J Genet Resour 2024; 10(2): 44-53 *Homepage: http://sc.journals.umz.ac.ir/*

RESEARCH ARTICLE *DOI: 10.22080/jgr.2024.27764.1401*

J|G|R

Identification of New Genes Regulating Nodule Development in *Medicago truncatula*: An *In-silico* Approach

Mahboobeh Azarakhsh*, Sara Eslami, Maryam Moghadas and Zahra Ataei

Cell and Molecular Biology Department, Kosar University of Bojnord, Bojnord, Iran

ARTICLE INFO

Article history:
Received 01 April 2024
Accepted 17 May 2024
Available 28 June 2024

Keywords:
Gene expression
Medicago truncatula
Nodule development
Symbiotic nodules
Transcription factors

Supplementary information:
Supplementary information for this article is available at *http://sc.journals.umz.ac.ir/*

Corresponding authors:
✉ M. Azarakhsh
mahboobeazarakhsh@gmail.com
mahboobeazarakhsh@kub.ac.ir

p-ISSN 2423-4257
e-ISSN 2588-2589

ABSTRACT

Biological nitrogen fixation is a process by which atmospheric nitrogen can be converted to bioavailable nitrogen forms for plants. Among plants, legumes have the ability to fix nitrogen through symbiotic interaction with a specific group of bacteria called rhizobia. Because of this interaction, organs named nodules form on the plant roots. Nodule organogenesis on plant roots begins with the perception of Nod factor by plant root cells. Downstream of Nod factor perception, activation of transcription factor (TF) signaling cascade in root tissue takes place. These results in the formation of a structure named infection thread, the initiation of cell division in the root cortex, and finally, the formation of functional nodules. Due to the importance of biological nitrogen fixation in improving soil fertility, the molecular mechanism of nodule development has been studied extensively. Here, in order to identify new possible regulators that affect the formation of nodules during symbiosis in *Medicago truncatula*, the most significant symbiotic-related TFs, including MtIPD3, MtNSP1, MtNSP2, MtNIN, MtERN1, MtERN2, and MtERN3 were studied. Analyzing co-expressed genes with the Phytozome database and examining interaction networks with the STRING database have indicated potential new regulators involved in nodule development. According to our data, there is a high physical interaction score between IPD3 with some splicing factors and cell cycle proteins. This shows that IPD3, through interaction with these proteins, could be involved in the regulation of gene expression and cell cycle. Besides, we found a possible cytokinin transport gene ABCG38 and showed the activation of its expression in nodules compared to root. Moreover, it was revealed that the auxin response factor *Medtr2g043250* could be a direct target of the NIN transcription factor. The results on potential regulators of nodule organogenesis will pave the way for additional research.

Please cite this paper as: Azarakhsh, M., Eslami, S., Moghadas, M., Ataei, Z. (2024). Identification of new genes regulating nodule development in *Medicago truncatula*: an in-silico approach. *Journal of Genetic Resources*, 10(2), 44-53. doi: 10.22080/jgr.2024.27764.1401

Introduction

Nitrogen is an essential element for the growth and productivity of plants. Legumes have the ability to establish partnerships with nitrogen-fixing bacteria. This symbiotic relationship between plants and bacteria converts atmospheric nitrogen into ammonia, a nitrogen form that plants absorb. Therefore, the process of biological nitrogen fixation contributes to soil fertility and reduces the dependence on chemical nitrogen fertilizer, so it is important to sustainable agriculture (Aasfar *et al*., 2024; Chaulagain and Frugoli, 2021; Simpson *et al*., 2017). Considering the importance of biological nitrogen fixation in agriculture, studying the molecular mechanism of nodule formation and investigating the possibility of its transfer to non-legume plants has always been of interest to scientists (Huisman *et al*., 2018). Therefore, this

research aimed to study the molecular mechanism of nodule organogenesis.

Nodulation is initiated by lipo-oligosaccharides called nodulation factors (Nod factor or NF) secreted by rhizobia. In the root cells, Nod factors (NFs) are detected by LysM receptor kinases (LYKs), including MtLYK3/MtLYK4 and MtNFP.

Downstream of Nod factor perception by receptors, calcium influx and chloride and potassium efflux take place, which results in cytoskeletal changes and root hair curling (Charpentier *et al*., 2016; Endre *et al*., 2002; Lévy *et al*., 2004; Oldroyd, 2013; Stracke *et al*., 2002). Rhizobia entraps within curled root hair and enters the cell by breaking down the cell wall, moving through the epidermal root hair to the cortex, forming an infection-threat structure. Infection threads help rhizobia colonize root tissues and trigger nodules forming in root tissues below the surface (Ferguson *et al*., 2010; Garrocho-Villegas *et al*., 2007; Roy *et al*., 2020).

