



Article

Expression and Variation of the Genes Involved in Rhizobium Nodulation in Red Clover

Randy D. Dinkins ^{1,*} , Julie A. Hancock ², Derek M. Bickhart ³, Michael L. Sullivan ³  and Hongyan Zhu ⁴

¹ Forage-Animal Production Research Unit, USDA-ARS, Lexington, KY 40506, USA

² College of Agriculture, Food and Environment, University of Kentucky, Lexington, KY 40508, USA

³ US Dairy Forage Research Center, USDA-ARS, Madison, WI 53706, USA

⁴ Department of Plant and Soil Sciences, University of Kentucky, Lexington, KY 40546, USA

* Correspondence: randy.dinkins@usda.gov

Abstract: Red clover (*Trifolium pratense* L.) is an important forage crop and serves as a major contributor of nitrogen input in pasture settings because of its ability to fix atmospheric nitrogen. During the legume-rhizobial symbiosis, the host plant undergoes a large number of gene expression changes, leading to development of root nodules that house the rhizobium bacteria as they are converted into nitrogen-fixing bacteroids. Many of the genes involved in symbiosis are conserved across legume species, while others are species-specific with little or no homology across species and likely regulate the specific plant genotype/symbiont strain interactions. Red clover has not been widely used for studying symbiotic nitrogen fixation, primarily due to its outcrossing nature, making genetic analysis rather complicated. With the addition of recent annotated genomic resources and use of RNA-seq tools, we annotated and characterized a number of genes that are expressed only in nodule forming roots. These genes include those encoding nodule-specific cysteine rich peptides (NCRs) and nodule-specific Polycystin-1, Lipoxygenase, Alpha toxic (PLAT) domain proteins (NPDs). Our results show that red clover encodes one of the highest number of NCRs and ATS3-like/NPDs, which are postulated to increase nitrogen fixation efficiency, in the Inverted-Repeat Lacking Clade (IRLC) of legumes. Knowledge of the variation and expression of these genes in red clover will provide more insights into the function of these genes in regulating legume-rhizobial symbiosis and aid in breeding of red clover genotypes with increased nitrogen fixation efficiency.

Keywords: forage legumes; nitrogen fixation; gene expression; nodulation; nodule-specific cysteine rich peptides



Citation: Dinkins, R.D.; Hancock, J.A.; Bickhart, D.M.; Sullivan, M.L.; Zhu, H. Expression and Variation of the Genes Involved in Rhizobium Nodulation in Red Clover. *Plants* **2022**, *11*, 2888. <https://doi.org/10.3390/plants11212888>

Academic Editors: Stacy D. Singer, Abdelali Hannoufa and Bill Biligetu

Received: 21 September 2022

Accepted: 25 October 2022

Published: 28 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The root nodule symbiosis between legumes and nitrogen-fixing soil bacteria is vital for the supply of biologically fixed nitrogen to soil and plants in agricultural systems. Development of nitrogen-fixing nodules involves simultaneous differentiation of both symbiotic partners. During this process, the bacteria become adjusted to an endosymbiotic lifestyle and develop into mature bacteroids that are capable of nitrogen fixation. Depending on the host, the morphology and physiology of bacteroids can be strikingly different. In *phaseoloid* legumes (e.g., soybeans and common beans), the bacteroids mostly retain the same morphology and can revert to the free-living form. In contrast, galegoid legumes (e.g., alfalfa, clovers, and peas), the bacteria often undergo terminal differentiation, which is characterized by cell enlargement coupled with genome amplification, increased membrane permeability, and loss of reproductive ability [1,2]. It was reported that terminal differentiation may lead to more efficient nitrogen fixation compared to non-terminal differentiation [3]. The terminal differentiation of bacteroids in *Medicago* and related legumes is associated with the expression of nodule-specific peptides that are targeted to bacteroids in the symbiosome, which include nodule-specific cysteine-rich peptides (NCRs) [4], proline-rich peptides (PRPs) [5], calmodulin-like proteins (CaMLs) [6],

glycine-rich proteins (GRPs) [7], LEED..PEED signature proteins (LPs) [8], and nodule-specific PLAT (Polycystin-1, Lipoygenase, Alpha-Toxin) domain proteins (NPDs) [9,10]. Some of these peptides are essential for development of high-efficiency nitrogen-fixing nodules [4,11], while others are toxic to specific bacterial strains and thus lead to the development of nodules incapable of fixing nitrogen [11,12]. Some of these peptides appear to be species-specific, such as the LPs that have only been identified in *Medicago* to date [8], whereas others appear to be widespread yet have very little sequence homology between species and highly differentially expanded in the different species [13]. An example of the latter are the NCR peptides that are only present in inverted-repeat lacking chloroplast (IRLC) galeoid legumes and absent in *phaseoloid* legumes. It has been proposed that these peptides optimize nitrogen fixation performance as different legume species contain similar, but different, sets of NCR gene products that target different rhizobial species [14,15].

Clovers and other forage legumes are commonly utilized in pastures as they offer an opportunity to reduce nitrogen fertilizer use when inter-seeded with forage grasses [16]. Perennial clovers in a dense grass-clover pasture can contribute more than 200 kg nitrogen per ha/year to the cropping system [17]. Red clover is an ideal legume pasture choice for many areas of the world in that it tolerates soil acidity and poor drainage and is well suited for hay production and to grazing when inter-seeded with common forage-grasses [18,19]. Molecular work on red clover has lagged behind many other cultivated crop species despite its relatively small genome [20], primarily due to the fact that it is an outcrossing species making genetic gains more difficult [21]. Genomic resources have been developed that now allow for molecular characterization and analysis [22–26]. The high heterozygosity in red clover, while complicating the assembly of contiguous chromosome molecules, has allowed for the identification of chromosomal regions and genes involved in various functions using single nucleotide polymorphic loci (SNPs) mapping [26–30], including identification of red clover genotypes differing in their nitrogen fixation potential [31]. However, still lacking is information on the molecular response and interaction of red clover with rhizobium that has been primarily done on model legumes such as soybeans (*Glycine max*), *Lotus japonicus*, and *Medicago truncatula*. For example, it has been postulated that red clover encodes roughly 425 NCRs, suggesting that this species contains one of the higher number of these peptides [9], although there is no information on the expression of the genes encoding these peptides in red clover due to the lack of annotation of the genes in the databases. Thus, the goal of the current work was to identify, and monitor expression of, genes in the interaction of red clover plants with rhizobium by comparing gene expression in roots of inoculated with uninoculated plants.

2. Results and Discussion

In order to characterize the red clover genes associated with nodulation, analysis was done on roots where plants were inoculated (Nod+) to those that were not inoculated (Nod−). The plants for the current experiment were derived from the Kenland variety and from crosses derived from a putative *ifs1* mutant generated by CRISPR/Cas9 [32]. As the knock-down of the *TpIFS1* gene had no effect on nodulation, and gene expression analysis revealed no interaction between nodulation and the *IFS1* knockdown (see Figure 5 in Dinkins et al. [32]), comparisons of differential gene expression between the Nod+ and Nod− roots were done across both WT and *ifs1/ifs1* genotypes (Figure 1a—dendrogram). Of the 43,827 red clover gene models annotated in the ARS1 genome assembly, 32,285 genes had at least one mapped read from the root-derived mRNA. For the current study, a minimum of five reads was used as a cutoff for comparisons between the Nod+ and Nod− root samples, resulting in the root transcriptome for the current analysis of 26,128 red clover gene models. Furthermore, for the statistical analysis of differential expression based on the four replicates (individual plants) for each Nod+ and Nod− RNA sample, the criteria used for designating significant differential expression between the Nod+ and Nod− roots was greater than two-fold expression difference, and $-\log_{10} p$ -value of 2.12 (corresponding to p -value of 0.003 and based on an FDR p -value 0.1). Using the above criteria, 1474 red clover genes were differentially expressed (DEGs) between the Nod+ and

