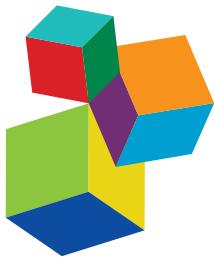


COOPERATIVE ADAPTATIONS AND EVOLUTION IN PLANT-MICROBE SYSTEMS

EDITED BY: Tatiana Matveeva, Nikolai Provorov and Jari P. T. Valkonen

PUBLISHED IN: *Frontiers in Plant Science*





Frontiers Copyright Statement

© Copyright 2007-2018 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714
ISBN 978-2-88945-599-7
DOI 10.3389/978-2-88945-599-7

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

COOPERATIVE ADAPTATIONS AND EVOLUTION IN PLANT-MICROBE SYSTEMS

Topic Editors:

Tatiana Matveeva, St. Petersburg State University, Russia

Nikolai Provorov, All-Russia Research Institute for Agricultural Microbiology, Russia

Jari P. T. Valkonen, University of Helsinki, Finland



Morphological diversity of bacterial colonies associated with plants.

Image: Nikolai Provorov.

Ecological and evolutionary genetics of plant-microbe interactions is of high importance for developing the plant science since the plants originated symbiotically (via incorporation of a phototrophic cyanobacterium into a heterotrophic eukaryon) and further evolve as the multipartite symbiotic systems, harboring the enormously diverse microbial communities. The Research Topic has integrated the top-level research on the genetic interactions in the plant-microbial associations required to develop the novel evolutionary approaches in the molecular and ecological genetics of different kinds of symbioses.

Citation: Matveeva, T., Provorov, N., Valkonen, J. P. T., eds. (2018). Cooperative Adaptations and Evolution in Plant-Microbe Systems. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-599-7

Table of Contents

- 05 Editorial: Cooperative Adaptation and Evolution in Plant-Microbe Systems**
Tatiana Matveeva, Nikolai Provorov and Jari P. T. Valkonen

SECTION I

NEXT GENERATION SEQUENCING FOR PLANT-MICROBE INTERACTION STUDY

- 07 Re-analyses of "Algal" Genes Suggest a Complex Evolutionary History of Oomycetes**
Qia Wang, Hang Sun and Jinling Huang
- 21 De novo Transcriptome Assembly of Phomopsis liquidambari Provides Insights into Genes Associated With Different Lifestyles in Rice (*Oryza sativa L.*)**
Jun Zhou, Xin Li, Yan Chen and Chuan-Chao Dai
- 40 Natural Agrobacterium Transformants: Recent Results and Some Theoretical Considerations**
Ke Chen and Léon Otten
- 56 Horizontal Gene Transfer Contributes to Plant Evolution: The Case of Agrobacterium T-DNAs**
Dora G. Quispe-Huamanquispe, Godelieve Gheysen and Jan F. Kreuze

SECTION II

COOPERATIVE ADAPTATIONS AND EVOLUTION IN MUTUALISTIC PLANT-MICROBE SYSTEMS

- 62 The Symbiosome: Legume and Rhizobia Co-evolution Toward a Nitrogen-Fixing Organelle?**
Teodoro Coba de la Peña, Elena Fedorova, José J. Pueyo and M. Mercedes Lucas
- 88 Purification and In Vitro Activity of Mitochondria Targeted Nitrogenase Cofactor Maturase NifB**
Stefan Burén, Xi Jiang, Gema López-Torrejón, Carlos Echavarri-Erasun and Luis M. Rubio
- 104 Diversity and Functional Analysis of the FeMo-Cofactor Maturase NifB**
Simon Arragain, Emilio Jiménez-Vicente, Alessandro A. Scandurra, Stefan Burén, Luis M. Rubio and Carlos Echavarri-Erasun
- 116 Selection Signatures in the First Exon of Paralogous Receptor Kinase Genes From the Sym2 Region of the *Pisum sativum L.* Genome**
Anton S. Sulima, Vladimir A. Zhukov, Alexey A. Afonin, Aleksandr I. Zhernakov, Igor A. Tikhonovich and Ludmila A. Lutova
- 136 Identification and Expression Analysis of *Medicago truncatula* Isopentenyl Transferase Genes (IPTs) Involved in Local and Systemic Control of Nodulation**
Mahboobeh Azarakhsh, Maria A. Lebedeva and Lyudmila A. Lutova

SECTION III

COOPERATIVE ADAPTATIONS AND EVOLUTION IN PLANT-PARTHOGEN SYSTEMS

147 *Arabidopsis ALA1 and ALA2 Mediate RNAi-Based Antiviral Immunity*

Biyun Zhu, Hua Gao, Gang Xu, Dewei Wu, Susheng Song, Hongshan Jiang, Shuifang Zhu, Tiancong Qi and Daoxin Xie

156 *Convergent Evolution of Pathogen Effectors Toward Reactive Oxygen Species Signaling Networks in Plants*

Nam-Soo Jwa and Byung Kook Hwang

168 *Interplant Aboveground Signaling Prompts Upregulation of Auxin Promoter and Malate Transporter as Part of Defensive Response in the Neighboring Plants*

Connor Sweeney, Venkatachalam Lakshmanan and Harsh P. Bais

SECTION IV

MICROBIAL IMPACT ON HORMONAL REGULATION OF STRESS TOLERANCE

178 *Bacillus Amyloliquefaciens Confers Tolerance to Various Abiotic Stresses and Modulates Plant Response to Phytohormones Through Osmoprotection and Gene Expression Regulation in Rice*

Shalini Tiwari, Vivek Prasad, Puneet S. Chauhan and Charu Lata

191 *Endophytic Paecilomyces Formosus LHL10 Augments Glycine Max L. Adaptation to Ni-Contamination Through Affecting Endogenous Phytohormones and Oxidative Stress*

Saqib Bilal, Abdul L. Khan, Raheem Shahzad, Sajjad Asaf, Sang-Mo Kang and In-Jung Lee



Editorial: Cooperative Adaptation and Evolution in Plant-Microbe Systems

Tatiana Matveeva^{1*}, Nikolai Provorov² and Jari P. T. Valkonen³

¹ Department of Genetics and Biotechnology, St. Petersburg State University, St. Petersburg, Russia, ² All-Russia Research Institute for Agricultural Microbiology, St. Petersburg, Russia, ³ Department of Agricultural Sciences, University of Helsinki, Helsinki, Finland

Keywords: plant-microbe interaction, coadaptation, coevolution, horizontal gene transfer, ecological genetics, beneficial and antagonistic symbioses, signal-receptor interactions, symbiotic N₂ fixation

Editorial on the Research Topic

Cooperative Adaptation and Evolution in Plant-Microbe Systems

Evolutionarily, plant-microbe interactions range from beneficial symbioses to the molecular arms race between pathogens and the immune systems of plants. Expanding our knowledge on ecological and evolutionary genetics of plant-microbe interactions is of high importance. Plants coevolve symbiotically with enormously diverse microbial communities, which has been pivotal since colonization of land by plants. The fungal and bacterial associates provide plants with important nutritional, protective and growth regulatory functions.

The well-studied mutualists (legume N₂-fixing nodules, arbuscular mycorrhizae) and antagonists (biotrophic and necrotrophic) represent only a minority of symbioses between plants and associated microbial communities. The endophytic and epiphytic microbiomes exceed their hosts greatly in terms of genetic information potentially useful for extending the plant ecological amplitude and improving crop production.

A major breakthrough in conceptualizing the role of plant-microbe interactions in evolution has become possible largely owing to the new research methods. Next generation sequencing (NGS) opens up new prospects for studies in inter-species interactions. On the one hand, analysis of the accumulating data makes it possible to approach macroevolution from a new angle. The study presented by Wang et al. in this issue of *Frontiers in Plant Science* lays ground for discussions about evolution of stramenopiles and more complex scenarios for the evolution of oomycetes, including the supposed ancestral endosymbioses or independent horizontal gene transfer events involving red and green algae, oomycetes and other stramenopiles.

NGS provides opportunities for deeper study of genomes and transcriptomes of species, new gene combinations and differentially expressed genes during the symbiotic interactions. For instance, *Phomopsis liquidambari* studied by Zhou et al. is established in endophytic and saprophytic systems with rice (*Oryza sativa* L.). Most genes for amino acids and carbohydrate metabolism, fatty acid biosynthesis, and secondary metabolism are up-regulated in endophytic fungi. Most pathways of xenobiotic biodegradation and metabolism are upregulated in saprophytic systems, demonstrating the genetic regulation of adaptation to various ecological niches.

Symbiotic relationships contribute not only to changes in the pattern of gene expression, but also to the exchange of genes between symbionts. NGS helps to find footprints of such exchanges. Five different types of T-DNA of *Rhizobium rhizogenes* (formerly *Agrobacterium rhizogenes*) were identified in *Nicotiana* during the analysis of genome sequence data, supplementing former information about T-DNA in *Nicotiana* and *Linaria* species with new types of T-DNA. Homologues

OPEN ACCESS

Edited and reviewed by:

Mari-Anne Newman,
University of Copenhagen, Denmark

*Correspondence:

Tatiana Matveeva
radishlet@gmail.com

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 13 June 2018

Accepted: 05 July 2018

Published: 14 August 2018

Citation:

Matveeva T, Provorov N and
Valkonen JPT (2018) Editorial:
Cooperative Adaptation and Evolution
in Plant-Microbe Systems.
Front. Plant Sci. 9:1090.
doi: 10.3389/fpls.2018.01090

of T-DNAs of agrobacteria were fortuitously found in the genome of sweetpotato while assembling small interfering RNAs (siRNAs) for metagenomic analysis. All this suggests that horizontal gene transfer between bacteria and plants is more widespread in evolution than previously thought (Chen and Otten; Quispe-Huamanquispe et al.). However, to understand the function of the transferred genes, additional studies are required.

Another interesting and actively studied aspect of the interaction of plants and rhizobia is nitrogen-fixing symbiosis. Evolution of N₂-fixing symbioses may be addressed as a conduit of constructing N₂-fixing plants. Construction of “ammonioplasts” by N₂-fixing intracellular symbionts was reviewed by de la Peña et al. who outlined that symbiosomes of galegoid legumes represent functional and structural analogs of regular cellular organelles. Moreover, in some legumes, the symbosome formation is compatible with the host cell divisions suggesting a possibility for regeneration of plants stably maintaining the ammonioplasts.

N₂-fixing organelles can be constructed by introducing *nif* genes into regular organelles since their free-living ancestors probably possessed those genes. Burén et al. have demonstrated that the functional *nifB* may be expressed in mitochondria of yeast and tobacco, whereas Arragain et al. found broad phylogenetic diversity of *nifB* in thermo- and mesophytic microbes. Hence, it seems possible to select optimal *nifB* alleles for expression in eukaryotic cells.

For developing the novel N₂-fixing systems, it is important to dissect the evolutionary mechanisms operating in the extant symbioses, wherein the evolution of N₂-fixing bacteria is directed by plant hosts. An important approach is to analyze the evolution of plant receptor genes and signatures of different types of natural selection, characterized by nonsynonymous/synonymous nucleotide substitutions. Their functionality was demonstrated for the family of LysM-RLK receptors in *Pisum sativum* recognizing the *Rhizobium leguminosarum* Nod factors, since different selection types (purifying, positive, balancing) are attributed to the functionally diverse plant gene domains (Sulima et al.). Evolution of symbiotically specialized plant genes may be sufficiently intensified due to their cytokinin-dependent local and systemic regulation which was for the first time demonstrated in *Medicago truncatula* (Azarakhsh et al.).

An alternative, antagonistic strategy of plant-microbe interactions may be represented by defense against pathogens. Co-evolution of plants and pathogens has affected the various mechanisms of their interaction. Defense against viruses in plants is largely based on recognition and degradation of double-stranded viral RNA by a mechanism called RNA interference (RNAi). New plant genes that participate in the

RNAi pathway and enhance its function are reported by Zhu et al. Plants undergo molecular arms race also with pathogens that interfere with other immune systems, e.g., by inhibiting signaling molecules such as reactive oxygen species (ROS) triggering defense responses upon infection. Jwa and Hwang provide updates of the mechanisms by which pathogen effectors interfere with ROS signaling in plants. Plants can also warn neighboring plants by signaling via volatile organic compounds (VOCs). The plants wounded by insects or pathogens produce VOCs triggering defense mechanisms in undamaged plants, induce root growth and enhance beneficial root microbes, and hence improve overall resistance of the plant to pathogens and pests (Sweeney et al.).

Besides the biotic stresses caused by pathogens and pests, plants face abiotic stresses such as drought and high or low temperatures that may be most limiting for growth. Non-pathogenic microbes may alleviate such stresses. Tiwari et al. report that a strain of *Bacillus amyloliquefaciens* living in the plant root system is able to reduce abiotic stresses in rice via cross-talk with pathways regulating stresses and phytohormones. Similarly, studies of Bilal et al. show that soil microbes such as *Paecilomyces formosus* can help alleviating the stress of plants caused by heavy metals such as nickel in plants.

The studies published in this thematic issue contribute to a broad range of fundamental and applied themes and integrate top level research on genetic interactions in plant-microbe associations. The novel information reported is required to develop the prospective evolutionary approaches for the study of molecular and ecological genetics of symbioses.

AUTHOR CONTRIBUTIONS

TM, NP, and JV edited the topic and wrote an editorial.

ACKNOWLEDGMENTS

Topic editors acknowledge support of RSF grants 14-26-00094 to NP and 16-16-10010 to TM and grant 1276136 of Academy of Finland to JV.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Matveeva, Provorov and Valkonen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Re-analyses of “Algal” Genes Suggest a Complex Evolutionary History of Oomycetes

Qia Wang^{1,2}, Hang Sun^{1*} and Jinling Huang^{1,3,4*}

¹ Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China, ² University of Chinese Academy of Sciences, Beijing, China, ³ State Key Laboratory of Cotton Biology, Institute of Plant Stress Biology, Henan University, Kaifeng, China, ⁴ Department of Biology, East Carolina University, Greenville, NC, United States

OPEN ACCESS

Edited by:

Tatiana Matveeva,
Saint Petersburg State University,
Russia

Reviewed by:

David John Studholme,
University of Exeter, United Kingdom
Huan Qiu,
Rutgers University, United States

*Correspondence:

Hang Sun
sunhang@mail.kib.ac.cn
Jinling Huang
huangj@ecu.edu

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 25 May 2017

Accepted: 22 August 2017

Published: 06 September 2017

Citation:

Wang Q, Sun H and Huang J (2017) Re-analyses of “Algal” Genes Suggest a Complex Evolutionary History of Oomycetes. *Front. Plant Sci.* 8:1540.
doi: 10.3389/fpls.2017.01540

The spread of photosynthesis is one of the most important but constantly debated topics in eukaryotic evolution. Various hypotheses have been proposed to explain the plastid distribution in extant eukaryotes. Notably, the chromalveolate hypothesis suggested that multiple eukaryotic lineages were derived from a photosynthetic ancestor that had a red algal endosymbiont. As such, genes of plastid/algal origin in aplastidic chromalveolates, such as oomycetes, were considered to be important supporting evidence. Although the chromalveolate hypothesis has been seriously challenged, some of its supporting evidence has not been carefully investigated. In this study, we re-evaluate the “algal” genes from oomycetes with a larger sampling and careful phylogenetic analyses. Our data provide no conclusive support for a common photosynthetic ancestry of stramenopiles, but show that the initial estimate of “algal” genes in oomycetes was drastically inflated due to limited genome data available then for certain eukaryotic lineages. These findings also suggest that the evolutionary histories of these “algal” genes might be attributed to complex scenarios such as differential gene loss, serial endosymbioses, or horizontal gene transfer.

Keywords: plastid evolution, stramenopiles, endosymbiosis, horizontal gene transfer, eukaryotic evolution

INTRODUCTION

How photosynthesis evolved in eukaryotes has been a subject of tremendous scientific interest. Oxygenic photosynthesis was first invented by cyanobacteria (Gould et al., 2008). During the early evolution of eukaryotes, a cyanobacterial cell was engulfed by a heterotrophic eukaryote (Margulis, 1970; Martin and Kowallik, 1999; McFadden, 2001; Palmer, 2003), spawning the origin of primary plastids and Plantae (also called Archaeplastida, including green plants, red algae, and glaucophytes) (Cavalier-Smith, 1981; Delwiche and Palmer, 1997; Gould et al., 2008). This process was accompanied by massive cyanobacterial gene loss and transfer to the host nucleus. Subsequently, the photosynthetic capacity was spread to multiple other eukaryotic lineages through higher-order endosymbioses (secondary, tertiary, or quaternary) (Delwiche, 1999), that is, these eukaryotes acquired plastids by engulfing another photosynthetic eukaryote instead of a cyanobacterial cell. Although it is clear that the spread of photosynthetic capacity in eukaryotic lineages represents a history of reticulate evolution involving multiple endosymbiotic events, the exact number and the nature of historical endosymbioses remain controversial.

Among eukaryotic lineages involved in higher-order endosymbioses, it was generally accepted that plastids of euglenids and chlorarachniophytes are derived from green algal endosymbionts (Gibbs, 1978; Ludwig and Gibbs, 1989; Van de Peer et al., 1996), whereas plastids of cryptophytes, alveolates, stramenopiles and haptophytes (CASH lineages) are from red algal endosymbionts (Cavalier-Smith, 1995; Cavalier-Smith et al., 1996; Palmer and Delwiche, 1998; Delwiche, 1999). For a long period of time, plastid gains through endosymbiotic events were considered to be extremely rare and plastid losses, on the other hand, were thought to be relatively common. Such a belief also formed the foundation of the Cabozoa hypothesis and the chromalveolate hypothesis (Cavalier-Smith, 1999; Cavalier-Smith and Chao, 2003). The Cabozoa hypothesis argued that plastids of euglenids and chlorarachniophytes could be traced back to a common secondary endosymbiotic event involving a green alga (Cavalier-Smith, 1999; Cavalier-Smith and Chao, 2003). Similarly, the chromalveolate hypothesis proposed that plastids in CASH lineages were vertically derived from a common ancestor that engulfed a red algal endosymbiont and, as such, aplastidic organisms in these lineages were interpreted as resulting from secondary plastid losses (Cavalier-Smith, 1999). However, multiple lines of evidence (Baldauf et al., 2000; Archibald et al., 2003; Leander, 2004; Gilson et al., 2006), including the complete chloroplast genome of chlorarachniophyte *Bigelowiella natans* (Rogers et al., 2007), rejected the Cabozoa hypothesis. Thus far, it is commonly believed that the plastids of euglenids and chlorarachniophytes were acquired from two independent green algal endosymbiotic events (Gould et al., 2008; Keeling, 2013). The chromalveolate hypothesis has long been under debate and now in jeopardy in face of recent data (Bodył, 2005; Burki et al., 2008, 2012b, 2016; Kim and Graham, 2008; Yoon et al., 2008; Bodył et al., 2009). Several other hypotheses have been proposed, each of which is supported by different lines of evidence (Bodył and Moszczyński, 2006; Sanchez-Puerta and Delwiche, 2008; Okamoto et al., 2009; Stiller et al., 2014). Therefore, how plastids evolved in red plastid lineages remains unsettled.

Stramenopiles (also known as heterokonts), as a major eukaryotic clade, include a wide variety of organisms (Cavalier-Smith, 1986; Patterson, 1989). This lineage contains not only many important algae, such as diatoms that are a major producer of oxygen and consumer of carbon dioxide in marine ecosystems, but also a significant fraction of aplastidic or heterotrophic organisms, including pathogens like *Phytophthora infestans*, the causative agent of potato late blight that triggered the Great Irish Famine in the 1840s. Whether these diverse organisms originated from a common photosynthetic ancestor is crucial for understanding the evolution of stramenopiles as well as eukaryotes in general. This in turn led to many studies on the existence of potential historical plastids in heterotrophic stramenopiles (Keeling, 2013).

Oomycetes are fungus-like eukaryotic microorganisms that often have a saprophytic or pathogenic lifestyle. Oomycetes were once placed within fungi in earlier classification systems, but are now widely considered as part of stramenopiles (Baldauf et al., 2000; Yoon et al., 2002). Although there are different

views about the phylogenetic relationships within stramenopiles (Brown and Sorhannus, 2009; Riisberg et al., 2009; Yang et al., 2012; Cavalier-Smith and Scoble, 2013; Ševčíková et al., 2015), the most recent phylogenomic analyses suggest that oomycetes form a clade closely related to ochrophytes, a monophyletic group of photosynthetic stramenopiles (Derelle et al., 2016). Unlike ochrophytes, oomycetes do not contain plastids (Tyler et al., 2006; Derelle et al., 2016), not even vestigial ones like those in apicomplexan parasites (called apicoplast) (Maréchal and Cesbron-Delauw, 2001). If all stramenopiles are derived from a single photosynthetic ancestor, plastids would have been lost in oomycetes.

In 2006, draft genome sequences of two oomycete species, *Phytophthora sojae* and *P. ramorum*, were published (Tyler et al., 2006). In this study, 855 genes of putatively algal origin ("algal" genes hereafter) were identified based on their unusually high similarities to sequences from algae and/or cyanobacteria, 30 of which were considered the most likely cases after detailed phylogenetic analyses. These "algal" genes were interpreted as the relic from a red algal endosymbiont (plastid) and subsequent endosymbiotic gene transfer (EGT) or endosymbiotic gene replacement (EGR). As key evidence for historical plastids in oomycetes, these "algal" genes were further used to support the hypothesis that all stramenopiles were derived from a photosynthetic ancestor. Such evidence, however, has been called into question by a more recent statistical genomic analysis that found no unusual contribution from a red algal endosymbiont to *Phytophthora* genomes (Stiller et al., 2009).

Like in many earlier studies on gene transfer, insufficient taxonomic sampling was a potential caveat for the identification of "algal" genes in oomycetes. This is evidenced by the fact that, although the identified "algal" genes were interpreted as derived from a red algal endosymbiont, their sequences, on the other hand, were often found to be closely related to green algal homologs, presumably due to the lack of sufficient sequence data from red algae. As more genome sequence data from various major eukaryotic lineages become available in recent years, we now revisit the "algal" genes identified in oomycete genomes. Our goal is to provide a better understanding of the nature of these genes and the potential interactions of oomycetes/stramenopiles with other organisms, particularly primary photosynthetic eukaryotes.

MATERIALS AND METHODS

Data Sources

In the original *Phytophthora* genome analyses, 855 genes were considered to be of algal or cyanobacterial origin, and 30 most likely candidates were subject to further detailed analyses (Tyler et al., 2006). We downloaded the protein sequences of these 30 genes of *Phytophthora ramorum* from <http://www.jgi.doe.gov/Pramorum>, and used them as queries to search the National Center for Biotechnology Information (NCBI) non-redundant (nr) protein sequences database (*E*-value cutoff $1e-7$). Additional searches were also performed against over 650 transcriptomes in the Marine Microbial Eukaryote

Transcriptome Sequencing Project (MMETSP) (Keeling et al., 2014), the fungal genome database at the Joint Genome Institute¹, and our internal customized database (Supplementary Table 1). Particularly, a total of six red algal genomes of five genera were used to search for *P. ramorum* gene homologs, including *Cyanidioschyzon merolae*, *Porphyridium purpureum*, *Chondrus crispus*, *Galdieria sulphuraria*, *G. phlegrea*, and *Pyropia yezoensis*. Complete genome sequence data of multiple photosynthetic stramenopiles (including *Aureococcus anophagefferens*, *Ectocarpus siliculosus*, *Fragilariaopsis cylindrus*, *Nannochloropsis gaditana*, *Phaeodactylum tricornutum*, *Saccharina japonica*, and *Thalassiosira pseudonana*) were also searched in our analyses.

Re-analyses of BLAST Results

In the original *Phytophthora* genome paper (Tyler et al., 2006), the search of “algal” genes was based on significant matches to sequences from Plantae and cyanobacteria (that is, these sequence matches had the highest bit scores and the lowest *E*-values outside the stramenopiles). The *Phytophthora* “algal” genes identified from the BLAST search were shared by other stramenopiles (or chromalveolates), and they had stronger BLAST matches to homologous genes of red algae and/or cyanobacteria than to sequences from archaea, opisthokonts or non-cyanobacterial bacteria. Additionally, a complementary approach based on Smith–Waterman alignment was also used to identify candidates with significantly higher similarities to red algal or green plant homologs than to those from opisthokonts or amoebozoans. Because *Cyanidioschyzon merolae*, which also happens to have a streamlined genome, was the only red alga whose complete nuclear genome sequence was then available, the matches to green plant homologs were included and interpreted as resulting from the lack of sufficient red algal genome data (Tyler et al., 2006).

In the current study, BLAST search was performed against NCBI nr, MMETSP and our internal customized databases for each of the 30 most likely “algal” genes, followed by re-analyses of its phylogenetic distribution and gene structure. Following the criteria used by the *Phytophthora* genome paper (Tyler et al., 2006), we compared the best BLAST matches (represented by the highest bit scores) between homologs from red algae, cyanobacteria, photosynthetic stramenopiles and those from archaea, opisthokonts, amoebozoans, and non-cyanobacterial bacteria. Because more red algal genomes and transcriptomic data were included in our analyses, the matches to green plant homologs were no longer included and used as proxy for red algal homologs.

Phylogenetic Analyses

For each of the 30 “algal” genes identified in *Phytophthora* genome sequencing project, we performed further phylogenetic analyses. In order to attain a broad and balanced sampling, we selected protein sequences from representative groups of three domains of life (eukaryotes, bacteria, and archaea). The same sampling strategy was also used to ensure sufficient coverage of representative taxa within each major eukaryotic group. This

was done using a Perl script followed by manual inspection and additional sequence sampling if needed. Particular attention was paid to groups under-sampled in the previous analyses, such as chromalveolates and other protists. Multiple alignments of sampled sequences were performed using MUSCLE (Edgar, 2004), followed by careful manual inspection of alignment quality, gene structure, shared insertions/deletions (indels), and conserved amino acid residues. Gaps, ambiguously aligned sites, and sequences whose real identity could not be confirmed were removed from alignments. Phylogenetic analyses were performed with maximum likelihood method using PhyML 3.1 (Guindon et al., 2010) and distance method using neighbor of PHYLIP-3.695 (Felsenstein, 2013). The optimal model of protein substitution and rate heterogeneity were chosen based on the result of ModelGenerator (Keeane et al., 2006). Bootstrap analyses were performed using 1,000 replicates.

RESULTS

The Identity of “Algal” Genes in Oomycetes

If the previously identified “algal” genes in *Phytophthora* are indeed derived from a red algal endosymbiont acquired by the ancestor of stramenopiles, their homologs might also be found in photosynthetic stramenopiles. Given their presumably red algal nature, these stramenopile sequences theoretically should have a closer relationship to homologs from red algae (or red algae and other photosynthetic eukaryotes plus cyanobacteria) than to those from other organisms (e.g., opisthokonts, amoebozoans, non-cyanobacterial bacteria, and archaea). Our BLAST results with a larger taxonomic sampling showed that, for all of the 30 most likely “algal” genes previously identified in *Phytophthora*, only 10 of them (about 33%) were more similar to sequences of red algae, photosynthetic eukaryotes and/or cyanobacteria (Table 1); these 10 genes were also the viable candidate genes of red algal origin.

Of the 30 “algal” genes in *Phytophthora*, nine (30%) showed stronger BLAST matches (represented by higher bit scores) to homologs from opisthokonts, amoebozoans, non-cyanobacterial bacteria or archaea than to those from photosynthetic stramenopiles (Table 1). Another gene encoding methylthioadenosine phosphorylase (*P. ramorum* Gene ID 86425) had no detectable homologs in sequenced photosynthetic stramenopiles. Although the possibility of differential gene losses cannot be ruled out, genes with such a distribution pattern may also suggest an independent origin in oomycetes, such as horizontal gene transfer (HGT) from prokaryotes or other eukaryotes to oomycetes. Moreover, 16 of these 30 genes (about 53%) had more significant matches to homologs from opisthokonts, amoebozoans, non-cyanobacterial bacteria or archaea than to those from red algae and cyanobacteria (Table 1). Particularly, four genes had the strongest BLAST matches in non-cyanobacterial bacteria, and two in opisthokonts or amoebozoans. This observation based on simple pairwise comparisons suggests that many of these “algal” genes in oomycetes have no stronger similarity to photosynthetic

¹<http://genome.jgi.doe.gov/programs/fungi/index.jsf>

TABLE 1 | BLAST search results of 30 most likely “algal” genes in *Phytophthora*.

<i>P. ramorum</i> Gene ID	Putative gene product	Target	Best BLASTp match (represented by bit scores)						
			Cyanobacteria	Red algae	Photosynthetic stramenopiles	Opisthokonts/ amoebozoans	Non-cyanobacterial bacteria	Archaea	
72019	Cobalamin-independent methionine synthase	8056	975	861	759	808	555	Figure 1	
54177	Prolyl oligopeptidase II	6333	822	733	493	630	214	Supplementary Figure 1	
75281	2'-Isopropylmalate synthase	*	593	543	545	618	528	Figure 6	
54068	Threonine ammonia-lyase	*	482	469	571	503	218	Supplementary Figure 2	
79142	Anthranilate synthase	*	580	498	543	474	258	Supplementary Figure 3	
38584	NCAIR mutase	*	234	none	224	126	193	Figure 3	
74880	3'-phosphoadenosine 5'-phosphosulfate reductase	2683	217	281	135	256	120	Figure 7	
51635	Uroporphyrinogen-III methyltransferase	*	217	174	161	<u>199</u>	<u>213</u>	Supplementary Figure 4	
95818	tRNA (guanine-N(7))-methyltransferase-like	*	176	148	149	91	<u>162</u>	none	
80275	Phosphatidate cytidyltransferase	108	87	85.9	102	<u>130</u>	none	Supplementary Figure 6	
87801	Ketol-acid reductoisomerase	*	599	639	665	190	640	212	
80380	Phosphoserine aminotransferase	*	393	407	436	405	417	Supplementary Figure 7	
72085	Asparaginyl tRNA synthetase	*	478	482	474	472	<u>486</u>	Supplementary Figure 8	
75838	SAICAR synthetase	*	215	383	427	172	388	256	
72293	Glucokinase	2556	278	328	265	237	127	Figure 4	
75742	Histidino-phosphate aminotransferase	99	332	491	426	116	86.7	Supplementary Figure 10	
78949	Zinc carboxypeptidase A	none	259	379	71	75	64	Supplementary Figure 11	
79657	cAMP-binding mitochondrial solute carrier	60.8	313	131	157	107	81	Figure 2	
45002	Enoyl-(acyl-carrier-protein) reductase	*	204	209	186	183	223	144	
77863	Sulfur transferase + methyl transferase fusion	166	296	288	186	176	127	Supplementary Figure 13	
86425	Methylthioadenosine phosphorylase	78	167	none	168	<u>127</u>	<u>142</u>	Supplementary Figure 14	
54215	Ribonuclease HII	160	185	230	154	199	81.3	Supplementary Figure 15	
71442	Nitrate reductase	141	717	703	625	224	115	Supplementary Figure 16	
71783	6-phosphogluconate dehydrogenase	549	545	799	559	538	538	Supplementary Figure 17	
833828	Aspartate kinase/homoserine dehydrogenase	180	222	214	157	233	71	Supplementary Figure 18	
73217	Galactonolactone dehydrogenase	158	418	417	409	147	140	Supplementary Figure 19	
85610	Cobalamin synthesis protein	314	319	379	389	325	167	Supplementary Figure 20	
78464	tRNA dihydrouridine synthase	236	240	332	187	253	88	Supplementary Figure 21	
72218	Folate-bipterin transporter	253	239	331	65	211	none	Supplementary Figure 22	
82990	Prephenate dehydratase family	*	95.9	127	199	157	160	148	Supplementary Figure 23

Phytophthora ramorum gene IDs are from the JGI database. Genes whose protein products are presumably localized in mitochondria in *Phytophthora*, but localized in plastids of photosynthetic eukaryotes according to the original *Phytophthora* genome paper (Tyler et al., 2006) are indicated by asterisks in Target column. Bit scores of sequence matches from opisthokonts, amoebozoans, non-cyanobacterial bacteria or archaea are underlined if higher than those from photosynthetic stramenopiles, and in bold if higher than those from cyanobacteria and red algae.

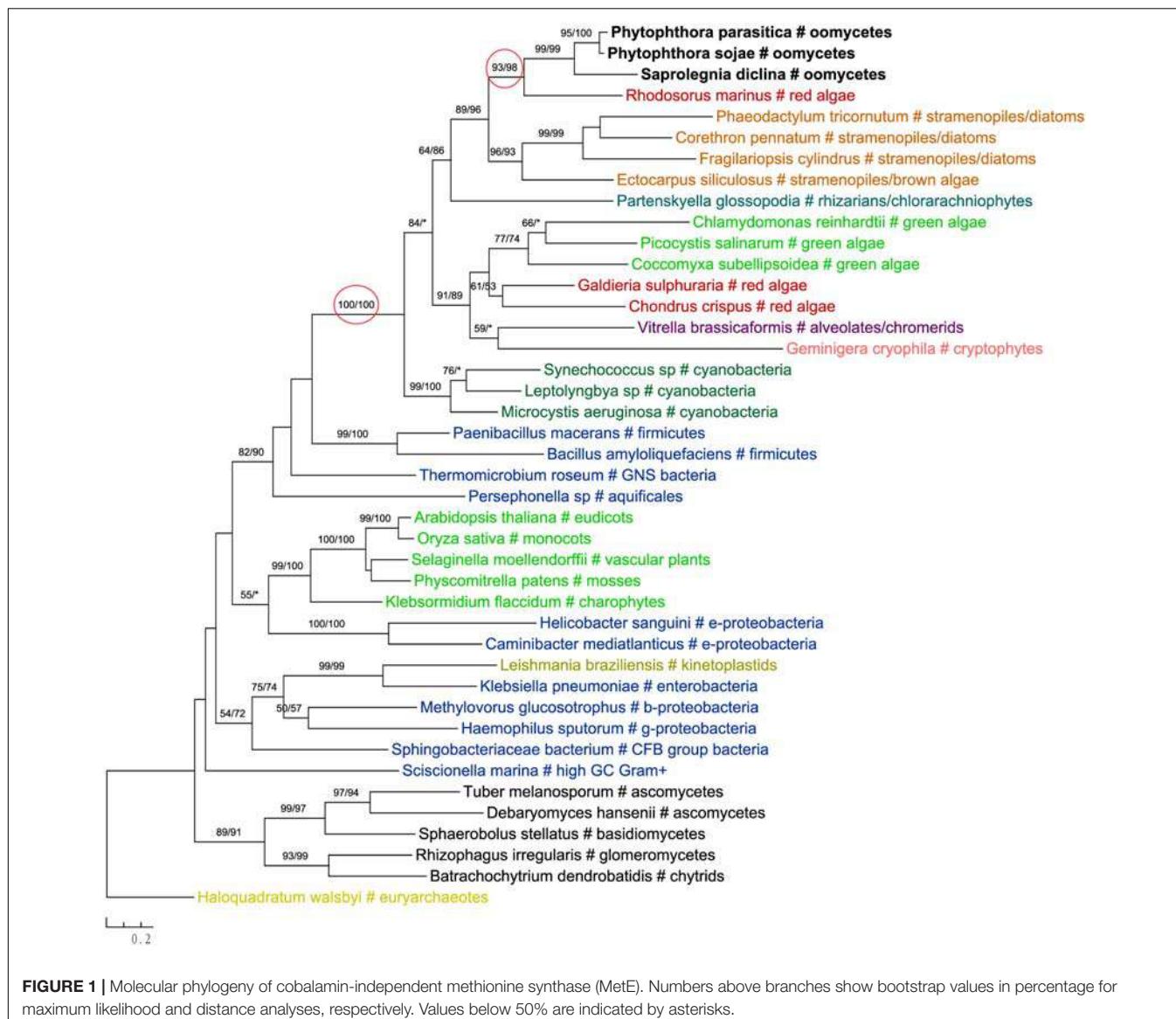


FIGURE 1 | Molecular phylogeny of cobalamin-independent methionine synthase (MetE). Numbers above branches show bootstrap values in percentage for maximum likelihood and distance analyses, respectively. Values below 50% are indicated by asterisks.

stramenopile, red algal or cyanobacterial sequences. If sequence similarity is largely correlated with sequence relatedness, as commonly believed, the nature of these “algal” genes might be seriously questioned.

We further performed phylogenetic analyses on each of these 30 “algal” genes to evaluate its origin. If an oomycete gene is of red algal origin, the gene and its stramenopile (or chromalveolate) homologs are expected to form a clade sister to red algal and/or cyanobacterial sequences. This, however, was not the pattern uncovered in our study. Tree topologies for 21 (70%) genes were poorly supported overall (or the position of oomycete sequences couldn’t be confidently determined), thus providing no sufficient evidence for any evolutionary scenarios (Supplementary Materials). These poorly supported tree topologies might be caused by multiple issues, for example, insufficient phylogenetic signal or heterogeneity in evolutionary rates. Nevertheless, such topologies, combined with

the information of phylogenetic distribution from BLAST search, should not be interpreted as evidence for a red algal origin of involved *Phytophthora* genes. The remaining genes had relatively well-resolved phylogenies and will be detailed in the following sections.

Algal or Cyanobacterial Genes in Oomycetes

In our analyses, several of these “algal” genes indeed showed a close affinity with algal or cyanobacterial sequences. In addition, for 12 “algal” genes previously identified in *Phytophthora*, the protein products of their plant and/or algal homologs are localized in plastids (Tyler et al., 2006), as predicted by TargetP (Emanuelsson et al., 2000) (Table 1). It is well known that proteins of organelles-derived genes are often re-imported into the original organelles (mitochondria or

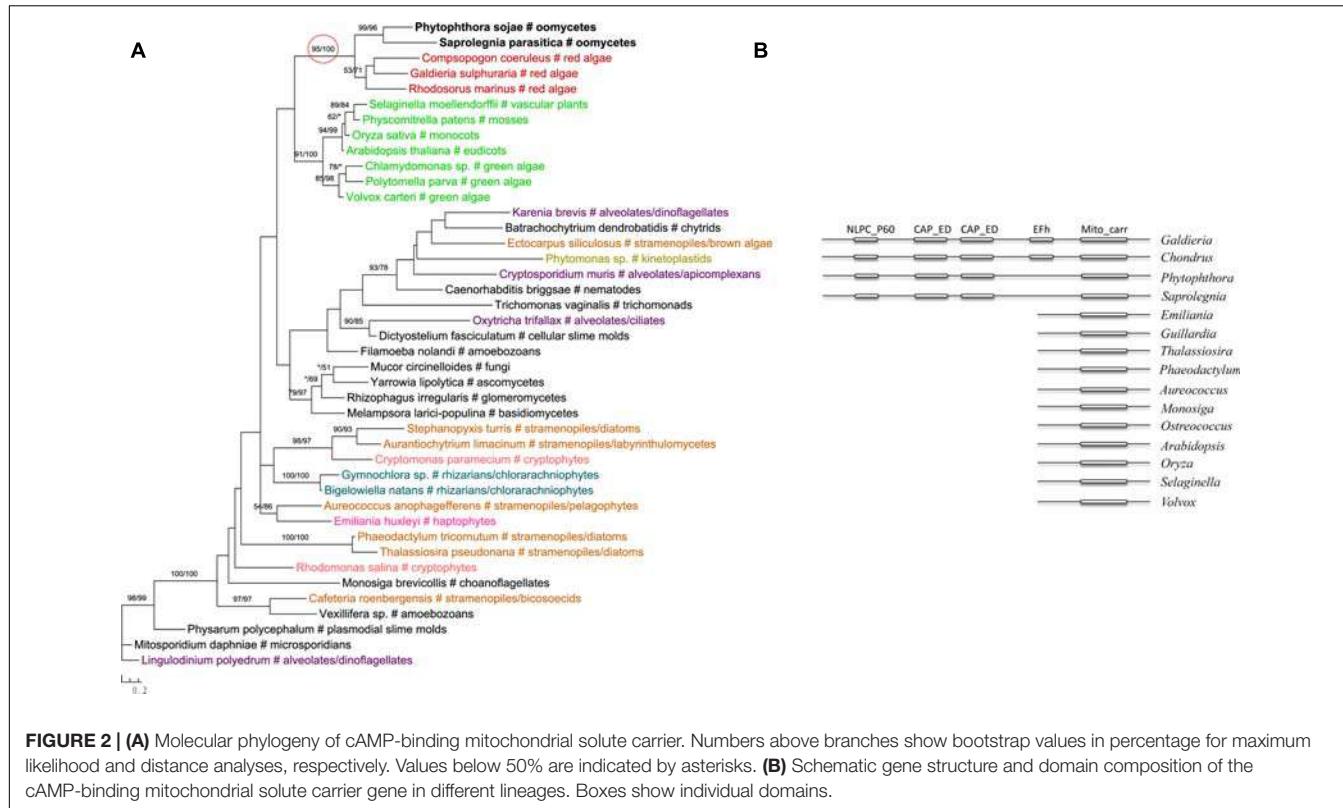


FIGURE 2 | (A) Molecular phylogeny of cAMP-binding mitochondrial solute carrier. Numbers above branches show bootstrap values in percentage for maximum likelihood and distance analyses, respectively. Values below 50% are indicated by asterisks. **(B)** Schematic gene structure and domain composition of the cAMP-binding mitochondrial solute carrier gene in different lineages. Boxes show individual domains.

plastids) to participate in related biochemical activities (Bogorad, 1975; Ellis, 1981; Weeden, 1981; Timmis et al., 2004). This information has been frequently used as supplemental evidence for genes of organellar origin. However, such affinity with algal/cyanobacterial sequences or functionality in other plastids might not necessarily support the suggestion of a historical red algal endosymbiont in the ancestral stramenopile.

The most likely “algal” gene uncovered in our current study encodes cobalamin-independent methionine synthase (*MetE*). Our phylogenetic analyses indicated that *MetE* sequences from oomycetes, photosynthetic stramenopiles, chlorarachniophytes, chromerids and cryptophytes formed a large group with homologs of red algae, green algae, and cyanobacteria (Figure 1). Within this group, oomycete *MetE* sequences formed a strongly supported clade with red algal instead of other stramenopile homologs. Although the overall molecular phylogeny of *MetE* is consistent with an algal origin of oomycetes and, to a certain extent, an algal/cyanobacterial origin of all stramenopiles, the strength of this evidence is somewhat compromised by the fact that oomycete and other stramenopile sequences didn’t form a monophyletic group (see Discussion). Two other similar cases are related to the genes encoding prolyl oligopeptidase II (Supplementary Figure 1) and cAMP-binding mitochondrial solute carrier (Figure 2). For both genes, their molecular phylogenies showed that oomycete and red algal sequences were closely related. Particularly in the latter case, a NLPC_P60 and two CAP_ED domains are uniquely shared by oomycetes and red algae, but are absent from other stramenopiles (Figure 2).

A parsimonious explanation for these findings would be that oomycetes obtained this gene from red algae directly or vice versa.

Two genes in our analyses were found to be specifically related to green plant sequences, which is in disagreement with the suggestion of a red algal plastid in the ancestral stramenopile. The gene encoding NCAIR mutase does not have detectable homologs in red algae. Phylogenetic analyses of NCAIR mutase supported a monophyletic group including sequences from oomycetes, photosynthetic stramenopiles, dinoflagellates, green algae and cyanobacteria (Figure 3). Because of the lack of detectable NCAIR mutase homologs in red algae, a red algal origin of this gene in all stramenopiles cannot be concluded. On the other hand, an independent green algal endosymbiont in stramenopiles might potentially explain such a distribution pattern (Moustafa et al., 2009; Dorrell and Smith, 2011). The other green plants-related gene in oomycetes encodes a probable folate-biopterin transporter (Supplementary Figure 22). Our analyses showed that sequences from oomycetes, diatom *Thalassionema frauenfeldii* and land plants formed a strongly supported clade, whereas other photosynthetic sequences, including red algae and cyanobacteria, formed another large group with only modest support.

In addition to primary algae and cyanobacteria, several groups of eukaryotes that have secondary plastids through higher-order endosymbioses might also be potential donors for genes in oomycetes. For instance, phylogenetic analyses of glucokinase indicated that sequences from oomycetes, haptophytes and ciliates formed a well-supported clade, which in turn grouped

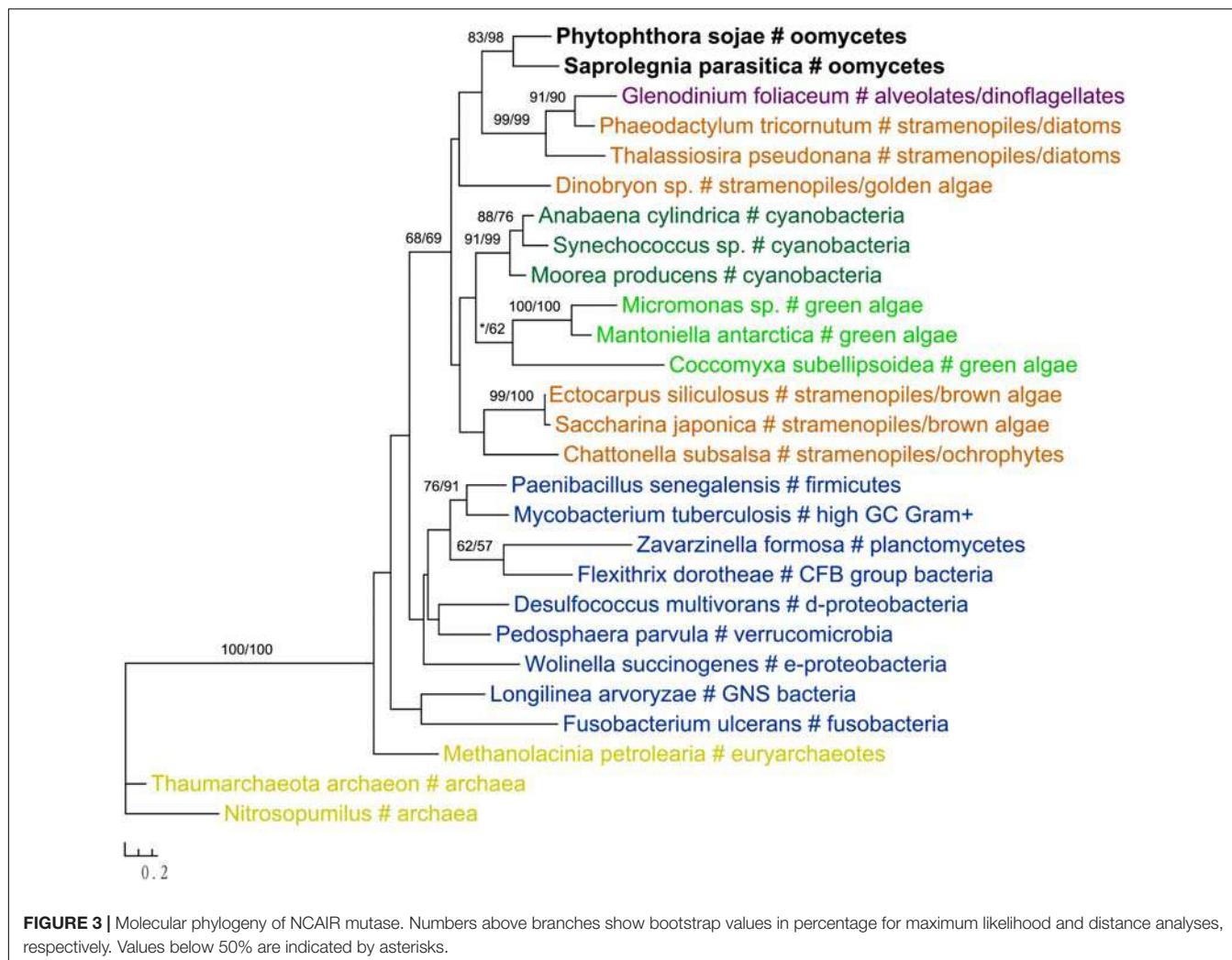


FIGURE 3 | Molecular phylogeny of NCAIR mutase. Numbers above branches show bootstrap values in percentage for maximum likelihood and distance analyses, respectively. Values below 50% are indicated by asterisks.

with homologs from photosynthetic stramenopiles, red algae, green algae, choanoflagellates and dinoflagellates (Figure 4). A similar case was also observed for the gene encoding ketol-acid reductoisomerase (Figure 5). As it is known that ciliates contain sequences of algal origin (Reyes-Prieto et al., 2008), this topology might suggest HGT from haptophytes-related groups to oomycetes, and again provides no support for a common photosynthetic origin between oomycetes and other stramenopiles.

Other Potential Evolutionary Scenarios

As indicated above, a large fraction of “algal” genes in oomycetes,stramenopiles showed stronger matches in our BLAST search to homologs from opisthokonts, amoebozoans or non-cyanobacterial bacteria rather than those from red algae and cyanobacteria. For several of these genes, this relationship was also confirmed by subsequent phylogenetic analyses.

One of these genes encodes 2-isopropylmalate synthase in leucine biosynthesis and was previously detailed in the *Phytophthora* genome paper (Tyler et al., 2006). According to the authors, this gene was subject to at least two transfer

events in eukaryotes: sequences of primary photosynthetic eukaryotes and stramenopiles (including oomycetes) were derived from cyanobacteria, whereas sequences of fungi were from α -proteobacteria. Specifically, diatom sequences were found to group with green plant rather than red algal homologs, which was interpreted as a separate ancestry or artifacts due to incomplete sampling (Tyler et al., 2006). Our current analyses support the previous conclusion that this gene in stramenopiles might have different origins, but also suggest a potentially more complicated evolutionary scenario. While sequences from brown algae and cryptophytes indeed grouped with red algal homologs, those from diatoms and *Aureococcus* with green plant sequences instead (Figure 6). The relationships between brown algae, cryptophytes and red algae uncovered here is in line with the suggestion of serial endosymbioses by Stiller et al. (2014), where a red alga was first adopted by a cryptophyte that was in turn engulfed by ochrophytes. The sequence affiliation between diatoms, *Aureococcus* and green algae might point to separate origins of this gene in other photosynthetic stramenopiles [e.g., from a potential green algal endosymbiont (Moustafa et al., 2009; Dorrell and Smith, 2011)]

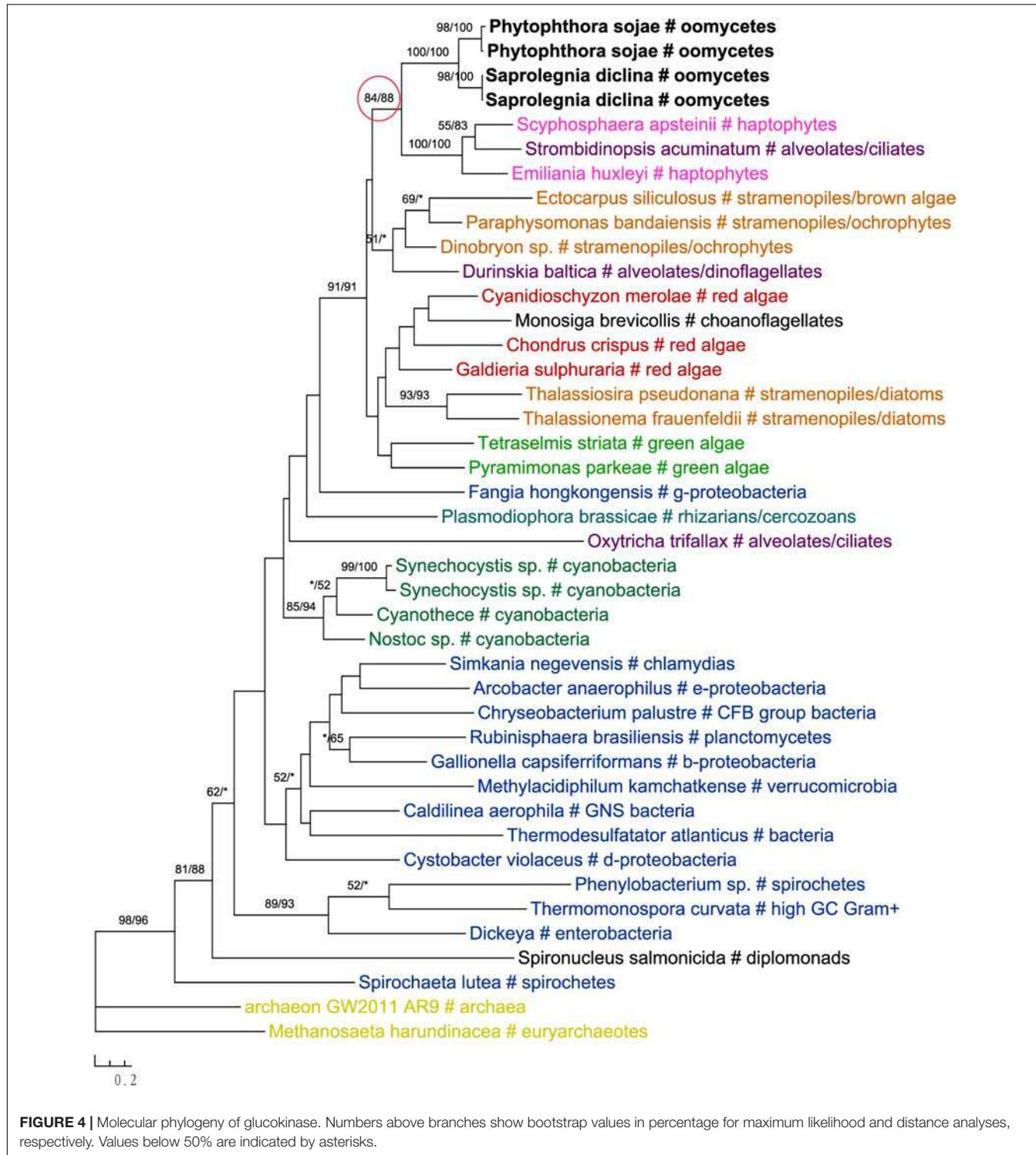


FIGURE 4 | Molecular phylogeny of glucokinase. Numbers above branches show bootstrap values in percentage for maximum likelihood and distance analyses, respectively. Values below 50% are indicated by asterisks.

or an independent HGT event]. Nevertheless, unlike previously reported in the *Phytophthora* genome paper (Tyler et al., 2006), oomycete sequences grouped with labyrinthulomycetes, another group of heterotrophic stramenopiles, and other eukaryotes, rather than being affiliated with diatoms, primary photosynthetic eukaryotes and cyanobacteria (Figure 6).

The gene encoding 3'-phosphoadenosine 5'-phosphosulfate reductase (PAPR), an enzyme in the sulfate assimilation pathway, is another example highlighting the potential pitfalls of insufficient sampling. PAPR and adenosine 5'-phosphosulfate reductase (APR) are homologous proteins and have a complex evolutionary history in eukaryotes (Kopriva et al., 2002; Kopriva

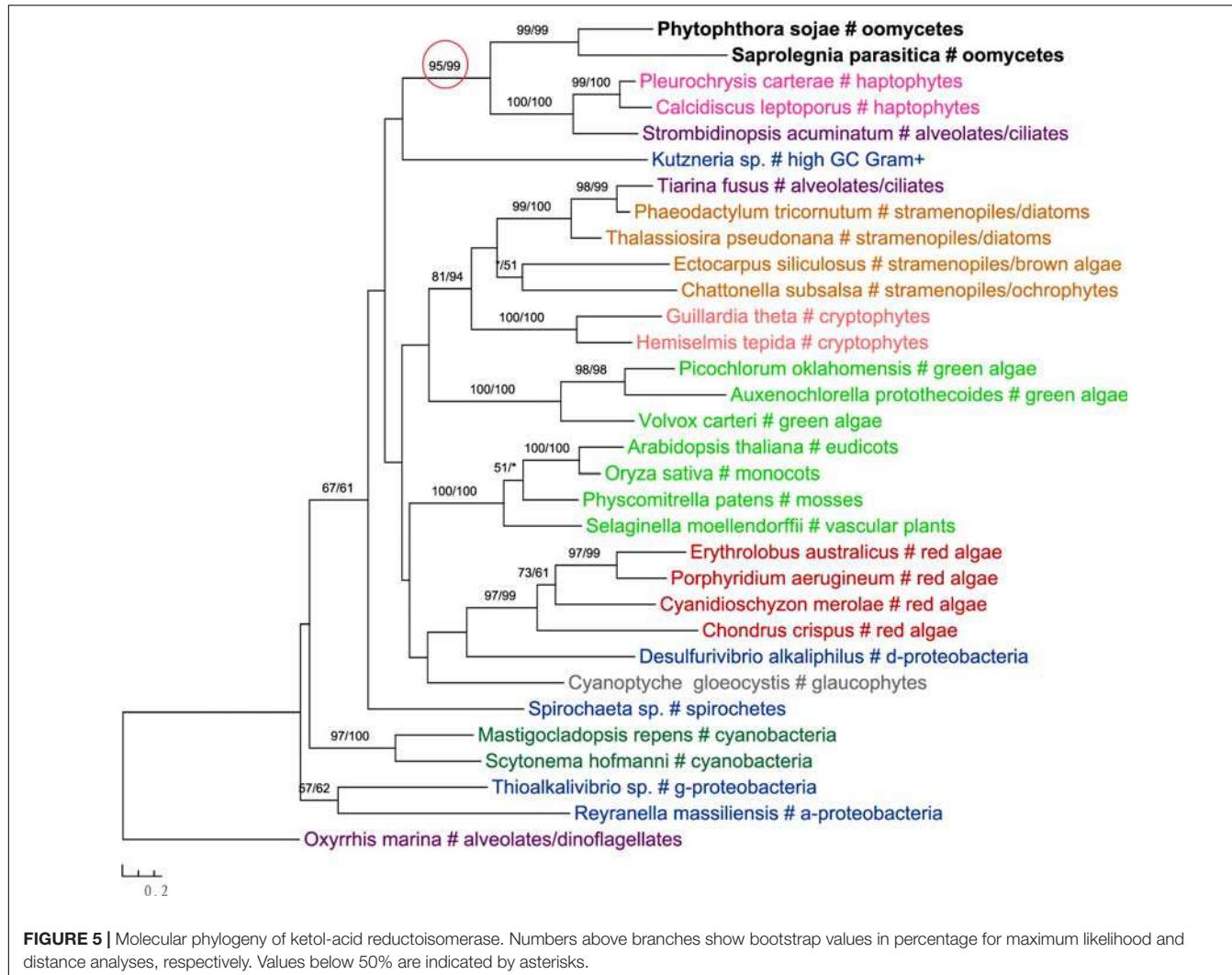


FIGURE 5 | Molecular phylogeny of ketol-acid reductoisomerase. Numbers above branches show bootstrap values in percentage for maximum likelihood and distance analyses, respectively. Values below 50% are indicated by asterisks.

and Koprivova, 2004; Patron et al., 2008). The *APR* gene was previously thought to exist in land plants, algae, and phototrophic bacteria. *PAPR*, on the other hand, was initially identified mainly in fungi and bacteria (Kopriva et al., 2002). Several more recent studies reported *PAPR* sparingly in phototrophic eukaryotes, suggesting potential HGT events (Kopriva and Koprivova, 2004; Kopriva et al., 2007; Patron et al., 2008). Particularly, the study of Patron et al. (2008) indicated a potential bacterial origin of *PAPR* in *P. sojae*. With a much larger taxonomic sampling, our analyses showed that sequences from some stramenopiles (including oomycetes), bacteria (both cyanobacteria and non-cyanobacteria) and *Paulinella chromatophora* formed a major *PAPR* clade (Figure 7). The cyanobacterial origin of *PAPR* in *P. chromatophora* is somewhat expected, as this species contains an independently evolved plastid organelle (cyanobacterial endosymbiont) (Marin et al., 2005). As the overall topology of this clade is poorly supported, whether *PAPR* in stramenopiles was derived from a red algal endosymbiont or a separate HGT event could not be answered by our study.

DISCUSSION

As evidence of historical plastids in oomycetes, the “algal” genes identified in *Phytophthora* genomes were used to support a common photosynthetic ancestry of stramenopiles, and the chromalveolate hypothesis in general. In the *Phytophthora* genome paper (Tyler et al., 2006), the identification of “algal” genes was heavily based on significant matches to sequences from red algae or cyanobacteria. Because the identification of foreign genes in eukaryote can be affected by taxonomic samplings and methods of analyses, studies on algal genes in different eukaryotes sometimes led to different interpretations after re-analyses. For example, 263 red algal genes and 250 green plant genes were reported in *Chromera velia* (Woehle et al., 2011), but only 23 and nine of them, respectively, were confirmed after re-evaluation (Burki et al., 2012a). When more stringent criteria were applied, the number of putative green algal genes in diatoms decreased from 1,700 (Moustafa et al., 2009) to only 144 (Dorrell and Smith, 2011). While an algal endosymbiont in the common ancestor of stramenopiles or any other lineages could certainly

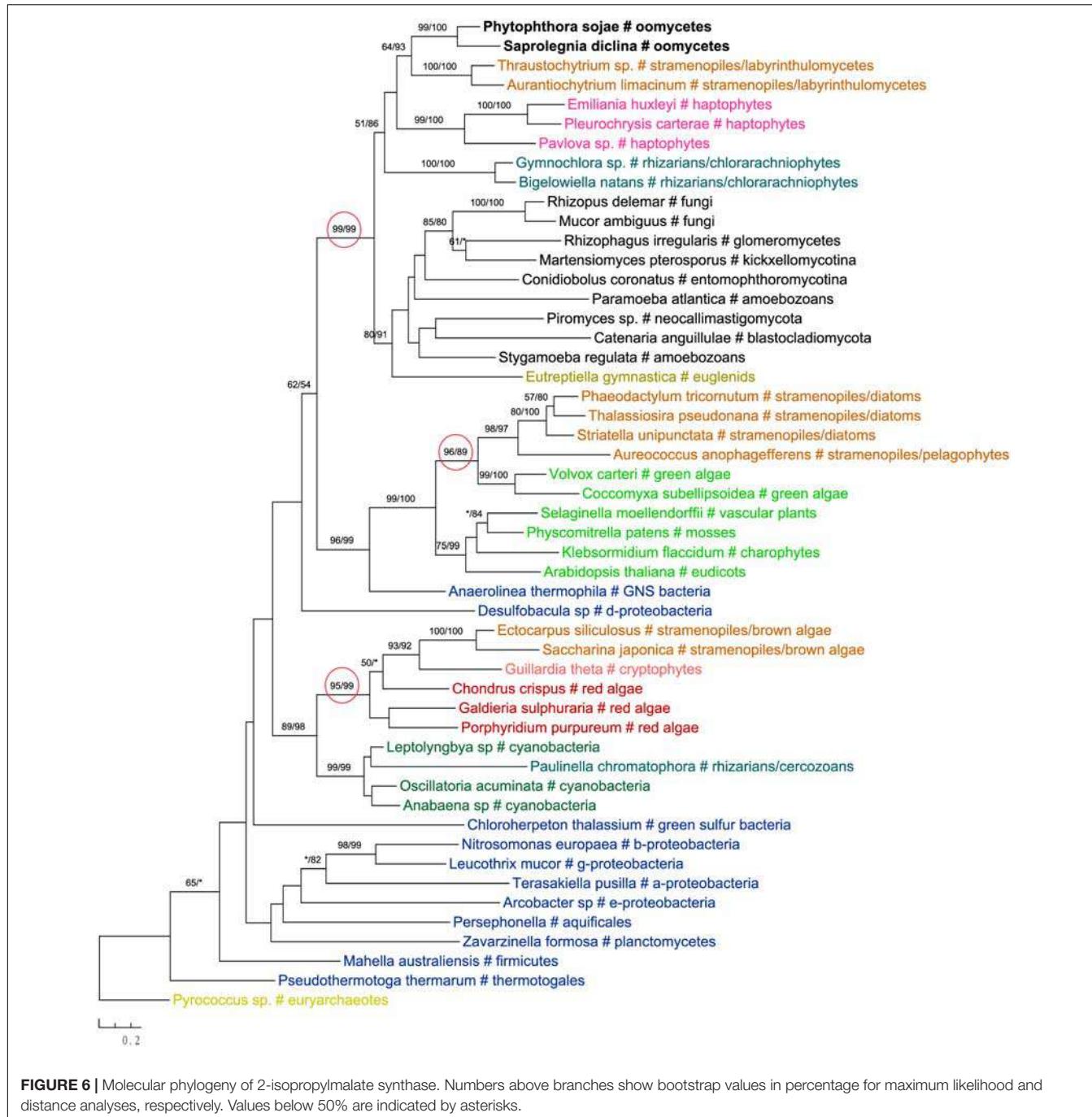


FIGURE 6 | Molecular phylogeny of 2-isopropylmalate synthase. Numbers above branches show bootstrap values in percentage for maximum likelihood and distance analyses, respectively. Values below 50% are indicated by asterisks.

be a significant source of foreign genes, other issues, notably phylogenetic artifacts, insufficient sampling, differential gene losses and independent HGT events, could all lead to the same or similar atypical gene distributions or relationships.

With a much larger sampling and careful phylogenetic analyses, we revisited the 30 most likely “algal” genes identified in the *Phytophthora* genomes (Tyler et al., 2006). Our results show that the identification of these “algal” genes, to a great extent, was affected by limited genome data then available for certain eukaryotic lineages. Almost none of these 30 genes

confidently supports the hypothesis of a red algal endosymbiont in the common ancestor of stramenopiles. Although the molecular phylogeny of *MetE* is indeed consistent with the suggestion of a photosynthetic ancestry of stramenopiles, its topology does not strictly support a historical red plastid in this lineage. As such, our current study is largely consistent with the statistical genome analyses of Stiller et al. (2009), which found no evidence for a red algal endosymbiont in the ancestral stramenopile. However, we should also caution here that, because the parasitic nature of oomycetes, the possibility

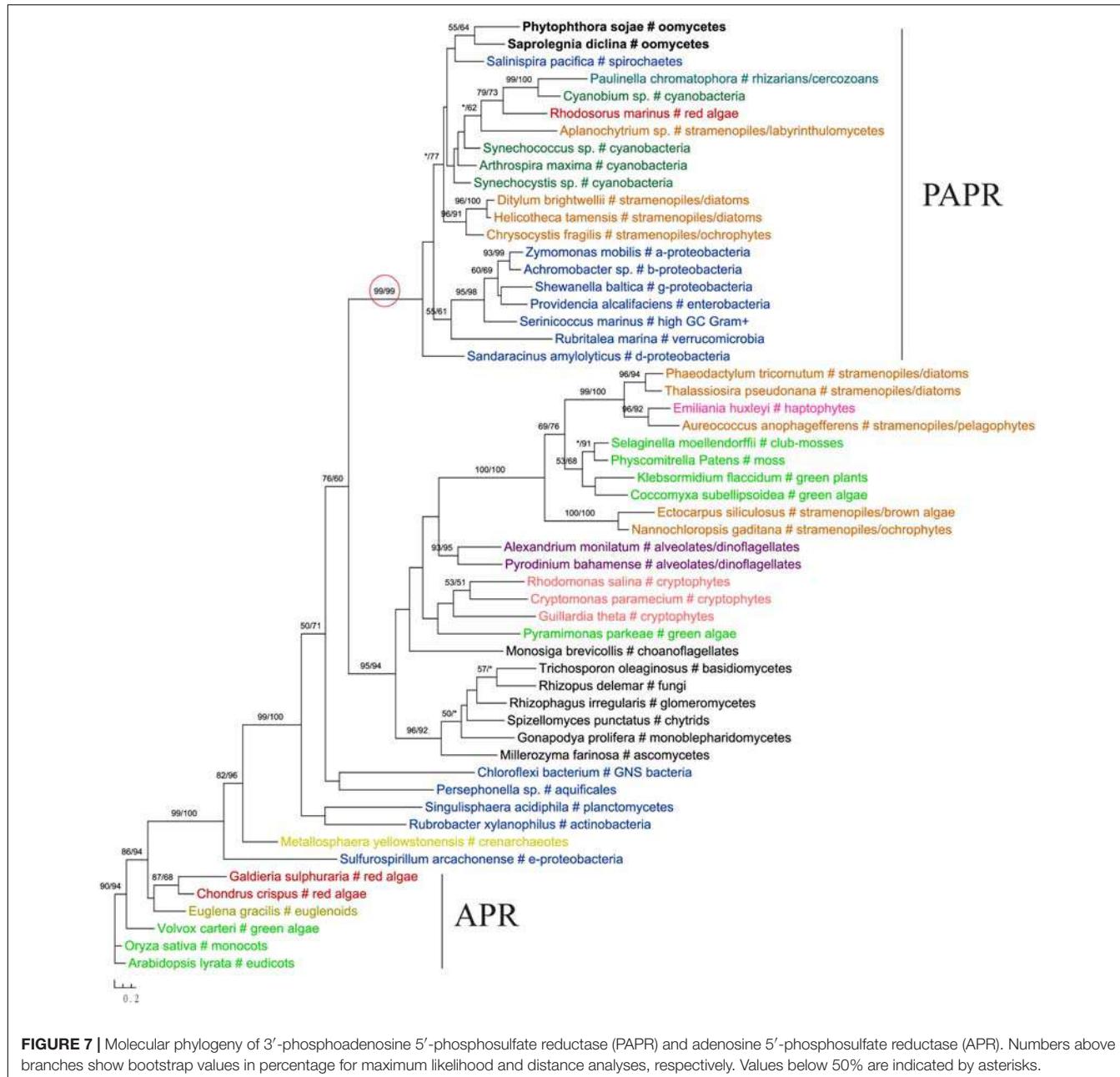


FIGURE 7 | Molecular phylogeny of 3'-phosphoadenosine 5'-phosphosulfate reductase (PAPR) and adenosine 5'-phosphosulfate reductase (APR). Numbers above branches show bootstrap values in percentage for maximum likelihood and distance analyses, respectively. Values below 50% are indicated by asterisks.

of plastid loss during oomycete evolution cannot be entirely excluded based on our data. Furthermore, given the fact that many of the sampled sequences in our analyses were from transcriptomic data, it is unclear whether and how the data quality, for example potential sequencing contamination, might have affected our results. Additional investigations are needed to resolve this significant, nevertheless controversial, issue of eukaryotic evolution.

On the other hand, our results also indicate that the abnormal phylogenetic signal of these “algal” genes might be caused by a complex evolutionary history of oomycetes or stramenopiles. Although the origins of these 30 genes in oomycetes or

stramenopiles are not always clear, several of them were found to be related to miscellaneous algae. To a certain extent, such sequence relatedness to various lineages might be attributed to other potential historical endosymbioses or independent HGT events involving oomycetes or stramenopiles. For instance, in lieu of the chromalveolate hypothesis, serial endosymbioses between different photosynthetic lineages have been proposed to explain the evolution of red algal plastids (Sanchez-Puerta and Delwiche, 2008; Stiller et al., 2014). A potential green algal endosymbiont was also suggested in stramenopiles (Moustafa et al., 2009; Dorrell and Smith, 2011). Furthermore, horizontally acquired genes have been reported in different eukaryotic

lineages (Richardson and Palmer, 2007; Keeling and Palmer, 2008; Andersson, 2009; Dunning Hotopp, 2011; Huang and Yue, 2012; Soucy et al., 2015; Qiu et al., 2016), even though some of the earlier reports might turn out to be false positives as suggested by our current study. Especially for microbial eukaryotes, the importance and frequency of HGT in their evolution is increasingly being appreciated (Keeling and Palmer, 2008; Andersson, 2009), and there is evidence that microbial eukaryotes might have frequently acquired genes from various organisms, instead of a specific source of endosymbiotic relationship (Huang et al., 2004; Loftus et al., 2005; Carlton et al., 2007; Bowler et al., 2008; Sun et al., 2010; Yue et al., 2013). Oomycetes originated in marine environments and gradually spread to freshwater and terrestrial environments (Beakes and Sekimoto, 2009; Beakes et al., 2012). Bacteria, miscellaneous algae or other organisms in a common habitat could be potential sources of foreign genes in oomycetes. Additionally, feeding activities of their ancestors in aquatic environments or the parasitic feature of modern species [many oomycetes are parasites; for instance, the early diverging species *Eurychasma dicksonii* is an obligate parasite of marine brown algae (Küpper and Müller, 1999; Gachon et al., 2009; Strittmatter et al., 2009)] might have also facilitated genes acquisition in oomycetes. Indeed, several studies have already reported gene acquisition events in oomycetes and other stramenopiles, including fungi to oomycetes (Richards et al., 2006, 2011), bacteria to diatoms (Bowler et al., 2008), and different prokaryotic or eukaryotic sources to *Blastocystis*

REFERENCES

- Andersson, J. O. (2009). Gene transfer and diversification of microbial eukaryotes. *Annu. Rev. Microbiol.* 63, 177–193. doi: 10.1146/annurev.micro.091208.073203
- Archibald, J. M., Longet, D., Pawłowski, J., and Keeling, P. J. (2003). A novel polyubiquitin structure in Cercozoa and Foraminifera: evidence for a new eukaryotic supergroup. *Mol. Biol. Evol.* 20, 62–66. doi: 10.1093/molbev/msg006
- Baldauf, S. L., Roger, A. J., Wenk-Siefert, I., and Doolittle, W. F. (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* 290, 972–977. doi: 10.1126/science.290.5493.972
- Beakes, G. W., Glockling, S. L., and Sekimoto, S. (2012). The evolutionary phylogeny of the oomycete “fungi”. *Protoplasma* 249, 3–19. doi: 10.1007/s00709-011-0269-2
- Beakes, G. W., and Sekimoto, S. (2009). “The evolutionary phylogeny of Oomycetes—insights gained from studies of holocarpic parasites of algae and invertebrates,” in *Oomycete Genetics and Genomics: Diversity, Interactions, and Research Tools*, eds K. Lamour and S. Kamoun (New York, NY: Wiley), 24–24.
- Bodyl, A. (2005). Do plastid-related characters support the chromalveolate hypothesis? *J. Phycol.* 41, 712–719. doi: 10.1111/j.1529-8817.2005.00091.x
- Bodyl, A., and Moszczyński, K. (2006). Did the peridinin plastid evolve through tertiary endosymbiosis? A hypothesis. *Eur. J. Phycol.* 41, 435–448. doi: 10.1080/09670260600961080
- Bodyl, A., Stiller, J. W., and Mackiewicz, P. (2009). Chromalveolate plastids: direct descent or multiple endosymbioses? *Trends Ecol. Evol.* 24, 119–121. doi: 10.1016/j.tree.2008.11.003
- Bogorad, L. (1975). Evolution of organelles and eukaryotic genomes. *Science* 188, 891–898. doi: 10.1126/science.1138359
- Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., et al. (2008). The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456, 239–244. doi: 10.1038/nature07410
- Brown, J. W., and Sorhannus, U. (2009). A molecular genetic timescale for the diversification of autotrophic stramenopiles (Ochrophyta): substantive underestimation of putative fossil ages. *PLoS ONE* 5:e12759. doi: 10.1371/journal.pone.0012759
- Burki, F., Flegontov, P., Oborník, M., Cihlář, J., Pain, A., Lukeš, J., et al. (2012a). Re-evaluating the green versus red signal in eukaryotes with secondary plastid of red algal origin. *Genome Biol. Evol.* 4:evs049. doi: 10.1093/gbe/evs049
- Burki, F., Kaplan, M., Tikhonenkov, D. V., Zlatogursky, V., Minh, B. Q., Radaykina, L. V., et al. (2016). Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proc. Biol. Sci.* 283, 20152802. doi: 10.1098/rspb.2015.2802
- Burki, F., Okamoto, N., Pombert, J.-F., and Keeling, P. J. (2012b). The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. *Proc. Biol. Sci.* 279, 2246–2254. doi: 10.1098/rspb.2011.2301
- Burki, F., Shalchian-Tabrizi, K., and Pawłowski, J. (2008). Phylogenomics reveals a new ‘megagroup’ including most photosynthetic eukaryotes. *Biol. Lett.* 4, 366–369. doi: 10.1098/rsbl.2008.0224
- Carlton, J. M., Hirt, R. P., Silva, J. C., Delcher, A. L., Schatz, M., Zhao, Q., et al. (2007). Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 315, 207–212. doi: 10.1126/science.1132894
- Cavalier-Smith, T. (1981). Eukaryote kingdoms: seven or nine? *Biosystems* 14, 461–481.
- Cavalier-Smith, T. (1986). The kingdom Chromista: origin and systematics. *Prog. Phycol. Res.* 4, 309–347.
- Cavalier-Smith, T. (1995). “Membrane heredity, symbiogenesis, and the multiple origins of algae,” in *Biodiversity and Evolution*, eds R. Arai, M. Kato, and Y. Doi (Tokyo: The National Science Museum Foundation), 75–114.
- Cavalier-Smith, T. (1999). Principles of protein and lipid targeting in secondary Symbiogenesis: Euglenoid, Dinoflagellate, and Sporozoan plastid origins and the eukaryote family tree1, 2. *J. Eukaryot. Microbiol.* 46, 347–366. doi: 10.1111/j.1550-7408.1999.tb04614.x
- Cavalier-Smith, T., and Chao, E. E.-Y. (2003). Phylogeny of choanozoa, apusozoa, and other protozoa and early eukaryote megaevolution. *J. Mol. Evol.* 56, 540–563. doi: 10.1007/s00239-002-2424-z
- Cavalier-Smith, T., Couch, J., Thorsteinse, K., Gilson, P., Deane, J., Hill, D., et al. (1996). Cryptomonad nuclear and nucleomorph 18S rRNA phylogeny. *Eur. J. Phycol.* 31, 315–328. doi: 10.1080/0967026960651541
- (Tsaousis et al., 2012; Eme et al., 2017). In this regard, our finding of multiple foreign genes in oomycetes might reflect the interactions among red/green algae, oomycetes/stramenopiles, and other microbes, as well as their ensuing genetic integration.

AUTHOR CONTRIBUTIONS

JH conceived the study and wrote the manuscript. QW performed the analyses and wrote the manuscript. HS contributed to the analyses.

FUNDING

This work is supported by the Major Program of National Natural Science Foundation of China (grant no. 31590823) to HS, CAS “Light of West China” Program to JH, and a joint Ph.D. training program of University of Chinese Academy of Sciences (UCAS[2015]37) to QW.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01540/full#supplementary-material>

- Cavalier-Smith, T., and Scoble, J. M. (2013). Phylogeny of Heterokonta: *incisomonas marina*, a uniciliate gliding opalozoan related to *Solenicola* (Nanomonadea), and evidence that Actinophryida evolved from raphidophytes. *Eur. J. Protistol.* 49, 328–353. doi: 10.1016/j.ejop.2012.09.002
- Delwiche, C. F. (1999). Tracing the thread of plastid diversity through the tapestry of life. *Am. Nat.* 154, S164–S177.
- Delwiche, C. F., and Palmer, J. D. (1997). *The Origin of Plastids and their Spread via Secondary Symbiosis*. Vienna: Springer.
- Derelle, R., López-García, P., Timpano, H., and Moreira, D. (2016). A phylogenomic framework to study the diversity and evolution of stramenopiles (= heterokonts). *Mol. Biol. Evol.* 33, 2890–2898. doi: 10.1093/molbev/msw168
- Dorrell, R. G., and Smith, A. G. (2011). Do red and green make brown?: perspectives on plastid acquisitions within chromalveolates. *Eukaryot. Cell* 10, 856–868. doi: 10.1128/EC.00326-10
- Dunning Hotopp, J. C. (2011). Horizontal gene transfer between bacteria and animals. *Trends Genet.* 27, 157–163. doi: 10.1016/j.tig.2011.01.005
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Ellis, R. J. (1981). Chloroplast proteins: synthesis, transport, and assembly. *Annu. Rev. Plant Physiol.* 32, 111–137. doi: 10.1146/annurev.pp.32.060181.000551
- Emanuelsson, O., Nielsen, H., Brunak, S., and Von, H. G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005–1016. doi: 10.1006/jmbi.2000.3903
- Eme, L., Gentekaki, E., Curtis, B., Archibald, J. M., and Roger, A. J. (2017). Lateral gene transfer in the adaptation of the anaerobic parasite *Blastocystis* to the gut. *Curr. Biol.* 27, 807–820. doi: 10.1016/j.cub.2017.02.003
- Felsenstein, J. (2013). *PHYLIP Version 3.695*. Seattle, WA: University of Washington.
- Gachon, C. M., Strittmatter, M., Müller, D. G., Kleinteich, J., and Küpper, F. C. (2009). Detection of differential host susceptibility to the marine oomycete pathogen *Eurychasma dicksonii* by real-time PCR: not all algae are equal. *Appl. Environ. Microbiol.* 75, 322–328. doi: 10.1128/AEM.01885-08
- Gibbs, S. P. (1978). The chloroplasts of *Euglena* may have evolved from symbiotic green algae. *Can. J. Bot.* 56, 2883–2898. doi: 10.1139/b78-345
- Gilson, P. R., Su, V., Slamovits, C. H., Reith, M. E., Keeling, P. J., and McFadden, G. I. (2006). Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9566–9571. doi: 10.1073/pnas.0600707103
- Gould, S. B., Waller, R. F., and McFadden, G. I. (2008). Plastid evolution. *Annu. Rev. Plant Biol.* 59, 491–517. doi: 10.1146/annurev.arplant.59.032607.092915
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Huang, J., Mullapudi, N., Lancto, C. A., Scott, M., Abrahamsen, M. S., and Kissinger, J. C. (2004). Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. *Genome Biol.* 5, R88.
- Huang, J. L., and Yue, J. P. (2012). Horizontal gene transfer in the evolution of photosynthetic eukaryotes. *J. Syst. Evol.* 51, 13–29. doi: 10.1111/j.1759-6831.2012.00237.x
- Keane, T. M., Creevey, C. J., Pentony, M. M., Naughton, T. J., and McInerney, J. O. (2006). Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* 6:29. doi: 10.1186/1471-2148-6-29
- Keeling, P. J. (2013). The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annu. Rev. Plant Biol.* 64, 583–607. doi: 10.1146/annurev-arplant-050312-120144
- Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., et al. (2014). The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* 12:e1001889. doi: 10.1371/journal.pbio.1001889
- Keeling, P. J., and Palmer, J. D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* 9, 605–618. doi: 10.1038/nrg2386
- Kim, E., and Graham, L. E. (2008). EEF2 analysis challenges the monophyly of Archaeplastida and Chromalveolata. *PLoS ONE* 3:e2621. doi: 10.1371/journal.pone.0002621
- Kopriva, S., Büchert, T., Fritz, G., Suter, M., Benda, R., Schünemann, V., et al. (2002). The presence of an iron-sulfur cluster in adenosine 5'-phosphosulfate reductase separates organisms utilizing adenosine 5'-phosphosulfate and phosphoadenosine 5'-phosphosulfate for sulfate assimilation. *J. Biol. Chem.* 277, 21786–21791. doi: 10.1074/jbc.M202152200
- Kopriva, S., Fritzemeier, K., Wiedemann, G., and Reski, R. (2007). The putative moss 3'-phosphoadenosine-5'-phosphosulfate reductase is a novel form of adenosine-5'-phosphosulfate reductase without an iron-sulfur cluster. *J. Biol. Chem.* 282, 22930–22938. doi: 10.1074/jbc.M702522200
- Kopriva, S., and Koprivova, A. (2004). Plant adenosine 5'-phosphosulphate reductase: the past, the present, and the future. *J. Exp. Bot.* 55, 1775–1783. doi: 10.1093/jxb/erh185
- Küpper, F. C., and Müller, D. G. (1999). Massive occurrence of the heterokont and fungal parasites *Anisopodium*, *Eurychasma* and *Chytridium* in *Pyliella littoralis* (Ectocarpales, Phaeophyceae). *Nova Hedw.* 69, 381–389.
- Leander, B. S. (2004). Did trypanosomatid parasites have photosynthetic ancestors? *Trends Microbiol.* 12, 251–258.
- Loftus, B., Anderson, I., Davies, R., Alsmark, U. C. M., Samuelson, J., Amedeo, P., et al. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433, 865–868. doi: 10.1038/nature03291
- Ludwig, M., and Gibbs, S. P. (1989). Evidence that the nucleomorphs of Chlorarachnion reptans (Chlorarachniophyceae) are vestigial nuclei: morphology, division and DNA-DAPI fluorescence. *J. Phycol.* 25, 385–394. doi: 10.1111/j.1529-8817.1989.tb00135.x
- Maréchal, E., and Cesbron-Delauw, M.-F. (2001). The apicoplast: a new member of the plastid family. *Trends Plant Sci.* 6, 200–205. doi: 10.1016/S1360-1385(01)01921-5
- Margulis, L. (1970). *Origin of Eukaryotic Cells: Evidence and Research Implications for a Theory of the Origin and Evolution of Microbial, Plant, and Animal Cells on the Precambrian Earth*. New Haven, CT: Yale University Press.
- Marin, B., Nowack, E. C., and Melkonian, M. (2005). A plastid in the making: evidence for a second primary endosymbiosis. *Protist* 156, 425–432. doi: 10.1016/j.protis.2005.09.001
- Martin, W., and Kowallik, K. (1999). Annotated English translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche'. *Eur. J. Phycol.* 34, 287–295. doi: 10.1080/09670269910001736342
- McFadden, G. I. (2001). Primary and secondary endosymbiosis and the origin of plastids. *J. Phycol.* 37, 951–959. doi: 10.1046/j.1529-8817.2001.01126.x
- Moustafa, A., Beszteri, B., Maier, U. G., Bowler, C., Valentini, K., and Bhattacharya, D. (2009). Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* 324, 1724–1726. doi: 10.1126/science.1172983
- Okamoto, N., Chantangsi, C., Horák, A., Leander, B. S., and Keeling, P. J. (2009). Molecular phylogeny and description of the novel katablepharid *Roombia truncata* gen. et sp. nov., and establishment of the Hacrobia taxon nov. *PLoS ONE* 4:e7080. doi: 10.1371/journal.pone.0007080
- Palmer, J. D. (2003). The symbiotic birth and spread of plastids: how many times and whodunit? *J. Phycol.* 39, 4–12. doi: 10.1046/j.1529-8817.2003.02185.x
- Palmer, J. D., and Delwiche, C. F. (1998). "The origin and evolution of plastids and their genomes," in *Molecular Systematics of Plants II*, eds D. E. Soltis, P. S. Soltis, and J. J. Doyle (Boston, MA: Springer), 375–409.
- Patron, N. J., Durnford, D. G., and Kopriva, S. (2008). Sulfate assimilation in eukaryotes: fusions, relocations and lateral transfers. *BMC Evol. Biol.* 8:39. doi: 10.1186/1471-2148-8-39
- Patterson, D. (1989). "Stramenopiles: chromophytes from a protistan perspective," in *The Chromophyte Algae: Problems and Perspectives*, eds J. C. Green, B. S. C. Leadbeater, and W. I. Diver (Oxford: Clarendon Press), 357–379.
- Qiu, H., Cai, G., Luo, J., Bhattacharya, D., and Zhang, N. (2016). Extensive horizontal gene transfers between plant pathogenic fungi. *BMC Biol.* 14:41. doi: 10.1186/s12915-016-0264-3
- Reyes-Prieto, A., Moustafa, A., and Bhattacharya, D. (2008). Multiple genes of apparent algal origin suggest ciliates may once have been photosynthetic. *Curr. Biol.* 18, 956–962. doi: 10.1016/j.cub.2008.05.042

- Richards, T. A., Dacks, J. B., Jenkinson, J. M., Thornton, C. R., and Talbot, N. J. (2006). Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms. *Curr. Biol.* 16, 1857–1864. doi: 10.1016/j.cub.2006.07.052
- Richards, T. A., Soanes, D. M., Jones, M. D. M., Olga, V., Guy, L., Konrad, P., et al. (2011). Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15258–15263. doi: 10.1073/pnas.1105100108
- Richardson, A. O., and Palmer, J. D. (2007). Horizontal gene transfer in plants. *J. Exp. Bot.* 58, 1–9. doi: 10.1093/jxb/erl148
- Riisberg, I., Orr, R. J. S., Kluge, R., Shalchian-Tabrizi, K., Bowers, H. A., Patil, V., et al. (2009). Seven gene phylogeny of heterokonts. *Protist* 160, 191–204. doi: 10.1016/j.protis.2008.11.004
- Rogers, M. B., Gilson, P. R., Su, V., McFadden, G. I., and Keeling, P. J. (2007). The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol. Biol. Evol.* 24, 54–62. doi: 10.1093/molbev/msl129
- Sanchez-Puerta, M., and Delwiche, C. F. (2008). A hypothesis for plastid evolution in chromalveolates. *J. Phycol.* 44, 1097–1107. doi: 10.1111/j.1529-8817.2008.00559.x
- Sevčíková, T., Horák, A., Klimeš, V., Zbránková, V., Demirhilton, E., Sudek, S., et al. (2015). Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte? *Sci. Rep.* 5:10134. doi: 10.1038/srep10134
- Soucy, S. M., Huang, J., and Gogarten, J. P. (2015). Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* 16, 472–482. doi: 10.1038/nrg3962
- Stiller, J. W., Huang, J., Ding, Q., Tian, J., and Goodwillie, C. (2009). Are algal genes in nonphotosynthetic protists evidence of historical plastid endosymbioses? *BMC Genomics* 10:484. doi: 10.1186/1471-2164-10-484
- Stiller, J. W., Schreiber, J., Yue, J., Guo, H., Ding, Q., and Huang, J. (2014). The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nat. Commun.* 5, 5764. doi: 10.1038/ncomms6764
- Strittmatter, M., Gachon, C. M., and Küpper, F. C. (2009). *Ecology of Lower Oomycetes. Oomycete Genetics and Genomics: Diversity, Interactions and Research Tools*. Hoboken, NJ: John Wiley and Sons Inc, 25–46. doi: 10.1002/9780470475898.ch2
- Sun, G., Yang, Z., Ishwar, A., and Huang, J. (2010). Algal genes in the closest relatives of animals. *Mol. Biol. Evol.* 27, 2879–2889. doi: 10.1093/molbev/msq175
- Timmis, J. N., Ayliffe, M. A., Huang, C. Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* 5, 123–135. doi: 10.1038/nrg1271
- Tsaousis, A. D., Ollagnier, D. C. S., Gentekaki, E., Long, S., Gaston, D., Stechmann, A., et al. (2012). Evolution of Fe/S cluster biogenesis in the anaerobic parasite *Blastocystis*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 10426–10431. doi: 10.1073/pnas.1116067109
- Tyler, B. M., Sucheta, T., Xuemin, Z., Paramvir, D., Jiang, R. H. Y., Andrea, A., et al. (2006). Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261–1266. doi: 10.1126/science.1128796
- Van de Peer, Y., Rensing, S., Maier, U., and De Wachter, R. (1996). Substitution rate calibration of small subunit ribosomal RNA identifies chlorarachniophyte endosymbionts as remnants of green algae. *Proc. Natl. Acad. Sci. U.S.A.* 93, 7732–7736. doi: 10.1073/pnas.93.15.7732
- Weeden, N. F. (1981). Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J. Mol. Evol.* 17, 133–139. doi: 10.1007/BF01733906
- Woehle, C., Dagan, T., Martin, W. F., and Gould, S. B. (2011). Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related *Chromera velia*. *Genome Biol. Evol.* 3, 1220–1230. doi: 10.1093/gbe/evr100
- Yang, E. C., Boo, G. H., Kim, H. J., Cho, S. M., Boo, S. M., Andersen, R. A., et al. (2012). Supermatrix data highlight the phylogenetic relationships of photosynthetic stramenopiles. *Protist* 163, 217–231. doi: 10.1016/j.protis.2011.08.001
- Yoon, H. S., Grant, J., Tekle, Y. I., Wu, M., Chaon, B. C., Cole, J. C., et al. (2008). Broadly sampled multigene trees of eukaryotes. *BMC Evol. Biol.* 8:14. doi: 10.1186/1471-2148-8-14
- Yoon, H. S., Hackett, J. D., Pinto, G., and Bhattacharya, D. (2002). From the cover: the single, ancient origin of chromist plastids. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15507–15512. doi: 10.1073/pnas.242379899
- Yue, J., Sun, G., Hu, X., and Huang, J. (2013). The scale and evolutionary significance of horizontal gene transfer in the choanoflagellate *Monosiga brevicollis*. *BMC Genomics* 14:1. doi: 10.1186/1471-2164-14-729

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Wang, Sun and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



De novo Transcriptome Assembly of *Phomopsis liquidambari* Provides Insights into Genes Associated with Different Lifestyles in Rice (*Oryza sativa* L.)

