Chapter 2
REVIEW OF LITERATURE

2.1 Relevance of Biological Nitrogen Fixation Research

All the living organisms required nitrogen for their daily needs like preparation of proteins, enzymes etc. It is a triple bonded, stable, almost inert colourless, odorless gas. In Earth's atmosphere there is about 77% of nitrogen which is conserved through nitrogen cycle (Rosswall, 1981). The cycle includes various steps which includes nitrogen assimilation, mineralization, nitrification, denitrification and biological nitrogen fixation. Most plants and microbes preferably assimilate the ammonium salts as a source of nitrogen through either glutamate dehydrogenase (GDS) or glutamine synthetase/ glutamate synthase (GS/GOGAT) cycle. This GS/ GOGAT cycle is widely and preferably used by organisms (Miflin and Lea, 1977). The ability to convert the molecular nitrogen into cellular nitrogen is called biological nitrogen fixation. This is an unique property of a group of free living, associative and symbiotic prokaryotic microorganisms. This fixed nitrogen is the primary source of available nitrogen in the nature.

In the modern technological era, scientists invented another industrial method, called Haber-Bosch process to produce nitrogenous fertilizer as an artificial source of nitrogen. This process is very energy consuming and complicated. However it is very demanding as the demands for nitrogenous fertilizer increases day by day as nitrogen (N) is the most commonly used limiting nutrient used for plant growth. Ladha and Reddy (1995) estimated the average energy used in ammonia production in Haber-Bosch process. It is well known that

the efficiency of the industrial process varies considerably, but it is estimated that about 1.3 ton of oil (or equivalent amount of energy) is required to fix 1 ton of NH_3. The annual production of about $77x10^6$ ton of NH_3- N in the year 1996 requires about 0.1 x 10^9 ton oil or equivalent. This is about 1.4% of the total fossil fuel consumption in the year 1995. (Patyk and Rheihardt, 1997; Scholz *et al.*, 1997). The combustion of such oil or coal add a huge amount of pollutant to the environment. Unfortunately our chemically produced nitrogenous fertilizer dependent green revolution can not afford any minimization of such figure which may result in food shortages due to inability of the industry to produce nitrogenous fertilizers.

After the Second World War, colonial powers lost their control over major parts of the globe and either democratic or socialist governments came to the power almost throughout the world. The main challenge of these responsible governments was to fight against hunger and poverty and feed their population that is to give the minimum food security to them. In the 20^{th} century grain crop yields have constantly increased and reached a remarkable level in many areas of agriculture throughout the world. In India it is generally known as "Green revolution". This green revolution occurred due to introduction of genetically improved cultivars of crop plants as well as use of chemical pesticides and of course use of huge amount of nitrogen fertilizers. These cultivars are very nitrogen-responsive varieties and produce large reproductive structures which can accommodate more grain. This is the reason for increase of use of total nitrogenous fertilizer between 1950 and 1990. The increase was about 10 fold. Recent studies show that the yields of the major cereal crops and utilization of nitrogenous fertilizer did not increase in the last decade. In some instances like Japan, United States of America, Western Europe it decreased slightly. On the other hand, farmers in Asian countries like India and China uses huge quantity of nitrogen fertilizer. It is estimated that there is a massive gap between the nutrient inputs from external sources and the constant drain of nutrients from the soil due to crop removals and soil erosion. In India, the estimated nutrient deficit was about 10 million tons in 1990, and is likely to touch new summit with more intensive cropping systems and

increasing soil degradation and finally leading to loss of soil fertility (Bumb and Baanante, 1996).

If the fertilizer demands of next few years are considered, it is found that the food needs will increase and it will come from a nearly static land base. Thus, a more intensive agriculture with more and more use of fertilizer is expected. Bumb and Baanante (1996) based on an econometric model estimated 1.2% increase in the world fertilizer use from 1990 to 2020. It is estimated that the expected fertilizer use will rise from 143.6 million tons to 208 million tons from 1990 to 2020. It is very interesting to note here that the rate of increase of nitrogen fertilizer is much lower than that compared with the period of 1960-1990 (Buckley, 2000). Bumb and Baanante (1996) explained this lower rate as 'a higher base, limited potential for further growth and changing policy environment. It is because of the already high application rates, environmental concerns, reduction in farm support programme and trade liberalization'. Buckley (2000) supported the calculation of Bumb and Baanante (1996) and estimated an annual 1.1% increase in nitrogenous fertilizer use from 1990 to 2070.

So, it is established that nitrogen fertilizer reached its peak and left very limited potential for growth in the present century and it is the time to search an alternate source of fertilizer which would be environment friendly and have higher potential for growth. In a global scenario of reaching nitrogen resource to a plateau, rising concerns over possible environmental effects of chemical fertilizers, as well as their cost for small-scale farmers in developing countries, it is essential to expand the use of the biological nitrogen fixation (BNF) technologies that offer the greatest environmental and economic benefits for each specific agro-ecosystem.

Biological nitrogen fixation is generally found in prokaryotic system. *Eriphorum vaginatum*, an eukaryotic plant, is the only exception (Chapin *et al.*, 1993). It is very difficult to calculate any direct correlation among nitrogen fixing bacteria, however the process of biological nitrogen fixation and its enzyme system is very similar in all nitrogen fixers. They are distributed among 27 families and 80 genera of eubateria (including cyanobacteria) and 3 thermophilic genera of archaebacteria (Elkan, 1992). This process has four universal

requirements, which are as follows:

1. An enzyme system, called nitrogenase enzyme complex.

2. A high energy requirement and availability (ATP).

3. Anaerobic condition for nitrogenase activity.

4. A strong reductant.

During the nitrogen fixation process nitrogenase complex catalyzes the ATP dependent reduction of atmospheric nitrogen to ammonium and is composed of two protein called dinitrogenase and dinitrogenase reductase. Dinitrogenase is the most important enzyme of the system, which reduces the atmospheric dinitrogen into ammonia at appropriate temperature and pressure (Peters and Szilagyi, 2006).The dinitrogenase protein is a $\alpha 2\beta 2$ tetramer of 240 KDa that are encoded by the gene *nif*KD. The complex also contains two types of metal centres, the P clusters (8Fe-8S), which bridges the α and β subunits and an unique iron –molybdenum co-factor (FeMo-Co), which is buried within the α subunit and is the site of substrate reduction (Dean *et al*, 1993). The α and β subunits of nitrogenase is linked with each other by cysteine residues of molybdo-ferro protein (Rubio and Ludden, 2008). Dinitrogenase reductase is a 60KDa homodimer that is encoded by the gene *nif*H. It specifically reduces dinitrogenase, apparently transferring elements to the P-cluster, which then channels them to FeMo-Co metalocluster. In addition to this catalytic role, dinitrogenase reductase is also involved in the biosynthesis of FeMo-Co and the maturation of dinirtogenase (Filler *et al.*, 1986).

The biological nitrogen fixing organisms can be grouped into three categories- freeliving, associative and symbiotic. Free living diazotrophs was first described by Winogradrodsky. They can fix sustainable amount of nitrogen either by any specialized structure like heterospore of cyanobacteria or in anaerobic or partial anaerobic condition in nature. The symbiotic organisms form an association with a particular host plant and fix nitrogen with in a specialized structure called nodule. Associative nitrogen fixing bacteria are newly invented group of micro organisms. They are the predominant microbial flora of root rhizosphere of particular host plant, but do not enter into the host or host do not produce any specialized structure for them (Elkan,

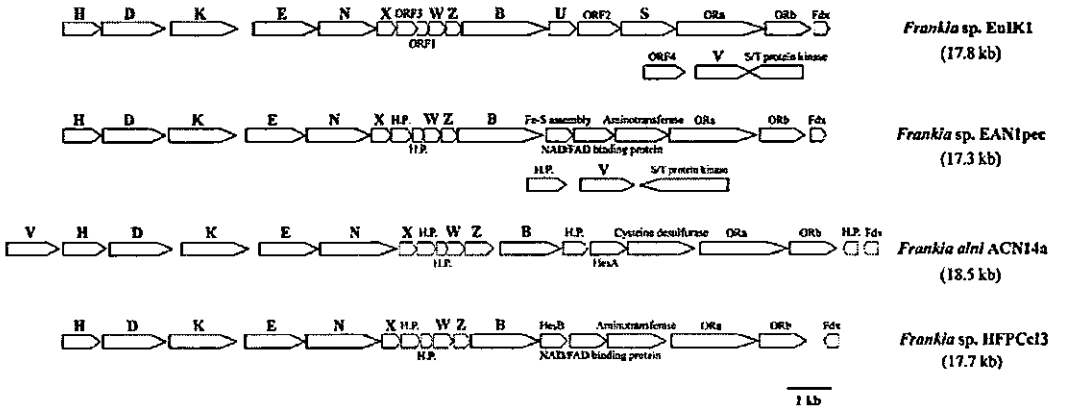


Fig 2.1 Comparison of the physical organization of *nif* and *nif*-associated genes from EulK1 strain with those from the 3 full-genome sequenced *Frankia* strains. Gene organization in the *nif* clusters are highly conserved among *Frankia* strains, except with regards to *nif*V location. In *Alnus* infecting strains, *nif*V was located just upstream of *nif*H, whereas the gene was found in a different genomic location in *Casuarina* and *Elaeagnus* infecting strains, including EuIK1 strains. Bold letters represents the nif gebnes, H.P = Hypothetical protein, Fe-S assembly accessory proteins (in EAN 1pec strains)
(Adopted from: "Organization of nif gene cluster in *Frankia* sp. EuIK1 strain, a symbiont of Elaeagnus umbellate' Chang Jae Oh • Ho Bang Kim • Jitae Kim •Won Jin Kim • Hyoungseok Lee • Chung Sun. An Arch Microbiol (2012) 194:29–34DOI 10.1007/s00203-011-0732-7)

1992).

2.1.1. Bacterial Genes Responsible For Biological Nitrogen Fixation

Primarily three gene families are associated with the process of biological nitrogen fixation. They are *nif, nod* and *fix* gene families. The *nif* genes are responsible for the production of nitrogenase (*nif*D and *nif*K) and dinitrogenase reductase (*nif*H) enzymes. Several *nif* genes are identified and are sequenced from various groups of microorganisms, which include *nif*E, *nif*N, *nif*V, *nif*B, *nif*S, *nif*A etc. They are structurally conserved (Fischer, 1994). The *fix* genes are those genes which have a definite role in nitrogen fixation and found in rhizobia, but their homologous counterpart in *Klebsiella pneumoniae* is not found. They are first identified in the symbiotic organism *Rhizobium meliloti* (Corbiun *et al.*, 1983; Earl *et al.*, 1987; Ruvkun *et al.*, 1982). Their presence was noted in several other rhizobia latter on, which includes *Bradyrhizobium japonicum,* (Fuhrmann *et al.*, 1985; Gubler and Hennecke, 1986) *R. leguminosarum bv. trifolii,* (Iismaa *et al.*, 1989) etc. Unlike the *nif* genes the *fix* genes are situated on a single operon and associated electron transport to nitrogenase. Important *fix* genes are *fix*A, *fixB*, *fix*H, *fix*N, *fix*O, *fix*X, *fix*R etc (Fischer, 1994). The *nod* genes are associated with the process of nodulation in legume- rhizobia system. *Frankia* genome contains some putative nod like genes, however, true *nod* gene is

Table 2.1 List of Actinorhizal plants ((Benson DR and Silvester WB, 1993; Sen 1996).

Plant Order	Plant family	Host plant genus
Fagales	Betulaceae	*Alnus*
Protales	Elaeagnaceae	*Elaeagnus, Hippophae, Shepherdia*
Rhamnales	Rhamnaceae	*Ceanothus, Colletia, Kentrothamnus, Retanilla, Telguenea, Trevoa, Discaria,*
Myricales	Myricaceae	*Myrica, Comptonia**
Casuarinales	Casuarinaceae	*Casuarina, Allocasuarina, Gymnostoma, Ceuthostoma*
Ranunculales	Coriariaceae	*Coriaria*
Rosales	Rosaceae	*Drayas, Purshia, Cercocapus, Cowania, Chamaebatia*
Violales	Datiscaceae	*Datisca*

* *Frankia* was first isolated from this plant

Table 2.2: Classification of nodulated plants of Rosid I linages (Wall. 2000).