Nod[−] roots (Figure 1b—volcano plot). Of these, 512 genes were down-regulated in Nod⁺ roots and 962 genes were up-regulated in Nod⁺ roots (Supplementary Materials, Table S1). Of the 962 DEGs showing enhanced expression in Nod⁺ roots, 398 were associated with the nodule-specific cysteine rich (NCR) peptides and cysteine-rich defensin peptides (described in more detail below). The ten most highly expressed genes in the Nod⁺ root samples include those encoding four leghemoglobin Lb120-1 homologues, two nodulin-25 like proteins, an early nodulin-75-like protein, a nodulin-specific embryo expressed 3-like protein, a nodule-specific glycine-rich protein and a putative NMS32/34 nodulin-specific protein (Table 1). The high expression of these genes is similar to results of the most highly expressed genes found during nodulation in *Medicago* species (Burghardt et al., 2017; Huang et al., 2022). The expression of these genes were also some of the most highly differentially expressed genes when comparing the Nod⁺ and Nod[−] root samples (i.e., no or very low expression in Nod[−] roots), confirming that the high expression is due to the presence of the nodules in these samples. The list of Nod⁺ greater than Nod[−] DEGs is presented in the supplemental file (Supplementary Materials, Table S2).

Table 1. Highest expressing red clover genes in Nod⁺ root samples.

Feature ID	Normalized Reads (Means) Nod [−]	Normalized Reads (Means) Nod ⁺	Log2 Diff Express (Nod ⁺) / (Nod [−])	Product
LOC123882175	381.52	74,205.38	7.60	leghemoglobin_Lb120-1
LOC123882181	31.62	63,699.05	10.98	leghemoglobin_Lb120-1
LOC123916815	15.55	57,593.44	11.86	nodulin-25-like isoform X10
LOC123902448	328.64	38,271.05	6.86	early nodulin-75-like
LOC123882180	7.62	28,608.07	11.88	leghemoglobin_Lb120-1
LOC123917408	97.65	27,409.22	8.13	nodulin-26-like isoform X2
LOC123882176	17.47	24,305.17	10.44	leghemoglobin_Lb120-1
LOC123899073	4.50	22,133.74	12.26	Embryo-specific protein ATS3B-like; NPD
LOC123884281	5.08	20,011.87	11.95	ctenidin-1-like
LOC123916883	6.72	17,961.34	11.38	NMS32/34 protein, putative
LOC123915658	3049.67	13,975.22	2.20	bifunctional aspartate aminotransferase
LOC123893140	4110.65	13,859.39	1.75	legume-specific_protein
LOC123918940	4.23	13,807.55	11.67	NMS32/34 protein, putative
LOC123891069	18.14	13,225.90	9.51	leghemoglobin_Lb120-1
LOC123888452	57.99	12,010.79	7.69	early_nodulin_ENOD18
LOC123913320	2502.68	11,384.90	2.19	glutamine synthetase nodule isozyme
LOC123909356	6.27	10,611.03	10.72	basic blue protein-like
LOC123895631	2916.89	10,575.69	1.86	inactive beta-amylase 9
LOC123921269	829.79	10,351.60	3.64	chaperone protein dnaJ 8, chloroplastic
LOC123911535	4108.67	10,330.98	1.33	heavy metal-associated isoprenylated plant protein 6-like

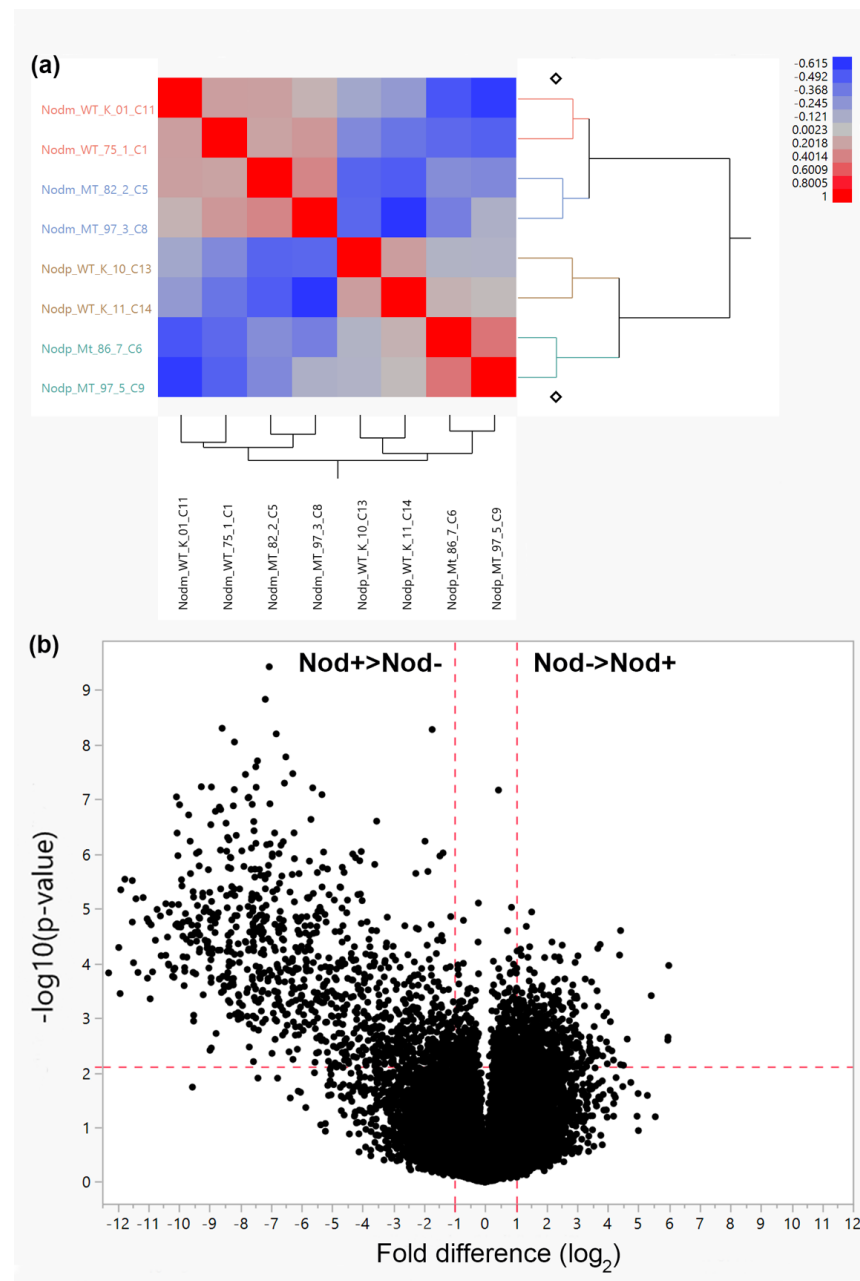


Figure 1. (a) Dendrogram and heat map of Nod+ (labeled as Nodp for the samples displayed in the figure vertical axis) and Nod− (labeled as Nodm for the samples displayed in the vertical axis) expression profiles for each individual sample and (b) volcano plot showing differential expression between Nod+ (NodP) and Nod− (NodM) root samples.

Here, we highlight an example of genes or gene families that were differentially expressed between the Nod+ and Nod− roots, with a focus on the genes that are up-regulated by nodulation.