Jun Zhou, Xin Li, Yan Chen and Chuan-Chao Dai *

Jiangsu Key Laboratory for Microbes and Functional Genomics, Jiangsu Engineering and Technology Research Center for Industrialization of Microbial Resources, College of Life Sciences, Nanjing Normal University, Nanjing, China

OPEN ACCESS

Edited by:

Tatiana Matveeva,
Saint Petersburg State University,
Russia

Reviewed by:

Zonghua Wang,
Fujian Agriculture and Forestry
University, China
Jihong Jiang,
Jiangsu Normal University, China

*Correspondence:

Chuan-Chao Dai
daichuanchao@njnu.edu.cn

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 25 November 2016

Accepted: 20 January 2017

Published: 06 February 2017

Citation:

Zhou J, Li X, Chen Y and Dai C-C
(2017) De novo Transcriptome
Assembly of *Phomopsis liquidambari*
Provides Insights into Genes
Associated with Different Lifestyles in
Rice (*Oryza sativa* L.).
Front. Plant Sci. 8:121.
doi: 10.3389/fpls.2017.00121

The mechanisms that trigger the switch from endophytic fungi to saprophytic fungi are largely unexplored. Broad host range *Phomopsis liquidambari* is established in endophytic and saprophytic systems with rice (*Oryza sativa* L.). Endophytic *P. liquidambari* promotes rice growth, increasing rice yield and improving the efficiency of nitrogen fertilizer. This species's saprophytic counterpart can decompose rice litterfall, promoting litter organic matter cycling and the release of nutrients and improving the soil microbial environment. Fluorescence microscopy, confocal laser scanning microscopy and quantitative PCR investigated the colonization dynamics and biomass of *P. liquidambari* in rice *in vivo*. *P. liquidambari* formed infection structures similar to phytopathogens with infected vascular tissues that systematically spread to aerial parts. However, different from pathogenic infection, *P. liquidambari* colonization exhibits space restriction and quantity restriction. Direct comparison of a fungal transcriptome under three different habitats provided a better understanding of lifestyle conversion during plant-fungi interactions. The isolated total RNA of Ck (pure culture), EP (endophytic culture) and FP (saprophytic culture) was subjected to Illumina transcriptome sequencing. To the best of our knowledge, this study is the first to investigate *Phomopsis* sp. using RNA-seq technology to obtain whole transcriptome information. A total of 27,401,258 raw reads were generated and 22,700 unigenes were annotated. Functional annotation indicated that carbohydrate metabolism and biosynthesis of secondary metabolites played important roles. There were 2522 differentially expressed genes (DEGs) between the saprophytic and endophytic lifestyles. Quantitative PCR analysis validated the DEGs of RNA-seq. Analysis of DEGs between saprophytic and endophytic lifestyles revealed that most genes from amino acids metabolism, carbohydrate metabolism, fatty acid biosynthesis, secondary metabolism and terpenoid and steroid biosynthesis were up-regulated in EP. Secondary metabolites of these pathways may affect fungal growth and development and contribute to signaling communication with the host. Most pathways of xenobiotic biodegradation and metabolism were upregulated in FP. Cytochrome P450s

play diverse vital roles in endophytism and saprophytism, as their highly specialized functions are evolutionarily adapted to various ecological niches. These results help to characterize the relationship between fungi and plants, the diversity of fungi for ecological adaptations and the application prospects for fungi in sustainable agriculture.

Keywords: *Phomopsis liquidambari*, rice, endophyte, saprophyte, transcriptome, ecological adaptation

INTRODUCTION

In farmland ecosystems, intensive agriculture weakens the self-cycle capacity of soil nutrients, forcing farmers to devote effort to collecting and burning plant litter and applying fertilizer to farmland. This leads to the waste of natural organic resources, atmospheric and environmental pollution and soil quality deterioration, which are unfavorable for sustainable agriculture. Therefore, it is necessary to find an appropriate medium to decompose plant residues in farmland in order to increase soil available nutrients.

Endophytic fungi and saprophytic fungi usually play important ecological functions in living plant tissue and dead plant material. Many studies have investigated the relationship between endophytes and saprophytes and have hypothesized that endophytes become saprophytic after senescence of host tissues (Hyde et al., 2007). This may be due to modification of host tissues during senescence, allowing fungal hyphae to penetrate the epidermis and colonize the surface of the host. Promputtha et al. (2007) isolated common endophytes from *Magnolia liliifera*, *Phomopsis*, *Guignardia*, *Fusarium*, and *Colletotrichum* that have a high degree of sequence similarity and are phylogenetically relevant to the corresponding saprophyte. These results suggest that some endophytes might alter their ecological strategies and adopt a saprophytic lifestyle. Promputtha et al. (2010) also reported that nine endophytes, *Phomopsis* sp. 2, *Phomopsis* sp. 6, *Phomopsis* sp. 10, *Guignardia mangiferae*, *Corynespora cassiicola*, *Leptosphaeria* sp., *Fusarium* sp. 1, *Colletotrichum gloeosporioides* and *Colletotrichum* sp. 2 were morphologically similar and phylogenetically related to saprophytes. These endophytes and their saprobic counterparts produce the same degrading enzymes and a similar isoform of β -mannanase. Fungal succession is relevant to enzyme production patterns during leaf decomposition, and the occurrence of saprophytes is related to enzyme production from endophytes. This provides further convincing evidence that endophytes can change their lifestyle to become saprophytes.

Lipids are an important component of all living cells that offer a structural basis for cell membranes and fuel for metabolism and have a role in cell signaling. Membrane lipid synthesis is a prerequisite of symbiosis, and the performance of the

membrane depends on lipid composition (Wewer et al., 2014). Fatty acids and modified fatty acids are important molecules for pathogen colonizing plants whose functions include signaling, energy sources and virulence factors (Uranga et al., 2016). The oxylipins are a vast diversified family of secondary metabolites derived from oxidation of unsaturated fatty acids or further conversion (Tsitsigiannis and Keller, 2007). In fungi, precursors of oxylipins are usually linoleic acid, oleic acid and α -linolenic acid (Pohl and Kock, 2014). Fungal oxylipins can be used as secondary metabolites that participate in infection processes, biotrophy and necrotrophy (Oliw et al., 2016). Fungi produce a series of secondary metabolites and small molecules that may not be directly required for growth, but play important roles in signal transduction, development and organism interaction. The cytochrome P450 enzyme system is thought to play various functions in biosynthesis of secondary metabolites and participated in biodegradation of lignin and various xenobiotics (Martinez et al., 2009).

Though most endophytes depend on readily available compounds such as soluble sugar to grow, xylariaceous endophytes can degrade cellulose and lignin (Promputtha et al., 2010). Hence, endophytes that produce enzymes to decompose lignin and cellulose could decompose host tissue and persist as saprophytes following host senescence. Our research shows that the endophyte *Phomopsis liquidambari* B3 can establish a symbiotic relationship with rice (*Oryza sativa* L.), systematically colonizing roots and aerial parts (Yang et al., 2014b), which promotes the growth of rice, increasing yield and significantly reducing application of nitrogen fertilizer (Yang et al., 2014a, 2015; Siddikee et al., 2016). In the saprophytic phase, the fungus can decompose rice straw, promote litter organic matter cycling and the release of nutrients, improve soil microbial environments (Chen et al., 2013a) and secrete laccase, cellulase and polyphenol oxidase. Dai et al. (2010a) investigated the capability to form cavities on the straw surface and the condition for laccase production in *P. liquidambari*, suggested that endophytes can form a series of cavities on straw to decompose plant materials by producing laccase. In addition, *P. liquidambari* also secretes degradative enzymes for phenanthrene (Dai et al., 2010b), indole (Chen et al., 2013c; Wang et al., 2014), ferulic acid (Xie and Dai, 2015), 4-hydroxybenzoic acid (Chen et al., 2011) and phytoestrogen luteolin (Wang et al., 2015) among others. The fungus utilizes these compounds as the sole carbon source for growth. Chen et al. (2013c) shows that the degradation rate of indole in endophyte B3 over 120 h in a pure culture condition was 41.7%. The exogenous addition of plant litter significantly increased the ratio of indole degradation within 60 h to 99.1%, indicating the utility of litter-induced fungi to produce laccase and lignin

Abbreviations: *P. liquidambari*, *Phomopsis liquidambari*; GFP, green fluorescent protein; Ck, pure culture/control treatment; EP, endophytic culture/ fungal mycelium of the *P. liquidambari* callus culture; FP, saprophytic culture/ litterfall-cultivated fungal mycelium of *P. liquidambari*; DEGs, differentially expressed genes; DGE, digital gene expression profiling; dai, days after inoculation; NR, non-redundant databases; KEGG, Kyoto Encyclopedia of Genes and Genomes; COG, Clusters of Orthologous Groups of proteins; GO, Gene Ontology; FDR, false discovery rate; CDS, Protein coding region.

peroxidase to non-specifically decompose nitrogen heterocyclic compounds.

However, comparisons of different types of plant-fungal interactions in the same plant species are limited because saprophytic systems and mutualistic systems are separated in various plants. Therefore, it will be valuable to perform experiments studying saprophytism and mutualism in a single plant species to directly compare endophytic and saprophytic plant-fungi interactions. Few studies have directly compared two different plant-microbe interactions in a single plant species. Furthermore, because only a small part of plant cells are colonized and it is difficult to accurately detect expression levels of fungal genes in colonized tissue, the detection of gene expression profiles and elucidation of interactional mechanisms during the endophytic lifestyle transition at different stages in associated host tissues remain poorly understood. Recently, “omics” approaches have been used to better understand endophyte-plant interactions (Kaul et al., 2016).

Rice is a representative gramineous plant, the staple food for approximately half the global population and a model material in agricultural microbiology research. *P. liquidambari* is a broad-spectrum endophyte that is typically used to study the switch from endophytism to saprophytism. We have established experimental systems to study the endophytic and saprophytic interaction of *P. liquidambari* B3 with a rice single host. Colonization dynamics and distribution in rice *in vivo* were monitored by green fluorescent protein (GFP)-tagged *P. liquidambari*. To discuss the differences in gene expression in *P. liquidambari* B3 interactions with rice under the two conditions of endophytism and saprophytism transcriptome sequencing technology and digital gene expression profiles were used. Endophytes are known to establish a symbiosis with the host through a series of regulatory mechanisms, as there is a sizable difference in performance between mutualistic fungi in the host and the common saprophyte. However, there is insufficient research into the switch from endophyte to saprophyte function in senescent plant litter. Clarification of this scientific problem has great significance for understanding the relationship between endophytes and plants and for documenting the diversity of endophytes.

MATERIALS AND METHODS

Fungal Strain and Transformation

P. liquidambari B3 was isolated from the inner bark of *Bischofia polycarpa*. It was stored on a potato dextrose agar slant (200 g potato extract, 20 g glucose, 20 g agar per liter, pH 7.0) at 4°C. The fungus was activated in potato dextrose broth (200 g potato extract, 20 g glucose per liter, pH 7.0) and cultured for 48 h at 28°C with a rotation speed of 180 rpm.

The transformation vector plasmid for filamentous fungi pCT74 expresses sGFP under the control of the *ToxA* promoter. It contains the hygromycin B resistance gene *hph*, which encodes an aminoglycosidic antibiotic and is derived from *Streptomyces hygroscopicus*. This gene has been used for selection and maintenance of transformed prokaryotic and

eukaryotic cells. Protoplast preparation and transformation were performed as described by Yang et al. (2014b) with some modifications.

Inoculation, Co-culture and Microscopy

The rice cultivar used in this study was “Wuyunjing 21”. Rice seeds were dehusked and surface-sterilized in 75% ethanol for 15 min, bleached in a 6% sodium hypochlorite solution (6% available chlorine) for 10 min, rinsed repeatedly in sterile distilled water, and planted in 1/2 Murashige & Skoog (MS) 0.7% agarose medium supplemented with 30 mM sucrose for 4 days. The seedlings were kept vertical at 25°C under a 16 h of light at 22°C and 8 h of dark. Seedlings of roughly the same size were transferred to 1/2 MS in a square petri dish (13 cm in width, 13 cm in length). Each plate of five plants inoculated with 7-mm GFP-B3 mycelial disks were placed near the plant roots on the medium. Potato dextrose agar disks of equal size were used as a control. All treatments were replicated five times. Rice shoots and roots were sampled and processed for microscopy.

An Axio Imager A1 fluorescence microscope (Zeiss, Jena, Germany) was used for observing the fungal structures as described previously by Yang et al. (2014b). Confocal laser scanning microscopy was performed using a Ti-E microscope with an A1 confocal system (Nikon, Tokyo, Japan) to monitor the infection process. GFP and FITC images of rice shoots and roots were captured simultaneously using 488 nm excitation with an argon laser and fluorescence detection at 543.5 nm. Images were processed using Adobe Photoshop CS6 (Adobe, San Jose, CA, USA).

Quantification of *P. liquidambari* Biomass in Rice Roots and Shoots by Quantitative PCR

P. liquidambari-infected roots and shoots were harvested at 0, 3, 7, 14, 21, and 28 dai (days after inoculation). The biomass of *P. liquidambari* in the infected plant tissue was quantified using quantitative PCR (qPCR) according to Yang et al. (2014b). DNA was extracted after grinding tissue powder with a Multisource Genomic DNA Miniprep Kit (Axygen). A primer set suitable for qPCR was designed based on a *P. liquidambari*-specific ITS locus (Bf1 and Br1) (Table S1). PCR amplified products were cloned into the pMD® 19-T vector (Takara, Otsu, Japan) and expressed in competent DH5α cells. Positive clones were screened and plasmids were extracted using a SanPrep Column Plasmid Mini-Prep Kit. A dilution range of the plasmids from 1.3×10^2 to 1.3×10^8 copies was used to make a standard curve. In the qPCR reaction system (20 μL), gDNA was mixed with SYBR® Green Master Mix (Vazyme, Nanjing, China), primers, and ddH₂O. The PCR procedure was as follows: 94°C (1 min) for one cycle; 94°C (15 s), 60°C (45 s), and 72°C (30 s) for 45 cycles; melting curve analysis from 72 to 60°C in 0.5°C decrements. The amplification of a single PCR product was validated using 1.5% gel electrophoresis.

Endophytic and Saprophytic Systems of *P. liquidambari* Interact with Rice

Activated *P. liquidambari* was filtered with 8-layer gauze. Fungal mycelia were cleaned three times with sterile deionized water and 0.10 g mycelia (0.01 g dry biomass) were weighed and used for callus inoculation. In addition, 0.30 g mycelia (0.03 g dry biomass) were weighed and inoculated in 20 mL sterile water as a fungus seed solution for litterfall inoculation; 0.30 g mycelia were weighed and inoculated in 1 × NB solid medium for 3 days and regarded as a control treatment (Ck).

The endophytic lifestyle was studied using tissues cultures of host rice in dual culture *in vitro* based on previous research (Sieber et al., 1990; Peters and Schulz, 1998; Nawrot-Chorabik et al., 2016). Inoculated rice seeds at the surface were sterilized in callus solid medium (1 × NB solid medium, 2 mg L⁻¹ 2,4-D, 30 g L⁻¹ sucrose, 10 g L⁻¹ agar, pH 7) and cultivated for 20 days at 28°C. The callus was stripped from rice seeds at germination and a yellow soft callus appeared at the base of medium. This was placed on new callus solid medium and cultivated for 50 days under 28°C. When a callus formed (Figure S1), *P. liquidambari* mycelium was inoculated on top of the callus with tweezers and cultivated for 3 days at 28°C. At this stage, the fungus was checked to ensure that it grew on the surface of the callus but did not contact the medium. The mycelium was stripped from the surface of the callus with delicate tweezers, and this sample was used as the fungal mycelium of the *P. liquidambari* callus culture (EP).

The saprophytic system was derived using the culture fungus method of Chen et al. (2013a) with litter. Collected rice litterfall completely withered from the ground of the rice experimental plot. The moisture content of rice litterfall was 7.6%. The litterfall surface was washed with sterile deionized water and cut into 1 cm × 1 mm segments. Weighed 0.5-g litterfall samples were added to a 250-mL triangular flask with 100 mL 1 × NB liquid medium (pH 5.5) and sterilized for 20 min at 121°C. A 2-mL sample of *P. liquidambari* was inoculated into the seed solution and cultivated for 3 days at 28°C and 160 rpm. The mycelium pellet was removed from the liquid medium with tweezers and washed clean. This sample was used as a litterfall-cultivated fungal mycelium of *P. liquidambari* (FP).

RNA Extraction

Total RNA was isolated from Ck, EP and FP using a Fungal RNA extraction kit (E.Z.N.A. Total RNA Kit I, OMEGA, USA) and treated with DNase I. The quality and concentration of extracted RNA were examined using agarose gel electrophoresis and a spectrophotometer (OneDropTM OD-2000+, China), and eligible groups were used for Illumina sequencing.

cDNA Library Construction and Sequencing

The mRNA of the total RNA was purified using magnetic Oligo (dT) beads. This mRNA was mixed with the fragmentation buffer, the mRNA was fragmented into short fragments. The cDNA was synthesized using mRNA fragments as templates. Short fragments were purified and resolved with EB buffer

for sticky-end preparation and single nucleotide A addition. Subsequently, the short fragments were connected with adapters, and suitable fragments were selected as templates for PCR amplification. Quantification and qualification of the sample library was performed using an ABI StepOnePlus Real-Time PCR System and an Agilent 2100 Bioanalyzer. The library was sequenced using Illumina HiSeqTM 2000.

Sequence Annotation

Image data output from Illumina sequencing was transformed by base calling into raw reads. Clean reads were obtained by removing dirty reads that contained adapters or unknown or low quality bases. Transcriptome *de novo* assembly was carried out with Trinity (Grabherr et al., 2011) and a k-mer library was constructed. The highest frequency k-mer was selected to assemble contigs and then mapped with clean reads. Paired-end reads were used to fill gaps in the scaffolds to assemble contigs to unigenes. Non-redundant unigenes were acquired by further processing of sequence splicing and removal of redundancy. All-unigenes were generated after gene family clustering. Unigene sequences were aligned with blastx ($e < 0.00001$) to protein databases including non-redundant databases (NR), Swiss-Prot, the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Clusters of Orthologous Groups of proteins (COG), and aligned by blastn ($e < 0.00001$) to the nucleotide databases nt. Gene Ontology (GO) functional annotation was achieved using NR annotation by Blast2GO (Conesa et al., 2005). GO functional classification was achieved using WEGO software (Ye et al., 2006).

Identification of Differentially Expressed Genes

The FPKM method was used to calculating unigene expression (Mortazavi et al., 2008). An algorithm to identify differentially expressed genes (DEGs) between the two samples was used according to the method of Audic and Claverie (1997). In our analysis, the genes with false discovery rates (FDR) ≤ 0.001 and ratios larger than 2 were regarded as significant DEGs. We mapped all DEGs to terms in the GO database and KEGG database for enrichment analysis.

Quantitative Real-Time PCR Analysis

To validate the DEGs obtained by Solexa RNA-seq, 20 genes (Table S1) were subjected to quantitative real-time PCR analysis using an ABI PRISM 7500 Real-time PCR System (Applied Biosystems). *P. liquidambari* β-actin (Table S1) was used as the endogenous control. cDNA synthesis was carried out using the same RNA samples as those used for digital gene expression profiling (DGE) experiments. The corresponding primers were designed using Primer Premier 6.0 and listed in Table S1. The reaction mixture (20 μl) contained 10 μl of SYBR[®] Green Master Mix (Vazyme, Nanjing, China), 0.4 μM of forward and reverse primers, and 0.2 μl of cDNA template. The amplification programs were performed with the following parameters: 95°C for 30 s; 95°C for 5 s, 60°C for 40 s (40 cycles) and followed by melting curve analysis from 60 to 95°C in 0.5°C increments. Each reaction was run in triplicate, including a negative control.

The relative expression levels of genes were calculated using the $2^{-\Delta\Delta CT}$ method.

RESULTS

Infection and Colonization Process of Rice Plants Systemically by *P. liquidambari*

To visualize the infection process of *P. liquidambari* in planta, we first obtained transgenic fungal strains by constitutively expressing cytoplasmic GFP (GFP-B3). In the early stage of infection (1–3 dai), *P. liquidambari* hyphae were only distributed on the root surface, especially the root-hair zone, infected root-hairs and infected epidermis (Figures 1A–C). A large number of runner hypha interwoven together to form a hyphal network (Figures 1D–F and Video S1). Runner hyphae were distributed in the low-lying area between the cells of root epidermal layers, growing along the longitudinal axis direction of the root and invading cells using a penetration peg (Figure 1G). In the middle stage of infection (4–15 dai), hyphae beginning intracellular and intercellular infection spread from epidermal layers to the cortex and finally to the endodermis (Figures 1H–J). In the epidermis and cortex, hyphae can undergo intracellular and intercellular growth along the direction parallel to the root spindle spread from one cell to another, branching in the intercellular space, and then continue to grow (Figures 1K–M). When it penetrated the cell wall, hyphae appeared with neck-like constrictions (Figure 1N). Parts of strong hypha entered the center of the root spindle and then penetrated the vascular bundle into the aerial part (Figures 1O–R). In late-stage colonization (>15 dai), which is associated with programmed cell death, the vast majority of epidermal cells and parts of outer cortex cells were crowded with mycelium and sclerotium (Figures 1S,T). Colonization of *P. liquidambari* was still observed in the senescence root (>50 dai) (Figures 1U–W).

Quantification of *P. liquidambari* Biomass in Rice Tissues

The concentration of *P. liquidambari* within plantlets is expressed as the number of *P. liquidambari*-specific ITS copies per ng total (plantlet + fungal) genomic DNA in the qPCR analysis. The concentration of endophytes in roots was always higher than in shoots from 0 to 28 dai. In roots, a significant increase from 0 to 7 dai was followed by a moderate decrease. A moderate increase in shoots occurred from 0 to 21 dai and then reached a steady state (Figure 2).

Illumina RNA-Sequencing and Read Assembly

To identify DEGs related to lifestyle, we mixed total RNA extracted from Ck, FP and EP of *P. liquidambari* hyphae equally for transcriptome sequencing. A total of 27,401,258 raw reads were generated. After filtering, 26,109,074 clean reads were obtained, for a total of 2,349,816,660 bp clean nucleotides. The Q20 percentage was 96.47% and the GC percentage was 56.45%. After editing and quality checking, 26 million 90 bp clean reads were assembled into 51,120 contigs with a mean

length of 487 bp. The N50 of contigs was 1240 bp, where larger numbers are better for the quality of assembly. Using paired-end joining and gap filling, the contigs were further assembled into 32,424 unigenes with a mean length of 945 bp, including 7946 distinct clusters and 24,478 distinct singletons. The N50 of unigenes was 1574 bp, indicating that the assembly results were desirable (Table 1). The assembled sequence length is one evaluative criteria of assembly quality. The size distribution of the contigs and unigenes are shown in Figure 3. By comparing the length distribution proportion of contigs and unigenes, we found that contigs from 100 to 200 bp accounted for 50.62%, greater than 500 bp that accounted for 22.65% (Table S2). However, the length of unigenes obtained from further assembly, of which of 100–500 bp accounted for just 46.48%, were all greater than 200 bp and the proportion greater than 1000 bp was over 31.51% (Table S3), indicating that the assembly quality of unigenes that assembled from contigs was high.

Functional Annotation of Predicted Proteins

We matched unigene sequences against NR, NT, Swiss-Prot, KEGG, GO and COG databases using blastx ($E < 10^{-5}$). Of these, 22,700 unigenes (70.0% of total) were annotated, and most could be annotated to protein functional information in the NR database. A total of 22,382 unigenes were annotated to NR database (Table S4). We have analyzed the *E*-value distribution (Figure 4A), similarity distribution (Figure 4B) and species distribution (Figure 4C) of the NR annotation. The *E*-value distribution showed that 64.8% of the mapped sequences displayed a high level of homology ($E < 10^{-30}$); 51.6% of the mapped sequences displayed a higher level of homology ($E < 10^{-45}$). The similarity distribution has a comparable pattern, with 50.5% of sequences having similarity higher than 60 and 12.5% of the sequences having similarity higher than 80%. For species distribution, 8.2% of the distinct sequences had top matches with sequences from *Colletotrichum higginsianum*, followed by the *Glomerella graminicola* M1.001 (8.0%), *Gaeumannomyces graminis* var. *tritici* R3-111a-1 (7.8%), *Thielavia terrestris* NRRL 8126 (7.7%), *Magnaporthe oryzae* 70-15 (6.3%), *Mycelopithora thermophila* ATCC 42464 (5.7%), *Nectria haematococca* mpVI 77-13-4 (4.4%) and 48.1% of the unigene sequences matched to the seven species.

GO and COG Classification

GO functional annotation was obtained according to NR annotation information. GO assignments were used to classify the functions of the predicted *P. liquidambari* unigenes. A total of 10,209 unigenes were assigned to 50 functional groups in each of the three main categories according to sequence homology (Figure 5 and Table S5). In the cellular component, the majority was “cell,” “cell part,” “membrane,” “organelle” and “membrane part” unigenes associated with cell membranes and organelles. In the section for molecular function, the dominant functions were “catalytic activity,” “binding,” “transporter activity” and “structural molecule activity.” In the section for biological processes, unigenes were mainly involved in metabolic processes

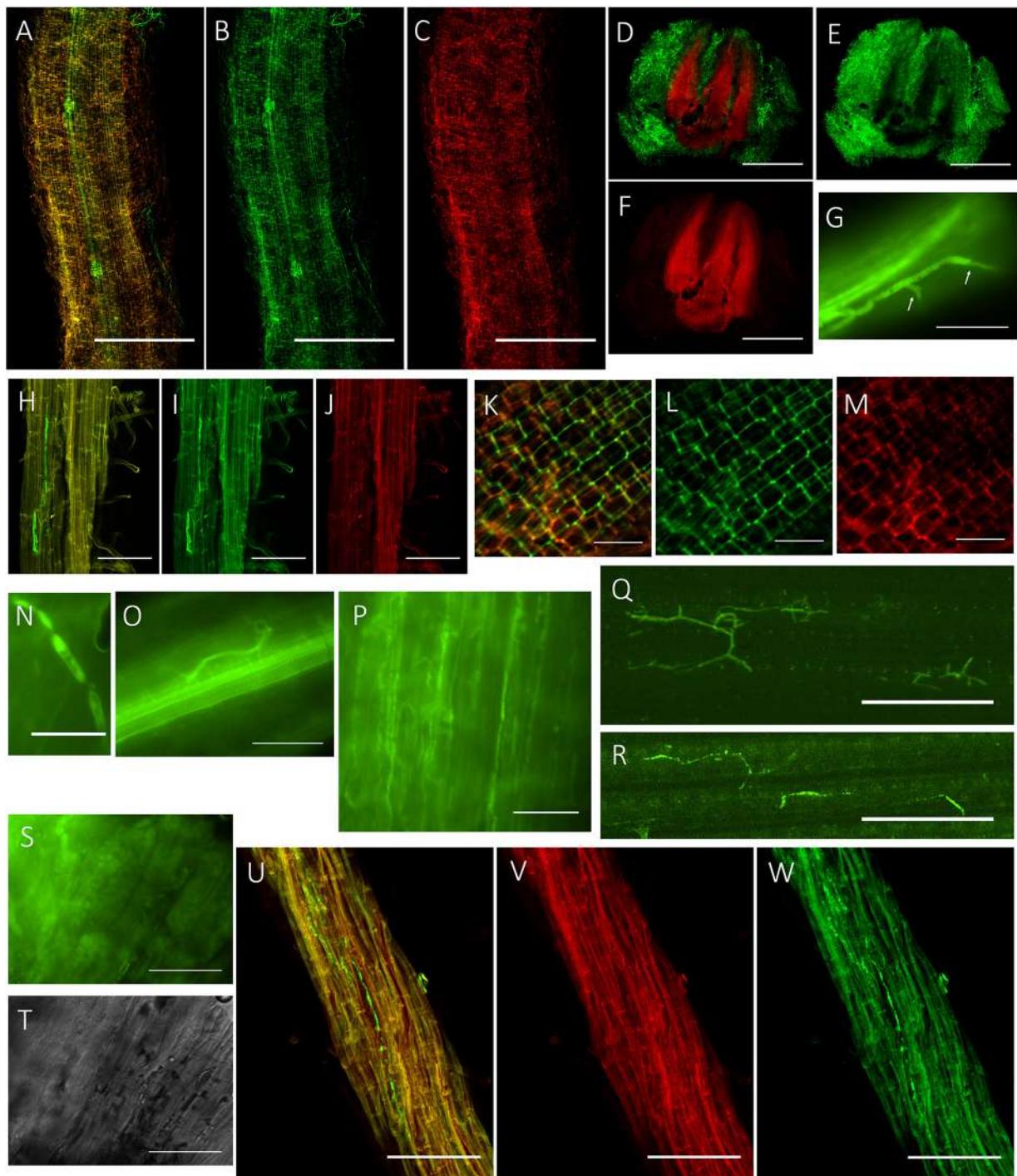


FIGURE 1 | Colonization pattern of *P. liquidambari* in rice roots and shoots. (A–C) Distribution of GFP- *P. liquidambari* hyphae on the root surface and infection of root-hair; 1–3 dai. **(A)** Overlay channel. **(B)** Green fluorescence channel. **(C)** Red fluorescence channel. Bar, 100 μm . **(D–F)** Cross-section of root tip, runner hypha interweaved to form hyphal network. **(D)** Overlay channel. **(E)** Green fluorescence channel. **(F)** Red fluorescence channel. Bars, 100 μm . **(G)** Penetration peg (arrow). Bar, 25 μm . **(H–J)** Intracellular and intercellular growth of hypha along the direction parallel to the root spindle spread from one cell to another; 4–15 dai. Bar, 25 μm . **(H)** Overlay channel. **(I)** Green fluorescence channel. **(J)** Red fluorescence channel. **(K–M)** Hyphae branching in the intercellular space. **(K)** Overlay channel. **(L)** Green fluorescence channel. **(M)** Red fluorescence channel. Bar, 50 μm . **(N)** Neck-like constriction, Bar, 10 μm . **(O)** Penetration of the center of the root spindle. Bar, 25 μm . **(P–R)** Colonization of hyphae in aerial parts. **(P)** Colonization of hyphae in stem cells. **(Q,R)** Colonization of hyphae in leaf cells. Bar, 50 μm . **(S,T)** Fluorescence microscopy of root cells were crowded with mycelium and sclerotium (>15 dai). **(S)** DIC channel. **(T)** bright field channel. Bar, 25 μm . **(U,W)** Colonization of hyphae in senescence rice root (>50 dai). **(U)** Overlay channel. **(V)** Red fluorescence channel. **(W)** Green fluorescence channel. Bar, 50 μm .

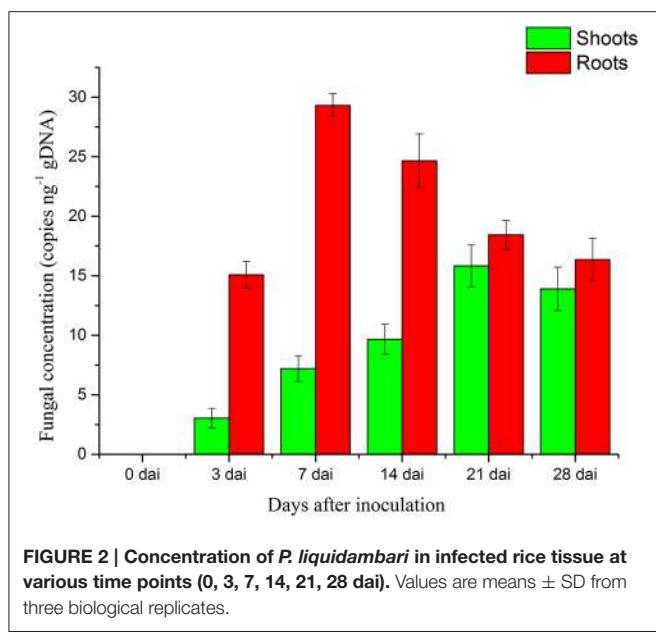


FIGURE 2 | Concentration of *P. liquidambari* in infected rice tissue at various time points (0, 3, 7, 14, 21, 28 dai). Values are means \pm SD from three biological replicates.

TABLE 1 | Statistics of *P. liquidambari* transcriptome.

Total raw reads	27,401,258
Total clean reads	26,109,074
Total clean nucleotides (nt)	2,349,816,660
Q20 percentage	96.47%
N percentage	0.01%
GC percentage	56.45%
Total number of contigs	51,120
Total length of contigs (nt)	24,882,464
Mean length of contigs (nt)	487
N50 of contigs	1240
Total number of unigenes	32,424
Total length of unigenes (nt)	30,646,255
Mean length of unigenes (nt)	945
N50 of unigenes	1574
Distinct clusters	7,946
Distinct singlettons	24,478

and cellular processes, in agreement with the results of cellular components and molecular function.

In total, 10,327 unigenes have a COG classification based on sequence homology. Among the 25 COG categories (Figure 6), “general function prediction only” (3698) was the largest group, followed by “carbohydrate transport and metabolism” (1998), “transcription” (1982), and “translation, ribosomal structure and biogenesis” (1887). The groups for “nuclear structure” (5), “extracellular structure” (25) and “RNA processing and modification” (74) were smallest.

KEGG Analysis

A total of 22,700 annotated unigenes of *P. liquidambari* were blasted to the KEGG database and annotated further. In all, 14,791 sequences were found to be involved in 108 signal

pathways. The number of sequences ranged from 4 to 4785 (Table S6). The first 20 pathways with the greatest number of sequences are indicated in Table 2, and the pathways that were most represented were metabolic pathways (4785) and biosynthesis of secondary metabolites (2154). These annotations provided important clues for further studying the specific development, function and pathways of *P. liquidambari*. The top 10 metabolic pathways were as follows: starch and sucrose metabolism (1481), amino sugar and nucleotide sugar metabolism (814), purine metabolism (630), pyrimidine metabolism (408), lysine degradation (365), tyrosine metabolism (329), glycolysis/gluconeogenesis (276), fructose and mannose metabolism (249), butanoate metabolism (236) and tryptophan metabolism (226). We believe that genes in carbohydrate metabolism and biosynthesis of secondary metabolites play significant roles in *P. liquidambari* endophytism and saprophytism.

Protein Coding Region (CDS) Prediction

In total, 22,300 and 1551 unigenes were predicted by BLASTx and ESTScan, respectively. The histogram seen in Figure S2 shows the length distribution of CDS predicted by BLAST and ESTScan. In general, as sequence length increased, the number of CDS was gradually reduced. This is consistent with unigene assembly results.

Analysis of Differentially Expressed Genes

To detect the DEGs between EP and FP, we screened differentially expressed tags between samples according to the method described by Audic and Claverie (1997). As shown in Figure 7, there were 2869 genes that were differentially expressed between Ck and FP. Among these genes, 1502 were up-regulated and 1367 were down-regulated in response to the FP switch. There were 2277 genes differentially expressed between Ck and EP. Among these genes, 1382 genes were up-regulated and 895 were down-regulated in response to the EP switch. There were 2522 genes differentially expressed between FP and EP. Among these genes, 1415 genes were up-regulated and 1107 were down-regulated in response to the switch between EP and FP. There were 491 genes co-expressed among the three expression patterns (Figure 7). DEGs were further categorized into different functional groups by GO and KEGG pathway enrichment analysis. Compared with Ck, “starch and sucrose metabolism” and “amino sugar and nucleotide sugar metabolism” were the most enriched pathways in FP, and “butanoate metabolism” was the most enriched pathway in EP. Compared with FP, the “ribosome” group was the most enriched pathway in EP ($P < 0.05$) (Table S7).

Validation of RNA-seq Data by qPCR

To validate the DEGs obtained by Solexa RNA-seq, we further performed quantitative real-time PCR analysis on 20 representative genes involved in the three lifestyles (Figure 8). We found that fold-change values of most DEGs using real-time qRT-PCR exhibited trends similar to RNA-Seq samples. Differential expression was observed for all candidate genes, suggesting that they are involved in regulatory networks that are active during the three environmental conditions. Only

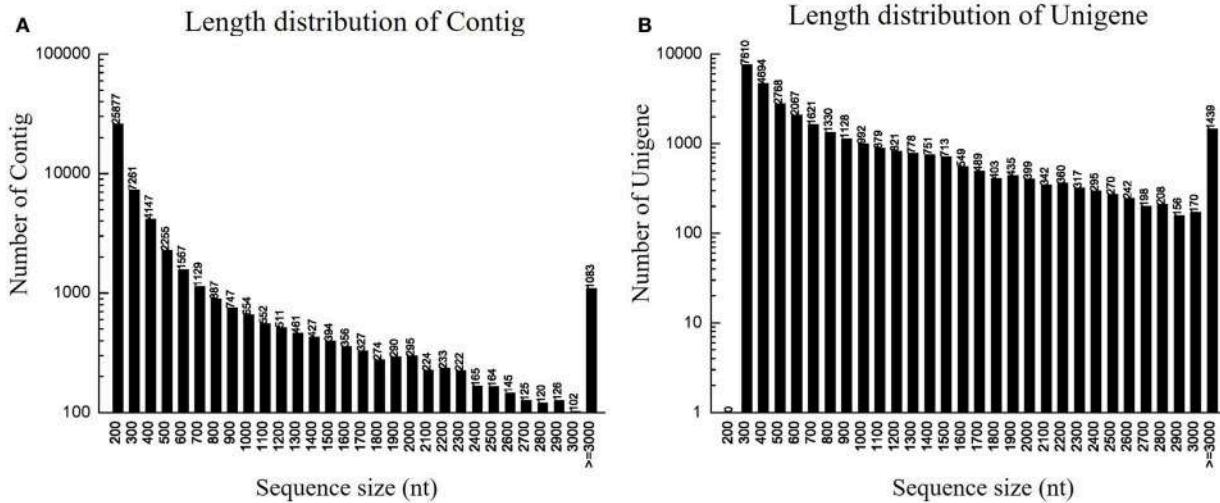


FIGURE 3 | Length distribution of contigs and unigenes. (A) Length distribution of contigs. Horizontal coordinate is contig length and vertical coordinate is contig number. **(B)** Length distribution of unigenes. Horizontal coordinate is unigene length and vertical coordinate is number of unigenes.

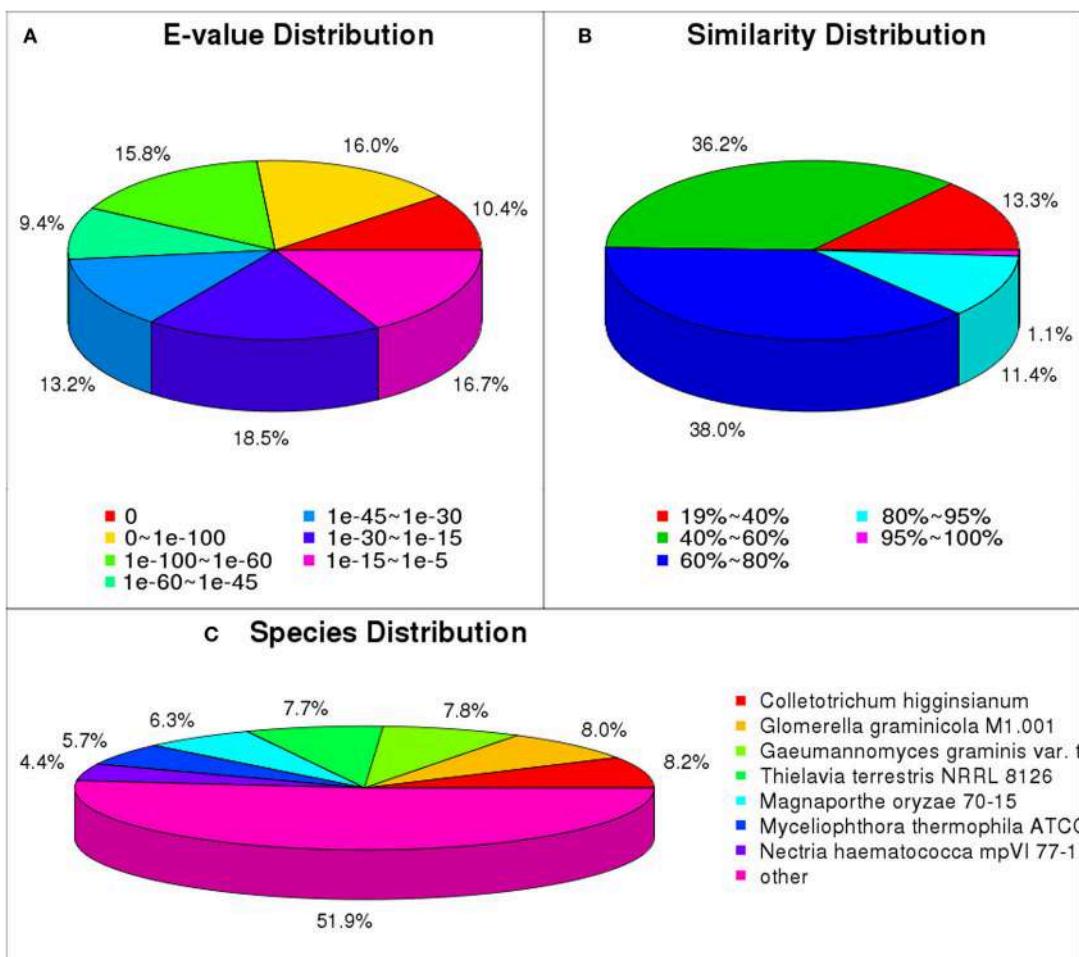


FIGURE 4 | NR classification of *P. liquidambari* unigenes. (A) E-value distribution. **(B)** Identity distribution. **(C)** Species distribution.

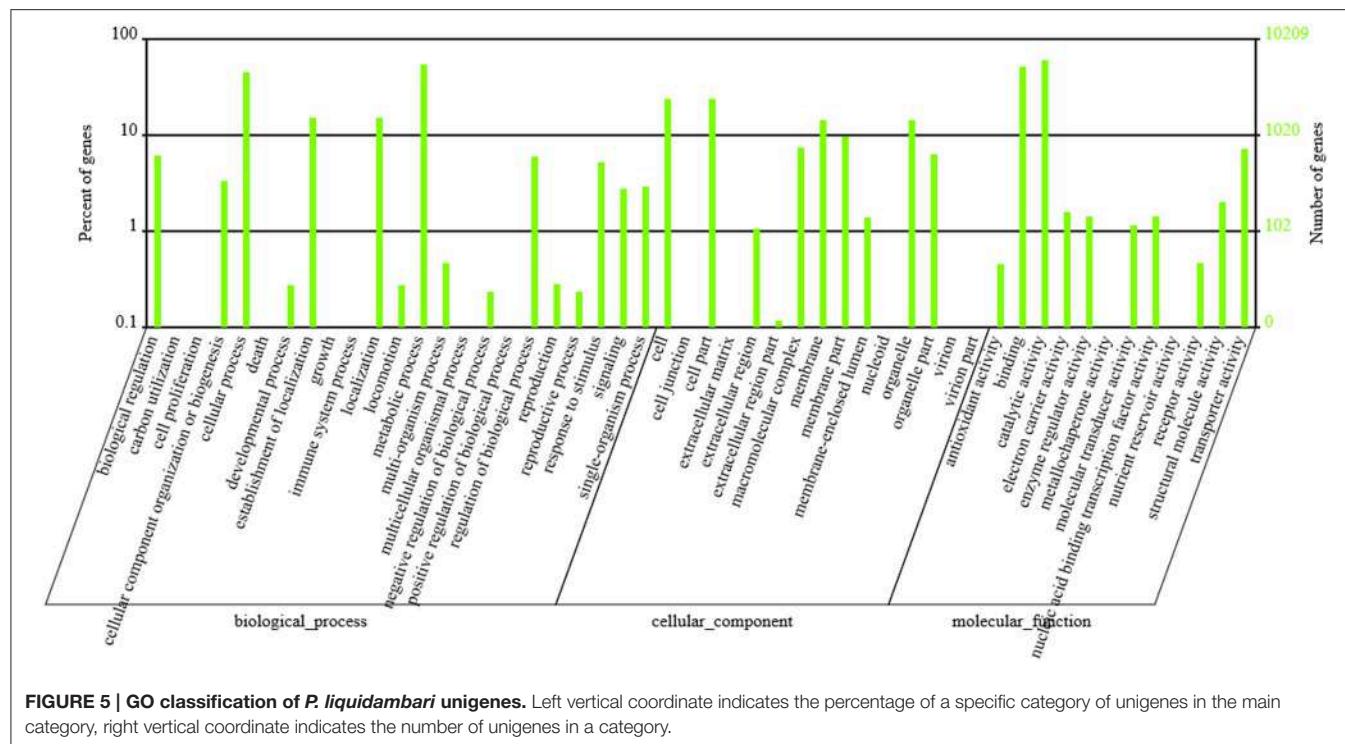


FIGURE 5 | GO classification of *P. liquidambari* unigenes. Left vertical coordinate indicates the percentage of a specific category of unigenes in the main category, right vertical coordinate indicates the number of unigenes in a category.

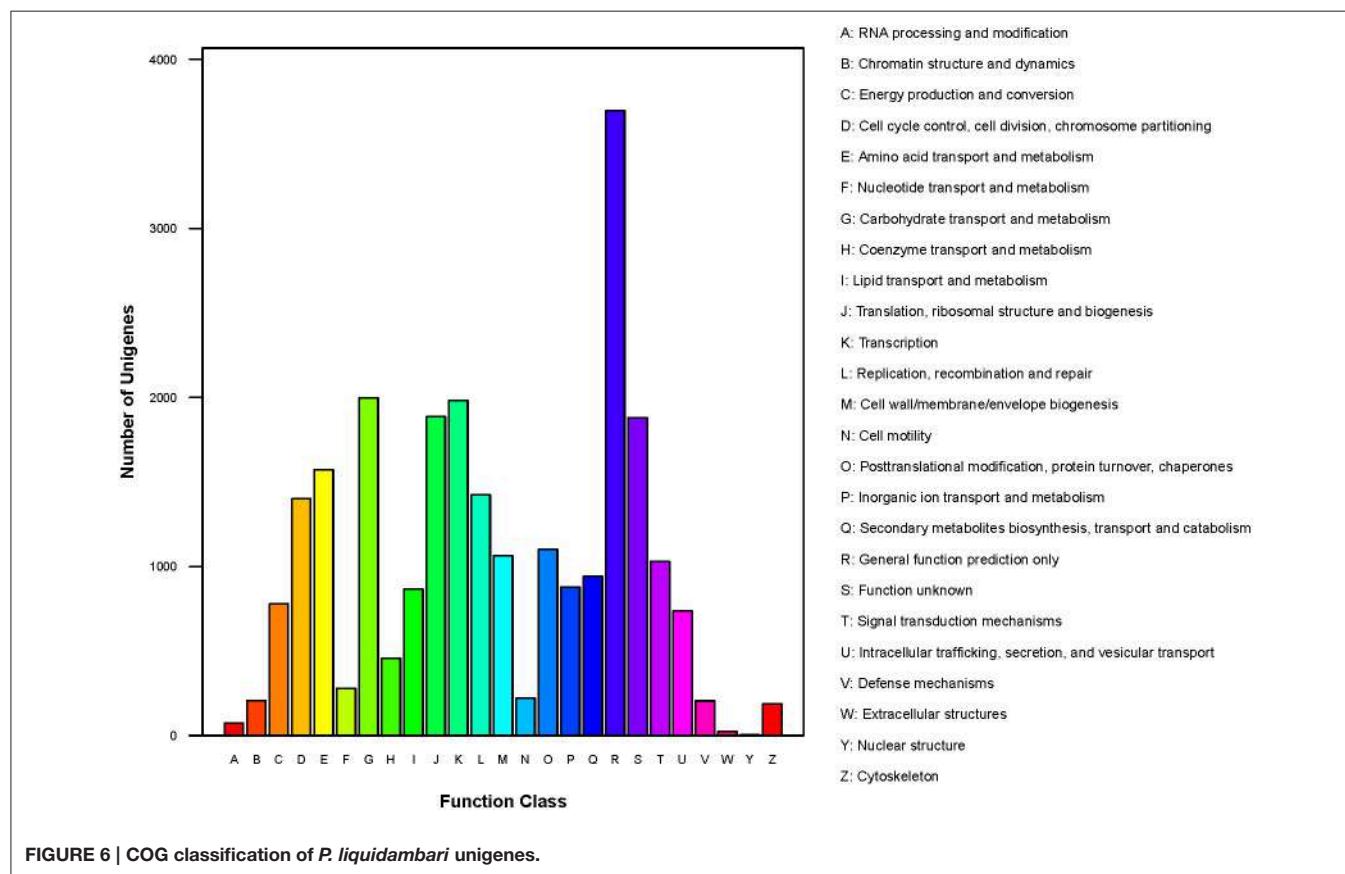


FIGURE 6 | COG classification of *P. liquidambari* unigenes.

TABLE 2 | List of first 20 pathways with highest sequence numbers.

Number	Pathway	All genes with pathway annotation (14,791)	Pathway ID
1	Metabolic pathways	4785 (32.35%)	ko01100
2	Biosynthesis of secondary metabolites	2154 (14.56%)	ko01110
3	Starch and sucrose metabolism	1481 (10.01%)	ko00500
4	RNA transport	817 (5.52%)	ko03013
5	Amino sugar and nucleotide sugar metabolism	814 (5.5%)	ko00520
6	Purine metabolism	630 (4.26%)	ko00230
7	MAPK signaling pathway - yeast	596 (4.03%)	ko04011
8	Protein processing in endoplasmic reticulum	543 (3.67%)	ko04141
9	Cell cycle - yeast	467 (3.16%)	ko04111
10	mRNA surveillance pathway	445 (3.01%)	ko03015
11	Meiosis - yeast	414 (2.8%)	ko04113
12	Spliceosome	409 (2.77%)	ko03040
13	Pyrimidine metabolism	408 (2.76%)	ko00240
14	RNA degradation	370 (2.5%)	ko03018
15	Lysine degradation	365 (2.47%)	ko00310
16	Tyrosine metabolism	329 (2.22%)	ko00350
17	Endocytosis	281 (1.9%)	ko04144
18	Glycolysis / Gluconeogenesis	276 (1.87%)	ko00010
19	Ribosome biogenesis in eukaryotes	256 (1.73%)	ko03008
20	Fructose and mannose metabolism	249 (1.68%)	ko00051

three genes (putative alcohol dehydrogenase, cytochrome P450 monooxygenase and glycoside hydrolase family 72 protein) did not show consistent expression between qRT-PCR and RNA-seq data sets. Comparison of data from Solexa sequencing analysis methods with data obtained from qRT-PCR indicates high credibility for these sequencing methods.

DISCUSSION

Colonization of *P. liquidambari* in Rice

To investigate the fate and behavior of *P. liquidambari* in rice *in situ*, B3 was tagged with the *gfp* gene. Colonization patterns of *P. liquidambari* were roughly divided into three successive time-space stages. First, extracellular colonization of runner hyphae outside the root, mainly concentrated in the base of the root hair, gradually formed a hyphal network on the root surface (<3 dai) (Figures 1A–F and Video S1). Next, entering the biotrophic phase, intracellular and extracellular hyphae underwent branching growth along the root axis, and hyphae were extruded in a deformed fashion (4–10 dai) (Figures 1H–M), meaning that fungal infection began to be restricted. Finally, in the stage of colonization associated with programmed cell death, the vast majority of epidermal cells and part of the outer cortex cells were crowded with a large number of mycelium and sclerotium; the fungal structure

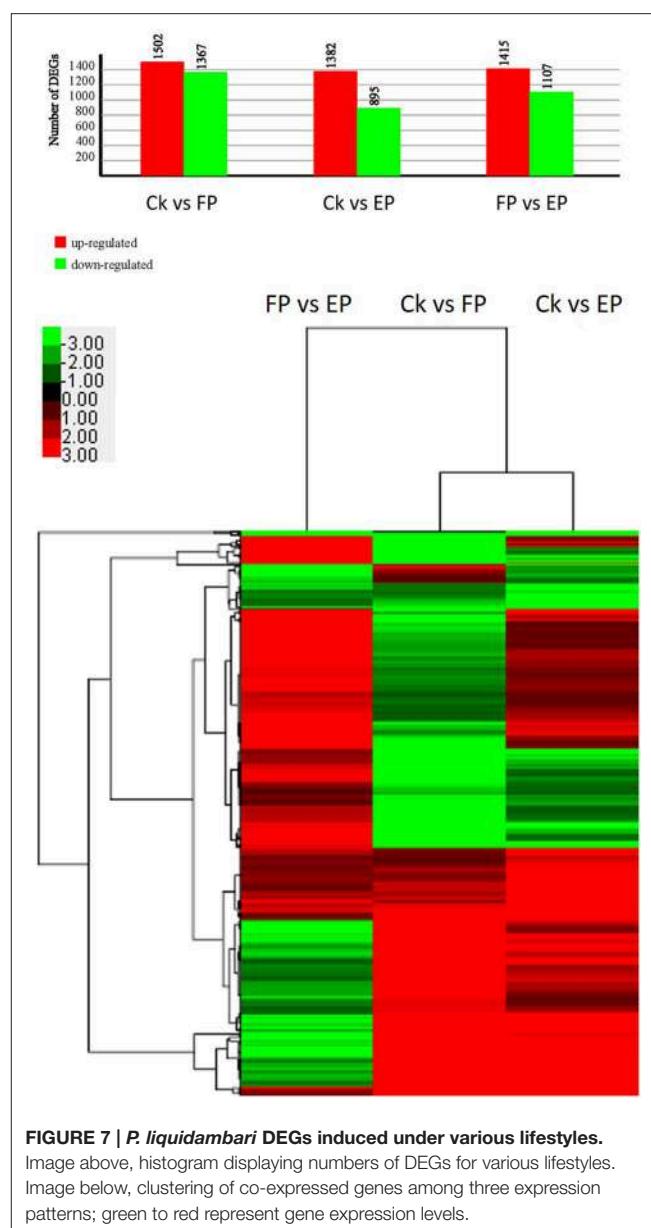
**FIGURE 7 |** *P. liquidambari* DEGs induced under various lifestyles.

Image above, histogram displaying numbers of DEGs for various lifestyles. Image below, clustering of co-expressed genes among three expression patterns; green to red represent gene expression levels.

indicated that fungal infection was further blocked (>15 dai) (Figures 1S,T). Interestingly, *P. liquidambari* still colonized when the host aged or died (Figures 1U–W). It is possible that *P. liquidambari* activated saprophytic programs to adapt to this variation. Colonization patterns of *P. liquidambari* in rice were different from *Harpophora oryzae* (Su et al., 2013), basidiomycete endophyte *Piriformospora indica* (Zuccaro et al., 2011; Lahrmann et al., 2013), soil invaders *Fusarium equiseti*, and *Pochonia chlamydosporia* (Macía-Vicente et al., 2009), which belong to strict root endophytes. It was similar to *Colletotrichum tofieldiae*, in that a fraction of strong hypha penetrated and entered the root axis center, through the vascular bundle into the acral part (Hiruma et al., 2016). During root infection, *P. liquidambari* generated a fungal structure similar to phytopathogen, with neck-like constriction (Figure 1N), showing that any phytopathogen

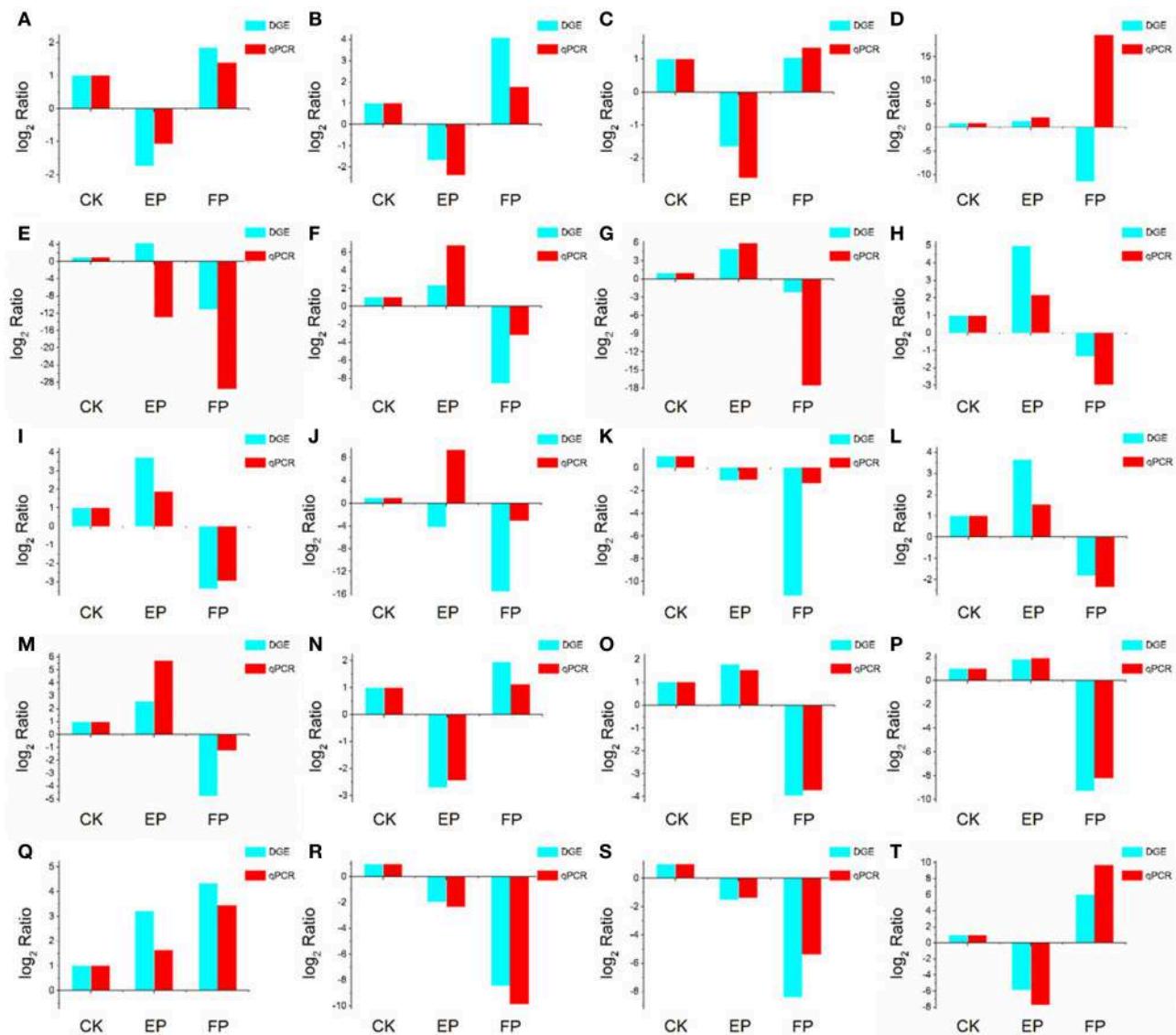


FIGURE 8 | Confirmation of DEGs by qRT-PCR. The x-axis indicates treatment method. The y-axis indicates relative expression level. The following genes were tested (with description). **(A)** glutaminase a (Unigene11948_A), **(B)** probable lysosomal cobalamin transporter (Unigene16976_A), **(C)** NAD(P)-binding protein (Unigene17203_A), **(D)** putative alcohol dehydrogenase (Unigene6694_A), **(E)** cytochrome P450 monooxygenase (Unigene14342_A), **(F)** beta-glucosidase (Unigene19402_A), **(G)** pyruvate decarboxylase (Unigene13421_A), **(H)** cyanide hydratase (Unigene13743_A), **(I)** major facilitator superfamily transporter (CL403.Contig1_A), **(J)** glycoside hydrolase family 72 protein (Unigene7006_A), **(K)** transmembrane amino acid transporter (Unigene19132_A), **(L)** epi1 protein (Unigene2645_A), **(M)** heavy metal translocating P-type ATPase (Unigene13596_A), **(N)** glutathione S-transferase Gst3 (Unigene2773_A), **(O)** phosphoadenosine phosphosulfate reductase (Unigene7219_A), **(P)** nucleolar protein nop-58 (CL1651.Contig1_A), **(Q)** laccase-like protein (Unigene11698_A), **(R)** high affinity copper transporter (Unigene9769_A), **(S)** 3-ketoacyl-CoA reductase (Unigene3887_A), **(T)** glucan 1,4-alpha-maltohexaosidase (Unigene17276_A).

or endophyte can form similar infection structures during root infection. This phenomenon is related to infected tissue and thus belongs to tissue-specific infection. The most obvious similarity between *P. liquidambari* and phytopathogens is that they can infect vascular tissues and systematically spread to acrual parts through vascular tissue (Figures 10–R).

In the colonization process of *P. liquidambari*, a large number of hyphae were limited to the epidermal layer and rhizosphere, and only a fraction of hyphae penetrated to the cortex. This fully

demonstrates that *P. liquidambari* colonization is restricted by space and quantity. In late infection, *P. liquidambari* biomass remained at a steady state after hyphae entered the cortex cells, which can be explained by a reproduction rate of endophyte that was controlled within a certain range (Figure 2). In contrast, the reproduction of pathogen in the root was unrestricted and spread from roots to acrual parts, increasing biomass in an unrestricted fashion and inducing plant disease (Marcel et al., 2010; Su et al., 2013). In contrast, with this exploding

pathogenic infection, the biomass of *P. liquidambari* in the host was maintained in an appropriate range without excessive reproduction. This is reminiscent of endophyte *H. oryzae* symbiosis with rice, initially showed moderate proliferation, subsequently colonization increased rapidly, finally reaching a steady-state level in rice roots (Su et al., 2013). Endophyte *C. tofieldiae* biomass was significantly increased in Trp-derived metabolites mutant, resulted in a severe negative effect on the growth of this mutant and eventually killed the plants (Hiruma et al., 2016). Likewise, the increased colonization of indolic glucosinolates mutant by root-associated fungi *P. indica* and *Sebacina vermifera* led in turn to plant death, suggesting compromised mutualism (Lahrmann et al., 2015). Therefore, another important difference between hostile interactions and mutualistic interactions is quantity rather than quality.

Endophytic and Saprophytic Systems and Transcriptome Sequencing of *P. liquidambari*

The transcriptome is the subset of genes active in tissues and species. To understand the dynamic of the transcriptome it is key to explain the phenotypic changes caused by combinations of genotype and environmental factors (Rockman and Kruglyak, 2006). Recently, Illumina RNA-seq has been used to identify genes of microbes related to plant interactions (Kawahara et al., 2012; O'Connell et al., 2012; Alkan et al., 2015). RNA-seq can be used not only to detect organismic transcripts in existing genomic sequences but also to sequence non-model organisms lacking genomic information. To our knowledge, this is the first study of *Phomopsis* sp. using RNA-seq technology to obtain whole transcriptome information. Our experimental results provide more resources and sequences for studying filamentous endophyte *P. liquidambari*.

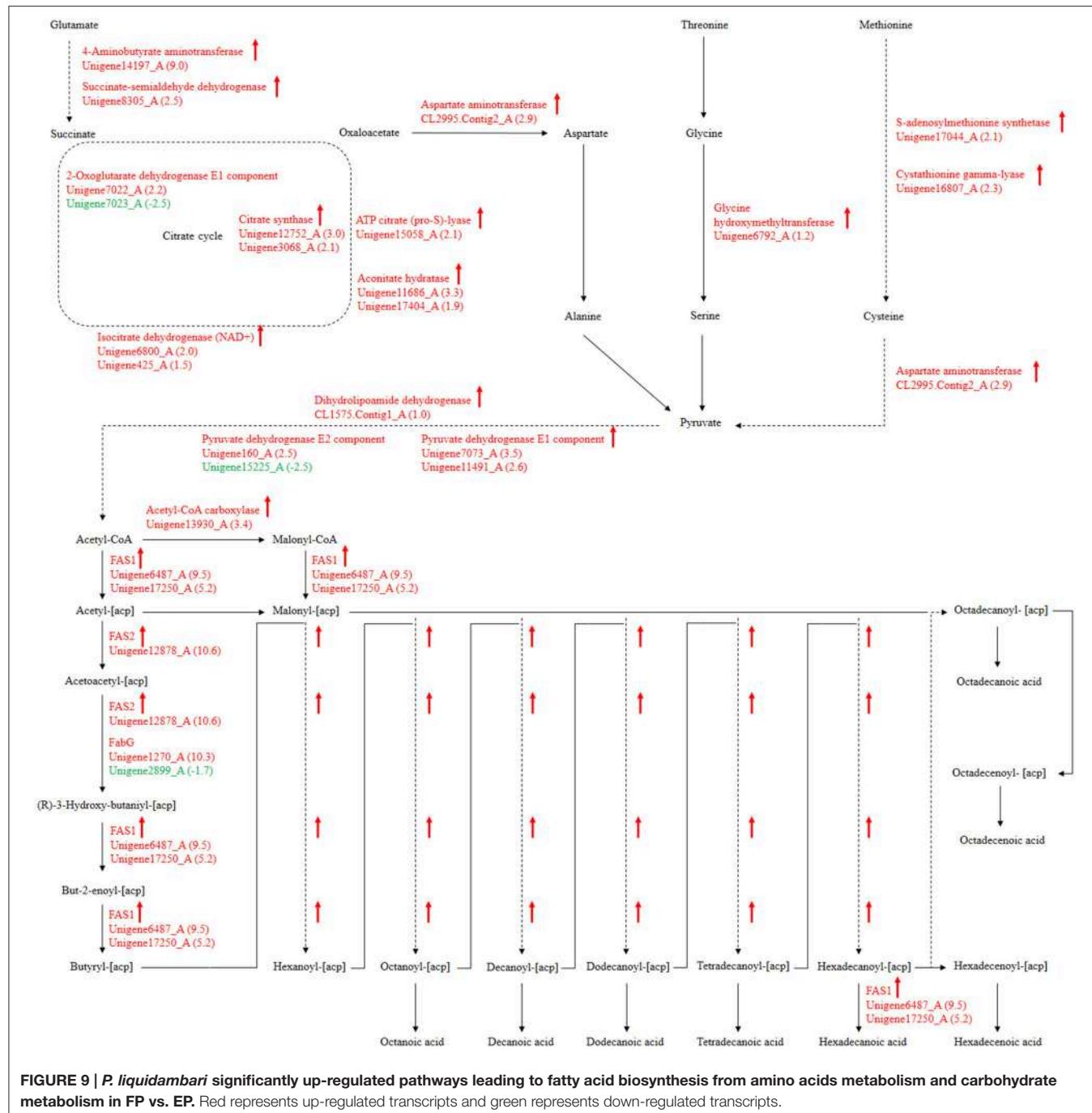
In the past, researchers have investigated the molecular genome of endophytes *in vivo*, but are challenged in retrieving endophyte gene information from high genetic background from the host plant. Compared to unrestrained pathogen reproduction after infection, endophytes steadily undergo symbiosis with the host and establish a subtle counterbalancing relationship that largely limits endophyte growth in the host. When host plants are used as vectors for transcriptome level research, a small number of expressed endophyte genes are typically disregarded due to the significant background of plant genes (Porras-Alfaro and Bayman, 2011). The callus is an undifferentiated living cell structure of plants that contains a set of defense systems similar to host plants (Nawrot-Chorabik et al., 2016). In this study, we used dual cultures of rice callus and *P. liquidambari* to simulate an endophytic environment in which organisms can release signals to recognize each other and form a relationship of simulated antagonism balance. Fungi growth on the callus surface can provide fungal hyphae directly, avoiding the interference from host cell. In litter culture, the humic acid substances of litter will significantly affect RNA quality. Thus, we adopted a litter liquid culture to suspend fungi in a liquid, collected the hyphae by flushing with sterile water and extracted RNA. The quality was improved using this method.

All transcriptome sequencing items were fit with the measurable indicators; Q20 percentage >80%, N percentage <0.5% and GC percentage was 35–65%, showing that the sequencing output and sequence quality were of good quality and could be further analyzed (Table 1).

Divergent Expression Patterns from Amino Acid Metabolism to Fatty Acid Biosynthesis

For our research, 108 biological pathways including the starch and sucrose metabolism pathway, amino sugar and nucleotide sugar metabolic pathways, the fatty acid biosynthesis pathway, and many others were identified by KEGG pathway analysis of unigenes. A total of 1638, 1330, and 1503 DEGs with pathway annotations were identified in the three respective contrast groups. From those pathways, we selected the fatty acid biosynthesis pathway, which is connected to amino acid metabolism and involved in carbohydrate metabolism, for deep analysis. The citrate cycle is the key pathway to energy metabolism. For *P. liquidambari* in the EP compared with FP, genes of the citrate cycle were differentially expressed, significantly enriched ($P < 0.05$), and nearly up-regulated (Figure 9 and Table S7). Most genes for oxidative phosphorylation related to the citrate cycle were also up-regulated. Carbon flux in the fatty acid biosynthesis pathway not only determines the component but also the content of fatty acids in fungi (Hao et al., 2014). Because alanine, aspartate and glutamate metabolism are closely connected to the citrate cycle, and thus genes for these pathways are also enriched and mostly up-regulated. The genes of glycine, serine and threonine metabolism and cysteine and methionine metabolism were up-regulated. The common product of these amino acids metabolisms is pyruvate, and a supply of acetyl-CoA plays a more important role in fungal fatty acid biosynthesis. In *Mortierella alpina*, tyrosine and phenylalanine were considered to contribute NADPH and acetyl-CoA to lipid metabolism through a phenylalanine-hydroxylating system (Wang et al., 2013).

Fatty acid synthesis and transformation plays an important role in fungal growth and development. During intraradical growth, much fatty acid synthesis is required for lipid storage and membrane proliferation of fungi. Adaptation of lipid metabolism may be the prerequisite for symbiosis to achieve function compatibility between fungi and periarbuscular membrane. Fungi were forced to change their membrane lipid composition to allow nutrient exchange between fungal arbuscular and plant periarbuscular membranes (Wewer et al., 2014). In the EP, compared to FP, nine DEGs associated with fatty acid biosynthesis were co-expressed. Interestingly, two identified unigenes were identical to those identified from the GO analysis. Among those nine DEGs, only one unigene (Unigene2899_A) was down-regulated; this unigene encodes 3-oxoacyl-[acyl-carrier protein] reductases (FabG). The other eight unigenes were up-regulated and may be positively regulated genes in the fatty acid biosynthesis pathway. Figure 9 shows the locations of DEGs in the fatty acid biosynthesis pathway. The many up-regulated genes indicate that more positively controlled



genes than negatively regulated genes function in fatty acid biosynthesis.

Polyunsaturated fatty acids or oxylipins can trigger extensive cellular responses, such as pathogenicity arsenals, defense and stress response, secondary metabolism, oxylipin synthesis and cell wall formation. This indicates that generation and recognition are important for coordinating these responses, which can guide pathogen adaptation to host response (Tsitsigiannis and Keller, 2007). Fungal oxylipin repertoire may

participate in the competition between pathogen and host and is also involved in reproduction and development (Oliw et al., 2016). Recent evidence has showed that fatty acids also appear and play a role in beneficial plant-fungi interactions. The endophyte *Fusarium incarnatum* in the embryo of *Aegiceras corniculatum* can produce archetypal plant defense oxylipins that can protect the embryo and are derived from linoleic acid (Pohl and Kock, 2014). Esterified fatty acids of *Lasiodiplodia theobromae* can be used as plant growth regulators in tobacco

and have similar activity to gibberellic acid (Uranga et al., 2016). In addition, commensal *Candida albicans* produce a low-level of resolvin E1, an eicosanoids that works as an effective anti-inflammatory lipid and can inhibit adaptive immune responses and protect commensal yeast from host immune attacks (Pohl and Kock, 2014). DGE data show that genes encoding the FAS1, FAS2, and Acetyl-CoA carboxylase of fatty acid synthesis were up-regulated in EP. FabG genes were both up- and down-regulated (**Figure 9**). There are considerable differences in fatty acid synthesis between the symbiotic and asymbiotic states. Lipid compounds play a key role in symbiotic signals and are likely involved in signaling communication between plants and endophytes. Oxidized fatty acids as signaling molecules have an ancient evolutionary origin and are ideal candidates for inter-kingdom signaling communication (Pohl and Kock, 2014). Oxylipins as intracellular and intercellular communication signals showed vital bioactivities in fungi, plants and animals. The oxylipin signature profile of fungi serves an adjusting function as a “master switch” under different environmental conditions and provides the appropriate mechanisms to microbes by balancing meiospore and mitospore development temporarily. On the basis of *Aspergillus*-seed pathosystems, as supported by data, oxylipin cross-talk is reciprocal. The structural similarity of plant and fungal oxylipins has given rise to a hypothesis that they are important molecules in cross-kingdom communication (Tsitsigiannis and Keller, 2007).

Differences in Secondary Metabolism from Terpenoid to Steroid Biosynthesis

Secondary metabolism is strictly regulated by fungi and is often closely related to asexual reproduction. Secondary metabolites produced from *Trichoderma* include pyrone, antimicrobial peptides and terpenoids that can inhibit the growth of phytopathogens. Several studies have reported that endophytes are involved in the synthesis of plant secondary metabolites

during symbiosis with plants. For example, as *Phoma Medicaginis* switches from the endophytic stage to the saprophytic stage, a large increase in the production of brefeldin A contributes to host defense competitive saprophytes. Low levels of these compounds will inhibit the defense system to maintain the endophytic state of *P. Medicaginis* (Weber et al., 2004). Endophytes can affect growth processes by influencing secondary metabolism under different habitats. Genes encoding secondary metabolism in *Epichloe* spp., *C. tofieldiae* and *H. oryzae* were significantly expressed, but strongly reduced in sebacinales, indicating convergent adaptation to a life inside living host cells (Zuccaro et al., 2011; Fesel and Zuccaro, 2016). Cytochrome P450 monooxygenases play diverse and vital roles in various metabolisms and when fungi adapt to specific ecological niches (Chen et al., 2014). As shown in **Table 3**, compared with FP, the corresponding genes of secondary metabolites synthesis pathways (e.g., stilbenoid, diarylheptanoid and gingerol biosynthesis, phenylpropanoid biosynthesis, terpenoid backbone biosynthesis, ubiquinone and other terpenoid-quinone biosynthesis), were up-regulated. Most of these genes coded for the cytochrome P450 enzyme family and revealed that various cytochrome P450s are involved when filamentous fungi generate a large number of secondary metabolites. Similarly, a phylogenetic analysis revealed the specific expansion of secondary metabolite synthesis genes in *H. oryzae*, as well as cytochrome P450 monooxygenases (Xu et al., 2014). Cytochrome P450 catalyzes biosynthetic metabolism of endogenous substances with important physiological functions such as fatty acids, terpenoids and hormones, and thus P450 plays an important role in the modification of secondary metabolites. In tryptophan metabolism, tryptophan is converted into indole derivatives through these cytochrome P450 enzymes and further forms various secondary metabolites. Several secondary metabolisms originating from tryptophan were essential for beneficial symbiosis with *C. tofieldiae*. Mutation of the organism not only ended this beneficial symbiotic relationship, but

TABLE 3 | Upregulation of genes involved in biosynthesis of secondary metabolites in FP vs. EP.

Category or unigene ID	Gene ID	Annotation	Accession no.	Log ₂ ratio
Unigene14342_A	145246720	cytochrome P450 monooxygenase	XP_001395609.1	15.43
Unigene113_A	310795890	cytochrome P450	EFQ31351.1	14.26
Unigene13908_A	302415555	cytochrome P450 monooxygenase	XP_003005609.1	14.20
Unigene6720_A	310789565	oxidoreductase NAD-binding domain-containing protein	EFQ25098.1	12.81
Unigene10987_A	322705062	branched-chain-amino-acid aminotransferase	EFY96651.1	12.29
Unigene3485_A	16416002	probable cytochrome P450 monooxygenase (lovA)	CAB91316.2	12.04
Unigene19402_A	302423734	beta-glucosidase	XP_003009697.1	10.93
Unigene2980_A	347839081	similar to cytochrome P450 monooxygenase	CCD53653.1	10.88
Unigene57_A	52548220	C-5 sterol desaturase C-like	AAU82099.1	10.77
Unigene14360_A	400600304	Cytochrome P450 CYP505D4	EJP67978.1	10.50
Unigene10045_A	380493536	trichothecene C-8 hydroxylase	CCF33809.1	10.47
Unigene1270_A	358369448	3-ketoacyl-acyl carrier protein reductase	GAA86062.1	10.34
Unigene19132_A	380492609	transmembrane amino acid transporter	CCF34481.1	10.13
Unigene7296_A	240282093	cytochrome P450	EER45596.1	9.89
Unigene17476_A	380488167	CTP synthase	CCF37564.1	9.41
Unigene19143_A	380479107	benzoate 4-monooxygenase cytochrome P450, partial	CCF43218.1	8.47

increased colonization of *C. tofieldiae* such that it eventually killed the host plant (Hiruma et al., 2016). DGE showed that the corresponding genes in the pathway of sesquiterpenoid and triterpenoid biosynthesis were significantly up-regulated ($P < 0.05$) (Figure 10 and Table S7). This is consistent with a previous study that showed that endophyte *Gilmaniaella* sp. AL12 can establish symbiosis with *Atractylodes lancea* and greatly promotes terpenoids accumulation in the herb (Yuan et al., 2016). We found two up-regulated genes in EP: farnesyl-diphosphate farnesyl transferase (Unigene14286_A), involved in isoprenoid biosynthesis, and squalene monooxygenase (Unigene13115_A). Both enzymes have oxidoreductase activity and effect secondary metabolites synthesis and plant-endophyte symbiosis. They also participated in steroid biosynthesis in lipid metabolism. Cytochrome P450 also plays housekeeping functions in fungi. For instance, CYP51 takes part in sterol biosynthesis and is a popular antifungal target to control fungal disease in humans and crops (Becher and Wirsel, 2012). Previous studies have shown that CYP51 and CYP61 play housekeeping functions in the sterol biosynthesis of filamentous fungi (Kelly et al., 2009). The personalized cytochrome P450 components of fungi indicate that highly specialized functions enable evolutionary adaptation to ecological niches.

