New Phytologist

Triarabinosylation is required for nodulation-suppressive CLE peptides to systemically inhibit nodulation in *Pisum sativum*

Journal:	*New Phytologist*
Manuscript ID	Draft
Manuscript Type:	MS - Regular Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Hastwell, April; University of Queensland, Centre for Integrative Legume Research, School of Agriculture and Food Science Corcilius, Leo; University of Sydney, School of Chemistry Williams, James; University of Sydney, School of Chemistry Gresshoff, Peter; University of Queensland, Centre for Integrative Legume Research, School of Agriculture and Food Science Payne, Richard; University of Sydney, School of Chemistry Ferguson, Brett; The University of Queensland, Centre for Integrative Legume Research, School of Agriculture and Food Science
Key Words:	Glycopeptide, Legume, Peptide Hormone, Plant Signalling and Development, Symbiosis, CLE Peptide

Triarabinosylation is required for nodulation-suppressive CLE peptides to systemically inhibit nodulation in *Pisum sativum*

Authors: April H. Hastwell[^a], Leo Corcilius[^b], James Williams[b], Peter M. Gresshoff[a], Richard J. Payne[b*] and Brett J. Ferguson[a*]

Affiliation: [a]Centre for Integrative Legume Research, School of Agriculture and Food Sciences, The University of Queensland, St Lucia, Brisbane, Queensland, 4072, Australia. [b]School of Chemistry, The University of Sydney, Sydney, NSW 2006, Australia. [^]These authors contributed equally to this manuscript.

***Corresponding Authors**: Brett J. Ferguson, address as above, Phone: +61-7-3346 9951, Fax: +61-7-3365-3556, email: b.ferguson1@uq.edu.au. Richard J Payne, address as above, Phone: +61-2-9351 5877, email: richard.payne@sydney.edu.au.

Running Head: Tri-arabinosylated CLE peptides in nodulation

Keywords: Glycopeptide, Legume, Peptide Hormone, CLE peptide, Plant Signalling and Development, Symbiosis

Word Count: 8079

Total: 3930

Introduction: 899

Materials and Methods: 629

Results: 1748

Discussion: 771

Acknowledgements: 102

Figures: 6 Figures, Figures 1,2, 5 and 6 coloured

Supplementary Figures and Tables: 6 Supplementary Figures, 2 Supplementary Tables, 1 Supplementary Information

Summary

- Legume plants form root nodules to house beneficial nitrogen-fixing rhizobia bacteria. Forming and maintaining nodules is resource demanding; hence, legumes evolved a systemic signalling mechanism, called Autoregulation of Nodulation (AON), to control nodule numbers. AON begins with the production of CLE peptides in the root, which are predicted to be glycosylated, transported to the shoot, and perceived by a receptor complex. We synthesised variants of nodulation-suppressing CLE peptides to test their activity in feeding studies.
- A novel petiole feeding method was used to introduce nodulation-suppressing CLE peptides into the shoot. Hydroxylated, monoarabinosylated and triarabinsoylated variants of soybean GmRIC1a and GmRIC2a were chemically synthesised and fed into recipient plants. *Pisum sativum* (pea) plants were used due to the availability of key mutants in the AON pathway.
- Triarabinosylated GmRIC1a and GmRIC2a suppressed nodulation of wild-type pea plants. No other variant tested suppressed nodulation. Suppression also occurred in the supernodulating hydroxyproline *O*-arbinosyltransferase mutant, *Psnod3*, but not in the supernodulating receptor mutants, *Pssym29*, and to some extent, *Pssym28*. Bioinformatic analyses identified 40 CLE peptide-encoding pea genes, including nodulation-suppressive orthologues which are highly similar to those of soybean.
- GmRIC1a and GmRIC2a require triarabinosylation to exert their activity, and this modification is likely facilitated by PsNOD3.

Introduction

Legumes are important in agriculture systems as a means to alleviate nitrogen fertiliser inputs, thus reducing fossil fuel use, fertiliser run-off and toxic gas emissions (Gresshoff *et al*., 2015; Foyer *et al*., 2016). They also promote soil health by increasing nitrogen levels through a mutualistic symbiotic relationship with bacteria (collectively known as rhizobia) that can convert atmospheric nitrogen gas (N_2) into a form of nitrogen the plant can use (NH_4^+). Agricultural practices take advantage of this, with legumes often used as rotation or cover crops (Jensen *et al*., 2012). Although the symbiosis is beneficial, the host plant regulates the number of nodules it forms as a means of balancing its need for nitrogen with its ability to expend resources forming and maintaining nodule structures. Thus, legumes have complex molecular signalling cascades to control nodulation (Ferguson *et al*., 2010; Reid *et al*., 2011a).

A systemic negative feedback signalling pathway that provides legumes with control over their nodule numbers is known as Autoregulation of Nodulation (AON; Kosslak and Bohlool 1984; Delves *et al*., 1986; Reid *et al*., 2011a). The AON pathway begins in response to initial rhizobia infection events, with the production of CLAVATA3/Endosperm Surrounding Region (ESR) related (CLE) peptides. In soybean, these peptides are GmRIC1 and GmRIC2 (Reid *et al*., 2011b), with orthologues in other legumes having also been identified (Okamoto *et al*., 2009; Mortier *et al*., 2010 Reid *et al*., 2011b; Ferguson *et al*., 2014; Nishida *et al*., 2016). While there is no clear distinction between the biological role of GmRIC1 and GmRIC2, there is some temporal separation in their expression patterns (Reid *et al*., 2011a). The AON CLE peptides are produced in the root, post-translationally modified (Okamoto *et al*., 2013; Kassaw *et al*., 2017), then transported to the shoot where they are perceived by a leucine-rich repeat receptor kinase, called GmNARK in soybean (known orthologues include PvNARK, LjHAR1, MtSUNN, PsSYM29, and GsNARK; Krusell *et al*., 2002; Nishimura *et al*., 2002; Searle et al, 2003; Schnabel *et al*., 2005; Ferguson *et al*., 2014). CLV2/SYM29 and KLAVIER are proposed to form a heterodimeric complex with NARK (which might also form a homodimer complex) to perceive the CLE peptides, with mutations in either NARK or its dimerisation partners resulting in supernodulation (Miyazawa *et al*., 2010; Ferguson *et al*., 2010; Krusell *et al*., 2011). Interestingly, the homeologous duplicate of GmNARK, called GmCLV1A, has no role in nodulation control, but instead functions in regulating shoot architecture, indicating that one of the genes has undergone the process of neofunctionalisation (Mirzaei *et al*., 2017). Following ligand binding by GmNARK, a shoot-derived signal that is transported to the root to inhibit further nodulation events is differentially regulated (Lin *et al*., 2010; Ferguson *et al*., 2010; Sasaki *et al*., 2014). This signal might act through the Kelch-Repeat F-box factor Too Much Love (TML), to regulate nodulation, as mutations in its gene also lead to a lack of nodulation control (Magori *et al*., 2009).

CLE peptides are 12-13 amino acids long, with the majority containing a central proline residue that is post-translationally hydroxylated and further modified with a triarabinose moiety containing β1,2 linkages (Ferguson and Mathesius 2014; Hastwell *et al*., 2015a). When synthetic CLE peptides possess this glycan, binding efficiency is increased (AtCLE3; Shinohara *et al*., 2013) and they exhibit increased biological activity (LjCLE-RS2, Okamoto *et al*., 2013; GmCLE40a, Corcilius *et al*., 2017). This modification is facilitated by an arabinosyltransferase (called MtRDN1/PsNOD3 in the case of the AON CLE peptides; Schnabel *et al*., 2011). Interestingly, only one rhizobia-induced CLE peptide of *M. truncatula*, MtCLE12, appears to require arabinosylation by MtRDN1, whereas MtCLE13 does not (Kassaw *et al*., 2017).

A similar mechanism to AON, called the nitrate-regulation of nodulation pathway, acts locally and is induced by soil nitrate to enable the plant to inhibit nodulation when ample nitrogen is available (Reid *et al*., 2011a). This nitrate-regulation of nodulation pathway begins with the production of

nitrate-induced CLE peptides (called GmNIC1a and its duplicate GmNIC1b in soybean) which are perceived by the GmNARK receptor located in the root (Reid *et al*., 2011b; Lim *et al*., 2014). CLE peptides induced specifically by nitrate have not been reported in other legumes, but in *L*. *japonicus* the rhizobia-induced CLE peptides *LjCLE-RS2*, *LjCLE-RS3* and *LjCLE40* are reported to exhibit increased expression with nitrate application (Okamoto *et al*., 2009; Nishida *et al*., 2016).

