcCGtagGCTG
653	12/18	+	GTAaTACaGaGGaTGCaa
797	11/18	+	GaATTtCCGgtGgaGCgG
857	11/18	-	GTgcTttCGCcGtTGgTG

ALTERNATE ANNEALING SITES OF PRIMER 8 & 16 FOUND ON THE CHLOROPLAST 16S rRNA GENE OF *ALNUS GLUTINOSA*.

TABLE 4.13 PRIMER 8 (FGPS6)

Strict Position	Bases Matched	Strand	Subseq. Found
514	12/21	-	GtccAcGgcTcAaCCCTGGa
871	12/21	-	TGGtAAGgTTctTCgCTttg
21	11/21	+	TGGAAAcaccacTtCCcGtC
180	11/21	-	TGGtAAGCTattgCCtcacC
391	11/21	-	gccgAtGCTTatTCCCcaGa
403	11/21	-	gGcAcAGagTtAgCCgatGC
439	11/21	-	gatAAcGCTTGcatCCTctg
494	11/21	+	TcccAgGgcTcAaCCCTGGa
609	11/21	+	GcGAAAGCactcTgCtgGGC
622	11/21	-	CaGAgtGCTTtcgCCgTtGg
696	11/21	+	GtaAAcGaTgGATaCtaGGC
785	11/21	-	TGcgAAcgTactcCCCaGGC
843	11/21	-	TaaAccaCaTGcTCCaccGC
924	11/21	+	gGacAcaggTGgTgCaTGGC

TABLE 4.14 PRIMER 16 (FGPS 505')

Strict Position	Bases Matched	Strand	Subseq. Found
420	18/18	-	GTATTACCGCGGCTGCTG
89	12/18	+	CTAaTACCcCGtagGCTG
414	12/18	+	GTAaTACaGaGGaTGCaa
558	11/18	+	GaATTtCCGgtGgaGCgG
618	11/18	-	GTgcTttCGCcGtTGgTG
980	11/18	-	GgATTACCcCaaCacCTt

site it showed alternate annealing sites as well. Primer FGPS 6 did not completely match with either of the sequences. However, if a few mismatches were allowed, multiple annealing sites were found for this primer. Tables 4.11-4.14 show the alternate annealing sites for primers FGPS 6 and FGPS 505'. The annealed primer will get extended only at those sites where only a few mismatches are present towards the 3' end. In such a condition, stable binding would occur at 3' end of the primer at lower stringency conditions giving rise to multiple bands. However, since the sequence for *Alnus nepalensis* was not available, it could only be hypothesized that the additional bands were a consequence of similar annealing sites in our samples. A look at the possible alternate site pairs does explain the appearance of additional bands. In fact many more bands were possible. Apparently, the three bands obtained were due to the stringency of the conditions of amplification. The variability observed in this region is in accordance to previous reports on the 16S rRNA gene suggesting that the initial part of this gene showed more variations (Harry *et al.*, 1991).

4.5.2 AFLP analysis of the middle part of the 16S rRNA gene:

To search for polymorphism in the amplification profiles of different samples, the amplicons were electrophoresed on a 2% agarose gel at low voltage (50 volts) for 4 hours and then observed and photographed using the Gel Doc 1000. No differences were seen in the amplification profiles of different samples (Fig. 4.19), thus suggesting that this is a relatively less variable region. However, the two bands could be explained by the annealing of primer FGPS 485 at position 642 and primer FGPS 910' at 1029 to give the approximately 400bp band, and primer FGPS 485 annealing at position 983 and primer FGPS 910' annealing at 1602 giving rise to the approximately 550bp band.

4.5.3 AFLP analysis of the *trn*T-*trn*L region:

To get a better resolution of all the bands, the amplicons of different samples were subjected to electrophoresis on a 2% agarose gel at low voltage (50 volts) for four hours. When observed all the samples showed multiple bands (Fig. 4.13). To search for alternate annealing sites for both primers, computer analysis of the published sequence of the *trn*L gene of *Alnus viridis* was done. The *trn*T gene was not analyzed since no published sequence of alder was available for this region. It was found that primer A49855 did not completely match with the alder sequence and therefore, it would anneal with the alder template only if the stringency conditions were low. Under these conditions it would not anneal to one site but had a number of alternate annealing sites. Unexpectedly, primer B48557 (*trn*T gene) also showed multiple low stringency annealing sites, though not exact matches, in this region (Table 4.15 & 4.16). Since, the primers were not perfect matches of our template, increasing the annealing temperature was not possible because it would even eliminate the band of interest. The AFLP gels showed too many bands and these bands were too close to each other. Hence, it was difficult to categorize the different profiles. This region was therefore, not used while correlating the different PCR-AFLP/RFLP profiles to the nitrogenase activity (Section 4.10).

4.5.4 AFLP analysis of the 18S-28S *rrn* operon:

For better resolution of all the bands, the amplicons of the 18S-28S *rrn* region produced using the primer pair ITS1-PLANT/ITSC26A were subjected to electrophoresis on 1.5% agarose gels at 50 volts for four hours. The gels were visualized and captured using the Gel Doc 1000 (Fig. 4.20 & 4.21). Analysis of the amplification profiles was done using the Multi-Analyst software® and the approximate band sizes were estimated. The profiles obtained for all hundred samples were compared and nine different amplification profiles (A-I) were revealed. Table 4.17 shows the banding patterns of these profiles.

ALTERNATE ANNEALING SITES OF PRIMER 19 FOUND ON THE CHLOROPLAST *trn*L GENE OF *ALNUS VIRIDIS*.

TABLE 4.15 PRIMER 19(A49855)

Strict Position	Bases Matched	Strand	Subseq. Found
50	11/20	-	GGGttTctctGaAtTTGAAa
146	11/20	+	GtGcAgAGActcAaTgGAAg
370	11/20	-	TtGGATcGAttcACaacAAC
485	11/20	-	AtGtAgAatGGGACTctAtc
42	10/20	+	GaGaAaccctGGAaTTaAAa
93	10/20	-	CGGaAaAcAGGattTgGctC
94	10/20	-	TcGGAaAacaGGAtTTGgct
97	10/20	+	AcaaATAaAacaAaTTGAAg
118	10/20	+	GttcATAaAGcGAgactAAa
120	10/20	+	TcatAaAGcGaGACTaaAAa
222	10/20	-	TttGATGAaGGAtTacttt
348	10/20	+	aGaGtTgttGtGAaTcGAtC

TABLE 4.16 PRIMER 21(B48557)

Strict Position	Bases Matched	Strand	Subseq. Found
397	13/20	-	TAaTAaAtATtCGATtCTtT
54	12/20	+	AATTAaAAATGgGcaatcCT
64	12/20	-	CATTtttAATtCcAgGgTtT
486	12/20	-	CATgtagAATGgGActCTaT
170	11/20	+	TcTaACAAtTGgGgTtgaCT
196	11/20	+	CgTTAttAAaGtaATcCTtc
102	10/20	+	TAaaACAAATtgaAgGgTtc
117	10/20	+	GgTTcatAAaGCGAgaCTaa
131	10/20	+	GAcTAaAAAgGatAgGtgCa
206	10/20	-	CtTTAatAAcGgaAcGaagT
222	10/20	+	AAcTACAtAaagGATGaagg
261	10/20	+	TATacgtAcTGaaATaCTaT
271	10/20	-	CAgTACgtATaCatatgTaT
339	10/20	+	CtgaAtgAAaGaGtTGtTgT
390	10/20	+	AtTTAttAATcaaATtaTtT
476	10/20	+	CATTctAcATGtcAatacCg
516	10/20	-	tAcTAtAAATttcATtgTtg

Fig 4.20-4.23

AFLP ANALYSIS OF THE 18S-28S *rrn* OPERON:

Fig.4.20 & 4.21: AFLP profiles obtained for the 18S-28S ITS region using the primers ITS1-PLANT/ITSC26A.

Fig.4.20

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 6, 7, 10, 13 & 15: Profile A.
Lanes 2, 3 & 5: Profile B.
Lanes 12: Profile C.
Lane 4: Profile D.
Lanes 1, 8 & 9: Profile G.
Lane 14: Profile H.

Fig.4.21

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf* I Digest).
Lane 1: Profile A.
Lane 2: Profile E.
Lane 4: Profile F.

PURIFICATION OF AMPLICONS FOR SEQUENCING:

Fig.4.22: Amplicons of the 18S-28S ITS region used for sequencing, prior to purification.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1: Amplicon of sample 1.
Lane 2: Amplicon of sample 2.
Lane 3: Amplicon of sample 3.

Fig.4.23: Amplicons of the 18S-28S ITS region used for sequencing, after purification.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1: Amplicon of sample 1.
Lane 2: Amplicon of sample 2.
Lane 3: Amplicon of sample 3.

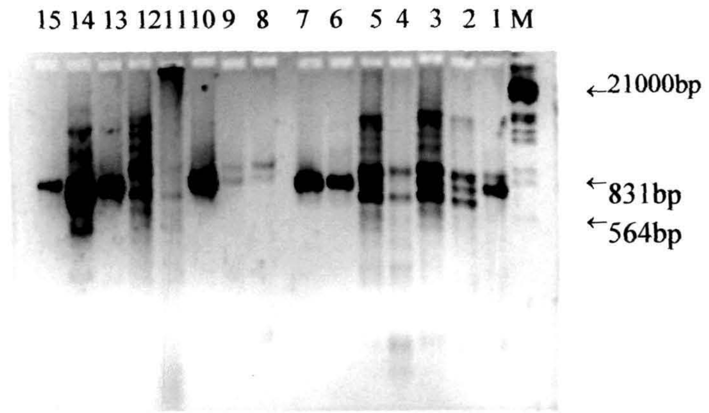


Fig.4.20

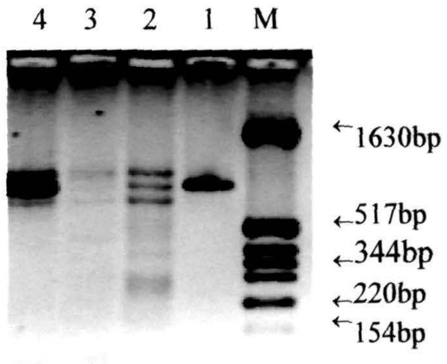


Fig.4.21

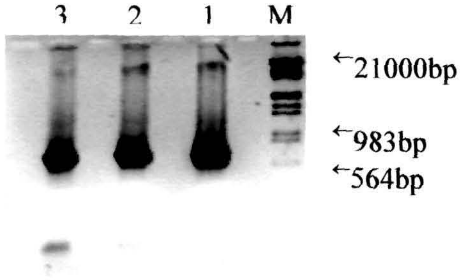


Fig.4.22

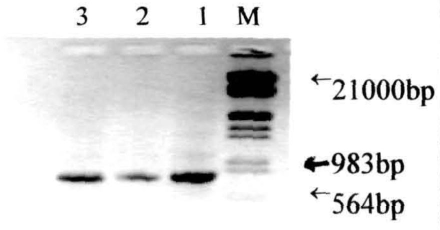


Fig.4.23

TABLE 4.17 COMPARISON OF THE BANDING PATTERNS OF THE AMPLIFICATION FRAGMENT LENGTH PROFILES OBTAIND FOR THE 18S-28S *rrn* OPERON USING THE PRIMERS ITS1-PLANT/ITSC26A

PROFILE No.	APPROXIMATE BAND SIZE (BASE PAIRS) 2000	1800	1550	1490	1210	1075	925	750	600	550	520	475	390	280
A	-	-	-	-	-	-	-	+	-	-	-	-	-	-
B	+	+	+	-	+	-	+	+	+	+	+	-	-	-
C	+	+	+	-	+	-	+	-	+	+	+	-	-	-
D	-	-	-	-	-	-	+	-	+	-	-	+	+	+
E	-	-	-	-	-	-	+	+	+	-	-	+	+	+
F	-	-	-	-	-	-	+	+	+	-	-	-	-	-
G	-	-	-	-	-	-	+	+	-	-	-	-	-	-
H	-	-	-	+	-	+	-	+	-	-	-	+	-	-
I	-	-	-	-	-	-	-	-	-	-	-	-	-	-

\+ = Band presence
\- = Band absence

Approximate band sizes were found out using the Multi-Analyst software®.

4.6 NUCLEOTIDE SEQUENCING:

As a first step towards designing of specific primer, we decided to sequence three of our samples. We selected the 18S-28S *rrn* operon for the purpose, primarily because of the availability of two highly variable regions in the form of ITS I and ITS II.

After electrophoresis the 750bp DNA band was eluted out from the gel and purified in order to eliminate the excess primers and dNTPs (Fig. 4.22 & 4.23). The purified samples were then sent to M/S Bangalore Genei for sequencing. Sequencing was done using the primers used for amplification (Primer ITS1-PLANT and ITSC26A). Each primer was used singly in a sequencing cycle. Sequences were obtained in the form of electropherograms where the four bases were represented by different colours (Fig. 4.24). Sequencing the template with primer ITS1-PLANT gave sequence information of about 130-140bp including the distal region of the 18S rRNA gene and ITS I. Since the molecular weight of the PCR product was approximately 750bp and the sequence obtained was only about 130bp, it was hypothesized that the enzyme stuttered and failed to proceed further due to the presence of several GC residues. We therefore, used the reverse primer (primer ITSC26A) along with DMSO to sequence the three samples. This gave us sequence information for about 467, 318 and 200bp respectively. The expected location of primer ITSC26A is in the initial part of the 28S rRNA gene and therefore, this primer should give sequence information of the opposite strand i.e., the strand complementary to the one sequenced earlier. A computer simulated analysis of the sequence of the 18S-28S ITS region of *Alnus matsumurae* was done using Mac Vector software®. This analysis revealed that 413bp downstream the start of the sequence of *Alnus matsumurae,* an alternate annealing site with a stable 3' end was present on the positive strand for this primer. However, the primer would anneal at this site only at low stringency conditions as this site was not a complete match. While amplifying the DNA the primer would not anneal to this site because of the high annealing temperature. Since sequencing is done at lower

Fig 4.24 : Electopherogram showing partial sequence of 18S-28S ITS region of *Alnus nepalensis* sample1. The different bases have been represented by different colours.

