Evolution of Lipochitooligosaccharide Binding to a LysM-RLK for Nodulation in Medicago truncatula

Julie Cullimore, Judith Fliegmann, Virginie Gasciolli, Chrystel Gibelin-Viala, Noémie Carles, Thi-bich Luu, Ariane Girardin, Marie Cumener, Fabienne Maillet, Stéphanie Pradeau, et al.

Evolution of lipochitooligosaccharide binding to a LysM-RLK for nodulation in *Medicago truncatula*

Julie Cullimore[a], Judith Fliegmann[a,1], Virginie Gasciolli[a], Chrystel Gibelin-Viala[a], Noémie Carles[a], Thi-Bich Luu[a,2], Ariane Girardin[a], Marie Cumener[a], Fabienne Maillet[a], Stéphanie Pradeau[b], Sébastien Fort[b], Jean-Jacques Bono[a], Clare Gough[a], Benoit Lefebvre[a]

Author affiliations:

[a]Laboratory of Plant-Microbe Interactions and Environment (LIPME), University Toulouse III, INRAE, CNRS, 31326 Castanet-Tolosan Cedex, France

[b]Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France

Running Title: Evolution of the LYR-IA LysM-RLK LCO binding site

Author contributions: J.C., J.F., A.G., C.G., J.-J.B, S.F. and B.L designed research; J.C, J.F., V.G., C.G.-V. N.C., T.-B.L, A.G., M.C., F.M. performed research; S.F., S.P. contributed new reagents/analytic tools; J.C., J.F., V.G., C.G.-V., N.C., A.G., J.-J.B, C.G. and B.L analyzed data; and J.C., C.G. and B.L. wrote the paper.

The authors declare no competing interest.

[1] present address: Centre for Plant Molecular Biology (ZMBP), Eberhard-Karls-Universität Tübingen, D-72076 Tübingen, Germany;

[2] present address: Department of Molecular Biology and Genetics, Aarhus University, Denmark.

Corresponding author: Benoit Lefebvre, Chemin de Borde Rouge CS 52627, 31326 Castanet-Tolosan cedex, France, 0033(0)5612853, Benoit.Lefebvre@inrae.fr

Keywords: Lysin motif receptor like kinases, plant symbiosis, lipochitooligosaccharide perception

ABSTRACT

Lysin motif receptor like kinases (LysM-RLKs) are involved in the perception of chitooligosaccharides (COs) and related lipochitooligosaccharides (LCOs) in plants. Expansion and divergence of the gene family during evolution have led to various roles in symbiosis and defence. By studying proteins of the LYR-IA subclass of LysM-RLKs of the *Poaceae*, we show here that they are high affinity LCO binding proteins with a lower affinity for COs, consistent with a role in LCO perception to establish arbuscular mycorrhiza (AM). In Papilionoid legumes whole genome duplication has resulted in two LYR-IA paralogs, MtLYR1 and MtNFP in *Medicago truncatul*a, with MtNFP playing an essential role in the root nodule symbiosis with nitrogen-fixing rhizobia. We show that MtLYR1 has retained the ancestral LCO binding characteristic and is dispensable for AM. Domain swapping between the three Lysin motifs (LysMs) of MtNFP and MtLYR1 and mutagenesis in MtLYR1 suggest that the MtLYR1 LCO binding site is on the second LysM, and that divergence in MtNFP led to better nodulation, but surprisingly with decreased LCO binding. These results suggest that divergence of the LCO binding site has been important for the evolution of a role of MtNFP in nodulation with rhizobia.

INTRODUCTION

N-acetyl glucosamine containing oligosaccharides are important signalling molecules in plants and can elicit defence and symbiotic responses to microorganisms, depending on their structure (Chiu and Paszkowski 2020; Khokhani et al. 2021; Zipfel and Oldroyd 2017). Short chain chitooligosaccharides (CO4-CO5) are produced by symbiotic arbuscular mycorrhizal (AM) fungi and elicit symbiotic responses in a variety of plants, whereas longer chain COs (CO7-CO8) elicit defence responses (Carotenuto et al. 2017; Feng et al. 2019; Kaku et al. 2006). Lipochitooligosaccharides (LCOs), in which the terminal non-reducing sugar is *N*-acylated generally with a C16 – C20 fatty acid, are produced by a variety of fungi, including AM fungi (Cope et al. 2019; Maillet et al. 2011; Rush et al. 2020), and also by rhizobial bacteria (D'Haeze and Holsters 2002). The production of these rhizobial LCOs, called Nod factors (NFs), is essential for most rhizobia to establish the nitrogen-fixing, root-nodule symbiosis with legume plants, and species-specific chemical substitutions enable interaction with the correct symbiotic host. For example, *Sinorhizobium meliloti* typically produces tetrameric NFs, that are *O*-sulphated on the reducing sugar, *N*-acylated with predominantly a C16:2 acyl chain on the terminal non-reducing sugar (termed LCO-IV(C16:2,S); Fig. S1), and partially *O*-acetylated on this sugar. These decorations allow nodulation of *Medicago* (Lerouge et al. 1990). NFs are very potent signals, eliciting symbiotic activities at pM to nM concentrations on roots of the legume host.

The discrimination of CO and LCO structures by plants is largely determined by lysin motif (LysM) containing proteins (Chiu and Paszkowski 2020; Khokhani et al. 2021; Zipfel and Oldroyd 2017). The LysM is a small (less than 65 amino acid) protein domain, which occurs in Bacteria and many Eukaryotes and generally has a conserved $\beta\alpha\alpha\beta$ fold. In plants three divergent tandem LysMs (LysM1, LysM2, LysM3) are associated in the extracellular region (ECR) with a transmembrane spanning helix and an intracellular region (ICR) containing a kinase domain to form plasma-membrane (PM) located LysM receptor like kinases (LysM-RLKs). Phylogenetically these LysM-RLKs can be divided into two major classes, the LYRs (generally with an inactive, pseudokinase domain) and the LYKs (in which the kinase domain is predicted to be active). These two classes can be further divided into sub-classes (Buendia et al. 2018).

Studies on CO perception, particularly in rice (*Oryza sativa* (Os)) and *Arabidopsis thaliana* (At), have identified a gene of the LYK-I subclass, CERK1, which is essential for CO-elicited responses (Yang et al. 2022). In rice, OsCERK1 interacts with a LysM protein, OsCEBIP, for defence responses whereas it interacts with a LYR-IA subclass protein, OsMYR1 (also called OsNFR5) for AM (Zhang et al. 2021). Crystal structures of the AtCERK1 and OsCEBIP ECRs have identified that COs bind in a groove in LysM2 (Liu et al. 2016; Liu et al. 2012). In dicots, such as tomato and *Petunia hybrida*, the LYR-IA protein, SlLYK10 and PhLYK10 respectively, is important for establishing AM and binds LCOs, rather than COs (Girardin et al. 2019). Thus plants have adopted a variety of ways in which to recognise CO/LCO signals via LysM proteins.

Legumes contain a highly expanded LysM-RLK family due to both segmental duplication to form clusters and by a whole genome duplication (WGD) event, which occurred in the Papilionoid legumes (Young et al. 2011). Thus there is more scope for variation, particularly to accommodate NF signalling. Both a LYK protein belonging to the LYK-I subclass (termed MtLYK3 in *M. truncatula* (Mt), LjNFR1 in *Lotus japonicus* (Lj), and PsSYM37, in *Pisum sativum* (Ps)) and a LYR protein belonging to the LYR-IA subclass (termed MtNFP, LjNFR5 and PsSYM10) are implicated in NF recognition and are essential for rhizobial nodulation (Krönauer and Radutoiu 2021). By domain swapping and mutagenesis, the LysM1 of MtLYK3/LjNFR1 (Bozsoki et al. 2020) and the LysM2 of MtNFP/LjNFR5 (Bensmihen et al. 2011; Gysel et al. 2021) have been shown to be particularly important for nodulation. Structural and biochemical studies on the ECRs, expressed in insect-cells, show that the proteins adopt a similar conformation to CO-binding LysM proteins and bind specific NFs with *K*ds in the μM range (Bozsoki et al. 2020; Gysel et al. 2021). The slow kinetics of LCO binding, compared to fast kinetics of CO binding, has been suggested to be more important than affinity for receptor activation and signalling specificity, through a kinetic proofreading mechanism (Gysel et al. 2021).

We have worked on the perception of LCOs for many years, initially through using a radiolabelled ligand corresponding to a major NF of *S. meliloti* (LCO-IV(C16:2,^{35}S)). We identified a

high affinity (nM range) LCO binding protein in the PM fraction of cell cultures as the LYR-IIIA protein, MtLYR3 (Fliegmann et al. 2013; Gressent et al. 1999). *MtLYR3* is located next to and in inverse orientation to *MtNFP* (Buendia et al. 2018; Gough et al. 2018), and legume orthologs of MtLYR3 bind LCOs (Bouchiba et al. 2022; Malkov et al. 2016). We also used a radiolabelled LCO-V ligand (LCO-V(C18:1,NMe,^{35}S)), which has more similarity to a Myc-LCO than the *S. meliloti* NFs, to characterise LCO binding to the tomato and *P. hybrida* LYR-IA protein, and showed the relevance of these proteins for AM (Girardin et al. 2019). In view of the physiological importance of LYR-IA, in this article we have investigated the binding properties of this subclass of protein from the *Poaceae* and *M. truncatula*, leading to a hypothesis on its ancestral role in plants. We also show that the two LYR-IA paralogs in *M. truncatula* have different LCO binding properties, which are unexpected in view of the physiological roles of these proteins.