Here, we report that triarabinosylated GmRIC1a and GmRIC2a of soybean suppress nodulation in pea. This was demonstrated using petiole feeding of peptides that were synthesised by solid-phase peptide synthesis (SPPS) using a synthetic β1,2 triarabinosylated hydroxyproline glycosylamino acid building block (Corcilius *et al*., 2017) to site selectively incorporate the glycan at position seven of the CLE domain. Using mutant plants, we showed that the suppressive activity required the PsSYM28 and PsSYM29 receptors, but acted downstream of the PsNOD3 arabinosyltransferase that post-translationally glycosylates the peptides. Chemically synthesised variants of GmRIC1a and GmRIC2a that were either hydroxylated-only or partial glycosylated were unable to suppress nodulation, confirming that triarabinosylation is required for the peptides to function in AON. Subsequently, pea orthologues of the nodulation-suppressive CLE peptides were determined from 40 CLE peptide-encoding gene family members identified in this study. The CLE peptide domains of these pea orthologues were almost identical to those of the soybean peptides fed in this study. Taken together, our findings demonstrate a clear requirement for GmRIC1a and GmRIC2a to be post-translationally modified with a triarabinosylated hydroxyproline moiety to exert their nodulation-suppressive activity.

Materials and Methods

Plant and bacterial growth

Wild-type and mutant *Pisum sativum* (pea) cv Frisson seeds (Postma *et al*., 1988; Duc and Messager *et al*., 1989; Sagan and Duc 1996; Li *et al*., 2009) were sterilised with 70% w/v ethanol before being imbibed with autoclaved Milli-Q® water. Imbibed seeds were germinated in 4 L euro pots with sterile Grade 3 vermiculite topped with approximately 3 cm of autoclaved UQ23 Mix (Central Glasshouse Services, University of Queensland, Australia) to assist germination. All plants were grown in either a E-75L1 or PGC-9/2 growth chamber (Percival Scientific, Perry, IA, USA) under 25°C:23°C, 12 hour day:night conditions. The short-day length condition induced longer internodes to assist with petiole feeding. Plants were watered as required with B & D nutrient solution (Broughton and Dilworth, 1971), supplemented with 1 mM KNO_3.

Rhizobium leguminosarum RLV248 was grown in liquid yeast mannitol broth (Somerville and Kahn, 1983) at 28°C for 36 hours and diluted to OD=0.1 with either ddH_2O or B & D nutrient solution. Approximately 250 mL of inoculum was applied to each pot 48 hours after petiole feeding commenced and nodule number was counted 14 days after inoculation.

Petiole feeding

Petiole feeding was carried out as per Lin *et al*., (2010; 2011) with the following modifications. The second petiole of three-week old pea plants was used in the first instance to attach the petiole feeding apparatus. The apparatus consisted of a 3 mL syringe barrel attached to 20 mm of clear silicone tubing having a 2.6 mm internal diameter. This was subsequently connected to 4 cm of silicone tubing having a 1.6 mm internal diameter, which was an appropriate size for attaching to the petiole of the pea plants. The petiole was severed behind the first leaflet, and the basal stipules

were left intact and used to help seal the petiole-tubing junction. After one week of feeding, the petioles became chlorotic and the feeding solution (control or peptide) ceased to be taken up by the plant. Thus, a fresh feeding apparatus was attached to a new petiole (usually two higher than the originally-fed petiole). To prevent any loss of peptide solution due to leakage, approximately 500 µL of autoclaved Milli-Q® water was injected into the silicone tubing of the newly attached feeding apparatus and left for 30 minutes prior to adding peptide solutions. Blue food colouring was used in preliminary studies to visualise uptake and ensure solutions were distributed throughout the plant.

Chemical synthesis of GmRIC1a and GmRIC2a (glyco)peptides

GmRIC1a and GmRIC2a peptides were synthesized via solid-phase peptide synthesis (SPPS) according to a previously reported procedure (Corcilius *et al*., 2017). Six synthetic peptides were prepared in total, each containing hydroxyproline at position 4, and either hydroxyproline, *O*-(β-L-arabinofuranosyl) hydroxyproline (monoarabinosylated hydroxyproline) or *O*-[β-(β1,2-tri-L-arabinofuranosyl)] hydroxyproline (triarabinosylated hydroxyproline) at position 7 of the CLE domain. Synthetic peptides were purified by reversed phase HPLC and characterized by analytical HPLC and both low and high resolution ESI-MS (+ve ion) (see Supporting Information for synthetic peptide characterization data).

Sequence identification and bioinformatic analysis

CLE peptide encoding genes in *Pisum sativum* were identified using BLAST searches of known legume genes identified in Hastwell *et al*., (2015b and 2017) as well as those from *Arabidopsis thaliana* (Cock and McCormick 2001) with E value = 1 (Altschul *et al*., 1997 and 2005). The searches were conducted in The Pea RNA-Seq gene atlas (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi; Alves-Carvalho *et al*., 2015). Multiple Sequence Alignments, logo diagrams, signal peptide and phylogenetic analyses were performed as per Hastwell *et al*., (2015b and 2017).

Statistical analyses

Student's *t*-tests were used to determine statistical differences between treatments and were calculated in GraphPad Prism 7.01 (La Jolla California, USA; *P<0.5, **P<0.01, ***P<0.001). Data are expressed as a mean ±SEM, with n = 6 to 8 plants per treatment, except for untreated plants where n = 14.

Results

Establishment of petiole feeding as a method to introduce solutions into pea plants

During AON, root-derived CLE peptides travel in the xylem to the shoot, where they are perceived by an LRR receptor kinase (Searle *et al*., 2003; Reid *et al*., 2011b; Okamoto *et al*., 2013). However, feeding CLE peptides to the root can have unwanted false-positive effects, with many inhibiting root growth due to functional redundancy and interacting with other receptors (Whitford *et al*., 2008; Shinohara and Matsubayashi 2015). Thus, a direct-feeding method to introduce the peptide closer to its correct receptor was desired. Petiole feeding achieves this (Lin *et al*., 2010, 2011), and pea was selected as the recipient species due to the availability of multiple pea mutants in the AON pathway. When this study commenced, CLE peptide sequences of pea were not available. We therefore focused on GmRIC1a and GmRIC2a of soybean as they have been shown to act interspecifically in other legume species using overexpression studies (Ferguson *et al*., 2014).

Preliminary experiments feeding water or dye revealed no observable differences in shoot or root weight, shoot height or node number between intact and petiole-fed pea plants (Figure 1A, Supplementary Figure 1). This confirmed that petiole feeding could be used to introduce and translocate solutions throughout the plant, and did not induce unwanted effects, which is consistent with previous reports using other plant species (Lin *et al*., 2010, 2011).

Chemical synthesis of GmRIC1a and GmRIC2a glycopeptide variants

Methods to extract and purify sufficient quantities of endogenous CLE glycopeptides have not been established and therefore chemical synthesis is the only tool available to access CLE glycopeptides for feeding studies. However, this is a considerable undertaking when post-translational modifications are taken into account because of the synthetically-challenging nature of the glycan (Kaeothip and Boons 2013). Despite this challenge, two successful syntheses of an SPPS-compatible triarabinosylated hydroxyproline 'building block' have been reported (Shinohara and Matsubayashi 2013; Kaeothip *et al*., 2013) along with examples of its incorporation into native CLE peptides (Shinohara and Matsubayashi 2013, Okamoto *et al*., 2013; Xu *et al*., 2015). An improved protocol for the synthesis of this triarabinosylated hydroxyproline building block (Figure 2, in box) was recently reported (Corcilius *et al*., 2017), and used in this study to access multi-milligram quantities of homogeneous hydroxyproline-7 triarabinosylated GmRIC1a and GmRIC2a glycopeptides. Briefly, the building block was incorporated into conventional Fmoc-SPPS protocols to obtain the resin-bound and side chain-protected glycopeptides, which were subsequently liberated from the resin and deprotected through treatment with an acidic cleavage cocktail containing trifluoroacetic acid (TFA), triisopropylsilane and water. After deacetylation of the glycan with sodium methoxide in methanol, the residues were purified by preparative reversed phase HPLC affording GmRIC1a and GmRIC2a glycopeptides as their corresponding trifluoroacetate salts in 17% and 28% overall yield, respectively (yield based on initial resin loading of the C-terminal amino acid). The corresponding hydroxyproline-7 monoarabinosylated and unglycosylated variants were also synthesised in order to probe the functional importance of the triarabinosylation modification (Figure 2). All variants were prepared with hydroxyproline at position 4 in analogy with the structures of known CLE peptides.