ABI PRISM™

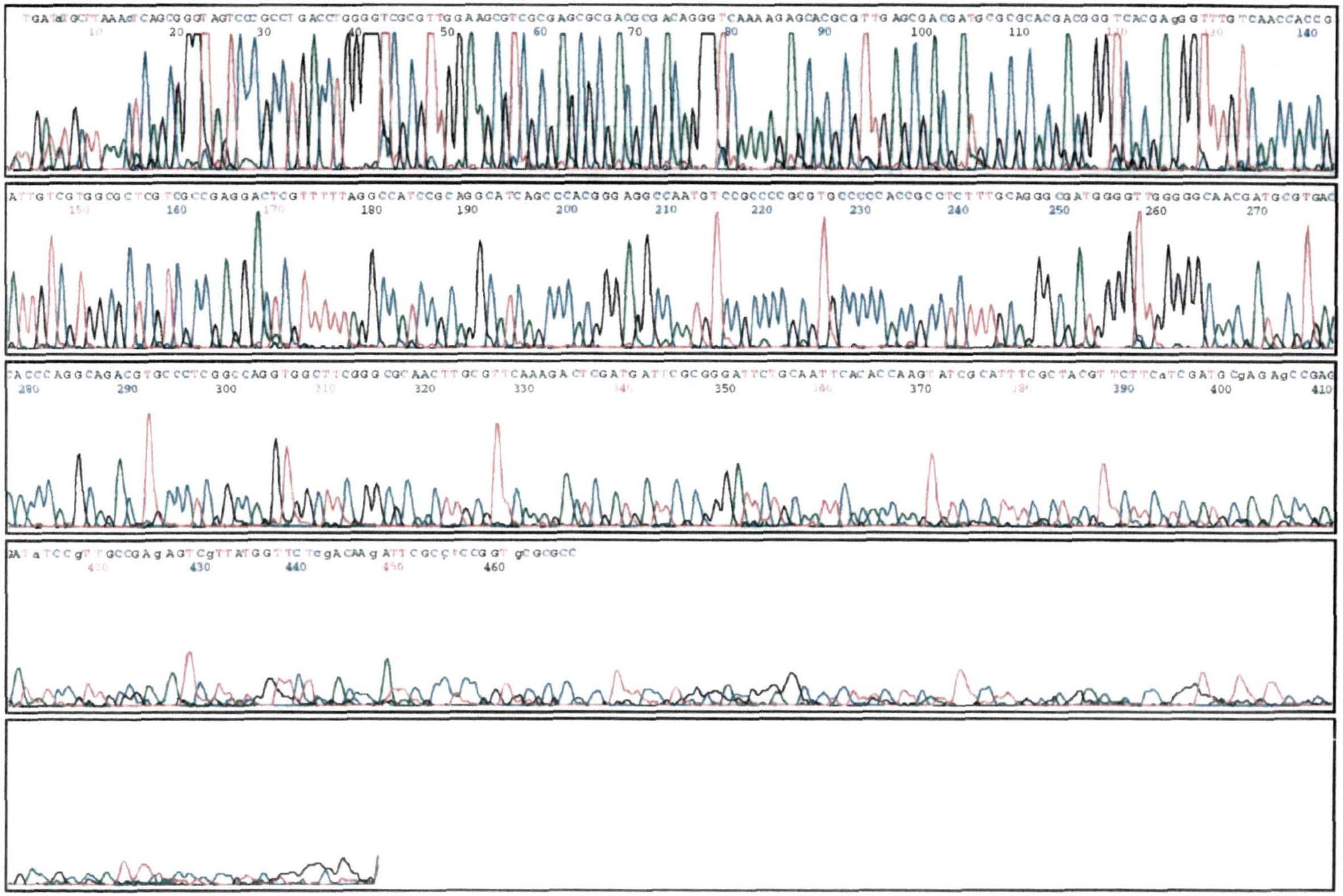
Model 377
Version 3.0
ABI200
Version 3.0

25•BG 23

BG 23
Lane 25

Signal G:73 A:133 T:119 C:90
DT4%Ac{A Set-AnyPrimer}
377 matrix/1
Points 917 to 7000 Base 1: 917

Mon, Mar 22, 1999 12:40 PM
Mon, Mar 20, 2000 4:32 PM
Spacing: 11.13{11.13}



temperature, the stringency would be low. Therefore, the chances of the primer annealing at this site were high. It was thus possible for us to sequence the same strand as that sequenced using primer ITS1-PLANT. When these sequences were compared with sequences of other species of alder and the 130-140bp sequence of *Alnus nepalensis* , it was found that primer ITSC26A had indeed sequenced the strand in continuation to that sequenced using primer ITS1-PLANT leaving unsequenced a stretch of about 290bp in between . The sequence obtained included the distal part of ITS I, 5.8S rRNA gene, ITS II and the initial part of the 28S rRNA gene. The sequence information was utilized for designing a genus specific probe as a first step for the identification of superior genotypes of alder (Section 4.8). The sequences of the species of alder showed a lot of similarity. Therefore, these sequences were utilized to infer phylogenetic relationships (Section 4.7).

The 18S-28S *rrn* operon sequences of a few other species of *Alnus* and a few members of the family Betulaceae were retrieved from the GenBank (ncbi.nlm.nih.gov) (Table. 4.18) after Blast Search. The sequences for the 18S-28S *rrn* operon of the three samples of *Alnus nepalensis* were aligned with these sequences using the multiple sequence alignment program CLUSTAL W (1.75) (http://www.es.embnet.org/Doc/phylodendron/clustal-form.html) (Fig. 4.25 & 4.26).

4.7 PHYLOGENETIC ANALYSIS:

For Phylogenetic studies, the aligned sequences as described in section 4.6 were used to construct phylogenetic trees. Parsimony analysis was carried out using the DNAPARS (parsimony program), SEQBOOT (bootstrap) and CONSENSE (consensus tree program) programmes of the PHYLIP (version 3.5c) inference package (Felsenstein, 1993 site http://bioweb.pasteur.fr/seqanal/phylogeny/phylip-uk.html). To be confident of the parsimony trees constructed, 500 bootstrap replicates were considered. Neighbor Joining trees were constructed using the program PHYLODENDRON (version 0.8d, beta by D.G.

TABLE 4.18 SEQUENCES OF 18S-28S INTERNALLY TRANSCRIBED SPACER, USED FOR PHYLOGENETIC ANALYSIS

Tree Genus & species	GenBank Accession Number	Source/ Reference
Alnus nepalensis (S.1)	-	Present study
Alnus nepalensis (S.2)	-	Present study
Alnus nepalensis (S.3)	-	Present study
Alnus matsumurae	-	Normand (Personal Communication)
Alnus incana	X68138	Savard *et al.*, 1993
Alnus crispa	X68137	Savard *et al.* ,1993
Alnus glutinosa	AF081529	Chen *et al.*, 1999 *(Unpublished)
Corylus chinensis	AF081520	Chen *et al.*, 1999 *(Unpublished)
Carpinus turczaninowii	AF081518	Chen *et al.*, 1999 *(Unpublished)
Betula pendula	AJ006445	Leskinen E, 1999 *(Unpublished)

* Sequences are direct submissions to the GenBank and are available at the site http.ncbi.nlm.nih.gov

Fig.4.25 Sequences of the distal part of the 18S rRNA gene and part of ITS I of *Alnus nepalensis* aligned to the sequences retrieved from the GenBank. The site where primer 24 is located has been highlighted by underlining and the bases written in bold highlight the variations at this site.

CLUSTAL W (1.75) multiple sequence alignment

ALIGNED SEQUENCES OF THE DISTAL PART OF THE18S rRNA GENE & PART OF ITS1

```
A.nepalensisS1    GACTTATCATTTAGAGGNAGGAGAAGTCGTAACAAGGTT-TCCGTAGGTGAACCTGCGGA
A.nepalensisS2    -ACTTATCATTTAGAGGAAGGAGAAGTCGTAACAAGGTTATCCGTAAGTGAACCTGCGGA
A.nepalensisS3    -ACTTATCATTTAGAGGAAGGAGAA-TCGNAACAAGGTT-TCCGTAGGTGAACCNGCGGA
A.matsumurae      --------------------------TCGTAACAAGGTT-TCCGTAGGTGAACCTGCGGA
A.incana          ------------------------------------------------------------
Betula            -----------------------------------------------GTGAAC-TGCGGA
Corylus           ------------------------------------------------------------
Carpinus          ------------------------------------------------------------
A.glutinosa       ------------------------------------------------------------
A.crispa          ------------------------------------------------------------


A.nepalensisS1    AGGATCATTGTCGAAACCTGCCCAGCAGAACGACCCGCGAACCTGTCACAACAACCGGGG
A.nepalensisS2    AGGATCATTGTCGAAACCTGCCCAACAGAACGACCCGCGAACCTGTCACAACAACCGGGG
A.nepalensisS3    AGGATCATTGTCGAAACCTGCCCAGCANAACAACCCGCGAACCTGTCACAACAACCGGGG
A.matsumurae      AGGATCATTGTCGAAACCTGCCCAGCAGAACGACCCGCGAACCTGTCACAACAACTGGGG
A.incana          -----------NNAAACCTGCCCAGCAGAACGACCCGCGAACCTGTCACAACAACTGGGG
Betula            AGGATCATTGTCGAAACCTGCCCAGCAGAACGACCCGTGAACCTGTTGAAACAACTGGGG
Corylus           ----------TCGAAACCTGCCCAGCAGAACGACCCGCGAACTTGTATAAACAACTGGGG
Carpinus          ----------TCGAAGCCTGCCCAGCAGAACGACCCGCGAACTTGTATAAACAACCGGGG
A.glutinosa       ------------------------------------------------------------
A.crispa          -----------NNNNNNNNNNNNNNNNGAAGCACCCGCGAAC-TGTCACAACAACTGGGG


A.nepalensisS1    G-CGGGGGGCTTCGCC-GCGCCCCGCCCCC------------------------------
A.nepalensisS2    G-CGGGGGGCTTCGCC-GCGCCCCGCCCC-------------------------------
A.nepalensisS3    G-CGGGGGGCTTC-----------------------------------------------
A.matsumurae      G-CGGGGGGCGATCTC-GCGCCCCGCCCTCGAACGGCAGGG-AGACACTCGTGCCTTCCT
A.incana          G-CGGGGGGCGATCTC-GCGCCC-GCCCTCGAACGGCAGGA-AGACACTCGTGCCTTCCT
Betula            G-TGTGGGGCGATCTC-GCCCCTTGCCCCCGAACGGTAGGG-AGACACTTGTGCATCCCT
Corylus           GGCGGGGGGCGTTCTCTGCCCCGTGCCCCCGAACGGCGGGGCAGACACTCGTGCCTTCTT
Carpinus          G--CAGGGGCGATCTC-GCCCCGTGCCCTCGAACGGCAGGG-AGACACTCGTGCCTTCTT
A.glutinosa       ------------------------------------------------------------
A.crispa          G-CGGNNGGCGATCTC-GCACCCCGCCCTCGAACGGCAGGG-AGA-ACTCGTGCCTTCCT


A.nepalensisS1    ------------------------------------------------------------
A.nepalensisS2    ------------------------------------------------------------
A.nepalensisS3    ------------------------------------------------------------
A.matsumurae      GCCGAACAACGTACCCCGGCGCGG-TCCGCGCCAAGGAACATGAACGAAAGAGTGCCTCC
A.incana          GCCGAACAACGTACCCCGGCGCGG-TCTGCGCCAAGGAACATGAACGAAAGAGTGCCTCC
Betula            GCCGAACAACGAACCCCGGCGCGG-TCCGCGCCAAGGAACTTTAACGAAAGA-TGCCTCC
Corylus           GCCGAACAACGAACCCCGGCGCGG-TCTGCGCCAAGGAACTTCAACTAAAGAGTGCCTCC
Carpinus          GTCGAACAACGAACCCCGGCGCGGGTCTGCGCCAAGGAACTTCAATTAAAGAGTGCCTCC
A.glutinosa       ------------------------------------------------------------
A.crispa          GCCGAACAACGTACCCCGGCGCGG-TCTGCGCCAAGGAACATTAACGAAAGAGTGCCTCC
```

```
A.nepalensisS1     ------------------------------------------------------------
A.nepalensisS2     ------------------------------------------------------------
A.nepalensisS3     ------------------------------------------------------------
A.matsumurae       GGTCGCCTCGGAAACGCTGCGCGCGCCGGAGGCGAATCTTGTCTAGAACCATAACGACTC
A.incana           GGTCGCCTCGGAAACCGTGCGCGCACCGGAGGCGAATCTTGTCTAGAACCAATCGTTGCC
Betula             GGCCGCCTCGGAAACGGTGTGCGTGCGGGAGGTGAATCTTGTCTAGAACCATAACGACTC
Corylus            GGTCGCCTCGGAAACGGCGTGCGTGCCGGAGGCGAATCTTGTGCAAAACCATAACGACTC
Carpinus           GGTCGCCTCGGAAACG-TGCGCGTGTCGGAGGCGAATCTTGTACAAAACCATAACGACTC
A.glutinosa        ------------------------------------------------------------
A.crispa           GGTAGCCTCGGAAACGCTGTGCTTGCCGGAGGCGAATCTTGTCTANNNNNNATCGTTGCC
```

```
A.nepalensisS1     ------------------------------------------------------------
A.nepalensisS2     ------------------------------------------------------------
A.nepalensisS3     ------------------------------------------------------------
A.matsumurae       TCGG---CAACGGATATCTCGGCTCTCGCATCGATG-AAGAACGTAGCGAAATGCGATAC
A.incana           CCCAACCCCATCGCCCTGCAAAGAGGCGGTGGGGGC-ATGCGGGGCGGACATTGGCCTCC
Betula             TCGG---CAACGGATATCTCGGCTCTCGCATCGATG-AAGAACGTAGCGAAATGCGATAC
Corylus            TCGG---CAACGGATATCTCGGCTCTCGCATCGATG-AAGAACGTAGCGAAATGCGATAC
Carpinus           TCGG---CAACGGATATCTCGGCTCTCGCATCGATG-AAGAACGTAGCGAAATGCGATAC
A.glutinosa        ------------------------------------------------------------
A.crispa           CCCAACCCCATCGCCCTGCAAAGAGGCGGTTGGGGGCATGCGGGGCGGACATTGGCCTCC
```

```
A.nepalensisS1     ------------------------------------------------------------
A.nepalensisS2     ------------------------------------------------------------
A.nepalensisS3     ------------------------------------------------------------
A.matsumurae       TTGGTGTGAATTGCAGAATCCCGCGAATCATCGAGTCTTTGAACG-CAAGTTGCGCCCGA
A.incana           CGTGGGCTGATGCCTGCGGCTGGCCTAAAAACGAGTCCTCGGCGA-CGATCGCCACGACA
Betula             TTGGTGTGAATTGCAGAATCCCGCGAATCATCGAGTCTTTGAACG-CAAGTTGCGCCCGA
Corylus            TTGGTGTGAATTGCAGAATCCCGCGAATCATCGAGTCTTTGAACG-CAAGTTGCGCCCGA
Carpinus           TTGGTGTGAATTGCAGAATCCCGCGAATCATCGAGTCTTTGAACG-CAAGTTGCGCCCGA
A.glutinosa        ------------------------------------------------------------
A.crispa           CGTGGGCTGATGCCTGCGGCTGGCCTAAAAACGAGTCCTCGGCCGACGAGCGCCACGACA
```

```
A.nepalensisS1     ------------------------------------------------------------
A.nepalensisS2     ------------------------------------------------------------
A.nepalensisS3     ------------------------------------------------------------
A.matsumurae       AGCCACCTGGCCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGCCCCCAACCCCATC
A.incana           ATCGGTGGTTGACAAACCTTCGTGACCCGTCGTGCGCGCATCGCCGCTCAACGCGTGCTC
Betula             AGCCACCTGGCCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGCCCCCAACCCCATC
Corylus            AGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTTGCCCCCAACCCCATC
Carpinus           AGCCATCTGGTCGAGGGCACGTCTGCCTGGGTGTCACGCATCGTCGCCCCCAACCCCATC
A.glutinosa        ---------------------------------------ATCGTTGCCCCCAACCCCATC
A.crispa           ATCGGTGGTTGACAAACCCTCGTNNNNNNCGTG-CGTGCATCGTCGCTCAATGTGTGCTC
```

```
A.nepalensisS1      ------------------------------------------------------------
A.nepalensisS2      ------------------------------------------------------------
A.nepalensisS3      ------------------------------------------------------------
A.matsumurae        GCCCTGCAAAGAGGCGGCGGGGGCGCGC--GGGGGCGGACATTGGCCTCCCGTGGGCTGA
A.incana            TTTTGACCCTGTCGCGTCGCGCTCGCGA--CGTTCCAACGCGA-----------------
Betula              TCCTTGCAAAGGGACGAGGGGGCCTGT----GGGGCAGAAATTGGCCTCCCGTGAGCTCA
Corylus             ATCGCCTCTCCAAGAGACGAGGGCGGTCTGCGGGGCGGACATTGGCCTCCCGTGAGCTTT
Carpinus            GCCTCTCCAAGA---GACGAGGGCAGTTTGCGGGGCGGACATTGGCCTCCCGTGAGCTTC
A.glutinosa         GCCCTGCAAAGAGGCGGTGGGGGCATG---CGGGGCGGACATTGGCCTCCCGTGGGCTGA
A.crispa            TTTTGACCCTGTCGCGTCGCTCGCGAN----NNNNNNNNNNNNN----------------

A.nepalensisS1      ------------------------------------------------------------
A.nepalensisS2      ------------------------------------------------------------
A.nepalensisS3      ------------------------------------------------------------
A.matsumurae        TGCCTGCGGCTGGCCTAAAAACGAGTCCTCGGCGACGATCGCCACGACAATCGGTGGTTG
A.incana            ------------------------------------------------------------
Betula              TGCATGCGGTTGGCCTAAAAGCGAGTCCTCGGCGACGCGCGCCACGACAATCGGTGGTTG
Corylus             CGCTCGCGGCTGGCCTAAAAGCGAGTCCTCGGCGACGAGCGCCACGACAATCGGTGGTTG
Carpinus            CACTTGCGGTTGGCCTAAAAGCGAGTCCTAGGCGACGAGCGCCACGACAATCGGTGGTTG
A.glutinosa         TGCCTGCGGTTGGCCTAAAAACGAGTCCTCGGCTGACGATCGCACDACAATCGGTGGTTG
A.crispa            ------------------------------------------------------------

A.nepalensisS1      ------------------------------------------------------------
A.nepalensisS2      ------------------------------------------------------------
A.nepalensisS3      ------------------------------------------------------------
A.matsumurae        ACAAA-CCCTCGTGACCCGTCGTGCGCGCACCGTCGCTCAACGCGTGCTCTTTTGACCCT
A.incana            ------------------------------------------------------------
Betula              TCAAA-CCCTCGTGTCCCGTCGTGCGTGCCGCGTCGCTCATCGTGTGCTCCTT-GACCCT
Corylus             ACAAA-CCCTCGTGTCCCGTCGTGCGGGCT-CGTCGCTCGTCTTGTGCTCTGT-GACCCT
Carpinus            CCAAAACCCTCGTGTCCCGTCGTGCGTGCCTCGTTGCYCATCCTGTGCTCTGT-GACCCT
A.glutinosa         ACAAA-CCTTCGTGACCCGTCGTGCGCGCATCG-CGCTCAACGCGTGCTCTTTTGACCCT
A.crispa            ------------------------------------------------------------

A.nepalensisS1      ------------------------------------------------------------
A.nepalensisS2      ------------------------------------------------------------
A.nepalensisS3      ------------------------------------------------------------
A.matsumurae        GTCGCGTCGCGCTCGCGACGCTTCCAACGCGACCCCAGGTCAGGCGGGACTACCCGCTGA
A.incana            ------------------------------------------------------------
Betula              GTTGTGTCGCGCTAGCGACGCTTCCAACGCGACCCCAGGTCAGGCGGGACTACCCGTGAA
Corylus             GTAGCGTCGCGCTCGCGACTCTTCCAACGCGA----------------------------
Carpinus            ATAGCGTCGCGATCGCGACTCTTCCAATGCGA----------------------------
A.glutinosa         GTCGCGTYGYGCTCGCGACGCTTCCAACGCGA----------------------------
A.crispa            ------------------------------------------------------------

A.nepalensisS       ------------------------------------
A.nepalensisS2      ------------------------------------
A.nepalensisS3      ------------------------------------
A.matsumurae        GTTTAAGCATATCAATAAGCGGAGGAAAAGAAAACA
A.incana            ------------------------------------
Betula              TTTAA-------------------------------
Corylus             ------------------------------------
Carpinus            ------------------------------------
A.glutinosa         ------------------------------------
A.crispa            ------------------------------------
```

Fig.4.26: Sequences of the region between the distal part of ITS I and initial part of 28S rRNA gene of *Alnus nepalensis* aligned to the sequences retrieved from the GenBank. The asterisks mark the sites where all the samples showed similarity.

