Analysis of the structure and function of the LYK cluster of Medicago truncatula A17 and R108

Thi-Bich Luu, Noémie Carles, Louis Bouzou, Chrystel Gibelin-Viala, Céline Remblière, Virginie Gasciolli, Jean-Jacques Bono, Benoit Lefebvre, Nicolas Pauly, Julie Cullimore

Analysis of the structure and function of the *LYK* cluster of *Medicago truncatula* A17 and R108

Thi-Bich Luu[a,1], Noémie Carles[a], Louis Bouzou[a], Chrystel Gibelin-Viala[a], Céline Remblière[a], Virginie Gasciolli[a], Jean-Jacques Bono[a], Benoit Lefebvre[a], Nicolas Pauly[a,b,2,*] and Julie Cullimore[a,2]

[a]Laboratory of Plant-Microbe-Environment Interactions (LIPME), Université de Toulouse, INRAE, CNRS, Castanet-Tolosan, France.

[b]Institut Sophia Agrobiotech, Université Côte d'Azur, INRAE, CNRS, Sophia Antipolis Cedex, France.

[1]Present address: Department of Molecular Biology and Genetics, Aarhus University, Denmark.

[2] Co-last authors

* Corresponding author:

Email: nicolas.pauly@inrae.fr

ABSTRACT

The establishment of the Legume-Rhizobia symbiosis is generally dependent on the production of rhizobial lipochitooligosaccharidic Nod factors (NFs) and their perception by plant Lysin Motif Receptor-Like Kinases (LysM-RLKs). In this study, we characterized a cluster of LysM-RLK genes implicated in strain-specific recognition in two highly divergent and widely-studied *Medicago truncatula* genotypes, A17 and R108. We then used reverse genetic approaches and biochemical analyses to study the function of selected genes in the clusters and the ability of their encoded proteins to bind NFs. Our study has revealed that the *LYK* cluster exhibits a high degree of variability among *M. truncatula* genotypes, which in A17 and R108 includes recent recombination events within the cluster and a transposon insertion in A17. The essential role of *LYK3* in nodulation in A17 is not conserved in R108 despite similar sequences and good nodulation expression profiles. Although, *LYK2, LYK5* and *LYK5bis* are not essential for nodulation of the two genotypes, some evidence points to accessory roles in nodulation, but not through high-affinity NF binding. This work shows that recent evolution in the *LYK* cluster provides a source of variation for nodulation, and potential robustness of signaling through genetic redundancy.

Key words: *Medicago truncatula*, legume-Rhizobia symbiosis, LysM receptor-like kinase, Nod factor perception.

1. Introduction

The Fabaceae is one of the largest families among flowering plants and greatly contributes to animal and human nutrition through being second, behind the Poaceae, in terms of world crop production [1]. This family is well-documented for its ability to form root-nodule, endophytic symbiotic interactions with soil nitrogen-fixing bacteria termed rhizobia, leading to their importance in sustainable agriculture and the natural environment.

The Legume-Rhizobia Symbiosis (LRS) is thought to have evolved from the more-widespread arbuscular mycorrhiza endosymbiosis (AM) [2] and to have a common origin with other nitrogen–fixing nodule symbioses, including the rhizobia symbiosis with the non-legume *Parasponia* [3,4]. Legumes and rhizobia exhibit partner specificity that is generally highly dependent on the structure of lipochitooligosaccharide (LCO) signals, called Nod factors (NFs) produced by bacteria under the control of a set of *Nodulation* (*nod*) genes [5]. NFs are generally composed of three to five residues of *N*-acetyl glucosamine (chitin backbone), *N*-acylated on the terminal non-reducing sugar and they manifest strain-specific, diverse decorations at both ends of the molecule [6]. Analysis of nodulation of different *Sinorhizobium meliloti* 2011 *nod* mutants affected in the structure of the long fatty acid chain (specified by *nodFE*) and the *O*-acetylation of the non-reducing sugar (specified by *nodL*) has revealed the significant effects of these substitutions for both infection and nodulation with various *Medicago* hosts [7]. Purified or synthetic NFs elicit symbiotic responses on legume roots down to pM concentrations, suggesting that they are perceived by high affinity receptors [7].

From the plant side, the perception of NF involves Lysin Motif Receptor-Like Kinases (LysM-RLKs), a plant-specific RLK family [8]. Members of this family have been shown to be located at the plasma membrane and are composed of an extracellular region (ECR) containing three LysMs, a single transmembrane-spanning helix (TM), which connects to an intracellular region (ICR) with a kinase-like domain [9]. Over the past two decades, the role of LysM-RLKs in LRS has been particularly studied in the model legumes *Medicago truncatula* (Mt) and *Lotus japonicus* (Lj), but also in legume crops such as pea (*Pisum sativum* – Ps) and soybean (*Glycine max* – Gm), and the non-legume *Parasponia* [10]. These studies suggest that legumes and *Parasponia* spp. use

orthologous LysM-type receptors to perceive rhizobium LCOs, suggesting a shared evolutionary origin of LCO-driven nodulation [10]. In each species, two pairs of orthologous LysM-RLKs, corresponding to MtNFP/LjNFR5 and MtLYK3/LjNFR1, have been shown to be essential for nodulation [11]. In *M. truncatula NFP* is required for all NF responses [12] and can be partially substituted by its orthologs from legumes and non-legumes [13,14] while *LYK3* stringently regulates the perception of specific NFs for infection but is not essential for early NF responses [15–18]. NFP and LYK3 have been shown to physically interact in nodules [19], and co-expression in *Nicotiana benthamiana* leaves leads to a cell-death response [19–21], suggesting that they may function as a heteromer in plants. Recently, NFP and LYK3 have been shown to individually bind NFs with an affinity (*K*d) in the µM range [18,22], which may suggest that a specific conformation of the two receptors or additional components is required for perception of NF.

In legumes, the *MtLYK3/LjNFR1* gene is part of the orthogroup called *LYK*-I, which includes the *CERK1* gene from non-legumes [9]. In rice (*Oryza sativa*), *OsCERK1* plays a major role in the perception of chitin-oligomers (COs) produced by arbuscular mycorrhizal fungi (AMF) and fungal pathogens and is required for both AM and for defence responses [23–25] The capacity to form AM has been lost in the Brassicaceae [26] but studies on *Arabidopsis thaliana* has shown that *CERK1* plays a major role in defence and in the perception of chitin-oligomers and other glucans [26,27]. In *Parasponia andersonii* duplication of the *LYKI* gene has led to *PanLYK3* and *PanLYK1*, both of which are essential for AM and control nodulation with rhizobia, with *PanLYK3* also retaining a role in CO perception [10]. In legumes the number of *LYK-I* genes has increased (9 genes in Mt and 5 in Lj) with many of them being located in a *LYK* gene cluster [9,28,29], which was initially identified through genomic synteny with the pea (*Pisum sativum*) *SYM2* locus, in which variation allows this species to nodulate with specific rhizobia strains [30]. Evidence from molecular phylogeny strongly indicates that neofunctionalization of the *MtLYK3/LjNFR1* gene in the *LYK* cluster, prior to legume speciation, allowed legumes to gain the ability to interact with specific rhizobia [9,31,32].

M. truncatula has been chosen as one of the models for legume biology studies due to its relatively small genome (compared to other legumes) and the availability of a

large number of genotypes, displaying genetic and geographic diversities [33,34]. *M. truncatula* spp. *truncatula* cv Jemalong A17 (referred to here as A17) and *M. truncatula* spp. *tricycla* R108 (referred to here as R108) are the two most commonly-used genotypes for research as their genomes have been sequenced and tools have been developed for genetic and functional studies [35,36]. The two genotypes are very distinct as shown by phylogenomics of sequenced *Medicago* accessions [37], and by analysis of various traits, including capacity to nodulate with different rhizobia strains [33,34]. Recently, we have reported a variation in the *LYK* cluster of the two genotypes, involving the presence of a R108-specific gene, *LYK2bis*, which extends the nodulation specificity of R108 to *S. meliloti* 2011 *nodL* mutant producing non-*O*-acetylated NFs and to some natural strain isolates [29]. In this article, we report on the divergence of the whole *LYK* cluster in the two genotypes and functional studies of specific *LYK* genes in the cluster.

2. Material and methods

2.1. Seed germination and growth conditions

M. truncatula seeds were extracted from pods and scarified in 95% sulfuric acid for 5 min followed by two washes with distilled water; then surface-sterilized in 3.2% chlorine bleach for 3 min then washed three times. After that, the seeds were kept in distilled water for 1 h and placed on 1% agar plates, supplemented with 1 µg ml^{-1} GA3. The plates were left upside-down for 5 d at 4ºC and then put at 16ºC overnight for germination. Seedlings were transferred to either pots or tubes and grown in a growth chamber at 22ºC with a 16 h photoperiod.

2.2. Medicago truncatula *mutants*

For CRISPR/Cas9 gene editing, the first exons of *LYK2*, *LYK5* and *LYK5bis* were used to design protospacers using CRISPOR (http://crispor.tefor.net/). For each gene, a polycistronic RNA [38] containing two or three guide RNAs (gRNAs) targeting the NGG PAM sites at positions from the start codon at 81, 109 and 436 (for *LYK2*), 335, 382 and 549 (for *LYK5*) and 376 and 518 (for *LYK5bis*), were produced either by DNA synthesis (Invitrogen) or by PCR/ligation and cloned under the control of either the MtU6.6 promoter (for *LYK2* and *LYK5*) or the MtU6.1 promoter (for *LYK5bis*). The CRISPR module(s) was, combined by Golden Gate cloning, into the pL2V-1 binary vector with modules for

Kanamycin resistance, DsRed and Cas9, essentially as described [39]. The constructs were then transformed into *Agrobacterium tumefaciens* AGL1 strain and used to transform and regenerate *M. truncatula* line 2HA essentially as described [40]. Regenerated plantlets were genotyped by PCR and sequencing, using primers listed in Table **S1**. First generation seeds were germinated and genotyped to select homozygous mutants of *lyk2* and double *lyk5/lyk5bis* and lacking Cas9, the progeny of which was used for phenotyping.

Tnt1 insertional mutants of lines, *lyk2-1R* (NF13076), *lyk3-1R* (NF2752), *lyk5-1R* (NF11221) and *lyk5bis-1R* (NF2619) of R108 and the fast neutron mutant of A17, FN9805, were obtained from the Noble Research Institute (USA). The genotyping primers for analysing the original mutants and backcrosses are provided in Table **S1**. All genotypes are described in Table **S2.**

2.3. Complementation assays

For *M. truncatula* hairy root transformation, full length coding sequences (CDS) of *LYK2*, *LYK2bis* and *LYK3* from R108, and *LYK2*, *LYK3* and *LYK4* from A17 were tagged with mCherry under the control of the Ubiquitin promoter from *L. japonicus* (ProLjUbi) by using Golden gate cloning [21]. Using *Agrobacterium rhizogenes*-mediated transformation [41], seedlings of *lyk3-1* were transformed with strains containing either empty vector (EV) or different constructs (Table **S3**), and transformants were selected on medium containing 25 µg ml^{-1} kanamycin, and after two-weeks growth, by expression of the DsRed marker.

2.4. Nodulation assays

The nodulation tests were performed using various rhizobial strains (Table. **S4**) either in pots or tubes as previously described [29]. R108, 2HA, A17 and wild type-like (WTL) siblings were used as controls. In each experiment, at least 10 plants per line per strain were analysed.

