JEEAR, Vol. 2, 2020

RESEARCH ARTICLE

DOI: 10.34002/jeear.v2i0.44

Characterization of Genes Involved in the Autoregulation of Nodulation in Chickpea (*Cicer arietinum* L.)

Muhammad Shah Jahan[1], Muhammad Ibrahim Tahir[1], Muhammad Inam-Ul-Haq[1], Umbreen Shahzad[2], and Phoebe Nemenzo Calica[3]*

[1]Department of Plant Pathology, PMAS-Arid Agriculture University, Rawalpindi, Pakistan
[2]College of Agriculture, Bahauddin Zakariya University, Bahadur Sub-Campus Layyah, Pakistan
[3]Environmental Science Department, Ateneo de Davao University, Davao City, Philippines
*Corresponding author: pncalica@addu.edu.ph

Received: September 16, 2019; Accepted: August 20, 2020

Abstract

Establishment of an excess number of nodules markedly affects plant growth and development due to overconsumption of photosynthates for nitrogen fixation. Hence, the total root nodule number must be tightly controlled via a negative feedback mechanism (AON) to maintain an optimal nitrogen and carbon balance in chickpea plants. Chickpea genes which play important regulatory roles in root nodulation in chickpea *(Cicer arietinum)* were identified by aligning these genes with known genes of legumes available at phytozome through BLAST search. Chickpea nodulation genes identified and characterized in this study include *CaNFR1/5, CaENOD40, CaNARK, CaRIC1, CaRIC2, CaNIC1, CaRDN1, CaRDN2, CaRDN3, CamiRNA172* and *CaNNC1*. These genes were orthologous to *Medicago truncatula* nodulation genes and were significantly expressed by inoculating chickpea plants with *Mesorhizobium ciceri* analysed by qRT-PCR using RNA isolated from the root and leaf tissues of inoculated chickpea plants at 0, 2 and 6 days after inoculation. Non-inoculated plants served as control. *CaNARK* genes were expressed in roots and leaves while the rest of the nodulation genes were expressed only in the roots. The nodulation ability of chickpea may be controlled by an internal AON mechanism which involves several genes that are orthologues with other legumes.

Keywords: *chickpea, nodulation genes, Autoregulation of Nodulation (AON), gene expression*

Chickpea (*Cicer arietinum* L.) (2n=2x=16) is an annual legume of the family Fabaceae, sub-family Faboideae. It is a member of the Papilionoid subfamily of legumes, a clade that contains essentially all of the important legume crops. Within this subfamily, chickpea is most closely related to crops such as *Medicago sativa*, *Trifolium spp.*, *Pisum sativum*, *Lens culinaris*, *Lotus japonicus*, and *Medicago truncatula*, which has a very similar genome with chickpea [1]. Chickpea grows up to 20–50 cm high and has small, feathery leaves on either side of the stem. It is a type of pulse that bears feathery pinnately compound leaves. The small white or reddish flowers often have distinctive veins in blue, purple or pink and are usually self-pollinated. The yellow-brown or dark green beans are borne one or two to a pod. There are large and small-seeded varieties and are generally grown on residual soil moisture in the arid and semiarid regions of the world [2]. The cultivated chickpea, was one of the first grain legumes to be domesticated in the world and is the world's second most widely grown legume after soybean, accounting for a substantial proportion of human dietary nitrogen intake and playing a crucial role in food security in developing

countries [3]. Despite its potential for high production, the yield of chickpea can be low due to biotic (pest and pathogen infection) and abiotic stresses (drought, high and low temperature).

Chickpea, as a legume, can produce nodules on their roots which are the sites of symbiotic nitrogen fixation by rhizobia [4]. Atmospheric nitrogen gas is plentiful but is unavailable to most organisms. However, chickpea can overcome this limitation through symbiotic nitrogen fixation where atmospheric nitrogen (N2) is fixed by the nitrogenase enzyme complex of the endocytotic bacteria when they reside inside the nodules. The process of forming nodules containing symbiotic rhizobacteria is called nodulation. Through nodulation and the subsequent nitrogen fixation process, chickpea can fix its own nitrogen (140 kg N per hectare) from the atmosphere, which meets 80% of its nitrogen (N) requirement and partially benefit the following crops of the system by enriching soil through its substantial amount of residual N in the leaves and by adding plenty of organic matter [5, 6]. Nitrogen is an important nutrient requirement of plants which is utilized by plants for the synthesis of organic macromolecules and amino acids to form different proteins making its availability critical to sustained plant growth and reproduction [7].

Nodulation in legumes such as chickpea is initiated when (Nodulation) Nod factor (NF) signals secreted by the rhizobia are perceived by root hairs, which initiate curling of root hair followed by initiation of cell division and nodule primordium formation which finally develops into nodules [8].

Autoregulation of Nodulation (AON) starts with the production of a root-derived signal [9]. This is expressed in response to a transcription factor, which was reported to be involved in cortical cell division during early nodulation process. These kinds of signals were named CLAVATA/Embryo or CLE [10]. Meanwhile, GmRIC1 and GmRIC2 in soybeans are reported as root-derived signal peptides via the xylem to the shoot, where specialized LRR receptor kinase recognized them for proper functioning [11-15]. In soybean, these LRR receptor kinases are called *GmNARK* (Nodulation Autoregulation Receptor Kinase) gene [13]. Ferguson et al. revealed that *Phaseolus vulgaris* CLE peptides-PvRIC1, PvRIC2 and PvNIC1 were recognized by PvNARK which are similar to soybean CLE peptides-GmRIC1, GmRIC2 and GmNIC1 [16]. Soybean CLE peptides which were recognized by *GmNARK* were expressed in the roots. Unlike GmRIC1 and GmRIC2, GmNIC1 is not induced by rhizobia *NIC* gene but is specifically induced by available nitrate compounds. Its expression is not systemic and is only expressed in roots [17]. Over-expression of *GmNIC1* causes a significant reduction in nodulation [18].

Mechanism of nodulation suppression through CLE peptide signal exported from the root and transported via the xylem was previously unknown until Nontachaiyapoom et al. revealed that these signals move from root to leaf via phloem parenchyma cells where these are perceived by a leucine-rich repeat (LRR) serine-threonine receptor kinase [19]. All CLE peptides have resemblance with peptides CLV3 and ESR: small amino acid based proteins (less than 15 kD) with highly conserved C terminal CLE domain (12–14) amino acids and supposed secretion signal at N terminus [20]. The most studied model to understand the interaction between CLE peptides and LRR-RLK is CLAVATA functioning in shoot apical meristem (SAM), which involves CLE peptide of CLV3 and its receptors: functionally active LRR-RLK CLV1 (CLAVATA1) and kinase receptor CLV2 (CLAVATA2) with mechanized complex with receptor kinase CRN (CORYNE) for CLV3 binding [21].

In soybean, miR172c regulates nodule formation by repressing target gene, that is Nodule Number Control 1 (*NNC1*), which encodes for a specific protein that directly targets and triggers the promoter of the early nodulating gene (*ENOD40*). This stimulates the transcriptional levels of *miR172c* which were regulated by both *NFR1/5* mediated activation and by autoregulation of nodulation-mediated inhibition [22]. Bensmihen et al. studied the interaction between membrane-bound *NFR1* and *NFR5* detected as bimolecular fluorescence complementation which was observed only when a kinase-inactive *NFR1* was expressed together with *NFR5* [23].

This study aimed to identify and characterize the different genes that control the nodulation of chickpea using comparative genomic approach and analyzed the expression of these genes by inoculating chickpea plants with *Mesorhizobium ciceri* using qRT-PCR. Establishing the presence and expression of these nodulation genes will help us understand the mechanism of AON in chickpea which led us to propose a possible mechanism of AON specifically for chickpea which has never been done before.