CLUSTAL W (1.75) multiple sequence alignment

ALIGNED SEQUENCES OF THE 18S-28S INTERNALLY TRANSCRIBED SPACER REGION.

```
A.incana          --------------------------------------------NNAAACCTGCCCAGCA
A.matsumurae      TCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTGTCGAAACCTGCCCAGCA
Betula            --------------------GTGAAC-TGCGGAAGGATCATTGTCGAAACCTGCCCAGCA
Corylus           -------------------------------------------TCGAAACCTGCCCAGCA
Carpinus          -------------------------------------------TCGAAGCCTGCCCAGCA
A.glutinosa       ------------------------------------------------------------
A.crispa          --------------------------------------------NNNNNNNNNNNNNNNN
A.nepalensisS3    ------------------------------------------------------------
A.nepalensisS2    ------------------------------------------------------------
A.nepalensisS1    ------------------------------------------------------------


A.incana          GAACGACCCGCGAACCTGTCACAACAACTGGGGG-CGGGGGGCGATCTC-GCGCCC-GCC
A.matsumurae      GAACGACCCGCGAACCTGTCACAACAACTGGGGG-CGGGGGGCGATCTC-GCGCCCCGCC
Betula            GAACGACCCGTGAACCTGTTGAAACAACTGGGGG-TGTGGGGCGATCTC-GCCCCTTGCC
Corylus           GAACGACCCGCGAACTTGTATAAACAACTGGGGGGCGGGGGGCGTTCTCTGCCCCGTGCC
Carpinus          GAACGACCCGCGAACTTGTATAAACAACCGGGGG--CAGGGGCGATCTC-GCCCCGTGCC
A.glutinosa       ------------------------------------------------------------
A.crispa          GAAGCACCCGCGAAC-TGTCACAACAACTGGGGG-CGGNNGGCGATCTC-GCACCCCGCC
A.nepalensisS3    ------------------------------------------------------------
A.nepalensisS2    ------------------------------------------------------------
A.nepalensisS1    ------------------------------------------------------------


A.incana          CTCGAACGGCAGGA-AGACACTCGTGCCTTCCTGCCGAACAACGTACCCCGGCGCGG-TC
A.matsumurae      CTCGAACGGCAGGG-AGACACTCGTGCCTTCCTGCCGAACAACGTACCCCGGCGCGG-TC
Betula            CCCGAACGGTAGGG-AGACACTTGTGCATCCCTGCCGAACAACGAACCCCGGCGCGG-TC
Corylus           CCCGAACGGCGGGGCAGACACTCGTGCCTTCTTGCCGAACAACGAACCCCGGCGCGG-TC
Carpinus          CTCGAACGGCAGGG-AGACACTCGTGCCTTCTTGTCGAACAACGAACCCCGGCGCGGGTC
A.glutinosa       ------------------------------------------------------------
A.crispa          CTCGAACGGCAGGG-AGA-ACTCGTGCCTTCCTGCCGAACAACGTACCCCGGCGCGG-TC
A.nepalensisS3    ------------------------------------------------------------
A.nepalensisS2    ------------------------------------------------------------
A.nepalensisS1    ------------------------------------------------------------


A.incana          TGCGCCAAGGAACATGAACGAAAGAGTGCCTCCGGTCGCCTCGGAAACCGTGCGCGCACC
A.matsumurae      CGCGCCAAGGAACATGAACGAAAGAGTGCCTCCGGTCGCCTCGGAAACGCTGCGCGCGCC
Betula            CGCGCCAAGGAACTTTAACGAAAGA-TGCCTCCGGCCGCCTCGGAAACGGTGTGCGTGCG
Corylus           TGCGCCAAGGAACTTCAACTAAAGAGTGCCTCCGGTCGCCTCGGAAACGGCGTGCGTGCC
Carpinus          TGCGCCAAGGAACTTCAATTAAAGAGTGCCTCCGGTCGCCTCGGAAACG-TGCGCGTGTC
A.glutinosa       ------------------------------------------------------------
A.crispa          TGCGCCAAGGAACATTAACGAAAGAGTGCCTCCGGTAGCCTCGGAAACGCTGTGCTTGCC
A.nepalensisS3    ------------------------------------------------------------
A.nepalensisS2    ------------------------------------------------------------
A.nepalensisS1    ------------------------------------------------------------
```

```
A.incana        GGAGGCGAATCTTGTCTAGAACCA------------------------------------
A.matsumurae    GGAGGCGAATCTTGTCTAGAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCA
Betula          GGAGGTGAATCTTGTCTAGAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCA
Corylus         GGAGGCGAATCTTGTGCAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCA
Carpinus        GGAGGCGAATCTTGTACAAAACCATAACGACTCTCGGCAACGGATATCTCGGCTCTCGCA
A.glutinosa     ------------------------------------------------------------
A.crispa        GGAGGCGAATCTTGTCTANNNNNN------------------------------------
A.nepalensisS3  ------------------------------------------------------------
A.nepalensisS2  ------------------------------------------------------------
A.nepalensisS1  ------------------------------------------------------------


A.incana        ------------------------------------------------------------
A.matsumurae    TCGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGCGAATCAT
Betula          TCGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGCGAATCAT
Corylus         TCGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGCGAATCAT
Carpinus        TCGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGCGAATCAT
A.glutinosa     ------------------------------------------------------------
A.crispa        ------------------------------------------------------------
A.nepalensisS3  ------------------------------------------------------------
A.nepalensisS2  ------------------------------------------------------------
A.nepalensisS1  ------------------------------------------------------------


A.incana        ------------------------------------------------------------
A.matsumurae    CGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCACCTGGCCGAGGGCACGTCTGCCTGGGT
Betula          CGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCACCTGGCCGAGGGCACGTCTGCCTGGGT
Corylus         CGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGT
Carpinus        CGAGTCTTTGAACGCAAGTTGCGCCCGAAGCCATCTGGTCGAGGGCACGTCTGCCTGGGT
A.glutinosa     ------------------------------------------------------------
A.crispa        ------------------------------------------------------------
A.nepalensisS3  ------------------------------------------------------------
A.nepalensisS2  ------------------------------------------------------------
A.nepalensisS1  ------------------------------------------------------------


A.incana        -------ATCGTTGCCCCCAACCCCATC---GCCCTGCAAAGAGGCGGT-GGGGGCATGC
A.matsumurae    GTCACGCATCGTTGCCCCCAACCCCATC---GCCCTGCAAAGAGGCGGC-GGGGGCGCGC
Betula          GTCACGCATCGTTGCCCCCAACCCCATC---TCCTTGCAAAG-GGACGA-GGGGGCCTGT
Corylus         GTCACGCATCGTTGCCCCCAACCCCATCATCGCCTCTCCAAGAGACGAG-GGCGGTCTGC
Carpinus        GTCACGCATCGTCGCCCCCAACCCCATC---GCCTCTCCAAGAGACGAG-GGCAGTTTGC
A.glutinosa     -------ATCGTTGCCCCCAACCCCATC---GCCCTGCAAAGAGGCGGT-GGGGGCATGC
A.crispa        -------ATCGTTGCCCCCAACCCCATC---GCCCTGCAAAGAGGCGGTTGGGGGCATGC
A.nepalensisS3  -----------GATAGCTTAAACTCAGCG--GGCGAGTCACG--CCTGACCTGGGGTCGC
A.nepalensisS2  -------TTTGATATGCTTAAA-TCAGC---GGGTAGTCCCG--CCTGACCTGGGGTCGC
A.nepalensisS1  ---------TGATATGCTTAAACTCAGC---GGGTAGTCCCG--CCTGACCTGGGGTCGC
                                *  **   ** *             *            *   *
```

```
A.incana        GGGG-CGGACATTGGCCTCCCGTGGGCTGATGCCTGCGGCTGGCCTAAAAACGAGTCCTC
A.matsumurae    GGGGGCGGACATTGGCCTCCCGTGGGCTGATGCCTGCGGCTGGCCTAAAAACGAGTCCTC
Betula          GGGG-CAGAAATTGGCCTCCCGTGAGCTCATGCATGCGGTTGGCCTAAAAGCGAGTCCTC
Corylus         GGGG-CGGACATTGGCCTCCCGTGAGCTTTCGCTCGCGGCTGGCCTAAAAGCGAGTCCTC
Carpinus        GGGG-CGGACATTGGCCTCCCGTGAGCTTCCACTTGCGGTTGGCCTAAAAGCGAGTCCTA
A.glutinosa     GGGG-CGGACATTGGCCTCCCGTGGGCTGATGCCTGCGGTTGGCCTAAAAACGAGTCCTC
A.crispa        GGGG-CGGACATTGGCCTCCCGTGGGCTGATGCCTGCGGCTGGCCTAAAAACGAGTCCTC
A.nepalensisS3  GTTG---GAGA---GCGTCGCGAGCGC-GACGCGACAGGGTCAAAAGAGCACGCGTCGAG
A.nepalensisS2  GTTG---GAA----GCGTCGCGAGCGC-GACGCGACAGGGTCAAAAGAGCACGCGTTGAG
A.nepalensisS1  GTTG---GAA----GCGTCGCGAGCGC-GACGCGACAGGGTCAAAAGAGCACGCGTTGAG
                *  *   **     ** ** ** * **     *    ** *      *   ** **

A.incana        GGC-GACGATCGCC--ACGACAATCGGTGGTTGACAAA-CCTTCGTGACCCGTCGTGCGC
A.matsumurae    GGC-GACGATCGCC--ACGACAATCGGTGGTTGACAAA-CCCTCGTGACCCGTCGTGCGC
Betula          GGC-GACGCGCGCC--ACGACAATCGGTGGTTGTCAAA-CCCTCGTGTCCCGTCGTGCGT
Corylus         GGC-GACGAGCGCC--ACGACAATCGGTGGTTGACAAA-CCCTCGTGTCCCGTCGTGCGG
Carpinus        GGC-GACGAGCGCC--ACGACAATCGGTGGTTGCCAAAACCCTCGTGTCCCGTCGTGCGT
A.glutinosa     GGCTGACGATCGC---ACDACAATCGGTGGTTGACAAA-CCTTCGTGACCCGTCGTGCGC
A.crispa        GGCCGACGAGCGCC--ACGACAATCGGTGGTTGACAAA-CCCTCGTNNNN-NNCGTGCGT
A.nepalensisS3  CGACGATGCGCGCACGACGGGTCCCGAGGGTTTGTCAA-CCACCGATTGTCGTGGTGT--
A.nepalensisS2  CGACGATGCGCGCACGACGGGTCACGAGGGTTTGTCAA-CCACCGATTGTCGTGGCGC--
A.nepalensisS1  CGACGATGCGCGCACGACGGGTCACGAGGGTTTGTCAA-CCACCGATTGTCGTGGCGC--
                 *  ** *  ***   **      **  ****    ** **  **         * *

A.incana        GCATCGCCGCTCAACGCGTGCTCTTTTGACCCTGTC---GCGTC-GCGCTCGCGACG---
A.matsumurae    GCACCGTCGCTCAACGCGTGCTCTTTTGACCCTGTC---GCGTC-GCGCTCGCGACGC--
Betula          GCCGCGTCGCTCATCGTGTGCTCCTT-GACCCTGTT---GTGTC-GCGCTAGCGACGC--
Corylus         GC-TCGTCGCTCGTCTTGTGCTCTGT-GACCCTGTA---GCGTC-GCGCTCGCGACTC--
Carpinus        GCCTCGTTGCYCATCCTGTGCTCTGT-GACCCTATA---GCGTC-GCGATCGCGACTC--
A.glutinosa     GCATCG-CGCTCAACGCGTGCTCTTTTGACCCTGTC---GCGTY-GYGCTCGCGACGC--
A.crispa        GCATCGTCGCTCAATGTGTGCTCTTTTGACCCTGTC---GCGTC-GC--TCGCGANNN--
A.nepalensisS3  ---TCGTCGC-CGACGACTCGTTTTTTGGCCAT---------TC-CTGCTGGTCATCC--
A.nepalensisS2  ---TCGTCGC-CGAGGACTCGTTTTTAGGCCATCCGCAGGCATCAGCCCACGGGAGGCCA
A.nepalensisS1  ---TCGTCGC-CGAGGACTCGTTTTTAGGCCATCCGCAGGCATCAGCCCACGGGAGGCCA
                    **  ** *      *  *   * * ** *         *        *  *

A.incana        --TTCCAACGCGA-----------------------------------------------
A.matsumurae    --TTCCAACGCGACCCCAGGTCAGGCGGGACTACCCGCTGAGTTTAAGCATATCAATAAG
Betula          --TTCCAACGCGACCCCAGGTCAGGCGGGACTACCCG-TGAATTTAA-------------
Corylus         --TTCCAACGCGA-----------------------------------------------
Carpinus        --TTCCAATGCGA-----------------------------------------------
A.glutinosa     --TTCCAACGCGA-----------------------------------------------
A.crispa        --NNNNNNNNNNN-----------------------------------------------
A.nepalensisS3  --GTCC------------------------------------------------------
A.nepalensisS2  ATGTCCGCCCCGCGTGCCCCCACCGCCTCTTTGCATGGCGATGGGGTTGGGGGCCACGAT
A.nepalensisS1  ATGTCCGCCCCGCGTGCCCCCACCGCCTCTTTGCAGGGCGATGGGGTTGGGGGCAACGAT
```

```
A.incana          ------------------------------------------------------------
A.matsumurae      CGGAGGAAAAGAAAACA-------------------------------------------
Betula            ------------------------------------------------------------
Corylus           ------------------------------------------------------------
Carpinus          ------------------------------------------------------------
A.glutinosa       ------------------------------------------------------------
A.crispa          ------------------------------------------------------------
A.nepalensisS3    ------------------------------------------------------------
A.nepalensisS2    GCGTGACACCCATGCNNANTTGCCCCCGGGCATGTGGTTTCGGGCG--------------
A.nepalensisS1    GCGTGACACCCAGGCAGACGTGCCCTCGGCCAGGTGGCTTCGGGCGCAACTTGCGTTCAA

A.incana          ------------------------------------------------------------
A.matsumurae      ------------------------------------------------------------
Betula            ------------------------------------------------------------
Corylus           ------------------------------------------------------------
Carpinus          ------------------------------------------------------------
A.glutinosa       ------------------------------------------------------------
A.crispa          ------------------------------------------------------------
A.nepalensisS3    ------------------------------------------------------------
A.nepalensisS2    ------------------------------------------------------------
A.nepalensisS1    AGACTCGATGATTCGCGGGATTCTGCAATTCACACCAAGTATCGCATTTCGCTACGTTCT

A.incana          ------------------------------------------------------------
A.matsumurae      ------------------------------------------------------------
Betula            ------------------------------------------------------------
Corylus           ------------------------------------------------------------
Carpinus          ------------------------------------------------------------
A.glutinosa       ------------------------------------------------------------
A.crispa          ------------------------------------------------------------
A.nepalensisS3    ------------------------------------------------------------
A.nepalensisS2    ------------------------------------------------------------
A.nepalensisS1    TCATCGATGCGAGAGCCGAGATATCCGTTGCCGAGAGTCGTTATGGTTCTCGACAAGATT

A.incana          -----------------
A.matsumurae      -----------------
Betula            -----------------
Corylus           -----------------
Carpinus          -----------------
A.glutinosa       -----------------
A.crispa          -----------------
A.nepalensisS3    -----------------
A.nepalensisS2    -----------------
A.nepalensisS1    CGCCTCCGGTGCGCGCC
```

Gilbert, 1999, site http://www.es.embnet.org/Doc/phylodendron/clustal-form.html). The most consensus tree was obtained after considering 1000 bootstrap replicates.

