International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

# The Phylogenetic Tree Of The 'Nod D' Host-Specific Factor Is Incongruent With The Rhizobial Species Tree And Shows a Possible Hgt Event Between Highly Unrelated Rhizobial Bacteria

**Vanitha T[1], Pratiksha B[1], Hemalata M S[1], Sahana B R[1], Poornima M[1], Manohar G M[2*]**

[1]*Department of PG studies in Biotechnology, Nrupatunga University, Bangalore-01.*
[2]*halliLabs, Ragihalli,Anekal Tq, Bangalore – 83*

**ABSTRACT:**
*Rhizobium* is a symbiotic biological nitrogen-fixing (BNF) bacteria. and has high diversity in different geographical regions worldwide. The BNF associated with the legume trees in tropical environments improves the efficiency of nitrogen and increases the soil organic matter, and soil fertility. Rhizobium consists of *Mesorhizobium, Sinorhizobium, and Bradyrhizobium,* the species belonging to each genera infecting a specific host. This host specificity is dictated by several factors among which Nod D was secreted by the bacterium and plays an important role. We investigated how Nod D gene sequences may have evolved to confer host specificity. We constructed phylogenetic trees of 16srRNA and Nod D trees and compared the trees with a manually built host tree. We find that the Nod D tree is incongruent with the 16srRNA tree and the topology does not show any specific pattern correlative of the host tree. We also uncover a possible HGT of the Nod D gene between highly unrelated bacterial species. This finding may have implications for the development of Rhizobium strains specific to a particular host for Agricultural purposes.
**Keywords:** Rhizobial bacteria, Phylogeny of 16srRNA and NodD, HGT.

## INTRODUCTION :

*Rhizobia* are gram-negative bacteria that are involved in the formation of highly specialized nitrogen-fixing organs called nodules on the roots of leguminous plants[1]. Eg: The Rhizobium strain NGR234 nodulates plants in more than 110 legume genera[2]. The root nodule formation is a result of the interaction of signals between the host and the *rhizobia*. The signals involved are Nod (nodulation) factor, flavonoids, etc. The *Rhizobium* species such as *Bradyrhizobium, Mesorhizobium,* and *Sinorhizobium* are host specific. The plants differentiate between pathogenic and symbiotic bacteria by transmitting symbiotic and immune signals enabling them to distinguish symbiotic from pathogenic bacteria[3) ]. The plant possesses plasma membrane-localized pattern recognition receptors (PRRs) that recognize symbiotic pathogens which are associated with molecular patterns (MAMPs/PAMPs) [4]. During the need for nitrogen, flavonoids are secreted by host plants. *Rhizobia* will sense these molecules and activate the product of the "NodD" (nodulation) gene, a transcription factor that controls the expression of genes involved in the synthesis of the Nod Factor (NF)[5]. Flavonoids are the secondary metabolites found in plants. The role of Flavonoids in the selection of *Rhizobial* symbionts is to induce transcription of the genes for biosynthesis of the *rhizobial* signaling molecules like Nod factors. The production of 'infection' flavonoids is highly seen at infection sites[6]."NodD" genes are encoded by the DNA binding proteins that activate the transcription of

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

Nod operons (structural genes). The core structural genes contain Nod A, B, and C genes required for the synthesis of Nod factors,
Nod A gene: N-acetylation of amino sugar backbone.
Nod B gene :D acetylase at non reducing end.
Nod C: synthesis of chitin oligomer backbone[7].
The Nod factor acts as a signal molecule, these signal molecules induce the activation of host genes which results in the curling of the route in the host plant. Nodulation factors that are synthesized by all the species of rhizobia are Lipo-chitooligosaccharide(LCOs). The structure of LCO varies in terms of the side chains and these side chains are added by different Nod genes[8]. These LCOs are implicated in host specificity. Here we tested whether host specificity dictates " Nod D" gene evolution by building phylogenetic trees of Nod D and a manually constructed Host tree. We find that Nod D gene evolution has varied considerably from 16srRNA which we used as a control for establishing the evolutionary relationship between the different species of Rhizobial bacteria. During the analysis, we uncover an event of horizontal transfer between distantly related Bradyrhizobium and Rhizobium genera.

