Chapter 2
Review of Literature

2.1. Actinorhizal plants:

Symbiosis is a mutual association between two different organisms living in close physical association typically for taking advantages from each other. These relationships are essential to many organisms and ecosystem to provide a balance that can only be achieved by working together. There are mainly two kinds of symbiotic relations. First one is mutualism, where both the symbiont and host benefit, and second one is commensalism, where symbiont benefits with little effect on the host. Two types of mutualism i.e. 1) Obligate: one organism cannot survive without other and 2) Facultative: each organism can survive independently but it benefits both when they remain together were found.

Symbiosis are observed between eukaryotes-eukaryotes, eukaryotes-prokaryotes and prokaryotes-prokaryotes system. Present study is confined to one such eukaryote-prokaryote facultative mutualism i.e. *Alnus*-*Frankia* actinorhizal symbiosis.

Actinobacteria *Frankia* develops root nodules in actinorhizal plants and fixes atmospheric nitrogen. An actinomycete *Frankia* and the location of their nitrogen fixation i.e. root nodules, is termed together as actinorhiza. A few angiosperms are able to have symbiosis with *Frankia* in terms of fix ing nitrogen by the modification of their root hair. this group of dicotyledonous plants that are nodulated by *Frankia* is called actinorhizal plants. Till date about 260 species 24 genera belonging to 8 families of 7 orders have been reported to have actinorhizal associatios (Froussart *et al*., 2016) (Figure 2.1).

Actinorhizal plants are found in different geographical location of the earth, that covers arctic Tundra (*Dryas* spp.) and alpine forest (*Alnus* sp., *Coriaria* sp. etc) to costal and xeric regions (*Casuarina* sp.). In Nepal, *Alnus nepalensis*, found at a wide altitudinal range, descending as low as

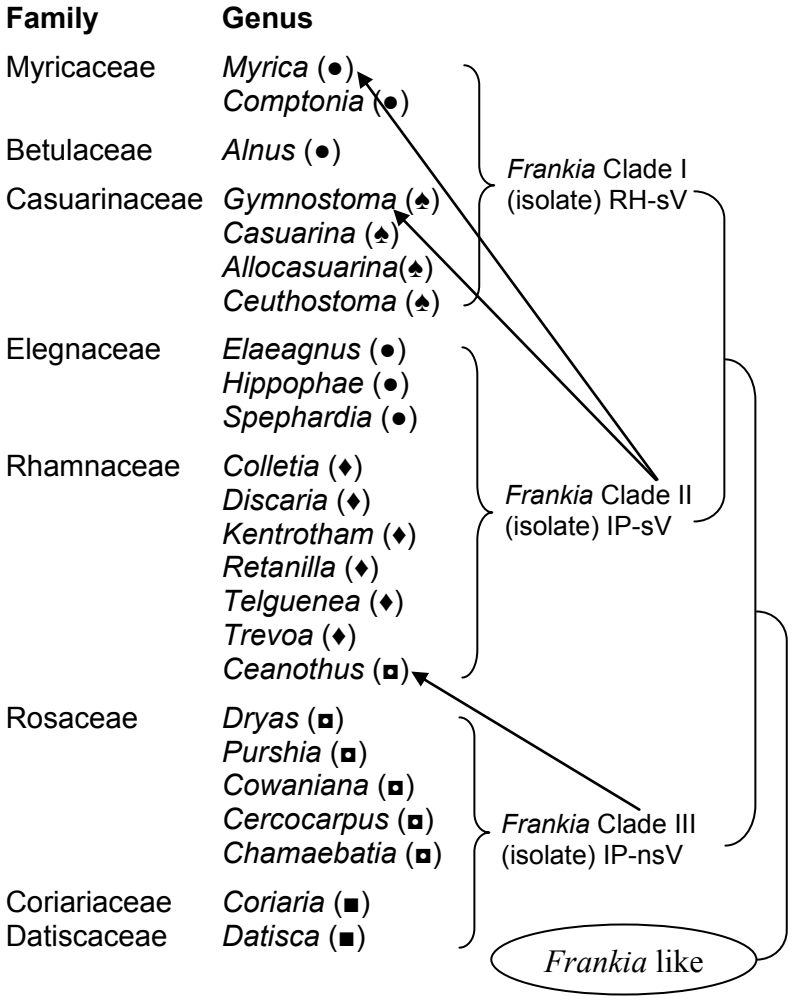

Figure 2.1: Phylogenetic grouping of actinorhizal plants (Symbol between brackets of plant genus indicated native geographical distribution – (●) to most continents; (♠) to Australia and Western Pacific; (♦) to South America and southern New Zealand; (◘) Western North America; (■) disjunction distribution in northern and southern temperate zones; RH – root hair infection; IP – intercellular penetration; sV – septate vescicles in nodule; nsV – nonseptated vesicles in nodule) (Benson and Clawson 2000, Jeong *et al*. 1999, Huss-Danell 1997, Saltis *et al*. 1995, Swensen and Mullin 1977, L G Wall 2000)

500 m, but most common from 900 m and above up to 2700m. At lower altitude, *A. nepalensis* mostly found in moist regions, such as riverside but also found abundant in rough rocky land exposed by landslides, and cultivable land (Sharma, 2012). It is liable to be damaged by browsing animals when young, but seedlings over 50 cm high are relatively immune.

A record regarding the presence of *A. nepalensis* in sub-himalayan West Bengal and Sikkim during the British period are available in the Bengal District Gazetteers by L S S O'MALLEY (1907).

Frankia in symbiosis with actinorhizal plants fix 240-350 kg N_2 $ha^{-1}y^{-1}$ range of molecular nitrogen, which is almost equivalent 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 economically important while used as timber, fire-wood, food, chemicals, etc (Benoit and Berry, 1990; Diem and Dommergues, 1990; Hibbs and Cromack, 1990; Myrold, 1994). In India, the actinorhizal plants are reported to be growing as dominant species under different environmental condition. The morphological characteristics, habit, habitat, distribution and possible ecological significance of actinorhizal plants found in Kumaun Himalayan region are studies and reported by Bargali (2011). Swensen, (1996) identified three major phylogenetic subgroups of actinorhizal plants. The first subgroup includes symbiotic taxa from the families Betulaceae, Myricaceae and Casuarinaceae. The second subgroup includes the families Datiscaceae and Coriariaceae from symbiotic taxa, and the third subgroup also includes symbiotic taxa from the families Rhamnaceae, Rosaceae and Elaeagnaceae (Swensen, 1996).

Morphology of nodules, formed by different actinorhizal plants differ considerably depending upon organization of infected cells in the cortex (i.e. symbiotic, nitrogen-fixing cells), oxygen protection mechanisms for nitrogenase, infection mechanisms, patterns of *Frankia* differentiation, and organization of carbon and nitrogen metabolism (Pawlowski and Bisseling, 1996; Berry *et al*., 2011; Schubert *et al*., 2011). The actinorhizal root nodules are conventionally formed through a series of interactions between the host actinorhizal plant and the symbiont ie. *Frankia*. The interaction takes place in the cortical cells of lateral root in actinorhizal plants and subsequently formed nodules. The actinorhizal root nodules are continuing structure, consists of multiple lobes. The lobes in *Alnus* are morphologically different from *Myrica* and *Ceanothus* lobes. *Myrica* and *Ceanothus* lobes are distinct but in *Alnus*, they are compactly crowded.

The perfect condition for root infection process and subsequent development of nodule is not well understood in *Frankia*-Actinorhizal system. The hyphae of *Frankia* are embedded in the mucilage layer of the root hair of actinorhizal plants and established the

infection precess. It is established that a single root-hair infection is good enough for development of nodule. However, more than one infection occurs simultaneously in the root-hair, for the nodule formation, in natural environment. It has been also well established that, under laboratory conditions the occurrence of nodulation is directly proportional with the amount of inoculums given (Newcomb and Wood, 1987). Knowlton *et al*., (1980) reported that, in actinorhizal plants, several non-symbiotic soil bacteria involve in root hairs deformation process. Those are called 'helper', which play an important role in the *Frankia*-actinorhizal interactions.

Under controlled condition, *Pseudomonas* sp. helps in the nodule formation in *Alnus* and *Casuarina* (Knowlton *et al*., 1980). The signal molecule of host plants allows same essential chemical changes and undergoes regulation of nitrogen fixing symbiosis with *Frankia*. This involves as a minimum of two different and successive signal molecules that leads to a root nodule formation mechanism in actinorhizal plants (Wall, 2000). Two different pathways are reported for infection of the host tissue. It takes place either by deforming the root hair as found in *Alnus*, *Casuarina*, *Myrica* etc (Callaham *et al*., 1979; Berry and Torry, 1983) or by intracellularly as found in case of *Ceanothus*, *Elaeagnus*, etc. Some strains can pursue both the pathways for successive infection (Miller and Baker, 1986).