4.7.1 Phylogenetic analysis using Parsimony method:

Parsimony analysis was done using the aligned sequences of the 18S-28S *rrn* region. The sequence information generated using each of the primers was utilized in isolation for phylogenetic analysis. The most parsimonious tree derived from the sequences of the distal region of the 18S rRNA gene and a part of ITS I has been presented in Fig.4. 27. The tree constructed using sequences of the region between the distal part of ITS I and the proximal part of the 28S rRNA gene is depicted in Fig. 4.28 Both trees were more or less similar except for a few minor differences. *Alnus matsumurae* appeared closest to the non-nitrogen fixing genus *Betula*. The three samples of *Alnus nepalensis* clustered together and the European alder *Alnus glutinosa* appeared closest to *Alnus nepalensis*. Varghese (2000) had reported a similar observation for the microsymbiont *Frankia*. While assessing the diversity of the *Frankia* populations of different parts of India and comparing it with a sample of Tuebingen, Germany, she found that one of the samples collected from the Indian alder (*Alnus nepalensis*) was closest to a sample collected from *Alnus glutinosa* tree growing at Tuebingen, Germany. The diversity that has been observed in the natural populations of *Alnus nepalensis*, from PCR-AFLP/RFLP studies, was also visible in the parsimony trees constructed and sample 3 seemed to have diversified earlier than sample 1 and 2.

4.7.2 Phylogenetic analysis using Neighbour joining method:

The neighbour joining method of the program PHYLODENDRON was utilized to assess the lineages. For this analysis also, sequences of the distal part of 18S rRNA gene and a part of ITS I and the region between the distal part of ITS I and the proximal part of the 28S rRNA gene were used separately. Bootstrap analysis was used to assess the robustness of the phylogenetic tree. The most consensus tree in either case was selected

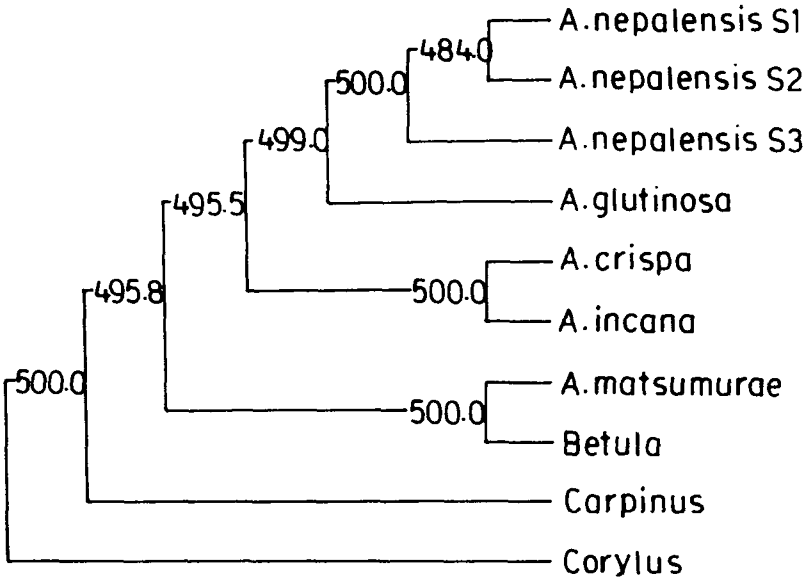


Fig.4.27. **DNA parsimony strict consensus phylogenetic tree derived from aligned sequences of the distal part of the 18S rRNA gene and part of ITS I. The numbers at each node represent the bootstrap values out of 500 replicates.**

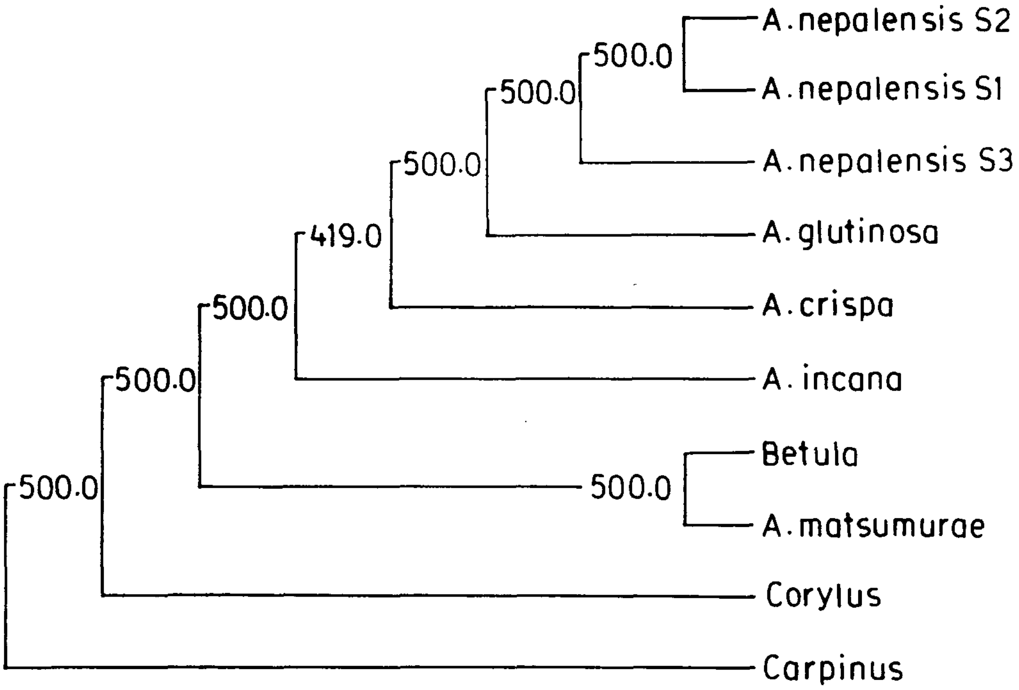


Fig.4.28. DNA parsimony strict consensus phylogenetic tree derived from aligned sequences of the region between the distal part of ITS I and initial part of 28S rRNA gene. The numbers at each node represent the bootstrap values out of 500 replicates.

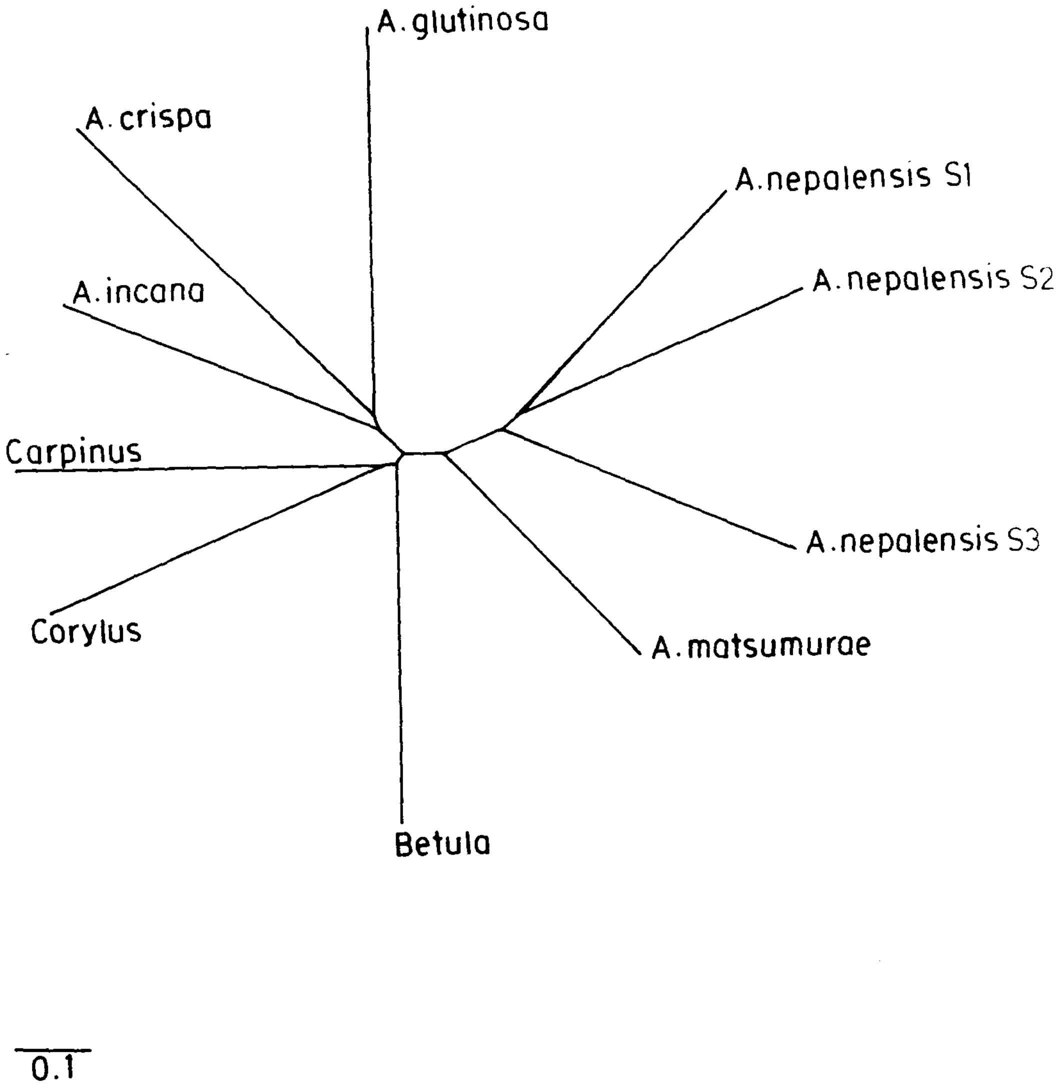


Fig.4.29 Neighbour joining phylogenetic tree constructed using aligned sequences of the distal part of the 18S rRNA gene and part of ITS I. The tree has been selected after considering 1000 bootstrap replicates. 0.1 represents the substitution rate per 100bp.

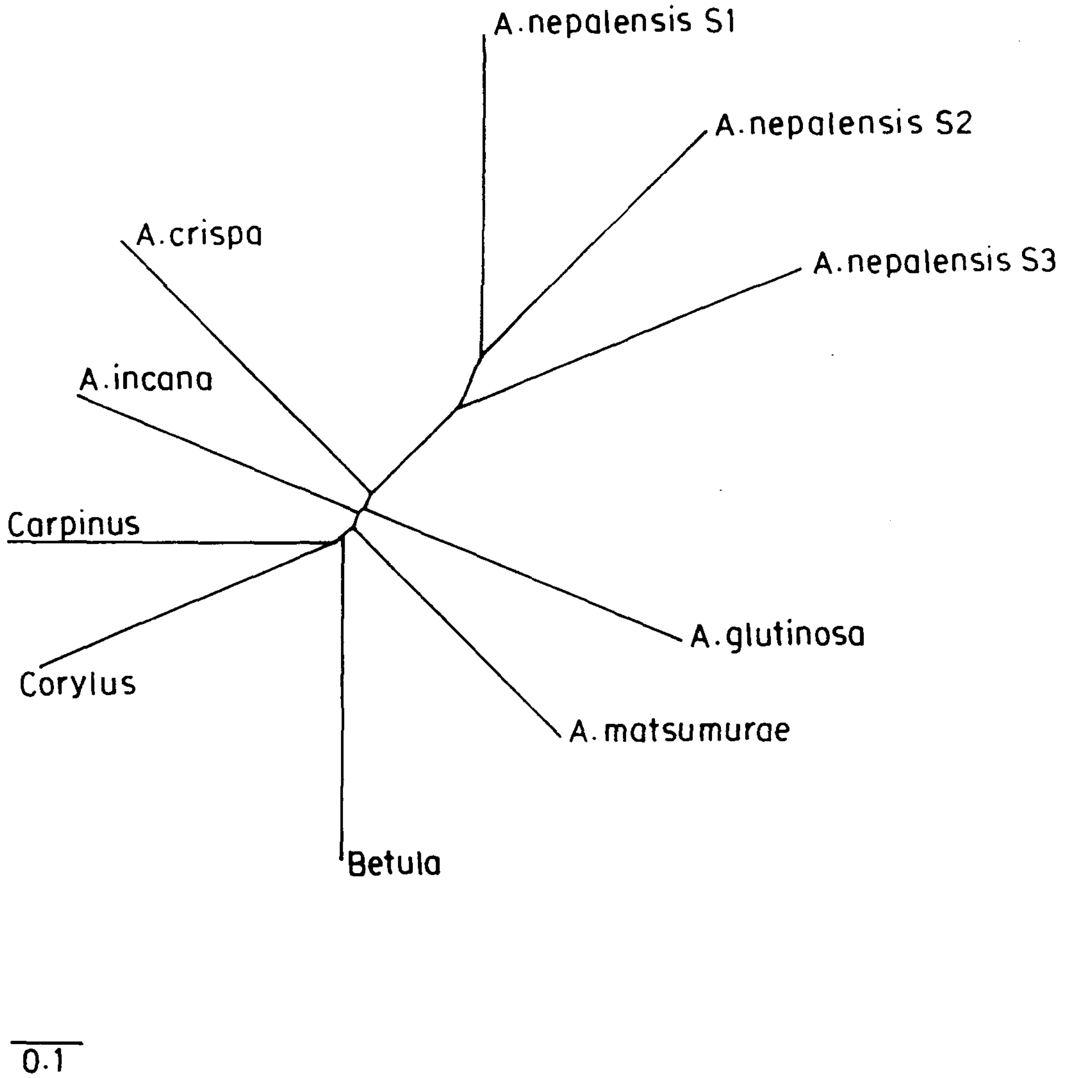


Fig.4.30 Neighbour joining phylogenetic tree constructed using aligned sequences of the region between the distal part of ITS I and initial part of 28S rRNA gene. The tree has been selected after considering 1000 bootstrap replicates. 0.1 represents the substitution rate per 100bp.

after considering 1000 replicates of bootstrap analysis. The tree constructed using the aligned sequences of the distal part 18S rRNA gene and a part of ITS I showed *Betula* at its base. The first bifurcation gave rise to *Corylus* and *Carpinus*. Following this a second bifurcation was seen (Fig. 4.29), which gave rise to two branches. One of the branches further divided and gave rise to the alders found in the Eastern region while the other branch divided and gave rise to the alders found in the Western region. This tree showed a distinct demarcation of the alders based on their geographical distribution.

The topology of the tree constructed using the aligned sequences of the region between the distal part of ITS I and the proximal part of the 28S rRNA gene did not show distinct demarcation of the alders based on geographical distribution. This tree too showed *Betula* at its base and then the first bifurcation gave rise to *Corylus* and *Carpinus*. Apparently, during the course of evolution the main branch divided at regular intervals diversifying into the different species of *Alnus*.. *Alnus matsumurae* appeared closest to the non-actinorhizal genus *Betula*. Both the neighbour joining trees clearly showed the diversification of sample 3 of *Alnus nepalensis* prior to that of sample 1 and 2. However, all the three samples of *Alnus nepalensis* clustered together (Fig. 4.30).

