
Chapter 4 : Molecular effects of *C. cajan* & *Z. mays* root exudates on NGR234 and its interactions with monocropped and intercropped *Z. mays* plants

4.1 Introduction

Rhizobia are among the best-studied root-associated microbes, with a drastic lifestyle change from saprophytic to endosymbiotic bacteroids, by adopting an oligotrophic lifestyle inside the host plant (Poole et al., 2018). Rhizobia-legume symbiosis is also one of the best-studied for the signal exchange process mediated through the flavonoids (secondary metabolites) released by the legume root (Masson-Boivin and Sachs, 2018). Rhizobia recognize root exudates signals using a positively acting transcription factor, encoded by *nodD*. The activated NodD protein binds to highly conserved bacterial promoters (nod boxes) and induce the expression of several genes involved in the formation of the differentiated structure, the nodule, in which the bacteria reside as nitrogen fixing endosymbionts (Downie, 2010).

The specificity in their symbiotic associations varies greatly amongst the different symbionts for instance *Sinorhizobium meliloti* can initiate nodule formation on a few host plants (*Medicago*, *Melilotus*, and *Trigonella*), whereas broad host range *Ensifer fredii* sp. NGR234 nodulates more than 112 genera of legumes, as well as the non-legume *Parasponia andersonii*, and fixes nitrogen in association with plants that form either determinate or indeterminate types of nodules (Pueppke and Broughton, 1999; Broughton et al., 2003). The genome of NGR234 is subdivided into three replicons: a chromosome, a mega- plasmid of ~ 2.2 Mb (pNGR234 b) that encodes several symbiotic loci such as the *exo-* genes that are involved in the biosynthesis of exopolysaccharides (EPSs), and; the symbiotic plasmid (pNGR234 a) that carries most of the genes required for symbiosis. The response of NodD1 to the flavonoids induces nod factor biosynthesis (Fig.4.1). Also, the NodD1 responds to from different plant extracts of cereals such as maize, wheat, and rice. Thus, NodD1 of NGR234 shows a versatile response and is less specific compared to other *nodD* alleles of other strains (Le strange et al.1990).

Rhizobia are also known to colonize the rhizosphere and promote the growth of non-legume plants such as maize and rice (Rosenblueth, 2004; Wu et al., 2018; Cavalcanti et al., 2020). Besides, the carbohydrate-rich mucilage (high molecular weight root exudates) of the aerial roots of maize aids nitrogen fixation by the diazotrophic bacteria in Sierra Mixe, Mexico (Van Deynze et al., 2018). However, the signaling mechanism that facilitates the interaction of rhizobia with low molecular weight root exudates from maize and/or rice, is not yet known.

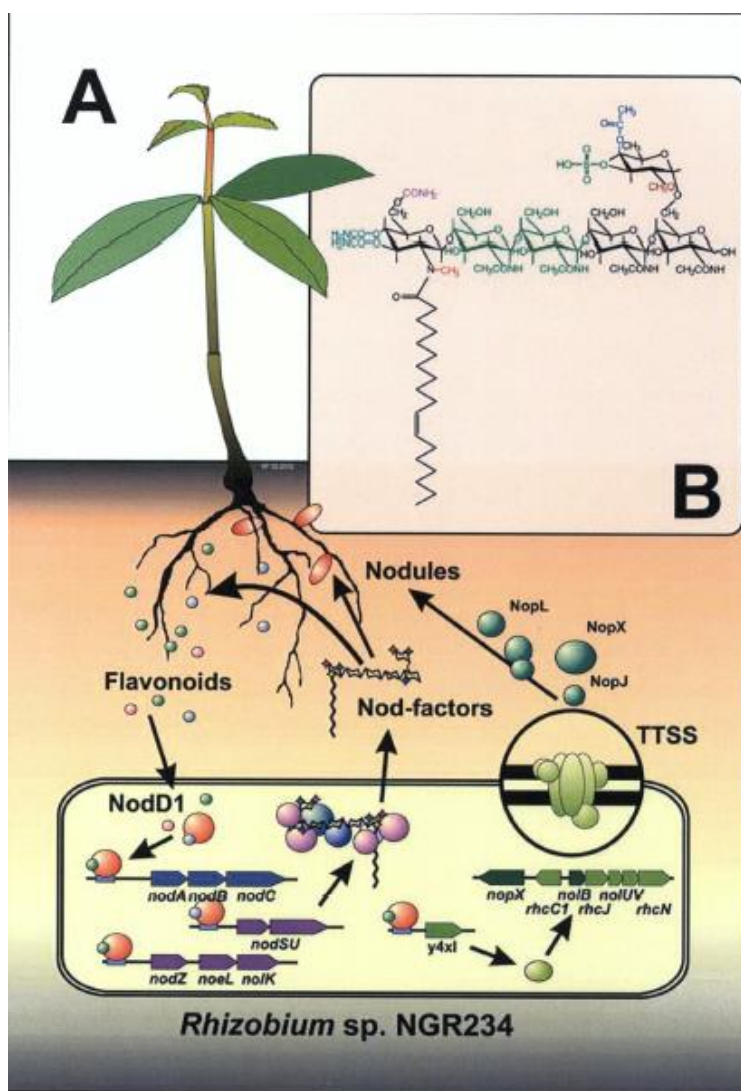


Fig. 4-1 Flavonoid-inducible determinants of nodulation in *Ensifer fredii* NGR234 (A) Expression of the rhizobial nodulation genes (*nod*, *not* and *noe*) induced by the release of flavonoids, (B) Structure of Nod factor (Broughton et al., 2003).

About up to 20% of plant photosynthate can be released from roots in the form of mucilage and metabolites (Walker et al., 2003). Among metabolites, the primary metabolites in the root exudates serve as carbon, nitrogen, or sulfur sources for free-living and/or symbiotically associated rhizobia (Dunn, 2015; Iyer et al., 2016). Consequently, attachment to roots and root hairs provides an excellent niche for microbial growth. Therefore, it appears that rhizobia have multiple mechanisms that enable them to attach to roots and these include surface polysaccharides and

secreted/surface proteins (Rodriguez-Navarro et al., 2007). However, the role of primary metabolites in the root exudates of legume-cereal intercrop in the epiphytic interactions of rhizobia needs to be characterized.

The recent development of omics-based technologies has bridged the gap in understanding the molecular aspects of beneficial rhizobial –plant symbiosis. Especially, proteomics has helped us to identify and characterize several novel symbiosis-specific and symbiosis-related proteins and post-translational modifications that play a critical role in mediating symbiotic plant-microbe interactions (Khatabi et al., 2019). The contribution of different nitrogen-fixing bacteria to maize rootlets was revealed by metatranscriptomics (Gómez-Godínez et al., 2019), while the combination of metagenomics and metaproteomics enabled the identification of functional nitrogenase protein in the *Bradyrhizobium* found on the roots of field-grown sorghum (Hara et al., 2019). But, more detailed studies on the interactions of rhizobia, with cereal plants intercropped with the legume plants, using an omics approach are not known.

4.2. Materials and Methods

Root exudate collection protocol from monocropped *C. cajan* and *Z. mays* plants and intercropped *Z. mays* were similar to that of mentioned in chapter 2-section 2.2.5.

4.2.1. Quantification of relative gene expression of selected genes of NGR234 in the presence of *C. cajan* & *Z. mays* monocrops root exudates

To determine the molecular effect of monocrop plants root exudates on NGR234, cells were grown in 5 ml broth of RMM medium with 55mM mannitol and 6.5mM glutamate as a carbon source (Chapter 3-section 3.2.5.2) at 30°C in the presence of 0.5 mg/ml concentration of *C. cajan* & *Z. mays* root exudates separately and in the absence of root exudates (untreated). The flavonoid, 0.2 µM Apigenin was taken as a positive control. Cells were grown for up to 6 h (early log phase) & 24 h (mid-log phase). There were a total of 12 genes (Table 4.1) studied for quantitative real-time polymerase chain reaction (qRT-PCR). Among them, 9 genes were selected based on the known interactions of rhizobia with the host plants, and 3 housekeeping genes were taken as an internal control. Fold change was calculated by considering treated/untreated samples. Three biological replicates were taken for each treated and untreated sample.

4.2.1.1. RNA extraction and cDNA synthesis

For expression analysis, mRNA was extracted from treated and untreated samples at 6 h and 24 h using the RNA Plus Kit (Macherey-Nagel, Germany). Eluted RNA samples were dissolved in 60µl of Milli Q water. To avoid DNA contamination, RNA samples were given an additional step of DNase (New England Biolabs, USA) treatment at 37°C for 10 min. Further, precipitation of RNA was carried out by using the LiCl method (60µl of 5M LiCl and 120 µl of absolute ethanol). After washing with 70% ethanol, RNA was resuspended in the final vol. of 40 µl Milli Q water.

To carry out the first-strand cDNA synthesis, along with the quantified RNA using Nanodrop (Thermo Scientific, USA) random hexamers primers, dNTPs, and RevertAid Reverse transcriptase (Thermo Scientific, USA) were added to a reaction volume of 20 µl. The protocol was similar to that mentioned in the catalog of Thermo Scientific RevertAid Reverse Transcriptase (# EP0442).

4.2.1.2. Transcription analysis of NGR234 genes

Gene-specific primers were constructed and their PCR products were confirmed on 2% agarose gel electrophoresis. The annealing temperature for each gene is mentioned in Table 4.1. A Lightcycler480 instrument and SYBR Green I Master (Roche Diagnostics GmbH) were used for the real-time quantitative PCR studies. The reaction mixtures consisted of a final concentration of 5 µl SYBR Green I Master, 0.5 µl each primer (10 µM), 1 µl template cDNA obtained above, and 3 µl sterile water. Reactions were performed using the Lightcycler480 instrument with the following conditions: 10 s at 95 °C → 95 °C for 5 s and 60 °C for 34 s of 45 cycles. The $2^{-\Delta\Delta CT}$ method reported by Livak et al., (2001) was used to calculate the fold change in the real-time gene expression from the Ct values of the PCR data. To get accurate expression, targeted gene expression was divided against all three housekeeping genes.