In rhizobia–legume symbioses nodulation process established by help of NOD signal molecules and forms the basis of host specificity (Oldroyd *et al*., 2009). These kinds of NOD factors and the bioassay of reporter genes are reported to be absent In *Frankia* (Ceremonie *et al*., 1999). On the basis of the actinorhizal root hair deformation study (Gherbi *et al*., 2008; Markmann *et al*., 2008) and signal factors present in arbuscular mycorrhizal fungi (Maillet *et al*., 2011), it is predicted that a chitin-based molecule in *Frankia*, act equivalents to the NOD factor of rhizobia–legume symbiosis. Recently a LysM-type mycorhizal receptor was reported to be found in infection process of *Rhizobium* symbiosis which involve in of the nonlegume *Parasponia andersonii* (Den Camp *et al*., 2011). This finding suggested that *Frankia* might have gone through a series of

unique pathways of synthesis of a novel chitin-based signal molecule distinct from *Rhizobium*. These molecules allow independent infection mechanism and act as NOD factor alike to rhizobia–legume symbiosis (Pawlowski *et al*., 2011). In nodule *Frankia* follow a unique pathway, for nitrogen utilization and the primary nitrogen assimilation reported to be stored as arginine (Berry *et al*., 2011).

2.1.1. Evolution of actinorhizal plants by studying the facultative mutualism:

Frankia and actinorhizal symbiosis evolved almost at similar time. Two different hypotheses regarding the evolution of the actinorhizal symbiosis have been proposed. According to the first hypothesis, plant genera obtained the ability of making symbiotic relationship with *Frankia* to achieve certain selective advantages related to ecological niches. So, the divergence of families into their respective genera had taken place after acquisition of the symbiotic character. The second hypothesis stated that a group of plant was forced to make association with nitrogen fixing soil bacteria *Frankia*, as the available soil nitrogen was limited in the early cretaceous period (Bond, 1983) which subsequently survives in struggle for existence with other angiosperms mostly belonging to Magnoliidae. Some of these plants progressively evolved afterwards into the actinorhizal plants. Nitrogen fixing microbes fix nitrogen to the soil and increase the availability of nitrogen during the last 100M years. So, due to stable amount of availability of nitrogen in the soil, some of the ancient plants lost their symbiotic nitrogen fixing ability. However this primordial property remains active in some selective pioneer plants i.e. actinorhizal plants. The DNA hybridization study on actinorhizal plants supports the second hypothesis (Normand and Bousquet, 1989). The nitrogen fixation process in nodules might have evolved at a definite time in the earth's evolution (Doyle, 1998). The nitrogen fixing actinorhizal plants are grouped in Rosid I linages of the seed plants. This Rosid I group may be further subgrouped into four linages, of which three are actinorhizal and the fourth one is Fabaceae (Soltis *et al*., 1995). Geographical distribution, fossil records, anatomical and morphological studies extend supports to this hypothesis (Wall, 2000).

2.2. Biological Nitrogen Fixation:

Nitrogen is an important nutrient used by living organisms for continued their

existence. It is most commonly deficient element, responsible for reduced agricultural yields throughout the world. Dinitrogen (N_2) or Molecular nitrogen makes up four-fifths of the atmosphere but is unavailable metabolically directly to higher plants or animals. Nitrogen is available to some microorganism through Biological Nitrogen Fixation (BNF) in which atmospheric nitrogen is converted to ammonia by help of the enzyme nitrogenase. The ammonia from microorganisms is then transferred to the plants to meet its nutritional need of nitrogen for the synthesis of proteins, enzymes, nucleic acids, chlorophyll, etc. thus nitrogen enters the food chain from atmosphere through microorganisms and plants. Thus all eukaryotes, including higher plants and animals unsurprisingly depend on the N-fixing microbes for their nitrogen supply followed by BNF activity. The organism that capable to grow without external sources and fixed nitrogen in soil is called diazotroph. According to literature regarding BNF only prokaryotes (members of the domains Archaea and Bacteria) are capable of performing it. The nitrogen fixing ability is extensively distributed across both the bacterial and archaeal domains (Raymond *et al*., 2004).

BNF provides a means to meet the needs of a growing population with a nutritious, environmentally friendly, sustainable food supply. This makes the need for BNF research very convincing in the present situation. In the last two decades, various interesting discoveries regarding nitrogen fixation have been reported. Genome sequence approaches, 'omics' approaches for microorganisms, genetically modifies crops have taken common platform in agricultural biology. The advance level nitrogen fixation research is mainly focused on the enzyme complex called nitrogenise. This enzyme complex has involved in various biochemical processes, apart from its usual functions. Those includes signal transduction, inter and intra molecular electron transfer, protein-protein interaction, involvement in enzymatic catalysis etc (Peters *et al*., 1995).

2.3. The microsymbiont - *Frankia:*

The research on root nodules of non-leguminous plants was getting focused from early nineteenth century. The physiological nature of nodules and the cause of nodule formation first speculated by Meyen (1829) to understand the parasitic infection in the

root. Detail anatomy of the nodules was studied by Woronin (1866). He observed hyphae like structures and some round vesicle within it. He also observed some intracellular region where the hyphae were passing through. He also found that some hyphae continued with round vesicular swelling tips. Woronin considered the unknown organism as fungus and named *Schinzia alni* as this endophyte showed resemblance with a fungus called *Schinzia cellulicolia* (Sen, 1996). Brunchorst (1886), studied the cytological difference between leguminous and non-leguminous roots and named the unknown organism as *Frankia subtilis* to honor his teacher A.B. Frank. Frank was a swiss microbiologist and ironically considered the structures as protein granules and did not believe in the presence of living microorganism in any kind of nodules. Afterwards Frank changed his idea and considers *Frankia* as a fungus along with Brunchorst (1885). In 1895, the name *Frankia alni* was also coined by Von Tubeuf (1895) as a tribute to A.B. Frank. Several names of *Frankia* have been proposed later on, those are *Plasmodiophora alni* (Moller, 1885), *Frankiella alni* (Maire and Tison, 1909), *Aktinomyces alni* (Peklo, 1910), *Actinomyces alni* and *Nocardia alni* by Von Plotho (1941) and Waksman (1941) respectively, *Proactinomyces alni* (Krassilnikov, 1949), *Streptomyces alni* (Fiuczek, 1959; Normand & Fernandez, 2006).

Hiltner (1898), identified the endophyte as member of actinomycetes for the first time, while studying the roots of *Alnus* and *Elaeagnus* and found close ally with *Streptomyces* (Hirsch, 2009). Hellriegel and Wilfarth, (1886-1888) (Quispel, 1988) studied on the atmospheric nitrogen fixation by the bacteria residing in the cortical layer in leguminous nodules. They introduced two terminologies during their research 'nitrogen user' and 'nitrogen accumulator' and described the differences between them. They identified actinorhizal plant alders as nitrogen accumulator, and also described those plants accumulate atmospheric nitrogen for plants growth. Their findings opened up a new era in plant microbial research (Bottomley, 1912; Quispel, 1988; Sen, 1996). Later Hiltner in his work demonstrated that young alders need root nodules to survive in nitrogen free soils. This result pointed out that though alders are certainly not nitrogen fixer but they are nitrogen accumulator on the soil and another organism must be involved in

the nitrogen fixation process. In another study, Brewin (2002) isolated bacteria from root nodules of leguminous plants which failed to infect non leguminous plants, which proved that leguminous microsymbionts and non-leguminous microsymbionts are two different organisms (Brewin, 2002; Pawlowski and Bisseling, 1996). Finally actinomecetes was identified as non-leguminous microsymbiont (Quispel, 1990). Pommer (1956) first successfully isolated a slow-growing actinobacteria from nodules of *Alnus glutinosa* in pure culture with unique morphological features like multilocular sporangia, hyphae and vesicles. After 2-3 weeks of growth in glucose-asparagin agar the bacteria was obtained as 0.6mm colonies described by Waksman (1950) for actinomycetes. Unfortunately this strain was lost before further studies in different independent laboratories. In the year1964 (Becking *et al*., 1964; Silver, 1964) scientists observed prokaryotic structure by electron microscope in root nodules of *Alnus glutinosa* and *Myrica cerifera*. This fact re-established that these root nodules were inhabited by actinomycetes bacteria. Finally *Frankia* was isolated for the first time by Callaham *et al.,* (1978) 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 demonstrated its ability to reinfect the host plant and established its symbiotic behaviour. The morphological character of the bacterium isolated by Callaham *et al*., (1978) shows unexpected resemblance with the lost strain of Pommer which he isolated in 1959 from *A. glutinosa* nodules (Benson and Silvester, 1993). Because of slow growing nature of *Frankia*, frequent contamination occurred by various other fast growing organisms, fungi and other actinomycetes. These are often mistaken as *Frankia*. To overcome this problem Lechevalier and Lechevalier, (1984) proposed the 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 reisolated from within the nodules of that plant, b) produce sporangia containing nonmotile 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." Scientists (Myrold, 1994, Benson and Silvester, 1993; and Akkermans and Hirsch, 1997) reported that three different kinds of cellular structure have been produced by *Frankia* in pure culture or in symbiotic condition i.e. vegetative hyphae, multilocular sporangia containing vesicle and spores. However, *Frankia* infected by *Casuarina* does not produce vesicle in root system of plants (Sen, 1996).