## MATERIALS AND METHODS:

### 1). Retrieval OF 16sr RNA:

The 16sr RNA gene is conserved in organisms and it is used in phylogenetic studies.16srRNA gene sequences in fasta format were collected from the NCBI database using the Keywords nucleotide and organism name from a list of *Rhizobium* species. The length of the sequences ranged from 1100-1400 in a notepad.
By using MEGA Software, We collected 34 *Rhizobial* species accession numbers for the 16srRNA gene from the NCBI using a database as the Nucleotide. Then we retrieved the fasta sequence of each *Rhizobial* species and saved it in the notepad. In MEGA software, the saved file was uploaded, and then selected all the sequences of *Rhizobial* species using Ctrl+A. After selecting the muscle alignment was done. Then using the Phylo analysis tree option the maximum likelihood tree was constructed by applying the saved MSA (Multiple Sequence Alignment) with the bootstrap values and later *Bradyrhizobium japonicum* was made as an outgroup in the tree by selecting the subtree option in the mega tool and by pointing the root on the outgroup then by clicking on the desired species as an outgroup branch, the outgroup could be constructed.
We performed Blastn "search" to obtain the 16sr RNA nucleotide sequences of the desired rhizobial species, by considering *R.leguminosarum* 16sr RNA gene Accession: ON428643 as a query sequence. Multiple rhizobial species were added to the organism search bar and subjected to the blast. As per the obtained results, the species are selected based on the parameters such as E value- 0.0, Query coverage - 95-100%, and Percent Identity 95-100%.

### 2)Retrieval of NodD protein sequences:

"12" NodD protein and "20" LysR protein (Ly's regulator) sequences in fasta format of listed Rhizobium species "were retrieved from NCBI" protein "database".Few species without "NodD" protein had LysR protein which was taken into consideration. "These sequences were ranging from 250-350 amino acids in length".We performed blastp to obtain the "NodD" protein sequences of the rhizobial species, by considering *R.leguminosarum* Accession: P04681.1 as a query sequence. Multiple rhizobial species were added in the parameters such as the organism search bar and subjected to the blastp. As per the obtained results, the species were selected based on the value- 0.0, Query coverage - 95-100%, and Percent Identity 95-100%.

**Multiple sequence alignment and Phylogenetic tree construction**-Version 11.0.13

MEGA was used for multiple sequence alignment and building of the respective phylogenetic tree. In the MEGA software choose to edit or build alignment using the MUSCLE program. "The resulting MSA was trimmed for variable regions at the ends". The obtained alignment was used for the construction of the phylogenetic tree (Maximum likelihood tree ) certain parameters are taken into

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

consideration which are: model/method→Jones- Taylor- Thornton, No.of threads, Data treatment - pairwise alignment.

**RESULTS:**

Nod D protein is a determinant of host specificity between the leguminous plants and Rhizobium bacteria. This protein is secreted by the Rhizobium bacteria. We wanted to see the evolutionary patterns in the Nod D gene in different genera and compare them with the evolutionary relationship of the host plants. So we constructed phylogenetic trees of the 16srRNA sequence of the selected bacteria and a phylogenetic tree of the Nod D gene from the same list of bacteria. We also draw a phylogenetic tree of the host tree manually using the ETEtool Kit. Then we compared the topology of all the trees with each other to see any interesting patterns.