Table 4-1 List of genes used for qRT-PCR studies in NGR234

Genes	Function	Sequence	Tm (°C)	Annealing temp (°C)	Product Size
<i>rpoD</i>	RNA polymerase sigma factor RpoD (Housekeeping gene)	F-GATCTCGAGACCACCTATTC R-CATTGGTGATGTCGTCGTCATTG	57.3 60.6	60	128
<i>recA</i>	Recombinase functioning in bacterial DNA repair (Housekeeping gene)	F-CCGGTGAAGTATCGACCTC R-ATCTCGATCTCGCGCAGAAG	61.4 59.4	60	136
<i>gyrA</i>	Bacterial enzyme that catalyses negative supercoiling of DNA (Housekeeping gene)	F-CGAGATTCCCTATCAGGTGAAC R-GCATTGGCATCGCGCTTCAG	60.3 61.4	61	145

<i>cheA2</i>	Chemotaxis protein	F-GCTCGCCAATGGTGAATTGC R-GTCGAAGGAGAAGGCAACGG	59.4 61.4	60	142
<i>cmrp</i>	Chemotaxis methyl-accepting protein	F-GAGACGCAGGTGAACGATCT R-CGGATCGTCTGTCCCAACTC	59.4 61.4	60	149
<i>flgB</i>	Flagellar basal body rod protein	F-AAGCGTACTCCAGAACACCG R-CGCCCCGTCTTCATCATTTC	59.4 59.4	58	167
<i>rhcN</i>	Type III secretion protein	F-ACACCTATTACCCGACGCAC R-GAAAGCAGCGAGGATTTGCC	59.4 59.4	61	169
<i>virB cluster</i>	Type IV secretion	F-GACTACGATCAGGACGGCG R-GTACCACACCCTCTGAGTGC	61.0 61.4	60	133
<i>ngrI</i>	Autoinducer synthase of quorum sensing	F-GCAACTGGGAAATCGACCAC R-ATTTCCGGATTTCCGCCGAG	59.4 59.4	61	163
<i>exoI</i>	Succinoglycan biosynthesis protein	F-ACCGTTATGGTCGCTACGTC R-GATCGTGCCATTTCTGTGC	59.4 59.4	59	142
<i>cpaE</i>	Pilus assembly protein	F-GGCTGGACGAGGTCTTTCTC R-TGCAGGATTTGAGGATCGG	61.4 59.4	61	130
<i>nodA</i>	N-acetyltransferase required for nod factor synthesis	F-GACCACACTGAACTCGCTGA R-GAGCGTCGTAGCCGATTACA	59.4 59.4	60	145

4.2.2 Plasmid and Genomic DNA isolation

Escherichia coli EPI300 strain carrying plasmid pFAJ1700 (Stable RK2-derived cloning vector, Ap^R, Tc^R) for constructing translational in-frame promoter GFP fusion. High-quality plasmid DNA was isolated using commercial kits. Plasmid were isolated using the QIAprep Spin Miniprep kit (QIAGEN cat# 27106) or the Qiagen Midi kit (QIAGEN cat# 12143). Plasmid extraction was performed according to the provided manufacturer protocols.

The genomic DNA of *Ensifer fredii* NGR234 was isolated by a modification of the ultra-quick genomic DNA preparation method described by Gonzalez-y-Merchand *et al.*, (1996). Cultures were grown in Tryptone Yeast extract (TY) broth of 5 mL for 48 h at 28°C with shaking.

Cultures (3 mL) were harvested by centrifugation at 10K for 10 min and the bacterial pellets were then resuspended in 500 µL lysis buffer (4 M guanidinium thiocyanate, 1 mM 2-mercaptoethanol, 10 mM EDTA, 0.1% [w/v] Tween-80). The lysate was then snap-frozen in a dry ice/ethanol bath before incubation at 65°C for 10 min. The snap-freezing/heating process was repeated twice followed by chilling the tubes on ice for 5 min. The lysate was then extracted once in chloroform, once in phenol/chloroform, then once more in chloroform. Genomic DNA was then precipitated with 0.5 ml of 100% ethanol, washed in 70% ethanol then resuspended in 50 µL of filter-sterile Milli-Q water.

4.2.3. Assembly of *nodA* promoter region-GFP fusion construct

Construct for detecting the GFP expression under the *nodA* gene of NGR234 *PnodA::gfp* was assembled by using overlap extension PCR to form a translational fusion between the *nodA* promoter regions and *gfp* at the ATG start site. Initially, the promoter regions and *gfp* were PCR-amplified as separate products using NGR234 genomic DNA and pSKGFP plasmid DNA (Kelly et al., 2013) respectively as a template. Primer sets were designed with overlapping sequences on the primers at the 3' end of the promoter regions and the 5' end of the *gfp* PCR product. Primer design incorporated restriction enzyme sites at the outermost ends of the left (LL) and right (RR) arm PCR products and 20 bp of overlapping sequence between the internal regions of the two arms (LR & RL) (Fig. 4.2). Details of the primers used are provided in Table 4.2. The left-arm promoter region of the *nodA* gene of 425bp was amplified by using primers of 20bp (*nodA* EcoRI-LL & *nodA* LR) and the right arm product size of 700 bp was obtained by using a set of primers of 20bp (*nodA* RL & *PnptIIgfp* Hind III-RR). To obtain a *PnodA::gfp* construct overlap extension PCR was then performed using the promoter region and *gfp* PCR products as templates with the outermost primers to amplify fusion product size of 1.1 kb and outward primers were added after 3 cycles of PCR.

The PCR reaction was carried out in 25µl system reaction mixtures containing 50ng of template DNA, 0.5 µl of 10 µM of each primer, 1 µl of 2.5 mM dNTPs, 0.25U of Phusion DNA polymerase (Thermo Fischer Scientific), and 0.25 µl of 5X HF buffer. Amplification was carried out with an initial denaturation at 98°C for 30 seconds, followed by 35 cycles, each consisting of denaturation at 98°C for 10 seconds, annealing temperature (Table 4.2.) for 30 seconds, extension at 72°C for 30 secs (left and right arms) and 1 min (Fusion product), final extension at 72 °C for 10

mins, and stored at 4°C. The promoter region-GFP fusion PCR product was digested with EcoRI/HindIII and cloned into similarly digested vector pFAJ1700 to form *PnodA::gfp* (*nodA* promoter region). PCR products were purified using the High Pure PCR Purification kit (Roche cat# 1732688) or extracted from agarose gel and purified using the QIAEX-II gel extraction kit (QIAGEN cat# 20021). The purification method was performed according to the manufacturer's instructions.



Fig. 4-2 Construction of 1.12 kb size *nodA* promoter region- GFP fusion Schematic of the *nod* promoter region-gfp overlap extension PCR products with the primers used indicated. Primers LR and RL contained overlapping sequences to allow the individually PCR amplified promoter regions and gfp PCR products to be joined at the ATG start codon by overlap extension PCR.

4.2.4. Plasmid transformation

4.2.4.1 Transformation into *Escherichia coli* strain

Plasmid transformation in to the *E. coli* strain EPI300 was transformed by CaCl₂ mediated competent cell preparation followed plasmid transfer by heat shock as given in Sambrook and Russell (2001). The *PnodA::gfp* construct was transformed in to the *E. coli* strain EPI300 (genotype *F- mcrA* Δ (*mrr-hsdRMS-mcrBC*) Φ 80*dlacZ* Δ *M15* Δ *lacX74* *recA1* *endA1* *araD139* Δ (*ara*, *leu*) 7697 *galU* *galK* λ - *rpsL* *nupG* *trfA* *tonA*, Str^R, Epicentre). Transformed *E. coli* EPI300 competent cells were plated onto Luria Broth containing tetracycline to select for clones harboring the plasmid DNA. Clone identified as harboring the desired construct by restriction analysis was confirmed by sequencing.

4.2.4.2 Triparental mating into NGR234

Triparental spot-matings incorporated the helper plasmid pRK2013 to aid in the transfer of non-self-transmissible plasmid from *E. coli* EPI3000. For spot-matings *E. coli* and NGR234 strains were grown to stationary phase in Tryptone Yeast extract (TY) broth. Aliquots of each culture (30 μ L) were dispensed together as a spot onto the surface of a 0.45 μ m filter (Type HA, 47 mm,

Millipore Corporation, USA) placed on a TY agar plate and incubated at 28°C overnight. The resultant bacterial growth was then streaked onto selective media containing Rifampicin + Tetracycline antibiotics.

Table 4-2 List of primer sequences used to construct *PnodA::gfp* construct

Primers	Sequences	Tm (°C)	Annealing temp. (°C)	Product Size
nodA - LL	AATTGAATTCCAACCTTCCTTGT TTTCGCGAGC	64.4	65	425 bp
nodA- LR	AGTTCTTCTCCTTTCCTCATATC CAAAGAACTCCACATCTTGC	70.4		
nodA- RL	GCAAGATGTGGAGTTCTTTGGA TATGAGTAAAGGAGAAGAAGT	70.4	63	700 bp
PnptII gfpR	ATATATAAGCTTGATGCCTGGA ATTAATTC	68.1		

4.2.5. Visualization of *PnodA::gfp* expression in NGR234 on *C. cajan* and *Z. mays* plants by CLSM

NGR234 cells carrying *PnodA::gfp* plasmid were grown into TY broth containing tetracycline (30µg/ml) as an antibiotic. The culture was freshly grown into TY broth with an inoculum of 0.1 O.D. and was allowed to grow up to 0.8 O.D. (10⁸ CFU/ml). Cells were coated onto the plant seedlings (Chapter 2, Section 2.2.2). Confocal microscopy was performed to visualize *C. cajan* plant roots using a Zeiss LSM 510 upright confocal microscope, with excitation at 488 nm for GFP and a BP505-530 emission filter for GFP. Images were recorded with LSM image browser software. While on *Z. mays* plant roots were observed through CLSM (LSM 700 Carl Zeiss, GmbH).