2.4. Systematic approach of *Frankia*-actinorhizal symbiosis:

16s r-DNA sequence study revealed that *Frankia* can be divided into three clades. Other than those three clades, there has been another clade, which one referred as '*Frankia* like clade'. The existence of this clade was proven by Benson and Clawson, (2000) in Polymerase chain reaction (PCR) amplification studies of the 16s-rDNA. The isolation of *Frankia* was successfully done from Clade I and Clade II, however, the no isolation of *Frankia* from clade III. This data revealed that *Frankia*-Actinorhizal symbiosis evolved from more than one common evolutionary ancestor or ancestral group and that to at least three or four times in the history of evolution. This evidence clearly rejected the Doyle (1998) hypothesis of nitrogen fixation in nodules might have evolved at a definite time in the earth's history. Swenson (1996), Benson and Clawson, (2000) and Jeong *et al*., (2000) supported the above idea for further study. The measurement of evolutionary distance was done by ribosomal RNA-encoding gene (rrn gene), however there are a few rrn genes to study the evolutionary distance. Normand *et al*., (1996), Jeong *et al*., (1999) and Clawson *et al*., (2004) made this types of study with with different conclusions. For this type of work (Normand and Fernandez, 2008) the Ochman's metric test (Ochman and Wilson, 1987) was known to be the best. According to this metric, *Frankia* ancestor were evolved at about 350 MY ago from a group of soil actinomycetes. The first trace of land plant was found in this period. *Frankia* was thought to be appeared from this ancestor after a second phase of evolution, at about 100MY ago. The first dicotyledonous plant families started appearing during this period in the earth. This findings revealed the fact that, *Frankia* clusters emerged at 100-200MY ago. During this period

the appearance of the oldest actinorhizal plant genera like *Myrica* and *Alnus* (Normand and Fernandez, 2008) takes place and they communicated with genus *Frankia* to make symbiosis with each other. *Frankia* has relaxed host specificity, compared to rhizobia. Host range for *Frankia* was found to be extremely broad and limited in clade-to-clade interaction. The plants belong to Hammamelidae clade is nodulated *by the* clade- I *Frankia*. clade II *Frankia* nodulate the plants belong to Elaeagnaceae and Rhamnaceae clade and Rosaceae plants are nodulated by clade III *Frankia* (Benson and Clawson, 2000). So, it can be concluded that that this specificity largely inhabits on a group of signal molecules or a group of specific molecule that have same common characters as well as chemical backbone. The *Myrica* and *Gymnostoma* type of primitive actinorhizal plants, produce nodules in their root system, when infected by both Clade I and Clade II *Frankia* (L G Wall 2000). This observation established the primitiveness of these plants which subsequently considered them as most primitive actinorhizae. *Gymnostoma* and *Casuarina* may consider as a probable explanation of this infection because of their habitat. *Casuarina* is inhabitant to drier Australian continent while *Gymnostoma* branched out in the wetter Melanesia region from the tertiary era beginning 65 MY ago (Pawlowski and Demchenko, 2009).

2.5. Haemoglobin – a key factor in global nitrogen balance:

Kubo first identified haemoglobin (Hb) as the haemoprotein like red pigment in the root nodule of soybean (Kubo, 1939). Further research revealed this protein has remarkable properties of oxygenation and oxido-reductase properties as well, where the central iron can change its valency (Burris and Haas, 1944). $Hb+O_2 \acute{o} HbO_2$ (Oxygenation). This property of valency change helps the protein to reversibly bind with oxygen and carbon monoxide (Jokipii-Lukkari, 2016) and is therefore function as an excellent oxygen carrier. The presence of Hb in root nodules of all leguminous and actinorhizal plants with very low partial pressure of carbon monoxide incline to the fact that Hb is somehow connected with BNF (Santi *et al.,* 2013) by buffering the free oxygen concentration at a low tension to protect the nitrogenase from oxygen-inactivation (Appleby, 1992).

2.5.1. Importance of haemoglobin in actinorhizal symbiosis:

Vescicle is the special structure produced by *Frankia* to protect its nitrogenase enzyme from O_2. In *Casuarina* nodules, *Frankia* do not form vesicle. *Frankia* have been synthesize Hb to protect nitrogenise from the oxygen (Beckwith *et al*., 2002). Symbiotic haemoglobin (sHb) in *Casuarina* nodule provides a partially anoxic environment, which subsequently help *Frankia* to fix nitrogen (Berg and McDowell, 1987; Jacobsen-Lyon *et al*., 1995). However, in pure culture *Casuarina – Frankia,* produce vesicle to protect their nitrogenase from the oxygen (Myrold *et al*., 1994). *Datisca glomerata* produces truncated haemoglobin (ptHb) in *Frankia* infected root nodules (Pawlowski *et al*., 2007). It is not involved in O_2 transport as class I Hb, but helps in nitric oxide (NO) detoxification. The occurrence of huge amounts of a class I haemogllobin in alder (Sasakura *et al*., 2006) and *Myrica gale* (Heckmann *et al*., 2006), along with the evidence of ptHb in the *Datisca glomerata* root nodule with *Frankia*, indicates that, big amounts of nitric oxide are produced in actinorhizal root nodules, as seen in legume nodules (Horchani *et al*., 2011), which further leads to high levels of stress. Reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2) and superoxide anions (O_2 –) are produces by Class II Hbs in root nodules which lead oxidative stress (Becana *et al*., 2000). RNAi inhibition of leghaemogllobin gene transcription in root nodules of the legume *Lotus japonicas* confirms this fact. It also helps to reduces H_2O_2 contents (Gunther *et al*., 2007), along with maintenance of free O_2, loss of nitrogenase and nitrogen fixation in root nodules. Tavares *et al*., (2007) have shown that *Frankia* contributes to production of ROS in nodules. So, both actinorhizal and legume nodules have to manage with high levels of stress. Although research revealed that actinorhizal plants can grow at in relatively adverse conditions than legumes, and therefore, it is also possible that they have improved antioxidant based defence mechanism (Pawlowski *et al*., 2011). The intercellular infection takes place at a higher rate compare to root hair infection. There are a number of bacterial factors and ecological conditions involved in nodulation process. The ecological factors involve

water, light, nitrogen availability, soil pH, phosphate, pO2 and pCO2. The bacterial factors are physiological state of the strain, amount and concentration of the inoculum, and nitrogen fixing ability of the bacterium. Those factors play crucial role in the regulation, development and function of the root nodules in actinorhizal plants. Valverde and Wall, (1999) showed the tap root system consists of a short-term gap of susceptibility for nodulation.

2.5.2. Multipurpose nature of haemoglobin:

Hb is a complex spherical shaped protein, containing a haem beta prosthetic group within the alpha helical globular fold, attached with a proximal histidine side chain by covalent bonding. The prosthetic group of Hb contains a characteristic iron atom popularly known as the haem component. It is crucial part of the molecule that plays key role in reversible binding of oxygen with four of the six coordination sites occupied by the haem pyrrole nitrogens. The protein part that encompasses and protects the haem component is known as the globin segment of the molecule (Goodman *et al*., 1988).

Usually, Hb reminds us of mammalian blood present in the circulatory system, but Hb is ubiquitously in eukaryotes as well as in many bacteria playing significant role (Dordas *et al*., 2003a; Dordas *et al*., 2003b). In fact, Hb is available in almost all forms of life and has been reported to be critically involved in oxygen transport and circulation mechanisms in eukaryotic forms of life including mammals (Dordas *et al*., 2003a; Dordas *et al*., 2003b).

Regarding plant Haemoglobin (pHb), it indeed play a crucial role in nitrogen fixation, diffusion and buffer of the concentration of free oxygen, supply of oxygen to the developing tissues, regulates nitric oxide scavenging activities, seed dormancy, transition to flowering and root development etc (Hebelstrup *et al*., 2007). Nevertheless, pHb largely remains behind the curtains and there are only a few review articles about the structure and functionality of plant related Hb.

2.5.3. History (Evolutionary Trend of haemoglobin from very beginning):

The magic molecule “Hb” has been in existence and circulation right from antiquity and is found to be very stable in nature, after being subject to severe forces of selection. Hb is one of the probable earliest proteins of ‘Hadean era’, thought to be present in Last

Universal Common Ancestor (LUCA). LUCA was a single cell organism and had a propensity to be much more flexible, and perform their metabolic processes. LUCA was found to be composed of Hb-like protein called protoglobin (Vinogradov *et al*., 2007).

In anoxygenic time period, protoglobin in LUCA established itself as anoxygenic photosynthesizer and employed oxygen even before it existed in free state (Vinogradov and Moens, 2008). Subsequently, protoglobin converted itself from anoxygenic to oxygenic photosynthesizer and released oxygen as byproduct. So in 'Archean era' protoglobin was one of the most important proteins which was actually responsible for transformation of our atmosphere from anoxygenic to oxygenic (Vinogradov *et al*., 2013).

By the end of the 'Archean era' protoglobin was thought to be used by LUCA in order to balance the oxygen homeostasis and for protecting them from antioxidant enzyme like by-products. Thus, they avoided probable interaction with more oxygen and acted like the sHb of modern day nitrogen fixing bacteria. So probably in the junction of 'Archean' and 'Proterozoic' eras, single domain protoglobin got diversified into the oxygenic photosynthesizer and the oxygen homeostasis (Vinogradov and Moens, 2008).

Oxygenic photosynthesizer single domain globin has been characterised by single or repetitive events with subsequently developed 2/2 fold. There was also marked evolutionary transition from monomeric to oligomeric protein state (D'Alessio, 1999).

Some ancient eukaryotes, bacteria and archaea acquired this protein by lateral gene transfer (Marti *et al*., 2006). In eukaryotic system this globin subsequently developed into neuroglobins, cytoglobins, myoglobins, and Hbs (Freitas *et al*., 2005).