Nitrogen fixing plant lineage	Nitrogen fixing plant Families
Lineage I	Myricaceae, Casuarinaceae and Betulaceae
Lineage II	Elaeagnaceae, Rhamnaceae and Rosaceae as well as *Parasponia* (Ulmaceae member, being infected by *Bradyrhizobium*)
Lineage III	Coriariaceae and Datiscaceae
Lineage IV	Fabaceae (Includes non actinorhizal legume plants)

absent in its genome (Franche *et al.*, 2009). The *nif* genes of *Frankia* are highly conserved. The size varies from 17.3kb (*Frankia* sp. EulK1) to 18.5 kb (*Frankia* strain ACN14a) (Chang *et al.*, 2012). The physical organization of *nif* and *nif*-associated proteins of various *Frankia* genomes are shown in Fig 2.1.

2.2 The Actinorhizal Plants

A group of dicotyledonous plants that are nodulated by *Frankia* is called actinorhizal plants. They are distributed among eight orders, eight families, 24 genera and 194 species of land plant (Table 2.1) (Benson and Silvester, 1993, Normand and Fernandez, 2006). These plants are found in different hemisphere of the earth ranging from arctic Tundra (*Dryas* spp.) and alpine forest (*Alnus* sp., *Coriaria sp.* etc) to costal and xeric condition (*Casuarina sp.*)(Fig 2.2). *Frankia* in symbiosis with actinorhizal plants fixed sustainable amount of molecular nitrogen, which ranges from

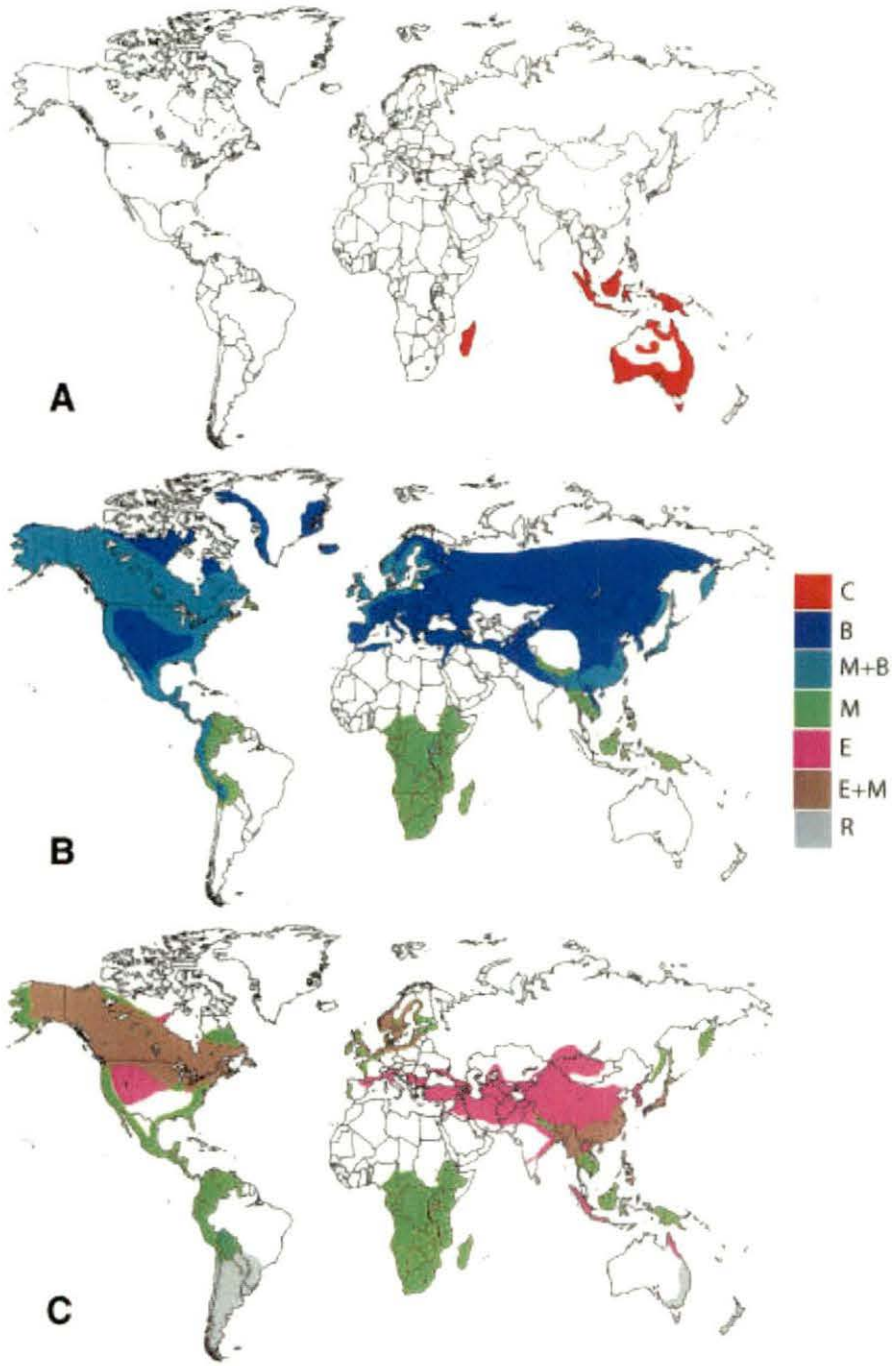


Fig 2.2 Present-day native distribution of actinorhizal plant hosts. (*A*) Distribution of plant hosts for CcI3, including *Casuarina* and *Allocasuarina* of the Casuarinaceae (C). (*B*) Distribution of plant hosts for ACN, including *Alnus* sp. in the Betulaceae (B) and Myricaceae (M) and their overlap (M+B). (*C*) Distribution of plant hosts for EAN including members of the Elaeagnaceae (E), Myricaceae (M), and the actinorhizal Tribe Colletieae of the Rhamnaceae in South America, Australia, and New Zealand (R). Elaeagnaceae and Myriceae (E+M) overlap in some areas. Maps were drawn with information from Silvester (1977) and from the Missouri Botanical Garden Web site (www.mobot.org). (Adopted from 'Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography' Philippe Normand *et al*, 2007 Genome Res. 2007 January; 17(1): 7–15.)

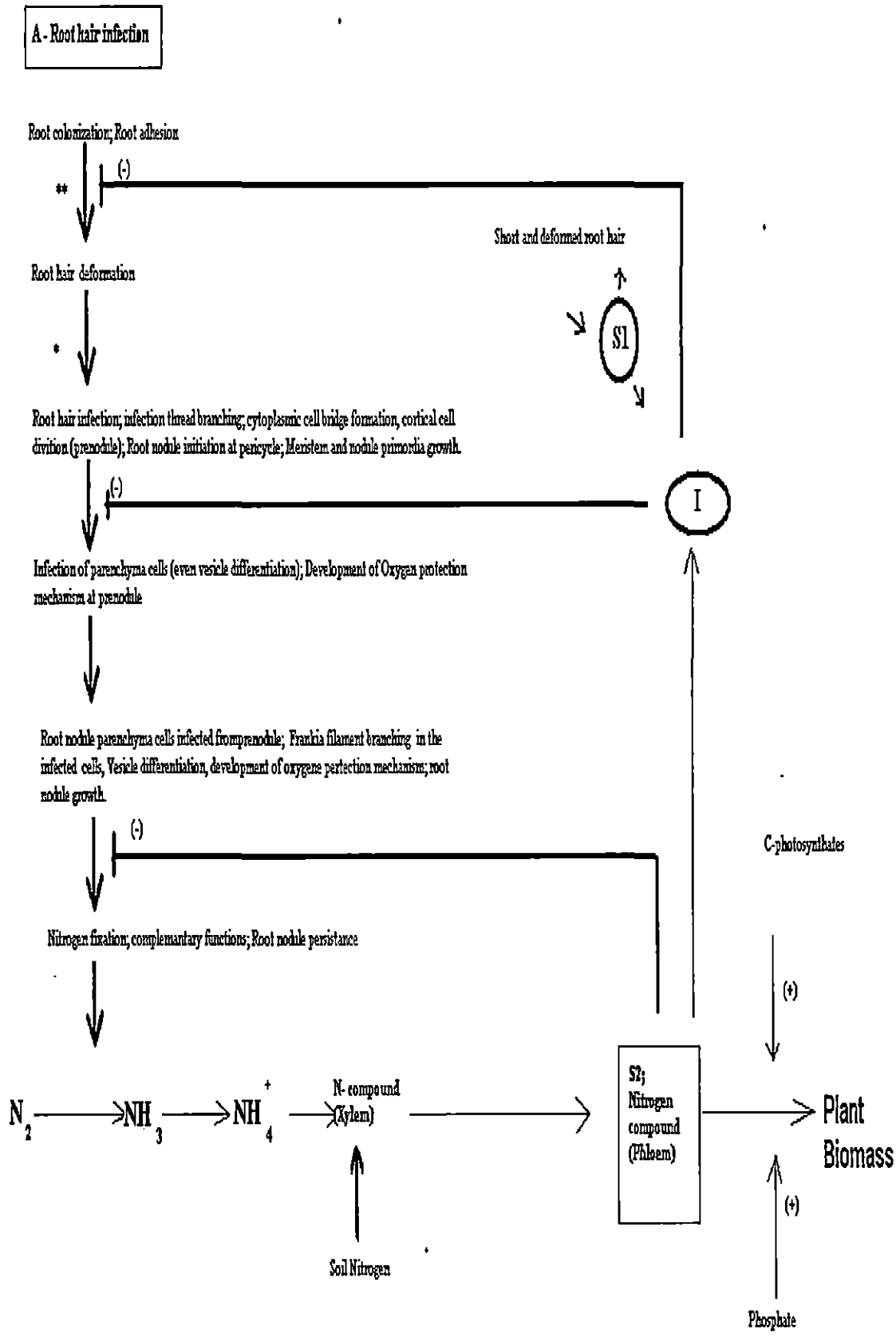


Fig 2.3 Model A for regulation of nodulation in actinorhizal symbiosis in root hair infection in *Frankia*. Thin lines means transient pathway. * = Second symbiotic interaction and molecular recognition step involving flavonoids. ** = First molecular signal involving the root hair deformation factor. (-) = Inhibition, (+) = Activation. S = Signal molecule, I = Inhibitor. (Wall, 2000)