2.5. Root hair deformation assays

Germinated seedlings were transferred to square plates (12x12 cm), 5 plants/plate, containing Fahraeus agar medium supplemented with 1 mM $CaCl_2$ and 0.2

mM NH_4NO_3 and kept in a growth chamber. After 7d, each plant root system was treated with 200 µL of either 10^{-8} M of NFs purified from *S. meliloti* 2011 WT, or the *S. meliloti* 2011 *nodF/nodL* mutant (GMI 6630) at OD_{600}= 0.1, or water as a control. After 1 h, the excess of liquid was removed, and plates were kept in the same growth condition for 24 h (NF treatment) or 8 days (bacteria inoculation). Plant roots were then collected, stained with trypan blue, and scored for root hair deformation (RHD) index by light microscopy. The scoring was based on the percentage of root hairs that deformed at the susceptible zone (about 1.0-1.5 cm from the root tip): 0: 0%; 1: <10%; 2: 10-50%; 3: >50%. Each experiment used 10 plants per line per treatment.

2.6. Gene expression analysis by qRT-PCR

A17 and FN9805 root samples treated for 24 h with NFs were collected and frozen in liquid nitrogen. The RNAs were extracted using Nucleospin plant RNA extraction kit (Macherey-Nagel GmbH & Co. KG, Germany) and quantified using Nanodrop. cDNAs were synthesized using SuperScript IV Reverse Transcriptase kit (ThermoFisher Scientific, USA) and used as templates for qRT-PCR analysis using LightCycler 480 (Roche, Switzerland). Early NF-induced genes *ENOD11*, *PUB1*, *VAPYRIN* and *LYK10* were amplified using primers described in [42].

2.7. Mycorrhiza assays

Mycorrhiza tests were performed using the gridline intersect method, as described in [32]. Briefly, FN9805 and A17 seedlings were inoculated with 200 spores per plant of *Rhizophagus irregularis* DAOM197198 (Agronutrition, Toulouse, France) and colonization was assessed at 2 and 4 wk post inoculation (wpi) using 10 plants per genotype per time-point.

2.8. Protein expression in Nicotiana benthamiana *leaves and ligand binding assays*

Golden Gate cloning was used to produce constructs of either the full-length CDS (for LYK2, LYK2bis, LYK3 or the CDS of the predicted ECR and TM (LYK5, LYK5bis), fused to a fluorescent protein (YFP, TagGFP or mCherry) with a strong constitutive promoter (ProCaMV 35S or ProLjUbi) in a pCambia vector (Table **S3**), as previously described [43]. *Agrobacterium tumefaciens LBA4404* strains containing different fusion

constructs were used to agro-infiltrate the three to four oldest leaves of *N. benthamiana*. Three days post infiltration, the fluorescent protein expression in leaves was assessed by confocal microscopy and the expressing leaves were then frozen in liquid nitrogen. Membrane fractions were isolated as previously described [43] and the amount of protein was quantified using the Pierce™ BCA Protein Assay Kit ThermoFisher Scientific, USA). Equilibrium binding assays were performed using radiolabelled LCOs, corresponding to NFs, as described [14,43].

2.9. Protein expression in E. coli *and kinase assays*

The predicted ICR of LYK2 from A17 (termed LYK2A-KD) was cloned and fused with Glutathione-S-transferase (GST). The fusion was then expressed in *E. coli* DH5α and purified using glutathione resin (GE Healthcare, USA) as described [21]. The LYK2A-KD was released from the GST using PreScission Protease (GE27-0843-01, Sigma Aldrich, Germany). The *in vitro* phosphorylation assays using ^{32}P-ATP were performed as previously described [29].

3. Results

3.1. Structure of the LYK cluster on chromosome 5 of M. truncatula *A17 and R108*

The sequence of the *LYK* cluster region was extracted from the genome of A17 (v4.0) and the more recently released genome of R108 [35,36]. The introns/exons of the LysM-RLK related genes were corrected manually by comparison to transcriptomic data (Fig. **1**).

In A17, the cluster contains the seven complete genes, *LYK1* to *LYK7,* described previously [30] (Table **S5**). In addition, we identified a truncated LysM gene with an ECR+TM and very short ICR, located in between *LYK5* and *LYK4*, which we designated as *LYK5bis*, as its sequence is very similar to LYK5 (Fig. **1,** Fig. **S1,** Table **S6)**.

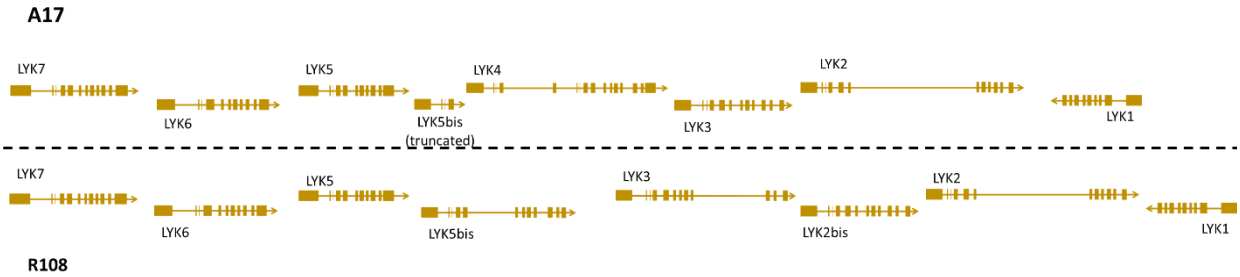

Fig. 1. Structure of the *LYK* cluster on chromosome 5 of *Medicago truncatula* A17 and R108. Scheme based on gene browsers of A17 v 4.0 and 5.0 (https://medicago.toulouse.inra.fr/MtrunA17r5.0-ANR/) and of R108 r1 (https://medicago.toulouse.inra.fr/MtrunR108r1/). Coding sequences are shown in filled blocks; non-coding sequences are shown in lines. Gaps between genes are not to scale.

In R108, the cluster shows some differences compared to the one in A17. We have already reported on the extra gene in R108 located between *LYK3* and *LYK2* and designated as *LYK2bis,* which is a chimera between *LYK2* and *LYK3* [29]. Another striking difference is that there is no *LYK4* gene in R108 but a complete *LYK5bis* gene was found (Fig. **1**). Pairwise identities of the whole proteins, ECR and ICR (Table **S6**) and phylogenetic analysis (Fig. **S2**) show that the LysM-RLK genes of A17 and R108 with the same designation and the same relative location in the cluster are indeed very similar, and probably orthologous. LYK4 however is a chimeric protein with an ECR similar to LYK3 and an ICR similar to LYK5bis of R108 (Fig. **S1**). Between the truncated LYK5bis of A17 and LYK4 of A17 a transposon sequence is present, annotated as RLX_singleton_family289_LTR_retrotransposon2745.

In the following text, the genotype is added to the gene and protein names, as in Fig. **S2**, when necessary to avoid ambiguity.

Analysis of the gene structure of the *LYK* cluster *LysM-RLKs* from A17 and R108 shows that most orthologous pairs have similar numbers of exons and that the introns are of a similar size (Fig. **1**). However, *LYK3-R108* contains a much larger intron 9 (about 3 kbp) than *LYK3-A17* (Fig. **1**). Also, intron 1 of *LYK3-R108* does not have the predominant GT-AG splice site but has a non-canonical GC-AG splice site, which we have confirmed by examination of RNA-seq reads and by cDNA cloning. The LYK3-R108 protein sequence, in relation to LYK3-A17, shows a cluster of polymorphisms in LysM3, more

similar to LYK2 [29], suggesting that there may have been a recombination event in the *LYK3-R108* gene.

It should be noted that the *LYK* cluster region (from *LYK7* to *LYK1*) extends over 192 kbp in A17 and 228 kbp in R108 (Fig. **S3a**) and that other non-related genes are interspersed between the *LysM-RLKs* (Fig. **S3b**). Analysis of the genome sequencing data of 15 different *M. truncatula* genotypes [37] revealed considerable variation in the predicted *LysM-RLKs* and other genes in the cluster (Fig. **S3b**). *LYK3* and *LYK2* are present in most genotypes, whereas *LYK2bis* is only present in R108, as shown previously [29]. Most genomes show predicted *LYK5* and *LYK5bis* genes. Another related gene but without a predicted ECR is present in both A17 and R108 between *LYK6* and *LYK5* and is termed *LYK5like KIN* (Fig. **S3b**).

Phylogenetic analysis (Fig. **S2**) shows that the *M. truncatula* LYK cluster proteins are together in a clade and are most closely related to LYK9 and then to LYK8 (both LYK-I class proteins) and are clearly separable to the LYK-II (LYK10) and LYK-III (LYK11) proteins, as defined previously [9,10]. The similarities between the genes suggests that there could be redundancy in function of the *LYK* cluster and *LYK8* and *LYK9* genes.

3.2. Expression of the LYK cluster genes

The expression data of *LYK* cluster genes, and for comparison, *NFP*, *LYR3* and *LYK10*, were extracted from transcriptomic data [44] in which the susceptible region of roots of Jemalong or R108 were analysed 24h post spot inoculation with *S. meliloti* (Fig. **S4**). As expected from work on A17 [45], *LYK10* is induced following inoculation in both genotypes. Of the *LYK* cluster genes, *LYK3* is by far the most highly expressed gene in these parts of the roots in Jemalong. In R108, *LYK2bis*, shows highest expression before inoculation, whereas both *LYK3* and *LYK5bis* show induction at 24h post-inoculation with *S. meliloti*, like *NFP*. In both genotypes *LYK2*, showed low expression which was decreased following *S. meliloti* inoculation, as shown previously for root hairs in A17 following NF treatment [8].

3.3. Roles in nodulation of selected LYK cluster genes in R108

In view of the expression patterns of the genes during symbiosis, and their conservation or divergence between genotypes, we chose *LYK2*, *LYK3*, *LYK5* and

LYK5bis for functional studies. Both A17 and R108 nodulate with *S. meliloti* 2011 but show differences in nodulation with NF mutants. In particular, A17 showed much reduced or abolished nodulation after being inoculated with *S. meliloti* 2011 *nodL* or *nodF/nodL* mutants respectively [7,17] whereas the presence of *LYK2bis* allows R108 to nodulate well with both mutants, whereas it is not essential for nodulation with the *S. meliloti* 2011 WT strain [29]. This evidence suggests that other genes in the R108 cluster play important roles in nodulation with *S. meliloti* 2011. Therefore, to determine the roles of other genes in nodulation in R108, *Tnt1* insertional mutants of *LYK2*, *LYK3*, *LYK5* and *LYK5bis*, designated as *lyk2-1R*, *lyk3-1R*, *lyk5-1R* and *lyk5bis-1R,* were identified from the Noble Foundation library (Fig. **2a**, Table **S2**). These mutants were phenotyped for nodulation at 21 dpi, following inoculation with either *S. meliloti* 2011 WT or *nodF/nodL* strains (Fig. **2b, c**).

A similar number of nodules per plant was obtained in all tested lines inoculated with *S. meliloti* 2011 suggesting that none of these genes play an essential role in nodulation with this WT strain. However, when inoculated with the *S. meliloti* 2011 *nodF/nodL* strain, the *lyk5-1R* mutant formed significantly less nodules compared to R108 and the other lines, indicating that *LYK5-R108* may play a role in controlling nodule formation of R108 with the 2011 *nodF/nodL* mutant but with a mean of 23.5 nodules/plant its role is minor in comparison to the essential role of *LYK2bis* (zero nodules/plant in the same conditions [29]).