2.5.4. Variation of haemoglobin in different domain of life form:

The discovery of Hbs in virtually all forms of life has revealed that besides, transport of oxygen between tissues, they execute additional functions which range from intracellular oxygen transport to catalysis of redox reactions (Vinogradov *et al*., 2006). These functional modifications of Hbs illustrate its acquisition of new roles through which changes in the coding regions as well as in the regulatory

elements of the same genes. Although, low oxygen concentrations induces many of these diverse Hbs, to date, none of the molecular mechanisms for their hypoxic induction reveal common regulatory proteins (Fago *et al*., 2004).

There are various sub-groups of globin family that include myoglobin, erythrocruorin, Hb beta, leghaemoglobin (LHb), Hb alpha, globin-nematode type, myoglobin-trematode type, globin-lamprey-hagfish type, Hb-extracellular type and globin-annelid type. (Vinogradov *et al*., 2007). Eight types of globins including Hb, neuroglobin, myoglobin, cytoglobin, androglobin, globin E, globin X and globin Y are found in Vertebrates. It is noteworthy that Hbs and myoglobins have the ability to bind oxygen reversibly and they are the only active members of globin family that can bind to the heme prosthetic group (Pesce *et al*., 2003). Nevertheless, in case of mammals, non-coding DNA sequences alignments did not reveal significant matches between mammalian α- and β-globin gene clusters, which thought to be diverged approximately 450 million years ago and are still expressed in certain coordinated fashion (Anderson *et al*., 1996).

The α-globin gene is characterised by a CpG island, which is actually responsible for functionality, whereas the β-globin gene tends to be AT-rich (Dikshit *et al*., 1990; Hardison, 1998).

Gene organization study revealed the presence of three exons in mammalian Hbs (Hardison, 1998). Plant system contains similar type of Hb sequence with other globin proteins; perhaps fascinatingly splitting of middle exon in LHb has given that particular Hb to four exon organizations (Anderson *et al*., 1996).

Bacterial Haemoglobin (bHb) binds oxygen at low concentration and significantly involves in delivering of oxygen to the terminal respiratory oxidase like cytochrome O (Freitas *et al*., 2005). Although various other functions like, sensing oxygen concentration, detoxification of NO, passing the signal to transcription factors has also been found. Apart from above functions, btHb directly involves in the mechanism of heme degradation and iron release (Ott *et al*., 2005).

Hb has also been found in another lineage of life form, namely archaeal system. The main function includes protection from nitrosative and oxidative stress (Vinogradov *et al*.,

2007). However, orientation of B10 distal tyrosine in archaeal system is bit different from other globin proteins and is parallel to the heme plane which subsequently helps to decrease the binding capacity of oxygen (Vinogradov *et al*., 2006).

2.5.5. Structural evolution of haemoglobin in various domains of life:

Hb in different domains of life has enlarged their amino acid length by addition of C- terminal region. Accumulation event of C-terminal domain occurred with the second segment of single domain globin.

This globin was acquired by unicellular organisms and a few lower multicellular organisms like green algae, lower eukaryotes, some protozoans, some algae like phytoplanktons, archaea and bacteria. Those events occurred during early 'Proterozoic era' (Vinogradov *et al*., 2007).

Consequently, this single domain globin developed into 3/3 folds structures, also termed as 'chimeric globins'. Higher plants acquired this chimerical globin probably by lateral gene transfer events and further diversified into flavoHb and globin couple sensor. Significant lateral gene transfer also occurred between bacterial to eukaryotic system in the middle of the 'Proterozoic era' (Vinogradov and Moens, 2008). A vivid display of the pattern of diversification of Hbs in different living forms has been illustrated in Figure 2.2.

At the end of the 'Proterozoic era' bacterial system included:

- 3/3 flavoHbs, 3/3 single domain globins
- 3/3 globin couple sensor, 3/3 protoglobins
- 2/2 truncated Hb.

Eukaryotic system comprised of:

- 3/3 flavoHbs, 3/3 single domain globins
- 2/2 truncated Hbs and lacks 3/3 globin couple sensor, 3/3 protoglobins; and

Archaean system contained:

- 3/3 globin couple sensor, 3/3 protoglobins
- 2/2 truncated Hb but lacks 3/3 flavoHbs, 3/3 single domain globins.

Such patterns of differential existence of the various types of Hbs in diverse forms of life are persistent till date (Kavanaugh *et al*., 1992). 2/2 truncated

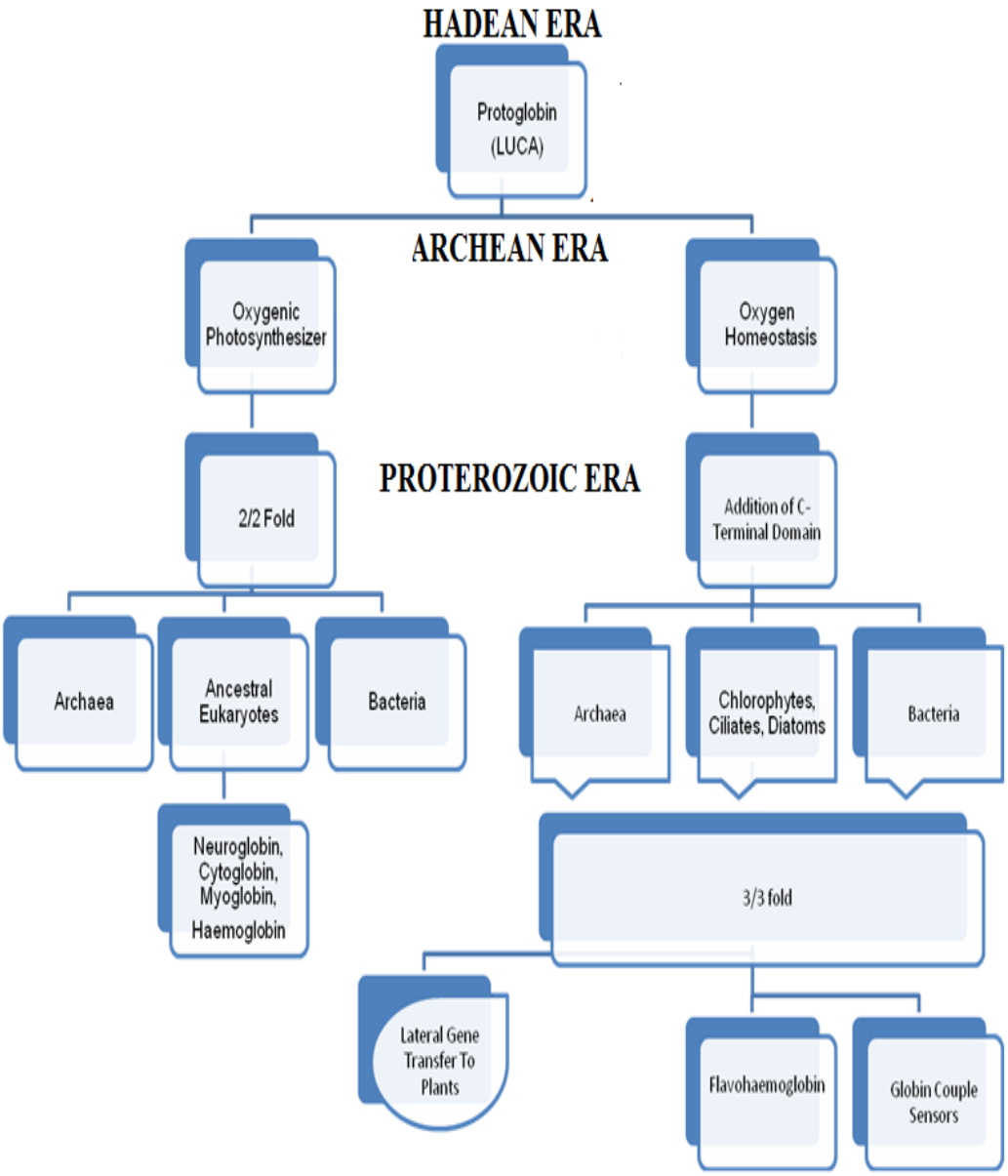

Figure 2.2: Evolution of haemoglobins from the ancient earth with the time period

Hb have been reported to exist in prokaryotic as well as eukaryotic clades of life and are clearly evident from Figure 2.3.

2.5.6. Discovery of plant haemoglobins:

Protoglobins are considered to be present in LUCA (Last Universal Common Ancestor) depending on the determined globin domain length, position of the proximal histidine and distal residues and the chemical nature of the heme pocket (Vinogradov *et al.*, 2007).