4.2.6. Protein expression profiling of NGR234 by label free proteomic approach

4.2.6.1 Extraction of total proteins from NGR234

To study differential protein expression in the presence of root exudates, NGR234 was grown in 5.0 ml of RMM broth with 55mM mannitol and 6.5mM glutamate as carbon sources for 48 h (~ 2.5 O.D) at 30°C in the presence of 1.0 mg ml⁻¹ of root exudates collected from monocrop *Z. mays*

and intercrop *Z. mays* individually for extraction of total proteins. The RMM broth without added root exudate served as control. Three biological replicates were taken for each sample. Total proteins were extracted as described by Gomes *et al.* (2012) with minor modifications wherein cells were harvested and centrifuged in a 2.0 ml microcentrifuge tube at 14,000 x g for 4°C and they were carefully washed with phosphate-buffered saline (3 mM KCl; 1.5 mM KH₂PO₄; 68 mM NaCl; and 9 mM NaH₂PO₄). For whole-cell protein extraction, washed cells were resuspended in 500 µl lysis buffer (8M urea; 2% CHAPS; 50mM Tris-Cl; and 50mM DTT) and lysed by using a sonicator (9.9 sec on and off cycles, for 5 min at 35% amplitude). The lysates were separated from particulate material at 16,000 x g for 15 min, at 4°C. An additional step of concentration with phenol was done to remove any associated nucleic acids (Faurobert, Pelpoir, & Chaïb, 2007). Briefly, aliquots (500 µL) of the lysates were mixed with a solution containing 0.8 ml of Tris-buffered phenol pH 8.0, and 0.8 mL of SDS buffer (0.1 M Tris-HCl pH 8.0; 2% SDS; 5% β-mercaptoethanol; 30% sucrose). The samples were homogenized for 5 min and centrifuged at 16,000 x g for 15 min at 4°C, and the top phenol layer (500 µL) was transferred to a new tube. Proteins were precipitated for 1 h, at -20°C with three volumes of pre-cooled 0.1 M ammonium acetate in absolute methanol, and then centrifuged (16,000 x g for 15 min at 4°C). The pellet was washed once with pre-cooled methanol and once with pre-cooled 80% v/v acetone. Samples were air-dried and stored at -20°C till further use.

4.2.6.2 In-solution digestion of proteins

Trypsin digestion was carried out using 15 µg of NGR234 protein from each sample. Protein concentration was determined by Bradford's method (Bradford, 1976). Protein samples were reduced with 10 mM dithiothreitol (DTT) and incubated at 37°C for 30 min followed by the addition of 30 mM iodoacetamide and incubated in the dark (30°C, 30 min.) to alkylate the free cysteine residues. The samples were diluted with 1M urea and pH was adjusted to 8.0 before the addition of trypsin (1:50 (w/w) and incubated at 37°C for overnight digestion. The final concentration of 0.1% formic acid was used to quench the reaction. The digested peptides were dried under vacuum and desalted using a C-18 zip tip (Pierce C18 Tips, 10 µL bed). The eluted peptides were dried and re-dissolved in 2% acetonitrile/0.1% formic acid.

4.2.6.3 LC-MS/MS conditions and analysis

The digested peptide sample was injected into the Thermo Orbitrap fusion tribird mass spectrometer coupled with an EASY-nLC 1200 series system. The peptides were injected into

reverse-phase C18 pre-column (Acclaim PepMap 100, C18, 3 μm , 75 $\mu\text{m} \times 2 \text{ cm}$ nanoviper) and then separated on C18 analytical column (Easy spray Pepmap RSLC, C18, 2 μm , 15 $\text{cm} \times 75 \mu\text{m}$) for a resolved separation. Peptides were eluted using a linear gradient for 168 min. from 5% to 45% solvent B (80% acetonitrile in 0.1% formic acid) at a flow rate of 300 nL/min, 45% B to 98% B for 12 min and solvent A (0.1% formic acid in LC grade water). The mass spectrometer was operated with a positive ionization voltage of 1900 V and 273 °C temperature for the ion transfer tube. MS spectra were acquired in the orbitrap with a resolution of 120000 over a mass range of 375-1700 m/z, automatic gain control (AGC) value was set to 4.0 e5 and a maximum injection time was kept as 50 ms. The 20 highly intense ions for fragmentation were selected by Top-Speed acquisition mode which was isolated by quadrupole with an isolation width of 1.2 Da. These ions with charge states ranging from 2+ to 7+ were fragmented by high energy collision-induced dissociation (HCD) with an optimized collision energy of 30% with step energy ± 5 . The fragmented ion spectra were acquired by the ion trap in centroid mode, AGC value was set to 1×10^4 and a maximum injection time of 35 s was used. The data acquisition was done with Xcalibur software (Version- 4.1.31.9).

The MaxQuant software (v. 1.6.7.0) with the Andromeda search engine was used to analyze the LC-MS/MS data. The spectra were searched against the *Ensifer fredii* protein sequences which were downloaded from Uniprot in Aug 2018. A precursor mass tolerance of 10 ppm, MS/MS fragment ion tolerance of 0.6 Da were used for database search. Besides, enzyme specificity for trypsin (for up to two missed cleavages), methionine oxidation (M), and acetylation (protein N-term) were set as variable modifications while cysteine carbamidomethylation (C) was set as a fixed modification. Peptide identification was performed using the criteria that proteins should contain two unique proteins at a 1% false discovery rate (FDR). A 1% FDR was set for both peptides spectral match and protein identification. The Perseus software (version 1.6.2.3) was used to carry out the bioinformatics analysis. Proteins identified in at least two out of the three replicates and with at least 2 unique peptide hits were included for further peptide quantitation analysis. Differential expression analysis was performed using LFQ intensities. After Log₂ transformation of the intensities and normalization based on "with adjustment", a two-sample Student's T-test was used to determine differentially abundant proteins. Scatter plots were used to determine the correlation between biological replicates. The Log₂ FC values (Student's T-test difference between Log₂ intensities of treated (root exudates) and untreated (Control)). Following

criteria was used to consider up-regulation ($\text{Log}_2 \text{FC} \geq 1.5$ and p values ≤ 0.05) and down-regulation ($\text{Log}_2 \text{FC} \leq -1.0$ and p -values ≤ 0.05).

4.2.7. Bioinformatic analysis of biological pathways and protein-protein interactions

Upregulated proteins of NGR234 in the presence of intercrop *Z. mays* root exudates were analyzed using STRING (Szklarczyk et al., 2015) to identify the protein-protein interactions and proteins that are part of KEGG pathways. The initial cut-off false discovery rate used for identifying upregulated proteins associated with KEGG pathways was 0.05. Proteins that showed the expression of the above 2 Fold change (Treated/Untreated) and those matched with NGR234 proteins were identified and used for protein-protein interaction analysis. Each node in a network represents an identified protein with a three-dimension structure and each network edge represents an interaction.

4.3 Results

4.3.1. Differentially expressed NGR234 genes in the presence of *C. cajan* and *Z. mays* monocrop plants root exudates

Root exudates released from plant roots are known to be key mediators for bacterial attachment. To understand the molecular mechanisms involved in the interactions of NGR234 with the *Z. mays* plants, the relative gene expression studies of NGR234 by qRT-PCR was carried out in the presence of *C. cajan* and *Z. mays* monocrop plants root exudates. In total, 13 genes (Fig. 4.3 & 4.4) were selected based on their involvement in plant-microbe interactions such as attachment and nodulation with host plant legume. The 9 genes selected (*nodA*, *cmrp*, *cheA*, *rhcN*, *virB*, *ngrI*, *flgB*, *exoI*, *cpaE*) are involved in the interactions with *Z.mays* plants and 3 housekeeping genes (*rpoD*, *gyrA*, *recA*) were used as an internal control. The differential genotypic response of NGR234 was noted in the presence of root exudates at a concentration (0.5 mg/ml) of *C. cajan* and *Z. mays* and in its absence at 6 h and 24 h.

Our results demonstrated that the genes related to chemotaxis, flagellar motility, and attachment (Fig.4.3. A-E) were found to be induced at a similar level in both the root exudates at 6 h. A significant ($p < 0.05$) induction was obtained with *cmrp* at 24 h with the *Z. mays* root exudates. Also, a significant induction at 24 h was observed in *exoI* (Fig. 4.3.D) in the presence of *Z. mays* root exudates. Interestingly, we also observed that the *Z.mays* root exudates stimulate nodulation gene (*nodA*) to a similar level compared to *C.cajan* (host) plant at both 6 h and 24 h (Fig.4.4.A). Our data also suggest that autoinducer (AI) synthase molecule (*ngrI*) for quorum sensing released by NGR234 is significantly ($p < 0.001$) induced with the root exudates of *Z.mays* plant at 24 h. Further, root exudates of *Z.mays* also induced genes at 24 h which are involved in Type III secretion system (*rhcN*) and Type IV secretion system (*virB*) at a significant level of $p < 0.01$ and $p < 0.001$ respectively. Therefore, our qRT-PCR results suggest that both *C. cajan* and *Z. mays* root exudates significantly affect the NGR234 strain and their interactions involved through chemotaxis, secretion systems, quorum sensing systems.