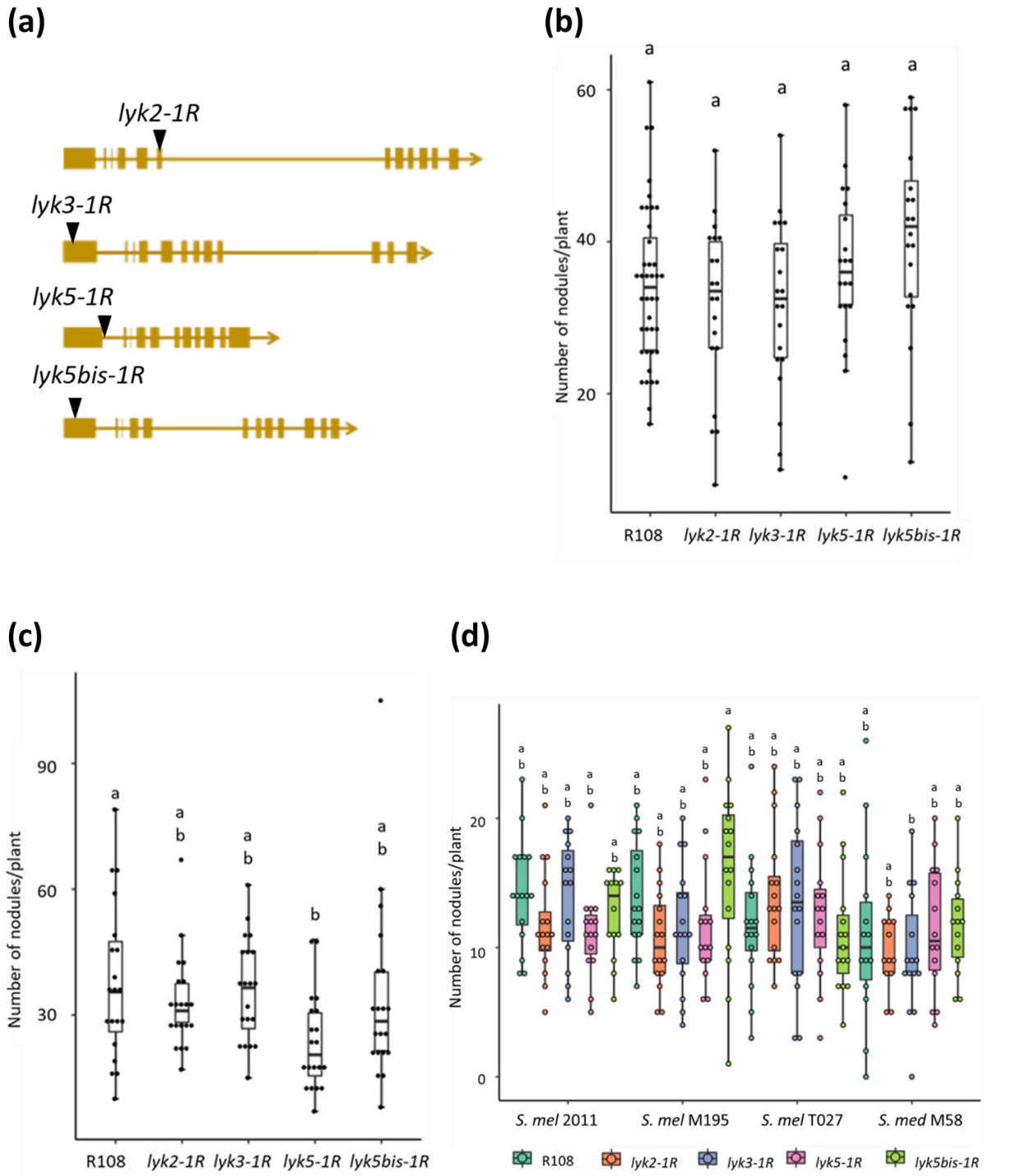

Fig. 2. Analysis of nodulation of mutants in selected R108 *LYK* cluster genes. (a) *Tnt1* insertional mutants in R108 *LYK2*, *LYK3*, *LYK5* and *LYK5bis*. The black arrows indicate the position of the identified insertion. Number of nodules at 21 d post inoculation with (b) *Sinorhizobium meliloti* 2011, (c) *S. meliloti* 2011 *nodF/nodL* of plants grown in pots and (d) Other natural *S. meliloti* or *S. medicae* strains of plants grown in tubes. At least 16 plants of each genotype/inoculum were analysed. Statistical analyses were done using ANOVA, Turkey ($P < 0.05$). Lowercase letters indicate statistical significance.

In [29], some natural *Sinorhizobium* strains were found to be dependent on *LYK2bis* for nodulation of R108, whereas *S. meliloti* 2011 and some other *S. meliloti* or *S. medicae* strains were found to be *LYK2bis*-independent. To determine whether the *LYK2bis*-independent rhizobial strains are dependent on other *LYK* genes in this cluster, we performed nodulation tests with these strains and the different mutant lines (Fig. **2d**, Table **S4)**. Each of the tested mutants, showed a similar nodulation capacity as R108 with each of the four *Sinorhizobium* strains, suggesting that none of the *LYK2*, *LYK3*, *LYK5* or *LYK5bis* genes is individually critical for nodulation of R108 by the tested strains.

3.4. Roles in nodulation of selected LYK cluster genes in A17

In A17, *LYK3* was reported not to be essential for early NF responses but plays an important role in infection and nodulation [15,17]. In addition, analysis of a weak mutant of *lyk3* with the *S. meliloti* 2011 *nodL* or *nodFE strains*, [17] found that *LYK3* is involved in controlling nodulation in a NF structure-dependent manner. Using the CRISPR-Cas9 gene editing system, we generated single *LYK2* and double *LYK5/LYK5bis* mutants and tested their nodulation phenotype with either *S. meliloti* 2011 WT or *nodL* strains (Table **S4**). Note that the *nodF/nodL* strain used on R108 does not nodulate A17 [7], so we used the weaker *nodL* mutant [17] on A17 and 2HA. For each mutant, one wildtype-like (WTL) sibling and two mutant lines were used for nodulation tests (Table **S2**). There was no significant difference between WTL and mutants in number of nodules per plant and nodule morphology following inoculation with *S. meliloti* 2011 WT strain (Fig. **3a, Fig. S5a**). When inoculated with *S. meliloti* 2011 *nodL*, both alleles of *lyk5/lyk5bis* double mutant showed significant reduction in number of nodules compared to WTL while *lyk2* ones did not (Fig. **3b**). However, no difference in overall nodule morphology was observed (Fig. **S5b)** and nodules on all lines were pink, suggesting expression of leghemoglobin and infection by rhizobia. In addition, the mutants were tested for root hair deformation (RHD) in response to *S. meliloti* 2011 *nodF/nodL* mutant: this mutant although unable to nodulate Jemalong can induce exaggerated RHD [7]. No difference in RHD index was observed in mutants compared to 2HA WT (Fig. **3c**), suggesting that none of the tested genes is individually essential for recognizing the *S. meliloti* 2011 *nodF/nodL* NFs.

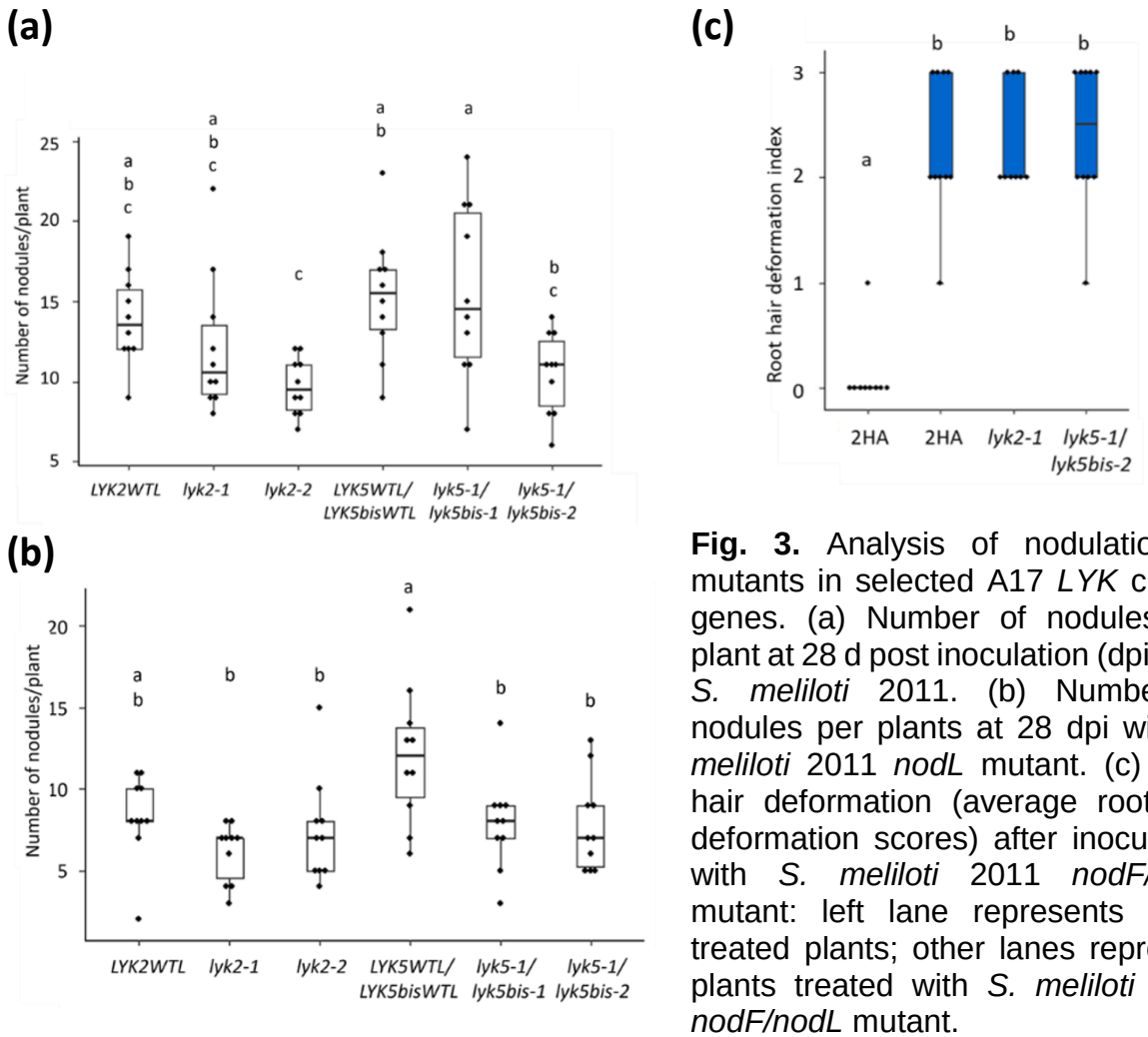

Fig. 3. Analysis of nodulation of mutants in selected A17 *LYK* cluster genes. (a) Number of nodules per plant at 28 d post inoculation (dpi) with *S. meliloti* 2011. (b) Number of nodules per plants at 28 dpi with *S. meliloti* 2011 *nodL* mutant. (c) Root hair deformation (average root hair deformation scores) after inoculation with *S. meliloti* 2011 *nodF/nodL* mutant: left lane represents mock treated plants; other lanes represent plants treated with *S. meliloti* 2011 *nodF/nodL* mutant.