GmRIC1a and GmRIC2a glycopeptides suppress nodulation in pea

Petiole feeding was used to determine whether GmRIC1a and GmRIC2a peptide variants could inhibit nodulation in pea. Soybean CLE peptides were used, rather than those of pea, as the transcriptome database enabling identification of the pea CLE peptide-encoding gene sequences was not available when this study commenced. The variants tested had the proline residues at positions four and seven hydroxylated, with or without triarabinosylation at position seven (Figure 2), and were fed at concentrations from 1 pM to 10 µM. CLE peptides with no modifications have previously been reported to have no nodulation-suppressive activity and were not used in this study (Okamoto *et al*., 2009; Mortier *et al*., 2010).

Nodule inhibition was observed in plants fed with 1 µM or higher of the triarabinosylated variant of GmRIC1a or GmRIC2b (Figure 1B, Supplementary Figure 2). In contrast, no significant difference in nodule number was observed with any concentration of the hydroxylated-only variant (Figure 1C, Supplementary Figures 2 and 3). This indicates that triarabinosylation is required for the peptide to exert its activity.

The extent of glycosylation can affect the efficacy of CLE peptide activity

All CLE peptides identified to date have been modified with three linked arabinose sugars at their central proline residue. To determine whether these three arabinose sugars are required to suppress

nodulation, wild-type pea plants were fed with either the triarabinosylated or monoarabinosylated variant of GmRIC1a. While the triarabinosylated variant significantly suppressed nodulation (Figure 1D), the monoarabinosylated variant was unable to do so (P>0.5). This further demonstrates that post-translational triarabinosylation is essential for activity.

Nodulation suppressing CLE peptides act downstream of PsNOD3 but require PsSYM28 and PsSYM29 to exert their activity

PsNOD3 encodes a hydroxyproline *O*-arabinosyltransferase (Schnabel *et al*., 2011) that might be required to post-translationally glycosylate mature, nodulation-suppressing CLE peptides in the root. *PsSYM28* and *PsSYM29* encode for receptors that likely form a complex to perceive nodulation-suppressing CLE peptide ligands in the shoot (Krusell et al 2002, 2011). Overexpression of rhizobia-induced CLE peptide-encoding genes results in complete suppression of nodulation in wild-type plants of several legumes (Okamoto *et al*., 2009; Mortier *et al*., 2010; Reid *et al*., 2011b), but does not alter nodule numbers in supernodulating receptor mutants (Okamoto *et al*., 2009; Reid *et al*., 2011b; Osipova *et al*., 2012; Ferguson *et al*., 2014). Interestingly, *MtCLE13* overexpression suppresses nodulation in *Mtrdn1-2*, the orthologue of *PsNOD3*, but not when interspecifically overexpressed in *Psnod3* (Osipova *et al*., 2012; Kassaw *et al*., 2017). To establish whether triarabinosylated GmRIC1a or GmRIC2a can suppress nodulation in supernodulating pea mutants, plants were fed via petiole feeding and nodule numbers determined. Soybean was not utilised as there are currently no lines containing mutations in SYM28 and NOD3 orthologues.

Nodule numbers were not affected in *Pssym29* plants fed with GmRIC1a (Figure 3A) and only a slight but significant reduction in nodulation was observed in *Pssym28* plants (Figure 3A). In contrast, nodulation was significantly reduced when feeding GmRIC2a into *Psnod3* plants (Figure 3B). Together, this indicates that SYM29, and to a lower extent SYM28, are required for perception of the nodulation CLE peptides, and that NOD3 is indeed likely responsible for arabinosylation of the peptides, which is required for their function.

Functional redundancy enables other CLE peptide family members to function as nodulation-suppressing CLE peptides

To determine whether other CLE glycopeptides could mimic the activity of the nodulation suppressing CLE peptides, petiole feeding was used to introduce hydroxylated-only or triarabinosylated GmCLE40a variants into wild-type pea plants. GmCLE40a acts to regulate the stem cell population of the root apical meristem (Corcilius *et al*., 2017) and would not normally come into contact with receptors of the nodulation suppressing CLE peptides. The CLE domain of GmCLE40a contains six amino acid residues that differ from the GmRIC1a or GmRIC2a CLE domain. Only two of these residues (positions three and twelve) affected GmRIC1a activity when modified via site-directed mutagenesis (Reid *et al*., 2013). This reduction in activity was only minor at position three, and the residue at position 12 of GmCLE40a would only be considered a conservative change from that of GmRIC1a (Asp>His), and thus not likely to have a large impact on activity.

The hydroxylated GmCLE40a variant was not able to suppress nodulation (Figure 4), similar to what was observed with hydroxylated GmRIC1a. However, triarabinosylated GmCLE40a did suppress nodulation. In fact, it suppressed nodulation to nearly the same extent as triarabinosylated GmRIC1a (Figure 1). These findings demonstrate functional redundancy can occur amongst CLE peptides, and further support the conclusion that triarabinosylation of the nodulation suppressing CLE peptides is required to suppress nodulation in pea.

Identification of CLE peptide-encoding genes of Pisum sativum

The complete genome of pea is not yet available and so we used the nodulation suppressing CLE peptides of soybean in this study. However, since commencing our work, several transcriptome analyses have become available that could be used to identify CLE peptide encoding genes of pea (Alves-Carvalho *et al*., 2015; Tayeh *et al*., 2015). To identify CLE peptide orthologues of pea, BLAST searches of the UniGene set in The Pea RNA-Seq gene atlas were conducted using CLE peptide-encoding gene sequences of *Medicago truncatula*, *Lotus japonicus*, *Phaseolus vulgaris* and *Arabidopsis thaliana* (Cock and McCormick 2001; Alves-Carvalho *et al*., 2015; Hastwell *et al*., 2015b and 2017). The search yielded 40 unique CLE peptide-encoding gene candidates of pea (Figure 5, Supplementary Table 1) and a further eight sequences with unclear gene structures and/or analogous CLE peptide domains (Supplementary Table 2). Three of the identified sequences contain multiple CLE domains (Supplementary Table 1, Supplementary Figure 4).

An initial phylogenetic tree was constructed using the 40 newly identified CLE prepropeptide sequences of pea, along with those previously identified in *M*. *truncatula*, *L*. *japonicus*, *P*. *vulgaris* and *A*. *thaliana* (Supplementary Figure 5) (Cock and McCormick 2001; Hastwell *et al*., 2015b and 2017). This enabled homologous sequences of pea to be identified. PsCam040153 and PsCam040702 grouped closely with rhizobia-induced CLE peptides, and PsCAM041632 grouped with nitrate-induced CLE peptides (Supplementary Figure 5). An additional phylogenetic tree focusing on nodulation-suppressing CLE peptides was then generated, which included both rhizobia- and nitrate-induced CLE peptides of *G. max* and other legumes (Figure 6A) (Reid *et al*., 2011b; Okamoto *et al*., 2015). Unsurprisingly, PsCam040153 and PsCam040702 formed a distinct branch with the rhizobia-induced CLE peptide orthologues, whereas no clear branch was observed with the nitrate-induced CLE peptides, despite it grouping in the phylogenetic tree designed using with the complete family of pea CLE prepropeptides (Figure 6A).

Based on the sequence and phylogenetic analyses, PsCam040153 and PsCam040702 are the likely orthologues of the rhizobia-induced CLE peptides (GmRIC1, GmRIC2, PvRIC1, PvRIC2, MtCLE12, MtCLE13, LjCLE-RS1, LjCLE-RS2 and LjCLE-RS3). Given that the CLE domain within the prepropeptide represents the functional ligand, the amino acid sequences within that domain were compared to those of previously identified orthologues (Figure 6B). PsCam040153 and PsCam040702 have CLE domains that are conserved at six and seven of the eight residues, respectively, that were identified by Reid *et al*., (2013) as being critical to the activity of nodulation-suppressive CLE peptides (Supplementary Figure 6). However, the two non-conserved residue changes at positions three and eight of the CLE domain are conservative, Ala3>Ser3 in both sequences and Asn8>Asp8 in only PsCam040702 (Supplementary Figure 6). The former is an amino acid found at position three of the majority of CLE domains from the nodulation-suppressive CLE peptides of *M. truncatula* and *L. japonicus*. Hence, these differences seem very unlikely to impact function.