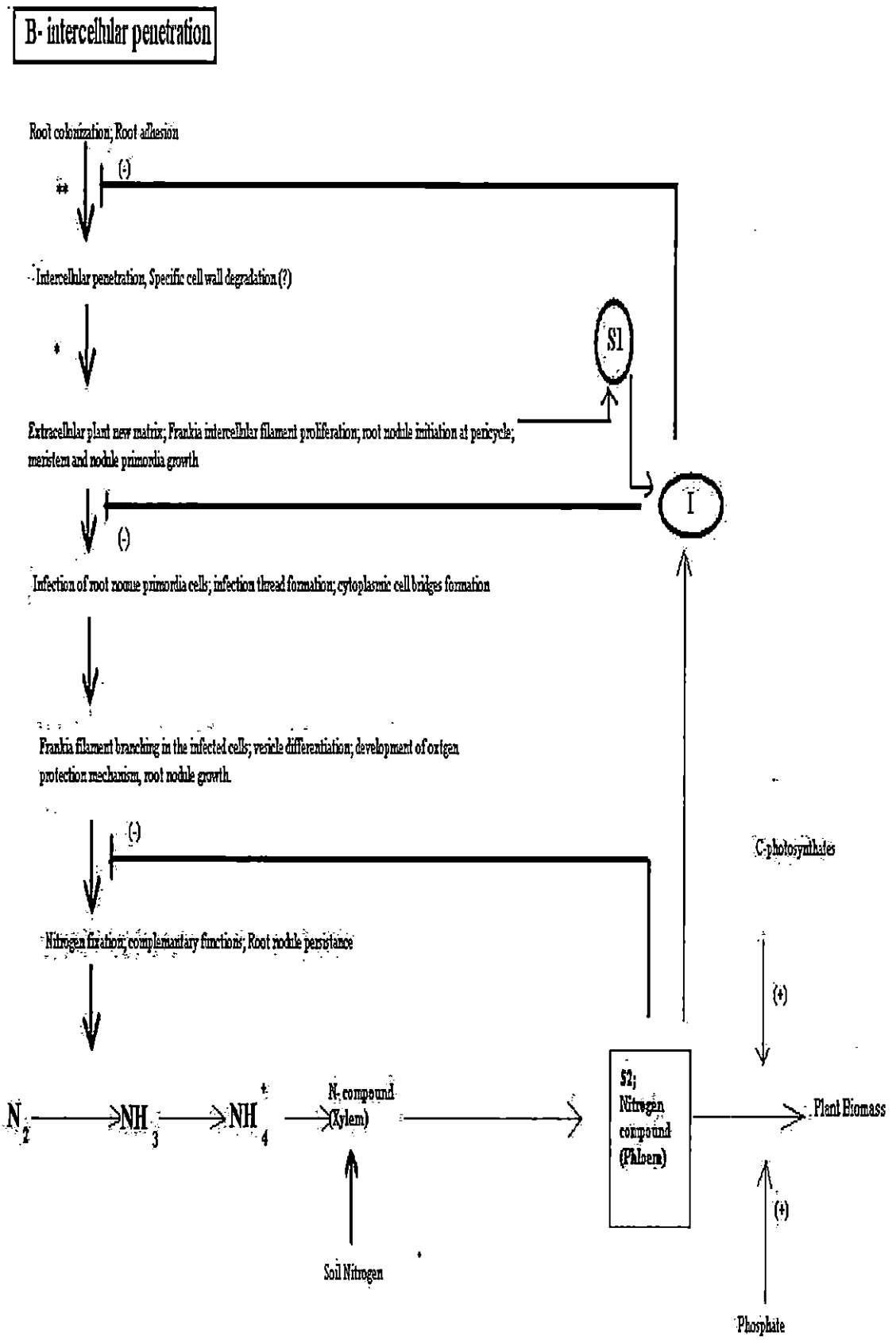


Fig 2.4 Model B for regulation of nodulation in actinorhizal symbiosis intercellular penetration of *Frankia*. Thin lines means transient pathway. * = Second symbiotic interaction and molecular recognition step involving flavonoids. ** = First molecular signal involving the root hair deformation factor. (-) = Inhibition, (+) = Activation. S = Signal molecule, I = Inhibitor. . (Wall, 2000)

240-350 kg N_2 $ha^{-1}y^{-1}$, which is comparable to those of leguminous plants (Dawson, 1990; Hibbs and Cromack, 1990; Wheeler and Miller, 1990; Wall, 2000). Since actinorhizal plants are the early visitors of marginal soils, they are considered as pioneer species in the landslides and other threatened areas. These plants are also used as wood, fiber, food, chemicals, etc (Benoit and Berry, 1990; Diem and Dommergues, 1988, 1990; Hibbs and Cromack, 1990; Myrold, 1994). In India, research on actinorhizal plants is limited. An excellent work has been done by Basistha *et al.* (2010) on *Hippophae salicifolia* growing in Lachen and Lachung valleys of North Sikkim where snowfall, heavy rain and landslides are very common. They found that this actinorhizal plant is growing as dominant species in different ecological and topological condition in that region. Recently, Bargali (2011) studied habit, habitat, distribution and possible ecological significance of eight actinorhizal plants found in Kumaun Himalayan region.

Three major phylogenetic subgroups of actinorhizal plants have been identified (Swensen, 1996). The first subgroup, includes symbiotic taxa from the families Betulaceae, Myricaceae and Casuarinaceae. The second subgroup, includes symbiotic taxa from the families Datiscaceae and Coriariaceae. The third subgroup, includes symbiotic taxa from the families Rhamnaceae, Rosaceae and Elaeagnaceae (Tab 2.2) (Swensen, 1996). The nodules formed by members of different actinorhizal plants differ considerably regarding organization of infected cells in the cortex (i.e. symbiotic, nitrogen-fixing cells), oxygen protection mechanisms for nitrogenase, patterns of *Frankia* differentiation, infection mechanisms, and organization of carbon and nitrogen metabolism (Pawlowski and Bisseling, 1996; Berry *et al.*, 2004, 2011; Schubert *et al.*, 2010, 2011).

The actinorhizal root nodules are established through a series of interactions between the host actinorhizal plant and the bacterium, *Frankia.* The actinorhizal root nodules are perennial with coralloid structure. It consists of multiple lobes. The *Myrica* and *Ceanothus* lobes are discrete and in *Alnus*, they are densely packed. Each nodule is a modified lateral root. The actinorhizal roots contain vascular cylinder, cortical tissue and periderm. The cortical cells get infected with the *Frankia.* From the apex of the mature

root nodule negatively geotrophic nodule root is developed. This nodule root does not take part in the nitrogen fixation process (Bajwa, 2004).

The ideal condition for root infection and subsequent nodule development is not well understood in *Frankia*-Actinorhizal system. *Frankia* hyphae were embedded in the mucilage layer of the root hair. It is established that a single infection is enough for nodule development. However, in the natural condition more than one infection occur in the nodule formation. In laboratory conditions the frequency of nodulation is directly proportional to the amount of inocula (Newcomb and Wood, 1987). It is interesting to note that root hairs of actinorhizal plants are deformed by several non-symbiotic soil bacteria (Knowlton *et al.*, 1980). These soil bacteria play an important role in the *Frankia*-actinorhizal interaction. They are called 'Helper' organisms. *Pseudomonas sp.* helps in the nodule formation in *Alnus* and *Casuarina* in controlled condition (Knowlton *et al.*, 1980).

The bacterium follows two different pathways for infection of the host tissue. It takes place either by root hair deformation as found in *Alnus, Casuarina, Myrica* etc (Callaham *et al.*, 1978, 1979; Berry and Torry, 1983) or intracellularly as in the case of *Ceanothus, Elaeagnus,* etc. Some strains can follow both the pathways for infection (Miller and Baker, 1986).

The signal molecule of these plants bear same basic chemical feature. Actinorhizal plants undergo feedback regulation of nitrogen fixing symbiosis with *Frankia,* which involves at least two different and consecutive signal molecules that leads to a mechanism controlling root nodulation in actinorhizal plants (Wall, 2000).

The mechanism of infection in actinorhizal plants follows two different ways –

i) root hair infection (*Casuarina sp.*) and

ii) intercellular infection (*Shepherdia sp.*).(Fig 2.3 and 2.4)

Frankia has very little or no host specificity as found in case of other nitrogen fixing bacterium like *Rhizobium.* In rhizobia–legume symbioses *nod* signal molecules helps in the nodulation and form the basis of host specificity (Oldroyd *et al.*, 2009) These kind of *nod* factors are absent in *Frankia.* The bioassay of reporter gene is also absent in *Frankia* (Céremonie *et al.*, 1999). On the basis of the study of

actinorhizal root hair deformation (Gherbi *et al.*, 2008; Markmann *et al.*, 2008) and signal factors of arbuscular mycorrhizal fungi (Maillet *et al.*, 2011), it is predicted that *Frankia nod* factor equivalents are a chitin-based molecule. Recently a LysM-type mycorhizal receptor was found in the infection process of *Rhizobium* symbiosis of the nonlegume *Parasponia andersonii* (Op den Camp *et al.*, 2011). This finding suggested that *Frankia* might have developed a unique pathway of synthesis of a novel chitin-based signal molecule unrelated to *Rhizobium* which functions in a *nod* factor-independent infection mechanism (Pawlowski *et al.*, 2011). In nodule, *Frankia* followed a unique pathway for nitrogen utilization. Unlike rhizobial symbiosis, the primary nitrogen assimilation might take place in *Frankia* and stored as arginine (Berry *et al.*, 2011).

Frankia is also able to produce a special structure called vesicle to protect its nitrogenase from O_2. Exception is *Casuarina* nodules, which provide a partially anoxic environment for *Frankia* to fix nitrogen by producing oxygen diffusion barrier by class II hemoglobin (Berg and McDowell, 1987, 1988; Jacobsen-Lyon *et al.*, 1995). However *Casuarina* infecting *Frankia* strains are capable enough to produce vesicle. In fact *Casuarina* - *Frankia* does produce vesicle to protect their nitrogenase from the toxic effect of oxygen, when growing in pure culture (Sen, 1996). *Frankia* strains have been shown to synthesize truncated hemoglobins for protection of nitrogenase from oxygen (Tjepkema *et al.*, 2002; Beckwith *et al.*, 2002; Coats *et al.*, 2009). Pawlowski *et al.* (2007) showed that in *Datisca glomerata* nodules, the plant also produces truncated hemoglobin in infected cells. It is involved in nitric oxide (NO) detoxification, not in O_2 transport as class I hemoglobin. The presence of truncated hemoglobin in the *D. glomerata* cells infected with *Frankia*, and the presence of large amounts of a class I hemoglobin in alder (Sasakura *et al.*, 2006) and *Myrica gale* nodules (Heckmann *et al.*, 2006) indicates that, as in legume nodules (Horchani *et al.*, 2011), large amounts of nitric oxide are produced in actinorhizal nodules, leading to high levels of stress. Class II haemoglobins are generally believed to contribute to oxidative stress in nodules by producing reactive oxygen species (ROS) such as superoxide anions (O^{2-})

and hydrogen peroxide (H_2O_2) (Becana *et al.*, 2000). This has been confirmed by the fact that RNAi inhibition of *leghemoglobin* gene transcription in nodules of the legume *Lotus japonicus*. It not only increase free O_2 and the loss of nitrogenase and nitrogen fixation in nodules, but also reduces H_2O_2 contents (Günther *et al.*, 2007). For *C. glauca*, it has been suggested that CgMT1, a class I type 1 metallothionein protein, could be part of the antioxidant system to prevent ROS accumulation in the nitrogen-fixing cells of the nodule (Obertello *et al.*, 2007). Additionally, Tavares *et al.* (2007) have shown that *Frankia* contributes to ROS production in nodules. Thus, both legume and actinorhizal nodules have to cope up with high levels of stress. Altogether, it should be pointed out that actinorhizal plants have a greater ability to tolerate extreme environmental conditions. Progress on research on growth condition of actinorhizal plants indicates that they grow in relatively adverse conditions than legumes, and therefore, it is also possible that they possess enhanced antioxidant based defense systems (Pawlowski *et al.*, 2011). Goyal *et al.* (2011) studied various extracts of *H. salicifolia* and found that it possess high antioxidant activities, which is helpful in preventing or slowing the progress of oxidative stress.

From the above discussion it is apparent, that there are two different ways of mechanism with two different sets of receptor and signal molecules. The intercellular infection takes place at a much faster rate than that of root hair infection. A flavonoid compound is involved in this process (Hughes *et al.*, 1999). However, the identity of the compound is not yet clear. There are several environmental and bacterial factors involved in nodulation. The environmental factors involve light, water, availability of nitrogen and phosphate, soil pH, pCO_2 and pO_2. The bacterial factors are concentration of the inocula, physiological state of the strain and nitrogen fixing ability of the bacterium. These factors play important role in the function, regulation and development of the nodules in actinorhizal plants (Fig 2.5). From the work of Valverde and Wall (1999), it is found that the tap root system consists of a temporary window of susceptibility for nodulation. Wall (2000) proposed two different models for the regulation of nodulation in actinorhizal plants which involved two