Materials and Methods

Identification and Molecular Characterization of Nodulation Genes

Nodulation genes were identified and named as "Ca+insert gene name" where Ca stands for the scientific name of chickpea, *Cicer arietinum* while the name of genes was based from the known genes of other legumes. Several nodulation genes of chickpea (*CaNFR1, CaNFR5, CaENOD40, CaNARK, CaRIC1, CaRIC2* and *CaNIC1*), (*CaRDN1, CaRDN2, CaRDN3, Ca-miRNA172* and *CaNNC1*) were identified by aligning the chickpea gene sequences with the known sequences of other legumes: *Glycine max, Phaseolus vulgaris, Arabidopsis thaliana* and *Medicago truncatula* available at phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). The known sequences of similar genes of other legumes were compared with chickpea genes through BLAST search (http://www.nipgr.res.in/CGWR/home.php) by Varshney et al. (2013). Initially, the sequences of Phvul.005G096900 (*PvRIC1*), Phvul.011G135900 (*PvRIC2*), Phvul.005G097000 (*PvNIC1*) and Phvul.011G042000 (*PvNARK*) were compared with chickpea genes to identify *CaNARK* genes (*CaRIC1*, *CaRIC2*, *CaNIC1*) and were further aligned with CLE peptides of *Glycine max, Medicago truncatula,* and *Arabidopsis thaliana* using MEGA 6 software. The same program was used for making the dendrogram (Construct/Test-Neighbor,Joining -Tree). The Medtr5g089520 (*MtRDN1*), Medtr8g039290 (*MtRDN2*), and Medtr1g012920.1 (*MtRDN3*) were used to identify *CaRDN1*, *CaRDN2* and *CaRDN3*. The glyma12g07800, glyma02g43860, glyma11g06740, glyma01g03470 and MI0010727 were used to find the sequences of *CaNNC1*, *CaNFR1*, *CaNFR5*, *CaENOD40* and *Ca-miRNA172*, respectively. For the alignment of genes, *PsNOD3* was used. Amino acid sequences for all genes were read by SMART sequence analysis program (http://smart.embl-heidelberg.de/).

For the phylogenetic analysis of the nodulation genes, the genes were aligned with nodulation genes of *Glycine max, Medicago truncatula,* and *Arabidopsis thaliana* using MEGA 6 software. The dendrograms were constructed using the Construct/Test-Neighbor, Joining-Tree.

Accession Numbers

In this study, chickpea nodulation gene sequences were obtained from Chickpea Genomic Web Resource (http://www.nipgr.res.in/CGWR/home.php). The gene sequences of other legumes (*Glycine max, Phaseolus vulgaris Arabidopsis thaliana* and *Medicago truncatula*) were obtained from Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) (see Supplementary Table 1).

Growth of Chickpea Plants and Culture Conditions of *M. ciceri*

In all experiments conducted, wild type chickpea (*C. arietinum* L.) was used. Seeds provided by the Department of Primary Industries, New South Wales, Australia were surface-sterilized using 70% (v /v) ethanol for 10 seconds followed by rinsing five times with sterile water, then were sown in sterile vermiculite in 5.3 L pots. Plants were grown in controlled glasshouse conditions (28 ˚C and 24˚C, day and night, respectively, with a 16-h day length). They were watered daily and supplemented with a B & D nutrient solution (Broughton and Dilworth, 1971) twice per week [24]. The B & D nutrient solution was composed of the following elements: Ca (in the form of 1000 µM of CaCl2·2H20), P (500 µM of KH2P04), Fe (10 µM of Fe-Citrate), Mg (250 µM of MgSO4·7H20), K (1500 µM of K2SO4), S (500 µM), Mn (1 µM of MnSO4·H20), B (2 µM of H3BO4), Zn (0.5 µM of ZnSO4·7H20), Cu (0.2 µM of CuSO4·5H20), Co (0.1 µM of CoSO4·7H20) and Mo (0.1 µM of Na2MoO4·2H20). The volume of B&D nutrient solution per pot was 150 ml. Pots were filled with 4L of vermiculite. The weight of dry vermiculite was 0.075 kg per liter. The vermiculite was placed in autoclavable plastic bags and autoclaved at 121° C for 20 minutes under 15 psi of pressure. The vermiculite was purchased from a garden shop in Pakistan.

M. ciceri (isolated from the chickpea experimental area at the Department of Primary Industries, New South Wales, Australia) was grown for 48 h at 28 °C in Yeast Mannitol Broth (YMB) [25]. The isolated strain was confirmed by 16s rDNA sequence. Cultures were diluted with water to a final concentration of OD600 = 0.01 prior to inoculating plants. Approximately 150 mL of this final concentration was applied per pot.

Gene Expression of Nodulation Genes

To determine the gene expression, chickpea seeds were surface-sterilized and soaked overnight in sterilized water and sown after 24 hours in sterile vermiculite medium placed in

sterile pots and grown under a temperature-controlled glasshouse at The University of Queensland (Australia). The plants were watered on a daily basis with sterile distilled water. The plants were treated with 150 ml of *M. ciceri* suspension (final concentration OD600=0.01) while the control plants were not treated with the rhizobium. Kant et al mentioned that chickpea (*Cicer arietinum* L.), forms symbiosis with *M. ciceri* forming nodules which leads to its capability to convert atmospheric nitrogen (N2) into ammonia (NH3) through this symbiotic association [8]. An average of 98 nodules per plant ($p<0.05$) was produced by inoculated chickpea plants which are comparable to other nodulation studies [26, 27]. Artificial seed inoculation of chickpea with *M. ciceri* in soils lacking native effective rhizobia is a very useful practice for improving root nodulation and yield of chickpea [28, 29]. The plants were harvested at 0, 2 and 6 days of inoculation and the roots and shoots of harvested chickpea plants were snap-frozen in liquid nitrogen and stored at -80^{0}C for RNA extraction. RNA extraction from root and shoot samples was done by grinding the samples into powder using liquid nitrogen by pestle and mortar. The homogenous powder were transferred into small tubes and stored at -80^{0}C. Samples were later centrifuged. The 20 uL of each sample was transfered into another tube and mixed with the master mix following the protocol of Maxwell® 16 LEV simplyRNA Cells Kit and Maxwell® 16 LEV simplyRNA Tissue Kit from Promega.

The RNA concentration was determined by nanodrop and the RNA tubes were stored at -80^{0}C. A 1 ug of RNA was used to synthesized the cDNA after 24 hours. The sample concentration was adjusted using the Rnase-free water based from the nanodrop results. A 1 uL of oligo T and 10 mM dNTPs were mixed and heated at 65^{0}C for 5 minutes. Tubes were placed for 3 minutes in ice. The first standard buffer (4 µl 5x) was added with 1 uL 0.1M DTT Rnase and 1 µl Superscript III. The samples were incubated at 50^{0}C for 60 minutes and then heated for 15 minutes at 70^{0}C. The cDNA was stored at -20^{0}C.

To determine the successful synthesis of cDNA, the housekeeping gene *CaEF1* was used. A 1ul of undiluted cDNA was added to the mixture of 2 uL dNTP, 2 uL 10x buffer with Mgcl2, 1 uL forward primer of *CAEF1* and 1 uL reverse primer of *CaEF1* (see Table 2). The stock concentrations of the PCR reagents were Primers - 5 uM, dNTP - 100 uM, Primers: 0.2 - 1.0 uM, Dimethyl sulphoxide (DMSO) - 10% (v/v), and Taq polymerase - 1.0 Unit/50ul. A 0.25 uL of Taq polymerase and 12.72 uL of MQ H2O were added. Samples were centrifuged followed by PCR using the PCR profile: initial step at 95^{0}C for 5 minutes, denaturation at 94^{0}C for 30 seconds, annealing at 56^{0}C for 30 seconds, extension at 72^{0}C for 30 seconds and final step at 72^{0}C for 5 minutes for 35 cycles). PCR products were subjected to gel electrophoresis using a 2.5% gel and was run for 30 minutes at 150 volts. A 50 bp molecular ladder was added and the positioning of molecular bands further confirmed the successful synthesis of cDNA.

Quantitative real-time PCR (qRT-PCR) was performed according to Hayashi et al (2012) to evaluate the gene expressions [30]. PCR was carried out using SYBR® Green PCR Master Mix Applied Biosystems), and the 384-well plates for qRT-PCR analysis were set up using an Eppendorf® epMotionTM 5075 Robotics System. The reactions were run on an ABI Prism® 7900 Sequence Detection System (Applied Biosystems). Each plate contained no template (water) controls. Genomic DNA contamination of the cDNA samples was verified by including the reverse transcription negative (RT−) controls in the qRT-PCR. All qRT-PCR were carried out in duplicate (technical replicates) and run for 40 cycles using an annealing temperature of 60°C. The qRT-PCR was run using the following cycle conditions: 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. A dissociation stage of 95°C for 2 minutes was added at the end of the cycle in order to assess the specificity of the PCR. PCR efficiency for each sample was calculated using LinRegPCR 7.5 software, and the relative expression for gene of interest was measured relative to *CaEF1* gene.

Primers Used in the Study

Table 1 shows the list of all the primers used in the study.

Results and Discussion

Identification and Molecular Characterization of Nodulation Genes

The identity and molecular characteristics of the nodulation genes of chickpea were established.

(1) *CaNFR1* and *CaNFR5*. This study identified *NFR1* and *NFR5* in chickpea which resemble genes from other legumes. Nod Factor helps in the symbiosis with rhizobacteria to

Table 1. The sequences, temperature and GC percentage of the forward and reverse primers used in qRT-PCR.