Discussion

The fundamental mechanisms that provide legumes with control over nodulation requires a better understanding to enable agricultural advances. Using synthetic variants of the nodulation-suppressing CLE peptides, GmRIC1a and GmRIC2a, we show that post-translational modification with a triarabinose moiety is required for activity. These findings are consistent with the literature that demonstrates that glycosylation is essential for the activity of nodulation-suppressive CLE peptides in *L. japonicus*, CLV3 orthologues in *A. thaliana* and tomato (Shinohara and Matsubayashi 2013, Okamoto *et al*., 2013; Xu *et al*., 2015), and CLE40 in soybean (Corcilius *et al*., 2017).

Activity of the triarabinosylated GmRIC1a and GmRIC2a peptides is dependent on *PsSYM29*, and to some extent *PsSYM28*, which are the proposed receptors of the nodulation suppressing CLE peptides. This is consistent with over-expression studies, where the peptides act through these receptors to inhibit nodule numbers (Krusell *et al*., 2011; Osipova et al 2012; Ferguson *et al*., 2014). These findings agree with the proposed AON pathway, where the CLE peptides are perceived by a receptor complex consisting of NARK/CLV2/KLV, which triggers regulation of a downstream signal that induces nodule number regulation.

In addition to suppressing nodulation in wild-type pea, the glycosylated GmRIC1a and GmRIC2a peptides also inhibited nodulation in *Psnod3* mutants. NOD3 orthologues are hydroxyproline *O*-arbinosyltransferases proposed to be responsible for catalysing the glycosylation of some CLE peptides in AON (Kassaw *et al*., 2017). Our study supports the requirement for arabinosylation by NOD3; however, the precise role of the modification remains unknown and may be required for structure, perception and/or stability of the peptide (Shinohara and Matsubayashi 2013). It is also possible that another mechanism for modification is required for other CLE peptides in nodulation (Kassaw *et al*., 2017).

The CLE peptide-encoding genes identified here considerably enhance our knowledge of the CLE peptide family of pea. The 40 genes identified include orthologues of the well-characterised CLE peptides RIC1, RIC2, NIC1, TDIF and multiple CLE domain containing prepropeptides of other species. Official gene nomenclature was not assigned to the newly identified pea genes as it is highly likely that new CLE peptide-encoding genes will be identified once the pea genome is released. When this occurs, a much more comprehensive study will be required, similar to Hastwell *et al*., (2015, 2017), as current gene-identifying resources are limited to tissues and treatments that were captured to generate the Pea RNA-Seq gene atlas.

Within a species, different CLE peptide-encoding genes can encode for the same mature peptide sequence, with functional specificity arising from temporal and spatial separation of gene expression in conjunction with divergent receptors. Our findings indicate that synthetic triarabinosylated GmCLE40a, which has high sequence similarity to GmRIC1a and GmRIC2a, can function in AON to suppress nodulation; however, endogenous GmCLE40a is highly unlikely come into contact with AON receptors as it is a component of root apical meristem development (Yamaguchi *et al*., 2016; Corcilius *et al*., 2017). This finding highlights the need to plan feeding studies and interpret their results with great care, similar to what has been reported for receptor binding studies that can generate false-positive outcomes (Shinohara and Matsubayashi 2015).

The available germplasm in pea made this study possible as there are multiple mutants available in the AON pathway. In contrast, the duplicated genome of soybean results in functional redundancy that makes selecting mutant lines a challenge. Typically, mutations in soybean are only isolated for duplicate genes that have undergone neofunctionalisation, such as GmNARK and GmCLV1a (Mirzaei *et al*., 2017). Difficulty in creating stable mutant lines also restricts mutant availability. At the beginning of this study the CLE peptide-encoding genes of pea had not been identified and the pea genome was not available to identify them. However, interspecific studies had concluded that the AON mechanism is conserved across legumes (Osipova *et al*., 2012; Ferguson *et al*., 2014) and so the nodulation suppressing CLE peptides of soybean were used with the AON mutants of pea. Subsequently, we were able to identify the likely othologues of the nodulation suppressing CLE peptides of pea, and established that their mature peptide sequences are highly similar to those of soybean.

CLE peptides are important plant hormones that may provide targets for agricultural advances in nodulation as well as other aspects of plant growth and development. It is important to better understand their patterns of expression, post-translational modifications, and function in molecular signalling pathways. Using recently advanced chemical methods (Corcilius *et al*., 2017), we demonstrate that this can be achieved using homogeneously modified CLE peptides coupled with precise delivery techniques to reduce off-target effects. Further understanding CLE peptides and how they are post-translationally modified by NOD3 is pertinent to expand our knowledge of AON and associated pathways.

Acknowledgements

This work was funded by the Hermon Slade Foundation, and Australian Research Council Discovery Project grants (DP130103084 and DP130102266) to BJF and PMG. This work was supported by an Australian Research Council Future Fellowship (FT130100150) to RJP. The Fellowship Fund Inc. is thanked for provision of a Molly-Budtz Olsen PhD fellowship to AHH. We also gratefully acknowledge the funding provided to LC by the John A. Lamberton research scholarship and the Agnes Campbell postgraduate prize. We would also like to thank Andreas Brust, Rob Capon, Dongxue Li, Laura Haaima, Huanan Su, Xitong Chu and Mengbai Zhang for their assistance with this study.

Author Contribution

A.H.H., P.M.G., R.J.P., and B.J.F. conceived the synthetic peptide targets. A.H.H., P.M.G., and B.J.F. designed the plant experiments. L.C. and J.W. synthesised peptide variants and performed compound characterisation. A.H.H. conducted the plant experiments. A.H.H., L.C., R.J.P., and B.J.F. wrote the manuscript with the assistance of all authors. All authors participated in data analysis and discussions.

Figure legends

Figure 1. Nodule number 14 days after inoculation of wild-type pea plants fed via petiole feeding. **A** Image of pea plants with petiole feeding apparatus attached (arrow). **B** 1 µM triarabinosylated GmRIC1a, GmRIC2a and water control. **C** 1 pM to 10 µM of hydroxylated (Hyd) or triarabinosylated (Tri) GmRIC1a. **D** 1 µM triarabinosylated (Tri), monoarabinosylated (Mono) or hydroxylated (Hyd) GmRIC1a, and water control. Statistical differences determined using Student's *t*-test. n = 7 to 10 plants per treatment.

Figure 2. Structures of synthetic GmRIC1a and GmRIC2a peptides, and triarabinosylated hydroxyproline building block (in box). Proline 4 is hydroxylated in all variants. Proline 7 is either hydroxylated only (R = H), or further modified by arabinosylation (R = monoarabinose or triarabinose).

Figure 3. Nodule number 14 days after inoculation of wild-type and nodulation-mutant pea plants fed via petiole feeding with either 1 µM triarabinosylated GmRIC1a, triarabinosylated GmRIC2a, or water control. **A** Wild type, *sym28* and *sym29* plants fed with GmRIC2a. **B** Wild type and *nod3* plants fed with GmRIC1a. Statistical differences determined using students t-test. n = 5-8 plants per treatment.

Figure 4. Nodule number 14 days after inoculation of wild-type pea plants fed via petiole feeding with 1 µM hydroxylated (Hyd) or triarabinosylated (Tri) GmRIC1a or GmCLE40a, or water control. Statistical differences determined using students t-test. n = 7 to 8 plants per treatment.

Figure 5. Multiple Sequence Alignment of the CLE prepropeptides of *P. sativum*. Shaded nucleotides indicate conservation. Not shown are the multi CLE domain containing prepropeptides.

Figure 6. Nodulation-suppressive CLE prepropeptides of *P. sativum* and their orthologues in *G. max*, *P. vulgaris*, *L. japonicus*, and *M. truncatula*. **A** Multiple Sequence Alignment where grey residues indicate conserved residues. **B** Phylogenetic tree, with Bootstrap confidence values expressed as a percentage of 1,000 bootstrap replications, using AtCLV3 as an outgroup.

Supplementary Figure 1. Nodule number of untreated or petiole-fed plants 14 days after inoculation with rhizobia. No statistical differences were observed using the Student's *t*-test (P=0.52) n = 8 to 14 plants per treatment.

Supplementary Figure 2. Phenotype 14 days after inoculation of wild-type pea plants fed via petiole feeding with 1 µM triarabinosylated GmRIC1a, GmRIC2a or water control. **A** Nodule number per g of fresh shoot weight. **B** Fresh shoot weight **C** Dried root weight. Statistical differences determined using Student's *t*-test. n = 8 to 10 plants per treatment.

Supplementary Figure 3. Nodule number 14 days after inoculation of wild-type pea plants fed via petiole feeding with 1 µM or 100 nM triarabinosylated GmRIC2a compared with water control. Statistical differences determined using Student's *t*-test. n = 8 to 14 plants per treatment.