RESULTS

MtLYR1 is a high affinity LCO binding protein with selectivity for LCOs versus short-chain COs

In order to analyse the LCO binding capability of the *M. truncatula* LYR-IA proteins, (MtNFP and MtLYR1), we made fusions of their ECRs with the transmembrane (TM) and ICR of MtNFP tagged with YFP, as previously done for MtLYR3 and tomato SlLYK10 (Fliegmann et al. 2013; Girardin et al. 2019). The aim was to ensure good expression of the proteins when expressed in *Nicotiana benthamiana* leaves, whilst leaving the ECRs intact. Membrane fractions of the leaves expressing the LysM-RLKs and of the negative control, mock-treated leaves (MOCK), were then prepared and analysed by Western blotting using anti-YFP antibodies. These membrane fractions were analysed for high-affinity binding to the *S. meliloti* NF ligand (LCO-IV(C16:2,^{35}S); Fig. S1) using about 1 nM of the labelled ligand. Specific LCO binding was deduced by subtracting the non-specific binding (not competed by high (2 µM) concentrations of unlabelled ligand) from the total binding. Only MtLYR3 extracts, and not those of MtNFP or MtLYR1, showed specific binding to this NF ligand compared to binding to MOCK (Fig. 1A). This lack of specific binding is particularly unexpected for MtNFP in view of its essential role in the perception of NFs during symbiosis (Arrighi et al. 2006) and the recent binding studies on MtNFP (Gysel et al. 2021). In terms of expression in this system, MtNFP localises to the PM as expected (Lefebvre et al. 2012) and its expression level is better than MtLYR3 (Fig 1A). We then tested the remaining 8 LYR proteins (MtLYR2, MtLYR4-MtLYR10) and although there was considerable variation in the level of expression, all proteins, except the MtLYR6 fusion, were expressed at least as well as MtLYR3 (Fig. S2), but in comparison to binding to MOCK, only MtLYR3 extracts showed specific binding to this NF ligand (Fig. S2).

In tomato and *Petunia*, LYR-IA are represented by a single ortholog (called SlLYK10 and PhLYK10), which binds LCO-V(C18:1,NMe,^{35}S) (Fig. S1), an LCO structure different from the NF ligand,

and able to reveal the Myc-LCO binding site (Girardin et al. 2019). We thus analysed binding to MtNFP and to its LYR-IA paralog, MtLYR1, with this ligand. The MtLYR1 extracts showed binding to the ligand (Fig. 1B), whereas despite similar expression level of the MtNFP protein, no specific LCO binding to the MtNFP extracts was detected.

To determine the parameters of LCO binding to MtLYR1, a ligand saturation experiment using LCO-V(C18:1,NMe,S) was subjected to Scatchard analysis (Fig. 1C). The results are consistent with the presence of a single class of binding site, exhibiting a high affinity for the ligand (*K*d = 6.5 nM). The LCO-V(C18:1,NMe,S) binding could be competed with the unlabelled NF LCO-IV(C16:2,S) (Fig. 1D) but with an IC_{50} greater than 500 nM, suggesting at least an 80-fold lower affinity than for LCO-V(C18:1,NMe,S). This observation could explain why binding to MtLYR1 of the radiolabelled *S. meliloti* NF could not be detected (Fig. 1A). Competition experiments showed that an LCO mixture of LCO-V(C18:1,Fuc)/LCO-V(C18:1,MeFuc) recently-described as abundant Myc factor structures (Rush et al. 2020), could efficiently compete for binding while both CO4 and CO8 could only partially compete the LCO-V(C18:1,NMe,^{35}S) binding to MtLYR1 (Fig. 1E). To test whether MtLYR1 has affinity for short-chain COs on a binding site independent of LCOs, we developed a crosslinkable biotinylated CO (thereafter called CO5-biot, Fig. S1), which can form a covalent bond with proteins in its vicinity. It allowed us to measure affinity of proteins for short-chain COs, following expression in *N. benthamiana*, incubation of membrane extracts with CO5-biot and detection by Western blotting using Streptavidin. A saturation experiment showed MtLYR1 labelling with CO5-biot, but only at high concentration of the ligand and not saturated between 500 nM and 5 µM, showing that the affinity for CO5 is in the µM range (Fig. 1F). A similar result was found for OsCEBIP, whereas no CO5-biot binding was detected to MtNFP (Fig. S3). Stronger competition of CO5-biot binding to MtLYR1 by LCOs compared to COs (Fig. 1G), shows that CO5-biot binds to the LCO binding site, which has higher affinity for LCOs than COs.

We previously found that *Solanaceae* LYR-IA genes could complement a *Mtnfp* mutant for nodulation but with a lower efficiency compared to *MtNFP* (Girardin et al. 2019). *MtLYR1* showed a similar low ability to complement *Mtnfp*, which could not be increased significantly by replacing the TM+ICR with that of NFP. Both the efficiency of nodulation and the nodule number were much lower with the *MtLYR1* constructs than with *MtNFP* (Fig. 1H).

Transcriptomic data show that *MtLYR1* is induced in roots during AM and expressed particularly in arbuscule-containing cells, whereas it is not induced during nodulation (Fig. S4; Gomez et al. 2009). To study the potential symbiotic role(s) of *MtLYR1*, we tested *Mtlyr1* mutants for both AM and nodulation. Two independent knockout alleles, were obtained; *Mtlyr1-1*, an EMS mutant in *M. truncatula* A17 (C58T, introducing an early stop codon), and *Mtlyr1-2*, a *Tnt*1 insertion in *M. truncatula* R108 (before the region encoding LysM1) (Fig. 2A).

When these mutants were tested with spores of *Rhizophagus irregularis*, no differences between mutant and WT plants were observed either in proportions of colonized roots at 6 wpi (Fig. 2B) or in arbuscule development (Fig. S5). As *Mtnfp* mutants show no observable AM phenotype (Ben Amor et al. 2003), we tested the possibility of redundancy between *MtLYR1* and *MtNFP*, and generated a double *Mtlyr1-1*/*Mtnfp-2* mutant. No differences were observed between the *R. irregularis* colonization levels of these plants compared to plants of the single mutants or WT (Fig. 2B).

We then quantified the number of nodules in *Mtlyr1* mutant plants inoculated by WT *S. meliloti*. No differences in nodule number between mutants and their corresponding WT plants were observed (Fig. 2C). Since MtLYR1 has higher affinity for LCO-V(C18:1,NMe,S) than LCO-IV(C16:2,S) we also tested whether MtLYR1 plays a role in nodulation through perception of NFs bearing a C18:1 fatty acid rather than a C16:2 fatty acid, as produced by the *S. meliloti nodFE* mutant (Ardourel et al. 1994). Again, no differences in nodule number were observed (Fig. 2D). With either rhizobia strain, the *Mtnfp* mutant was devoid of nodules (Nod^-), as expected (Ben Amor et al. 2003).

Our results show that, despite an induction of its expression during AM, *MtLYR1* is dispensable both for AM and for nodulation.

***Poaceae* LYR-IA proteins have similar LCO/CO binding properties as dicot LYR-IA proteins.**

We previously showed that *Solanaceae* LYR-IA proteins have high affinity for LCO-V(C18:1,NMe,S) and selectivity for LCOs versus COs (Girardin et al. 2019). To determine whether these properties are conserved among plant species, we characterized LYR-IA from monocotyledons, *Brachypodium distachyon (Bd)*, *Oryza sativa (Os)* and *Triticum aestivum (Ta)*. A single gene was found in the diploid species *B. distachyon* and *O. sativa* (*BdLYR1* and *OsMYR1/OsNFR5* respectively), while three genes were found in the hexaploid *T. aestivum* (*TaLYR1*, *TaLYR1B* and *TaLYR1D*; Fig. S6). We tested extracts of all 5 proteins, expressed in *N. benthamiana* leaves, and found specific LCO-V(C18:1,NMe,^{35}S) binding to all of them (Fig. 3A). Scatchard analysis of the BdLYR1, OsMYR1 and TaLYR1D binding, which in each case correlated to a single class of high affinity binding sites (Fig. S7), yielded affinities (*K*d) from 9 to 19 nM (Fig. 3B), which are in the same range as those found for MtLYR1, SlLYK10 and PhLYK10. We also tested the selectivity of these LCO binding sites for COs and found that CO4 and CO8 are poor competitors of the LCO-V(C18:1,NMe,^{35}S) binding (Fig. 3C) for each of the proteins. Using CO5-biot, absence of high affinity for short-chain COs was confirmed for OsMYR1 as, like MtLYR1, OsMYR1 was labelled only at a high concentration of the ligand (Fig. 3D). We characterized BdLYR1 in more detail. We found that synthetic LCO-V(C18:1,Fuc/MeFuc) (Fig. S1), but not LCO-IV(C16:2,S), can fully compete LCO-V(C18:1,NMe,^{35}S) binding to BdLYR1 (Fig. 3E). High affinity for LCO-V(C18:1,NMe,S), but not for Sm LCO-IV(C16:2,S) and selectivity for LCOs versus CO4 was

confirmed by using a radiolabelled crosslinkable LCO derivative (Fig. S1) on BdLYR1 purified from *N. benthamiana* leaves (Fig. 3F, Fig. S8). This LCO can form a covalent bond with proteins in its vicinity and allowed us to measure affinity of proteins for LCO, following detection by Phosphorimaging. Finally, we also tested the ability of *BdLYR1* to complement the *Mtnfp* mutant for nodulation. As for *MtLYR1*, *BdLYR1* allowed the formation of a few nodules with much lower efficiency than *MtNFP* (Fig. 3G). The nodules were well colonized by *S. meliloti* (Fig. S9).

LysM domain swaps reveal the critical role of LysM2 in the functions of MtLYR1 and MtNFP

We have shown that MtLYR1 has high-affinity LCO-binding properties and lower ability to complement *Mtnfp* for nodulation than MtNFP. To pinpoint the regions responsible for the differences, we made domain swaps between the three LysMs of the MtNFP (NNN) and MtLYR1 (LLL) ECRs in fusion proteins to MtNFP-TM+ICR/YFP (Fig. 4A).