Distinct Xenobiotic Biodegradation and Metabolism

Our previous studies have reported the capacity of *P. liquidambari* to decompose phenolic acids, cellulose, N-heterocyclic indole and the polycyclic aromatic hydrocarbon phenanthrene (Dai et al., 2010a,b; Chen et al., 2011, 2013b,c). The cytochrome P450 family also contributes to ecological functions as a decomposer or saprophyte. For example, the cytochrome P450s of white-rot fungi *Phanerochaete chrysosporium* is involved in vast xenobiotic biodegradation of extensive environmental toxic chemicals and the natural aromatic polymer lignin (Syed and Yadav, 2012). The diversity of cytochrome P450s may be closely related to fungal survival environment. For example, the cytochrome P450s of white-rot and brown-rot fungi break down plant materials in the environment (Eastwood and Watkinson, 2011; Chen et al., 2014). As DGE showed that cytochrome P450s are involved in xenobiotic biodegradation, we speculated that endophyte *P. liquidambari* appropriately biodegraded harmful xenobiotics or used them as carbon sources to adapt to the host after entering. In addition, endophytes promote production of secondary metabolites that are beneficial to the host and enable both the endophytes and host plants both to grow.

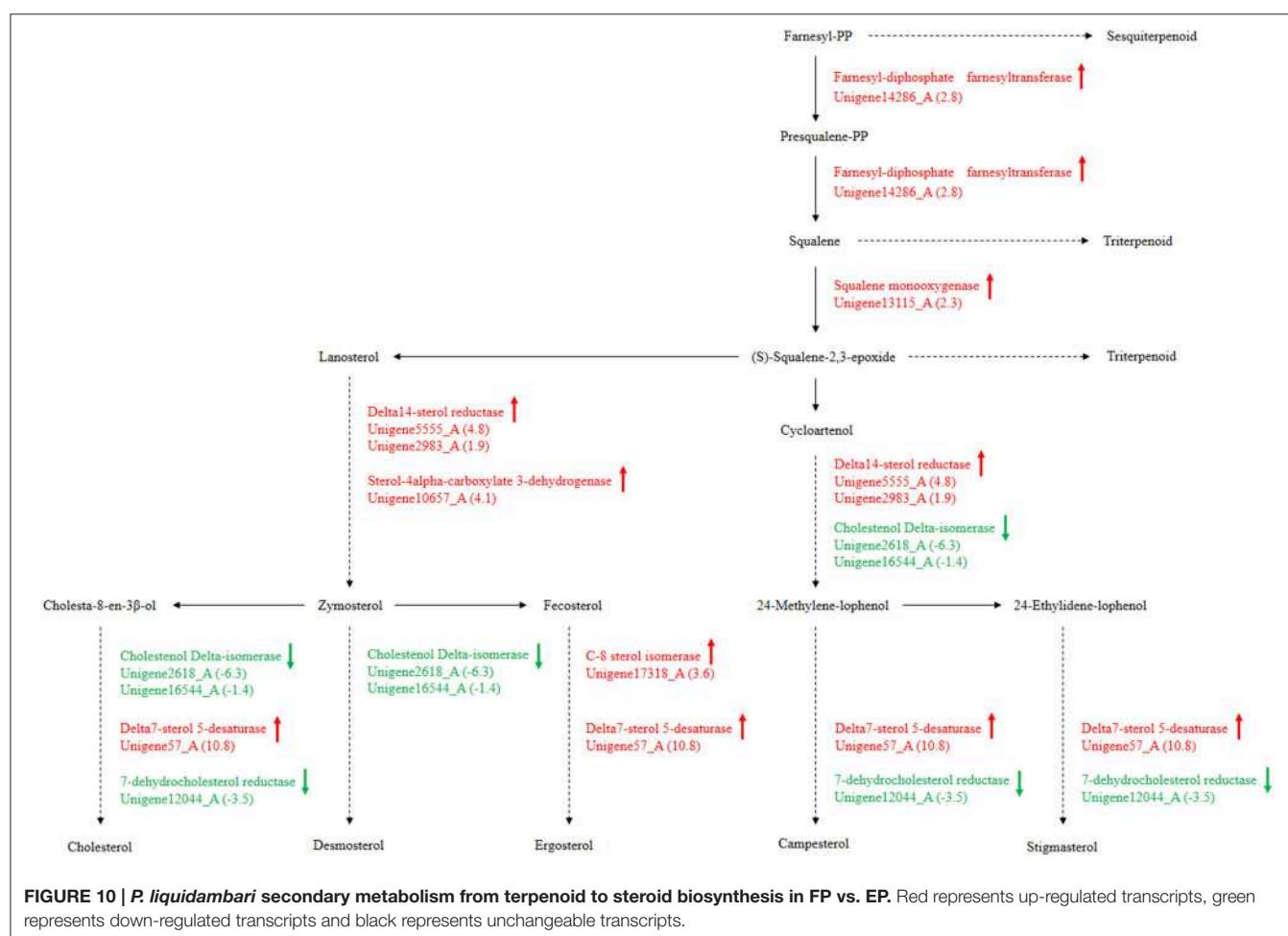


FIGURE 10 | *P. liquidambari* secondary metabolism from terpenoid to steroid biosynthesis in FP vs. EP. Red represents up-regulated transcripts, green represents down-regulated transcripts and black represents unchangeable transcripts.

TABLE 4 | Expression profiles of pathways involved in xenobiotics biodegradation in FP vs. EP.

Enzyme name	Orthology numbers of unigene	Numbers of up-regulated gene	Numbers of down-regulated gene	Expression
BISPHENOL DEGRADATION				
1.14.-.	K00517	23	14	9 Up
4.2.1.-	K01726	1	0	1 Down
1.14.13.-	K00492	4	2	2
1.13.-.	K05915	1	0	1 Down
1.1.1.-	K00100	9	1	8 Down
3.1.1.-	K01066	2	1	1
CHLOROALKANE AND CHLOROALKENE DEGRADATION				
1.1.1.1	K13953; K00121	6	1	5 Down
1.2.1.3	K00128	1	0	1 Down
3.8.1.-	K01564	2	1	1
1.1.1.-	K00100	9	1	8 Down
3.8.1.2	K01560	1	1	0 Up
CAPROLACTAM DEGRADATION				
1.14.13.22	K03379	1	0	1 Down
3.1.1.17	K01053	1	0	1 Down
1.1.1.2	K00002	2	0	2 Down
POLYCYCLIC AROMATIC HYDROCARBON DEGRADATION				
1.3.1.-	K00224	3	0	3 Down
1.14.-.	K00517	23	14	9 Up
1.14.13.1	K00480	4	1	3 Down
1.14.13.-	K00492	4	2	2
2.1.1.-	K00599	2	1	1
NAPHTHALENE DEGRADATION				
1.14.13.-	K00492	4	2	2
1.14.13.1	K00480	4	1	3 Down
1.13.-.	K05915	1	0	1 Down
4.2.1.-	K01726	1	0	1 Down
1.2.1.-	K00155	1	0	1 Down
1.1.1.1	K13953; K00121	6	1	5 Down
CHLOROCYCLOHEXANE AND CHLOROBENZENE DEGRADATION				
3.1.1.45	K01061	3	1	2
1.14.13.-	K00492	4	2	2
3.8.1.2	K01560	1	1	0 Up
1.14.13.7	K03380	1	0	1 Down
1.3.1.-	K00224	3	0	3 Down
AMINOBENZOATE DEGRADATION				
3.5.1.4	K01426	1	0	1 Down
1.14.13.12	K07824	3	3	0 Up
1.14.-.	K00517	23	14	9 Up
1.14.13.7	K03380	1	0	1 Down
3.1.1.-	K01066	2	1	1
3.1.3.1	K01113; K01077	3	0	3 Down
3.1.3.2	K01078	1	0	1 Down
1.14.13.-	K00492	4	2	2
2.3.1.-	K00680	5	4	1 Up

(Continued)

TABLE 4 | Continued

Enzyme name	Orthology numbers of unigene	Numbers of up-regulated gene	Numbers of down-regulated gene	Expression
1.14.14.1	K14338; K00493	2	1	1
STYRENE DEGRADATION				
3.5.1.4	K01426	1	0	1 Down
1.13.11.5	K00451	1	0	1 Down
3.7.1.2	K01555	1	0	1 Down
FLUOROBENZOATE DEGRADATION				
3.1.1.45	K01061	3	1	2
ATRAZINE DEGRADATION				
3.8.1.-	K01564	2	1	1
DIOXIN DEGRADATION				
1.14.13.1	K00480	4	1	3 Down
TOLUENE DEGRADATION				
1.14.13.7	K03380	1	0	1 Down
1.14.13.-	K00492	4	2	2
3.1.1.45	K01061	3	1	2
BENZOATE DEGRADATION				
1.14.13.12	K07824	3	3	0 Up
4.2.1.-	K01726	1	0	1 Down
2.3.1.16	K00632	1	1	0 Up
2.3.1.-	K00680	5	4	1 Up
2.3.1.9	K00626	1	0	1 Down
ETHYLBENZENE DEGRADATION				
2.3.1.16	K00632	1	1	0 Up
2.3.1.-	K00680	5	4	1 Up
METABOLISM OF XENOBIOTICS BY CYTOCHROME P450				
3.3.2.9	K01253	1	0	1 Down
2.5.1.18	K00799	1	1	0 Up
1.1.1.1	K13953; K00121	6	1	5 Down
1.2.1.5	K00129	1	1	0 Up
DRUG METABOLISM - CYTOCHROME P450				
1.14.13.8	K00485	5	1	4 Down
2.5.1.18	K00799	1	1	0 Up
1.1.1.1	K13953; K00121	6	1	5 Down
1.2.1.5	K00129	1	1	0 Up

However, endophytes affected by defensive responses and host plant metabolites cannot overgrow at large scales, resulting in a subtle symbiotic relationship.

It has been increasingly demonstrated that endophytes that quickly decompose plant litter *in vitro* can initiate saprophytic effects when the endophytic survival environment is destroyed, for example through plant senescence or falling. This saprophytic effect is similar to a saprophytic lifestyle that maintains survival and growth by metabolizing compounds in litter that are usually difficult to decompose. In this study, we found that the expression of partial genes involved in xenobiotic biodegradation and metabolism of *P. liquidambari* in FP was up-regulated by

DGE. As shown in **Table 4**, in EP compared with FP, genes of pathways concerning bisphenol degradation, chloroalkane and chloroalkene degradation, caprolactam degradation, polycyclic aromatic hydrocarbon degradation, naphthalene degradation, chlorocyclohexane and chlorobenzene degradation, aminobenzoate degradation, styrene degradation, fluorobenzoate degradation, atrazine degradation, dioxin degradation, toluene degradation, benzoate degradation, ethylbenzene degradation, metabolism of xenobiotics by cytochrome P450 and drug metabolism by cytochrome P450 were both up- and down-regulated. This indicates that *P. liquidambari* can decompose heterocyclic compounds in both endophytic and saprophytic lifestyles. However, most genes are down-regulated in FP vs. EP, indicating that the ability of *P. liquidambari* to degrade aromatic or phenolic compounds was enhanced in a saprophytic lifestyle. Endophyte *P. liquidambari* in a simulated saprophytic environment that lacked nutrition that can be utilized directly forced the fungus to decompose residual organic matter in the litter. Lignin, a main component of litter, was also utilized by the fungus because the corresponding genes for biodegradation of xenobiotics such as bisphenol via lignin degradation were up-regulated. This further verified previous research by Dai et al. (2010b), who found that endophytes can decompose polycyclic aromatic hydrocarbons *in vitro*. Chen et al. (2013a) reported that application of endophyte *P. liquidambari* to soil observably promoted the release of inorganic nitrogen through organic matter degradation. Co-culture of *P. liquidambari* with indole and litter increased indole degradation significantly: 99.1% of indole was removed after 60 h of cultivation, and residual indole levels were below the detection threshold at the 84 h time point (Chen et al., 2013c). Zhou et al. (2014a,b) utilized food waste and wheat straw as nutrient sources in a simulated saprophytic system of cultured *P. liquidambari*. The fermentation product was applied to continuously cropped peanut soil, and the concentrations of vanillic acid, coumaric acid, and 4-hydroxybenzoic acid in soil had decreased by 52.5, 49.4, and 57.4%, respectively, after 28 days. The bacterial and fungal community structures in the rhizosphere soil were affected by changes in phenolic acid concentration and promoted peanut seedling growth and nodulation. These changes demonstrate the application prospects for *P. liquidambari* in the decomposition of difficult-to-decompose organic compounds and environmental remediation. In addition, the results indicate the advantages of nutrient restoration to successive cropping of farmland,

when plant residue exempt from plowing can thoroughly decompose.

AUTHOR CONTRIBUTIONS

JZ performed most of the work, including experimental design and operation, data analysis and manuscript writing. XL and YC prepared samples and extracted fungal RNA. CD supervised all work.

ACKNOWLEDGMENTS

We are grateful to the National Natural Science Foundation of China (NSFC NO. 31570491), a project funded by the Priority Academic Program Development of the Jiangsu Higher Education Institutions, the Research Fund of the State Key Laboratory of Soil and Sustainable Agriculture, Nanjing Institute of Soil Science, Chinese Academy of Science (Y412201435) and the Graduate Research and Innovation Project of Jiangsu Province (KYLX16_1282) for their financial support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.00121/full#supplementary-material>

Table S1 | qPCR primer sequences.

Table S2 | Length distribution of contigs.

Table S3 | Length distribution of unigenes.

Table S4 | Functional annotation of unigenes.

Table S5 | GO classification of *P. liquidambari* unigenes.

Table S6 | KEGG annotation of *P. liquidambari* transcriptome.

Table S7 | Significantly enriched pathway in Ck vs. EP, Ck vs. FP, and FP vs. EP.

Figure S1 | Rice callus.

Figure S2 | Length distribution of CDS for unigenes predicted by BLAST and ESTScan. Length distribution of nucleotide sequence (**A**) and protein sequence (**B**) predicted by BLAST and length distribution of nucleotide sequence (**C**) and protein sequence (**D**) predicted by ESTScan. Horizontal coordinates are sequence size and vertical coordinates are numbers of unigenes.

Video S1 | Three-dimensional scanning for cross-section of root tip, runner hypha interwoven to form hyphal network.

REFERENCES

- Alkan, N., Friedlander, G., Ment, D., Prusky, D., and Fluhr, R. (2015). Simultaneous transcriptome analysis of *Colletotrichum gloeosporioides* and tomato fruit pathosystem reveals novel fungal pathogenicity and fruit defense strategies. *New Phytol.* 205, 801–815. doi: 10.1111/nph.13087
- Audic, S., and Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Res.* 7, 986–995. doi: 10.1101/gr.7.10.986
- Becher, R., and Wirsel, S. G. R. (2012). Fungal cytochrome P450 sterol 14 α -demethylase (CYP51) and azole resistance in plant and human pathogens. *Appl. Microbiol. Biotechnol.* 95, 825–840. doi: 10.1007/s00253-012-4195-9
- Chen, W., Lee, M. K., Jefcoate, C., Kim, S. C., Chen, F., and Yu, J. H. (2014). Fungal cytochrome p450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin. *Genome Biol. Evol.* 6, 1620–1634. doi: 10.1093/gbe/evu132
- Chen, Y., Peng, Y., Dai, C.-C., and Ju, Q. (2011). Biodegradation of 4-hydroxybenzoic acid by *Phomopsis liquidambari*. *Appl. Soil Ecol.* 51, 102–110. doi: 10.1016/j.apsoil.2011.09.004
- Chen, Y., Ren, C. G., Yang, B., Peng, Y., and Dai, C. C. (2013a). Priming effects of the endophytic fungus *Phomopsis liquidambari* on soil mineral N transformations. *Microb. Ecol.* 65, 161–170. doi: 10.1007/s00248-012-0093-z

- Chen, Y., Wang, H. W., Li, L., and Dai, C. C. (2013b). The potential application of the endophyte *Phomopsis liquidambari* to the ecological remediation of long-term cropping soil. *Appl. Soil Ecol.* 67, 20–26. doi: 10.1016/j.apsoil.2013.02.004
- Chen, Y., Xie, X. G., Ren, C. G., and Dai, C. C. (2013c). Degradation of N-heterocyclic indole by a novel endophytic fungus *Phomopsis liquidambari*. *Bioresour. Technol.* 129, 568–574. doi: 10.1016/j.biortech.2012.11.100
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610
- Dai, C. C., Chen, Y., Tian, L. S., and Shi, Y. (2010a). Correlation between invasion by endophytic fungus *Phomopsis* sp and enzyme production. *Afr. J. Agric. Res.* 5, 1324–1330. doi: 10.5897/AJAR09.541
- Dai, C. C., Tian, L. S., Zhao, Y. T., Chen, Y., and Xie, H. (2010b). Degradation of phenanthrene by the endophytic fungus *Ceratobasidium stevensii* found in *Bischofia polycarpa*. *Biodegradation* 21, 245–255. doi: 10.1007/s10532-009-9297-4
- Eastwood, D. C., and Watkinson, S. C. (2011). The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science* 333, 762–765. doi: 10.1126/science.1205411
- Fesel, P. H., and Zuccaro, A. (2016). Dissecting endophytic lifestyle along the parasitism/mutualism continuum in *Arabidopsis*. *Curr. Opin. Microbiol.* 32, 103–112. doi: 10.1016/j.mib.2016.05.008
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Hao, G., Chen, H., Wang, L., Gu, Z., Song, Y., Zhang, H., et al. (2014). Role of malic enzyme during fatty acid synthesis in the oleaginous fungus *Mortierella alpina*. *Appl. Environ. Microbiol.* 80, 2672–2678. doi: 10.1128/AEM.00140-14
- Hiruma, K., Gerlach, N., Sacristan, S., Nakano, R. T., Hacquard, S., Kracher, B., et al. (2016). Root endophyte *colletotrichum tofieldiae* confers plant fitness benefits that are phosphate status dependent. *Cell* 165, 464–474. doi: 10.1016/j.cell.2016.02.028
- Hyde, K. D., Bussaban, B., Paulus, B., Crous, P. W., Lee, S., McKenzie, E. H. C., et al. (2007). Diversity of saprobic microfungi. *Biodivers. Conserv.* 16, 7–35. doi: 10.1007/s10531-006-9119-5
- Kaul, S., Sharma, T., and Dhar, M. K. (2016). “Omics” tools for better understanding the plant-endophyte interactions. *Front. Plant Sci.* 7:955. doi: 10.3389/fpls.2016.00955
- Kawahara, Y., Oono, Y., Kanamori, H., Matsumoto, T., Itoh, T., and Minami, E. (2012). Simultaneous RNA-seq analysis of a mixed transcriptome of rice and blast fungus interaction. *PLoS ONE* 7:e49423. doi: 10.1371/journal.pone.0049423
- Kelly, D. E., Kraševc, N., Mullins, J., and Nelson, D. R. (2009). The CYPome (Cytochrome P450 complement) of *Aspergillus nidulans*. *Fungal Genet. Biol.* 46, S53–S61. doi: 10.1016/j.fgb.2008.08.010
- Lahrmann, U., Ding, Y., Banbara, A., Rath, M., Hajirezaei, M. R., Döhlemann, S., et al. (2013). Host-related metabolic cues affect colonization strategies of a root endophyte. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13965–13970. doi: 10.1073/pnas.1301653110
- Lahrmann, U., Strehmel, N., Langen, G., Frerigmann, H., Leson, L., Ding, Y., et al. (2015). Mutualistic root endophytism is not associated with the reduction of saprotrophic traits and requires a noncompromised plant innate immunity. *New Phytol.* 207, 841–857. doi: 10.1111/nph.13411
- Maciá-Vicente, J. G., Jansson, H.-B., Talbot, N. J., and Lopez-Llorca, L. V. (2009). Real-time PCR quantification and live-cell imaging of endophytic colonization of barley (*Hordeum vulgare*) roots by *Fusarium equiseti* and *Pochonia chlamydosporia*. *New Phytol.* 182, 213–228. doi: 10.1111/j.1469-8137.2008.02743.x
- Marcel, S., Sawers, R., Oakeley, E., Angliker, H., and Paszkowski, U. (2010). Tissue-adapted invasion strategies of the rice blast fungus *Magnaporthe oryzae*. *Plant Cell* 22, 3177–3187. doi: 10.1105/tpc.110.078048
- Martinez, D., Challacombe, J., Morgenstern, I., Hibbett, D., Schmoll, M., Kubicek, C. P., et al. (2009). Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1954–1959. doi: 10.1073/pnas.0809575106
- Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628. doi: 10.1038/nmeth.1226
- Nawrot-Chorabik, K., Grad, B., Kowalski, T., and Woodward, S. (2016). Interactions between callus cultures of *Pinus sylvestris* and pine fungi with different trophic properties. *For. Pathol.* 46, 179–186. doi: 10.1111/efp.12240
- O’connell, R. J., Thon, M. R., Hacquard, S., Amyotte, S. G., Kleemann, J., Torres, M. F., et al. (2012). Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nat. Genet.* 44, 1060–1065. doi: 10.1038/ng.2372
- Oliw, E. H., Aragó, M., Chen, Y., and Jernerén, F. (2016). A new class of fatty acid allene oxide formed by the dioxygenase-p450 fusion proteins of human and plant pathogenic fungi, *Coccidioides immitis* and *Zymoseptoria tritici*. *J. Lipid Res.* 57, 1518–1528. doi: 10.1194/jlr.M068981
- Peters, S., and Schulz, B. (1998). Interactions in dual cultures of endophytic fungi with host and nonhost plant calli. *Mycologia* 90, 360–367. doi: 10.2307/3761393
- Pohl, C. H., and Kock, J. L. (2014). Oxidized fatty acids as inter-kingdom signaling molecules. *Molecules* 19, 1273–1285. doi: 10.3390/molecules19011273
- Porras-Alfaro, A., and Bayman, P. (2011). Hidden fungi, emergent properties: endophytes and microbiomes. *Annu. Rev. Phytopathol.* 49, 291–315. doi: 10.1146/annurev-phyto-080508-081831
- Promputtha, I., Hyde, K. D., McKenzie, E. H. C., Peberdy, J. F., and Lumyong, S. (2010). Can leaf degrading enzymes provide evidence that endophytic fungi become saprobes? *Fungal Divers.* 41, 89–99. doi: 10.1007/s13225-010-0024-6
- Promputtha, I., Lumyong, S., Dhanasekaran, V., McKenzie, E. H., Hyde, K. D., and Jeewon, R. (2007). A phylogenetic evaluation of whether endophytes become saprotrophs at host senescence. *Microb. Ecol.* 53, 579–590. doi: 10.1007/s00248-006-9117-x
- Rockman, M. V., and Kruglyak, L. (2006). Genetics of global gene expression. *Nat. Rev. Genet.* 7, 862–872. doi: 10.1038/nrg1964
- Siddique, M. A., Zereen, M. I., Li, C. F., and Dai, C. C. (2016). Endophytic fungus *Phomopsis liquidambari* and different doses of N-fertilizer alter microbial community structure and function in rhizosphere of rice. *Sci. Rep.* 6:32270. doi: 10.1038/srep32270
- Sieber, T. N., Siebercanavesi, F., and Dorworth, C. E. (1990). Simultaneous stimulation of endophytic *Cryptodiaporthe hystrix* and inhibition of *acer macrophyllum* callus in dual culture. *Mycologia* 82, 569–575. doi: 10.2307/3760047
- Su, Z. Z., Mao, L. J., Li, N., Feng, X. X., Yuan, Z. L., Wang, L. W., et al. (2013). Evidence for biotrophic lifestyle and biocontrol potential of dark septate endophyte *Harpophora oryzae* to rice blast disease. *PLoS ONE* 8:e61332. doi: 10.1371/journal.pone.0061332
- Syed, K., and Yadav, J. S. (2012). P450 monooxygenases (P450ome) of the model white rot fungus *Phanerochaete chrysosporium*. *Crit. Rev. Microbiol.* 38, 339–363. doi: 10.3109/1040841X.2012.682050
- Tsitsigiannis, D. I., and Keller, N. P. (2007). Oxylipins as developmental and host-fungal communication signals. *Trends Microbiol.* 15, 109–118. doi: 10.1016/j.tim.2007.01.005
- Uranga, C. C., Beld, J., Mrse, A., Cordova-Guerrero, I., Burkart, M. D., and Hernández-Martínez, R. (2016). Fatty acid esters produced by *Lasiodiplodia theobromae* function as growth regulators in tobacco seedlings. *Biochem. Biophys. Res. Commun.* 472, 339–345. doi: 10.1016/j.bbrc.2016.02.104
- Wang, H., Chen, H., Hao, G., Yang, B., Feng, Y., Wang, Y., et al. (2013). Role of the phenylalanine-hydroxylating system in aromatic substance degradation and lipid metabolism in the oleaginous fungus *Mortierella alpina*. *Appl. Environ. Microbiol.* 79, 3225–3233. doi: 10.1128/AEM.00238-13
- Wang, H. W., Dai, C. C., Zhu, H., and Wang, X. X. (2014). Survival of a novel endophytic fungus *Phomopsis liquidambari* B3 in the indole-contaminated soil detected by real-time PCR and its effects on the indigenous microbial community. *Microbiol. Res.* 169, 881–887. doi: 10.1016/j.mires.2014.05.006
- Wang, H. W., Zhang, W., Su, C. L., Zhu, H., and Dai, C. C. (2015). Biodegradation of the phytoestrogen luteolin by the endophytic fungus *Phomopsis liquidambari*. *Biodegradation* 26, 197–210. doi: 10.1007/s10532-015-9727-4

- Weber, R. W., Stenger, E., Meffert, A., and Hahn, M. (2004). Brefeldin A production by *Phoma medicaginis* in dead pre-colonized plant tissue: a strategy for habitat conquest? *Mycol. Res.* 108, 662–671. doi: 10.1017/s0953756204000243
- Wewer, V., Brands, M., and Dörmann, P. (2014). Fatty acid synthesis and lipid metabolism in the obligate biotrophic fungus *Rhizophagus irregularis* during mycorrhization of *Lotus japonicus*. *Plant J.* 79, 398–412. doi: 10.1111/tpj.12566
- Xie, X. G., and Dai, C. C. (2015). Degradation of a model pollutant ferulic acid by the endophytic fungus *Phomopsis liquidambari*. *Bioresour. Technol.* 179, 35–42. doi: 10.1016/j.biortech.2014.11.112
- Xu, X. H., Su, Z. Z., Wang, C., Kubicek, C. P., Feng, X. X., Mao, L. J., et al. (2014). The rice endophyte *Harpophora oryzae* genome reveals evolution from a pathogen to a mutualistic endophyte. *Sci. Rep.* 4:5783. doi: 10.1038/srep05783
- Yang, B., Ma, H. Y., Wang, X. M., Jia, Y., Hu, J., Li, X., et al. (2014a). Improvement of nitrogen accumulation and metabolism in rice (*Oryza sativa* L.) by the endophyte *Phomopsis liquidambari*. *Plant Physiol. Biochem.* 82, 172–182. doi: 10.1016/j.plaphy.2014.06.002
- Yang, B., Wang, X.-M., Ma, H.-Y., Jia, Y., Li, X., and Dai, C.-C. (2014b). Effects of the fungal endophyte *Phomopsis liquidambari* on nitrogen uptake and metabolism in rice. *Plant Growth Regul.* 73, 165–179. doi: 10.1007/s10725-013-9878-4
- Yang, B., Wang, X. M., Ma, H. Y., Yang, T., Jia, Y., Zhou, J., et al. (2015). Fungal endophyte *Phomopsis liquidambari* affects nitrogen transformation processes and related microorganisms in the rice rhizosphere. *Front. Microbiol.* 6:982. doi: 10.3389/fmicb.2015.00982
- Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., et al. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* 34, W293–W297. doi: 10.1093/nar/gkl031
- Yuan, J., Sun, K., Deng-Wang, M. Y., and Dai, C. C. (2016). The mechanism of ethylene signaling induced by Endophytic Fungus *Gilmania* sp. AL12 mediating sesquiterpenoids biosynthesis in *Atractylodes lancea*. *Front. Plant Sci.* 7:361. doi: 10.3389/fpls.2016.00361
- Zhou, J., Kang, L., Wang, H.-W., Yang, T., and Dai, C.-C. (2014a). Liquid laccase production by *Phomopsis liquidambari* B3 accelerated phenolic acids degradation in long-term cropping soil of peanut. *Acta Agric. Scand. Sec. B — Soil Plant Sci.* 64, 683–693. doi: 10.1080/09064710.2014.953987
- Zhou, J., Yang, T., Mei, Y.-Z., Kang, L., and Dai, C.-C. (2014b). Laccase production by *Phomopsis liquidambari* B3 cultured with food waste and wheat straw as the main nitrogen and carbon sources. *J. Air Waste Manage. Assoc.* 64, 1154–1163. doi: 10.1080/10962247.2014.930077
- Zuccaro, A., Lahrmann, U., Guldener, U., Langen, G., Pfiffi, S., Biedenkopf, D., et al. (2011). Endophytic life strategies decoded by genome and transcriptome analyses of the mutualistic root symbiont *Piriformospora indica*. *PLoS Pathog.* 7:e1002290. doi: 10.1371/journal.ppat.1002290

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Zhou, Li, Chen and Dai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Natural *Agrobacterium* Transformants: Recent Results and Some Theoretical Considerations

Ke Chen¹ and Léon Otten^{2*}

¹ Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China, ² Institut de Biologie Moléculaire des Plantes, Centre National de la Recherche Scientifique (CNRS), Strasbourg, France

OPEN ACCESS

Edited by:

Nikolai Provorov,
All-Russian Research Institute of Agricultural Microbiology of The Russian Academy of Agricultural Sciences, Russia

Reviewed by:

Igor Kovalchuk,
University of Lethbridge, Canada
Stanton B. Gelvin,
Purdue University, United States

*Correspondence:

Léon Otten
leon.otten@ibmp-cnrs.unistra.fr

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 26 May 2017

Accepted: 31 August 2017

Published: 13 September 2017

Citation:

Chen K and Otten L (2017) Natural *Agrobacterium* Transformants: Recent Results and Some Theoretical Considerations. *Front. Plant Sci.* 8:1600.
doi: 10.3389/fpls.2017.01600

Agrobacterium rhizogenes causes hairy root growth on a large number of plant species. It does so by transferring specific DNA fragments (T-DNA) from its root-inducing plasmid (pRi) into plant cells. Expression of T-DNA genes leads to abnormal root growth and production of specific metabolites (opines) which are taken up by the bacterium and used for its growth. Recent work has shown that several *Nicotiana*, *Linaria*, and *Ipomoea* species contain T-DNA genes from *A. rhizogenes* in their genomes. Plants carrying such T-DNAs (called cellular T-DNA or cT-DNA) can be considered as natural transformants. In the *Nicotiana* genus, seven different T-DNAs are found originating from different *Agrobacterium* strains, and in the *Tomentosae* section no <4 successive insertion events took place. In several cases cT-DNA genes were found to be expressed. In some *Nicotiana tabacum* cultivars the opine synthesis gene TB-mas2' is expressed in the roots. These cultivars were found to produce opines. Here we review what is known about natural *Agrobacterium* transformants, develop a theoretical framework to analyze this unusual phenomenon, and provide some outlines for further research.

Keywords: *Agrobacterium rhizogenes*, *Nicotiana tabacum*, hairy roots, natural transformation, T-DNA

INTRODUCTION

Agrobacterium is well-known for its capacity to transfer part of its DNA to plants during a natural infection process leading to tumors (Crown gall) or abnormal roots (Hairy roots, HR) (Gelvin, 2012; Christie and Gordon, 2014; Kado, 2014). The genus *Agrobacterium* contains different species such as *A. tumefaciens*, *A. rhizogenes* (Riker, 1930), *A. vitis* (Ophel and Kerr, 1990), and *A. rubi* (Hildebrand, 1940). Another classification uses biotypes (Kerr and Panagopoulos, 1977). The transferred DNA (T-DNA) is located on a large plasmid (tumor-inducing or Ti plasmid) or root-inducing plasmid (pRi plasmid). Strains can carry one, two, or three T-DNAs on their pTi/pRi plasmid (Canaday et al., 1992). T-DNAs are surrounded by direct repeats of about 25 nucleotides (called borders). The transfer starts from the so-called right border and proceeds to the left border. Often, the integrated T-DNAs are incomplete and truncated at the left part. They can occur as single copies or as tandem or inverted repeats.

Genes located on the T-DNA are expressed in the plant cells and lead to growth changes (Binns and Costantino, 1998) and opine synthesis. Sterile Crown gall and HR tissues contain opines (Bielmann et al., 1960; Ménagé and Morel, 1964). They constitute different families of conjugated structures, the nature of which depends on the inciting bacterium. Opines often accumulate to very large quantities as they cannot be metabolized by the plant (Scott, 1979). Uptake and degradation

of opines by *Agrobacterium* are encoded by specific genes located on the pTi or pRi plasmid, outside the T-DNA region(s), and agrobacteria can be attracted to opine sources by chemotaxis (Kim and Farrand, 1998). pTi/pRi plasmids can be transferred to other *Agrobacterium* strains by a conjugation process which can be induced by opines. Much has already been learnt about the way *Agrobacterium* transfers its T-DNA to plants (Gelvin, 2012; Christie and Gordon, 2014; Kado, 2014).

In 1983 it was discovered by Southern blot analysis (White et al., 1983) that *N. glauca* (Solanaceae family, *Noctiflorae* section of the *Nicotiana* genus) carries *A. rhizogenes*-like sequences in its nuclear genome. These sequences were called cellular T-DNAs (cT-DNAs). A more extensive study (Furner et al., 1986) involving other members of the *Nicotiana* genus revealed cT-DNA sequences in *N. tabacum*, *N. tomentosiformis*, *N. tomentosa*, and *N. otophora* (all belong to the *Tomentosae* section). Although *N. benavidesii* (section *Paniculatae*) was also mentioned as carrying a cT-DNA, there is no strong evidence for this.

A partial map of the *N. glauca* cT-DNA was obtained showing two dissimilar T-DNA copies linked together as an inverted repeat (called left and right arm). This map was later completed (Suzuki et al., 2002). In the case of *N. tabacum*, a few cT-DNA fragments were sequenced (Meyer et al., 1995; Fründt et al., 1998a,b; Intrieri and Buiatti, 2001; Suzuki et al., 2002; Mohajel-Shoja et al., 2011). It has been reported that *C. arvensis* and carrot contain T-DNA sequences (D. Tepfer, cited in Matveeva and Lutova, 2014 and elsewhere), but this could not be confirmed by others (Matveeva and Lutova, 2014).

In 2012, a large-scale survey led to the discovery of cT-DNA sequences in *Linaria vulgaris*, a member of the Plantaginaceae family (Matveeva et al., 2012). In 2014, deep sequencing revealed four cT-DNAs (TA, TB, TC, and TD) in *N. tomentosiformis* and their distribution was studied in related species of the section *Tomentosae*. An additional type of cT-DNA sequence (TE) was found in *N. otophora* (Chen et al., 2014). In 2015, cT-DNA sequences were reported for *Ipomoea batatas* (Convolvulaceae family), a common crop. This species contains two cT-DNAs, *Ib*T-DNA1 and *Ib*T-DNA2. *Ib*T-DNA1 was found in cultivated sweet potatoes but not in wild relatives, whereas *Ib*T-DNA2 was found in both (Kyndt et al., 2015). Thus, gene transfer from agrobacteria to various plant species (natural genetic transformation) had occurred under natural circumstances. This led to genetically stable transformants, which we will call «natural transformants».

Although the study of natural transformants is still in its infancy, we would like to summarize recent observations and develop several theoretical considerations that may be useful for further investigations. We will start by having a close look at the agent that introduced the cT-DNAs: *A. rhizogenes*.

AGROBACTERIUM RHIZOGENES STRAINS AND THEIR VARIABILITY

Fründt et al. (1998a) speculated that cT-DNAs were initially normal plant sequences that were later captured by agrobacteria and employed for tumor and HR induction. We believe this

is very unlikely because of the following reasons: cT-DNAs are absent from most plant species, their phylogenies do not match plant phylogenies, and the cT-DNAs end at the classical pRi T-DNA right borders as expected for transfer by *Agrobacterium*. Thus, there is little doubt that plants with cT-DNAs were indeed transformed by *Agrobacterium*.

The published cT-DNA structures all seem to be derived from *A. rhizogenes*-like T-DNAs. We know relatively little about *A. rhizogenes* strains, their Ri plasmids, and their T-DNA structures. Only a few strains have been studied and classified into mikimopine, cucumopine, agropine, and mannopine strains (represented by strains MAFF03-01724, NCPPB2659, ATCC15834, and NCIB8196 respectively) according to the opines they induce in the transformed roots. Their host ranges are very broad (De Cleene and De Ley, 1981).

The opine-based *A. rhizogenes* classification has no phylogenetic value because opine genes can be exchanged between different agrobacteria by horizontal gene transfer. Frequent horizontal gene transfer makes the construction of phylogenetic trees for T-DNA structures, pTi/pRi plasmids, and whole genomes practically impossible. Even if thousands of *Agrobacterium* genomes were available, it might still be impossible to establish phylogenetic trees (Van Nuenen et al., 1993). This was illustrated by a detailed analysis of *A. vitis*, the only *Agrobacterium* species for which a large number of isolates were compared. Three very different pTi types were found, but no intermediate structures, making it impossible to construct a tree. These studies suggested the selection of particular T-DNA gene combinations, loss of intermediates, and expansion of efficient strains into a few dominant groups (Burr and Otten, 1999).

Horizontal gene transfer also leads to chimeric T-DNAs. Examples are the pRi1724, pRiA4, and pRi2659 T-DNAs: their central parts are very similar, but close to the right border pRi1724 carries a mikimopine synthase (*mis*) gene, pRiA4 has an ornithine cyclodeaminase gene (*rolD*, Trovato et al., 2001), and pRi2659 a cucumopine synthase (*cus*) gene. These differences are most likely due to recombinations between different Ti plasmids (Otten and De Ruffray, 1994).

WHICH TYPES OF AGROBACTERIUM STRAINS INTRODUCED THE cT-DNAS?

Because pRi plasmids can be exchanged between *Agrobacterium* strains and are often chimeric, it is very difficult (if not impossible) to attribute a cT-DNA to a particular type of *Agrobacterium* strain. For example, the *N. glauca* cT-DNA strongly resembles part of the pRi1724 T-DNA, but the bacterium that introduced the cT-DNA is not necessarily derived from a 1724-like *A. rhizogenes* strain, since the remaining genome might be completely different. Unless natural transformation can be directly observed to occur in nature (see below), it will be impossible to identify the strain responsible for a natural transformation event on the sole basis of a cT-DNA sequence. In order to get a better idea of the pRi and T-DNA gene repertoire of *A. rhizogenes*, more isolates will have

to be investigated. The variation in *A. rhizogenes* T-DNA structures is probably quite large, as shown by the new cT-DNA sequences. In *N. tomentosiformis*, six previously unknown T-DNA genes were found: two (in TA and TD) are distantly related to *orf14*, one codes for a protein with weak similarity to agrocinopine synthase (Acs, TB), another for a protein with weak similarity to octopine synthase (Ocs, TC), one for a C-like protein (*c*-like gene, TC), and one for a large, completely unknown protein (Orf511, TD). It is noteworthy that octopine synthase-like genes are normally only found in *A. tumefaciens* or *A. vitis*. In *N. otophora*, vitopine synthase (*vis*)-like sequences (distantly related to *ocs*) and 6b genes with low similarity to their counterparts in *A. tumefaciens* and *A. vitis* were found alongside typical *A. rhizogenes* T-DNA genes such as *rolC*, *orf13*, and *orf14* (Chen et al., 2014). *Ib*T-DNA2 of *I. batatas* carries typical *A. rhizogenes* genes (*orf13*, *orf14*, *rolB*, *orf17n*, *orf18*) but with an unusual organization and an unusual *rolB*-like gene. *Ib*T-DNA1 carries *iaaM*, *iaaH*, C-protein, and *acs* genes (Kyndt et al., 2015). The latter gene combination has been found in *A. tumefaciens* strain C58 and in the *A. vitis* strain Tm4 TB region (Otten et al., 1999), but not in *A. rhizogenes*. These unusual T-DNA structures and genes were introduced by unknown *Agrobacterium* strains which might possess unusual root-inducing properties. However, if transformation happened long ago, strains might have evolved toward other forms or disappeared altogether.

In the next three sections we will discuss when the different transformation events could have taken place and how they relate to the evolutionary history of the recipient plants.

ACCUMULATION OF cT-DNAS BY SUCCESSIVE TRANSFORMATIONS

When it was discovered that different *Nicotiana* species carry cT-DNAs in their genomes (Furner et al., 1986), it was suggested that this could result from the transformation of a common ancestor species. In a later report, two possibilities were proposed to explain the presence of T-DNA genes in *N. glauca* (*Noctiflorae* section), but at that time considered part of the *Paniculatae* section) and *N. tomentosiformis* (*Tomentosae* section). First, a T-DNA was inserted in an ancestor of these sections (part of the *Nicotiana* Cestroid ancestral complex) and inherited by the descendants. Second, the two cT-DNAs were inserted separately and independently, after the split between the two sections (Meyer et al., 1995). When the genome sequences of *N. tomentosiformis* (Chen et al., 2014), *N. otophora*, and three cultivars of *N. tabacum* (Sierro et al., 2014) became available, the situation turned out to be considerably more complex. The *N. tomentosiformis* genome was found to contain four cT-DNAs, each from a different *Agrobacterium* strain and different from the *N. glauca* cT-DNA. A fifth cT-DNA (TE) was discovered in *N. otophora* (*Tomentosae* section); its structure has not yet been assembled. The unexpected presence of related genes located on different cT-DNAs (such as the three *orf14* genes of TA, TB, and TD in *N. tabacum*) implied that phylogenetic analysis of partial cT-DNA sequences from different species

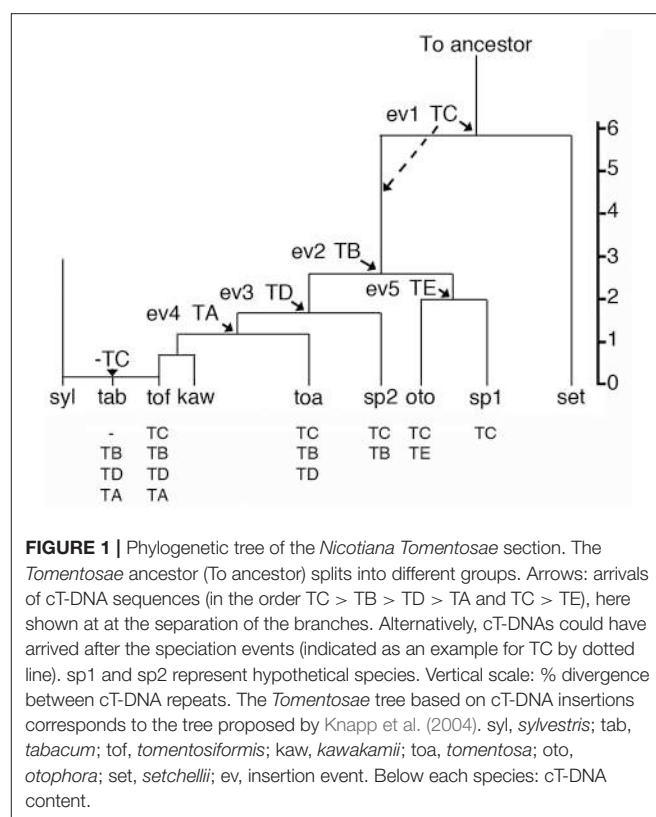


FIGURE 1 | Phylogenetic tree of the *Nicotiana Tomentosae* section. The *Tomentosae* ancestor (To ancestor) splits into different groups. Arrows: arrivals of cT-DNA sequences (in the order TC > TB > TD > TA and TC > TE), here shown at the separation of the branches. Alternatively, cT-DNAs could have arrived after the speciation events (indicated as an example for TC by dotted line). sp1 and sp2 represent hypothetical species. Vertical scale: % divergence between cT-DNA repeats. The *Tomentosae* tree based on cT-DNA insertions corresponds to the tree proposed by Knapp et al. (2004). syl, *sylvestris*; tab, *tabacum*; tof, *tomentosiformis*; kaw, *kawakamii*; toa, *tomentosa*; oto, *otophora*; set, *setchellii*; ev, insertion event. Below each species: cT-DNA content.

(Intriери and Buiatti, 2001) can only be carried out after it has been established whether they belong to the same cT-DNA or not.

If one assumes that the four *N. tomentosiformis* inserts were introduced by successive transformations (and did not accumulate through crosses between different transformants), five different types of plants can be expected (Figure 1). In the *Tomentosae* section, the relative order of the insertion events (ev1 to ev4) can be estimated from the divergence values of the cT-DNA repeats (Chen et al., 2014, Table 1). Events 1, 2+3 (probably in the order TB > TD because of the differences in the repeat divergence), and 4 correspond to the introduction of TC, TB+TD, and TA. *N. setchellii* probably lacks a cT-DNA, as shown by the fact that its transcriptome contains no cT-DNA sequences (Long et al., 2016). *N. otophora* has two cT-DNAs (TC and TE, the latter being specific for *N. otophora* and introduced at event 5), *N. tomentosa* three (TC, TB, and TD), *N. kawakamii* and *N. tomentosiformis* four (TC, TB, TD, and TA). *N. tabacum* has three cT-DNAs, but its TC region has been completely deleted (including 1 kb of flanking DNA on each side, Chen et al., 2014). The remarkable loss of TC in *N. tabacum* shows the importance of investigating cT-DNA insertion sites (Chen et al., 2014; Chen, 2016). According to Figure 1, two intermediate *Nicotiana* forms (sp1 and sp2) are lacking in the *Tomentosae* section: one with TC, but without TE, and one with TC and TB (Figure 1). Possibly, they do occur as variants of existing species, as yet undetected species, or became extinct.

USE OF cT-DNA INSERTS AS MARKERS TO RECONSTRUCT NICOTIANA EVOLUTION

Transferred DNA (T-DNA) insertion events provide interesting clues to reconstruct plant evolution. All species with a cT-DNA at the same insertion site derive from a common ancestor in which the original insertion took place. The divergence between the repeats of such shared cT-DNAs should be consistent with the overall genome divergence between the species, but this has still to be tested.

Gemini viruses such as Geminivirus-Related DNA sequence (GRD, Murad et al., 2004) or retrotransposons such as the TS retrotransposons in tobacco (Wenke et al., 2011) can also provide clues for plant evolution. In the case of the *Tomentosae* section, it may be possible to date the different insertion events, since *Nicotiana* evolutionary trees have been calibrated, with an estimated DNA divergence of about 28% per 5 Mio years (Clarkson et al., 2005). The most diverged *Nicotiana* cT-DNA (TC) shows 5.8% divergence between the repeats which leads to an estimated age of 1 Mio years.

cT-DNAS AND EVOLUTION OF IPOMOEA AND LINARIA

In the case of *Ipomoea*, *orf13* sequences (from *Ib*T-DNA2) were detected in *I. batatas* and in *I. trifida* (Kyndt et al., 2015). This suggests that as in *Nicotiana*, cT-DNAs were introduced in an ancestor species and transmitted across speciation events. However, *Ib*T-DNA2 could have been transferred by interspecific hybridization, known to occur between *I. batatas* and *I. trifida* (Rouillier et al., 2013). Whether *Ib*T-DNA1 and *Ib*T-DNA2 were introduced by one or two transformation events is not clear, because both could be derived from a single *Agrobacterium* strain. The origin of the cultivated hexaploid (6x) species *I. batatas* is much debated. Two independent origins have been proposed which led to the socalled Northern and Southern lineages. The 6x genome has probably arisen in two steps, from 2x to 3x or 4x, and then to 6x. Possibly, *I. trifida* contributed to *I. batatas*, but it has also been proposed that *I. batatas* is derived from wild polyploid *I. batatas* plants (Rouillier et al., 2013). The distribution of cT-DNAs within *I. batatas* (both cultivated and wild forms) and *I. trifida* could shed new light on these questions. For *Linaria*, a calculation has been made on the basis of sequence divergence between *orf14-mis* sequences of *L. vulgaris*, *L. dalmatica*, and *L. acutiloba*. Assuming that the *orf14-mis* sequences are located on the same cT-DNA insert, the insert was estimated to be 1 Mio years old (Kovacova et al., 2014).

In none of the known cases, cT-DNA repeat divergence is more than 10% (see Table 1). This may indicate that cT-DNA insertions did not occur earlier than 1.5 Mio years ago. Alternatively, it may be that within this time span, the statistical probability of a complete cT-DNA deletion became sufficiently high, so that more diverged structures had little chance to survive.

TABLE 1 | Sequence divergence between repeats within different cT-DNA structures.

	cT-DNA	Accession	% Divergence	References
<i>N. tomentosiformis</i>	TA	KJ599826	1.2	Chen et al., 2014
	TB	KJ599827	2.6	Chen et al., 2014
	TC	KJ599828	5.8	Chen et al., 2014
	TD	KJ599829	1.7	Chen et al., 2014
<i>N. glauca</i>	Ng	AB071334.1 AB071335.1	3	Suzuki et al., 2002
<i>Ipomoea batatas</i>	<i>Ib</i> T-DNA1	KM052616.1 KM113766.1	0.9	Kyndt et al., 2015
	<i>Ib</i> T-DNA2	KM052617.1	0.7	Kyndt et al., 2015
<i>Linaria vulgaris</i>	T-DNA	EU735069	8.5	Matveeva et al., 2012

COULD cT-DNA INSERTIONS LEAD TO PLANT SPECIATION ?

It has been proposed that cT-DNA insertions may have led to new species (Martin-Tanguy et al., 1996; Fründt et al., 1998a; Chen et al., 2014). In the case of the *Nicotiana Tomentosae* section different cT-DNA combinations were found in different species, and the order of cT-DNA entry corresponds to the proposed branching order of the species (Knapp et al., 2004; Chen et al., 2014, Figure 1). This pattern is consistent with the idea of speciation by transformation. Speciation could be favored by the strong effects of *A. rhizogenes* T-DNA genes on development (for example by changing flower morphology or flowering time), but this has not been investigated for natural or artificial HR transformants. The speciation hypothesis can be tested by comparing normal plants with their HR transformants obtained from *A. rhizogenes* infection under laboratory conditions. If indeed HR plants no longer hybridize with the ancestor and therefore have become new species, further studies could be carried out to identify the T-DNA genes responsible for introducing the change that leads to the reproductive barrier. Alternatively, cT-DNA sequences of natural transformants may be removed by CRISPR and the resulting plants compared with the unmodified natural transformant. However, the function of those genes that led to a reproductive barrier at an early stage might have been lost in later steps.

In the next section we will investigate in more detail what is known about the structures of cT-DNAs and their evolution.

STRUCTURAL ORGANIZATION OF cT-DNAs

In 8 out of 9 cases, cT-DNA structures are partial inverted repeats, inserted in a single site. The *Linaria* cT-DNA is an exception, being a partial direct repeat (Matveeva et al., 2012). In Figures 2A–D the four *N. tomentosiformis* cT-DNAs (TA, TB, TC, and TD) are shown with the original contigs constructed from small reads obtained by deep sequencing. Highly similar

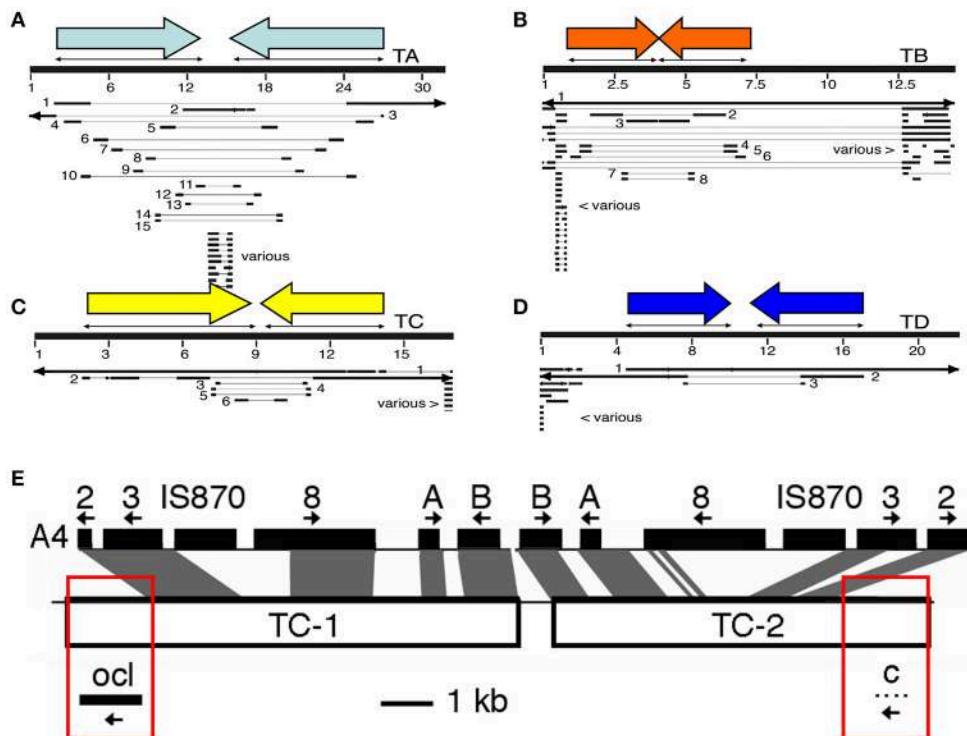


FIGURE 2 | Maps of the four cT-DNAs from *N. tomentosiformis*. **(A–D)**: TA, TB, TC, and TD. Inverted repeats are indicated by colored arrows. Renumbered contigs from *N. tomentosiformis* (AWOL series, Sierro et al., 2014) are mapped on the final cT-DNA maps. The large numbers of contigs are due to difficulties in automatic assembly for these closely related inverted repeats. Various: repeated plant sequences, only part of these contigs are shown. **(E)**: An example of a cT-DNA, the TC region from *N. tomentosiformis*. The figure shows the similarity between the repeated part of TC and part of the *A. rhizogenes* A4 region with its T-DNA genes (A, B, C; *rolA*, *rolB*, *rolC*). Boxed in red are sequences (*ocl*: octopine synthase-like, and gene *c*) that have so far only been found in *A. tumefaciens* or *A. vitis*. Thus, the *A. rhizogenes* strain that inserted the TC-region belongs to a new type of strain that combines features of *A. rhizogenes* and *A. tumefaciens* or *A. vitis*.

repeats can cause problems for the assembly of reads into contigs. This leads to many small contigs which must be linked by PCR amplification and sequencing. In **Figures 2A–D** the published *N. tomentosiformis* contigs (Sierro et al., 2014, AWOL series, renumbered) are shown aligned with the four assembled cT-DNA sequences. The TC region is shown in more detail (**Figure 2E**). The inverted repeat of TC partially aligns with TL from *A. rhizogenes* strain A4. At both ends of the repeat unique regions are found with an *ocl* gene on the left and a protein-C gene on the right. The T-DNA that gave rise to the TC region is unknown, and it is unclear how the inverted repeat and the single copy fragments were assembled. Further progress may require identification of *A. rhizogenes* strains with the relevant T-DNA genes.

All cT-DNAs seem to be truncated. In experimental infections with present-day *Agrobacterium* strains, T-DNA insertions can occur in different ways: in single sites (with a complete or truncated T-DNA, with direct or inverted repeats, with complete or incomplete repeats) or in multiple sites (with combinations of different structures). Some strains carry two different T-DNAs on their Ti/Ri plasmid, such as the TL and TR regions of *A. rhizogenes* strain A4 (Bouchez and Tourneur, 1991) and can introduce them separately or combined as a single insert. Potentially, this leads to a large variety of cT-DNA structures.

The fact that most natural transgenic plants carry a single insert consisting of a partial inverted cT-DNA repeat is therefore probably not coincidental. No simple hypothesis can be proposed why this is so, but the following factors might be considered. cT-DNA inserts in multiple sites will segregate during sexual propagation, favoring single inserts. Repeat structures are more tolerant to mutations, thus facilitating preservation of important genes. Because T-DNA transfer starts at the right border and proceeds to the left, incomplete T-DNA structures will tend to have intact right borders and break off on the left. Studies on experimentally obtained regenerants or with additional natural transformants may show whether some structures are indeed preferred and what could be the underlying reasons.

In the next section we will discuss cT-DNA evolution and variability.

EVOLUTION OF cT-DNAs

After stable integration, cT-DNAs will evolve through point mutations, insertions, and deletions, in the same way as normal plant DNA. Many cT-DNA genes in natural transgenic plants are interrupted by stop codons or are partially deleted (**Table 2**, see also below). *NgroB* of *N. glauca* is inactive but was converted

TABLE 2 | cT-DNA genes in different natural transformants.

Species	cT-DNA	cT-DNA genes	Intact	Expr.	Biol. activity	References
<i>Nicotiana glauca</i>	T-DNA	<i>rolB</i>	—	+	+ (after restoration)	Aoki and Syono, 1999c
		<i>rolC</i>	+	+	+	Aoki and Syono, 1999c
		<i>orf13</i>	+	+	+	Aoki and Syono, 1999b
		<i>orf14</i>	+	+	+	Aoki and Syono, 1999b
		<i>mis</i>	+	+	+ (in <i>E. coli</i>)	Suzuki et al., 2002
<i>N. tomentosiformis</i> and <i>N. tabacum</i> (tob)	TA	<i>orf8</i>	—	nt	nt	Chen et al., 2014
		<i>rolA</i>	—	nt	nt	Chen et al., 2014
		<i>rolB</i>	—	nt	nt	Chen et al., 2014
		<i>rolC</i>	+	+	+ (tob)	Mohajel-Shoja et al., 2011
		<i>orf13</i>	+	+	+ (tob)	Fründt et al., 1998b
		<i>orf14-like</i>	—	+	nt	Chen et al., 2014
		<i>mis</i>	—	nt	nt	Chen et al., 2014
	TB	<i>orf14</i>	+	nt	nt	Chen et al., 2014
		<i>mis</i>	—	nt	nt	Chen et al., 2014
		<i>ags</i>	—	nt	nt	Chen et al., 2014
		<i>mas1'</i>	—	nt	nt	Chen et al., 2014
		<i>mas2'</i>	+	+	+ (tob)	Chen et al., 2016
	TC	<i>ocs-like</i>	+	nt	—	Chen et al., 2014
		2	—	nt	nt	Chen et al., 2014
		3	—	nt	nt	Chen et al., 2014
		8	—	nt	nt	Chen et al., 2014
		<i>rolA</i>	—	nt	nt	Chen et al., 2014
	TD	<i>rolB</i>	—	nt	nt	Chen et al., 2014
		Gene c	—	nt	nt	Chen et al., 2014
		<i>orf18</i>	—	nt	nt	Chen et al., 2014
		<i>orf14-like</i>	+	nt	nt	Chen et al., 2014
		<i>orf15</i>	+	nt	nt	Chen et al., 2014
	TE	<i>orf511</i>	?	nt	nt	Chen et al., 2014
		<i>vis</i>	?	nt	nt	Chen et al., 2014
		6b	?	nt	nt	Chen et al., 2014
		<i>mas1'</i>	?	nt	nt	Chen et al., 2014
		<i>mas2'</i>	?	nt	nt	Chen et al., 2014
		<i>rolB</i>	?	nt	nt	Chen et al., 2014
		<i>rolC</i>	?	nt	nt	Chen et al., 2014
		<i>orf13</i>	?	nt	nt	Chen et al., 2014
		<i>orf14</i>	?	nt	nt	Chen et al., 2014
		<i>iaaH</i>	?	nt	nt	Chen et al., 2014
<i>Linaria vulgaris</i>	T-DNA	<i>iaaM</i>	?	nt	nt	Chen et al., 2014
		<i>acs</i>	?	nt	nt	Chen et al., 2014
		<i>orf2</i>	—	nt	nt	Matveeva et al., 2012
		<i>orf3</i>	—	nt	nt	Matveeva et al., 2012
		<i>orf8</i>	—	nt	nt	Matveeva et al., 2012
		<i>rolA</i>	—	nt	nt	Matveeva et al., 2012
		<i>rolB</i>	—	—	nt	Matveeva et al., 2012
		<i>rolC</i>	+	—	nt	Matveeva et al., 2012
		<i>orf13</i>	—	—	nt	Matveeva et al., 2012
		<i>orf14</i>	—	—	nt	Matveeva et al., 2012
		<i>mis</i>	—	nt	nt	Matveeva et al., 2012

(Continued)

TABLE 2 | Continued

Species	cT-DNA	cT-DNA genes	Intact	Expr.	Biol. activity	References
<i>L. dalmatica</i>	T-DNA	<i>rolC</i>	+	nt	nt	Matveeva and Lutova, 2014
<i>Ipomoea batatas</i>	<i>Ib</i> T-DNA1	<i>acs</i>	+	+	nt	Kyndt et al., 2015
		Gene c	+	+	nt	Kyndt et al., 2015
		<i>iaaH</i>	+	+	nt	Kyndt et al., 2015
		<i>iaaM</i>	+	+	nt	Kyndt et al., 2015
		<i>orf14</i>	-	nt	nt	Kyndt et al., 2015
	<i>Ib</i> T-DNA2	<i>orf17n</i>	-	nt	nt	Kyndt et al., 2015
		<i>rolB</i> -like	+	+	nt	Kyndt et al., 2015
		<i>orf13</i>	+	+	nt	Kyndt et al., 2015
		<i>orf18/orf17n</i>	+	nt	nt	Kyndt et al., 2015
		<i>orf13</i>	+	nt	nt	Kyndt et al., 2015
<i>Ipomoea trifida</i>	<i>Ib</i> T-DNA2	<i>orf13</i>	+	nt	nt	Kyndt et al., 2015

Adapted from Matveeva and Lutova (2014) No distinction is made between copies on repeats of the same cT-DNA. Since the *orf511* gene from the TD region has no equivalent in the databases, it is unknown whether it is intact. As *N. otophora* contigs and reads have not yet been assembled, it is still unknown whether there are intact cT-DNA gene copies or not in this species. nt, not tested.

to an active form by removal of two stop codons (Aoki, 2004). However, it is not clear whether the active form really corresponds to the original *rolB* gene. As expected, cT-DNA sequence variation can also occur within the same species. In early Southern blot experiments, cT-DNA variants were reported for *N. glauca* (Furner et al., 1986). Among *N. tabacum* cultivars, three TA variants occur (Chen et al., 2014).

cT-DNA evolution in *Ipomoea*, *Linaria*, and *Nicotiana* might be influenced by interspecific hybridization. *I. batatas* hybridizes with *I. trifida* (its closest natural relative) under natural conditions, although probably with low efficiency (Rouillier et al., 2013). The *Ib*T-DNA2 genes of *I. batatas* and *I. trifida* (Kyndt et al., 2015) could have been transferred by interspecific crosses. This could also apply to *L. vulgaris* and *L. dalmatica*, both of which contain cT-DNA sequences (Matveeva and Lutova, 2014) and are known to hybridize (Ward et al., 2009).

Interspecific crosses can have other consequences for cT-DNAs. *N. tabacum* results from an interspecific cross between *N. sylvestris* and *N. tomentosiformis* accompanied by massive genome reorganization (Lim et al., 2004). Whether this reorganization follows certain rules and reproducibly leads to the loss of the TC region, might be investigated with artificial hybrids.

When trying to understand cT-DNA evolution, one needs to reconstruct the original structures. This might be attempted by comparing the sequences of the cT-DNA repeats, both within the same species and between related species, favoring variants which correspond to intact open reading frames, are expressed and show biological activity.

In the next section we will investigate the important question of cT-DNA gene expression and regulation.

cT-DNA EXPRESSION AND REGULATION

Although some studies have described cT-DNA gene expression and regulation, this field is still at its beginning and much remains to be done. Table 2 contains a list of expressed cT-DNA genes. Expression patterns depend on the insertion site and on

the regulatory properties of the promoters. Promoter properties can be measured in different ways, either directly by mRNA analysis, or by using reporter genes. In reporter gene constructs promoters are linked to genes for visible markers, such as β -glucuronidase (GUS, Jefferson, 1987). Although much research has been carried out on T-DNA gene promoters (Maurel et al., 1990; Capone et al., 1991, 1994; Leung et al., 1991; Yokoyama et al., 1994; Di Cola et al., 1997; Hansen et al., 1997; Handayani et al., 2005), these studies should be extended in order to get a more detailed description of tissue-specificity, and to identify the corresponding plant transcription factors. Since T-DNA genes of Ri plasmids are expressed in hairy roots, it can be expected that cT-DNA genes are also expressed in roots. However, the properties of their promoters could have evolved, especially if expression in other plant parts would provide some selective advantage. Expression studies show that several cT-DNA genes have maintained their expression patterns in natural transgenic plants (Table 2). How and why T-DNA/cT-DNA genes are regulated the way they are, will need more research on T-DNA/cT-DNA function in hairy roots and natural transformants. It will be important to study those promoter properties in the right context. *A. rhizogenes* T-DNA reporter genes have rarely been studied in hairy roots. Likewise, cT-DNA promoters should be studied in the corresponding natural transformants. However, there is a danger that promoter constructs interfere with the expression of the genes from which they are derived, either by gene silencing or by competing for transcription factors.

The expression of *N. glauca* cT-DNA genes have received special attention because of their possible role in tumor formation. Interspecific hybridization between *N. glauca* and *N. langsdorffii* leads to so-called GGLL plants that spontaneously form tumors. It has been proposed that the *N. glauca* cT-DNA genes play a role in the abnormal growth of these tumors. Expression of Ngorf13 and Ngorf14 (Aoki et al., 1994; Udagawa et al., 2004), and NgrolB and NgrolC (Nagata et al., 1995, 1996) is enhanced in tumor tissues, possibly by a kind of inverted gene dosage effect

(Martin-Tanguy et al., 1996). Up to now it has not been demonstrated that *N. glauca* T-DNA genes are indeed required for tumorous growth. For this they will need to be silenced or removed.

Another cT-DNA gene regulation study involved the TB-*mas2'* gene of *N. tomentosiformis* and *N. tabacum*. Most tobacco cultivars and their paternal ancestor *N. tomentosiformis* have low TB-*mas2'* expression levels (LE cultivars), but a few show high expression levels (HE cultivars). HE cultivars do indeed produce the expected *mas2'* product desoxyfructosylglutamine (DFG) and are the only known cases so far of natural transformants which synthesize opines (Chen et al., 2016). The TB-*mas2'* promoter sequences from HE and LE cultivars are identical, and P*mas2'-GUS* constructs are highly expressed in *N. benthamiana* roots, suggesting that TB-*mas2'* can be silenced and re-activated. Silenced tobacco lines carrying artificially introduced *mas* genes could be re-activated by 5-azacytidine (Van Slogteren et al., 1984), but this was not the case for TB-*mas2'* in LE cultivars (Chen et al., 2016). Mendelian inheritance of the LE/HE phenotype (Chen et al., 2016) suggested that activation and silencing of TB-*mas2'* are due to a *cis* element linked to the TB insert.

Once it is established that cT-DNA genes are actively transcribed in natural transformants it will be necessary to investigate their influence on plant growth and metabolism.

ROLE AND ACTIVITY OF GROWTH-MODIFYING GENES IN NATURAL TRANSFORMANTS

The most interesting question concerning natural *Agrobacterium* transformants is undoubtedly whether they are mere accidents of evolution (by-products of hairy roots as it were, without any selective advantage), or whether cT-DNA integration led to new plant types with particular advantages compared to the non-transformed ancestors (Tepfer, 1983; Meyer et al., 1995). Since at least some natural transformants produce opines, they could also be of advantage to agrobacteria, without special advantages to the plants (sse below).

At the moment of writing, no direct evidence exists for a particular role for any of the cT-DNA genes within their normal context. However, some indirect arguments clearly indicate that they could influence the growth of natural transformants. The T-DNAs from *A. rhizogenes* carry genes known to induce hairy roots and these roots can be regenerated into plants with characteristic phenotypes, called the hairy root or HR phenotype. HR plants generally have a short stature with short internodes and wrinkled leaves (Tepfer, 1990; Christey, 2001; Lütken et al., 2012). Enhanced root growth could possibly improve survival under dry conditions. Among the *A. rhizogenes* T-DNA genes, the «root locus» (*rol*) genes *rolA*, *rolB*, *rolC*, and *rolD* influence hairy root induction on *Kalanchoe daigremontiana* leaves (White et al., 1985), and *rolA*, *rolB*, and *rolC* are sufficient to induce roots on several species. The *rolB* and *rolC* genes belong to the *plast* gene family, a large family of mostly T-DNA-located genes which includes *orf13*, *orf14*, *6a*, and *6b* (Levesque et al.,

1988; Studholme et al., 2005). *rolB* has a more general meristem-inducing activity (Altamura et al., 1994; Koltunow et al., 2001). In addition, *rolB* induces necrosis in tobacco leaves (Schmülling et al., 1988; Mohajel-Shoja, 2010). *orf13* has been considered to be non-essential for root induction although capable of stimulating HR induction by *rolABC* genes (Cardarelli et al., 1987; Capone et al., 1989; Aoki and Syono, 1999a). However, *orf13* expression in tobacco (Hansen et al., 1993; Lemcke and Schmülling, 1998), tomato (Stieger et al., 2004), and *Arabidopsis* (Kodahl et al., 2016) led to various growth changes up to extreme dwarfism in *Arabidopsis* (Kodahl et al., 2016). The *rolA* gene has strong morphogenetic effects (Dehio and Schell, 1993; Guivarc'h et al., 1996). Thus, expression of *rol* genes and *orf13* in natural transformants can be expected to influence their growth.

Linaria, *Ipomoea*, and *N. otophora* contain *iaaH* and *iaaM* genes. Together these encode indole acetic acid synthesis and could have been active in early stages of transformation. It is noteworthy that the *iaaM* and *orf8* (Lemcke et al., 2000) T-DNA genes carry a *rolB*-like part at the 5' end and a bacterial *iaaM* part at the 3' end (Levesque et al., 1988). Both can be separated and retain their function (Otten and Helfer, 2001; Umber et al., 2002, 2005). Thus, an intact *rolB* part in an otherwise mutated *orf8* or *iaaM* gene might still influence the growth of natural transformants.

Ngorf13, *NgrolC*, *trolC*, and *torf13* are expressed in the corresponding *Nicotiana* species. When overexpressed in tobacco, *Ngorf13* leads to dark-green rounded leaves (Aoki and Syono, 1999b), *NgrolC* (Aoki and Syono, 1999c), and *trolC* (Mohajel-Shoja et al., 2011) to a dwarf phenotype and lanceolate, pale leaves, whereas *torf13* induces green callus on carrot disks (Fründt et al., 1998b). In natural transformants, *rolC*, *orf13*, *orf14* are frequently intact (Table 2).

It is generally assumed that each type of T-DNA/cT-DNA gene has a specific effect, so that a cT-DNA-located *rolC* gene will have the same activity as a T-DNA-located *rolC* gene. However, variants of a given gene type can encode different biological activities. The *rolB* genes from 1,855 and 2,659 are less dependent on auxin for root induction on carrot disks as *rolB* from A4 (Schmülling et al., 1993; Serino et al., 1994). Six different *6b* genes from *A. tumefaciens* and *A. vitis* differ in their capacity to induce tumors (Helfer et al., 2002). Thus, functional differences between a cT-DNA gene and a related T-DNA gene (as noted by Aoki and Syono, 2000) might result from differences between the model strain and the strain that introduced the cT-DNA, rather than from divergent evolution after transfer to the plant.

The oldest cT-DNA (from *Linaria*) has lost all open reading frames except *LvrolC*, suggesting positive selection of this gene. Inactivation of the *rolC*, *orf13*, and *orf14* genes in various natural transformants are obvious targets for the future.

It is possible that some (or even most) cT-DNA genes only played a role in the initial transformation/regeneration event, by allowing HR regeneration and the establishment of a new species (see above). After that, they could have lost their function either because of detrimental effects (like dwarfing by *rolA* or *orf13*, or necrosis by *rolB*) or because they were selectively neutral. In that case cT-DNA gene inactivation would show no effects and could lead to the wrong conclusion that these genes had no function.

in the evolution of the natural transformants. If cT-DNA genes induce significant morphological changes in other plants upon strong and constitutive expression, their expression in natural transgenics will probably also lead to changes, although these might be more restricted.

In the case of the widely cultivated tobacco and sweet potato, cT-DNA structures and expression patterns could have been subjected to selection during domestication. This hypothesis can be tested by careful comparison between certain cultivars and their isogenic cT-DNA mutants.

ROLE OF OPINE SYNTHESIS GENES IN NATURAL TRANSFORMANTS

T-DNA/cT-DNA regions generally contain opine genes. Opines are conjugation products of common metabolites such as amino acids, α -keto acids, and sugars, and cannot be metabolized by plants. Often, opine enzymes use multiple substrates (as in the case of lysopine dehydrogenase, Otten et al., 1977) thereby potentially sequestering a large amount of metabolites which might affect plant growth. Thus, it is important to know where T-DNA/cT-DNA opine genes are expressed, and to what extent they are regulated. The *rolD* gene strongly inhibits growth of transgenic carrot (Limami et al., 1998). In tomato, it does not affect morphology (the reason for the difference with carrot is unknown), but flowering occurs earlier with increased numbers of flowers and fruits (Bettini et al., 2003). Opines in crown galls and hairy roots are assumed to be secreted, in order to make them available to the agrobacteria, but this important process has not been studied in detail. It is unknown whether there are specific mechanisms for opine secretion, and whether T-DNA/cT-DNA genes play a role in this. It has been proposed that the *A. tumefaciens* 6a gene (a member of the plast gene family) stimulates secretion of octopine and nopaline (Messens et al., 1985), but unfortunately this interesting study has not been followed up.

Additional genes such as gene *c* and *orf511* (coding for a large, 511 amino acid protein) also remain to be studied. Gene *c* from *A. tumefaciens* strain C58 has shoot-inducing properties (Otten et al., 1999). Interestingly, it is also found in organisms other than plants (see below).

The morphological effects of various cT-DNA genes (expressed to different extents in different tissues) add up in complex ways. For example, *rolA* and *rolB* gene are antagonistic in tomato (Van Altvorst et al., 1992). *rolA*, *rolB*, and *rolC* (Spena et al., 1987), and *rolB*, *rolC*, *orf13*, and *orf14* act synergistically (Nilsson and Olsson, 1997; Aoki and Syono, 1999b). It will therefore be a particularly challenging task to establish the contribution of each gene in the context of their combined expression in natural transformants. In addition, two *Agrobacterium* T-DNA genes which are also found in natural transformants, can produce growth effects at a distance: *orf13* (Hansen et al., 1993) and *6b* (Helfer et al., 2003). This means that their effects might extend beyond their domains of expression.

Apart from changing plant growth, cT-DNA gene expression may confer immunity to *Agrobacterium* by silencing incoming T-DNA (for an experimental example of such T-DNA silencing, see Escobar et al., 2001). However, in the *Tomentosae* section agrobacteria were able to re-infect already transformed species, arguing against this possibility.

We will now investigate the question whether cT-DNA gene expression in natural transformants could influence the growth and evolution of *Agrobacterium*.

DOES AGROBACTERIUM BENEFIT FROM NATURAL TRANSGENIC PLANTS ?

Natural transformants which synthesize opines could influence the growth and evolution of *Agrobacterium* (Chen et al., 2016). In HE tobacco cultivars (see above) TB-mas2' is expressed at high levels in root tips, and leads to production of significant amounts of DFG, a well-known opine (Chen et al., 2016). DFG can be used by agrobacteria and other microbes (Moore et al., 1997; Baek et al., 2005), but it has not yet been tested whether the DFG of HE cultivars is secreted and whether it might accumulate in the rhizosphere. Studies on artificial symbiosis based on opine utilization (Guyon et al., 1993; Dessaux et al., 1998; Savka et al., 2002; Mondy et al., 2014) provide experimental models to test this idea. Controlled inoculation of HE cultivars and isogenic CRISPR mutants with DFG-metabolizing and non-metabolizing *Agrobacterium* mutants could show whether DFG production by HE cultivars confers a selective advantage on DFG-using bacteria. If so, this could have some interesting implications. It has been postulated that the genetic modification of plant cells allows *Agrobacterium* to take control of its host, by re-directing its growth and metabolism to its own benefit. This process has been called \ll genetic colonization \gg (Schell et al., 1979). If it could be shown that opine production by HE plants favors *Agrobacterium* growth it would take the genetic colonization theory one step further. In that case the role of the pRi plasmid is not only (or even not at all) to induce hairy roots, but to create transgenic plants. Such plants could provide a genetically stable and much increased opine production, as compared to opine synthesis by relatively small numbers of non-permanent hairy roots growing from infected plants. If *Agrobacterium* benefits from opine production by natural transformants, hairy roots might be considered as mere intermediates on the way to transgenic plants. Opine production might be detrimental to plant growth, but reproductive isolation of the initial transformants could ensure their survival. Subsequently, cT-DNA functions might be selected against and growth might revert to normal. Thus, natural transformants could be transient plant species with various levels of genetic stability.

So far, it is not known how much *A. rhizogenes* benefits from opines produced in hairy roots growing in nature. Opine sources can attract Agrobacteria (Kim and Farrand, 1998) *in vitro*, but do agrobacteria also accumulate and multiply on hairy roots or on roots of natural transformants? What are

the dynamics of these interactions? Do the bacteria concentrate around areas of highest production? Are opines stable in soil and do they accumulate over time? Do the modified growth properties of hairy roots increase opine production or secretion (for example by stimulating lateral root formation)? Experimental HR induction is generally done by infecting stems in the greenhouse or leaf disks *in vitro*, and the hairy roots develop in agar or in air. It would be interesting to know how hairy roots grow in soil and whether their growth is favored over that of normal roots. All these questions merit attention when one considers the effects of opine-producing plants on agrobacteria.

Apart from TB-*mas2'*, other opine synthesis enzymes (encoded by *acs*, *vis*, *ocl*, *mis*, *rolD*) should be investigated for their opine synthesis properties. Different forms with different substrate preferences may exist, as in the case of octopine dehydrogenase (Ocs, Otten and Szegedi, 1985).

Unusual growth characteristics of hairy roots and HR-derived plants could stimulate growth of agrobacteria independently from opines, for example if some T-DNA genes favor secretion of common root metabolites. When exploring the structure, expression and biological function of cT-DNA genes, it should be realized that some of these genes could have played a role in the first steps of the transformation/regeneration processes and that these events are still unknown. In the next section we will therefore look at a possible scenario for the evolutionary origin of natural transformants.

A SCENARIO FOR THE ORIGIN OF NATURAL TRANSFORMANTS

The details of the origin of natural transformants are still unclear. Different types of *Agrobacterium* strains with different T-DNAs were involved, as mentioned above. These could have induced different types of hairy roots, depending on their cT-DNA genes. In general, it is assumed that individual hairy roots represent clones growing from a single transformed cell (Tepfer, 1984; McKnight et al., 1987). A particular *A. rhizogenes* strain may induce hairy roots with different T-DNA structures (complete or incomplete) and different gene expression levels depending on the insertion sites, which probably leads to different types of roots. It is often assumed that hairy roots represent a single, well-defined type of roots, but this seems highly unlikely in view of the many combinations of T-DNA genes and expression levels expected to occur in individual hairy root clones. The occurrence of different agrobacteria strains, each with their own combination of T-DNA genes, increases the problem of HR variability. *A. rhizogenes*-induced roots have not yet been systematically investigated in terms of growth rate, cell division, elongation, differentiation, and root branching patterns. Plants regenerated from HR have not only modified roots, but also aberrant, wrinkled leaves and stunted growth. The conspicuously wrinkled leaves of HR plants have not yet been analyzed at the developmental level. Possibly they result from changes in

vascular development. We suspect that a whole gradient of HR phenotypes may exist and that the expression \ll hairy root phenotype \gg is an oversimplification. Detailed cellular analysis of HR plants carrying T-DNA genes with inducible promoters will be of great use to understand how T-DNA genes affect growth (for an example using the *6b* gene, see Pasternak et al., 2017).

In the case of the natural transformants, there could have been a selection for HR types with T-DNA gene combinations that allowed plant regeneration. Some genes could be detrimental to regeneration (possibly *rolA*: inhibition of flowering, Martin-Tanguy et al., 1996; and *rolB*: necrosis, Schmülling et al., 1988), whereas others might favor this process.

In the case of the *Toomentosae* section, plants carrying the first cT-DNA (TC, carrying *rolA* and *rolB* genes) may have acquired a better regeneration capacity compared to the non-transformed ancestor. Thus, when TC-carrying plants were infected with another *A. rhizogenes* strain, the resulting hairy roots (carrying TC and TB) could more easily regenerate into plants, and the process could repeat itself several times. Tobacco plants transformed by *A. rhizogenes* A4 spontaneously formed shoots from roots when grown in pots, contrary to normal tobacco (Tepfer, 1984). We need more research on the shoot regeneration properties of hairy roots in different species, the role of the different T-DNA genes in this process, and the underlying molecular mechanisms.

When considering the origin of natural transformants, it is worth noting that *A. tumefaciens* nopaline strains T37 and C58 (Yang and Simpson, 1981) or 82.139 (Drevet et al., 1994) can induce abnormal shoots (called shooty teratomas, Figures 3c,d). These are due to expression of the T-DNA-located isopentenyltransferase (*ipt*) gene, but shoot growth is probably also influenced by other T-DNA genes. It would be worth investigating whether teratomas could lead to rooting plants under natural conditions and eventually give rise to natural transformants.

Some plant species may have special regeneration abilities, so that hairy roots induced on such plants could easily produce fertile plants. *Linaria* carries buds on its roots, which may greatly facilitate plant regeneration from hairy roots (Figures 3a,b). *L. vulgaris* (but not *L. maroccana*) internode fragments easily form shoots and callus *in vitro*, even on hormone-free medium (Matveeva et al., 2012). It remains to be shown whether this is an intrinsic property of some *Linaria* species or due to cT-DNA genes. *I. batatas* shoot fragments (called slips) easily form roots, whereas root pieces carry dormant buds which easily produce plants (George et al., 2011). Re-transformation events may be favored if opine-producing plants attract agrobacteria. These could then introduce additional cT-DNAs (Chen et al., 2014).

In order to definitely establish themselves, the new transgenic plants had to transmit their cT-DNA to their progeny and reproduce successfully in the same environment as the ancestors. It is questionable whether a presumably very rare natural transgenic plant could have survived without reproductive isolation (sympatric speciation, see below). Later, the need for reproductive isolation might have

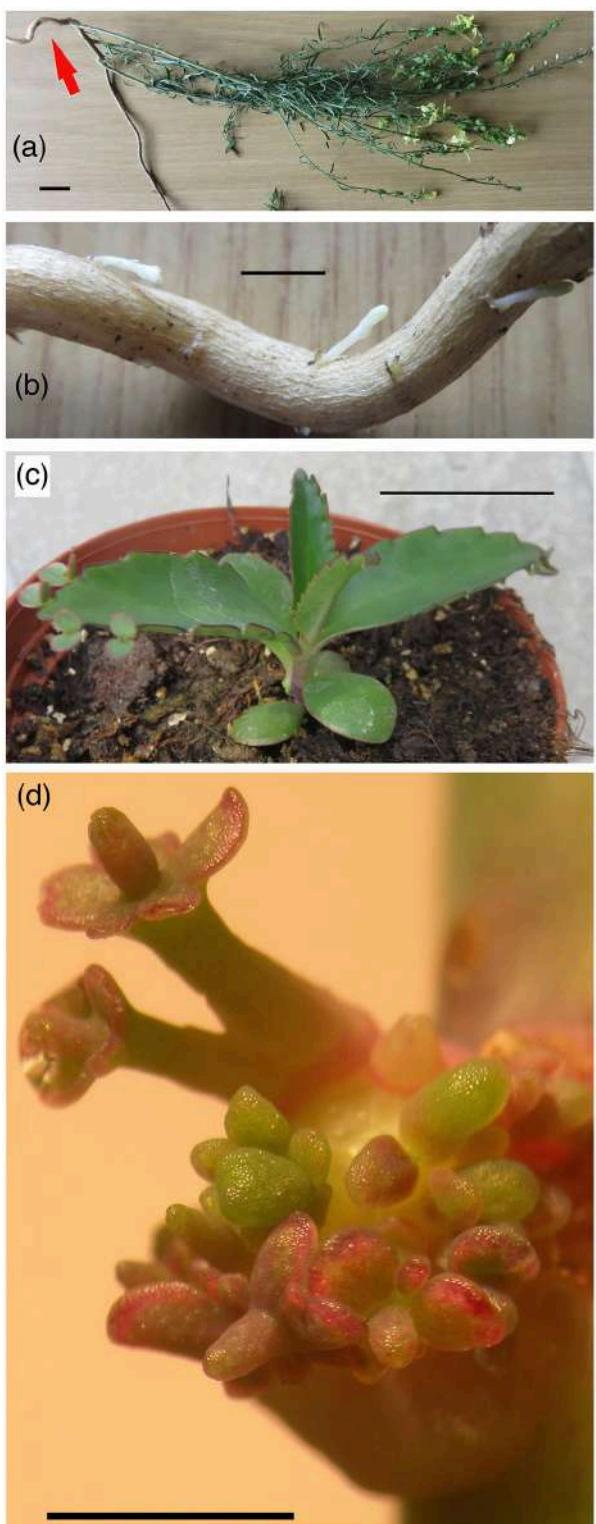


FIGURE 3 | Regeneration of buds from *Linaria vulgaris* roots and of shoots from *Kalanchoe daigremontiana* tumors. **(a)** *L. vulgaris*, overview. Scale: 2 cm. **(b)** Detail buds. Scale 5 mm. **(c)** Normal *K. daigremontiana* plantlet. Scale: 3 cm. **(d)** Teratoma formation on *K. daigremontiana* stems infected with *A. tumefaciens* strain Tm4. The *Kalanchoe* teratoma structures are abnormal, but structured. Scale: 1 cm.

disappeared, when sufficient differences had accumulated to prevent hybridization with the ancestral species. This could have led to the counterselection of cT-DNA genes that were important for speciation, especially if they reduced growth and reproduction. Selection to reduce negative cT-DNA effects could also have occurred elsewhere in the plant genome.