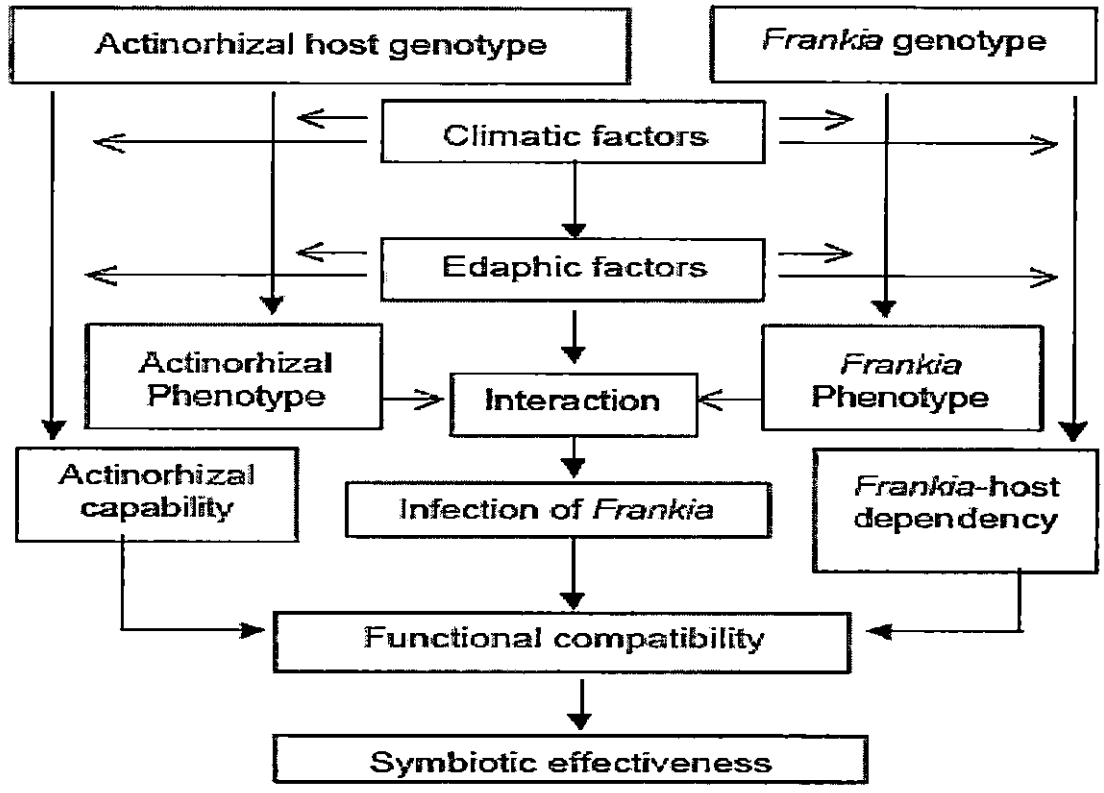


Fig 2.5 Complexity of host–*Frankia* interaction and symbiotic. (Adopted from "*Frankia*–actinorhizal symbiosis with special reference to host–microsymbiont relationship." By Verghese and Misra, 2002)

unknown signal molecule (S1, S2) and an unknown inhibitor molecule (I). These are two-step models and the signal and inhibitor molecules regulate the whole process by three independent pathways as follows:-

- Opening and closing of susceptibility window for interaction.
- Arresting nodule primordia at stages before host cell invasion by *Frankia* and vascular bundle differentiation.
- Inhibiting nodule development in the growing root (Fig 2.3 & 2.4).

Normand and his co-workers had proved the existence of the bacterium *Frankia* in more than 100 Million years ago on the basis of *nifH* gene sequence (Normand and Bousquet, 1989). The actinorhizal symbiosis may evolve almost in similar time (Sen, 1996). There are two different hypotheses that have been projected regarding the evolution of the actinorhizal symbiosis. According the first hypothesis, widely related plant genera acquired the capacity of making symbiotic relationship with *Frankia* to achieve selective advantages in certain

ecological niches. So, acquisition of the symbiotic character had hardly taken place before the divergence of families into their respective genera. The second hypothesis stated that the available soil nitrogen was limited in the early Cretaceous period (Bond, 1983) and a group of plant was forced to make association with nitrogen fixing soil bacteria *Frankia* to survive in struggle for existence with other wind pollinated angiosperms largely belonging to Magnoliidae. Some of these plants later evolved into the actinorhizal plants. During the last 100M years the availability of nitrogen in the soil increased due to nitrogen fixation and some of the ancient plants lost their symbiotic nitrogen fixing habit. Although this ancient property remains as a selective advantage in some pioneer plants which are now called actinorhizal plants. The DNA hybridization study by Bousquet *et al.* (1989) on host plants support the second hypothesis.

The *rbcL* gene sequence shows that the actinorhizal nitrogen fixing plants are grouped in Rosid I linages of the seed plants (Fig 2.6). The property of nitrogen fixation in nodules may evolve in a single time in the earth's history (Doyle, 1998). This Rosid I group may be subgrouped into four linages, of which three are actinorhizal and the fourth one is Fabaceae (Soltis *et al.,* 1995). The fossil records, geographical distribution, morphological and anatomical studies extended supports to this hypothesis (Wall, 2000). However, on the basis of 16s r-DNA sequence, *Frankia* can be divided into three clades and one '*Frankia* like clade'. The *Frankia* strains have been successfully isolated from Clade I and Clade II; however there is no isolates from Clade III *Frankia*. The existence of this clade is proved by PCR amplification of the 16s-rDNA. The analysis of data shows that the *Frankia*-Actinorhizal symbiosis follows a polyphyletic line of evolution and this occur at least three or four times in the history of evolution which clearly rejects the idea of Doyle (1998). This new idea was supported by Swenson (1996), Benson and Clawson (2000) and Jeong *et al* (2000). There are a few *Frankia rrn* gene sequences available for the study. From these sequences the evolutionary distances can be measured. This type of study was made in *Frankia* by Normand *et al.* (1996), Jeong *et al.* (1999) and Clawson *et al.* (2004) with different conclusions. The Ochman's

metric (Ochman and Wilson, 1987) is recognized as the best estimate for this type of work (Normand and Fernandez, 2006) and according to this metric, ancestor of *Frankia* evolved at about 350 MY ago from a group of soil actinomycetes. During this period the first trace of land plant was found. This ancestors of *Frankia* underwent a second radiation about 100MY ago giving birth of *Frankia*. During this period the first dicotyledonous plant families started to appear in the earth. This estimation that is very close to the fact that *Frankia* clusters emerged at 100-200MY ago, which corresponds well to the appearance of the oldest actinorhizal plant genera like *Myrica* and *Alnus* (Normand and Fernandez, 2006).

Compared to rhizobia, it is found that *Frankia* has relaxed host specificity. *Frankia* has very broad host range and restricted in clade-to-clade interaction. The plants belong to Hammamelidae clade is nodulated by the *Frankia* strains belong to clade-I. Elaeagnaceae and Rhamnaceae clade is nodulated by clade II *Frankia* and clade III *Frankia* nodulate the Rosaceae plants. So in conclusion, it can be said that this specificity largely resides on a group of signal molecules or a group of molecule that have same chemical backbone as well as common characters. The primitive actinorhizal plants, namely *Myrica* and *Gymnostoma* produce nodules when infected by both Clade I and Clade II *Frankia*. This is a notable exception and this observation established primitiveness of these plants and on this basis they can also be considered as the most primitive actinorhizae. A probable explanation of this infection lies in the habitat of *Gymnostoma* and *Casuarina*. *Casuarina* is native to drier Australian continent while *Gymnostoma* diversified in the wetter Melanesia region from the tertiary era beginning 65 MY ago.

2.3. The *Frankia*

The root nodules of non-leguminous plants invite attentions of researchers from early nineteenth century. Meyen (1829) first speculate on the nature of the cause of nodule and stated that it is a result of parasitic infection in the root. Woronin (1866) studied the detailed anatomy of the nodules and found some round vesicle and hyphae like structure within it. He found that the hyphae were passing through the intracellular region. He also noticed that there are some hyphae with round vesicular swelling tips. Woronin

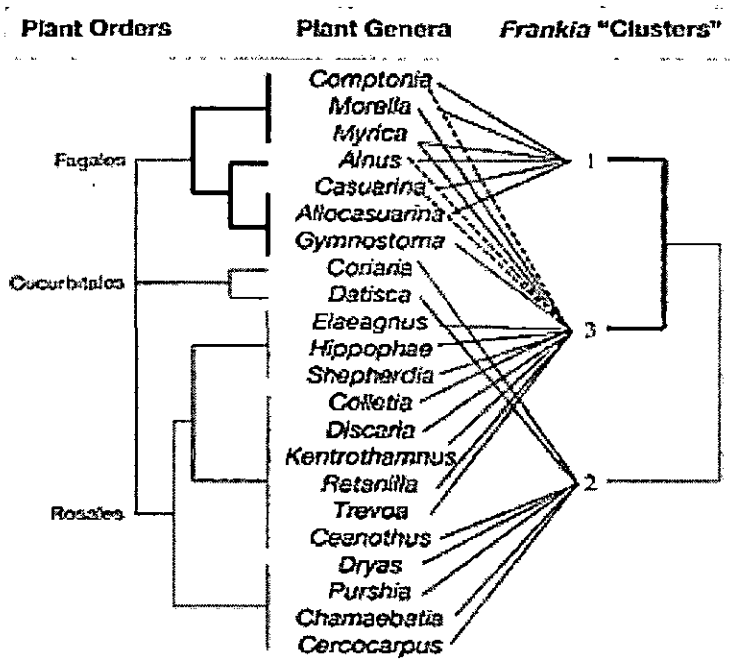


Figure 2.6. Correspondence between the plant genera RbcL (left) and *Frankia* 16S rRNA (right) phylogenies. Heavy lines - infection via root hair. Lighter lines - root penetration. ((Clawson *et al.*, 2004);web.uconn.edu/mcbstaff/benson/Frankia/PhylogenyFrankia.htm). The oblique lines join the plant and bacterial taxa when they can establish symbiotic structures, dashed lines indicate laboratory only symbioses not found in nature. (photo curtsey: Prof. Phillippe Normand, University of Lyon1, France).

considered the endophyte as a fungus and coined the name *Schinzia alni* as this unknown organism showed resemblance with a fungal parasite called *Schinzia cellulicolia* (Sen, 1996). Brunchorst, (1886) studied the cytological difference of leguminous and non-leguminous roots and named the endophyte as *Frankia subtilis* to honor his teacher eminent Swiss microbiologist A.B. Frank (1885), who, ironically did not believe in the presence of living microorganism in any kind of nodules and considered the structures as protein granules. Later Frank changed his idea on the nature of endoiphyte and along with Brunchorst considers *Frankia* as a fungus. On the other hand, the name *Frankia alni* was also coined by Von Tubeuf (1895) as a tribute to A.B. Frank. Several synonyms of *Frankia* have been proposed later on, which includes *Plasmodiophora alni* (Woronin) (Möller 1885), *Frankiella alni* (Woronin) (Maire and Tison 1909), *Aktinomyces alni* (Peklo 1910), *Actinomyces alni* (von Plotho 1941), *Nocardia alni* (Waksman 1941), *Proactinomyces alni* (Krassil'nikov,

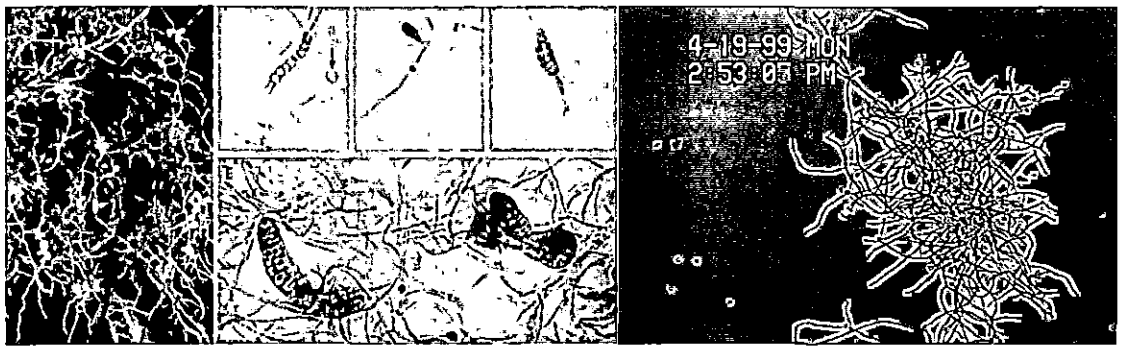
1949, 1959), *Streptomyces alni* (Fiuczek, 1959) (Normand & Fernandez, 2006).

Hiltner (1898) for the first time identified the endophyte as member of actinomycetes which is a close alley of *Streptomyces* while studying the roots of *Alnus* and *Elaeagnus*. (http:// www.mcdb.ucla.edu/Research/Hirsch/ images b / HistoryDiscoveryN2fixingOrganisms.pdf; Quispel, 1990). Hellriegel and Wilfarth published two papers during 1886-1888 on the fixation of atmospheric nitrogen by leguminous nodules with the help of the bacteria residing in the cortical layer. In their paper they introduced two terminology 'nitrogen user' and 'nitrogen accumulator' and showed differences between them. They identified Alders as nitrogen accumulator, and undoubtedly established source of nitrogen for growth of plants. Their work opened up a new avenue of research in plant microbial science (Bottomley, 1912; Quispel, 1988; Sen, 1996).