4.8 DESIGNING OF ALDER SPECIFIC PROBE:

4.8.1 Designing and testing the genus specific probe:

Alignment of the *Alnus nepalensis* 18S-28S ITS sequences with sequences of other species of alder, *Corylus*, *Carpinus* and *Betula* (Section 4.6) showed areas where all the samples analyzed were similar as well as areas where the alder sequences were similar while the non-alder genera showed differences. One such site was found at 94-110bp downstream of the *Alnus nepalensis* sequence in the distal region of the 18S rRNA gene. This region has been highlighted in Fig. 4.25 and it was here that the primer designed in the present study (primer 24) was located. The differences were mainly in the 3' region of

Corylus, Carpinus and *Betula*. Therefore, the chances of the primer annealing and amplifying these templates were feeble, especially if the stringency conditions were high. The primer was found possibly located partly in the 18S rRNA gene and partly in the ITS region. This primer in conjunction with primer ITSC26A was utilized for amplifying the two internally transcribed spacers and the 5.8S rRNA gene lying between the 18S and 28S rRNA genes of the nuclear DNA.

4.8.2 Standardization of conditions to eliminate chances of amplification of exogenous fungal DNA:

Amplification of the 18S-28S rRNA region, using the primer pair 24/ ITSC26A yielded the expected band of approximately seven hundred base pairs at an annealing temperature of 45°C. In addition to this band, other smaller bands were also present. One of the probable reasons for the occurrence of the additional bands was the presence of alternate annealing sites for the primer. The other reason could have been the presence of exogenous DNA of fungal origin. This was quite likely since several fungal hyphae and spores are normally present on leaf surfaces. It was for this reason, very young leaves were taken for isolation of DNA. Moreover, care was taken to thoroughly surface sterilize the leaves with 30% H_2O_2 prior to isolation of DNA. However, to confirm that the contaminating fungal DNA did not get amplified an experiment was conducted.

Fungal DNA was isolated from a mixed culture of fungi. DNA was isolated by rapid sonication of the fungal culture in an eppendorf tube. This was followed by chloroform:isoamyl alcohol (24:1) extraction and ethanol precipitation. To confirm the presence of DNA 2μl of the DNA was subjected to electrophoresis on a 0.8% agarose gel (Fig. 4.31).

This DNA and alder DNA were then amplified using the primers 24/ ITSC26A under identical conditions at 45°C annealing temperature. Electrophoresis of the amplicons revealed several bands for alder DNA. The fungal DNA also showed multiple bands but these were fainter than those of alder (Fig. 4.32). Therefore, the annealing temperature was increased to 50°C for both the DNAs. At this temperature too alder showed multiple bands but fungal DNA yielded a single faint band (Fig. 4.33). Obviously the additional bands were not of fungal origin. However, to eliminate the multiple bands and also to completely discount the possibility of fungal DNA amplification, the annealing temperature was increased to 60°C. At this temperature neither DNA got amplified. So the annealing temperature was decreased to 55°C. At 55°C fungal DNA did not amplify but alder DNA also showed very faint single band (Fig. 4.34). The annealing temperature was further decreased to 52°C. At this temperature a single strong band of approximately 700bp was produced for alder and the fungal DNA did not amplify (Fig. 4.35). We therefore, amplified all other samples of *Alnus nepalensis* at this annealing temperature (52°C). Fig. 4.36-4.40 show the amplification of alder DNA at different annealing temperatures (45°C, 50°C. 55°C, 52°C).

4.8.3 Test to confirm genus specific nature of the primer:

Sequence analysis showed that the primer designed would be alder specific. To confirm our contention we decided to test amplification of the DNA of other actinorhizal trees and some members of the alder family, Betulaceae, using these primers. Only two other actinorhizal trees *Elaeagnus* and *Myrica* and two members of the family Betulaceae, *Betula* and *Carpinus* were tested. Total genomic DNA was isolated from the leaves of all the genera tested (Fig. 4.41-4.43). All the DNAs were amplified under identical conditions along with alder DNA. To be sure of the results a negative control was always kept. After electrophoresis of the amplicons on a 0.8% agarose gel, the gel was stained with ethidium bromide and photographed using the Gel Doc 1000 (BioRad). It was observed that alder samples got amplified but *Elaeagnus*, *Myrica*, *Betula* and *Carpinus* did not show any amplification (Fig. 4.44 and 4.45). Thus, on the basis of this experiment we can say that

Fig 4.31-4.35

STANDARDIZATION OF AMPLIFICATION CONDITIONS FOR THE 18S-28S ITS REGION, USING THE PRIMER PAIR 24/ITSC26A, IN ORDER TO ELIMINATE CHANCES OF AMPLIFICATION EXOGENOUS FUNGAL DNA:

Fig.4.31: Isolation of Fungal DNA.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-7: Fungal DNA.

Fig.4.32: Amplification of alder and fungal DNA at an annealing temperature of 45°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-3: Negative amplification of alder DNA.
Lane 4: Positive amplification of alder DNA.
Lane 5: Positive amplification of fungal DNA.
Lane 6: Negative control.

Fig.4.33: Amplification of alder and fungal DNA at an annealing temperature of 50°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1 & 2: Negative amplification of alder DNA.
Lanes 3-5: Positive amplification of alder DNA.
Lane 6: Positive amplification of fungal DNA.
Lane 7: Negative control.

Fig.4.34: Amplification of alder and fungal DNA at an annealing temperature of 55°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-4: Negative amplification of alder DNA.
Lane 5: Positive amplification of alder DNA.
Lane 6: Negative amplification of fungal DNA.
Lane 7: Negative control.

Fig.4.35: Amplification of alder and fungal DNA at an annealing temperature of 52°C.

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest).
Lanes 1-4: Negative amplification of alder DNA.
Lane 5: Positive amplification of alder DNA.
Lane 6: Negative amplification of fungal DNA.

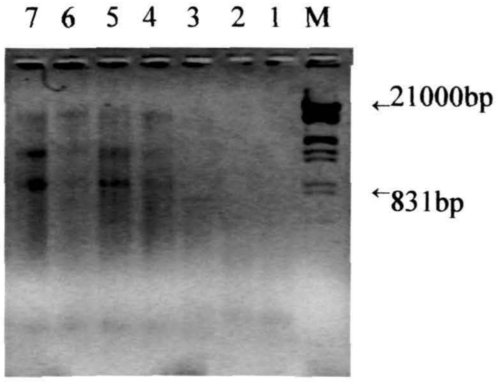


Fig.4.31

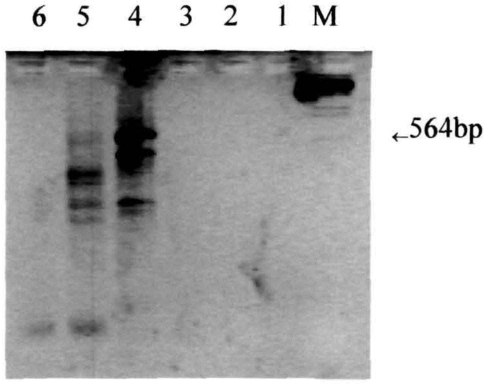


Fig.4.32

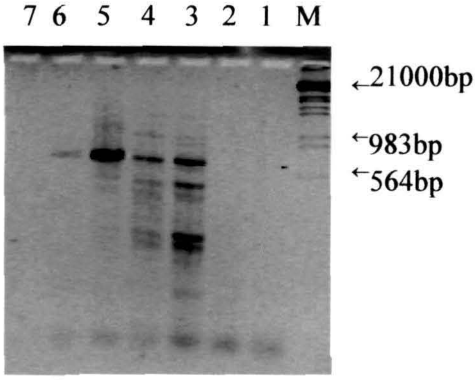


Fig.4.33

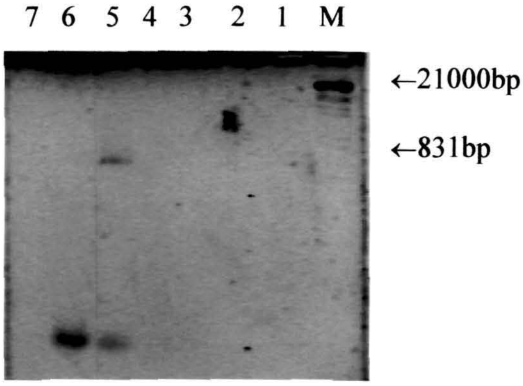


Fig.4.34

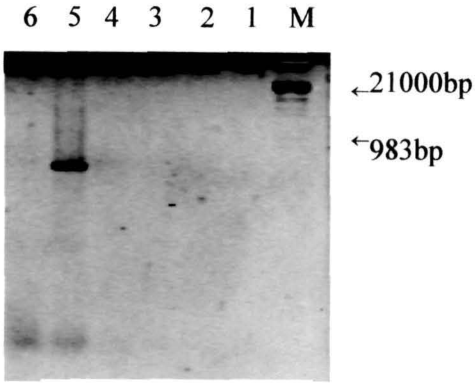


Fig.4.35

Fig 4.36-4.40

AMPLIFICATION OF THE 18S-28S *rrn* OPERON AT DIFFERENT ANNEALING TEMPERATURES USING THE PRIMERS 24/ITSC26A:

Fig.4.36: Amplification of alder DNA from different trees at 45°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-6: Positive amplification of different samples.
Lane 7: Negative control.

Fig.4.37: Amplification of alder DNA from different trees at 50°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-6: Positive amplification of different samples.
Lane 7: Negative control.

Fig.4.38: Amplification of alder DNA from different trees at 55°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1: Positive amplification.
Lane 2-6: Negative amplification.
Lane 7: Negative control.

Fig.4.39: Amplification of alder DNA from different trees at 52°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1-7: Positive amplification.

Fig.4.40: Amplification of alder DNA from different trees at 52°C.

Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1, & 3-7: Positive amplification.
Lane 2: Negative amplification.

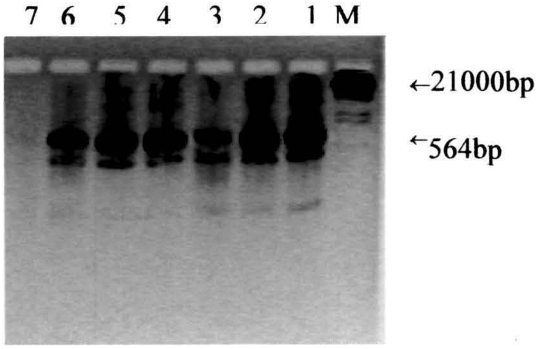


Fig.4.36

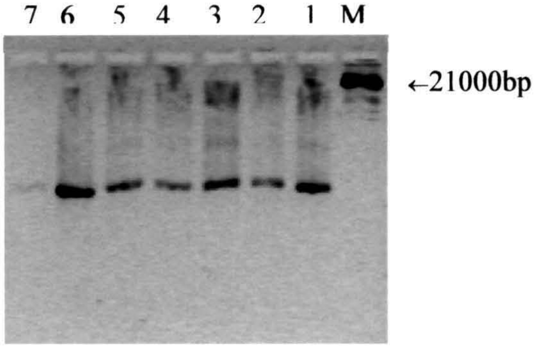


Fig.4.37

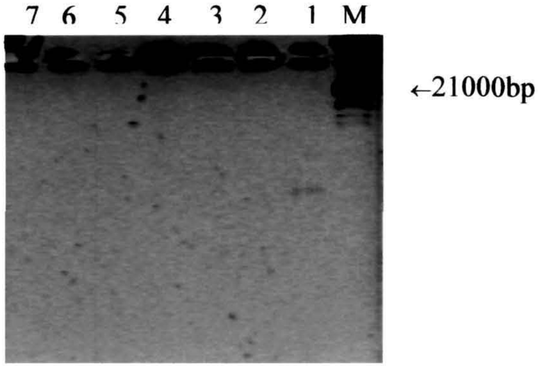


Fig.4.38

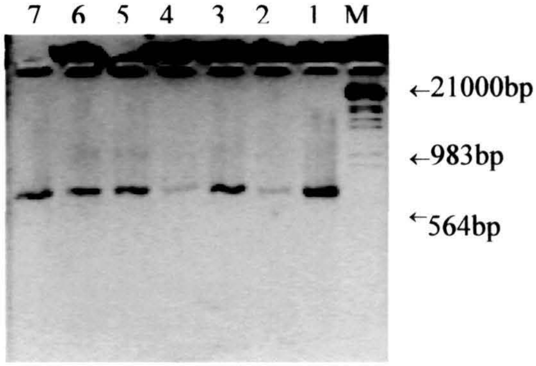


Fig.4.39

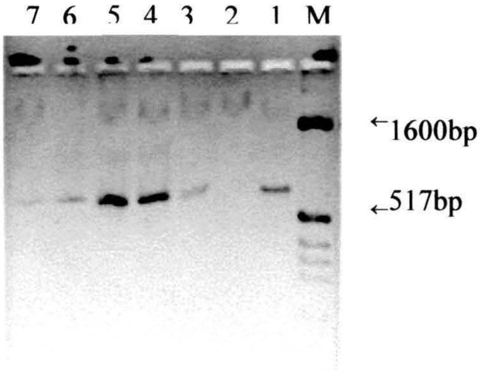


Fig.4.40

Fig 4.41-4.45

CONFIRMATION OF THE GENUS SPECIFIC NATURE OF PRIMER 24:

Fig.4.41: Isolation of total genomic DNA from the leaves *Alnus nepalensis* trees.
Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-5: Total genomic DNA of *Alnus nepalensis*.

Fig.4.42: Isolation of total genomic DNA from the leaves *Elaeagnus* and *Carpinus*.
Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1: Total genomic DNA of *Elaeagnus*.
Lane 2: Total genomic DNA of *Carpinus*.

Fig.4.43: Isolation of total genomic DNA from the leaves *Betula* and *Myrica*.
Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1: Total genomic DNA of *Betula*.
Lane 2: Total genomic DNA of *Myrica*.

AMPLIFICATION OF *ALNUS NEPALENSIS*, *ELAEAGNUS*, *CARPINUS*, *BETULA* AND *MYRICA* DNA UNDER IDENTICAL CONDITIONS, USING THE PRIMER PAIR 24/ITSC26A.

.**Fig.4.44: Amplification of *Alnus nepalensis* and *Carpinus* DNA at 52°C.**
Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lanes 1-4 & 6: Positive amplification of *Alnus nepalensis*.
Lanes 5 & 7: Negative amplification of *Alnus nepalensis*.
Lanes 8 & 9: Negative amplification of *Carpinus*.
Lane 10: Negative control.

Fig.4.45: Amplification of *Alnus nepalensis*, *Betula*, *Myrica* and *Elaeagnus* DNA at 52°C.
Lane M: Molecular weight marker (Lambda DNA/*Eco*RI + *Hind*III Double Digest).
Lane 1-6: Positive amplification of *Alnus nepalensis*.
Lanes 7, 8 & 9: Negative amplification of *Betula*, *Myrica* and *Elaeagnus* respectively.
Lane 10: Negative control

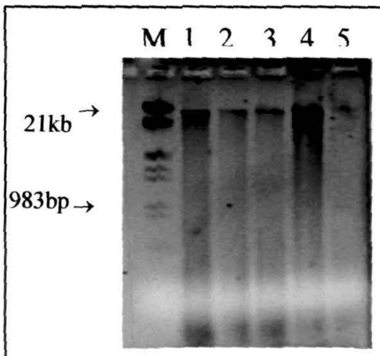


Fig.4.41

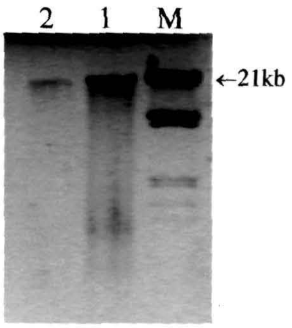


Fig.4.42

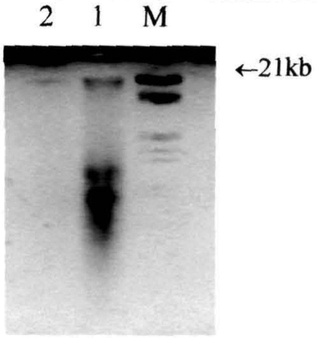


Fig.4.43

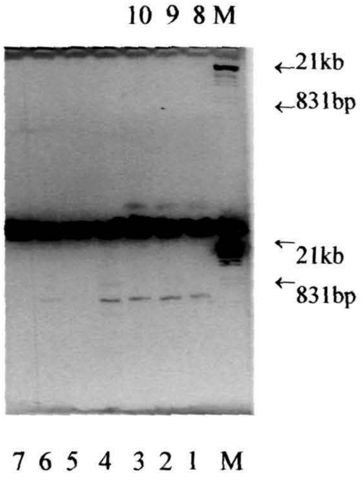


Fig.4.44

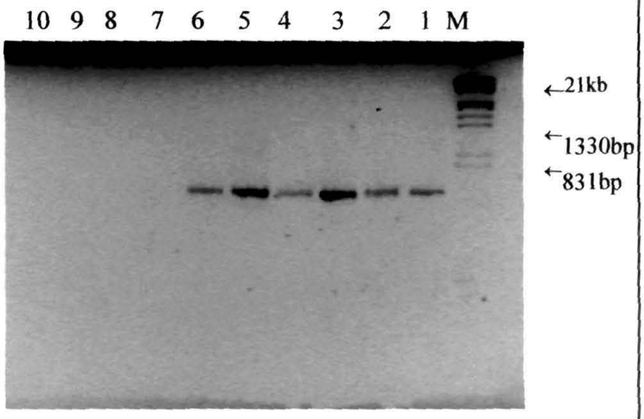


Fig.4.45

primer 24 is *Alnus* specific. However, based on sequence analysis it is likely that it was a genus specific primer. Other species of *Alnus* remain to be practically tested before declaring this primer as truly genus specific. We could not test other *Alnus* species.

