Understanding of Root Nodule Development at Level of System Biology as Obtained by High Throughput Transcriptomic Approach

7

Akanksha Bhardwaj and Senjuti Sinharoy

Abstract

Nodule development starts with dedifferentiation of the already differentiated root tissues. Dedifferentiation involves a complex but finely tuned, coherent host, and symbiont crosstalk that ultimately leads to the development of an extraordinary new organ called "root nodule." This developmental process is regulated by innumerable interconnected transcriptional networks. Medicago generates a cylindrical-shaped indeterminant nodule where a developmental gradient persists from the nodule tip to the base. Both plant and symbiont undergo a huge transcriptomic change as they adjusted to the mutualistic lifestyle. Transcriptome analysis on nodule developmental time series, mutant nodules, laser-capture microdissection (LCM) of separate nodule zones, and single-cell transcriptomics of root hair cells allowed researchers to understand the complex developmental circuit during nodule development at a spatio-temporal resolution. In this chapter, we will focus on the advancement in the transcriptomics study that leads to the understanding of Medicago nodule development in finer details.

7.1 Introduction

Medicago truncatula generates the indeterminate type of nodules that retain a persistent apical meristem. Mutualistic interaction of Medicago with its symbiotic partner *Ensifer meliloti* leads to the establishment of nitrogen-fixing root nodules. Nodule development starts in Medicago with the recognition of the rhizobial Nod Factor (NF) in the plant epidermis. Almost instantly with the application of the NF, calcium oscillation has been observed in the root hair cell nuclei. This calcium oscillation leads to the activation of the downstream signaling pathway (see Chap. 6). Within 24-h post-inoculation (hpi), transcriptomic changes have been observed in the root epidermis and cortex (Jardinaud et al. 2016). Root nodule development (RNS) is a complex process. It is characterized by several cellular, morphological, and physiological events. The precise transcriptomic analysis leads to the understanding of the nodule development in finer details. The transcriptomics era started with the EST sequencing and microarray analysis in 1998 and geared up with the advancement of the high throughput next-generation sequencing approaches and release of the Medicago genome sequences (Covitz et al. 1998; Benedito et al. 2008; Pecrix et al. 2018). In this chapter, we will highlight the transcriptomic studies and how these studies generated a system biology level understanding of the Medicago nodule development.

A. Bhardwaj · S. Sinharoy (✉)
National Institute of Plant Genome Research,
New Delhi 110067, India
e-mail: ssinharoy@nipgr.ac.in

7.2 EST Sequencing and Microarray – First Generation of Nodule Transcriptomics

EST sequencing was the first high throughput approach for the identification of genes and facilitate functional genomic studies during RNS. As root hairs are the primary site of rhizobial recognition and entry, the first EST sequencing was conducted to detect root hair cells-enriched cDNA. In this study, root hair cells were mechanically harvested from the whole root. This study is the first example of EST sequencing as well as single-cell transcriptomics in Medicago. Sequencing of the root hair-enriched cDNA library produced ∼890 ESTs (Covitz et al. 1998). EST sequencing sets a platform for the development of other genomic tools such as DNA microarray.

The next major upliftment in this area was the design of custom-made oligonucleotides probe-based chips in collaboration with Affymetrix (http://www.affymetrix.com). This Affymetrix GeneChip contains ∼10,000 plant probe sets and the probes from the complete *Ensifer meliloti* genome. The first use of Affymetrix GeneChip identified differential expression of 584 Medicago and 1288 *Ensifer* genes (Barnett et al. 2004). This study was followed by another microarray-based transcriptome profiling of wild type and three "*defective in nitrogen fixation*" (*dnf*) series mutants (*dnf1*, *dnf2* and *dnf7*) to understand the global transcriptional changes between wild type and these mutants. Noteworthy, *dnf* mutants generate non-nitrogen fixing nodules (Starker et al. 2006).

This Affymetrix GeneChip array was further used to develop a centralized platform with a web server called "The *Medicago truncatula* gene expression atlas" (MtGEA) described in Chap. 1. Parallel to the development of Affymetrix microarray, a second oligonucleotide probe-based microarray was designed with ∼6000 probes identified by a separate EST sequencing study from uninfected roots, mycorrhizal roots, and young root nodules named as

root interaction transcriptome (Mt6k-RIT) (Küster et al. 2004). This array was further upgraded in collaboration with Samuel Roberts Noble Foundation to 16086–70 mer oligonucleotides-based Medicago Genome Oligo Set 1.0 (Mt16kOLI) microarrays and subsequently upgraded to Mt16kOLI1Plus by adding 384 new genes, primarily transcription factors (Tellström et al. 2007; Moreau et al. 2011). The transcriptomic resources have seen continuous development between 1998–2010. In between this period, few microarray chips were also designed independently for root nodule developmental studies (see Table 7.1). During this period, several microarray-based studies enriched our knowledge about RNS. As mentioned earlier, many of these microarray chips contain plant and bacterial genes, in this section, we are only highlighting the results obtained from the plant part. The bacterial gene expression was summarized in Sect. 7.5.