Primer Name	Sequence	Annealing Temperature	% GC
CaRIC1	TCGTATGGCAAGATCGAGTA F	56.90	45.00
	TCTGGACCTCCTGGACTTAG R	56.88	55.00
CaRIC2	TCTTGACTTTGCAAGCTCGT F	58.05	45.00
	GGCCTCCTGGTGAGAGTCTA R	60.03	60.00
CaNIC1	ACAACCCTTGGCTTGGATCTTF	59.85	47.62
	AAAGGGACACGGGGAGTATCR	58.80	55.00
CaRDN1	GGCTTTTGTTCAGTGGCTGG F	59.97	55.00
	GGGAGTTGCCAATGGGATCA R	60.03	55.00
CaRDN2	CAACAAATGGCAGTGTCGCA F	59.97	50.00
	AACCACTGAACAAAGGCCCA R	60.0	50.00
CaNARK	CTTGTTGAATGGATGAGCAGAGT F	58.99	43.48
	GTGGGGAAGATTGGTGAGCA R	59.96	55.00
CaRDN3	TGGGGATCTGGAAAAGTAATCGT F	59.48	43.48
	CCCTATGACGACCAGTTCGG R	59.90	60.00
CamiRNA172	TGGATCATCATGGAgGTGAAF	59.85	45.00
	GCAGCAGCATGAGCTGTATC R	59.73	55.00
CaNNC	TTGCAGTGGCAGAATACGTCF	59.87	50.00
	CCCATCTTCCAGTCCTTCTGR	59.65	55.00
CaNFR1	AGAGGCAGTGGGATTGTGTTF	59.58	50.00
	CACTACCAGATGCGCTTTGAR	60.01	50.00
CaNFR5	GCTCTTTCGCCAATAACACC F	59.71	50.00
	GAGTCTTTGAAGGGCACCTG R	59.84	55.00
CaENOD40	GGACGTTTACCACTCCTTCTTC F	60.05	47.83
	GCCAAGATAGCATTTGTAGGAGA R	59.77	43.48
CaEF1	CTGTAACAAGATGGATGCCAC F	62.0	47.60
	CAGTCAAGGTTAGTGGACCT R	58.0	50.00

trigger nodulation using Lys motifs that act as receptor kinase, which induce nodulation. Glyma.02G270800 (*GmNFR1*), Medtr5g086090.1 (*MtNFR1*) were compared with chickpea genome to identify *CaNFR1* (Ca_19889.1) while *Glyma11g06740* and *Medtr5g019040.1* were used to identify the presence of *CaNFR5* (Ca_13116.1). *CaNFR1* was located in chromosome 2 and was orthologous to MtNRF1 which was located in chromosome 5. Meanwhile, *CaNFR5* was located in chromosome 8 which was orthologous to *MtNFR5* and also located in chromosome 5. In *CaNFR1* sequence, two Lys motifs were found: one at 51-98 bp and the other at 113-160 bp. However, in *CaNFR5*, only one motif was found which was at 98-149 bp. Lys motif and STKyc domain were also analyzed in *CaNFR1* and *CaNFR5* which were found to have the same domain and motifs reported in other legumes as shown in Figure 1.

(2) *CaENOD40.* *GmENOD40* was compared with chickpea genes to identify *CaENOD40* which was further confirmed by gene expression. *CaENOD40* lacks a long reading frame (ORF) and its activity was known to be RNA mediated. The transcripts share 2 regions of high sequence similarity where short ORFS reside. These two conserved boxes are present in all *ENOD40* genes of legumes. A 10-13 amino acid oligopeptide encoded by *CaENOD40* is conserved among plant species.

(3) *CaNARK, CaRIC1, CaRIC2* and *CaNIC1.* Peptide sequences of CaNARK showed several motifs and domain similar to those in *Medicago truncatula*, *Pisum sativum* and other legumes (Figure 2A). *CaNARK* was present in chickpea and located in chromosome 6 with a size of 3.726 kb. It has introns of 602 bp (Figure 2B). As shown in Figure 2C, *CaNARK* showed close resemblance with *M. truncatula, P. sativum,* and *L.*

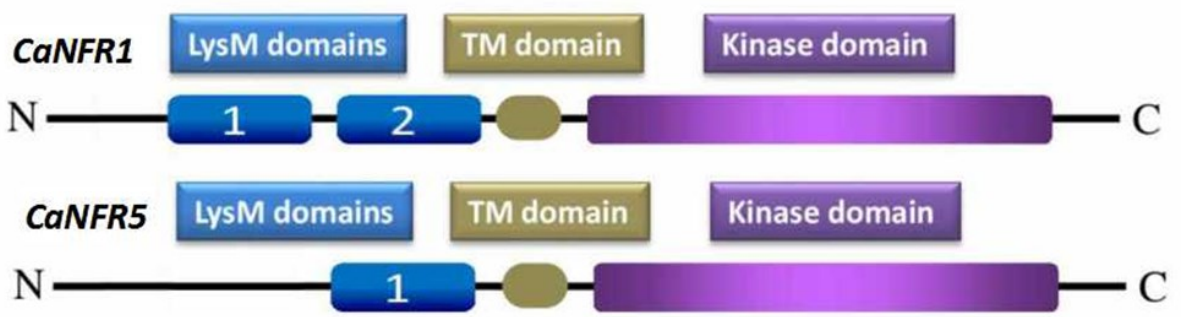

Figure 1. Schematic diagram of *CaNFR1* gene showing the two Lys motifs and *CaNFR5*, with only one Lys motif, and the STKyc domain.

japonicus.

CaRIC1 and *CaRIC2* which consist of 12 amino acid CLE peptides were responsible for the symbiosis of chickpea with *M. ciceri.* These genes were similar in function with PvRIC1 and PvRIC2 CLE peptides responsible for the symbiosis of *Rhizobium phaseoli* with *Phaseolus vulgaris* [16] and *LjCLE-RS1* and *LjCLE-RS2* genes which were responsible for *Mesorhizobium loti* mAFF303099 and *L. japonicus* symbiosis [31]. *GmRIC1*, *GmRIC2* and *GmNIC1* were compared with *CaRIC1*, *CaRIC2* and *CaNIC1* [17]. These chickpea CLE peptides were also aligned with CLE of *Phaseolus vulgaris*, *Lotus japonicus* and *Medicago truncatula*. *CaRIC1* and *CaRIC2* were found to be orthologous to *RIC* genes in *M. truncatula* (*MtCLE12*, *MtCLE13*). The 12 amino acid CLE peptides at N terminal showed highly