We tested the constructs for their ability to complement the *Mtnfp* mutant for nodulation. The results show that the introduction of MtNFP LysM2 in MtLYR1 (LNL), but not of LysM1 (NLL) or LysM3 (LLN), led to a gain in nodulation efficiency to a level similar to that of NNN while conversely introducing MtLYR1 LysM2 into MtNFP (NLN), led to a loss of function in nodulation (Fig. 4B). To further explore the structure-function of the LysM2 of the MtLYR1 protein, three mutants were designed in which MtLYR1 residues differing from MtNFP were converted to the corresponding MtNFP residue(s) in the LLL construct (Fig. 4D). Analysis of the constructs for nodulation revealed that the LLL-S114N and LLL-K147P mutants, but not the double LLL-N142S/Y144T mutant, gained some ability to complement the *Mtnfp* mutant (Fig. 4B), thus suggesting that these MtNFP residues in LysM2 are important for nodulation.

We then tested LCO-V(C18:1,NMe,^{35}S) binding to membrane extracts of the various chimeras/mutated proteins, after expression in *N. benthamiana*. The LysM2 (LNL) but not the LysM3 (LLN) domain swap, showed almost complete loss of binding to LCO-V(C18:1,NMe,S), whereas replacing the LysM2 of MtNFP with that of MtLYR1 (NLN) led to gain of LCO-V(C18:1,NMe,S) binding. Binding assays of the 3 LLL proteins mutated in LysM2 to corresponding MtNFP residues, showed that LLL-S114N and the double mutant LLL-N142S/Y144T, and to a lesser extent LLL-K147P, have reduced LCO-V(C18:1,NMe,S) binding, supporting the importance of these LysM2 residues for LCO-V(C18:1,NMe,S) binding. Although the NLL construct showed reduced binding, this construct showed poor plasma membrane localization of the protein (Fig. S10), suggesting that it may not be folded correctly (Lefebvre et al. 2012).

These results show that residues of the LysM2 are very important for the role of MtLYR1 in LCO binding and of MtNFP in nodulation.

MtLYR1 ECR modelling suggests that residues critical for LCO binding lie on the putative CO binding groove and near the hydrophobic patch

A model of the MtLYR1-ECR has previously been produced, based on the MtNFP-ECR crystal structure, with placement of a CO in a groove (Gysel et al. 2021), similar to the one found by AtCERK1-ECR/CO co-crystallization (Liu et al. 2012). Due to the higher similarity of MtLYR1-ECR with the LjLYS11-ECR, we made models of MtLYR1 ECR, based on the available crystal structure of LjLYS11-ECR in order to place the key binding residues found here in relation to the proposed binding site. The model of MtLYR1-ECR based on LjLYS11 (Fig. S11) is very similar to the one based on MtNFP (Gysel et al. 2021). The key LCO binding residues discussed in the previous section, S114 and N142/Y144 lie on opposite sides of the proposed CO-binding groove in LysM2, while K147 lies at the end of the groove closest to the proposed non-reducing end of the CO chain (Fig. 4E, Gysel et al. 2021). The hydrophobic patch, described by Gysel *et al* (2021) is in a position to accommodate the acyl chain of the LCO, if the CO moiety is located in the same orientation and location as the CO in AtCERK1/OsCEBIP. Although we have not shown whether both the N142S and Y144T mutations are responsible for the decrease in MtLYR1 binding in the double mutant, N142 is close to the hydrophobic patch, formed by L145, L146 and possibly I148. Superposition of the crystal structure of the MtNFP-ECR on the model of the MtLYR1-ECR reveals differences in the structure of the two proteins in the key residues that we identified to be required for the LLL mutant proteins to complement for nodulation (Fig. 4F). The two Leu residues, identified by Gysel *et al* (2021) in which a double mutation L147D/L154D reduces the ability of MtNFP to complement the *Mtnfp* mutant for nodulation, form part of the hydrophobic patch, and are conserved in MtLYR1 (L141, I148).

In conclusion, residue positions, critical for high affinity LCO-V(C18:1,NMe,S) binding to MtLYR1, lie in a groove, conserved in CO binding proteins, and near a hydrophobic patch, described by Gysel *et al* (2021). The corresponding residues of MtNFP change the structure of the groove and are critical for permitting the MtLYR1 ECR construct to complement for nodulation of *Mtnfp*.

DISCUSSION

An ancestral role of LYR-IA as high affinity LCO binding proteins involved in AM?

We previously found that two *Solanaceae* LYR-IA proteins have high affinity for LCOs and selectivity for LCOs versus COs. Here we extended our analysis to *Poaceae* LYR-IA proteins (rice, wheat and *Brachypodium*) and found that these proteins have similar binding properties to those of the *Solanaceae* LYR-IA, with a *K*d for LCOs in the nM range (Fig. 3). Our binding results suggest that the LYR-IA ancestor was an LCO binding protein, with a lower ability to bind COs.

Our data on the rice LYR-IA protein, OsMYR1 (also called OsNFR5), differ from those of He *et al* (2019), who measured a high affinity for CO4 (*K*d of 90 nM) and did not detect LCO binding, using

the *S. meliloti* NF as LCOs. We found that a major *S. meliloti* NF, LCO-IV(C16:2,S) binds poorly to OsMYR1 compared to LCO-V(C18:1,NMe,S), which gives a possible explanation of how they may have missed high-affinity LCO binding to this protein. Moreover, we have shown that short chain COs bind with lower affinity than LCOs to OsMYR1, by both direct binding of the CO5-biot ligand and by poor CO competition of LCO binding. Our results also suggest that the LCOs and COs bind at the same site. The lack of high affinity CO binding is consistent with the observation of Carotenuto *et al* (2017) that CO4-induced Ca^{2+} spiking responses are not affected in a *Osmyr1* (*Osnfr5*) mutant, although conflicting results were found by He *et al* (2019).

The LYR-IA gene has generally been lost in non-mycorrhizal plants suggesting a role in AM (Gough et al. 2018). Indeed, analysis of the expression of LYR-IA genes and phenotyping of mutants in tomato, *Petunia* and rice (Girardin et al. 2019; He et al. 2019; Miyata et al. 2016) suggest that Myc-LCO binding LYR-IA proteins play an important and widespread role in AM, which in view of the antiquity of this symbiosis (established over 400 MYA) is probably an ancestral role. However, we cannot exclude that they may also be involved in interaction with other LCO-producing fungi or other microorganisms (Rush et al. 2020).

Duplication and divergence of LYR-IA in *M. truncatula*

In dicotyledonous plants, the evolution of LYR-IA genes is associated with the evolution of nodulation and nitrogen fixation (Rutten et al. 2020). A duplication of LYR-IA associated with the origin of the so-called nitrogen fixing clade of plants (orders *Fabales*, *Fagales*, *Cucurbitales* and *Rosales*) led to two ancestral paralogs, termed *NFP-I* and *NFP-II.* LCO-driven nodulation is correlated with a functional *NFP-II* type gene (the orthogroup of *MtNFP/*MtLYR1) and in legume plants the *NFP-I*-type gene has been lost (Rutten et al. 2020). In Papilionoid legumes, a WGD event occurred more recently (about 58 MYA), leading to two *NFP-II* genes. While *MtNFP/LjNFR5* is present in the majority of Papilionoids, the paralogous *MtLYR1/LjLYS11* is found in Galegoid and Robinoid legume species, but has apparently been lost in others, including in most of the related Millettioids, which include *Glycine* and *Phaseolus* spp. (Fig. S6; Rutten et al. 2020).

The *MtNFP/LjNFR5* orthogroup has clearly become specialized for an essential and paramount role in rhizobial NF perception and nodulation (Krönauer and Radutoiu 2021), whereas it was speculated that *MtLYR1/LjLYS11* retained an analogous role in AM (Young et al. 2011). Transcriptomic data support this hypothesis as both *MtLYR1* and *LjLYS11* are induced during AM and in arbuscule-containing cells (Fig. S4; Gomez et al. 2009; Rasmussen et al. 2016), like the tomato LYR-IA gene (Girardin et al. 2019). However, analysis of *Mtlyr1* mutants from two divergent *M. truncatula* genotypes, A17 and R108, did not reveal a symbiotic role for this protein either in AM or rhizobial nodulation (Fig. 2). This lack of AM phenotype is not due to redundancy with *MtNFP* as the double

Mtlyr1/*Mtnfp* mutant also had no observable AM phenotype (Fig. 2). Similar results have been obtained in *L. japonicus*, using *Ljlys11* and *Ljnfr5* mutants (Rasmussen et al. 2016).

The role of LYR-IA in AM might be masked by redundancy with other LysM-RLKs, particularly in legumes where this family is considerably expanded (about 22 genes in *M. truncatula* versus 10 or 11 in rice, depending on the genotype). This hypothesis is supported by the AM phenotype of *Mtlyk9* mutants (LYK-I subclass; Gibelin-Viala et al. 2019), which is enhanced in a *Mtlyk9*/*Mtnfp* double mutant (Feng et al. 2019). Other contenders for a role in AM in *M. truncatula* are MtLYR3 (LYR-IIIA subclass) since loss of LCO binding in MtLYR3 orthologs can be associated with loss of AM (Bouchiba et al. 2022; Malkov et al. 2016) and MtLYR8 (LYR-IB subclass) since, like LYR-IA, this subclass is absent in the non-AM species *A. thaliana* (Buendia et al. 2018) and shown to be involved in AM in barley (Li et al. 2022). Clearly, analysis of double, triple and higher order mutants are required to unravel possible LysM-RLK redundancy in AM.

Using a LCO-V(C18:1,NMe,S) ligand, we found high affinity binding to MtLYR1, but not to MtNFP (Fig. 1). Binding to MtLYR1 showed similar affinity and specificity as the non-legume proteins (Fig. 3; Girardin et al. 2019). Neither protein directly bound the *S. meliloti* NF ligand (LCO-IV(C16:2,S)) to high affinity although this LCO can compete for LCO-V(C18:1,NMe,S) binding to MtLYR1 at high concentrations, as shown for non-legume LYR-IA proteins (Fig. 3; Girardin et al. 2019).