2.1. Up-Regulated Genes by Nodulation

2.1.1. NOD25 and N22

NOD25 orthologs in *Medicago truncatula* (MtNOD25), *Medicago sativa* (MsNOD25) and *Vicia faba* (VfNOD28/32) are expressed only in root nodules and encode a nodule-specific protein designated nodulin-25. Nodulin-25 homologs consist of different repetitive modules that are highly diverged between legume species but the N-terminal signal peptide (SP) and the C-terminus are highly conserved. The N-terminal SP is required and sufficient

to target the protein to the symbiosome of infected root nodule cells but its function in symbiosis development remains unknown [33]. In red clover, we identified two *NOD25* loci on chromosome 3 in the current genome assembly but they most likely represent two allelic (haplotype) versions of the same gene. Nevertheless, there are over twenty-five different putative nodulin-25 protein isoforms predicted via alternative splicing from the two *NOD25* genes. The two *NOD25* genes had very high expression in the Nod+ roots but no expression was detected in Nod− roots. The RNA-seq data support most, although not all, of the different splicing mRNA products (Figure S1), suggesting that these might be expressed at different stages during the symbiosome development not observed in the current study, or that these predicted alternative splicing products may not occur in these plants.

The *NOD25* locus is highly syntenic in galegoid legumes. Immediately adjacent to the *NOD25* gene is another nodule-specific gene called *MtN22* in *M. truncatula* and *NMS32/34* in *M. sativa*. In red clover, the *MtN22* ortholog appears to be duplicated; there are three copies in one putative haplotype (*Tp3g123916883*, *Tp3g123916884*, and *Tp3g123916889*) and two copies (*Tp3g123918940* and *Tp3g123918941*) in another. All the five genes were expressed only in the Nod+ roots (Table 1). *Tp3g123916883*, *Tp3g123918940* and *Tp3g123918941* are postulated to produce three protein isoforms each via differential splicing (Figure S2). From the RNA-seq data, it is difficult to differentiate differential splicing due to sequence similarities that results in similar protein products (Figures S3 and S4). *MtN22* also produces three putative protein isoforms and has the highest similarity to the red clover *Tp3g123918940*. Interestingly, the only other *MtN22*-like protein that is currently predicted in other species are in *Cicer arietinum* where three genes are found. In all probability, this lack of the matches to other legumes, such as to the *Viciaea*, may be due to the lack of gene information at this time.

2.1.2. NCRs

Of the 962 DEGs up-regulated in the Nod+ roots, 393 encode the nodule-specific cysteine rich (NCR) peptides. Many of the NCR genes in the AR-100 genome assembly were annotated as non-coding RNA (ncRNA) genes. Thus, we manually re-annotated the genes based on known conserved NCR sequences and gene structure. Specifically, these genes were identified by the presence of two (or sometimes three) exons, where the first exon primarily encodes the signal peptide and the second exon codes the mature peptide containing four or six cysteine residues as described further below. Where a ncRNA locus was associated with the NCR gene, the locus number was maintained, however in instances where no locus was identified, we manually annotated according to chromosome number and start location of the gene on the chromosome (Supplementary Materials, Table S3). Expression of the NCR genes varied in the Nod+ roots, but essentially no expression was observed in the uninoculated root samples (Supplementary Materials, Table S3).

The distinguishing feature of these small peptides are the 4–6 conserved cysteine residues [15]. The four cysteine peptides were conserved C-(X4)-DC-(X11-17)-C-(X4)-C, the six cysteine peptides were conserved C-(X3-6)-C-(X4-7)-C-(X4-7)-C-(X4-17)-C-(X1)-C (Supplementary Materials Figure S5A,B). The similarity of the C-terminal CXC motif of the six-cysteine group with the cysteine defensin family of proteins suggests that disulfide bridging is a component of the protein conformation for mode of action as an antimicrobial peptide.

A total of 460 putative NCR genes were identified in the assembly of which 425 were expressed in the current study. Genome location and expression of the putative NCR genes are listed in Supplemental File (Supplementary Materials, Table S3). The predominant gene structure of the NCR genes in red clover was two exons interrupted by an intron near the signal peptide cleavage site. However, 32 genes also contain a second intron at the 3' end that included one to a few codons prior the stop codon. The inclusion of the third exon was annotated with the aid of 3' Tag-sequencing data [34]. There were a few (18) that appear to be intronless. A predicted signal peptide was found for 425 putative NCR peptides using the SignalP 4.1 software in CLC Genomics Workbench, while the remaining 34 did not have

a predicted signal peptide but had the conserved cysteine amino acids and are included as they were expressed in the current study.

The distribution of the number of NCR peptides based on the isoelectric point ranges is presented in Figure 2. It has been suggested that cationic peptides, those with the isoelectric point above 9.0, can function as antimicrobial peptides [15,35,36] of which the MtNCR247 has been specifically shown to be a strong antimicrobial peptide [35]. There are 149 expressed cationic NCR peptides in red clover which represents 32% of the red clover NCR peptides, while *M. truncatula* proportion is roughly 15%, alfalfa (*Medicago sativa*) is currently postulated to have 10%, and sweet clover (*Melilotus officinalis*) and pea (*Pisum sativum*), 9% and 8.5%, respectively. Interestingly, red clover only has a single anionic NCR with an isoelectric point less than 4.0, whereas the other species above contain roughly 6%.

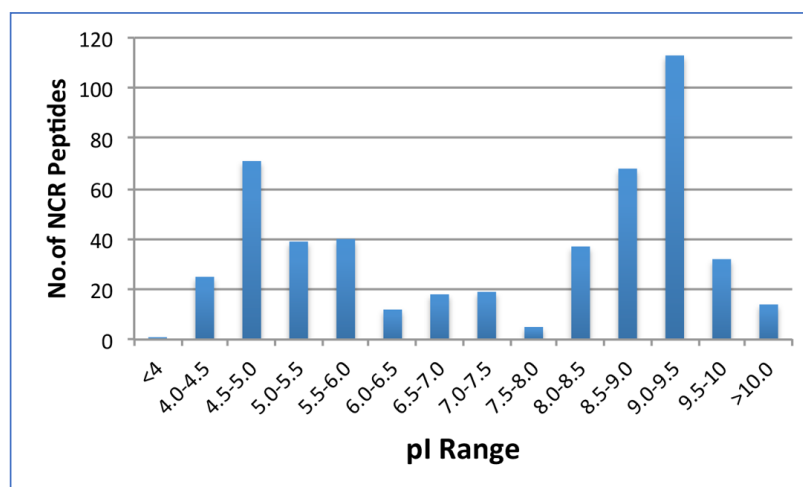


Figure 2. Number of the red clover NCR peptides distributed by isoelectric point (pI).

There are 25 genes that encode proteins containing eight, or more, cysteine residues. These are not listed as NCR genes and likely encode putative defensins [37,38]. Nevertheless, these genes are also observed to be primarily expressed in the Nod⁺ roots, have an N-terminal signal peptide, and thus likely targeted to the symbiosome.