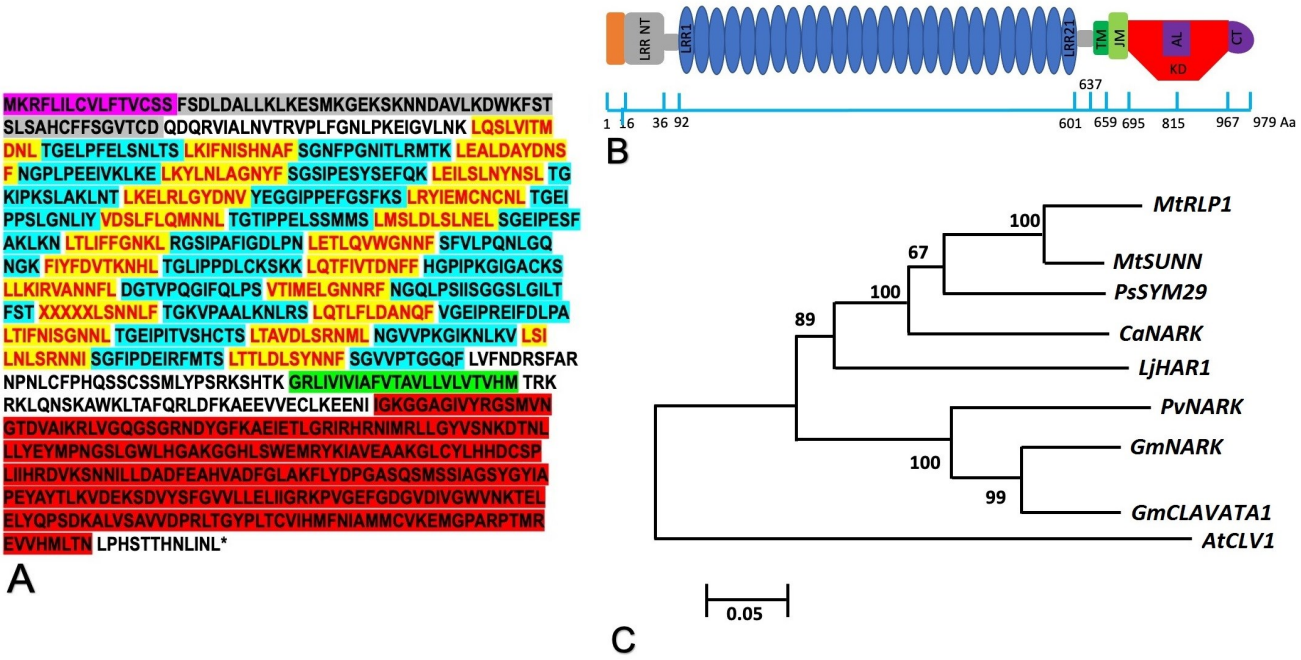

Figure 2. Nodulation Autoregulation Receptor Kinase of chickpea. (A) *CaNARK* is comprised of four domains: N-terminal signal peptide (pink), Leucine-rich repeat N-terminal domain (gray), Leucine-rich repeats domain (blue, LRR motif marked as yellow background with red letters), transmembrane region (green) and serine/threonine kinase (red). (B) The diagram of *CaNARK* domain structure showing the Leucine-rich repeats (LRR), single-pass transmembrane region (TM), Juxtamembrane region (JM), Kinase domain (KD), Kinase domain activation loop (AL) and C-terminal (CT) region. (C) Phylogenetic tree of *CaNARK* and its orthologs in soybean (*GmNARK*), *Lotus japonicus* (*LjHAR1*), *Medicago truncatula* (*MtSUNN*) and pea (*PsSYM29*), while *Arabidopsis thaliana* (*AtCLV1*) was the outgroup. Also included are the truncated copies of *MtSUNN* in *Medicago truncatula* (*MtRLP1*) and the homologous copy of *GmNARK* (formerly *GmCLV1B*) in soybean (*GmCLV1A*). Bootstrap confidence values were expressed as a percentage from 1000 bootstrap replications.

conserved region and enclosed in a red box as shown in Figure 3.

CaRIC1 which consists of 80 amino acids was found in chromosome 6 while its ortholog *MtCLE12* (Medtr4g079630) has 81 amino acids and found in chromosome 4. *CaRIC2* has 86 amino acids and found in chromosome 4 while its ortholog *MtCLE13* (Medtr4g079610) has 84 amino acids and was also found in chromosome 4. *CaNIC1* which has 82 amino acids showed close resemblance with *GmNIC1* (Glyma.12G208900) and with *PvNIC1* gene, which was located in chromosome 12 (Supplementray Figure 1, showing the position of *CaRIC1*, *CaRIC2,* and *CaNIC1*, and their orthologs in other legumes.

(4) *CaRDN1*, *CaRDN2* and *CaRDN3*. *CaRDN1*, *CaRDN2* and *CaRDN3* are present in chickpea which are commonly expressed in the roots. These genes were initially identified by genomic BLAST technique, using the gene sequences of Medtr5g089520 (*MtRDN1*), Medtr8g039290 (*MtRDN2*) and Medtr1g0129201 (*MtRDN3*) as reported by Kassaw (2012). The BLAST search and further alignment of predicted *CaRDN1* (Ca_16652.1), *CaRDN2* (Ca_08825.2), *CaRDN3* (Ca_03948.1) revealed that these genes were orthologous to *GmRDN* and *LjRDN* genes, which showed highly conserved domain. *CaRDN1*, which encodes 358 amino acids was located in chromosome 2 with multiple introns and was an ortholog of *MtRDN1* which also has the same number of amino acids (358) located in chromosome five (5). *LjRDN1*, *GmRDN1A* and *GmRDN1B* have 361, 358 and 364 amino acids, respectively (Supplementary Figure 2, dendogram for *CaRDN1*, *CaRDN2*, and *CaRDN3*, and their orthologs in other legumes.

CaRDN2 was comprised of 361 amino acids, and was found in chromosome 7, while its ortholog *MtRDN2* (Medtr8g039290; with 360 amino acids) was found in chromosome 8 in *M. truncatula. RDN2* of other legumes which were aligned with *CaRDN2* were *LjRDN2*, *GmRDN2A* and *GmRDN2B*, have 360, 359 and 363 amino acids, respectively. *CaRDN3* which consists of 363 amino acids was located in chromosome 4 while its ortholog was located in chromosome 1.

(5) *CamiRNA172*. Medtr2g093060.1 (*Mt_miRNA172*) sequence which was located in chromosome 2 was used to identify *Ca_miRNA172* in chickpea genome. The gene sequences were also compared with Glyma.12G073300 (*Glycine max*) and Phvul.005G138300 (*Phaseolus vulgaris*). The amino acid alignment of chickpea miRNA, located in chromosome 1, showed highly conserved region with miRNA amino acids of *Medicago* and other legumes. *Ca_miRNA172* was found to have ethylene response AP2 factor which is essential for nodulation. The region of AP2 is shown in Supplementary Figure 3.

(6) *CaNNC1*. For the identification of *CaNNC1*, Glyma12g07800 (*GmNNC1*), Med tr4g061200.2(*MtNNC1*), Phvul.011G071100.1(*Pv NNC1*) and At4g36920.1 (*AtmiRNA172*) were compared with chickpea genes. CaNNC1 was an ortholog of *Medicago* NNC1 (see Supplementary Figure 4). *CaNNC1* was located in chromosome 6 and its ortholog in *Medicago* was found in

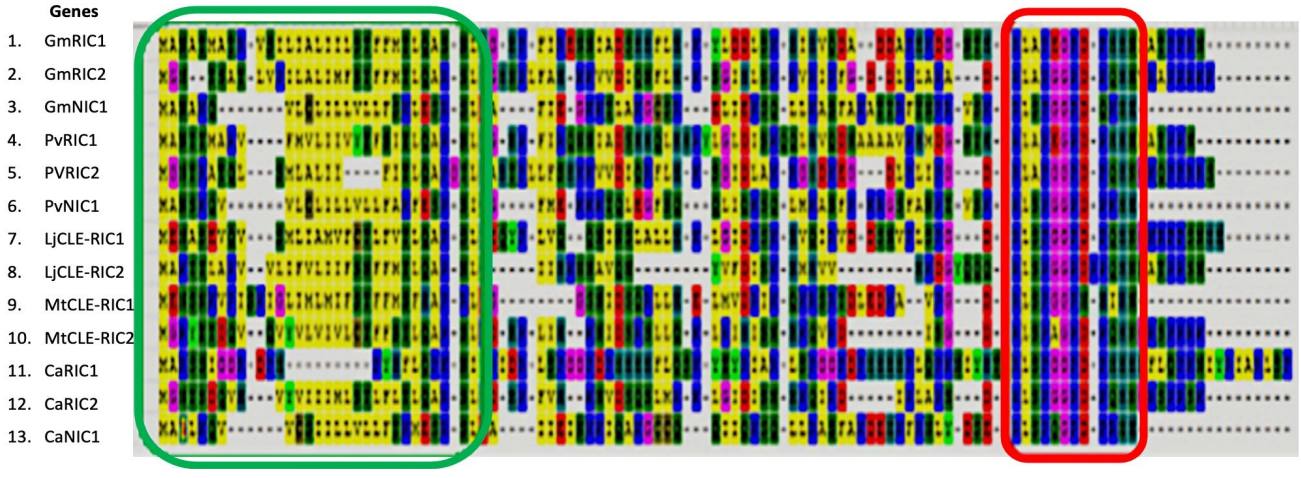

Figure 3. Multiple sequence alignment of inoculation-(RIC1 and 2) and nitrate-(NIC1) responsive CLE peptide-encoding genes of chickpea, soybean, bean, lotus and medicago. Predicted amino acid sequences showed a highly conserved signal peptide domains (enclosed in a green box) and the 12 amino acid CLE domains (enclosed in the red box).

chromosome 4.

Expression of Nodulation Genes in Chickpea

Gene expression of chickpea nodulated genes were determined by analyzing their transcript abundance from root tissues using qRT-PCR at 0, 2, and 6 days after inoculation. As shown in Figure 4, *CaNARK*, *CaRIC1*, *CaRIC2*, *CaNIC1*, *CaRDN1*, *CaRDN3*, *CaENOD40*, *CaNNC1*, *CaNFR1* and *CaNFR5* were expressed after 2 days of rhizobium inoculation. On the other hand, *CaRDN2* was expressed from the second day up to sixth day after inoculation while *CamiRNA172* was expressed after 6 days of inoculation. There were no gene expression observed in the non-nodulated chickpea plants.