Several hypotheses have been proposed regarding history and evolution of globin. The LHbs and animal Hb genes are the result of convergent evolution

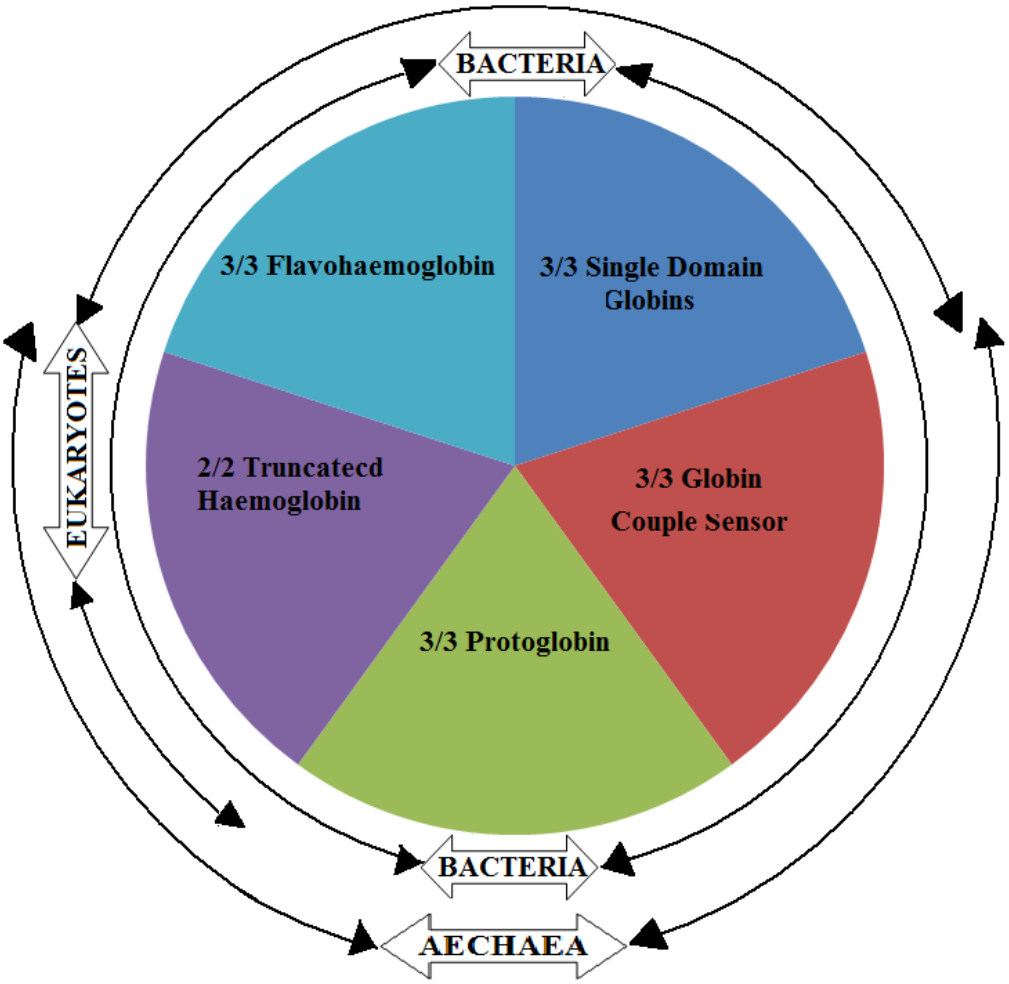

Figure 2.3: Distribution of haemoglobins in different life systems

story, where each evolved independently and acquire identical characters (Freitas *et al*., 2005). Another hypothesis is that it arose in the distant past prior to the common ancestor of living organisms. Then a major portion lost their Hb genes or perhaps, they simply mutated to a great extent (Vinogradov *et al*., 2013). Another significant possibility has been the lateral gene transfer. The trait acquired laterally and subsequently passed on from the ancestors to descendants. However, there have also been reports that claim a vertical shift of Hb genes at the end of 'Proterozoic era' (Roesner *et al*., 2005).

PHbs are structurally similar to animal Hbs and myoglobins and were first characterized from the root nodules of leguminous plants (Kubo, 1939). Further investigation revealed that presence of Hbs accounted for the red pigmentation and these proteins were termed as LHb.

PHbs are abundant in both dicots and monocots. The intricacies of monocot

associated Hb functions in non-symbiotic tissues still remain in dark. Primarily, pHb proteins were thought to be restricted to plant species carrying out symbiotic nitrogen fixation but further scientific investigations established their presence in non-nodulating plants as well. In the two types of Hbs, Hb-1 is expressed in leaves and roots while Hb -2 expresses only in leaves. The biochemical properties suggest that this Hb protein probably does not function to facilitate the diffusion of oxygen (Arredondo-Peter *et al*., 1997).

SHb helps in diffusion of oxygen and buffer the concentration of free oxygen to protect the nitrogenase from oxygen-inactivation and thus aid to improve soil fertility. In legumes, symbiotic nitrogen fixation occurs in specific organs called nodules. The LHb are exclusively required for symbiotic nitrogen fixation, but not for cellular functions of plants. Physiological analysis of nodules revealed the crucial contribution of LHb towards establishing low free-oxygen concentrations but high energy state in nodules, condition that is necessary for effective symbiotic nitrogen fixation. Such an observation demonstrates the strength of Hbs in plants and helps in elucidating the complexities of Hb functions in plants versus animals (Ott *et al*., 2005).

The ratio of oxygen and nitric oxide concentration in symbiotic nodules is maintained by the non-symbiotic haemoglobin (nsHb). Flavins present in prosthetic group reduce globins by transferring electrons from NAD (P) H to the heme iron of globin protein. This reduction mechanism depends on various factors like pyridine nucleotide, flavin coenzyme and different classes of globins by means of a specific combination with NADH or NADPH. FAD reduces class 1 globins at fast reduction rate while class 2 globin reduce slowly. Significantly Class 3 globins reduce rapidly by FMN than FAD (Sainz *et al*., 2013).

2.5.7. Types of plant haemoglobin:

PHb has been classified by various ways. Based on the symbiotic attributes, Hb may be broadly classified into class ‘0’, sHb, nsHb and ptHb (Ohwaki *et al*., 2005) (Figure 2.4). The functional intricacies of various types of Hbs have been discussed later in this review.

2.5.7.1. Class ‘0’ haemoglobin:

In ancient Bryophytes, Hb is often coined as class ‘0’. This course group

contains both penta and hexa coordinated structures (Garrocho-Villegas *et al*., 2007). They contain no intron in their gene sequences like other pHbs (Reddy, 2006).

2.5.7.2. Symbiotic haemoglobin:

SHbs were evolved from nsHbs (Trevaskis *et al*., 1997; Sturms *et al*., 2011a) among non-legume nodulating plants. (Kubo, 1939, Appleby, 1984, Bhattacharya *et al*., 2013). SHbs, in plant system are found in various forms like nitrosyl-LHb, oxy Hb, ferryl Hb, ferrous Hb etc. (Herold and Puppo, 2005; Ji *et al*., 1994). The major function of sHb is to facilitate oxygen diffusion, thus aiding bacterial endosymbiosis (Perazzolli *et al*., 2006). It also buffers the free oxygen concentration, facilitating nitrogenase

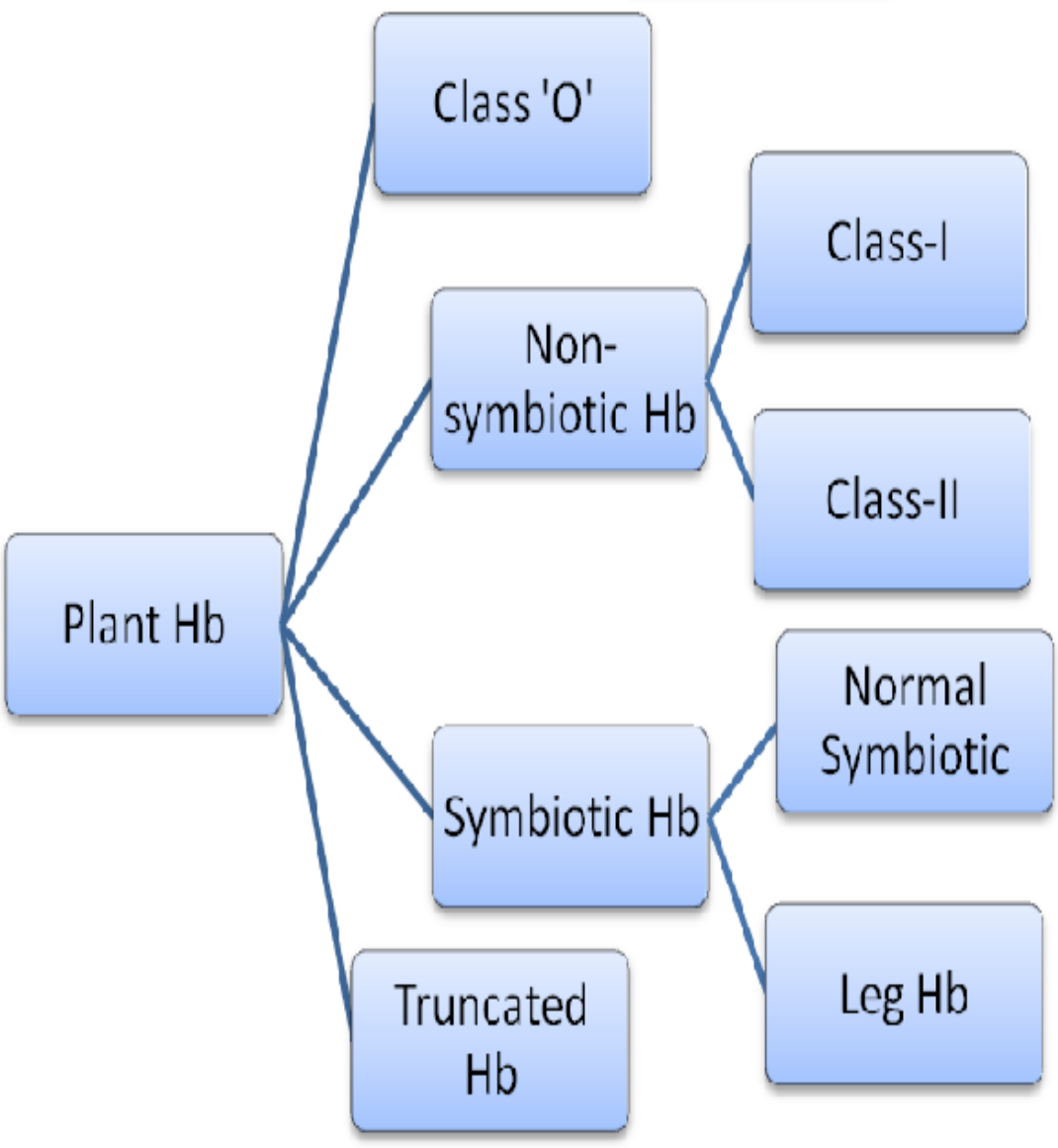

Figure 2.4: Classification of haemoglobins in plant system

to remain at low tension state and in turn protect nitrogenase from oxygen-inactivation (Ohwaki *et al*., 2005; Appleby, 1992). Higher oxygen association and lower oxygen dissociation rate constants help to achieve binding and release of oxygen gradient within the nodule (Igamberdiev *et al*., 2011; Kundu *et al*., 2003).