2.1.3. ATS3-like/NPDs

Also observed in the highest expressed Nod⁺ gene list (Table 1) is a gene encoding a class of cysteine-rich peptides with homology to the *Arabidopsis thaliana* embryo-specific protein-3, or ATS3-like. In *M. truncatula* the nodule-specific ATS3-like genes are comprised of five closely related members (Figure S6), encoding proteins that have been named as nodule-specific Polycystin-1, Lipoxigenase, Alpha toxic (PLAT) domain proteins, or NPD's [9]. In red clover this gene family has expanded to 12 genes that are all tightly linked on chromosome 7 (Supplementary Figure S6). All of the red clover ATS3-like/NPD genes were expressed only in the Nod⁺ roots, most at high levels, including the *TpNPD5* (*Tp7g123899073*) that was one of the most highly expressed genes in this study (Table 1). Analysis using SignalP 4.1 software suggests that these proteins are cleaved and most likely targeted to the symbiosome as previously found for MtNPD1 [10]. The *TpNPD7* gene, while actively transcribed, probably encodes a non-function protein in the sequenced cultivar as it lacks a motif encoded by exon 2 and a small deletion in the C-terminus (Supplementary Materials Figure S7). Additionally, the *TpNPD9* locus, while also apparently transcribed, may be a pseudogene as the third exon is not proximal to the first two, although RNA-seq expression data suggests possible inclusion of the exon 3 sequence shared with *TpNPD12* (Supplementary Materials Figure S6). *TpNPD12* appears to encode two protein isoforms with a two amino acid difference due to alternative splicing in exon 2, both supported in the RNA-seq mapping results. The rationale for the expansion of this gene family in red clover is not known. In *M. truncatula* it appears that these proteins may function in the

selection for, or against, different strains of rhizobium [10,39], and thus it will be of interest to characterize this gene family in red clover as to whether it could aid in identifying better interacting *R. leguminosarum* strains with different red clover genotypes.

2.1.4. NAD1/NIN

Other genes involved in early nodulation include *NODULES WITH ACTIVATED DEFENSE 1* (*NAD1*) and *NODULE INCEPTION* (*NIN*). *MtNAD1* is essential for the control of plant defense during the colonization of the nitrogen-fixing nodule and is required for bacteroid persistence. *NIN* orthologs are required to initiate nodule formation in the cortex [40] and to transcriptionally activate NF-YA1 (Nuclear Factor Y), a transcription factor [41,42] involved in various nodulation steps [41,43,44]. In the epidermis, *NIN* is necessary for the onset of rhizobium infection [45,46] but restricts the expression of *EARLY NODULIN11* (*ENOD11*), a marker of the pre-infection and infection steps [47,48]. In red clover, the putative orthologs of *NAD1* (*Tp6g23891533*) and *NIN* (*Tp4g123881749*) both were up-regulated in the Nod+ roots. The *TpNIN* promoter region has the same upstream elements that regulate transcription as described for *MtNIN*, suggesting that regulation is probably similar [49].

The *MtNF-YA1* that controls meristem function is repressed in nodules through interaction with miR169, which is mediated by a small 62-aa peptide encoded by an alternatively spliced 5'-end of the *MtNF-YA1* gene that reduces expression [50]. The putative NF-YA1 ortholog in red clover (*Tp1g123902956*) was expressed as a significant DEG in the current study. Additionally, the alternative splicing of the first intron seen in *MtNF-YA1* was also observed in *TpNF-YA1*, suggesting a conserved regulation mechanism via miRNA169 [51].

2.1.5. DNF1-like and SYP132

Other components of the secretory pathway that are targeted to the symbiosome in *M. truncatula* include components of the signal peptidase complex (SPC) that include *dnf1* (*defective in nitrogen fixation 1*). Putative red clover genes that could encode putative orthologues of *dnf1* are: *Tp6g123890890*, *Tp6g123917546* or *Tp1g123900039*. None of the three genes were observed to be expressed only in the Nod+ roots, although *Tp6g123917546* was observed to have a two-fold increase in expression in the Nod+ roots, although not declared significantly different. No expression differences between the Nod+ and Nod− roots were observed in the other putative orthologues of the other subunits to the signal peptidase complex (*DAS12*, *DAS18* and *DAS25*) [52].

Another protein involved in the transport and protein targeting to the symbiosome is SYNTAXIN 132 (*SYP132*). In *M. truncatula*, the *SYP132* gene encodes two isoforms that distinguish host plasma membrane and symbiosome membrane via alternative splicing and polyadenylation [53,54]. While no significant differences in overall gene expression of the putative red clover *SYP132* (*Tp7g123897430*) between Nod+ and Nod− roots was observed, only reads from Nod+ root samples mapped to the alternative spliced isoform suggesting that *Tp7g123897430* gene is the red clover *SYP132* ortholog. Complicating the analysis of the putative *TpSYP132* gene is the presence of an overlapping gene (*Tp7g123897432*) at the 3' end with the alternatively spliced isoform (Figure S8). The *Tp7g123897432* gene is currently labeled as a ncRNA gene and is expressed in both Nod+ and Nod− roots. RNA-seq data suggests that this (*Tp7g123897432*) gene also produces alternative spliced mRNA isoforms but, as the ncRNA designations suggests, no open reading frames have been found. Additional analysis will be necessary to determine the function of this ncRNA and whether there might be any involvement in regulation of the red clover *SYP132* gene as suggested for *M. truncatula* [55].

2.1.6. Transcription Factor-Encoding Genes

One of the most highly expressed transcription factor-encoding genes was *Tp6g123889481* that encodes a C₂H₂ zinc finger protein that is orthologous to the *M. truncatula* REGULATION OF SYMBIOSOME DEVELOPMENT (*RSD*) [56]. Another Nod+-specific C₂H₂ zinc finger

gene, *Tp6g123890030* is orthologous *MtrunA17_Chr7g0231181*, which is also expressed only in nodules in *M. truncatula* [57]. The role of this putative C₂H₂ LOC123890030 gene product in nodulation is still unknown.

2.2. Down-Regulated Genes by Nodulation

A list of DEGs that were down-regulated in the Nod+ roots is presented in Excel file S1. These include a number of genes involved in pathogen response and biosynthesis of terpenoids (Table 2). This is consistent with the observations that the immune responses were suppressed during the symbiosis development and maintenance. Gene Ontology, Biological Process, Molecular Function and Cellular Component terms associated with these down-regulated genes are presented in Table 3. Expression of genes with molecular function GO:0016705 (oxidoreductase activity) was associated with a number of cytochrome P450's. These genes are likely involved in terpenoid biosynthesis as previously highlighted as highly differentially expressed.

Table 2. Most differentially expressed red clover genes in Nod− root samples compared to Nod+ samples.

Feature ID	Normalized Reads (Means) Nod−	Normalized Reads (Means) Nod+	Log2 Diff Express (Nod+)/(Nod−)	Product
LOC123899143	201.60	3.08	6.03	wound-responsive_family_protein
LOC123898241	210.61	3.43	5.94	wound-responsive_family_protein
LOC123899195	171.98	2.86	5.91	structural_constituent_of_cell_wall_protein,_putative
LOC123900972	110.72	2.63	5.39	sesquiterpene_synthase
LOC123923258	99.89	3.96	4.66	sesquiterpene_synthase
LOC123908919	312.53	13.93	4.49	geranylgeranyl_pyrophosphate_synthase
LOC123891221	66.37	3.09	4.42	disease_resistance_protein_(TIR-NBS-LRR_class)
LOC123920096	45.48	2.20	4.37	replication_factor-A_carboxy-terminal_domain_protein
LOC123917771	62.76	3.61	4.12	translation_elongation_factor_EF1B,_gamma_chain
LOC123908518	518.67	30.81	4.07	benzyl_alcohol_O-benzoyltransferase
LOC123924362	826.36	52.79	3.97	polygalacturonase_plant-like_protein
LOC123921773	1639.61	108.42	3.92	LysM_receptor_kinase_K1B
LOC123899196	98.73	6.92	3.83	1-aminocyclopropane-1-carboxylate_oxidase-like_protein
LOC123891479	241.74	18.20	3.73	cysteine/histidine-rich_C1_domain_protein
LOC123921461	230.24	17.49	3.72	cytochrome_P450_family_71_protein
LOC123912953	25.87	2.06	3.65	cytochrome_P450_family_71_protein
LOC123909630	39.54	3.18	3.64	amino acid transporter AVT1H
LOC123919510	15,190.12	1224.73	3.63	alpha-copaene synthase-like
LOC123891304	59.92	4.84	3.63	zinc finger protein 6-like
LOC123919667	210.91	17.06	3.63	disease_resistance-responsive,_dirigent_domain_protein

Table 3. Gene ontology (GO) Biological Process (P), Molecular Function (F) and Cellular Component (C) terms associated with differentially expressed genes (DEGs) of red clover Nod[−] plants greater than Nod⁺ plants.