Gene expression of chickpea nodulated genes were also determined by analyzing their transcript abundance from leaf tissues using qRT-PCR at 0, 2, and 6 days after inoculation. Only *CaNARK* gene was expressed in the leaves while *CaRIC1*, *CaRIC2*, *CaNIC1*, *CaRDN1*, *CaRDN2*, *CaRDN3*, *CaENOD40*, *CaNNC1*, *CaNFR1*, *CaNFR5* and *CamiRNA172* were not expressed in the leaves. The expression level of *CaNARK* in the leaves of inoculated chickpea was compared with non-inoculated chickpea (Figure 5). Statistical analysis revealed that there was a significant difference between the expression levels of nodulated ($M = 1.05 \times 10^{-3}$) versus non-nodulated chickpea plants (1.47×10^{-3}) with a *p*-value of $< .00001$.

The mechanism of nodule formation in legumes, including *C. arietinum*, is controlled by a long distance complex signaling system in which several genes are involved in the whole pathway. Root nodules are formed simultaneously with onset of root hair infection in a process that requires the perception of specialized lipochito-oligosaccharide signals Nod Factors (NF) by plant root cells [32]. NFs are recognized by LysM-type receptor kinases (NF receptors) CaNFR1/5. In this study, *CaNFR1* and *CaNFR5* in chickpea resembled genes from other legumes [22, 33, 34]. Lys motif and STKyc domain which were present in *CaNFR1* and *CaNFR5* have the same domain and motifs

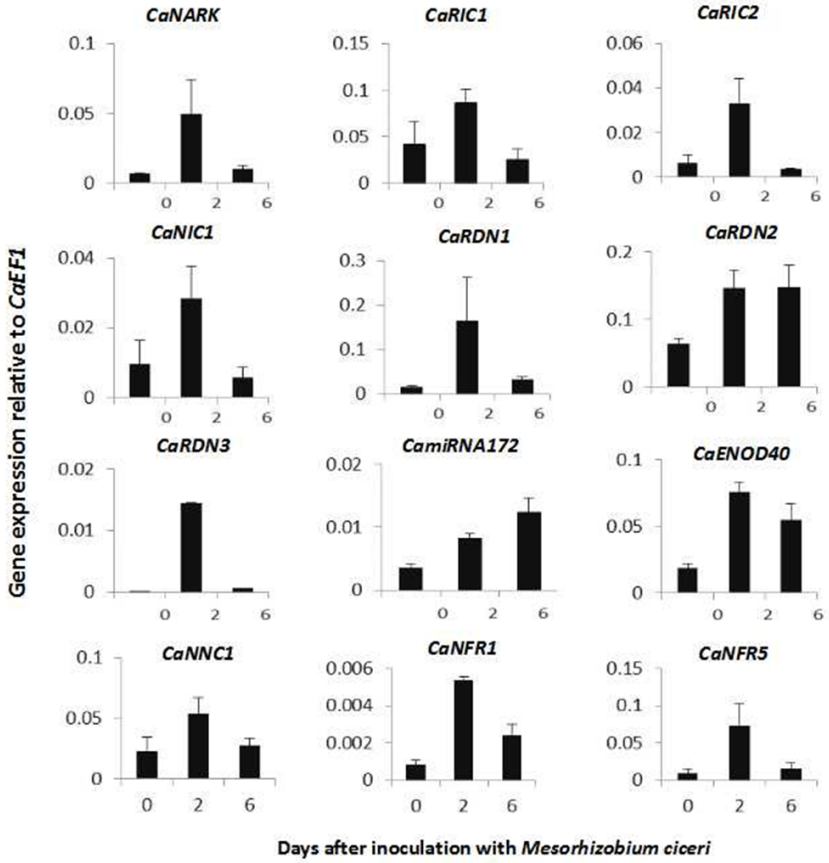

Figure 4. Transcript abundance of chickpea nodulation genes analyzed by qRT-PCR using RNA isolated from the root tissues of *Mesorhizobium ciceri*-inoculated chickpea at 0, 2 and 6 days after inoculation. Error bars indicate the standard error of the mean resulting from three independent experiments.

reported in other legumes. Upon recognition of NFs by *NFR1* and *NFR2*, legumes initiate a series of biochemical cascades, such as regular calcium oscillations in and around the nuclei of root epidermal cells to trigger the activation of downstream signaling components, such as calcium/calmodulin-dependent protein kinases and multiple transcription factors, which activate nodulation-related genes, *ENOD40*. Activation of *ENOD40* results in the initiation of root hair deformation, infection thread formation, cortical/ pericycle cell division, and nodule primordia formation, which collectively result in nodule formation and symbiotic nitrogen fixation [32]. *CaENOD40* was identified and characterized by this study which was similar with other legumes.

Nodulation genes which play important regulatory roles in root nodulation in chickpea were identified. The different functions of these genes can be deduced from the functions of their orthologous genes in other legumes. *CaNFR1* and *CaNFR5* (*C. arietinum* Nod Factor Receptor) may be involved in the perception of NOD factor signals secreted by rhizobia through the root hair to initiate nodulation just like in *Medicago* [35]. *CaENOD40* (*C. arietinum* Early Nodulation) may induce the division of root cortical cells, the nodule primordium and the pericycle of the root vascular bundle at early stages of nodule development just like in soybeans [36]. *CaNARK* (*C. arietinum* Nodule Autoregulation Receptor Kinase) which is a group of CLE (CLAVATA3/ Embryo) peptides found in plants is reported to be involved with cell signaling that includes *CaRIC1* (*C. arietinum* Rhizobia-Induced CLE) and *CaRIC2*. They are known to be involved in the AON signaling transduction that initiate expression solely in the roots with NOD factor-producing rhizobia and are considered as root-derived signals in AON pathway [37], and CaNIC1 (*C. arietinum* Nitrate-Induced CLE) that are reported to act locally in the roots to suppress nodulation by nitrate inhibition [38]. *CaRND1*, *CaRDN2* and *CaRDN3* (*C. arietinum* Root-Determined Nodulation) may serve as receptors for root-derived signals and trigger biosynthesis and release of shoot-derived inhibitor (SDI) just like in *Medicago* [39]. *CamiRNA172* (*C. arietinum* microRNA), which is known to be regulated by *CaNFR1* and *CaNFR5*, may activate *CaRIC1* and *CaRIC2* by removing the transcriptional repression of these genes by *CaNNC1* leading to the activation of AON pathway [22]. *CaNNC1* (*C. arietinum* Nodule Number Control 1) is reported to hamper the transcriptional activation of *CaRIC1* and *CaRIC2* and transcriptionally repress *CamiRNA172* expression adding a negative feedback loop in *CaNNC1* regulatory network [22].

These nodulation genes which are obviously involved in the mechanism of AON

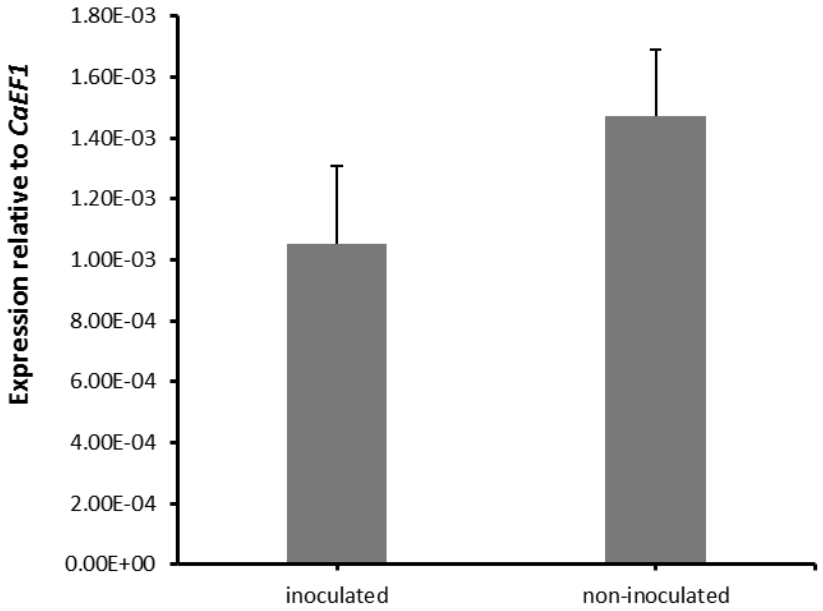

Figure 5. The expression of *CaNARK* gene in the leaves of chickpea. *CaNARK* expression level was analysed by qRT-PCR using RNA isolated from leaf tissues of *Mesorhizobium ciceri*-inoculated chickpea after 6 days of inoculation. Non-inoculated plants served as control. Error bars indicate the standard error of the mean. The *t*-value is 51.43929. The *p*-value is < .00001.

pathway in chickpea were orthologous to *M. truncatula* nodulation genes. These genes were significantly expressed in the root or leaf tissues. *CaNARK*, *CaRIC1*, *CaRIC2*, *CaNIC1*, *CaRDN1*, *CaRDN3*, *CaENOD40*, *CaNNC1*, *CaNFR1* and *CaNFR5* were expressed in the roots after 2 days of rhizobium inoculation. On the other hand, *CaRDN2* was expressed in the roots from the second day up to sixth day after inoculation while *CamiRNA172* was expressed in the roots after 6 days of inoculation. *CaNARK* genes were expressed in roots and leaves while the rest of the nodulation genes were expressed only in the roots. This means that *CaNARK* may have functions in the shoot in the AON pathway to regulate nodulation in the root. It may act in the leaf vascular tissue to perceive root-derived peptides produced by *CaRIC1* and *CaRIC2* that originate during nodule primordium formation. This perception may result in the production of a shoot-derived inhibitor (SDI), which travels to the roots to inhibit further nodule development. Excitingly, recent evidence has revealed that cytokinins in the shoot may function as an SDI to systemically suppress nodulation in chickpea [40].