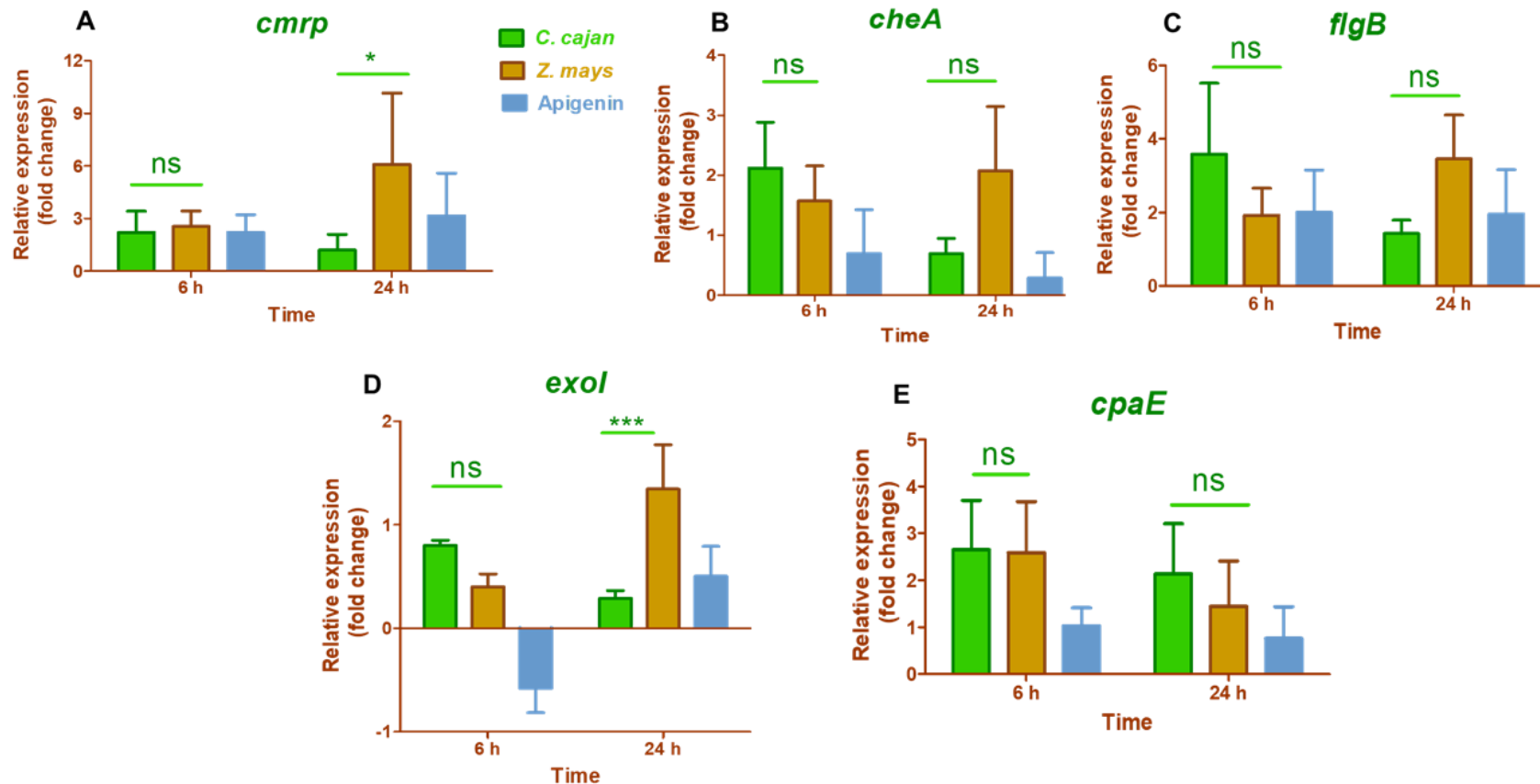


Fig. 4-3 Relative gene expression levels studies of NGR234 inducing the genes of chemotaxis, flagellar motility, and attachment with root exudates. Fold change is with respect to the mean of treated samples with untreated samples. The NGR234 strain was treated with or without *C. cajan* and *Z. mays* root exudates. Apigenin (flavonoid) was taken as a positive control. Error bars indicate standard deviation based on three independent values. Significance was measured by two-way ANOVA followed by Bonferroni's multiple comparison post hoc test. 'ns' if non-significant, * $p < 0.05$, *** $p < 0.001$.

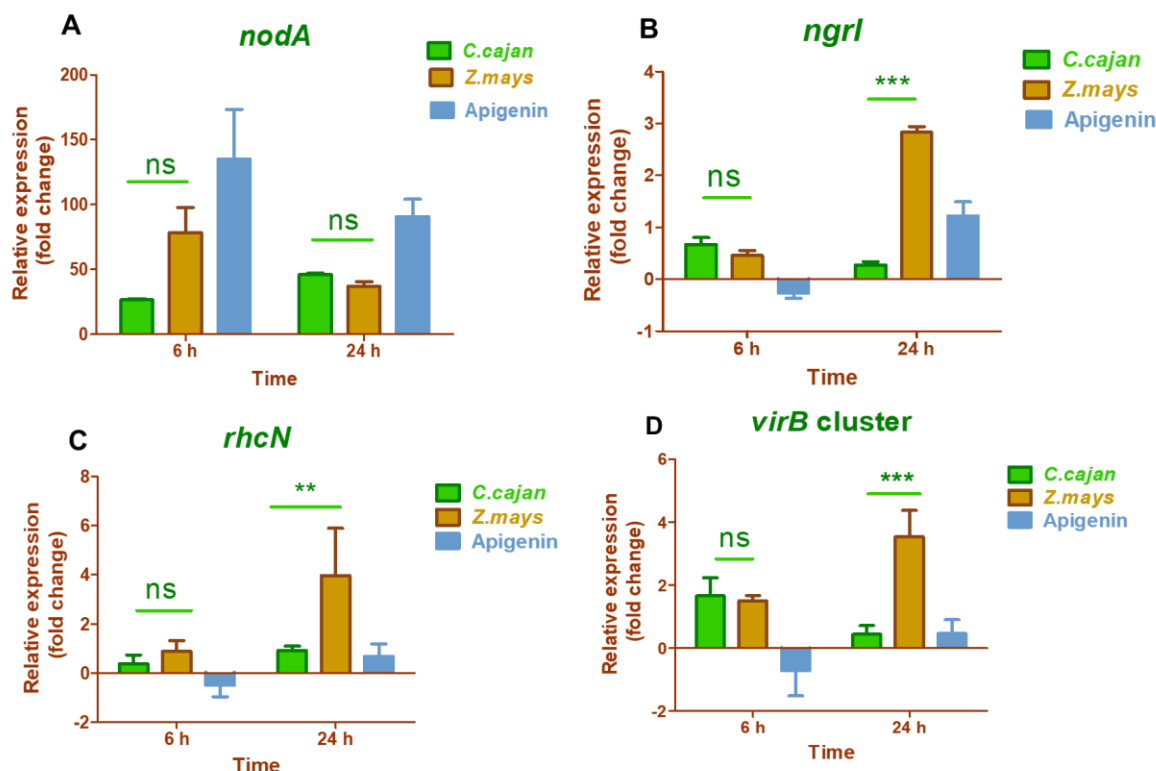


Fig. 4-4 Relative gene expression levels studies of NGR234 inducing the genes of nodulation, quorum sensing, and type III-IV secretions systems with root exudates Fold change is with respect to the mean of treated samples with untreated samples. The NGR234 strain was treated with or without *C. cajan* and *Z. mays* root exudates. Apigenin (flavonoid) was taken as a positive control. Error bars indicate standard deviation based on three independent values. The data were subjected to two-way ANOVA followed by Bonferroni's multiple comparison post hoc test. 'ns' if non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.3.2. Visualization of *PnodA::gfp* expression on *C. cajan* and *Z. mays* roots

To confirm the expression of *nodA* in the presence of *C. cajan* and *Z. mays* root exudates from the qRT-PCR studies, a translational fusion between the *nodA* promoter regions and *gfp* at the ATG start site was prepared (Fig. 4.2.) and clone was confirmed and digested with EcoRI/HindIII restriction enzymes (Fig. 4.5). A 1.1 kb of the fused product was obtained through the overlap extension PCR approach (Fig 4.5. B). Further, results strikingly reflected the expression of *nodA* onto the roots of *C. cajan* and *Z. mays* plants (Fig. 4.6). It was interesting to note that the flavonoids or phenolic-like compounds activated the expression of *nodA* along with the root hairs of *C. cajan*

plants (Fig. 4.6 A) after 5 days of inoculation. While on *Z. mays* plants (Fig.4.6 B), its expression was found after 7 days of inoculation along the elongation region of the roots.

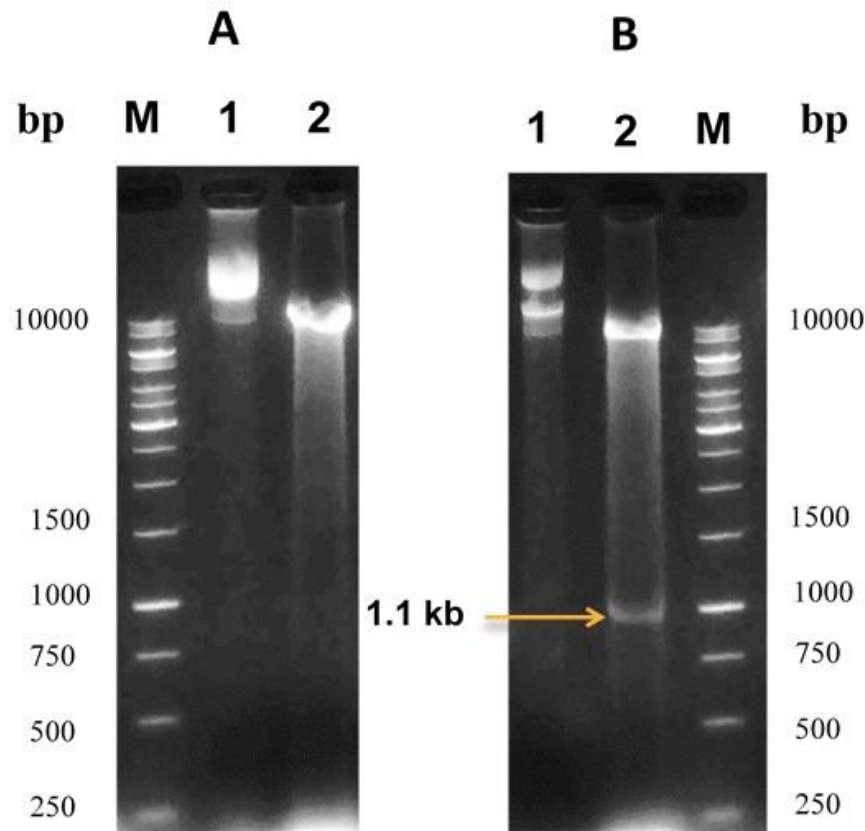


Fig. 4-5 Confirmation of *PnodA::gfp* clone from *E. coli* EPI300 on 1% agarose gel electrophoresis A) Undigested & Digested empty vector pFAJ700 with a size of 10.5 kb. Lane 1: Undigested vector, Lane 2: Digested with EcoR1/HindIII restriction enzymes and M represents ladder of 1 kb lambda DNA; B) Undigested & digested clone. Lane 1: Undigested clone of size 11.6 kb, Lane 2: Insert release of 1.1 kb fragment and vector 10.5 kb