GO Term	Ontology	Description	Number in Input List	Number in BG/Ref	p-Value	FDR
GO:0071555	P	cell wall organization	8	62	6.10×10^{-5}	0.016
GO:0045229	P	external encapsulating structure organization	8	62	6.10×10^{-5}	0.016
GO:0071554	P	cell wall organization or biogenesis	8	84	5.10×10^{-4}	0.077
GO:0042545	P	cell wall modification	6	48	0.0006	0.077
GO:0030599	F	pectinesterase activity	6	48	0.0006	0.059
GO:0020037	F	heme binding	20	384	0.00042	0.059
GO:0046906	F	tetrapyrrole binding	20	385	0.00044	0.059
GO:0052689	F	carboxylic ester hydrolase activity	7	68	0.00071	0.059
GO:0016705	F	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	16	292	0.00086	0.059
GO:0030312	C	external encapsulating structure	9	76	0.000043	0.0021
GO:0005618	C	cell wall	9	76	0.000043	0.0021

While not highlighted in the GO pathways, *Tp4g123921773*, a gene encoding a protein with homology to the LysM/CERK Nod Factor Receptors from *Lotus japonicus* (NFR5), soybeans (NFRa and NFRb) and pea (SYM37A) was also significantly down-regulated in the Nod[−] roots (Table 2). Other putative LysM/CERK homologs, *Tp4g123921715* and *Tp2g123904164*, showed a similar expression pattern. While these gene products have been shown to be the receptors for rhizobium and determinants that specify species-specific plant-rhizobium interactions [58] and it has not been demonstrated that these loci are the specific receptors in the interactions of red clover and rhizobia. Down-regulation of these genes in Nod⁺ roots may also be associated with suppression of immune responses because LysM-domain containing receptors also serve as immune receptors that perceive common microbe-associated molecular patterns such as chitin to trigger immunity signaling [59].

3. Materials and Methods

3.1. Plant Growth

Seeds of the Kenland cultivar and from crosses derived from a putative *isoflavone synthase 1* or *ifs1* mutant [32] were germinated on sterile moist filter paper and the germinated seedlings were individually transferred to 5-cm pots containing autoclaved sand. Seedlings were grown for two weeks in a growth chamber programmed for 16 h light at 22 °C and 8 h dark at 20 °C. The plants were then inoculated with a commercial strain of *Rhizobium leguminosarum biovar trifolii* sold as N-DureTM True Clover (Verdesian, Cary, NC, USA). The un-inoculated plants were used as controls. Water was applied twice daily using an overhead sprinkler system for six weeks. Following six weeks of growth, the nodulated roots (including both nodules and roots, designated as Nod⁺) and uninoculated roots (Nod[−]) were collected for transcriptome analysis.

3.2. RNA Extraction and Library Preparation

RNA was isolated and processed according to methods described in Dinkins et al. [32] and the frozen total RNA samples were submitted to NovoGene (Seattle, WA, USA) for the construction of 150-base paired-end mRNA libraries and subsequent sequencing following the manufacture procedures (<https://en.novogene.com/services/research-services/transcriptome-sequencing/eukaryote-mrna-sequencing/>, accessed on November 28 2019). Sequencing was done using Illumina Hi-Seq sequencing, resulting in a minimum 6 GB of

sequence data per library. The raw RNA-seq reads have been deposited in the Short Read Archive database at NCBI under BioProject PRJNA657870.

3.3. Analysis of RNA-Seq Reads

Raw reads were filtered to remove reads less than 30 bases and then mapped to the red clover genome assembly (version ARS_RC1.1; GenBank Accession No. GCA_020283565.1) [25] using the RNA-Seq tool, CLC Genomics Workbench (v22.0; CLC Genomics Workbench, Qiagen Aarhus A/S, Denmark). Minimum length fraction was set to 0.5 and minimum similarity fraction set to 0.8. Reads mapped to each gene were transformed by addition of one (1) to eliminate zeros and normalized by the Quantile method in CLC Genomics Workbench. Only genes that had a minimum of five average reads over either the Nod+ or Nod− samples were retained for differential expression analysis.

Manual annotation was performed for some of red clover genes based on the RNA-seq exon/intron expression profiles. This included manual annotation of the red clover NCR genes, many of which were annotated as ncRNAs in the current assembly, in order to identify the open reading frames. Thus we manually re-annotated the genes based on known conserved NCR gene structure and from sequences obtained from genes derived using the SPADA software [60] in the red clover genome assembly Tpv2.1 (GenBank Accession No. GCA_900079335.1). Following re-annotation, the RNA-seq tool in CLC Workbench was re-run to obtain the corrected read expression data. The SignalP 4.1 software in CLC Genomics Workbench [61,62] was used to predict the signal peptide cleavage site.

For differential expression analysis, the transformed read number in each sample was analyzed using JMP Genomics (version 9.0—SAS Institute, Cary, NC, USA). The normalized reads for each sample were log transformed (log2) using the JMP Basic RNA-Seq Workflow analysis of variance (ANOVA). Comparisons were done between Nod+ and Nod− samples over four replications (samples from four plants each). Based on the error estimated by ANOVA for each gene-by-gene comparison, significant differential expression was calculated in the comparisons between the *IFS1* wild-type and *ifs1* mutant expression means. A false discovery rate (FDR) multiple testing method at $-\log_{10}$ *p*-value of 2, that corresponds to an FDR *p*-value of 0.1, and a 2-fold difference were chosen as criteria for significance.

Pathway analysis of the differentially expressed genes was done using the annotation and gene ontology term from the *Trifolium pratense* Gene Ontology list in Legume Information Service (<https://legumeinfo.org>, accessed on 7 September 2021) using the web-based AgriGOv2 (<http://systemsbiology.cau.edu.cn/agriGOv2/>, accessed on 25 July 2022) [63].