Together, our and other results suggest that the *MtLYR1* gene has retained many characteristics, in terms of enhanced expression in arbuscules and LCO binding of the encoded protein, to an ancestral gene, although its role in AM remains enigmatic and is probably redundant with other LysM proteins. A redundant role for MtLYR1 in the rhizobia symbiosis cannot be excluded, although there is no indication of this from the expression pattern.

Has MtNFP a high affinity for *S. meliloti* LCO?

In our assays, no evidence was found for high affinity LCO or CO binding to MtNFP, using either the LCO-IV(C16:2,S) *S. meliloti* NF ligand, the LCO-V(C18:1,NMe,S) ligand, or indeed the CO5-biot ligand (Fig. 1). Our results on absence of high affinity NF binding to MtNFP are robust as they have been confirmed in many experiments over a long period of time, using different MtNFP constructs and expression systems. Moreover, a clear difference is seen, using the same techniques, to the previously described high affinity LCO binding protein MtLYR3 (Fliegmann et al. 2013, positive control) whereas extracts of MtNFP, other LYR proteins and MOCK (negative control) showed no high-affinity LCO-IV(C16:2,S) NF binding (Fig. 1, Fig. S2). It should be noted that our work used a non-*O*-acetylated *S. meliloti* NF as ligand, whereas at least part of the NFs from this species is *O*-acetylated on the terminal non-reducing sugar (specified by the *NodL* gene; Ardourel et al. 1994). This decoration is not essential for *MtNFP*-dependent signalling as a *S. meliloti nodL* mutant is able to

nodulate *M. truncatula*, although with a slightly lower efficiency than the WT (Smit et al. 2007) and *nodL* NFs are active on *Medicago* roots down to pM concentrations in biological assays (Ardourel et al. 1994; Journet et al. 1994; Wais et al. 2002). Even if such NFs have a 10-fold lower affinity for MtNFP compared to *O*-acetylated NFs, as suggested by Gysel *et al* (2021), we should still be able to detect affinities in the 100 nM range, as shown in previous studies on other proteins (Fliegmann et al. 2013; Girardin et al. 2019). We cannot however measure low affinity binding in the µM range.

Using biolayer interferometry assays, Gysel *et al* (2021) have reported that the soluble MtNFP-ECR, expressed and purified from insect cells, binds to LCO chips with a specificity expected from genetic studies, but with an affinity in the µM range. They suggest that the slow kinetics of LCO binding to MtNFP, rather than affinity, is important for receptor activation and signalling specificity. It is thus possible that MtNFP has a lower affinity for Sm LCOs compared to other LCO binding proteins such as MtLYR3, and that this is sufficient for mediating NF perception.

Evolution in the LCO binding site in MtNFP allowed higher efficiency for nodulation

All tested non-legume LYR-IA genes showed ability to complement the *Mtnfp* mutant for nodulation and rhizobial colonization (This study and Girardin et al. 2019), but with low efficiency since the proportion of nodulating plants and the nodule number were much lower than by complementation with *MtNFP.* Similarly, *MtLYR1* showed poor complementation of *Mtnfp* for nodulation, consistent with the protein retaining properties inherited from the ancestral gene.

By analysing LysM domain swaps of the MtLYR1 (LLL) and MtNFP (NNN) proteins we have confirmed a critical role of the MtNFP LysM2 for a role in nodulation (Fig. 4; Bensmihen et al. 2011; Gysel et al. 2021). This was found by both gain of function in the LLL construct (LNL) and loss of function in NNN (NLN). Moreover replacing some of the MtLYR1 LysM2 residues in LLL by those of MtNFP has highlighted the critical role of changing MtLYR1/MtNFP residues S114/N122 and K147/P153 in conferring nodulation ability. In MtLYR1, these residues are important for LCO binding (Fig. 4), as are N142 (replaced by S) and/or Y144 (replaced by T). Through homology modelling to the crystal structure of the LjLYS11-ECR (Fig. 4), our work is consistent with the CO moiety of the LCO binding in a groove in LysM2, similar to the one where CO binds in the crystal structure of Arabidopsis AtCERK1 and rice OsCEBIP (Liu et al. 2016; Liu et al. 2012). Our work is also compatible with the acyl chain of the LCO being accommodated in a hydrophobic patch, identified on certain LYR proteins (Gysel et al. 2021). However, because there is no crystal structure of a LCO ligand in a LysM protein, this location and orientation of the LCO remain speculative. It should be noted that the residues identified in this work are important for the difference between MtLYR1 and MtNFP in terms of high affinity LCO binding (and nodulation) but are not fully conserved between the LYR-IA LCO binding proteins of different species (Fig. S12). This suggests that substitutions in this groove and

the associated hydrophobic patch render the protein still capable of accommodating the notoriously flexible lipid chain and oligosaccharide of the LCO (Groves et al. 2005).

Model of NF perception and nodulation

As MtNFP, in addition to its essential role in rhizobial nodulation, plays a role in immunity to pathogens and in Myc-LCO responses, it has been suggested that it may participate in different protein complexes to lead to different responses (Gough and Jacquet 2013). For nodulation, its essential role in mediating sulfated *S. meliloti* NF perception at pM to nM concentrations, suggests that MtNFP is a key component of a NF receptor, although, as shown here, not as a high affinity NF binding protein. We propose that it could either share NF binding with another LysM protein (Moling et al. 2014; Igolkina et al. 2018; Solovev et al. 2021) or be more involved in the signalling role of the NF receptor. Because of the exquisite sensitivity to sulfated NFs, it is possible that NF perception in *Medicago* shows some differences compared to other legumes. A precedent for species-specific differences in LysM-RLK function exists from work on CO perception in rice/Arabidopsis (Yang et al. 2022).

It should be noted, that during nodulation in the natural environment the role of high affinity LCO binding proteins, such as MtLYR1 and MtLYR3, could be to prevent infection by non-compatible rhizobia or other LCO producing microorganisms.

In conclusion, our results on different plant families suggest that the ancestor of the LYR-IA subclass of LysM-RLKs was an LCO binding protein, with a role in AM, properties that have been retained throughout evolution. Duplication of the gene in legumes led to specialisation of one paralog (corresponding to *MtNFP*) for interacting with rhizobia. The other paralog (*MtLYR1*) retained many characteristics of the original protein in terms of LCO binding and a potential, but probably redundant, role in AM. In *M. truncatula*, changes in the expression pattern of *MtNFP* coupled with evolution of the putative ancestral LCO binding site have been important for the function of MtNFP in nodulation with *S. meliloti*, and our results suggest that the latter has been through a decrease in LCO binding, compared to MtLYR1.

MATERIALS AND METHODS

Plant material

Medicago truncatula Gaertn. A17 was used as the gene source and for phenotyping experiments, unless otherwise stated. *Brachypodium distachyon* Bd21.3 and wheat Chinese spring were used as the gene source.

Mutants and phenotyping

The A17 *Mtnfp* mutant (*Mtnfp-2*), carrying the ProMtENOD11:GUS construct, has been previously described (Arrighi et al. 2006). Mutant *Mtlyr1-1* (C58/T that leads to a stop codon), was derived from EMS mutagenesis of A17 (Le Signor et al. 2009). Mutant *Mtlyr1-2* (NF3951) is a *Tnt1* insertional mutant of *M. truncatula* ssp. *tricycla* R108-1 (Tadege et al. 2008). Putative alternative translation would lead to proteins without signal-peptide and thus not functional.

Phenotyping of the mutants for nodulation was done in test tubes on Fahraeus agar slants (Luu et al. 2022) using about 10^6 bacteria/plant of *Sinorhizobium meliloti* 2011 (GMI6526) or the *nodFE* mutant (GMI6528), containing a constitutive LacZ fusion on plasmid pXLGD4 (Ardourel et al. 1994). Mycorrhiza analysis was performed on mycorrhizal roots using the gridline intersect method after plants were grown for 6 weeks, as described (Gibelin-Viala et al. 2019) in the presence of 20 spores *Rhizophagus irregularis* (DAOM197198, Agronutrition, Toulouse, France) / plant, with 7-20 plants in total per genotype.

Complementation assays

Using *Agrobacterium rhizogenes*-mediated transformation (Boisson-Dernier et al. 2001), seedlings of the *Mtnfp-2* or *Mtnfp-2*/ProMtENOD11:GUS lines were transformed using strains containing either empty vector (EV) or the appropriate construct, and transformants were selected on media containing 25 µg ml^{-1} kanamycin. After two-weeks growth, the plants were transferred to pots (Luu et al. 2022) and analysed for nodulation following inoculation with 10^6 bacteria/plant of *S. meliloti* 2011.

Constructs for *in planta* protein expression

The predicted extracellular region (ECR) of the *M. truncatula* LYR genes (Lefebvre et al. 2012) was amplified from genomic DNA using the primers listed in Fig. S13 and cloned behind either the Pro35S or the ProLjUb and in front of the TM+KIN domains of MtNFP, followed by YFP, in a pcambia vector modified for Golden Gate cloning (Fliegmann et al. 2016). The MtLYR3 and MtLYR4 constructs are described in Fliegmann et al (2013). The complete coding sequences (CDS) of *BdLYR1* (Bradi1g69290), *TaLYR1A* (TraesCS4A03G0122100), *TaLYR1B* (TraesCS4B03G0646400) and *TaLYR1D* (TraesCS4D03G0574900) were similarly cloned in a pbin vector modified for Gateway cloning (Lefebvre et al. 2010), in fusion with YFP, using the primers listed in Fig. S13. The *OsMYR1*/*OsNFR5* (LOC_Os03G13080) sequence was optimized with the *N. benthamiana* codon usage (Fig. S14) by gene synthesis and cloned in the pcambia vector modified for Golden Gate cloning. For the domain swapping experiments the ECR of MtNFP and MtLYR1 was each divided into three domains at the conserved CXC sequences (between the LysMs) to form LysM1, LysM2 and LysM3 modules. Each module was amplified by PCR using the primers listed in Fig. S13 and cloned. Full length proteins were reconstituted by Golden Gate cloning using the same modules described above. Mutations

were introduced in the LysM2 of MtLYR1 by site-directed mutagenesis. All constructs were verified by sequencing.