A path-breaking study conducted by Eva Kondorosi's group (Maunoury et al. 2010) on wild type and different non-functional (fix−) nodules developed by plant and bacterial mutant's demarcated two waves of transcriptional switches. The first wave of transcriptional programming is required during nodule organogenesis. During this phase, mainly, transient activation of cell cycle and protein synthesis takeplace. The second wave of transcriptional programming is required during bacteroids differentiation. At this stage, mainly, the genes belonging to the secretory pathway, transmembrane, secretory proteins, or peptides were induced. The same studies compared the transcriptome of wild type and mutant symbiotic partners and found that different symbiotic mutants are showing three different transcriptomic signatures- (i) plants or bacterial mutants that are devoid of infection or contain only ITs show root-like transcriptomic signature, (ii) mutants where plant cells were differentiated and infected but the bacteroids did not differentiate passed the first transcriptome switch but not the second one, (iii) mutants nodules where both plant cell and bacteroids were fully differentiated

Table 7.1 Microarray-based transcriptomics approaches conducted to study *Medicago truncatula-Ensifer meliloti* symbiosis transcriptional reprogramming

Microarray developed	Technology	Probes	Year	References
Affymetrix GeneChip	Photolithographic masking oligonucleotide microarray	Complete *Ensifer meliloti* genome, 10 k from plant	2004, 2015	Barnett et al. (2004), Lang and Long (2015)
Affymetrix GeneChip	Photolithographic masking oligonucleotide microarray	9935 from plant	2004, 2006	Mitra et al. (2004), Starker et al. (2006)
Affymetrix GeneChip	*Mediacgo* Genome array (MtGEA v1) and functional genomics annotation to genes (MtGEA v2)	50,900 from plant	2008, 2009	Benedito et al. (2008), He et al. (2009)
Unigene microarray	cDNA microarray	2366 from plant	2010	Maunoury et al. (2010)
Mt6k-RIT	cDNA microarray (accession number A-MEXP-80; http://www.ebi.ac.uk/arrayexpress)	6359 from plant	2004	Küster et al. (2004)
Mt16kOLI1Plus, Mt16KPlus	Oligonucleotide microarray (accession number A-MEXP-138; http://www.ebi.ac.uk/arrayexpress)	16,470 from plant	2007,2011	Tellström et al. (2007), Moreau et al. (2011)

but are non-functional passed both transcriptome switches similar to the wild type nodule. These transcriptomic studies postulated that nodule development goes through different developmental phases. Interestingly, few transcription factors are involved in almost every developmental phase and they govern different transcriptional hubs (see Chap. 6).

7.3 Medicago Genome Versions and Next Generation Sequencing-Based Transcriptomes

Publication of Medicago whole-genome sequence provided a platform for a better and comprehensive understanding of nodule development. To date, four *M. truncatula* genome sequences and corresponding gene annotations have been released. The effort of the International Medicago Genome Annotation Group (IMGAG) led to the first release of the Medicago draft genome (Mt3.5.1) (Young et al. 2011). This sequencing was conducted based on BAC assembly and Illumina shotgun sequence, producing a genome sequence with ∼94% of coverage. IMGAG further published an improved version of the Medicago genome (Mt4.0). This version was generated by de novo whole-genome shotgun assembly using Illumina and 454 reads. The scaffolds generated by the de novo whole-genome shotgun method was anchored onto the previously generated pseudomolecules (Tang et al. 2014). In a parallel timeframe, a laser dissection RNA-Seq experiment was conducted on nodule tissue (see Sect. 7.4.2). While conducting this study, the group developed a genome version which was similar to Mt4.0 termed as *Mt20120830-LIPM*. This version was created by the Laboratory of Plant-Microbe Interactions (LIPM) Toulouse, INRA, France, based on Mt3.5.1 release (Young et al. 2011) and combining other available *M. truncatula* genome

sequencing data (Roux et al. 2014). The latest and most advanced version of the *M. truncatula* genome is published by LIPM, and this version was created based on PacBio sequencing. The long PacBio reads led to a substantially improved *M. truncatula* genome (Pecrix et al. 2018). The most interesting feature of Mt5.0 is the annotation of the long non-coding RNAs (lncRNAs).