4. Conclusions

Red clover has not been a major contributor in legume/rhizobium studies primarily due to the issues associated with self-incompatibility and difficulties obtaining homozygous plants for genetic analysis compared to the model species currently used. Using recently developed red clover genomic resources and RNA-seq expression analysis, we describe the genes that are expressed in nodule forming roots. As has been observed for other species, for red clover many of the genes involved in symbiosis are conserved with those of other legumes, while other genes are plant species-specific with little or no homology across species. These include the nodule-specific cysteine rich proteins (NCRs) that are found only in the IRLC legume clade, that while have common features, i.e., conserved cysteine residues, lack sequence homology and likely regulate the specific plant genotype/symbiont strain interactions. With the knowledge of the expression of these genes and genetic variation in red clover, breeding for increased nitrogen fixation efficiency by selection of genotypes with increased rhizobium specificity should be feasible.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants11212888/s1>, Figure S1. Mapped reads from Nod+ and Nod− red clover plants to putative Nod25-like genes on Chromosome 3. (A) *Tp3g123916815*; (B) *Tp3g123918937*; Figure S2. Gene structure and overall expression of the red clover NMS32/34-like genes. (A) NMS32/34-like *Tp3g123916883*; (B) NMS32/34-like *Tp3g123916884* and *Tp3g123916889*; (C) NMS32/34-like *Tp3g123918940*; (D) NMS32/34-like *Tp3g123918941*; Figure S3. Sequence alignment of the different mRNA isoforms encoding the red clover TP NMS32/34-like protein isoforms. Figure S4. Protein Sequence alignment of the red clover NMS32/34-like gene products. Figure S5. (A) Sequence alignment of red clover nodule-specific cysteine rich (NCR) mature peptides that contain four conserved cysteine residues. Figure S5. (B) Sequence alignment of a few of the red clover nodule-specific cysteine rich (NCR) mature peptides that contain six conserved cysteine residues. Figure S6. Location of the red clover embryo-specific ATS3-like/Nodule-specific PLAF domain (NPD) gene on Linkage Group 7 and mapped reads from the Nod+ and Nod− root samples. Figure S7. Protein sequence alignment of the red clover, *Medicago truncatula* and *Arabidopsis* embryo-specific ATS3-like/Nodule-specific PLAF domain (NPD) proteins. Figure S8. Chromosome location, expression and alternate splicing of the red clover putative SYP132 (*Tp7g123897430*) and ncRNA locus (*Tp7g123897432*). Supplemental Excel file: Excel Table S1: List of differentially expressed genes Nod− > Nod+. Excel Table S2: List of differentially expressed genes Nod+ > Nod−. Excel Table S3: List of red clover nodule-specific cysteine rich (NCR) peptides.

Author Contributions: R.D.D. planned the experiment, analyzed the data and wrote the manuscript. J.A.H. performed the experiments and analyzed the data; D.M.B. and M.L.S. provided genomic data; and H.Z. co-planned the experiment and co-wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the United States Department of Agriculture USDA—ARS CRIS project 5042-21000-002-00D and Non-Assistance Cooperative Agreement 5850428003 (H.Z.).

Data Availability Statement: The raw RNA-seq reads are deposited in the Short Read Archive database at GenBank under BioProject PRJNA657870.

Acknowledgments: The authors are thankful to Troy Bass for technical support with the greenhouse plant maintenance, and with DNA and RNA isolations; and to Diana Trujillo and Nevin Young for sharing their red clover genomic information.

Conflicts of Interest: The authors declare that there are no competing interest.

Disclaimer Note: Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

References

1. Kereszt, A.; Mergaert, P.; Kondorosi, E. Bacteroid development in legume nodules: Evolution of mutual benefit or of sacrificial victims? *Mol. Plant-Microbe Interact.* **2011**, *24*, 1300–1309. [[CrossRef](#)] [[PubMed](#)]
2. Haag, A.F.; Arnold, M.F.; Myka, K.K.; Kerscher, B.; Dall'Angelo, S.; Zanda, M.; Mergaert, P.; Ferguson, G.P. Molecular insights into bacteroid development during Rhizobium–legume symbiosis. *FEMS Microbiol. Rev.* **2013**, *37*, 364–383. [[CrossRef](#)]
3. Oono, R.; Denison, R.F. Comparing Symbiotic Efficiency between Swollen versus Nonswollen Rhizobial Bacteroids. *Plant Physiol.* **2010**, *154*, 1541–1548. [[CrossRef](#)] [[PubMed](#)]
4. Van de Velde, W.; Zehirov, G.; Szatmari, A.; Debreczeny, M.; Ishihara, H.; Kevei, Z.; Farkas, A.; Mikulass, K.; Nagy, A.; Tiricz, H.; et al. Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **2010**, *327*, 1122–1126. [[CrossRef](#)] [[PubMed](#)]
5. Wilson, R.C.; Long, F.; Maruoka, E.M.; Cooper, J.B. A new proline-rich early nodulin from *Medicago truncatula* is highly expressed in nodule meristematic cells. *Plant Cell* **1994**, *6*, 1265–1275. [[CrossRef](#)]
6. Liu, J.; Miller, S.S.; Graham, M.; Bucciarelli, B.; Catalano, C.M.; Sherrier, D.J.; Samac, D.A.; Ivashuta, S.; Fedorova, M.; Matsumoto, P.; et al. Recruitment of Novel Calcium-Binding Proteins for Root Nodule Symbiosis in *Medicago truncatula*. *Plant Physiol.* **2006**, *141*, 167–177. [[CrossRef](#)] [[PubMed](#)]
7. Kevei, Z.; Vinardell, J.M.; Kiss, G.B.; Kondorosi, A.; Kondorosi, E. Glycine-Rich Proteins Encoded by a Nodule-Specific Gene Family Are Implicated in Different Stages of Symbiotic Nodule Development in *Medicago* spp. *Mol. Plant-Microbe Interact.* **2002**, *15*, 922–931. [[CrossRef](#)] [[PubMed](#)]
8. Trujillo, D.I.; Silverstein, K.A.T.; Young, N.D. Genomic Characterization of the LEED..PEEDs, a Gene Family Unique to the *Medicago* Lineage. *G3 Genes | Genomes | Genet.* **2014**, *4*, 2003–2012. [[CrossRef](#)] [[PubMed](#)]
9. Trujillo, D.I.; Silverstein, K.A.T.; Young, N.D. Nodule-specific PLAT domain proteins are expanded in the *Medicago* lineage and required for nodulation. *New Phytol.* **2019**, *222*, 1538–1550. [[CrossRef](#)] [[PubMed](#)]