Supplementary Figure 4. Multiple Sequence Alignment of the multi CLE domain containing CLE prepropeptides identified in *P. sativum*. The CLE domains of the prepropeptides are highlighted by a red box.

Supplementary Figure 5. Phylogenetic tree of the CLE prepropeptides of *P. sativum*, *M. truncatula*, *L. japonicus*, *A. thaliana* and *P. vulgaris,* including AtRGF1 as an outgroup. Bootstrap confidence values are expressed as a percentage from 100 bootstrap replications.

426 **Supplementary Figure 6**. Sequence Logo Diagrams of orthologous CLE domains from *G. max*, *P.*
427 *vulgaris*, *L. japonicus*, *M. truncatula* and *P. sativum*. **A** Rhizobia-induced CLE peptides. **B** Nitrate-
428 induced CLE peptide. LjCLE5, MtCLE34 and GmCLE32 are not included.

429 **Supplementary Table 1.** Features of the *P. sativum* CLE genes.

430 **Supplementary Table 2.** Sequences with a CLE domain from PsCAmour Low Copy Number that are
431 not included in analyses reported here.

References

Alves-Carvalho S, Aubert G, Carrère S, Cruaud C, Brochot AL, Jacquin F, Klein A, Martin C, Boucherot K, Kreplak J. 2015. Full-length de novo assembly of RNA-seq data in pea (*Pisum sativum* L.) provides a gene expression atlas and gives insights into root nodulation in this species. *The Plant Journal* **84**(1): 1-19.

Broughton W, Dilworth M. 1971. Control of leghaemoglobin synthesis in snake beans. *Biochemical journal* **125**(4): 1075-1080.

Cock JM, McCormick S. 2001. A large family of genes that share homology with CLAVATA3. *Plant Physiology* **126**(3): 939-942.

Corcilius L, Hastwell AH, Zhang M, Williams J, Mackay JP, Gresshoff PM, Ferguson BJ, Payne RJ. 2017. Arabinosylation modulates the growth-regulating activity of the peptide hormone CLE40a from Soybean. *Cell Chemical Biology*. In Press

Delves AC, Mathews A, Day DA, Carter AS, Carroll BJ, Gresshoff PM. 1986. Regulation of the soybean-Rhizobium nodule symbiosis by shoot and root factors. *Plant Physiology* **82**(2): 588-590.

Duc G, Messager A. 1989. Mutagenesis of pea (*Pisum sativum* L.) and the isolation of mutants for nodulation and nitrogen fixation. *Plant Science* **60**: 207–213.

Elise S, Etienne-Pascal J, de Fernanda C-N, Gérard D, Julia F. 2005. The *Medicago truncatula* SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. *Plant Molecular Biology* **58**(6): 809-822.

Ferguson BJ, Indrasumunar A, Hayashi S, Lin MH, Lin YH, Reid DE, Gresshoff PM. 2010. Molecular analysis of legume nodule development and autoregulation. *Journal of Integrative Plant Biology* **52**(1): 61-76.

Ferguson BJ, Li D, Hastwell AH, Reid DE, Li Y, Jackson SA, Gresshoff PM. 2014. The soybean (*Glycine max*) nodulation-suppressive CLE peptide, GmRIC1, functions interspecifically in common white bean (*Phaseolus vulgaris*), but not in a supernodulating line mutated in the receptor PvNARK. *Plant Biotechnology Journal* **12**(8): 1085-1097.

Ferguson BJ, Mathesius U. 2014. Phytohormone regulation of legume-rhizobia interactions. *Journal of Chemical Ecology* **40**(7): 770-790.

Foyer CH, Lam H-M, Nguyen HT, Siddique KH, Varshney RK, Colmer TD, Cowling W, Bramley H, Mori TA, Hodgson JM. 2016. Neglecting legumes has compromised human health and sustainable food production. *Nature Plants* **2**: 16112.

Gresshoff PM, Hayashi S, Biswas B, Mirzaei S, Indrasumunar A, Reid D, Samuel S, Tollenaere A, van Hameren B, Hastwell A. 2015. The value of biodiversity in legume symbiotic nitrogen fixation and nodulation for biofuel and food production. *Journal of Plant Physiology* **172**: 128-136.

Hastwell AH, de Bang TC, Gresshoff PM, Ferguson BJ. 2017. CLE peptide-encoding gene families *in Medicago truncatula* and *Lotus japonicus*, compared with those of soybean, common bean and Arabidopsis. *Scientific Reports* **7**.

Hastwell AH, Gresshoff PM, Ferguson BJ. 2015b. Genome-wide annotation and characterization of CLAVATA/ESR (CLE) peptide hormones of soybean (*Glycine max*) and common bean (*Phaseolus vulgaris*), and their orthologues of *Arabidopsis thaliana*. *Journal of Experimental Botany* **66**(17): 5271-5287.

Hastwell AH, Gresshoff PM, Ferguson BJ. 2015a. The structure and activity of nodulation-suppressing CLE peptide hormones of legumes. *Functional Plant Biology* **42**(3): 229-238.

Jensen ES, Peoples MB, Boddey RM, Gresshoff PM, Hauggaard-Nielsen H, Alves BJ, Morrison MJ. 2012. *C*egumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review. *Agronomy for Sustainable Development* **32**(2): 329-364.

Kaeothip S, Boons G-J. 2013. Chemical synthesis of β-arabinofuranosyl containing oligosaccharides derived from plant cell wall extensins. *Organic & Biomolecular Chemistry* **11**(31): 5136-5146.

Kaeothip S, Ishiwata A, Ito Y. 2013. Stereoselective synthesis of Arabidopsis CLAVATA3 (CLV3) glycopeptide, unique protein post-translational modifications of secreted peptide hormone in plant. *Organic & Biomolecular Chemistry* **11**(35): 5892-5907.

Kassaw T, Nowak S, Schnabel E, Frugoli JA. 2017. ROOT DETERMINED NODULATION1 is required for M. truncatula CLE12, but not CLE13 peptide signaling through the SUNN receptor kinase. *Plant Physiology*: pp. 00278.02017.

Kosslak RM, Bohlool BB. 1984. Suppression of nodule development of one side of a split-root system of soybeans caused by prior inoculation of the other side. *Plant Physiology* **75**(1): 125-130.

Krusell L, Madsen LH, Sato S, Aubert G, Genua A, Szczyglowski K, Duc G, Kaneko T, Tabata S, de Bruijn F. 2002. Shoot control of root development and nodulation is mediated by a receptor-like kinase. *Nature* **420**(6914): 422-426.

Krusell L, Sato N, Fukuhara I, Koch BE, Grossmann C, Okamoto S, Oka-Kira E, Otsubo Y, Aubert G, Nakagawa T. 2011. The Clavata2 genes of pea and *Lotus japonicus* affect Autoregulation of Nodulation. *The Plant Journal* **65**(6): 861-871.

Li D, Kinkema M, Gresshoff PM. 2009. Autoregulation of nodulation (AON) in *Pisum sativum* (pea) involves signalling events associated with both nodule primordia development and nitrogen fixation. *Journal of Plant Physiology*. **166**: 955-67.

Lim CW, Lee YW, Lee SC, Hwang CH. 2014. Nitrate inhibits soybean nodulation by regulating expression of CLE genes. *Plant Science* **229**: 1-9.

Lin YH, Ferguson BJ, Kereszt A, Gresshoff PM. 2010. Suppression of hypernodulation in soybean by a leaf-extracted, NARK-and Nod factor-dependent, low molecular mass fraction. *New Phytologist* **185**(4): 1074-1086.

Lin Y-H, Lin M-H, Gresshoff PM, Ferguson BJ. 2011. An efficient petiole-feeding bioassay for introducing aqueous solutions into dicotyledonous plants. *Nature Protocols* **6**(1): 36-45.

Magori S, Oka-Kira E, Shibata S, Umehara Y, Kouchi H, Hase Y, Tanaka A, Sato S, Tabata S, Kawaguchi M. 2009. TOO MUCH LOVE, a root regulator associated with the long-distance control of nodulation in *Lotus japonicus*. *Molecular Plant-Microbe Interactions* **22**(3): 259-268.

Mirzaei S, Batley J, El-Mellouki T, Liu S, Meksem K, Ferguson BJ, Gresshoff PM. 2017. Neodiversification of homeologous CLAVATA1-like receptor kinase genes in soybean leads to distinct developmental outcomes. *Scientific Reports* **7**.

Miyazawa H, Oka-Kira E, Sato N, Takahashi H, Wu G-J, Sato S, Hayashi M, Betsuyaku S, Nakazono M, Tabata S. 2010. The receptor-like kinase KLAVIER mediates systemic regulation of nodulation and non-symbiotic shoot development in *Lotus japonicus*. *Development* **137**(24): 4317-4325.