Moreover, downstream of Nod factor perception, a Ca^{2+}channel from the MtCNGC15 family (cyclic nucleotide-gated channel) form a complex, with the potassium-gated ion channel known as MtDMI1 (does not make infections 1) on the nuclear envelope (Charpentier *et al*., 2016). The formation of this complex leads to calcium oscillation in the nucleus. Calcium oscillation could be decoded by the nuclear-localized calcium and calmodulin-dependent protein kinase (CCaMK) MtDMI3, which will result in the activation of the transcription factor (TF) signaling cascade. Besides, the MtDMI3 protein, which was activated through calcium spiking, interacts with the MtIPD3 TF (interacting protein of DMI3) and activates it through phosphorylation (Messinese *et al*., 2007; Yano *et al*., 2008). The CYCLOPS, which is the orthologue of MtIPD3, could activate the NIN (nodule inception) and ERN1 (Ethylene Responsive Factor Required for Nodulation 1) gene by binding to their promotors (Liu *et al*., 2019; Singh *et al*., 2014). Downstream of Calcium response, there are MtNSP1 and MtNSP2 (nodulation signaling pathway 1 and 2) TF of GRAS family (Horváth *et al*., 2011; Ovchinnikova *et al*., 2011), which function as homodimers. It was shown that NSP1 could bind to the promoter of NIN and induce its expression (Hirsch *et al*., 2009). TFs, like ERN1, ERN2, and nuclear factor Y, which consists of NF-YA, NF-YB, and NF-YC, are located genetically after NIN. MtENOD11 and MtENOD12 are regulated by MtERN1 and MtNIN in a way that influences the growth of infection threads by controlling cell wall-related gene expression (Andriankaja *et al*., 2007).

Besides, NIN is sufficient for the activation of cytokinin receptor MtCRE1 in the root cortex (Vernié *et al*., 2015). MtCRE1 initially work in dividing cortical cells and later facilitates the transition between zones responsible for growth and specialization within mature nodules (Plet *et al*., 2011; Tirichine *et al*., 2007). The signaling pathway involving MtCRE1 triggers the activation of TFs related to nodule formation, including MtERN1, MtNSP2, and MtNIN (Fig. 1), as well as regulating the presence of MtPIN carriers that transport auxin (Cerri *et al*., 2016; Liu *et al*., 2019; Mortier *et al*., 2012; Murray *et al*., 2007; Ötvös *et al*., 2021; Plet *et al*., 2011; Tirichine *et al*., 2007).

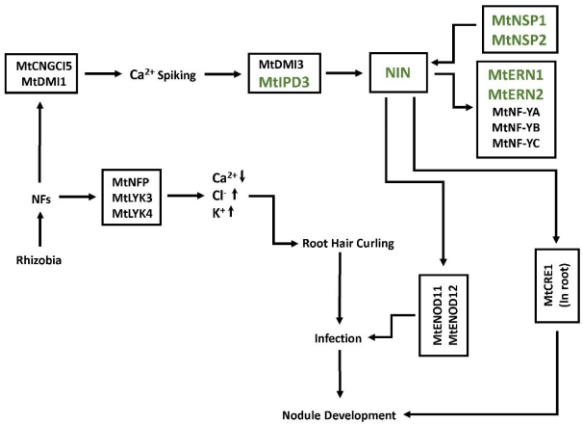

Fig. 1. The schematic representation of transcription factors signaling cascade during nodule organogenesis.

The plants' hormones cytokinin and auxin and their interaction with TFs are required for the development of symbiotic nodule development (Azarakhsh and Lebedeva, 2023; Lebedeva *et al*., 2021).

The aim of this work was to identify a new possible regulator of symbiotic nodule development. Therefore, the most significant TFs during nodule organogenesis, including MtIPD3, MtNSP1, MtNSP2, MtNIN, MtERN1, MtERN2, and MtERN3 were selected. Then the

most co-expressed genes were studied with this TFs using phytozome database. Secondly, the interaction of these proteins was investigated. Taken together, the results will provide ideas for further research.

Materials and Methods

Transcription factors and Co-expression

The main TFs involved in nodulation in *Medicago truncatula*, according to Chakraborty *et al*. (2022), were selected for the co-expression and interaction analysis (Table 1). Co-expressed genes with the selected TFs (Table 1) were obtained using the Phytozome database (https://phytozome-next.jgi.doe.gov/). The Phytozome database provides access to sequenced and annotated *Medicago truncatula* whole genome, as well as information about co-expressed genes, homolog genes, and metabolic pathways (Goodstein *et al*., 2012).

For more information about specific criteria used for building co-expressed clusters in Phytozome, please see Sreedasyam *et al*. (2023). The annotation and gene description of the co-expressed genes, together with their expression level in nodule and leaf (symbiotic condition), were listed in the supplement Excel file.

Table 1. TFs that were studied in this work.