In addition, the A17 *lyk3-1* mutant, which is Nod$^-$, was used in complementation assays using constructs overexpressing LYK3-R108, LYK2bis-R108, LYK2-R108, LYK2-A17 and LYK4-A17 for nodulation with *S. meliloti* 2011 (Fig. **4**). The empty vector (EV) and LYK3-A17 were used as controls. All the gene constructs tested, except LYK4-A17, formed nodules after inoculation with *S. meliloti*. This result suggests that the role of *LYK3* in nodulation with *S. meliloti* in A17 can be substituted by overexpressing *LYK2* and *LYK3* from either genotype. LYK2 and LYK3 are highly similar proteins, whereas LYK4 is almost identical to LYK3-A17 in the ECD but contains a LYK5-like kinase (Table **S6**, Fig. **S1**). This evidence suggests that it may be the divergent kinase of the chimeric LYK4-A17 which prevents it from substituting for LYK3-17 in nodulation of A17 with *S. meliloti* 2011 as kinase domain structure was previously reported to be important for plant complementation using a kinase domain swapping approach [46,47].

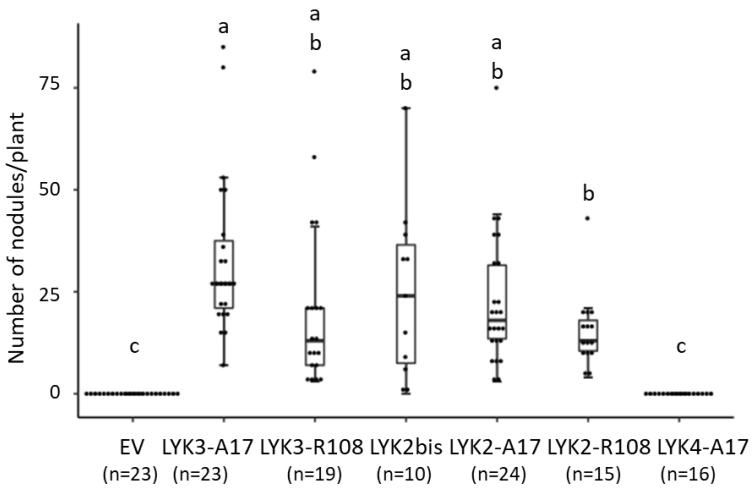

Fig. 4. *Complementation of lyk3-1 using different LYK genes from both A17 and R108.* The mutant roots were transformed with constructs of empty vector (EV), LYK3-A17, LYK3-R108, LYK2bis-R108, LYK2-A17, LYK2-R108 and LYK4-A17 using *A. rhizogenes*. Plants with transformed roots were analysed at 4 wk post inoculation. Statistical analyses were performed using ANOVA, Turkey (P<0.05). Lowercase letters indicate statistical significance. Numbers (n) below indicate number of transformed plants.

As genes in this cluster have a high potential for redundancy in nodulation, we also identified and characterized a fast neutron mutant in A17, called FN9805, which contains a large deletion from *LYK7* to *LYK3* (Table **S7,** Table **S8**). After inoculation with *S. meliloti* WT, the mutant showed a complete Nod$^-$ phenotype, which is expected as the single *lyk3* mutant is Nod$^-$ (Fig. **5a**). By overexpressing *LYK3*-A17, the nodulation in this mutant was only partially restored (Fig. **5b**). This result confirms the essential function of *LYK3* in nodulating A17 of *S. meliloti* but also suggests a role of genes in the deleted region in optimizing nodulation. Furthermore, the mutant was either treated with NF or inoculated with *S. meliloti* 2011 *nodF/nodL* to determine whether any early NF responses could be observed. In both cases, the mutant showed a similar level of RHD compared to A17 (Fig. **5c, d**). In addition, root samples after being treated with NF were collected and analysed for NF-induced gene expression. All tested genes including *ENOD11, PUB1, VAPYRIN* and *LYK10* were induced by NF at a similar level in both A17 and FN9805 (Fig. **S5**). These results indicate that early NF responses including RHD and gene expression are not dependent on *LYK3-LYK7* but must depend on other genes. Finally, the mutant was tested for colonization by the mycorrhizal fungus *Rhizophagus irregularis* at 2 and 4 wk post inoculation (wpi). A significant reduction in percentage of colonization in FN9805

compared to A17 was obtained at both time points suggesting that gene(s) in the deleted region is required for optimal AM (Fig. **5e**).

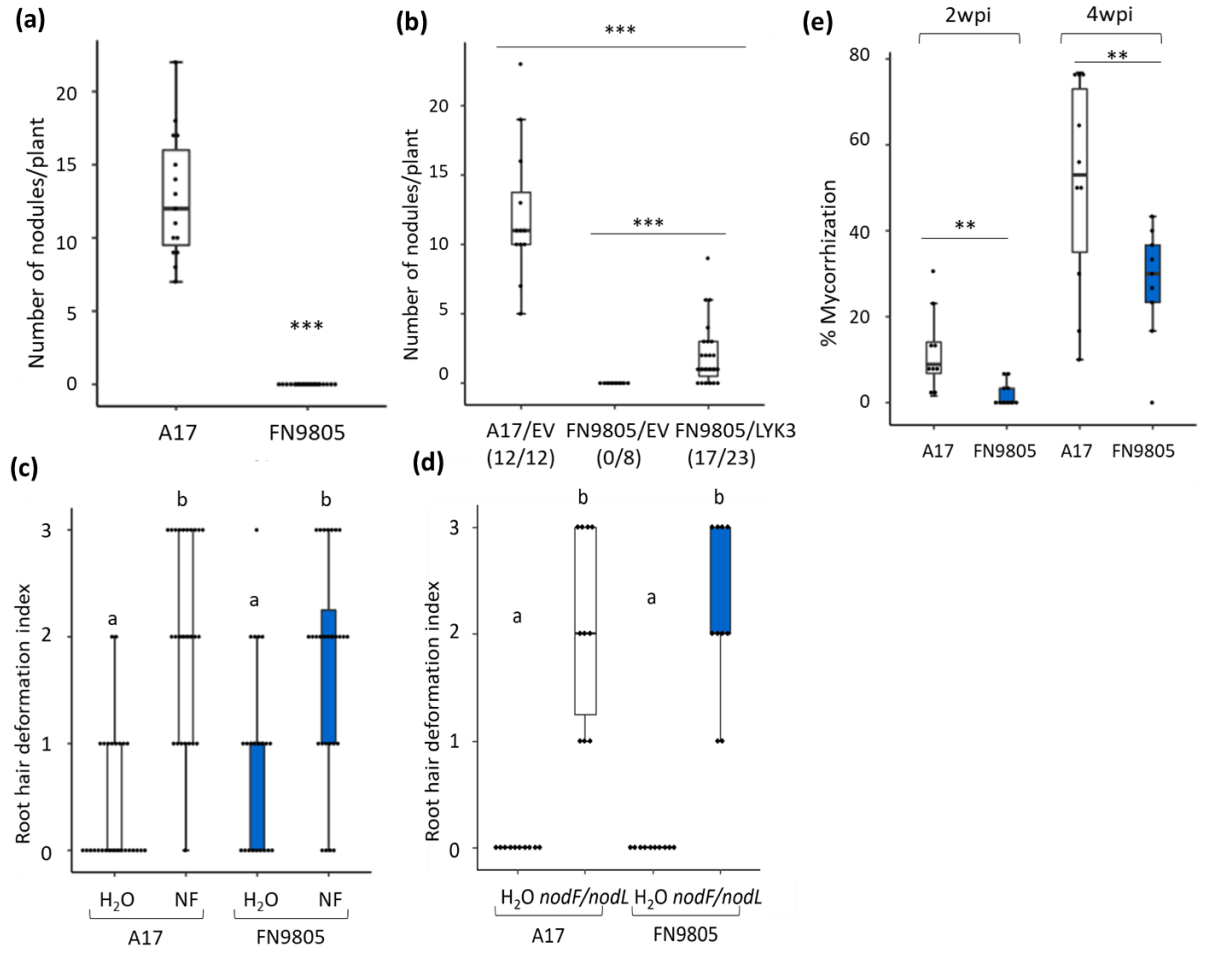

Fig. 5 Characterization of a fast neutron mutant of *LYK* cluster in A17. The FN9805 mutant is deleted from *LYK7* to *LYK3*. (a) Nodulation of FN9805 mutant with *Sinorhizobium meliloti* 2011 at 21 d post inoculation in tubes. Statistical analyses were done using Student's *t*-test (***, $P < 0.001$). (b) Complementation for nodulation of FN9805 mutant with *S. meliloti* 2011 using *Agrobacterium rhizogenes* transformation. Statistical analyses were done using Student's *t*-test (***, $P < 0.001$). Root hair deformation assay in FN9805 roots (average root hair deformation scores) either (c) treated with Nod factors or (d) inoculated with *S. meliloti* 2011 *nodF/nodL* mutant. Statistical analyses were done using ANOVA, Turkey ($P<0.05$). (e) Mycorrhizal phenotype (% of root length colonization) of FN9805 mutant at 2 wk and 4 wk after inoculation with *Rhizophagus irregularis*. Statistical analyses were done using Student's *t*-test (**, $P < 0.01$). wpi: week post inoculation.

3.5. Biochemical properties of selected LYK proteins

Recently, the ECR of LYK3-A17 was reported to bind NF [18,22] indicating the ability of a LYK protein to directly interact with a NF ligand. In this work, to study the binding ability of selected LYK proteins, we performed equilibrium binding assays in which a radiolabelled *S. meliloti* NF (LCO-IV (C16:2, S) was incubated with the membrane fraction of *N. benthamiana* leaves expressing the proteins of interest [21].