Genome sequencing revealed many interesting features such as the Medicago genome consisting of more synteny blocks compared to other legumes. Synteny blocks are conserved regions within two sets of chromosomes that depict the ancestry. The phylogenomic analyzes suggest that the Medicago genome has undergone whole-genome duplication (WGD) ~ 58 million years ago (MYA). RNA sequencing revealed 963 WGD-derived gene pairs in the *M. truncatula* genome. Among these duplicated pairs, many got recruited in root nodule symbiosis. For example, nod factor receptor *Nod Factor Perception* (*NFP*) and a transcription factor *ERF required for nodulation 1* (*ERN1*) both have paralogues that participate in mycorrhizal signaling as well. These phylogenomic analyzes based on transcriptomic data and genome sequences suggest sub- and/or neo-functionalization of Medicago genes for nodulation (Young et al. 2011).

7.4 High-Resolution Transcriptomics Studies and Understanding of the Root Nodule Symbiosis

Transcriptomics approaches with multicellular tissues and organs like nodules or whole roots possessing heterogeneous cells lead to dilution of the transcriptome. Transcriptome dilutions restrict the identification of both low-abundant transcripts and moderate-to-high abundant transcripts when its expression is restricted to a single-cell layer. Homeotic regulators, transcription factors, and chromatin regulators are the control hubs and usually express transiently,

hence transcriptome coupled with cell type-specific expression or based on laser microdissection (LCM) has the potential to provide much accuracy.

7.4.1 Isolated Root Hair Cell-Based Transcriptomics

Root hair cells are the first contact and entry point of rhizobia. Rhizobia recognition leads to activation of a chain of events in the root hair cells, a. curling, b. entrapment of rhizobia and micro-colony formation, and c. infection thread (IT) formation. Mechanical dissection and isolation of the root hair cells are not complicated to achieve. The first root hair-specific single-cell transcriptomics were attempted in 1998 using a microarray chip (Covitz et al. 1998). More recently after the publication of the Medicago genome sequence, microarray and NGS-based root hair transcriptomics were conducted in wild type and transcription factor mutants (Breakspear et al. 2014). The root hairs (single-cell outgrowth from the epidermis) were mechanically isolated before and after infection with *Ensifer meliloti,* and in some cases, the same strain was defective in nod factor synthesis. Profiling of root hair cells after *Ensifer* inoculation gave a comprehensive understanding of genetic events during early stages of rhizobial infection (Nod factor perception particularly), termed as "Infectome." The major finding from this transcriptomic analysis is that phytohormones, primarily auxin, and cytokinin (CK) regulates the transcriptional reprogramming of epidermal cells. Auxin promotes rhizobial infection. An *Auxin Response Factor* (*ARF16a*) was found to be essential for IT progression. *Lonely guy* genes (*LOG*) get induced during rhizobial inoculation which generate active CK and this further trigger CK signaling pathway by response regulators and two-component system (TCS) (Breakspear et al. 2014). ABCG family of the transporter (*ABCG56*) gets activated in the root hair cells
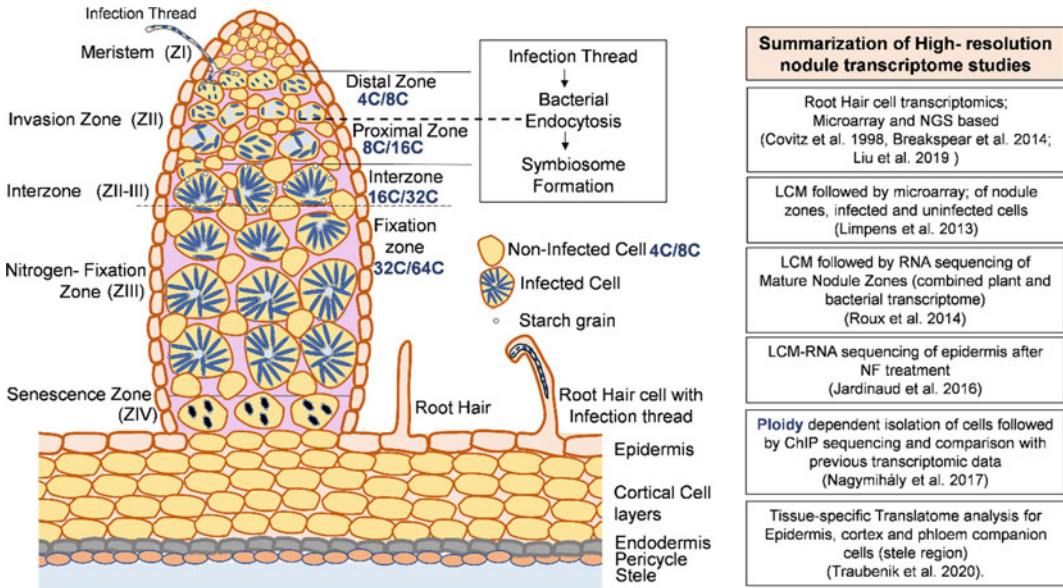
Fig. 7.1 Diagrammatic representation of the mature indeterminate Medicago nodule with the summarization of the single-cell or tissue-specific transcriptome and translatome studies

and this transporter plays a pivotal role in active CK transport during symbiosis (Jarzyniak et al. 2021) (see Chap. 6, Fig. 6.2).