Establishment of an excess number of nodules (supernodulation) markedly affects plant growth and development due to overconsumption of photosynthates for nitrogen fixation [41]. Therefore, the total root nodule number is tightly controlled via a negative feedback mechanism (AON) to maintain an optimal nitrogen and carbon balance in the host [32]. AON mechanism in chickpea involves several genes that are orthologues with other legumes including *G. max*, *P. vulgaris*, and *M. truncatula.* It has been reported that AON is activated in root cortical cells during rhizobial infection and remains active during nodule primordium formation and nodule maturation. It starts with the production of root-derived RIC1 and RIC2 in following the first induced cortical cell divisions during rhizobial infection, nodule development, and the onset of nodule functionality. In this study, *CaRIC1* and *CaRIC2* were found to be orthologous with RIC genes in *M. truncatula* (*MtCLE12*, *MtCLE13*) and may have similar functions with *PvRIC1* and *PvRIC2* [16].

This study strongly suggests that AON in chickpea may be facilitated by a feedback loop involving root to shoot and shoot to root signaling. Short peptides are synthesised by *CaRIC1* and *CaRIC2* of the root and travel to the shoot where they interact with *CaNARK*. Homologous genes of *CaNARK* were previously reported in different legumes: *LjHAR1* [42], *PsSYM29* [11], *GmNARK* [13], *MtSUNN* [15] and *PvNARK* [16]. In this study, *CaNARK* showed close resemblance to *M. truncatula, P. sativum,* and *L. japonicus* [3]. *CaNARK* comprised of four domains: N-terminal signal peptide, Leucine-rich repeat N-terminal domain, leucine-rich repeats domain, LRR motif, transmembrane region and serine/threonine kinase. The structure of *CaNARK* domain consists of Leucine-rich repeats, single-pass transmembrane region, Juxtamembrane region, kinase domain, kinase domain activation loop and C-terminal region. *CaNARK* has orthologues in soybean, lotus, *Medicago* and pea. *CaNARK* has a central role in the AON of chickpea and mutations in these genes can result in hyper/supernodulation.

In the root, events associated with nodulation generate specific rhizobia-induced CLE peptides which in some cases appear to be arabinosylated via action of the enzyme produced by *RDN1* genes [43]. *CaRDN1* may encode a protein of unknown function that appears to be expressed at low levels in vasculature such as in the case of M. truncatula [44, 45]. Although *CaRDN1* is involved in the legume AON signaling pathway, it was reported to have high level of conservation of *RDN* family genes throughout the green plant lineage [46]. Transgenic *RDN1/NOD3* gene in *M. truncatula* and *P. sativum* produced mutant plants with higher number of nodules compared with the wild type variety [46]. *MtRDN1*, *MtRDN2* and *MtRDN3* were reported in *Medicago* by Kassaw et al. [47] and Kassaw and Frugoli [48]. *CaRDN1*, *CaRDN2* and *CaRDN3* are present in chickpea and are expressed in the roots.

Meanwhile in the leaves, activation of *CaNARK* leads to the suppression of a specific microRNA species, which is phloem-translocated to the root, causing a suppression of further nodulation induction. Chickpea genome has been found to have *CamiRNA172* with an ethylene response AP2 factor (Apetala 2, a member of a large family of transcription factors) which is helpful for nodulation. It is significantly expressed in the roots as a response to rhizobia application. *CamiR172c* may activate *CaRIC1* and *CaRIC2* by removing the transcriptional repression of these genes by *CaNNC1* that may directly trigger *CaENOD40* [30,49,50]. *CaNNC1* has a role in the direct activation of *CaENOD40*: it encodes a protein that directly targets the promoter of the early nodulin gene, *ENOD40* [22]. The stimulated transcriptional level of *CamicroRNA172* is

regulated by *CaNFR1* and *CaNFR5*. *CamiR172c-NNC1* network may be a master switch that coordinately regulates and optimizes NF and AON signaling, supporting the balance between nodulation and AON in chickpea. *CaNIC1* acts locally in the roots to suppress nodulation by nitrate inhibition [38]. *CaNIC1* is likely a local inducer of NARK-dependent nodule regulation independent of the systemic AON mechanism [17]. The reduced nodule inhibition response of *CaNIC1* may be due to this local role in nodule inhibition and a reflection of successful transformation events or to differences in the signal peptide or CLE domain that may reduce the effectiveness of processing or activity of *CaNIC1* [17].

MicroRNAs (miRNAs) are small non-coding RNAs (22 nt in length) that regulate eukaryotic development by repressing the expression of target genes, particularly at the post-transcriptional level. It is known as a master regulator in the nodulation process of legumes. MicroRNA plays an important role in plant growth, differentiation, transduction and development [51-54]. It functions by targeting many types of transcription factors such as auxin response transcriptional factor family (ARF) [55-57], phosphorus starvation response transcription factors [58], and root apical meristem transcription factors [59]. Currently, a very limited number of miRNAs have been validated for their roles in the nodulation of legumes and none in chickpea. The reduction in activity and over-expression of *CamiR172c* gene may cause dramatic changes in nodulation of chickpea [60,61].

Based from our findings, a hypothesis on a possible mechanism of AON specifically for chickpea was formulated, as shown in Figure 6.

Conclusion and Recommendations

Nodulation genes involved in the AON mechanism in chickpea were identified by aligning chickpea genes with known genes of legumes. Chickpea nodulation genes characterized in this

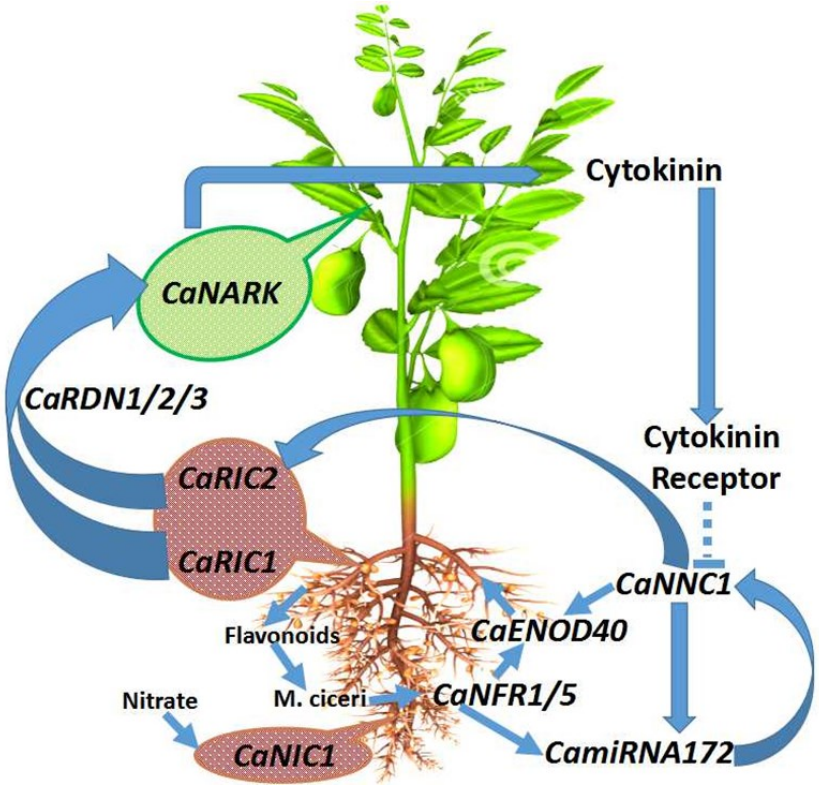

Figure 6. Proposed mechanism of Autoregulation of Nodulation (AON) in chickpea, a hypothesis formulated based on our findings. Flavonoids are secreted by the roots which are perceived by *M. ciceri* that releases Nod Factors. Nod Factors are perceived by *CaNFR1/5* which triggers CaENOD40 to form nodules. Meanwhile, *CaNFR1/5* regulates *CamiRNA172* which activates *CaRIC1* and *CaRIC2* by *CaNNC1*. *CaRIC1* and *CaRIC2* are then perceived by *CaRDN1/2/3* and triggers *CaNARK* to initiate Shoot Derived Inhibitors, presumably cytokinin. Cytokinin perceived by CK Receptor triggers *CaNNC1*. *CaNNC1* transcriptionally represesses *CamiRNA172* and directly triggers *CaENOD40*. Meanwhile, *CaNIC1* is activated by nitrate present in soil which suppresses the nodule formation by nitrate inhibition.