4.9 PCR-RFLP STUDIES:

PCR based techniques have proved to be valuable for the molecular characterization of innumerable organisms. One such tool is PCR-RFLP, where PCR products are digested with a combination of restriction endonucleases and polymorphism detected due to differences in restriction endonuclease sites. This is a highly versatile and indispensable tool. Its utility lies in its ability to detect variability, at times allowing detection of very minor changes in the nucleotide sequence. Further, compared to Restriction Fragment Length Polymorphism of total DNA, it is quicker and easier to interpret.

4.9.1 Selection of Restriction Endonucleases:

To best utilize the RFLP technique, selection of the appropriate restriction enzyme is needed. For this, prior knowledge of the DNA sequence to be analyzed is helpful. In the present study, published sequences of alder were subjected to restriction digestion analysis with a number of restriction endonucleases using the MacVector software®. This gave the locations of the restriction sites and the expected fragment size for the different enzymes. The enzymes that would allow the detection of maximum variability by way of possessing maximum number of restriction sites on the template were selected. Table 3.2 shows a list of the restriction endonucleases used for restriction digestion of the different regions.

4.9.2 PCR-RFLP analysis of the proximal part of the 16S rRNA gene:

Restriction digestion of amplicons of the proximal part of the 16S rRNA gene was done using the four base cutter restriction endonuclease *Msp*1. Restriction digestion was carried out as detailed in section 3.15. Electrophoresis was done on a 3% agarose gel at 40 volts for eight hours. Fig 4.46 shows the differences in the restriction fragment profiles. Most of the samples gave rise to two bands of approximately 340bp and 75bp. Some samples in addition to the 340bp and 75bp bands yielded four additional bands of approximately 250bp, 170bp, 130bp and 100bp (Table 4.19). A few samples did not get digested. Electrophoresis of digested DNA revealed that the samples differed in susceptibility towards a given enzyme. This difference was an outcome of variability in the DNA base composition of different trees. Hence, the profiles generated by restriction digestion of this region, could categorize the alder trees into three groups. Those having profile Ms(I)1 and Ms(I)2 and the third group comprising samples which did not show any digestion. These findings have reinforced previous reports stating the initial part of the gene to be more variable in nature (Harry *et al.*, 1991).

4.9.3 PCR-RFLP analysis of the middle part of the 16S rRNA gene:

For restriction analysis of the middle part of the 16S rRNA gene three restriction endonucleases, *Alu*I, *Mbo*I and *Msp*I, were used. When the two bands produced as a result of amplification of the middle part of the 16S rRNA gene of *Alnus nepalensis* were restricted with the enzyme *Alu*I, they did not get digested. Restriction digestion with *Mbo*I yielded two bands of approximately 160bp and 200bp. The same samples, when digested with *Msp*I gave rise to three bands of approximately 135bp, 230bp and 340bp. The restriction patterns obtained for this region using the two restriction enzymes *Mbo*I and *Msp*I have been listed in table 4.20. However, restriction patterns obtained for different trees were similar (Fig. 4.47) and so could not be used for intraspecific differentiation of *Alnus nepalensis*. AFLP analysis of this region had revealed similar results. Hence, it can be said that the middle part of the 16S rRNA gene is a more conserved region. Previous reports also support this view with a few exceptions. Verghese (1999) had detected variability in this region for *Frankia*.

Fig 4.46-4.47

RFLP ANALYSIS OF THE CHLOROPLAST 16S rRNA GENE:

Fig.4.46: RFL profiles obtained for the proximal part of the 16S rRNA gene using the restriction endonuclease *Msp*1.

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest).
Lanes 1, 3 & 5: Undigested DNA.
Lanes 2 & 6: Digested DNA representing profile Ms(I)1.
Lane 4: Digested DNA representing profile Ms(I)2

Fig.4.47: RFL profiles obtained for the middle part of the 16S rRNA gene using the restriction endonuclease's Mbo1 and Msp1.

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest).
Lanes 1, 4 & 7: Undigested DNA.
Lanes 2, 5 & 8: DNA digested with the enzyme *Mbo*1.
Lanes 3, 6 & 9: DNA digested with the enzyme *Msp*1.

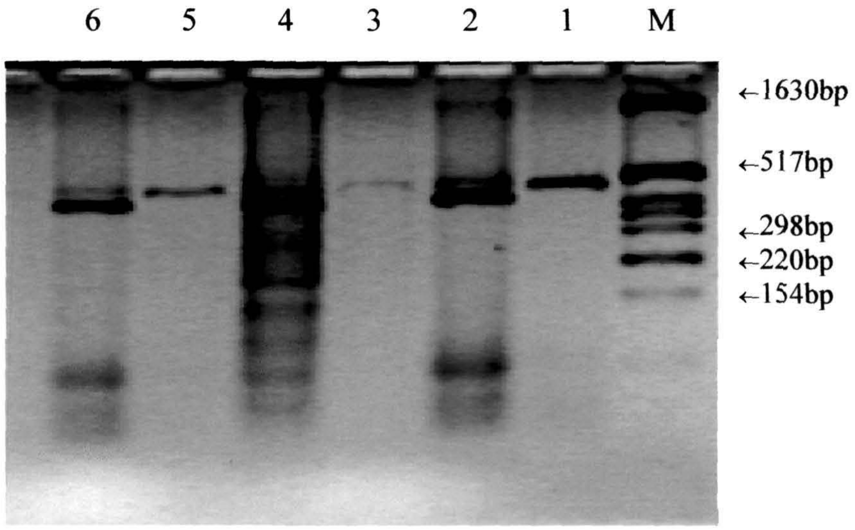


Fig.4.46

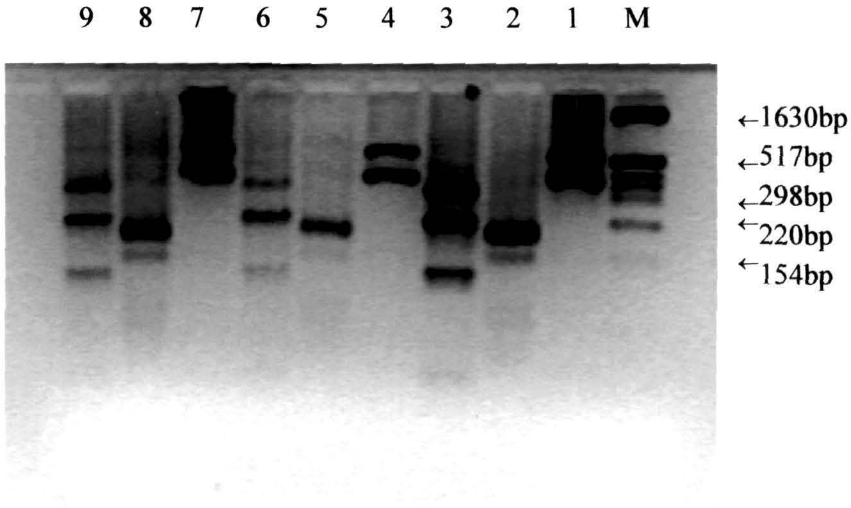


Fig.4.47

TABLE 4.19 COMPARISON OF THE BANDING PATTERNS OF THE THREE RESTRICTION FRAGMENT LENGTH PROFILES OBTAIND FOR THE INITIAL PART OF THE 16S rRNA GENE USING THE RESTRICTION ENDONUCLEASE *Msp* 1.

PROFILE	APPROXIMATE BAND SIZE (BASE PAIRS)					
	340	250	170	130	100	75
Ms(I) 1	+	-	-	-	-	+
Ms(I) 2	+	+	+	+	+	+
Ms(I) 3	-	-	-	-	-	-

TABLE 4.20 RESTRICTION DIGESTION PROFILES OBTAINED FOR THE MIDDLE PART OF THE 16S rRNA GENE USING THE RESTRICTION ENZYMES *Msp* 1 & *Mbo* 1.

BAND SIZE (BASE PAIRS)	ENZYMES	
	Mbo 1	*Msp* 1
340	-	+
230	-	+
200	+	-
160	+	-
135	-	+

+ = Band presence
- = Band absence

Approximate band sizes were found out using the Multi-Analyst software®.

4.9.4 PCR/RFLP analysis of the 18S-28S *rrn* region:

The non-coding regions normally harbor greater genetic variability since they are not subjected to selection pressure as much as the coding regions are. The internally transcribed spacer between the 18S-28S rRNA gene is one such region where chances of detecting diversity are high. When this region was amplified using the primers ITS1-PLANT / ITSC26A, a lot of variability was observed within the *Alnus nepalensis* population. However, when this region was amplified using primer 24 instead of primer ITS1-PLANT, in conjunction with primer ITSC26A at 52°C, all the samples showed the same amplification profile. Minor differences in the band size may have been there but they were not detectable by agarose gel electrophoresis. To investigate this region more precisely and to search for differences, if any, we carried out RFLP analysis of the amplicons. The enzyme used for restriction digestion of the amplified products was *Scr*F1. Restriction digestion of the amplicons was performed as described earlier. The restriction patterns of the different samples were resolved by agarose gel electrophoresis on 3% gels at 40 volts for 8 hours. The gels were photographed and analyzed using the Multi-Analyst software® (Fig. 4.48-4.51). Results obtained permitted the separation of the alder trees into nine different groups. Table 4.21 shows the banding patterns of these profiles. To predict the fragments that would be generated on digestion with *Scr*F1 a computer simulated analysis of the 18S-28S ITS sequence was done for *Alnus matsumurae* and *Alnus nepalensis*. On the basis of this analysis a restriction map was prepared (Fig. 4.52) which, predicted the generation of six bands of approximately 20bp, 60bp, 110bp, 166bp, 220bp and 245bp. Most of the bands represented in our profiles were in accordance with the predictions. However, all the bands generated were not exactly as expected. The basic bands were present most often, yet, additional bands were also seen. This can be explained on the basis of the restriction map. If one restriction site was absent it would give a fragment that was bigger in size. For example, absence of site 'D' represented in the restriction map (Fig. 4.52) would generate a band of 265bp. If two cleavages, 'C' and 'D', do not take place

Fig 4.48-4.51

RFLP ANALYSIS OF THE 18S-28S ITS REGION:

Fig.4.48-4.51: RFL profiles obtained for the 18S-28S ITS region when amplicons of primers 24/ITSC26A were digested using the restriction endonuclease *Scr*F1.

Fig.4.48

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest).
Lanes 1, 3, 7 & 9: Undigested DNA.
Lane 2: Digested DNA representing profile Sc4.
Lanes 4 & 8: Digested DNA representing profile Sc3.
Lane 8: Digested DNA representing profile Sc8.

Fig.4.49

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest).
Lanes 1, 3, 5, 7, 9& 11: Undigested DNA.
Lanes 2 & 4: Digested DNA representing profile Sc6.
Lanes 6, 8, 10 & 12: Digested DNA representing profile Sc2.

Fig.4.50

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest).
Lanes 1 & 3: Undigested DNA.
Lane 2: Digested DNA representing profile Sc5.
Lane 4: Digested DNA representing profile Sc6.

Fig.4.51

Lane M: Molecular weight marker (pBR-322 DNA/*Hinf*I Digest).
Lanes 1, 3, & 5: Undigested DNA.
Lanes 2 & 4: Digested DNA representing profile Sc1.
Lane 6: Digested DNA representing profile Sc7.

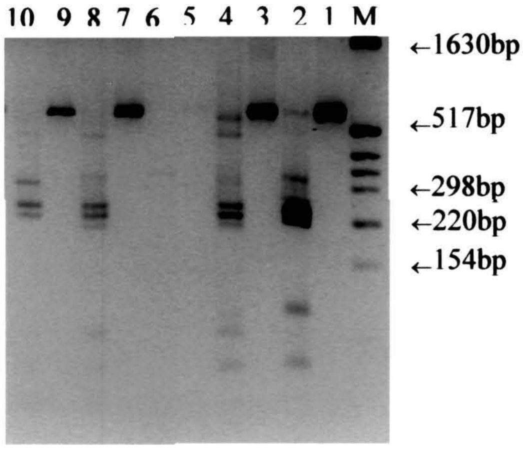


Fig.4.48

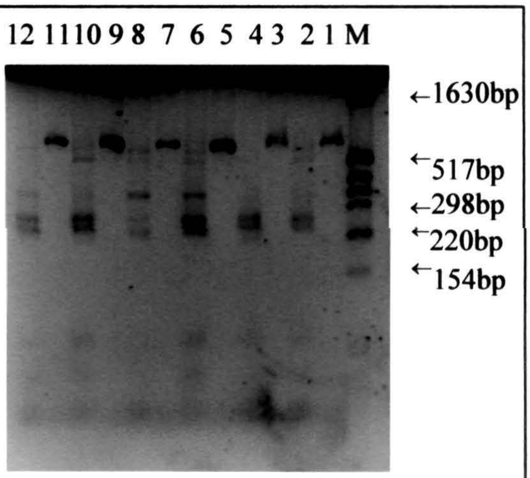


Fig.4.49

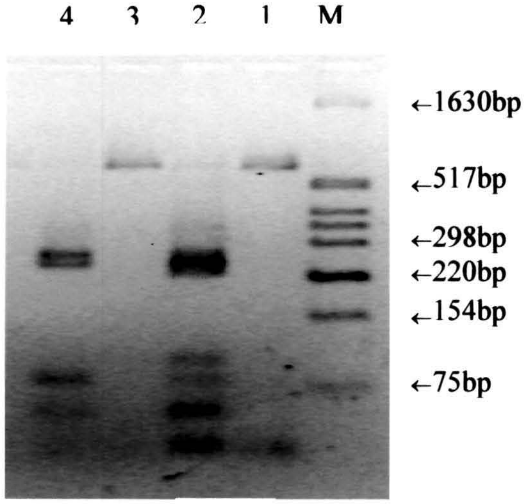


Fig.4.50

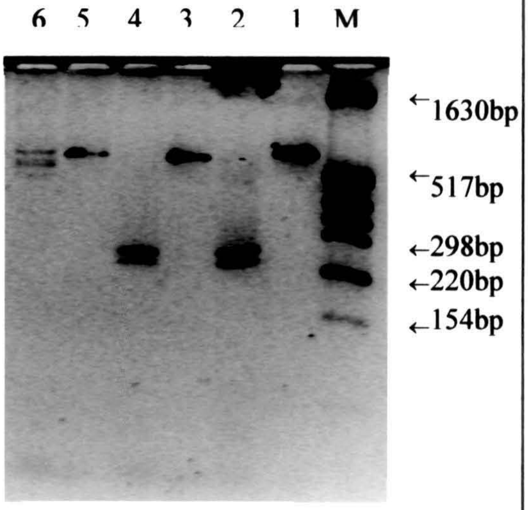


Fig.4.51

TABLE 4.21 COMPARISON OF THE BANDING PATTERNS OF THE RESTRICTION FRAGMENT LENGTH PROFILES OBTAIND FOR THE 18S-28S ITS REGION USING THE RESTRICTION ENDONUCLEASE *Scr* F 1.

PROFILE	APPROXIMATE BAND SIZE (BASE PAIRS)									
	630	520	330	265	245	180	130	110	60	20
Sc 1	-	-	+	+	-	-	-	-	-	-
Sc 2	-	-	+	+	+	-	-	+	+	+
Sc 3	-	-	+	+	+	+	-	+	+	+
Sc 4	-	-	+	+	+	+	-	+	-	+
Sc 5	-	-	-	+	+	+	+	+	+	-
Sc 6	-	-	-	+	+	-	-	+	+	-
Sc 7	+	+	-	-	-	-	-	-	-	-
Sc 8	-	-	+	+	+	-	-	-	-	-
Sc 9	-	-	-	-	-	-	-	-	-	-

+ = Band presence

- = Band absence

Approximate band sizes were found out using the Multi-Analyst software®.