SHb mostly falls under class-II Hb other than *Parasponia* Hb, which is originates from class-I and structurally & functionally similar to nsHb (Bhattacharya *et al*., 2013).

2.5.7.3. Leg haemoglobin:

LHb is distributed widely in leguminous plants and function to modulate NO during symbiosis. LHbs were evolved after diversification of sHbs from nsHbs (Appleby, 1984; Perazzolli *et al*., 2006).

2.5.7.4. Non-symbiotic haemoglobin:

Initially it was assumed that Hb exists only in symbiotic plant systems due to their abundance in symbiotic root nodules (Arredondo-Peter *et al*., 1997; Ross *et al*., 2001; Appleby *et al*., 1988; Heckmann *et al*., 2006). However a series of research, involving various plant species showed that Hb is virtually present in all kinds of plant system. Though most of these Hbs are non-symbiotic (Appleby *et al*., 1988; Heckmann *et al*., 2006).

NsHbs were found to dwell mainly in leaf and stem. Nevertheless, Taylor *et al*., (1994) and Sowa *et al*., (1998) reported its presence in roots of the seedlings, particularly the root tissues exposed to flooding stress and oxygen-stressed aleurone tissues. The presence of different types of Hb in different plant tissues indicates its function other than oxygen transportation (Arredondo -Peter *et al*., 1997; Ross *et al*., 2001).

NsHbs are structurally hexacoordinates where distal histidine reversibly binds the iron at the sixth coordination site (Arredondo-Peter *et al*.; 1997; Duff *et al*., 1997).

NsHbs have also been observed to actively participate in mitochondrial ATP synthesis (Nie and Hill, 1997). NsHb expression is modulated by the calcium-ATPase concentration and also found to interfere with calcium modulated signal transduction (Nie *et al*., 2006; Ross *et al*., 2004). Excess occurrence of nsHb helps to maintain the growth of root portion, cytosolic ATP levels and lower nitric oxide levels under hypoxia (Dordas *et al*., 2003a; Igamberdiev and Hill, 2004). Nitrate, nitrite, and nitric oxide donors

induce nsHb proteins to associate with NADH-nitrate reductase and enhance its function (Ohwaki *et al*., 2005; Igamberdiev *et al*., 2006a).

Phylogenetic analyses reflect two distinct classes of nsHb, (Class-I and Class-II) (Trevaskis *et al*., 1997; Gopalasubramaniam *et al*., 2008) where evolution was observed from hexacoordinate to pentacoordinate transitional state. Further study revealed, this classification also associated with the affinity of Hbs towards oxygen (Ohwaki *et al*., 2005). Monocotyledonous plants contain exclusively the class-I nsHbs, whereas, both the classes of nsHbs (class-I and II) are reported in dicots (Hunt *et al*., 2002).

2.5.7.5. Non-symbiotic class-I haemoglobin:

Class-I nsHb shows high affinity to oxygen, (Hargrove *et al*., 2000; Smagghe *et al*., 2009) and also regulates nitric oxide scavenging activities, thus correlates crucially to major developmental processes like seed dormancy, transition to flowering, root development (Mendelson *et al*., 1994; Gupta *et al*., 2011) and adaptation to abiotic and biotic stress (Baudouin, 2011). Nitrate reductase produce NO, which helps in plant adaptation, cold tolerance and also regulates the gene expression involved in phosphatidic acid synthesis and sphingolipid phosphorylation (Cantrel *et al*., 2011).

Class-I nsHbs are characterized by low values of the hexacoordination equilibrium constant (KH) in comparison to other classes of Hbs. (Garrocho-Villegas *et al*., 2007; Hebelstrup *et al*., 2007).

The first class-I nsHb crystal structure to be studied thoroughly was the *Oryza sativa* Hb I (Hargrove, 2000; Halder *et al*., 2007). Usually class-I nsHbs constitute one or two cysteine residues per monomer which helps in reduction of ferric ions and protects the protein from autoxidation (Hoy and Hargrove, 2008; Igamberdiev *et al*., 2006b). Cystein residues along with sulfahydryl reagents such as reduced glutathione, facilitate the transition of ferric to ferrous state of the protein (Igamberdiev *et al*., 2005, Sturms *et al*., 2010) and allow nsHb I to serve as soluble electron transport proteins, in the enzymatic system (Sturms *et al*., 2011b; Dordas *et al*., 2003a). Class-I nsHb is also found to assist the maintenance of the ATP level, at low oxygen concentrations (Arechaga-Ocampo *et al*., 2001; Trevisan *et al*.,

2011; Dordas *et al*., 2004; Sowa *et al*., 1998). However, over expression of class-I nsHb leads to vigorous growth which however, protects the plant from hypoxia and modulate survival rate (Ohwaki *et al*., 2005; Hunt *et al*., 2002). Furthermore, highly expressed levels of class-I nsHb was found to exhibit a wide domain of functions that included hydrogen peroxide scavenging activities under hypoxia (Perazzolli *et al*., 2004; Igamberdiev *et al*., 2014; Yang *et al*., 2005), peroxidase-like activities, NADH oxidation, and S-nitrosoHb nitric oxide scavenging (Perazzolli *et al*., 2004; Sakamoto *et al*., 2004).

Expression of class-I nsHbs in root region are sometimes repressed by fungal infection (Gupta *et al*., 2011; Uchiumi *et al*., 2002) but enhanced by different conditions like hypoxia, cold stress, rhizobial infection, plant hormones, nitric oxide and sucrose levels, etc (Sereglyes *et al*., 2000; Shimoda *et al*., 2005; Bidon-Chanal *et al*., 2007; Marti *et al*., 2006; Igamberdiev *et al*., 2005; Qu *et al*., 2006).

Hordeum vulgare class-I nsHb has been observed to be bound with cyanide ligand (Smagghe *et al*., 2008; Ioanitescu *et al*., 2005). It has been evident from the *Hordeum vulgare* class-I nsHb structure that the ‘‘piston ‘movement of the E-helix along the helical axis is responsible for ligand binding properties (Smagghe *et al*., 2006; Bykova *et al*., 2006). Class-I nsHbs have been found to display lower necrotic symptoms and higher nitric oxide scavenging activity in *Nicotiana tabacum* plants (Sereglyes *et al*., 2004; Sereglyes *et al*., 2000). Distal histidine of nsHb I has been reported to interact with a conserved phenylalanine of the β-helix (PheB10) and result in disorder in the CD-region (Sasakura *et al*., 2006; Roesner *et al*., 2005, Smagghe *et al*., 2006). Ligand migration in class-I nsHb is regulated by suitable interaction between hydrophobic channel and internal ligand docking sites (Bruno *et al*., 2007).

2.5.7.6. Non-symbiotic class-II haemoglobins:

Class-II nsHb displays a comparatively lower likeness to oxygen (Dordas *et al*., 2003b; Vigeolas *et al*., 2011) and involved in supplying of oxygen to the developing tissues (Hill 2012; Vigeolas *et al*., 2011).

Generally it has been reported that dicot plants usually contain class-II nsHb. However, the only monocot

plant to contains class-II nsHb is *Zea mays* (Garrocho-Villegas and Arredondo-Peter, 2008).

Hexacoordination arrangement in class -I nsHbs appear to be much more stable compared to class-II nsHbs. This arrangement actually helps class-II nsHbs to increase the oxygen storage and diffusion affinities and also lower the NO scavenging function (Kakar *et al*., 2011a; Garrocho-Villegas and Arredondo-Peter, 2008).

NO and oxygen occupy two ligand binding sites of class-I nsHbs while class-II nsHbs offer only one ligand binding site as another site is always occupied by NO (Nienhaus *et al*., 2010). Protein instability regulate the exchange of ligand to the distal pocket in class II ns Hbs (Bruno *et al*., 2007).

The oxygen binding properties of class -II nsHbs are more similar to sHbs and display affinity to oxygen, resembling the rate of cytochrome oxidase. Such a tendency of predilection for oxygen binding actually suggests their role towards oxygen diffusion (Spyrakis *et al*., 2011; Vigeolas *et al*., 2011).