Mortier V, Den Herder G, Whitford R, Van de Velde W, Rombauts S, D'haeseleer K, Holsters M, Goormachtig S. 2010. CLE peptides control *Medicago truncatula* nodulation locally and systemically. *Plant Physiology* **153**(1): 222-237.

Nishida H, Handa Y, Tanaka S, Suzaki T, Kawaguchi M. 2016. Expression of the *CLE-RS3* gene suppresses root nodulation in *Lotus japonicus*. *Journal of Plant Research* **129**(5): 909-919.

Nishimura R, Hayashi M, Wu G-J, Kouchi H, Imaizumi-Anraku H, Murakami Y, Kawasaki S, Akao S, Ohmori M, Nagasawa M. 2002. HAR1 mediates systemic regulation of symbiotic organ development. *Nature* **420**(6914): 426-429.

Ogawa-Ohnishi M, Matsushita W, Matsubayashi Y. 2013. Identification of three hydroxyproline O-arabinosyltransferases in *Arabidopsis thaliana*. *Nature Chemical Biology* **9**(11): 726-730.

Okamoto S, Ohnishi E, Sato S, Takahashi H, Nakazono M, Tabata S, Kawaguchi M. 2009. Nod factor/nitrate-induced CLE genes that drive HAR1-mediated systemic regulation of nodulation. *Plant and Cell Physiology* **50**(1): 67-77.

Okamoto S, Shinohara H, Mori T, Matsubayashi Y, Kawaguchi M. 2013. Root-derived CLE glycopeptides control nodulation by direct binding to HAR1 receptor kinase. *Nature Communications* **4**.

Okamoto S, Suzuki T, Kawaguchi M, Higashiyama T, Matsubayashi Y. 2015. A comprehensive strategy for identifying long-distance mobile peptides in xylem sap. *The Plant Journal* **84**(3): 611-620.

Osipova MA, Mortier V, Demchenko KN, Tsyganov VE, Tikhonovich IA, Lutova LA, Dolgikh EA, Goormachtig S. 2012. *WUSCHEL-RELATED HOMEOBOX5* gene expression and interaction of CLE peptides with components of the systemic control add two pieces to the puzzle of Autoregulation of Nodulation. *Plant Physiology* **158**(3): 1329-1341.

Postma JG, Jacobsen E, Feenstra WJ. 1988. Three pea mutants with an altered nodulation studied by genetic analysis and grafting. *Journal of Plant Physiology* **132**: 424–430.

Reid DE, Ferguson BJ, Gresshoff PM. 2011a. Inoculation-and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation. *Molecular Plant-Microbe Interactions* **24**(5): 606-618.

Reid DE, Ferguson BJ, Hayashi S, Lin Y-H, Gresshoff PM. 2011b. Molecular mechanisms controlling legume Autoregulation of Nodulation. *Annals of Botany* **108**(5): 789-795.

Reid DE, Li D, Ferguson BJ, Gresshoff PM. 2013. Structure–function analysis of the GmRIC1 signal peptide and CLE domain required for nodulation control in soybean. *Journal of Experimental Botany*: ert008.

Sagan M, Duc G. 1996. *Sym28* and *Sym29*, two new genes involved in regulation of nodulation in pea (*Pisum sativum* L.). *Symbiosis* **20**: 229–245.

Sasaki T, Suzaki T, Soyano T, Kojima M, Sakakibara H, Kawaguchi M. 2014. Shoot-derived cytokinins systemically regulate root nodulation. *Nature Communications* **5**.

Schnabel E, Kassaw T, Smith L, Marsh J, Oldroyd GE, Long SR, Frugoli J. 2011. ROOT DETERMINED NODULATION 1 regulates nodule number in *M. truncatula* and defines a highly conserved, uncharacterized plant gene family. *Plant Physiology*: pp. 111.178756.

Searle IR, Men AE, Laniya TS, Buzas DM, Iturbe-Ormaetxe I, Carroll BJ, Gresshoff PM. 2003. Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. *Science* **299**(5603): 109-112.

Shinohara H, Matsubayashi Y. 2013. Chemical synthesis of Arabidopsis CLV3 glycopeptide reveals the impact of hydroxyproline arabinosylation on peptide conformation and activity. *Plant and Cell Physiology* **54**(3): 369-374.

Shinohara H, Matsubayashi Y. 2015. Reevaluation of the CLV3-receptor interaction in the shoot apical meristem: dissection of the CLV3 signaling pathway from a direct ligand-binding point of view. *The Plant Journal* **82**(2): 328-336.

Somerville JE, Kahn ML. 1983. Cloning of the glutamine synthetase I gene from *Rhizobium meliloti*. *Journal of Bacteriology* **156**(1): 168-176.

Tayeh N, Aluome C, Falque M, Jacquin F, Klein A, Chauveau A, Bérard A, Houtin H, Rond C, Kreplak J. 2015. Development of two major resources for pea genomics: the GenoPea 13.2 K SNP Array and a high-density, high-resolution consensus genetic map. *The Plant Journal* **84**(6): 1257-1273.

Whitford R, Fernandez A, De Groodt R, Ortega E, Hilson P. 2008. Plant CLE peptides from two distinct functional classes synergistically induce division of vascular cells. *Proceedings of the National Academy of Sciences* **105**(47): 18625-18630.

Xu C, Liberatore KL, MacAlister CA, Huang Z, Chu Y-H, Jiang K, Brooks C, Ogawa-Ohnishi M, Xiong G, Pauly M. 2015. A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nature Genetics* **47**(7): 784-792.