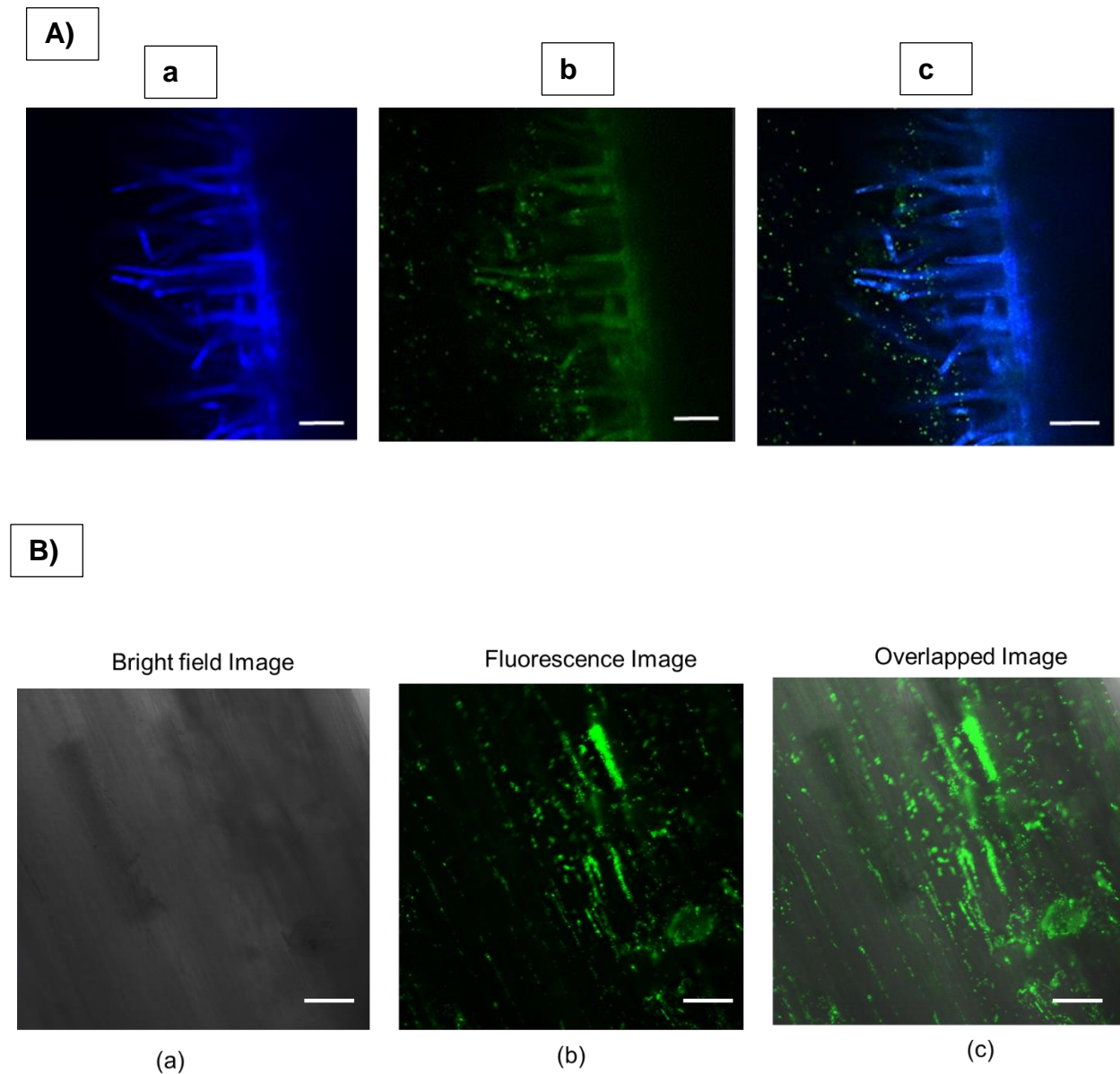


Fig. 4-6 Detection of *PnodA::gfp* expression in NGR234 on the roots of *C. cajan* and *Z. mays* plants. A) *C. cajan* plants: NGR234 carrying a *PnodA::gfp* in the pFAJ1700 vector on *C.cajan* imaged by dual laser CLSM at 5days post-inoculation. Dual laser CLSM images showing a) Excitation at 405nm, b) Excitation at 561nm, c) overlapped images. Scale bar represents 20µm. B) *Z. mays* plants: NGR234 carrying a *PnodA::gfp* in the pFAJ1700 vector on *Z. mays* imaged by single laser CLSM at 7days post-inoculation. Scale bar represents 20µm.

4.3.3. Differentially expressed NGR234 proteins in presence of monocrop and intercrop *Z. mays* root exudates

To unravel the early molecular signaling of NGR234 initiated during colonization on roots of intercrop and monocrop *Z.mays*, total proteins of NGR234 were isolated and were analyzed by a label-free quantitative proteomic approach. A total of 2570 proteins were identified across all three samples i.e. untreated cells (control), intercrop *Z. mays*, and monocrop *Z. mays* root exudates treated NGR234 cells as represented in the Venn diagram present in 2 or 3 biological replicates (Fig. 4.7). There were around 50 % of proteins in each sample that could be annotated in the UniProt database based on biological function. There were 50 unique proteins in control, 100 in intercrop *Z. mays*, and 51 in monocrop *Z. mays*. Functional classification based on gene ontology biological function (Fig. 4.8) using Uniprot showed the highest number of proteins were uniquely present in NGR234 treated with intercrop *Z. mays* root exudates. The majority of them are related to signal transduction, nodulation, nitrogen fixation, and cellular organization. The cells treated with monocrop *Z. mays* root exudates indicated the presence of a higher number of proteins responsible for post-translational modification, while the DNA replication and repair proteins decreased compared to control. The proteins related to nodulation arranged in nod-boxes like NB-2(**NoeL**), NB-8(NodA, NodI, **NoIO**, NoeI), NB-12 (**NodU**), and NB-19 (SyrM2) were notably found in treated NGR234 samples (Bold in parenthesis proteins were common in both root exudates treated samples; rest were specifically found by intercrop root exudates treated samples). It was also interesting to note that both GMD and FCL proteins, part of fucose biosynthesis on nod factors, were expressed in the presence of intercrop *Z. mays* root exudates in NGR234.

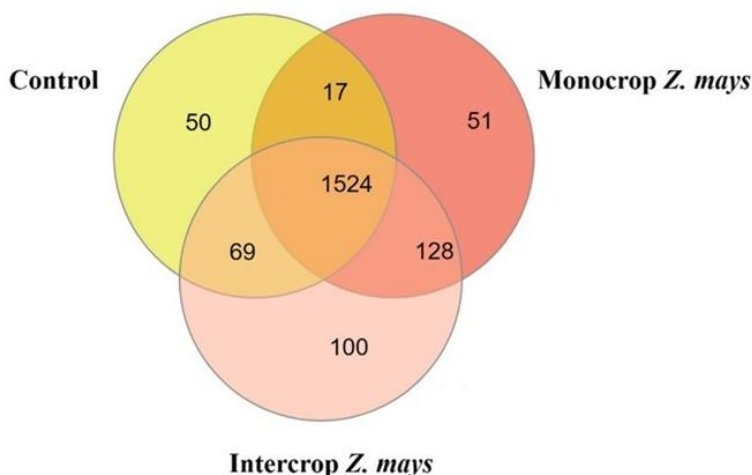


Fig. 4-7 Venn diagram representing the proteome analysis of NGR234 Cells were treated with root exudates of monocrop *Z.mays* and intercrop *Z.mays* (with *C. cajan*) analyzed by LC/MS/MS.