It is often assumed that natural transformants are homozygous for cT-DNA sequences, but it is possible that different cT-DNA gene alleles occur in natural populations (for intraspecific cT-DNA variants, see above). Selectively neutral genes would gradually be eroded and finally disappear. In extreme cases, the complete insert could be lost, as observed for the *N. tabacum* TC-DNA. TB-mas2' seems to have been silenced in *N. tomentosiformis* and subsequently re-activated in some *N. tabacum* cultivars (Chen et al., 2016) which might constitute a case of evolutionary « reversion ».

Thus, to ensure the transition from a hairy root clone to the many successful populations of present-day natural transformants such as *Nicotiana glauca* or *Linaria vulgaris*, many steps might have been necessary. For a summary of these steps, see **Figure 4A**. The following section suggests some experiments to investigate this scheme (summarized in **Figure 4B**).

EXPERIMENTAL EVIDENCE FOR EVOLUTIONARY SCENARIOS

What kind of experimental evidence could lend support to theoretical evolutionary scenarios as described above? It seems impossible to reconstruct the exact transformation events and the subsequent evolution leading to present-day natural transformants. However, if similar events still occur in nature, one might learn more about them. In the case of the natural *Nicotiana* transformants, it could be investigated whether *Nicotiana* species from the *Tomentosae* or *Noctiflorae* section are infected by *A. rhizogenes* in their natural South-American environment, and one could try to isolate and characterize *A. rhizogenes* strains from their rhizosphere.

The next question concerns the capacity of hairy roots to spontaneously produce transgenic plants under natural conditions. This may be studied by challenging different plant species with different *A. rhizogenes* strains under controlled conditions, preferably using plants growing in soil. Regeneration of plants from hairy roots under laboratory conditions has been reported for 53 plant species (Christey, 2001). However, nothing is known about conditions that favor regeneration in nature, such as climate, humidity, age of the plant, type of soil, type of wounding, or site of infection. Starting with a system of robust HR induction on plants growing in soil, it might be possible to study plant regeneration from such roots. Several ornamental plants have been transformed with natural *A. rhizogenes* strains in order to obtain dwarfed forms, a desirable trait in horticulture (Lütken et al., 2012). Such applied HR research could address several of the questions raised above (HR types, effects of cT-DNA genes, regeneration capacity, reproductive isolation). A significant potential exists for plant improvement using A.

A Questions on the origin and evolution of natural transformants	B Experimental approaches to study questions raised in A.
What are likely target plant species?	Study plant species for their HR regeneration properties
What kind of <i>Agrobacterium</i> strains are involved?	Isolate and compare large numbers of <i>A. rhizogenes</i> strains for HR induction and HR properties
Can different types of HR clones arise from one infection?	Study growth properties of HR clones derived from infections with one or more <i>Agrobacterium</i> -plant combinations
How do plants regenerate (and form new species) from a HR clone?	Study role of cT-DNA genes in regeneration from HR clones. Study sexual compatibility between regenerants and ancestor plants
How do cT-DNA structures and their expression evolve?	Compare structure and expression of a particular cT-DNA insertion event across different species, compare genes from repeat structures
How do cT-DNA variants segregate in populations?	Search for plants with heterozygous cT-DNA structures and analyze segregation patterns
Can cT-DNAs be completely lost and what traces do they leave?	Search for natural variants that have lost their cT-DNA

FIGURE 4 | Theoretical steps in the origin of natural transformants. **(A)** Questions on the origin and evolution of natural transformants. **(B)** Experimental approaches to study the questions raised in A.

rhizogenes T-DNA genes (Christey, 2001; Casanova et al., 2005; Guillou et al., 2006) which probably also applies to cT-DNA genes.

In order to study possible ancestor phenotypes, cT-DNA genes might be silenced or removed by CRISPR. Compared to the CRISPR approach, silencing may have an interesting advantage: placed under control of an inducible promoter, a silencing construct could reduce expression of a target gene to different levels and in a spatially and temporally controlled way.

Naturally transformed plants have so far been found in the genus *Nicotiana*, *Linaria*, and *Ipomoea*. In the next part we will discuss how to search for additional transformants.

SEARCH FOR ADDITIONAL NATURAL TRANSFORMANTS

In order to search for natural transformants, three approaches can be used. First, deep sequencing of many plant species is yielding vast numbers of DNA sequences, both from genomic DNA and from transcriptomes. These sequences can be regularly analyzed for T-DNA-like sequences by automatic search robots. The cT-DNAs of the *Nicotiana* group have revealed the presence of genes that were thought to be specific for *A. tumefaciens* or *A. vitis* (6b, ocl, vis, Chen et al., 2014). Therefore, query sequences should not only include all known *A. rhizogenes* T-DNA sequences, but *A. tumefaciens* and *A. vitis* T-DNAs as well. In order to increase the chance of finding sequences with weak homology to model sequences, nucleotide data bases can be interrogated with protein query sequences (NCBI, tblastn search).

Second, plant species with close affinity to natural transformants or different accessions of the same species should be investigated, in order to define the distribution limits of the cT-DNA sequences within a group of species, and to explore their structural and functional variability.

Third, species that easily form plants from root fragments, have wrinkled leaves or other HR characteristics, might be candidates and could be tested by PCR experiments or deep sequencing.

We believe that the search for cT-DNA sequences should not be limited to plants. The capacity of *Agrobacterium* to introduce T-DNA genes into fungi under laboratory conditions has been well documented (de Groot et al., 1998; Michielse et al., 2008). It seems possible that this also occurs in nature, especially in the mycosphere (Zhang et al., 2014). Regeneration of transformed cells might be easy in such organisms, since single cells can be transformed. No *bona fide* cT-DNA sequences have yet been found outside the plant world. However, protein searches led to the discovery of several T-DNA-like protein sequences in fungi (Mohajjal-Shoja et al., 2011; Chen et al., 2014). Among these, opine enzyme-like sequences were found in *Nectria hematococca* (Acs), *Aspergillus nidulans* (Ocl), and Sus-like proteins are relatively widespread in various fungi. Plast proteins were detected in *Laccaria bicolor*. Protein C sequences were found in *Melampsora larici-populina* and *Pestalotiopsis fici* (Chen et al., 2014). These fungal T-DNA-like sequences are more divergent with respect to known T-DNA sequences than the plant cT-DNA plast sequences (Table 3) and could be derived from other types of *Agrobacterium* strains. Their patchy distribution among fungi argues in favor of horizontal gene transfer. Some fungi (such as *Pestalotiopsis* and *Melampsora*) contain several T-DNA-like genes. If such genes are grouped (as expected in the case of T-DNA transfer), this would provide an argument for ancient T-DNA transfer. Further investigations should concentrate on the chromosomal sequences around these genes and their comparison with relatives lacking such genes. Finally, it will be important to investigate their expression and function.

CONCLUSIONS

Natural *Agrobacterium* transformants represent special cases of horizontal gene transfer, as they result from a highly adapted process aimed at the transfer and insertion of functional genes in plants. The bacteria responsible for the insertion of the cT-DNAs were probably related to *A. rhizogenes*. The natural variability of this bacterium and the capacity of various *A. rhizogenes* types to induce hairy roots in nature (and not only under laboratory conditions), both on aerial parts and in soil, is still largely

TABLE 3 | T-DNA-like protein sequences in fungi.

Protein category	Organism	Accession number	Closest T-DNA relative	Accession number of closest relative	% Identity
Opine enzymes	<i>Nectria haematococca</i> , <i>Fusarium oxysporum</i>	XP_003047010.1	ChsA (not on T-DNA), Ags	AAK08601.1	60
		XP_018252422.1		P27875.1	42
	<i>Aspergillus niger</i> and many others	EHA20957.1	TC-Ocl	XP_009611266	40
	<i>Melampsora larici-populina</i>	EGG11641.1	TC-Ocl	XP_009611266	34
Plast proteins	<i>Laccaria bicolor</i>	XP_001884962	Protein 5 (Tm4)	AAB41873	21
	<i>Laccaria bicolor</i>	XP_001884963	Protein 5 (Tm4)	AAB41873	20
	<i>Laccaria bicolor</i>	XP_001884964 (409aa, first part)	Protein 5 (Tm4)	AAB41873	19
	<i>Laccaria bicolor</i>	XP_001884861 (451aa, last part)	Protein 5 (Tm4)	AAB41873	22
	<i>Laccaria bicolor</i>	XP_001881215 (491 aa)	C' protein	NP_053417.1	19
C protein	<i>Melampsora larici-populina</i>	EGG11381.1	C (C58)	AAD30491.1	30
	<i>Pestalotiopsis fici</i>	XP_007840635.1 (540aa)	C (C58)	AAD30491.1	32

unexplored. Spontaneous regeneration of natural hairy roots may depend on the properties of the non-transformed hosts, but probably also involves cT-DNA genes. More studies are required on the function and molecular mechanism of the T-DNA genes, in order to explain how and why natural transformants differ from their ancestors, and how they managed to establish themselves. An important direction for future research will be the removal or silencing of cT-DNA genes. The *plast* genes, opine genes, *rolA*, gene *c*, and *orf511* all require detailed analysis by themselves. Opine synthesis by natural transformants and its potential to favor *Agrobacterium* growth should be investigated under natural conditions, and should include studies on the influence of opine synthesis on plant metabolism, and on the mechanisms and specificities of opine secretion. The *plast* genes constitute an especially challenging subject as 30 years of research have not been able to convincingly reveal their basic function. They seem to be involved in the transport of plant metabolites and in the induction of abnormal growth. Studies on cell division

and differentiation of various types of hairy roots and HR plants will be essential to understand how T-DNA/cT-DNA genes redirect the growth of roots and other plant organs. In view of their strong morphogenetic activities, both T-DNA and cT-DNA genes may be used for applications in horticulture and agriculture. Such research would undoubtedly benefit from a better understanding of their functions.

AUTHOR CONTRIBUTIONS

LO wrote the basic structure of the paper. KC participated in writing and correcting the paper.

ACKNOWLEDGMENTS

KC was supported by doctoral grant 2011679003 from the Chinese Scholarship Council. We thank Sonia Sokornova for correcting the manuscript.

REFERENCES

- Altamura, M. M., Capitani, F., Gazza, L., Capone, I., and Costantino, P. (1994). The plant oncogene *rolB* stimulates the formation of flower and root meristemoids in tobacco thin cell layers. *New Phytol.* 126, 283–293. doi: 10.1111/j.1469-8137.1994.tb03947.x
- Aoki, S. (2004). Resurrection of an ancestral gene: functional and evolutionary analyses of the Ngrol genes transferred from *Agrobacterium* to *Nicotiana*. *j. Plant Res.* 117, 329–337. doi: 10.1007/s10265-004-0163-5
- Aoki, S., Kawaoka, A., Sekine, M., Ichikawa, T., Fujita, T., Shinmyo, A., et al. (1994). Sequence of the cellular T-DNA in the untransformed genome of *Nicotiana glauca* that is homologous to ORFs 13 and 14 of the Ri plasmid and analysis of its expression in genetic tumors of *N. glauca* x *N. langsdorffii*. *Mol. Gen. Genet.* 243, 706–710.
- Aoki, S., and Syono, K. (1999a). Synergistic function of *rolB*, *rolC*, ORF13, and ORF14 of TL-DNA of *Agrobacterium rhizogenes* in hairy root induction in *Nicotiana tabacum*. *Plant Cell Physiol.* 40, 252–256. doi: 10.1093/oxfordjournals.pcp.a029535
- Aoki, S., and Syono, K. (1999b). Function of Ngrol genes in the evolution of *Nicotiana glauca*: conservation of the function of NgORF13 and NgORF14 after ancient infection by an *Agrobacterium rhizogenes*-like ancestor. *Plant Cell Physiol.* 40, 222–230. doi: 10.1093/oxfordjournals.pcp.a029531
- Aoki, S., and Syono, K. (1999c). Horizontal gene transfer and mutation: ngrol genes in the genome of *Nicotiana glauca*. *Proc. Natl. Acad. Sc. U.S.A.* 96, 13229–13234. doi: 10.1073/pnas.96.23.13229
- Aoki, S., and Syono, K. (2000). The roles of Riro and Ngrol genes in hairy root induction in *Nicotiana debneyi*. *Plant Sci.* 159, 183–189. doi: 10.1016/S0168-9452(00)00333-2

- Baek, C. H., Farrand, S. K., Park, D. K., Lee, K. E., Hwang, W., and Kim, K. S. (2005). Genes for utilization of deoxyfructosyl glutamine (DFG), an amadori compound, are widely dispersed in the family Rhizobiaceae. *FEMS Microbiol. Ecol.* 53, 221–233. doi: 10.1016/j.femsec.2004.12.008
- Bettini, P., Michelotti, S., Bindu, D., Giannini, R., Capuana, M., and Buiatti, M. (2003). Pleiotropic effect of the insertion of the *Agrobacterium rhizogenes* rolD gene in tomato (*Lycopersicon esculentum* Mill.). *Theor. Appl. Genet.* 107, 831–836. doi: 10.1007/s00122-003-1322-0
- Biemann, K., Lioret, C., Asselineau, K., Lederer, E., and Polonski, J. (1960). Sur la structure chimique de la lysopine, nouvel acide aminé isolé du tissu de crown-gall. *Bull. la Soc. Chim. Biol.* 42, 979–991.
- Binns, A. N., and Costantino, P. (1998). "The *Agrobacterium* oncogenes," in *The Rhizobiaceae: Molecular Biology of Model Plant-Associated Bacteria*, eds H. P. Spaink, A. Kondorosi, and P. J. Hooykaas (Dordrecht: Kluwer Academic Publishers), 251–266.
- Bouchez, D., and Tourneur, J. (1991). Organization and nucleotide sequence of the agropine synthesis region on the T-DNA of the Ri plasmid from *Agrobacterium rhizogenes*. *Plasmid* 25, 27–39. doi: 10.1016/0147-619X(91)90004-G
- Burr, T. J., and Otten, L. (1999). Crown gall of grape: biology and disease management. *Annu. Rev. Phytopathol.* 37, 53–80. doi: 10.1146/annurev.phyto.37.1.53
- Canaday, J., Gérard, J. C., Crouzet, P., and Otten, L. (1992). Organization and functional analysis of three T-DNAs from the vitopine Ti plasmid pTiS4. *Mol. Gen. Genet.* 235, 292–303. doi: 10.1007/BF00279373
- Capone, I., Cardarelli, M., Mariotti, D., Pomponi, M., De Paolis, A., and Costantino, P. (1991). Different promoter regions control level and tissue specificity of expression of *Agrobacterium rhizogenes* rolB gene in plants. *Plant Mol. Biol.* 16, 427–436. doi: 10.1007/BF00023993
- Capone, I., Frugis, G., Costantino, P., and Cardarelli, M. (1994). Expression in different populations of cells of the root meristem is controlled by different domains of the rolB promoter. *Plant Mol. Biol.* 25, 681–691. doi: 10.1007/BF00029606
- Capone, I., Spano, L., Cardarelli, M., Bellincampi, D., Petit, A., and Costantino, P. (1989). Induction and growth properties of carrot roots with different complements of *Agrobacterium rhizogenes* T-DNA. *Plant Mol. Biol.* 13, 43–52. doi: 10.1007/BF00027334
- Cardarelli, M., Mariotti, D., Pomponi, M., Spano, L., Capone, I., and Costantino, P. (1987). *Agrobacterium rhizogenes* T-DNA genes capable of inducing hairy root phenotype. *Mol. Gen. Genet.* 210, 111–115. doi: 10.1007/BF00331152
- Casanova, E., Trillas, M. I., Moysset, L., and Vainstein, A. (2005). Influence of rol genes in floriculture. *Biotechnol. Adv.* 23, 3–39. doi: 10.1016/j.biotechadv.2004.06.002
- Chen, K. (2016). *Sequencing and Functional Analysis of cT-DNAs in Nicotiana*. Thesis of University of Strasbourg.
- Chen, K., Dorlhac de Borne, F., Julio, E., Obszynski, J., Pale, P., and Otten, L. (2016). Root-specific expression of opine genes and opine accumulation in some cultivars of the naturally occurring GMO *Nicotiana tabacum*. *Plant J.* 87, 258–269. doi: 10.1111/tpj.13196
- Chen, K., Dorlhac de Borne, F., Szegedi, E., and Otten, L. (2014). Deep sequencing of the ancestral tobacco species *Nicotiana tomentosiformis* reveals multiple T-DNA inserts and a complex evolutionary history of natural transformation in the genus *Nicotiana*. *Plant J.* 80, 669–682. doi: 10.1111/tpj.12661
- Christey, M. C. (2001). Invited review: use of Ri-mediated transformation for production of transgenic plants. *In Vitro Cell. Dev. Biol. Plant* 37, 687–700. doi: 10.1007/s11627-001-0120-0
- Christie, P. J., and Gordon, J. E. (2014). The *Agrobacterium* Ti Plasmids. *Microbiol. Spectr.* 2, 1–29. doi: 10.1128/microbiolspecPLAS-0010-2013
- Clarkson, J. J., Yoong Lim, K., Kovarik, A., Chase, M. W., Knapp, S., and Leich, A. R. (2005). Long-term genome diploidization in allopolyploid *Nicotiana* section Repandae (Solanaceae). *New Phytol.* 168, 241–252. doi: 10.1111/j.1469-8137.2005.01480.x
- De Cleene, M., and De Ley, J. (1981). The host range of infectious hairy root. *Bot. Rev.* 47, 147–194. doi: 10.1007/BF02868853
- de Groot, M. J. A., Bundock, P., Hooykaas, P. J. J., and Beijersbergen, A. G. M. (1998). *Agrobacterium tumefaciens* mediated transformation of filamentous fungi. *Nat. Biotechnol.* 16, 839–842. doi: 10.1038/nbt0998-839
- Dehio, C., and Schell, J. (1993). Stable expression of a single-copy rolA gene in transgenic *Arabidopsis thaliana* plants allows an exhaustive mutagenic analysis of the transgene-associated phenotype. *Mol. Gen. Genet.* 241, 359–366. doi: 10.1007/BF00284689
- Dessaix, Y., Petit, A., Farrand, S. K., and Murphy, P. J. (1998). "Opines and opine-like molecules involved in plant-Rhizobiaceae interactions," in *The Rhizobiaceae: Molecular Biology of Model Plant-Associated Bacteria*, eds H. P. Spaink, A. Kondorosi, and P. J. J. Hooykaas (Dordrecht: Kluwer Academic Publishers), 173–197.
- Di Cola, A., Poma, A., and Spano, L. (1997). rolB expression pattern in the early stages of carrot somatic embryogenesis. *Cell Biol. Int.* 21, 595–600. doi: 10.1006/cbir.1997.0190
- Drevet, C., Brasileiro, A. C., and Jouanin, L. (1994). Oncogene arrangement in a shooty strain of *Agrobacterium tumefaciens*. *Plant Mol. Biol.* 25, 83–90. doi: 10.1007/BF00024240
- Escobar, M. A., Civerolo, E. L., Summerfelt, K. R., and Dandekar, A. M. (2001). RNAi-mediated oncogene silencing confers resistance to crown gall tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13437–13442. doi: 10.1073/pnas.241276898
- Fründt, C., Meyer, A. D., Ichikawa, T., and Meins, F. (1998a). "Evidence for the ancient transfer of Ri plasmid T-DNA genes between bacteria and plants," in *Horizontal Gene Transfer*, eds M. Syvanen and C. I. Kado (London: Chapman and Hall), 94–106.
- Fründt, C., Meyer, A. D., Ichikawa, T., and Meins, F. (1998b). A tobacco homologue of the Ri-plasmid orf13 gene causes cell proliferation in carrot root discs. *Mol. Gen. Genet.* 259, 559–568. doi: 10.1007/s004380050849
- Furner, I. J., Huffman, G. A., Amasino, R. M., Garfinkel, D. J., Gordon, M. P., and Nester, E. W. (1986). An *Agrobacterium* transformation in the evolution of the genus *Nicotiana*. *Nature* 319, 422–427. doi: 10.1038/319422a0
- Gelvin, S. B. (2012). Traversing the cell: *Agrobacterium* T-DNA's journey to the host genome. *Front. Plant Sci.* 26:52. doi: 10.3389/fpls.2012.00052
- George, N. A., Pecota, K. V., Bowen, B. D., Schultheis, J. R., and Yencho, G. C. (2011). Root piece planting in sweetpotato-A synthesis of previous research and directions for the future. *HortTechnology* 21, 703–711.
- Guillon, S., Trémouillaux-Guiller, J., Pati, P. K., Rideau, M., and Gantet, P. (2006). Hairy root research: recent scenario and exciting prospects. *Curr. Opin. Plant Biol.* 9, 341–346. doi: 10.1016/j.pbi.2006.03.008
- Guivarc'h, A., Carneiro, M., Vilaine, F., Pautot, V., and Criqui, D. (1996). Tissue-specific expression of the rolA gene mediates morphological changes in transgenic tobacco. *Plant Mol. Biol.* 30, 125–134. doi: 10.1007/BF00017807
- Guyon, P., Petit, A., Tempé, J., and Dessaix, Y. (1993). Transformed plants producing opines specifically promote growth of opine-degrading agrobacteria. *Mol. Plant Microbe Interact.* 6, 92–98. doi: 10.1094/MPMI-6-092
- Handayani, N. S. N., Moriuchi, H., Yamakawa, M., Yamashita, I., Yoshida, K., and Tanaka, N. (2005). Characterization of the rolB promoter on mikimopine-type pRi1724 T-DNA. *Plant Sci.* 108, 1353–1364. doi: 10.1016/j.plantsci.2005.01.016
- Hansen, G., Vaubert, D., Cléröt, D., and Brevet, J. (1997). Wound-inducible and organ-specific expression of ORF13 from *Agrobacterium rhizogenes* 8196 T-DNA in transgenic tobacco plants. *Mol. Gen. Genet.* 254, 337–343. doi: 10.1007/s004380050424
- Hansen, G., Vaubert, D., Heron, J. H., Cléröt, D., Tempé, J., and Brevet, J. (1993). Phenotypic effects of overexpression of *Agrobacterium rhizogenes* T-DNA ORF13 in transgenic tobacco plants are mediated by diffusible factor(s). *Plant J.* 4, 581–585. doi: 10.1046/j.1365-313X.1993.04030581.x
- Helper, A., Clément, B., Michler, P., and Otten, L. (2003). The *Agrobacterium* oncogene AB-6b causes a graft-transmissible enation syndrome in tobacco. *Plant Mol. Biol.* 52, 483–493. doi: 10.1023/A:1023962121894
- Helper, A., Pien, S., and Otten, L. (2002). Functional diversity and mutational analysis of *Agrobacterium* 6B oncproteins. *Mol. Genet. Genomics* 267, 577–586. doi: 10.1007/s00438-002-0707-0
- Hildebrand, E. M. (1940). Cane gall of brambles caused by *Phytomonas rubi* n. sp. *J. Agric. Res.* 61, 685–696.
- Intrieri, M. C., and Buiatti, M. (2001). The horizontal transfer of *Agrobacterium rhizogenes* genes and the evolution of the genus *Nicotiana*. *Mol. Phylogenet. Evol.* 20, 100–110. doi: 10.1006/mpev.2001.0927
- Jefferson, R. A. (1987). Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol. Biol. Rep.* 5, 387–405. doi: 10.1007/BF02667740
- Kado, C. I. (2014). Historical account on gaining insights on the mechanism of crown gall tumorigenesis induced by *Agrobacterium tumefaciens*. *Front. Microbiol.* 5:340. doi: 10.3389/fmicb.2014.00340

- Kerr, A., and Panagopoulos, C. G. (1977). Biotypes of *Agrobacterium radiobacter* var. *tumefaciens* and their biological control. *Phytopathol. Z.* 90, 172–179. doi: 10.1111/j.1439-0434.1977.tb03233.x
- Kim, H., and Farrand, S. K. (1998). Opine catabolic loci from *Agrobacterium* plasmids confer chemotaxis to their cognate substrates. *Mol. Plant Micr. Interact.* 2, 131–143. doi: 10.1094/MPMI.1998.11.2.131
- Knapp, S., Chase, M. W., and Clarkson, J. J. (2004). Nomenclatural changes and a new sectional classification in *Nicotiana* (*Solanaceae*). *Taxon* 53, 73–82. doi: 10.2307/4135490
- Kodahl, N., Müller, R., and Lütken, H. (2016). The *Agrobacterium rhizogenes* oncogenes *rolB* and *ORF13* increase formation of generative shoots and induce dwarfism in *Arabidopsis thaliana* (L.) Heynh. *Plant Sci.* 252, 22–29. doi: 10.1016/j.plantsci.2016.06.020
- Koltunow, A. M., Johnson, S. D., Lynch, M., Yoshihara, T., and Costantino, P. (2001). Expression of *rolB* in apomictic *Hieracium piloselloides* Vill. causes ectopic meristems in planta and changes in ovule formation, where apomixis initiates at higher frequency. *Planta* 214, 196–205. doi: 10.1007/s004250100612
- Kováčová, V., Zluvová, J., Janousek, B., Talianova, M., and Vyskot, B. (2014). The evolution of the horizontally transferred agrobacterial mikimopine synthase gene in the genera *Nicotiana* and *Linaria*. *PLoS ONE* 9:e113872. doi: 10.1371/journal.pone.0113872
- Kyndt, T., Quispe, D., Zhai, H., Jarret, R., Ghislain, M., Liu, Q., et al. (2015). The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: an example of a naturally transgenic food crop. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5844–5849. doi: 10.1073/pnas.1419685112
- Lemcke, K., Prinsen, E., van Onckelen, H., and Schmülling, T. (2000). The ORF8 gene product of *Agrobacterium rhizogenes* TL-DNA has tryptophan 2-monooxygenase activity. *Mol. Plant Microbe Interact.* 13, 787–790. doi: 10.1094/MPMI.2000.13.7.787
- Lemcke, K., and Schmülling, T. (1998). Gain of function assays identify non-*rol* genes from *Agrobacterium rhizogenes* TL-DNA that alter plant morphogenesis or hormone sensitivity. *Plant J.* 15, 423–433. doi: 10.1046/j.1365-313X.1998.00223.x
- Leung, J., Fukuda, H., Wing, D., Schell, J., and Masterson, R. (1991). Functional analysis of cis-elements, auxin response and early developmental profiles of the mannopine synthase bidirectional promoter. *Mol. Gen. Genet.* 230, 463–474. doi: 10.1007/BF00280304
- Levesque, H., Delepelaire, P., Rouzé, P., Slightom, J., and Tepfer, D. (1988). Common evolutionary origin of the central portion of the Ri TL-DNA of *Agrobacterium rhizogenes* and the Ti T-DNAs of *Agrobacterium tumefaciens*. *Plant Mol. Biol.* 11, 731–744. doi: 10.1007/BF00019514
- Lim, K. Y., Skalicka, K., Koukalova, B., Volkov, R. A., Matyasek, R., Hemleben, V., et al. (2004). Dynamic changes in the distribution of a satellite homologous to intergenic 26–18S spacer in the evolution of *Nicotiana*. *Genetics* 166, 1935–1946. doi: 10.1534/genetics.166.4.1935
- Limami, M. A., Sun, L. Y., Douat, C., Helgeson, J., and Tepfer, D. (1998). Natural genetic transformation by *Agrobacterium rhizogenes*. *Plant Physiol.* 118, 543–550. doi: 10.1104/pp.118.2.543
- Long, N., Ren, X., Xiang, Z., Wan, W., and Dong, Y. (2016). Sequencing and characterization of leaf transcriptomes of six diploid *Nicotiana* species. *J. Biol. Res.* 23, 1–12. doi: 10.1186/s40709-016-0048-5
- Lütken, H., Clarke, J. L., and Müller, R. (2012). Genetic engineering and sustainable production of ornamentals: current status and future directions. *Plant Cell Rep.* 31, 1141–1157. doi: 10.1007/s00299-012-1265-5
- Martin-Tanguy, J., Sun, L. Y., Burtin, D., Vernoy, R., Rossin, N., and Tepfer, D. (1996). Attenuation of the phenotype caused by the root-inducing, left-hand, transferred DNA and its *rolA* gene. *Plant Physiol.* 111, 259–267. doi: 10.1104/pp.111.1.259
- Matveeva, T. V., Bogomaz, D. I., Pavlova, O. A., Nester, E. W., and Lutova, L. A. (2012). Horizontal gene transfer from genus *Agrobacterium* to the plant *Linaria* in nature. *Mol. Plant Microbe Interact.* 25, 1542–1551. doi: 10.1094/MPMI-07-12-0169-R
- Matveeva, T. V., and Lutova, L. A. (2014). Horizontal gene transfer from *Agrobacterium* to plants. *Front. Plant Sci.* 5:326. doi: 10.3389/fpls.2014.00326
- Maurel, C., Brevet, J., Barbier-Brygoo, H., Guern, J., and Tempé, J. (1990). Auxin regulates the promoter of the root-inducing *rolB* gene of *Agrobacterium rhizogenes* in transgenic tobacco. *Mol. Gen. Genet.* 223, 58–64. doi: 10.1007/BF00315797
- McKnight, T. D., Lillis, M. T., and Simpson, R. B. (1987). Segregation of genes transferred to one plant cell from two separate *Agrobacterium* strains. *Plant Mol. Biol.* 8, 439–445. doi: 10.1007/BF00017989
- Ménagé, A., and Morel, G. (1964). Sur la présence d'octopine dans les tissus de crown-gall cultivés *in vitro*. *Comptes Rendus de l'Académie des Sciences de Paris* 259, 4795–4796.
- Messens, E., Lenaerts, A., Van Montagu, M., and Hedges, R. W. (1985). Genetic basis for opine secretion from crown gall tumor cells. *Mol. Gen. Genet.* 199, 344–348. doi: 10.1007/BF00330278
- Meyer, A. D., Ichikawa, T., and Meins, F. (1995). Horizontal gene transfer: regulated expression of tobacco homologue of the *Agrobacterium rhizogenes* *rolC* gene. *Mol. Gen. Genet.* 249, 265–273. doi: 10.1007/BF00290526
- Michielse, C. B., Hooykaas, P. J. J., van den Hondel, C. A. M. J. J., and Ram, A. F. J. (2008). *Agrobacterium* mediated transformation of the filamentous fungus Aspergillus awamori. *Nat. Protoc.* 3, 1671–1678. doi: 10.1038/nprot.2008.154
- Mohajel-Shoja, H. (2010). Contribution to the Study of the *Agrobacterium rhizogenes* Plast Genes, *rolB* and *rolC*, and their Homologs in *Nicotiana tabacum*. Thesis of University of Strasbourg.
- Mohajel-Shoja, H., Clément, B., Perot, J., Alioua, M., and Otten, L. (2011). Biological activity of the *Agrobacterium rhizogenes*-derived *rolC* gene of *Nicotiana tabacum* and its functional relationship to other plast genes. *Mol. Plant Microbe Interact.* 24, 44–53. doi: 10.1094/MPMI-06-10-0139
- Mondy, S., Lenglet, A., Beury-Cirou, A., Libanga, C., Ratet, P., Faure, D., et al. (2014). An increasing opine carbon bias in artificial exudation systems and genetically modified plant rhizospheres leads to an increasing reshaping of bacterial populations. *Mol. Ecol.* 23, 4846–4861. doi: 10.1111/mec.12890
- Moore, L. W., Chilton, W. S., and Canfield, M. L. (1997). Diversity of opines and opine-catabolizing bacteria isolated from naturally occurring crown gall tumors. *Appl. Environ. Microbiol.* 63, 201–207.
- Murad, L., Bielawski, J. P., Matyasek, R., Kovarik, A., Nichols, R. A., Leitch, A. R., et al. (2004). The origin and evolution of geminivirus-related DNA sequences in *Nicotiana*. *Heredity (Edinb)* 92, 352–358. doi: 10.1038/sj.hdy.6800431
- Nagata, N., Kosono, S., Sekine, M., Shinmyo, A., and Syono, K. (1995). The regulatory functions of the *rolB* and *rolC* genes of *Agrobacterium rhizogenes* are conserved in the homologous genes (Ngrol) of *Nicotiana glauca* in tobacco genetic tumors. *Plant Cell Physiol.* 36, 1003–1012. doi: 10.1093/oxfordjournals.pcp.a078842
- Nagata, N., Kosono, S., Sekine, M., Shinmyo, A., and Syono, K. (1996). Different expression patterns of the promoters of the NgrolB and NgrolC genes during the development of tobacco genetic tumors. *Plant Cell Phys.* 37, 489–498. doi: 10.1093/oxfordjournals.pcp.a028971
- Nilsson, O., and Olsson, O. (1997). Getting to the root: the role of the *Agrobacterium rhizogenes* *rol* genes in the formation of hairy roots. *Physiol. Plant.* 100, 463–473. doi: 10.1111/j.1399-3054.1997.tb03050.x
- Ophel, K., and Kerr, A. (1990). *Agrobacterium vitis*-new species for strains of *Agrobacterium* biovar 3 from grapevine. *Int. J. Syst. Bacteriol.* 40, 236–241. doi: 10.1099/00207713-40-3-236
- Otten, L., and De Ruffray, P. (1994). *Agrobacterium vitis* nopaline Ti plasmid pTiAB4: relationship to other Ti plasmids and T-DNA structure. *Mol. Gen. Genet.* 245, 493–505. doi: 10.1007/BF00302262
- Otten, L., and Helfer, A. (2001). Biological activity of the *rolB*-like 5' end of the A4-orf8 gene from the *Agrobacterium rhizogenes* TL-DNA. *Mol. Plant Microbe Interact.* 14, 405–411. doi: 10.1094/MPMI.2001.14.3.405
- Otten, L., Salomon, J. Y., Helfer, A., Schmidt, J., Hammann, P., and De Ruffray, P. (1999). Sequence and functional analysis of the left-hand part of the T-region from the nopaline-type Ti plasmid, pTiC58. *Plant Mol. Biol.* 41, 765–776. doi: 10.1023/A:1006370207379
- Otten, L., and Szegedi, E. (1985). Crown galls induced by octopine-degrading biotype 3 strains of *Agrobacterium tumefaciens* contain a new form of lysopine dehydrogenase. *Plant Sci.* 40, 81–85. doi: 10.1016/0168-9452(85)90045-7
- Otten, L., Vreugdenhil, D., and Schilperoort, R. A. (1977). Properties of D⁺-lysopine dehydrogenase from crown gall tumour tissue. *Biochim. Biophys. Acta* 485, 268–277. doi: 10.1016/0005-2744(77)90163-2
- Pasternak, T., Haser, T., Falk, T., Ronneberger, O., Palme, K., and Otten, L. (2017). A 3D digital atlas of the *Nicotiana tabacum* root tip and its use to investigate changes in the root apical meristem induced by the *Agrobacterium* 6b oncogene. *Plant J.* doi: 10.1111/tpj.13631. [Epub ahead of print].

- Riker, A. J. (1930). Studies on infectious hairy root of nursery apple trees. *J. Agri. Res.* 41, 507–540.
- Rouillier, C., Duputié, A., Wennekes, P., Benoit, L., Bringas, V., Rossel, G., et al. (2013). Disentangling the origins of cultivated sweet potato (*Ipomoea batatas* (L.) Lam.). *PLoS ONE* 8:e62707. doi: 10.1371/journal.pone.0062707
- Savka, M. A., Dessaux, Y., Oger, P., and Rossbach, S. (2002). Engineering bacterial competitiveness and persistence in the phytosphere. *Mol. Plant Microbe Interact.* 15, 866–874. doi: 10.1094/MPMI.2002.15.9.866
- Schell, J., Van Montagu, M., De Beuckeleer, M., De Block, M., Depicker, A., De Wilde, M., et al. (1979). Interactions and DNA transfer between *Agrobacterium tumefaciens*, the Ti-plasmid and the plant host. *Proc. R. Soc. Lond. B Biol. Sci.* 204, 251–266. doi: 10.1098/rspb.1979.0026
- Schmülling, T., Fladung, M., Grossmann, K., and Schell, J. (1993). Hormonal content and sensitivity of transgenic tobacco and potato plants expressing single rol genes of *Agrobacterium rhizogenes* T-DNA. *Plant J.* 3, 371–382. doi: 10.1046/j.1365-313X.1993.t01-20-00999.x
- Schmülling, T., Schell, J., and Spena, A. (1988). Single genes from *Agrobacterium rhizogenes* influence plant development. *EMBO J.* 7, 2621–2629.
- Scott, I. M. (1979). Opine content of unorganized and teratomatous tobacco crown gall tissues. *Plant Sci. Lett.* 16, 239–248. doi: 10.1016/0304-4211(79)90034-8
- Serino, G., Cléröt, D., Brevet, J., Costantino, P., and Cardarelli, M. (1994). rol genes of *Agrobacterium rhizogenes* cucumopine strain: sequence, effects and pattern of expression. *Plant Mol. Biol.* 26, 415–422. doi: 10.1007/BF00039550
- Sierro, N., Battey, J. N., Ouadi, S., Bakaher, N., Bovet, L., Willig, A., et al. (2014). The tobacco genome sequence and its comparison with those of tomato and potato. *Nat. Commun.* 5:3833. doi: 10.1038/ncomms4833
- Spena, A., Schmülling, T., Koncz, C., and Schell, J. (1987). Independent and synergistic activity of rol A, B and C loci in stimulating abnormal growth in plants. *EMBO J.* 6, 3891–3899.
- Stieger, P. A., Meyer, A. D., Kathmann, P., Fründt, C., Niederhauser, I., Barone, M., et al. (2004). The orf13 T-DNA gene of *Agrobacterium rhizogenes* confers meristematic competence to differentiated cells. *Plant Physiol.* 135, 1798–1808. doi: 10.1104/pp.104.040899
- Studholme, D. J., Downie, J. A., and Preston, G. M. (2005). Protein domains and architectural innovation in plant-associated Proteobacteria. *BMC Genomics* 16:17. doi: 10.1186/1471-2164-6-17
- Suzuki, K., Yamashita, I., and Tanaka, N. (2002). Tobacco plants were transformed by *Agrobacterium rhizogenes* infection during their evolution. *Plant J.* 32, 775–787. doi: 10.1046/j.1365-313X.2002.01468.x
- Tepfer, D. (1983). "The biology of genetic transformation of higher plants by *Agrobacterium rhizogenes*," in Molecular Genetics of the *Bacteria-Plant Interaction*, ed A. Puehler (Berlin: Springer Verlag), 248–258.
- Tepfer, D. (1984). Transformation of several species of higher plants by *Agrobacterium rhizogenes*: sexual transmission of the transformed genotype and phenotype. *Cell* 37, 959–967. doi: 10.1016/0092-8674(84)90430-6
- Tepfer, D. (1990). Genetic transformation using *Agrobacterium rhizogenes*. *Physiol. Plant.* 79, 140–146. doi: 10.1111/j.1399-3054.1990.tb05876.x
- Trovato, M., Maras, B., Linhares, F., and Constantino, P. (2001). The plant oncogene rolD encodes a functional ornithine cyclodeaminase. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13449–13453. doi: 10.1073/pnas.231320398
- Udagawa, M., Aoki, S., and Syono, K. (2004). Expression analysis of the NgORF13 promoter during the development of tobacco genetic tumors. *Plant Cell Physiol.* 45, 1023–1031. doi: 10.1093/pcp/pch123
- Umber, M., Clément, B., and Otten, L. (2005). The T-DNA oncogene A4-orf8 from *Agrobacterium rhizogenes* A4 induces abnormal growth in tobacco. *Mol. Plant Microbe Interact.* 18, 205–211. doi: 10.1094/MPMI-18-0205
- Umber, M., Voll, L., Weber, A., Michler, P., and Otten, L. (2002). The rolB-like part of the *Agrobacterium rhizogenes* orf8 gene inhibits sucrose export in tobacco. *Mol. Plant Microbe Interact.* 15, 956–962. doi: 10.1094/MPMI.2002.15.9.956
- Van Altvorst, A. C., Bino, R. J., van Dijk, A. J., Lamers, A. M. J., Lindhout, W. H., van der Mark, F., et al. (1992). Effects of the introduction of *Agrobacterium rhizogenes* rol genes on tomato plant and flower development. *Plant Sci.* 83, 77–85. doi: 10.1016/0168-9452(92)90064-S
- Van Nuenen, M., de Ruffray, P., and Otten, L. (1993). Rapid divergence of *Agrobacterium vitis* octopine-cucumopine Ti plasmids from a recent common ancestor. *Mol. Gen. Genet.* 240, 49–57. doi: 10.1007/BF00276883
- Van Slooteren, G. M., Hooykaas, P. J. J., and Schilperoort, R. A. (1984). Silent T-DNA genes in plant lines transformed by *Agrobacterium tumefaciens* are activated by grafting and by 5-azacytidine treatment. *Plant Mol. Biol.* 3, 333–336. doi: 10.1007/BF00033379
- Ward, S. M., Fleischmann, C. E., Turner, M., and Singh, S. E. (2009). Hybridization between invasive populations of dalmatian toadflax (*Linaria dalmatica*) and yellow toadflax (*Linaria vulgaris*). *Invasive Plant Sci. Manag.* 2, 369–378. doi: 10.1614/IPSM-09-031.1
- Wenke, T., Döbel, T., Sörensen, T. R., Junghans, H., Weisshaar, B., and Schmidt, T. (2011). Targeted identification of short interspersed nuclear element families show their widespread existence and extreme heterogeneity in plant genomes. *Plant Cell* 23, 3117–3128. doi: 10.1105/tpc.111.088682
- White, F. F., Garfinkel, D. J., Huffman, G. A., Gordon, M. P., and Nester, E. W. (1983). Sequence homologous to *Agrobacterium rhizogenes* T-DNA in the genomes of uninfected plants. *Nature* 301, 348–350. doi: 10.1038/301348a0
- White, F. F., Taylor, B. H., Huffman, G. A., Gordon, M. P., and Nester, E. W. (1985). Molecular and genetic analysis of the transferred DNA regions of the root-inducing plasmid of *Agrobacterium rhizogenes*. *J. Bacteriol.* 164, 33–44.
- Yang, F., and Simpson, R. B. (1981). Revertant seedlings from crown gall tumors retain a portion of the bacterial Ti plasmid DNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* 78, 4151–4155. doi: 10.1073/pnas.78.7.4151
- Yokoyama, R., Hirose, T., Fujii, N., Aspuria, E. T., Kato, A., and Uchimiya, H. (1994). The rolC promoter of *Agrobacterium rhizogenes* Ri plasmid is activated by sucrose in transgenic tobacco plants. *Mol. Gen. Genet.* 244, 15–22. doi: 10.1007/BF00280182
- Zhang, M., Pereira e Silva Mde, C., Chaib De Mares, M., and van Elsas, J. D. (2014). The mycosphere constitutes an arena for horizontal gene transfer with strong evolutionary implications for bacterial-fungal interactions. *FEMS Microbiol. Ecol.* 89, 516–526. doi: 10.1111/1574-6941.12350

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Chen and Otten. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Horizontal Gene Transfer Contributes to Plant Evolution: The Case of *Agrobacterium* T-DNAs

Dora G. Quispe-Huamanquispe^{1,2}, Godelieve Gheysen¹ and Jan F. Kreuze^{2*}

¹ Department of Molecular Biotechnology, Ghent University, Ghent, Belgium, ² International Potato Center (CIP), Lima, Peru

OPEN ACCESS

Edited by:

Tatiana Matveeva,
Saint Petersburg State University,
Russia

Reviewed by:

Katharina Pawłowski,
Stockholm University, Sweden
Nikolai Ravin,
Institute of Bioengineering, Research
Center for Biotechnology Russian
Academy of Sciences, Russia

***Correspondence:**

Jan F. Kreuze
j.kreuze@cgiar.org

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 24 May 2017

Accepted: 13 November 2017

Published: 24 November 2017

Citation:

Quispe-Huamanquispe DG, Gheysen G and Kreuze JF (2017) Horizontal Gene Transfer Contributes to Plant Evolution: The Case of *Agrobacterium* T-DNAs. *Front. Plant Sci.* 8:2015. doi: 10.3389/fpls.2017.02015

Horizontal gene transfer (HGT) can be defined as the acquisition of genetic material from another organism without being its offspring. HGT is common in the microbial world including archaea and bacteria, where HGT mechanisms are widely understood and recognized as an important force in evolution. In eukaryotes, HGT now appears to occur more frequently than originally thought. Many studies are currently detecting novel HGT events among distinct lineages using next-generation sequencing. Most examples to date include gene transfers from bacterial donors to recipient organisms including fungi, plants, and animals. In plants, one well-studied example of HGT is the transfer of the tumor-inducing genes (T-DNAs) from some *Agrobacterium* species into their host plant genomes. Evidence of T-DNAs from *Agrobacterium* spp. into plant genomes, and their subsequent maintenance in the germline, has been reported in *Nicotiana*, *Linaria* and, more recently, in *Ipomoea* species. The transferred genes do not produce the usual disease phenotype, and appear to have a role in evolution of these plants. In this paper, we review previous reported cases of HGT from *Agrobacterium*, including the transfer of T-DNA regions from *Agrobacterium* spp. to the sweetpotato [*Ipomoea batatas* (L.) Lam.] genome which is, to date, the sole documented example of a naturally-occurring incidence of HGT from *Agrobacterium* to a domesticated crop plant. We also discuss the possible evolutionary impact of T-DNA acquisition on plants.

Keywords: HGT (horizontal gene transfer), *Agrobacterium*, *Ipomoea batatas* (L.) Lam., T-DNAs, evolution

INTRODUCTION

Horizontal gene transfer (HGT) can be defined as the acquisition of genetic material from another organism without being its offspring. It contrasts with vertical gene transfer, which is the acquisition of genetic material from an ancestor. HGT is a universal phenomenon and occurs frequently among prokaryotes. Bacteria have acquired a variety of important traits including antibiotic resistance, pathogenesis and metabolic pathways, via HGT. These horizontal gene acquisitions enabled bacteria to explore new habitats and hence facilitated their rapid evolution (Maiden, 1998; Ochman et al., 2000; Gogarten et al., 2002; Hotopp et al., 2007). In contrast to its rather common occurrence in prokaryotes, examples of HGT in eukaryotes have been reported only infrequently. However, that appears to be changing as recent discoveries indicate the possible contribution of HGT to the acquisition of traits with adaptive significance, suggesting that HGT is an important driving force in the evolution of eukaryotes, as well as prokaryotes. In this paper, we review HGT in higher organisms emphasizing examples involving *Agrobacterium* species and plants. We also discuss the possible evolutionary impact of the transferred genes on their respective hosts.

HGT IN EUKARYOTES

Horizontal gene transfer played a pivotal role in the origin of eukaryotes. Endosymbiosis and the subsequent genetic integration of entire organisms gave rise to the mitochondria and plastids (Talianova and Janousek, 2011). Advances in sequencing technologies in combination with ever increasing amounts of sequence data have facilitated the identification of additional examples of HGT in eukaryotes. In most instances, these were identified by chance while sequencing for other purposes or as a result of phylogenetic incongruences while attempting to establish evolutionary relationships. Examples include DNA transfer from bacteria, fungi and plants to bdelloid rotifers (Gladyshev et al., 2008), bacteria to insects (Hotopp et al., 2007), bacteria and fungi to nematodes (Noon and Baum, 2016), fish to fish (Graham et al., 2008), bryophytes to ferns (Li et al., 2014), and bacteria to plants (White et al., 1983; Matveeva and Lutova, 2014; Kyndt et al., 2015).

A particularly interesting example of HGT is the transfer of fungal genes to the pea aphid (*Acyrthosiphon pisum*) (Moran and Jarvik, 2010). Genes coding for carotenoid synthase/cyclase and carotenoid desaturase enzymes had not been reported in animals until their discovery in the pea aphid genome. Carotenoid biosynthesis genes are responsible for the body pigmentation in pea aphids. Body color is considered an ecologically important trait that influences the susceptibility of pea aphids to predators and/or parasites. Ladybugs (Coccinellidae) prefer to attack red aphids while parasitic wasps are more likely to lay their eggs in green aphids. Phylogenetic analysis of these genes in the pea aphid (Nováková and Moran, 2012) indicated that they had been obtained from fungi. Further analyses suggested that these genes were acquired early in the radiation of this group (Nováková and Moran, 2012).

Despite being one of the oldest groups of land plants, the majority of living ferns resulted from a relatively recent diversification following the arrival of angiosperms. In order to exploit the new understory habitats created by angiosperm dominated ecosystems, ferns evolved strategies to thrive under the low light conditions created by the angiosperm canopy. In adapting to these conditions, ferns acquired an unconventional chimeric photoreceptor, called neochrome, that fuses red sensing phytochrome and blue sensing phototropin modules into a single gene, thereby optimizing phototropic responses (Li et al., 2014). The recent analysis of 434 transcriptomes and 40 genomes of plants and algae demonstrated that ferns acquired this gene from hornworts (a bryophyte lineage) via HGT about 179 million year ago (Li et al., 2014).

Horizontal gene transfer seems to have played an important role also in the transition from the aquatic to the terrestrial environment. A novel genome analyses of the moss *Physcomitrella patens* reveal that 57 families of nuclear genes were acquired by HGT from prokaryotes, fungi or viruses. These genes have strong implications on plant-specific activities, such as xylem formation, plant defense, and hormone biosynthesis. The study suggests that many of these genes were transferred to the ancestors of green or land plants (Yue et al., 2012). Other examples of HGT in plants involve the case of

parasitic plants (Yang et al., 2016). The transcriptome analyses of three parasitic members of Orobanchaceae family show the occurrence of 52 high-confidence HGT events. Genes acquired by HGT are preferentially expressed in the haustorium, the host connecting organ of parasitic plants, proposing that these genes are contributing to the unique adaptive feeding structure of parasitic plants.

HGT FROM *Agrobacterium* SPECIES TO PLANTS (THE CLASSICAL MODEL)

Agrobacterium-mediated plant genetic transformation is probably the best studied and best understood system of transkingdom gene transfer. *Agrobacterium* is a plant pathogenic bacterium that causes neoplastic growth, i.e., uncontrolled cell division in host plants resulting in crown galls or in proliferating roots following the transfer of a segment of its DNA into the host cell genome.

Most of the bacterial genes necessary for the DNA transfer are located in a large tumor- or root-inducing plasmid (Ti/Ri plasmid) which also contains that part of the plasmid that is transferred (T-DNA). During *Agrobacterium* infection, plant-derived phenolics trigger the expression of the bacterium's virulence genes, and the encoded proteins subsequently mediate the T-DNA transfer to the host plant cell. The final destiny of the T-DNA in the host cell is dependent on various interactions between *Agrobacterium* and plant proteins. Several host cell pathways are utilized to ensure that the T-DNA is imported to the nucleus and integrated into the host genome (Lacroix and Citovsky, 2016). Expression of the T-DNA genes in the plant can alter the physiology to stimulate cell division and root growth. *IaaM* and *iaaH* encode enzymes for the biosynthesis of auxin that is essential for crown gall development (Zhang et al., 2015). Several *rol* (root loci) genes are involved in root formation while the function of several T-DNA genes such as *C-prot* is still unknown (Otten et al., 1999). Opines are also encoded on the T-DNAs, they are utilized as carbon and nitrogen sources by invading bacteria and their presence can alter the biological root environment, particularly, root associated bacterial populations (Oger et al., 1997). *Acs* encodes the key enzyme for the biosynthesis of the opine called *agrocinopine* while *mis* is a mikimopine synthase and *mas* a mannopine synthase.

The ability of *Agrobacterium* to transform plants has been exploited for decades as a means to introduce foreign genes of interest into crop plants (Tzfira and Citovsky, 2006; Gelvin, 2009). However, *Agrobacterium*-mediated HGT is not restricted to the production of genetically modified crops. Evidence of the naturally occurring transfer of T-DNA genes from *Agrobacterium* into plant genomes and their subsequent maintenance in the germline has been documented in *Nicotiana*, *Linaria*, and more recently *Ipomoea* species **Figure 1** (White et al., 1983; Intrieri and Buiatti, 2001; Matveeva et al., 2012; Pavlova et al., 2014; Kyndt et al., 2015). In these examples, the transferred genes are fixed and are expressed in the host plant's lineage suggesting that they might have a functional role.

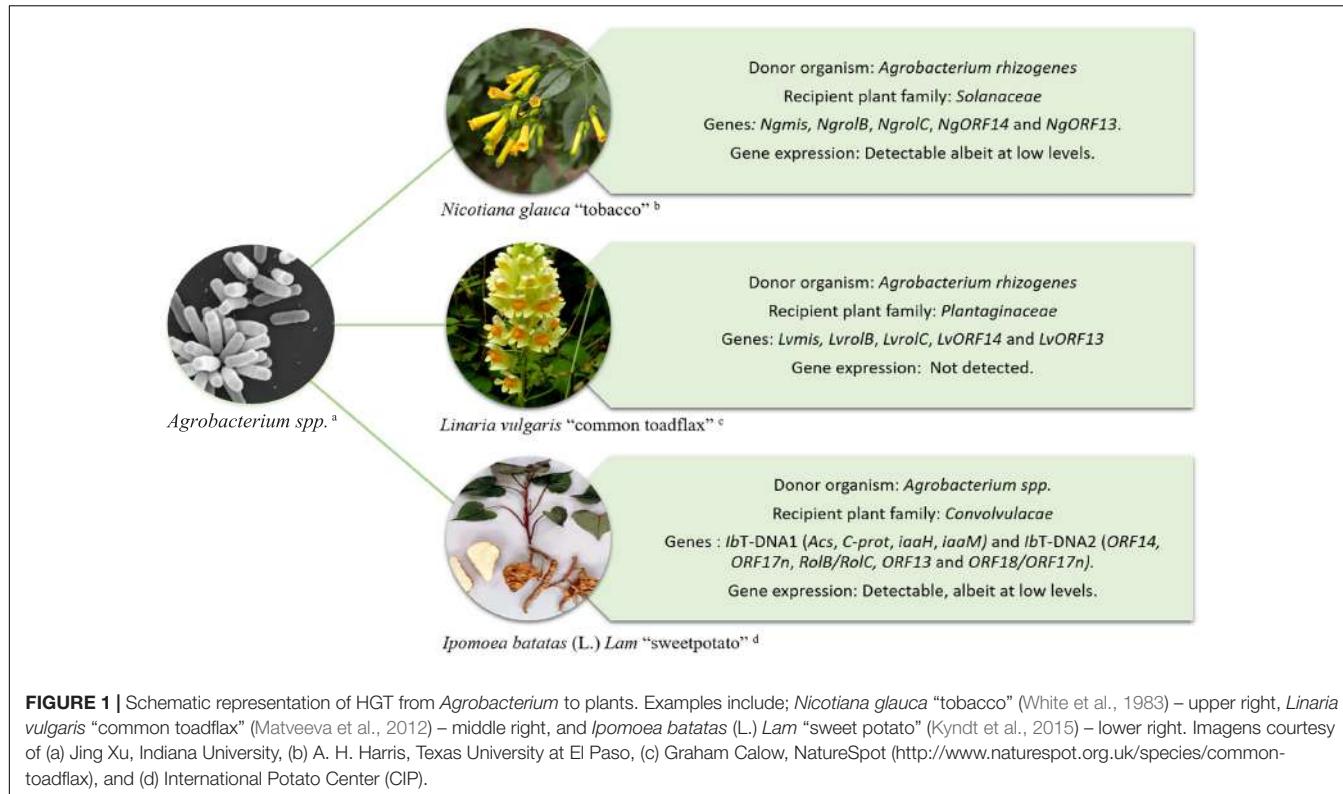


FIGURE 1 | Schematic representation of HGT from *Agrobacterium* to plants. Examples include; *Nicotiana glauca* "tobacco" (White et al., 1983) – upper right, *Linaria vulgaris* "common toadflax" (Matveeva et al., 2012) – middle right, and *Ipomoea batatas* (L.) Lam "sweet potato" (Kyndt et al., 2015) – lower right. Imagens courtesy of (a) Jing Xu, Indiana University, (b) A. H. Harris, Texas University at El Paso, (c) Graham Calow, NatureSpot (<http://www.naturespot.org.uk/species/common-toadflax>), and (d) International Potato Center (CIP).

HGT FROM *Agrobacterium rhizogenes* TO *Nicotiana* AND *Linaria*

More than three decades ago White et al. (1983) detected a region in the genome of *Nicotiana glauca* homologous to regions in the T-DNA of the Ri plasmid of *Agrobacterium rhizogenes*. The region was called cellular T-DNA (cT-DNA) (White et al., 1983). The cT-DNA in the *N. glauca* genome was initially described as an imperfect inverted repeat that contained two homologs to *rol* genes, *NgrolB* and *NgrolC* (Ng, *N. glauca*). Later, the cT-DNA was found to contain two additional genes corresponding to open reading frames *ORF13* and *ORF14* (Aoki et al., 1994). The discovery of mikimopine synthase (*mis*) sequences (*NgmisL* and *NgmisR*) in the *N. glauca* cT-DNA indicated that it originated from a mikimopine-type Ri plasmid (Suzuki et al., 2002).

PCR analysis and southern hybridization confirmed the acquisition of cT-DNA by *N. glauca* (Furner et al., 1986). Intrieri and Buiatti (2001) screened a total of 42 *Nicotiana* species for the presence of *rolB*, *rolC*, *ORF13* and *ORF14*, and at least one of those genes was detected in the genome of 15 species. Phylogenetic analyses concluded that the *rol* genes seemed to follow the evolution of the genus *Nicotiana*. This study (Intrieri and Buiatti, 2001) also suggested that more than one independent infection of *Nicotiana* by *A. rhizogenes* occurred in ancient times. This hypothesis was recently corroborated through deep sequencing of the genome of the ancestral tobacco species *Nicotiana tomentosiformis* (Chen et al., 2014). The genome of *N. tomentosiformis* contains four cT-DNAs all of which are derived from different *Agrobacterium* strains. These cT-DNAs,

TA, TB, TC, and TD, each contain an incomplete inverted-repeat structure. The TB region contains an intact mannopine synthase 2' gene (TB-*mas2'*) that is highly expressed in roots of some *N. tabacum* cultivars. These results suggest that the TB-*mas2'* gene could have been selected in some tobacco populations by nature or by tobacco growers, as a result of changes in the root metabolism of these plants (Chen et al., 2016).

cT-DNA sequences are not restricted to the genus *Nicotiana*. Indeed, they have also been found in species belonging to the genus *Linaria*, primarily within sections *Linaria* and *Speciosae* (Pavlova et al., 2014). Two copies of cT-DNA are present in *Linaria vulgaris* and are imperfect direct repeats. The *Linaria* cT-DNA appears to have originated from an ancient infection by a mikimopine strain of *A. rhizogenes*. Among the cT-DNA genes, *rolC* is the most conserved gene in the *Linaria* group and it contains an intact ORF. However, reverse transcriptional (RT) real-time PCR assays carried out using *L. vulgaris* internodes, leaves and roots under *in vitro* conditions have shown that *rolC* and the other cT-DNA genes are not expressed in these tissues (Matveeva et al., 2012).

HGT FROM *Agrobacterium* TO *Ipomoea* spp.

Sweet potato [*Ipomoea batatas* (L.) Lam.] belongs to genus *Ipomoea*. *Ipomoea* is the largest genus in the family Convolvulaceae and contains 600–700 species. Over half of *Ipomoea* spp. are concentrated in the Americas, where they are

distributed as cultigens, medicinal plants and weeds (Huaman, 1992). Series Batatas is a small group of taxa within the genus *Ipomoea* that contains 13 species that are considered to be closely related to sweet potato (Nimmakayala et al., 2011). Members of this series include *Ipomoea cordatotriloba*, *I. cynanchifolia*, *I. grandiflora*, *I. lacunosa*, *I. leucantha*, *I. littoralis*, *I. ramosissima*, *I. umbraticola*, *I. tabascana*, *I. tenuissima*, *I. tiliacea*, *I. trifida*, and *I. triloba*. The basic chromosome number of the series is 15 whereas the cultivated sweet potato is a hexaploid species ($2n = 6x = 90$). However, tetraploid ($2n = 4x = 60$) variants of *I. batatas* have also been reported (Bohac et al., 1993; Roullier et al., 2013) and these are sometimes referred to as tetraploid *I. trifida* or wild sweet potatoes in the scientific literature. Today, sweet potato is a staple food crop in many areas of the world. However, its botanical origins and the details concerning its domestication remain obscure.

The discovery of *Agrobacterium* genes *IbT*-DNA1 and *IbT*-DNA2 in the sweet potato genome represents the only known example of an ancient HGT that occurred in, what is today, a domesticated crop (Kyndt et al., 2015). Both regions, *IbT*-DNA1 and *IbT*-DNA2 were fortuitously detected during an analysis of small interfering RNA (*siRNA*) in the sweet potato cultivar Huachano. Plants of cv. Huachano contain an *IbT*-DNA1 with at least 4 ORFs with significant homology to the bacterial genes tryptophan-2-monooxygenase (*iaaM*), indole-3-acetamide hydrolase (*iaaH*), C-protein (*C-prot*) and agrocinopine synthase (*Acs*) and an *IbT*-DNA2 containing at least five ORFs with significant homology to *ORF14*, *ORF17n*, *RolB/RolC*, *ORF13*, and *ORF18/ORF17n*. The insertion of *IbT*-DNA1 has been corroborated by sequence analysis of a bacterial artificial chromosome (BAC) clone of sweet potato cv. Xu781. The BAC sequence revealed that the complete *IbT*-DNA1 encompassed 21,564 bp and consisted of an inverted repeat. *IbT*-DNA1 and 2 are transmitted from parent to progeny and the genes are expressed at detectable levels in different sweet potato tissues suggesting that they may have a function (Kyndt et al., 2015).

THE EVOLUTIONARY IMPACT OF THE ACQUISITION OF T-DNAs IN PLANTS

In general, for any foreign gene to be acquired by a host and stably inherited by its offspring (i) it must enter a cell and be integrated into the recipient genome, (ii) the DNA should not be lost after genome rearrangements during subsequent cell divisions, (iii) the transformed cell must enter the germ line, and finally (iv) the integrated sequence must be preserved in the course of evolution, which is most likely to happen if the gene confers a selective advantage to the recipient organism (Huang, 2013; Lacroix and Citovsky, 2016). In the case of T-DNA genes one could assume another specific requirement which is that the inserted genes must somehow be modified or controlled from their ‘natural’ expression pattern to avoid vigorous cell growth that would be detrimental to survival of the plant. The gold standard to determine gene function resulting from HGT is the existence of a phenotype that is correlated with the presence of those genes.

However, changes to the phenotype are not always so obvious and may in fact be difficult to detect.

The phenotypic effect of the *Agrobacterium* *rol* genes present in *Nicotiana* and *I. batatas* is likely associated with root traits (Matveeva et al., 2012; Kyndt et al., 2015). The suite of genes *rolA*, *rolB*, and *rolC* when transformed in tobacco plants, is able to induce the full “hairy root syndrome” (Maurel et al., 1991). *RolA* and *rolB* are mutated in *N. tabacum*, while *rolC* is intact. Transgenic tobacco plants bearing only *rolC* display phenotypic changes such as reduced apical dominance, dwarfism, shortened internodes, lanceolate leaves, and early flowering. They also exhibit increased root production when compared to untransformed plants (Shoja, 2010). The exact role of *IbT*-DNA genes remains to be elucidated, although it is known that the larger part of *IbT*-DNA1 and *IbT*-DNA2 genes are intact and are expressed.

How plants have avoided the *Agrobacterium* programmed expression of T-DNA sequences after insertion into their genomes to avoid *Agrobacterium* programmed cell proliferation is not clear yet, but several options can be considered. The T-DNA may have integrated in a region of the genome that is transcriptionally inactive, a property which is subsequently imparted on the inserted T-DNA. On the other hand, in sweetpotato the *IbT*-DNAs were originally discovered by small RNA sequencing and assembly, indicating that the genes - arranged in an inverted repeat- are targeted by the RNA silencing mechanism of the plant and in that way suppressed in their expression, even if they integrated in a transcriptionally active region of the genome (Kyndt et al., 2015). In the case of TB-*mas2'*, evidence suggests that originally it was a functional gene but has lost its expression in *N. tomentosiformis*, perhaps due to gene silencing; whereas it is active in *N. tabacum* (Chen et al., 2016).

The production of storage roots and the ability to easily propagate via rooted vine cuttings are major traits associated with the domestication of the sweet potato. Considering that *IbT*-DNAs appear to be inherited from wild relatives and some *IbT*-DNA genes have the potential to change plant physiology (auxin biosynthesis or sensitivity), it is tempting to speculate that *IbT*-DNAs have conferred an adaptive advantage to the host. A possible association between *rolB/rolC* genes and root parameters (total root yield, dry matter content, and harvest index) was evaluated in a population segregating for *IbT*-DNA2. No association between the occurrence of these genes and the noted root characteristics was detected, except for root yield at one location (Kyndt et al., 2015). Further study is required to establish the role, if any, of *IbT*-DNA2 genes in root development. To this end, a functional analysis using CRISPR-Cas9 to knockout single genes, combinations of genes, or the whole *IbT*-DNA1 and/or 2, would result in plants that could be analyzed in detail for their phenotype and developmental characteristics.

Knowledge of the timing of the ancestral infection as well as details about the infection process (such as whether it occurred as a single event or as multiple independent events), could shed light on the evolutionary impact of the cT-DNA and the *IbT*-DNA sequences on *Nicotiana* and *Ipomoea* spp. In *Nicotiana*, the data suggest that *Agrobacterium* spp. infected this group multiple

times, independently. Incongruences during the phylogenetic analyses of *rolB* in *Nicotiana* were the first evidence of this hypothesis (Intrieri and Buiatti, 2001; Suzuki et al., 2002) which is now gaining general acceptance. Indeed, the four cT-DNAs found in the ancestral tobacco species *N. tomentosiformis* appear to be derived from different *Agrobacterium* strains (Chen et al., 2016). In the sweet potato, *Ib*T-DNA 1 and 2 are present at different loci and segregate independently, i.e., *Ib*T-DNA1 seems to be fixed, while *Ib*T-DNA2 is restricted to only some accessions and segregates at random depending which genes are being analyzed. These differences may reflect different infection events. However, *A. rhizogenes* plasmids typically have two T-DNAs corresponding to *Ib*T-DNA1 and 2 that are transferred independently but often simultaneously.

FUTURE PERSPECTIVES

Investigations about the role of *Agrobacterium* T-DNAs in the evolution of plants are only just beginning. Screening of additional *Ipomoea* species in our labs will demonstrate if T-DNA genes are confined to the cultivated sweet potato, or are also present in some of its wild relatives. The pattern of possible acquisition of *Ib*T-DNAs by other *Ipomoea* species may help to formulate a hypothesis on the role that these sequences have

played in the evolution of this crop – and its related species. Although these genes are expressed at detectable levels in sweet potato, and some of them (*rolB/rolC*) are associated with root parameters, further analyses are needed in order to clarify their function(s).

AUTHOR CONTRIBUTIONS

DQ-H wrote the first draft and JK and GG subsequently contributed to produce the final version.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided to DQ-H from the Special Research Fund (BOF) of Ghent University, Belgium (01W02112) and Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC) of the government of Peru. We acknowledge Robert L. Jarret and León Otten for critical review of the manuscript. Research by JK was undertaken as part of, and funded by, the CGIAR Research Program on Roots, Tubers and Bananas (RTB) and supported by CGIAR Fund Donors (<http://www.cgiar.org/about-us/our-funders/>).

REFERENCES

- Aoki, S., Kawaoka, A., Sekine, M., Ichikawa, T., Fujita, T., Shinmyo, A., et al. (1994). Sequence of the cellular T-DNA in the untransformed genome of *Nicotiana glauca* that is homologous to ORFs 13 and 14 of the Ri plasmid and analysis of its expression in genetic tumors of *N. glauca* x *N. langsdorffii*. *Mol. Gen. Genet.* 243, 706–710. doi: 10.1007/BF00279581
- Bohac, J. R., Austin, D. F., and Jones, A. (1993). Discovery of wild tetraploid sweetpotatoes. *Econ. Bot.* 47, 193–201. doi: 10.1007/BF02862022
- Chen, K., Borne, F. D., Julio, E., Obszynski, J., Pale, P., and Otten, L. (2016). Root-specific expression of opine genes and opine accumulation in some cultivars of the naturally occurring genetically modified organism *Nicotiana tabacum*. *Plant J.* 87, 258–269. doi: 10.1111/tpj.13196
- Chen, K., Dorlac de Borne, F., Szegedi, E., and Otten, L. (2014). Deep sequencing of the ancestral tobacco species *Nicotiana tomentosiformis* reveals multiple T-DNA inserts and a complex evolutionary history of natural transformation in the genus *Nicotiana*. *Plant J.* 80, 669–682. doi: 10.1111/tpj.12661
- Furner, I. J., Huffman, G. A., Amasino, R. M., Garfinkel, D. J., Gordon, M. P., and Nester, E. W. (1986). An *Agrobacterium* transformation in the evolution of the genus *Nicotiana*. *Nature* 319, 422–427. doi: 10.1038/319422a0
- Gelvin, S. B. (2009). *Agrobacterium* in the genomics age. *Plant physiol.* 150, 1665–1676. doi: 10.1104/pp.109.139873
- Gladyshov, E. A., Meselson, M., and Arkhipova, I. R. (2008). Massive horizontal gene transfer in bdelloid rotifers. *Science* 320, 1210–1213. doi: 10.1126/science.1156407
- Gogarten, J. P., Doolittle, W. F., and Lawrence, J. G. (2002). Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* 19, 2226–2238. doi: 10.1093/oxfordjournals.molbev.a004046
- Graham, L. A., Lougheed, S. C., Ewart, K. V., and Davies, P. L. (2008). Lateral transfer of a lectin-like antifreeze protein gene in fishes. *PLOS ONE* 3:e2616. doi: 10.1371/journal.pone.0002616
- Hotopp, J. C. D., Clark, M. E., Oliveira, D. C., Foster, J. M., Fischer, P., Torres, M. C. M., et al. (2007). Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317, 1753–1756. doi: 10.1126/science.1142490
- Huaman, Z. (1992). *Systematic Botany and Morphology of the Sweetpotato Plant*. Lima: International Potato Center (CIP).
- Huang, J. (2013). Horizontal gene transfer in eukaryotes: the weak-link model. *Bioessays* 35, 868–875. doi: 10.1002/bies.201300007
- Intrieri, M. C., and Buiatti, M. (2001). The horizontal transfer of *Agrobacterium rhizogenes* genes and the evolution of the genus *Nicotiana*. *Mol. Phylogenet. Evol.* 20, 100–110. doi: 10.1006/mpve.2001.0927
- Kyndt, T., Quispe, D., Zhai, H., Jarret, R., Ghislain, M., Liu, Q., et al. (2015). The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: an example of a naturally transgenic food crop. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5844–5849. doi: 10.1073/pnas.1419685112
- Lacroix, B., and Citovsky, V. (2016). Transfer of DNA from bacteria to eukaryotes. *mBio* 7:e00863-16. doi: 10.1128/mBio.00863-16
- Li, F. W., Villarreal, J. C., Kelly, S., Rothfels, C. J., Melkonian, M., Frangedakis, E., et al. (2014). Horizontal transfer of an adaptive chimeric photoreceptor from bryophytes to ferns. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6672–6677. doi: 10.1073/pnas.1319929111
- Maiden, M. C. J. (1998). Horizontal genetic exchange, evolution, and spread of antibiotic resistance in bacteria. *Clin. Infect. Dis.* 27, S12–S20. doi: 10.1086/514917
- Matveeva, T. V., Bogomaz, D. I., Pavlova, O. A., Nester, E. W., and Lutova, L. A. (2012). Horizontal gene transfer from genus *Agrobacterium* to the plant *Linaria* in nature. *Mol. Plant Microbe Interact.* 25, 1542–1551. doi: 10.1094/MPMI-07-12-0169-R
- Matveeva, T. V., and Lutova, L. A. (2014). Horizontal gene transfer from *Agrobacterium* to plants. *Front. Plant Sci.* 5:326. doi: 10.3389/fpls.2014.00326
- Maurel, C., Barbier-Brygoo, H., Spena, A., Tempé, J., and Guern, J. (1991). Single *rol* genes from the *Agrobacterium rhizogenes* TL-DNA alter some of the cellular responses to auxin in *Nicotiana tabacum*. *Plant Physiol.* 97, 212–216.
- Moran, N. A., and Jarvik, T. (2010). Lateral transfer of genes from fungi underlies carotenoid production in aphids. *Science* 328, 624–627. doi: 10.1126/science.1187113
- Nimmakayala, P., Vajja, G., and Reddy, U. K. (2011). “Ipomoea,” in *Wild Crop Relatives: Genomic and Breeding Resources*, ed. C. Kole (Berlin: Springer), 123–132.

- Noon, J. B., and Baum, T. J. (2016). Horizontal gene transfer of acetyltransferases, invertases and chorismate mutases from different bacteria to diverse recipients. *BMC Evol. Biol.* 16:74. doi: 10.1186/s12862-016-0651-y
- Nováková, E., and Moran, N. A. (2012). Diversification of genes for carotenoid biosynthesis in aphids following an ancient transfer from a fungus. *Mol. Biol. Evol.* 29, 313–323. doi: 10.1093/molbev/msr206
- Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304. doi: 10.1038/35012500
- Oger, P., Petit, A., and Dessaix, Y. (1997). Genetically engineered plants producing opines alter their biological environment. *Nat. Biotechnol.* 15, 369–372. doi: 10.1038/nbt0497-369
- Otten, L., Salomone, J.-Y., Helfer, A., Schmidt, J., Hammann, P., and De Ruffray, P. (1999). Sequence and functional analysis of the left-hand part of the T-region from the nopaline-type Ti plasmid, pTiC58. *Plant Mol. Biol.* 41, 765–776. doi: 10.1023/A:1006370207379
- Pavlova, O., Matveeva, T., and Lutova, L. (2014). Genome of *Linaria dalmatica* contains *Agrobacterium rhizogenes* *RolC* gene homolog. *Russ. J. Genet. Appl. Res.* 4, 461–465. doi: 10.1134/S2079059714050116
- Roullier, C., Duputié, A., Wennekes, P., Benoit, L., Bringas, V. M. F., Rossel, G., et al. (2013). Disentangling the origins of cultivated sweet potato (*Ipomoea batatas* (L.) Lam.). *PLOS ONE* 8:e62707. doi: 10.1371/journal.pone.0062707
- Shoja, H. M. (2010). *Contribution to the Study of the Agrobacterium rhizogenes Plast Genes, RolB and RolC, and Their Homologs in Nicotiana tabacum*. Strasbourg: Universite de Strasbourg.
- Suzuki, K., Yamashita, I., and Tanaka, N. (2002). Tobacco plants were transformed by *Agrobacterium rhizogenes* infection during their evolution. *Plant J.* 32, 775–787. doi: 10.1046/j.1365-313X.2002.01468.x
- Talianova, M., and Janousek, B. (2011). What can we learn from tobacco and other Solanaceae about horizontal DNA transfer? *Am. J. Bot.* 98, 1231–1242. doi: 10.3732/ajb.1000370
- Tzfira, T., and Citovsky, V. (2006). *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. *Curr. Opin. Biotechnol.* 17, 147–154. doi: 10.1016/j.copbio.2006.01.009
- White, F. F., Garfinkel, D. J., Huffman, G. A., Gordon, M. P., and Nester, E. W. (1983). Sequences homologous to *Agrobacterium rhizogenes* T-DNA in the genomes of uninfected plants. *Nature* 301, 348–350. doi: 10.1038/301348a0
- Yang, Z., Zhang, Y., Wafula, E. K., Honaas, L. A., Ralph, P. E., Jones, S., et al. (2016). Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 113, E7010–E7019. doi: 10.1073/pnas.1608765113
- Yue, J., Hu, X., Sun, H., Yang, Y., and Huang, J. (2012). Widespread impact of horizontal gene transfer on plant colonization of land. *Nat. Commun.* 3, 1152. doi: 10.1038/ncomms2148
- Zhang, Y., Lee, C.-W., Wehner, N., Imdahl, F., Svetlana, V., Weiste, C., et al. (2015). Regulation of oncogene expression in T-DNA-transformed host plant cells. *PLOS Pathog.* 11:e1004620. doi: 10.1371/journal.ppat.1004620

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Quispe-Huamanquispe, Gheysen and Kreuze. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Symbiosome: Legume and Rhizobia Co-evolution toward a Nitrogen-Fixing Organelle?

Teodoro Coba de la Peña ^{1,2}, Elena Fedorova ^{1,3}, José J. Pueyo ¹ and M. Mercedes Lucas ^{1*}

¹ Instituto de Ciencias Agrarias ICA-CSIC, Madrid, Spain, ² Centro de Estudios Avanzados en Zonas Áridas (CEAZA), La Serena, Chile, ³ K. A. Timiryazev Institute of Plant Physiology, Russian Academy of Science, Moscow, Russia

OPEN ACCESS

Edited by:

Nikolai Provorov,
All-Russian Research Institute of
Agricultural Microbiology of the
Russian Academy of Agricultural
Sciences, Russia

Reviewed by:

Stefanie Wienkoop,
University of Vienna, Austria
Marc Libault,
University of Oklahoma, United States
Oksana Yurievna Shtark,
All-Russian Research Institute of
Agricultural Microbiology of the
Russian Academy of Agricultural
Sciences, Russia

*Correspondence:

M. Mercedes Lucas
mlucas@ica.csic.es

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 25 August 2017

Accepted: 19 December 2017

Published: 22 January 2018

Citation:

Coba de la Peña T, Fedorova E, Pueyo JJ and Lucas MM (2018) The Symbiosome: Legume and Rhizobia Co-evolution toward a Nitrogen-Fixing Organelle? *Front. Plant Sci.* 8:2229.
doi: 10.3389/fpls.2017.02229

In legume nodules, symbiosomes containing endosymbiotic rhizobial bacteria act as temporary plant organelles that are responsible for nitrogen fixation, these bacteria develop mutual metabolic dependence with the host legume. In most legumes, the rhizobia infect post-mitotic cells that have lost their ability to divide, although in some nodules cells do maintain their mitotic capacity after infection. Here, we review what is currently known about legume symbiosomes from an evolutionary and developmental perspective, and in the context of the different interactions between diazotroph bacteria and eukaryotes. As a result, it can be concluded that the symbiosome possesses organelle-like characteristics due to its metabolic behavior, the composite origin and differentiation of its membrane, the retargeting of host cell proteins, the control of microsymbiont proliferation and differentiation by the host legume, and the cytoskeletal dynamics and symbiosome segregation during the division of rhizobia-infected cells. Different degrees of symbiosome evolution can be defined, specifically in relation to rhizobial infection and to the different types of nodule. Thus, our current understanding of the symbiosome suggests that it might be considered a nitrogen-fixing link in organelle evolution and that the distinct types of legume symbiosomes could represent different evolutionary stages toward the generation of a nitrogen-fixing organelle.

Keywords: endosymbiosis, legumes, rhizobia, nodule, symbiosome, lupin, nitrogen fixation, organelle evolution

INTRODUCTION

Symbiosis between different organisms has played a key role in evolution and in fact, the term “symbiogenesis” is an evolutionary concept that refers to “the appearance of new physiologies, tissues, organs, and even new species as a direct consequence of symbiosis” (Chapman and Margulis, 1998; Margulis and Chapman, 1998; O’Malley, 2015). Endosymbiosis is a reciprocal advantageous association in which one organism lives inside another and it has a pivotal importance in symbiogenesis. Endosymbiotic theories to explain the origin of eukaryote cells and their organelles have been proposed and discussed for more than a century (Zimorski et al., 2014; Martin et al., 2015; O’Malley, 2015). Mitochondria and chloroplasts of eukaryotic cells, key organelles for respiration and photosynthesis, are thought to result from the evolution of an ancient endosymbiosis in which ancient bacterial-like organisms were engulfed into an ancient prokaryotic or eukaryotic-like cell (Dyall et al., 2004; Kutschera and Niklas, 2005; Zimorski et al., 2014; Archibald, 2015).

The endosymbiosis that leads to organelle formation follows distinct key processes and stages: recognition between symbionts, engulfment, the failure of defense systems to eliminate

the endosymbiont by defense reaction, physiological integration and finally, genetic integration (Margulis and Chapman, 1998). It is commonly accepted that during the transition from an endosymbiont to an organelle, cyclical endosymbiosis becomes permanent or obligate endosymbiosis by the transfer of endosymbiont genes to the nucleus of the host cell, establishment of a protein targeting system to reimport the products of these genes, division of the endosymbiont inside the macrosymbiont and the vertical transmission to the macrosymbiont's offspring (Cavalier-Smith and Lee, 1985; Chapman and Margulis, 1998; McFadden, 1999; Parniske, 2000; Douglas and Raven, 2003; Dyall et al., 2004). Therefore, is it obvious what differentiates an endosymbiont from an organelle? It has been suggested that "the boundaries between these terms can blur" and that it might be necessary to employ other criteria to distinguish an endosymbiont from an organelle (Keeling and Archibald, 2008). Thus, studies focusing on more modern endosymbioses might reveal how organelles came to be and why they look the way they do (Keeling et al., 2015; McCutcheon, 2016).

The oxygen respiration and photosynthetic capacity of ancestral mitochondria and chloroplasts, respectively, was the key driving force for endosymbiosis and co-evolution toward organelle formation. As nitrogen is an important component of biomolecules and frequently a limiting nutrient, nitrogen fixation is a fundamental process in ecosystems (Tyrrell, 1999). The capacity to fix atmospheric nitrogen (diazotrophy) is exclusive to prokaryotic organisms that contain the nitrogenase enzyme complex. Diazotrophs include some archaea and within the eubacteria, they include proteobacteria, cyanobacteria, and actinobacteria. Eukaryotic organisms are unable to fix nitrogen and thus, different types of symbiotic relationships have been established between eukaryotes and diazotrophic bacteria to fulfill this function, ranging from loose interactions to highly regulated intracellular symbioses (Kneip et al., 2007). In these interactions, eukaryotic organisms supply nutrients and energy to the diazotrophs in exchange for fixed nitrogen.

In plants, there are two types of associations with soil diazotroph eubacteria that are relevant to the symbiotic fixation of atmospheric nitrogen in a new organ developed in plant, the nodule. The filamentous Gram-positive bacteria *Frankia* are nitrogen-fixing endosymbionts of plants that are collectively called actinorhizal plants. By contrast, Gram-negative bacteria known as rhizobia, fix nitrogen in root nodules of legumes and of the non-legume *Parasponia*. Nitrogen-fixing symbiosis in legume root nodules is the best studied to date and it is significantly important for the nitrogen input in both agricultural and natural ecosystems. The legume root nodule was considered as "the best example of symbiospecific morphogenesis" (Chapman and Margulis, 1998). Specific recognition between symbionts takes place through the exchange of signaling molecules. For example, legume roots secrete flavonoids and other compounds to the rhizosphere, generally inducing the synthesis and secretion of rhizobial lipo-chito-oligosaccharides (LCOs, Nod factors). These molecules act as mitogens inducing cell division in the root cortex, and the formation of the root nodule through the progressive differentiation of specialized cells and tissues (Pueppke, 1996; Geurts et al., 2005; Cooper, 2007). Concomitant

with nodule primordium development, bacteria enter the root cortex and infect cells of the nodule primordium (Brewin, 1991; Jones et al., 2007).

Two main types of symbiotic nodules have been described as a function of the type of growth: indeterminate and determinate. The typical indeterminate nodule is originated by proliferation of inner root cortical cells; it has a persistent apical meristem and adopting a cylindrical shape. The typical determinate nodule originates by proliferation of outer cortical cells and it has a lateral meristem that remains active for some days. After the arrest of meristematic activity, the nodule grows by cell expansion and it adopts a spherical shape (Patriarca et al., 2004).

Rhizobia can use intracellular or intercellular routes to infect legume roots. In the former, infection occurs at root hairs where infection threads (IT) form. IT grows inwardly until it reaches the nodule primordium cells. The intracellular mode of infection occurs in most of the rhizobia-legume symbioses studied and it is tightly controlled by the host. Intercellular infection may take place via natural wounds, where lateral roots emerge through epidermal breaks (crack infection), or it may occur directly between epidermal cells or between an epidermal cell and an adjacent root hair (Gualtieri and Bisseling, 2000; Vega-Hernández et al., 2001; González-Sama et al., 2004; reviewed in Sprent, 2009; and in Ibáñez et al., 2017). At least 25% of all legume genera may undergo non-hair rhizobia infection and their nodules lack ITs (Sprent, 2007). Rhizobia that enter the nodule host cell are surrounded by a host-derived membrane called the peribacteroid membrane or symbiosome membrane (SM). This new cellular compartment formed by the intracellular bacteria (bacteroid) enclosed within a SM is referred to as the symbiosome (**Figure 1**). Bacteria can divide within the symbiosome and whole symbiosomes can also divide inside the host cell, both these types of division being carried out synchronously or not (Whitehead and Day, 1997; Oke and Long, 1999). After rhizobia division ceases, the bacteria differentiate into nitrogen-fixing bacteroids. Plant defense reactions are suppressed or attenuated during the infection process (Mithöfer, 2002; Luo and Lu, 2014) or evaded (Saeki, 2011).