10. Pislariu, C.I.; Sinharoy, S.; Torres-Jerez, I.; Nakashima, J.; Blancaflor, E.B.; Udvardi, M.K. The Nodule-Specific PLAT Domain Protein NPD1 Is Required for Nitrogen-Fixing Symbiosis. *Plant Physiol.* **2019**, *180*, 1480–1497. [[CrossRef](#)] [[PubMed](#)]
11. Wang, D.; Yang, S.; Tang, F.; Zhu, H. Symbiosis specificity in the legume: Rhizobial mutualism. *Cell. Microbiol.* **2012**, *14*, 334–342. [[CrossRef](#)]
12. Yang, S.; Wang, Q.; Fedorova, E.; Liu, J.; Qin, Q.; Zheng, Q.; Price, P.A.; Pan, H.; Wang, D.; Griffiths, J.S.; et al. Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6848–6853. [[CrossRef](#)]
13. Roy, P.; Achom, M.; Wilkinson, H.; Lagunas, B.; Gifford, M.L. Symbiotic Outcome Modified by the Diversification from 7 to over 700 Nodule-Specific Cysteine-Rich Peptides. *Genes* **2020**, *11*, 348. [[CrossRef](#)] [[PubMed](#)]
14. Downie, J.A.; Kondorosi, E. Why Should Nodule Cysteine-Rich (NCR) Peptides Be Absent From Nodules of Some Groups of Legumes but Essential for Symbiotic N-Fixation in Others? *Front. Agron.* **2021**, *3*, 654576. [[CrossRef](#)]
15. Montiel, J.; Downie, J.A.; Farkas, A.; Bihari, P.; Herczeg, R.; Bálint, B.; Mergaert, P.; Kereszt, A.; Kondorosi, É. Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 5041–5046. [[CrossRef](#)]
16. Evers, G.W. Forage legumes: Forage quality, fixed nitrogen, or both. *Crop Sci.* **2011**, *51*, 403–409. [[CrossRef](#)]
17. Carlsson, G.; Huss-Danell, K. Nitrogen fixation in perennial forage legumes in the field. *Plant Soil* **2003**, *253*, 353–372. [[CrossRef](#)]
18. Ball, D.M.; Hoveland, C.S.; Lacefield, G.D. *Southern Forages*; International Plant Nutrition Group: Peachtree Corners, GA, USA, 2007; 332p.
19. Taylor, N. A century of clover breeding developments in the United States. *Crop Sci.* **2008**, *48*, 1–13. [[CrossRef](#)]
20. Vizintin, L.; Javornik, B.; Bohanec, B. Genetic characterization of selected *Trifolium* species as revealed by nuclear DNA content and ITS rDNA region analysis. *Plant Sci.* **2006**, *170*, 859–866. [[CrossRef](#)]
21. Riday, H.; Krohn, A.L. Genetic map-based location of the red clover (*Trifolium pratense* L.) gametophytic self-incompatibility locus. *Theor. Appl. Genet.* **2010**, *121*, 761–767. [[CrossRef](#)] [[PubMed](#)]
22. Chakrabarti, M.; Dinkins, R.D.; Hunt, A.G. De novo Transcriptome Assembly and Dynamic Spatial Gene Expression Analysis in Red Clover. *Plant Genome-US* **2016**, *9*, plantgenome2015.06.0048. [[CrossRef](#)] [[PubMed](#)]
23. Istvanek, J.; Jaros, M.; Kreněk, A.; Repkova, J. Genome assembly and annotation for red clover (*Trifolium pratense*; *Fabaceae*). *Am. J. Bot.* **2014**, *101*, 327–337. [[CrossRef](#)] [[PubMed](#)]
24. De Vega, J.J.; Ayling, S.; Hegarty, M.; Kudrna, D.; Goicoechea, J.L.; Ergon, A.; Rognli, O.A.; Jones, C.; Swain, M.; Geurts, R.; et al. Red clover (*Trifolium pratense* L.) draft genome provides a platform for trait improvement. *Sci. Rep.* **2015**, *5*, 17394. [[CrossRef](#)] [[PubMed](#)]
25. Bickhart, D.M.; Koch, L.M.; Smith, T.P.L.; Riday, H.; Sullivan, M.L. Chromosome-scale assembly of the highly heterozygous genome of red clover (*Trifolium pratense* L.), an allogamous forage crop species. *Gigabyte* **2022**, *2022*, 1–13. [[CrossRef](#)]
26. Li, W.; Riday, H.; Riehle, C.; Edwards, A.; Dinkins, R. Identification of Single Nucleotide Polymorphism in Red Clover (*Trifolium pratense* L.) Using Targeted Genomic Amplicon Sequencing and RNA-seq. *Front. Plant Sci.* **2019**, *10*, 1257. [[CrossRef](#)]
27. Osterman, J.; Hammenhag, C.; Ortiz, R.; Geleta, M. Insights Into the Genetic Diversity of Nordic Red Clover (*Trifolium pratense*) Revealed by SeqSNP-Based Genic Markers. *Front. Plant Sci.* **2021**, *12*, 748750. [[CrossRef](#)]
28. Jones, C.; De Vega, J.; Lloyd, D.; Hegarty, M.; Ayling, S.; Powell, W.; Skøt, L. Population structure and genetic diversity in red clover (*Trifolium pratense* L.) germplasm. *Sci. Rep.* **2020**, *10*, 8364. [[CrossRef](#)] [[PubMed](#)]
29. Istvanek, J.; Dluhosova, J.; Dluhos, P.; Patkova, L.; Nedelnik, J.; Repkova, J. Gene Classification and Mining of Molecular Markers Useful in Red Clover (*Trifolium pratense*) Breeding. *Front. Plant Sci.* **2017**, *8*, 367. [[CrossRef](#)]
30. Benevenuto, J.; Bhakta, M.; Lohr, D.A.; Ferrão, L.F.V.; Resende, M.F.R.; Kirst, M.; Quesenberry, K.; Munoz, P. Cost-effective detection of genome-wide signatures for 2,4-D herbicide resistance adaptation in red clover. *Sci. Rep.* **2019**, *9*, 20037. [[CrossRef](#)]
31. Trněný, O.; Vlk, D.; Macková, E.; Matoušková, M.; Řepková, J.; Nedělník, J.; Hofbauer, J.; Vejražka, K.; Jakešová, H.; Jansa, J.; et al. Allelic Variants for Candidate Nitrogen Fixation Genes Revealed by Sequencing in Red Clover (*Trifolium pratense* L.). *Int. J. Mol. Sci.* **2019**, *20*, 5470. [[CrossRef](#)]
32. Dinkins, R.D.; Hancock, J.; Coe, B.L.; May, J.B.; Goodman, J.P.; Bass, W.T.; Liu, J.; Fan, Y.; Zheng, Q.; Zhu, H. Isoflavone levels, nodulation and gene expression profiles of a CRISPR/Cas9 deletion mutant in the isoflavone synthase gene of red clover. *Plant Cell Rep.* **2021**, *40*, 517–528. [[CrossRef](#)] [[PubMed](#)]
33. Hohnjec, N.; Lenz, F.; Fehlberg, V.; Vieweg, M.F.; Baier, M.C.; Hause, B.; Küster, H. The Signal Peptide of the *Medicago truncatula* Modular Nodulin MtNOD25 Operates as an Address Label for the Specific Targeting of Proteins to Nitrogen-Fixing Symbiosomes. *Mol. Plant-Microbe Interact.* **2009**, *22*, 63–72. [[CrossRef](#)] [[PubMed](#)]
34. Chakrabarti, M.; Dinkins, R.D.; Hunt, A.G. Genome-wide atlas of alternative polyadenylation in the forage legume red clover. *Sci. Rep.* **2018**, *8*, 11379. [[CrossRef](#)] [[PubMed](#)]
35. Jenei, S.; Tiricz, H.; Szolomájer, J.; Tímár, E.; Klement, É.; Al Bouni, M.A.; Lima, R.M.; Kata, D.; Harmati, M.; Buzás, K.; et al. Potent Chimeric Antimicrobial Derivatives of the *Medicago truncatula* NCR247 Symbiotic Peptide. *Front. Microbiol.* **2020**, *11*, 270. [[CrossRef](#)] [[PubMed](#)]
36. Lima, R.M.; Kylarová, S.; Mergaert, P.; Kondorosi, É. Unexplored Arsenal of Legume Peptides with Potential for Their Applications in Medicine and Agriculture. *Front. Microbiol.* **2020**, *11*, 1307. [[CrossRef](#)]