Three transcription factors (TFs) activated immediately after rhizobial infection are *Nodule INception* (*NIN*), *Nuclear Factor Y1* (*NF-YA1*), and *ERF required for Nodulation 1* (*ERN1*). Among them, NIN is the master regulator of infection and organogenesis. NIN directly activates *NF-YA1* expression. Further, NIN and ERN1 signaling pathways are intertwined (see Chap. 6). Root hair cell transcriptomic analysis had been conducted on *nin-1, ern1* and *nf-ya1* mutants after inoculating with wild type and non-NF-producing rhizobium. This study highlighted that among 1124 genes upregulated after rhizobium inoculation in root hair cells, 43% get down-regulated in *nin-1* root hair cells, whereas only 9.1% are down-regulated in the *ern1*. Hence, established NIN as a central regulator of root hair transcriptome. This study also identified the putative direct common targets of NIN, NF-YA1, and ERN1 (Liu et al. 2019). Thus, these transcriptomic studies drew a broader image of rhizobial infection through root hair cells.

7.4.2 Laser-Capture Microdissection (LCM) coupled with Microarray and RNA-Sequencing

The second set of high-resolution transcriptome data came from laser-capture microdissection (LCM) of different nodule zones coupled with microarray/RNA sequencing. In one attempt nodule zones, infected and uninfected cells were isolated using LCM (Fig. 7.1). A special development gradient exists from the top to the bottom across the cylindrical indeterminate nodule of Medicago. The nitrogen fixation zone of Medicago nodules contains highly endoreduplicated, bigger, rhizobium-containing cells, and less endoreduplicated, smaller cells that are devoid of rhizobia. Nodule zones and cells present in each zone are at different developmental phases. Hence, it is expected that different nodule zones should have different transcriptomic signatures (Limpens et al. 2013). The first zone-specific transcriptomic analysis was conducted using Affymetrix Medicago GeneChips containing 50, 900 probe sets. This analysis presented cell type-specific gene enrichment where infected and

non-infected cells were separated and further dissected, revealing transcriptional dynamics from infected to non-infected cells, meristem to infection zone, and distal to proximal infection zone (Limpens et al. 2013).

LCM of the mature nodule zone followed by RNA-Seq was conducted (Roux et al. 2014) to characterize gene expression across Medicago indeterminate nodule zones (Vasse et al. 1990). They used five nodule zones representing (i) the apical meristem region with small cells (ZI); (ii) distal invasion zone, where endocytosis of rhizobia takes place (ZIIa); (iii) proximal invasion zone, where differentiation of bacteroids takes place (ZIIb); (iv) interzone II–III, where major starch accumulation takes place; and (v) nitrogen fixation zone ZIII (Fig. 7.1). This sequencing data were compiled in a database, SYMbiMICS (https://iant.toulouse.inra.fr/symbimics/), which is a vital platform for exploring gene expression and understanding its role during Medicago nodule development (Roux et al. 2014).

Another LCM-RNA sequencing analysis was conducted after 4 and 24 h post Nod Factors (NFs) treatment followed by LCM of root regions responsive to NF coupled to RNA sequencing (Jardinaud et al. 2016). This study identified around 1,070 genes being regulated by NF. Further, approximately 300 genes showed 10-folds upregulation after NF treatment. Among them, 44 genes were symbiosis specific. These 44 genes include *NCRs* that are common to mycorrhizal and nodulated roots. This approach identified two important receptor kinases, *SYM-BIOSIS LEUCINE-RICH RECEPTOR KINASE (SYMRK)* and *LYSINE MOTIF RECEPTOR KINASE (LYK3)*. These two were highly upregulated in the epidermis. Additionally, receptor kinase *LYK10* is also expressed in epidermal cells. Hormone signaling-related genes such as CK and strigolactone pathway genes and flavonoids biosynthesis genes show NF-induced expression. This study identified strong but transient activation of genes after NF treatment in the root epidermis. In summary, epidermal regulation of NF responsive genes is highly complex (Jardinaud et al. 2016).