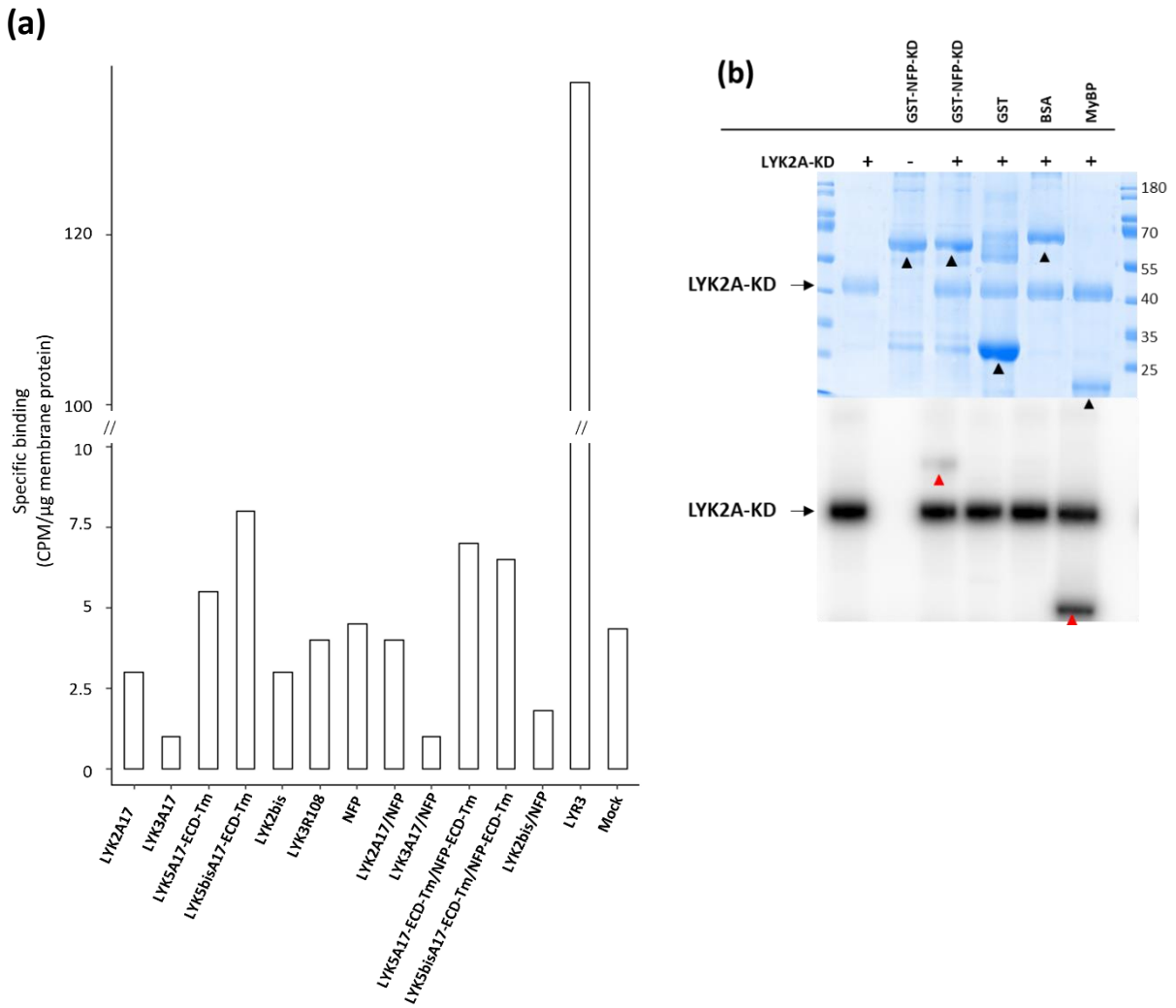

Fig. 6 Biochemical properties of selected LYK proteins. (a) Binding of *S. meliloti* Nod factors LCO-IV (C16:2Δ2,9,^{35}S) to membrane fractions containing indicated proteins using equilibrium binding assays. Specific binding was determined using membrane fraction extracts of *N. benthamiana* leaves expressing the indicated proteins. (b) Kinase activity of LYK2-A17. The ICR of LYK2-A17 (LYK2A-KD) was expressed in *E. coli* and purified. *In vitro* phosphorylation assays using radiolabelled ATP were performed. LYK2A-KD was incubated individually or co-incubated with GST-NFP-KD or other test proteins in the presence of [ɣ-^{32}P] ATP. Assays were analysed by

SDS-PAGE, followed by Coomassie staining (above) and phosphor imaging (below) The position of test proteins are marked on the Coomassie gel (black arrowheads). Trans-phosphorylated proteins are marked by red arrowheads on the phosphor-image.

As LYR3 was shown to be a high affinity LCO binding protein in the same type of assay, LYR3 and mock-treated leaf extracts were used as positive and negative controls, respectively. In some samples, NFP was co-expressed with the LYK protein. In contrast to LYR3, none of the tested protein extracts showed clear NF binding in this assay (Fig. **6a)**. We also tested the extracts with a LCO ligand (LCO-V (C18:1, NMe, S) which has been used to characterize receptors involved in AM [14]. Again, no clear binding was found to any of the LYK extracts, in comparison to LYR3 (Fig. **S7**).

NFP has a pseudokinase which was previously shown to be trans-phosphorylated by the active kinase domain (KD) of LYK3-A17 and LYK2bis [12,21,29]. In this study, we tested the kinase activity of LYK2-A17 using an *in vitro* phosphorylation assay. The ICR containing the KD of LYK2 (termed LYK2-KD) was purified after expression in *E. coli* and incubated with [γ-^{32}P] ATP in the absence or presence of other proteins. LYK2-KD showed a strong autophosphorylation activity and could trans-phosphorylate the GST fusion of NFP-KD and the model kinase substrate myelin basic protein (MyBP) but not GST and BSA (Fig. **6b**). In addition, the specificity of LYK2-KD was tested by co-incubating with LYR2-KD, LYR3-KD, LYR4-KD and LYK3-deadKD (G334E mutation) (Fig. **S8**). Although LYK2-KD poorly phosphorylated the LYR2-KD, it could clearly phosphorylate the other proteins. These results indicate that LYK2-A17 has an active KD and can trans-phosphorylate NFP, some other kinase domains and the model kinase substrate, MyBP.

4. Discussion

The LysM-RLK gene family is highly extended in legumes and plays important roles in controlling plant responses to both symbionts and pathogens [9]. In *M. truncatula* genotype A17, 22 LysM-RLKs have been found and can be phylogenetically divided into two subfamilies: LYR (with predicted inactive kinase) and LYK (with predicted active kinase) [8]. In this study, we have identified and characterized the *LYK* gene cluster located on Chromosome 5 in the two most commonly-used *M. truncatula* genotypes, A17

(derived from the cultivar Jemalong) and R108 (spp. *tricycla*). We have shown that the two clusters share many similarities but also some divergences (Fig. **1**). Functional analysis has been performed using various mutants in the *LYK* gene cluster of both genotypes in order to study the role of this variation for both nodulation and AM (Fig. **2, 3, 4, 5**).

4.1. Evolution and divergence of the LYK gene cluster

LysM-RLK genes in the legume *LYK* cluster are members of the LYK-I clade, which is represented by a single gene (*AtCERK1*) in Arabidopsis and two genes (*OsCERK1* and *OsRLK10*) in rice [9]. In *Parasponia* gene duplication led to two *LYK-I* genes, *PanLYK3* and *PanLYK1,* which is the ortholog of *MtLYK8* [10]. In legumes further rounds of duplication led to *MtLYK9* and the *LYK* gene cluster, which are on different chromosomes [9,10]. The presence of both closely-related genes in the *LYK* gene cluster of different legumes and species-specific differences in their number, suggest that the cluster has expanded by tandem duplication both before and after legume speciation. In *M. truncatula* we have shown that there is considerable variation in this cluster between different genotypes (Fig. **S3**). Although, both A17 and R108 contain complete copies of *LYK1, LYK2, LYK3, LYK5, LYK6* and *LYK7,* we show here that the region between *LYK2* and *LYK5* has undergone genotype-specific recombination events. The most parsimonious explanation for the differences is that ancestral *LYK5* and *LYK3* genes both duplicated to form a second functional copy of *LYK5* (*LYK5bis)* and separate *LYK2* and *LYK3* genes. These two duplications are ancestral to the divergence of the A17 and R108 genotypes (Fig. **7**). *LYK5* and *LYK5bis* both exist as complete genes in R108. In A17, insertion of a transposon in *LYK5bis* led to the truncation of the *LYK5bis* gene. Duplication of the ECR of *LYK3* and recombination with the ICR of *LYK5bis* then led to the formation of the chimeric *LYK4* gene. In R108 another duplication and recombination event led to the evolution of *LYK2bis*, which has a LYK2-like ECR and a LYK3-like ICR [29]. There is evidence that recombination and divergence has also affected *LYK3* in R108, as it contains a cluster of *LYK2*-like polymorphisms in LysM3, but these are not in the region (in LysM1) implicated in NF perception [18]. The gene is well expressed in R108, and the

transcript predicts a LYK3-R108 protein with 87/99% identity in the ECR/ICR respectively with LYK3-A17.

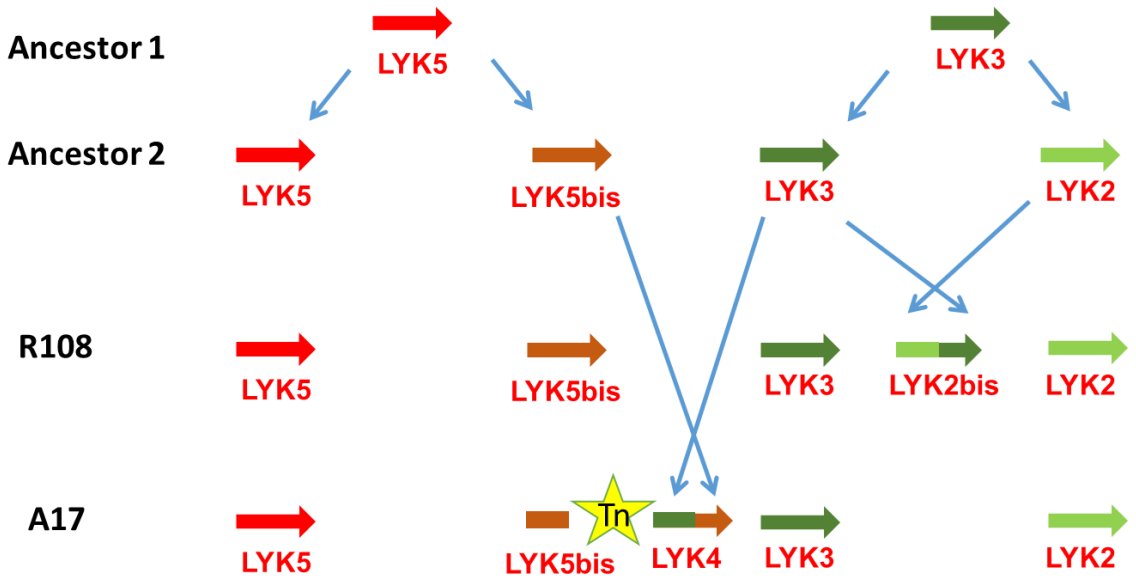

Fig. 7. Hypothetical scheme for evolution of part of the *LYK* cluster in A17 and R108. Colors represent the relatedness among genes. Tn represents a transposon. The scheme was hypothesized based on the structure of the cluster and the relatedness between the LysMs and KD regions of those proteins.

4.2. Roles of LYK-I genes in AM

In rice the two *LYK-I* genes, *OsCERK1* and *OsRLK10*, have been shown to play a role in AM [23,24,48]. It was initially reported that *LYK3* is required for optimal AM formation [24], however this has not been confirmed in more recent work from the same group [49]. In this study, we examined the FN9805 mutant, deleted from *LYK3* to *LYK7*, and observed a clear AM phenotype (Fig. **5e**). This result is compatible with a hypothesis that *LYK3* may play a redundant role in AM with genes in the cluster from *LYK4* to *LYK7* or with other genes missing due to this deletion (Table **S7**). Through comparative studies with the non-symbiotic *AtCERK1* gene it has been suggested that the presence of a YAQ motif in the kinase domain of LYK-I proteins is an indication of a symbiotic role: this motif has been lost in the non-symbiotic AtCERK1 [46]. In *M. truncatula* four LYK-I proteins contain the YAQ motif, LYK2, LYK3, LYK8, LYK9 [9]. In *M. truncatula,* and pea, *LYK9* plays a clear role in AM [32,50] whereas in the non-legume *Parasponia andersonii* a double mutant in *PanLYK3* and *PanLYK1* (the ortholog of *LYK8*) has a strong AM

phenotype [10]. It would be interesting to see whether double mutants of the *M. truncatula* FN9805 mutant and *lyk9* and/or *lyk8* produce a stronger AM phenotype.