**Phylogenetic tree using 16srRNA:** We constructed a 16srRNA tree to understand the evolutionary relationships between different rhizobial species. Also, this will give us a species tree against which we can compare the evolution of other genes like the "NodD". 16srRNA sequences ranging from 1,100 to 1,400bp were retrieved from NCBI Genbank of 34 species ( Tab 1). The multiple sequence alignment was done using muscle alignment in MEGA software. This MSA was used to construct a Maximum likelihood phylogenetic tree with a bootstrap value of 100.

| SL.NO | HOST NAME | BACTERIA NAME | ACCESSSION NO.OF 16srRNA | ACCESSION NO.OF NodD PROTEIN |
|---|---|---|---|---|
| 1 | Melilotus | S.Meliloti | LT614644 | AAB95383 |
| 2 | Glycine max | S.fredii | LT615092 | WP_042777972 |
| 3 | Glycine max | S.xinjangense | MT036085 | ABD75272 |
| 4 | Sesbania | S.sahelli | KT715731 | WP_153435557 |
| 5 | Sesbania | S.terangae | MT534141 | WP_153439631 |
| 6 | Acacia | S.arboris | KF580872 | WP_028002355 |
| 7 | Acacia,Prosopis | S. kostiense | MT533830 | WP_209606289 |
| 8 | Kummerowia stipulacea | S.kummerowiae | AF364067 | AET31456 |
| 9 | Acacia | S.americanum | JN624739 | WP_132078501 |
| 10 | Glycine max | S.glycinis | HQ174477 | OAP43783 |
| 11 | Lotus,Lupinus | M.loti | MH780220 | ACT34148 |
| 12 | Astragalus | M.huakuii | MT197404 | WP_183454981 |
| 13 | Cicer arietinum | M.ciceri | KY515341 | AMZ00216 |
| 14 | Glycine , Glycyrrhiza | M.tianshanense | FM203303 | WP_145716791 |
| 15 | Cicer arietinum | M.mediterraneum | KF709113 | ACT34153 |
| 16 | Acacia | M.plurifarium | LC515499 | AFJ42542 |
| 17 | Amorpha fruticosa | M.amorphae | KJ556357 | ACT34168 |
| 18 | Prosopis | M.chacoense | MW183078 | ACT34147 |
| 19 | Astragalus adsurgens | M.septentrionale | KM068066 | ACT34158 |
| 20 | Astragalus adsurgens | M.temperatum | KJ953896 | ACT34146 |
| 21 | Rhizosphere of Clitoria ternatea | M.thiogangeticum | MT197364 | - |
| 22 | Phaseolus vulgaris | R.giardinii | EU399693 | WP_018328749 |
| 23 | Oryza alta | R.oryzae | KM672535 | WP_085421349 |
| 24 | Phaseolus vulgaris | R.gallicum | MN181175 | WP_074070743 |
| 25 | Medicago | R.mongolense | EU256427 | WP_145611500 |
| 26 | Phaseolus vulgaris | R.etli | MW958080 | AAM54780 |
| 27 | Hedysarum,hedysari | R.sullae | FJ785219 | WP_027513895 |
| 28 | Phaseolus vulgaris | R.phaseoli | KP875549 | RUM15637 |
| 29 | Leucaena,Phaeolus vulgaris | R.tropici | FN178365 | BAU45324 |
| 30 | Desmodium | R.hainanense | MT409523 | WP_075853022 |
| 31 | Indigofera spp | R. indigoferae | MK872365 | WP_193446036 |
| 32 | Pisum sativam | R.leguminosarum bv | ON428643 | AUW47541.1 |
| 33 | Glycine max | Bradyrhizobium japonicum | HG518577 | BAU45315 |
| 34 | Phaseolus vulgaris | A.undicola | DQ648578 | WP_027488673 |

Table 1: List of the 34 *Rhizobial* species selected with their respective hosts the accession numbers of 16srRNA and "NodD" genes of the same *Rhizobial* species.In the table column of ACCESSION

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

NO.OF Nod D PROTEIN the highlighted yellow color indicates the Nod D protein of respective *Rhizobium* species and the other indicates the LysR regulatory protein of *Rhizobium* species.