	TF	Gene ID in Phytozome	Gene ID in NCBI
1	MtIPD3 (Interacting Protein of DMI3)	*Medtr5g026850*	AES72209.1
2	MtNSP1 (Nodulation Signaling Pathway 1)	*Medtr8g020840.1*	KEH36774.1
3	MtNSP2 (Nodulation Signaling Pathway 2)	*Medtr4g068000*	AES89009.1
4	MtNIN (Nodule INception)	*Medtr5g099060*	AES01067.1
5	MtERN1 (Ethylene Responsive Factor Required for Nodulation 1)	*Medtr7g085810*	AES80852.2
6	MtERN2 (Ethylene Responsive Factor Required for Nodulation 2)	*Medtr6g029180*	AES75163.1
7	MtERN3 (Ethylene Responsive Factor Required for Nodulation 3)	*Medtr8g085960*	AES82167.1

Interaction analysis

Using the NCBI gene ID (Table 1), proteins interacting with the target TFs were found on the STRING database (https://string-db.org/). The physical and functional association scores in the STRING database originated from various sources, including automated text mining of the scientific literature, computational interaction predictions, conserved genomic context, databases of interaction experiments, and known complexes/pathways from curated sources (Szklarczyk *et al*., 2023). The gene ID of interacting proteins, together with their functional and physical interaction scores extracted from STRING, are provided in the supplement Excel file. Then, the annotation corresponding to each gene was found on the Phytozome website (https://phytozome-next.jgi.doe.gov/). Functional and physical interaction scores are indicators of confidence; they show how likely an interaction could take place according to the available evidence. These scores do not indicate the strength or the specificity of the interaction. Confidence limits are as follows: low confidence is 0.15-0.4, medium confidence is 0.4-0.7, high confidence is 0.7-0.9, and highest confidence is 0.9-1.

Motif search in the promoter of interested genes

Using the Phytozome database (https://phytozome-next.jgi.doe.gov/), the 2000 bp upstream of the coding sequences of the interested genes were selected for the motif search. The promoters are usually 100-1000bp long and located upstream of the transcription start site (Le *et al*., 2019). Besides, some experimental data show that regulatory sequences could also be located in 5'UTR (Srivastava *et al*., 2018). To ensure that all promoter sequences were analyzed, we selected 2000bp upstream of the ATG site, which includes the 5'UTR region and upstream regulatory sequences.

The NIN binding sites (NBS) were identified before using the Chromatin immunoprecipitation (or ChIP) technique (Laffont *et al*., 2020; Soyano *et al*., 2014; Soyano *et al*., 2013). All identified NBS were mentioned by Chakraborty et al. (2022). We used these NBSs and found the following consensus motif in these sequences by MEME suit (https://meme-suite.org/): ACYYTTKRRSNTHANMAARGG (Bailey and Elkan, 1994), which showed the Nucleotide

IUPAC Alphabet on MEME suite web site (https://meme-suite.org/meme/doc/iupac.html).

Table 2. Interaction network of MtIPD3 and MtNIN.

	Gene identifier	Gene ID	Functional Interaction*	Physical Interaction	Annotation
IPD3	NSP2	Medtr3g072710	0.922 (very high)	0.356 (exploratory)	Scarecrow-like protein 26
	CCAMK	Medtr8g043970	0.969 (very high)	0.826 (high)	Serine/Threonine-Protein Kinase CG17528
	G7JHS5_MEDTR	Medtr4g116430	0.932 (very high)	0.602 (medium)	U5 small nuclear ribonucleoprotein component
	A0A072VS66	Medtr1g038815	0.952 (very high)	0.798 (high)	Intron-binding protein Aquarius
	G7J6U4_MEDTR	Medtr3g091250	0.933 (very high)	0.582 (medium)	Splicing factor 3A subunit 3
	G7K4N4_MEDTR	Medtr5g095510	0.976 (very high)	0.799 (high)	mRNA splicing factor
	G7IMA2_MEDTR	Medtr2g104440	0.923 (very high)	0.601 (medium)	Small nuclear ribonucleoprotein E
	G7J609_MEDTR	Medtr3g079200	0.973 (very high)	0.876 (high)	Cell cycle control protein
	A0A072VRJ1	Medtr1g112640	0.928 (very high)	0.600 (medium)	GCIP-interacting protein P29
	G7JIW9_MEDTR	Medtr4g100970	0.944 (very high)	0.776 (high)	Cell cycle control protein CWF15
	G7LCC3_MEDTR	Medtr8g106150	0.693 (medium)	-	Histidine kinase / Protein kinase (histidine) // Protein-serine/threonine phosphatase / Serine/threonine-specific protein phosphatase
	G7I6D5_MEDTR	Medtr1g011850	0.690 (medium)	-	Histone 2A
	A0A072VLP8	Medtr1g070100	0.777 (high)	-	Hyoscyamine (6S)-dioxygenase / Hyoscyamine 6-beta-hydroxylase
	G7I8J1_MEDTR	Medtr1g056530	0.728 (high)	-	Nuclear transcription factor Y subunit A-10-related
NIN	G7IDR2_MEDTR	Medtr1g087140	0.690 (medium)	-	Histidine kinase CKI1
	A0A072VZ96	Medtr1g090850	0.690 (medium)	-	Response regulator receiver domain (Response_reg) // Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase (HATPase_c)
	A0A072V405	Medtr3g110840	0.698 (medium)	-	Non-specific serine/threonine protein kinase / Threonine-specific protein kinase
	G7JYE9_MEDTR	Medtr5g066450	0.693 (medium)	-	Hydroxyproline-rich Glycoprotein family Protein
	A0A072V7K3	Medtr2g043250	0.690 (medium)	-	Auxin response factor 19- related
	A0A072VC05	Medtr2g093740	0.690 (medium)	-	Auxin response factor 4

*Medium, high and very high interaction scores correspond to 0.4-0.7, 0.7-0.9 and 0.9-1 scores, respectively. The interaction results for other TFs can be found in the supplement Excel file.