4.3. Role of the LYK gene cluster in Nod factor responses and nodulation

In *M. truncatula* A17, it has been known for a long time that *LYK3* is essential for nodulation but is not required for early NF responses [12,15,17], whereas in *L. japonicus*, its ortholog *NFR1* is essential for both types of response [51]. An explanation is that there is redundancy in *M. truncatula LYK-I* gene(s) for early NF responses. RNA interference has suggested that *LYK4* could be involved in nodulation, although the data is not conclusive [30]. We have shown that *LYK4* is unable to complement the *lyk3-1* mutant (Fig. **4**), perhaps due to its divergent kinase domain (Table **S6**, Fig. **S1**), and that the FN9805 *LYK* cluster mutant (deleted in *LYK3-LYK7*) still shows early NF responses (Fig. **5,** Fig. **S5**) suggesting other *LYK-I* gene(s) are responsible for such redundancy. As *LYK2* shows an overall higher homology to *LYK3* (Table **S6**, Fig. **S2**), is able to substitute for *LYK3* for nodulation in the *lyk3-1* mutant (Fig. **4**), and is present in the FN9805 *LYK* cluster mutant, it is a better candidate for this role. Moreover, we have also shown that the kinase domain of LYK2 is active and can weakly trans-phosphorylate NFP, similarly to that of LYK3 [43]. Finally, the duplication that led to *LYK2* and *LYK3* is found in *M. truncatula* and not in *L. japonicus* [28] and might explain the difference in the regulation of early NF responses between these species.

We predicted that the ECR of LYK2 as it is very similar to that of LYK2bis of R108, could be involved in perceiving non-*O*-acetylated NF, produced by a *S. meliloti* 2011 *nodL* mutant [29]. Mutants of *LYK2* were tested for nodulation with *S. meliloti* 2011 *nodL* and showed no defect (Fig. **3c, d**), which did not confirm our hypothesis. However, as *lyk3* mutants show reduced nodulation with 2011 *nodL* [17], there may be redundancy with *LYK3* for this function. In addition, in A17, *LYK5bis* encodes a truncated LysM-RLK, lacking the kinase domain and has been found to be co-regulated with *NFP* during nodulation in different studies [52,53]. Mutation of *LYK5bis* (together with *LYK5*) did not lead to any observable nodulation phenotype with *S. meliloti* WT, but a reduction in nodulation, compared to the WTL line, was observed with the 2011 *nodL* mutant (Fig. **3**),

suggesting that *LYK5/LYK5bis* may be involved in improving the efficiency of nodulation between A17 and particular *Sinorhizobium* strains.

In R108, reverse genetics has revealed the essential role of *LYK2bis* in nodulation with *S. meliloti* 2011 *nodF/nodL* and some tested natural *S. meliloti* and *S. medicae* strains [29]. Despite LYK3-R108 showing about 95% identity to its orthologue in A17 and an expression pattern compatible with nodulation, it surprisingly does not show a key role in nodulation with any of the tested strains ([29], Fig. **2**). In addition, nodulation of R108 with *S. meliloti* 2011 and the natural strains which are independent of *LYK2bis* is shown here to not require either *LYK2*, *LYK3*, *LYK5* or *LYK5bis*, individually (Fig. **2**) suggesting that there may be redundancy in the function of these closely related genes stronger than in A17. Such stronger redundancy may be explained by the presence and high expression of *LYK2bis* in R108 compared to *LYK2* in A17 (Fig **S4**).

In this study, we have tested several LYK proteins for high affinity binding to lipochitooligosaccharides (LCOs)/NFs but did not detect any clear binding to these proteins (LYK2, LYK3, LYK4, LYK5, LYK5bis when expressed individually (Fig. **6**). Although these negative results cannot be conclusive on the ability of these proteins to bind LCOs, it contrasts to the high affinity LCO binding detected to the LYR-IIIA protein LYR3 [43] and the LYR-IA protein LYR1, but not NFP, (Cullimore et al. submitted) when using the same expression system and technique. We suggested that high affinity NF binding could require both NFP and another LysM protein (Luu et al), and LYK3 is an obvious candidate. Although co-expression of LYK3 or other LYK proteins with NFP did not lead to a clear increase in LCO binding, the results remain inconclusive due to co-expression inducing a cell death response in the *N. benthamiana* leaf expression system [20,21]. By using bio-layer interferometry assays on insect cell expressed and purified proteins, the LYK3 ECD [18] and the NFP-ECD [18,22] were shown individually to bind NF with an affinity in the µM range, which cannot be measured in our assays. Although speculating that NF perception may require heteromeric receptors, no studies were done on the two proteins together. Altogether, the LCO binding studies contribute to two possible hypotheses that either a low affinity NF binding is sufficient for activating NF responses or that a high affinity NF receptor complex requires two or more components (probably including NFP and a LYK)..

4.4. Conclusion

In conclusion, we have shown that the *LYK* cluster exhibits a high degree of variability amongst *M. truncatula* genotypes and has been subjected to recent recombination events within the cluster leading to the truncated *LYK5bis* gene and the chimeric *LYK4* gene in A17 and the chimeric *LYK2bis* gene in R108. The essential role of *LYK3* for nodulation of A17 [15,17] is not conserved in R108, where the *LYK2bis* gene is essential for nodulation of R108 with certain rhizobial strains [29]. No such essential roles in nodulation were found for *LYK2*, *LYK5* and *LYK5bis* although the similarity between some of the *LYK* cluster genes suggests that they might be mutually redundant for such roles. Redundancy with other *LYK-I* class genes, such as *LYK9* and *LYK8*, is also a possibility, particularly for AM. Fast neutron induced large deletions can permit the analysis of redundancy between clustered genes but depends on the fortuitous deletion of the region to be studied. Unfortunately, the deletion identified in the A17 *LYK* cluster in the FN9805 mutant did not extend to *LYK2*. As found previously in legumes [54,55], we have shown that CRISPR-Cas9 is an alternative and efficient way of creating mutants in genes of interest and we have successfully used it to produce single mutants in *LYK2* and double mutants in *LYK5/LYK5bis*. It should be possible to make mutants in multiple genes of the *LYK* cluster using this technique, which would not be possible by crossing of single mutants due to the close gene linkage. However, crossing provides a viable strategy for creating mutants with other LYK-I class genes such as *LYK8* and *LYK9.*

Funding

Funding for part of our work was gratefully received from the Fédération de Recherche Agrobiosciences, Interactions et Biodiversité (FR AIB – project CHAIN, coordinators J. Cullimore and C. Jacquet) and the Agence Nationale de la Recherche (ANR – project DUALITY, ANR-20-CE20-0017-01, coordinator C. Gough). This study is set within the framework of the "Laboratoires d'Excellences (LABEX)" TULIP (ANR 10 LABX 41) and of the "École Universitaire de Recherche (EUR)" TULIP GS (ANR 18 EURE 0019). T-B.L. gratefully acknowledges receipt of a Bourse d'Excellence from the Ambassade de France au Vietnam, to fund her PhD in France.

Credit authorship contribution statement

Conceptualization, T-B.L., J-J.B, B.L., N.P., J.C.; Investigation T-B.L., N.C., L.B., C.G-V., C.R., V.G., J.C.; Roles/Writing – original draft, T-B.L, J.C.; Writing -review & editing T-B.L., N.P., B.L., J.C.; Visualization, T-B.L.; Supervision, J-J.B., B.L., N.P., J.C. All authors have read and agreed to the published version of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and materials will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at

Acknowledgements

We thank Clare Gough for critical reading of the manuscript and advice during the project; Fabienne Maillet for NFs, protocols and advice, Jerome Gouzy for analysis of transcriptomic data, Judith Fliegmann and Nikita Malkov for help with LysM-RLK constructs, Marie-Françoise Jardinaud and Yann Pecrix for advice on CRISPR-Cas9 work. For the *Tnt1* mutants, the *Medicago truncatula* plants utilized in this research project, which are jointly owned by the Centre National de la Recherche Scientifique, were obtained, with the fast neutron mutant, from the Noble Research Institute, and were created through research funded, in part, by a grant from National Science Foundation, NSF-0703285.

References

[1] P.H. Graham, C.P. Vance, Update on Legume Utilization Legumes: Importance and Constraints to Greater Use, Plant Physiol. 131 (2003) 872–877. https://doi.org/10.1104/pp.017004.

[2] M. Parniske, Arbuscular mycorrhiza: the mother of plant root endosymbioses, Nature Reviews Microbiology 6 (2008) 763–775. https://doi.org/10.1038/nrmicro1987.

[3] R. van Velzen, R. Holmer, F. Bu, L. Rutten, A. van Zeijl, W. Liu, L. Santuari, Q. Cao, T. Sharma, D. Shen, Y. Roswanjaya, T.A.K. Wardhani, M.S. Kalhor, J. Jansen, van den H. Johan, B. Güngör, M. Hartog, J. Hontelez, J. Verver, W.C. Yang, E. Schijlen, R. Repin, M. Schilthuizen, M.E. Schranz, R. Heidstra, K. Miyata, E. Fedorova, W. Kohlen, T. Bisseling, S. Smit, R. Geurts, Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses, Proc Natl Acad Sci U S A. 115 (2018) E4700–E4709.

[4] M. Griesmann, Y. Chang, X. Liu, Y. Song, G. Haberer, M.B. Crook, B. Billault-Penneteau, D. Lauressergues, J. Keller, L. Imanishi, Y.P. Roswanjaya, W. Kohlen, P. Pujic, K. Battenberg, N. Alloisio, Y. Liang, H. Hilhorst, M.G. Salgado, V. Hocher, H. Gherbi, S. Svistoonoff, J.J. Doyle, S. He, Y. Xu, S. Xu, J. Qu, Q. Gao, X. Fang, Y. Fu, P. Normand, A.M. Berry, L.G. Wall, J.M. Ané, K. Pawlowski, X. Xu, H. Yang, M. Spannagl, K.F.X. Mayer, G.K.S. Wong, M. Parniske, P.M. Delaux, S. Cheng, Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis, Science 361 (2018) eaat1743.

[5] M. Andrews, M.E. Andrews, Specificity in Legume-Rhizobia Symbioses, Int J Mol Sci. 18 (2017) 705–744. https://doi.org/10.3390/IJMS18040705.

[6] W. D'Haeze, M. Holsters, Nod factor structures, responses, and perception during initiation of nodule development, Glycobiology. 12 (2002) 79R-105R. https://doi.org/10.1093/GLYCOB/12.6.79R.

[7] M. Ardourel, N. Demont, F. Debelle, F. Maillet, F. de Billy, J.C. Prome, J. Denarie, G. Truchet, *Rhizobium meliloti* lipooligosaccharide nodulation factors: Different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses, Plant Cell. 6 (1994) 1357–1374. https://doi.org/10.1105/tpc.6.10.1357.

[8] J. Bono, J. Fliegmann, C. Gough, J. Cullimore, Expression and function of the *Medicago truncatula lysin motif receptor-like kinase* (LysM - RLK) gene family in the legume–rhizobia symbiosis, in: The model legume *Medicago truncatula*, Wiley, 2020: pp. 439–447. https://doi.org/10.1002/9781119409144.ch55.

[9] L. Buendia, A. Girardin, T. Wang, L. Cottret, B. Lefebvre, LysM receptor-like kinase and lysM receptor-like protein families: An update on phylogeny and functional

characterization, Front Plant Sci. 9 (2018) 1531. https://doi.org/10.3389/fpls.2018.01531.

[10] L. Rutten, K. Miyata, Y.P. Roswanjaya, R. Huisman, F. Bu, M. Hartog, S. Linders, R. van Velzen, A. van Zeijl, T. Bisseling, W. Kohlen, R. Geurts, Duplication of symbiotic lysin motif receptors predates the evolution of nitrogen-fixing nodule symbiosis, Plant Physiol. 184 (2020) 1004–1023. https://doi.org/10.1104/PP.19.01420.