**Maximum likelihood tree: 16srRNA tree:**

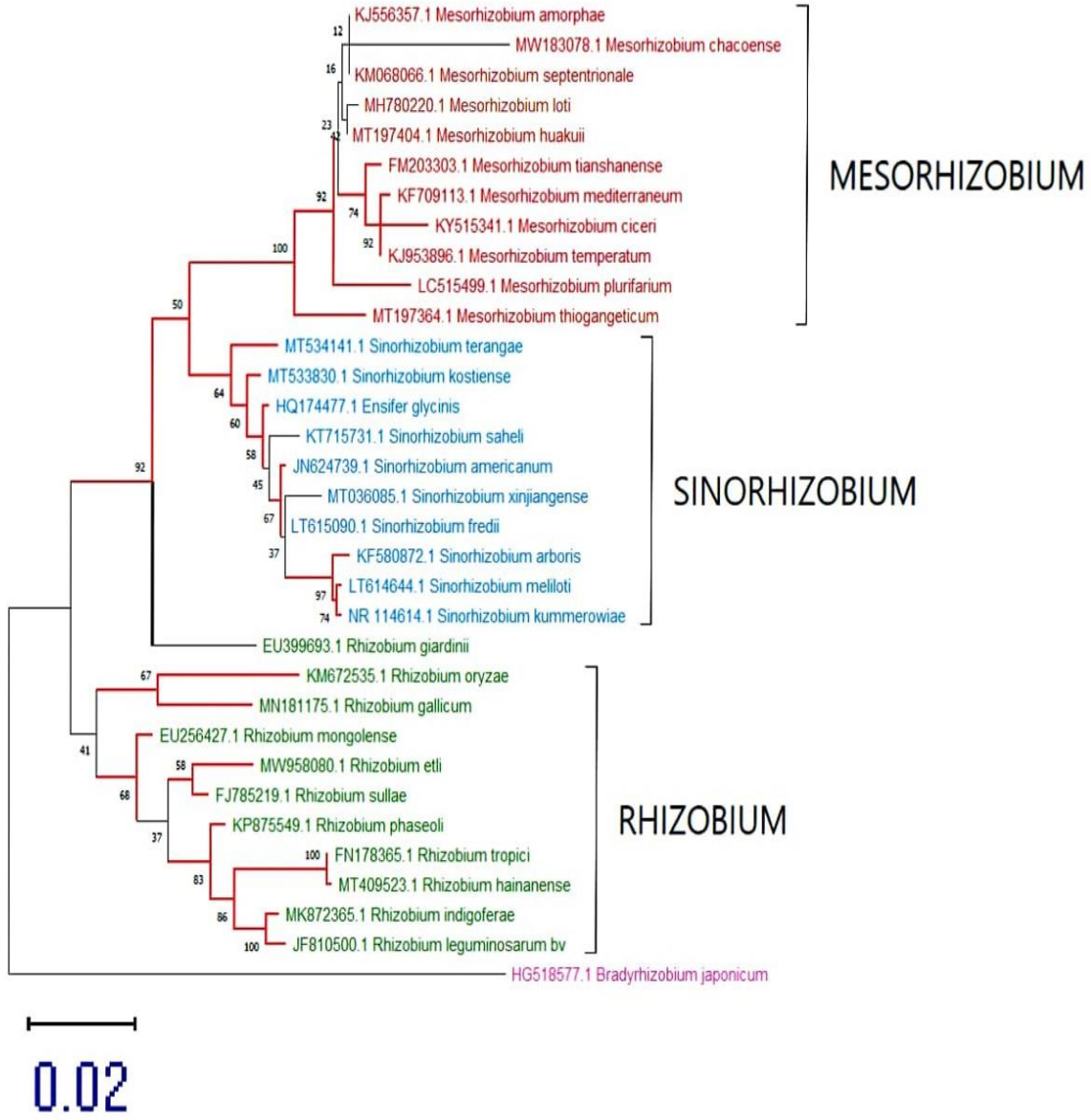


**Fig.1** Maximum likelihood tree of 16srRNA of the selected *Rhizobium* bacteria with *Bradyrhizobium japonicum* as an outgroup. Tree branches with more than 50 bootstrap valves are colored red and are chosen for further analysis.