Hiltner in his later work with Alders demonstrated that young Alders can not survive in nitrogen free soils without root nodules. This result pointed out that though Alders are nitrogen accumulator they are certainly not the nitrogen fixer on the soil and another organism must be involved in the process of nitrogen fixation.

In another land mark study, Beijerinck isolated bacteria from legume root nodules which failed to infect non leguminous nitrogen accumulator and other plants proved that leguminous microsymbionts and non-leguminous microsymbionts are two different organism. (Brewin, 2002; Pawlowski and Bisseling,1996). Finally Krebber identified non-leguminous microsymbiont as actinomycetes in 1932 (Quispel, 1990).

Pommer (1956) was probably the first worker who was successful to isolate a slow-growing actinobacteria from nodules of *Alnus glutinosa* (Fig:2.7) which had unique morphological features like hyphae, multilocular sporangia and vesicles in pure culture. About 0.6mm colonies of that bacterium were obtained after 2-3 weeks of growth in glucose-asparagin agar described by Waksman for actinomycetes (1950). Unfortunately this strain was lost before independent studies in different laboratories. In the year 1964, electron microscopic study of root nodules of *Alnus glutinosa* and *Myrica cerifera* reestablished that these



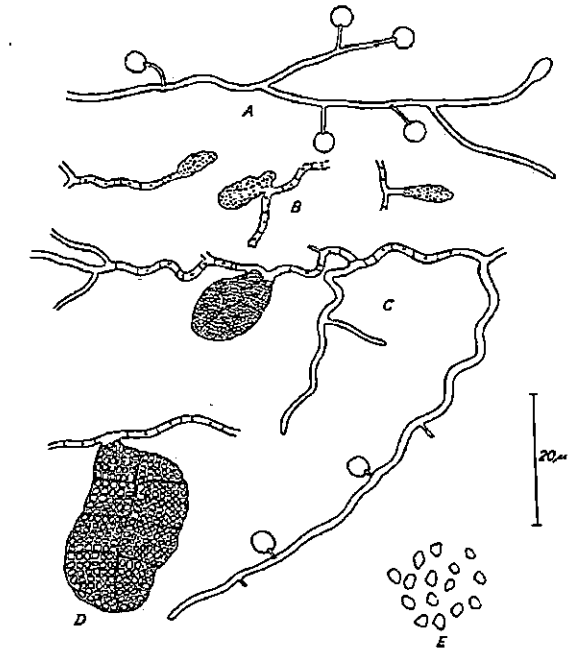


Fig 2.7 Some of the early photographs of Frankia. (photos provided by Prof. Phillippe Normand, University of Lyon1, France and Prof Louis S Tisa, University of New Hampshire, USA

Abb. 1. Der Wurzelknöllchen-Endophyt von *Alnus glutinosa* auf Glukose-Asparagin-Agar. Erläuterungen im Text.

root nodules were inhabited by actinomycetes bacteria as the electron microscope found prokaryotic structure in them (Becking *et al.*, 1964; Silver, 1964). From then efforts were on to isolate this unknown prokaryotic endophyte in pure culture. At last Callaham *et al* (1978) first isolated *Frankia* from *Comptonia peregrina* in pure culture followed by other workers (Diem and Dommergues, 1983; Diem *et al.,* 1983 and Sarma *et al.,* 1998). Lalonde (1978) worked on *Frankia* strain CpI1 isolated by *Callaham et al* (1978) and proved its ability to reinfect the host plant and established its symbiotic nature. The morphological nature of the bacterium isolated by *Callaham et al* (1978) shows unexpected resemblance with the lost strain of Pommer which he isolated from *Alnus glutinosa* nodules in 1959 (Benson and Silvester, 1993).

Because of slow growth of *Frankia* and frequent contamination by other fast growing organisms, fungi and other

actinomycetes are often mistaken as *Frankia*. To overcome this problem Lechevalier and Lechevalier, (1984) proposed following definition of *Frankia*:

"Actinomycetic, nitrogen fixing, nodule forming endophytes or endoparasite that have grown in pure culture *in vitro* and that:

a. induce effective or ineffective nodules in a host plant and may be re-isolated from within the nodules of that plant, and

b. produce sporangia containing non-motile spores in submerged liquid culture, and may also form vesicles.

c. free living actinomycetes having no known nodule forming or nitrogen fixing capacity, but that show the morphology described above."

Frankia produces three different kinds of cellular structure during its growth in pure culture or in symbiotic condition i.e. vegetative hyphae, multilocular sporangia containing spores and vesicle (Myrold, 1994, Benson and Silvester, 1993; and Akkermans and Hirsch, 1997). However, *Casuarina* infecting *Frankia* does not produce vesicle *in planta* (Sen, 1996).

2.3.1 Parts of Frankia

A. Hyphae

The actively growing cell type of *Frankia* is hyphae (0.5-2μ diameter) which form a mycelial mat by repeated branching. The vesicles and sporangia are produced from hyphae by differentiation. Aerial hyphae are absent on solid medium. The structure of *Frankia* hyphae, both *in vitro* and in symbiosis, was extensively reviewed by Newcomb and Wood (1987). Free-living organisms under the light microscope show branched hyphae. Under phase contrast and dark-field microscopes a large number of bright areas are seen within it (Benson and Silvester, 1993). They have type III cell wall and type I phospholipid content (Lechevalier *et al.*, 1982). Cell wall in chemically fixed materials appears to be composed of electron dense base layer (Lalonde, 1979; Baker *et al.*, 1980; Horriere *et al.*, 1983; Lancelle *et al.*, 1985). Cross walls originate from the base layer. A membranous layer some time may be visualized outside the outer wall layer (Newcomb *et al.*, 1979; Horriere *et al.*, 1983; Newcomb and Wood, 1987) and the cell wall, that may also contain spherical or

ovoid inclusions which can be visualized in both chemically fixed and freeze-substituted materials (Lancelle *et al.*, 1985). Internally hyphae cells contain numerous rosette-shaped glycogen granules (Benson and Eveleigh, 1979; Lancelle *et al.*, 1985) and lipid droplets. Individual ribosomes and polyribosomes are seen in freeze-substituted hyphal cells cytoplasm as relatively large (300nm) globular bodies. A large number of cytoplasmic tubules are present at the periphery of the freeze-substituted hyphal cell cytoplasm. They are circular in cross section (45 nm in diameter). These structures underlie the cell membrane both at cell septum and at the outside wall (Benson and Silvester, 1993; Lancelle *et al.*, 1985). An extracellular multilayered envelope was first identified in hyphae of free-living *Frankia alni* HFPCpII (Newcomb *et al.*, 1979). This structure is also present in vesicle of *Frankia* and this multilayered envelope has been identified in symbiotic hyphae of *Frankia* (Lalonde *et al.*, 1976; Berg and McDowell, 1987; Abeysekera *et al.*, 1990). This indicates that lipid enveloped hyphae may be common among the *Frankia* (Benson and Silvester, 1993).

B. Vesicles

Vesicles are the most definitive and characteristic structure of *Frankia* which is a unique developmental structure designed for physiological compartmentation and is totally absent in any other prokaryotic group (Benson and Silvester, 1993). In pure culture, vesicles are spherical, thick-walled and short stalked structure (2-6μ diameter) and *in planta*, they have variable shapes and their shapes are determined by their host. The multilaminate thick wall of the vesicle acts as a barrier to oxygen and protects the nitrogenase enzyme from toxic effect of oxygen (Parson *et al.*, 1987; Silvester *et al.*, 1990). These vesicles are the site of nitrogen fixation in free-living and symbiotic *Frankia* (Tisa and Ensign, 1987; Huss-Danell and Bergmann, 1990). The vesicles show a characteristic metabolism (Tisa and Ensign, 1988; Tisa, 1998). Although spores form the main mode of propagation, the vesicles can also give rise to vegetative hyphae (Schultz and Benson, 1989). They are lipid encapsulated, born either terminally or laterally to hyphae by a short encapsulated stalk and generally produced in nitrogen deficient medium

(Benson and Silvester, 1993).

Vesicles developed in nitrogen free medium first as a terminal swelling of hyphae or on short side branches. These early structures are separated by a septum near the base and are termed provesicle, which rapidly develop in to mature vesicles (2-4 μ in diameter) (Fontaine *et al.*, 1984). Provesicles are spherical cells, with dense cytoplasm (1.5-2μ in diameter), and they may show the initiation of internal compartmentation. The provesicles are unable to convert molecular nitrogen into ammonia nitrogen (Fontaine *et al.*, 1984). However some workers have reported the formation of vesicle in presence of available nitrogen in the medium (Gauthier *et al.*, 1981; Meesters *et al.*, 1987). Vesicles are observed as a bright structure under phase contrast microscope (Fontaine *et al.*, 1984) and also shows birefringence under polarized light microscope (Torrey and Callaham,1982), and can be seen as a bright halo under dark field microscope (Parsons *et al.*, 1987). The vesicle envelope and internal septations are the necessary prerequisite for nitrogenase activity and identified as two important and characteristic structural elements of mature vesicle in *Frankia* cells growing in culture, but neither of them appears essential in symbiosis (Benson and Silvester, 1993).

C. Sporangia and spores

Frankia strains are readily recognized by production of multilocular sporangia located either terminally or in an intercalary position (Newcomb *et al.*, 1979) on the hyphae in liquid culture. This morphology differentiates *Frankia* from other actinomycetes (Myrold, 1994). Sporangia (10 x 30-40 μ) contain multiple ovoid spores and are generally produced in stationary phase of growth (Myrold, 1994). They are also formed inside nodules of some actinorhizal plants (Schwintzer, 1990)

The *Frankia* sporangia were first identified in the nodules of *Alnus glutinosa* (Van Dijk and Merkus, 1976). They remain infective in dry soil for a long period (Tortosa and Cusato, 1991). On the basis of presence or absence of sporangia within a root nodule, *Frankia* strains have been classified as either spore$^+$ or spore$^-$ (Schwintzer, 1990) (Fig 2.8). Spore$^+$ strains appear to be much more infective than spore$^-$ strains (Normand and Lalonde, 1982), both spore$^+$ or spore$^-$ strains have been characterized at the molecular level (Simonet *et al.*, 1994). Filaments,

vesicles and sporangia have the infection potential; however the spores are the best material to infect. They are the major means for *Frankia* propagation in nature (Wall, 2000).

Chemically fixed developing spores show an electron translucent nucleoid region with dispersed fibrils, like the hyphae residue of a laminate envelope and numerous lipid droplets (Newcomb *et al.*, 1979). However the nucloid region is not evident in freeze-substituted preparation. The cytoplasm of the developing sporangia dispersed like that of the hyphae. The tubules are absent in the developed spores (Newcomb *et al.*, 1979). The sporangia developed by hyphal thickening followed by segmentation by septa originating from the inner layer of a double layered sporangial cell wall (Horriere *et al.*, 1983). This type of sporogenesis has been termed as enterothalic sporogenesis (Locci and Sharples, 1984). The immature spores are densely packed and are arranged in the middle portion of the sporangium, where as the mature spores are arranged in the peripheral region of the sporangia (Benson and Silvester, 1993). Mature spores are nonmotile, spherical or ovoid and measure about 1 -5μ in diameter (Benson and Silvester, 1993) (Fig:2.8).

2.3.2 Growth and physiology of *Frankia*

Frankia grow as filamentous colony on agar plates and they are generally grown in liquid culture In static batch culture the bacterium grow as submerged colonies with out any floating or aerial growth. They are obligate heterotrophic microaerophilic bacterium. *Frankia* is a slow growing bacterium with a doubling time of about 15hr or more in culture (Benson and Silvester, 1993); however their growth in symbiotic condition is seems to be unrestricted, since timing of root infection, nodule development, and host cell infection are similar to that of rhizobia-legume system. Thus difficulties to grow the bacterium in culture or isolating the bacterium are due to lack of knowledge in their growth conditions and growth requirements (Wall, 2000).