It has also been revealed that under low external oxygen concentration, the expression of class-II nsHb increases up to 5-fold and prevents fermentation (Hunt *et al*., 2001; Sakamoto *et al*., 2004; Thiel *et al*., 2011) which signify that class-II nsHbs are deeply involved in seed oil production and supplying oxygen to the developing seeds (Hoy and Hargrove, 2008; Trevaskis *et al*., 1997; Roesner *et al*., 2005).

2.5.7.7. Comparative relationship between symbiotic and non-symbiotic haemoglobins:

SHb differ structurally from nsHbs and myoglobin, as their points of ligand orientations are completely diverse. NsHbs and myoglobin actually depends on distal pockets whereas, sHb depends on its proximal pockets for ligand regulation (Kakar *et al*., 2011b; Kundu *et al*., 2003). In case of myoglobin the distal histidine is found to be bound with a stabilised water molecule by a strong hydrogen bond (Kundu and Hargrove, 2003) which is subsequently removed paving way for other ligand to bind. On the contrary, in sHb water is not stabilized in the distal pocket. Proximal histidine has been found to bear an eclipsed orientation with haem pyrole nitrogens in nsHbs and myoglobins whereas, in case of sHb proximal histidine increase ligand affinity by maintaining a definite orientation (Quillin *et al*., 1993; Takano, 1977; Kundu *et al*.,

2003; Tarricone *et al*., 1997).

2.5.7.8. Plant truncated (Class-III) haemoglobin:

PtHbs are found to be distributed, almost entire plant kingdom, from the ancient algal ancestor to the recent higher plants (Couture *et al*., 1994). They have very low affinity to oxygen therefore classified as class III Hb (Hunt *et al*., 2001) however, their function is somewhat obscure as of now. ptHb shows very low similarities with class-I and class-II Hbs and possess 40-45 % sequence similarities with structural motif' of btHbs. BtHb sequences are usually shorter than usual vertebrate and non-vertebrate Hbs and myoglobins though their length differs extensively for each case (Ascenzi *et al*., 2007). However ptHbs possess genes that are longer than genes encoding 3-on-3 Hbs (Jokipii-Lukkari *et al*., 2009), where 2-on-2 sandwich system were found, instead of 3-on-3 classical Hb fold (Vinogradov *et al*., 2006).

According to researchers, the '2 on 2 Hb fold' originated by the division of classical '3 on 3 alpha globin fold' (Holm and Sander, 1993; Baudouin, 2011; Cantrel *et al*., 2011).

PtHb possess glycine-glycine motifs in the region of globin fold with glutamine E7 and tyrosine B10 which basically stabilize the ligand (Pesce *et al*., 2003). The surface of the protein is connected to the distal haem pocket by a polar tunnel with a xenon binding site at the position of tunnel entry and tunnel branches (Milani *et al*., 2004a; Milani *et al*., 2004b).

Some plant species contains two ptHbs - the first one usually found in the root region and shows similarities in their expression with sHbs, whereas second one found in the base of the nodules, vascular tissues and mycorhizal roots. However, function of these Hbs still remain an area unexplored (Vieweg *et al*., 2005; Watts *et al*., 2001). Based on the functional patterns, scientists proposed that the function of ptHb might be related to suppression of defence processes against symbiosis and also detoxification of nitric oxide (Vieweg *et al*., 2005; Pawlowski *et al*., 2007).

In oxygenated state ptHb shows pentacoordinate arrangement (Watts *et al*., 2001) while hexacoordinate arrangement has also been seen during reduction and deoxygenation (Couture *et al*., 1999; Das *et al*., 1999). Conversion of hexa to penta coordination arrangement in ptHbs,

takes around 20 minutes which is extremely slow compared to other classes of Hbs (Wittenberg *et al*., 2002; Gupta *et al*., 2011).

2.5.8. Resemblance of plant haemoglobins with different forms of globin protein:

Characterization of pHbs, showed similarities of class-I and II Hbs with animal myoglobin and class-III Hbs with btHb (Watts *et al*., 2001).

Structural analysis and intron-exon arrangement revealed that Barley (class -I) and Rice (class-II) Hb resembled animal myoglobin and Hb (Holm and Sander, 1993; Hoy *et al*., 2007; Perutz, 1979) which further established that there are two conserved introns prevalent in both in animal system (myoglobin and Hb genes) as well as plant system (sHb and nsHb) (Mathieu *et al*., 1998).

Additionally plant symbiotic and nsHbs have been reported to contain a third intron that resembles neuroglobin and cytoglobin genes of vertebrates and Hb genes from invertebrates (Burmester *et al*., 2002; Roesner *et al*., 2005). Such occurrence indicates that three intron arrangements is the primitive version of the Hb genes in which the central intron might have been deleted in course of evolution.

Comprehensive sequence and structure based studies suggest that class-I and II plant nsHb in animal myoglobin and Hb might have been derived from a common unicellular ancestor, estimated to have lived around 1500 million years ago (Moens *et al*., 1996).

2.5.9. Structural rearrangement of plant haemoglobin:

Hbs may also be classified on the basis of coordination of the haem iron. SHbs and nsHbs are homologous to 3/3 fold of animal Hb and the remaining pHbs consist of 2/2 fold called 'ptHb' (Vazquezlim *et al*., 2012). 3/3 Hbs had emerged from green algae before even the evolution of embryophytes (Meakin *et al*., 2007). 2/2 Hbs originated from ancient organic photosynthesizers which further interacted with photo-system I and II and such an interaction has continued till date (Vinogradov *et al*., 2006). Duplication, diversification, and functional adaptations have been attributed to evolutionary patterning in the plant kingdom. In the middle of the 'Proterozoic era' diversification of plant nsHb genes into the nsHb-1 and nsHb-2 occurred and then it emerged gradually in monocot and dicot plants

respectively, followed by subsequent evolution into sHbs at 600 mya (Arredondo-Peter, 2011). Rearrangement of heme-Fe coordination paved way for successful transition of hexa- to penta-coordination which subsequently lead non-symbiotic to symbiotic arrangement of pHbs (Nakajima *et al*., 2005). Systematic order of Hbs in plant system from the Algal ancestor till date was shown in Figure 2.5.

In plants, pentacoordinate and hexacoordinate Hbs are predominant. Erythrocyte Hb and other oxygen transporter Hb possess pentacoordinate structure. The haem prosthetic group contains an iron atom with six coordination sites of which four coordination sites are occupied by the haem pyrrole nitrogens. Usually two histidine side chains are always attached to the prosthetic group of Hb (Trent and Hargrove, 2002).

When the proximal histidine coordinates the fifth position and leaving the sixth position free to bind the exogenous ligands of diatomic gases like nitric oxide and oxygen, pentacoordinate structure of Hb is formed. Pentacoordinate Hb has the ability to transport its bound ligands thus, scavenging, sensing, detoxification and electron transfer (Arredondo-Peter *et al*., 1998).

In case of hexacoordination, both distal and proximal histidine occupies the haem active site. Proximal histidine occupies the fifth position as usual while the binding of distal histidine is reversible which also allow the gaseous exogenous ligands to bind. Therefore, pentacoordination state provides an optimal condition for storage and transportation of oxygen over the hexacoordination state (Gupta *et al*., 2011).

2.6. Computational biology of haemoglobin proteins associated with nitrogen fixation:

Along with the developments of various computational techniques, a large amount of data made available in the public domain. This data also includes nucleic-acid and amino-acid sequences of Hbs from a wide range of plant system and microbes as well. However, very little is known about the structure and role of all these Hb proteins. X-ray and NMR crystallography techniques are most commonly used method to determine the pHb and bHb protein structures experimentally. However the tertiary structures of large number of Hb proteins from different plants and