Expression in *N. benthamiana* leaves and preparation of protein extracts.

Agrobacterium tumefaciens LBA4404 strains containing expression cassettes for the fusion constructs and/or the silencing inhibitor P19 were used to agro-infiltrate the three to four oldest leaves of *N. benthamiana* plants, such that P19 was co-expressed with the LysM-RLKs. Leaves expressing only P19 were used as control (MOCK). Three days post infiltration, the protein expression in leaves was assessed by confocal microscopy and the leaves were harvested, frozen in liquid nitrogen and stored at -70 °C until extraction. Membrane extracts of the MOCK or fusion-expressing leaves (sedimenting between 3 000 – 80 000 *g*) were prepared as described (Fliegmann et al. 2013; Girardin et al. 2019) and then stored at -70 °C in 30 % glycerol in binding buffer until use.

BdLYR1-YFP was solubilized from membrane fraction in IP buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 % glycerol with protease and phosphatase inhibitors) with 0.2 % dodecylmaltoside (DDM, ratio detergent:proteins 1:4) and purified with magnetic anti-GFP beads (Chromotek, gtm-20).

CO5-biot synthesis

The CO5-biot was derived from the CO5-propargyl (Masselin et al. 2019). CO5-propargyl (5 mg, 4 µmol) and azide-Peg3-biotin conjugate (Aldrich) (2.7 mg, 6 µmol) were mixed in 0.1 M phosphate buffer (600 µL, pH 7.4), and copper sulfate (100 µL of an aqueous solution at 10 mg/mL, 4 µmol) and sodium ascorbate (100 µL of an aqueous solution at 10 mg/mL, 5 µmol) were successively added. The reaction was stirred overnight at 20 °C and purified by C18 reversed-phase chromatography (1 g SPE cartridge) with a gradient of MeOH (0 % to 100 %) in water. After freeze−drying, CO5-biot was isolated with 65 % yield (4.4 mg). HRMS (ESI) *m/z* $[M+2H]^{2+}$ calcd for $C_{69}H_{107}O_{33}N_{15}S$ 853.85110, found 853.85189.

LCO and CO binding assays

Synthetic LCO-IV(C16:2Δ2EΔ9Z), LCO-V(C18:1Δ11Z,NMe) purified from the supernatant of a *Rhizobium tropici nodH* mutant and synthetic azido photoactivatable LCO derivative (crosslinkable LCO; Fliegmann et al. 2013), were enzymatically sulfated with ^{35}S and then purified by HPLC to separate out the non-sulfated fraction (see Gressent et al. 1999) for details). After sulfation the ligand generally has a specific radioactivity of 100 to 200 Ci/mmol. Binding assays generally used 5-60 µg protein of membrane fraction and 0.5 to 1 nM labelled ligand in binding buffer (25 mM sodium cacodylate pH 6, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 250 mM saccharose, 1 mg/L antipain, 1 mg/L leupeptin, 1 mg/L aprotinin, 1 mg/L pepstatin, 1 mg/L chymostatin and 100 µM AEBSF). Non-specific binding was assessed by competition with 2 µM unlabelled ligand. Bound ligand was separated from un-bound by filtration (Gressent et al. 1999). Specific binding (non-specific subtracted from total) was calculated. Each LCO binding assay consisted of 3 technical replicates. *K*d and *K*i were determined in cold ligand

saturation and competitive inhibition experiments respectively, with concentrations of the molecules ranging from 10 pM to 2 µM ligand, and calculated using the RadLig software. Competition experiments with LCOs and COs used 0.5 to 1 nM of labelled ligand and 1 to 2 µM of competitor. Crosslink of the azido photoactivatable LCO was performed on purified proteins attached to the anti-GFP beads, incubated with 3 nM radioactive photoactivatable ligand in binding buffer upon UV irradiation as described in Fliegmann et al (2013). Samples were irradiated for 10 min at 254 nm.

For CO5-biot crosslink 50 µg of membrane protein was incubated in binding buffer with 5 µM CO5-biot.

Bioinformatic analysis and structural modelling

The percentages of sequence identity (% ID) between proteins was calculated by EMBOSS Needle and protein sequence alignments used Multalin. Homology modelling used SWISS-MODEL and the crystal structure of LjLYS11 ECD (PDB: 7BAX) or MtNFP ECD (7AU7) as templates. Electrostatic surfaces were calculated using PDB2PQR and APBS webservers. The structures were visualised in PyMOL.

Data, Materials and Software Availability

All plant lines, microbial strains and constructs are available upon request. All study data are included in the article and/or SI Appendix.

ACKNOWLEDGEMENTS

We thank Cécile Pouzet (FRAIB-TRI Imaging Platform Facilities) for confocal imaging, Ton Timmers for nodule sectioning and imaging and Sandra Bensmihen for critical reading of the manuscript. For the *Mtlyr1* mutants we are very grateful to Christine Le Signor (Agroécologie, Dijon), Pascal Ratet (Paris Saclay), and the FP6 EU Project GLIP Grain Legumes FOOD-CT-2004-506223. 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 from Noble Research Institute, LLC and were created through research funded, in part, by grants National Science Foundation NSF-0703285 and IOS-1127155. This work was supported by the Agence Nationale de la Recherche (ANR) projects NodBindsLysM (ANR-05-BLAN-0243-01, coordinator J. Cullimore/C. Gough), NICE CROPS (ANR-14-CE18-0008-01, coordinator G. Bécard), SYMNALING (ANR-12-BSV7-01, coordinator J. Cullimore), MycSignalling (ANR-09-BLAN-0241, coordinator C. Gough), DUALITY (ANR-20-CE20-0017-01, coordinator C. Gough) and WHEATSYM (ANR-16-CE20-0025-01, coordinator B. Lefebvre). 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). TBL gratefully acknowledges receipt of a Bourse d'Excellence from the Ambassade de France au Vietnam. SF acknowledges Labex ARCANE and CBH-EUR-GS (ANR-17-EURE-0003) for partial support and NanoBio ICMG (UAR 2607) for providing facilities for mass spectrometry (A. Durand, L. Fort, R. Gueret) and NMR analyses (I. Jeacomine).

Figure legends

Fig. 1. MtLYR1 has a high affinity for Myc-LCOs but not for COs. A) Binding of the radiolabelled LCO-IV(C16:2,S) ligand to membrane extracts of *N. benthamiana* leaves expressing the indicated *M. truncatula* LYR proteins (upper panel). Proteins correspond to ECR of the MtLYR proteins fused to the TM and ICR of MtNFP and YFP. MOCK corresponds to membrane extracts of leaves expressing only the silencing inhibitor P19. Bars represent the specific LCO binding/mg membrane protein (means and SD between at least 5 independent experiments). Statistical differences were calculated using a Kruskal-Wallis test. Different letters indicate significant difference ($P < 0.05$). Western blot of 10 µg total protein of the membrane extracts using anti-GFP antibodies (lower panel). B) Binding of the radiolabelled LCO-V(C18:1,NMe,S) ligand to membrane extracts of *N. benthamiana* leaves expressing the same fusion proteins as in (A). Bars represent the specific LCO binding/mg membrane protein (means and SD between at least 5 independent experiments). Statistical differences were calculated using a Kruskal-Wallis test. Different letters indicate significant difference ($P < 0.05$). C) Affinity of MtLYR1 for LCO-V(C18:1,NMe,S). Scatchard plot analysis of a cold saturation experiment using a membrane extract expressing MtLYR1, the radiolabelled LCO-V(C18:1,NMe,S) ligand and a range of concentrations of unlabeled LCO-V(C18:1,NMe,S). D) Affinity of MtLYR1 for LCO-IV(C16:2,S). Competitive inhibition of the radiolabelled LCO-V(C18:1,NMe,S) binding using a range of concentration of unlabeled LCO-IV(C16:2,S) as competitor. E) Specificity of the MtLYR1 LCO binding site for LCOs versus COs. Bars represent the percentage competition of specific binding of radiolabelled LCO-V(C18:1,NMe,S) in presence of 2 µM of the indicated competitor (means and SD between 2 independent experiments, each with 3 technical replicates). Statistical differences were calculated using pairwise Student tests. Different letters indicate significant difference ($P < 0.05$). F) Affinity of MtLYR1 for short-chain COs. Saturation experiment using a range of concentration of a crosslinkable biotinylated CO5 (CO5-biot). Western blots using sequentially anti-GFP antibodies and Streptadivin were performed on the same membrane. The arrowhead indicates the position of MtLYR1, whereas * indicates an endogenously (biotinylated) protein in the extracts. G) Specificity of the MtLYR1 CO binding site for LCOs versus COs. CO5-biot binding was competed using CO4 or LCO-V(C18:1,NMe,S). Western blots using sequentially anti-GFP antibodies and Streptadivin were performed on the same membrane. The arrowhead indicates the position of MtLYR1 and the * an endogenously (biotinylated) protein. H) Number of nodulated plants and mean number of nodules per plant on the *Mtnfp* mutant complemented with the indicated constructs. Data from 2 independent experiments. Statistical differences were calculated using a Chi2 test for the number of nodulated plants and a Kruskal-Wallis test for the mean number of nodules. Different letters in each column indicate significant difference ($P < 0.05$).