Motif scanning was performed by FIMO (Grant *et al*., 2011) with the provided consensus motif in the promoter of possible NIN targets (Fig. 2). FIMO identifies motif occurrences with a p-value less than 0.0001.

Result and discussion

Identification of co-expressed genes

To identify new elements that impact the formation of nodules during symbiosis in *Medicago truncatula*, we selected the most significant known symbiotic TFs, including MtIPD3, MtNSP1, MtNSP2, MtNIN, MtERN1, MtERN2, and MtERN3 (Chakraborty *et al*., 2022). Next, the most co-expressed genes with these TFs were identified (Supplement Excel file).

NIN is a nodulation-specific TF that regulates diverse programs during nodule development (Vernié *et al*., 2015). Among NIN co-expressing genes, there is a nuclear transcription factor *Y subunit* gene (*Medtr8g019540*). In the Symbimics database (https://iant.toulouse.inra.fr/symbimics/), this gene was annotated as *MtNF-YA8* (Roux *et al*., 2014). Besides, according to Symbimics expression profile, *MtNF-YA8* demonstrate activation of expression in nodules in comparison with roots (Fig. 3A).

In *Lotus japonicus*, *LjNF-YA1* and *LjNF-YB1* were identified as direct targets of NIN TF (Soyano *et al*., 2013). In *Medicago truncatula*, it was shown that *MtNF-YA1* could be the transcriptional target of NIN (Laffont *et al*., 2020). We searched in the promoter of the *MtNF*-YA8 gene for the possible consensus NIN binding site (Fig. 2). The consensus NBS was not found in the promoter of the *MtNF-YA8*

gene, so we cannot classify this gene as a direct target of the NIN TF.

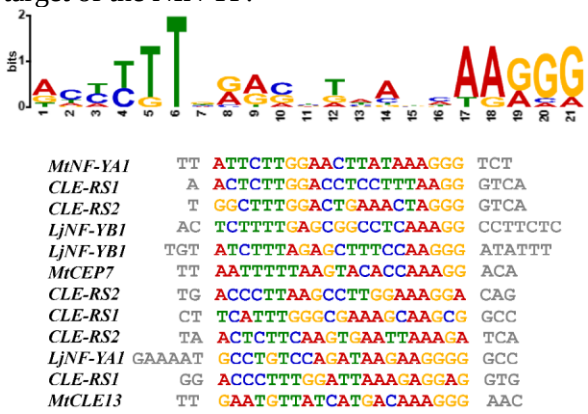

Fig. 2. Consensus NIN binding site was identified by MEME using NBSs previously identified in *Medicago truncatula* and *Lotus japonicas* (All the previously identified NBSs in the promoter of target genes were used to identify the consensus NBS, then this consensus motif was used to identify new NBS in the promoter of the possible target genes).

Among NSP1 co-expressing genes, there is an ABC transporter G family member 38 (MtABCG38), which was shown in the supplementary data file. We further analyzed ABCG38 expression in nodules in comparison with roots according to Symbimics and found induction of ABCG38 in nodules in comparison with the root (Fig. 3B). MtABCG56 transporter, which is the paralog of MtABCG38, mediates cytokinin transport in early stages of nodule organogenesis (Jarzyniak *et al*., 2021). Cytokinin could induce cortical cell division required for nodule development (Bauer *et al*., 1996). Nevertheless, cytokinin-induced nodule formation is absent in *nsp1* mutants (Heckmann *et al*., 2011), suggesting that cytokinin-induced primordia formation relies on NSP1 TF.

While cytokinin is an important hormone for nodule formation, little is known about its transport during nodule formation (Azarakhsh and Lebedeva, 2023). Considering the diverse and extensive role of cytokinin during nodule development, it could be expected that its transport should be regulated precisely. According to these pieces of evidence, it can be hypothesized that ABCG38, like its paralog gene (ABCG56), might be involved in cytokinin transport.

Identification of interaction network of TFs

This study also investigated the interaction network of symbiotic TFs IPD3, NSP1, NSP2, NIN, ERN1, ERN2, and ERN3 by STRING (supplement Excel file).