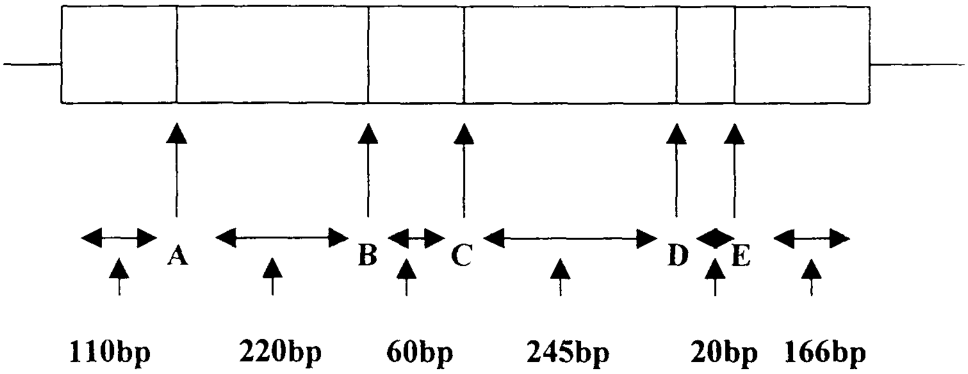


Fig. 4.52 RESTRICTION MAP OF THE 18S-28S ITS REGION HIGLIGHTING THE RESTRICTION SITES FOR THE ENZYME *Scr*F1.

then the 330bp band would be generated. Similarly the occurrence of other bigger bands of 520bp and 630bp, can be explained. However, if the 520bp and 630bp bands were present then other smaller bands should not be present as the bigger bands arise because of the absence of restriction sites. This has been observed in one of our profiles (Sc7). In some of the samples it was observed that in addition to the smaller bands the bigger bands were also present. This anomaly could be a consequence of either partial digestion or the presence of more than one operons. The restriction sites in these operons may differ. A particular restriction site may be present in one of the operons but absent in the other. However, this is only a hypothesis and remains to be proved. We have assumed that the bigger bands were present due to partial digestion. Therefore, we have ignored these feeble bands.

4.10 RELATIONSHIP BETWEEN THE PCR-AFL/RFL PROFILES AND THE NITROGENASE ACTIVITY:

The acetylene reduction assay values representing the nitrogenase activity showed a wide range. To make comparison easier the values were grouped into classes with each class representing a range of ARA values (0-5, 5-10, 10-15, 15-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80 and 80-90). The frequency distribution of each class was found out and has been graphically represented in Fig. 4.53. Most trees fell in the range of 0-5 followed by those belonging to the 5-10. 74% trees had nitrogenase activity below 15 n Moles ethylene produced/g fresh wt./hour. Few trees showed very high nitrogenase activity. Trees having an ARA value exceeding 15 n Mol ethylene produced/g fresh wt./hr were considered as high nitrogen fixers and those below this were considered as low nitrogen fixers.

The PCR-AFL/RFL profiles were thoroughly scanned and efforts were made to co-relate them with the ARA values. Tables 4.22 & 4.23 show the different PCR-AFL/RFL profiles for each tree. Most of the profiles were common to both high and low nitrogen fixing groups. However, it was observed that AFL profile 3 for the initial part of the 16S

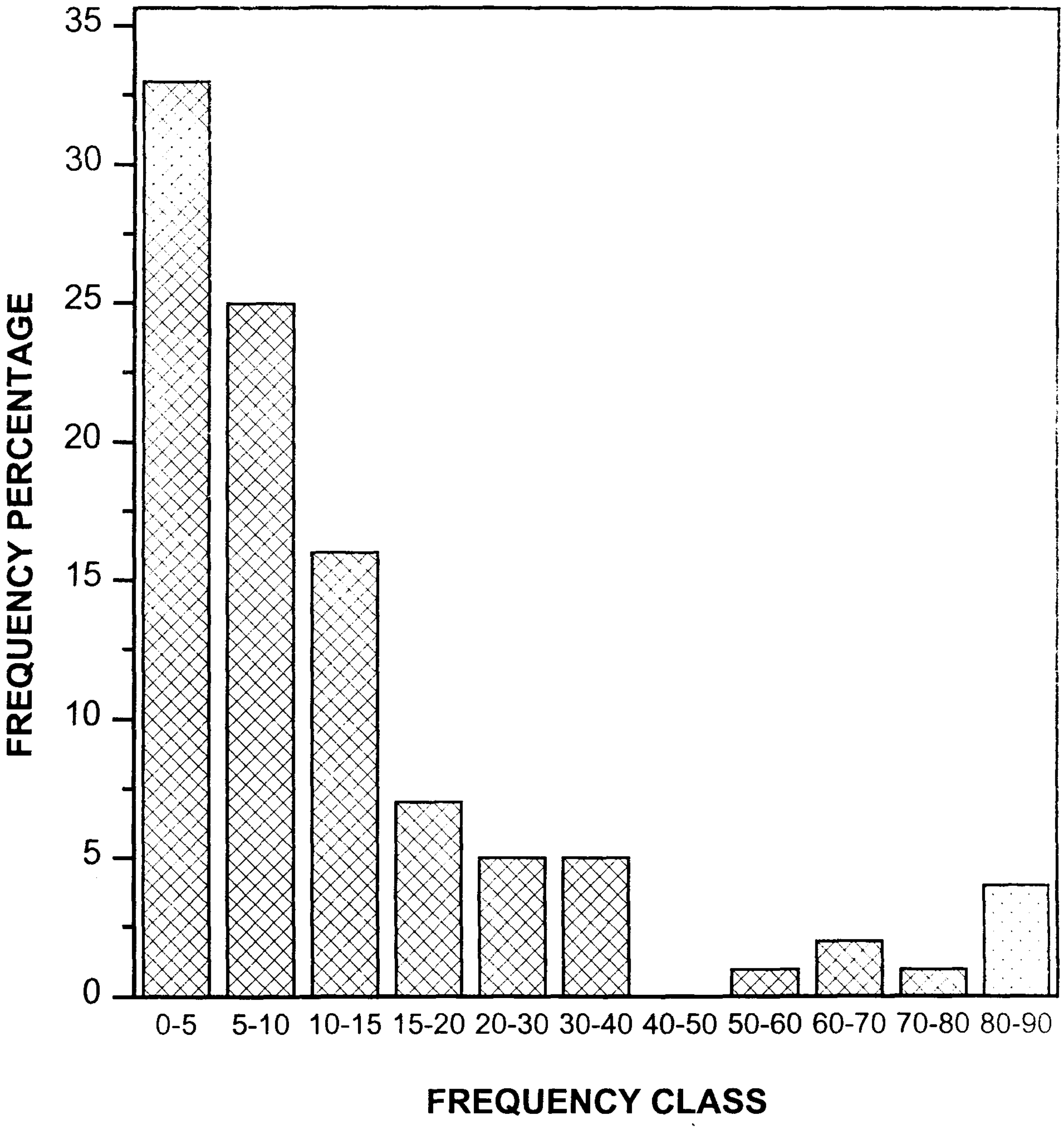


Fig.4.53 GRAPHICAL REPRESENTATION OF THE NUMBER OF TREES BELONGING TO EACH FREQUENCY CLASS

TABLE 4.22 PCR-AFLP/RFLP PROFILES FOR REGIONS SHOWING VARIATIONS

	PROFILES			
SAMPLE No.	AFLP (8/16)	RFLP (8/16)	AFLP (22/23)	RFLP (24/23)
P.C:1	1	Ms(I) 1	B	Sc6
P.C:2	1	Ms(I) 1	-	Sc3
P.C:3	2	Ms(I) 1	-	Sc2
P.C:4	2	Ms(I) 1	A	Sc6
P.C:5	2	N.D	-	Sc6
P.C:6	2	N.D	A	Sc6
P.C:7	1	Ms(I) 1	A	Sc3
P.C:8	3	Ms(I) 2	-	Sc7
P.C:9	1	Ms(I) 1	E	N.D
P.C:10	1	Ms(I) 1	B	N.D
P.C:11	1	Ms(I) 1	B	Sc6
P.C:12	2	Ms(I) 1	B	Sc6
P.C:13	1	N.D	B	Sc6
P.C:14	1	N.D	B	Sc3
P.C:15	1	Ms(I) 1	B	Sc2
P.C:16	2	Ms(I) 2	B	Sc2
P.C:17	1	Ms(I) 1	D	Sc6
P.C:18	1	Ms(I) 2	C	Sc2
P.C:19	1	Ms(I) 1	C	N.A
P.C:20	1	Ms(I) 1	-	Sc2
P.C:21	2	N.A	B	Sc6
P.C:22	1	Ms(I) 1	A	N.D
P.C:23	1	Ms(I) 1	H	Sc3
P.C:24	2	Ms(I) 1	D	Sc6
P.C:25	2	Ms(I) 2	D	Sc6
P.C:26	2	Ms(I) 1	E	N.D
P.C:27	1	Ms(I) 2	-	Sc6
P.C:28	2	Ms(I) 2	A	Sc6
P.C:29	1	Ms(I) 2	-	Sc7
P.C:30	1	Ms(I) 1	A	Sc6
P.C:31	2	Ms(I) 2	-	Sc5
P.C:32	2	Ms(I) 2	A	Sc6
P.C:33	2	Ms(I) 2	A	Sc2
P.C:34	1	Ms(I) 2	A	Sc2
P.C:35	1	Ms(I) 1	-	Sc2
P.C:36	1	Ms(I) 1	-	Sc6
P.C:37	1	Ms(I) 1	-	Sc6
P.C:38	1	Ms(I) 2	-	Sc2
P.C:39	1	Ms(I) 1	B	Sc1
P.C:40	1	Ms(I) 1	B	Sc6
P.C:41	2	N.A	-	Sc6
P.C:42	1	Ms(I) 2	A	Sc1
P.C:43	1	Ms(I) 2	A	Sc6
P.C:44	1	Ms(I) 2	-	N.A
P.C:45	1	Ms(I) 2	C	Sc1
P.C:46	1	Ms(I) 1	D	Sc1
P.C:47	1	N.A	B	Sc2
P.C:48	1	Ms(I) 1	E	N.D
P.C:49	1	Ms(I) 1	-	Sc2
P.C:50	3	Ms(I) 2	-	Sc8

N.A: Data not available
- : Result negative
N.D: No Digestion

TABLE 4.23 PCR-AFLP/RFLP PROFILES FOR REGIONS SHOWING VARIATIONS

SAMPLE No.	PROFILES			
	AFLP (8/16)	RFLP (8/16)	AFLP (22/23)	RFLP (24/23)
U.S:1	3	Ms(I) 2	G	Sc3
U.S:2	3	Ms(I) 2	B	Sc6
U.S:3	1	Ms(I) 1	B	Sc3
U.S:4	2	Ms(I) 2	D	N.A
U.S:5	2	Ms(I) 1	B	Sc3
U.S:6	1	Ms(I) 1	-	Sc2
U.S:7	1	Ms(I) 2	A	Sc7
U.S:8	1	Ms(I) 2	A	N.A
U.S:9	1	Ms(I) 2	G	Sc2
U.S:10	3	Ms(I) 1	G	Sc6
U.S:11	3	Ms(I) 1	-	Sc6
U.S:12	1	Ms(I) 1	A	Sc8
U.S:13	1	Ms(I) 1	D	N.D
U.S:14	1	Ms(I) 1	C	Sc3
U.S:15	1	N.A	A	Sc1
U.S:16	2	Ms(I) 2	-	Sc8
U.S:17	1	Ms(I) 1	H	N.A
U.S:18	3	Ms(I) 2	A	Sc2
U.S:19	1	Ms(I) 2	B	Sc1
U.S:20	1	Ms(I) 2	D	Sc4
U.S:21	1	Ms(I) 1	C	Sc8
U.S:22	3	Ms(I) 2	A	Sc1
U.S:23	2	Ms(I) 1	D	Sc3
U.S:24	2	Ms(I) 2	A	Sc2
U.S:25	2	Ms(I) 1	F	Sc1
U.S:26	2	Ms(I) 1	F	Sc1
U.S:27	N.A	N.A	D	Sc2
U.S:28	2	Ms(I) 2	-	Sc2
U.S:29	2	Ms(I) 1	G	Sc1
U.S:30	2	Ms(I) 2	G	Sc2
U.S:31	3	Ms(I) 1	D	Sc2
U.S:32	2	Ms(I) 1	D	Sc6
U.S:33	1	Ms(I) 1	D	Sc6
U.S:34	2	Ms(I) 2	F	N.D
U.S:35	1	Ms(I) 1	A	Sc6
U.S:36	3	Ms(I) 1	A	Sc1
U.S:37	1	Ms(I) 1	N.A	N.A
U.S:38	1	Ms(I) 1	N.A	N.A
U.S:39	2	Ms(I) 2	B	Sc2
U.S:40	2	Ms(I) 1	B	Sc3
U.S:41	2	Ms(I) 2	E	Sc3
U.S:42	2	Ms(I) 2	-	Sc7
U.S:43	2	Ms(I) 1	-	Sc2
U.S:44	2	Ms(I) 2	F	N.A
U.S:45	2	N.A	F	Sc3
U.S:46	2	Ms(I) 1	E	Sc2
U.S:47	2	Ms(I) 2	D	N.A
U.S:48	2	Ms(I) 1	F	N.A
U.S:49	2	Ms(I) 2	D	N.A
U.S:50	2	Ms(I) 2	D	Sc2

N.A: Data not available
- : Result negative
N.D: No Digestion

rRNA gene predominantly occurred in low nitrogen fixing trees. There was only one exception to this i.e., of the ten trees showing this profile only one tree (tree no. P.C:50) exhibited a high nitrogen fixation rate (38.63 n Mol ethylene produced/g fresh wt./hr).

Comparison of the RFL profiles obtained for the 18S-28S ITS region to the nitrogenase activities of the trees revealed that Profile No. Sc3 was found to occur in twelve trees. All these twelve trees had low ARA values, the least being 1.20 n Mol ethylene produced/g fresh wt./hr and the maximum was 12.35 n Mol ethylene produced/g fresh wt./hr. Profile Sc5 was unique to a single tree having low ARA value. The other RFL profiles obtained for this region could differentiate between trees but they were not strictly restricted to either of the nitrogen-fixing groups. Nevertheless, profile Sc3 can definitely be considered as a molecular marker to weed out the low nitrogen fixing trees.

Overall analysis, clubbing all the profiles for a particular tree, could not clearly differentiate the two nitrogen fixing groups. However, an interesting observation was made. Tree numbers P.C:27 and P.C:39 had very high nitrogenase activity, exceeding 80 n Mol ethylene produced/g fresh wt./hr. They had PCR-AFL/RFL profiles that were unique to each. The former had the profile (1,Ms(I)2,I,Sc6) and the latter (1,Ms(I)1,B,Sc1). It is noteworthy that these two profile groups were very rare. It is difficult to say at this stage whether the high nitrogenase activity and these profile groups were associated by chance or not. Further analysis of a larger population of alder trees would be required to clinch the issue.

A question may be raised about the presence of different *Frankia* genotypes in the nodules. We have taken that into account. One hundred fifty nodules/ tree were randomly collected. Therefore, it is likely that the entire range of *Frankia* genotypes was represented. The average nitrogenase activity estimates, therefore, represented all the different *Frankia* genotypes associated with the given tree. Therefore, a tree supporting high/ low nitrogenase activity nodules is more likely to be a function of the genotype of the tree.

Thus, the present study has been able to develop molecular markers for eliminating the *Alnus nepalensis* genotypes that support low nitrogenase activity. In addition to this, the study also identified two profiles as likely markers for *Alnus nepalensis* genotypes supporting high nitrogenase activity.