Some *Frankia* strains may produce red, yellow, orange, pink brown, greenish and black pigments in the medium. However, their pigment production depends on the strain, medium used and the age of the culture (Lechevalier and Lechevalier, 1984). A common red pigment has been identified as 2-methy

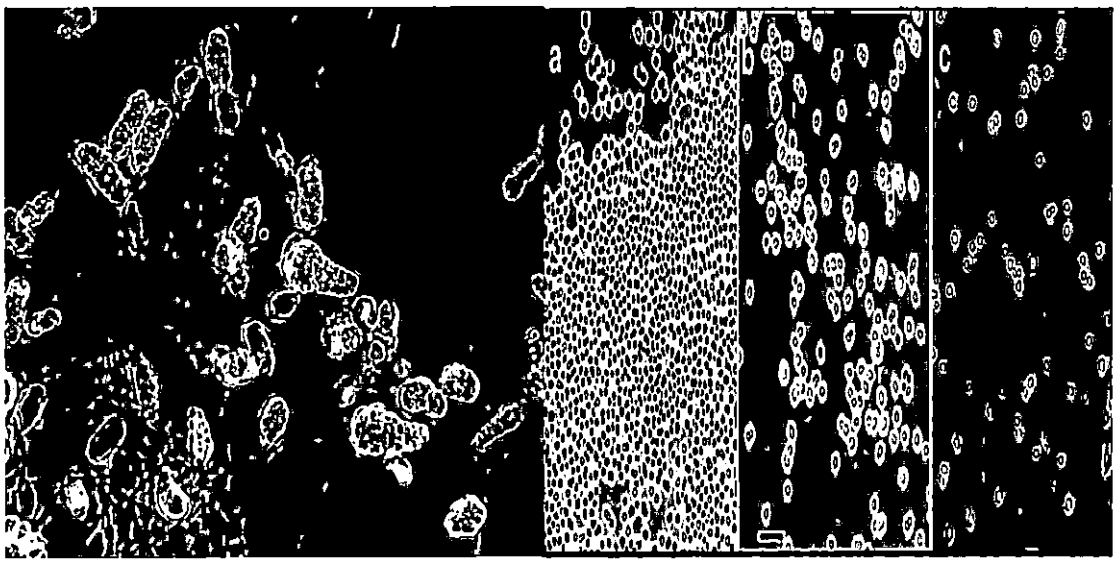


Fig 2.8 Micrograph of *Frankia* spores and sporangia. (Photographs are kindly provided by Prof Louis S Tisa, University of New Hampshire, USA)

1-4, 7, 9, 12, tetrahydroxy-5, 6-dihydrobenzo [a] naphthacene-8, 13-dione, a carboxylated derivative and a characteristic sugar 2-*O*-methyl-D-mannose (Mort *et al.*, 1983) have been identified in *Frankia* (Gerber and Lechevalier, 1984). Since pigmentation is highly variable, it is a poor criterion for placement in to genus but may also prove useful as a phenotypic trait for species recognition (Benson and Silvester, 1993).

Lechevalier divided the *Frankia* in two physiological groups (Lechevalier *et al.*, 1983; Lechevalier and Ruan, 1984); namely Gr-A and Gr-B. The Gr-A *Frankia* is more saprophytic, more aerobic and can be maintained in the slants. They show comparative rapid growth and diversity on serological and genetical basis. They also show protease and amylase activity and utilize various carbon sources which include arabinose, glucose, maltose, sucrose, trehalose, xylose, acetate, propionate, pyruvate and succinate. The Gr-B *Frankia* are more symbiotic and physiologically less active.

2.3.3 Heavy Metal Resistance of *Frankia*

Heavy metals are essential trace elements as they act as a cofactor of many enzymes and essential for growth in plants and microbes, but generally they are toxic at higher concentrations to all forms of life (Silver and Phung, 1996). The heavy metal resistance pattern arose much earlier than evolution of human beings on earth. This system arose soon after the life began in the world already polluted by volcanic activities and other geographical sources (Silver and

Phung, 1996). As this system has evolved over a long time so it attained a good diversity. (Bose *et al.*, 2007). These heavy metal resistance systems are important survival strategy for the bacteria. Bacterial chromosomes and plasmids contain genes for resistance to metal ions. They are frequently generating strong reactive oxygen species (ROS) and directly or indirectly causing gene mutations or protein damage. (Spain and Alm, 2003). More than fifty elements are available in the periodic table, which are designated as heavy metals. But the microbiologists are interested to only 17 of them. This is because rests are either not available or available in trace to the biological system. Weast (1984) estimated that elements at an average concentration lower than 1 nM are very un-likely to have any useful or toxic effect in biological system, and a resistance gene will not evolved for these metals. Tri- or tetravalent cations have very low solubility in water and because of the low solubility they have no biological significance. The remaining 17 heavy metals are important for study and grouped into three subgroups. The first subgroup contains Fe, Mo and Mn, which are important trace elements with low toxicity. The second subgroup contains Zn, Ni, Cu, V, Co, W and Cr, which are generally toxic but acts also as trace elements. Third subgroup contain well known toxic Ag, Sb, Cd, Hg, Pb and U ions which have little or no trace elemental activity (Nies, 1999).

The importance of study of heavy metal resistance pattern is three fold, firstly they acts as a genetic marker, secondly in bio-mining of expensive metals and thirdly heavy-metal-resistant bacteria may be used for bioremeditation of heavy metal contaminated environments (Nice, 1999).

The mechanisms of heavy metal resistance are generally resides in efflux pumping and enzymatic detoxification of metal ions. Occasionally, bio-accumulation or sequestration is a mechanism of resistance. Efflux pumps are the major currently known group of resistance systems with both plasmid and chromosomal systems. They can be either ATPase (e.g., the Cadmium and Copper ATPase of gram positive and Arsenite ATPase of plasmids of gram negative bacteria) or chemiosmotic (e.g., the divalent cation efflux systems of soil Alcaligenes and the Arsenite efflux system of the chromosome of

the gram negative bacteria and of plasmid of gram positive bacteria) (Silver and Phung, 1996).

In *Frankia* research the use of heavy metal resistance is very limited and focused on into genetic marker study only as there is a paucity of genetic marker in *Frankia*. In order to determine a genetic marker in *Frankia* system beside heavy metal resistance pattern (Richards *et al.*, 2002, Bose and Sen , 2006, 2007) and antibiotic resistance pattern (Tisa *et al.*, 1999) of *Frankia* have also been determined. *Frankia* strain ACN1^{AG}, Cc1.17, CcI3, CN3, CpI1-P, CpI1-S, DC12, EAN1pec, EI5c, EuI1c, EUN1f, and QA3 were studied for heavy metal resistance pattern. It is found that most of the *Frankia* strains are resistant to elevated levels of several heavy metal ions. The heavy metal resistance pattern is proved to be more effective than antibiotic resistance pattern due to the non degradable nature of heavy metal ions for long time of incubation (Richards *et al.*, 2002).

2.3.4 Taxonomy and Diversity of *Frankia*

Becking (1970) was for the first time tried to classify the family Frankiaceae with 10 species of unisolated symbionts in genus *Frankia* (Tab 2.3). His tentative classification was based on *in planta* morphology; vesicles morphology and he also considered this bacterium as an obligate symbiont. However many subsequent workers challenged Becking's view. After the successful isolation of *Frankia*, a good number of isolates were available for classical taxonomical approaches for classifying this bacterium. Lechevalier (Lechevalier and Lechevalier, 1984; Lechevalier and Ruan, 1984) proposed several taxonomic criteria for classification of *Frankia* which were based on ecology, infectivity, morphology, cell chemistry, physiology, serology, DNA homology and 16S rRNA profile. He proposed two *Frankia* types, "A" and "B". The type "A" correspond to Elaeagnaceae-infective (i.e. cluster 3) *Frankia* strains and the type "B" corresponds to *Alnus*-infective (cluster 1) strains. However this classification scheme did not meet satisfaction. Baker (1987) proposed a classification scheme (Fig 2.9) based on host infectivity which is as follows:-

- *Alnus* and *Myrica* infective group,
- *Casuarina* and *Myrica*, infective group
- Elaeagnaceae and *Myrica* infective group and
- Those infective only on *Elaeagnus*

Table 2.3. List of species of genus *Frankia* proposed by Becking (1970). (Normand and Fernandez, 2006) (Prof. Phillippe Normand, University of Lyon, France kindly provides this figure).

Plant Order	Plant family	Host plant genus	*Frankia* species	Vesicles (size,shape)
Fagales	Betulaceae	*Alnus*	*alni* (type)	3-8µm, spherical
Rhamnales	Elaeagnaceae	*Elaeagnus*	*elaeagni*	2-4µm, spherical
"	Rhamnaceae	*Ceanothus*	*ceanothi*	1.5-3µm, spherical
"	Discariaceae	*Discaria*	*discariae*	4µm, sphaerical
Myricales	Myricaceae	*Myrica*	*brunchorstii*	7.5-12.5x1.6-2.4µm, club
Casuarinales	Casuarinaceae	*Casuarina*	*casuarinae*	3-4x0.6-1.5µm, club
Coriariales	Coriariaceae	*Coriaria*	*coriariae*	9-12x1.2µm, club
Rosales	Rosaceae	*Dryas*	*dryadis*	1.5-5x1.5-2µm, club
"	"	*Purshia*	*purshiae*	ND
"	"	*Cercocarpus*	*cercocarpi*	ND

(Fig 2.9) .

This classification was based on the finding that *Myrica* is a promiscuous host, which was infected by most of the *Frankia*, with exception with *Myrica gale*, this species can be nodulated by only *Alnus*-infective strains (Huguet *et al.*, 2005).

Lalonde *et al.* (1988) proposed another taxonomical scheme based on the system proposed by Lechevalier, but he used the names *alni* and *elaeagni* to refer to host plants. He further subdivided *F. alni* into two species namely *pommerii* and *vandickyi*, which was based on the spore positive (sp^+) and spore negative (sp^-) characters. But this idea was later discarded due to uncertainties on spore characters and

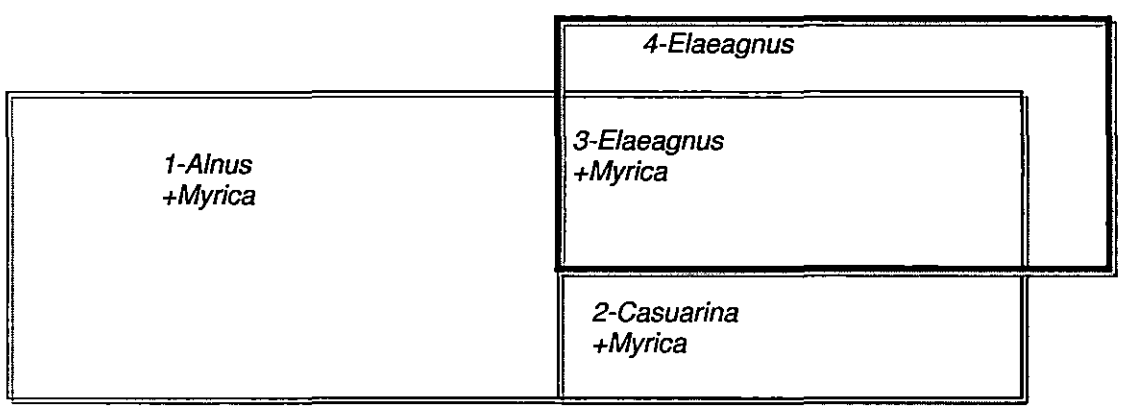


Fig 2.9 Schematic of strains grouping based on infection of host plants (Baker, 1987). (Prof. Phillippe Normand, University of Lyon, France kindly provides this picture). (Normand and Fernandez, 2006)

Table 2.4. Synopsis of strains tested for DNA homology (Fernandez *et al.*, 1989; Lumini *et al.*, 1996), and 16S sequence is available for phylogeny approaches. (Prof. Phillippe Normand, University of Lyon, France kindly provides this scheme). (Normand and Fernandez, 2006)

Species by Fernandez *et al.*, 1989	**Ref. strains**	**Other strains tested**	**16S available**	**Infectivity**
Genomic species 1	AcoN24d	CpI1, ArI3, $ACN1^{AG}$	ACN14a	*Alnus, Morella*
Genomic species 2	AV22C	AVN17o	AVN17s	"
Genomic species 3	$ARgP5^{Ag}$		$ARgP5^{Ag}$	"
Genomic species 4	Ea1-12	$HR27_{14}$, $Ea1_2$, $Ea2_6$, $Ea3_3$ $Ea1_2Ea33$	$Ea1_2$,	Elaeagnaceae, *Morella*
Genomic species 5	$TX31e^{HR}$	EAN1pec, HRX401a	EAN1pec	*Morella*
Genomic species 6	EUN1f			*Morella+Alnus*
Genomic species 7	HRN18a			*Morella+Alnus*
Genomic species 8	Ea50-1			*Morella*
Genomic species 10				*Morella+Alnus*
Genomic species 11				*Morella+Alnus*
Genomic species 12			SCN10a	*Morella*
Genomic species 9	CeD (ORS020606)	ORS020607, Cj1-82, AllI1	CeD, CcI3	*Casuarina*

absence of molecular evidence in favor of spore characters.