37. Parisi, K.; Shafee, T.M.A.; Quimbar, P.; van der Weerden, N.L.; Bleackley, M.R.; Anderson, M.A. The evolution, function and mechanisms of action for plant defensins. *Semin. Cell Dev. Biol.* **2019**, *88*, 107–118. [\[CrossRef\]](#)
38. Terras, F.R.; Eggermont, K.; Kovaleva, V.; Raikhel, N.V.; Osborn, R.W.; Kester, A.; Rees, S.B.; Torrekens, S.; Van Leuven, F.; Vanderleyden, J. Small cysteine-rich antifungal proteins from radish: Their role in host defense. *Plant Cell* **1995**, *7*, 573–588. [\[CrossRef\]](#)
39. Burghardt, L.T.; Trujillo, D.I.; Epstein, B.; Tiffin, P.; Young, N.D. A Select and Resequencing Approach Reveals Strain-Specific Effects of Medicago Nodule-Specific PLAT-Domain Genes. *Plant Physiol.* **2020**, *182*, 463–471. [\[CrossRef\]](#)
40. Schauser, L.; Roussis, A.; Stiller, J.; Stougaard, J. A plant regulator controlling development of symbiotic root nodules. *Nature* **1999**, *402*, 191–195. [\[CrossRef\]](#)
41. Soyano, T.; Kouchi, H.; Hirota, A.; Hayashi, M. Nodule Inception Directly Targets NF-Y Subunit Genes to Regulate Essential Processes of Root Nodule Development in Lotus japonicus. *PLoS Genet.* **2013**, *9*, e1003352. [\[CrossRef\]](#)
42. Laloum, T.; Baudin, M.; Frances, L.; Lepage, A.; Billault-Penneteau, B.; Cerri, M.R.; Ariel, F.; Jardinaud, M.-F.; Gamas, P.; de Carvalho-Niebel, F.; et al. Two CCAAT-box-binding transcription factors redundantly regulate early steps of the legume-rhizobia endosymbiosis. *Plant J.* **2014**, *79*, 757–768. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Combier, J.-P.; Frugier, F.; de Billy, F.; Boualem, A.; El-Yahyaoui, F.; Moreau, S.; Vernié, T.; Ott, T.; Gamas, P.; Crespi, M.; et al. MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev.* **2006**, *20*, 3084–3088. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Xiao, T.T.; Schilderink, S.; Moling, S.; Deinum, E.E.; Kondorosi, E.; Franssen, H.; Kulikova, O.; Niebel, A.; Bisseling, T. Fate map of *Medicago truncatula* root nodules. *Development* **2014**, *141*, 3517–3528. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Fournier, J.; Teillet, A.; Chabaud, M.; Ivanov, S.; Genre, A.; Limpens, E.; de Carvalho-Niebel, F.; Barker, D.G. Remodeling of the Infection Chamber before Infection Thread Formation Reveals a Two-Step Mechanism for Rhizobial Entry into the Host Legume Root Hair. *Plant Physiol.* **2015**, *167*, 1233–1242. [\[CrossRef\]](#)
46. Xie, F.; Murray, J.D.; Kim, J.; Heckmann, A.B.; Edwards, A.; Oldroyd, G.E.D.; Downie, J.A. Legume pectate lyase required for root infection by rhizobia. *Proc. Natl. Acad. Sci. USA* **2011**, *109*, 633–638. [\[CrossRef\]](#)
47. Marsh, J.F.; Rakocevic, A.; Mitra, R.M.; Brocard, L.; Sun, J.; Eschstruth, A.; Long, S.R.; Schultze, M.; Ratet, P.; Oldroyd, G.E.D. *Medicago truncatula* NIN Is Essential for Rhizobial-Independent Nodule Organogenesis Induced by Autoactive Calcium/Calmodulin-Dependent Protein Kinase. *Plant Physiol.* **2007**, *144*, 324–335. [\[CrossRef\]](#)
48. Vernié, T.; Kim, J.; Frances, L.; Ding, Y.; Sun, J.; Guan, D.; Niebel, A.; Gifford, M.L.; de Carvalho-Niebel, F.; Oldroyd, G.E.D. The NIN Transcription Factor Coordinates Diverse Nodulation Programs in Different Tissues of the *Medicago truncatula* Root. *Plant Cell* **2015**, *27*, 3410–3424. [\[CrossRef\]](#)
49. Liu, J.; Rutten, L.; Limpens, E.; van der Molen, T.; van Velzen, R.; Chen, R.; Chen, Y.; Geurts, R.; Kohlen, W.; Kulikova, O.; et al. A Remote cis-Regulatory Region Is Required for NIN Expression in the Pericycle to Initiate Nodule Primordium Formation in *Medicago truncatula*. *Plant Cell* **2019**, *31*, 68–83. [\[CrossRef\]](#)
50. Combier, J.P.; de Billy, F.; Gamas, P.; Niebel, A.; Rivas, S. Trans-regulation of the expression of the transcription factor MtHAP2-1 by a uORF controls root nodule development. *Genes Dev.* **2008**, *22*, 1549–1559. [\[CrossRef\]](#)
51. Sorin, C.; Declerck, M.; Christ, A.; Blein, T.; Ma, L.; Lelandais-Brière, C.; Njo, M.F.; Beeckman, T.; Crespi, M.; Hartmann, C. A miR169 isoform regulates specific NF-Y targets and root architecture in Arabidopsis. *New Phytol.* **2014**, *202*, 1197–1211. [\[CrossRef\]](#)
52. Wang, D.; Griffiths, J.; Starker, C.; Fedorova, E.; Limpens, E.; Ivanov, S.; Bisseling, T.; Long, S. A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* **2010**, *327*, 1126–1129. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Huisman, R.; Hontelez, J.; Mysore, K.S.; Wen, J.; Bisseling, T.; Limpens, E. A symbiosis-dedicated SYNTAXIN OF PLANTS 13II isoform controls the formation of a stable host–microbe interface in symbiosis. *New Phytol.* **2016**, *211*, 1338–1351. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Pan, H.; Oztas, O.; Zhang, X.; Wu, X.; Stonoha, C.; Wang, E.; Wang, B.; Wang, D. A symbiotic SNARE protein generated by alternative termination of transcription. *Nat. Plants* **2016**, *2*, 15197. [\[CrossRef\]](#)
55. Yu, L.; Huang, T.; Qi, X.; Yu, J.; Wu, T.; Luo, Z.; Zhou, L.; Li, Y. Genome-Wide Analysis of Long Non-coding RNAs Involved in Nodule Senescence in *Medicago truncatula*. *Front. Plant Sci.* **2022**, *13*, 917840. [\[CrossRef\]](#)
56. Sinharoy, S.; Torres-Jerez, I.; Bandyopadhyay, K.; Kereszt, A.; Pislariu, C.I.; Nakashima, J.; Benedito, V.A.; Kondorosi, E.; Udvardi, M.K. The C₂H₂ Transcription Factor REGULATOR OF SYMBIOSOME DIFFERENTIATION Represses Transcription of the Secretory Pathway Gene VAMP721a and Promotes Symbiosome Development in *Medicago truncatula*. *Plant Cell* **2013**, *25*, 3584–3601. [\[CrossRef\]](#)
57. Roux, B.; Rodde, N.; Jardinaud, M.-F.; Timmers, T.; Sauviac, L.; Cottret, L.; Carrère, S.; Sallet, E.; Courcelle, E.; Moreau, S.; et al. An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J.* **2014**, *77*, 817–837. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Bisseling, T.; Geurts, R. Specificity in legume nodule symbiosis. *Science* **2020**, *369*, 620–621. [\[CrossRef\]](#)
59. Bozsoki, Z.; Cheng, J.; Feng, F.; Gysel, K.; Vinther, M.; Andersen, K.R.; Oldroyd, G.; Blaise, M.; Radutoiu, S.; Stougaard, J. Receptor-mediated chitin perception in legume roots is functionally separable from Nod factor perception. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8118–E8127. [\[CrossRef\]](#)
60. Zhou, P.; Silverstein, K.A.T.; Gao, L.; Walton, J.D.; Nallu, S.; Guhlin, J.; Young, N.D. Detecting small plant peptides using SPADA (Small Peptide Alignment Discovery Application). *BMC Bioinform.* **2013**, *14*, 335. [\[CrossRef\]](#)

-
61. Nielsen, H.; Engelbrecht, J.; Brunak, S.; von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng. Des. Sel.* **1997**, *10*, 1–6. [[CrossRef](#)]
 62. Bendtsen, J.D.; Nielsen, H.; von Heijne, G.; Brunak, S. Improved Prediction of Signal Peptides: SignalP 3.0. *J. Mol. Biol.* **2004**, *340*, 783–795. [[CrossRef](#)] [[PubMed](#)]
 63. Tian, T.; Liu, Y.; Yan, H.; You, Q.; Yi, X.; Du, Z.; Xu, W.; Su, Z. agriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **2017**, *45*, W122–W129. [[CrossRef](#)] [[PubMed](#)]