The symbiosome is the basic nitrogen-fixing unit of the nodule and the nitrogen fixed by bacteroids is exported as ammonium to the host plant cytoplasm, where it is assimilated and transported toward the rest of the plant. Conversely, reduced carbon compounds from the plant are transported to the nodule, and many other metabolites may also be exchanged between the host cell and symbiosome (Udvardi and Day, 1997; Hinde and Trautman, 2002). In 1997, it was first postulated that "symbiosomes can be interpreted as special nitrogen-fixing organelles within the host cell" (Whitehead and Day, 1997).

In most of the legumes studied, nodule host cells stop dividing upon rhizobia infection (Brewin, 1991), although young infected cells can still undergo cell division in several determinate nodules but this process is not sustained for long (Patriarca et al., 2004). Nevertheless, rhizobia-infected cell division does occur in some specific cases, such as the peculiar indeterminate nodule of *Lupinus* known as lupinoid nodule (González-Sama et al., 2004; Fedorova et al., 2007), and it is a key event in forming the infected tissue in which nitrogen will be fixed.

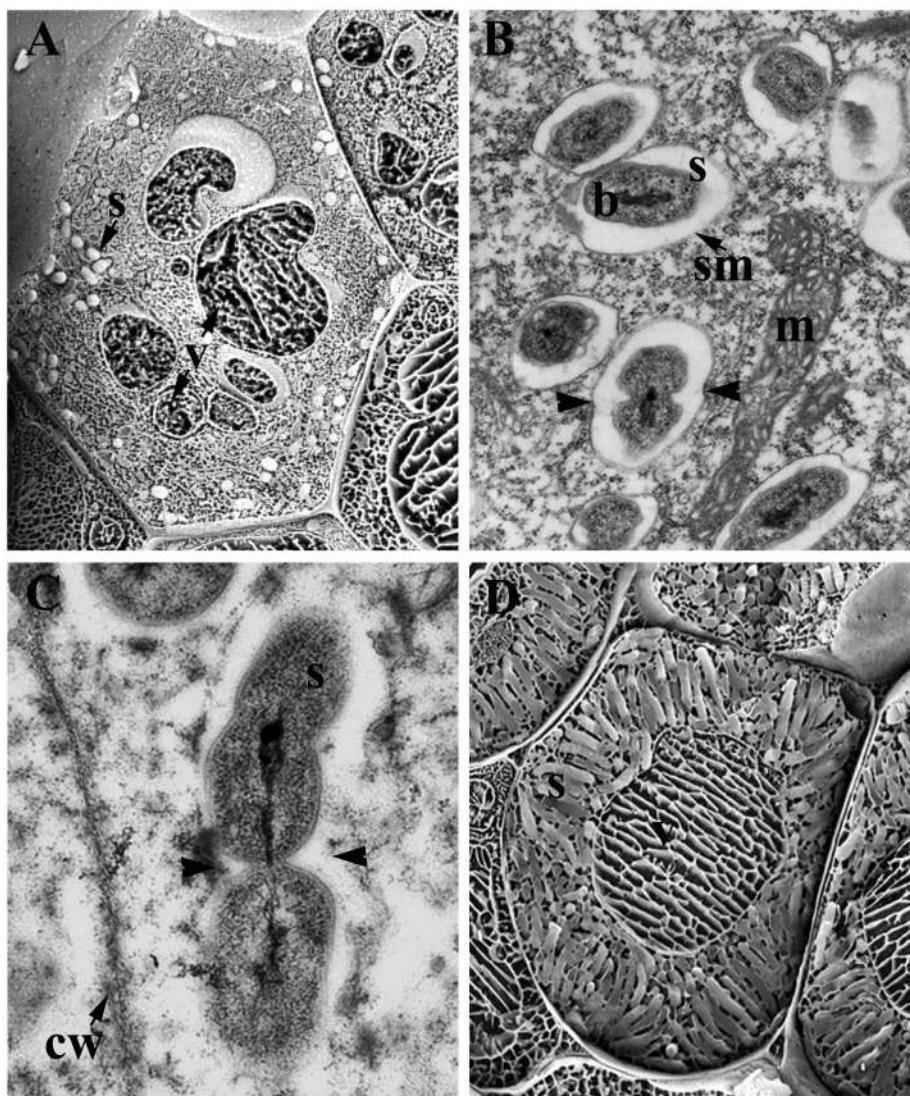


FIGURE 1 | Infected cells of legume nodules. **(A)** Young infected cells showing few symbiosomes and vacuole disintegration; *Medicago sativa* nodule. **(B,C)** Symbiosomes in young infected cells; *M. sativa* **(B)** and *Lupinus albus* **(C)** nodules. Note the dividing symbiosomes (arrowheads). **(D)** Mature infected cells harboring mature symbiosomes; *M. sativa* nodule. Low temperature-scanning electron microscopy image **(A,D)**. Transmission electron microscopy image **(B,C)**. b, Bacteroid; cw, cell wall; m, mitochondria; s, symbiosome; sm, symbiosome membrane; v, vacuole.

Here, we will present some evolutionary considerations regarding rhizobia-legume symbioses in general, and about *Lupinus* symbiosis in particular, leading us to suggest that different legume symbiosomes could represent some different stages in an evolutionary process toward a nitrogen-fixing organelle. First, we will introduce some evolutionary considerations about the origin of mitochondria and chloroplast, contrasting this with the apparent absence of diazotrophic organelles. We will compare the different degrees of association between diazotrophs and eukaryotes, and we will detail a number of evolutionarily relevant features of rhizobia-legume symbiosis. Finally, we will analyse the various organelle-like characteristics of the symbiosome, providing evidence suggesting that the

symbiosome might be considered a nitrogen-fixing link in organelle evolution.

THE ORIGIN OF MITOCHONDRIA AND CHLOROPLASTS AS A MODEL OF ORGANELLE EVOLUTION. EVOLUTIONARY CONSIDERATIONS ON THE ABSENCE OF NITROGEN-FIXING ORGANELLES

Biochemical, genetic, phylogenetic, and structural studies indicate that mitochondria are derived from an

α -proteobacterium-like ancestor that was engulfed as a microsymbiont by an Archaea-type host between 2.2 and 1.5 Bya (Table 1; Dyall et al., 2004; Kutschera and Niklas, 2005; Gray, 2012). This specific symbiotic association was linked to the appearance of the first heterotrophic unicellular eukaryotes. Similarly, the primary origin of plastids is due to a symbiotic association between an ancient cyanobacterium and a mitochondrial carrying eukaryote, which took place between 1.5 and 1.2 Bya, giving rise to photosynthetic unicellular eukaryotes (Table 1; McFadden, 1999; Dyall et al., 2004; Kutschera and Niklas, 2005; Keeling, 2010).

The distinction between an endosymbiont and an organelle remains a matter of debate. It has been postulated that key aspects to distinguish an organelle from an endosymbiont include the transfer of genes from the symbiont to the host nucleus, together with the establishment of a protein import apparatus in order to reimport the products of the transferred genes back into the compartment where they originally acted (Cavalier-Smith and Lee, 1985; Theissen and Martin, 2006; Keeling and Archibald, 2008; Archibald, 2015). Thus, a key event in the evolution from endosymbiont to organelle is the loss of autonomy of the microsymbiont as a free-living organism. This loss of autonomy is generally a consequence of microsymbiont genome reduction due to gene transfer to the host genome and gene loss (Dyall et al., 2004; Archibald, 2015). Such reduction is a continuous process (Douglas and Raven, 2003; Bock and Timmis, 2008) and the relocation of proto-organelle genes to the host genome may occur to avoid harboring duplicate sets of microsymbiont genes. Moreover, DNA transfer from organelles to the nucleus may drive gene and genome evolution (Kleine et al., 2009). An additional criterion thought to define an organelle is the host's control of organelle division and segregation (Keeling and Archibald, 2008). In the proposed major transitions approach, the evolution of symbiotic partnerships in the newly integrated organism is thought to be driven by the vertical transmission of symbionts into the host's offspring, a key event for the integration of both partners (Kiers and West, 2015).

Mitochondria and plastids, double membrane-surrounded cell organelles of endosymbiotic origin, fit with these criteria of reduced genome size, gene transfer to the host cell's nucleus, the presence of a protein import machinery, and host-driven division and segregation (Keeling, 2010; Strittmatter et al., 2010; Gray, 2012; Dudek et al., 2013). It is interesting to note that putative intermediate stages in mitochondrial and plastid evolution have been proposed. A heterotrophic flagellate of the genus *Reclinomonas* is reported to contain a minimally-derived mitochondrial genome with 67 protein encoding genes, many more than the mitochondrial genes conserved in yeast (8) and humans (13). Moreover, ancestral bacterial protein transport routes coexist with the evolving mitochondrial protein import machinery in *R. americana*. Accordingly, *Reclinomonas* mitochondria may represent a “connecting link” between the metazoan mitochondria and their ancestral bacterial progenitors (Lang et al., 1997, 1999; Tong et al., 2011).

The thecate amoeba *Paulinella chromatophora* contains obligate subcellular plastid-like photosynthetic bodies called chromatophores. It was estimated that these chromatophores

evolved from free-living *Synechococcus* cyanobacteria 200–60 Mya (Nowack, 2014), although it is unclear whether these subcellular bodies should be considered as endosymbionts or organelles (Keeling and Archibald, 2008). Some years ago, the cyanobacterium-like plastids of the amoeba *P. chromatophora* were believed to represent intermediate forms in the transition from endosymbiont to plastids, these chromatophores retaining a prokaryotic peptidoglycan cell wall that is lost in current plastids (Keeling, 2004). These subcellular bodies have a smaller genome than their free-living relatives and they are metabolically dependent on their host. Indeed, several chromatophore genes have been transferred to the host nucleus and at least some of the proteins encoded by these genes are targeted to the chromatophores. Moreover, these subcellular bodies divide in synchrony with their host. Thus, in accordance to the aforementioned criteria, the chromatophores of *Paulinella* can be considered an early stage photosynthetic organelle that is the result of a relatively recent endosymbiotic event (Nowack et al., 2008; Nakayama and Ishida, 2009; Nakayama and Archibald, 2012; Nowack and Grossman, 2012; Archibald, 2015).

In contrast to mitochondria and chloroplasts, nitrogen-fixing organelles are absent in extant organisms, raising questions as to why these organelles have not yet appeared in the course of evolution (McKay and Navarro-González, 2002). Based on the close phylogenetic relationship between current diazotrophic bacteria (α -proteobacteria rhizobia and cyanobacteria) and the most likely free-living ancestors of mitochondria or chloroplasts, there doesn't appear to be any fundamental incompatibility of diazotrophic predecessors for endosymbiosis and for the transfer of nitrogen-fixing genes to the host cell's nucleus (Allen and Raven, 1996).

Nitrogenase is inhibited by oxygen, so nitrogen-fixing organisms might have appeared before the Great Oxidation Event more than 2 Bya (i.e., the accumulation of oxygen in the atmosphere) (Raymond et al., 2004). Nitrogen-fixing organisms probably originated in a time when there was a shortage in the availability of fixed-nitrogen. Three nitrogen crises have been proposed during evolution: the first just after the origin of life (more than 3.5 Bya); the second possibly due to a strong reduction in atmospheric CO₂ (about 2.5 Bya); and the third, possibly induced by the action of pluricellular plant-based ecosystems (500 Mya; McKay and Navarro-González, 2002).

Isotope studies suggest that biological nitrogen fixation first took place about 3.2 Bya (Stüeken et al., 2015). It was also postulated that biological nitrogen fixation appeared later than the genesis of the eukaryotic cell (McKay and Navarro-González, 2002, and references therein) and molecular dating suggested that the origin of biological nitrogen fixation was between 2.2 and 1.5 Bya (Fani et al., 2000; Boyd and Peters, 2013). Thus, biological nitrogen fixation could have appeared during the second nitrogen crisis. Eukaryogenesis had been completed by then and it was a single event. Thus, for whatever reason, an opportunity for new endosymbiosis between the unicellular eukaryotic cell and diazotrophs did not arise. If nitrogen-fixing organisms appeared during the third crisis, higher plants already existed, and thus, incorporation and vertical transmission in multicellular organisms was much more difficult (McKay

TABLE 1 | Cellular organelles derived from endosymbionts and putative connecting-link or intermediate stages in organelle evolution (adapted from Lang et al., 1997; Douglas and Raven, 2003; Kutschera and Niklas, 2005; Marin et al., 2005).

Organelle	Ancestor	Age (million years ago)	Organelle function	Eukaryotic host	Morphological, physiological and molecular modifications
Mitochondria	Ancient α -proteobacterium-like	2,200–1,500	Aerobic Respiration	Eukaryotes	High rates of gene loss (just 8–13 genes retained) Gene transfer to the host cell nucleus Protein import machinery Rapid sequence evolution No bacterial-like division
Atypical mitochondria	Ancient α -proteobacterium-like	More recent than typical mitochondria	Aerobic respiration	<i>Reclinomonas americana</i>	67 protein-encoding genes retained Eubacterial-like gene transcription Eubacterial-like protein sorting coexisting with evolving mitochondrial protein import machinery
Chloroplast	Ancient cyanobacterium-like	1,500–1,200	Photosynthesis	Photosynthetic eukaryotes	High rates of gene loss Gene transfer to the nucleus of the host cell Rapid sequence evolution Bacterial-like division
Chromatophores	Cyanobacterium-like plastids	200–60 More recent than typical chloroplast	Photosynthesis	<i>Paulinella chromatophora</i>	Reduced genome and gene transfer to the nucleus Protein targeting from the nucleus Division in synchrony with the host Peptidoglycan cell wall retained Similar pigmentation to many cyanobacteria Bacterial-like β -carboxysomes

and Navarro-González, 2002). Indeed, it has been postulated that organelle development does not occur in differentiated multicellular organisms (McKay and Navarro-González, 2002).

A very interesting case of co-evolution involving a permanent nitrogen-fixing endosymbiont can be found in diatoms of the Rhopalodiaceae family, protists that contain the so-called spheroid bodies (SB) in their cytoplasm. As in the case of rhizobia-legume symbioses, the host and microsymbiont are strictly separated by a host-derived membrane in these species (Drum and Pankratz, 1965; Prechtl et al., 2004; Bothe et al., 2010). Moreover, phylogenetic analyses showed that these SBs are derived from a group of cyanobacteria and that their genome is closely related to that of nitrogen-fixing bacteria of the genus *Cyanothece* (Adler et al., 2014). This is a case of obligate symbiosis with vertical transmission, because SBs cannot survive outside the host cells (Prechtl et al., 2004). Indeed, this seems to be a case of recent symbiosis induced by a loss of photosynthetic capacity of the cyanobacteria-derived symbiont (Prechtl et al., 2004). This endosymbiosis was proposed to have occurred in the middle Miocene epoch, ~12 Mya (Nakayama et al., 2011). The complete genome of a SB from one of these diatom species was recently sequenced (Nakayama et al., 2014), confirming the reduced size and gene repertoire of the SB relative to their closer free-living relatives. Furthermore, the presence of pseudogenes and gene fusions suggest an ongoing process of genome reduction. Interestingly, the genome of SBs contains a set of genes for nitrogen fixation and isotope analysis indicated that the host diatoms use the nitrogen fixed by the SBs (Nakayama and Inagaki, 2014). However, genes for functional photosynthesis are lacking in its genome and thus, SBs depend on their diatom hosts for their energy requirements. To date, SBs have not been considered as organelles *stricto sensu*, as gene transfer to the host nucleus and protein import machinery have not yet

been detected. Moreover, little is known about endosymbiont division and segregation to host daughter cells (Adler et al., 2014; Nakayama and Inagaki, 2014).

It is interesting to note that some unicellular nitrogen-fixing cyanobacteria of the oceanic picoplankton, termed UCYN-A, have suffered a more pronounced reduction of their genome than that observed in SBs. These cyanobacteria lack genes that code for several metabolic pathways, yet they are evolutionarily related to SBs. It has been proposed that these cyanobacteria may enter into symbiosis with prymnesiophyte photosynthetic unicellular algae, supplying fixed nitrogen to the host and receiving fixed carbon in return (Thompson et al., 2012; Nakayama and Inagaki, 2014). Like SBs, this relationship between UCYN-A cyanobacteria and unicellular algae can be considered another stage in the evolution of symbiosis involving nitrogen-fixation.

DIFFERENT LEVELS OF INTERACTION BETWEEN DIAZOTROPH BACTERIA AND EUKARYOTES

Only some diazotroph bacteria are known to establish symbiotic interactions with eukaryotes, be they animal, plant, fungus, or protist. These interactions range from loose associations to highly specific intracellular symbioses, involving different molecular, physiological, and morphological modifications. As such, the co-evolutionary status of these associations can be estimated by considering the degree of interdependence (facultative or obligate symbiont), the extra- or intracellular location of the microsymbiont, the presence or absence of segregation to daughter cells and of vertical transmission (Kneip et al., 2007). Some examples of diversity of interactions between diazotroph and plants or photosynthetic protists are shown in Table 2.

TABLE 2 | Some associations of diazotrophs with photosynthetic eukaryotes.

Diazotroph	Eukaryotic host	Microsymbiont location	Degree of dependence	Infection mode	Host niche	Infected cell division	Vertical transmission	Characteristic features of bacteria	References
<i>Azospirillum</i> sp.	Poaceae <i>Zea mays</i> <i>Oryza sativa</i>	Extracellular	Facultative	None	Intercellular spaces in roots and other plant tissues No nodules	No	No	None	Hurek et al., 1994 Reinhold-Hurek and Hurek, 1998; Hurek, 2011 Stehouw and Vanderleyden, 2000 Endelin and Meeks, 1993 Adams and Duggan, 2008 Bergman et al., 2008
<i>Nostoc</i> sp.	Bryophyta <i>Anthoceros punctatus</i>	Extracellular	Facultative	None	Cavities of the gametophyte No nodules	No	No	Increased heterocyst frequency	Vessey et al., 2004
<i>Nostoc</i> sp. or <i>Anabaena</i> sp.	Monilophyta <i>Azolla</i> sp.	Extracellular	Obligate	Infection of sexual megasporangium	Cavities in the dorsal leaf that are obligately infected by filamentous cyanobacteria No nodules	No	Yes	Cyanobiont genome degradation	Bergman et al., 2008
<i>Nostoc</i> sp. (predominant)	Cycads <i>Encephalatros</i> sp. <i>Macrozamia</i> sp. Others	Extracellular	Facultative	Infection of coraloid roots (somewhat comparable to crack entry)	Coralloid roots No nodules	No	No	Irreversibly modified coraloid roots	Rasmussen and Nilsson, 2002 Vessey et al., 2004
<i>Nostoc</i> sp.	Angiosperm <i>Gunnera</i> L.	Intracellular	Facultative	Plant stem glands	Specialized plant stem glands No nodules	No	No	Differentiation of <i>Nostoc</i> filaments	Rasmussen et al., 1994 Bergman et al., 2008
<i>Frankia</i> sp.	Dicotyledonous actinorhizal plants Casuarinaceae and others (<i>Aleurites</i> sp., <i>Casuarina</i> sp., etc.)	Intracellular	Facultative	Infection thread-like structures or penetration between root epidermal cells	Symbiotic nodule Indeterminate Multi-lobed Central vasculature	No	No	Miller and Baker, 1986 Vessey et al., 2004 Pawlowski and Sprent, 2008 Kucho et al., 2010	
<i>Frankia</i> sp. of cluster II	Actinorhizal Rosales and Cucurbitales	Intracellular	Likely obligate	Unknown	Symbiotic nodule Indeterminate Multi-lobed	No	Unknown	High percentage of pseudogenes Proposed genome reduction	Pawlowski and Sprent, 2008 Person et al., 2011
<i>Bradyrhizobium</i> sp. <i>Rhizobium</i> sp.	Cannabaceae (<i>Paraspasia</i> sp.)	Intracellular	Facultative	Crack entry Root erosion (intercellular infection-thread)	Symbiotic nodule Indeterminate Central vasculature (similar to actinorrhizal nodules)	No	No	Fixation threads	Trinick, 1973 Becking, 1979 Vessey et al., 2004
<i>Ensifer</i> sp. <i>Rhizobium</i> sp.	<i>Medicago</i> sp. <i>Pisum</i> sp. <i>Vicia</i> sp.	Intracellular	Facultative	Root hair Infection thread	Symbiotic nodule Indeterminate, Cylindrical Peripheral vasculature	No	No	Irreversible differentiation into bacteroids	Newcomb, 1981 Rae et al., 1992

(Continued)

TABLE 2 | Continued

Diazotroph	Eukaryotic host	Microsymbiont location	Degree of dependence	Infection mode	Host niche	Infected cell division	Vertical transmission	Characteristic features of bacteria	References
<i>Mesorhizobium</i> sp.	<i>Lotus</i> sp. <i>Phaseolus</i> sp.	Intracellular	Facultative	Root hair Infection thread	Symbiotic nodule Determinate	Limited	No		Newcomb, 1981 Parkhurst et al., 1987
<i>Rhizobium</i> sp.	<i>Chamaecytisus proliferus</i>	Intracellular	Facultative	Crack entry	Symbiotic nodule Indeterminate; Cylindrical	Limited	No	Aborted infection threads	Vega-Hernández et al., 2001
<i>Bradyrhizobium</i> sp.	<i>Glycine max</i>	Intracellular	Facultative	Root hair Infection thread	Symbiotic nodule Determinate	Limited	No		Bhuvaneswari et al., 1980
<i>Bradyrhizobium</i> sp.	<i>Dalbergioids Arachis hipogaea</i> <i>Stylosanthes</i> sp.	Intracellular	Facultative	Crack entry	Symbiotic nodule Determinate Aeschynomeneoid	Yes	No		Chandler, 1978 Chandler et al., 1982 Saijaga et al., 2001 Saeki, 2011
<i>Bradyrhizobium</i> sp.	<i>Sarothamnus scoparius</i>	Intracellular	Facultative	Intercellular	Symbiotic nodule Indeterminate	Yes	No		Kalita et al., 2006
<i>Bradyrhizobium</i> sp.	<i>Genistoid Genista tinctoria</i> .	Intracellular	Facultative	Intercellular	Symbiotic nodule Indeterminate lupinoid	Yes	No		Tang et al., 1993 Lotocka et al., 2000 González-Sáma et al., 2004
<i>Bradyrhizobium</i> sp.	<i>Lupinus</i> sp.	Intracellular	Facultative	Intercellular					Fedorova et al., 2007
<i>Cyanothece</i> sp.	Diatom <i>Rhopalodia gibba</i>	Intracellular	Obligate	Obligate endosymbiont	Spheroid bodies No nodules	Yes	Yes	Loss of genes Accumulation of deleterious mutations	Drum and Pankratz, 1985 Kneip et al., 2008 Bothe et al., 2010

Examining these diazotroph-plant interactions has enabled different degrees of specialization to be defined. For example, *Azospirillum* sp., *Azoarcus* sp., and some other free-living diazotroph bacteria are plant-growth promoting bacteria that can establish interactions with different cereals by root colonization or endophytic association, and they profit from microaerobic environments to fix nitrogen while obtaining nutrients from the plant's roots (Reinhold-Hurek and Hurek, 1998, 2011; Steenhoudt and Vanderleyden, 2000; Pérez-Montaña et al., 2014). Another example of a relative loose association involving diazotrophs is the symbiosis established between the cyanobacteria *Nostoc* sp. and the bryophyte *Anthoceros punctatus* L. (Adams and Duggan, 2008). In this case, the microsymbiont is located extracellularly in the cavities of the gametophyte and one physiological adaptation of this is that the heterocyst frequency in *Nostoc* sp. is higher than in free-living conditions (Endelin and Meeks, 1983).

In the symbiosis between cyanobacteria (*Nostoc* or *Anabaena*) and the fern *Azolla*, the diazotroph microsymbiont resides extracellularly in a mucilaginous sheath in the dorsal cavities of *Azolla* leaves. The cyanobacteria's filaments enter into the fern's sexual megasporangium, allowing the microsymbiont to be transferred vertically to the next plant generation. While it retains its photosynthetic capacity, it seems that these diazotroph cyanobacteria have lost their capacity to survive as free-living organisms (Bergman et al., 2008). Indeed, there are signs of reductive genome evolution or degradation of the cyanobiont, i.e., the presence of a high proportion of pseudogenes and a high frequency of transposable elements (Larsson, 2011). As such, it has been proposed that this cyanobiont may be at the initial phase of the transition from a free-living organism to a nitrogen-fixing plant entity, similar to chloroplast evolution (Ran et al., 2010). Moreover, it is possible that this *Nostoc* symbiosis may have persisted for 200 million years (Bergman et al., 2008).

All gymnosperm cycads can establish root symbioses with *Nostoc* sp. and with other cyanobacteria (Thajuddin et al., 2010). Cyanobacteria invade a particular root type, the cycad coralloid roots, provoking irreversible morphological modifications. The cyanobacteria remain extracellular in this symbiosis, which could have originated up to 250 Mya (Vessey et al., 2004 and references therein). A different strategy is adopted in the symbiosis between *Nostoc* sp. and the angiosperm *Gunnera* L. These bacteria infect specialized plant stem glands to become intracellular. Indeed, these glands secrete a specific signaling molecule that induces the differentiation of *Nostoc* filaments into a specialized form that is essential for infection (Rasmussen et al., 1994; Bergman et al., 2008). Moreover, *Nostoc* filaments are always surrounded by a host plasma membrane. In these examples, cyanobacteria fix nitrogen in both free-living and symbiotic conditions, and symbiosis is facultative and there has been no vertical transmission observed (Bonnett and Silvester, 1981; Rasmussen et al., 1994; Santi et al., 2013).

Root-nodule symbioses can be established between higher plants and soil bacteria, and it was estimated that nitrogen-fixing root nodule symbioses evolved 50–100 Mya (Kistner and Parniske, 2002). Symbiosis of the actinorhiza *Frankia* originated about 70–90 Mya (Doyle, 1998, 2011; Hoher et al., 2011), while

legume-rhizobia symbiosis originated about 55–60 Mya (Lavin et al., 2005), *Parasponia*-rhizobia symbiosis is much more recent (less than 10 million years; Op den Camp et al., 2011). In actinorhizal symbioses, soil actinobacteria of the genus *Frankia* induce nodules in the roots of about 260 plant species from eight different families of dicotyledonous plants (Vessey et al., 2004; Benson and Dawson, 2007). *Frankia* can fix nitrogen as a free-living organism and it can enter the host plant root either intracellularly (through root hairs) or intercellularly, depending on the host plant species. *Frankia* induces the formation of multi-lobed, indeterminate nodules, which are modified adventitious secondary roots formed from the root pericycle. Nodule infected cells become full of branching *Frankia* hyphae surrounded by a perimicrobial membrane of host origin, forming vesicles in which nitrogen fixation takes place (Vessey et al., 2004; Pawłowski and Sprent, 2008; Kucho et al., 2010; Froussart et al., 2016). This symbiosis is usually facultative but *Frankia* strains of cluster II, which form symbiosis with actinorhizal Rosales and Cucurbitales, still cannot be cultured and thus, these actinobacteria are probably obligate symbionts (Pawłowski and Sprent, 2008). The failure to culture these microbial strains may be related with atypical patterns of auxotrophy (Gtari et al., 2015). The genome of a member of this cluster is small and with a relatively high proportion of pseudogenes, suggesting that this strain underwent a process of genome reduction and that genome degradation is ongoing (Persson et al., 2011). However, this genome reduction does not involve physiological impairment, as no metabolic pathways appear to be incomplete. Notably, it also contains fewer genes involved in stress responses.

The symbiosis established between rhizobia and legumes is very specific and it involves a more complex exchange of signals and the development of a root nodule. This structure is not a modified root (as in the case in cycads, actinorhizal plants and *Parasponia*) but rather, it arises from unique zones of cell division in the root cortex (Vessey et al., 2004). Most rhizobia can only fix nitrogen in symbiotic conditions, when the bacteria have differentiated into bacteroids (the nitrogen-fixing form) inside the symbiosomes within the nodule's host cells (Brewin, 1991; Whitehead and Day, 1997). In most symbioses, legume host cells do not further divide once infected by the bacteria. This is the case for thread-infected indeterminate nodules formed by *Pisum* or *Medicago*. It has been suggested that young cells in thread-infected determinate nodules, such as those formed by *Glycine*, *Lotus*, or *Phaseolus*, undergo cell division but not in a sustained manner (Patriarca et al., 2004). In the case of the symbiosis established between *Bradyrhizobium* and *Arachis* or *Stylosantes*, giving rise to determinate nodules, infected cells can divide (Chandler, 1978; Chandler et al., 1982). In lupinoid nodules formed by *Lupinus albus*, infected host cells continue to divide for several cycles (Fedorova et al., 2007) and indeed, the lupinoid nodule grows continually and maintains an active lateral meristem with infected dividing cells (Figure 2), allowing the segregation of symbiosomes between daughter cells (Figure 3). Nevertheless, legume symbiosis is facultative and no vertical transmission occurs, such that new infection by rhizobia must occur for each new plant generation and no gene transfer from micro- to macro-symbiont has been reported.

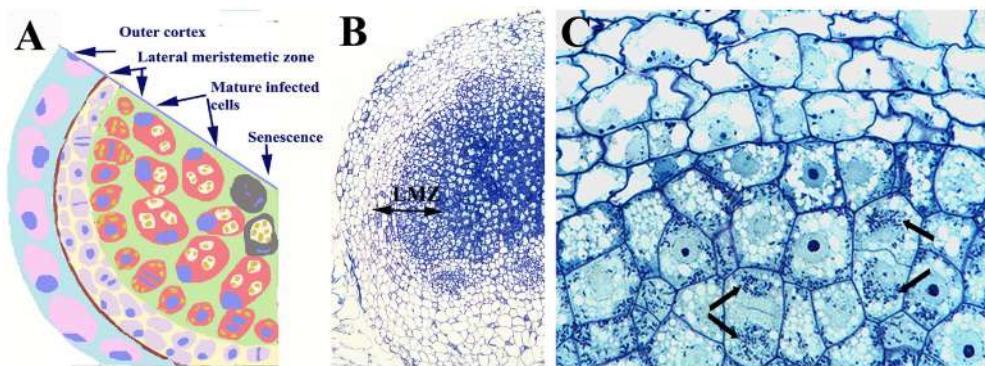


FIGURE 2 | Nodule of *Lupinus albus* showing dividing infected cells. **(A)** Scheme of a nodule section and **(B)** light microscopy image showing the outer cortex, and the lateral meristematic zone (LMZ) composed of infected and uninfected dividing cells, as well as the central zone composed of infected cells. **(C)** Detail of the LMZ in which the arrows label the symbiosomes. Note the symmetric distribution of symbiosomes between daughter cells. Images **(B,C)** modified from Fedorova et al. (2005); they are being reproduced with permission from the copyright holder.

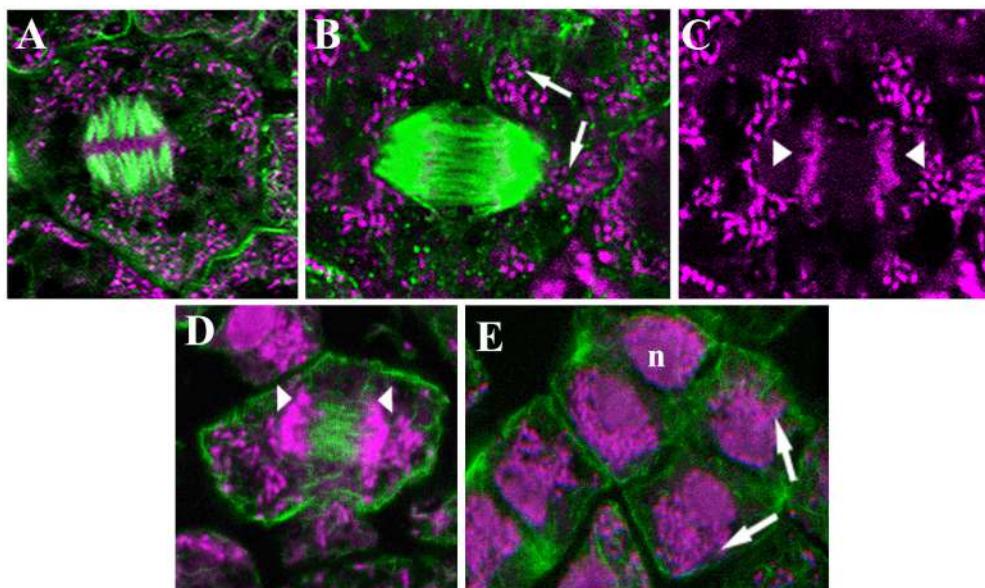


FIGURE 3 | Confocal laser images of dividing infected cells of *Lupinus albus* nodules showing the cytoskeletal elements in green, and the DNA of bacteroids (arrow) and chromosomes (arrowhead) in magenta. **(A–D)** Metaphase, microtubules **(A,B)** and actin microfilaments **(D)**. **(E)** Different anaphase–telophase stages showing the actin microfilaments. n, Nuclei. Images modified from Fedorova et al. (2007); they are being reproduced with permission from the copyright holder.

Parasponia (Cannabaceae, order Rosales) is the only non-legume plant that can establish effective nodule symbiosis with rhizobia. This symbiosis is a case of convergent evolution and it occurred more recently than that of legumes. From a phylogenetic and taxonomic point of view, *Parasponia* is closer to some actinorhizal plants that belong to the Rhamnaceae, Elaeagnaceae, and Rosaceae families, than to legumes (Soltis et al., 1995; Geurts et al., 2012). *Parasponia* nodules are modified lateral roots that originate from the pericycle, and they are indeterminate and more similar to actinorhizal nodules than to legume nodules. The entry of symbiotic bacteria (*Rhizobium*, *Bradyrhizobium*) does not involve root hairs but rather, crack

entry or root erosion and an intercellular IT. This IT protrudes into the host plant cell by plant membrane invagination, forming the so-called fixation thread. Fixation-thread, that remains in contact with the plasma membrane, are the equivalent to a symbiosome in legumes and to arbuscules in arbuscular mycorrhizal (AM) roots (Vessey et al., 2004; Pawłowski and Sprent, 2008; Behm et al., 2014). AM symbiosis preceded root nodule symbioses and the interactions of plants with AM fungi probably originated more than 400 Mya (Bonfante and Genre, 2008). This symbiosis is wide spread, involving more than 80% of all terrestrial plants, and fungi from order Glomales (Harrier, 2001). In this symbiosis, AM fungi enter the roots

and spread into the inner cortex by invagination of the plasma membrane. Invading hyphae branch and they develop the arbuscule, a specialized structure that is subsequently enveloped by the periarbuscular membrane, an extension of the host plant's plasma membrane. A symbiotic interface between the arbuscule and the periarbuscular membrane controls the efficient exchange of nutrients between both symbionts, including the transfer of phosphorus and nitrogen from the fungus in return for photosynthates from the plant (Smith and Read, 2008). It is notable that some components of the signaling pathway required to establish rhizobia-legume symbiosis and the symbiotic interface are also present in AM symbiosis (Kouchi et al., 2010; Harrison and Ivanov, 2017).

As described above, the endosymbiosis of SBs related to the cyanobacterium *Cyanothecae* sp., with the diatom *Rhopalodia gibba* and some other species, seems to be a unique case of obligate nitrogen-fixing endosymbiosis, involving genome reduction, a lack of metabolically essential genes and vertical transmission. As indicated above, the microsymbiont is currently not considered a real organelle due to the lack of gene transfer to the host nucleus and of a protein import machinery (Nakayama and Inagaki, 2014).

EVOLUTIONARY CONSIDERATIONS ABOUT INDIVIDUAL SYMBIONTS IN RHIZOBIA-LEGUME SYMBIOSSES

Some Genetic and Evolutionary Characteristics of the Microsymbiont

In general terms, rhizobia are defined as soil bacteria that fix nitrogen in symbiotic association with legumes and *Parasponia*. The Proteobacteria is an important phylum that contains diazotrophic organisms and phylogenetic studies using 16S ribosomal RNA sequences indicate that the best-known rhizobial genera are from the α -proteobacteria group (Rogel et al., 2011; Weir, 2016), including the genera: *Rhizobium*, *Mesorhizobium*, *Sinorhizobium* (renamed *Ensifer*, Martens et al., 2007; Judicial Commission of the International Committee on Systematics of Prokaryotes, 2008), *Bradyrhizobium*, *Azorhizobium*, and *Allorhizobium*. Some other α -proteobacteria genera also contain one or more rhizobial species, such as *Aminobacter*, *Methylobacterium*, *Devosia*, *Ochrobactrum*, *Phyllobacterium*, *Microvirga*, and *Shinella* (Rogel et al., 2011; Ormeño-Orrillo et al., 2015; Weir, 2016; ICSP Subcommittee on the taxonomy of *Rhizobium* and *Agrobacterium* <http://edzna.ccg.unam.mx/rhizobial-taxonomy/>). Recently, *Neorhizobium* and *Pararhizobium* have been proposed as new genera (Mousavi et al., 2014, 2015). Several rhizobial species belong to the β -proteobacteria genera, including *Burkholderia*, *Cupriavidus*, and *Herbaspirillum* (Moulin et al., 2001; Chen et al., 2003; Lloret and Martínez-Romero, 2005; Masson-Boivin et al., 2009; Rogel et al., 2011; Weir, 2016). Indeed, the taxonomy of rhizobia has recently been revised (Peix et al., 2015; Shamseldin et al., 2017).

In a first instance, a comparison of glutamine synthetase (GS) genes I and II allowed the time of divergence among the α -proteobacteria genera of rhizobia to be estimated (Turner

and Young, 2000). The data from GSII sequences suggest that *Rhizobium* and *Ensifer* are the most recent genera, and *Bradyrhizobium* and *Mesorhizobium* the most ancient. Based on GS, *Rhizobium*, *Ensifer*, and *Mesorhizobium* genera appear to have separated at the same time, and *Bradyrhizobium* is the most ancient genus. Based on the analysis of GS genes and the amino acid substitution rates in their orthologs, the *Bradyrhizobium* genus probably diverged from the last common ancestor of all rhizobia some 500 Mya, before the appearance of land plants (about 400 Mya). Similarly, the most recent genus *Ensifer* diverged about 200 Mya (Turner and Young, 2000; Morton, 2002; Lloret and Martínez-Romero, 2005), before the appearance of Angiosperms (dated more than 150 Mya; Martin et al., 1989) and legumes (about 70 Mya; Lavin et al., 2005). When phylogenetic analyses of the 16S rRNA gene and the intergenic spacer region was combined, slightly but not significantly more recent divergence times were found for rhizobia: about 385 Mya for *Bradyrhizobium*, 344 Mya for *Mesorhizobium*, 201 Mya for *Ensifer*, 145 Mya for *Rhizobium/Agrobacterium*, and 54 Mya for *Neorhizobium* (Chriki-Adeeb and Chriki, 2016).

Evolutionary studies of α -proteobacteria indicate that while the evolution of a reductive genome has been observed in intracellular animal-associated bacteria, genome expansion is observed in plant symbionts (as well as in several animal and plant pathogens, such as *Rickettsia*, *Brucella*, or *Bartonella*). Rhizobia are among the α -proteobacteria with the largest genomes (MacLean et al., 2007) and genes involved in nitrogen fixation and nodulation (or pathogenicity) have become integrated for symbiosis, often arranged on auxiliary replicons in genomic islands (mobile elements). The genome size and the diversity among rhizobia are due to the presence of these highly dynamic auxiliary replicons and to a high degree of paralogy (Batut et al., 2004).

Genome plasticity and instability in rhizobia is due to largescale recombination events (the presence of repeated DNA sequences, insertion elements and multiple replicons), and in fact, lateral gene transfer is the primary source of genetic diversity in rhizobia (Flores et al., 2000; Guo et al., 2003; MacLean et al., 2007; Provorov et al., 2008). It has been proposed that the genomes of rhizobia have evolved by expansion as a means to adjust to the challenges imposed by their multiphase lifestyle, principally through horizontal gene transfer and gene duplication (Batut et al., 2004; MacLean et al., 2007; Provorov and Andronov, 2016). In some rhizobia-legume symbioses, up to 15–20% of the rhizobial genome is activated in symbiosis (Udvardi et al., 2004; Tikhonovich and Provorov, 2009). Different models of co-evolution in the rhizobia-legume symbiosis have been proposed or are under study (but they are still controversial); especially in relation to the selection of rhizobial symbiotic traits by the host legume (Provorov et al., 2008; Martínez-Romero, 2009).

An evolutionary step from free-living diazotrophs related to *Rhodopseudomonas* to the symbiotic diazotroph *Bradyrhizobium* through the acquisition of *fix* genes was proposed as the first stage of rhizobial evolution (Provorov, 2015; Provorov and Andronov, 2016). It is noteworthy that when compared to other well-known rhizobia, *Bradyrhizobium* displays several particular

genomic and physiological characteristics related to diazotrophy and symbiosis. For example:

- a) *Bradyrhizobium japonicum* strains have some of the largest bacterial chromosomes sequenced to date (9.1–9.6 Mb), and the largest of all rhizobia (Kündig et al., 1993; Kaneko et al., 2002, 2011; Batut et al., 2004; Siqueira et al., 2014).
- b) The nodulation genes in *Bradyrhizobium* are located in a chromosomal segment that could be a mobile element, whereas they are located in symbiotic plasmids in other genera such as *Rhizobium*, *Ensifer*, and several strains of *Mesorhizobium* (Minamisawa et al., 1998; Sessitsch et al., 2002).
- c) Some *Bradyrhizobium* strains do not have genes for Nod factors but they can induce nodulation in certain some legumes of the *Aeschynomene* genus using an alternative triggering molecule (Giraud et al., 2007). This is a unique case of nodulation by rhizobia that does not involve Nod factors.
- d) *Bradyrhizobium* is the only genus of rhizobia in which some species (those nodulating some species of the *Aeschynomene* genus by a Nod factor-independent mechanism) can perform photosynthesis and fix nitrogen in symbiosis or in free-living conditions (Molouba et al., 1999). Other rhizobia can only fix nitrogen in symbiotic conditions, with the exception of *Azorhizobium caulinodans* and some strains of *Burkholderia* (Sprent et al., 2017).
- e) *Bradyrhizobium* displays an atypical two-component regulatory system, NodV and NodW, which is involved in controlling *nod* gene expression (Stacey, 1995; Loh et al., 1997) and in activating type III secretion system (Deakin and Broughton, 2009).
- f) In certain legumes belonging to the dalbergioid and genistoid genera, *Bradyrhizobium* induces nodules in which the host cells divide for several cycles after infection (Vega-Hernández et al., 2001; González-Sama et al., 2004; Kalita et al., 2006; Fedorova et al., 2007). This peculiarity makes the symbiosome of nodules formed by *Bradyrhizobium* in these legumes a prominent candidate in the evolutionary pathway toward “genetically obligatory symbiosis.”

Some Evolutionary and Phylogenetic Considerations about the Macrosymbiont

All angiosperms that perform symbiotic nitrogen-fixing symbioses (except *Gunnera*) are included in the Rosid I clade (Soltis et al., 2000). This clade includes actinorhizal plants and plants that are nodulated by rhizobial bacteria. Recent phylogenetic and molecular data suggest that these nitrogen-fixing plants are derived from a common ancestor of the Rosid I clade with a genetic predisposition for nodulation (Soltis et al., 1995; Pawłowski and Sprent, 2008; Hocher et al., 2011). It was proposed that rhizobial symbioses has evolved four times independently within the Rosid I clade, three times for legumes and once for *Parasponia* (Doyle, 1998; Pawłowski and Sprent, 2008; Sprent, 2008). More recently, it was postulated that there might have been six to seven separate origins of nodulation in legumes (Doyle, 2011).

All plants nodulated by rhizobia are included in the family Leguminosae, except *Parasponia*. Leguminosae comprises more than 700 genera and about 20,000 species (Doyle, 2011), divided into three subfamilies: Caesalpinoideae, Mimosoideae, and Papilionoideae, although the legume taxonomy is currently under revision (Sprent et al., 2017). A key evolutionary study of the Leguminosae family has been performed taking into account molecular and fossil data (Lavin et al., 2005), concluding that legumes evolved about 60 Mya. It was postulated that nodulation could have developed due to an important climatic change at that time, involving an important increase in CO₂ levels that made nitrogen limiting for plant growth (Sprent, 2007). A crucial first step in rhizobia-legume symbiosis is the capacity for mutual recognition and it is thought that this capacity derived from ancient arbuscular mycorrhizal symbiosis (Szczęglowski and Amyot, 2003). In fact, arbuscular mycorrhizal fungi secrete soluble LCO signals (Gough and Cullimore, 2011; Maillet et al., 2011) that are essential for arbuscular mycorrhiza development in legumes, indicating there is a common signaling pathway for both rhizobia-legume and arbuscular mycorrhizal symbioses (Capoen et al., 2009; Markmann and Parniske, 2009; Genre and Russo, 2016).

The macrosymbiont determines the mode of root infection by rhizobia, and the structure and morphology of the nodule. The way of infection has been related to the evolution of legume nodulation, while the structure and morphology of nodules are different among legume clades and may be markers of legume phylogeny (Sprent, 2007, 2009; Sprent et al., 2013). It was considered that the infection processes and nodule structure are more important taxonomic characteristics of legumes than their ability or inability for nodulation (Sprent et al., 2017). An evolutionary scheme of the different rhizobia infection types and the nodule structure of extant legumes has been proposed (for details of this scheme and for examples of the legume nodules on which the model is based see: Sprent, 2007, 2008, 2009; Sprent and James, 2007; and Ibáñez et al., 2017). In this scheme the origin of rhizobia infection could either be through direct epidermal infection or crack infection, which would produce two distinct branches of nodule evolution (**Figure 4**). The more complex evolutionary line involves the formation of transcellular ITs and their entry into some daughter cells of the meristem. In a further evolutionary step, bacteria could be retained in a modified IT (no bacteria released into the host cell and consequently, no symbiosome is formed), as observed in Caesalpinoideae and Papilionoideae legumes. Alternatively, bacteria could be released into the host cell to form the symbiosome. The infection of root hair would be a later, key event in the evolution of the determinate and indeterminate nodules found in Mimosoideae, and in some Papilionoideae and Loteae legumes. All nodules originated in this evolutionary line contain infected and uninjected cells in their nitrogen-fixing zone. About 75% of nodulated legumes, including almost all mimosoids and Caesalpinoideae, and more than 50% of papilionoids, would have followed this strategy.

In the other branch of nodule evolution a few cells are infected by rhizobia and they divide repeatedly (**Figure 4**). The bacteria enter the host cytoplasm in symbiosomes but not via an IT because no such structure is formed. The most

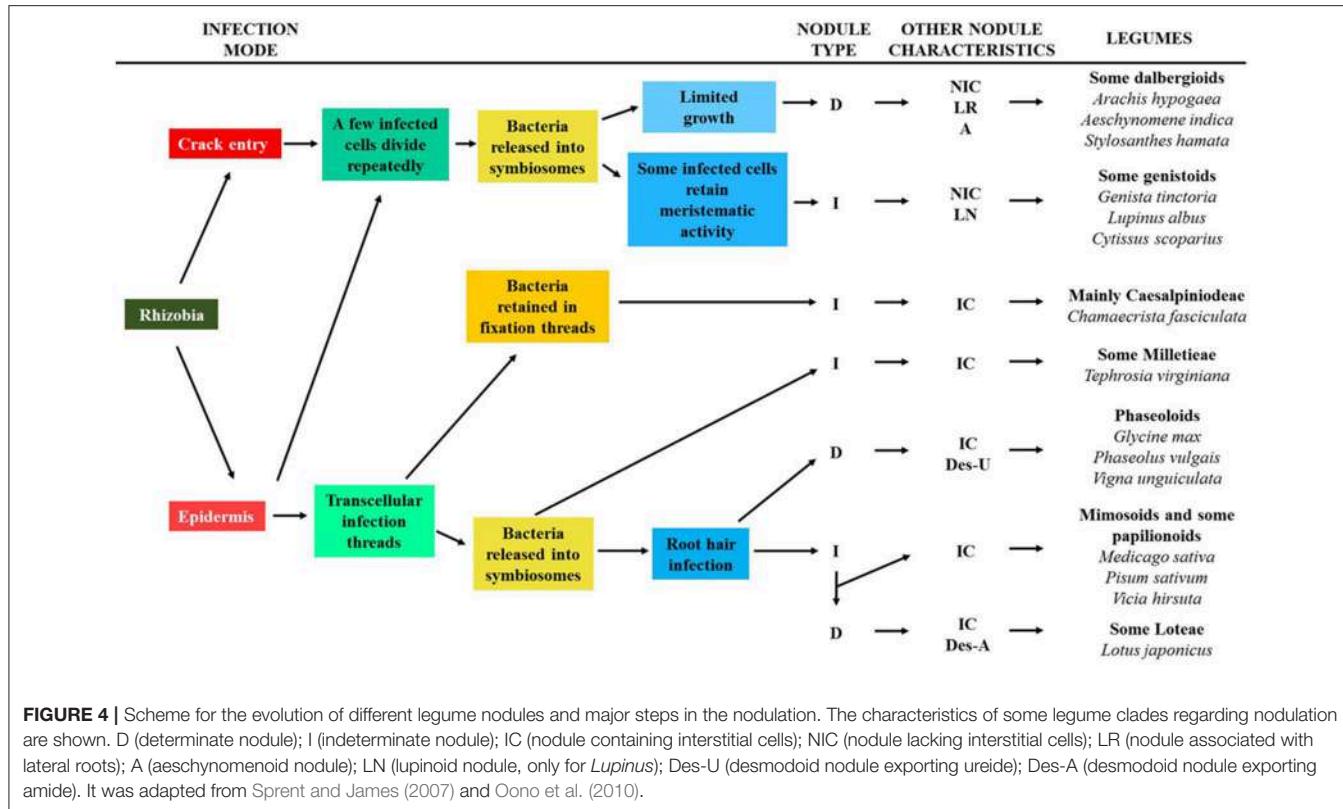


FIGURE 4 | Scheme for the evolution of different legume nodules and major steps in the nodulation. The characteristics of some legume clades regarding nodulation are shown. D (determinate nodule); I (indeterminate nodule); IC (nodule containing interstitial cells); NIC (nodule lacking interstitial cells); LR (nodule associated with lateral roots); A (aeschnomenoid nodule); LN (lupinoid nodule, only for *Lupinus*); Des-U (desmodoid nodule exporting ureide); Des-A (desmodoid nodule exporting amide). It was adapted from Sprent and James (2007) and Oono et al. (2010).

distinctive structural feature of these nodules is that the infected zone is composed of only infected cells. Nodules evolved in this way are only found in Papilionoidae legumes and they include the determinate dalbergioid nodules (crack infection, aeschnomenoid nodules), and those of many Genistae and some Crotalarieae legumes (epidermal infection and some infected cells with meristematic activity: indeterminate nodules and lupinoid nodules).

The Papilionoid crown node arose about 58 Mya, while the genistoid and dalbergioid nodes date to about 56 and 55 Mya, respectively. In comparison, galegoid legumes (a clade that includes *Medicago*, *Vicia*, and *Pisum*) began their spread about 39 Mya and thus, it is the genistoid and dalbergioid that have the oldest origin within the papilionoids (Sprent, 2007; Hane et al., 2017). All legumes that originated later than 40 Mya form their nodules by root hair infection (Sprent, 2009).

In the framework of this review, it is interesting to note some features of the genistoid legume lupin. The *Lupinus* genus includes about 300 species that can be found all over the world. Although they predominantly exist on the American continent and in the Mediterranean area, some Mediterranean species have been introduced into Australia and South Africa. *Lupinus* species colonize different environments and they have particular agronomic potential as they are more tolerant to certain abiotic stresses than other legumes (Fernández-Pascual et al., 2007). These legumes can grow in nitrogen and phosphate depleted soils, and their capability to exploit poor, degraded, contaminated or stress-affected soils, and produce safe, protein-rich seeds make

Lupinus a legume of great interest (Lucas et al., 2015). The *Lupinus* genus has the fastest evolution rate in plants and species from the Andes evolved less than 2 Mya (Hughes and Eastwood, 2006). Moreover, it is the only legume genus known to be unable to establish mycorrhizal symbiosis. A draft genome sequence of *L. angustifolius* was recently obtained (Hane et al., 2017), showing that all mycorrhiza-symbiotic specific genes have been lost, although this species has retained genes commonly required for mycorrhization and for nodulation. The lupin nodule has unique peculiarities (lupinoid) in which a lateral meristem allows the nodule to grow and surround the root (Figure 1). Beyond *Lupinus* spp., this type of nodule has only been found in some species of *Listia* to our knowledge (Yates et al., 2007; Ardley et al., 2013; Sprent et al., 2017). Using *L. albus* and *Bradyrhizobium* as a model, we described the mode of rhizobia infection of lupin roots and other early steps of nodule development in detail (González-Sama et al., 2004). Bacteria infect the root intercellularly, at the junction between the root hair base and an adjacent epidermal cell, and they invade a sub-epidermal outer cortical cell through structurally altered cell wall regions. This infected cell divides repeatedly and together with uninfected dividing cells, the nodule primordium is formed. Thus, the infected zone of the nodule originates through the division of a single infected cortical cell and therefore, the central zone of the lupin nodules has no uninfected cells.

Despite the advantages associated with the colonization of nitrogen poor environments, the ability of many legumes to nodulate may have evolutionary benefits in terms of alleviating

abiotic stress. However, this issue has been little explored. Accordingly, a range of nodulated legumes are found in desert ecosystems and in high altitude areas, suggesting that nitrogen-fixing symbiosis confers an advantage in these ecosystems (Spren and Gehlot, 2010). Nitrogen-fixing legumes make more efficient use of the available water and their fitness is enhanced in arid and semi-arid climates relative to non-fixing plants (Adams et al., 2016). Some putative adaptations of symbiosis to the environment have been reported and for example, some *Mimosa* species prefer to nodulate with certain rhizobia species rather than others, a preference that may be influenced by soil fertility and pH (Elliot et al., 2009; Garau et al., 2009). The semiaquatic legume *Sesbania rostrata* displays phenotypic plasticity for legume nodulation driven by environmental conditions. Thus, *Sesbania* can develop nodules of the indeterminate or determinate type depending on the environmental conditions (Fernández-López et al., 1998). Similarly, rhizobia infection is via an IT in non-flooding conditions whereas flooding switches the infection mechanism to crack entry, favoring nodulation in conditions of water stress in this legume (Goormachtig et al., 2004). On the other hand, the mode of infection may also be determined by the rhizobia in certain legumes. For example the intercellular via was used by a *S. fredii* strain in *Lotus burttii* (Acosta-Jurado et al., 2016) as well as by a strain of *R. leguminosarum* (Gossmann et al., 2012), whereas a *M. loti* strain enters by IT (Gossmann et al., 2012).

ORGANELLE-LIKE CHARACTERISTICS OF THE SYMBIOSOME

Composite Origin and Differentiation of the Symbiosome Membrane Complex

Several biochemical, genetic, and proteomic studies have set out to characterize the composition of the symbiosome (or peribacteroid) membrane and the peribacteroid space (Whitehead and Day, 1997; Panter et al., 2000; Hinde and Trautman, 2002; Saalbach et al., 2002; Wienkoop and Saalbach, 2003; Catalano et al., 2004; Limpens et al., 2009; Clarke et al., 2014, 2015; Emerich and Krishnan, 2014).

Some membrane microdomain-associated proteins can be found in the SM and they seem to play a key role in the regulation of the nodulation process. Flotillins are markers for membrane microdomains called “lipid rafts.” Flotillin genes are induced during early nodulation events in *M. truncatula* (Haney and Long, 2010). Some of these proteins are involved in infection thread invagination and elongation and they could be involved in endocytosis and trafficking of bacteria and nodule organogenesis (Haney and Long, 2010). Flotillin-like genes are induced in soybean nodules (Winzer et al., 1999) and flotillin-like peptides have been identified and isolated from SM of soybean and pea nodules (Panter et al., 2000; Saalbach et al., 2002). A remorin gene encoding another membrane microdomain-associated protein (MtSYMREM1) is specifically and strongly induced during the rhizobial infection and nodule organogenesis of *M. truncatula* (Lefebvre et al., 2010). This protein was located in plasma membrane of ITs and in the SM and may

be a scaffolding protein required for infection and bacterial release into the host cytoplasm (Lefebvre et al., 2010). FWL1 is another interesting membrane microdomain-associated protein identified in soybean symbiosomes (Clarke et al., 2015). FWL1 interacts with remorins, flotillins and other proteins associated with membrane microdomains, regulating legume nodulation (Qiao et al., 2017).

Even at early stages of formation the SM has particular characteristics (Whitehead and Day, 1997), and both the composition and the function of the SM change as it develops (Hinde and Trautman, 2002). In principle, the SM is derived from the plant cell membrane and several plasma membrane markers can be found in the peribacteroid membrane, such as a plasma membrane H⁺-ATPase (Wienkoop and Saalbach, 2003) and the SNARE (N-ethylmaleimide-sensitive factor attachment protein receptor) protein SYP132 (Catalano et al., 2007; Limpens et al., 2009). It is noteworthy that the activation of H⁺-ATPases was also detected in the arbuscular membrane at the AM symbiosis interface (Harrier, 2001).

Symbiosome formation and division induces the activation of the endomembrane system of the host cell (Roth and Stacey, 1989), and it has been proposed that the endoplasmic reticulum (ER) and Golgi vesicles fuse with the SM (Whitehead and Day, 1997; Ivanov et al., 2010; Gavrin et al., 2017). Several proteins from the endomembrane system can be detected in the SM (e.g., cytochrome P450 and a luminal binding protein), and calreticulin, a disulphide-isomerase protein, and some chaperonin-like proteins of the ER have also been identified in symbiosomal fractions and they are probably located in the symbiosome lumen (Saalbach et al., 2002; Wienkoop and Saalbach, 2003; Catalano et al., 2004; Verhaert et al., 2005). Other endomembrane-related proteins in the symbiosome are annexin and syntaxin, which are involved in vesicle transport and secretion, as well as small GTPases involved in the regulation of membrane fusion (Wienkoop and Saalbach, 2003; Catalano et al., 2004; Limpens et al., 2009; Ivanov et al., 2012; Gavrin et al., 2017). It is interesting to note that many of these ER and Golgi proteins, as well as small Rab7 GTPases, have also been found in phagosomes, an organelle compartment of macrophages (Garin et al., 2001; Verhaert et al., 2005), suggesting that symbiosome and phagosome membranes may form in a similar way. Carbohydrate epitopes associated with Golgi-derived glycoproteins and glycolipids have been identified in the inner face of the SM (Perotto et al., 1991). These glycoconjugated molecules, collectively known as the glycocalyx, are involved in physical interactions with the bacterial surface inside the symbiosome and they are important in symbiosome development (Bolaños et al., 2004).

Tonoplast proteins have also been identified in the SM, including a vacuolar H⁺-pyrophosphatase, a vacuolar type H⁺-ATPase (V-ATPase) and an intrinsic tonoplast protein of the Nod26 group (Saalbach et al., 2002; Wienkoop and Saalbach, 2003; Catalano et al., 2004). The presence of active H⁺-ATPases in the SM drives proton accumulation and the establishment of a membrane potential (Whitehead and Day, 1997; Fedorova et al., 1999; Hinde and Trautman, 2002; Clarke et al., 2014). A vacuolar cysteine protease that could be involved in protein

turnover and/or the adaptation to changes in cell turgor was also identified in the symbiosome lumen (Vincent and Brewin, 2000; Vincent et al., 2000). This cysteine protease is also involved in nodule organogenesis and function (Sheokand et al., 2005). The vacuolar SNAREs SYP22 and VTI11 were also found in senescent symbiosomes (Limpens et al., 2009; Emerich and Krishnan, 2014; Gavrin et al., 2014).

Several proteins originating from mitochondria and chloroplasts are also associated with the SM. Among the chloroplast proteins identified are the peripheral membrane protein F1 ATPase α - and β - subunits, the chloroplast outer envelope protein 34 and a chloroplast nucleoid DNA-binding protein. Mitochondrial membrane proteins have also been found, such as a membrane anion channel (porin) and a nucleotide translocator (malate dehydrogenase), as well as mitochondrial processing peptidases, probably located in the symbiosome lumen (Panter et al., 2000; Saalbach et al., 2002; Wienkoop and Saalbach, 2003; Catalano et al., 2004). Bacterial proteins can also be detected in SMs and the peribacteroid lumen, including several nitrogenase components, chaperones, the α -subunit of bacteroid ATP synthase and others (Whitehead and Day, 1997; Saalbach et al., 2002; Catalano et al., 2004; Emerich and Krishnan, 2014).

The SM is a regulated interface with a key role in nutrient exchange between both symbiotic partners, and different types of proteins and transporters are specifically located at this membrane (White et al., 2007; Clarke et al., 2014; Emerich and Krishnan, 2014). The SM has specific integral membrane proteins, such as nodulin 24 (a glycine-rich protein; Sandal et al., 1992), nodulin 26 (an aquaporin; Dean et al., 1999), and others (Clarke et al., 2015). The sulfate transporter gene (*Sst1*) that is expressed in a nodule-specific manner in *Lotus japonicus*, is essential for nodule symbiosis (Krusell et al., 2005). This transporter seems to reside in the SM (Wienkoop and Saalbach, 2003) and it is thought to transport sulfate from the plant cell cytoplasm to the bacteroids (Krusell et al., 2005). Similarly, a proteomic analysis of the SM from nodules of *L. japonicus* revealed the presence of a putative sucrose transporter of the SUC family (Wienkoop and Saalbach, 2003). More recently, another sucrose transporter (MtSWEET11) was proposed to be located at the symbiosome membrane in *M. truncatula* nodules (Kryvoruchko et al., 2016), suggesting the possible transport of sucrose toward the rhizobia. However, specific transporters for some crucial molecules for nitrogen fixation seem not to be located in the SM. For example molybdenum is a key element for the bacteroidal nitrogenase but the molybdate transporter has not been identified in the SM (Tejada-Jiménez et al., 2017). The sulfate transporter *Sst1* (Krusell et al., 2005) could be involved in molybdenum delivery to the symbiosome, as some sulfate transporters can transfer molybdate across membranes (González-Guerrero et al., 2016). Similarly, specific ammonium transporters have not yet been identified in the SM. Although a symbiotic ammonium transporter1 (SAT1) was seen to localize to the SM (Kaiser et al., 1998), it was recently shown that this protein to actually be a membrane-localized basic helix-loop-helix DNA-binding transcription factor involved in ammonium transport (Chiasson et al., 2014). However, ammonium may enter

the symbiosome via the aquaporin-like nodulin 26 channel, or through a cation channel that transports K and Na, as well as by diffusion (Tyerman et al., 1995; Hwang et al., 2010; Courty et al., 2015).

The roles and functions of several proteins located at the symbiosome membrane and the peribacteroid space remain unknown (Kereszt et al., 2011; Emerich and Krishnan, 2014). However, the information available provides some markers of the symbiosome membrane identity. Evidence suggests that secretory pathways play an important role in the formation of the symbiosome and perimicrobial compartments, i.e., an exocytosis-related pathway already present in arbuscular mycorrhizal symbiosis. In fact, an exocytotic pathway for endosymbiosis was defined (Ivanov et al., 2012), providing the first evidence that symbiosomes are generated through exocytosis and that they could therefore be considered apoplastic compartments rather than endocytotic compartments. Rhizobia are confined to plasma membrane protrusions, compartments that rapidly increase in surface area and volume due to microsymbiont expansion. Because the plasma membrane is not elastic and it is unable to stretch more than 3%, exocytosis of new membrane material is crucial to increase the membrane's surface area (Grefen et al., 2011). Membrane fusion is achieved through the action of SNARE proteins in the targeted compartment (t-SNAREs) and the vesicle-associated membrane protein (VAMP or v-SNAREs) that form a SNARE complex, small GTPases of the Rab family that control the transport and docking of vesicles to their target membrane, and Ca^{2+} -sensors from the synaptotagmin group involved in membrane repair (Catalano et al., 2007; Limpens et al., 2009; Ivanov et al., 2010, 2012; Wang et al., 2010; Gavrin et al., 2016, 2017; Harrison and Ivanov, 2017). Briefly, a plasma membrane t-SNARE (SYP123) is present in the SM throughout the life of the symbiosome (from when the rhizobia is released from the IT to symbiosome senescence) and only when the symbiosome has stopped dividing does the SM acquire a late endosomal/vacuolar marker (Rab7), which persists until senescence. At the onset of senescence, the SM acquires a lytic vacuolar identity due to the appearance of the two vacuolar t-SNAREs (SYP22S, VTI11). These SNAREs allow the symbiosome to fuse and form lytic compartments in which the rhizobia are eventually killed. On the other hand, transporters may have a third, new identity for SM (Emerich and Krishnan, 2014) and it could be speculated that a sulfate transporter like-Sst1 should be considered at this point.

The Symbiosome as a Derivative of a Lytic Compartment

The activity of the vacuolar H^+ -ATPase in the symbiosome membrane leads to the accumulation of protons, which should generate an acidic pH in the symbiosome (Whitehead and Day, 1997; Hinde and Trautman, 2002). Several symbiosome enzymes have an acidic optimum pH, including the proteases, acid trehalase, protein protease inhibitor, and alpha-mannosidase isoenzyme II that are typically found in vacuoles (Mellor, 1989; Panter et al., 2000). In fact, certain mutant and senescent bacteroids are degraded by these proteases and glycosidases,

suggesting that the survival of these bacteroids is dependent on them avoiding acid digestion in the symbiosome compartment (Mellor, 1989; Parniske, 2000). As mentioned above, a functional cysteine protease with proteolytic activity has been characterized in the symbiosome lumen (Vincent and Brewin, 2000; Vincent et al., 2000). In 1989, it was proposed that since symbiosomes (which can be considered to be “temporary but independent organelles”) are morphologically different from the plant central vacuole, they may represent organ-specific modifications of lysosomes, analogous to the protein bodies of seeds (Mellor, 1989). Nitrogen activity counteracts the tendency of the ATPase to acidify the lumen of the symbiosome and thus, if the bacteroids stop fixing nitrogen the pH will drop to a level that favors the lysis of the symbiosome. Again, this phenomenon would support the notion of the symbiosome as a modified lysosomal compartment (Brewin, 1991; Hinde and Trautman, 2002).

Symbiosomes do not fuse with lytic vacuoles but they remain as individual units within the cytosol. In fact, it was suggested that vacuolar formation is altered in nodule infected cells in order to allow the expansion of the bacteria in the cytoplasm (Gavrin et al., 2014). Indeed, the vacuoles in infected cells are non-functional and have a neutral pH, or they are degraded (Gavrin et al., 2014, 2016). This facilitates the maintenance of symbiosomes as individual nitrogen-fixing organelles (Limpens et al., 2009; Emerich and Krishnan, 2014).

Rab7 GTPase is thought to be required for the formation of lytic compartments in different organisms (Bucci et al., 2000). In nodules, the plant late endosomal marker Rab7 has been localized in symbiosomes after division stops and it persists until the symbiosome reaches the senescence stage. Therefore, it seems to be involved in symbiosome maintenance (Cheon et al., 1993; Son et al., 2003; Limpens et al., 2009; Clarke et al., 2015). Symbiosome senescence occurs when symbiosomes fuse and form lytic compartments (Hernández-Jiménez et al., 2002; Van de Velde et al., 2006). During senescence, symbiosomes acquire a lytic vacuolar identity, evident through the presence of vacuolar SNAREs and the vacuolar proteins of the HOPS complex at the symbiosome membrane, making it competent for trafficking similar to that of a lytic vacuole (Gavrin et al., 2014).

The Symbiosome Behaves Like a Metabolic Organelle

It has been postulated that metabolic innovations may be important for organelle-producing endosymbiosis (O’Malley, 2015). Rhizobia-legume symbiosis depends on the highly regulated exchange of carbon and nitrogen sources, and nutrients, across the bacteroid and SMs. Specific transporters in these membranes that are critical for symbiosis have been identified through transcriptome and proteome analyses (Udvardi et al., 1988; Vincill et al., 2005; White et al., 2007; Clarke et al., 2014). Most rhizobial species only exhibit highly efficient nitrogen fixation when they are endosymbiotic in the host nodule cells. This suggests that the host plant controls rhizobial nitrogen fixation. It was reported that the host plant has overcome the lack of a bacterial gene necessary for symbiotic nitrogen fixation, a homocitrate synthase gene, a key genetic adaptation

needed to establish efficient nitrogen-fixing symbiosis in legumes and rhizobia. In *L. japonicus*, a legume host nodule-specific homocitrate-synthase is exclusively expressed in infected cells and it supplies homocitrate to the symbiosome. This tricarboxylic acid is an essential component of the iron-molybdenum cofactor of nitrogenase, although it is not itself required for plant metabolism and it is absent from almost all rhizobia species. This homocitrate makes the nitrogen-fixing activity of the endosymbiont possible and it represents an example of the co-evolution of metabolic pathways in the two symbiotic partners (Hakoyama et al., 2009; Terpolilli et al., 2012). It is interesting to note that photosynthetic bradyrhizobia interacting with *Aeschynomene* legumes can synthesize bacterial homocitrate for free-living and symbiotic nitrogen fixation, and that the plant enzyme is not usually induced. *A. caulinodans*, which form nodules with *S. rostrata*, also has this enzyme. These data suggest that different rhizobia-legume symbioses could have co-evolved differently.

A complex amino acid cycle has been observed in pea nodules, whereby the plant cell supplies amino acids to the symbiosome, which can shut down nitrogen fixation, and in return the latter acts like a plant organelle supplying amino acids back to the plant cell for asparagine synthesis. It has been postulated that this exchange induces mutual dependence, preventing the symbiotic relationship from being dominated by the plant and generating selective pressure for the evolution of mutualism (Ludwig et al., 2003). Further studies into amino acid metabolism suggest that symbiosomes in the indeterminate nodules of pea (carrying *Rhizobium leguminosarum* bv. *viciae* as a microsymbiont) and alfalfa (*E. meliloti*), and in the determinate nodules of soybean (*Bradyrhizobium japonicum*), display metabolic dependence on the host for branched-chain amino acids (Prell et al., 2009, 2010; Dunn, 2014). Thus, symbiosomes become symbiotic auxotrophs and they behave like facultative plant organelles. It was suggested that this enabled the plant to control the degree of bacterial infection (Prell et al., 2009, 2010; Terpolilli et al., 2012; Haag et al., 2013).

Nitrogen fixation is uncoupled from bacterial nitrogen stress metabolism in rhizobia-legume symbiosis, such that bacteria generate “excess” ammonia and release this ammonia to the plant, a case of metabolic integration in this symbiosis (Yurgel and Kahn, 2008). The switching to ammonia synthesis by symbiosomes is accompanied by the switching off of ammonia assimilation into amino acids (Patriarca et al., 2002). Because mature bacteroids deplete nitrogen and release ammonia to the plant without assimilation, it was proposed they could be considered as ammoniaplasts (Oldroyd et al., 2011; Downie, 2014).

Processing and Targeting of Symbiosome Proteins

The appearance of an organelle-specific protein import mechanism is considered a key step in the conversion of a symbiont into a permanent organelle (Cavalier-Smith and Lee, 1985; Cavalier-Smith, 1992; Theissen and Martin, 2006; Archibald, 2015). Indeed, chloroplasts and mitochondria

have developed the specific TIC/TOC and TIM/TOM protein transport systems, respectively. The presence of a signal peptide specific for protein targeting is a distinctive trait of cell organelles. Although strictly referring to targeting in order to reimport proteins back from organelle genes that were transferred to the nucleus, it is interesting to consider the specific targeting of protein products to symbiosomes as an organelle-related process. N-terminal sequence comparisons of some SM proteins, like nodulin 26B and HSP60, suggest that N-terminal signal sequences have been removed from these proteins (Panter et al., 2000). Mitochondrial processing peptidases, homologs of which have been identified in the symbiosome, catalyse the cleavage of leader peptides in precursor proteins, although their function in symbiosomes remains unknown (Catalano et al., 2004). The N-terminal processing of proteins may target them to the symbiosome (Panter et al., 2000; Catalano et al., 2004), although these proteins might be targeted to the ER or Golgi, loosing their signal peptide and later being delivered to the SM via the endomembrane system (Panter et al., 2000). A N-terminal signal peptide in nodulin MtNOD25 specifically translocates this protein to the symbiosomes (Hohnjec et al., 2009), the first clear role for a signal peptide in protein targeting to the symbiosome in nodule infected cells. Other nodulins and calcium-binding proteins from *Medicago*, *Vicia*, and *Lupinus* carry signal peptides (Hohnjec et al., 2009; Meckfessel et al., 2012), although no conserved N-targeting signal for SM or symbiosome space proteins has yet been identified. Moreover, these symbiosome targeting signal peptides cannot account for the majority of proteins identified in symbiosomes (Hohnjec et al., 2009). Thus, other targeting systems must be available for protein translocation to the symbiosome (Catalano et al., 2004; Clarke et al., 2014).

Vesicle trafficking to the symbiosome via the endomembrane system is not fully understood. It has been postulated that protein delivery to the symbiosome relies on the plant secretory system (Catalano et al., 2007; Limpens et al., 2009; Ivanov et al., 2010; Maunoury et al., 2010; Mergaert and Kondorosi, 2010; Wang et al., 2010) and it is interesting that proteins lacking plastid-targeting signals might also be targeted to the chloroplast via the secretory system (Bhattacharya et al., 2007 and references therein; Mergaert and Kondorosi, 2010). The syntaxin SNARE SYP132, which localizes to the SM (Catalano et al., 2004), may be involved in site-specific vesicle fusion for the delivery of cargo vesicles to the SM (Catalano et al., 2007). Indeed, some tonoplast proteins involved in symbiosome maturation appear to be retargeted to the symbiosome by a mechanism that involves membrane fusion, as observed in infected cells of *Medicago truncatula* nodules (Gavrin et al., 2014, 2017).

The Host Legume Controls Microsymbiont Differentiation and Proliferation

In *M. truncatula*, the *DMI2* gene that encodes a receptor kinase plays a critical role in the Nod factor signaling cascade during the early stages of nodulation, and it is also a key regulator of symbiosome formation, allowing bacteria to be released from the infection thread into the host cell. In nodules, this kinase is

found in the host cell plasma membrane and in the membrane surrounding the ITs. If *DMI2* expression is compromised in plants, infected nodule cells are occupied by large intracellular ITs that do not release the bacteria rather than organelle-like symbiosomes, a phenotype that is reminiscent of the nodules of primitive legumes and *Parasponia* (Limpens et al., 2005; Op den Camp et al., 2011).

In galegoid legumes of the Inverted Repeat Lacking Clade (IRLC), all of which form indeterminate nodules (like *Medicago* and *Pisum*), a legume family of nodule-specific cysteine-rich (NCR) peptides are targeted to the endosymbiotic bacteria. These peptides are responsible for the bacteroid differentiation that involves the induction of endopolyploidy, cell cycle arrest, terminal differentiation, and a loss of bacterial viability. It was recently demonstrated that a nodule specific thyoredoxin (Trx s1) is targeted to the bacteroid, controlling NCR activity and bacteroid terminal differentiation (Ribeiro et al., 2017). The NCR gene family is estimated to have appeared between 51 and 25 Mya, the time at which IRLC legumes separated from the other legumes (Lavin et al., 2005; Alunni et al., 2007; Yokota and Hayashi, 2011).

All IRLC species tested induce terminal differentiation of their rhizobia endosymbionts, resulting in different morphotypes. NCR genes were also identified in all these species, although the number of NCR peptides was highly variable, ranging from over 630 in *M. truncatula* to only 7 in the most basal IRLC legume *Glycyrrhiza uralensis* (Montiel et al., 2016, 2017). The nodules of this latter legume lack cationic NCR peptides, which could indicate that the ancestral NCRs were neutral or anionic and that they originated from a single evolutionary event in IRLC legumes (Montiel et al., 2017).

It was proposed that the differentiated polyploid bacteroids might have a more efficient metabolism, like polyploid eukaryotic cells (Van de Velde et al., 2010). NCR peptides are derived from antimicrobial, defensin-related peptides, and these antimicrobial peptides have different mechanisms of action and drive different states of bacteroid differentiation (Haag et al., 2013; Maróti and Kondorosi, 2014; Pan and Wang, 2017). This may be an evolved mechanism by which the host legume dominates microsymbiont proliferation (Mergaert et al., 2006; Mergaert and Kondorosi, 2010; Van de Velde et al., 2010; Maróti and Kondorosi, 2014; Yang et al., 2017). NCR peptides optimize bacteroid metabolism and the nitrogen fixation process (Van de Velde et al., 2010), and they control discrimination against incompatible microsymbionts (Yang et al., 2017). It has also been suggested that this control of bacteroid proliferation by the host plant can avoid the spreading of rhizobia to tissues other than the nodule (Mergaert et al., 2006).

Until recently, it was thought that bacteroids of non-galegoid, non-IRLC legumes, do not undergo terminal differentiation nor is their replication restricted. Indeed, they are comparable to free-living bacteria in cell size, DNA content and proliferation (Mergaert et al., 2006). It is noteworthy that in indeterminate nodules of the mimosoid legume *Leucaena glauca* elicited by *Bradyrhizobium*, no NCR peptides have been detected and the bacteroids display a moderate differentiation phenotype; it is an “intermediate” state relative to that of IRLC and non-IRLC

legumes with determinate nodules (Ishihara et al., 2011). The presence of swollen (differentiated) bacteroids has been noted in five out of the six major papilionoid subclades, although each of these subclades also includes species with non-swollen or non-differentiated bacteroids (Oono et al., 2010). Moreover, there was no consistent relationship between nodule type and the host's effects on bacteroid differentiation. Accordingly, it would appear that legumes inducing bacteroid differentiation have evolved independently on five occasions from an ancestral papilionoid legume that hosts non-swollen (non-differentiated) bacteroids (Oono et al., 2010). This repeated evolution of the host's legume traits suggests a possible advantage for the plant in terms of fitness. It has been hypothesized that differentiated bacteroids fix nitrogen more efficiently than non-differentiated bacteroids (Oono et al., 2009, 2010). In fact, Oono and Denison (2010) demonstrated that swollen bacteroids confer net benefits to the host legume due to their more efficient nitrogen fixation and the higher return on the cost of nodule construction (host biomass growth per total nodule mass growth).

It was recently shown that NCR antimicrobial peptides are involved in the permeability of the SM to diverse metabolites. NCR peptides might contribute to metabolic integration between the symbiosome and plant host and in the past, similar antimicrobial peptides may have contributed to the metabolic integration and organellogenesis of mitochondrial and plastid ancestors (Mergaert et al., 2017). This hypothesis emphasizes the importance of metabolic integration in organelle development (see O'Malley, 2015). It was recently discovered that nodules of dalbergioid legume species of the *Aeschynomene* genus (which establish symbiosis with *Bradyrhizobium* spp.) carry polyploid and enlarged bacteroids, and that these plants also express NCR peptides. However, these peptides are not homologous to NCR peptides from IRLC legumes, suggesting an independent evolutionary origin (Czernic et al., 2015).

New plant and bacterial factors that induce bacteroid differentiation remain to be identified (Mergaert et al., 2006; Oono and Denison, 2010; Oono et al., 2010; Van de Velde et al., 2010; Ishihara et al., 2011). A bacterial conserved BacA (bacteroid development factor A) protein that forms an ABC transporter system is produced by rhizobia, and it is required for bacteroid development and survival in IRLC and *Aeschynomene* legumes. BacA may protect rhizobia and bacteroids from the antimicrobial activities of NCR peptides, antagonizing NCR peptides, or it may be involved in the uptake of these antimicrobial peptides by bacteroids (Haag et al., 2013; Guefrachi et al., 2015; Pan and Wang, 2017, and references therein).

It is now assumed that the fate of bacteroids is controlled by the host plant (Mergaert et al., 2006; Maróti and Kondorosi, 2014), although some data suggest that a particular genotype of the microsymbiont might be required, most probably related to their surface polysaccharides. Terminal bacteroid differentiation of *Ensifer fredii* strain HH103 does not take places in nodules of the IRLC legume *Glycyrrhiza uralensis*, (Crespo-Rivas et al., 2016), whereas it does occur when *Mesorhizobium tianshanense* forms the nodules (Montiel et al., 2016). Notably *G. uralensis* is the IRLC legume with the fewest NCR peptides reported to date (Montiel et al., 2017).

Interestingly, species within the genus *Lupinus* may host either swollen (*L. angustifolius*) or non-swollen (*L. albus*, *L. diffusus*, and *L. bicolor*) bacteroids, suggesting that the effects on bacteroid differentiation might have changed during the evolution of the *Lupinus* genus (Oono et al., 2010). Thus, it is possible that the host legumes have regained non-differentiated bacteroids in these latter three species, because bacteroid differentiation is no longer beneficial (for some unknown reason). Alternatively, some rhizobial strains that nodulate *Lupinus* may have evolved traits to overcome host-induced swelling and the loss of reproductive viability (Oono et al., 2010). To our knowledge, there is no data currently available about NCR peptides or any other similar molecules in the nodules of *Lupinus* (or in other legume nodules with dividing infected cells).

Division of Rhizobia-Infected Host Cells

Infected nodule cells are usually post-mitotic and do not divide further. However, one of the most interesting and quite unusual traits for eukaryotic cells is found in certain legume nodules whose host cells can divide after being infected by rhizobia (Figures 2, 3). The division of infected cells containing symbiosomes has been observed in nodules of *Lupinus* spp. and *Genista tinctoria* (genistoid legumes), and also in certain dalbergioid legumes (e.g., *Arachis hypogaea*, *Stylosanthes* spp., *Sarothamnus scoparius*). All these legumes are infected by *Bradyrhizobium* spp. through epidermal infection or crack infection, and the infected zone of their nodules has no uninfected cells (Chandler, 1978; Chandler et al., 1982; Sprent and Thomas, 1984; Tang et al., 1993; Lotocka et al., 2000; Sajnaga et al., 2001; González-Sama et al., 2004; Kalita et al., 2006; Fedorova et al., 2007). Nodules elicited by *Bradyrhizobium* in the genistoid *Chamaecytisus proliferus* (renamed as *Cytisus proliferus*) also contain dividing infected cells (Vega-Hernández et al., 2001). This is the only elongated indeterminate nodule reported to date without uninfected cells in the central infected zone. Root infection of this legume occurs by a singular intercellular mechanism, and ITs are aborted and do not contribute to infection (Vega-Hernández et al., 2001). Therefore, the division of infected cells appears to be a trait restricted to nodules in which infection is independent of ITs, rather than it being influenced by the type of nodule growth (determinate/indeterminate).

Mitochondria and plastids divide in the plant cytoplasm, and cytoskeletal elements not only secure their distribution and movement but also, their correct partitioning between the daughter cells at cytokinesis (King, 2002; Sheahan et al., 2004). Symbiosomes also have the ability to divide in the host cytoplasm (Figure 1), and the accommodation of endosymbionts in host cells involves microtubule and actin microfilament rearrangements (Whitehead et al., 1998; Davidson and Newcomb, 2001a,b; Fedorova et al., 2007; Timmers, 2008; Gavrin et al., 2015; Kitaeva et al., 2016). The conformation of the cytoskeleton in dividing infected cells of legume nodules has only been studied in *L. albus* (Fedorova et al., 2007). We showed that in the infected cells of *L. albus* nodules, symbiosomes are segregated equally between the two daughter cells when

the host plant cell divides, just like other cell organelles, e.g., mitochondria (González-Sama et al., 2004; Fedorova et al., 2007). The cytoskeletal dynamics of infected nodule cells during the cell cycle appear to be relatively normal. In interphase cells, thick cortical arrays of microtubules form a radial network of strands perpendicular to the cell wall to facilitate the migration of organelles and symbiosomes toward the cell periphery (Fedorova et al., 2007). During cell division, symbiosomes concentrate at opposite poles of the cell and do not interfere with the arrangement of microtubules and microfilaments, segregating evenly between the two daughter cells. These cytoskeletal rearrangements in dividing infected cells, along with the detection of an antigen of the molecular motor myosin, suggests that lupin symbiosomes are in contact with and they are driven by the cytoskeleton. Thus, the positioning of symbiosomes in lupin nodule cells seems to depend on the same mechanisms used to segregate genuine plant cell organelles during mitosis (Fedorova et al., 2007). Therefore, in this regard the symbiosome displays significant organelle-like characteristics, unlike symbiosomes from nodules in which infected cells do not divide.

Considerations about Rhizobial Genome Reduction and Gene Transfer to the Nucleus

It has been established that a key event in the evolution from a free-living bacteria to an organelle is the loss of bacterial genes and their transfer to the nucleus of the plant host, a fate that occurred during mitochondrial and chloroplast evolution (Douglas and Raven, 2003; Archibald, 2015). In rhizobia-legume symbiosis, the presence of duplicated prokaryotic genes in the host genome has yet to be reported, although this possibility cannot be overlooked (Raven, 1993).