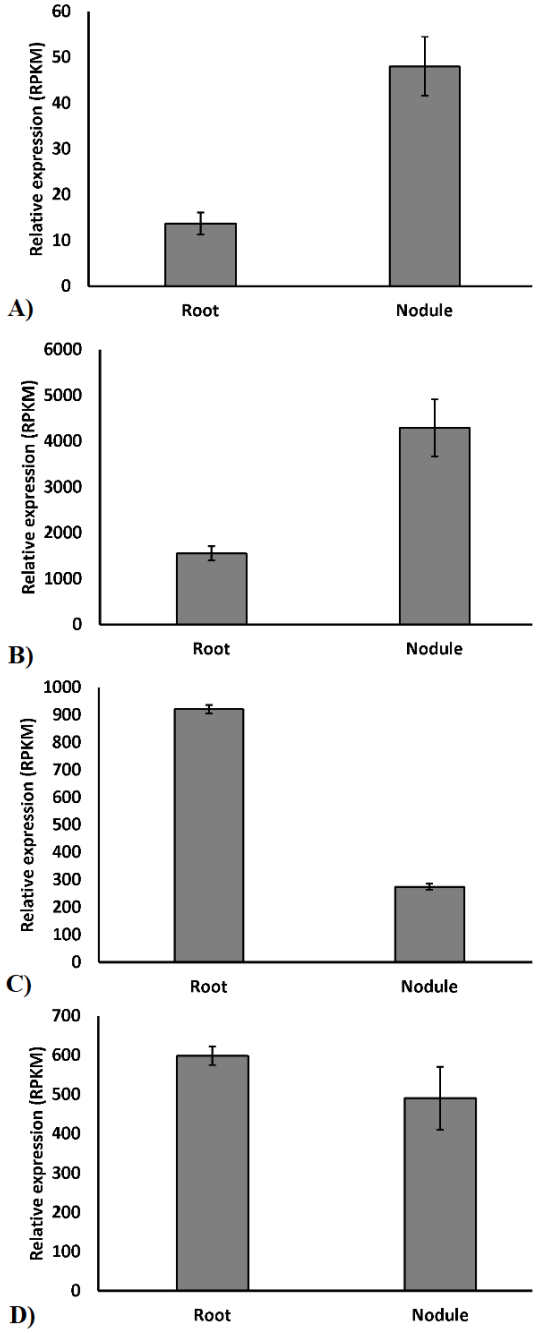

Fig. 3. Gene expression in nodule in comparison with the root (control) in *Medicago truncatula*: A) *Medtr8g019540 (NF-YA8)*; B) *Medtr7g091880 (ABCG38)*; C*) Medtr2g043250 (ARF19)* and D) *Medtr2g093740 (*ARF4). RPKM is reads per kilobase of transcript per Million mapped reads (The adjusted P value threshold is 0.01).

MtIPD3 and its orthologue in *Lotus japonicus* are master regulators of symbiotic nodule development (Horváth *et al*., 2011; Ovchinnikova *et al*., 2011). The interaction of MtIPD3 with DMI3, which is a calcium calmodulin-dependent kinase (CCAMK), was shown previously (Messinese *et al*., 2007). These data are in line with our results on MtIPD3 interactions (Table 2). MtIPD3 and DMI3 (CCAMK) demonstrate a physical interaction score of 0.826. Interestingly, MtIPD3 also demonstrate a high physical interaction score with splicing factors Medtr5g095510, Medtr3g091250, and Medtr2g104440 with a physical interaction score of 0.799, 0.582, and 0.601 (Table 2). According to the Phytozome database, these splicing factors are highly expressed in symbiotic nodules (data are not shown). Therefore, MtIPD3, through interaction with the splicing factors, could be involved in the regulation of gene expression during nodule organogenesis.

Besides, MtIPD3 show high physical interaction scores of 0.876 and 0.776 with cell cycle proteins Medtr3g079200 and Medtr4g100970, respectively. Medtr3g079200 and Medtr4g100970 also express in nodules (data are not shown). Therefore, MtIPD3 could be involved in the control of cell cycle during nodule development.

According to the results, MtIPD3 also demonstrate very high functional interaction with NSP2 (0.922), while the possibility of their physical interaction is relatively low. It was shown previously that MtDELLA protein could bridge IPD3 and NSP2 proteins (Jin et al., 2016), which is in agreement with the findings on low physical interaction score of IPD3 and NSP2.

As it was mentioned before, LjNF-YA1, LjNF-YB1, and MtNF-YA1 were identified as direct targets of NIN (Laffont *et al*., 2020; Soyano *et al*., 2013). Here, MtNF-YA1 TF (Medtr1g056530) was identified among the genes with high functional interaction with NIN (0.728 scores). Interestingly, two possible NBS were found in the promoter of Mt*NF-YA1*, which is in agreement with previous findings (Table 3). This observation suggests that the conserved motif used in this research (see materials and method section) could be properly used for the identification of NBSs.

Table 3. Searching for NIN binding site (NBS) in the promoter of *Medtr1g056530* and *Medtr2g043250* genes using FIMO.