7.4.3 Ploidy Dependent Isolation of Nodule Cells and Transcriptional Dynamics

The third high-resolution study was conducted based on the ploidy of nodule cells (Nagymihály et al. 2017). As the dividing cells leave the meristem, they start endoreduplication. In the distal invasion zone, cells are usually present in 4C and get infected with rhizobium. In the invasion zone, the bacteroid (rhizobium inside the plant-bound membrane) also gets endoreduplicated. During this stage, the plant cell nucleus keeps on endoreduplicating and reaches up to 32C in the interzone. Further, the ploidy level of these cells is 64C in the nitrogen fixation zone. Infected nodule cells were sorted based on flow cytometry and cells with different ploidy (4C, 8C, 16C, and 32C) were isolated. Chromatin immunoprecipitation (ChIP) was conducted on the DNA that was isolated from these cells against H3K27me3 and/or anti-H3K9ac histone marks. The ChIP data were further compared with the LCM nodule zone-specific RNA-Seq-based transcriptomics data. This study revealed that methylation and histone modification mediate epigenetic control of nodule-specific genes like *NCRs* and *ENOD12*. The increasing ploidy levels alter chromatin accessibility thus gene expression is affected. Symbiotic cells undergo successive rounds of endoreduplication, parallel to the terminal differentiation of endosymbionts. Expression of those genes responsible for terminal bacteroid differentiation (TBD) such as *NCRs* (see Chap. 6) were found to be epigenetically regulated (Nagymihály et al. 2017). Hence, this study opened the avenue of research in the area of chromatin dynamics during nodule development, a largely unknown area.

7.4.4 Translating Ribosome Affinity Purification

Another technique for examining cell-specific expression is translating ribosome affinity purification (TRAP). In this technique, ribosomal proteins tagged with an epitope (usually FLAG)

are expressed in plant tissues. Immunopurification of polysomes (i.e., transcripts bound to two or more ribosomes) followed by RNA-Seq estimates translational control of gene expression (translatome). The first Medicago nodule translatome was conducted by Traubenik's group (Traubenik et al. 2020). To understand the translational control of gene expression, *p35S:: FLAG-RPL18* was transformed into Medicago followed by both transcriptome and translatome analysis. Comparison between transcriptome and translatome highlighted ∼ 65% of the differentially upregulated genes identified by transcriptome, which did not show enrichment in nodule translatome. Underlining, transcriptional and translational responses are partially uncoupled during nodule development. In the same study, TRAP has been combined with Medicago epidermal cell-specific *EXPANSIN7* (*MtEXP7*), Arabidopsis cortex specific *CORTEX SPECIFIC TRANSCRIPT (AtCO2)*, and phloem companion cell-specific *SUCROSE TRANS- PORTER 2 (AtSUC2)* promoters to understand tissue-specific gene expression during Medicago nodule development. These tissue-specific ribosome-pulldown experiments have been designed specially to understand the early recognition and nodule primordium initiation. The analysis highlighted the high accumulation of *DMI2*-mRNA in the polysomes during rhizobial infection, suggesting a large amount of DMI2 protein is getting synthesized during infection. Some transcripts show cell-specific abundance such as *ERN1*, *NF-YA1,* and *NPL* show more enrichment in epidermal cells compared to cortical cells (Reynoso et al. 2013; Pan et al. 2018; Traubenik et al. 2020). Cell type-specific translatome analysis holds high potential and has proven to be a method of choice in animal models or Arabidopsis as TRAP can be performed on intact tissues as opposed to LCM or FACS mediated tissue isolation. Hence, a high amount of mRNA can be obtained without cumbersome methodology or techniques that can change gene expression. Cell/tissue-specific TRAP is going to be a method of choice to answer many unanswered questions in nodule biology.

Medicago nodule possess a) meristematic zone (ZI) containing mitotically active cells; b) invasion zone (ZII) where bacterial endocytosis takes place and is subdivided into two subzones vis distal (ploidy level 4C/8C) and proximal (ploidy level 4C/8C) invasion zone; c) Interzone (ploidy level 16C/32C), a zone distinctly visualized due to accumulation of starch; d) nitrogen fixation zone (ZIII) containing a mixture of infected cells (ploidy level 32C/64C) and uninfected cells (ploidy level 4C/8C) where enzymatic fixation of di-nitrogen takes place; and e) senescence zone (ZIV) presents only in older nodules. Based on the nodule developmental gradient, several tissue-specific transcriptomic studies and ploidy-based ChIP sequencing have been conducted. Infected root hair cell transcriptomic study was conducted in wild type and three transcription factor mutants (Breakspear et al. 2014; Liu et al. 2019), and epidermal cells after NF treatment were dissected by LCM (Jardinaud et al. 2016) to understand the early stages of plant-rhizobia recognition. A translatome study was conducted using different tissue-specific promoters *EXPANSION7*, *COR- TEX SPECIFIC TRANSCRIPT,* and *SUCROSE TRANSPORTER 2,* respectively, for epidermis, cortex, and phloem-companion cells (Fig. 7.1) (Traubenik et al. 2020).