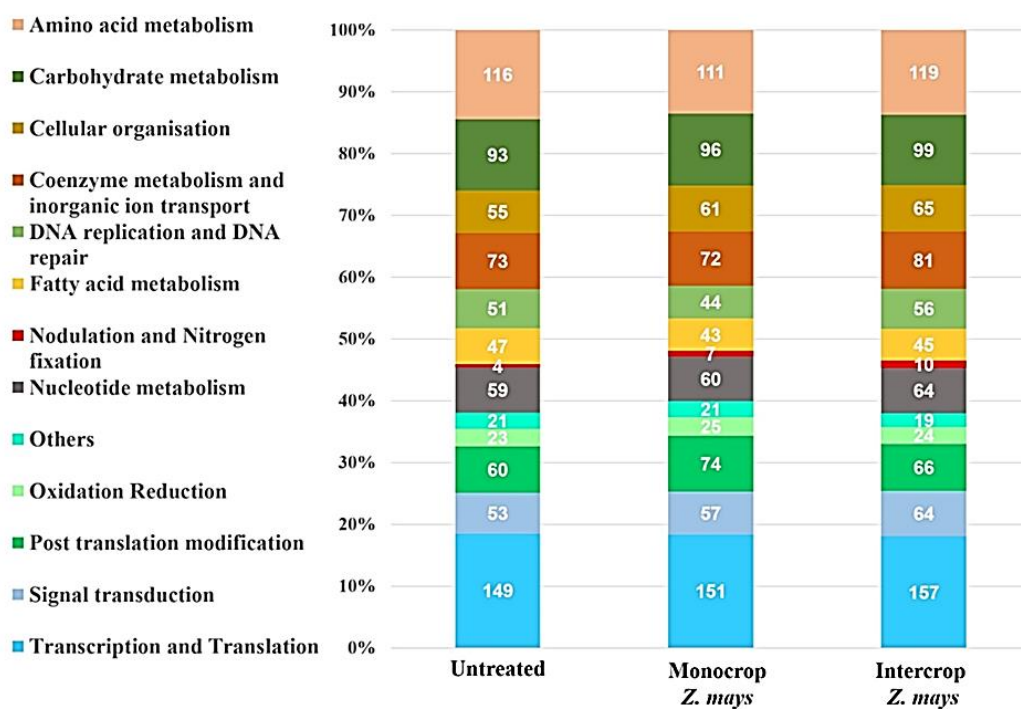
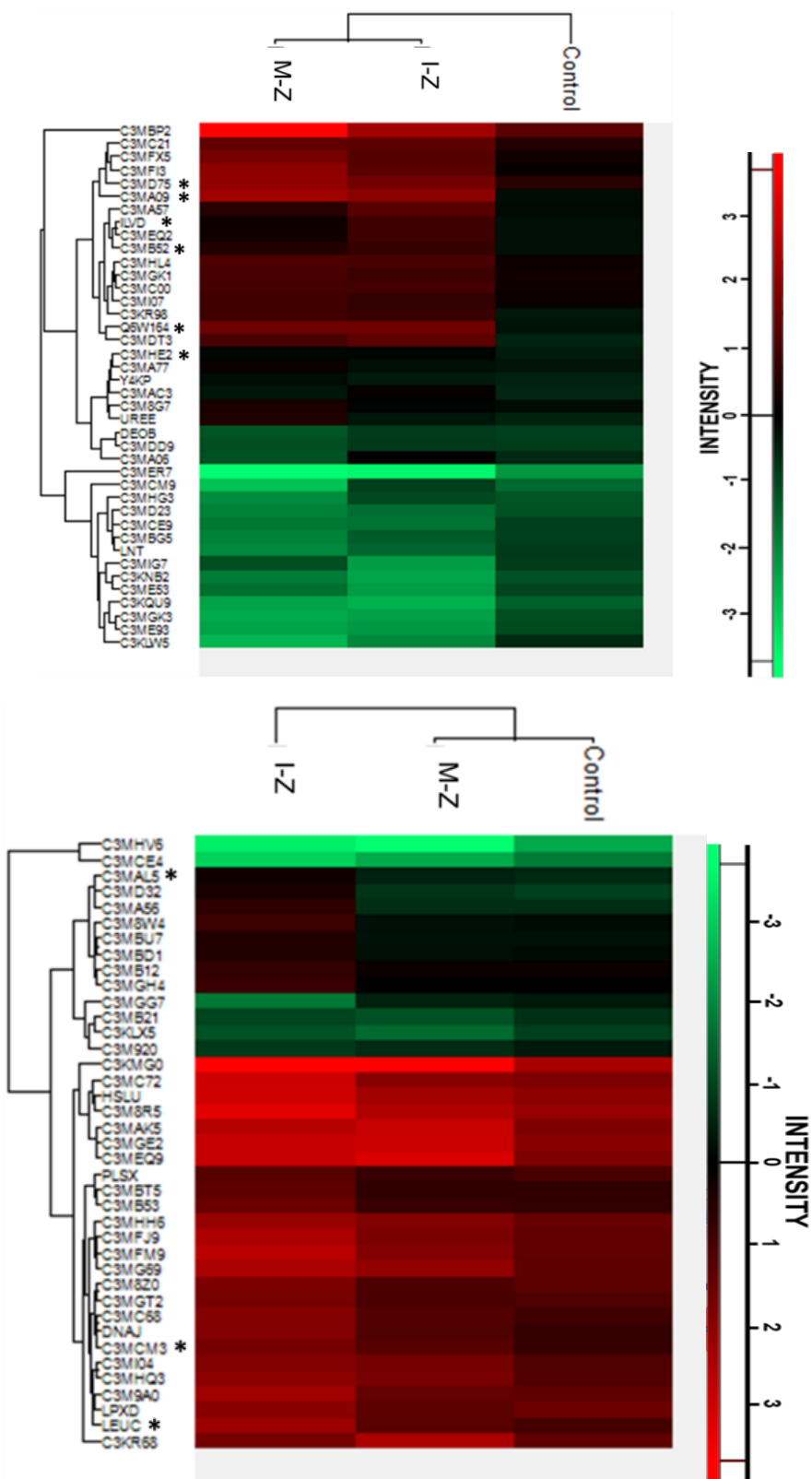


Fig. 4-8 Gene Ontology-based biological classification of the identified proteins through Uniprot

4.3.4. Quantitative expression of proteins in presence of monocrop and intercrop *Z. mays* root exudates

Unsupervised hierarchical cluster analysis of NGR234 expressed proteins demonstrated a clear separation between intercrop and monocrop *Z. mays* root exudates presence when compared to control. Out of 1519 proteins, 161 proteins showed differential expression across the three conditions which were represented in the heat map (Fig. 4.9; Appendix Table A2). Overall, 18, 52, and 35 proteins of NGR234 were expressed (Log_2 fold increase >1.5) in the case of control, and the presence of intercrop *Z. mays*, and monocrop *Z. mays* root exudates respectively. Whereas, the number of down-regulated (Log_2 fold decrease >1.0) proteins were 24, 38, and 34, in the three treatments, respectively (Appendix Table A3).

From the data analysis, it can be inferred that a total of 10 proteins are likely to participate in the adaptation of rhizobia with the *Z. mays* plants at a molecular level (Table 4.3). Among these were 4 proteins as glutaredoxins are involved in redox reactions, Metallo beta-lactamase in quorum quenching of autoinducer molecules, pseudoazurins electron donor to copper-containing nitrite reductase, and GroES2 chaperonin proteins were found in the presence of both intercrop and monocrop *Z. mays*. Proteins like dihydroxy-acid dehydratase (IlvD), 3-isopropyl malate dehydratase large subunit (LeuC); 3-isopropyl malate dehydratase small subunit (LeuD) (Fig. 4.10), putative nitroreductase, and bacterioferritin co-migratory protein (BCP) and GroEL- Cpn60 proteins were only up-regulated in presence of intercrop *Z. mays* root exudates. On the other hand monocrop, *Z. mays* induced the expression of PurD (Table 4.3). Proteins like AracA2 involved in the arginine deiminase pathway were significantly downregulated in both the treated samples, while HutU involved in histidine catabolism was downregulated in both treatments but significant only in the presence of monocrop *Z. mays* root exudates compared to control. NGR234 proteins like PheA and HisC expression were not significantly different from control. However, there was a significant upregulation ($p < 0.05$) by the root exudates of intercrop *Z. mays* compared to monocrop *Z. mays*. Nucleoside diphosphate (NDK) protein, involved in quorum sensing, was upregulated by root exudates of *Z. mays* (Fig. 4.10).