The 16srRNA phylogenetic tree yielded three major clades where the genera *Mesorhizobium, Sinorhizobium,* and *Rhizobium* species were clustered in distinctly separate clades and were monophyletic. *Bradyrhizobium japonicum* was used as the outgroup and rooted in this branch.

Clades with higher bootstrap values are more likely to support the nodes in the phylogenetic tree. We used a cut-off of 50 percent bootstrap value to select the nodes for further analysis and ignored the nodes which showed lesser bootstrap values. **For example,** A node for the clade having *Rhizobium indigoferae* and *Rhizobium leguminosarum bv* has a bootstrap value of 100 and has strong support higher in the phylogenetic tree but a node having the clade having *Mesorhizobium huakii, Mesorhizobium loti,* and *Mesorhizobium chacoense* has a bootstrap value of 12 and therefore are likely candidates to be ignored.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

**NodD Phylogenetic tree: Boot strap tree of "NodD" (ML Tree):**


**Fig -2:** Maximum likelihood tree of Nod D tree of the *Rhizobium* consisting of three genera - *Rhizobium, Mesorhizobium, Sinorhizobium* species. In the phylogeny tree, the brown color tree indicates - *Mesorhizobium* species, the green color tree indicates - *Rhizobium* species, and the blue color indicates the *Sinorhizobium* species which are randomly arranged in the tree. The pink color indicates the *Bradyrhizobium japonicum*, it is an outgroup in the phylogeny tree. The tree branches with more than 50% values are colored with red color and analyzed further which strongly supports the phylogenetic tree.

We wanted to see how the "NodD" gene might have been evolving as compared to the 16srRNA tree. Since we know that the "Nod D" gene confers host specificity we wanted to find molecular changes in "NodD" sequences that could explain the reported host specificities. "NodD" gene sequences ranging from 250-350 amino acids were retrieved from Blastp taking *R.leguminosarum bv. viciae* as a query sequence. The sequences were trimmed at the beginning from 1-9bp and at the end from 287-316bp. Multiple sequence alignment was done using muscle alignment. This MSA was used to construct a maximum likelihood tree with a bootstrap value of 100.

The phylogenetic tree of the "NodD" protein has three unusual branch lengths of the species named *Rhizobium phaseoli, Mesorhizobium tianshanense, and Rhizobium undicola* compared to the "NodD" protein of other species. This might be because of two major possibilities. One is that the branch length of the "NodD" gene in these species is longer because they are not orthologous genes compared with the rest of the sequences in the tree. We had used sequences of the LysR "gene also in the analysis since blast results retrieved LysR sequences. LysR belongs to the transcriptional regulatory family. Second is the "NodD" gene might have been horizontally transferred between *Bradyrhizobium japonicum and Rhizobium tropici*.

**HGT**: First we analyzed the possibility of horizontal gene transfer. From the phylogeny of the "NodD" gene, we observed that the "NodD" gene virtually has no branch length difference between two different genera such as *R.tropici and Bradyrhizobium japonicum*. It is impossible that being two different genera the "NodD" gene could not have accumulated even a little variation. To test the possibility of horizontal gene transfer, we used a blastp tool to check the horizontal gene transfer between *Bradyrhizobium japonicum* and *Rhizobium tropici.*

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

First, we took the fasta sequence of the Nod D gene (Acc.No.BAU45324) of *Bradyrhizobium japonicum* and used it as a query, and blasted it against the organism *R.tropici*. From the results, we observed several different strains of *R .tropici* were reported for the same gene. This indicates that the gene is shared between many evolutionarily related strains and that the gene is vertically inherited among all the strains. This also proves that the "NodD" gene of *R. tropici* is not a mistake in the annotation. Secondly, we retrieved the fasta sequence of Nod D (Acc.no.BAU45324) of *R.tropici* and used it as a query, and blasted against the organism *Bradyrhizobium japonicum*. We obtained the same results as in the earlier case. The results listed several different strains of *Bradyrhizobium japonicum*, which means that the gene is shared between many evolutionary-related strains and is not a mistake in the annotation. Therefore we have strong reasons to believe that the "NodD" gene may have been horizontally transferred between *Rhizobium tropici* and *Bradyrhizobium japonicum.*