7.5 Understanding from the Bacterial Part

In root nodule symbiosis, symbiotic bacteria play the solo part in the enzymatic fixation of the gaseous di-nitrogen. Host cell physioxia plays a pivotal role in the onset of nitrogen fixation. In a mature indeterminate nodule, leghemoglobin accumulates only in the nitrogen fixation zone. Accumulation of leghemoglobin leads to the hypoxic condition in the nodule. Hypoxia induces bacterial nitrogen fixation operon (nif and fix genes), thus promoting enzymatic fixation of nitrogen. Like the plant, sequencing of the symbiont genome accelerated the SNF research. *Ensifer meliloti* whole-genome consists of a chromosome (3.65 Mb) and two megaplasmids

(pSymA and pSymB). This genome sequence has been assembled in 2001 (Galibert et al. 2001). The bacterial genome sequence leads to the incorporation of *S. meliloti* specific probe sets in the Affymetrix-based Medicago GeneChip. In the gene chip, a total of 9,935 probe sets are present that correspond to the Ensifer genome. Hence, whenever a microarray has been attempted using Affymetrix GeneChip, consolidated analysis of plant and/or bacterial gene expression can be obtained depending on the sample preparation. Bacteroid transcriptome analysis highlighted that transcription factors that activate nod operon (nodD), NF biosynthesis genes, transporter genes, such as sitA, exopolysaccharide abundance regulatory genes syrM and syrA, adenylate cyclases genes (cyaE, cyaF1, cyaF6), genes responsible for nitrogen fixation (nifA, nifH nifN, fixK1, fixN1, fixL etc.), and phenylacetic acid catabolism genes paaABCD, were highly induced inside nodules. Additionally, nodule transcriptomic studies revealed that ∼33% of symbiosis-associated genes were located on the pSymA plasmid (Barnett et al. 2004).

Nodule transcriptome study was conducted using bacterial mutants defective in different stages of nodule development keeping the plant background as wild type to understand how plant gene expression is coupled with the bacterial gene expression. Bacterial *exoY* mutant that is defective in succinoglycan production cannot initiate successful IT formation (Cheng and Walker 1998) and produces small nodule bumps. Transcriptomics of these nodules resemble plant roots. Further, inoculation of *bacA* mutant forms small white nodules where bacterial endocytosis is normal, but the bacteroids are small and undifferentiated. The transcriptome signature of this mutant resembles partly with root but shows the first wave of nodule-specific gene expression (see Sect. 7.2). Further, *E. meliloti* mutants (such as *nifH*, *nifA*, *fixG*, *fixJ*, and *fixK*) where nodule organogenesis per se gets completed but cannot start the fixation of atmospheric di-nitrogen in spite of showing signature of complete nodule organogenesis. This suggests enzymatic fixation of nitrogen in the root nodule happens after the

completion of the nodule development (Maunoury et al. 2010).

The advancement in the understanding of bacterial gene expression came from the LCM study that has already been described above (Sect. 7.4.2). The LCM coupled with RNA-Seq analysis was conducted for both plant and bacterial genes (Roux et al. 2014). In the LCM analysis, 7799 bacterial genes were detected to be expressed inside the nodule. These genes were categorized in 13 different clusters as per their expression patterns. The integrated analysis showed fine coordination between plant and bacterial gene expression profiles, for example, bacterial genes for cell division (ftsK, minCDE), DNA replication genes (repC1, repC2), and genes controlling cell cycles (ctrA, divJ) were highly expressed in proximal and distal invasion zone where host and bacterial cells were continuing their development, whereas these genes show negligible expression in the nitrogen fixation zone. The LCM analysis also raised many unanswered questions. The master regulator of the flagellar operon (visN) shows expression from the interzone, and its expression is the highest in the nitrogen fixation zone, indicating the whole operon is functional in this zone. In the invasion and nitrogen fixation zone, the rhizobium is present inside the plant-derived membrane envelope (Roux et al. 2014). What is the function of flagella at this stage? Why does not the pathogen-associated molecular patterns (PAMPs) trigger immunity (PTI) during this stage (See Chap. 6)? There are definitely more questions than answers.

7.6 Transcriptomic Studies Unfold Many Unanswered Questions— The Future Direction of Nitrogen Fixation Research

Dedifferentiation of already differentiated root cells was the first evolutionary innovation that nodule-forming plants must have learned. Hence, it can be visualized easily that chromatin remodeling and change in the epigenetic mark were the first few steps that would have happened inside the plant nucleus concomitantly with the perception