[11] C. Krönauer, S. Radutoiu, Understanding Nod factor signalling paves the way for targeted engineering in legumes and non-legumes, Curr Opin Plant Biol. 62 (2021) 102026. https://doi.org/10.1016/J.PBI.2021.102026.

[12] J.F. Arrighi, A. Barre, B. Ben Amor, A. Bersoult, L.C. Soriano, R. Mirabella, F. De Carvalho-Niebel, E.P. Journet, M. Ghérardi, T. Huguet, R. Geurts, J. Dénarié, P. Rougé, C. Gough, The *Medicago truncatula* lysine motif-receptor-like kinase gene family includes *NFP* and new nodule-expressed genes, Plant Physiol. 142 (2006) 265–279. https://doi.org/10.1104/pp.106.084657.

[13] S. Bensmihen, F. de Billy, C. Gough, Contribution of NFP LysM domains to the recognition of Nod Factors during the *Medicago truncatula/Sinorhizobium meliloti* symbiosis, PLoS One. 6 (2011) e26114. https://doi.org/10.1371/JOURNAL.PONE.0026114.

[14] A. Girardin, T. Wang, Y. Ding, J. Keller, L. Buendia, M. Gaston, C. Ribeyre, V. Gasciolli, M.-C. Auriac, T. Vernié, A. Bendahmane, M. Katharina Ried, M. Parniske, P. Morel, M. Vandenbussche, M. Schorderet, D. Reinhardt, P.-M. Delaux, J.-J. Bono, B. Lefebvre, LCO receptors involved in arbuscular mycorrhiza are functional for rhizobia perception in legumes, Current Biology. 29 (2019) 4249–4259. https://doi.org/10.1016/j.cub.2019.11.038.

[15] R. Catoira, A.C.J. Timmers, F. Maillet, C. Galera, R. v Penmetsa, D. Cook, J. Dénarié, C. Gough, The *HCL* gene *Medicago truncatula* controls *Rhizobium*-induced root hair curling, Development. 128 (2001) 1507–1518.

[16] R.M., Mitra, S.L., Shaw, S.R. Long, Six nonnodulating plant mutants defective for Nod factor-induced transcriptional changes associated with the Legume-Rhizobia Symbiosis, Proc Natl Acad Sci U S A. (2004) 10217–10222.

[17] P. Smit, E. Limpens, R. Geurts, E. Fedorova, E. Dolgikh, C. Gough, T. Bisseling, Medicago LYK3, an entry receptor in rhizobial nodulation factor signaling, Plant Physiol. 145 (2007) 183–191. https://doi.org/10.1104/pp.107.100495.

[18] Z. Bozsoki, K. Gysel, S.B. Hansen, D. Lironi, C. Krönauer, F. Feng, N. de Jong, M. Vinther, M. Kamble, M.B. Thygesen, E. Engholm, C. Kofoed, S. Fort, J.T. Sullivan, C.W. Ronson, K.J. Jensen, M. Blaise, G. Oldroyd, J. Stougaard, K.R. Andersen, S. Radutoiu, Ligand-recognizing motifs in plant LysM receptors are major

determinants of specificity., Science. 369 (2020) 663–670. https://doi.org/10.1126/science.abb3377.

[19] S. Moling, A. Pietraszewska-Bogiel, M. Postma, E. Fedorova, M.A. Hink, E. Limpens, T.W.J. Gadell, T. Bisseling, Nod Factor Receptors form heteromeric complexes and are essential for intracellular infection in *Medicago* nodules, Plant Cell. 26 (2014) 4188–4199. https://doi.org/10.1105/TPC.114.129502.

[20] A. Pietraszewska-Bogiel, B. Lefebvre, M.A. Koini, D. Klaus-Heisen, F.L.W. Takken, R. Geurts, J. V. Cullimore, T.W.J. Gadella, Interaction of *Medicago truncatula* Lysin Motif Receptor-Like Kinases, NFP and LYK3, produced in *Nicotiana benthamiana* induces defence-like responses, PLoS One. 8 (2013) e65055. https://doi.org/10.1371/journal.pone.0065055.

[21] J. Fliegmann, A. Jauneau, C. Pichereaux, C. Rosenberg, V. Gasciolli, A.C.J. Timmers, O. Burlet-Schiltz, J. Cullimore, J.J. Bono, LYR3, a high-affinity LCO-binding protein of *Medicago truncatula*, interacts with LYK3, a key symbiotic receptor, FEBS Lett. 590 (2016) 1477–1487. https://doi.org/10.1002/1873-3468.12191.

[22] K. Gysel, M. Laursen, M.B. Thygesen, D. Lironi, Z. Bozsóki, C.T. Hjuler, N.N. Maolanon, J. Cheng, P.K. Bjørk, M. Vinther, L.H. Madsen, H. Rübsam, A. Muszyński, A. Ghodrati, P. Azadi, J.T. Sullivan, C.W. Ronson, K.J. Jensen, M. Blaise, S. Radutoiu, J. Stougaard, K.R. Andersen, Kinetic proofreading of lipochitooligosaccharides determines signal activation of symbiotic plant receptors, Proc Natl Acad Sci U S A. 118 (2021) e2111031118. https://doi.org/10.1073/PNAS.2111031118.

[23] K. Miyata, T. Kozaki, Y. Kouzai, K. Ozawa, K. Ishii, E. Asamizu, Y. Okabe, Y. Umehara, A. Miyamoto, Y. Kobae, K. Akiyama, H. Kaku, Y. Nishizawa, N. Shibuya, T. Nakagawa, The bifunctional plant receptor, OsCERK1, regulates both chitin-triggered immunity and arbuscular mycorrhizal symbiosis in rice, Plant Cell Physiol. 55 (2014) 1864–1872. https://doi.org/10.1093/PCP/PCU129.

[24] X. Zhang, W. Dong, J. Sun, F. Feng, Y. Deng, Z. He, G.E.D. Oldroyd, E. Wang, The receptor kinase CERK1 has dual functions in symbiosis and immunity signalling, Plant Journal. 81 (2015) 258–267. https://doi.org/10.1111/tpj.12723.

[25] G. Carotenuto, M. Chabaud, K. Miyata, M. Capozzi, N. Takeda, H. Kaku, N. Shibuya, T. Nakagawa, D.G. Barker, A. Genre, The rice LysM receptor-like kinase OsCERK1 is required for the perception of short-chain chitin oligomers in arbuscular mycorrhizal signaling, New Phytologist. 214 (2017) 1440–1446. https://doi.org/10.1111/NPH.14539.

[26] P.M. Delaux, K. Varala, P.P. Edger, G.M. Coruzzi, J.C. Pires, J.M. Ané, Comparative Phylogenomics Uncovers the Impact of Symbiotic Associations on

Host Genome Evolution, PLoS Genet. 10 (2014). https://doi.org/10.1371/JOURNAL.PGEN.1004487.

[27] C. Yang, E. Wang, J. Liu, CERK1, more than a co-receptor in plant–microbe interactions, New Phytologist. 234 (2022) 1606–1613. https://doi.org/10.1111/NPH.18074.

[28] S. De Mita, A. Streng, T. Bisseling, R. Geurts, Evolution of a symbiotic receptor through gene duplications in the legume-rhizobium mutualism, New Phytologist. 201 (2014) 961–972. https://doi.org/10.1111/nph.12549.

[29] T. Luu, A. Ourth, C. Pouzet, N. Pauly, J. Cullimore, A newly-evolved chimeric lysin motif receptor-like kinase in *Medicago truncatula* spp. *tricycla* R108 extends its Rhizobia symbiotic partnership, New Phytologist. (2022). https://doi.org/10.1111/NPH.18270.

[30] E. Limpens, C. Franken, P. Smit, J. Willemse, T. Bisseling, R. Geurts, LysM Domain Receptor Kinases regulating rhizobial nod factor-induced infection, Science (1979). 302 (2003) 630–633. https://doi.org/10.1126/science.1090074.

[31] Z. Bozsoki, J. Cheng, F. Feng, K. Gysel, M. Vinther, K.R. Andersen, G. Oldroyd, M. Blaise, S. Radutoiu, J. Stougaard, Receptor-mediated chitin perception in legume roots is functionally separable from Nod factor perception, Proc Natl Acad Sci U S A. 114 (2017) E8118–E8127. https://doi.org/10.1073/pnas.1706795114.

[32] C. Gibelin-Viala, E. Amblard, V. Puech-Pages, M. Bonhomme, M. Garcia, A. Bascaules-Bedin, J. Fliegmann, J. Wen, K.S. Mysore, C. le Signor, C. Jacquet, C. Gough, The *Medicago truncatula* LysM receptor-like kinase *LYK9* plays a dual role in immunity and the arbuscular mycorrhizal symbiosis, New Phytologist. 223 (2019) 1516–1529. https://doi.org/10.1111/nph.15891.

[33] N.D. Young, F. Debellé, G.E.D. Oldroyd, R. Geurts, S.B. Cannon, M.K. Udvardi, V.A. Benedito, K.F.X. Mayer, J. Gouzy, H. Schoof, Y. Van De Peer, S. Proost, D.R. Cook, B.C. Meyers, M. Spannagl, F. Cheung, S. De Mita, V. Krishnakumar, H. Gundlach, S. Zhou, J. Mudge, A.K. Bharti, J.D. Murray, M.A. Naoumkina, B. Rosen, K.A.T. Silverstein, H. Tang, S. Rombauts, P.X. Zhao, P. Zhou, V. Barbe, P. Bardou, M. Bechner, A. Bellec, A. Berger, H. Bergès, S. Bidwell, T. Bisseling, N. Choisne, A. Couloux, R. Denny, S. Deshpande, X. Dai, J.J. Doyle, A.M. Dudez, A.D. Farmer, S. Fouteau, C. Franken, C. Gibelin, J. Gish, S. Goldstein, A.J. González, P.J. Green, A. Hallab, M. Hartog, A. Hua, S.J. Humphray, D.H. Jeong, Y. Jing, A. Jöcker, S.M. Kenton, D.J. Kim, K. Klee, H. Lai, C. Lang, S. Lin, S.L. MacMil, G. Magdelenat, L. Matthews, J. McCorrison, E.L. Monaghan, J.H. Mun, F.Z. Najar, C. Nicholson, C. Noirot, M. O'Bleness, C.R. Paule, J. Poulain, F. Prion, B. Qin, C. Qu, E.F. Retzel, C. Riddle, E. Sallet, S. Samain, N. Samson, I. Sanders, O. Saurat, C. Scarpelli, T. Schiex, B. Segurens, A.J. Severin, D.J. Sherrier, R. Shi, S. Sims, S.R. Singer, S. Sinharoy, L. Sterck, A. Viollet, B.B. Wang, K. Wang, M. Wang, X. Wang, J.

Warfsmann, J. Weissenbach, D.D. White, J.D. White, G.B. Wiley, P. Wincker, Y. Xing, L. Yang, Z. Yao, F. Ying, J. Zhai, L. Zhou, A. Zuber, J. Dénarié, R.A. Dixon, G.D. May, D.C. Schwartz, J. Rogers, F. Quétier, C.D. Town, B.A. Roe, The *Medicago* genome provides insight into the evolution of rhizobial symbioses, Nature 2011 480:7378. 480 (2011) 520–524. https://doi.org/10.1038/nature10625.