In the case of nitrogen-fixing rhizobia, the absence of gene transfer to the nucleus may be due to the low oxygen concentrations required by the nitrogenase enzyme, which would generate poor ROS production and mutation rates (Allen and Raven, 1996). Thus, mutation by ROS generation is unlikely to be an evolutionary driving force in the case of symbiosomes. However, strong ROS production has been detected in nodule host cells and in the symbiosome, the electron transport chain of bacteroids generates superoxide radicals and hydrogen peroxide (Matamoros et al., 2003). Oxidation of nitrogenase and ferredoxin in bacteroids also induces ROS generation (Matamoros et al., 2003). Lipid peroxidation has been detected in the SM during senescence (Puppo et al., 1991) and it could be due to the autoxidation of leghemoglobin (a protein controlling the accurate oxygen level in nodules) that is in direct contact with the SM, as well as to a decline in the activity of antioxidants like superoxide dismutase and catalase that are also present in the bacteroid (Puppo et al., 1991; Matamoros et al., 2003). Moreover, ROS generation induces ultrastructural alterations and senescence of symbiosomes (Puppo et al., 2005; Redondo et al., 2009).

No gene loss or genome reduction has been observed in viable symbiotic rhizobia. Symbiotic rhizobia that do not undergo

terminal differentiation are still capable of existing as free-living bacteria. Accordingly, they must be equipped with a number of genes to survive in different environments and to compete with other microorganisms. Moreover, in nodules containing swollen terminally differentiated bacteroids, some non-differentiated bacteria inhabit the apoplastic space and consequently, all the genes necessary for independent life are still retained (Stępkowski and Legocki, 2001). In fact, rhizobia underwent a genome expansion during evolution (MacLean et al., 2007).

An evolutionary pathway has been proposed in symbiotic systems to shift from free-living organisms to facultative symbiosis and to ecologically obligatory symbiosis, usually involving genome expansion (Provorov et al., 2008). The following step in this evolutionary pathway would be “genetically obligatory symbiosis,” which would involve microsymbiont genome simplification or reduction, and the last stage would be a new organism (Provorov et al., 2008). The availability of more recent molecular data from microbes has driven more in-depth studies into the evolutionary transitions in bacterial symbioses, including rhizobia-legume symbiosis (Sachs et al., 2011a,b). Based on phylogenetic analyses, it was hypothesized that transitions from horizontal to obligate vertical transmission of the microsymbiont are driven by the host, the partner that most benefits from these transitions (Sachs et al., 2011b).

CONCLUDING REMARKS

While it has been postulated that organelle development cannot occur in differentiated multicellular organisms (McKay and Navarro-González, 2002), the information presented in this review suggest that the symbiosome might well be considered a step in the co-evolution of legumes and rhizobia toward a nitrogen-fixing organelle. Symbiosomes display features that favor their consideration as nitrogen-fixing organelles, including the host cell’s control of microsymbiont proliferation and differentiation, the composite origin and differentiation of the symbiosome membrane, the retargeting of the host cell’s proteins, or their metabolic behavior. In some legume nodules, such as lupin nodules, host cells seem to perceive their symbiosomes as entities equivalent to their own real organelles. As such, division of infected cells involves the normal cytoskeletal arrangements of regular dividing plant cells, allowing symbiosome segregation into the daughter cells in the same manner as other cell organelles. Symbiosomes in nodules with dividing infected cells might represent a crucial step in the evolution toward real organelles. Nodules with dividing cells form in evolutionarily older legumes in which rhizobial infection does not occur via ITs. In this context, distinct evolutionary routes cannot be ruled out for nodules with non-dividing infected cells and their symbiosomes. In fact, the differences among nodules range from those with nitrogen-fixing ITs and no symbiosomes, to those with infected cells that are able to divide with an organelle-like segregation of the symbiosomes. These could be considered different events or steps in the evolution toward the nitrogen-fixing organelle. In any case, they represent different outcomes or stages in the co-evolution processes, which might or might not continue.

AUTHOR CONTRIBUTIONS

TC, EF, JP, and ML wrote the manuscript. All the authors read and approved the final version of the manuscript.

REFERENCES

- Acosta-Jurado, S., Rodríguez-Navarro, D. N., Kawahara, Y., Fernández Perea, J., Gil-Serrano, A., Jin, H., et al. (2016). *Sinorhizobium fredii* HH103 invades *Lotus burttii* by crack entry in a Nod-factor-and surface polysaccharide-dependent manner. *Mol. Plant Microbe Interact.* 29, 925–937. doi: 10.1094/MPMI-09-16-0195-R
- Adams, D. G., and Duggan, P. S. (2008). Cyanobacteria-bryophyte symbioses. *J. Exp. Bot.* 59, 1047–1058. doi: 10.1093/jxb/ern005
- Adams, M. A., Turnbull, T. L., Sprent, J. I., and Buchmann, N. (2016). Legumes are different: leaf nitrogen, photosynthesis and water use and efficiency. *Proc. Natl. Acad. Sci. U.S.A.* 113, 4098–4103. doi: 10.1073/pnas.1523936113
- Adler, S., Trapp, E. M., Dede, C., Maier, U. G., and Zauner, S. (2014). “*Rhopalodia gibba*: the first steps in the birth of a novel organelle?” in *Endosymbiosis*, ed W. Löffelhardt (Wien: Springer-Verlag), 167–179.
- Allen, F. J., and Raven, J. A. (1996). Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. *J. Mol. Evol.* 42, 482–492. doi: 10.1007/BF02352278
- Alunni, B., Kevei, Z., Redondo-Nieto, M., Kondorosi, A., Mergaert, P., and Kondorosi, E. (2007). Genomic organization and evolutionary insights on *GRP* and *NCR* genes, two large nodule-specific gene families in *Medicago truncatula*. *Mol. Plant Microbe Interact.* 20, 1138–1148. doi: 10.1094/MPMI-20-9-1138
- Archibald, J. M. (2015). Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.* 25, R911–R921. doi: 10.1016/j.cub.2015.07.055
- Arndley, J. K., Reeve, W. G., O’Hara, G. W., Yates, R. J., Dilworth, M. J., and Howieson, J. G. (2013). Nodule morphology, symbiotic specificity and association with unusual rhizobia are distinguishing features of the genus *Listia* within the southern African crotalariaoid clade *Lotononis* s. l. *Ann. Bot.* 112, 1–15. doi: 10.1093/aob/mct095
- Batut, J., Andersson, S. G. E., and O’Callaghan, D. (2004). The evolution of chronic infection strategies in the α-proteobacteria. *Nat. Rev. Microbiol.* 2, 933–945. doi: 10.1038/nrmicro1044
- Becking, J. H. (1979). Root-nodule symbiosis between *Rhizobium* and *Parasponia parviflora* (Ulmaceae). *Plant Soil* 51, 289–296. doi: 10.1007/BF02232892
- Behm, J. E., Geurts, R., and Kiers, E. T. (2014). *Parasponia*: a novel system for studying mutualism stability. *Trends Plant Sci.* 19, 757–763. doi: 10.1016/j.tplants.2014.08.007
- Benson, D. R., and Dawson, J. O. (2007). Recent advances in the biogeography and genealogy of symbiotic *Frankia* and its host plants. *Physiol. Plant.* 130, 318–330. doi: 10.1111/j.1399-3054.2007.00934.x
- Bergman, B., Zheng, W.-W., Klint, J., and Ran, L. (2008). On the origin of plants and relations to contemporary cyanobacterial-plant symbioses. *Plant Biotechnol.* 25, 213–220. doi: 10.5511/plantbiotechnology.25.213
- Bhattacharya, D., Archibald, J. M., Weber, A. P. M., and Reyes-Prieto, A. (2007). How do endosymbionts become organelles? Understanding early events in plastid evolution. *Bioessays* 29, 1239–1246. doi: 10.1002/bies.20671
- Bhuvaneswari, T. V., Turgeon, B. G., and Bauer, W. D. (1980). Early events in the infection of soybean (*Glycine max* L. Merr) by *Rhizobium japonicum*: I. Localization of infective root cells. *Plant Physiol.* 66, 1027–1031. doi: 10.1104/pp.66.6.1027
- Bock, R., and Timmis, J. N. (2008). Reconstructing evolution: gene transfer from plastids to the nucleus. *Bioessays* 30, 556–566. doi: 10.1002/bies.20761
- Bolaños, L., Redondo-Nieto, M., Rivilla, N., Brewin, N. J., and Bonilla, I. (2004). Cell surface interactions of *Rhizobium* bacteroids and other bacterial strains with symbiosomal and peribacteroid membrane components form pea nodules. *Mol. Plant Microbe Interact.* 17, 216–223. doi: 10.1094/MPMI.2004.17.2.216
- Bonfante, P., and Genre, A. (2008). Plants and arbuscular mycorrhizal fungi: an evolutionary-developmental perspective. *Trends Plant Sci.* 13, 492–498. doi: 10.1016/j.tplants.2008.07.001
- Bonnett, H. T., and Silvester, W. B. (1981). Specificity in the *Gunnera-Nostoc* endosymbiosis. *New Phytol.* 89, 121–128. doi: 10.1111/j.1469-8137.1981.tb04754.x
- Bothe, H., Tripp, H. J., and Zehr, J. P. (2010). Unicellular cyanobacteria with a new mode of life: the lack of photosynthetic oxygen evolution allows nitrogen fixation to proceed. *Arch. Microbiol.* 192, 783–790. doi: 10.1007/s00203-010-0621-5
- Boyd, E. S., and Peters, J. W. (2013). New insights into the evolutionary history of biological nitrogen fixation. *Front. Microbiol.* 4:201. doi: 10.3389/fmicb.2013.00201
- Brewin, N. J. (1991). Development of legume root nodules. *Annu. Rev. Cell Biol.* 7, 191–226. doi: 10.1146/annurev.cb.07.110191.001203
- Bucci, C., Thomsen, P., Nicoziani, P., McCarthy, J., and van Deurs, B. (2000). Rab7: a key to lysosome biogenesis. *Mol. Biol. Cell* 11, 467–480. doi: 10.1091/mbc.11.2.467
- Capoen, W., Den Herder, J., Sun, J., Verplancke, C., De Keyser, A., De Rycke, R., et al. (2009). Calcium spiking patterns and the role of the calcium/calmodulin-dependent kinase CaMK in lateral root base nodulation of *Sesbania rostrata*. *Plant Cell* 21, 1526–1540. doi: 10.1105/tpc.109.066233
- Catalano, C. M., Czymmek, K. J., Gann, J. G., and Sherrier, D. J. (2007). *Medicago truncatula* syntaxin SYP132 defines the symbiosome membrane and infection droplet membrane in root nodules. *Planta* 225, 541–550. doi: 10.1007/s00425-006-0369-y
- Catalano, C. M., Lane, W. S., and Sherrier, D. J. (2004). Biochemical characterization of symbiosome membrane proteins from *Medicago truncatula* root nodules. *Electrophoresis* 25, 519–531. doi: 10.1002/elps.200305711
- Cavalier-Smith, T. (1992). The number of symbiotic origins of organelles. *Biosystems* 28, 91–106. doi: 10.1016/0303-2647(92)90011-M
- Cavalier-Smith, T., and Lee, J. J. (1985). Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. *J. Protozool.* 32, 376–379. doi: 10.1111/j.1550-7408.1985.tb04031.x
- Chandler, M. R. (1978). Some observation on infected nodules of *Arachis hypogaea* L. by *Rhizobium*. *J. Exp. Bot.* 29, 749–755. doi: 10.1093/jxb/29.3.749
- Chandler, M. R., Date, R. A., and Roughley, R. J. (1982). Infection and root-nodule development in *Stylosanthes* species by *Rhizobium*. *J. Exp. Bot.* 33, 47–57. doi: 10.1093/jxb/33.1.47
- Chapman, M. J., and Margulis, L. (1998). Morphogenesis by symbiogenesis. *Int. Microbiol.* 1, 319–326.
- Chen, W. M., Moulin, L., Bontemps, C., Vandamme, P., Béna, G., and Boivin-Masson, C. (2003). Legume symbiotic nitrogen fixation by β-Proteobacteria is widespread in nature. *J. Bacteriol.* 185, 7266–7272. doi: 10.1128/JB.185.24.7266–7272.2003
- Cheon, C.-I., Lee, N.-G., Siddique, A.-B. M., Bal, A. K., and Verma, D. P. S. (1993). Roles of plant homologs of Rab1p and Rab7p in the biogenesis of the peribacteroid membrane, a subcellular compartment formed *de novo* during root nodule symbiosis. *EMBO J.* 12, 4125–4135. doi: 10.1002/j.1460-2075.1993.tb06096.x
- Chiasson, D. M., Loughlin, P. C., Mazurkiewicz, D., Mohammadidehcheshmeh, M., Fedorova, E. E., Okamoto, M., et al. (2014). Soybean SAT1 (Symbiotic Ammonium Transporter 1) encodes a bHLH transcription factor involved in nodule growth and NH₄⁺ transport. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4814–4819. doi: 10.1073/pnas.1312801111
- Chriki-Adeeb, R., and Chriki, A. (2016). Estimating divergence times and substitution rates in Rhizobia. *Evol. Bioinform. Online* 12, 87–97. doi: 10.4137/EBO.S39070
- Clarke, V. C., Loughlin, P. C., Day, D. A., and Smith, P. M. C. (2014). Transport processes of the legume symbiosome membrane. *Front. Plant Sci.* 5:699. doi: 10.3389/fpls.2014.00699

ACKNOWLEDGMENTS

This work was supported by grants from MINECO (AGL2013-40758-R, AGL2017-88381-R) and CSIC (i-COOP 2016SU0005).

- Clarke, V. C., Loughlin, P. C., Gavrin, A., Chen, C., Brear, E. M., Day, D. A., et al. (2015). Proteomic analysis of the soybean symbiosome identifies new symbiotic proteins. *Mol. Cell. Proteomics* 14, 1301–1322. doi: 10.1074/mcp.M114.043166
- Cooper, J. E. (2007). Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. *J. Appl. Microbiol.* 103, 1355–1365. doi: 10.1111/j.1365-2672.2007.03366.x
- Courty, P. E., Smith, P., Koegel, S., Redecker, D., and Wipf, D. (2015). Inorganic nitrogen uptake and transport in beneficial plant root-microbe interactions. *Crit. Rev. Plant Sci.* 34, 4–16. doi: 10.1080/07352689.2014.897897
- Crespo-Rivas, J. C., Guefrachi, I., Mok, K. C., Villaécija-Aguilar, J. A., Acosta-Jurado, S., Pierre, O., et al. (2016). *Sinorhizobium fredii* HH103 bacteroids are not terminally differentiated and show altered O-antigen in nodules of the Inverted Repeat-Lacking Clade legume *Glycyrhiza uralensis*. *Environ. Microbiol.* 18, 2392–2404. doi: 10.1111/1462-2920.13101
- Czernic, P., Gully, D., Cartieaux, F., Moulin, L., Guefrachi, I., Patrel, D., et al. (2015). Convergent evolution of endosymbiont differentiation in dalbergioid and IRLC legumes mediated by nodule-specific cysteine rich peptides. *Plant Physiol.* 169, 1254–1265. doi: 10.1104/pp.15.00584
- Davidson, A. L., and Newcomb, W. (2001a). Changes in actin microfilament arrays in developing pea root nodule cells. *Can. J. Bot.* 79, 767–776. doi: 10.1139/b01-046
- Davidson, A. L., and Newcomb, W. (2001b). Organization of microtubules in developing pea root nodule cells. *Can. J. Bot.* 79, 777–786. doi: 10.1139/cjb-79-7-777
- Deakin, W. J., and Broughton, W. J. (2009). Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat. Rev. Microbiol.* 7, 312–320. doi: 10.1038/nrmicro2091
- Dean, R. M., Rivers, R. L., Zeidel, M. L., and Roberts, D. M. (1999). Purification and functional reconstitution of soybean nodulin 26. An aquaporin with water and glycerol transport properties. *Biochemistry* 38, 347–353. doi: 10.1021/bi982110c
- Douglas, A. E., and Raven, J. A. (2003). Genomes at the interface between bacteria and organelles. *Phil. Trans. R. Soc. Lond. B* 358, 5–18. doi: 10.1098/rstb.2002.1188
- Downie, J. A. (2014). Legume nodulation. *Curr. Biol.* 24, R184–R190. doi: 10.1016/j.cub.2014.01.028
- Doyle, J. J. (1998). Phylogenetic perspectives on nodulation: evolving views of plants and symbiotic bacteria. *Trends Plant Sci.* 3, 473–478. doi: 10.1016/S1360-1385(98)01340-5
- Doyle, J. J. (2011). Phylogenetic perspectives on the origins of nodulation. *Mol. Plant Microbe Interact.* 24, 1289–1295. doi: 10.1094/MPMI-05-11-0114
- Drum, R. W., and Pankratz, S. (1965). Fine structure of an unusual cytoplasmic inclusion in the diatom genus *Rhopalodia*. *Protoplasma* 60, 141–149.
- Dudek, J., Rehling, P., and van der Laan, M. (2013). Mitochondrial protein import: common principles and physiological networks. *Biochim. Biophys. Act.* 1833, 274–285. doi: 10.1016/j.bbamcr.2012.05.028
- Dunn, M. F. (2014). Key roles of microsymbiont amino acid metabolism in rhizobia-legume interactions. *Crit. Rev. Microbiol.* 41, 411–451. doi: 10.3109/1040841X.2013.856854
- Dyall, S. D., Brown, M. T., and Johnson, P. J. (2004). Ancient invasions: from endosymbionts to organelles. *Science* 304, 253–257. doi: 10.1126/science.1094884
- Elliot, G. N., Chou, J.-H., Chen, W.-M., Bloemberg, G. V., Bontemps, C., Martínez-Romero, E., et al. (2009). *Burkholderia* spp. are the most competitive symbionts of Mimosa, particularly under N-limited conditions. *Environ. Microbiol.* 11, 762–778. doi: 10.1111/j.1462-2920.2008.01799.x
- Emerich, D. W., and Krishnan, H. B. (2014). Symbiosomes: temporary moonlighting organelles. *Biochem. J.* 460, 1–11. doi: 10.1042/BJ20130271
- Endelin, C. S., and Meeks, J. C. (1983). Pure culture and reconstitution of the *Anthoceros-Nostoc* symbiotic association. *Planta* 158, 157–165. doi: 10.1007/BF00397709
- Fani, R., Gallo, R., and Liò, P. (2000). Molecular evolution of nitrogen fixation: the evolutionary history of the *nifD*, *nifK*, *nifE*, and *nifN* genes. *J. Mol. Evol.* 51, 1–11. doi: 10.1007/s002390010061
- Fedorova, E. E., De Felipe, M. R., Pueyo, J. J., and Lucas, M. M. (2007). Conformation of cytoskeletal elements during the division of infected *Lupinus albus* L. nodule cells. *J. Exp. Bot.* 58, 2225–2236. doi: 10.1093/jxb/erm083
- Fedorova, E., Redondo F. J., Koshiba, T., Pueyo, J. J., de Felipe, M. R., and Lucas, M. M. (2005). Aldehyde oxidase (AO) in the root nodules of *Lupinus albus* and *Medicago truncatula*: identification of AO in meristematic and infection zones. *Mol. Plant-Microbe Interact.* 21, 405–413. doi: 10.1094/MPMI-18-0405
- Fedorova, E., Thomson, R., Whitehead, L. F., Maudoux, O., Udvardi, M. K., and Day, D. A. (1999). Localization of H⁺-ATPases in soybean root nodules. *Planta* 209, 25–32. doi: 10.1007/s004250050603
- Fernández-López, M., Goormachtig, S., Gao, M., D'haeze, W., Van Montagu, M., and Holsters, M. (1998). Ethylene-mediated phenotypic plasticity in root nodule development on *Sesbania rostrata*. *Proc. Natl. Acad. Sci. U.S.A.* 95, 12724–12728. doi: 10.1073/pnas.95.21.12724
- Fernández-Pascual, M., Pueyo, J. J., De Felipe, M. R., Golvano, M. P., and Lucas, M. M. (2007). “Singular features of the *Bradyrhizobium-Lupinus* symbiosis,” in *Dynamic Soil, Dynamic Plant*, eds J. A. Teixeira da Silva and K. Shima (Miki: Global Science Books), 1–16.
- Flores, M., Mavingui, P., Perret, X., Broughton, W. J., Romero, D., Hernández, G., et al. (2000). Prediction, identification, and artificial selection of DNA rearrangements in *Rhizobium*: toward a natural genomic design. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9138–9143. doi: 10.1073/pnas.97.16.9138
- Froussart, E., Bonneau, J., Franche, C., and Boqusz, D. (2016). Recent advances in actonrhizal symbiosis signaling. *Plant Mol. Biol.* 90, 613–622. doi: 10.1007/s11103-016-0450-2
- Garau, G., Yates, R. J., Deiana, P., and Howieson, J. G. (2009). Novel strains of nodulating *Burkholderia* have a role in nitrogen fixation with papilionoid herbaceous legumes adapted to acid, infertile soils. *Soil Biol. Biochem.* 41, 125–134. doi: 10.1016/j.soilbio.2008.10.011
- Garin, J., Diez, R., Kieffer, S., Dermine, J. F., Duclos, S., Gagnon, E., et al. (2001). The phagosome proteome: insight into phagosome functions. *J. Cell Biol.* 152, 165–180. doi: 10.1083/jcb.152.1.165
- Gavrin, A., Chiasson, D., Ovchinnikova, E., Kaiser, B. N., Bisseling, T., and Fedorova, E. E. (2016). VAMP721a and VAMP721d are important for pectin dynamics and release of bacteria in soybean nodules. *New Phytol.* 210, 1011–1021. doi: 10.1111/nph.13837
- Gavrin, A., Jansen, V., Ivanov, S., Bisseling, T., and Fedorova, E. (2015). ARP32/3-Mediated actin nucleation associated with symbiosomes membrane is essential for the development of symbiosomes in infected cells of *Medicago truncatula* nodules. *Mol. Plant Microbe-Interact.* 28, 605–614. doi: 10.1094/MPMI-12-14-0402-R
- Gavrin, A., Kaiser, B. N., Geiger, D., Tyerman, S. D., Wen, Z., Bisseling, T., et al. (2014). Adjustment of host cells for accommodation of symbiotic bacteria: vacuole defunctionalisation, HOPS suppression, and TIP1g retargeting in *Medicago*. *Plant Cell* 26, 3809–3822. doi: 10.1105/tpc.114.128736
- Gavrin, A., Kulikova, O., Bisseling, T., and Fedorova, E. E. (2017). Interface symbiotic membrane formation in root nodules of *Medicago truncatula*: the role of Synaptotagmins MtSyt1, MtSyt2 and MtSyt3. *Front. Plant Sci.* 8:201. doi: 10.3389/fpls.2017.00201
- Genre, A., and Russo, G. (2016). Does a common pathway transduce symbiotic signals in plant-microbe interactions? *Front. Plant Sci.* 7:96. doi: 10.3389/fpls.2016.00096
- Geurts, R., Fedorova, E., and Bisseling, T. (2005). Nod factor signalling genes and their function in the early stages of *Rhizobium* infection. *Curr. Opin. Plant Biol.* 8, 346–352. doi: 10.1016/j.pbi.2005.05.013
- Geurts, R., Lillo, A., and Bisseling, T. (2012). Exploiting an ancient signalling machinery to enjoy a nitrogen fixing symbiosis. *Curr. Opin. Plant Biol.* 15, 438–443. doi: 10.1016/j.pbi.2012.04.004
- Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., et al. (2007). Legume symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science* 316, 1307–1312. doi: 10.1126/science.1139548
- González-Guerrero, M., Escuderom, V., Saéz, Á., and Tejada-Jiménez, M. (2016). Transition metal transport in plants and associated endosymbionts: arbuscular mycorrhizal fungi and rhizobia. *Front. Plant Sci.* 7:1088. doi: 10.3389/fpls.2016.01088
- González-Sama, A., Lucas, M. M., De Felipe, M. R., and Pueyo, J. J. (2004). An unusual infection mechanism and nodule morphogenesis in white lupin (*Lupinus albus*). *New Phytol.* 163, 371–380. doi: 10.1111/j.1469-8137.2004.01121.x
- Goormachtig, S., Capoen, W., James, E. K., and Holsters, M. (2004). Switch from intracellular to intercellular invasion during water stress-tolerant

- legume nodulation. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6303–6308. doi: 10.1073/pnas.0401540101
- Gossmann, J. A., Markmann, K., Brachmann, A., Rose, L. E., and Parniske, M. (2012). Polymorphic infection and organogenesis patterns induced by a *Rhizobium leguminosarum* isolate from *Lotus* root nodules are determined by the host genotype. *New Phytol.* 196, 561–573. doi: 10.1111/j.1469-8137.2012.04281.x
- Gough, C., and Cullimore, J. (2011). Lipo-chitooligosaccharide signaling in endosymbiotic plant-microbe interactions. *Mol. Plant Microbe Interact.* 24, 867–878. doi: 10.1094/MPMI-01-11-0019
- Gray, M. W. (2012). Mitochondrial evolution. *Cold Spring Harb. Perspect. Biol.* 4:a011403. doi: 10.1101/cshperspect.a011403
- Grefen, C., Honsbein, A., and Blatt, M. R. (2011). Ion transport, membrane traffic and cellular volume control. *Curr. Opin. Plant Biol.* 14, 332–339. doi: 10.1016/j.pbi.2011.03.017
- Gtari, M., Ghodhbane-Gtari, F., Nouiou, I., Ktari, A., Hezbi, K., Mimouni, W., et al. (2015). Cultivating the uncultured: growing the recalcitrant cluster-2 *Frankia* strains. *Sci. Rep.* 5:13112. doi: 10.1038/srep13112
- Gualtieri, G., and Bisseling, T. (2000). The evolution of nodulation. *Plant Mol. Biol.* 42, 181–194. doi: 10.1023/A:1006396525292
- Guefrachi, I., Pierre, O., Timchenko, T., Alunni, B., Barrière, Q., Czernic, P., et al. (2015). *Bradyrhizobium* BclA is a peptide transporter required for bacterial differentiation in symbiosis with *Aeschynomene* legumes. *Mol. Plant Microbe Interact.* 28, 1155–1166. doi: 10.1094/MPMI-04-15-0094-R
- Guo, X., Flores, M., Mavingui, P., Fuentes, S. I., Hernández, G., Dávila, G., et al. (2003). Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Res.* 13, 1810–1817. doi: 10.1101/gr.1260903
- Haag, A. F., Arnold, M. F. F., Myka, K. K., Kerscher, B., Dall'Angelo, S., Zanda, M., et al. (2013). Molecular insights into bacteroid development during *Rhizobium*-legume symbiosis. *FEMS Microbiol. Rev.* 37, 364–383. doi: 10.1111/1574-6976.12003
- Hakoyama, T., Niimi, K., Watanabe, H., Tabata, R., Matsubara, J., Sato, S., et al. (2009). Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation. *Nature* 462, 514–518. doi: 10.1038/nature08594
- Hane, J. K., Ming, Y., Kamphuis, L. G., Nelson, M. N., Garg, G., Atkins, C. A., et al. (2017). A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*), an emerging health food: insights into plant-microbe interactions and legume evolution. *Plant Biotechnol. J.* 15, 318–330. doi: 10.1111/pbi.12615
- Haney, C. H., and Long, S. R. (2010). Plant flotillins are required for infection by nitrogen-fixing bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 107, 478–483. doi: 10.1073/pnas.0910081107
- Harrier, L. A. (2001). The arbuscular mycorrhizal symbiosis: a molecular review of the fungal dimension. *J. Exp. Bot.* 52, 469–478. doi: 10.1093/jexbot/52.suppl_1.469
- Harrison, M., and Ivanov, S. (2017). Exocytosis for endosymbiosis: membrane trafficking pathways for development of symbiotic membrane compartments. *Curr. Opin. Plant Biol.* 38, 101–108. doi: 10.1016/j.pbi.2017.04.019
- Hernández-Jiménez, M. J., Lucas, M. M., and de Felipe, R. (2002). Antioxidant defence and damage in senescing lupin nodules. *Plant Physiol. Biochem.* 40, 645–657. doi: 10.1016/S0981-9428(02)01422-5
- Hinde, R., and Trautman, D. A. (2002). “Symbiosomes” in *Symbiosis: Mechanisms and Model Systems*, ed J. Seckbach (Dordrecht: Staempfli Publishers Ltd.) 209–220.
- Hocher, V., Alloisio, N., Auguy, F., Fournier, P., Doumas, P., Pujic, P., et al. (2011). Transcriptomics of actinorhizal symbioses reveals homologs of the whole common symbiotic signaling cascade. *Plant Physiol.* 156, 700–711. doi: 10.1104/pp.111.174151
- Hohnjec, N., Lenz, F., Fehlberg, V., Vieweg, M. F., Baier, M. C., Hause, B., et al. (2009). The signal peptide of the *Medicago truncatula* modular nodulin MtNOD25 operates as an address label for the specific targeting of proteins to nitrogen-fixing symbiosomes. *Mol. Plant Microbe Interact.* 22, 63–72. doi: 10.1094/MPMI-21-1-0063
- Hughes, C., and Eastwood, R. (2006). Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10334–10339. doi: 10.1073/pnas.0601928103
- Hurek, T., Reinhold-Hurek, B., van Montagu, M., and Kellebenger, E. (1994). Root colonization and systemic spreading of *Azoarcus* sp. Strain BH72 in grasses. *J. Bacteriol.* 176, 1913–1923. doi: 10.1128/jb.176.7.1913-1923.1994
- Hwang, J. H., Ellingson, S. R., and Roberts, D. M. (2010). Ammonia permeability of the soybean nodulin 26 channel. *FEBS Lett.* 584, 4339–4343. doi: 10.1016/j.febslet.2010.09.033
- Ibáñez, F., Wall, L., and Fabra, A. (2017). Starting points in plant-bacteria nitrogen-fixing symbioses: intercellular invasion of the roots. *J. Exp. Bot.* 68, 1905–1918. doi: 10.1093/jxb/erw387
- Ishihara, H., Koriyama, H., Osawa, A., Zehirov, G., Yamaura, M., Kucho, K.-I., et al. (2011). Characteristics of bacteroids in indeterminate nodules of the leguminous tree *Leucaena glauca*. *Microbes Environ.* 26, 156–159. doi: 10.1264/jsmc2.ME11104
- Ivanov, S., Fedorova, E. E., Limpens, E., De Mita, S., Genre, A., Bonfante, P., et al. (2012). *Rhizobium*-legume symbiosis shares an exocytotic pathway required for arbuscule formation. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8316–8321. doi: 10.1073/pnas.1200470109
- Ivanov, S., Fedorova, E., and Bisseling, T. (2010). Intracellular plant microbe associations: secretory pathways and the formation of perimicrobial compartments. *Curr. Opin. Plant Biol.* 13, 372–377. doi: 10.1016/j.pbi.2010.04.005
- Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., and Walker, G. C. (2007). How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. *Nat. Rev. Microbiol.* 5, 619–633. doi: 10.1038/nrmicro1705
- Judicial Commission of the International Committee on Systematics of Prokaryotes (2008). The genus name *Sinorhizobium* Chen et al. 1988 is a later synonym of *Ensifer* Casida 1982 and is not conserved over the latter genus name, and the species name ‘*Sinorhizobium adhaerens*’ is not validly published. Opinion 84. *Int. J. Syst. Evol. Microbiol.* 58:1973. doi: 10.1099/ijst.0.2008/005991-0
- Kaiser, B. N., Finnegan, P. M., Tyerman, S. D., Whitehead, L. F., Bergersen, F. J., Day, D. A., et al. (1998). Characterization of an ammonium transport protein from the peribacteroid membrane of soybean nodules. *Science* 281, 1202–1206. doi: 10.1126/science.281.5380.1202
- Kalita, M., Stepkowski, T., Lotocka, B., and Malek, W. (2006). Phylogeny of nodulation genes and symbiotic properties of *Genista tinctoria* bradyrhizobia. *Arch. Microbiol.* 186, 87–97. doi: 10.1007/s00203-006-0124-6
- Kaneko, T., Maita, H., Hirakawa, H., Uchiike, N., Minamisawa, K., Watanabe, A., et al. (2011). Complete genome sequence of the soybean symbiont *Bradyrhizobium japonicum* strain USDA6. *Genes* 2, 763–787. doi: 10.3390/genes2040763
- Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., et al. (2002). Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res.* 9, 189–197. doi: 10.1093/dnarese/9.6.189
- Keeling, P. J. (2004). Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.* 91, 1481–1493. doi: 10.3732/ajb.91.10.1481
- Keeling, P. J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 729–748. doi: 10.1098/rstb.2009.0103
- Keeling, P. J., and Archibald, J. M. (2008). Organelle evolution: what's in a name? *Curr. Biol.* 18, R345–R347. doi: 10.1016/j.cub.2008.02.065
- Keeling, P. J., McCutcheon, J. P., and Doolittle, W. F. (2015). Symbiosis becoming permanent: survival of the luckiest. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10101–10103. doi: 10.1073/pnas.1513346112
- Kereszt, A., Mergaert, P., and Kondorosi, E. (2011). Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims? *Mol. Plant Microbe Interact.* 24, 1300–1309. doi: 10.1094/MPMI-06-11-0152
- Kiers, E. T., and West, S. A. (2015). Evolving new organisms via symbiosis. *Science* 348, 392–394. doi: 10.1126/science.aaa9605
- King, S. M. (2002). Dyneins motor on in plants. *Traffic* 3, 930–931. doi: 10.1034/j.1600-0854.2002.31208.x
- Kistner, C., and Parniske, M. (2002). Evolution of signal transduction in intracellular symbiosis. *Trends Plant Sci.* 7, 511–518. doi: 10.1016/S1360-1385(02)02356-7
- Kitaeva, A. B., Demchenko, K. N., Tikhonovich, I. A., Timmers, A. C. J., and Tsyanov, V. E. (2016). Comparative analysis of the tubulin cytoskeleton organization in nodules of *Medicago truncatula* and *Pisum sativum*: bacterial release and bacteroid positioning correlate with characteristic

- microtubule rearrangements. *New Phytol.* 210, 168–183. doi: 10.1111/nph.13792
- Kleine, T., Maier, U. G., and Leister, D. (2009). DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu. Rev. Plant Biol.* 60, 115–138. doi: 10.1146/annurev.arplant.043008.092119
- Kneip, C., Lockhart, P., Voß, C., and Maier, U.-G. (2007). Nitrogen fixation in eukaryotes – New models for symbiosis. *BMC Evol. Biol.* 7:55. doi: 10.1186/1471-2148-7-55
- Kneip, C., Voß, C., Lockhart, P. J., and Maier, U. G. (2008). The cyanobacterial endosymbiont of the unicellular algae *Rhopalodia gibba* shows reductive genome evolution. *BMC Evol. Biol.* 28:30. doi: 10.1186/1471-2148-8-30
- Kouchi, H., Imaizumi-Anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., et al. (2010). How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. *Plant Cell Physiol.* 51, 1381–1397. doi: 10.1093/pcp/pcq107
- Krusell, L., Krause, K., Ott, T., Desbrosses, G., Krämer, U., Sato, S., et al. (2005). The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. *Plant Cell* 17, 1625–1636. doi: 10.1105/tpc.104.030106
- Kryvoruchko, I. S., Sinharoy, S., Torres-Jerez, I., Sosso, D., Pislaru, C. I., Guan, D., et al. (2016). MtSWEET11, a nodule-specific sucrose transporter of *Medicago truncatula*. *Plant Physiol.* 171, 554–565. doi: 10.1104/pp.15.01910
- Kucho, K.-I., Hay, A.-E., and Normand, P. (2010). The determinants of the actinorhizal symbiosis. *Microbes Environ.* 25, 241–252. doi: 10.1264/jsme2.ME10143
- Kündig, C., Hennecke, H., and Göttfert, M. (1993). Correlated physical and genetic map of the *Bradyrhizobium japonicum* 110 genome. *J. Bacteriol.* 175, 613–622. doi: 10.1128/jb.175.3.613-622.1993
- Kutschera, U., and Niklas, K. J. (2005). Endosymbiosis, cell evolution, and speciation. *Theory Biosci.* 124, 1–24. doi: 10.1016/j.thbio.2005.04.001
- Lang, B. F., Burger, G., O'Kelly, C. J., Cedergren, R., Golding, G. B., Lemieux, C., et al. (1997). An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387, 493–497. doi: 10.1038/387493a0
- Lang, B. F., Gray, M. W., and Burger, G. (1999). Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.* 33, 351–397. doi: 10.1146/annurev.genet.33.1.351
- Larsson, J. (2011). *Cyanobacterial Genome Evolution Subsequent to Domestication by a Plant (Azolla)*. Doctoral Thesis, Stockholm University; Faculty of Science, 61.
- Lavin, M., Herendeen, P. S., and Wojciechowski, M. F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Syst. Biol.* 54, 574–594. doi: 10.1080/10635150590947131
- Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Hervé, C., Tóth, K., et al. (2010). A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2343–2348. doi: 10.1073/pnas.0913320107
- Limpens, E., Ivanov, S., van Esse, W., Voets, G., Fedorova, E., and Bisseling, T. (2009). *Medicago N₂-fixing symbiosomes acquire the endocytic identity marker Rab7 but delay the acquisition of vacuolar identity*. *Plant Cell* 21, 2811–2828. doi: 10.1105/tpc.108.064410
- Limpens, E., Mirabella, R., Fedorova, E., Franken, C., Franssen, H., Bisseling, T., et al. (2005). Formation of organelle-like N₂-fixing symbiosomes in legume root nodules is controlled by DM12. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10375–10380. doi: 10.1073/pnas.0504284102
- Lloret, L., and Martínez-Romero, E. (2005). Evolución y filogenia de *Rhizobium*. *Rev. Latinoam. Microbiol.* 47, 43–60.
- Lodwig, E. M., Hosie, A. H. F., Bourdès, A., Findlay, K., Allaway, D., Karunakaran, R., et al. (2003). Amino-acid cycling drives nitrogen fixation in the legume-*Rhizobium* symbiosis. *Nature* 422, 722–726. doi: 10.1038/nature01527
- Loh, J., Garcia, M., and Stacey, G. (1997). NodV and NodW, a second flavonoid recognition system regulating nod gene expression in *Bradyrhizobium japonicum*. *J. Bacteriol.* 179, 3013–3020. doi: 10.1128/jb.179.9.3013-3020
- Lotocka, B., Kopcińska, J., Gorecka, M., and Golinowski, W. (2000). Formation and abortion of root nodule primordial in *Lupinus luteus* L. *Acta Biol. Cracov. Ser. Bot.* 42, 87–102.
- Lucas, M. M., Stoddard, F. L., Annicchiarico, P., Frías, J., Martínez-Villaluenga, C., Sussmann, D., et al. (2015). The future of lupin as a protein crop in Europe. *Front. Plant Sci.* 6:705. doi: 10.3389/fpls.2015.00705
- Luo, L., and Lu, D. (2014). Immunosuppression during *Rhizobium-legume* symbiosis. *Plant Signal. Behav.* 9:e28197. doi: 10.4161/psb.28197
- MacLean, A. M., Finan, T. M., and Sadowsky, M. J. (2007). Genomes of the symbiotic nitrogen-fixing bacteria of legumes. *Plant Physiol.* 144, 615–622. doi: 10.1104/pp.107.101634
- 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. doi: 10.1038/nature09622
- Margulis, L., and Chapman, M. J. (1998). Endosymbioses: cyclical and permanent in evolution. *Trends Microbiol.* 6, 342–345. doi: 10.1016/S0966-842X(98)01325-0
- Marin, B., Nowack, E. C., and Melkonian, M. (2005). A plastid in the making: evidence for a second primary endosymbiosis. *Protist* 156, 425–430. doi: 10.1016/j.protis.2005.09.001
- Markmann, K., and Parniske, M. (2009). Evolution of root endosymbiosis with bacteria: how novel are nodules? *Trends Plant Sci.* 14, 77–86. doi: 10.1016/j.tplants.2008.11.009
- Maróti, G., and Kondorosi, E. (2014). Nitrogen-fixing *Rhizobium*-legume symbiosis: are polyploidy and host peptide-governed symbiont differentiation general principles of endosymbiosis? *Front. Microbiol.* 5:326. doi: 10.3389/fmicb.2014.00326
- Martens, M., Delaere, M., Coopman, R., De Vos, P., Gillis, M., and Willems, A. (2007). Multilocus sequence analysis of *Ensifer* and related taxa. *Int. J. Syst. Evol. Microbiol.* 57, 489–503. doi: 10.1099/ijts.0.64344-0
- Martin, W. F., Garg, S., and Zimorski, V. (2015). Endosymbiotic theories for eukaryote origin. *Phil. Trans. R. Soc. B* 370:20140330. doi: 10.1098/rstb.2014.0330
- Martin, W., Gierl, A., and Seadler, H. (1989). Molecular evidence for pre-Cretaceous angiosperm origins. *Nature* 339, 46–48. doi: 10.1038/339046a0
- Martínez-Romero, E. (2009). Coevolution in *Rhizobium*-legume symbiosis? *DNA Cell Biol.* 28, 361–370. doi: 10.1089/dna.2009.0863
- Masson-Boivin, C., Giraud, E., Perret, X., and Batut, J. (2009). Establishing nitrogen-fixing symbiosis with legumes: how many *Rhizobium* recipes? *Trends Microbiol.* 17, 458–466. doi: 10.1016/j.tim.2009.07.004
- Matamoros, M. A., Dalton, D. A., Ramos, J., Clemente, M. R., Rubio, M. C., and Becana, M. (2003). Biochemistry and molecular biology of antioxidants in the rhizobia-legume symbiosis. *Plant Physiol.* 133, 499–509. doi: 10.1104/pp.103.025619
- Maunoury, N., Redondo-Nieto, M., Bourcy, M., Van de Velde, W., Alunni, B., Laporte, P., et al. (2010). Differentiation of symbiotic cells and endosymbionts in *Medicago truncatula* nodulation are coupled to two transcriptome-switches. *PLoS ONE* 5:e9519. doi: 10.1371/journal.pone.0009519
- McCUTCHEON, J. P. (2016). From microbiology to cell biology: when intracellular bacterium becomes part of its host cell. *Curr. Opin. Cell Biol.* 41, 132–136. doi: 10.1016/j.jceb.2016.05.008
- McFadden, G. I. (1999). Endosymbiosis and evolution of the plant cell. *Curr. Opin. Plant Biol.* 2, 513–519. doi: 10.1016/S1369-5266(99)00025-4
- McKay, C. P., and Navarro-González, R. (2002). “The absence of nitrogen-fixing organelles due to timing of the nitrogen crisis,” in *Symbiosis: Mechanisms and Model Systems*, ed J. Seckbach (Dordrecht: Kluwer Academic Publishers), 221–228.
- Meckfessel, M. H., Blancaflor, E. B., Plunkett, M., Dong, Q., and Dickstein, R. (2012). Multiple domains in MtENOD8 protein including the signal peptide target it to the symbiosome. *Plant Physiol.* 159, 299–310. doi: 10.1104/pp.111.191403
- MELLOR, R. B. (1989). Bacteroids in the *Rhizobium*-legume symbiosis inhabit a plant internal lytic compartment: implications for other microbial endosymbiosis. *J. Exp. Bot.* 40, 831–839. doi: 10.1093/jxb/40.8.831
- Mergaert, P., and Kondorosi, E. (2010). “Role of the secretory pathway and antimicrobial peptides in the accommodation and differentiation of endosymbiotic *Rhizobium* bacteria in *Medicago* nodules,” in *Biological Nitrogen Fixation and Plant-Associated Microorganisms*, ed M. Becana (Zaragoza: Sociedad Española de Fijación de Nitrógeno), 109–111.
- Mergaert, P., Kikuchi, Y., Shigenobu, S., and Nowack, E. C. M. (2017). Metabolic integration of bacterial endosymbionts through antimicrobial peptides. *Trends Microbiol.* 25, 703–712. doi: 10.1016/j.tim.2017.04.007
- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., et al. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the

- Rhizobium-legume symbiosis.* *Proc. Natl. Acad. Sci. U.S.A.* 103, 5230–5235. doi: 10.1073/pnas.0600912103
- Miller, I. M., and Baker, D. D. (1986). Nodulation of actinorhizal plants by *Frankia* strains capable of both root hair infection and intercellular penetration. *Protoplasma* 131, 82–91. doi: 10.1007/BF01281689
- Minamisawa, K., Isawa, T., Nakatsuka, Y., and Ichikawa, N. (1998). New *Bradyrhizobium japonicum* strains that possess high copy numbers of the repeated sequence RS_a. *Appl. Environ. Microbiol.* 64, 1845–1851.
- Mithöfer, A. (2002). Suppression of plant defence in rhizobia-legume symbiosis. *Trends Plant Sci.* 7, 440–444. doi: 10.1016/S1360-1385(02)02336-1
- Molouba, F., Lorquin, J., Willems, A., Hoste, B., Giraud, E., Dreyfus, B., et al. (1999). Photosynthetic bradyrhizobia from *Aeschynomene* spp. are specific to stem-nodulated species and form a separate 16S ribosomal DNA restriction fragment length polymorphism group. *Appl. Environ. Microbiol.* 65, 3084–3094.
- Montiel, J., Downie, J. A., Farkas, A., Bihari, P., Herczeg, R., Bálint, B., et al. (2017). Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proc. Natl. Acad. Sci. U.S.A.* 19, 5041–5046. doi: 10.1073/pnas.1704217114
- Montiel, J., Szűcs, A., Boboescu, I. Z., Gherman, V. D., Kondorosi, E., and Kereszt, A. (2016). Terminal bacteroid differentiation is associated with variable morphological changes in legume species belonging to the Inverted-Repeat-Lacking clade. *Mol. Plant Microbe Interact.* 29, 210–219. doi: 10.1094/MPMI-09-15-0213-R
- Morton, R. A. (2002). “Comparison of chromosomal genes from *M. loti* and *S. meliloti* suggest an ancestral genome,” in *Nitrogen Fixation: Global Perspectives*, eds M. F. Tulrlough, M. R. O’Brien, D. B. Layzell, J. K. Vessey, and W. Newton (Wallingford: CBAI Publishing), 55–58.
- Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. (2001). Nodulation of legumes by members of the β-subclass of Proteobacteria. *Nature* 411, 948–950. doi: 10.1038/35082070
- Mousavi, S. A., Österman, J., Wahlberg, N., Nesme, X., Lavire, C., Vial, L., et al. (2014). Phylogeny of the *Rhizobium-Allorhizobium-Agrobacterium* clade supports the delineation of *Neorhizobium* gen. nov. *Syst. Appl. Microbiol.* 37, 208–215. doi: 10.1016/j.syapm.2013.12.007
- Mousavi, S. A., Willems, A., Nesme, X., de Lajudie, P., and Lindström, K. (2015). Revised phylogeny of Rhizobiaceae: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations. *Syst. Appl. Microbiol.* 38, 84–90. doi: 10.1016/j.syapm.2014.12.003
- Nakayama, T., and Archibald, J. M. (2012). Evolving a photosynthetic organelle. *BMC Biol.* 10:35. doi: 10.1186/1741-7007-10-35
- Nakayama, T., and Inagaki, Y. (2014). Unique genome evolution in an intracellular N₂-fixing symbiont of a rhopalodiacean diatom. *Acta Soc. Bot. Pol.* 83, 409–413. doi: 10.5586/asbp.2014.046
- Nakayama, T., and Ishida, K. (2009). Another acquisition of a primary photosynthetic organelle is underway in *Paulinella chromatophora*. *Curr. Biol.* 19, R284–R285. doi: 10.1016/j.cub.2009.02.043
- Nakayama, T., Ikegami, Y., Nakayama, T., Ishida, K.i., Inagaki, Y., and Inouye, I. (2011). Spheroid bodies in rhopalodiacean diatoms were derived from a single endosymbiotic cyanobacterium. *J. Plant Res.* 124, 93–97. doi: 10.1007/s10265-010-0355-0
- Nakayama, T., Kamikawa, R., Tanifugi, G., Kashiyama, Y., Ohkouchi, N., Archibald, J. M., et al. (2014). Complete genome of a non-photosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. *Proc. Natl. Acad. Sci. U.S.A.* 11, 11407–11412. doi: 10.1073/pnas.1405222111
- Newcomb, W. (1981). “Nodule morphogenesis and differentiation,” in *Biology of the Rhizobiaceae*, eds K. L. Giles and A. G. Atherly (New York, NY: Academic Press), 247–297.
- Nowack, E. C. (2014). *Paulinella chromatophora* – rethinking the transition from endosymbiont to organelle. *Acta Soc. Botan. Polon.* 83, 387–397. doi: 10.5586/asbp.2014.049
- Nowack, E. C., and Grossman, A. R. (2012). Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5340–5345. doi: 10.1073/pnas.1118800109
- Nowack, E. C., Melkonian, M., and Glöckner, G. (2008). Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr. Biol.* 18, 410–418. doi: 10.1016/j.cub.2008.02.051
- Oke, V., and Long, S. R. (1999). Bacteroid formation in the *Rhizobium-legume* symbiosis. *Curr. Opin. Microbiol.* 2, 641–646. doi: 10.1016/S1369-5274(99)00035-1
- Oldroyd, G. E., Murray, J. D., Poole, P. S., and Downie, J. A. (2011). The rules or engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* 45, 119–144. doi: 10.1146/annurev-genet-110410-132549
- O’Malley, M. A. (2015). Endosymbiosis and its implications for evolutionary theory. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10270–10277. doi: 10.1073/pnas.1421389112
- Ono, R., and Denison, R. (2010). Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids. *Plant Physiol.* 154, 1541–1548. doi: 10.1104/pp.110.163436
- Ono, R., Denison, R. F., and Kiers, E. T. (2009). Controlling the reproductive fate of rhizobia: how univesal are legume sanctions? *New Phytol.* 183, 967–979. doi: 10.1111/j.1469-8137.2009.02941.x
- Ono, R., Schmitt, I., Sprent, J. I., and Denison, R. F. (2010). Multiple evolutionary origins of legume traits leading to extreme rhizobial differentiation. *New Phytol.* 187, 508–520. doi: 10.1111/j.1469-8137.2010.03261.x
- Op den Camp, R., Streng, A., De Mita, S., Cao, Q., Polone, E., Liu, W., et al. (2011). LysM-type mycorrhizal receptor recruited for *Rhizobium* symbiosis in nonlegume *Parasponia*. *Science* 331, 909–912. doi: 10.1126/science.1198181
- Ormeño-Orrillo, E., Servín-Garcidueñas, L. E., Rogel, M. A., González, V., Peralta, H., Mora, J., et al. (2015). Taxonomy of rhizobia and agrobacteria from the Rhizobiaceae family in light of genomics. *Syst. Appl. Microbiol.* 38, 287–291. doi: 10.1016/j.syapm.2014.12.002
- Pan, H., and Wang, D. (2017). Nodule cysteine-rich peptides maintain a working balance during nitrogen-fixing symbiosis. *Nat. Plants* 3:17048. doi: 10.1038/nplants.2017.48
- Pankhurst, C. E., Hopcroft, D. H., and Jones, W. T. (1987). Comparative morphology and flavolam content of *Rhizobium loti* induced effective and ineffective root nodules on *Lotus* species, *Leuceana leucocephala*, *Carmichaelia flagelliformis*, *Ornithopus sativus*, and *Cianthus puniceus*. *Can. J. Microbiol.* 65, 2676–2685. doi: 10.1139/b87-358
- Panter, S., Thomson, R., de Bruxelles, G., Laver, D., Trevaskis, B., and Udvardi, M. (2000). Identification with proteomics of novel proteins associated with the peribacteroid membrane of soybean root nodules. *Mol. Plant Microbe Interact.* 13, 325–333. doi: 10.1094/MPMI.2000.13.3.325
- Parniske, M. (2000). Intracellular accommodation of microbes by plants: a common developmental program for symbiosis and disease? *Curr. Opin. Plant Biol.* 3, 320–328. doi: 10.1016/S1369-5266(00)00088-1
- Patriarca, E. J., Taté, R., and Iaccarino, M. (2002). Key role of bacterial NH₄⁺ metabolism in *Rhizobium*-plant symbiosis. *Microbiol. Mol. Biol. Rev.* 66, 203–222. doi: 10.1128/MMBR.66.2.203-222.2002
- Patriarca, E. J., Taté, R., Ferraioli, S., and Iaccarino, M. (2004). Organogenesis of the legume root nodules. *Int. Rev. Cytol.* 234, 201–262. doi: 10.1016/S0074-7696(04)34005-2
- Pawlowski, K., and Sprent, J. I. (2008). “Comparison between actinorhizal and legume symbiosis,” in *Nitrogen-fixing Actinorhizal Symbioses*, eds K. Pawlowski and W. E. Newton (New York, NY: Springer-Verlag), 261–288.
- Peix, A., Ramírez-Bahena, M. H., Velázquez, E., and Bedmar, E. J. (2015). Bacterial associations with legumes. *Crit. Rev. Plant Sci.*, 34, 17–42, doi: 10.1080/07352689.2014.897899
- Pérez-Montaño, F., Alías-Villegas, C., Bellogín, R. A., del Cerro, P., Espuny, M. R., Jiménez-Guerrero, I., et al. (2014). Plant growth promotion in cereal and leguminous agricultural important plants: from microorganism capacities to crop production. *Microbiol. Res.* 169, 326–336. doi: 10.1016/j.micres.2013.09.011
- Perotto, S., Vandebosch, K. A., Butcher, G. W., and Brewin, N. J. (1991). Molecular composition and development of the plant glycocalyx associated with the peribacteroid membrane of pea root nodules. *Development* 11, 763–773.
- Persson, T., Benson, D. R., Normand, P., Heuvel, B. V., Pujic, P., Chertkov, O., et al. (2011). Genome sequence of “*Candidatus Frankia datiscae*” Dg1, the uncultured microsympiont from nitrogen-fixing root nodules of the Dicot *Datisca glomerata*. *J. Bacteriol.* 193, 7017–7018. doi: 10.1128/JB.06208-11
- Precht, J., Kneip, C., Lockhart, P., Wenderoth, K., and Maier, U. G. (2004). Intracellular spheroid bodies of *Rhopalodia gibba* have nitrogen-fixation

- apparatus of cyanobacterial origin. *Mol. Biol. Evol.* 21, 1477–1481. doi: 10.1093/molbev/msh086
- Prell, J., Bourdès, A., Kumar, S., Lodwig, E., Hosie, A., Kinghorn, S., et al. (2010). Role of symbiotic auxotrophy in the *Rhizobium*-legume symbioses. *PLoS ONE* 5:e13933. doi: 10.1371/journal.pone.0013933
- Prell, J., White, J. P., Bourdès, A., Bunnewell, S., Bongaerts, R. J., and Poole, P. S. (2009). Legumes regulate *Rhizobium* bacteroid development and persistence by the supply of branched-chain amino acids. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12477–12482. doi: 10.1073/pnas.0903653106
- Provorov, N. A. (2015). Adaptive macroevolution of legume-rhizobia symbiosis. *Agric. Biol.* 50, 323–331. doi: 10.15389/agrobiology.2015.3.323eng
- Provorov, N. A., and Andronov, E. E. (2016). Evolution of root-nodule bacteria: reconstruction of the speciation processes resulting from genomic rearrangements in a symbiotic system. *Microbiology* 85, 131–139. doi: 10.1134/S0026261716020156
- Provorov, N. A., Vorobyov, N. I., and Andronov, E. E. (2008). Macro- and microevolution of bacteria in symbiotic systems. *Russ. J. Genet.* 44, 6–20. doi: 10.1134/S102279540801002X
- Pueppke, S. G. (1996). The genetic and biochemical basis of nodulation of legumes by rhizobia. *Crit. Rev. Biotechnol.* 16, 1–51. doi: 10.3109/07388559609146599
- Puppo, A., Groten, K., Bastian, F., Carzaniga, R., Soussi, M., Lucas, M. M., et al. (2005). Legume nodule senescence: roles for redox and hormone signalling in the orchestration of the natural aging process. *New Phytol.* 165, 683–701. doi: 10.1111/j.1469-8137.2004.01285.x
- Puppo, A., Herrada, G., and Rigaud, J. (1991). Lipid peroxidation in peribacteroid membranes from French-bean nodules. *Plant Physiol.* 96, 826–830. doi: 10.1104/pp.96.3.826
- Qiao, Z., Brechenmacher, L., Smith, B., Strout, G. W., Mangin, W., Taylor, C., et al. (2017). The *GmFWL1* (FW2-2-like) nodulation gene encodes a plasma membrane microdomain-associated protein. *Plant Cell Environ.* 40, 1442–1455. doi: 10.1111/pce.1294
- Rae, A. L., Bonfante-Fasolo, P., and Brewin, N. J. (1992). Structure and growth of infection threads in the legume symbiosis with *Rhizobium leguminosarum*. *Plant J.* 2, 385–395. doi: 10.1111/j.1365-313X.1992.00385.x
- Ran, L. A., Larsson, J., Vigil-Stenman, T., Nylander, J. A. A., Ininbergs, K., Zheng, W. W., et al. (2010). Genome erosion in a nitrate-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. *PLoS ONE* 5:e11486. doi: 10.1371/journal.pone.0011486
- Rasmussen, U., and Nilsson, M. (2002). “Cyanobacterial diversity and specificity in plant symbioses,” in *Cyanobacteria in Symbiosis*, eds A. N. Rai, B. Bergman, and U. Rasmussen (Dordrecht: Academic Publishers), 313–328.
- Rasmussen, U., Johansson, C., and Bergman, B. (1994). Early communication in the *Gunnera-Nostoc* symbiosis: plant induced cell differentiation and protein synthesis in the cyanobacterium. *Mol. Plant Microbe Interact.* 7, 696–702. doi: 10.1094/MPMI-7-0696
- Raven, J. A. (1993). Energy and nutrient adquisition by autotroph symbioses and their asymbiotic ancestors. *Symbiosis* 14, 33–60.
- Raymond, J., Siefert, J. L., Staples, C. R., and Blankenship, R. E. (2004). The natural history of nitrogen fixation. *Mol. Biol. Evol.* 21, 541–554. doi: 10.1093/molbev/msh047
- Redondo, F. J., Coba de la Peña, T., Morcillo, C. N., Lucas, M. M., and Pueyo, J. J. (2009). Overexpression of flavodoxin induces changes in antioxidant metabolism leading to delayed senescence and starch accumulation in alfalfa nodules. *Plant Physiol.* 149, 1166–1178. doi: 10.1104/pp.108.129601
- Reinhold-Hurek, B., and Hurek, T. (1998). Life in grasses: diazotrophic endophytes. *Trends Microbiol.* 6, 139–144. doi: 10.1016/S0966-842X(98)01229-3
- Reinhold-Hurek, B., and Hurek, T. (2011). Living inside plants: bacterial endophytes. *Curr. Opin. Plant Biol.* 1, 435–443. doi: 10.1016/j.pbi.2011.04.004
- Ribeiro, C. W., Baldacci-Cresp, F., Pierre, O., Larousse, M., Benyamina, S., Lambert, A., et al. (2017). Regulation of differentiation of nitrogen-fixing bacteria by microsymbiont targeting of plant thioredoxin s1. *Curr. Biol.* 27, 250–256. doi: 10.1016/j.cub.2016.11.013
- Rogel, M. A., Ormeño-Orillo, E., and Martinez Romero, E. (2011). Symbioses in rhizobia reflect bacterial adaptation to legumes. *Syst. Appl. Microbiol.* 34, 96–104. doi: 10.1016/j.syapm.2010.11.015
- Roth, L. E., and Stacey, G. (1989). Bacterium release into host-cells of nitrogen-fixing soybean nodules – the symbiosome membrane comes from 3 sources. *Eur. J. Cell Biol.* 49, 13–23.
- Saalbach, G., Erik, P., and Wienkoop, S. (2002). Characterisation by proteomics of peribacteroid space and peribacteroid membrane preparations from pea (*Pisum sativum*) symbiosomes. *Proteomics* 2, 325–337. doi: 10.1002/1615-9861(200203)2:3<325::AID-PROT325>3.0.CO;2-W
- Sachs, J. L., Essenberg, C. J., and Turcotte, M. M. (2011a). New paradigms for the evolution of beneficial infections. *Trends Ecol. Evol.* 26, 202–209. doi: 10.1016/j.tree.2011.01.010
- Sachs, J. L., Skophamer, R. G., and Regus, J. U. (2011b). Evolutionary transitions in bacterial symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10800–10807. doi: 10.1073/pnas.1100304108
- Saeki, K. (2011). Rhizobial measures to evade host defense strategies and endogenous threats to persistent symbiotic nitrogen fixation: a focus on two legume-rhizobium model systems. *Cell. Mol. Life Sci.* 68, 1327–1339. doi: 10.1007/s0018-011-0650-5
- Sajnaga, E., Malek, W., Lotocka, B., Stepkowski, T., and Legocki, A. (2001). The root-nodule symbiosis between *Sarrothamnus scoparius* L. and its microsymbionts. *Anton Leeuw. Int. J. G. Mol. Microbiol.* 79, 385–391. doi: 10.1023/A:1012010328061
- Sandal, N. N., Bojsen, K., Richter, H., Sengupta-Gopalan, C., and Marcker, K. A. (1992). The nodulin 24 protein family shows similarity to a family of glycine-rich plant proteins. *Plant Mol. Biol.* 18, 607–610. doi: 10.1007/BF00040679
- Santi, C., Bogusz, D., and Frache, C. (2013). Biological nitrogen fixation in non-legume plants. *Ann. Bot.* 111, 743–767. doi: 10.1093/aob/mct048
- Sessitsch, J., Howieson, X., Perret, H., Antoun, H., and Martinez-Romero, E. (2002). Advances in *Rhizobium* research. *Crit. Rev. Plant Sci.* 21, 323–378. doi: 10.1080/0735-260291044278
- Shamseldin, A., Abdelkhalek, A., and Sadowsky, M. J. (2017). Recent changes to the classification of symbiotic, nitrogen-fixing, legume-associating bacteria: a review. *Symbiosis* 71, 91–109. doi: 10.1007/s13199-016-0462-3
- Sheahan, M. B., Rose, R. J., and McCurdy, D. W. (2004). Organelle inheritance in plant cell division: the actin cytoskeleton is required for unbiased inheritance of chloroplasts, mitochondria and endoplasmic reticulum in dividing protoplasts. *Plant J.* 37, 379–390. doi: 10.1046/j.1365-313X.2003.01967.x
- Sheokand, S., Dahiya, P., Vincent, J. L., and Brewin, N. J. (2005). Modified expression of cysteine protease affects seed germination, vegetative growth and nodule development in transgenic lines of *Medicago truncatula*. *Plant Sci.* 169, 966–975. doi: 10.1016/j.plantsci.2005.07.003
- Siqueira, A. F., Ormeño-Orrillo, E., Souza, R. C., Rodrigues, E. P., Almeida, L. G. P., Barcellos, F. G., et al. (2014). Comparative genomics of *Bradyrhizobium japonicum* CPAC 15 and *Bradyrhizobium diazoefficiens* CPAC 7: elite model strains for understanding symbiotic performance with soybean. *BMC Genomics* 15:420. doi: 10.1186/1471-2164-15-420
- Smith, S. E., and Read, D. J. (2008). *Mycorrhizal Symbiosis*, 3rd Edn. London: Academic Press.
- Soltis, D. E., Soltis, P. S., Chase, M. E., Morr, M. E., Albach, D. C., Zanis, M., et al. (2000). Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *arpB* sequences. *Bot. J. Linn. Soc.* 133, 381–461. doi: 10.1006/bjol.2000.0380
- Soltis, D. E., Soltis, P. S., Morgan, D. R., Swensen, S. M., Mullin, B. C., Dowd, J. M., et al. (1995). Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc. Natl. Acad. Sci. U.S.A.* 92, 2647–2651. doi: 10.1073/pnas.92.7.2647
- Son, O., Yang, H. S., Lee, H. J., Lee, M. Y., Shin, K. H., Jeon, S. L., et al. (2003). Expression of *srbA7* and *ScAM* genes required for endocytosis of *Rhizobium* in root nodules. *Plant Sci.* 165, 1239–1244. doi: 10.1016/S0168-9452(03)00331-5
- Sprent, J. I. (2007). Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol.* 174, 11–25. doi: 10.1111/j.1469-8137.2007.02015.x
- Sprent, J. I. (2008). 60 Ma of legume nodulation: what's new? What's changing? *J. Exp. Bot.* 59, 1081–1084. doi: 10.1093/jxb/erm286
- Sprent, J. I. (2009). *Legume Nodulation: Global Perspective*. Oxford: Wiley-Blackwell. doi: 10.1002/9781444316384.fmatter
- Sprent, J. I., and Gehlot, H. S. (2010). Nodulated legumes in arid and semi-arid environments: are they important? *Plant Ecol. Divers.* 3, 211–219. doi: 10.1080/17550874.2010.538740

- Sprent, J. I., and James, E. K. (2007). Legume evolution: where do nodules and mycorrhizas fit in? *Plant Physiol.* 144, 575–581. doi: 10.1104/pp.107.096156
- Sprent, J. I., and Thomas, R. J. (1984). Nitrogen nutrition of seedling grain legumes: some taxonomic, morphological and physiological constraints. *Plant Cell Environ.* 7, 637–645. doi: 10.1111/1365-3040.ep11571523
- Sprent, J. I., Ardley, J. K., and James, E. K. (2013). From North to South: a latitudinal look at legume nodulation processes. *S. Afr. J. Bot.* 89, 31–41. doi: 10.1016/j.sajb.2013.06.011
- Sprent, J. I., Ardley, J., and James, E. K. (2017). Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol.* 215, 40–56. doi: 10.1111/nph.14474
- Stacey, G. (1995). *Bradyrhizobium japonicum* nodulation genetics. *FEMS Microbiol. Lett.* 127, 1–9. doi: 10.1111/j.1574-6968.1995.tb07441.x
- Steenhoudt, O., and Vanderleyden, J. (2000). *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. *FEMS Microbiol. Rev.* 24, 487–506. doi: 10.1111/j.1574-6976.2000.tb00552.x
- Stępkowski, T., and Legocki, A. B. (2001). Reduction of bacterial genome size and expansion resulting from obligate intracellular lifestyle and adaptation to soil habitat. *Acta Biochim. Pol.* 48, 367–381.
- Strittmatter, P., Soll, J., and Böltner, B. (2010). The chloroplast protein import machinery: a review. *Methods Mol. Biol.* 619, 307–321. doi: 10.1007/978-1-60327-412-8_18
- Stüeken, E. E., Buick, R., Guy, B. M., and Koehler, M. C. (2015). Isotopic evidence for biological nitrogen fixation by molybdenum-nitrogenase from 3.2 Gyr. *Nature* 520, 666–669. doi: 10.1038/nature14180
- Szczyglowski, K., and Amyot, L. (2003). Symbiosis, inventiveness by recruitment? *Plant Physiol.* 131, 935–940. doi: 10.1104/pp.017186
- Tang, C., Robson, A. D., Kuo, J., and Dilworth, M. J. (1993). Anatomical and ultrastructural observations on infection of *Lupinus agustifolius* L. by *Bradyrhizobium* sp. *J. Comput. Assist. Microsc.* 5, 47–51.
- Tejada-Jiménez, M., Gil-Díez, P., León-Medivilla, J., Wen, J., Mysore, K. S., Imperial, J., et al. (2017). *Medicago truncatula* Molybdate Transporter type 1 (MtMOT1.3) is a plasma membrane molybdenum transporter required for nitrogenase activity in root nodules under molybdenum deficiency. *New Phytol.* 216, 1223–1235. doi: 10.1111/nph.14739
- Terpolilli, J. J., Hood, G. A., and Poole, P. S. (2012). What determines the efficiency of N₂-fixing *Rhizobium*-legume symbioses? *Adv. Microb. Physiol.* 60, 325–389. doi: 10.1016/B978-0-12-398264-3.00005-X
- Thajuddin, N., Muralitharan, G., Sundaramoorthy, M., Ramamoorthy, R., Ramachandran, S., Akbarsha, M. A., et al. (2010). Morphological and genetic diversity of symbiotic cyanobacteria from cycads. *J. Basic Microbiol.* 50, 254–265. doi: 10.1002/jobm.200900343
- Theissen, U., and Martin, W. (2006). The difference between organelles and endosymbionts. *Curr. Biol.* 16, R1016–R1017. doi: 10.1016/j.cub.2006.11.020
- Thompson, A. W., Foster, R. A., Krupe, A., Carter, B. J., Musat, N., Vaulot, D., et al. (2012). Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science* 337, 1546–1550. doi: 10.1126/science.1222700
- Tikhonovich, I. A., and Provorov, N. A. (2009). From plant-microbe interactions to symbiogenetics: a universal paradigm for the interspecies genetic integration. *Ann. Appl. Biol.* 154, 341–350. doi: 10.1111/j.1744-7348.2008.00306.x
- Timmers, A. C. J. (2008). The role of the plant cytoskeleton in the interaction between legumes and rhizobia. *J. Microsc.* 231, 247–256. doi: 10.1111/j.1365-2818.2008.02040.x
- Tong, J., Dolezal, P., Selkirk, J., Crawford, S., Simpson, A. G. B., Noinaj, N., et al. (2011). Ancestral and derived protein import pathways in the mitochondrion of *Reclinomonas americana*. *Mol. Biol. Evol.* 28, 1581–1591. doi: 10.1093/molbev/msq305
- Trinick, M. (1973). Symbiosis between *Rhizobium* and the non-legume, *Trema aspera*. *Nature* 244, 459–460. doi: 10.1038/244459a0
- Turner, S. L., and Young, J. P. (2000). The glutamine synthetases of rhizobia: phylogenetics and evolutionary implications. *Mol. Biol. Evol.* 17, 309–319. doi: 10.1093/oxfordjournals.molbev.a026311
- Tyerman, S. D., Whitehead, L. F., and Day, D. A. (1995). A channel-like transporter for NH₄⁺ on the symbiotic interface of N₂-fixing plants. *Nature* 378, 629–632. doi: 10.1038/378629a0
- Tyrrell, T. (1999). The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature* 400, 525–531. doi: 10.1038/22941
- Udvardi, M. K., and Day, D. A. (1997). Metabolite transport across symbiotic membranes of legume nodules. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48, 493–523. doi: 10.1146/annurev.applant.48.1.493
- Udvardi, M., Bock, V., Colebatch, G., Desbrosses, G., Kloska, S., Kopla, J., et al. (2004). "Genetic reorganization of legume transport and metabolism during symbiotic nitrogen fixation," in *Biology of Plant-Microbe Interactions*, Vol. 4, eds L. A. Tikhonovich, N. A. Provorov, V. I. Romanov, and W. E. Newton (St. Petersburg: Biont), 490–492.
- Udvardi, M. K., Price, G. D., Gresshoff, P. M., and Day, D. A. (1988). A dicarboxylate transporter on the peribacteroid membrane of soybean nodules. *FEBS Lett.* 231, 36–40. doi: 10.1016/0014-5793(88)80697-5
- Van de Velde, W., Pérez Guerra, J. C., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., et al. (2006). Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. *Plant Physiol.* 141, 711–720. doi: 10.1104/pp.106.078691
- Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327, 1122–1126. doi: 10.1126/science.1184057
- Vega-Hernández, M. C., Pérez-Galdona, R., Dazzo, F. B., Jarabe-Lorenzo, A., Alfayate, M. C., and León-Barrios, M. (2001). Novel infection process in the indeterminate root nodule symbiosis between *Chamaecytisus proliferus* (tagasaste) and *Bradyrhizobium* sp. *New Phytol.* 150, 707–721. doi: 10.1046/j.1469-8137.2001.00120.x
- Verhaert, J., Vanderleyden, J., and Michiels, J. (2005). Bacterial endocytic systems in plants and animals: Ca²⁺ as a common theme? *Crit. Rev. Plant Sci.* 24, 283–308. doi: 10.1080/0735268050022105
- Vessey, J. K., Pawłowski, K., and Bergman, B. (2004). Root-based N₂-fixing symbioses: legumes, actinorhizal plants, *Parasponia* sp. and cycads. *Plant Soil* 266, 205–230. doi: 10.1007/s11104-005-0871-1
- Vincent, J. L., and Brewin, N. J. (2000). Immunolocalization of a cysteine protease in vacuoles, vesicles and symbiosomes of pea nodule cells. *Plant Physiol.* 123, 521–530. doi: 10.1104/pp.123.2.521
- Vincent, J. L., Knox, M. R., Ellis, T. H. N., Kalo, P., Kiss, G. B., and Brewin, N. J. (2000). Nodule expressed *Cyp15a* cysteine protease genes map to syntenic genomic regions in *Pisum* and *Medicago* spp. *Mol. Plant Microbe Interact.* 13, 715–723. doi: 10.1094/MPMI.2000.13.7.715
- Vincill, E. D., Szczyglowski, K., and Roberts, D. M. (2005). GmN70 and LjN70. Anion transporters of the symbiosome membrane of nodules with a transport preference for nitrate. *Plant Physiol.* 137, 1435–1444. doi: 10.1104/pp.104.051953
- Wang, D., Griffitts, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., et al. (2010). A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* 327, 1126–1129. doi: 10.1126/science.1184096
- Weir, B. S. (2016). *The Current Taxonomy of rhizobia*. New Zealand rhizobia website. Available online at: <http://www.rhizobia.co.nz/taxonomy/rhizobia.htm>
- White, J., Prell, J., James, E. K., and Poole, P. (2007). Nutrient sharing between symbionts. *Plant Physiol.* 144, 604–614. doi: 10.1104/pp.107.097741
- Whitehead, L. F., and Day, D. A. (1997). The peribacteroid membrane. *Physiol. Plant.* 100, 30–44. doi: 10.1111/j.1399-3054.1997.tb03452.x
- Whitehead, L. F., Day, D. A., and Hardham, A. R. (1998). Cytoskeleton arrays in the cells of soybean root nodules: the role of actin microfilaments in the organisation of symbiosomes. *Protoplasma* 203, 194–205. doi: 10.1007/BF01279476
- Wienkoop, S., and Saalbach, G. (2003). Proteome analysis. Novel proteins identified at the peribacteroid membrane from *Lotus japonicus* root nodules. *Plant Physiol.* 131, 1080–1090. doi: 10.1104/pp.102.015362
- Winzer, T., Bairl, A., Linder, M., Werner, D., and Müller, P. (1999). A novel 53-kDa nodulin of the symbiosome membrane of soybean nodules, controlled by *Bradyrhizobium japonicum*. *Mol. Plant Microbe Interact.* 12, 218–226. doi: 10.1094/MPMI.1999.12.3.218
- Yang, S., Wang, Q., Fedorova, E., Liu, J., Qin, Q., Zheng, Q., et al. (2017). Microsymbiont discrimination mediated by a host-secreted peptide in *Medicago truncatula*. *Proc. Natl. Acad. Sci. U.S.A.* 114, 6848–6853. doi: 10.1073/pnas.1700460114
- Yates, R. J., Howieson, J. G., Reeve, W. G., Nandasena, K. G., Law, I. J., Bräu, L., et al. (2007). *Lotononis angolensis* forms nitrogen-fixing, lupinoid nodules

- with phylogenetically unique, fast-growing, pink-pigmented bacteria, which do not nodulate *L. bainesii* or *L. listii*. *Soil Biol. Biochem.* 39, 1680–1688. doi: 10.1016/j.soilbio.2007.01.025
- Yokota, K., and Hayashi, M. (2011). Function and evolution of nodulation genes in legumes. *Cell. Mol. Life Sci.* 68, 1341–1351. doi: 10.1007/s00018-011-0651-4
- Yurgel, S. N., and Kahn, M. L. (2008). A mutant *GlnD* nitrogen sensor protein leads to a nitrogen-fixing but ineffective *Sinorhizobium meliloti* symbiosis with alfalfa. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18958–18963. doi: 10.1073/pnas.0808048105
- Zimorski, V., Ku, C., Martin, W. F., and Gould, S. B. (2014). Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* 22, 38–48. doi: 10.1016/j.mib.2014.09.008

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer OS and handling Editor declared their shared affiliation.

Copyright © 2018 Coba de la Peña, Fedorova, Pueyo and Lucas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Purification and *In Vitro* Activity of Mitochondria Targeted Nitrogenase Cofactor Maturase NifB

Stefan Burén, Xi Jiang, Gema López-Torrejón, Carlos Echavarri-Erasun and Luis M. Rubio *

Centro de Biotecnología y Genómica de Plantas (CBGP), Universidad Politécnica de Madrid (UPM), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain

OPEN ACCESS

Edited by:

Nikolai Provorov,
All-Russian Research Institute
of Agricultural Microbiology of the
Russian Academy of Agricultural
Sciences, Russia

Reviewed by:

Oswaldo Valdes-Lopez,
National Autonomous University
of Mexico, Mexico
Zonghua Wang,
Fujian Agriculture and Forestry
University, China

***Correspondence:**

Luis M. Rubio
lm.rubio@upm.es

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 31 May 2017

Accepted: 28 August 2017

Published: 12 September 2017

Citation:

Burén S, Jiang X, López-Torrejón G, Echavarri-Erasun C and Rubio LM (2017) Purification and *In Vitro* Activity of Mitochondria Targeted Nitrogenase Cofactor Maturase NifB. *Front. Plant Sci.* 8:1567.
doi: 10.3389/fpls.2017.01567

Active NifB is a milestone in the process of engineering nitrogen fixing plants. NifB is an extremely O₂-sensitive S-adenosyl methionine (SAM)-radical enzyme that provides the key metal cluster intermediate (NifB-co) for the biosyntheses of the active-site cofactors of all three types of nitrogenases. NifB and NifB-co are unique to diazotrophic organisms. In this work, we have expressed synthetic codon-optimized versions of NifB from the γ -proteobacterium *Azotobacter vinelandii* and the thermophilic methanogen *Methanocaldococcus infernus* in *Saccharomyces cerevisiae* and in *Nicotiana benthamiana*. NifB proteins were targeted to the mitochondria, where O₂ consumption is high and bacterial-like [Fe-S] cluster assembly operates. In yeast, NifB proteins were co-expressed with NifU, NifS, and FdxN proteins that are involved in NifB [Fe-S] cluster assembly and activity. The synthetic version of thermophilic NifB accumulated in soluble form within the yeast cell, while the *A. vinelandii* version appeared to form aggregates. Similarly, NifB from *M. infernus* was expressed at higher levels in leaves of *Nicotiana benthamiana* and accumulated as a soluble protein while *A. vinelandii* NifB was mainly associated with the non-soluble cell fraction. Soluble *M. infernus* NifB was purified from aerobically grown yeast and biochemically characterized. The purified protein was functional in the *in vitro* FeMo-co synthesis assay. This work presents the first active NifB protein purified from a eukaryotic cell, and highlights the importance of screening *nif* genes from different organisms in order to sort the best candidates to assemble a functional plant nitrogenase.

Keywords: nitrogen fixing plants, yeast, mitochondria, SAM-radical, iron-molybdenum cofactor

INTRODUCTION

Agricultural systems in developed countries are largely based on cereal crops, which provide most of the calories and proteins in the human diet (Borlaug, 2002). Nitrogen and water availability are the most important factors limiting cereal crop productivity. Over the last 100 years, cereal crop yields have been increased by addition of chemically synthesized nitrogen fertilizers (Smil, 2000; Galloway et al., 2008). However, the extensive use of these commercial nitrogen fertilizers in developed countries poses enormous and pressing environmental threats (Vitousek et al., 1997). On the other hand, the cost of chemical fertilizers is prohibitive for poor farmers, and they are scarcely used in most of Africa with consequence of poverty and hunger derived from

extremely low crop yields (Sanchez and Swaminathan, 2005). During the last years, the scientific community has paid considerable attention to this problem and acknowledged the need of disruptive technological changes. One way to tackle this problem could be the generation of so-called nitrogen-fixing plants, aimed to solve the nitrogen problem (Stokstad, 2016; Vicente and Dean, 2017).

Nitrogen fixation is the conversion of inert atmospheric N₂ into NH₃, a biologically active form of nitrogen. Only a few bacteria are naturally able to fix nitrogen through the activity of an enzyme called nitrogenase; collectively known as diazotrophs (nitrogen eaters). In this regard, transferring bacterial nitrogen fixation (*nif*) genes into the plant genome could result in plants able to fix nitrogen, and therefore in crops less dependent of external nitrogen fertilization. This direct approach eliminates the need to generate or optimize interactions of cereals with specific symbiotic or associative nitrogen fixing bacteria (Oldroyd and Dixon, 2014). However, two main barriers are believed to impair the direct *nif* gene transfer approach (Curatti and Rubio, 2014): the known sensitivity of nitrogenase to O₂ and the genetic complexity of nitrogenase biosynthesis. Two decades of genetic and biochemical analyses culminated with the unambiguous identification of the essential proteins required for nitrogenase cofactor biosynthesis (Curatti et al., 2007). On the contrary, overcoming the O₂ sensitivity barrier in plants remains largely unexplored.

Nitrogenases have two O₂-sensitive protein components: a dinitrogenase that catalyzes the nitrogen fixation reaction and a dinitrogenase reductase that serves as obligate electron donor to dinitrogenase (Bulen and LeComte, 1966). In the case of the widespread molybdenum nitrogenase these components are called Fe protein and MoFe protein. The Fe protein is a homodimer of the *nifH* gene product that contains a [4Fe-4S] cluster at the subunit interface. The MoFe protein is a heterotetramer of the *nifD* and *nifK* gene products that contains two complex iron-sulfur (Fe-S) clusters called iron-molybdenum cofactor (FeMo-co) and P-cluster. The type of [4Fe-4S] cluster found in NifH is ubiquitous in nature. In fact, plants carry [Fe-S] cluster assembly machineries in mitochondria, chloroplasts, and cytosol, which are all capable of synthesizing [4Fe-4S] clusters (Balk and Pilon, 2011). However, the P-cluster and FeMo-co are unique to diazotrophs. Their uniqueness implies that specialized cellular biosynthetic pathways, involving multiple *nif* gene products, are required for cofactor synthesis and NifDK maturation (Rubio and Ludden, 2008).

Successful expression and maturation of the prokaryotic nitrogenase protein in a eukaryotic host, in order to develop N₂ fixing cereal crops, could revolutionize agricultural systems worldwide. For this to succeed, a deeper understanding of the processes involved in the formation of active nitrogenase in a eukaryotic cell is required. In this regard, expression of *nif* genes in *Saccharomyces cerevisiae* have shown that: (1) active NifH can be achieved upon mitochondrial targeting (Lopez-Torrejon et al., 2016), providing a proof of concept that O₂-sensitive Nif proteins can be assembled in an eukaryotic cell organelle, and (2) that expression and mitochondria targeting of nine Nif proteins

(NifUSHMDKBEN) resulted in proper mitochondria targeting, processing and NifDK tetramer formation (Burén et al., 2017), an essential step of nitrogenase assembly. However, to obtain similar results in a plant cell background is likely to be more challenging, as the O₂ generated during photosynthesis could create even a harsher environment for nitrogenase proteins, especially in the chloroplast. This was recently suggested from work by Ivleva et al. (2016), where Fe protein activity from transplastomic *Nicotiana tabacum* plants only could be detected at very low levels in plants previously incubated at sub-ambient O₂ (Ivleva et al., 2016). Importantly, a recent study showed that 16 mitochondria targeted Nif proteins from *Klebsiella pneumoniae* (among them NifB) could be successfully expressed in leaves of *Nicotiana benthamiana* (Allen et al., 2017). Although no protein activities were reported, this work showed that most Nif proteins were well expressed and accumulated at their estimated sizes within the plant tissue, with the exception of NifD that appeared to be processed to a polypeptide of smaller size, as has also been observed in yeast (Burén et al., 2017).

A main hurdle to overcome in order to generate functional nitrogenase proteins is obtaining active NifB. NifB is an extremely O₂-sensitive S-adenosyl methionine (SAM)-radical enzyme (Curatti et al., 2006), that provides the key intermediate metal cluster (called NifB-co) in the biosynthesis of FeMo-co (Shah et al., 1994; Guo et al., 2016). As NifB-co also serves as precursor for FeV-co in the vanadium nitrogenase and for FeFe-co in the iron-only nitrogenase, NifB is required for all biological nitrogen fixation activity in nature (Bishop and Joerger, 1990; Curatti et al., 2007; Dos Santos et al., 2012). Because of its O₂ sensitivity and instability, and since it is unlikely that NifB or NifB-co can be replaced by components of plant origin, plant cell NifB-co accumulation is likely to be one of the main barriers in the generation of an active plant nitrogenase (Curatti and Rubio, 2014; Vicente and Dean, 2017).

In this work, two naturally occurring NifB proteins, from the model-diazotroph *Azotobacter vinelandii* and from the thermophilic methanogen *Methanocaldococcus infernus* (see accompanying paper by Arragain et al., submitted manuscript), were expressed in *S. cerevisiae* and targeted to mitochondria. Mitochondria was chosen due to the high rate of O₂ consumption, and the plentiful ATP and reducing power generated by respiration (Curatti and Rubio, 2014), in addition to the bacterial-like [Fe-S] cluster assembly machinery available (Lill and Muhlenhoff, 2008). NifB proteins were co-expressed with NifU, NifS, and FdxN proteins, involved in NifB [Fe-S] cluster formation and activity (Yuvaniyama et al., 2000; Johnson et al., 2005; Zhao et al., 2007; Jiménez-Vicente et al., 2014). Surprisingly, only NifB from the thermophile was found to accumulate in a soluble form, while NifB from *A. vinelandii* appeared to form aggregates. The soluble *M. infernus* NifB was purified and proven functional in the *in vitro* FeMo-co synthesis assay. *A. vinelandii* and *M. infernus* NifB were also targeted to the mitochondria in leaf cells of *N. benthamiana* (tobacco). As in yeast, the synthetic version of NifB from *M. infernus* was better expressed and accumulated as a soluble protein while the *A. vinelandii* NifB was mainly associated with the non-soluble cell fraction. These results underline the importance of screening

for functionality each one of the Nif proteins required to mature nitrogenase.