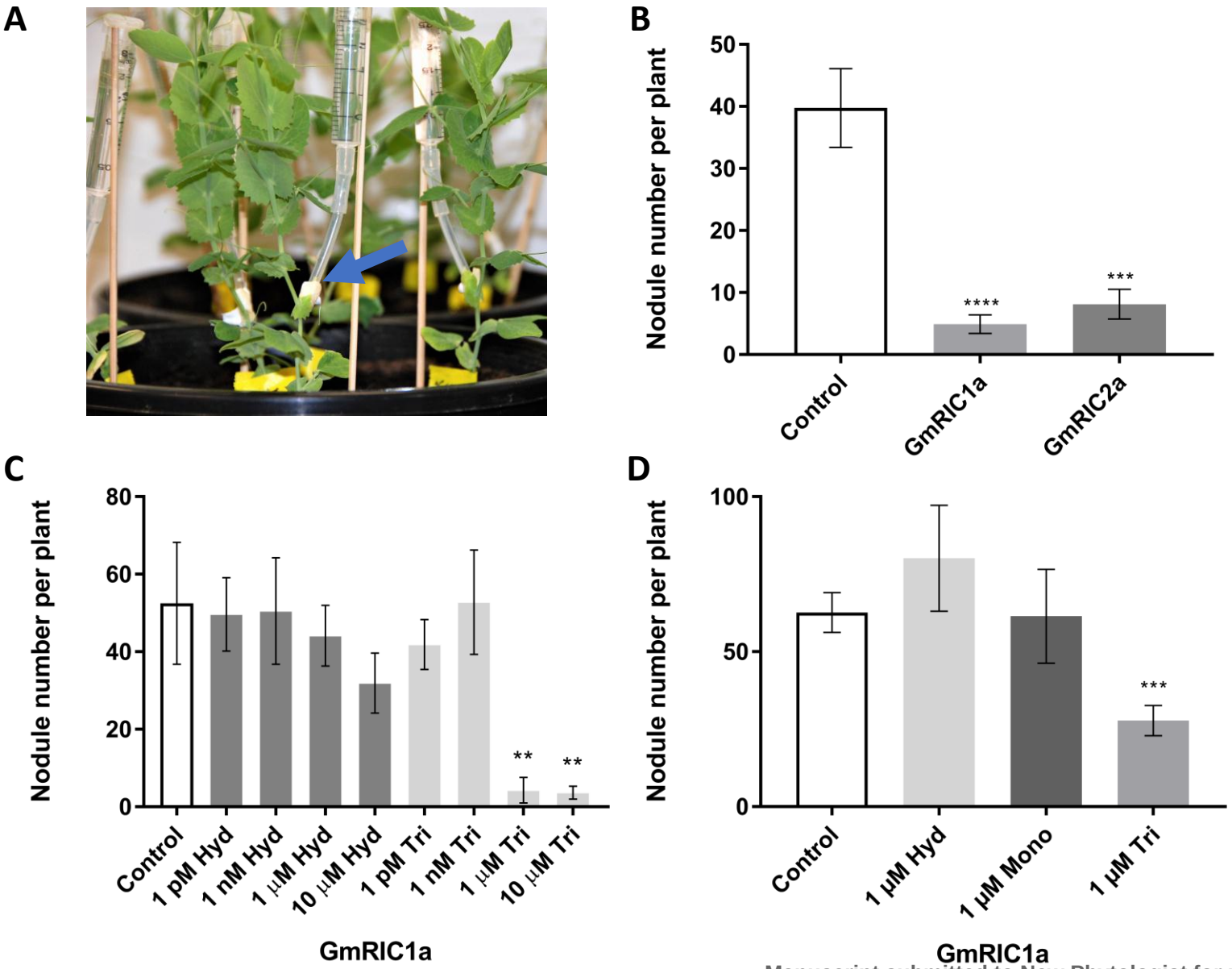


Figure 1. Nodule number 14 days after inoculation of wild-type pea plants fed via petiole feeding. **A** Image of pea plants with petiole feeding apparatus attached (blue arrow). **B** 1 µM triarabinosylated (Tri) GmRIC1a, GmRIC2a and water control. **C** 1 pM to 10 µM of hydroxylated (Hyd) or triarabinosylated (Tri) GmRIC1a. **D** 1 µM triarabinosylated (Tri), monoarabinosylated (Mono) or hydroxylated (Hyd) GmRIC1a, and water control. Statistical differences determined using Student's *t*-test. n = 7 to 10 plants per treatment.

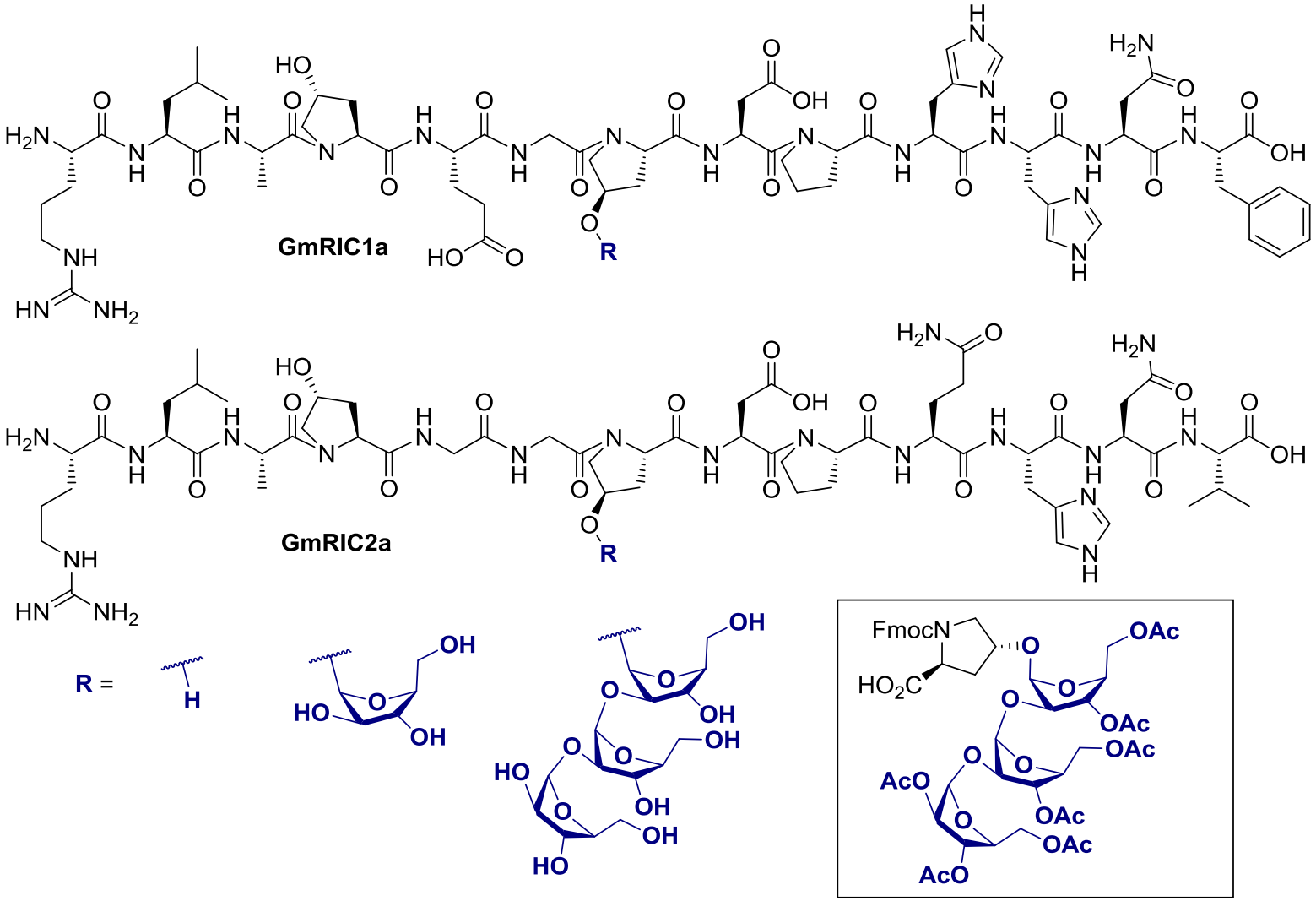


Figure 2. Structures of synthetic GmRIC1a and GmRIC2a peptides, and triarabinosylated hydroxyproline building block (in box). Proline 4 is hydroxylated in all variants. Proline 7 is either hydroxylated only (R = H), or further modified by arabinosylation (R = monoarabinose or triarabinose).

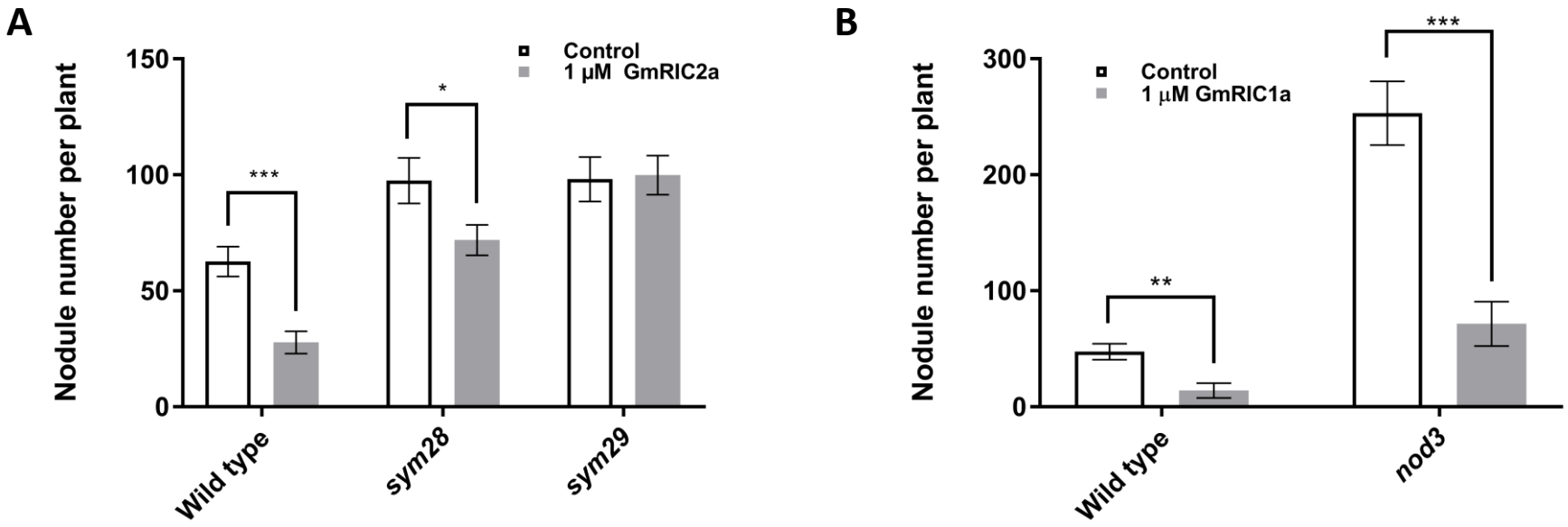


Figure 3. Nodule number 14 days after inoculation of wild-type and nodulation-mutant pea plants fed via petiole feeding with either 1 µM GmRIC1a, GmRIC2a, or water control. **A** Wild type, *sym28* and *sym29* plants fed with GmRIC2a. **B** Wild type and *nod3* plants fed with GmRIC1a. Statistical differences determined using students t-test. n = 5-8 plants per treatment.