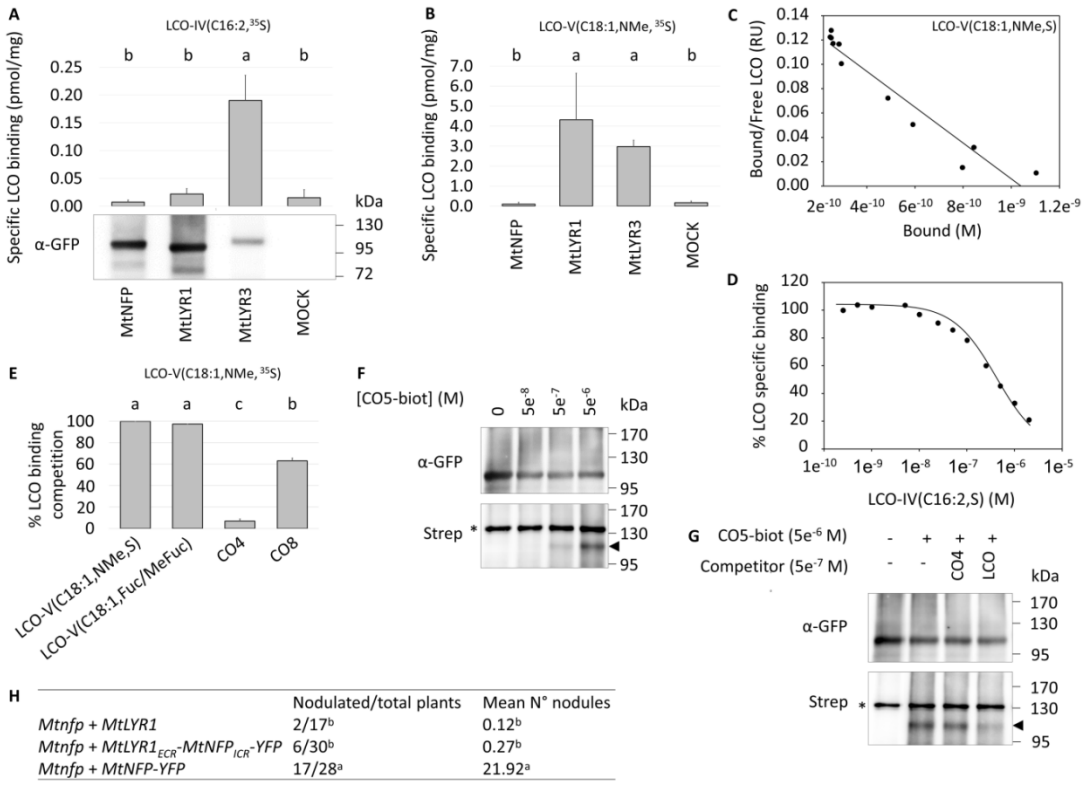
Fig. 2. MtLYR1 is dispensable for AM and nodulation. A) Position of point mutation and transposon (*Tnt*1) insertion in *Mtlyr1-1* (A17) and *Mtlyr1-2* (R108) mutants, respectively, relative to the encoded MtLYR1 protein structure. Grey: LysM, Black: transmembrane domain (TM), White: kinase domain. B) Root-length colonization at 6 wpi with *Rhizophagus irregularis* in *Mtlyr1-1* (A17), *Mtlyr1-2* (R108), *Mtnfp* (A17) and *Mtlyr1-1/Mtnfp* (A17) mutants and their respective controls (A17 or R108). Boxplots represent the distribution between 13-25 individuals from 2 independent experiments. No significant differences were found between each *Mtlyr1* mutant and its corresponding WT line using a Van der Waerden test ($P > 0.05$). C) Number of nodules at 28 dpi with *Sinorhizobium meliloti* in *Mtlyr1-1* (A17), *Mtlyr1-2* (R108) and *Mtnfp* (A17) mutants and their respective wild-type controls (A17 or R108). Boxplots represent the distribution between 10 individuals. Different letters indicate significant difference using a Van der Waerden test ($P < 0.05$). D) Same as in C) except that the plants were inoculated with *S. meliloti nodFE* mutant instead of WT *S. meliloti.*

Fig. 3. *Poaceae* LYR-IA have similar biochemical properties as MtLYR1. A) Binding of the radiolabelled LCO-V(C18:1,NMe,S) to membrane extracts of *N. benthamiana* leaves expressing the indicated fusion proteins (upper panel, means and SD between at least 2 independent experiments, each with 3 technical replicates). MOCK corresponds to membrane extracts of leaves expressing only the silencing inhibitor P19. Statistical differences were calculated using a Kruskal-Wallis test. Different letters indicate significant difference ($P < 0.05$). Proteins correspond to full length proteins fused to YFP. Western blot of 10 µg total protein of the membrane extracts using anti-GFP antibodies (lower panel). B) Affinities of the indicated proteins for radiolabelled LCO-V(C18:1,NMe,S) deduced from Scatchard plot analysis (Fig. S7). C) Selectivity of the indicated protein LCO binding sites to COs. Bars represent the percentage competition of specific radiolabelled LCO-V(C18:1,NMe,S) binding by 1 µM of the indicated competitors (mean and SD between 3 technical replicates). Statistical differences were calculated using pairwise Student tests. Different letters indicate significant difference ($P < 0.05$). D) Affinity of OsMYR1 for short-chain COs. Saturation experiment using a range of concentrations of a crosslinkable biotinylated CO5 (CO5-biot). Western blots using sequentially anti-GFP antibodies and Streptadivin were performed on the same membrane. The arrowhead indicates the position of OsMYR1 and the * an endogenously (biotinylated) protein. E) Specificity of the BdLYR1 LCO binding site for LCOs. Bars represent the percentage competition of specific LCO-V(C18:1,NMe,S) binding in presence of 2 µM of the indicated competitors (means and SD between 3 independent experiments). Statistical difference was calculated using a Student test. * indicates a significant difference ($P < 0.05$). F) Competition experiments on purified BdLYR1-YFP using a radiolabelled- and crosslinkable -LCO derivative and a range of concentrations of unlabeled LCO-V(C18:1,NMe,S) or CO4 as competitor. Phosphorimaging was performed on the same nitrocellulose

membrane used for immunodetection with anti-GFP antibodies. Note that lower LCO signal in the absence of CO4 competitor (left lane phosphorImage panel) is due to a lower quantity of BdLYR1 as visible in the α-GFP panel. G) Number of nodulated plants and mean number of nodules per plant on the *Mtnfp*/proENOD11:GUS mutant complemented with the indicated constructs. Data from at least four independent experiments. Statistical differences were calculated using a Chi2 test for the number of nodulated plants and a Kruskal-Wallis test for the mean number of nodules. Different letters within each column indicate significant difference ($P < 0.05$).

Fig. 4. The structure of MtLYR1 LysM2 is critical for high affinity LCO binding and for poor efficacy to complement *Mtnfp* mutant for nodulation. A) Cartoon of domain and AA swaps between MtLYR1 (green) and MtNFP (cyan). B) Number of nodules per plant on the *Mtnfp* mutant complemented with constructs of the indicated fusion proteins. EV corresponds to the empty vector control. Boxplots represent the distribution between 21-30 individuals from 2 independent experiments. Different letters indicate significant difference using a Van der Waerden test ($P < 0.05$). C) Binding of radiolabelled LCO-V(C18:1,NMe,S) to membrane extracts of *N. benthamiana* leaves expressing the indicated fusion proteins. Error bars show SD of the means of 2 independent experiments, each with 3 technical replicates. MOCK corresponds to membrane extracts of leaves expressing only the silencing inhibitor P19. Statistical differences were calculated using pairwise Student tests. Different letters indicate significant difference ($P < 0.05$). D) Sequence of LysM2 of MtLYR1 (green) and MtNFP (cyan) showing position of mutations. AA highlighted in yellow are conserved between the proteins. E) Structural model of the MtLYR1 LysM2 coloured for surface electrostatic potential to show the hydrophobic patch described in Gysel *et al*. (2021) (dotted line), the putative CO-binding groove and the position of the mutated residues. F) Superposition of the MtLYR1 LysM2 (green) on the MtNFP LysM2 (cyan) showing residues in MtLYR1 (red) and MtNFP (blue) important for LCO binding or nodulation activity respectively. The importance of MtNFP L147/L154 is described by Gysel et al (2021). Corresponding residues with unknown/unclear roles are shown in the same colour as the main protein.

Fig. 1.



H

	Nodulated/total plants	Mean N° nodules
Mtnfp + MtLYR1	2/17[b]	0.12[b]
Mtnfp + MtLYR1$_{ECR}$-MtNFP$_{ICR}$-YFP	6/30[b]	0.27[b]
Mtnfp + MtNFP-YFP	17/28[a]	21.92[a]

Fig. 2.