An *et al* published a series of papers on the use of DNA homology technique in classification of *Frankia* (An *et al.*, 1983; An *et al.*, 1985; An *et al.*, 1987;). This bacterium has a high G+C content of 68-72% (An *et al.*, 1983). Prof. An examined 19 isolates and proposed 1 species with 9 *Alnus*-infective isolates, 1 species with a single *Elaeagnus*-infective isolate and several unclustered isolates. The *Alnus*-infective isolates belong to the cluster 1 and the *Elaeagnus*-infective isolate belongs to the cluster 3 of this bacterium. Later Fernandez *et al.* (1989) worked in the same way with 43 isolates and proposed 9 *Frankia* species among them, in which 3

belongs to the group of *Alnus*-infective (i.e. belongs to cluster 1), 5 belongs to the group of *Elaeagnus*-infective (i.e. belongs to cluster 3) and the last one in the group of *Casuarina*-infective (also in cluster 1) strains (Tab 2.4). Dobritsa and her co-workers (Akimov *et al.*, 1991; Akimov and Dobritsa, 1992;) partially extended the study by Fernandez *et al.* and found 5 genospecies in the *Alnus* infective strains and four *Frankia* genospecies on *Elaeagnus* infective strains. However, in both the group tested few strains and therefore it is impossible to conclude if these numbers add to those previously described or not. The exception of this case is the *Frankia alni* strains CpI1 and ArI3. Lumini *et al* (1996) worked on DNA homology further with other *Elaeagnus*-infective strains, particularly with those strains that can infect both of *Elaeagnus* and *Alnus*, and they described three further species. These species have no pure culture existence and their differentiations are based only on DNA hybridization. They are called genomospecies or genospecies (Normand and Fernandez, 2006). Host infectivity or host species of origin are studied in this species. The *Casuarina* infective strains are closely homologous as evident by DNA homology values (above 70%; Fernandez *et al.*, 1989) and have a very conserved PCR-RFLP patterns in the *rrn* and *nif* operons (Rouvier *et al.*, 1992, 1996). When a *Frankia* strain is isolated from the *C. equisetifolia* out side Australia, the strains showed variations. Thus it can be concluded from the above discussion that there are about 3 to 7 or more genomospecies infective on *Alnus* and *Myrica* and from 8 to 12 or more genomospecies infective on Elaeagnaceae and *Gymnostoma*. So, a minimum of 12 species are available in isolates under the genus *Frankia*. So, it is possible to that there exists over hundred species of *Frankia* which are distributed in the four clusters.

Isolation of *Frankia* from Rosaceous members or from *Datisca* or from *Coriaria* is very difficult. According to Becking (1970) there were at least three *Frankia* species present in Rosaceous plants. Mirza *et al.* (1994) first showed that the *Frankia* isolated from *Datisca* could nodulate *Coriaria*, and both have the similar 16S rRNA sequence. He presumed that these bacterium infecting both *Datisca* and *Coriaria* may be the same bacterium. Bosco *et al.* (1994) on the basis of 16S

rRNA sequence concluded that the symbiont isolated from *Dryas* (a member of the family Rosaceae) are close to those of *Datisca* and *Coriaria* symbionts. These plants can also be nodulated by cluster 3 genomospecies under laboratory conditions. These cluster 3 strains are cosmopolitan and can cross infect with other actinorhizal plants while other strains are selective.

The prime obstacle of *Frankia* research lies in its difficulties in the isolation. With invention of polymerase chain reaction (PCR) technology (Mullis *et al.*, 1986) identification of *Frankia* become easier. First PCR amplification of *Frankia* was done with universal primer FGPS849 and FGPS1176 and the size of amplicon was 325bp (Nazaret *et al.*, 1991). The first 16s rDNA specific primers (FGPS989ac and FGPS989e) targeting the helix 31 of domain III of 16s rRNA gene was developed by Bosco *et al.* (1992). According to Misra and Verghese (2004) the ITS region of *rrn* genes shows good diversity among *Frankia* strains. Simonet *et al* (1991) and Mirza *et al* (1994) constructed primers for that region and successfully amplified the ITS region of *rrn* gene of *Frankia.* Normand *et al* (1994) studied the structure of *rrn* operon in *Frankia* and compared with corresponding gene of *Streptomyces* for taxonomic purpose. The length of 16s rDNA is 1513nt, 23s rDNA is 3099nt and 5s rDNA is 200nt in *Frankia.* The 16s rDNA region of *Frankia* is highly conserved (Misra and Verghese, 2004)

Hahn *et al* (1989) were the first to use 16S rRNA sequences of *Frankia* to investigate the phylogenetic relationship of *Frankia* with its neighbors. On the basis of their study it was revealed that *Frankia* was close to *Geodermatophilus*, a dry soil actinomycetes, in possessing similar multilocular sporangia and to *Blastococcus*, a sea microbe. But there is little description about the genus *Blastococcus* (Normand and Fernandez, 2006). Marechal *et al* (2000), based on their studies on, *recA* marker on this regard found an interesting finding that a thermal spring microbe, *Acidothermus cellulolyticus* is the closest phyletic neighbor of *Frankia* rather than *Geodermatophilus* . Recent studies also confirm that *Frankia* can be subdivided into four clusters (Fig .2.10 and 2.11)

Frankia is world wide distributed and faces various physiochemical soil properties like soil pH, aeration, moisture etc. Even this bacterium may

persist in soil in absence of the host plants (Huss-Danell and Frej, 1986; Smolander and Sundman, 1987). It has been found that actinorhizal plants gets nodulated far away from their original habitat. This indicates the presence of that particular *Frankia* in that area (Dawson *et al.*, 1989; Smolander, 1990; Jamann *et al.*, 1992, Huguet *et al.*, 2004, McCray Batzl *et al.*, 2004; Normand and Fernandez, 2006).

Jamann *et al.* (1992, 1993) working on *nifD-nifK* intergenic region of *Frankia* from *Elaeagnus* shows that the diversity which is measured by the Shannon-Weaver diversity index differ in accordance with soil pH. Soil pollutant also has a role in *Frankia* diversity, which was assessed by Ridgway *et al.* (2004) in *Frankia* strains associated with *Alnus incana.* It was found that different metal concentration play a major role in distribution of *Frankia* associated with *Gynmostoma* (Navarro *et al.*, 1999).

The recent whole genome sequencing data of the three strains of these bacterium shows that there are huge variation in the genome size (Fig 2.12). The *Frankia* strain EAN1pec has a genome size of 9,035,218,7976bp while the *Frankia* strain CcI3 has genome size of about 5,433,6284499bp. The *Frankia* strain ACN 14a has intermediate genome size of about 7,497,9346786bp (Normand *et al.*, 2007). Normand *et al.* (2007) found 2,810 genes common in the genome of CcI3, ACN14a and EAN1pec strain of *Frankia.* So it can be stated that the diversity of *Frankia* strain lies in their variable genome size. So the conclusion on the diversity of *Frankia* can be drawn from the word of Normand "The full genetic diversity of *Frankia* may not yet have been revealed due to limits brought about by the approaches used. Indeed, the time-consuming isolation step not only strongly limits sampling, it also induces biases, under representing non- or poorly nodulating strains and selecting those isolates that are easily culturable. A previous facilitating but bias-inducing step often seen is the use of an intermediate host inoculated in the laboratory by crushed nodules from the field. Indeed, it was reported that the genetic diversity observed in natural populations can be totally different from that observed following greenhouse inoculations with soil or field nodules (Huguet *et al.*, 2005). Direct detection has been used in several recent attempts to describe *Frankia* diversity in nodules or soil of

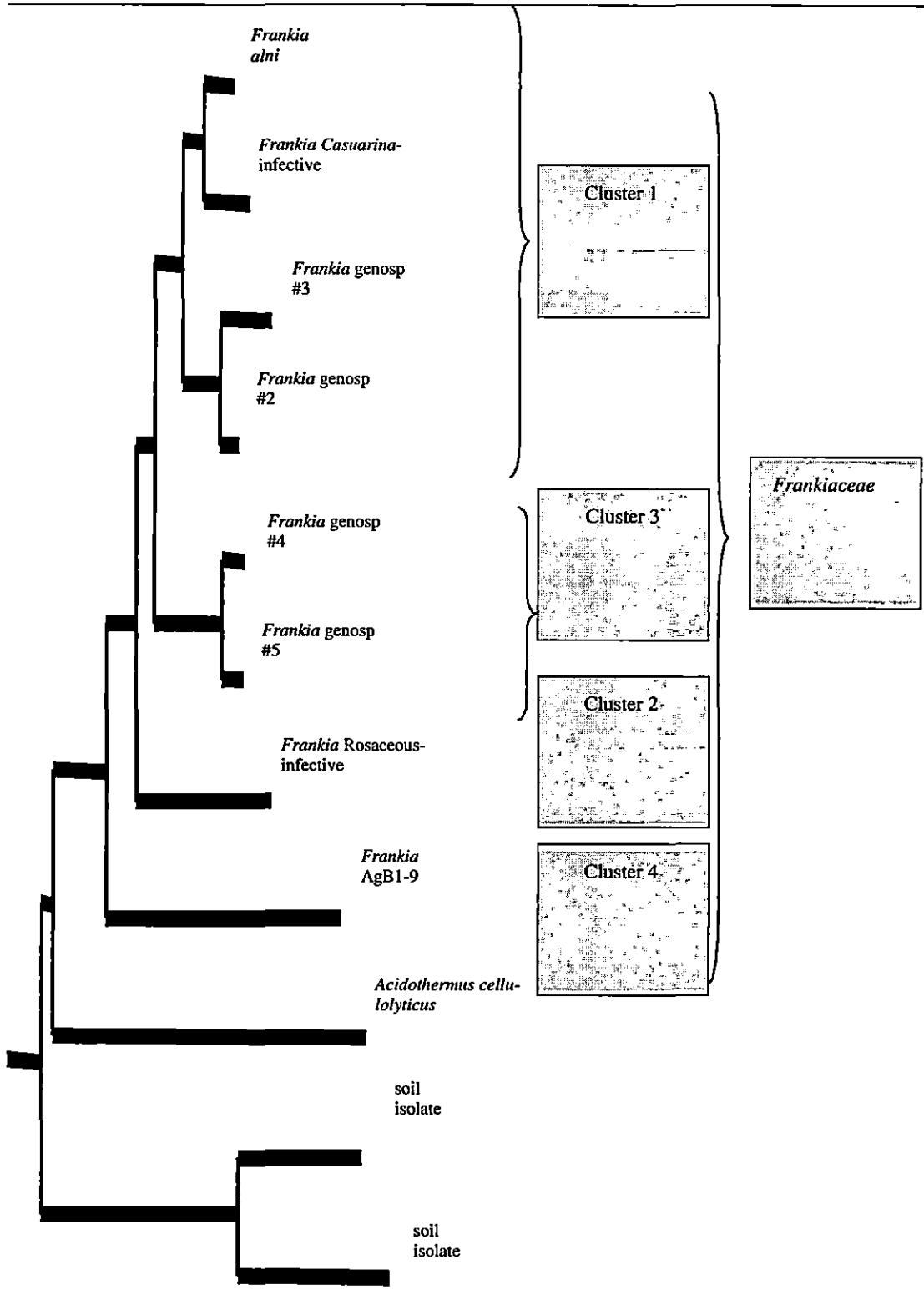


Fig 2.10: Phylogenetic tree (Saitou and Nei, 1987) of *Frankia* and closest neighbors comprising soil sequences. 1: *Alnus* and *Casuarina* infective strains; 2: *Rosaceaous-*, *Datiscaceae-* and *Coriariaceae*-unisolated symbionts; 3: *Elaeagnus*-infective strains; 4: a divergent group of ineffective or non-infective strains. (Normand and Fernandez, 2006)). ((Prof. Phillippe Normand, University of Lyon, France, kindly provides this dendogram)