of the NF. TFs might have played a vital role just after the opening/remodeling of chromatin. Nodule-specific upregulation of TFs would have happened either (a) under NF signaling or (b) under the regulation of hormones (Breakspear et al. 2014; Nagymihály et al. 2017). Further, the most recent version of the Medicago genome along with the LCM mediated transcriptome analysis highlighted that small and long non-coding RNAs (lncRNAs) are expressed specifically to different nodule zones (Pecrix et al. 2018). lncRNA is antisense non-coding transcripts. lncRNA-mediated gene regulation has been broadly studied in mammalian cells. They can play multiple roles such as a enhancer of transcription by interacting with the chromatin (recruits chromatin remodeling complex to activate transcription); (b) decoy to sequester transcription factors (sequester proteins with regulatory function); (c) as guide RNA (carries protein molecules to their target genes); and (d) modulator of transcriptional expression (acts as a scaffold to recruit the molecular complex to target genes) (Batley et al. 2020). Only a limited number of long non-coding RNAs have been functionally characterized in plants, especially in legumes. The recent release of the Medicago genome has annotated thousands of long non-coding RNAs showing positive or negative correlations with neighboring mRNAs and being induced in the nodule. Such profiles have been seen for many vital genes that have already been genetically characterized for their roles during nodule development such as *NSP2, IPD3, NIN, SymCRK, ERN2, DME, DNF1,* and *NCRs* (for the role of these genes see Chap. 6). The positively correlated lncRNA-mRNA pairs are *DME, NSP2, NIN, IPD3, RSD, SYMREM1,* and *SymCRK*. Interestingl,y sense and antisense transcripts of *EFD* show different expression profiles in nodule zones (Pecrix et al. 2018). Sense transcript of *EFD* shows maximum expression in early nodule stages and at distal and proximal invasion zone, whereas the antisense transcript expresses maximum in the interzone and fixation zone in mature nodules. It hints that *lncRNAs* might regulate these symbiosis-related genes' expression. In Medicago, *ENOD40* (*Early Nodulin 40*) was the first characterized lncRNA which binds to RNA-Binding protein (RBP) and promotes cytoplasmic relocalization of this protein during nodule formation. Thus, *lncRNA*-mediated regulation of nodule development needs further attention (Campalans et al. 2004; Mergaert et al. 2020).

7.7 Conclusion

Gradual advances from EST sequencing to next-generation sequencing built up a platform for the Medicago genome and the nodule-specific gene expression profile. Further, transcriptomics coupled with functional genomics studies have drawn a skeleton structure of the molecular mechanism of RNS. Over two decades of research also highlighted the complexity of the molecular interaction that guides this marvelous organ development. Recent advances in the next generation sequencing (NGS) technology combined with cell/tissue-specific transcriptomic, ChIP, and translatome analysis highlight the hidden complexity that guides nodule development. These new techniques hold enormous potential and the growing information that came out from these analyzes will help to create a superior understanding of RNS.

References

Barnett MJ, Toman CJ, Fisher RF, Long SR (2004) A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. Proc Natl Acad Sci U S A 101:16636–16641. https://doi.org/10.1073/pnas.0407269101

Batley J, Willmann MR, Budak H, et al (2020) Long Non-coding RNA in Plants in the Era of Reference Sequences. https://doi.org/10.3389/fpls.2020.00276

Benedito VA, Torres-Jerez I, Murray JD, et al (2008) A gene expression atlas of the model legume Medicago truncatula. Plant J 55:504–513. https://doi.org/10.1111/j.1365-313X.2008.03519.x

Breakspear A, Liu C, Roy S, et al (2014) The root hair "infectome" of medicago truncatula uncovers changes in cell cycle genes and reveals a requirement for auxin signaling in rhizobial infectionw. Plant Cell 26:4680–4701. https://doi.org/10.1105/tpc.114.133496

Campalans A, Kondorosi A, Crespi M (2004) Enod40, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in Medicago truncatula. Plant Cell 16:1047–1059. https://doi.org/10.1105/tpc.019406

Cheng HP, Walker GC (1998) 1998-Succinoglycan Is Required for Initiation and Elongation of Infection Threads.pdf. J Bacteriol 180:5183–5191

Covitz PA, Smith LS, Long SR (1998) Expressed sequence tags from a root-hair-enriched Medicago truncatula cDNA library. Plant Physiol 117:1325–1332. https://doi.org/10.1104/pp.117.4.1325

Galibert F, Finan TM, Long SR, et al (2001) The composite genome of the legume symbiont Sinorhizobium meliloti. Science (80) 293:668–672. https://doi.org/10.1126/science.1060966

He J, Benedito VA, Wang M, et al (2009) The Medicago truncatula gene expression atlas web server. BMC Bioinformatics 10. https://doi.org/10.1186/1471-2105-10-441

Jardinaud MF, Boivin S, Rodde N, et al (2016) A laser dissection-RNAseq analysis highlights the activation of cytokinin pathways by nod factors in the Medicago truncatula root epidermis. Plant Physiol 171:2256–2276. https://doi.org/10.1104/pp.16.00711