RESULTS

Generation of Yeast Platform Strains for NifB Expression

Synthetic versions of *A. vinelandii* *nifB* (*nifB_{Av}*) and *M. infernus* *nifB* (*nifB_{Mi}*) were codon-optimized for *S. cerevisiae* and cloned into expression vectors under the control of the galactose inducible promoters (Supplementary Figures S1, S2). As NifU, NifS, and FdxN participate in NifB maturation and activity (Zhao et al., 2007; Jiménez-Vicente et al., 2014), synthetic versions of the *A. vinelandii* *nifU*, *nifS*, and *fdxN* genes, codon-optimized for *S. cerevisiae*, were additionally cloned into expression vectors under the control of the galactose inducible promoters. To ensure mitochondria targeting of the expressed proteins, SU9 leader sequences, previously shown to deliver yeast expressed Nif proteins to the mitochondria (Burén et al., 2017) were added in-frame and upstream of each gene. To facilitate NifB purification by affinity chromatography, the coding sequence for 10 histidines were added 3' of each *nifB* gene. Expression vectors were co-introduced into *S. cerevisiae* strain W303-1a, generating yeast strains SB09Y and SB10Y for the expression of mitochondria targeted NifU, NifS, and FdxN, together with NifB_{Av}-His₁₀ or NifB_{Mi}-His₁₀ (hereafter called yNifB_{Av} and yNifB_{Mi}), respectively (Table 1).

NifB Expression, Mitochondria Targeting, and Solubility

Western blot analysis of yeast cell-free extracts using antibodies specifically recognizing NifU, NifS, and histidine-tag (for yNifB_{Av} or yNifB_{Mi}) confirmed expression of all these proteins in SB09Y and SB10Y strains grown aerobically with galactose as inducer. Protein migrations in SDS-PAGE were consistent with efficient mitochondria leader sequence processing (Table 1 and Figure 1A). Detection of FdxN in SB09Y and SB10Y was difficult, presumably due to the small size of FdxN (10 kDa) and/or to weak binding of FdxN antibodies generated for this study. To confirm that FdxN was successfully expressed from the GAL10 promoter, presence of *fdxN* transcripts were verified in SB09Y and SB10Y strains (Supplementary Figure S3). In addition, an epitope-tagged version of the protein where a C-terminal HA-tag was added to the SU9-FdxN construct was generated and expressed in strain SB12Y (Figure 1B).

Further analysis of the soluble fractions prepared from yeast cells lysed in absence of detergents indicated that most yNifB_{Mi} and nearly all yNifB_{Av} were of poor solubility (Figure 2). This suggested that the proteins were either forming insoluble aggregates upon strong expression or interacting with membranes. A similar behavior was observed for NifB_{Av} and NifB_{Ko} (*Klebsiella oxytoca* NifB) overexpressed in *Escherichia coli* (data not shown). Exchanging the C-terminal His-tag for an N-terminal variant, and addition of detergents during lysis (see Materials and Methods for details), did not improve

TABLE 1 | Yeast expression plasmids and yeast strains used in this work.

Strain	Transformed plasmids	Expressed proteins	Full-length (kDa)	Processed (kDa)
SB09Y	pN2GLT4	SU9-NifU	40.6	33.6
		SU9-NifS	51.0	44.0
SB10Y	pN2GLT4	SU9-FdxN	17.1	9.8
	pN2SB22	SU9-NifB _{Av} -His ₁₀	63.6	56.3
SB03Y	pN2GLT4	SU9-NifU	40.6	33.6
	pN2SB24	SU9-NifS	51.0	44.0
SB12Y	pN2GLT4	SU9-FdxN	17.1	9.8
	pN2SB15	SU9-His ₁₀ -NifB _{Av}	63.6	56.3
SB12Y	pN2SB39	SU9-FdxN-HA	18.2	10.9
		SU9-NifB _{Mi} -His ₁₀	43.8	36.4

Nitrogenase-related proteins and their expected sizes when expressed in yeast from expression vectors generated in this study. Processed forms refer to sizes following predicted SU9 cleavage (RAY-SS, (Vögtle et al., 2009)). See Supplementary Figures S1, S2 for details.

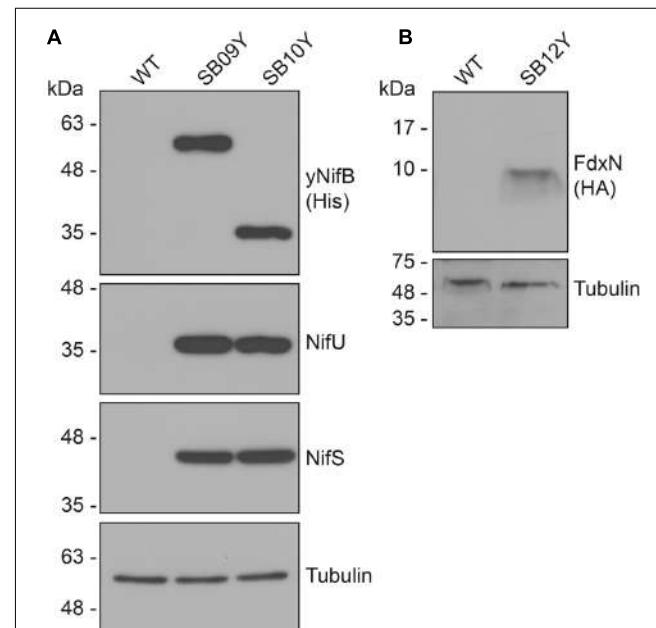
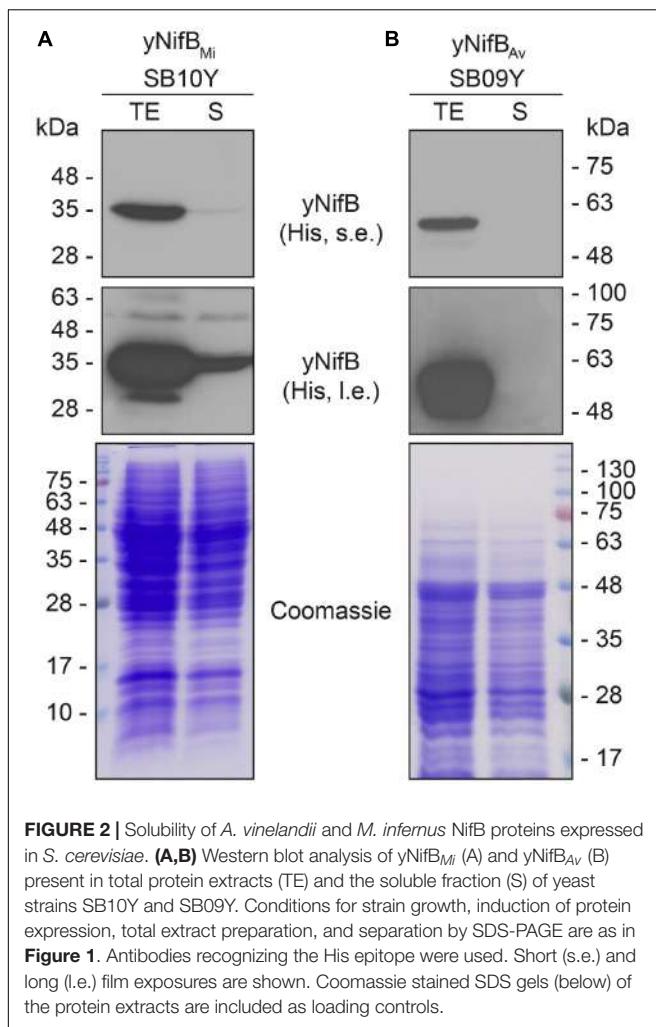


FIGURE 1 | Expression of NifB, NifU, NifS and FdxN proteins in *S. cerevisiae*. **(A)** Western blot analysis of yNifB_{Av} and yNifB_{Mi}, as well as NifU and NifS, in total protein extracts from strains SB09Y and SB10Y. **(B)** Western blot analysis of C-terminally HA-tagged FdxN in total protein extract from SB12Y strain. Extracts in (A) and (B) were prepared from aerobically grown cells following galactose induction, and proteins in the extract separated by SDS-PAGE before transferred to membranes. Antibodies recognizing NifB_{Av}, NifB_{Mi}, His epitope, HA epitope, and tubulin were used. Tubulin immunoblot signal intensity is used as loading control. See Table 1 for details about recombinant yeast strains.

solubility (Supplementary Figure S4). As NifB from the thermophile *M. infernus* previously showed heat-resistant properties (Wilcoxen et al., 2016), several distinct extraction



conditions were tested (including different temperatures) to screen yNifB_{Mi} solubility and to find a protocol for extraction and enrichment of yNifB_{Mi}. While increased concentration of glycerol did not improve solubility, the pH of the extraction buffer was important (Supplementary Figure S5). In addition, exposing the total yeast lysate to elevated temperatures before centrifugation not only reduced the amount of total yeast proteins remaining in solution, but also increased the levels of yNifB_{Mi} in the soluble fraction of the extract. Unfortunately, no similar improvement could be obtained for yNifB_{Av} (Figures 3A,B and Supplementary Figure S6) impairing yNifB_{Av} purifications. Further optimization confirmed that maximum solubility yNifB_{Mi} was obtained at pH 8 upon treatment at 60–65°C (Figure 3C). Therefore 65°C was chosen for the following yNifB_{Mi} extractions in order to minimize the complexity of the yeast cell-free extracts used for affinity chromatography.

Yeast-Expressed NifB_{Mi} Is Active in the In Vitro FeMo-co Synthesis Assay

Typical yeast NifB_{Mi} purification yielded about 4 mg/100 g cell pellet (4.4 ± 1.1 , mean and standard deviation from

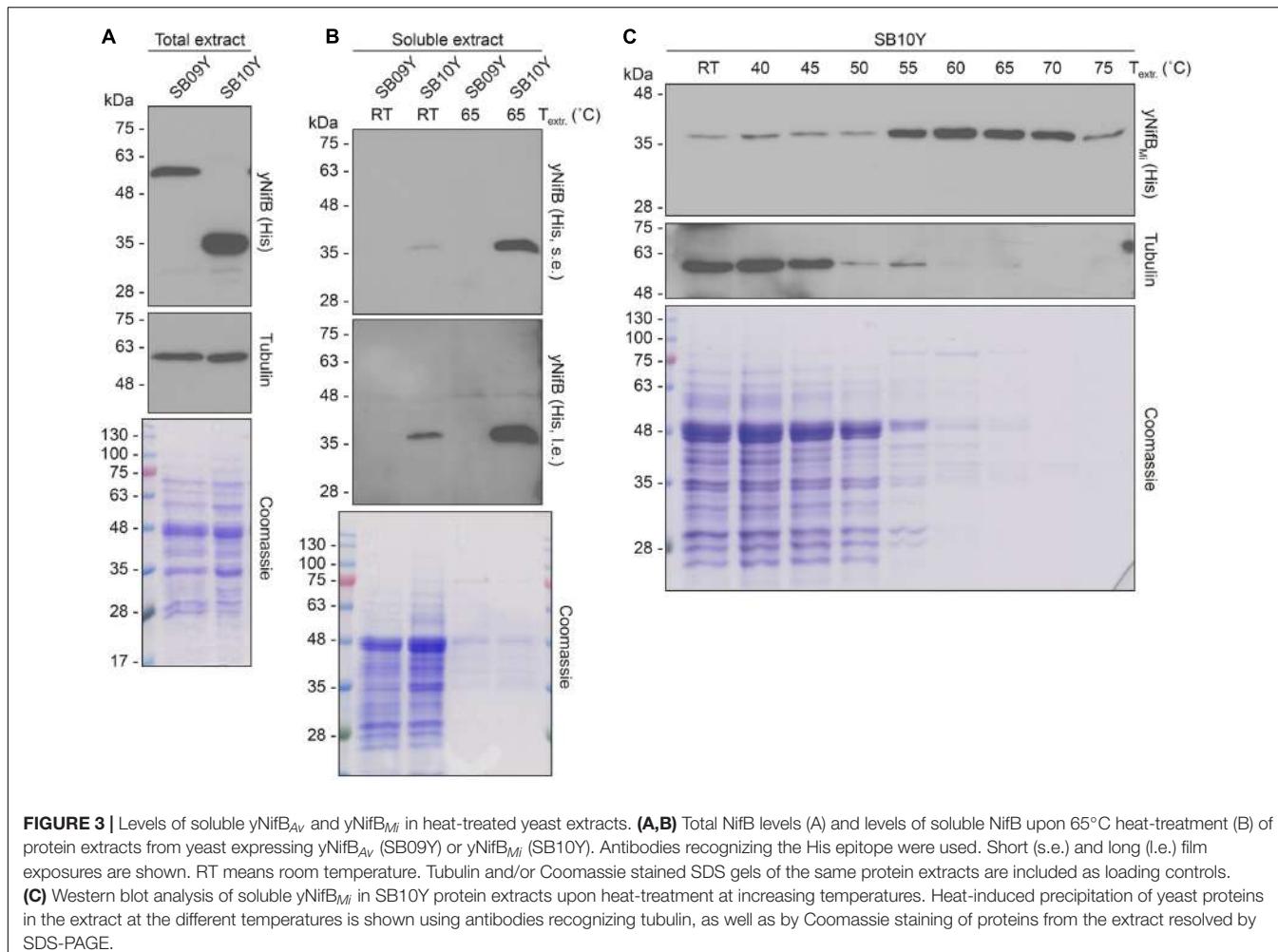
four individual purifications), and yNifB_{Mi} was at near purity as determined by SDS-PAGE analysis (Figure 4A). To confirm mitochondria import and functionality of the SU9 leader sequence, purified yNifB_{Mi} was subjected to N-terminal sequencing. Successful processing of the SU9 sequence was verified, and cleavage appeared at the site predicted from alignment of the SU9 peptide with a reported consensus sequence for yeast mitochondria proteins (Vögtle et al., 2009) (Figure 4B). While as isolated yNifB_{Mi} showed some color and UV-vis absorbance spectrum characteristic of Fe-S protein (3.3 ± 0.8 Fe atoms per monomer from four individual purifications, S not determined), *in vitro* reconstitution with Fe and S increased color intensity and the 320 and 420 nm features of the UV-vis spectrum indicative of [4Fe-4S] cluster formation (Figures 4C,D). Treatment with dithionite (DTH) reduced absorbance at 420 nm as expected for a redox responsive Fe-S protein. Fe and S content of reconstituted yNifB_{Mi} was consistent with the presence of, at minimum, two [Fe-S] clusters in addition to the SAM-binding [4Fe-4S] cluster (12.5 ± 2.8 Fe and 10.6 ± 3.1 S atoms per monomer; average \pm standard deviation from four individual purifications). All these features are typical of NifB proteins (Curatti et al., 2006; Wilcoxen et al., 2016).

NifB can be used for *in vitro* FeMo-co synthesis and nitrogenase reconstitution assays using cell-free extracts of *ΔnifB* *A. vinelandii*, supplemented with ATP-regenerating mixture, molybdenum (Mo), and homocitrate (Curatti et al., 2006). When *in vitro* FeMo-co synthesis occurs, *de novo*-synthesized FeMo-co is incorporated into apo-MoFe nitrogenase present in the extract and activity of reconstituted nitrogenase can be determined by the acetylene reduction assay. To test whether reconstituted yNifB_{Mi} was functional, 5 μM protein was added to UW140 extracts lacking NifB-co activity, but providing the rest of the protein components required for FeMo-co synthesis and activatable apo-MoFe nitrogenase. While extract without yNifB_{Mi} only showed negligible acetylene reduction, addition of yNifB_{Mi} resulted in 40-fold increase in ethylene formation (Table 2). Importantly, yNifB_{Mi} showed similar concentration-dependent activity as purified and reconstituted NifB from *A. vinelandii* (Curatti et al., 2006) (Figure 4E). The maximum activity appeared to occur at slightly higher concentration (5 μM vs. 1 μM), which could result from slight incompatibility between the yNifB_{Mi} and the other Nif components in the UW140 *A. vinelandii* extract, as has been shown for NifH (Emerich and Burris, 1976, 1978), or from the suboptimal reaction temperature for the thermophile *M. infernus* (optimal growth at 85°C) NifB protein (Jeanthon et al., 1998).

In summary, yNifB_{Mi} exhibits the spectroscopic and catalytic properties of active NifB proteins. Further studies will aim to determine whether the yNifB_{Mi} protein can support NifB-co synthesis *in vivo* in mitochondria of yeast.

Expression and Mitochondria Targeting of NifB in Plant Leaves

In a work by Allen et al. (2017), 16 HA-tagged nitrogenase proteins from *K. oxytoca* were separately expressed in



Agrobacterium tumefaciens infiltrated *N. benthamiana* leaves, and targeted to the mitochondria. *K. oxytoca* NifB was one of the Nif proteins that resulted in highest protein expression level (Allen et al., 2017). However, as leaf proteins were extracted in SDS buffer upon heating, the level of soluble NifB protein is difficult to estimate. In order to test whether differences in solubility of plant expressed and mitochondria targeted NifB proteins could be observed, as in yeast, NifB_{Av} and NifB_{Mi} were cloned into plant expression vectors under the control of the constitutive 35S promoter (Supplementary Figures S7, S8). As yeast and tobacco codon usage is similar, no further sequence optimization and gene synthesis was performed.

As SU9 is a mitochondria leader sequence from fungi without obvious plant homolog, the C-terminal His-tag of SU9-NifB_{Av} and SU9-NifB_{Mi} was replaced with GFP to track SU9 functionality in *N. benthamiana* cells. Confocal microscopy analysis showed that SU9 successfully targeted the two NifB variants to the mitochondria of *N. benthamiana*, as seen from co-localization with a red fluorescent mitochondria marker (Candat et al., 2014) (Figures 5A–D). Specific and individual detection of the fluorescent signals was verified from adjacent cells expressing only each one of the constructs (Figure 5E). Confocal microscopy

indicated that the expression level of SU9-NifB_{Av}-GFP was lower than SU9-NifB_{Mi}-GFP, which was confirmed by Western blot analysis (Figure 6A). Importantly, SU9-NifB_{Mi}-GFP was only detected in the soluble fraction of the extract, in contrast to SU9-NifB_{Av}-GFP that could also be seen in the pellet fraction (data not shown). Migration of the expressed fusion proteins was consistent with correct SU9 leader sequence processing in *N. benthamiana* cells (Figure 6A and Table 3).

Migration of the plant expressed C-terminally His-tagged versions of the SU9-NifB_{Av} and SU9-NifB_{Mi} proteins appeared identical to the corresponding proteins expressed in yeast, supporting that the SU9 leader sequence was processed correctly also in the plant mitochondria (Figures 6B,C and Supplementary Figure S9).

To enable simultaneous and comparative detection of the two *N. benthamiana* expressed NifB variants, and to exclude that solubility was affected by the C-terminal GFP moiety, new constructs were generated where the His-tag was exchanged for an N-terminal 28 amino acid Twin-Strep-tag (Schmidt et al., 2013) (Supplementary Figures S8, S10). The Twin-Strep-tag is an improved version of the eight amino acid Strep-tag II that was shown superior to His-tag for use with plant tissue

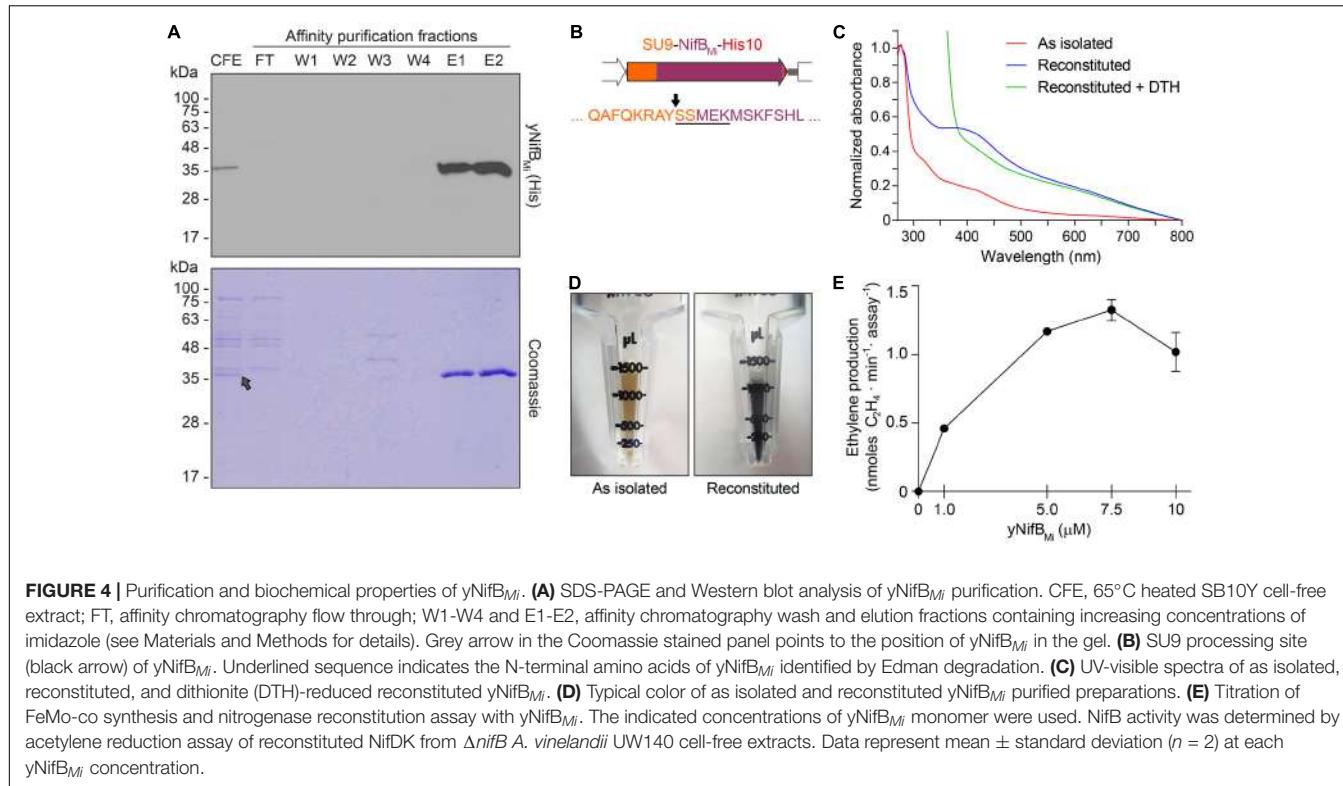


TABLE 2 | *yNifB_{Mi}*-dependent *in vitro* FeMo-co synthesis and nitrogenase reconstitution assays.

UW140	nmol C ₂ H ₄ (min ⁻¹ · assay ⁻¹)
– NifB-co	0.04
+ NifB-co	15.54 ± 0.23
+ <i>yNifB_{Mi}</i> (1)	1.33 ± 0.22
+ <i>yNifB_{Mi}</i> (2)	0.97 ± 0.02
+ <i>yNifB_{Mi}</i> (3)	1.92 ± 0.04
+ <i>yNifB_{Mi}</i> (4)	1.41 ± 0.06

Acetylene reduction assays of nitrogenase reconstituted in *ΔnifB A. vinelandii* extracts. UW140 extracts, or UW140 extracts supplemented with NifB-co, were used as negative and positive controls, respectively. Data represent mean ± standard deviation ($n = 2$) from four individual *yNifB_{Mi}* purifications (at 5 μM *yNifB_{Mi}*).

extracts (Witte et al., 2004). In addition, the SU9 signal was replaced by the first 29 amino acids of the yeast cytochrome c oxidase IV (COX4) protein, which has been shown to successfully target proteins to the mitochondria in tobacco and *Arabidopsis thaliana* (Köhler et al., 1997; Nelson et al., 2007; Pan et al., 2014). As cleavage of COX4 in yeast has been shown to occur between amino acids 25 and 26 (Vögtle et al., 2009), similar processing in *N. benthamiana* would leave only four amino acids in addition to the Twin-Strep-tag. To verify functionality of the COX4 peptide, and to confirm that the Twin-Strep-tag was not interfering with targeting or solubility, a COX4-twin-Strep-GFP construct was generated (Supplementary Figures S8, S10). As expected, COX4 efficiently targeted twin-Strep-GFP to mitochondria in *N. benthamiana* cells (Figures 7A–C). Specific

and individual detection of the fluorescent signals was verified using adjacent cells expressing only one of the constructs (Figure 7C).

Both COX4-twin-Strep-NifB_{Av} and COX4-twin-Strep-NifB_{Mi} were readily detected in total protein extracts of *A. tumefaciens* infiltrated *N. benthamiana* leaves (Figure 8A). To test the solubility of the expressed NifB proteins, total protein extracts were separated in soluble fractions and pellet associated fractions. COX4-twin-Strep-NifB_{Mi} was detected exclusively in the soluble fraction, even upon prolonged exposure (Figure 8B). On the contrary, COX4-twin-Strep-NifB_{Av} was more difficult to detect using the Strep-tag II antibody, and appeared to be in the non-soluble fraction. To verify the identity of the NifB_{Av} protein detected by the Strep-tag II antibody we used NifB_{Av} specific antibody, which confirmed that COX4-twin-Strep-NifB_{Av} was mainly present in the pellet associated fraction (Figure 8C).

In summary, we show that mitochondria targeting using SU9 and COX4 resulted in expression of both NifB_{Av} and NifB_{Mi} in leaves of *N. benthamiana*. Leader sequence processing of all proteins appeared efficient and correct, as only one band of the expected size was detected. Similar to yeast, the NifB_{Mi} protein was more soluble than the corresponding NifB_{Av} variant in *N. benthamiana*.

DISCUSSION

Expression of functional NifB is absolutely required to engineer nitrogenase in eukaryotic organisms (e.g., plants). NifB catalyzes the formation of NifB-co, a unique [Fe–S] cluster intermediate

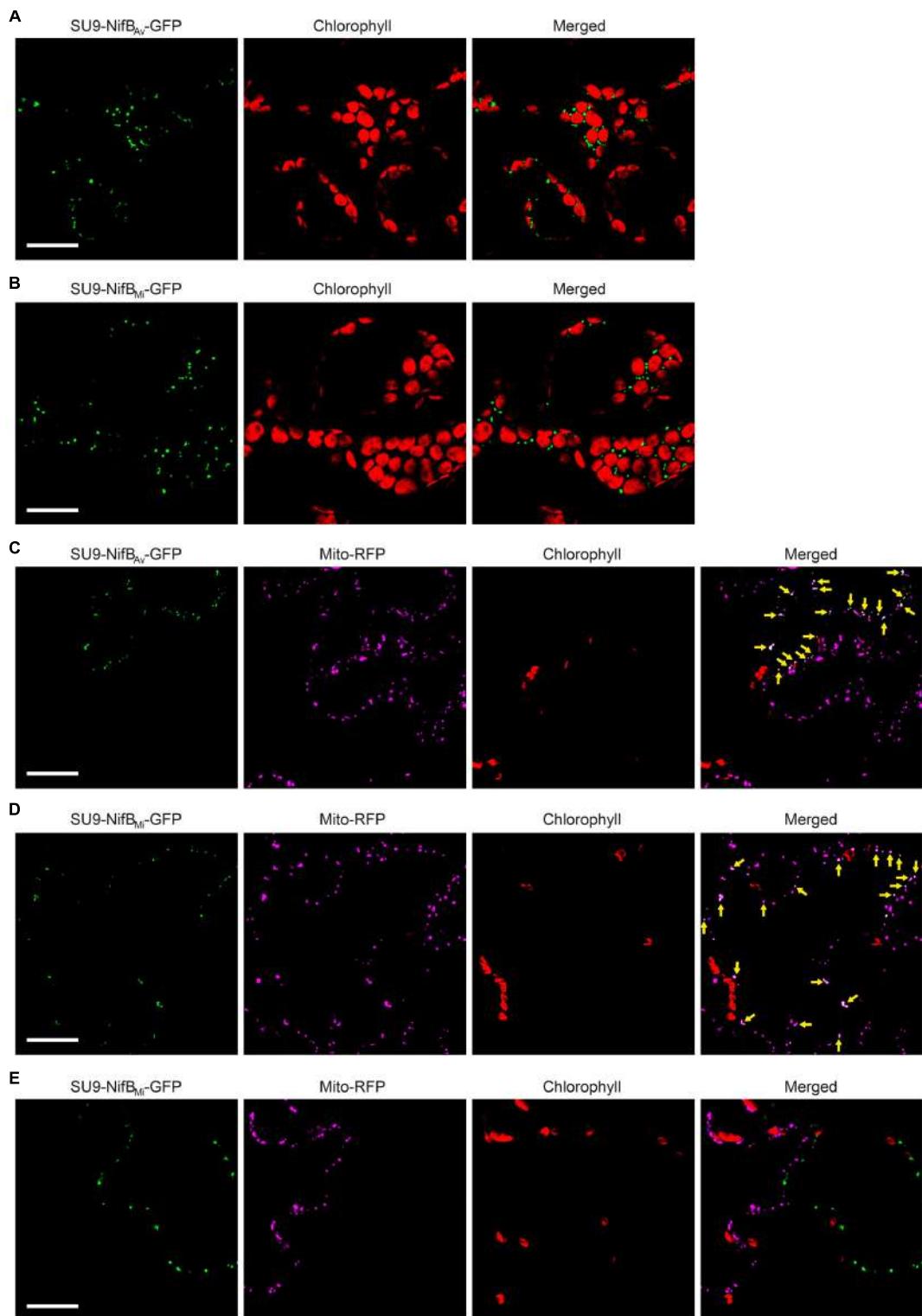


FIGURE 5 | Expression of mitochondria targeted (SU9) NifB_{Av} and NifB_{Mi} GFP fusions in *N. benthamiana* leaves. **(A,B)** Mesophyll cells expressing SU9-NifB_{Av}-GFP (A) or SU9-NifB_{Mi}-GFP (B). GFP (green) and chlorophyll autofluorescence (red) of chloroplasts is shown. **(C–E)** Epidermal cells expressing SU9-NifB_{Av}-GFP (C) and SU9-NifB_{Mi}-GFP (D,E), together with a fluorescent mitochondria marker (Mito-RFP). GFP (green), Mito-RFP (magenta) and chlorophyll autofluorescence (red) of chloroplasts is shown. Co-localization of SU9-NifB_{Av}-GFP or SU9-NifB_{Mi}-GFP constructs with Mito-RFP labeled structures is shown as white in the merged images, and highlighted with yellow arrows. Adjacent cells expressing SU9-NifB_{Mi}-GFP or Mito-RFP are shown as control to verify the specificity of the signal recorded in each channel (E). Scale bars show 30 μ m. Confocal Microscopy conditions are specified in Materials and Methods.

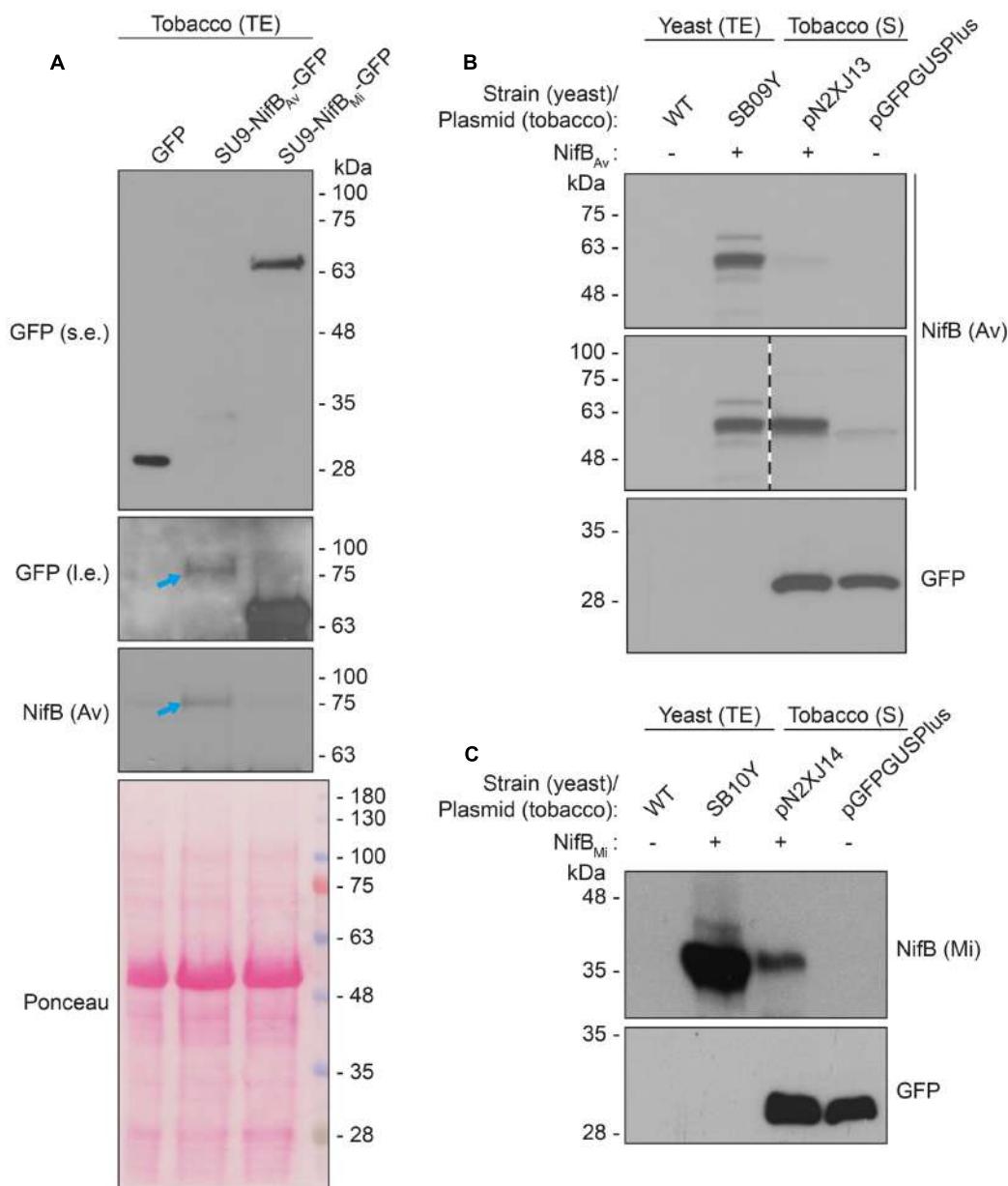


FIGURE 6 | Expression and solubility of mitochondria targeted (SU9) NifB_{Av} and NifB_{Mi} in *N. benthamiana* leaves. **(A)** Western blot analysis of total protein extracts (TE) prepared from infiltrated *N. benthamiana* leaves expressing GFP, SU9-NifB_{Av}-GFP or SU9-NifB_{Mi}-GFP. Blue arrows indicate the polypeptide recognized both by GFP and NifB_{Av} specific antibodies. Short (s.e.) and long (l.e.) film exposures of the GFP antibody probed membrane are shown. **(B)** Migration of SU9-NifB_{Av}-His₁₀ when expressed in *S. cerevisiae* and *N. benthamiana*. Migration in SDS-PAGE was determined after Western blot analysis using NifB_{Av} specific antibodies. Total protein extracts (TE) from W303-1a *S. cerevisiae* cells (WT) or cells expressing SU9-NifB_{Av}-His₁₀ (SB09Y) were prepared. Soluble protein extracts (S) from *N. benthamiana* leaf cells infiltrated with *A. tumefaciens* containing control vector (pGFPGUSPlus) or vector for expression of SU9-NifB_{Av}-His₁₀ (pN2XJ13). Dotted line indicate different exposures of the right part of the membrane. See Supplementary Figure S9 for entire gel of the cropped exposure. **(C)** Migration of SU9-NifB_{Mi}-His₁₀ when expressed in *S. cerevisiae* and *N. benthamiana*. Migration in SDS-PAGE was determined after Western blot analysis using NifB_{Mi} specific antibodies. Total protein extracts (TE) from W303-1a *S. cerevisiae* cells (WT) or cells expressing SU9-NifB_{Mi}-His₁₀ (SB10Y) were prepared. Soluble protein extracts (S) from *N. benthamiana* leaf cells infiltrated with *A. tumefaciens* containing control vector (pGFPGUSPlus) or vector for expression of SU9-NifB_{Mi}-His₁₀ (pN2XJ14). As control of *N. benthamiana* leaf infiltration, GFP expressed from the pGFPGUSPlus vector backbone was detected (B,C).

in the biosynthesis of FeMo-co of nitrogenase (Shah et al., 1994; Curatti et al., 2006; Wiig et al., 2012; Guo et al., 2016). All diazotrophs carry at least one *nifB* gene (Dos Santos et al., 2012), and it is not likely that NifB-co can be produced by any other

enzyme of plant origin (Vicente and Dean, 2017). As NifB-co also serves as precursor to the FeV-co of V-nitrogenase and the FeFe-co of Fe-only nitrogenase, it is required for all biological nitrogen fixation activity in nature (Bishop and Joerger, 1990;

TABLE 3 | Tobacco expressed nitrogenase related proteins and their expected sizes.

Plasmid	Expressed protein	Full-length (kDa)	Processed (kDa)
pN2XJ13	SU9-NifB _{Av} -His ₁₀	64.0	56.7
pN2XJ14	SU9-NifB _{Mi} -His ₁₀	44.3	37
pN2XJ15	SU9-NifB _{Av} -GFP	90.0	82.7
pN2XJ16	SU9-NifB _{Mi} -GFP	70.3	63
pN2XJ19	COX4-twinStrep-GFP	33.3	30.3
pN2XJ20	COX4-twinStrep- NifB _{Av}	61.1	58.1
pN2XJ21	COX4-twinStrep- NifB _{Mi}	41.3	38.3

Plant expression vectors generated in order to express nitrogenase related proteins. Expressed proteins and their expected sizes are indicated. Processed forms refer to sizes following SU9 (RAY-SSM, **Figure 4**) or COX4 (YLL-QQK, (Vögtle et al., 2009)) cleavage as in yeast. See Supplementary Figures S7, S8, S10 for details.

Curatti et al., 2007). Therefore, finding NifB proteins able to function in eukaryotic cells is of utmost importance.

In this work, we investigated eukaryotic expression and functionality of NifB proteins from two evolutionary distant organisms, the γ -proteobacterium *A. vinelandii* and the methanogen *M. infernus*. These NifB proteins differ in domain composition, quaternary structure, optimum temperature and stability, but both contain the same complement of [Fe-S] clusters, catalyze the SAM-radical dependent formation of NifB-co (Curatti et al., 2006; Wilcoxon et al., 2016), and have been shown to support FeMo-co biosynthesis *in vivo* (Arragain et al., submitted manuscript).

As most Nif proteins, NifB is extremely O₂-labile. Expressed NifB variants were therefore targeted to mitochondria for respiratory protection (Lopez-Torrejon et al., 2016). Mitochondria targeting in yeast was achieved using the SU9 leader sequence, which had proved efficient for Nif protein targeting and processing in a previous study (Burén et al., 2017). Expression of the NifB variants in *S. cerevisiae* was coordinated with expression of *A. vinelandii* NifU, NifS and FdxN, as these proteins are important for the assembly of NifB [Fe-S] clusters and for NifB functionality (Jiménez-Vicente et al., 2014; Zhao et al., 2007). Both NifB variants appeared mostly insoluble and the major protein pools were pelleting together with the cellular debris of the broken cells. The accumulation insoluble NifB might be a result of overexpression or NifB hydrophobicity. It is intriguing to note that overexpression of NifB proteins often result in insoluble aggregates and that purification of NifB from different organisms, as well as purification of NifB-co itself, requires detergents (Curatti et al., 2006; Echavarri-Erasun et al., 2014; Shah et al., 1994).

The two NifB variants were also expressed and targeted to mitochondria of *N. benthamiana* leaf cells. Although the extraction methods and buffers used to prepare protein extracts were slightly different in yeast and tobacco, the results obtained from both systems were similar, i.e. expression levels were slightly higher and solubility was significantly better for the *M. infernus* NifB. Intriguingly, solubility overall appeared better in tobacco than in yeast, perhaps due to lower expression levels from the 35S promoter compared to the very strong GAL

promoters. Mitochondria targeting was achieved both with SU9 and COX4 leader sequences, expanding the synthetic toolbox for Nif expression in plants. The identical migration in SDS-PAGE of each set of yeast- and tobacco-expressed NifB variants suggests that all of them underwent correct processing, which in the case of yeast SU9-NifB_{Mi} was demonstrated to occur at the predicted peptide bond between tyrosine and serine residues.

Taking advantage of the heat-resistant properties of *M. infernus* NifB, a protocol to enrich levels of the protein in the soluble fraction of yeast cell-free extracts was developed permitting further purification and biochemical analysis. Pure yNifB_{Mi} preparations exhibited properties characteristic of bacteria-purified NifB proteins (Curatti et al., 2006; Wilcoxon et al., 2016), including color, Fe and S content, and UV-vis spectra changes upon [Fe-S] cluster reconstitution and reduction. Importantly, yNifB_{Mi} showed activities similar to the NifB_{Av} protein (purified from *A. vinelandii* cells) when used to complement Δ nifB *A. vinelandii* extracts in FeMo-co synthesis assays.

Whether or not NifB proteins were functional *in vivo* in *S. cerevisiae* is not known, as we have not yet been able to establish an *in vivo* assay for NifB activity. In this regard, we tested whether yNifB_{Mi} in the heat-treated extract was functional without reconstitution, but failed to detect activity, indicating that yNifB_{Mi} has low or no activity prior to [Fe-S] cluster reconstitution, or that some of its [Fe-S] clusters are lost during yeast cell lysis and/or heat-treatment. Importantly, addition of NifB-co to UW140 extract, in the presence or absence of yeast cell-free extract, resulted in identical activity, suggesting that the heat-treated cell-free extract did not inhibit the *A. vinelandii* nitrogenase activity and that lack of detectable yNifB_{Mi} activity was not due to inhibition of the yeast extract *per se* (data not shown).

We recently used a synthetic biology approach to assemble a yeast library for the expression of nine *A. vinelandii* nif genes (*nifHDKUSMBEN*), where the gene products were targeted to the yeast mitochondria using distinct mitochondria leader sequences (Burén et al., 2017). That work highlighted that expression levels (using distinct promoter/terminator combinations) and mitochondria signals in many cases need to be empirically tested for each gene. We also learned that two isoforms of the NifD polypeptide accumulated in the yeast mitochondria, one of which was the result of N-terminal degradation of the full-length NifD. A similar result could be observed when NifD was expressed in leaf cells of tobacco (Allen et al., 2017), where their NifD degradation isoform showed a SDS-PAGE migration very similar to ours. A plausible explanation to this NifD degradation could be instability of NifDK precursors, as stability of the NifDK tetramer is improved with protein maturation and nitrogenase co-factor (FeMo-co) insertion. In this regard, expression of functional NifB will stabilize NifDK and help increase its levels.

In summary, we have purified NifB protein expressed in a eukaryotic cell background. Following NifH and NifU (Lopez-Torrejon et al., 2016), to our knowledge this is the third Nif protein that has been purified with specific *in vitro* activity. Using both yeast and tobacco as expression hosts, we observed that the monomeric NifB_{Mi} (Wilcoxon et al., 2016) was expressed at

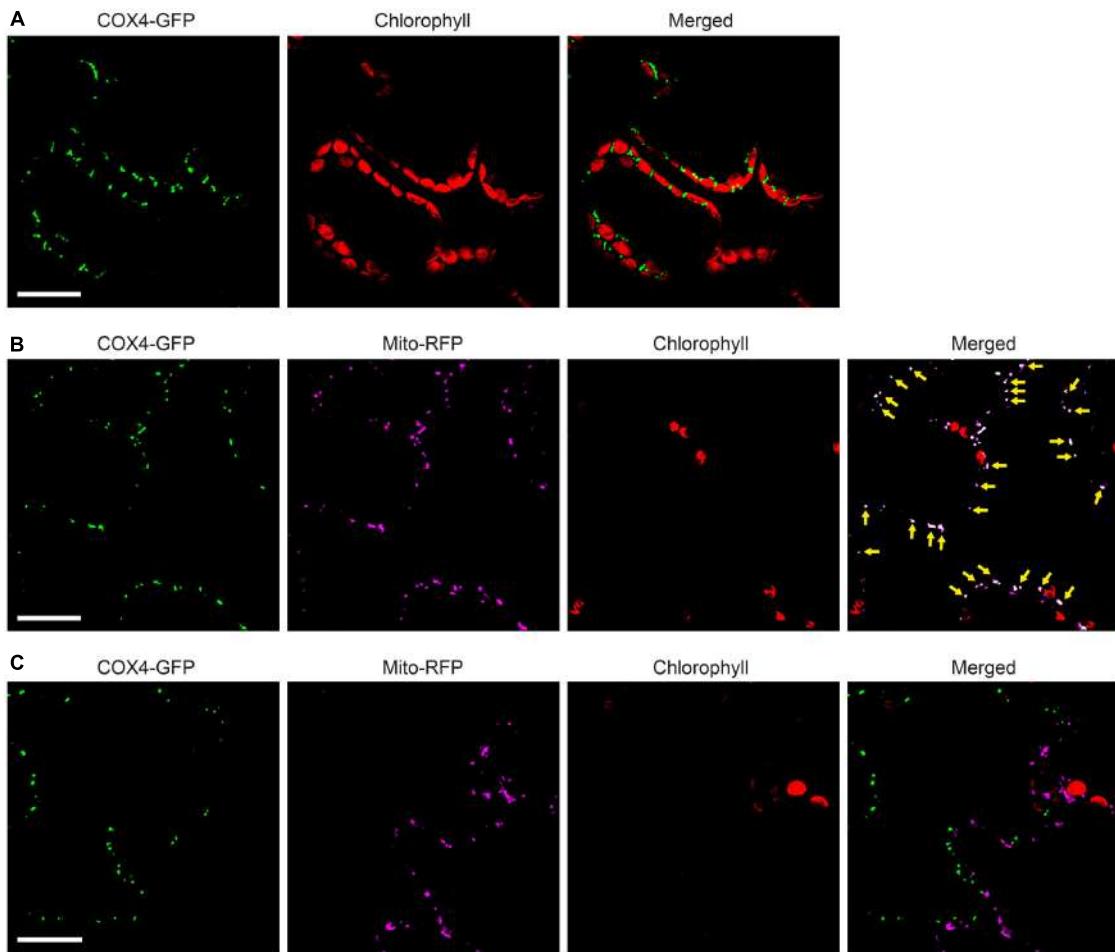


FIGURE 7 | Functionality of COX4 leader sequence for mitochondria targeting of GFP in *N. benthamiana* leaves. **(A)** Mesophyll cells expressing COX4-twinStrep-GFP. GFP (green) and chlorophyll autofluorescence (red) of chloroplasts is shown. **(B,C)** Epidermal cells expressing COX4-twinStrep-GFP together with a fluorescent mitochondria marker (Mito-RFP). GFP (green), Mito-RFP (magenta) and chlorophyll autofluorescence (red) of chloroplasts is shown. Co-localization of COX4-twinStrep-GFP with Mito-RFP labeled structures (B) is shown as white in the merged image, and highlighted with yellow arrows. Adjacent cells expressing COX4-twinStrep-GFP or Mito-RFP (C) are shown as control to verify the specificity of the signal recorded in each channel. Scale bars show 30 μ m. Confocal Microscopy conditions are specified in Materials and Methods.

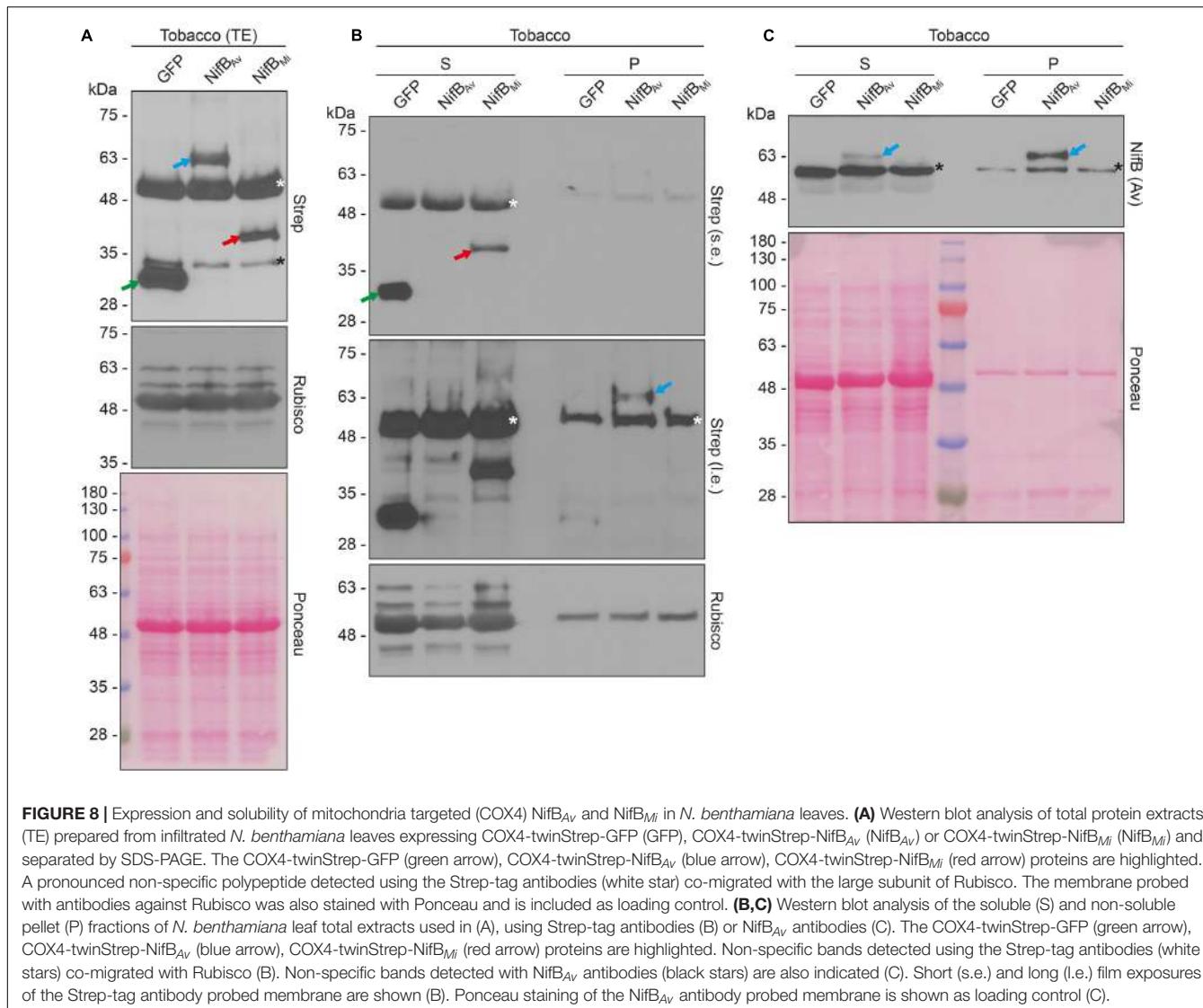
higher levels, and to a higher extent accumulated as a soluble protein, than the dimeric *NifB_{Av}*. This study emphasizes that simpler and more robust Nif protein variants could provide advantages in the ambitious goal of obtaining a functional plant expressed nitrogenase. Importantly, this study confirms that yeast synthetic biology provides a valuable tool in the initial designing and screening process for Nif protein expression and functionality, prior to performing more complex and time-consuming plant-based experiments.

MATERIALS AND METHODS

Generation of Plasmids for Galactose-Induced Yeast Expression

Escherichia coli DH5 α was used for storage and amplification of yeast expression vectors. *E.coli* was grown at 37°C in

Luria-Bertani (LB) medium supplemented with appropriate antibiotics. Yeast optimized coding sequences for *nifU*, *nifS*, *nifB* (*A. vinelandii* and *M. infernus*) and *fdxN* with in-frame su9 leader sequences (Westermann and Neupert, 2000) were generated by GenScript, or by overlapping PCR reactions as specified below, and cloned into pESC vectors (Agilent Technologies) using standard techniques. *su9-nifU* and *su9-nifS* were cloned into pESC-URA using *BamHI/HindIII* and *EcoRI/BglII*, respectively, generating pN2GLT4. *su9-fdxN* and *su9-His₁₀-nifB_{Av}* were cloned into pESC-TRP using *NotI/ClaI* and *BamHI/SalI*, respectively, generating pN2GLT18. *su9-nifB_{Av}-His₁₀* and *su9-nifB_{Mi}-His₁₀* were created using overlapping PCR, to add su9 and His₁₀ at the 5'- and 3'-termini of *nifB_{Av}* and *nifB_{Mi}*. Primers used for generating *su9-nifB_{Av}-His₁₀* were 5'-ATTTTCGGTTGTATTACTTC-3' and 5'-CATGGAAGA GTAGGCGC-3' (using pN2GLT18 as template), 5'-GCGCCT ACTCTCCATGGAATTGTCTGTTTGGGT-3' and 5'-ATG



ATGGTGTTGGTGTAGATGATGAGCCTAGCTTGCAAC-3' (using pN2GLT18 as template), 5'-ATCACCAACCACCATCAT CACCATTAAGTCGACATGGAACA-3' and 5'-GTACACCGCGT CTGTACAGAA-3' (using pN2GLT18 as template), to amplify su9, nifB_{Av} and His₁₀, respectively. 5'-ATTTCGGTTTGATA TTACTTC-3' and 5'-GTACACCGCGTCTGTACAGAA-3' were used for the overlapping PCR reaction. Primers used for generating su9-nifB_{Mi}-His₁₀ were 5'-ATTTCGGTTTGATT ACTTC-3' and 5'-CATGGAAGAGTAGGCAC-3' (using pN2GLT18 as template), 5'-GCGCCTACTCTCCATGGAGAA AATGTCTAAATT-3' and 5'-ATGATGGTGGTGGTGTGATG ATGATGGTGTGAGAAATGCTTC-3' (using nifB_{Mi} as template), 5'-ATCACCAACCACCATCATCACCATTAAGTCG ACATGGAACA-3' and 5'-GTACACCGCGTCTGTACAGAA-3' (using pN2GLT18 as template), to amplify su9, nifB_{Mi} and His₁₀, respectively. 5'-ATTTCGGTTGTATTACTTC-3' and 5'-GTA CACCGCGTCTGTACAGAA-3' were used for the overlapping PCR reaction. su9-nifB_{Av}-His₁₀ and su9-nifB_{Mi}-His₁₀ were

cloned into pN2GLT18, replacing su9-His₁₀-nifB_{Av} using BamHI/SalI, and generating pN2SB22 and pN2SB24, respectively. su9-fdxN-HA was created using overlapping PCR, to add HA at the 3'-terminus of su9-fdxN. Primers used for generating su9-fdxN-HA were 5'-GGTGGTAATGCC ATGTAATATG-3' and 5'-GCATAATCTGGAACATCATATGGG TACCTTGCCTGTATT-3' (using pN2SB22 as template), 5'-GATGTTCCAGATTATGCTTAAGAGCTCTTAATTAACAA TT-3' and 5'-AAAGTTAACCGCATCAGGAAATTGTAA-3' (using pN2SB22 as template), to amplify su9-fdxN and HA, respectively. 5'-GGTGGTAATGCCATGTAATATG-3' and 5'-AAAGTTAACCGCATCAGGAAATTGTAA-3' were used for the overlapping PCR reaction. su9-fdxN-HA was cloned into pN2SB24, replacing su9-fdxN using NotI/PacI, generating pN2SB39. To make pN2GLT18 (su9-fdxN and su9-His₁₀-nifB_{Av}) compatible with transformation into prototrophic *S. cerevisiae* CEN.PK113-7D clone DOE56, the LEU2 auxotrophic marker was replaced with the hygromycin marker hphMX4 (Burén et al.,

2017), generating pN2SB15. DNA and protein sequences of all constructs are listed in Supplementary Figure S2.

Generation of Yeast Strains, Growth, and Protein Expression

Saccharomyces cerevisiae W303-1a (*MATA leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*) was the host strain for expression vectors pN2GLT4 and pN2SB22 (to generate strain SB09Y), pN2GLT4 and pN2SB24 (to generate strain SB10Y), and pN2GLT4 and pN2SB39 (to generate strain SB12Y). CEN.PK113-7D (*MATA URA3 TRP1 LEU2 HIS3 MAL2-8c SUC2*) strain DOE56 (having constitutive expression of mitochondria targeted NifU and NifS) (Burén et al., 2017) was the host strain for expression vector pN2SB15 (to generate strain SB03Y). Yeast transformations were carried out according to the lithium acetate method (Gietz and Schiestl, 2007).

Saccharomyces cerevisiae were grown in flasks at 28°C and 200 rpm in synthetic drop-out (SD) medium (1.9 g/l yeast nitrogen base, 5 g/l ammonium sulfate, 20 g/l glucose, and Kaiser drop-out mixture (Kaiser et al., 1994) (SC -His-Leu-Trp-Ura, FORMEDIUM) supplemented with 20 mg/l adenine and 40 mg/l tryptophan, 40 mg/l histidine, 20 mg/l uracil, 60 mg/l leucine, depending on auxotrophic requirements). Plasmid for the inducible expression of SU9-FdxN and SU9-His₁₀-NifB_{Ay} in transformed DOE56 (SB03Y) was maintained by supplementing the inoculum growth media with 300 µg/l hygromycin. Galactose induction for small-scale protein extracts was performed in the above-described SD medium in which glucose was replaced by 20 g/l galactose, and additionally supplemented with 0.1% yeast extract and 1% peptone. Total yeast protein extracts to verify protein expression were performed using mild alkali treatment (Kushnirov, 2000). Similar loading on SDS-PAGE experiments was obtained by preparing samples according to optical density, and was confirmed by using either Commassie staining of polyacrylamide gels or Ponceau staining of nitrocellulose membranes. Additionally, immunoblotting with antibodies against tubulin was used as control of gel loading and sample precipitation.

Cultures for yeast expressed NifB purifications were grown following a previously reported procedures (Lopez-Torrejon et al., 2016; Burén et al., 2017), in a 4-L fermenter (BIO-STAT). Cultures were grown at 30°C in selective SD-medium for 16 h, followed by 8 h in rich medium (0.5% yeast extract, 0.5% bactopeptone, 0.5% bactotryptone, 2.5% glucose), supplemented with 25 mg/L ammonium iron(III) citrate, 1.25 mM magnesium sulfate, 1.5 mM calcium chloride, and trace element and vitamin solutions (Lopez-Torrejon et al., 2016). Finally, protein expression was induced by addition of 2.5% galactose for 16 h. The pH was automatically maintained around 5 using 0.8 M ammonium hydroxide. Air flow was maintained at 0.75 liter air/min per liter of culture, at 300 rpm. Dissolved oxygen dropped to zero (as measured by oxygen sensor, Mettler Toledo) before addition of galactose, and remained at zero during the rest of the process.

Reverse Transcription and Quantitative Real-time Polymerase Chain Reactions

Total yeast RNA was extracted from 25 ml cultures 6 h following galactose induction. Briefly, the cells were harvested by centrifugation (10 min at 3,000 x g), washed once in Milli-Q water and then resuspended in 1 ml TRIzol reagent (ThermoFisher Scientific). Cells were broken in 2 ml screw-cap tubes using 0.5 mm glass beads (BioSpec Products) in a mixer mill (Retsch MM300) operating at 30 Hz, in 5 cycles of 1 min at 4°C. Two hundred µl chloroform was added to the lysate, vortexed for 15 sec and incubated 5 min at room temperature. Samples were then centrifuged at full speed for 5 min at 4°C. The supernatant was transferred to new tube and re-extracted with 400 µl chloroform. The supernatant containing RNA was precipitated with 1 volume isopropanol at -20°C for 20 minutes and then pelleted by centrifugation at full speed for 5 min at 4°C. The pellet was washed twice in 1 ml 70% ethanol and finally resuspended in nuclease free water. RNA concentration was measured in a nanodrop apparatus and quality was analysed by agarose gel electrophoresis. RNA was then treated with DNase to remove eventual DNA contamination (TURBO DNA-free Kit, AM1907, Ambion). Absence of DNA was verified using polymerase chain reaction (PCR) (KAPA2G Fast HotStart ReadyMix, B4KK5609, KAPA2G). Four µg RNA was used for cDNA synthesis (High-Capacity cDNA Reverse transcription Kit, 4374966, Applied Biosystems). The presence of *fdxN* cDNA was verified by reverse transcription PCR (RT-PCR) using 4 distinct *fdxN* primer combinations (A, 5'-TGTGAAC TGCTGGCATGTG-3' and 5'-TCTCCATCACACTCGGTGCAT-3'; B, 5'-TGCACC GAGTGTGATGGAGA-3' and 5'-TGCCCTAGCCAATCTTC AGGT-3'; C, 5'-ACCGAGTGTGATGGAGACTAT-3' and 5'-GT AAGTGAACCAGGTGGTTAG-3'; D, 5'-CCTTGGCAGGTC CTCATTT-3' and 5'-TAGCACCTCACTGGACAATA-3'). Primers targeting housekeeping genes (*alg9*, 5'-CACGGATAG TGGCTTGGTGAACAATTAC-3' and 5'-TATGATTATCTGG CAGCAGGAAAGAACTTGGG-3'; *rdn18*, 5'-AACTCACCAGG TCCAGACACAATAAGG-3' and 5'-AAGGTCTCGTTCGTTA TCGCAATTAAAGC-3') were selected based on Teste and colleagues (Teste et al., 2009). Quantitative real-time PCR (qPCR) (KAPA SYBR FAST Universal qPCR Kit, KK4601, KAPA2G) was performed using two *fdxN* primer combinations in addition to primers targeting the housekeeping genes, using an Eco Real-Time PCR System (Illumina) following user instructions.

Solubility of Yeast-Expressed NifB

Saccharomyces cerevisiae cells expressing yNifB_{Ay} and yNifB_{Mi} were resuspended in 5 volumes of lysis buffer (100 mM Tris-HCl, 400 mM NaCl, 5 mM β-mercaptoethanol (β-ME), 1 mM phenylmethylsulfonyl fluoride (PMSF)), at pH 7 or 8 with 10% or 30% glycerol. Cells were broken in 2 ml tubes using 0.5 mm glass beads (BioSpec Products) in a mixer mill (Retsch MM300) operating at 30 Hz in 3 cycles of 1 min at 4°C. Lysates were incubated at room temperature (RT), or heated at 5°C temperature intervals from 40°C to 75°C, for 30 min. The supernatant after 20 min centrifugation at 20,000 x g and 4°C containing soluble proteins was analyzed by SDS-PAGE and immunoblot analysis.

Preparation of Yeast Anaerobic Cell-Free Extracts and NifB Purifications

Saccharomyces cerevisiae cells expressing $y\text{NifB}_{Av}$ and $y\text{NifB}_{Mi}$ were resuspended in anaerobic lysis buffer (100 mM Tris-HCl pH 8.0, 400 mM NaCl, 10% glycerol) supplemented with 2 mM dithionite (DTH), 5 mM β -ME, 1 mM PMSF, 1 $\mu\text{g}/\text{ml}$ leupeptin and 5 $\mu\text{g}/\text{ml}$ DNase I. The cells were lysed in an Emulsiflex-C5 homogenizer (Avestin Inc.) at 25,000 lb per square inch. Cell-free extracts (CFE) were obtained after heat-treatment in a water bath ($y\text{NifB}_{Mi}$ only, 65°C for 30 min), removal of cell debris and precipitated yeast proteins by centrifugation (50,000 $\times g$ for 1 h at 4°C) and filtration through a 0.2 μM pore size filter (Nalgene Rapid-Flow, Thermo Scientific). All procedures were performed under anaerobic conditions.

Saccharomyces cerevisiae cells expressing SU9-His₁₀-NifB_{Av} were lysed as described above, in buffer with detergents (50 mM Tris-HCl pH 8, 200 mM KCl, 10% glycerol, 5 mM β -ME, 0.05% n-dodecyl- β -D-maltopyranoside, 0.1% triton X-100 and 0.1% Tween 20) as previously described for purification of NifB_{Mi} from *E. coli* (Wilcoxen et al., 2016).

His-tagged $y\text{NifB}_{Mi}$ was purified by Co²⁺ affinity chromatography under anaerobic conditions (<0.1 ppm of O₂) using an AKTA Prime FPLC system (GE Healthcare) inside a glovebox (MBraun). All buffers were previously made anaerobic by sparging with N₂. Before loading the affinity column, the cell-free extract was diluted to reach 50 mM Tris-HCl, while maintaining other buffer components. Typically, anaerobic cell-free extract from 100 g of cell paste was loaded at 2 ml/min onto a column filled with 5 ml of IMAC resin (GE Healthcare) equilibrated with buffer A (50 mM Tris-HCl pH 8, 400 mM NaCl, 10% glycerol, 2 mM DTH, 5 mM β -ME) and washed with four successive washes of buffer A supplemented with 0, 10, 40 and 100 mM imidazole (10–15 column volumes per wash), respectively. Bound protein was eluted in two steps, with buffer A containing 200 and 500 mM imidazole, respectively. Eluted fractions showing the desired purity were pooled and concentrated using a 100 kDa cutoff pore centrifugal membrane device (Amicon Ultra-15, Millipore), and then desalting in PD10 columns (GE Healthcare) equilibrated with buffer A. Pure $y\text{NifB}_{Mi}$ was frozen and stored in liquid N₂.

In Vitro Reconstitution of $y\text{NifB}_{Mi}$ Fe-S Clusters, UV-Visible Spectroscopy, N-terminal Sequencing and Protein Methods

In vitro reconstitution of purified $y\text{NifB}_{Mi}$ was performed as previously described with modifications (Curatti et al., 2006). Pure $y\text{NifB}_{Mi}$ stored in buffer A was buffer-exchanged to buffer B (50 mM Tris-HCl pH 8, 400 mM NaCl, 10% glycerol, 5 mM β -ME) by using a PD10 column to recover “as isolated” protein. The desalted sample (20 μM NifB monomer) was incubated with 10 mM DTT at room temperature inside a glovebox (MBraun) for 10 min. (NH₄)₂Fe(SO₄)₂ and Na₂S were then added at 20-fold molar excess ratio and incubated at 35°C overnight. $y\text{NifB}_{Mi}$ was again desalting in buffer B to recover “reconstituted”

protein. As isolated and reconstituted proteins were used for colorimetric Fe (Fish, 1988) and S (Beinert, 1983) determination, *in vitro* FeMo-co synthesis and nitrogenase activity assays, and UV-visible spectroscopy. UV-visible absorption spectra were recorded under anaerobic conditions in septum-sealed cuvettes using a Shimadzu UV-2600 spectrophotometer. When indicated, 5 mM DTH was added to reconstituted $y\text{NifB}_{Mi}$. UV-visible absorption spectra were recorded against buffer B as baseline. Absorbance at 800 nm was subtracted and spectra were then normalized to 279 nm. The N-terminal amino acid sequence of purified $y\text{NifB}_{Mi}$ was determined by Edman degradation (Proteome Factory AG). Protein concentrations were measured using the BCA protein assay (PIERCE). NifB samples were pre-treated with iodoacetamide before performing the BCA assay to eliminate the interfering effect of DTH (Hill and Straka, 1988).

In Vitro Synthesis of FeMo-co and Nitrogenase Reconstitution Assay

In vitro $y\text{NifB}_{Mi}$ dependent FeMo-co synthesis and nitrogenase reconstitution reactions were performed in 9-ml serum vials sealed with serum stoppers (Curatti et al., 2006). Complete reactions contained 17.5 μM Na₂MoO₄, 175 μM homocitrate, 1.75 mM (NH₄)₂FeSO₄, 1.75 mM Na₂S, 880 μM SAM, 1.23 mM ATP, 18 mM phosphocreatine, 2.2 mM MgCl₂, 3 mM DTH, 40 $\mu\text{g}/\text{ml}$ creatine phosphokinase, 2.2 μM NifH (dimer), 2.9 mg/ml UW140 (*A. vinelandii* Δ nifB) proteins, 5 μM (or 0–10 μM titration) reconstituted $y\text{NifB}_{Mi}$ (monomer) in 22 mM Tris-HCl (pH 7.5). The reactions (total volume of 500 μl) were incubated at 30°C for 35 min to allow for FeMo-co synthesis and insertion reactions. NifB-co-dependent *in vitro* FeMo-co synthesis assays were performed using 2 μM NifB-co isolated from *K. oxytoca* (Shah et al., 1994). Following *in vitro* synthesis of FeMo-co, activation of apo-MoFe nitrogenase present in UW140 extract was analyzed following addition of excess NifH and ATP-regenerating mixture (total volume 1 ml) by acetylene reduction assay at 30°C for 30 min following standard procedures (Shah and Brill, 1973). Positive control reactions for acetylene reduction were carried out with pure preparations of *A. vinelandii* Fe protein and MoFe protein incubated with ATP-regenerating mixture at 30°C during 30 min.

Generation of Plant Expression Vectors and Protein Expression in Leaves of *N. benthamiana*

Escherichia coli DH5 α was used for storage and amplification of plant expression vectors. *E. coli* was grown at 37°C in LB medium supplemented with appropriate antibiotics. *su9-nifB_{Av}-His₁₀* and *su9-nifB_{Mi}-His₁₀* were PCR amplified using primers 5'-AAAA GGATCCAATGGCCTCCACTCGTGTCTCG-3' and 5'-TTTT CACGTGTTAATGGTGATGGTGGTG-3', with pN2SB22 and pN2SB24 as templates, respectively. *su9-nifB_{Av}-His₁₀* and *su9-nifB_{Mi}-His₁₀* were digested with *Bam*HI and *Pml*I, and inserted into pGFP-GUSplus vector (Vickers et al., 2007) (Addgene plasmid #64401) digested with *Bgl*II and *Pml*I, replacing GUS and generating pN2XJ13 (*su9-nifB_{Av}-His₁₀*) and pN2XJ14 (*su9-nifB_{Mi}-His₁₀*), respectively.

su9-nifB_{Av} was PCR amplified using primers 5'-AAAAGCTA GCATGGCCTCCACTCGTGCCTCG-3' and 5'-TTTGCT AGCGCCTTAGCTGCAACAAAGC-3', with pN2SB22 as template. *su9-nifB_{Av}* was digested with *Nhe*I and inserted into pGFPGUSplus vector digested with *Xba*I, generating pN2XJ15 for expression of *su9-nifB_{Av}-gfp*. *su9-nifB_{Mi}* was PCR amplified using primers 5'-AAAAGCTAGCATGGCC TCCACTCGTGCCTCG-3' and 5'-TTTGCTAGCGCGT GTGAGAAATGCTCAAGTCG-3', with pN2SB24 as template. *su9-nifB_{Mi}* was digested with *Nhe*I and inserted into pGFPGUSplus vector digested with *Xba*I, generating pN2XJ16 for expression of *su9-nifB_{Mi}-gfp*. DNA sequence encoding the enhanced 35S promoter and an in-frame fusion of the *cox4* mitochondria leader sequence (Köhler et al., 1997) with the 28 amino acid Twin-Strep-tag was generated by ThermoFisher. The E35S-cox4-twinStrep DNA sequence was flanked by *Hind*III and *Bgl*II, with a *Bam*HI site additionally added 5' of the *Bgl*II site. E35S-cox4-twinStrep was digested with *Hind*III and *Bgl*II, and inserted into pGFPGUSplus vector also digested with *Hind*III and *Bgl*II, to generate pN2SB41. DNA sequence encoding *egfp* was PCR amplified using primers 5'-AAAAAGGATCCATGGT GAGCAAGGGCGA-3' and 5'-AAAAAGGT CACCTTA CTTACTTG TACAGCTCGTCCATG-3', and pGFPGUSplus as template. *egfp* was digested with *Bam*HI and *Bst*EII, and inserted into pN2SB41 also digested with *Bam*HI and *Bst*EII, creating pN2XJ17. pN2XJ17 was digested with *Pst*I to remove the non-targeted EGFP, to generate pN2XJ19 (*cox4-twinStrep-gfp*). DNA sequences encoding *nifB_{Av}* and *nifB_{Mi}*, flanked by *Bam*HI and *Bst*EII, were generated by ThermoFisher. *nifB_{Av}* and *nifB_{Mi}* were digested with *Bam*HI and *Bst*EII, and inserted into pN2SB41 also digested with *Bam*HI and *Bst*EII, to generate pN2XJ20 (*cox4-twinStrep-nifB_{Av}*) and pN2XJ21 (*cox4-twinStrep-nifB_{Mi}*). DNA and protein sequences of all constructs are listed in Supplementary Figure S8.

Agrobacterium tumefaciens strain GV3101(pMP90) was transformed with plasmids pN2XJ13, pN2XJ14, pN2XJ15, pN2XJ16, pN2XJ19, pN2XJ20, pN2XJ21 and the silencing suppressor p19 (Huang et al., 2009). The pDCL-mito-mRFP1 mitochondria marker (Mito-RFP) in *A. tumefaciens* strain C58 (Candat et al., 2014) was kindly provided by Prof. Macherel and Prof. Logan at the Angers University (France). *A. tumefaciens* mediated infiltration of *N. benthamiana* leaves was essentially performed as described by Leuzinger and colleagues (Leuzinger et al., 2013). Three to four days post infiltration, plant tissue was used for protein extraction or confocal microscopy.

Protein extracts were prepared from infiltrated *N. benthamiana* leaf tissue in lysis buffer (100 mM Tris-HCl pH 8, 150 mM NaCl, 10 mM MgCl₂, 0.2% NP-40, 5% glycerol, 5 mM β-ME and 5 mM ethylenediaminetetraacetic acid (EDTA)). Two hours before use, 5% polyvinylpolypyrrolidone (PVPP) was added to lysis buffer and, just before use, 1 mM PMSF, 1 μg/ml leupeptin and 1x protease inhibitor cocktail (P8215, Sigma) were added. Extraction was performed at a 2:1 ratio of buffer to tissue. Ten leaf discs of 5 mm diameter each (approximate weight of 200 mg) were added to a 2-ml Eppendorf tube containing a 7-mm diameter steel ball. Tubes

were kept in liquid N₂ until use. Leaf tissue was broken using mixer mill (Retsch MM300) operating at 30 Hz for 1 min at 4°C. The dry tissue powder was supplemented with 400 μl lysis buffer and mixed for another 1 min at 30 Hz and 4°C. The broken tissue in lysis buffer was further incubated on an orbital shaker for 30 min at 4°C. One hundred μl extract were added to 100 μl 2x Laemmli buffer (2xLB) and heated for 10 min at 95°C to obtain the “total extract”. The rest of the extract was centrifuged at 20,000 × g for 30 min at 4°C to separate pellet from supernatant. The supernatant “soluble extract (S)” was mixed with 2xLB and heated for 10 min at 95°C. The pellet (P) was resuspended in 1 ml lysis buffer (no additional PVPP added) and centrifuged at 20,000 × g for 10 min at 4°C. Finally, the pellet was resuspended in 800 μl 2xLB and heated for 10 min at 95°C. Ten μl of each fraction were used for SDS-PAGE and immunoblot analysis. Similar sample loading on SDS-PAGE lanes was assessed either by Commassie staining of polyacrylamide gels, by Ponceau staining of transferred nitrocellulose membranes, or by immunoblotting with antibodies against Rubisco.

Confocal Microscopy of *N. benthamiana* Leaf Tissue

Subcellular localization of fluorescent protein tagged proteins was examined in leaves of *A. tumefaciens* infiltrated *N. benthamiana* using a Leica TCS SP8 laser scanning confocal microscope with a 40x/1.10 water immersion objective equipped with LAS X software (Leica). EGFP, RFP, and chlorophyll were excited with 488-, 561-, or 638-nm laser lines, respectively, with an emission band of 500 to 537 nm for EGFP detection, 585 to 620 nm for RFP detection, and 652 to 727 nm for chlorophyll autofluorescence. EGFP and chlorophyll was recorded simultaneously, while RFP was detected in a separate scan. Laser intensity and gain was maintained during each experiment. For each experiment, specificity of the recorded signals was verified using single transformed cells.

Antibodies

Antibodies used for immunoblotting in this study were as follows: polyclonal antibodies detecting NifU_{Av}, NifS_{Av}, NifB_{Av} and NifB_{Mi} were raised against purified preparations of the corresponding *A. vinelandii* or *M. infernus* proteins. Rubisco specific antibodies were a kind gift from Prof. Göran Samuelsson, Umeå University. His-tag (H-3, sc-8036, Santa Cruz), HA-tag (3F10, 12013819001, Roche), GFP (B-2, sc-9996, Santa Cruz), Strep-tag II (StrepMAB-Classic, 2-1507-001, IBA Lifesciences) specific antibodies are commercially available.

AUTHOR CONTRIBUTIONS

SB, XJ, GL-T, and CE-E carried out the experimental work. SB, XJ, GL-T, CE-E and LR contributed to experimental design and data analysis. SB and LR wrote the paper.

FUNDING

Funding for this research was provided by the Bill & Melinda Gates Foundation Grant OPP1143172 (LR).

ACKNOWLEDGMENTS

We thank Jose María Buesa for yeast fermentations and Marcel Veldhuizen for help with cloning. We thank Emilio Jiménez for fruitful discussions about NifB purifications

REFERENCES

- Allen, R. S., Tilbrook, K., Warden, A. C., Campbell, P. C., Rolland, V., Singh, S. P., et al. (2017). Expression of 16 nitrogenase proteins within the plant mitochondrial matrix. *Front. Plant Sci.* 8:287. doi: 10.3389/fpls.2017.00287
- Balk, J., and Pilon, M. (2011). Ancient and essential: the assembly of iron-sulfur clusters in plants. *Trends Plant Sci.* 16, 218–226. doi: 10.1016/j.tplants.2010.12.006
- Beinert, H. (1983). Semi-micro methods for analysis of labile sulfide and of labile sulfide plus sulfane sulfur in unusually stable iron-sulfur proteins. *Anal. Biochem.* 131, 373–378. doi: 10.1016/0003-2697(83)90186-0
- Bishop, P. E., and Joerger, R. D. (1990). Genetics and molecular biology of alternative nitrogen fixation systems. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41, 109–125. doi: 10.1146/annurev.pp.41.060190.000545
- Borlaug, N. E. (2002). Feeding a world of 10 billion people: the miracle ahead. *Vitr. Cell. Dev. Biol. Plant* 38, 221–228. doi: 10.1079/IVP2001279
- Bulen, W. A., and LeComte, J. R. (1966). The nitrogenase system from Azotobacter: two-enzyme requirement for N₂ reduction, ATP-dependent H₂ evolution, and ATP hydrolysis. *Proc. Natl. Acad. Sci. U.S.A.* 56, 979–986.
- Burén, S., Young, E. M., Sweeny, E. A., López-Torrejón, G., Veldhuizen, M., Voigt, C. A., et al. (2017). Formation of nitrogenase NifDK tetramers in the mitochondria of *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 6, 1043–1055. doi: 10.1021/acssynbio.6b00371
- Candat, A., Paszkiewicz, G., Neveu, M., Gautier, R., Logan, D. C., Avelange-Macherel, M.-H., et al. (2014). The ubiquitous distribution of late embryogenesis abundant proteins across cell compartments in *Arabidopsis* offers tailored protection against abiotic stress. *Plant Cell* 26, 3148–3166. doi: 10.1105/tpc.114.127316
- Curatti, L., Hernandez, J. A., Igarashi, R. Y., Soboh, B., Zhao, D., and Rubio, L. M. (2007). In vitro synthesis of the iron-molybdenum cofactor of nitrogenase from iron, sulfur, molybdenum, and homocitrate using purified proteins. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17626–17631. doi: 10.1073/pnas.0703050104
- Curatti, L., Ludden, P. W., and Rubio, L. M. (2006). NifB-dependent in vitro synthesis of the iron – molybdenum cofactor of nitrogenase. *Proc. Natl. Acad. Sci. U.S.A.* 103, 5297–5301. doi: 10.1073/pnas.0601115103
- Curatti, L., and Rubio, L. M. (2014). Challenges to develop nitrogen-fixing cereals by direct nif-gene transfer. *Plant Sci.* 225, 130–137. doi: 10.1016/j.plantsci.2014.06.003
- Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C., and Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* 13:162. doi: 10.1186/1471-2164-13-162
- Echavarri-Erasun, C., Arragain, S., Scandurra, A. A., and Rubio, L. M. (2014). “Expression and purification of NifB proteins from aerobic and anaerobic sources,” in *Metalloproteins: Methods and Protocols*, eds J. C. Fontecilla-Camps and Y. Nicolet (Totowa, NJ: Humana Press), 19–31. doi: 10.1007/978-1-62703-794-5_3
- Emerich, D. W., and Burris, R. H. (1976). Interactions of heterologous nitrogenase components that generate catalytically inactive complexes. *Proc. Natl. Acad. Sci. U.S.A.* 73, 4369–4373.
- Emerich, D. W., and Burris, R. H. (1978). Complementary functioning of the component proteins of nitrogenase from several bacteria. *J. Bacteriol.* 134, 936–943.
- and activity assays. We thank David Macherel and David Logan for providing the RFP mitochondria marker, and Göran Samuelsson for the anti-Rubisco specific antibodies.
- Fish, W. W. (1988). Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples. *Methods Enzymol.* 158, 357–364.
- Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., et al. (2008). Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science* 320, 889–892. doi: 10.1126/science.1136674
- Gietz, R. D., and Schiestl, R. H. (2007). Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* 2, 35–37. doi: 10.1038/nprot.2007.14
- Guo, Y., Echavarri-Erasun, C., Demuez, M., Jimenez-Vicente, E., Bominaar, E. L., and Rubio, L. M. (2016). The nitrogenase FeMo-cofactor precursor formed by NifB protein: a diamagnetic cluster containing eight iron atoms. *Angew. Chem. Int. Ed. Engl.* 55, 12764–12767. doi: 10.1002/anie.201606447
- Hill, H. D., and Straka, J. G. (1988). Protein determination using bicinchoninic acid in the presence of sulphydryl reagents. *Anal. Biochem.* 170, 203–208. doi: 10.1016/0003-2697(88)90109-1
- Huang, Z., Chen, Q., Hjelm, B., Arntzen, C., and Mason, H. (2009). A DNA replicon system for rapid high-level production of virus-like particles in plants. *Biotechnol. Bioeng.* 103, 706–714. doi: 10.1002/bit.22299
- Ivleva, N. B., Groat, J., Staub, J. M., and Stephens, M. (2016). Expression of active subunit of nitrogenase via integration into plant organelle genome. *PLOS ONE* 11:e0160951. doi: 10.1371/journal.pone.0160951
- Jeanthon, C., L'Haridon, S., Reysenbach, A. L., Vernet, M., Messner, P., Sleytr, U. B., et al. (1998). *Methanococcus infernus* sp. nov., a novel hyperthermophilic lithotrophic methanogen isolated from a deep-sea hydrothermal vent. *Int. J. Syst. Bacteriol.* 48(Pt 3), 913–919. doi: 10.1099/00207713-48-3-913
- Jiménez-Vicente, E., Navarro-Rodríguez, M., Poza-Carrión, C., and Rubio, L. M. (2014). Role of *Azotobacter vinelandii* FdxN in FeMo-co biosynthesis. *FEBS Lett.* 588, 512–516. doi: 10.1016/j.febslet.2013.12.018
- Johnson, D. C., Dean, D. R., Smith, A. D., and Johnson, M. K. (2005). Structure, function, and formation of biological iron-sulfur clusters. *Annu. Rev. Biochem.* 74, 247–281. doi: 10.1146/annurev.biochem.74.082803.133518
- Kaiser, C., Michaelis, S., and Mitchell, A. (1994). *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Köhler, R. H., Zipfel, W. R., Webb, W. W., and Hanson, M. R. (1997). The green fluorescent protein as a marker to visualize plant mitochondria in vivo. *Plant J.* 11, 613–621. doi: 10.1046/j.1365-313X.1997.11030613.x
- Kushnirov, V. V. (2000). Rapid and reliable protein extraction from yeast. *Yeast* 16, 857–860.
- Leuzinger, K., Dent, M., Hurtado, J., Stahnke, J., Lai, H., Zhou, X., et al. (2013). Efficient agroinfiltration of plants for high-level transient expression of recombinant proteins. *J. Vis. Exp.* 77:e50521. doi: 10.3791/50521
- Lill, R., and Muhlenhoff, U. (2008). Maturation of iron-sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. *Annu. Rev. Biochem.* 77, 669–700. doi: 10.1146/annurev.biochem.76.052705.162653
- Lopez-Torrejon, G., Jimenez-Vicente, E., Buesa, J. M., Hernandez, J. A., Verma, H. K., Rubio, L. M., et al. (2016). Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat. Commun.* 7:11426. doi: 10.1038/ncomms11426
- Nelson, B. K., Cai, X., and Nebenführ, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. *Plant J.* 51, 1126–1136. doi: 10.1111/j.1365-313X.2007.03212.x

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01567/full#supplementary-material>

- Oldroyd, G. E. D., and Dixon, R. (2014). Biotechnological solutions to the nitrogen problem. *Curr. Opin. Biotechnol.* 26, 19–24. doi: 10.1016/j.copbio.2013.08.006
- Pan, R., Kaur, N., and Hu, J. (2014). The Arabidopsis mitochondrial membrane-bound ubiquitin protease UBP27 contributes to mitochondrial morphogenesis. *Plant J.* 78, 1047–1059. doi: 10.1111/tpj.12532
- Rubio, L. M., and Ludden, P. W. (2008). Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annu. Rev. Microbiol.* 62, 93–111. doi: 10.1146/annurev.micro.62.081307.162737
- Sanchez, P. A., and Swaminathan, M. S. (2005). Hunger in Africa: the link between unhealthy people and unhealthy soils. *Lancet* 365, 442–444. doi: 10.1016/S0140-6736(05)17834-9
- Schmidt, T. G. M., Batz, L., Bonet, L., Carl, U., Holzapfel, G., Kiem, K., et al. (2013). Development of the Twin-Strep-tag and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr. Purif.* 92, 54–61. doi: 10.1016/j.pep.2013.08.021
- Shah, V. K., Allen, J. R., Spangler, N. J., and Ludden, P. W. (1994). In vitro synthesis of the iron-molybdenum cofactor of nitrogenase. Purification and characterization of NifB cofactor, the product of NIFB protein. *J. Biol. Chem.* 269, 1154–1158.
- Shah, V. K., and Brill, W. J. (1973). Nitrogenase. IV. Simple method of purification to homogeneity of nitrogenase components from *Azotobacter vinelandii*. *Biochim. Biophys. Acta* 305, 445–454.
- Smil, V. (2000). *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of World Food Production*. Cambridge, MA: MIT Press.
- Stokstad, E. (2016). The nitrogen fix. *Science* 353, 1225–1227. doi: 10.1126/science.353.6305.1225
- Teste, M.-A., Duquenne, M., François, J. M., and Parrou, J.-L. (2009). Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. *BMC Mol. Biol.* 10:99. doi: 10.1186/1471-2199-10-99
- Vicente, E. J., and Dean, D. R. (2017). Keeping the nitrogen-fixation dream alive. *Proc. Natl. Acad. Sci. U.S.A.* 114, 3009–3011. doi: 10.1073/pnas.1701560114
- Vickers, C. E., Schenk, P. M., Li, D., Mullineaux, P. M., and Gresshoff, P. M. (2007). pgFPGUSPlus, a new binary vector for gene expression studies and optimising transformation systems in plants. *Biotechnol. Lett.* 29, 1793–1796. doi: 10.1007/s10529-007-9467-6
- Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., et al. (1997). Human alteration of the global nitrogen cycle: sources and consequences. *Ecol. Appl.* 7, 737–750.
- Vögtle, F. N., Wortelkamp, S., Zahedi, R. P., Becker, D., Leidhold, C., Gevaert, K., et al. (2009). Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. *Cell* 139, 428–439. doi: 10.1016/j.cell.2009.07.045
- Westermann, B., and Neupert, W. (2000). Mitochondria-targeted green fluorescent proteins: convenient tools for the study of organelle biogenesis in *Saccharomyces cerevisiae*. *Yeast* 16, 1421–1427.
- Wiig, J. A., Hu, Y., Lee, C. C., and Ribbe, M. W. (2012). Radical SAM-dependent carbon insertion into the nitrogenase M-cluster. *Science* 337, 1672–1675. doi: 10.1126/science.1224603
- Wilcoxon, J., Arragain, S., Scandurra, A. A., Jimenez-Vicente, E., Echavarri-Erasun, C., Pollmann, S., et al. (2016). Electron paramagnetic resonance characterization of three iron-sulfur clusters present in the nitrogenase cofactor maturase NifB from *Methanocaldococcus infernus*. *J. Am. Chem. Soc.* 138, 7468–7471. doi: 10.1021/jacs.6b03329
- Witte, C. P., Noël, L. D., Gielbert, J., Parker, J. E., Romeis, T., Noe, L. D., et al. (2004). Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope. *Plant Mol. Biol.* 55, 135–147. doi: 10.1007/s11103-004-0501-y
- Yuvaniyama, P., Agar, J. N., Cash, V. L., Johnson, M. K., and Dean, D. R. (2000). NifS-directed assembly of a transient [2Fe-2S] cluster within the NifU protein. *Proc. Natl. Acad. Sci. U.S.A.* 97, 599–604. doi: 10.1073/pnas.97.2.599
- Zhao, D., Curatti, L., and Rubio, L. M. (2007). Evidence for nifU and nifS participation in the biosynthesis of the iron-molybdenum cofactor of nitrogenase. *J. Biol. Chem.* 282, 37016–37025. doi: 10.1074/jbc.M708097200

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Burén, Jiang, López-Torrejón, Echavarri-Erasun and Rubio. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Diversity and Functional Analysis of the FeMo-Cofactor Maturase NifB

OPEN ACCESS

Edited by:

Nikolai Provorov,

All-Russian Research Institute of Agricultural Microbiology of the Russian Academy of Agricultural Sciences, Russia

Reviewed by:

Oswaldo Valdes-Lopez, Universidad Nacional Autónoma de México, Mexico

Teresa Thiel, University of Missouri-St. Louis, United States

*Correspondence:

Carlos Echavarri-Erasun carlos.echavarri@upm.es

Luis M. Rubio lm.rubio@upm.es

†Present address:

Simon Arragain, Department of Chemistry, University of California, Davis, Davis, CA, United States

Emilio Jiménez-Vicente, Department of Biochemistry, Virginia Tech, Blacksburg, VA, United States

Alessandro A. Scandurra, Merck, Cambridge, United Kingdom

[‡]These authors have contributed equally to this work.