study include *CaNFR1/5*, *CaNARK*, *CaRIC1*, *CaRIC2*, *CaNIC1*, *CaRND1*, *CaRDN2*, *CaRDN3*, *CamiRNA172*, *CaNNC1* and *CaENOD40*. These genes were orthologous to *M. truncatula* nodulation genes. *CaNARK* is expressed in chickpea leaves and roots while the rest of the genes are only expressed in the roots. Chickpea nodulation is tightly controlled via a negative feedback mechanism (AON) to maintain an optimal nitrogen and carbon balance in chickpea plants.

AON mechanism in chickpea involves several genes. *CaNFR1/5* regulates *CamiRNA172* which activates *CaRIC1* and *CaRIC2* by *CaNNC1*. *CaRIC1* and *CaRIC2* are then perceived by *CaRDN1/2/3* which triggers *CaNARK* to initiate Shoot Derived Inhibitors, presumably cytokinin. Cytokinin perceived by CK Receptor triggers *CaNNC1*. *CaNNC1* transcriptionally represesses *CamiRNA172* and directly triggers *CaENOD40*. Meanwhile, *CaNIC1* is activated by nitrate present in soil which suppresses the nodule formation by nitrate inhibition.

For future research, it is recommended that studies that will validate the role of these genes in AON be conducted.

Acknowledgment

The authors would like to thank Dr. Jonathan M. Plett, Assistant Professor at Western Sydney University for providing *M. ciceri* pure cultures and Higher Education Commission (HEC) of the Government of Pakistan for the research grant.

References

[1] Nayak, S. N., Zhu, H., Varghese, N., Datta, S., Choi, H. K., Horres, R.,& Hoisington, D. A. (2010). Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with *Medicago truncatula* genome. *Theoretical and Applied Genetics, 120 (7),* 1415-1441.

[2] Sivasakthi, K., Zaman-Allah, M., Tharanya, M., Kholová, J., Thirunalasundari, T., & Vadez, V. (2017). Chickpea. *In Water-Conservation Traits to Increase Crop Yields in Water-deficit Environments* (pp. 35-45). Springer, Cham.

[3] Varshney, R. K., Song, C., Saxena, R. K., Azam, S., Yu, S., Sharpe, A. G., Cannon, S., Baek, J., Rosen, B. D., & Tar'An, B. (2013). Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nature biotechnology, 31,* 240-246.

[4] Valentine, A. J., Benedito, V. A., & Kang, Y. (2018). Legume nitrogen fixation and soil abiotic stress: from physiology to genomics and beyond. *Annual Plant Reviews online,* 207-248.

[5] Gogoi, N., Baruah, K. K., & Meena, R. S. (2018). Grain legumes: impact on soil health and agroecosystem. *In Legumes for Soil Health and Sustainable Management* (pp. 511-539). Springer, Singapore.

[6] Panchal, P., Patel, P. H., Patel, A. G., & Desai, A. (2017). Effect of Panchagavya on growth yield and economics of chickpea (*Cicer arietinum)*. *International Journal of Chemical Studies, 5(2),* 265-267.

[7] Calica, PN. (2017). Nodulation and Nitrogen Fixation of Pongamia pinnata. *Journal of Tropical Crop Science.* ISSN 2356-0177. Vol. 4 No. 1.pp. 1-12.

[8] Kant, C., Pradhan, S., & Bhatia, S. (2016). Dissecting the root nodule transcriptome of chickpea (*Cicer arietinum* L.). *PLoS One, 11(6)*.

[9] Ferguson, B. J., Mens, C., Hastwell, A. H., Zhang, M., Su, H., Jones, C. H., ... & Gresshoff, P. M. (2019). Legume nodulation: The host controls the party. *Plant, cell & environment, 42(1),* 41-51.

[10] Gresshoff, P. M., & Delves, A. C. (1986). Plant genetic approaches to symbiotic nodulation and nitrogen fixation in legumes. *In A Genetic Approach to Plant Biochemistry* (pp. 159-206): Springer.

[11] Krusell, L., Madsen, L. H., Sato, S., Aubert, G., Genua, A., Szczyglowski, K., Duc, G., Kaneko, T., Tabata, S., & De Bruijn, F. (2002). Shoot control of root development and nodulation is mediated by a receptor-like kinase. *Nature, 420*, 422-426.

[12] Nishimura, R., Hayashi, M., Wu, G.-J., Kouchi, H., Imaizumi-Anraku, H., Murakami, Y., Kawasaki, S., Akao, S., Ohmori, M., & Nagasawa, M. (2002). HAR1 mediates systemic regulation of

symbiotic organ development. *Nature, 420*, 426-429.

[13] Searle, I. R., Men, A. E., Laniya, T. S., Buzas, D. M., Iturbe-Ormaetxe, I., Carroll, B. J., & Gresshoff, P. M. (2003). Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. *Science, 299*, 109-112.

[14] Elise, S., Etienne-Pascal, J., De Fernanda, C.-N., Gérard, D., & Julia, F. (2005). The *Medicago truncatula* SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. *Plant Molecular Biology, 58*, 809-822.

[15] Schnabel, E., Etienne-Pascal, J., De Fernanda, C.-N., Gérard, D., & Julia, F. (2005). The Medicago truncatula SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. *Plant Molecular Biology, 58,* 809-822.

[16] Ferguson, B. J., Li, D., Hastwell, A. H., Reid, D. E., Li, Y., Jackson, S. A., & Gresshoff, P. M. (2014). The soybean (Glycine max) nodulation-suppressive CLE peptide, GmRIC1, functions interspecifically in common white bean (*Phaseolus vulgaris*), but not in a supernodulating line mutated in the receptor PvNARK. *Plant Biotechnology Journal, 12,* 1085-1097.

[17] Reid, D. E., Ferguson, B. J., & Gresshoff, P. M. (2011). Inoculation-and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation. *Molecular Plant — Microbe Interactions, 24*, 606-618.

[18] Reid, D. E., Ferguson, B. J., Hayashi, S., Lin, Y.-H., & Gresshoff, P. M. (2011). Molecular mechanisms controlling legume autoregulation of nodulation. *Annals of Botany, 108,* 789-795.

[19] Nontachaiyapoom, S., Scott, P. T., Men, A. E., Kinkema, M., Schenk, P. M., & Gresshoff, P. M. (2007). Promoters of orthologous Glycine max and Lotus japonicus nodulation autoregulation genes interchangeably drive phloem-specific expression in transgenic plants. *Molecular Plant-Microbe Interactions, 20,* 769-780.

[20] Goad, D. M., Zhu, C., & Kellogg, E. A. (2017). Comprehensive identification and clustering of CLV3/ESR-related (CLE) genes in plants finds groups with potentially shared function. *New Phytologist, 216(2),* 605-616.

[21] Müller, R., Bleckmann, A., & Simon, R. (2008). The receptor kinase CORYNE of Arabidopsis transmits the stem cell–limiting signal CLAVATA3 independently of CLAVATA1. *The Plant Cell, 20,* 934-946.

[22] Wang, Y., Wang, L., Zou, Y., Chen, L., Cai, Z., Zhang, S., Zhao, F., Tian, Y., Jiang, Q., & Ferguson, B. J. (2014). Soybean miR172c targets the repressive AP2 transcription factor NNC1 to activate ENOD40 expression and regulate nodule initiation. *The Plant Cell, 26,* 4782-4801.

[23] Bensmihen, S., De Billy, F., & Gough, C. (2011). Contribution of NFP LysM domains to the recognition of Nod factors during the *Medicago truncatula/Sinorhizobium meliloti* symbiosis. *PloS one, 6,* e26114.

[24] Broughton, W. J., & Dilworth, M. J. (1971). Control of leghaemoglobin synthesis in snake beans. *Biochemical journal, 125(4),* 1075-1080.

[25] Somerville, J. E., & Kahn, M. L. (1983). Cloning of the glutamine synthetase I gene from Rhizobium meliloti. *Journal of bacteriology, 156 (1)*, 168-176.