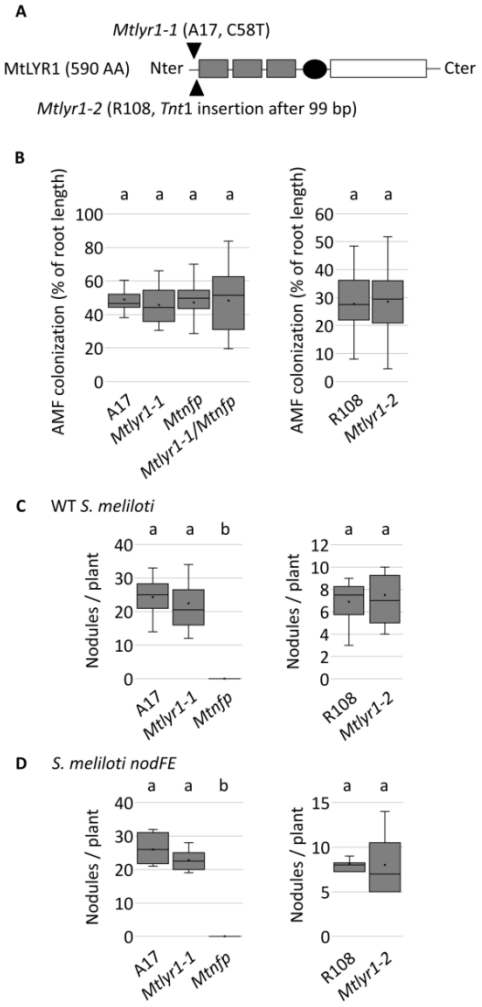
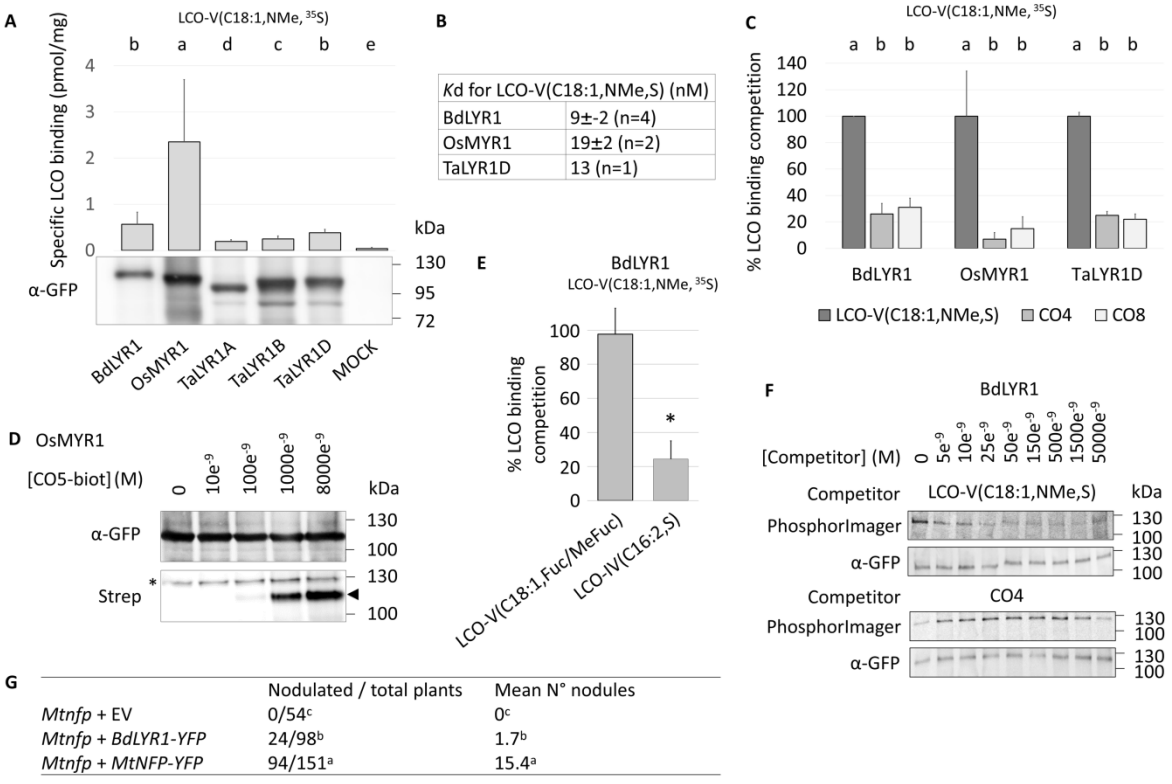

A
Mtlyr1-1 (A17, C58T)
MtLYR1 (590 AA)
Nter
Cter
Mtlyr1-2 (R108, Tnt1 insertion after 99 bp)
B
AMF colonization (% of root length)
a a a a
A17
Mtlyr1-1
Mtnfp
Mtlyr1-1/Mtnfp
AMF colonization (% of root length)
a a
R108
Mtlyr1-2
C
WT S. meliloti
Nodules / plant
a a b
A17
Mtlyr1-1
Mtnfp
Nodules / plant
a a
R108
Mtlyr1-2
D
S. meliloti nodFE
Nodules / plant
a a b
A17
Mtlyr1-1
Mtnfp
Nodules / plant
a a
R108
Mtlyr1-2

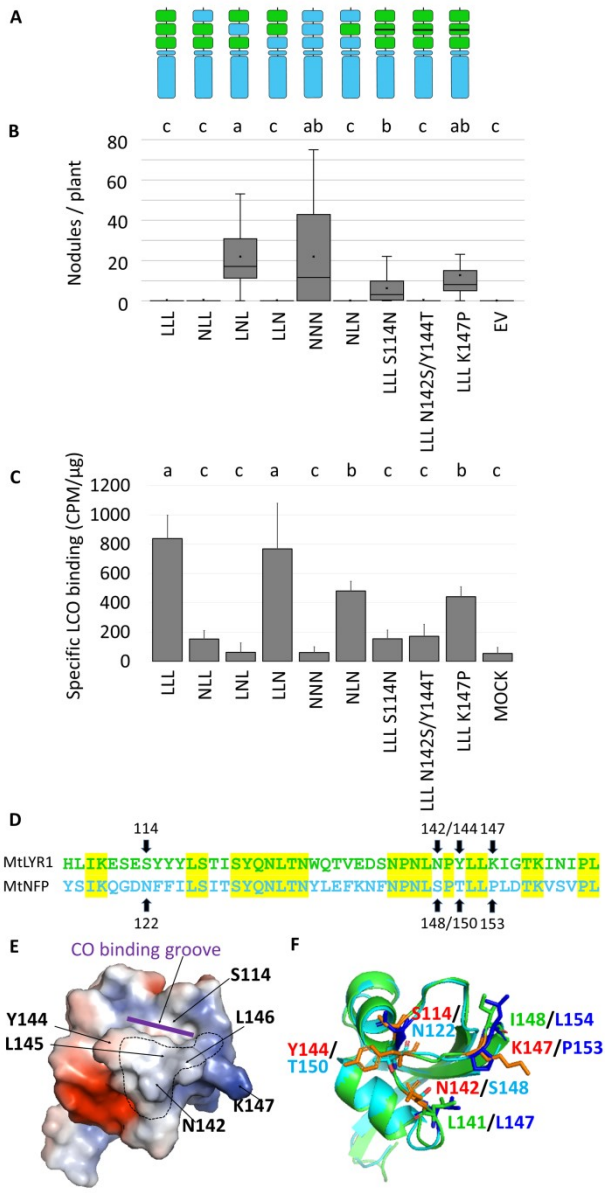
Fig. 3.



G

	Nodulated / total plants	Mean N° nodules
Mtnfp + EV	0/54[c]	0[c]
Mtnfp + *BdLYR1-YFP*	24/98[b]	1.7[b]
Mtnfp + *MtNFP-YFP*	94/151[a]	15.4[a]

Fig. 4.


A
B
c c a c ab c b c ab c
Nodules / plant
80
60
40
20
0
LLL
NLL
LNL
LLN
NNN
NLN
LLL S114N
LLL N142S/Y144T
LLL K147P
EV
C
a c c a c b c c b c
Specific LCO binding (CPM/µg)
1200
1000
800
600
400
200
0
LLL
NLL
LNL
LLN
NNN
NLN
LLL S114N
LLL N142S/Y144T
LLL K147P
MOCK
D
114
142/144 147
MtLYR1 HLIKESESYYYLSTISYQNLTNWQTVEDSNPNLNPYLLKIGTKINIPL
MtNFP YSIKQGDNFFILSITSYQNLTNYLEFKNFNPNLSPTLLPLDTKVSVPL
122
148/150 153
E
CO binding groove
S114
L146
Y144
L145
K147
N142
F
S114/
N122
I148/L154
Y144/
T150
K147/P153
N142/S148
L141/L147

REFERENCES

Ardourel, M., Demont, N., Debelle, F.D., Maillet, F., Debilly, F., Prome, J.C., et al. (1994) *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: 1357-1374.

Arrighi, J.F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L.C., Mirabella, R., et al. (2006) The *Medicago truncatula* lysin motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol* 142: 265-279.

Ben Amor, B., Shaw, S.L., Oldroyd, G.E.D., Maillet, F., Penmetsa, R.V., Cook, D., et al. (2003) The NFP locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant J* 34: 495-506.

Bensmihen, S., de Billy, F. and Gough, C. (2011) Contribution of NFP LysM domains to the recognition of Nod factors during the *Medicago truncatula*/*Sinorhizobium meliloti* symbiosis. *PLoS One* 6: e26114.

Boisson-Dernier, A., Chabaud, M., Garcia, F., Becard, G., Rosenberg, C. and Barker, D.G. (2001) Agrobacterium rhizogenes-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol Plant Microbe Interact* 14: 695-700.

Bouchiba, Y., Esque, J., Cottret, L., Maréchaux, M., Gaston, M., Gasciolli, V., et al. (2022) An integrated approach reveals how lipo-chitooligosaccharides interact with the lysin motif receptor-like kinase MtLYR3. *Protein Sci* 31: e4327.

Bozsoki, Z., Gysel, K., Hansen, S.B., Lironi, D., Krönauer, C., Feng, F., et al. (2020) Ligand-recognizing motifs in plant LysM receptors are major determinants of specificity. *Science* 369: 663-670.

Buendia, L., Girardin, A., Wang, T., Cottret, L. and Lefebvre, B. (2018) LysM receptor-like kinase and LysM receptor-like protein families: An update on phylogeny and functional characterization. *Front Plant Sci* 9: 1531.

Carotenuto, G., Chabaud, M., Miyata, K., Capozzi, M., Takeda, N., Kaku, H., et al. (2017) The rice LysM receptor-like kinase OsCERK1 is required for the perception of short-chain chitin oligomers in arbuscular mycorrhizal signaling. *New Phytol* 214: 1440-1446.

Chiu, C.H. and Paszkowski, U. (2020) Receptor-Like Kinases sustain symbiotic scrutiny. *Plant Physiol* 182: 1597-1612.

Cope, K.R., Bascaules, A., Irving, T.B., Venkateshwaran, M., Maeda, J., Garcia, K., et al. (2019) The Ectomycorrhizal fungus *Laccaria bicolor* produces lipochitooligosaccharides and uses the common symbiosis pathway to colonize *Populus* roots. *Plant Cell* 31: 2386-2410.

D'Haeze, W. and Holsters, M. (2002) Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiol* 12: 79-105.