Motif ID	Strand	Start	End	p-value*	Matched Sequence
Medtr1g056530					
ACYYTTKRRSNTHANMAARGG	+	2194	2214	9.66e-07	ATTCTTGGAACTTATAAAGG
ACYYTTKRRSNTHANMAARGG	+	164	184	4.34e-05	ACTTTTAGGGATGAACAAGAA
Medtr2g043250					
ACYYTTKRRSNTHANMAARGG	+	2868	2888	7.52e-05	AATCTTGTTGCTAAGAAAAAG

*p-value ≤ 0.0001

It is known that NIN is necessary for the induction of auxin biosynthesis genes STYLISH (STY) and YUCCAs (YUC) (Schiessl et al., 2019). Among NIN interacting proteins, two auxin response factors (*Medtr2g043250* and *Medtr2g093740*) were found. The expression of these two genes is downregulated in nodules (Fig. 3C and 3D). This study further investigated the possibility of direct binding of NIN as a TF to the promoter of auxin response factors (*Medtr2g043250* and *Medtr2g093740*). We used the consensus NIN binding sites (Fig. 2) and searched for them in the promoter of auxin response factors. Interestingly, a NIN binding site was found in the promoter of auxin response factor *Medtr2g043250* (*ARF19*) that was identified initially in the NIN interaction network (Table 3). These data suggest that auxin response factors *ARF19* might be a direct target of NIN TF. However, further *in-vivo* and *in-vitro* studies are required to confirm this finding.

Conclusion

Legumes are integral to sustainable farming practices due to their ability to improve soil fertility naturally. Therefore, studying the molecular mechanism of nodule formation is of great importance. The aim of this *in-silico* analysis was to identify new possible regulators of symbiotic nodule development in *Medicago*

truncatula. Among NSP1 co-expressing genes, we found a possible cytokinin transport gene ABCG38, whose expression was induced in nodules in comparison with the roots. A high possibility of physical interaction was also found between IPD3 and some splicing factors and cell cycle proteins, suggesting the possible involvement of IPD3 in gene expression and cell cycle control. Moreover, using motif search and expression analysis, we found the auxin response factor *Medtr2g043250* to be a possible target of NIN TF. The results, taken together, provide insight into possible new regulators of nodule organogenesis. Since the results of this study were based on *in-silico* studies, further *in-vivo* and *in-vitro* studies are required to confirm them.

Conflict of Interests

There is no conflict of interest associated with this paper.

Authors' Contribution

Conceptualization of the presented idea: MA; investigation, data curation: MA, SE, MM, and ZA; original draft preparation: MA, SE; writing-review and editing: MA.

References

Aasfar, A., Meftah Kadmiri, I., Azaroual, S. E., Lemriss, S., Mernissi, N. E., Bargaz, A., ... & Hilali, A. (2024). Agronomic advantage of bacterial biological nitrogen fixation on wheat plant growth under contrasting nitrogen and phosphorus regimes. *Frontiers in Plant Science*, 15, 1388775. https://doi.org/10.3389/fpls.2024.1388775

Andriankaja, A., Boisson-Dernier, A., Frances, L., Sauviac, L., Jauneau, A., Barker, D. G., & de Carvalho-Niebel, F. (2007). AP2-ERF transcription factors mediate Nod factor-dependent Mt ENOD11 activation in root hairs via a novel cis-regulatory motif. *The Plant Cell*, 19(9), 2866-2885. https://doi.org/10.1105tpc.107.052944

Azarakhsh, M., & Lebedeva, M. A. (2023). Lateral root versus nodule: the auxin-cytokinin interplay. *Journal of Plant Growth Regulation*, 42(11), 6903-6919. https://doi.org/10.1007/s00344-023-10983-4

Bailey, T. L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in bipolymers. *Proceedings - International Conference on Intelligent Systems for Molecular Biology*, 2, 28-36.

Bauer, P., Ratet, P., Crespi, M. D., Schultze, M., & Kondorosi, A. (1996). Nod factors and cytokinins induce similar cortical cell division, amyloplast deposition and MsEnod12A expression patterns in alfalfa roots. *The Plant Journal*, 10(1), 91-105. https://doi.org/10.1104/pp.116.1.53

Cerri, M. R., Frances, L., Kelner, A., Fournier, J., Middleton, P. H., Auriac, M. C., ... & de Carvalho-Niebel, F. (2016). The symbiosis-related ERN transcription factors act in concert to coordinate rhizobial host root infection. *Plant Physiology*, 171(2), 1037-1054. https://doi.org/10.1104/pp.16.00230

Chakraborty, S., Valdés-López, O., Stonoha-Arther, C., & Ané, J. M. (2022). Transcription factors controlling the rhizobium-legume symbiosis: integrating infection, organogenesis and the abiotic environment. *Plant and Cell Physiology*, 63(10), 1326-1343. https://doi.org/10.1093/pcp/pcac063

Charpentier, M., Sun, J., Martins, T. V., Radhakrishnan, G. V., Findlay, K., Soumpourou, E., ... & Oldroyd, G. E. (2016). Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science*, 352(6289), 1102-1105. https://doi.org/10.1126/science.aae0109

Chaulagain, D., & Frugoli, J. (2021). The regulation of nodule number in legumes is a balance of three signal transduction pathways. *International Journal of Molecular Sciences*, 22(3), 1117. https://doi.org/10.3390/ijms22031117

Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P., & Kiss, G. B. (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature*, 417(6892), 962-966. https://doi.org/10.1038/nature00842

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M. H., Lin, Y. H., Reid, D. E., & Gresshoff, P. M. (2010). Molecular analysis of legume nodule development and autoregulation. *Journal of Integrative Plant Biology*, 52(1), 61-76. https://doi.org/10.1111/j.1744-7909.2010.00899.x