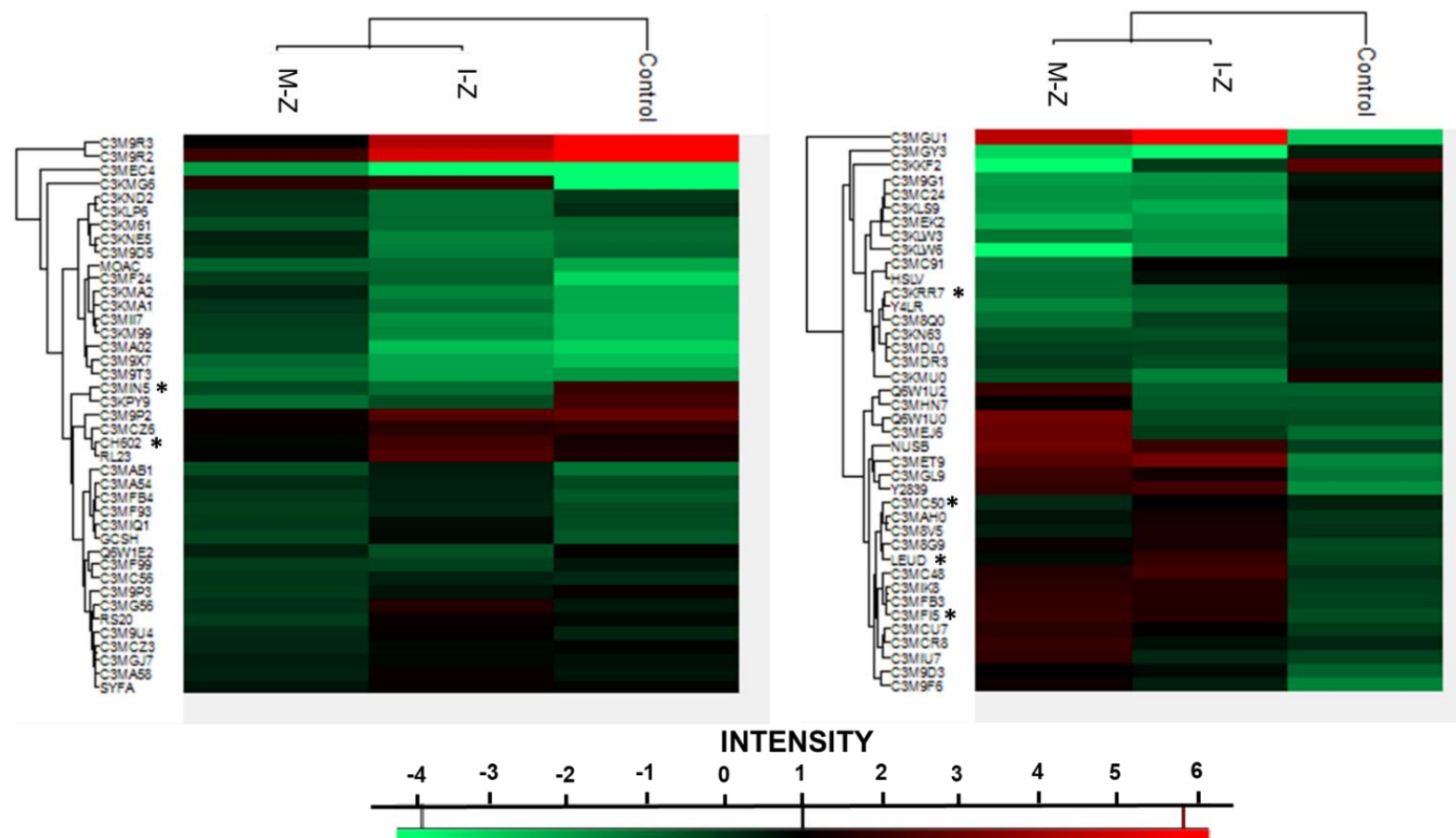


Fig. 4-9 Heat map of 161 proteins differentially expressed in the presence of *Z. mays* monocropped and intercropped root exudates. NGR234 proteins expressed in presence of intercrop *Z. mays* (I-Z) and monocrop *Z. mays* (M-Z) root exudates were compared to the untreated cells (Control). Proteins identified in at least two of the three biological replicates were considered for the analysis. A heat map was generated using Unsupervised Hierarchical clustering with Perseus software (version 1.6.2.3).

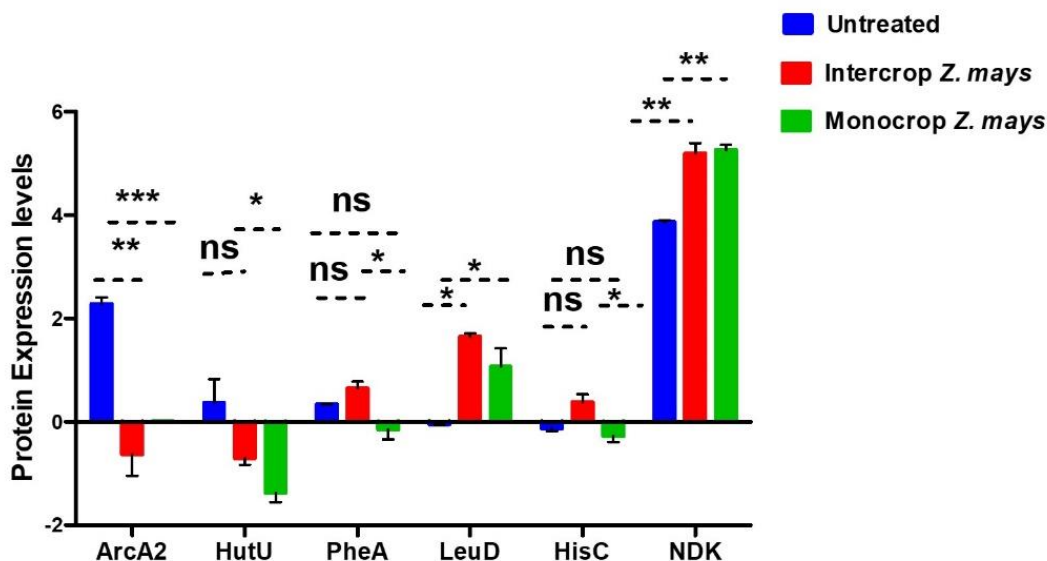


Fig. 4-10 Differential expression of proteins in the absence or presence of monocrop and intercrop *Z. mays* root exudates Error bars indicate the mean of standard deviation. Statistical analysis was performed by one way ANOVA followed by Bonferroni's multiple comparison post hoc test. 'ns' if non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Two biological replicates were considered for each treated and untreated sample.

4.3.5. Biological pathways and protein-protein interactions of upregulated proteins of NGR234 in the presence of intercrop *Z. mays* root exudates

Out of 161 proteins (Fig. 4.9), there were in total 45 proteins of NGR234 which showed upregulation by 2 Fold change in the presence of intercrop *Z. mays* root exudates. These proteins were further analyzed for biological interactions and protein-protein interactions by STRING (Fig. 4.11). It was interesting to note that the PPI enrichment p-value of this network analysis was 0.00855. This means that proteins have more interactions among themselves than what would be expected for a random set of proteins of similar size, drawn from the genome. Such an enrichment indicates that the proteins are at least partially biologically connected, as a group. For instance, proteins like ILVD, LEUC, and LEUD belongs to the branched-chain amino acid biosynthesis pathway were found to detect with a false discovery rate of 0.0066 indicating the significance of the enrichment. The line thickness (Fig. 4.11) between the three proteins (ILVD, LEUC, and LEUD) indicates the strength of the data support.

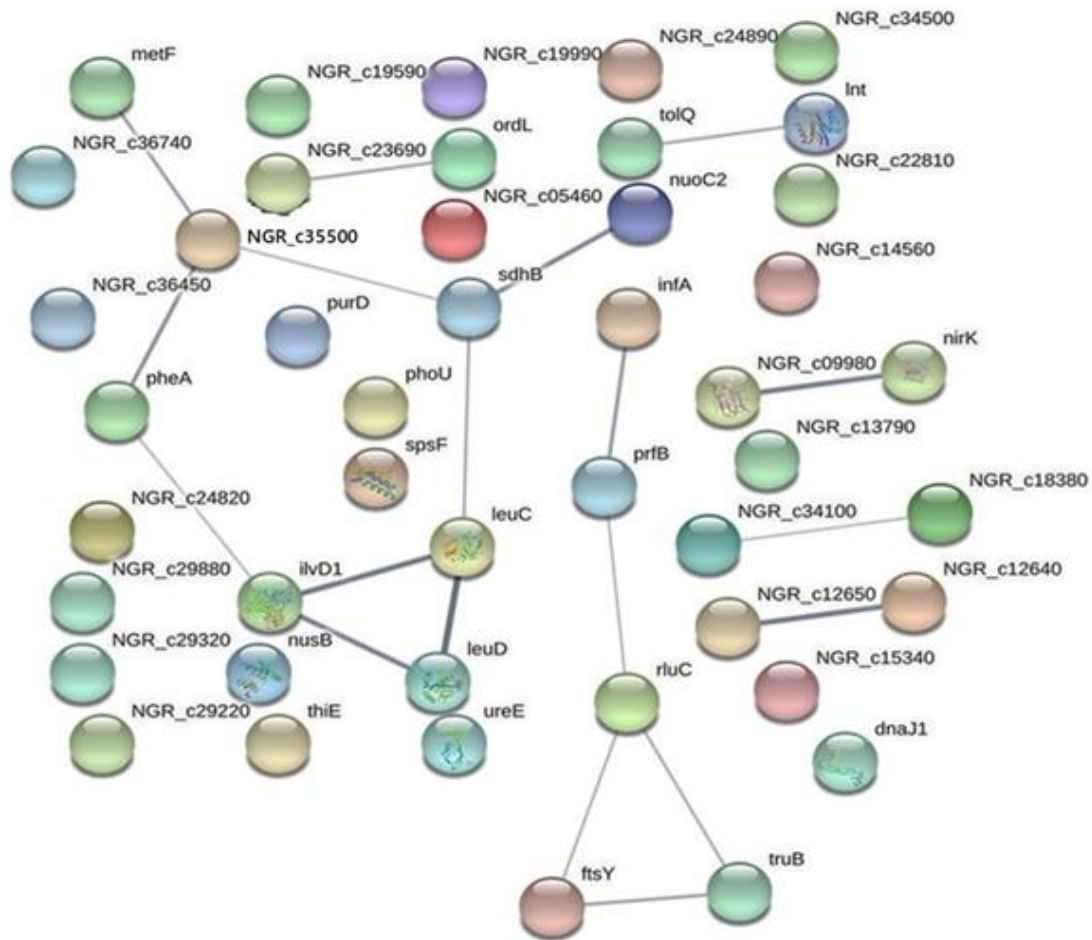


Fig. 4-11 Upregulated proteins in the presence of intercrop *Z. mays* root exudates associated with biological pathways were subjected to protein-protein interactions using STRING

Table 4-3 List of NGR234 proteins differentially expressed upon treatment with monocropped and intercropped *Z. mays* root exudates identified by label-free quantitative proteomics

Uniprot Protein ID	Name of the protein	Average Fold change (Intercrop <i>Z.mays</i> / Control)	p-value	Average Fold change (Monocrop <i>Z.mays</i> / Control)	p-value
C3MD75	Glutaredoxin	1.9	0.042	2.3	0.043
C3MFI5	Metallo-beta-lactamase family protein	1.7	0.01	2.2	0.042
C3MA09	Pseudoazurin	6.3	0.01	7.0	0.002
Q6W164	10 kDa chaperonin (GroES protein - Cpn10)	6.6	0.02	6.6	0.01
CH602	60 kDa chaperonin 2 (GroEL protein 2- Cpn60)	1.5	0.03	-	-
ILVD	Dihydroxy-acid dehydratase (ilvD)	4.1	0.01	-	-
C3MB52	Putative nitroreductase protein	3.8	0.043	-	-
LEUC	3-isopropylmalate dehydratase large subunit (leuC)	2.3	0.03	-	-
C3MCM3	Bacterioferritin comigratory protein	2.2	0.043	-	-
C3MHE2	Phosphoribosylamine--glycine ligase (purD)	-	-	2.3	0.03

Proteins up-regulated with Log₂ fold change ≥ 1.5 and p -values ≤ 0.05 and expressed in three biological replicates were considered. Here “-” denotes the expression of a protein with no significant fold change. A student t-test was applied to know the significant fold change in NGR234 proteins between REs treated and control samples.