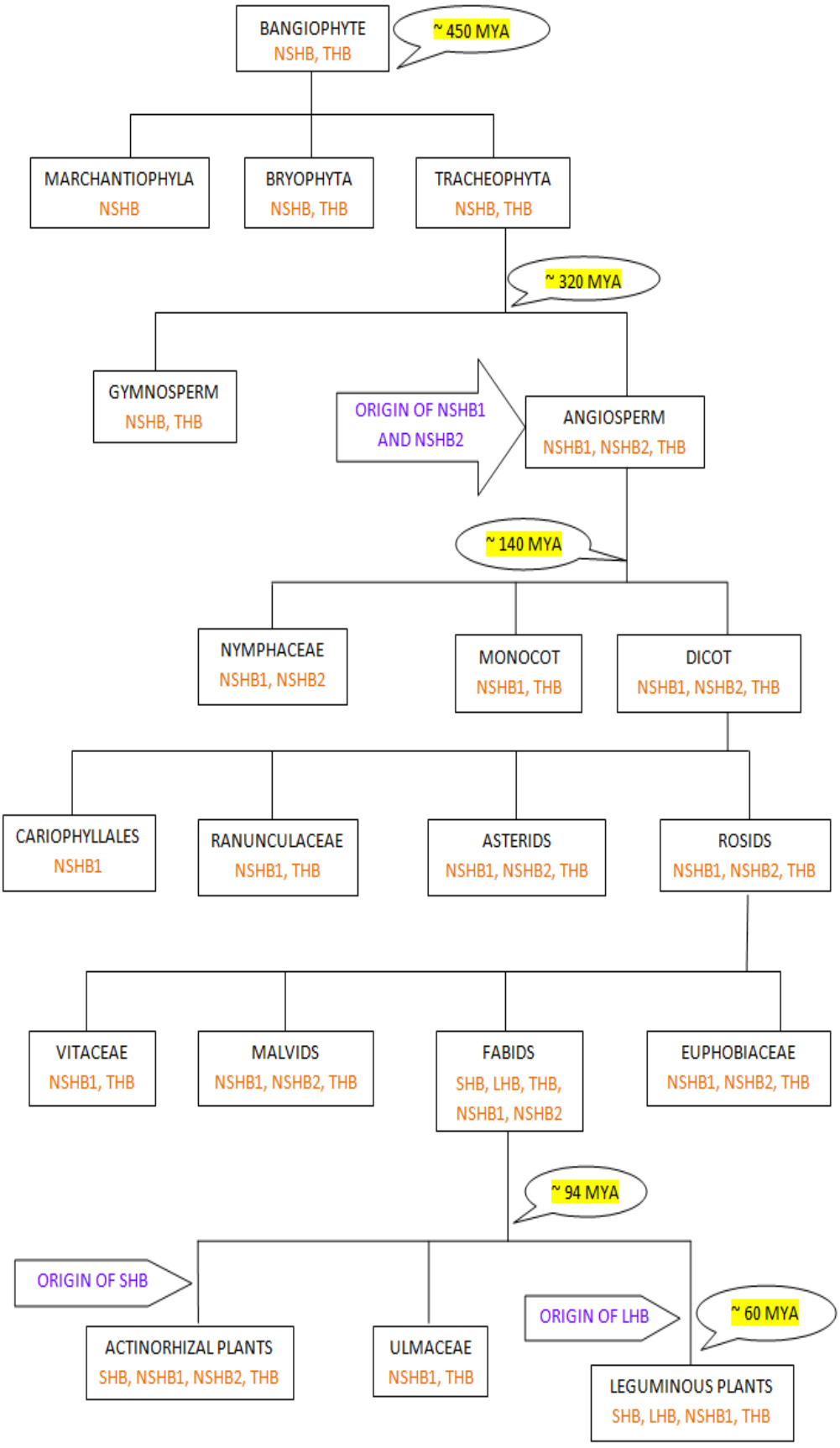

Figure 2.5: Systematic order of haemoglobins in plant system from the Algal ancestor till date

microbes particularly those of symbiotic, non-symbiotic and truncated has not been yet resolved. The exact mechanism of working of these plant and bHb proteins are also relatively unknown due to the difficulty in obtaining crystals of oxygenated Hbs. This is because the resting state of the protein does not bind oxygen, and also the competitive binding nature of Hb with O2 and NO. Additionally, a number of differences have also been come out regarding the X-ray crystallographic protein structures (Chang *et al*., 2012). In this regard a possible alternative approach to predict the three dimensional protein structure was introduced. This technique is the homology based virtual protein modeling technique. Homology modeling is a consistent technique that can predict the three dimensional structure of protein with accuracy similar to one obtained at low-resolution by experimental means (Marti-Renom *et al*., 2000). This computer based bioinformatics technique of protein modelling depends on the various parameters like selection of template of query protein, the alignment of a protein sequence of unknown structure (target) etc. This technique is very much useful mostly in case of slow growing organisms and their subsequent proteins which poses difficulties during protein purification.

Homology modeling technique introduced in the field of research by Browne *et al*., (1969). a large number of homology models of proteins with different folds and functions have been reported since the mid 1980s (Johnson *et al*., 1994; Sali, 1995). The protein models generated by homology modeling techniques are quite useful in providing conformational properties and structure-function relationship.

The three dimensional protein structure of plant as well as bHbs are often considered as ideal model system for the study of the oxygen binding properties, electron transfer mechanisms, complex metal cluster assembly, protein–protein interactions, structure-function relationships and last but not the least the plant microbe association specially the actinorhizal symbiosis.

Molecular Dynamics simulations offer detail information about the molecular motional properties of specific given time period and are widely used to study protein motions at the atomic level. First protein simulation for 9.2 ps was carried out by Mc Cammon *et al*.,

(1977). This research group run bovine pancreatic tyrpsin inhibitor (BPTI) in molecular dynamics simulation to understand the motional properties of that particular inhibitor. In the year 1979 Case and Karplus, (1979) worked on dynamics study of ligand binding heme protein. First application of normal modes to identify low frequency of the proteins was done by Brooks and Karplus, (1983). This was achieved by the oscillations using the energy minimization of the molecular mechanics force-field of protein. First simulation of protein in a virtual water box was done by Levitt and Sharon, (1988).

Metalloproteins like pHbs are responsible for many vital functions. Structure based studies focused mainly on storage role, involvement of electron-transfer process and binding properties of these proteins.

On the basis of the availability of their 3D structures, structural divergent of these protein was also been studied apart from molecular dynamics simulation. The studies revealed that various pHb and bHb proteins have structural, functional and mechanistic similarities as well as evolutionary relationships. The structures of various homologs of pHb and bHb proteins have been utilized for phylogenetic reconstruction based on a structural dissimilarity (RMSD) matrix (Boyd *et al*., 2011).

2.7. *In-silico* analysis involved in computational biology:

A phylogenetic approach is one of the best techniques in computational biology to understand the evolution of proteins. This technique is useful to construct the phylogenetic tree in case of divergently evolved proteins. Phylogenetic approaches have mainly come about on the basis of the similarities in the nucleic acis as well as protein sequences. But previous studies revealed that the origin and extending distribution of pHbs and bHbs are confusing from a phylogenetic perspective. This is because of factors that confused molecular phylogeny such as sequence divergence, paralogy, and horizontal gene transfer (Raymond *et al*., 2004). This confusion lead the postulation that only sequence based phylogeny is not enough to disclose the complex evolutionary trend in pHbs and bHbs. Many workers (Nadler, 1995; Qi *et al*., 2004; Sims *et al*., 2009) supported the belief of erroneous sequence based alignment methods. Hence the alternative approaches came out.

The suitable alternative approach found other than sequence alignment is structure based phylogenetic approach. It is well known fact that the features of protein conformational structures are highly conserved than that of amino acid sequences in case of homologous proteins (Chothia *et al*., 1986; Hubbard and Blundell, 1987). Various way of research demonstrated that the homologous proteins maintain similar protein structure as well as function, though they could diverge beyond recognition at the level of their amino acid sequences. Various protein families like short-chain alcohol dehydrogenases (Breitling *et al*., 2001) and metallo-β-lactamases (Garau *et al*., 2005) showed low level sequence similarities but retained the similar kind of protein folds as well as broad biochemical features. Currently, phylogenetic approach based on 3D structures has been utilized for characterization of functional properties of cupin folds (Agarwal *et al*., 2009). It was found that structure based phylogenetic analysis reflect functional clustering of cupin superfamily.

Therefore, structure based approaches can be employed to assess the phylogenetic relationships of pHb proteins with many other bHb proteins with diverse biological functions, which share low sequence similarity but high structural similarities. Along with the evolution of pHb and bHb, features that need attention is the functional divergence of the proteins involved in this biological process. Experimental works by various workers (Gu, 1999; Dermitzakis and Clark, 2001; Raes and Van de Peer, 2003) have shown that a gene duplication event leads to a shift in protein function from an ancestral function. As a result some residues are altered their functional constrains which subsequently leads into functional divergence. This is the reason behind the different evolutionary rates found at these sites which vary with different homologous genes of a gene family. Site-specific altered evolutionary rates can be detected by comparing the rate correlation between gene clusters, when the phylogeny is given (Gu, 1999). This approach has been earlier exploited by Gribaldo and his co-workers (2003) to trace the functional divergence in vertebrate Hbs. Alpha subunits of G-protein (Zheng *et al*., 2007), OPR gene family in plants (Li *et al*., 2009), anoctamin family of

membrane proteins (Milenkovic *et al*., 2010) were also been exploited by this method. But a broad picture on the functional divergence in the pHb and bHb protein family is still unavailable.

2.8. Challenges and future prospects:

Considerable progress has been made in understanding the machinery of pHbs in last decade. The major part of the research has been focused on the diversity of actinorhizal Hbs from other pHbs, elucidation of the compositions and functions of all of the pHb gene products. In the past, problems associated with detection of Hb gene from different types of plants specially symbiotic and truncated Hb of actinorhizal plants and subsequent crystallization of the protein has been the major problems in the Hb of plant kingdom research. But as we entered into the post genomics era, the major hurdles have been removed. So the challenge now is to deposit all the known information's together and with the help of combined application of biochemistry. genetics and bioinformatics techniques, determine the molecular properties and functions of the protein at the basic level. With the rapid increase in the number of pHb genes along with the complete genomes of varied bacteria in the public domain, the computational tools act as powerful weapon to handle the unresolved properties of actinoHb (actinorhizal pHbs and bHbs).

With the help of *in-silico* protein modelling techniques and new algorithms associated with structural divergence, the problems associated with interpretation of sequence data, functional evolution of pHbs and bHbs can be resolved in a better way and subsequently increases the new sights of research. Studies aided by the bioinformatics tools offer a global view of the expression, regulation, dynamics, evolution of the proteins, inter evolutionary rate and the capability in offering new opportunities to preserve and improve biotic resources.