Garrocho-Villegas, V., Gopalasubramaniam, S. K., & Arredondo-Peter, R. (2007). Plant hemoglobins: what we know six decades after their discovery. *Gene*, 398(1-2), 78-85. https://doi.org/10.1016/j.gene.2007.01.035

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., ... & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1), D1178-D1186. https://doi.org/10.1093/nar/gkr944

Grant, C. E., Bailey, T. L., & Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics*, 27(7), 1017-1018. https://doi.org/10.1093bioinformatics/btr064

Heckmann, A. B., Sandal, N., Bek, A. S., Madsen, L. H., Jurkiewicz, A., Nielsen, M. W., ... & Stougaard, J. (2011). Cytokinin induction of root nodule primordia in Lotus japonicus is regulated by a mechanism operating in the root cortex. *Molecular Plant-Microbe Interactions*, 24(11), 1385-1395. https://doi.org/10.1094MPMI-05-11-0142

Hirsch, S., Kim, J., Munoz, A., Heckmann, A. B., Downie, J. A., & Oldroyd, G. E. (2009). GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in *Medicago truncatula*. *The Plant Cell*, 21(2), 545-557. https://doi.org/10.1105/tpc.108.064501

Huisman, J., Codd, G. A., Paerl, H. W., Ibelings, B. W., Verspagen, J. M. H., & Visser, P. M. (2018). Cyanobacterial blooms. *Nature Reviews Microbiology*, 16(8), 471-483. https://doi.org/10.1038/s41579-018-0040-1

Horváth, B., Yeun, L. H., Domonkos, Á., Halász, G., Gobbato, E., Ayaydin, F., ... & Kaló, P (2011). *Medicago truncatula* IPD3 is a member of the common symbiotic signaling pathway required for rhizobial and mycorrhizal symbioses. *Molecular plant-microbe interactions*, 24(11), 1345-1358. https://doi.org/10.1094/MPMI-01-11-0015

Jarzyniak, K., Banasiak, J., Jamruszka, T., Pawela, A., Di Donato, M., Novák, O., ... & Jasiński, M. (2021). Early stages of legume–rhizobia symbiosis are controlled by ABCG-mediated transport of active cytokinins. *Nature Plants*, 7(4), 428-436. https://doi.org/10.1038/s41477-021-00873-6

Jin, Y., Liu, H., Luo, D., Yu, N., Dong, W., Wang, C., et al. (2016). DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. *Nature Communications*, 7(1), 12433. https://doi.org/10.1038/ncomms12433

Laffont, C., Ivanovici, A., Gautrat, P., Brault, M., Djordjevic, M. A., & Frugier, F. (2020). The NIN transcription factor coordinates CEP and CLE signaling peptides that regulate nodulation antagonistically. *Nature Communications*, 11(1), 3167. https://doi.org/10.1038/s41467-020-16968-1

Le, N. Q. K., Yapp, E. K. Y., Nagasundaram, N., & Yeh, H.-Y. (2019). Classifying promoters by interpreting the hidden information of DNA sequences via deep learning and combination of continuous fasttext N-grams. *Frontiers in Bioengineering and Biotechnology*, 7, 305. https://doi.org/10.3389/fbioe.2019.00305

Lebedeva, M., Azarakhsh, M., Sadikova, D., & Lutova, L. (2021). At the root of nodule organogenesis: conserved regulatory pathways recruited by rhizobia. *Plants*, 10(12), 2654. https://doi.org/10.3390/plants10122654

Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., ... & Debellé, F (2004). A putative Ca^{2+} and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science*, 303(5662), 1361-1364. https://doi.org/10.1126/science.1093038

Liu, C. W., Breakspear, A., Guan, D., Cerri, M. R., Jackson, K., Jiang, S., ... & Murray, J. D. (2019). NIN acts as a network hub controlling a growth module required for rhizobial infection. *Plant Physiology*, 179(4), 1704-1722. https://doi.org/10.1104/pp.18.01572

Liu, J., Rutten, L., Limpens, E., Van Der Molen, T., Van Velzen, R., Chen, R., ... & Bisseling, T. (2019). A remote cis-regulatory region is required for NIN expression in the pericycle to initiate nodule primordium formation in *Medicago truncatula*. *The Plant Cell*, 31(1), 68-83. https://doi.org/10.1105/tpc.18.00478

Messinese, E., Mun, J. H., Yeun, L. H., Jayaraman, D., Rougé, P., Barre, A., ... & Ané, J. M. (2007). A novel nuclear protein interacts with the symbiotic DMI3 calcium-

and calmodulin-dependent protein kinase of *Medicago truncatula*. *Molecular Plant-Microbe Interactions*, 20(8), 912-921. https://doi.org/10.1094/MPMI-20-8-0912

Mortier, V., Holsters, M., & Goormachtig, S. (2012). Never too many? How legumes control nodule numbers. *Plant Cell & Environment*, 35(2), 245-258. https://doi.org/10.1111/j.1365-3040.2011.02406.x