Jarzyniak K, Banasiak J, Jamruszka T, et al (2021) Early stages of legume–rhizobia symbiosis are controlled by ABCG-mediated transport of active cytokinins. Nat Plants 7:428–436. https://doi.org/10.1038/s41477-021-00873-6

Küster H, Hohnjec N, Krajinski F, et al (2004) Construction and validation of cDNA-based Mt6k-RIT macro- and microarrays to explore root endosymbioses in the model legume Medicago truncatula. J Biotechnol 108:95–113. https://doi.org/10.1016/j.jbiotec.2003.11.011

Lang C, Long SR (2015) Transcriptomic analysis of Sinorhizobium meliloti and Medicago truncatula symbiosis using nitrogen fixation-deficient nodules. Mol Plant-Microbe Interact 28:856–868. https://doi.org/10.1094/MPMI-12-14-0407-R

Limpens E, Moling S, Hooiveld G, et al (2013) Cell- and Tissue-Specific Transcriptome Analyses of Medicago truncatula Root Nodules. PLoS One 8:. https://doi.org/10.1371/journal.pone.0064377

Liu CW, Breakspear A, Guan D, et al (2019) NIN acts as a network hub controlling a growth module required for rhizobial infection. Plant Physiol 179:1704–1722. https://doi.org/10.1104/pp.18.01572

Maunoury N, Redondo-Nieto M, Bourcy M, et al (2010) Differentiation of symbiotic cells and endosymbionts in Medicago truncatula nodulation are coupled to two transcriptome-switches. PLoS One 5: https://doi.org/10.1371/journal.pone.0009519

Mergaert P, Kereszt A, Kondorosi E (2020) Gene Expression in Nitrogen-Fixing Symbiotic Nodule Cells in Medicago truncatula and Other Nodulating Plants. Plant Cell 32:42–68. https://doi.org/10.1105/tpc.19.00494

Mitra RM, Shaw SL, Long SR (2004) Six nonnodulating plant mutants defective for Nod factor-induced transcriptional changes associated with the legume-rhizobia symbiosis. Proc Natl Acad Sci U S A 101:10217–10222. https://doi.org/10.1073/pnas.0402186101

Moreau S, Verdenaud M, Ott T, et al (2011) Transcription reprogramming during root nodule development in Medicago truncatula. PLoS One 6:. https://doi.org/10.1371/JOURNAL.PONE.0016463

Nagymihály M, Veluchamy A, Györgypál Z, et al (2017) Ploidy-dependent changes in the epigenome of symbiotic cells correlate with specific patterns of gene expression. Proc Natl Acad Sci U S A 114:4543–4548. https://doi.org/10.1073/pnas.1704211114

Pan H, Stonoha-Arther C, Wang D (2018) Medicago plants control nodulation by regulating proteolysis of the receptor-like kinase DMI2. Plant Physiol 177:792–802. https://doi.org/10.1104/pp.17.01542

Pecrix Y, Staton SE, Sallet E, et al (2018) Whole-genome landscape of Medicago truncatula symbiotic genes. Nat Plants 4:1017–1025. https://doi.org/10.1038/s41477-018-0286-7

Reynoso MA, Blanco FA, Bailey-Serres J, et al (2013) Selective recruitment of mRNAs and miRNAs to polyribosomes in response to rhizobia infection in Medicago truncatula. Plant J 73:289–301. https://doi.org/10.1111/tpj.12033

Roux B, Rodde N, Jardinaud MF, et al (2014) An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. Plant J 77:817–837. https://doi.org/10.1111/tpj.12442

Starker CG, Parra-Colmenares AL, Smith L, et al (2006) Nitrogen fixation mutants of Medicago truncatula fail to support plant and bacterial symbiotic gene expression. Plant Physiol 140:671–680. https://doi.org/10.1104/pp.105.072132

Tang H, Krishnakumar V, Bidwell S, et al (2014) An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC Genomics 15:1–14. https://doi.org/10.1186/1471-2164-15-312

Tellström V, Usadel B, Thimm O, et al (2007) The lipopolysaccharide of Sinorhizobium meliloti suppresses defense-associated gene expression in cell cultures of the host plant Medicago truncatula. Plant Physiol 143:825–837. https://doi.org/10.1104/pp.106.090985

Traubenik S, Reynoso MA, Hobecker K, et al (2020) Reprogramming of root cells during nitrogen-fixing symbiosis involves dynamic polysome association of coding and noncoding RNAs. Plant Cell 32:352–373. https://doi.org/10.1105/tpc.19.00647

Vasse J, de Billy F, Camut S, Truchet G (1990) Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J Bacteriol 172:4295–4306

Young ND, Debellé F, Oldroyd GED, et al (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480:520–524. https://doi.org/10.1038/nature10625