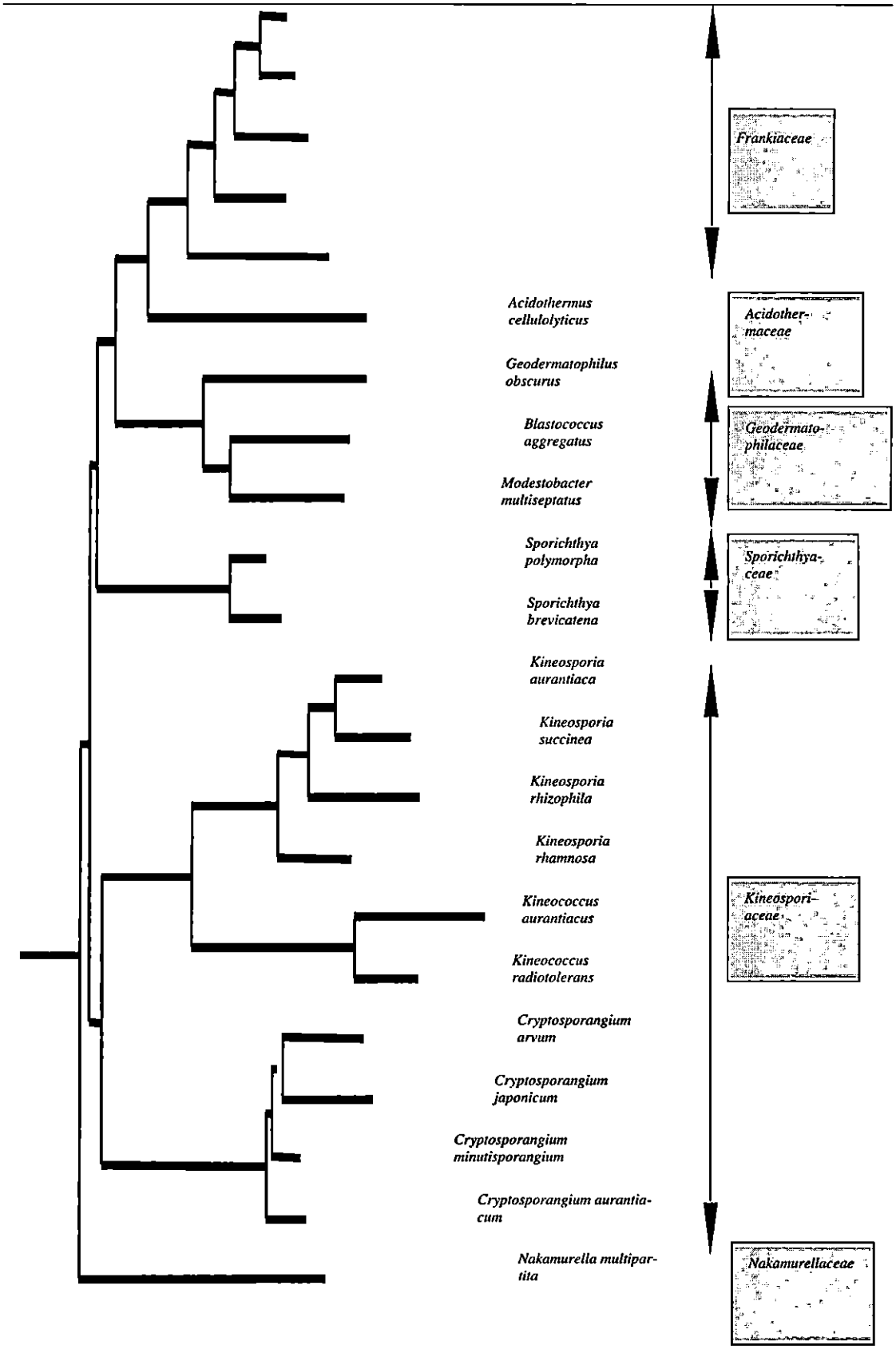


Fig 2.11: Phylogenetic tree of the Frankinae. by the Neighbor-Joining (Saitou and Nei, 1987) method including all validly described genera and species of suborder Frankinae (Prof. Phillippe Normand, University of Lyon, France kindly provides this dendogram). (Normand and Fernandez, 2006)

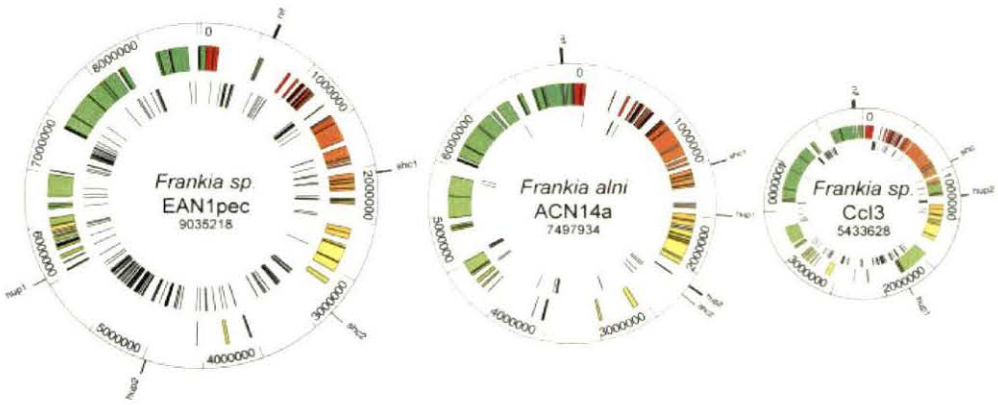


Fig 2.12 Genome maps of the three *Frankia* strains. Circles, from the outside in, show (1) gene regions related to symbiosis including *shc*1, *hup2*, *hup1*, and *nif*; (2) the coordinates in Mb beginning at 0 = oriC; (3) regions of synteny (syntons) calculated as a minimum of five contiguous genes present in all strains with an identity >30% over 80% of the length of the shortest gene (syntons are tagged with a spectrum-based [red-yellow-green] color code standardized on ACN to indicate regions where syntons have moved in the other strains); (4) IS elements and transposases. Circles were drawn using GenVision Software from DNAStar.(Adopted from 'Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography' Philippe Normand *et al*, 2007 Genome Res. 2007 January; 17(1): 7–15.

PCR-amplified DNA fragments, but these methods are often based on one or two conserved genes (generally *rrn* and *nif* genes) and cannot but fail to describe the full extent of biological diversity." (Normand and Fernandez, 2006). Thus according to the present taxonomic status of this bacterium belongs to the monogeneric family Frankiaceae (Normand *et al.,* 1996), which is a member of suborder Frankinae together with 5 other families namely Acidothermaceae, Geodermatophilaceae, Sporichthyaceae, Kineosporiaceae and Microsphaeraceae, and the suborder Frankineae is a member of the order Actinomycetales (Normand and Fernandez, 2006).

2.3.5 Bioinformatics of *Frankia*

Bioinformatics is essentially a marriage between biology, information technology and computer science. In the present era, bioinformatic analysis is virtually required in every aspect of bioscience and *Frankia* research is no exception. Perhaps the first *Frankia* partial sequence submitted to the public domain was a 243bp long sequence of rDNA (GenBank accession no S70679.1) done by Lyon group (Nick *et al.,* 1992). Since then on, several sequences have been submitted in various DNA depositories

like GenBank, EMBI database, DDBJ, etc. The first DNA sequence which appear from India was done by AK Misra group (GenBank: S70135.1).

Two groups came forward for *Frankia* whole genome sequencing project. These are LS Tisa and D. Benson from USA, who have done the sequencing of *Frankia* strains CcI3 and EAN1pec in collaboration with DOE, JGI, USA and P Normand, who did the sequencing of *Frankia* strain ACN14a with Genoscope, France. With the advent of *Frankia* genome project various other workers started working on bioinformatics of *Frankia*. Prominent among them are Sen and his group in University of North Bengal, Ventura *et al.* (2007) and McEwan and Gatherer (1999). McEwan and Gatherer (1999) first used codon indices to predict the expression of *nif* operon of *Frankia*. Sen and his group published a series of paper on bioinformatics of *Frankia* (Sur *et al.*, 2006, 2007; Bose *et al.*, 2007; and Sen *et al.*, 2008, 2010). They used certain parameters like GC content, GC3 content (G or C nucleotide in the third codon position), effective number of codon (Nc), relative synonymous codon usage (RSCU), Fop (frequency of optimal codon) and CAI (codon adaptation index) to investigate codon usage pattern of whole genome of *Frankia* with special reference to *nif* genes. Sur *et al.* (2006, 2007) studied the codon usage profiling and analyze the intergenic association of *Frankia* EuIK1 *nif* genes and compared the result with *Bradyrhizobium*. On the basis of codon usage data they concluded that *Frankia* strain show high level of codon preference. In another study (Sur *et al.*, 2007), this group found that the expression level of *nif* genes of *Frankia* are moderately high and show mutational bias. The life style pattern of various *Frankia* strains was studied using codon usage pattern analysis (Sen *et al.*, 2008). It was found that *in Frankia*, the highly expressed genes are more biased in comparison to the other protein coding genes. The analysis of COG profiles of the predicted highly expressed genes further highlighted the differences amongst the *Frankia* strains. The results obtained in their study further supported the hypothesis that *Frankia* CcI3 is slowly becoming a symbiotic specialist. Besides such sequence analysis studies, structurally analysis of proteins of *Frankia* have also been carried out. The three-dimensional structure of NifH protein of *Frankia*

was resolved using homology modeling technique (Sen *et al.*, 2010). The metal binding sites and important functional regions of this protein have also been determined. The structure provided valuable insight into the 3D framework and structure-function relation of this protein. Beside *nif* genes other genes of *Frankia* also got serious attention from the workers. Kosawang (2009) studied the regulation of hydrogenase using comparative genomics of *Frankia* strain CcI3, ACN14a and R43. *Frankia* strain R43 show an interesting feature in its hydrogenase protein. It works in bidirectional mode and evolved hydrogen which is absent in CcI3, ACN14a. In this decades the knowledge and technologies of structural bioinformatics, protrome and secretome research was generated (Sur *et al.*, 2010) which gave a boost in understanding the biological nitrogen fixation by *Frankia*. The proteome of *Frankia* was first studied by Alloisio *et al.* (2007) in both under nitrogen fixing and nitrogen depleted condition. This study reveals that total of 126 proteins is associated with nitrogen assimilation and oxidative defense system. Mastronunzio *et al.* (2008) studied the secretome of *Frankia* and unable to locate any polysaccharide degrading enzyme within it. They attributed property of secretome to a group of esterase, lipases and hydrolases which help in the process of nodulation. Niemann and Tisa (2008) found two truncated hemoglobin genes (*hboN* and *hboO*) in *Frankia* strain CcI3 genome. These genes help in the oxygen protection in *Frankia* (Misra and Sen, 2011). Mastronunzio and Benson (2010) used mass spectroscopy and two dimensional liquid chromatography in proteome research in *Frankia* and identified 1300 proteins of *Frankia* isolated from *Alnus incana*, 1,100 proteins from *Elaeagnus angustifolia* and 100 proteins from *Ceanothus americanus*. They found that the iron proteins are the most abundant proteins in *Frankia*

Recently Sen *et al.* (2011) explored the significance of rare TTA codon containing genes in *Frankia*. Amongst the high GC content genomes of Actinomycetes, *Frankia* genomes had the highest percentages of TTA containing genes. They have found that *Frankia* genomes retain large number of such genes in several important functional groups such as metabolism and cellular process. Many of these genes had ortholog in other *Frankia*

genomes. Their study on evolutionary significance and codon-anticodon interaction further point to the fact that so many of the *Frankia* genes hold on to this rare codon without compromising their translational efficiency.

The study of morphology and physiology of *Frankia* was started from the time of first isolation. In the last decade of the last century genetics and molecular biology based work of *Frankia* started and gathered a good chunk of information about this organism, which reached its peak with whole genome sequencing of the organism. In the recent past post sequencing work using bioinformatics tools started and *Frankia* research grew up. However, many opportunities of study have been left in this field, which invites serious attention of workers.