Murray, J. D., Karas, B. J., Sato, S., Tabata, S., Amyot, L., & Szczyglowski, K. (2007). A cytokinin perception mutant colonized by Rhizobium in the absence of nodule organogenesis. *Science*, 315(5808), 101-104. https://doi.org/10.1126/science.1132514

Oldroyd, G. E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology*, 11(4), 252-263. https://doi.org/10.1038/nrmicro2990

Ötvös, K., Marconi, M., Vega, A., O'Brien, J., Johnson, A., Abualia, R., ... & Benková, E. (2021). Modulation of plant root growth by nitrogen source-defined regulation of polar auxin transport. *The EMBO Journal*, 40(3), e106862. https://doi.org/10.15252/embj.2020106862

Ovchinnikova, E., Journet, E. P., Chabaud, M., Cosson, V., Ratet, P., Duc, G., ... & Limpens, E. (2011). IPD3 controls the formation of nitrogen-fixing symbiosomes in pea and Medicago Spp. *Molecular Plant-Microbe Interactions*, 24(11), 1333-1344. https://doi.org/10.1094/MPMI-01-11-0013

Plet, J., Wasson, A., Ariel, F., Le Signor, C., Baker, D., Mathesius, U., ... & Frugier, F. (2011). MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in *Medicago truncatula*. *The Plant Journal*, 65(4), 622-633. https://doi.org/10.1111/j.1365-313X.2010.04447.x

Roux, B., Rodde, N., Jardinaud, M. F., Timmers, T., Sauviac, L., Cottret, L., ... & Gamas, P. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *The Plant Journal*, 77(6), 817-837. https://doi.org/10.1111/tpj.12442

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., ... & Udvardi, M. K. (2020). Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. *The Plant Cell*, 32(1), 15-41. https://doi.org/10.1105/tpc.19.00279

Schiessl, K., Lilley, J. L., Lee, T., Tamvakis, I., Kohlen, W., Bailey, P. C., ... & Oldroyd, G. E. (2019). Nodule inception recruits the lateral root developmental program for symbiotic nodule organogenesis in *Medicago truncatula*. *Current Biology*, 29(21), 3657-3668. e3655. https://doi.org/10.1016/j.cub.2019.09.005

Simpson, C., Mur, L. A., Gupta, A. K., Kumari, A., & Gupta, K. J. (2017). Moving nitrogen to the centre of plant defence against pathogens. *Annals of Botany*, 119(5), 703-709. https://doi.org/10.1093/aob/mcw179

Singh, S., Katzer, K., Lambert, J., Cerri, M., & Parniske, M. (2014). CYCLOPS, a DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. *Cell Host and Microbe*, 15(2), 139-152. https://doi.org/10.1016/j.chom.2014.01.011

Soyano, T., Hirakawa, H., Sato, S., Hayashi, M., & Kawaguchi, M. (2014). NODULE INCEPTION creates a long-distance negative feedback loop involved in homeostatic regulation of nodule organ production. *Proceedings of the National Academy of Sciences*, 111(40), 14607-14612. https://doi.org/10.1073/pnas.1412716111

Soyano, T., Kouchi, H., Hirota, A., & Hayashi, M. (2013). Nodule inception directly targets NF-Y subunit genes to regulate essential processes of root nodule development in Lotus japonicus. *PLoS Genetics*, 9(3), e1003352. https://doi.org/10.1371/journal.pgen.1003352

Sreedasyam, A., Plott, C., Hossain, M. S., Lovell, J. T., Grimwood, J., Jenkins, J. W., ... & Schmutz, J. (2023). JGI Plant Gene Atlas: an updateable transcriptome resource to improve functional gene descriptions across the plant kingdom. *Nucleic acids research*, 51(16), 8383-8401. https://doi.org/10.1093/nar/gkad616

Srivastava, A. K., Lu, Y., Zinta, G., Lang, Z., & Zhu, J. K. (2018). UTR-dependent control of gene expression in plants. *Trends in Plant Science*, 23(3), 248-259. https://doi.org/10.1016/j.tplants.2017.11.003

Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., ... & Parniske, M. (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature*, 417(6892), 959-962. https://doi.org/10.1038/nature00841

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., ... & Von Mering, C. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, 51(D1), D638-D646. https://doi.org/10.1093/nar/gkac1000

Tirichine, L., Sandal, N., Madsen, L. H., Radutoiu, S., Albrektsen, A. S., Sato, S., et al. (2007). A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science*, 315(5808), 104-107. https://doi.org/10.1126/science.1132397

Vernié, T., Kim, J., Frances, L., Ding, Y., Sun, J., Guan, D., ... & Oldroyd, G. E. (2015). The NIN transcription factor coordinates diverse nodulation programs in different tissues of the *Medicago truncatula* root. *The Plant Cell*, 27(12), 3410-3424. https://doi.org/10.1105/tpc.15.00461

Yano, K., Yoshida, S., Müller, J., Singh, S., Banba, M., Vickers, K., ... & Parniske, M (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proceedings of the National Academy of Sciences*, 105(51), 20540-20545. https://doi.org/10.1073/pnas.0806858105