4.4 Discussion

Most plant species depend on the uptake of soil N while, certain clades, most notably the legumes, are capable of fixing N via a symbiotic relationship with rhizobia (Carranca 2013). To improve nitrogen supply in cereal cropping systems one strategy is by the inclusion of nitrogen-fixing legumes, which can provide nitrogen benefits to companion crops through belowground nitrogen transfer (Thilakarathna et al., 2016). A recent study carried out by Zhao et al., (2020) suggested that interspecific facilitation in the alfalfa/triticale intercropping system resulted in enhanced nodulation and N fixation ability, which thereby enhanced N use efficiency. A better understanding of the underlying mechanisms that govern rhizobia interactions with the intercropped plants is important to improve their association. So far studies carried out indicate the interactions of rhizobia with cereals plants however the detailed molecular mechanisms are less known. Thus the present work was focused on finding the commonalities in their association of NGR234 at a molecular level with a legume (*C. cajan*) and non-legume (*Z. mays*) plants and their molecular signaling with *Z. mays* plants in monocropping and intercropping setup.

4.4.1 Molecular effects displayed by *C. cajan* and *Z. mays* root exudates on NGR234

From qRT-PCR (Fig. 4.3 & 4.4) studies it was an interesting to note that the phenolic-like compounds present in the root exudates of *Z. mays* induced the expression of *nodA* (Fig. 4.4 A & Fig. 4.6 B) on plants as well. These results corroborate well with the previous study carried out by Le Strange (1990) in an in-vitro study with *nodA-lacZ* fusion. In addition, to the induction of nod genes, it's also known that the overall production of quorum sensing (QS) molecules like *nglI* (autoinducer-AI synthase) gets enhanced in the presence of flavonoids (Pérez-Montaña, F. et al., 2011; Grote et al., 2014). Subsequently, it induces transcription of the *exoI* responsible for the succinoglycan biosynthesis gene and which is known to enhance the transcription of the *Exo* cluster in the presence of AI molecules (Krysciak.D et al., 2014). Induction of type III and type IV secretion systems suggests their role in the infection of *Z.mays* plants as observed similarly in rice-*Bradyrhizobia* interactions (Piromyou, P. et al., 2015). Hence, our results from qRT-PCR studies signify that NGR234 recognizes and responds well to the *Z. mays* exudates compared to *C. cajan* exudates through common mechanisms like chemotaxis, secretion systems, quorum sensing in response to phenolic like compounds present in the *Z.mays* roots.

4.4.2. Differentially expressed proteins of NGR234 at the stationary phase in the presence of *Z. mays* root exudates of monocropped and intercropped plants

Our proteomic analysis revealed that the root exudates of *Z. mays* induced the expression of nod-box (NB-2, NB-8, NB-12) proteins in NGR234 (Fig. 4.7, 4.8). The nod-box proteins are known to be positively regulated by the flavonoid responsive protein, NodD1 by both legumes and non-legume (wheat) plants (Le Strange, 1990; Kobayashi et al. 2004). Besides, we could also find clear upregulation of GroES2 co-chaperone proteins by the root exudates of *Z. mays*. Additionally, GroEL2 protein is also upregulated by root exudates of intercrop *Z. mays*. The GroEL/GroES chaperonins are involved in the folding and/or assembly of active NodD proteins to facilitate the regulation of nod gene expression (Ogawa et al., 1995; Yeh et al., 2002). Our findings are in concurrence with Liang et al. (2013) who suggested that there is perception of signal and response of rhizobial nod factor towards non-legumes, including corn, suggesting the possibility of nod factor- mediated signaling in the intercropping system.

GMD protein induced in presence of both monocrop and intercrop *Z. mays* root exudates is a homolog of NoeL protein-encoding GDP-D-mannose dehydratase, an enzyme involved in the synthesis of GDP-L fucose. Additionally, GDP-L-fucose synthase (FCL) protein required for fucose biosynthesis induced only in the presence of intercrop *Z. mays* root exudates, is known to mediate fucosylation of Nod factors. Inactivation of *noeL* affects nodulation rate and competitiveness for nodulation in soybean. (Lamrabet et al., 1999). Besides, it was interesting to note the expression of NodU in the presence of both monocrop and intercrop root exudates of *Z. mays*. In contrast, the root exudates from rice plants did not show the activation of the *nodA* of *R. leguminosarum* bv. *trifolii* and *nodSU* of *Rhizobium*. sp. NGR234 (Reddy et al., 1997). Overall our results with respect to nod box proteins imply that flavonoids present in the root exudates of both monocrop and intercrop *Z. mays* (Li et al., 2016) modulated nodD1 gene-dependent nod boxes of NGR234. Furthermore, flavonoids are also known to regulate quorum sensing (QS)-biofilm formation which leads to the development of the symbiotic biofilm on legume roots (Rinaudi et al., 2010). A significant up-regulation of NDK protein (nucleoside diphosphate kinase) (Fig. 4.10) associated with QS in the planktonic cells helps in the initial metabolic changes that occur before surface attachment (Reyes-Perez et al., 2016). Therefore, from our physiological studies and protein identification, we surmised that QS-biofilm in NGR234 allowed the formation of the

symbiotic biofilm (Chapter 3- section 3.3.3) and successful root colonization (Chapter 2- section 2.3.3) on the *Z. mays* plants as well.

Further, our data suggested that BCP was upregulated in NGR234 during interaction with intercrop *Z. mays* plant root exudates and glutaredoxin (Fig. 4.9 & Table 4.3), part of the antioxidant defense mechanism were induced by both intercrop and monocrop *Z.mays* plant root exudates. It is known that during the legume-rhizobia association, ROS and antioxidant defense mechanisms play a crucial role in symbiosis with legumes (Chang et al., 2009; Benyamina et al., 2013). Also, to infect the plant successfully or to down-regulate the plant ROS-producing systems, the bacteria depend on ROS scavenging enzymes (Nanda et al., 2010). Recently, Liu et al., (2019a) also described the role of BCP in the bacterial defense against H₂O₂ at the free-living stage of rhizobia. Meanwhile, putative nitroreductase protein (known as oxidoreductase) was induced by root exudates of intercrop *Z. mays* in NGR234, possibly involved in the oxidative stress response (De Oliveira et al., 2010). Based on our results it can be speculated that to maintain cellular redox homeostasis in bacteria, a similar mechanism exists while interacting with the non-legume, *Z. mays*.

Interestingly, we also noticed upregulation of key proteins of branched-chain amino acids like IlvD, LeuC, LeuD in NGR234 by intercrop *Z. mays* root exudates and IMP pathways like PurD by monocrop *Z. mays* root exudates in NGR234. Also through the biological network analysis STRING (Fig. 4.11), we found significant interactions between branched-chain amino acids proteins (ILvD, LeuC, LeuD). The importance of these proteins has been described for the free-living growth of betarhizobia, as well as for their ability to form effective symbioses with their host plants (Prell et al., 2009; Chen et al., 2012; Okazaki et al., 2007). Recently, the role of leucine in the interaction between pathogens and hosts was found which highlighted that amino acids not only act as nutrients for bacterial growth but also play essential roles in the *Xanthomonas oryzae* pv. *oryzae* and rice interaction (Li et al., 2019). Based on these findings, we speculate that the branched-chain amino acids might play a crucial role in rhizobial interaction with *Z. mays* plants. Besides, there was an activation of phenylalanine pathway proteins (PheA and HisC) (Fig. 4.10) by root exudates of intercropped *Z. mays* which are reported to be induced in *R. etli* when infecting bean plants (Ferraioli et al., 2002). The catabolic pathway of histidine (HutU; urocanate hydratase) was found to be downregulated upon treatment of root exudates as similarly observed

in the case of *R. leguminosarum* bv. *viciae* 3841 in the rhizosphere of host or non-host plants (Ramachandran et al., 2011).

Furthermore, present data reflected upregulation of putative beta-lactamase family protein in NGR234 upon treatment with root exudates of both the plants, suggesting a probable role involved in fitness advantage on *Z. mays* plants as observed similarly on the cowpea roots (Krysciak et al., 2011). Interestingly, pseudoazurins proteins were found predominantly in case of exposure to both the intercrop and monocrop *Z.mays* plant root exudates. These are soluble blue copper proteins that serve as electron donors to nitrite reductase (NirK and NirS) present in the periplasmic region of the bacteria (Chen and Strous, 2013). One of the pseudoazurins was also strongly up-regulated in *Leucaena leucocephala* nodules infected by NGR234 compared to the free-living condition (Li et al., 2013). Till now, its role has been presumed in the denitrification process and no direct association with symbiotic interaction.

To summarize, the proteome of NGR234 was largely similar in the case of treatment with both monocrop and intercrop *Z. mays* as compared to the untreated sample. However, notable variance in proteins are related to chaperonin proteins (GroES2 and GroEL2), and efficient symbiosis/interaction (BCP, IlvD, LeuC, LeuD) are expressed significantly by intercropping *Z. mays* root exudates, as evident from our studies. The upregulation of branched-chain amino acid proteins by the root exudates of intercrop *Z. mays* provides an additional basis for understanding the rhizobial –*Z. mays* interaction in intercropped plants. This work, therefore, identified the crucial proteins of rhizobia that could be involved in positive interactions with *Z. mays* (non-legume) plants in both monocrop and intercrop conditions.