**REVERSE RECIPROCAL BLAST**: It might happen that "NodD" in some species might have been annotated as"LysR". So we tested whether LysR transcriptional regulator protein and "NodD" protein are the same or not. To find out we did a reverse reciprocal blast. We randomly choose two *Rhizobium* species - *M.loti* (Acc.no.ACT34148) which is annotated as "NodD" protein and another *A.undicola* (Acc.no.WP_027488673) which is annotated as LysR protein. We did a protein blast of the *M.loti* (taking as the query sequence) against the LysR protein of *A.undicola*.Then from the result of blastp, we took the first hit as the query sequence(Acc.no.WP_027487450) and blasted against the *M.loti*. From the blast results, we came to know that they are not the **orthologs** genes. It may be possible that Nod D might have been annotated as LysR because it belongs to the family of LysR transcriptional regulator family. However the pairwise alignment was done using the Needle (EMBOSS) which creates a global alignment of two sequences using the Needleman-Wunsch algorithm, showed an identity of 67% between LysR and NodD of *M.loti* and *A.undicola*-organism which is a definite indication that they are homologous.

**Host phylogenetic tree**

The host tree was constructed using the Newick format tool. The collected host plant species concerning the *Rhizobial* species were written in the following Newick format.

((Acacia, Prosopis)); ((Amorpha Fruticosa)); ((Indigofera),(Desmodium,Kummerwoia stipulacea, Glycine max) ,(Vigna, Phaseolus)); ((Sesbania, Lotus),(Astragalus,Hedysarum),( Galega, Cicer),(Medicago, Melilotus), ( Pisum sativum) );;((Oryza sativa));

Using the ETE toolkit, the Newick format was uploaded in the search bar and then viewed by the host phylogenetic tree. In the host tree, the Oryza sativa was made as the outgroup.

**Comparison of 16sr RNA and "NodD" tree**

1. In 16sr RNA the *Mesorhizobium, Sinorhizobium,* and *Rhizobium* clades are monophyletic whereas in the "NodD" tree the same three genera are found to be non-monophyletic. Each genus keeps interpolating with the other genera. We attempted to explain this mixing of different genera by comparing it with the host tree. But we did not find any definite pattern that is common to the topology between the Nod D and the host tree. Some host species were distantly related for which the "NodD " genes infecting bacteria appeared closely related, but the bootstrap values for such nodes were also less. For example, Species with low bootstrap values in the "NodD " ML tree such as *Mesorhizobium amorphae* and *Meosrhizobium septentrionale* is 46. Therefore this node is not strongly supported in the phylogenetic tree. In comparison with the host tree, it was found that the hosts namely; *Fruticosa ( Mesorhizobium amorphae), and Astragalus (Mesorhizobium Septentrionale)* were also found to be distantly related.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

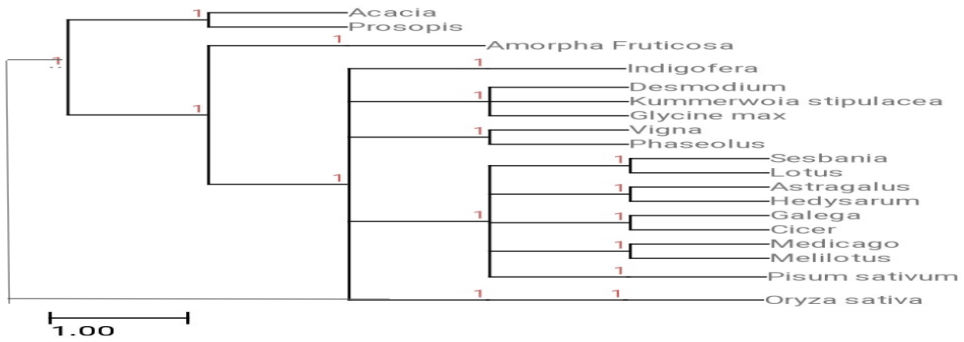


Fig-3: Tree constructed for the selected hosts using manually drawn Newick formats derived from published phylogenetic trees. The ETE toolkit was used to draw the host tree.