[26] Gul, R., Khan, H., Khan, N. U., & Khan, F. U. (2014). Characterization of chickpea germplasm for nodulation and effect of rhizobium inoculation on nodules number and seed yield. *Journal of Animal and Plant Sciences, 24,* 1421-1429.

[27] Pandey, R. P., Srivastava, A. K., Gupta, V. K., O'Donovan, A., & Ramteke, P. W. (2018). Enhanced yield of diverse varieties of chickpea (*Cicer arietinum L.)* by different isolates of *Mesorhizobium ciceri*. *Environmental Sustainability, 1(4),* 425-435.

[28] Khattak, S. G., Khan, D. F., Shah, S. H., Madani, M. S., & Khan, T. (2006). Role of rhizobial inoculation in the production of chickpea crop. *Soil & Environ, 25(2),* 143-145.

[29] Ahmad, M. J., Muhammad, K., & Amer, J. (2011). screening of Mesorhizobium ciceri isolates from rainfed area for improvement in growth and nodulation of chickpea seedlings under controlled

conditions. *Soil and Environment, 30(2),* 160-165.

[30] Hayashi, S., Reid, D. E., Lorenc, M. T., Stiller, J., Edwards, D., Gresshoff, P. M., & Ferguson, B. J. (2012). Transient Nod factor-dependent gene expression in the nodulation competent zone of soybean (Glycine max [L.] Merr.) roots. *Plant Biotechnology Journal, 10,* 995 -1010.

[31] Suzaki, T., & Kawaguchi, M. (2014). Root nodulation: a developmental program involving cell fate conversion triggered by symbiotic bacterial infection. *Current opinion in plant biology, 21,* 16-22.

[32] Wang L., Sun Z., Su C., Wang Y., Yan Q., Chen J., Ott T., And Li X. (2019). A GmNINa-miR172c-NNC1 Regulatory Network Coordinates the Nodulation and Autoregulation of Nodulation Pathways in Soybean. *Mol. Plant. 12,* 1211– 1226.

[33] Govers, F., Moerman, M., Downie, J. A., Hooykaas, P., Franssen, H. J., Louwerse, J., Van Kammen, A., & Bisseling, T. (1986). Rhizobium nod genes are involved in inducing an early nodulin gene. *Nature, 323,* 564-566.

[34] Gage, D. J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiology and Molecular Biology Reviews, 68,* 280-300.

[35] Geurts, R., Fedorova, E., & Bisseling, T. (2005). Nod factor signaling genes and their function in the early stages of Rhizobium infection. *Current opinion in plant biology, 8(4),* 346-352.

[36] Minami, E., Kouchi, H., Cohn, J. R., Ogawa, T., & Stacey, G. (1996). Expression of the early nodulin, ENOD40, in soybean roots in response to various lipo-chitin signal molecules. *The Plant Journal, 10(1),* 23-32.

[37] Hastwell, A. (2018). Functional characterisation of novel peptide hormones in legume nodulation and plant development.

[38] Lim, C. W., Lee, Y. W., Lee, S. C., & Hwang, C. H. (2014). Nitrate inhibits soybean nodulation by regulating expression of CLE genes. *Plant Science, 229,* 1-9.

[39] Suzaki, T., Yoro, E., & Kawaguchi, M. (2015). Leguminous plants: inventors of root nodules to accommodate symbiotic bacteria. *In International review of cell and molecular biology* (Vol. 316, pp. 111-158). Academic Press.

[40] Sasaki, T., Suzaki, T., Soyano, T., Kojima, M., Sakakibara, H., & Kawaguchi, M. (2014). Shoot-derived cytokinins systemically regulate root nodulation. *Nature communications, 5(1),* 1-9.

[41] Matsunami, T., Kaihatsu, A., Maekawa, T., Takahashi, M., & Kokubun, M. (2004). Characterization of vegetative growth of a supernodulating soybean genotype, Sakukei 4. *Plant production science, 7(2),* 165-171.

[42] Buzas, D. M., & Gresshoff, P. M. (2007). Short-and long-distance control of root development by LjHAR1 during the juvenile stage of Lotus japonicus. *Journal of Plant Physiology, 164,* 452-459.

[43] Wang, C., Reid, J. B., & Foo, E. (2018). The Art of Self-Control–Autoregulation of Plant–Microbe Symbioses. *Frontiers in plant science, 9,* 988.

[44] Saur, I. M., Oakes, M., Djordjevic, M. A., & Imin, N. (2011). Crosstalk between the nodulation signaling pathway and the autoregulation of nodulation in *Medicago truncatula*. *New Phytologist, 190,* 865-874.

[45] Schnabel, E., Mukherjee, A., Smith, L., Kassaw, T., Long, S., & Frugoli, J. (2010). The lss supernodulation mutant of *Medicago truncatula* reduces expression of the SUNN gene. Plant physiology, 154, 1390-1402.

[46] Schnabel, E. L., Kassaw, T. K., Smith, L. S., Marsh, J. F., Oldroyd, G. E., Long, S. R., & Frugoli, J. A. (2011). The Root Determined Nodulation1 gene regulates nodule number in roots of *Medicago truncatula* and defines a highly conserved, uncharacterized plant gene family. *Plant Physiology, 157,* 328-340.

[47] Kassaw, T., & Frugoli, J. (2015). Multiple autoregulation of nodulation (AON) signals identified through split root analysis of *Medicago truncatula* sunn and rdn1 mutants. *Plants, 4(2),* 209 -224.

[48] Kassaw, T., Nowak, S., Schnabel, E., & Frugoli, J. (2017). ROOT DETERMINED NODULATION1 is required for M. truncatula CLE12, but not CLE13, peptide signaling through the SUNN receptor kinase. *Plant Physiology, 174 (4),* 2445-2456.

[49] Mathesius, U., Charon, C., Rolfe, B. G., Kondorosi, A., & Crespi, M. (2000). Temporal and spatial order of events during the induction of cortical cell divisions in white clover by Rhizobium leguminosarum bv. trifolii inoculation or localized cytokinin addition. *Molecular Plant-Microbe Interactions, 13,* 617-628.

[50] Compaan, B., Yang, W.-C., Bisseling, T., & Franssen, H. (2001). ENOD40 expression in the pericycle precedes cortical cell division in Rhizobium-legume interaction and the highly conserved internal region of the gene does not encode a peptide. *Plant and Soil, 230,* 1-8.

[51] Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell, 116*, 281-297.

[52] Zhou, X., Wang, G., Sutoh, K., Zhu, J.-K., & Zhang, W. (2008). Identification of cold-inducible microRNAs in plants by transcriptome analysis. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1779,* 780-788

[53] Kapoor M., Arora R., Lama T., Nijhawan A., Khurana J.P., Tyagi A.K., & S., K. (2008). Genome-wide identification, organization and phylogenetic analy-sis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics, 9,* 1-17.

[54] Wahid, F., Shehzad, A., Khan, T., & Kim, Y. Y. (2010). MicroRNAs: synthesis, mechanism, function, and recent clinical trials. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1803,* 1231-1243.

[55] Mallory, A. C., Bartel, D. P., & Bartel, B. (2005). MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. *The Plant Cell, 17,* 1360-1375.

[56] Gomes, M. S., Cabral, F. J., Jannotti-Passos, L. K., Carvalho, O., Rodrigues, V., Baba, E. H., & Sá, R. G. (2009). Preliminary analysis of miRNA pathway in *Schistosoma mansoni. Parasitology International, 58,* 61-68.

[57] Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annual Review of Plant Biology, 64,* 137-159.

[58] Nilsson, L., Müller, R., & Nielsen, T. H. (2007). Increased expression of the MYB-related transcription factor, PHR1, leads to enhanced phosphate uptake in Arabidopsis thaliana. *Plant, Cell & Environment, 30,* 1499-1512.

[59] Ochando, I., Jover-Gil, S., Ripoll, J. J., Candela, H., Vera, A., Ponce, M. R., Martínez-Laborda, A., & Micol, J. L. (2006). Mutations in the microRNA complementarity site of the INCURVATA4 gene perturb meristem function and adaxialize lateral organs in Arabidopsis. *Plant physiology, 141,* 607-619.

[60] Srivastava, P. K., Moturu, T. R., Pandey, P., Baldwin, I. T., & Pandey, S. P. (2014). A comparison of performance of plant miRNA target prediction tools and the characterization of features for genome-wide target prediction. *BMC Genomics, 15,* 348.

[61] Zhang, B. H., Pan, X. P., Wang, Q. L., George, P. C., & Anderson, T. A. (2005). Identification and characterization of new plant microRNAs using EST analysis. *Cell Research, 15,* 336-360.