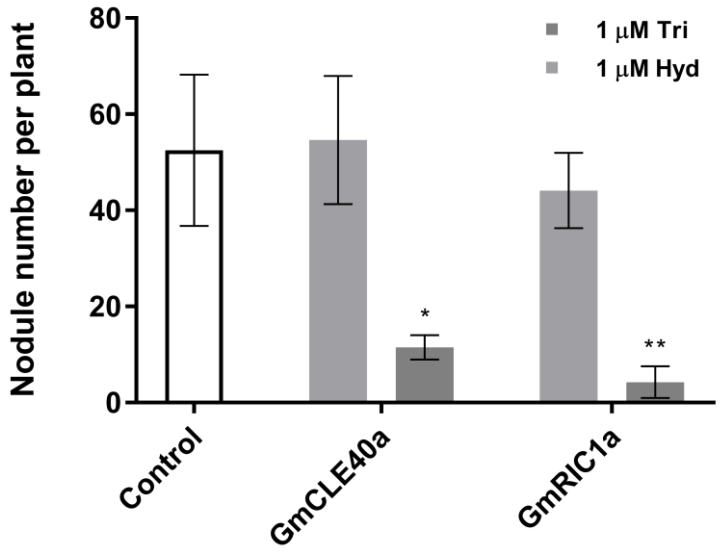


Figure 4. Nodule number 14 days after inoculation of wild-type pea plants fed via petiole feeding with 1 μM hydroxylated (Hyd) or triarabinosylated (Tri) GmRIC1a, GmCLE40a, or water control. Statistical differences determined using students t-test. n = 7 to 8 plants per treatment.

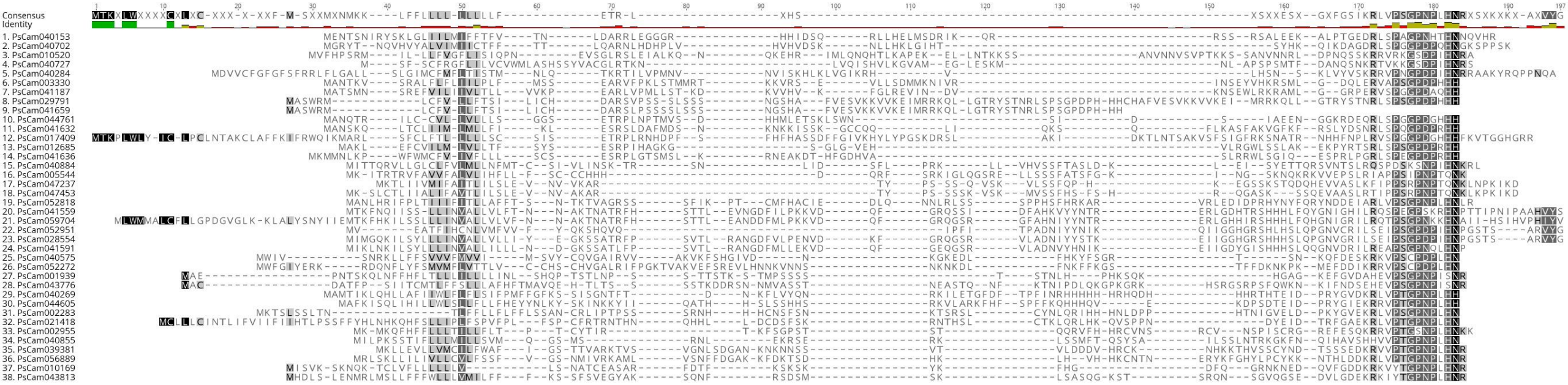
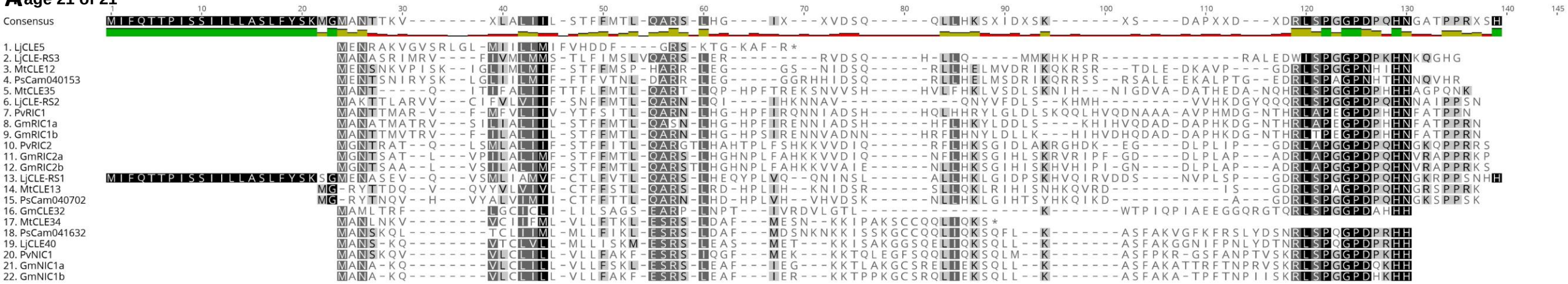


Figure 5. Multiple Sequence Alignment of the CLE prepropeptides of *P. sativum*. Shaded nucleotides indicate conservation. Not shown are the multi CLE domain containing prepropeptides.

A



B

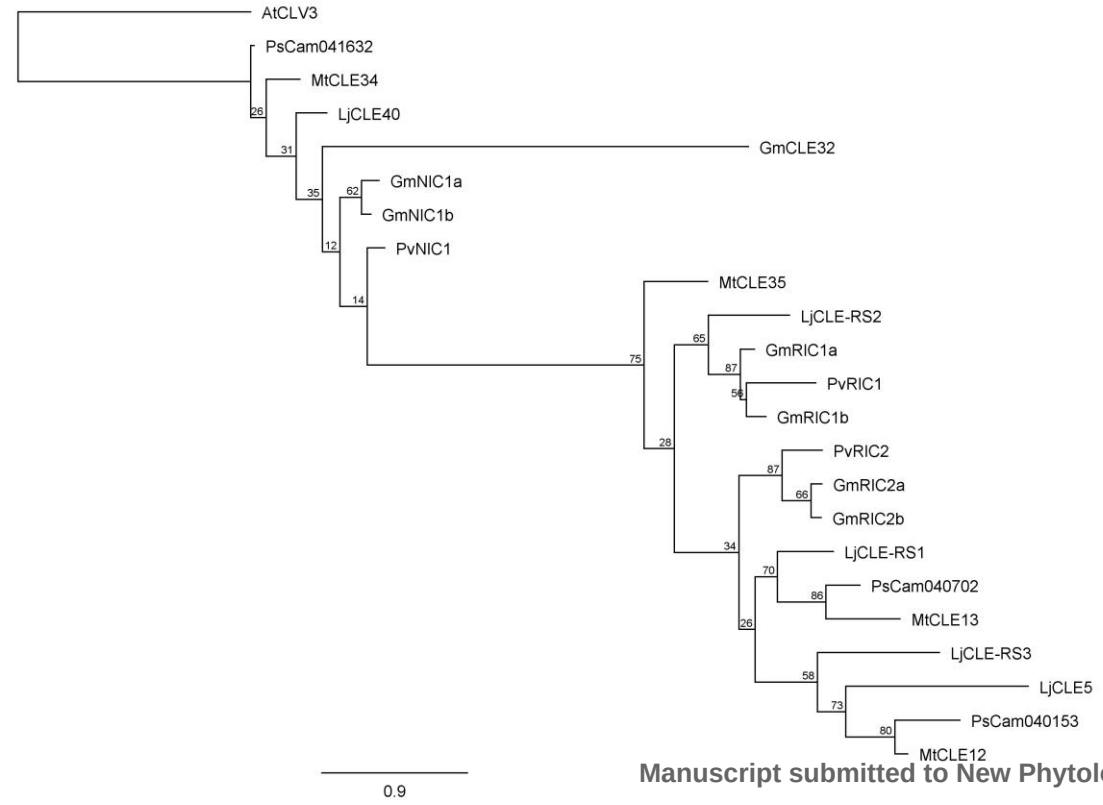


Figure 6. Nodulation-suppressive CLE prepropeptides of *P. sativum* and their orthologues in *G. max*, *P. vulgaris*, *L. japonicus*, and *M. truncatula*. **A** Multiple Sequence Alignment where grey residues indicate conserved residues. **B** Phylogenetic tree, with Bootstrap confidence values expressed as a percentage of 1,000 bootstrap replications, using AtCLV3 as an outgroup.