The above host phylogenetic tree contains the host species of *Rhizobium*, where *Oryza sativa* is made as an outgroup because it doesn't belong to the family of *Leguminosae*. We began by comparing the 16srRNA tree with that of the "NodD" tree. Then we attempted to rationalize any discrepancies found between the topology of the "NodD" tree, that of the 16 srRNA tree, and that of the host tree. The rationale here was that "NodD" genes being the determinants of host specificities the topology of the "NodD" tree may reflect the host tree. A long branch of the "NodD" gene in a species can be variation accumulated to adapt to a particular host which might be distantly related in its host tree.

2. *Rhizobium indigoferae* and *Rhizobium leguminosarum* have 100 bootstrap value as strong support in the phylogenetic tree of 16srRNA whereas in the "NodD" tree *Rhizobium indigoferae* have 78 and *Rhizobium leguminosarum* has 88 bootstrap values where these two species are placed distinctly.

**DISCUSSION :**

A two-gene comparative analysis was used to reconstruct phylogenetic relationships among 34 species of *Rhizobium*. Nucleotide sequences for two conserved genes (16S rRNA and Nod D gene) were analyzed, and phylogenetic relatedness estimates were developed with these sequences. Although certain groupings were consistent between the trees representing the different loci, some significant differences between the trees representing the different loci were evident. One of the most noticeable differences was the relative placement of the *Mesorhizobium tianshanense, Rhizobium phaseoli,* and *Allorhizobium undicola* in the different trees. 16sRNA phylogenetic tree yielded three major clades where the genera *Mesorhizobium, Sinorhizobium,* and *Rhizobium* species were clustered in distinctly separate clades and were monophyletic and further analyzed by bootstrap values above 50% that gives strong support to the phylogenetic study. In the phylogenetic tree of the "NodD " protein, the clades are randomly arranged and there are three unusual branch lengths of the species named *Rhizobium phaseoli, Mesorhizobium tianshanense, Rhizobium undicola.*

Through the study of the Horizontal gene transfer of *Bradyrhizobium japonicum,* we observed that genes are transferred with *Rhizobium tropici*. Therefore we have strong reasons to believe that the "NodD" gene may have been horizontally transferred between *Rhizobium tropici* and *Bradyrhizobium japonicum.* The NodD Phylogenetic tree was compared with the host tree(Newick tree) to study the evolutionary relationship between them. So to explain this mixing of different genera by comparing it with the host tree. But we did not find any definite correlation with topology patterns between the Nod D and the host tree.

**The implication of horizontal gene transfer:** Regional legume plants will have their specific repertoire of Rhizobium bacteria. This specificity is dictated by the selection process which is an outcome of the specific interactions between the host and the bacterial strain or species. However, these species might not be the best nitrogen fixers compared to the lab-developed strains. We can however genetically engineer a specific Nod D gene into the lab-developed Rhizobium inoculum. So

DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

that the host develops specificity to the introduced inoculum. Since we have already observed the Nod D gene horizontally transferred between unrelated species and thriving in them, we propose that a foreign Nod D gene could be very well accepted in the Rhizobium inoculum.

Our analysis only covers the Nod D gene. Along with the Nod D, other genes that are involved in the infection of the root nodules of the host plant such as Nod A, Nod B, and Nod C genes can also be used for further research in phylogenetic analysis and also for testing HGT scenarios.