CONCLUSION

CHAPTER 5

CONCLUSION

1. Variability was observed in the proximal part of the chloroplast 16S rRNA gene. AFLP analysis of this region could categorize *Alnus nepalensis* into three groups. RFLP analysis of the same region also revealed three groups. However, the three groups due to RFLP were independent of those due to AFLP. That is to say that all the trees belonging to one group of AFLP were not necessarily members of the same RFLP group.
2. AFLP and RFLP analysis of the middle part of the 16S rRNA gene did not reveal any diversity. Therefore, it is highly conserved region and was not suitable for detecting intra-specific variability.
3. A lot of genetic variability was observed in the 18S-28S ITS region. AFLP analysis could divide the alder trees into nine different groups (A-I). On the basis of RFLP analysis also the alder trees could be divided into nine different PCR-RFLP (Sc1-Sc9) groups. However, like the proximal part of the 16S rRNA gene, the groups identified were independent of each other.
4. Nucleotide sequencing was done for three representative samples of *Alnus nepalensis*. When aligned sequences were subjected to the phylogenetic studies, *Alnus nepalensis* was found closest to the European species *Alnus glutinosa.*
5. The genus specific probe in the distal region of the 18S rRNA gene was designed on the basis of the aligned nucleotide sequences of related genera.
6. RFLP analysis of the 18S-28S *rrn* operon revealed the molecular marker (Sc3) that could be used to weed out genotypes supporting lower nitrogen fixing rates.
7. Taking all the PCR-AFL/RFL profiles together, two groups could be identified as likely molecular markers for *Alnus nepalensis* genotypes supporting high nitrogenase activity.

8. Further analysis of a bigger fragment of DNA is needed for developing definite positive markers. This can be accomplished by designing an appropriate primer which in conjunction with the primer designed in the present study, would give a larger amplicon, facilitating analysis of a larger region.

REFERENCES

REFERENCES

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

Bassam B.J., G. Caetano-Anolles and P.M. Gresshoff. (1991). Fast and sensitive silver staining of DNA in Polyacrylamide gels. Anal. Biochem. 196: 80-83.

Benson D.R. and W.B. Silvester. (1993). Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. Microbiol. Rev. 57 (2): 293-319.

Boerwinkle E., W. Xiong, E. Fourest and L. Chan. (1989). Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: Application to the apolipoprotein B 3' hypervariable region. Proc. Natl Acad. Sci. USA. 86: 212-216.

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 (5): 1569-1576.

Bousquet J., S.H. Strauss and P. Li (1992a). Complete congruence between morphological and *rbc*L - based molecular phylogenies in birches and related species (Betulaceae). Mol. Biol. Evol. 9 (6): 1076-1088.

Bousquet J., S.H. Strauss, A.H. Doerksen and R.A. Price. (1992b). Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. Proc. Natl. Acad. Sci. USA. 89: 7844-7849.

Bult C., M. Kallersjo and Y. Suh. (1992). Amplification and sequencing of 16/ 18S rDNA from gel-purified total plant DNA. Pl. Mol. Biol. Rep. 10 (3): 273-284.

Byrnes L., B. Enenkel, F. Gannon and T. Smith. (1995). Preparation of DNA for use as probes. In Gene Probes 1: A Practical Approach. B.D. Hames and S.J. Higgins (eds.), IRL Press, Oxford. pp17-55.

Caetano-Anolles G. and P.M. Gresshoff. (1994). DNA amplification fingerprinting using arbitrary mini-hairpin oligonucleotide primers. Biotechnology. 12: 619-623.

Caetano-Anolles G. and P.M. Gresshoff. (1996). Generation of sequence signatures from DNA amplification fingerprints with mini-hairpin and microsatellite primers. BioTechniques. 20 (6): 1044-1056.

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.

Chase M.W., D.E. Soltis and R.G. Olmstead (1993). Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. Mo. Bot. Garden. 80: 528-580.

Chen Z.D., S.R. Manchester and H.Y. Sun. (1999). Phylogeny and evolution of the Betulaceae as inferred from DNA sequences, morphology and paleobotany. Am. J. Bot. 86 (8): 1168-1181.

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 (9): 3539-3543.

Cournoyer B., M. Gouy and P. Normand. (1993). Molecular phylogeny of the symbiotic actinomycetes of the genus *Frankia* matches host-plant infection processes. Mol. Biol. Evol. 10 (6): 1303-1316.

Diem H.G., D. Gauthier and Y.R. Dommergues. (1983). An effective strain of *Frankia* from *Casuarina* sp. Can. J. Bot. 61: 2815-2821.

Felsenstein J. (1993). PHYLIP (Phylogeny Inference Package) Version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, U.S.A.

Frascaria N., L. Maggia, M. Michaud and J. Bousquet. (1993). The *rbc*L gene sequence from chestnut indicates a slow rate of evolution in the Fagaceae. Genome: 36: 668-671.

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

Gielly L. and P. Taberlet. (1994). The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbc*L sequences. Mol. Biol. Evol. 11 (5): 769-777.

Gilbert D.G. (January, 1999). PHYLODENDRON version 0.8d, beta. Biology Department, Indiana University, USA.

Guan C., K. Pawlowski and T. Bisseling (1998). Interaction between *Frankia* and actinorhizal plants. Sub Cellular Biochemistry. 29: Plant-Microbe Interactions. Biswas and Das (eds)., Plenum Press, New York. pp165-189.

Hadrys H., M. Balick and B. Schierwater. (1992). Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. Molecular Ecology. 1: 55-63.

Hahn D., M. Dorsch, E. Stackbrandt and A.D.L. Akkermans. (1989). Synthetic oligonucleotide probes for identification of *Frankia* strains. Plant and Soil. 118: 211-219.

Hamby R.K. and E.A. Zimmer. (1991). Ribosomal RNA as a phylogenetic tool in plant systematics. In P.S. Soltis, D.E. Soltis and J.J. Doyle (Eds.), Molecular Systematics of plants.. Chapman and Hall, New York. pp 50-91.

Hamby R.K., L.E. Sims, L.E. Issel and E.A. Zimmer. (1988). Direct ribosomal RNA sequencing: optimization of extraction and sequencing methods for work with higher plants. Pl. Mol. Bio. Rep. 6: 175-192.

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

Hooker J.D. (1825). The Flora of British India. M/S Bishen Singh and Mahendra Pal Singh, Dehradun. vol 5 p 600.

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. Ecol. 2: 17-26.

Janssen P., R. Coopman, G. Huys, J. Swings, M. Bleeker, P. Wos, M. Zabeau and K. Kersters. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. Microbiology. 142: 1881-1893.

Kawashima H., M.J. Bazim and J.M. Lynch. (1997). A modeling study of world protein supply and nitrogen fertilizer demand in the 21st century. Environmental Conservation: 24 (1): 50-56.

Laguerre G., P. Mavingui, M. Allard, M. Charnay, P. Louvrier, S. Mazurier, L. Rigottier-Gois 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. Microbiol. 62 (6): 2029-2036.

Lalonde M., L. Simon, J. Bousquet and A. Seguin. (1988). Advances in the taxonomy of *Frankia*: recognition of species *alni* and *elaeagni* and novel sub species *pommerii* and *vandijkii*. *In* Nitrogen Fixation: Hundred Years After. Bothe H., F.J. de Bruijn and W.E. Newton (eds). Gustav Fisher, Stuttgart. pp 671-680.

Lawrence G.H.M. (1967). Taxonomy of vascular plants. Oxford and IBH Publ. Co., New Delhi.

Lechevalier M.P. and H.A. Lechevalier. (1989). Genus *Frankia* Brunchrost (1886), 174 AL. *In* Bergey's manual of systematic bacteriology. Williams S.T., M.E. Sharp and J.G. Holt (eds.). The Williams and Wilkins Co., Baltimore, USA. pp 2410-2417.

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

Ludovic G. and P. Taberlet. (1994). The use of chloroplast DNA to resolve plant phylogenies: non-coding versus *rbc*L sequences. Mol. Biol. Evol. 11 (5): 769-777.

Lumini. E., M. Bosco, G. Puppi, R. Isopi, M. Frattegiani, E. Buresti and F. Favilli. (1994). Field performance of *Alnus cordata* Loisel (Italian Alder) inoculated with *Frankia* and VA-Mycorrhizal strains in mine-spoil afforestation plots. Soil Biol. Biochem. 26: 659-661.

Lumini. E. and M. Bosco. (1996). PCR-Restriction Fragment Length Polymorphism identification and host range of single spore isolates of flexible *Frankia* sp. UFI32715. Appl. Environ. Microbiol. 62: 3026-3029.

Maggia L., and J. Bousquet. (1994). Molecular phylogeny of the actinorhizal Hamamelidae and relationship with host promiscuity towards *Frankia*. Mol. Ecol. 3: 459-467.

Margheri M.C., M. Bosco, L. Vagnoli and F. Favilli. (1989). Compatibility of *Frankia* strains belonging to different host-specificity groups with *Alnus glutinosa*, *A. incana* and *A. cordata*. Ann. Microbiol. 39: 247-255.

Miller I.M. and D.D. Baker. (1985). The initiation, development and structure of root nodules in *Elaeagnus angustifolia* L. (Elaeagnaceae). Protoplasma. 128: 107-119.

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.

Mullin B.C., P. Goetting-Minesky and P. Twigg. (1993). Differential gene expression in the development of actinorhizal root nodules, *In* New Horizons in Nitrogen Fixation. Palacios R., J. Mora and W.E. Newton (eds). Kluwer, Dordrecht, The Netherlands. pp 309-314.

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.

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. Bacteriol. 173: 4072-4078.

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., 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. Int. J. Syst. Bacteriol. 46 (1): 1-9.

Prat D. (1989). Effect of some pure and mixed *Frankia* strains on seedling growth in different *Alnus* species. Plant and Soil. 113: 31-38.

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.

Ramakrishnan P.S. (1994). The jhum agro ecosystem in north-eastern India: A case study of the biological management of soils in a shifting agricultural system; *In* The Biological Management of Tropical Soil Fertility. Woomer P.L. and M.J. Swift (eds). John Wiley and Sons; Chichester, U.K. pp189-207.

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. Mirobiol. 62: 979-985.

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

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. (1999). Use of molecular characteristics in *Frankia* for strain identity. Ph.D thesis, NEHU, Shillong, India.

Sarma G., A. Sen, R. Varghese and A.K. Misra (1998). A novel technique for isolation of *Frankia* and generation of single spore cultures. Canadian J. Microbiol. 44: 490-492.

Savard L., M. Michaud and J. Bousquet. (1993). Genetic diversity and phylogenetic relationships between birches and alders using ITS, 18S rRNA, and *rbc*L gene sequences. Mol. Phylogenet. Evol. 2 (2); 112-118.

Schwintzer C.R. and J.D. Tjepkema. (1990). The Biology of *Frankia* and Actinorhizal plants. Academic Press, California, USA.

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., R. Sharma and M. Pradhan. (1998). Ecology of Himalayan alder (*Alnus nepalensis* D. Don). PINSA B64 (1): 59-78.

Sheffield V.C., D.R. Cox, L.S Lerman and R.M. Myers. (1989). Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. Proc. Natl. Acad. Sci. USA. 86: 232-236.

Simon. L., A. Stein and M. Lalonde. (1985). Performance of *in vitro* propagated *Alnus glutinosa* (L) Gaertn: Clones inoculated with *Frankia.* Plant and Soil. 105: 33-40.

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

Simonet. P, M.C. Grosjean, A.K. Misra, S. Nazaret, B. Cournoyer and P. Normand. (1991). *Frankia* genus-specific characterization by polymerase chain reaction. Appl. Environ. Microbiol. 57 (11): 3278-3286.

Soltis D.E., P.S. Soltis, M.T. Clegg and M. Durbin. (1990). *rbc*L sequence divergence and phylogenetic relationships in Saxifragaceae sensu lato. Proc. Natl. Acad. Sci. USA. 87: 4640-4644.

Soltis D.E., P.S. Soltis, D.R. Morgan, S.M. Swensen, B.C. Mullin, J.M. Dowd and P.G. Martin. (1995). Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc. Natl. Acad. Sci. USA. 92: 2647-2651.

Sougoufara B., L. Maggia, E. Duhoux and Y.R. Dommergues. (1992). Nodulation and N_2 fixation in nine *Casuarina* clone-*Frankia* strain combinations. Acta. Oecologica. 13 (4): 497-503.

Sprent J.I. and P. Sprent. (1990). Legume and non-legume nodules. Nitrogen fixing organisms-pure and applied aspects. Chapman and Hall, London. pp 30-65.

Stewart W.P.D., G.P. Fitzgerald and R.H. Burris. (1968). Acetylene reduction in nitrogen fixing blue-green algae. Arch. Microbiol. 62: 336-348.

Suh Y., L. B. Thien, H.E. Reeve and E.A. Zimmer. (1993). Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. Am. J. Bot. 80 (9): 1042-1055.

Swensen S.M. (1996). The evolution of actinorhizal symbiosis: Evidence for multiple origins of the symbiotic association. Am. J. Bot. 83 (11): 1503-1512.

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 symbiosis. Physiol. Plantarum 99: 565-573.

Taberlet P., L. Gielly, G. Pautou and J. Bouvet. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. Pl. Mol. Biol. 17: 1105-1109.

Thomson 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.

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

Tjepkema J.D., C.R. Schwintzer and D.R. Benson. (1986). Physiology of actinorhizal nodules. Annu. Rev. Plant Physiol. 37: 209-232.

Tremblay F.M., P. Perinet and M. Lalonde. (1986). Tissue culture of *Alnus* spp. with regard to symbioses. *In* Biotechnology in Agriculture and Forestry. Bajaj Y.P.S. (ed.). Springer-Verlag, Berlin. pp 87-100.

Tohdoh N. and M. Sugiura. (1982). The complete nucleotide sequence of a 16S ribosomal RNA gene from tobacco chloroplasts. Gene. 17: 213-218.

Varghese R. (2000). Investigations on molecular diversity in alder compatible Frankiae. Ph.D thesis, NEHU, Shillong, India.

Verghese S.K. (1999). Development of molecular signatures for *Frankia* strains. Ph.D. thesis, NEHU, Shillong, India.

Verghese S.K. and A.K. Misra. (2000). PCR-RFLP based screening of *Frankia* in alder nodules having different levels of nitrogenase activity. Symbiosis: 28: 337-350.

Welsh J. and M. McClelland. (1990). Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18: 7213-7218.

Wen J. and E.A. Zimmer (1996). Phylogeny and biogeography of *Panax* L. (the Ginseng Genus, Araliaceae): Inferences from ITS sequences of nuclear ribosomal DNA. Mol. Phylogenet. Evol. 6 (2): 167-177.

Zurawski G., M.T. Clegg and A.H.D. Brown. (1984). The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. Genetics. 106: 735-749.

APPENDICES

APPENDIX 1

BUFFERS USED FOR EXTRACTION OF DNA AND AGAROSE GEL ELECTROPHORESIS

1. EXTRACTION BUFFER FOR ISOLATION OF DNA (pH 8.0)

COMPONENT	AMOUNT/100ML
Tris base (1M)	10ml
NaCl (5M)	28ml
EDTA (0.5M)	4ml
CTAB	2% (wt./vol.)
PVP	3% (wt./vol.)

CTAB: Cetyl Trimethyl Ammonium Bromide
PVP : Polyvinylpyrrolidone

2. TBE BUFFER (pH 8.0)

COMPONENT	AMOUNT/LITRE
Tris base	54.0g
Boric acid	27.5g
EDTA (0.5M)	20ml

3. TYPE III LOADING BUFFER (6X)

Bromophenol blue	0.25% (wt./vol.)
Xylene Cyanol FF	0.25% (wt./vol.)
Glycerol	30% in water (wt./vol.)

APPENDIX 2

BUFFERS USED FOR PCR AND RESTRICTION ANALYSIS

1. **10X PCR BUFFER** (pH 8.3 at 25°C)

Tris–HCl	100mM
KCl	500mM
$MgCl_2$	15mM
Gelatin	0.01% (wt./vol.)

2. **BUFFER B (BOEHRINGER MANNHEIM)** (pH 8.0 at 37°C)

Tris-HCl	10mM
NaCl	100mM
$MgCl_2$	5mM
2-mercaptoethanol	1mM

3. **1X NE *Nru*I BUFFER (NEW ENGLAND BIOLABS)** (pH 7.7 at 25°C)

KCl	100mM
Tris-HCl	50mM
$MgCl_2$	10mM

4. **BUFFER C (BANGALORE GENEI)** (pH 7.8)

Tris-HCl	10mM
NaCl	50mM
$MgCl_2$	10mM
2-mercaptoethanol	5mM

5. **BUFFER L (BANGALORE GENEI)** (pH 7.8)

Tris-HCl	10mM
$MgCl_2$	10mM
2-mercaptoethanol	5mM

6. **BUFFER B (BANGALORE GENEI)** (pH 8.0)

Tris-HCl	10mM
NaCl	100mM
$MgCl_2$	10mM
2-mercaptoethanol	5mM