Feng, F., Sun, J., Radhakrishnan, G.V., Lee, T., Bozsóki, Z., Fort, S., et al. (2019) A combination of chitooligosaccharide and lipochitooligosaccharide recognition promotes arbuscular mycorrhizal associations in *Medicago truncatula*. *Nat Commun* 10: 5047.

Fliegmann, J., Canova, S., Lachaud, C., Uhlenbroich, S., Gasciolli, V., Pichereaux, C., et al. (2013) Lipo-chitooligosaccharidic symbiotic signals are recognized by LysM receptor-like kinase LYR3 in the legume *Medicago truncatula*. *ACS Chem Biol* 8: 1900-1906.

Fliegmann, J., Jauneau, A., Pichereaux, C., Rosenberg, C., Gasciolli, V., Timmers, A.C., et al. (2016) LYR3, a high-affinity LCO-binding protein of *Medicago truncatula*, interacts with LYK3, a key symbiotic receptor. *FEBS Lett* 590: 1477-1487.

Gibelin-Viala, C., Amblard, E., Puech-Pages, V., Bonhomme, M., Garcia, M., Bascaules-Bedin, A., et al. (2019) The *Medicago truncatula* LysM receptor-like kinase LYK9 plays a dual role in immunity and the arbuscular mycorrhizal symbiosis. *New Phytol* 223: 1516-1529.

Girardin, A., Wang, T., Ding, Y., Keller, J., Buendia, L., Gaston, M., et al. (2019) LCO receptors involved in arbuscular mycorrhiza are functional for rhizobia perception in legumes. *Curr Biol* 29: 4249-4259.

Gomez, S.K., Javot, H., Deewatthanawong, P., Torres-Jerez, I., Tang, Y., Blancaflor, E.B., et al. (2009) *Medicago truncatula* and *Glomus intraradices* gene expression in cortical cells harboring arbuscules in the arbuscular mycorrhizal symbiosis. *BMC Plant Biol* 9: 10.

Gough, C., Cottret, L., Lefebvre, B. and Bono, J.J. (2018) Evolutionary history of plant LysM Receptor proteins related to root endosymbiosis. *Front Plant Sci* 9: 923.

Gough, C. and Jacquet, C. (2013) Nod factor perception protein carries weight in biotic interactions. *Trends Plant Sci* 18: 566-574.

Gressent, F., Drouillard, S., Mantegazza, N., Samain, E., Geremia, R.A., Canut, H., et al. (1999) Ligand specificity of a high-affinity binding site for lipo-chitooligosaccharidic Nod factors in *Medicago* cell suspension cultures. *Proc Natl Acad Sci USA* 96: 4704-4709.

Groves, P., Offermann, S., Rasmussen, M.O., Canada, F.J., Bono, J.J., Driguez, H., et al. (2005) The relative orientation of the lipid and carbohydrate moieties of lipochitooligosaccharides related to nodulation factors depends on lipid chain saturation. *Org Biomol Chem* 3: 1381-1386.

Gysel, K., Laursen, M., Thygesen, M.B., Lironi, D., Bozsóki, Z., Hjuler, C.T., et al. (2021) Kinetic proofreading of lipochitooligosaccharides determines signal activation of symbiotic plant receptors. *Proc Natl Acad Sci USA* 118: e2111031118.

He, J., Zhang, C., Dai, H., Liu, H., Zhang, X., Yang, J., et al. (2019) A LysM receptor heteromer mediates perception of arbuscular mycorrhizal symbiotic signal in rice. *Mol Plant* 12: 1561–1576.

Igolkina, A.A., Porozov, Y.B., Chizhevskaya, E.P. and Andronov, E.E. (2018) Structural insight into the role of mutual polymorphism and conservatism in the contact zone of the NFR5-K1 heterodimer with the Nod Factor. *Front Plant Sci* 9: 344.

Journet, E.P., Pichon, M., Dedieu, A., de Billy, F., Truchet, G. and Barker, D.G. (1994) *Rhizobium meliloti* Nod factors elicit cell-specific transcription of the ENOD12 gene in transgenic alfalfa. *Plant J* 6: 241-249.

Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K., et al. (2006) Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc Natl Acad Sci USA* 103: 11086-11091.

Khokhani, D., Carrera Carriel, C., Vayla, S., Irving, T.B., Stonoha-Arther, C., Keller, N.P., et al. (2021) Deciphering the Chitin Code in Plant Symbiosis, Defense, and Microbial Networks. *Annu Rev Microbiol* 75: 583-607.

Krönauer, C. and Radutoiu, S. (2021) Understanding Nod factor signalling paves the way for targeted engineering in legumes and non-legumes. *Curr Opin Plant Biol* 62: 102026.

Le Signor, C., Savois, V., Aubert, G., Verdier, J., Nicolas, M., Pagny, G., et al. (2009) Optimizing TILLING populations for reverse genetics in *Medicago truncatula*. *Plant Biotechnol J* 7: 430-441.

Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Herve, C., Toth, K., et al. (2010) A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proc Natl Acad Sci USA* 107: 2343-2348.

Lefebvre, B., Klaus-Heisen, D., Pietraszewska-Bogiel, A., Herve, C., Camut, S., Auriac, M.-C., et al. (2012) Role of N-Glycosylation sites and CXC motifs in trafficking of *Medicago truncatula* Nod Factor perception protein to plasma membrane. *J Biol Chem* 287: 10812-10823.

Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Prome, J.C., et al. (1990) Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* 344: 781-784.

Li, X.R., Sun, J., Albinsky, D., Zarrabian, D., Hull, R., Lee, T., et al. (2022) Nutrient regulation of lipochitooligosaccharide recognition in plants via NSP1 and NSP2. *Nat Commun* 13: 6421.

Liu, S., Wang, J., Han, Z., Gong, X., Zhang, H. and Chai, J. (2016) Molecular Mechanism for fungal cell wall recognition by rice chitin receptor OsCEBiP. *Structure* 24: 1192-1200.

Liu, T., Liu, Z., Song, C., Hu, Y., Han, Z., She, J., et al. (2012) Chitin-induced dimerization activates a plant immune receptor. *Science* 336: 1160-1164.

Luu, T.B., Ourth, A., Pouzet, C., Pauly, N. and Cullimore, J. (2022) A newly evolved chimeric lysin motif receptor-like kinase in *Medicago truncatula* spp. tricycla R108 extends its *Rhizobia* symbiotic partnership. *New Phytol* 235: 1995-2007.

Maillet, F., Poinsot, V., André, O., Puech-Pagès, V., Haouy, A., Gueunier, M., et al. (2011) Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature* 469: 58-63.

Malkov, N., Fliegmann, J., Rosenberg, C., Gasciolli, V., Timmers, A.C., Nurisso, A., et al. (2016) Molecular basis of lipo-chitooligosaccharide recognition by the lysin motif receptor-like kinase LYR3 in legumes. *Biochem J* 473: 1369-1378.

Masselin, A., Petrelli, A., Donzel, M., Armand, S., Cottaz, S. and Fort, S. (2019) Unprecedented affinity labeling of carbohydrate-binding proteins with s-Triazinyl Glycosides. *Bioconjug Chem* 30: 2332-2339.

Miyata, K., Hayafune, M., Kobae, Y., Kaku, H., Nishizawa, Y., Masuda, Y., et al. (2016) Evaluation of the role of the LysM receptor-like kinase, OsNFR5/OsRLK2 for AM symbiosis in rice. *Plant Cell Physiol* 57: 2283-2290.

Moling, S., Pietraszewska-Bogiel, A., Postma, M., Fedorova, E., Hink, M.A., Limpens, E., et al. (2014) Nod factor receptors form heteromeric complexes and are essential for intracellular infection in *Medicago* nodules. *Plant Cell* 26: 4188-4199.

Rasmussen, S.R., Füchtbauer, W., Novero, M., Volpe, V., Malkov, N., Genre, A., et al. (2016) Intraradical colonization by arbuscular mycorrhizal fungi triggers induction of a lipochitooligosaccharide receptor. *Sci Rep* 6: 29733.

Rush, T.A., Puech-Pagès, V., Bascaules, A., Jargeat, P., Maillet, F., Haouy, A., et al. (2020) Lipo-chitooligosaccharides as regulatory signals of fungal growth and development. *Nat Commun* 11: 3897.

Rutten, L., Miyata, K., Roswanjaya, Y.P., Huisman, R., Bu, F., Hartog, M., et al. (2020) Duplication of symbiotic Lysin motif receptors predates the evolution of nitrogen-fixing nodule symbiosis. *Plant Physiol* 184: 1004-1023.

Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., et al. (2007) Medicago LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiol* 145: 183-191.

Solovev, Y.V., Igolkina, A.A., Kuliaev, P.O., Sulima, A.S., Zhukov, V.A., Porozov, Y.B., et al. (2021) Towards understanding afghanistan pea symbiotic phenotype through the molecular modeling of the interaction between LykX-Sym10 receptor heterodimer and Nod Factors. *Front Plant Sci* 12: 642591.

Tadege, M., Wen, J., He, J., Tu, H., Kwak, Y., Eschstruth, A., et al. (2008) Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume *Medicago truncatula*. *Plant J* 54: 335-347.

Wais, R.J., Keating, D.H. and Long, S.R. (2002) Structure-function analysis of nod factor-induced root hair calcium spiking in *Rhizobium*-legume symbiosis. *Plant Physiol* 129: 211-224.

Yang, C., Wang, E. and Liu, J. (2022) CERK1, more than a co-receptor in plant-microbe interactions. *New Phytol* 234: 1606-1613.

Young, N.D., Debellé, F., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K., et al. (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480: 520-524.

Zhang, C., He, J., Dai, H., Wang, G., Zhang, X., Wang, C., et al. (2021) Discriminating symbiosis and immunity signals by receptor competition in rice. *Proc Natl Acad Sci USA* 118: :e2023738118.

Zipfel, C. and Oldroyd, G.E. (2017) Plant signalling in symbiosis and immunity. *Nature* 543: 328-336.