[34] M. Garmier, L. Gentzbittel, J. Wen, K.S. Mysore, P. Ratet, *Medicago truncatula*: Genetic and genomic resources, Curr Protoc Plant Biol. 2 (2017) 318–349. https://doi.org/10.1002/CPPB.20058.

[35] Y. Pecrix, S.E. Staton, E. Sallet, C. Lelandais-Brière, S. Moreau, S. Carrère, T. Blein, M.-F. Jardinaud, D. Latrasse, M. Zouine, M. Zahm, J. Kreplak, B. Mayjonade, C. Satgé, M. Perez, S. Cauet, W. Marande, C. Chantry-Darmon, C. Lopez-Roques, O. Bouchez, A. Bérard, F. Debellé, S. Muños, A. Bendahmane, H. Bergès, A. Niebel, J. Buitink, F. Frugier, M. Benhamed, M. Crespi, J. Gouzy, P. Gamas, Whole-genome landscape of *Medicago truncatula* symbiotic genes, Nature Plants. 4 (2018) 1017–1025. https://doi.org/10.1038/s41477-018-0286-7.

[36] P. Kaur, C. Lui, O. Dudchenko, R.S. Nandety, B. Hurgobin, M. Pham, E.L. Aiden, J. Wen, K.S. Mysore, Delineating the *Tnt1* insertion landscape of the model legume *Medicago truncatula* cv. R108 at the Hi-C resolution using a chromosome-length genome assembly, Int J Mol Sci. 22 (2021) 4326. https://doi.org/10.3390/IJMS22094326.

[37] P. Zhou, K.A.T. Silverstein, T. Ramaraj, J. Guhlin, R. Denny, J. Liu, A.D. Farmer, K.P. Steele, R.M. Stupar, J.R. Miller, P. Tiffin, J. Mudge, N.D. Young, Exploring structural variation and gene family architecture with *De Novo* assemblies of 15 *Medicago* genomes, BMC Genomics. 18 (2017) 261. https://doi.org/10.1186/S12864-017-3654-1.

[38] K. Xie, B. Minkenberg, Y. Yang, Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system, Proc Natl Acad Sci U S A. 112 (2015) 3570–3575. https://doi.org/10.1073/PNAS.1420294112.

[39] M.F. Jardinaud, J. Fromentin, M.C. Auriac, S. Moreau, Y. Pecrix, L. Taconnat, L. Cottret, G. Aubert, S. Balzergue, J. Burstin, S. Carrere, P. Gamas, MtEFD and MtEFD2: Two transcription factors with distinct neofunctionalization in symbiotic nodule development, Plant Physiol. 189 (2022) 1587–1607. https://doi.org/10.1093/PLPHYS/KIAC177.

[40] M. Chabaud, F. De Carvalho-Niebel, D.G. Barker, Efficient transformation of *Medicago truncatula* cv. Jemalong using the hypervirulent *Agrobacterium tumefaciens* strain AGL1, Plant Cell Rep. 22 (2003) 46–51. https://doi.org/10.1007/S00299-003-0649-Y.

[41] A. Boisson-Dernier, M. Chabaud, F. Garcia, G. Bécard, C. Rosenberg, D.G Barker, *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study

of nitrogen-fixing and endomycorrhizal symbiotic associations, Mol Plant Microbe Interact. 14 (2001) 695–700. https://doi.org/10.1094/MPMI.2001.14.6.695.

[42] V. Herrbach, X. Chirinos, D. Rengel, K. Agbevenou, R. Vincent, S. Pateyron, S. Huguet, S. Balzergue, A. Pasha, N. Provart, C. Gough, S. Bensmihen, Nod factors potentiate auxin signaling for transcriptional regulation and lateral root formation in *Medicago truncatula*, J Exp Bot. 68 (2017) 569–583. https://doi.org/10.1093/JXB/ERW474.

[43] J. Fliegmann, S. Canova, C. Lachaud, S. Uhlenbroich, V. Gasciolli, C. Pichereaux, M. Rossignol, C. Rosenberg, M. Cumener, D. Pitorre, B. Lefebvre, C. Gough, E. Samain, S. Fort, H. Driguez, B. Vauzeilles, J.M. Beau, A. Nurisso, A. Imberty, J. Cullimore, J.J. Bono, Lipo-chitooligosaccharidic symbiotic signals are recognized by LysM receptor-like kinase LYR3 in the legume *Medicago truncatula*, ACS Chem Biol. 8 (2013) 1900–1906. https://doi.org/10.1021/cb400369u.

[44] K. Schiessl, J.L.S. Lilley, T. Lee, I. Tamvakis, W. Kohlen, P.C. Bailey, A. Thomas, J. Luptak, K. Ramakrishnan, M.D. Carpenter, K.S. Mysore, J. Wen, S. Ahnert, V.A. Grieneisen, G.E.D. Oldroyd, NODULE INCEPTION recruits the lateral root developmental program for symbiotic nodule organogenesis in *Medicago truncatula*, Current Biology. 29 (2019) 3657-3668.e5. https://doi.org/10.1016/j.cub.2019.09.005.

[45] F. Maillet, J. Fournier, H.C. Mendis, M. Tadege, J. Wen, P. Ratet, K.S. Mysore, C. Gough, K.M. Jones, *Sinorhizobium meliloti* succinylated high-molecular-weight succinoglycan and the *Medicago truncatula* LysM receptor-like kinase MtLYK10 participate independently in symbiotic infection, The Plant Journal. 102 (2020) 311–326. https://doi.org/10.1111/tpj.14625.

[46] T. Nakagawa, H. Kaku, Y. Shimoda, A. Sugiyama, M. Shimamura, K. Takanashi, K. Yazaki, T. Aoki, N. Shibuya, H. Kouchi, From defense to symbiosis: Limited alterations in the kinase domain of LysM receptor-like kinases are crucial for evolution of legume–Rhizobium symbiosis, The Plant Journal. 65 (2011) 169–180. https://doi.org/10.1111/J.1365-313X.2010.04411.X.

[47] E. Murakami, J. Cheng, K. Gysel, Z. Bozsoki, Y. Kawaharada, C.T. Hjuler, K.K. Sørensen, K. Tao, S. Kelly, F. Venice, A. Genre, M.B. Thygesen, N. de Jong, M. Vinther, D.B. Jensen, K.J. Jensen, M. Blaise, L.H. Madsen, K.R. Andersen, J. Stougaard, S. Radutoiu, Epidermal LysM receptor ensures robust symbiotic signalling in *Lotus japonicus*, Elife. 7 (2018). https://doi.org/10.7554/eLife.33506.

[48] K. Miyata, S. Hasegawa, E. Nakajima, Y. Nishizawa, K. Kamiya, H. Yokogawa, S. Shirasaka, S. Maruyama, N. Shibuya, H. Kaku, OsCERK2/OsRLK10, a homolog of OsCERK1, has a potential role for chitin-triggered immunity and arbuscular mycorrhizal symbiosis in rice, Plant Biotechnology. 39 (2022) 119. https://doi.org/10.5511/PLANTBIOTECHNOLOGY.21.1222A.

[49] F. Feng, J. Sun, G. v. Radhakrishnan, T. Lee, Z. Bozsóki, S. Fort, A. Gavrin, K. Gysel, M.B. Thygesen, K.R. Andersen, S. Radutoiu, J. Stougaard, G.E.D. Oldroyd, A combination of chitooligosaccharide and lipochitooligosaccharide recognition promotes arbuscular mycorrhizal associations in *Medicago truncatula*, Nat Commun. 10 (2019) 1–12. https://doi.org/10.1038/s41467-019-12999-5.

[50] I. v. Leppyanen, V.Y. Shakhnazarova, O.Y. Shtark, N.A. Vishnevskaya, I.A. Tikhonovich, E.A. Dolgikh, Receptor-Like Kinase LYK9 in *Pisum sativum L.* Is the CERK1-Like receptor that controls both plant immunity and AM symbiosis development, Int J Mol Sci. 19 (2017) 8. https://doi.org/10.3390/IJMS19010008.

[51] S. Radutoiu, L.H. Madsen, E.B. Madsen, H.H. Felle, Y. Umehara, M. Grønlund, S. Sato, Y. Nakamura, S. Tabata, N. Sandal, J. Stougaard, Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases, Nature. 425 (2003) 585–592. https://doi.org/10.1038/nature02039.

[52] E. Larrainzar, B.K. Riely, S.C. Kim, N. Carrasquilla-Garcia, H.J. Yu, H.J. Hwang, M. Oh, G.B. Kim, A.K. Surendrarao, D. Chasman, A.F. Siahpirani, R. V. Penmetsa, G.S. Lee, N. Kim, S. Roy, J.H. Mun, D.R. Cook, Deep sequencing of the *Medicago truncatula* root transcriptome reveals a massive and early interaction between nodulation factor and ethylene signals, Plant Physiol. 169 (2015) 233–265. https://doi.org/10.1104/pp.15.00350.

[53] P. Gautrat, V. Mortier, C. Laffont, A. De Keyser, J. Fromentin, F. Frugier, S. Goormachtig, Unraveling new molecular players involved in the autoregulation of nodulation in *Medicago truncatula*, J Exp Bot. 70 (2019) 1407–1417. https://doi.org/10.1093/jxb/ery465.

[54] L. Wang, M.C. Rubio, X. Xin, B. Zhang, Q. Fan, Q. Wang, G. Ning, M. Becana, D. Duanmu, CRISPR/Cas9 knockout of leghemoglobin genes in *Lotus japonicus* uncovers their synergistic roles in symbiotic nitrogen fixation, New Phytol. 224 (2019) 818–832. https://doi.org/10.1111/NPH.16077.

[55] A. Bao, L.S.P. Tran, D. Cao, CRISPR/Cas9-based gene editing in soybean, Methods in Molecular Biology. 2107 (2020) 349–364. https://doi.org/10.1007/978-1-0716-0235-5_19/COVER.

Appendix A: List of Supporting information

Table S1 List of primers used in this study

Table S2 *Medicago truncatula* genotypes used in this study

Table S3 List of constructs used in this study

Table S4 Rhizobial strains used in this study

Table S5 Genes in *LYK* cluster in A17

Table S6 Pairwise identities of the whole proteins, ECD and ICD of selected LYK cluster proteins in A17 and R108

Table S7 Summary of genotyping of FN9805 mutant

Table S8 List of primers used for genotyping FN9805

Fig. S1 Amino acid alignment between LYK2-A17, LYK3-A17, LYK4, LYK5-A17, LYK5bis-A17, LYK5-R108 and LYK5bis-R108

Fig. S2 Phylogenetic tree of LYK proteins from A17 and R108

Fig. S3 Extent of the LYK cluster in A17 and R108 and structure of the *LYK* cluster of different *M. truncatula* genotypes

Fig. S4 Expression of the *LYK* cluster and other selected genes in the susceptible region of roots of Jemalong or R108, 24h post spot inoculation with *S. meliloti* or mock

Fig. S5 Nodule morphology of *lyk2* and *lyk5/lyk5bis* mutants after inoculating with (a) *S. meliloti* 2011 WT and (b) *S. meliloti* 2011 *nodF/nodL* mutant.

Fig. S6 Early NF-induced gene expression in FN9805 mutant treated with either H_2O or Nod factors (NFs) after 24h

Fig. S7 Binding of LCO-V (C18:1Δ11, NMe, ^{35}S) to membrane fractions containing indicated proteins using equilibrium binding assays.

Fig. S8 Trans-phosphorylation specificity of LYK2-A17 ICR (LYK2A-KD)