The Host phylogenetic tree was constructed using the Newick format(ETE toolkit) sliced from phylogenetic trees of earlier studies. However, it will be more accurate and useful to construct the tree host tree using any of the standard Phylogenetic tree-building methods. Also if our study could be extended to more Bacterial species and host species it could prove more useful

## CONCLUSION:

The phylogenetic tree of the 16srRNA and Nod D genes are incongruent in many places. The Nod D gene has undergone an accumulation of variations corresponding to their host specificity owing to the demands of the selection process while the 16srRNA gene is neutrally evolving. However, these trees uncover a possible case of HGT. Such comparisons with other genes involved in the infection process may shed more light on the host specificity and any other HGTs.

## ACKNOWLEDGMENT:

We deeply appreciate the help and Mentoring from Dr. Saurabh Mahajan, Ph.D., NCBS and Assistant professor, Atria University, Bangalore, for regular inputs for this research.

## References:

**1.** Fauvart, M., & Michiels, J. (2008). Rhizobial secreted proteins as determinants of host specificity in the rhizobium–legume symbiosis. *FEMS microbiology letters*, *285*(1), 1-9.
**2.** Fernandes, C., & Ravi, L. (2023). Screening of symbiotic ability of Rhizobium under hydroponic conditions. *Microbial Symbionts*, 327.
**3.** Calla B. (2022). Friend or foe: How plants discriminate between pathogenic and mutualistic bacteria. *Plant Physiology*, *189*(4), 1893–1895.
**4.** Newman, M. A., Sundelin, T., Nielsen, J. T., & Erbs, G. (2013). MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Frontiers in plant science*, *4*, 139.
**5.** Via, V. D., Zanetti, M. E., & Blanco, F. (2016). How legumes recognize rhizobia. *Plant signaling & behavior*, *11*(2), e1120396.
**6.** Kumar, S., & Pandey, A. K. (2013). Chemistry and biological activities of flavonoids: an overview. *TheScientificWorldJournal*, *2013*, 162750.
**7.** Roche, P., Maillet, F., Plazanet, C., Debellé, F., Ferro, M., Truchet, G., ... & Dénarié, J. (1996). The common nodABC genes of Rhizobium meliloti are host-range determinants. *Proceedings of the National Academy of Sciences*, *93*(26), 15305-15310.
**8.** Dénarié, J., Debellé, F., & Promé, J. C. (1996). Rhizobium lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. *Annual review of biochemistry*, *65*(1), 503-535.
**9.** Saad, M. M., Kobayashi, H., Marie, C., Brown, I. R., Mansfield, J. W., Broughton, W. J., & Deakin, W. J. (2005). NopB, a type III secreted protein of Rhizobium sp. strain NGR234, is associated with pilus-like surface appendages. *Journal of Bacteriology*, *187*(3), 1173-1181.
**10.** Delgado, M. J., Casella, S., & Bedmar, E. J. (2007). Denitrification in rhizobia-legume symbiosis. In *Biology of the nitrogen cycle* (pp. 83-91). Elsevier.
**11.** Slonczewski, J. L., & Foster, J. W. (2013). *Microbiology: An evolving science: Third international student edition*. WW Norton & Company.
**12.** Liu, C. W., & Murray, J. D. (2016). The role of flavonoids in nodulation host-range specificity: an update. *Plants*, *5*(3), 33.

DOI:10.46647/ijetms.2023.v07i03.129 ISSN: 2581-4621

**13.** Eckardt, N. A. (2006). The role of flavonoids in root nodule development and auxin transport in Medicago truncatula.
**14.** Garg, N., & Geetanjali. (2009). Symbiotic nitrogen fixation in legume nodules: process and signaling: a review. *Sustainable agriculture*, 519-531.
**15.** Limpens, E., van Zeijl, A., & Geurts, R. (2015). Lipochitooligosaccharides modulate plant host immunity to enable endosymbioses. *Annual review of phytopathology*, *53*, 311-334.

**SUPPLEMENTARY SECTION:**
**https://docs.google.com/document/d/1nY8Qe7zWhTAneSB_3BZxhyoSAg_w602gwns_PMP_c68/edit?usp=drivesdk**