Specialty section:

This article was submitted to Plant Microbe Interactions, a section of the journal Frontiers in Plant Science

Received: 31 May 2017

Accepted: 30 October 2017

Published: 14 November 2017

Citation:

Arragain S, Jiménez-Vicente E, Scandurra AA, Burén S, Rubio LM and Echavarri-Erasun C (2017) Diversity and Functional Analysis of the FeMo-Cofactor Maturase NifB.

Front. Plant Sci. 8:1947.

doi: 10.3389/fpls.2017.01947

Simon Arragain^{†‡}, Emilio Jiménez-Vicente^{†‡}, Alessandro A. Scandurra[†], Stefan Burén,
Luis M. Rubio^{*} and Carlos Echavarri-Erasun^{*}

Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain

One of the main hurdles to engineer nitrogenase in a non-diazotrophic host is achieving NifB activity. NifB is an extremely unstable and oxygen sensitive protein that catalyzes a low-potential SAM-radical dependent reaction. The product of NifB activity is called NifB-co, a complex [8Fe-9S-C] cluster that serves as obligate intermediate in the biosyntheses of the active-site cofactors of all known nitrogenases. Here we study the diversity and phylogeny of naturally occurring NifB proteins, their protein architecture and the functions of the distinct NifB domains in order to understand what defines a catalytically active NifB. Focus is on NifB from the thermophile *Chlorobium tepidum* (two-domain architecture), the hyperthermophile *Methanocaldococcus infernus* (single-domain architecture) and the mesophile *Klebsiella oxytoca* (two-domain architecture), showing *in silico* characterization of their nitrogen fixation (*nif*) gene clusters, conserved NifB motifs, and functionality. *C. tepidum* and *M. infernus* NifB were able to complement an *Azotobacter vinelandii* ($\Delta nifB$) mutant restoring the Nif⁺ phenotype and thus demonstrating their functionality *in vivo*. In addition, purified *C. tepidum* NifB exhibited activity in the *in vitro* NifB-dependent nitrogenase reconstitution assay. Intriguingly, changing the two-domain *K. oxytoca* NifB to single-domain by removal of the C-terminal NifX-like extension resulted in higher *in vivo* nitrogenase activity, demonstrating that this domain is not required for nitrogen fixation in mesophiles.

Keywords: nitrogenase, iron-molybdenum cofactor, SAM-radical, nitrogen fixation, Azotobacter, methanogens

INTRODUCTION

Although nitrogen is abundant on Earth, most of it is in the form of dinitrogen (N≡N or N₂). Due to the strength of its triple bound, N₂ shows very little reactivity and is therefore not easily available to living organisms (Hoffman et al., 2014). N₂ fixing organisms (diazotrophs) capable of converting N₂ into NH₃, an accessible form of nitrogen, probably appeared in the primordial Earth when the levels of combined nitrogen gradually depleted (Raymond et al., 2004; Canfield et al., 2010). Although evolution and fine-tuning of biological nitrogen fixation (BNF) had an immense impact on the Earth's nitrogen cycle and allowed life to prosper, only a few bacteria and archaea are actually capable of performing it (Boyd and Peters, 2013). The enzymes that catalyze N₂ fixation are called nitrogenases (Burris and Roberts, 1993). Nitrogenases are two-component protein complexes, with a catalytic Component I and a Component II acting as obligate electron donor (Bulen and Lecomte, 1966). Three genetically and biochemically distinct classes of nitrogenases have been described to date: the molybdenum nitrogenase, the vanadium nitrogenase, and the iron-only nitrogenase (Bishop and Joerger, 1990). All diazotrophs carry the Mo-nitrogenase and may or may not carry

the V or Fe-only ones, referred to as alternative nitrogenases (Dos Santos et al., 2012; McGlynn et al., 2013). In the case of the Mo-nitrogenase, the Component I is called MoFe protein and is a heterotetramer of the *nifD* and *nifK* gene products, whereas the Component II is called Fe protein and is a homodimer of the *nifH* gene product. A functional nitrogenase complex requires three metal cofactors embedded in the polypeptide chains to reduce N₂ to NH₃ (Peters et al., 2011). The NifH homodimer carries a [4Fe–4S] cluster located between the two NifH subunits (Georgiadis et al., 1992), while NifDK harbors an [8Fe–7S] P-cluster at the interface of each NifD (α) and NifK (β) subunits, and an iron-molybdenum cofactor (FeMo-co; [7Fe–9S–C–Mo-homocitrate]) embedded 10 Å beneath the surface of each NifD subunit (Einsle et al., 2002; Spatzal et al., 2011). Alternative nitrogenases contain a third type of subunit in Component I, encoded by *vnfG* (V-nitrogenase) or *anfG* (Fe-only nitrogenase), and either FeV or FeFe cofactors at the active site. These cofactors are proposed to be identical to FeMo-co except for containing V or Fe in place of Mo (Eady, 1996).

NifB stands out as the only protein essential for the activity of all nitrogenases (in addition to homocitrate synthase) (Joerger and Bishop, 1988; Dos Santos et al., 2012). NifB is an S-adenosyl methionine (SAM)-radical protein that converts [4Fe–4S] clusters into NifB-co, an [8Fe–9S–C] cluster that serves as precursor to FeMo-co, FeV-co and FeFe-co, thus catalyzing the first committed step in nitrogenase active-site cofactor biosynthesis (Shah et al., 1994; Allen et al., 1995; Curatti et al., 2006; George et al., 2008; Wiig et al., 2012) (Supplementary Figure 1). In contrast to FeMo-co, NifB-co is a diamagnetic cluster containing two spectroscopically distinct Fe sites (Guo et al., 2016).

NifB proteins were first purified from the model bacteria *Azotobacter vinelandii* (Curatti et al., 2006) and *Klebsiella oxytoca* (Zhao et al., 2007). The NifB_{Av} and the NifB_{Ko} proteins contain a C-terminal NifX-like extension that appears to result from gene fusions during evolution of *nifB* (Boyd et al., 2011). The NifX protein is known to bind and transfer NifB-co to the NifEN scaffold protein for further processing into FeMo-co (Hernandez et al., 2007), but the role of the NifX-domain of NifB is not known. NifB from the archaea *Methanocaldococcus infernus*, expressed and purified from recombinant *Escherichia coli* cells, was stable and enabled biochemical characterization (Echavarri-Erasun et al., 2014). Electron paramagnetic resonance (EPR) studies identified three [4Fe–4S] clusters in NifB_{Mi}; the SAM-binding [4Fe–4S] cluster and two auxiliary [4Fe–4S] clusters thought to act as substrates for NifB-co synthesis (Wilcoxon et al., 2016). Amino acid residues involved in the coordination of two of these metal clusters were identified by site-directed mutagenesis. NifB_{Mi} was found capable of FeMo-co synthesis *in vitro*, and exhibited both SAM radical chemistry and SAM demethylation reactions. Additionally, NifB proteins from the archaea *Methanosaerina acetivorans* and *Methanobacterium thermoautotrophicum* purified from recombinant *E. coli* cells were found to catalyze carbide insertion into the FeMo-co precursor (Fay et al., 2015). Importantly, none of the studied archaeal NifB proteins contained the NifX-like extension, showing its dispensability in the *in vitro* FeMo-co synthesis assays for this particular NifB subfamily.

In this work, we have compared the diversity, phylogeny, and domain architecture of 390 putative NifB proteins to understand the minimal requirements for NifB activity. We further used genetic complementation to investigate the *in vivo* functionality of NifB from a hyperthermophilic anaerobic Euryarchaeon, a thermophilic anaerobic green sulfur bacterium, and a mesophilic γ -proteobacterium, representing the three existing NifB protein architectures. Finally, NifB from *Chlorobium tepidum* was purified from a recombinant *A. vinelandii* strain and characterized *in vitro*.

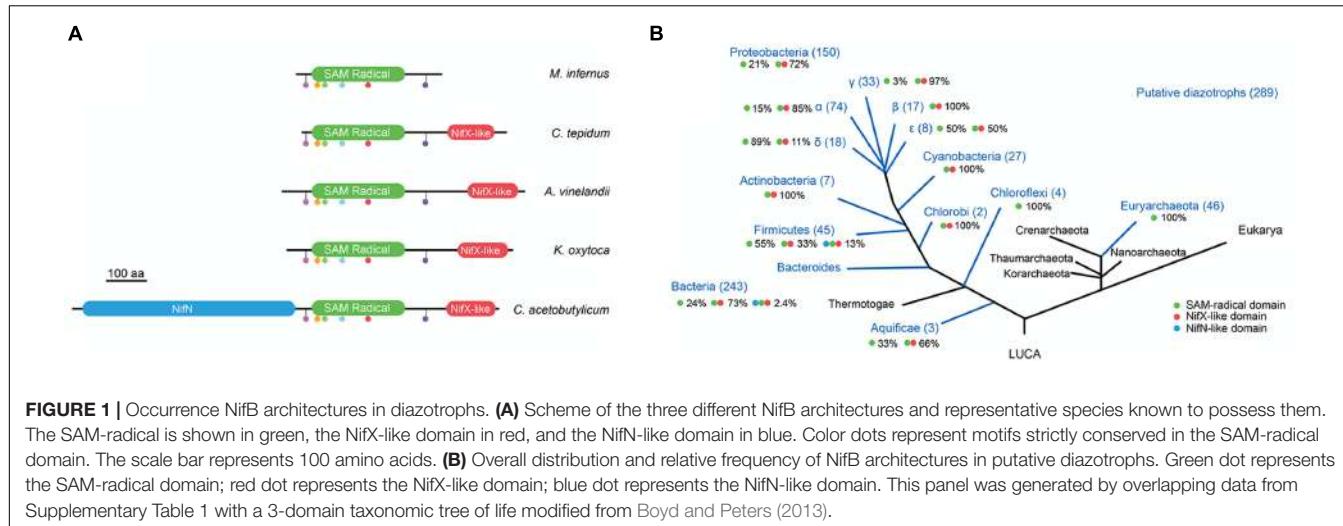
RESULTS

Generation of a Representative NifB Database

The 390 putative NifB sequences found in the Structure and Function Linkage Database (SFLD) (Akiva et al., 2014) are shown in the Supplementary Table 1. Since SFLD “relates specific sequence-structure features to specific chemical capabilities,” and is therefore not immune to faulty annotations, we identified specific NifB fingerprint motifs and applied them as filter to curate the database. By aligning experimentally proven NifB proteins from *A. vinelandii* (NifB_{Av}) (Curatti et al., 2006), *K. oxytoca* (NifB_{Ko}) (Zhao et al., 2007), *Clostridium acetobutylicum* (NifB_{Ca}) (Chen et al., 2001; Wiig et al., 2011), *M. infernus* (NifB_{Mi}) (Wilcoxon et al., 2016), *Methanosaerina acetivorans* (NifB_{Ma}) (Fay et al., 2015), *Methanobacterium thermoautotrophicum* (NifB_{Mt}) (Fay et al., 2015), and *C. tepidum* (NifB_{Ct}, this work), a number of conserved motifs were identified in the SAM-radical domain including an HPC motif, the AdoMet motif (Cx₃Cx₂C) common to all SAM-radical proteins, an ExRP motif, an AGPG motif, a TxTxN motif and a Cx₂CRxDAXG motif. Putative NifB proteins that did not present all these motifs were eliminated from the dataset, which was then reduced by 28% down to 289 sequences (Figure 1 and Supplementary Table 1).

Phylogenetic Distribution of Three Distinct NifB Domain Architectures

The most widely occurring NifB domain architecture consists of an N-terminal SAM-radical domain linked to a C-terminal NifX-like domain (Figure 1A). This protein configuration accounted for 73% of NifB sequences in the Bacteria domain of the curated database (Figure 1B and Supplementary Table 2). This configuration has been proposed to emerge after an ancestral gene fusion event (Curatti et al., 2006; Boyd et al., 2011). The functionality of this NifB subfamily has been demonstrated *in vivo* in many bacteria, and *in vitro* for NifB_{Av} and NifB_{Ko} (Curatti et al., 2006; Zhao et al., 2007). A second NifB subfamily that included an additional NifN-like domain was found in 6 NifB sequences in the Bacteria domain (corresponding to 2.4% of the curated database). This NifB subfamily was first described in Clostridia (Chen et al., 2001) and then proven functional *in vitro* using purified preparations of an *A. vinelandii* engineered NifN-B fusion that mimicked the Clostridium protein (Wiig et al., 2011). However, *in vivo* complementation of an *A. vinelandii*



ΔnifB mutant was not shown. Finally, a stand-alone SAM-radical domain was found in 104 NifB sequences, accounting for 100% of the Euryarchaeota and 24% of the Bacteria NifB proteins (Figure 1B). The functionality of this NifB subfamily has been demonstrated exclusively *in vitro* for *M. infernus* (*NifB_{MI}*) (Wilcoxen et al., 2016), *M. acetivorans* (*NifB_{MA}*) and *M. thermoautotrophicum* (*NifB_{MT}*) (Fay et al., 2015).

Importantly, the *Clostridium* genus of the Firmicutes phylum is unique in that it contains all three NifB architectures. The curated NifB database contains 45 Firmicutes species likely to be diazotrophic organisms. Among these, 55% carry the stand-alone SAM-radical domain, 33% carry the two-domain architecture, and 13% carry the three-domain architecture (Figure 1B).

NifB Phylogeny Provides Information about the Evolution of Diazotrophs

Using the curated NifB database, 28 organisms representing the diversity of phylogenetic groups having diazotrophic members (Boyd and Peters, 2013) were selected to construct a circular phylogenetic tree (Figure 2A and Supplementary Table 2) and used as a reference to further overlap NifB phylogenetic trees. In this phylogenetic tree Archaea clade together, as out-group to Bacteria, forming two different subclades: the Methanococci (*M. infernus* and *M. villosovorus*) and the Methanobacteria (*Methanobrevibacter smithii* and *Methanothermobacter thermoautotrophicus*). Bacteria diazotrophic species were distributed as follows: Aquaficaceae (*Thermocrinis albus* and *Hydrogenobaculum* sp.); Bacteroidetes (*Dysgonomonas gadei* and *Paludibacter propionigenes*), which clade with Chorobi (*Chlorobium ferrooxidans*, *Chlorobium parvum*, and *Chlorobaculum tepidum*); Actinobacteria (*Frankia alni*); Chloroflexi (*Dehalococcoides mccartyi* and *Roseiflexus castenholzii*); Cyanobacteria (*Anabaena* sp. and *Cyanothece* sp.); and Firmicutes (*Clostridium kluyveri*, *C. acetobutylicum*, and *C. pasteurianum*), all found in the same clade. Finally, α-proteobacteria (*Rhodopseudomonas palustris*, *Bradyrhizobium japonicum*, *Rhodospirillum rubrum*, and *Rhodobacter capsulatus*),

β-proteobacteria (*Azoarcus* sp.), γ-proteobacteria (*A. vinelandii*, *K. oxytoca* and *Pseudomonas stutzeri*), and δ-proteobacteria (*Arcobacter nitrofigilis*) were all in the same clade.

Because of the existence of three different NifB architectures, poorly aligned segments could potentially distort the phylogenetic tree analyses. Therefore, these regions were removed with Gblocks software (Talavera and Castresana, 2007) leaving a 315 contiguous amino acid sequence that was used to generate the SAM-radical domain tree (Figures 2B,C) and a 64 contiguous amino acid sequence used to generate two different NifX-like domain trees (Figures 2D,E).

The SAM-radical domain tree was rooted in *M. infernus* and is shown in Figure 2B. A derivative tree illustrating the distribution of NifB domain architecture is presented in Figure 2C. The Aquaficaceae, Actinobacteria, and Cyanobacteria did not clade with Firmicutes, as expected according to Figure 2A, but with Proteobacteria classes, leaving the Firmicutes as out-group to all of them. Interestingly, the γ-proteobacteria *NifB_{KO}* was found as out-group to all proteobacteria in agreement with previous analysis (Boyd et al., 2011). The Chlorobi and Bacteroidetes NifB claded as expected. However, Chloroflexi NifB rooted deeper in the Bacteria, being the closest relative to Archaea NifB. Previous studies proposed that the entire *nif* operon might have been laterally transferred to Chloroflexi from an ancestral methanogen co-existing in a common ecological niche (Eisen et al., 2002). Our data support this hypothesis.

The phylogenetic signal of the NifX-like domain of NifB was also analyzed (Figures 2D,E). No Archaea NifB with a NifX-like domain has to our knowledge been found. Chloroflexi also lack this domain, suggesting acquisition from Archaea by a lateral gene transfer event (LGT) as previously suggested (Eisen et al., 2002; Boyd et al., 2011). Distinct NifX proteins encoded in the genomes of some methanogens were then used to root the trees. Two substantially different phylogenetic trees were obtained depending on the NifX protein used as root. Since NifX was not found in any Methanococci (i.e., *M. infernus*), Figure 2D uses NifX from Methanobacteriales (*M. thermoautotrophicus*) and Figure 2E uses NifX from Methanosarcinales (*M. acetivorans*).

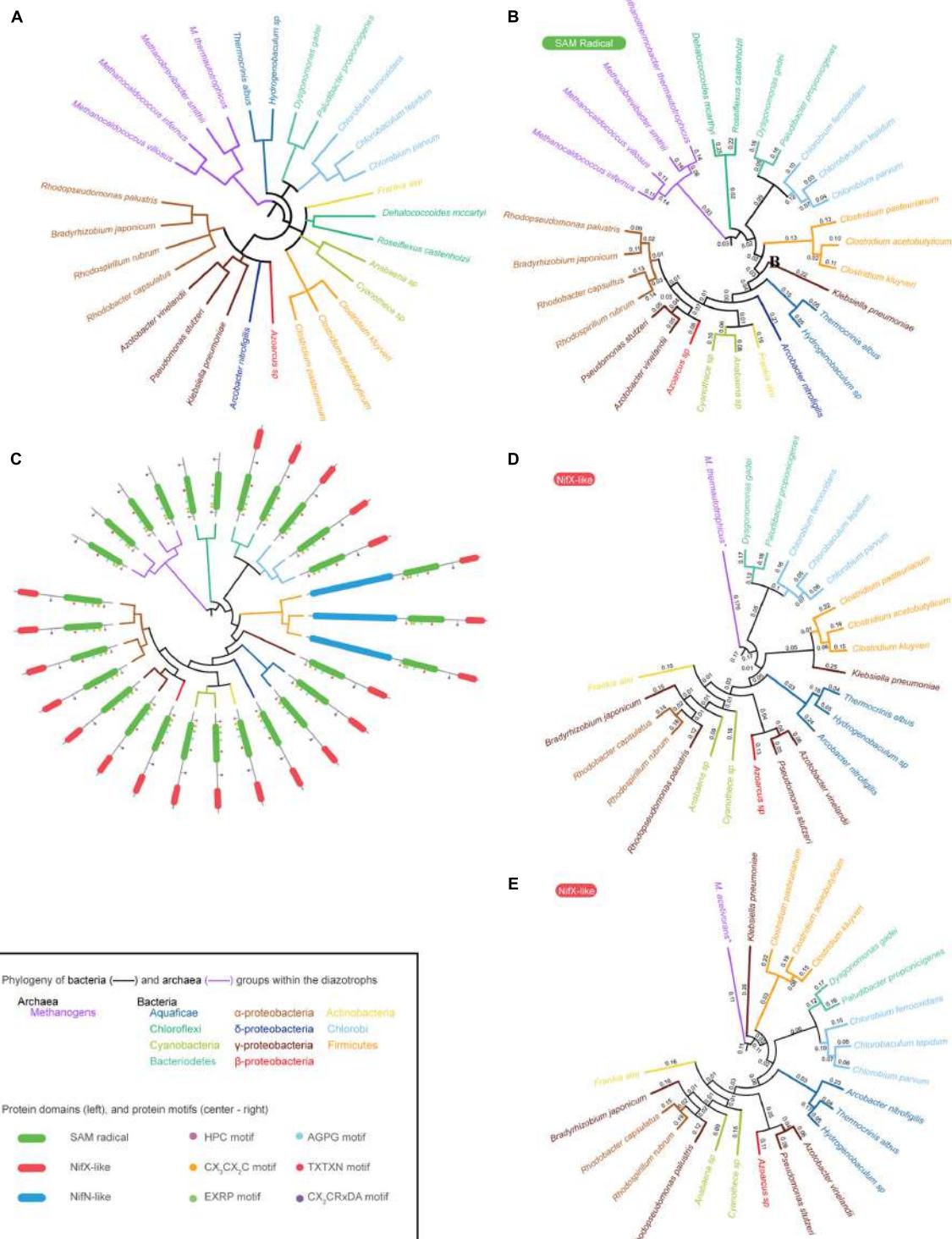


FIGURE 2 | NifB phylogeny. **(A)** Phylogenetic tree of twenty-eight selected species representing all Bacteria and Archaea groups reported to carry *nif* genes.

(B) Phylogenetic tree of twenty-eight NifB proteins based on their SAM-radical domain. **(C)** Overlay of the phylogenetic tree shown in **(B)** and existing NifB architectures. **(D)** Phylogenetic tree of twenty-two NifB proteins based on their NifX-like domain generated using *M. thermautrophicus* NifX as root.

(E) Phylogenetic tree of twenty-two NifB proteins based on their NifX-like domain generated using *M. acetivorans* NifX as root. The inset provides color code for the different Archaea and Bacteria groups with diazotroph members shown in the phylogenetic trees. It also details NifB protein domains and the strictly conserved motifs within the SAM-radical domain **(C)**.

The pattern of the first tree is similar to that of the SAM-radical domain tree, suggesting Chloroflexi as the bacterial ancestor from which the lineage emerged. The second tree, however, points to Firmicutes as the bacterial ancestor from which *nif* genes proliferated in Bacteria. This remains an interesting possibility given that Firmicutes present the three different NifB architectures known to date.

Organization of *nif* Genes in the Genomes of *C. tepidum* and *M. infernus*

In order to define essential and not essential domains for NifB function *in vivo* in an aerobic mesophilic host, we focused on NifB from the thermophile *C. tepidum* (two-domain architecture), the hyperthermophile *M. infernus* (single-domain architecture), and the mesophile *K. oxytoca* (two-domain architecture). NifB_{Ko} and NifB_{Mi} have previously been purified and characterized *in vitro* (Zhao et al., 2007; Wilcoxen et al., 2016) but not NifB_{Ct}, which is reported in this study.

Chlorobium tepidum is a well-described diazotroph (Wahlund and Madigan, 1993) with annotated genome (Eisen et al., 2002). Most of its *nif* genes are located in a single 20-kb cluster containing the Mo-nitrogenase structural genes (*nifH*, *nifD*, and *nifK*), FeMo-co biosynthetic genes (*nifB*, *nifE*, *nifN*, *nifV*, and *fdxN*), and regulatory genes (*nifA*, *nifI₁*, and *nifI₂*) (Supplementary Figure 2). Genome blast with individual *nif* genes from the model diazotroph *K. oxytoca* did not reveal anomalies, supporting the current *C. tepidum* annotation.

While NifB_{Mi} expressed in *E. coli* was shown to support FeMo-co synthesis *in vitro* (Wilcoxen et al., 2016), *M. infernus* has not yet been proven to be diazotrophic. The *nif* genes in the *M. infernus* genome consist of *nifH*, *nifD*, and *nifK* structural genes, *nifB* and *nifE* cofactor biosynthetic genes, and *nifI₁* and *nifI₂* regulatory genes. Intriguingly, a second *nifH* gene is located 17-kb apart from the *nif* cluster and *nifB* was found 470-kb apart with no apparent *nif* genes in close proximity.

NifB_{Ct} and NifB_{Mi} are Functional *in Vivo* When Expressed in the Aerobic Mesophilic Host *A. vinelandii*

Genetic complementation analyses were performed by expressing synthetic codon-optimized *nifB_{Ct}* and *nifB_{Mi}* genes in the *A. vinelandii* UW140 ($\Delta nifB$) strain under the control of the *nifH* promoter (Figure 3A). *A. vinelandii* is a strict aerobe with optimum growth temperature of 30°C and is used here to provide an initial screen of NifB functionality that will be useful for further screening and implementation in Eukaryotic hosts. Strains UW418 ($\Delta nifB$, *PnifH::nifB_{Mi}*) and UW422 ($\Delta nifB$, *PnifH::nifB_{Ct}*) exhibited diazotrophic growth both in solid and liquid culture media (Figures 3B,D), in contrast to the Nif⁻ phenotype of the parental strain UW140 ($\Delta nifB$). Calculated diazotrophic growth rates ($\ln 2/t_d$) were: 0.23 for the wild type, <0.001 for UW140, 0.015 for UW418, and 0.13 for UW422. This data shows that, although both NifB_{Ct} and NifB_{Mi} originate from strict anaerobic and thermophilic microbes, the proteins were functional and could complement the *A. vinelandii* $\Delta nifB$ mutant phenotype. However, whereas NifB_{Ct} supported similar growth

rate at 30°C as the *A. vinelandii* wild type strain, the recombinant NifB_{Mi} did not, possibly explained by the almost 40°C difference in optimal growth temperature between *C. tepidum* (48°C, Wahlund and Madigan, 1993) and *M. infernus* (85°C, Jeannot et al., 1998). No difference in growth rate could be observed when using NH₄⁺ as nitrogen source: 0.31 for the wild type, 0.29 for UW140, 0.30 for UW418, and 0.30 for UW422 (Figure 3C).

In vivo nitrogenase activities determined by the acetylene reduction assay showed significant activity in UW422 in the 8 h period following nitrogenase derepression (Figure 3E). No activity was detected in UW418 within this period of time, consistent with its significantly lower diazotrophic growth rate.

Purification and Biochemical Characterization of NifB_{Ct}

NifB_{Ct} was expressed and purified from a recombinant *A. vinelandii* strain (Figure 4A). The yield of pure NifB_{Ct} from *A. vinelandii* cells was 0.3 µg NifB_{Ct} per gram of cell, 15-fold higher than that of overexpressed NifB_{Av} (Curatti et al., 2006). Purity of the NifB_{Ct} preparations exceeded 95%, as determined by Coomassie stained SDS-gels, and the identity of NifB_{Ct} was confirmed by MALDI-TOF analysis with 60% sequence coverage (Supplementary Table 3). NifB_{Ct} migrated as a monomer of 46.5 kDa in anaerobic size exclusion chromatography (Figure 4B), in good agreement with theoretical mass determined by the amino acid sequence (46.8 kDa). As isolated NifB_{Ct} contained 3.05 Fe atoms per monomer. *In vitro* reconstitution of its [Fe-S] clusters under reducing conditions increased Fe contents to 10.1 ± 0.07 Fe atoms ($n = 3$). Consistently, features characteristic of [Fe-S] proteins (especially the broad shoulder at 400–420 nm) were more prominent in the reconstituted NifB_{Ct} UV-vis spectrum (Figure 4C). Reconstituted NifB_{Ct} was active in the *in vitro* FeMo-co synthesis and nitrogenase activation assay: 5.2 ± 2.2 nmol ethylene formed·min⁻¹·assay⁻¹ ($n = 2$) compared to 8.2 ± 1.5 nmol ethylene formed·min⁻¹·assay⁻¹ ($n = 2$) when using pure NifB-co.

The NifX-like Domain of NifB_{Ko} Is Not Essential for Nitrogenase Activity or Diazotrophic Growth

The capacity of NifB_{Mi} to complement the $\Delta nifB$ strain strongly suggests that the SAM-radical domain of NifB is the only one required for the synthesis of the FeMo-co precursor, but this could be a property specific to the stand-alone SAM-radical domain subfamily. To determine whether the NifX-like domain naturally present in the two-domain NifB architecture is required for NifB-co synthesis, a truncated NifB_{Ko} variant lacking the entire NifX-like domain (*nifB_{Ko}-ΔC*) was generated, introduced in *K. oxytoca* UC9 ($\Delta nifB$) and expressed under the control of a *tac* promoter (Figure 5A and Supplementary Figure 3). Additionally, as this truncated version would mimic a mesophilic single-domain NifB, we could test whether presence of the NifX-like domain is important for growth under moderate, non-thermophilic, temperatures. Diazotrophic growth and *in vivo* nitrogenase activity of UC28 ($\Delta nifB$, *Ptac::nifB_{Ko}-ΔC*) were measured at 3 h intervals in a

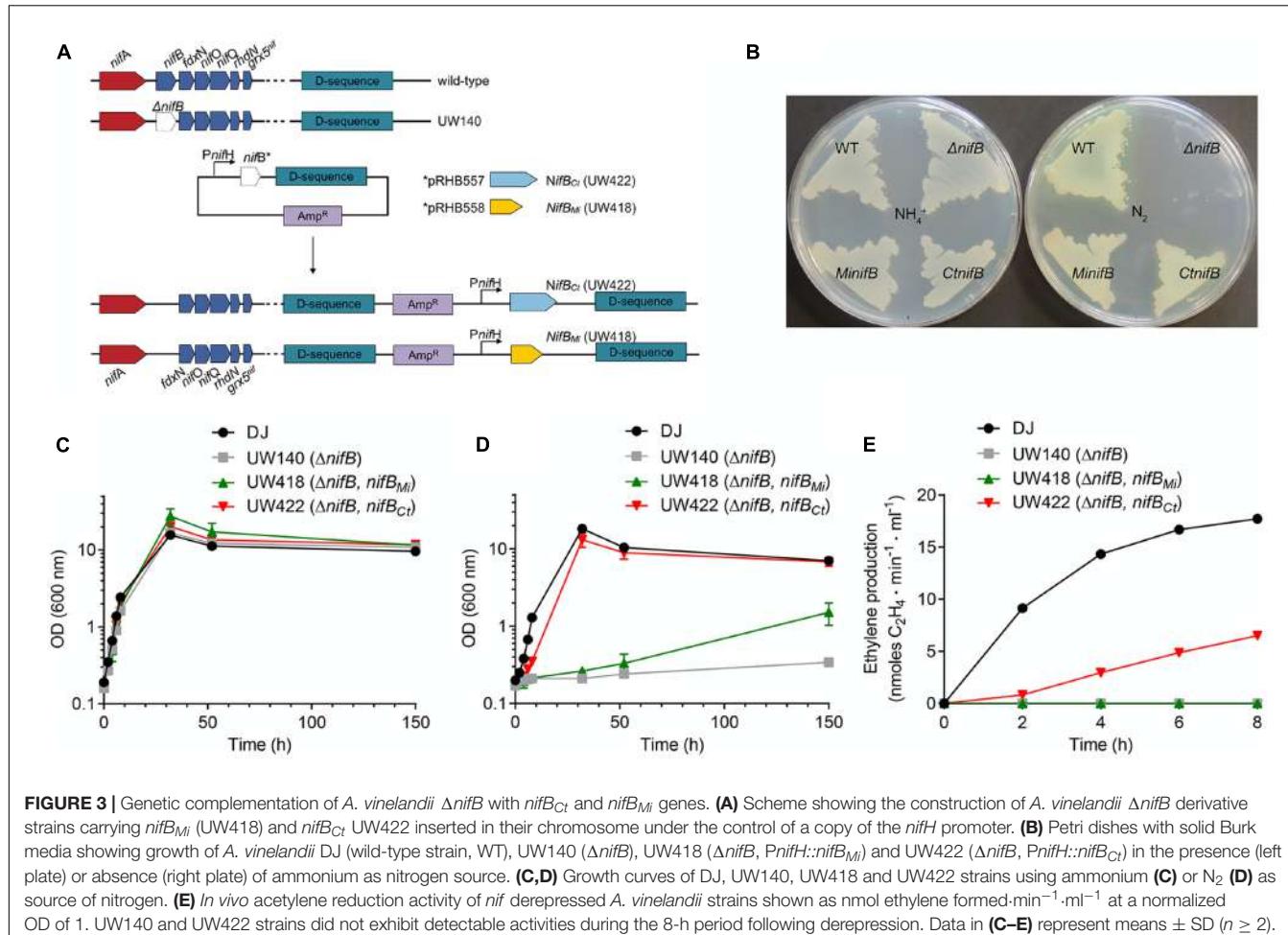


FIGURE 3 | Genetic complementation of *A. vinelandii* $\Delta nifB$ with *nifB_{Ct}* and *nifB_{Mi}* genes. **(A)** Scheme showing the construction of *A. vinelandii* $\Delta nifB$ derivative strains carrying *nifB_{Mi}* (UW418) and *nifB_{Ct}* (UW422) inserted in their chromosome under the control of a copy of the *nifH* promoter. **(B)** Petri dishes with solid Burk media showing growth of *A. vinelandii* DJ (wild-type strain, WT), UW140 ($\Delta nifB$), UW418 ($\Delta nifB, PnifH::nifB_{Mi}$) and UW422 ($\Delta nifB, PnifH::nifB_{Ct}$) in the presence (left plate) or absence (right plate) of ammonium as nitrogen source. **(C,D)** Growth curves of DJ, UW140, UW418 and UW422 strains using ammonium **(C)** or N_2 **(D)** as source of nitrogen. **(E)** In vivo acetylene reduction activity of *nif* derepressed *A. vinelandii* strains shown as nmol ethylene formed $\cdot min^{-1} \cdot ml^{-1}$ at a normalized OD of 1. UW140 and UW422 strains did not exhibit detectable activities during the 8-h period following derepression. Data in **(C–E)** represent means \pm SD ($n \geq 2$).

24 h time course following derepression and compared to those of UC16 ($\Delta nifB, Ptac::nifB_{Ko}$), a control strain expressing full-length NifB_{Ko}. Surprisingly, UC28 exhibited diazotrophic growth similar to UC16 and *in vivo* nitrogenase activity higher than UC16 (Figures 5B,C). The UC9 parental strain did not exhibit nitrogenase activity or diazotrophic growth, confirming that the functionality of the expressed NifB_{Ko} variants and suggesting that the NifX-like extension of NifB_{Ko} is not required for NifB-co synthesis, at least under the growth conditions tested in this study, and that this could be a general rule for the two-domain family of NifB proteins.

DISCUSSION

NifB Phylogeny and Architecture

To our knowledge, this work presents the largest compilation of NifB proteins described to date. The NifB database was stringently filtered to exclude faulty annotated proteins and the curated dataset provides insights about NifB origin, taxonomy and architecture that complement previous work (Soboh et al., 2010; Boyd et al., 2011; Boyd and Peters, 2013). In this study we demonstrate that the SAM-radical domain of NifB is sufficient

to support FeMo-co biosynthesis *in vivo* in the model organisms *A. vinelandii* and *K. oxytoca*.

A strict filter, based on motifs exhibited by experimentally confirmed NifB proteins, was applied to the initial database. As a result, 28% NifB sequences were excluded from further analysis. Although these criteria might be too strict, we reasoned that it was better to miss some true-positives than to risk including false-positives. Most excluded NifB proteins lacked the conserved Cx₃Cx₂C motif required for SAM-radical catalysis. In contrast, the NifX domain was identified in each one of them and we think that these faulty annotated NifB proteins are instead NifX. This confusion originates from the fact that the NifX domain is present in NifB, NafY, NifY as well as NifX proteins.

Three distinct NifB protein architectures exist. The most widespread in Bacteria consists of an N-terminal SAM-radical domain followed by a C-terminal NifX-like domain. However, this configuration is absent in Archaea, which present smaller NifB proteins consisting of a stand-alone SAM-radical domain. Boyd and collaborators investigated the lineage of the stand-alone SAM-radical domain in Archaea NifB proteins and compared it to the two-domain architecture favored in Bacteria (Boyd et al., 2011). The authors traced an event that suggested that a methanogen donated its *nif* cluster via LGT to a

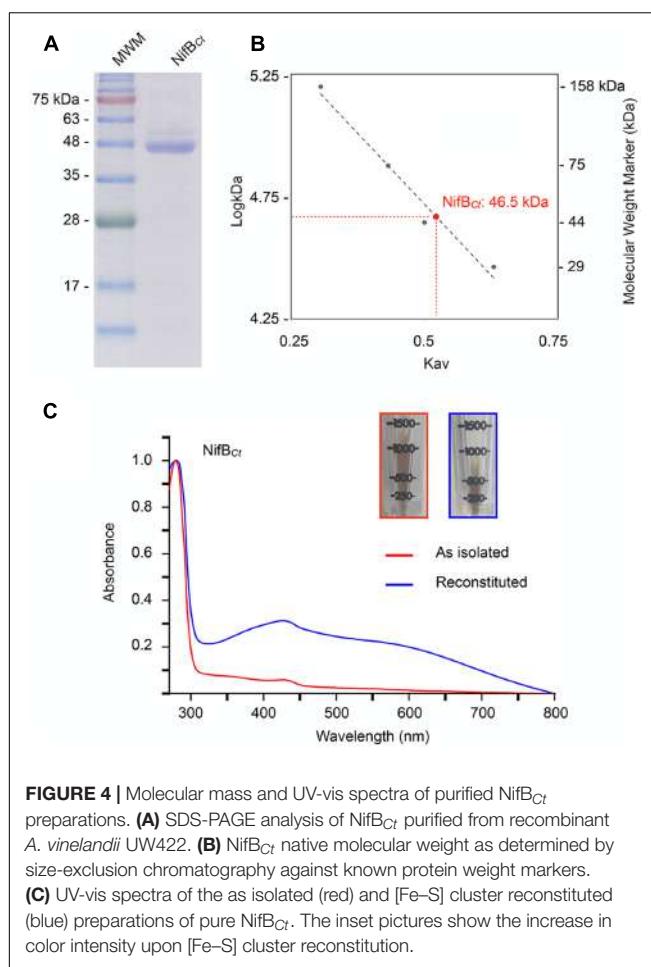


FIGURE 4 | Molecular mass and UV-vis spectra of purified NifB_{Ct} preparations. **(A)** SDS-PAGE analysis of NifB_{Ct} purified from recombinant *A. vinelandii* UW422. **(B)** NifB_{Ct} native molecular weight as determined by size-exclusion chromatography against known protein weight markers. **(C)** UV-vis spectra of the as isolated (red) and [Fe–S] cluster reconstituted (blue) preparations of pure NifB_{Ct}. The inset pictures show the increase in color intensity upon [Fe–S] cluster reconstitution.

Firmicutes ancestor that co-existed in the same ecological niche. Then, a fusion event happened that resulted in the *nifB-nifX* protein occurring in Firmicutes. It was later suggested that the wide spread of the *nifB-nifX* fusion protein in Bacteria was independent of the selective pressure associated with aerobic diazotrophy (Boyd et al., 2015). An additional fusion event between *nifN* and *nifB-nifX* also occurred in Firmicutes leading to the three-domain NifB architecture. This last event was confined to Firmicutes, which is the only phylum presenting all three types of NifB architecture. It is surprising that the three-domain NifB was not widespread in Bacteria. From knowledge gained through *in vitro* FeMo-co synthesis studies (Curatti et al., 2007), it could be assumed that a NifENB fusion protein would be beneficial by protecting labile NifB-co and streamlining FeMo-co synthesis. However, it is possible that a NifENB fusion might not allow fine-tuning of precursor biosynthesis.

Based on the phylogeny of independent NifX proteins, another early *nifB* LGT was detected between Methanomicrobiales and Chloroflexi. This event was also apparent in the SAM-radical domain phylogenetic tree, with Chloroflexi rooting deeper than any other group. The short distance between Methanomicrobiales and Chloroflexi NifB lineages was also observed by Boyd and colleagues (Boyd et al., 2011).

Ancestral NifB Proteins from Strict Anaerobic and Thermophilic Organisms that Function *in Vivo* in an Aerobic Mesophilic Host

Stand-alone SAM-radical domain NifB proteins catalyze NifB-co synthesis *in vitro* (Fay et al., 2015; Wilcoxen et al., 2016). However, they have not yet been proven capable of sustaining diazotrophic growth of *M. thermautotrophicus*, *M. acetivorans*, and *M. infernus* (which also are not yet experimentally confirmed to be diazotrophs). It was also not clear whether this NifB family would function in a mesophilic and aerobic environment, which could prevent their use for plant nitrogenase engineering. Therefore, the Nif⁺ phenotype exhibited by the *A. vinelandii* $\Delta nifB$ strain complemented with *nifB_Mi* presented in this study is convincing evidence of its *in vivo* functionality in a mesophilic and aerobic bacterium.

As expected, stronger Nif⁺ phenotype was achieved by complementation with NifB_{Ct}. *C. tepidum* is a mild thermophile with optimum growth temperature of 48°C and therefore much closer to the 30°C optimum of *A. vinelandii*. In addition, NifB_{Ct} has a two-domain NifB architecture similar to NifB_{Av}. Interestingly, NifB_{Ct} was a monomer, similar to the archaeal single-domain NifB proteins and different from the NifB_{Av} and NifB_{Ko} homodimers. Although constrained by the limited set of available experimental data, it appears that NifB monomers might be more stable and therefore favored in thermophilic organisms regardless of protein architecture. Importantly, both configurations are functional *in vivo* in a mesophilic host. The strong diazotrophic growth of UW418 in plates compared to liquid medium suggests that there are other factors limiting NifB_{Mi} activity *A. vinelandii* in addition to operational temperature. One possibility is that oxygen limitation during growth in plate has a positive effect on NifB_{Mi} that is not observed in liquid medium.

The NifX-like Domain of NifB_{Ko} May Have a Role Regulating the Flux of NifB-co during FeMo-Co Biosynthesis

It was suggested that the distinct NifB_{Av} domain architecture (the N-terminal SAM-radical domain and the C-terminal NifX-like domain) could be required to coordinate [Fe–S] cluster precursors prior to catalysis resulting in NifB-co synthesis (Curatti et al., 2006). This possibility was put into question when stand-alone SAM-radical domain archaeal NifB were found active *in vitro* (Arragain et al., 2014; Fay et al., 2015; Wilcoxen et al., 2016). Here, we demonstrate that the NifX-like domain of NifB_{Ko} is not essential for catalytic activity *in vivo*. A truncated NifB_{Ko} lacking the NifX-like domain supported *in vivo* nitrogenase (ethylene production) rates even higher than full-length NifB. It is thus reasonable to think that NifB catalysis only requires the SAM-radical domain, and that other domains may perform complementary functions that are beneficial but not essential for FeMo-co biosynthesis. A critical role in cofactor biosynthesis for alternative nitrogenases is not likely as this

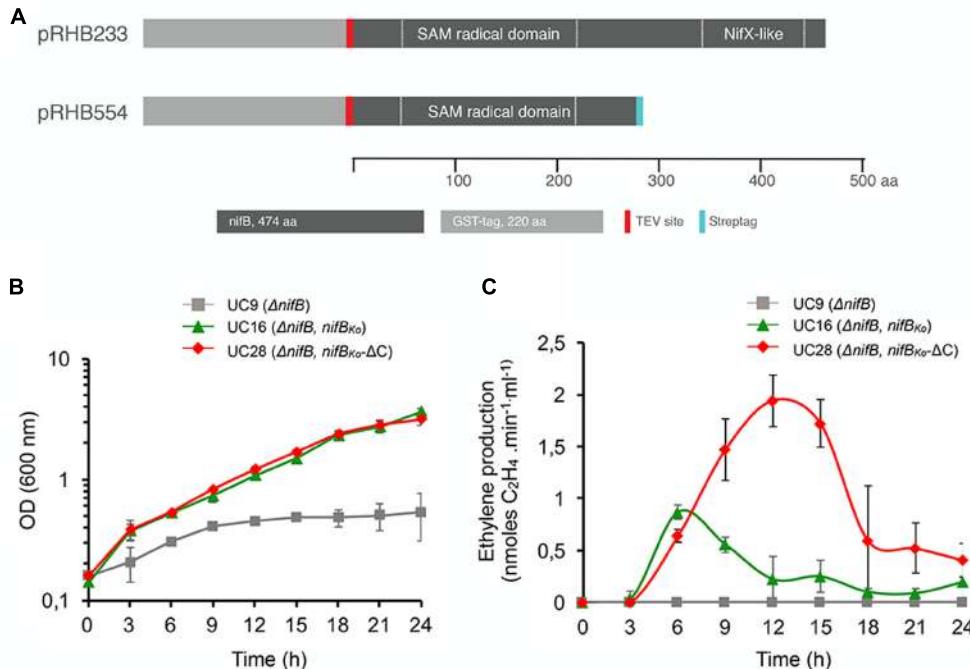


FIGURE 5 | (A) Scheme of full-length GST-NifB_{Ko} and GST-NifB_{Ko}-ΔC in expression plasmids carried by UC16 and UC28, respectively. **(B,C)** Diazotrophic growth (B) and *in vivo* acetylene reduction activity (C) of *K. oxytoca* UC16 and UC28 expressing full-length and truncated NifB variants, respectively, compared to the UC9 ($\Delta nifB$) strain. Data represent means \pm SD ($n = 3$). Maximum ethylene production of *K. oxytoca* UN (wild-type) was 12.6 nmol·min $^{-1}$ ·ml $^{-1}$ ($n = 3$).

domain is absent in NifB from *M. acetivorans*, which carries all three types of nitrogenase (Galagan et al., 2002).

Prospects to Implement NifB Activity in Eukaryotes

The successful purification of active NifH from yeast mitochondria, when co-expressed with NifU, NifS and NifM, represented a first advance toward implementing BNF in eukaryotic systems (Lopez-Torrejon et al., 2016). However, major steps are still required to engineer active nitrogenase in a eukaryote. In this regard, expression of functional NifB is expected to be a major barrier to overcome. This is not only because NifB catalyzes a reaction unique and essential to diazotrophs, but also because of the O₂-lability of its [Fe-S] clusters, including NifB-co.

NifB from well-established model organisms, such as *A. vinelandii* and *K. oxytoca*, might be difficult to use in the harsh environment provided by a eukaryotic cell. There is evidence that NifB catalysis makes it susceptible to proteolysis (Martinez-Noel et al., 2011). Screening for simpler, but more suitable variants from less “sophisticated” diazotrophs may be a rewarding strategy. In this aspect, the use of less labile, monomeric, and temperature-resistant NifB from Archaea or Bacteria, such as the two examples shown in this study, may help engineering FeMo-co biosynthesis in Eukaryotic (plant) cells. The accompanying paper (Burén et al., 2017) describes the first successful step in this direction.

MATERIALS AND METHODS

Data Mining and Phylogenetic Analysis

The 390 annotated NifB sequences retrieved from the Structure and Function Linkage Database (SFLD) (Akiva et al., 2014) and UniProt¹ are shown in Supplementary Table 1. To exclude potentially faulty annotated sequences, the following filtering procedure was applied to the dataset. First, amino acid sequences of experimentally proven NifB proteins, including *A. vinelandii* (NifB_{Av}) (Curatti et al., 2006), *K. oxytoca* (NifB_{Ko}) (Zhao et al., 2007), *Clostridium pasteurianum* (NifB_{Cp}) (Chen et al., 2001; Wiig et al., 2011), *M. infernus* (NifB_{Mi}) (Wilcoxen et al., 2016), *Methanosarcina acetivorans* (NifB_{Ma}) (Fay et al., 2015), *Methanobacterium thermoautotrophicum* (NifB_{Mt}) (Fay et al., 2015), and *C. tepidum* (this work) were aligned to determine conserved motifs. These NifB fingerprint motifs localized in the SAM-radical domain and included an HPC motif, the AdoMet Cx₃Cx₂C motif, an ExRP motif, an AGPG motif, a TxTxN motif, and a Cx₂CRxDAxG motif (Figure 1). The full NifB dataset was then analyzed for the presence of these fingerprints, reducing the initial 390 sequences to 289 (Supplementary Table 1). Protein domain architecture was analyzed using the PFAM database² (Finn et al., 2016). The frequency of appearance of each one of the different NifB domains in diazotrophic phyla shown in Figure 1 was represented by overlapping data from Supplementary Table 1

¹<http://uniprot.org>

²<http://pfam.xfam.org>

with a 3-domain taxonomic tree of life (modified from Boyd and Peters, 2013).

Twenty-eight NifB proteins representing all phylogenetic groups known to contain diazotrophs (Boyd and Peters, 2013) (Supplementary Table 2) were selected from the reduced list and used to investigate taxonomy versus architecture correlation. The taxonomy of diazotrophic groups was resolved using PhyloT³, an online tool that uses the full NCBI taxonomy to generate phylogenetic trees (**Figure 2A**).

Clustal Omega⁴ was used to generate protein alignments and neighbor joining (NJ) phylogenetic trees (Sievers et al., 2011). Maximum likelihood (ML) trees shown in **Figures 2B–E** were produced using the IQ-Tree web server⁵ (Trifinopoulos et al., 2016). Gblocks (Talavera and Castresana, 2007) was used to remove non-conserved aligned segments leaving a 315 contiguous amino acid sequence that was used to generate the SAM-radical domain tree (**Figures 2B,C**) and a 64 contiguous amino acid sequence used to generate the NifX-like domain trees (**Figures 2D,E**). Phylogenetic trees shown in **Figures 2B–E** were resolved using the Interactive Tree of Life online tool⁶ (Letunic and Bork, 2007) and FigTree.

Plasmids, Strains and Growth Conditions

The strains and plasmids used in this work are listed in Supplementary Table 4. *A. vinelandii* strains DJ (wild-type) (D.R. Dean, Virginia Tech) and UW140 ($\Delta nifB$) (Hernandez et al., 2007) have been described. *K. oxytoca* strains UC9 ($\Delta nifB$) and UC16 ($\Delta nifB$, Ptac::gst-*nifB_{Ko}*) (Zhao et al., 2007) have been described.

The *M. infernus* (*nifB_{Mi}*, accession number D5VRM1) and *C. tepidum* (*nifB_{Ct}*, accession number CT1540) *nifB* sequences were codon-optimized and synthesized by GenScript (Piscataway, NJ, United States) for expression in *E. coli*. Plasmids pRHB557 and pRHB558 contained the *nifB_{Ct}* and *nifB_{Mi}* genes, respectively, cloned into the *NdeI* and *EcoRI* sites of pRHB258 for the expression of His₉-tagged proteins under the control of the *nifH* promoter (Curatti et al., 2007). Plasmids pRHB557 and pRHB558 were inserted into the chromosome of *A. vinelandii* UW140 ($\Delta nifB$) by homologous recombination at the D-sequence, a 1.1-kb DNA fragment from the chromosomal region downstream of Avin02530 (Hernandez et al., 2008), to generate strains UW422 and UW418, respectively (**Figure 3A**). Transformants were selected in agar plates of NH₄⁺-free Burk's modified medium (Shah et al., 1972) containing 50 µg/ml ampicillin.

For diazotrophic growth rate *A. vinelandii* strains were grown at 30°C on N-free Burk's medium. When a fixed nitrogen source was required, ammonium acetate was added to a final concentration of 29 mM. Growth was estimated as OD₆₀₀ using an Ultrospec 3300 Pro spectrophotometer (Amersham). The exponential growth rate constant corresponds to ln2/td, where td represents the doubling time.

³<http://phylot.biobyte.de/>

⁴<http://www.ebi.ac.uk/Tools/msa/clustalo/>

⁵<http://www.iqtree.org>

⁶<http://itol.embl.de>

For *A. vinelandii* *in vivo* nitrogenase activity determinations strains were grown at 30°C on NH₄⁺ supplemented Burk's medium and then collected, washed and derepressed for nitrogenase as previously described (Shah et al., 1972). Acetylene reduction was determined as described in (Stewart et al., 1967).

Expression plasmid pRHB233 (Ptac::gst-*nifB_{Ko}*) is a derivative of pGEX-4T-3 (GE Healthcare) that contains the entire *nifB_{Ko}* gene (1404 nucleotides encoding a 468 amino-acid polypeptide; UniProt accession number P10390) fused to a *gst*-encoding gene (Zhao et al., 2007). Plasmid pRHB233 was used as template to amplify a truncated *nifB_{Ko}* variant using oligonucleotides 5'-C CCCATATGACTACTCCTGCTCCTCTTTCTGGCGGC-3' and 5'-GGGCTCGAGTCATGATGATGATGATGATGATGATGATGATGATGCGCGGGTCGCAATGCTGGCGTGCAG-3'. The resulting 1008 bp fragment, encoding a 336 amino acid NifB_{Ko} polypeptide that lacked the C-terminal NifX-like domain (NifB_{Ko}-ΔC), was cloned into the *NdeI* and *XhoI* sites of pGEX-4T-3 to generate plasmid pRHB554. *K. oxytoca* UC9 ($\Delta nifB$) strain was transformed with pRHB554 to generate strain UC28 ($\Delta nifB$, Ptac::gst-*nifB_{Ko}*-ΔC). Positive transformants were selected in LC agar plates containing ampicillin (150 µg/ml) and carbenicillin (800 µg/ml).

For diazotrophic growth rate and *in vivo* nitrogenase activity determinations, *K. oxytoca* strains were grown overnight at 30°C in minimal medium supplemented with 28.5 µM ammonium acetate (Shah et al., 1994). Cells were washed three times using N-free medium and finally resuspended at a final OD₆₀₀ value of 0.15 in N-free medium supplemented with 0.1% serine, 150 µg/ml ampicillin, 800 µg/ml carbenicillin, and 5 µM IPTG in dual-sealed 100-ml vials under O₂-free conditions. At 3-h intervals during a period of 24 h, culture growth was monitored by OD₆₀₀ using an Ultrospec 3300 Pro spectrophotometer (Amersham), and the *in vivo* nitrogenase activity was determined by ethylene production at 30°C for 30 min in 1-ml culture samples at a normalized OD₆₀₀ value of 1, as previously described (Stewart et al., 1967). The growth rate constant corresponds to ln2/td, where td represents the doubling time.

Purification of NifB_{Ct} from *A. vinelandii* Recombinant Cells

Azotobacter vinelandii UW422 cells overexpressing NifB_{Ct} under the control of a *nifH* promoter were grown in 32-l batches in a 300-l fermenter (Bioprocess Technology). Nitrogenase derepression and cell collection were carried out as described in (Echavarri-Erasun et al., 2014).

Purification of His-NifB_{Ct} from *A. vinelandii* cells was as follows: 150 g of cells were resuspended in 450 ml buffer A (50 µM Na₂HPO₄, pH 7.6, 4 M glycerol, 5 µM 2-mercaptoethanol and 2 µM Na₂S₂O₄) supplemented with protease inhibitors (200 µM PMSF and 1 µg/ml leupeptin) and 5 µg/ml DNase inside a Coy Labs glovebox for 30 min. Cells were pelleted at 14,000 × g for 10 min at 4°C and then transferred back inside the glovebox. Pellets were lysed by osmotic shock in 450 ml buffer B (50 µM Na₂HPO₄, pH 7.6, 0.05% *n*-dodecyl-β-D-maltoside, 5 µM 2-mercaptoethanol and 2 µM Na₂S₂O₄). A cell-free extract was obtained by collecting the supernatant after centrifugation at 70,000 × g for 1 h at

4°C. The cell-free extract was supplemented with NaCl to a final concentration of 180 μM and loaded onto a 25-ml IMAC column (GE Healthcare) previously charged with Co²⁺ and equilibrated in buffer C (50 μM Na₂HPO₄, pH 7.6, 180 μM NaCl, 0.05% n-dodecyl-β-D-maltoside, 5 μM 2-mercaptoethanol, 10% glycerol and 2 μM Na₂S₂O₄) at 4°C. Column was washed with 3 column volumes of buffer C, followed by 7 column volumes of buffer C supplemented with 50 μM imidazole. NifB_{Ct} was eluted using buffer C supplemented with 300 μM imidazole. Eluted fractions were analyzed by SDS-PAGE and Coomassie staining. Fractions containing pure NifB_{Ct} were pooled and desalting using a HiPrep 26/10 desalting column (GE Healthcare) previously equilibrated with buffer C. Purified NifB_{Ct} was stored in liquid N₂ as pellets.

Determination of NifB_{Ct} Native Molecular Weight

NifB_{Ct} Native Molecular Weight was determined by size-exclusion chromatography using a HiLoad 16/600 Superdex 200 column attached to an AKTA FPLC (GE Healthcare). The column was equilibrated with 50 μM Na₂HPO₄, pH 7.6, 180 μM NaCl, 10% glycerol, 5 μM 2-mercaptoethanol and 2 μM Na₂S₂O₄ and the chromatography was run with the same buffer at a flow rate of 1 ml/min. The column was calibrated for molecular mass determination by using the molecular weight standard proteins aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (44 kDa), and carbonic anhydrase (29 kDa) (GE Healthcare).

NifB_{Ct} [Fe–S] Cluster Reconstitution

As isolated NifB_{Ct} samples were diluted in 50 μM Tris-HCl (pH 8) buffer containing 200 mM KCl and 10% glycerol to a final concentration of 10 μM NifB_{Ct}. Samples were then incubated during 2 h at 37°C with a 12-fold molar excess of Fe²⁺ [(NH₄)₂Fe(SO₄)₂] and S²⁻ (Na₂S), in the presence of 10 μM DTT. The Fe and S excess was removed from reconstituted preparations by filtration in a HiPrep 26/10 desalting column (GE Healthcare) equilibrated in dilution buffer. After desalting, Fe content of reconstituted NifB_{Ct} samples was quantified as described by Fish (1988).

NifB_{Ct}-Dependent *In Vitro* Synthesis of FeMo-co

Azotobacter vinelandii UW140 (Δ nifB) cell-free extracts were obtained as described above and used for biochemical

REFERENCES

- Akiva, E., Brown, S., Almonacid, D. E., Barber, A. E. II, Custer, A. F., Hicks, M. A., et al. (2014). The structure-function linkage database. *Nucleic Acids Res.* 42, D521–D530. doi: 10.1093/nar/gkt1130
- Allen, R. M., Chatterjee, R., Ludden, P. W., and Shah, V. K. (1995). Incorporation of iron and sulfur from NifB cofactor into the iron-molybdenum cofactor of dinitrogenase. *J. Biol. Chem.* 270, 26890–26896. doi: 10.1074/jbc.270.45.26890
- Arragain, S., Scandurra, A. A., Jimenez-Vicente, E., Echavarri-Erasun, C., and Rubio, L. M. (2014). “*Methanocaldococcus infernus*” NifB: a new model to

complementation assays. Purified NifB_{Ct} (0.16 μM) was added to reaction mixtures containing 0.2 ml of UW140 cell-free extract (4.4 μg protein/ml) in 22 μM Tris-HCl (pH 7.4), 17.5 μM Na₂MoO₄, 175 μM R-homocitrate, 400 μM (NH₄)₂FeSO₄, 400 μM Na₂S, 880 μM SAM, 1.32 μM ATP, 18 μM phosphocreatine, 2.2 μM MgCl₂, 3 μM Na₂S₂O₄, 3.5% glycerol, 40 μg creatine phosphokinase, and 2 μM NifH at a final volume of 400 ml. Control reactions contained 1.4 μM pure NifB-co instead of NifB_{Ct}. Reactions were incubated at 30°C for 45 min inside an MBraun glovebox (O₂ < 0.1 ppm) to allow for FeMo-co synthesis and insertion into apo-NifDK present in the UW140 extract. Acetylene reduction activity of reconstituted NifDK protein was quantified after addition of 0.1 μg NifH, changing the vial gas phase to 100% argon, and finally injecting 0.5 ml acetylene. Reaction mixtures were incubated in a water bath at 30°C for 15 min and 600 rpm shaking and then stopped by addition of 0.1 ml of 8 M NaOH. Ethylene formation was measured in a Shimadzu GC-2014 gas chromatograph equipped with a Porapak N80/100 column.

AUTHOR CONTRIBUTIONS

CE-E, SA, EJ-V, and AS carried out experimental work; CE-E, SA, EJ-V, SB, and LR carried out experimental design and data analysis; CE-E, SB, and LR wrote the paper.

FUNDING

Funding for this research was provided by Bill & Melinda Gates Foundation OPP1143172, ERC Starting Grant 205442, and MINECO BIO2014-59131-R. AS was recipient of FPI Fellowship BES-2010-038322.

ACKNOWLEDGMENT

We thank Jose María Buesa for *A. vinelandii* fermentations.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2017.01947/full#supplementary-material>

- study NifB-co formation,” in *Proceedings of the XI European Nitrogen Fixation Conference*, Tenerife.
- Bishop, P. E., and Joerger, R. D. (1990). Genetics and molecular biology of alternative nitrogen fixation systems. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41, 109–125. doi: 10.1146/annurev.pp.41.060190.000545
- Boyd, E. S., Anbar, A. D., Miller, S., Hamilton, T. L., Lavin, M., and Peters, J. W. (2011). A late methanogen origin for molybdenum-dependent nitrogenase. *Geobiology* 9, 221–232. doi: 10.1111/j.1472-4669.2011.00278.x
- Boyd, E. S., Costas, A. M., Hamilton, T. L., Mus, F., and Peters, J. W. (2015). Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. *J. Bacteriol.* 197, 1690–1699. doi: 10.1128/JB.02611-14

- Boyd, E. S., and Peters, J. W. (2013). New insights into the evolutionary history of biological nitrogen fixation. *Front. Microbiol.* 4:201. doi: 10.3389/fmicb.2013.00201
- Bulen, W. A., and Lecomte, J. R. (1966). The nitrogenase system from *Azotobacter*: two enzyme requirements for N₂ reduction, ATP dependent H₂ evolution and ATP hydrolysis. *Proc. Natl. Acad. Sci. U.S.A.* 56, 979–986. doi: 10.1073/pnas.56.3.979
- Burris, R. H., and Roberts, G. P. (1993). Biological nitrogen fixation. *Annu. Rev. Nutr.* 13, 317–335. doi: 10.1146/annurev.nu.13.070193.001533
- Burén, S., Jiang, X., López-Torrejón, G., Echavarri-Erasun, C., and Rubio, L. M. (2017). Purification and *in vitro* activity of mitochondria targeted nitrogenase cofactor maturase NifB. *Front. Plant Sci.* 8:1567. doi: 10.3389/fpls.2017.01567
- Canfield, D. E., Glazer, A. N., and Falkowski, P. G. (2010). The evolution and future of Earth's nitrogen cycle. *Science* 330, 192–196. doi: 10.1126/science.1186120
- Chen, J. S., Toth, J., and Kasap, M. (2001). Nitrogen-fixation genes and nitrogenase activity in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. *J. Ind. Microbiol. Biotechnol.* 27, 281–286. doi: 10.1038/sj.jim.7000083
- Curatti, L., Hernandez, J. A., Igarashi, R. Y., Soboh, B., Zhao, D., and Rubio, L. M. (2007). *In vitro* synthesis of the iron-molybdenum cofactor of nitrogenase from iron, sulfur, molybdenum, and homocitrate using purified proteins. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17626–17631. doi: 10.1073/pnas.0703050104
- Curatti, L., Ludden, P. W., and Rubio, L. M. (2006). NifB-dependent *in vitro* synthesis of the iron-molybdenum cofactor of nitrogenase. *Proc. Natl. Acad. Sci. U.S.A.* 103, 5297–5301. doi: 10.1073/pnas.0601115103
- Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C., and Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* 13:162. doi: 10.1186/1471-2164-13-162
- Eady, R. R. (1996). Structure-function relationships of alternative nitrogenases. *Chem. Rev.* 96, 3013–3030. doi: 10.1021/cr950057h
- Echavarri-Erasun, C., Arragain, S., Scandurra, A. A., and Rubio, L. M. (2014). Expression and purification of NifB proteins from aerobic and anaerobic sources. *Methods Mol. Biol.* 1122, 19–31. doi: 10.1007/978-1-62703-794-5_3
- Einsle, O., Tezcan, F. A., Andrade, S. L., Schmid, B., Yoshida, M., Howard, J. B., et al. (2002). Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor. *Science* 297, 1696–1700. doi: 10.1126/science.1073877
- Eisen, J. A., Nelson, K. E., Paulsen, I. T., Heidelberg, J. F., Wu, M., Dodson, R. J., et al. (2002). The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9509–9514. doi: 10.1073/pnas.132181499
- Fay, A. W., Wiig, J. A., Lee, C. C., and Hu, Y. (2015). Identification and characterization of functional homologs of nitrogenase cofactor biosynthesis protein NifB from methanogens. *Proc. Natl. Acad. Sci. U.S.A.* 112, 14829–14833. doi: 10.1073/pnas.1510409112
- Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344
- Fish, W. W. (1988). Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples. *Methods Enzymol.* 158, 357–364. doi: 10.1016/0076-6879(88)58067-9
- Galagan, J. E., Nusbaum, C., Roy, A., Endrizzi, M. G., Macdonald, P., Fitzhugh, W., et al. (2002). The genome of *M. acetylavorans* reveals extensive metabolic and physiological diversity. *Genome Res.* 12, 532–542. doi: 10.1101/gr.223902
- George, S. J., Igarashi, R. Y., Xiao, Y., Hernandez, J. A., Demuez, M., Zhao, D., et al. (2008). Extended X-ray absorption fine structure and nuclear resonance vibrational spectroscopy reveal that NifB-co, a FeMo-co precursor, comprises a 6Fe core with an interstitial light atom. *J. Am. Chem. Soc.* 130, 5673–5680. doi: 10.1021/ja0755358
- Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J., and Rees, D. C. (1992). Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science* 257, 1653–1659. doi: 10.1126/science.1529353
- Guo, Y., Echavarri-Erasun, C., Demuez, M., Jimenez-Vicente, E., Bominaar, E. L., and Rubio, L. M. (2016). The nitrogenase FeMo-cofactor precursor formed by NifB is a diamagnetic 8 iron-containing cluster. *Angew. Chem. Int. Ed. Engl.* 55, 12764–12767. doi: 10.1002/anie.201606447
- Hernandez, J. A., Curatti, L., Aznar, C. P., Perova, Z., Britt, R. D., and Rubio, L. M. (2008). Metal trafficking for nitrogen fixation: NifQ donates molybdenum to NifEN/NifH for the biosynthesis of the nitrogenase FeMo-cofactor. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11679–11684. doi: 10.1073/pnas.0803576105
- Hernandez, J. A., Igarashi, R. Y., Soboh, B., Curatti, L., Dean, D. R., Ludden, P. W., et al. (2007). NifX and NifEN exchange NifB cofactor and the VK-cluster, a newly isolated intermediate of the iron-molybdenum cofactor biosynthetic pathway. *Mol. Microbiol.* 63, 177–192. doi: 10.1111/j.1365-2958.2006.05514.x
- Hoffman, B. M., Lukyanov, D., Yang, Z. Y., Dean, D. R., and Seefeldt, L. C. (2014). Mechanism of nitrogen fixation by nitrogenase: the next stage. *Chem. Rev.* 114, 4041–4062. doi: 10.1021/cr400641x
- Jeanthon, C., L'Haridon, S., Reysenbach, A. L., Vernet, M., Messner, P., Sleytr, U. B., et al. (1998). *Methanococcus infernus* sp. nov., a novel hyperthermophilic lithotrophic methanogen isolated from a deep-sea hydrothermal vent. *Int. J. Syst. Bacteriol.* 48, 913–919. doi: 10.1099/00207713-48-3-913
- Joerger, R. D., and Bishop, P. E. (1988). Nucleotide sequence and genetic analysis of the *nifB-nifQ* region from *Azotobacter vinelandii*. *J. Bacteriol.* 170, 1475–1487. doi: 10.1128/jb.170.4.1475-1487.1988
- Letunic, I., and Bork, P. (2007). Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127–128. doi: 10.1093/bioinformatics/btl529
- Lopez-Torrejon, G., Jimenez-Vicente, E., Buesa, J. M., Hernandez, J. A., Verma, H. K., and Rubio, L. M. (2016). Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat. Commun.* 7:11426. doi: 10.1038/ncomms11426
- Martinez-Noel, G., Curatti, L., Hernandez, J. A., and Rubio, L. M. (2011). NifB and NifEN protein levels are regulated by ClpX2 under nitrogen fixation conditions in *Azotobacter vinelandii*. *Mol. Microbiol.* 79, 1182–1193. doi: 10.1111/j.1365-2958.2011.07540.x
- McGlynn, S. E., Boyd, E. S., Peters, J. W., and Orphan, V. J. (2013). Classifying the metal dependence of uncharacterized nitrogenases. *Front. Microbiol.* 3:419. doi: 10.3389/fmicb.2012.00419
- Peters, J. W., Boyd, E. S., Hamilton, T. L., and Rubio, L. M. (2011). “Biochemistry of Mo-nitrogenase,” in *Nitrogen Cycling in Bacteria: Molecular Analysis*, ed. J. W. B. Moir (Norfolk, VA: Caister Academic Press).
- Raymond, J., Siebert, J. L., Staples, C. R., and Blankenship, R. E. (2004). The natural history of nitrogen fixation. *Mol. Biol. Evol.* 21, 541–554. doi: 10.1093/molbev/msh047
- Shah, V. K., Allen, J. R., Spangler, N. J., and Ludden, P. W. (1994). *In vitro* synthesis of the iron-molybdenum cofactor of nitrogenase. Purification and characterization of NifB cofactor, the product of NIFB protein. *J. Biol. Chem.* 269, 1154–1158.
- Shah, V. K., Davis, L. C., and Brill, W. J. (1972). Nitrogenase. I. Repression and derepression of the iron-molybdenum and iron proteins of nitrogenase in *Azotobacter vinelandii*. *Biochim. Biophys. Acta* 256, 498–511. doi: 10.1016/0005-2728(72)90078-3
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. doi: 10.1038/msb.2011.75
- Soboh, B., Boyd, E. S., Zhao, D., Peters, J. W., and Rubio, L. M. (2010). Substrate specificity and evolutionary implications of a NifDK enzyme carrying NifB-co at its active site. *FEBS Lett.* 584, 1487–1492. doi: 10.1016/j.febslet.2010.02.064
- Spatzal, T., Aksoyoglu, M., Zhang, L., Andrade, S. L., Schleicher, E., Weber, S., et al. (2011). Evidence for interstitial carbon in nitrogenase FeMo cofactor. *Science* 334, 940. doi: 10.1126/science.1214025
- Stewart, W. D., Fitzgerald, G. P., and Burris, R. H. (1967). In situ studies on N₂ fixation using the acetylene reduction technique. *Proc. Natl. Acad. Sci. U.S.A.* 58, 2071–2078. doi: 10.1073/pnas.58.5.2071
- Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577. doi: 10.1080/10635150701472164
- Trifinopoulos, J., Nguyen, L. T., Von Haeseler, A., and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44, W232–W235. doi: 10.1093/nar/gkw256
- Wahlund, T. M., and Madigan, M. T. (1993). Nitrogen fixation by the thermophilic green sulfur bacterium *Chlorobium tepidum*. *J. Bacteriol.* 175, 474–478. doi: 10.1128/jb.175.2.474-478.1993
- Wiig, J. A., Hu, Y., Lee, C. C., and Ribbe, M. W. (2012). Radical SAM-dependent carbon insertion into the nitrogenase M-cluster. *Science* 337, 1672–1675. doi: 10.1126/science.1224603

- Wiig, J. A., Hu, Y., and Ribbe, M. W. (2011). NifEN-B complex of *Azotobacter vinelandii* is fully functional in nitrogenase FeMo cofactor assembly. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8623–8627. doi: 10.1073/pnas.1102773108
- Wilcoxon, J., Arragain, S., Scandurra, A. A., Jimenez-Vicente, E., Echavarri-Erasun, C., Pollmann, S., et al. (2016). Electron paramagnetic resonance characterization of three iron-sulfur clusters present in the nitrogenase cofactor maturase NifB from *Methanocaldococcus infernus*. *J. Am. Chem. Soc.* 138, 7468–7471. doi: 10.1021/jacs.6b03329
- Zhao, D., Curatti, L., and Rubio, L. M. (2007). Evidence for *nifU* and *nifS* participation in the biosynthesis of the iron-molybdenum cofactor of nitrogenase. *J. Biol. Chem.* 282, 37016–37025. doi: 10.1074/jbc.M708097200

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Arragain, Jiménez-Vicente, Scandurra, Burén, Rubio and Echavarri-Erasun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Selection Signatures in the First Exon of Paralogous Receptor Kinase Genes from the Sym2 Region of the *Pisum sativum* L. Genome

Anton S. Sulima¹, Vladimir A. Zhukov^{1*}, Alexey A. Afonin¹, Aleksandr I. Zhernakov¹, Igor A. Tikhonovich^{1,2} and Ludmila A. Lutova²

¹ All-Russia Research Institute for Agricultural Microbiology, Saint-Petersburg, Russia, ² Department of Genetics and Biotechnology, Faculty of Biology, Saint-Petersburg State University, Saint-Petersburg, Russia

OPEN ACCESS

Edited by:

Jari Valkonen,
University of Helsinki, Finland

Reviewed by:

Clare Gough,
Institut National de la Recherche
Agronomique de Toulouse, France
Yangrong Cao,
Huazhong Agricultural University,
China

*Correspondence:

Vladimir A. Zhukov
vzhukov@arriam.ru

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 25 May 2017

Accepted: 30 October 2017

Published: 14 November 2017

Citation:

Sulima AS, Zhukov VA, Afonin AA, Zhernakov AI, Tikhonovich IA and Lutova LA (2017) Selection Signatures in the First Exon of Paralogous Receptor Kinase Genes from the Sym2 Region of the *Pisum sativum* L. Genome. *Front. Plant Sci.* 8:1957.
doi: 10.3389/fpls.2017.01957

During the initial step of the symbiosis between legumes (Fabaceae) and nitrogen-fixing bacteria (rhizobia), the bacterial signal molecule known as the Nod factor (nodulation factor) is recognized by plant LysM motif-containing receptor-like kinases (LysM-RLKs). The fifth chromosome of barrel medic (*Medicago truncatula* Gaertn.) contains a cluster of paralogous LysM-RLK genes, one of which is known to participate in symbiosis. In the syntenic region of the pea (*Pisum sativum* L.) genome, three genes have been identified: *PsK1* and *PsSym37*, two symbiosis-related LysM-RLK genes with known sequences, and the unsequenced *PsSym2* gene which presumably encodes a LysM-RLK and is associated with increased selectivity to certain Nod factors. In this work, we identified a new gene encoding a LysM-RLK, designated as *PsLykX*, within the Sym2 genomic region. We sequenced the first exons (corresponding to the protein receptor domain) of *PsSym37*, *PsK1*, and *PsLykX* from a large set of pea genotypes of diverse origin. The nucleotide diversity of these fragments was estimated and groups of haplotypes for each gene were revealed. Footprints of selection pressure were detected via comparative analyses of SNP distribution across the first exons of these genes and their homologs *MtLYK2*, *MtLYK3*, and *MtLYK4* from *M. truncatula* retrieved from the Medicago Hapmap project. Despite the remarkable similarity among all the studied genes, they exhibited contrasting selection signatures, possibly pointing to diversification of their functions. Signatures of balancing selection were found in LysM1-encoding parts of *PsSym37* and *PsK1*, suggesting that the diversity of these parts may be important for pea LysM-RLKs. The first exons of *PsSym37* and *PsK1* displayed signatures of purifying selection, as well as *MtLYK2* of *M. truncatula*. Evidence of positive selection affecting primarily LysM domains was found in all three investigated *M. truncatula* genes, as well as in the pea gene *PsLykX*. The data suggested that *PsLykX* is a promising candidate for *PsSym2*, which has remained elusive for more than 30 years.

Keywords: pea (*Pisum sativum* L.), molecular evolution, legume-rhizobial symbiosis, LysM-containing receptor-like kinases, Sym2, Nod factor perception, *Medicago truncatula*

INTRODUCTION

One of the hallmark features of legumes (Fabaceae) is their ability to form beneficial symbioses with nitrogen-fixing soil bacteria collectively known as rhizobia. The establishment of such symbioses is a very complex process that involves many genes, as it requires proper recognition of the symbiotic partners, with subsequent development of novel symbiotic organs and partial integration of metabolic pathways in organisms belonging to different domains (Oldroyd et al., 2011; Suzuki et al., 2015; Zipfel and Oldroyd, 2017). The main function of these symbioses is the fixation of atmospheric nitrogen, making it a significant evolutionary advantage for legumes (Wheatley and Sprent, 2010; Werner et al., 2015).

The establishment of the legume–rhizobial symbiosis begins with mutual recognition of the partners. During this initial step, a bacterial lipo-chitooligosaccharide signaling molecule known as the Nod factor (nodulation factor; Denarie et al., 1996) is recognized by plant receptors in the LysM-RLK protein family (Limpens et al., 2003; Madsen et al., 2003; Broghammer et al., 2012). LysM-RLKs are receptor-like kinases (RLKs) with three lysin motifs (LysMs) in the ligand-binding extracellular region (Nakagawa et al., 2011; Mesnage et al., 2014). Although LysMs are widespread among prokaryotes and eukaryotes (with the exception of the Archaea), they are only connected to a kinase domain in plants (Ponting et al., 1999; Buist et al., 2008). Various LysM-RLKs are required to recognize molecular signals from rhizobia, arbuscular-mycorrhizal fungi (Buendía-Clavería et al., 2003; Oldroyd, 2013; Gobbato, 2015; Kawaharada et al., 2015; Buendia et al., 2016; Rasmussen et al., 2016), and some pathogenic microorganisms (Zhang et al., 2015). Thus, they occur also in non-legumes such as, *Arabidopsis thaliana*, which is unable to form nitrogen-fixing nodules or mycorrhizal associations (Veiga et al., 2013). Nonetheless, LysM-RLKs play a crucial role in regulating the legume–rhizobial symbiosis, providing plant selectivity toward microsymbiotic partners. Consequently, LysM-RLKs contribute to the selection of the most effective combinations of micro- and macrosymbionts (Provorov and Vorobyov, 2013).

Many genes encoding LysM-RLKs have been described in model legumes, such as, barrel medic (*Medicago truncatula* Gaertn; Arrighi et al., 2006), and *Lotus japonicus* (Regel) K. Larsen (Lohmann et al., 2010). Some are known to participate in symbiosis, while others have non-symbiotic functions. Interestingly, each legume species seems to have its own nuances in the organization and function of symbiotic LysM-RLKs; for example, mutations in the orthologous LysM-RLK genes *NFR1* (NOD FACTOR RECEPTOR 1) from *L. japonicus* and *MtLYK3* (LysM DOMAIN-CONTAINING RECEPTOR-LIKE KINASE 3) from *M. truncatula* manifested differently, blocking the penetration of rhizobia into plant cells at distinct stages of symbiotic development (Smit et al., 2007).

Garden pea (*Pisum sativum* L.) is a valuable pulse crop (<http://www.fao.org/faostat>). Studies on its symbiosis with rhizobia are very important (Borisov et al., 2004, 2007; Smýkal et al., 2012; Zhukov et al., 2016), since effective symbioses can increase the yield and quality of pea crops. Because of the long history of pea

domestication and its wide distribution (Jing et al., 2010; Smýkal et al., 2012), many research centers around the world have large collections of pea genetic resources, making it possible to study the natural variability of particular pea genes and corresponding traits using sample sets close to genetic saturation (Jing et al., 2010). The N. I. Vavilov Institute of Plant Genetic Resources (VIR) in St. Petersburg, Russia, maintains a uniquely valuable collection of crop germplasm that includes pea and other legumes (Plekhanova et al., 2017). Studies on large ecotype collections and the use of association genetics may help uncover valuable alleles related to responses to environmental stresses (Gentzbittel et al., 2015). Recent genome-wide association (GWA) studies of the chickpea (*Cicer arietinum* L.) germplasm collection of the VIR revealed potential candidate genes likely to affect traits of agricultural importance (Plekhanova et al., 2017). Unfortunately, many molecular genetic methods are unsuitable for analyses of the pea genome because it is very large, congested with repetitive elements (Macas et al., 2007), without even a draft assembly. However, the high degree of synteny between the genomes of pea and *M. truncatula*, a well-studied model legume, makes it possible to investigate pea genes using the *M. truncatula* genome as a reference (Young and Udvardi, 2009).

The first discovered symbiosis-related pea gene was *PsSym2*, which determines increased selectivity toward rhizobia in the pea cv. Afghanistan (Govorov, 1928; Razumovskaya, 1937; Lie, 1984). Although the phenotype conferred by the allelic state of *PsSym2* is clear, the molecular basis of its function is not. It is thought to encode a LysM-RLK, because the selectivity inherent to cv. Afghanistan is associated with the ability to distinguish among the structural features of Nod factors produced by different (effective, ineffective, or even pathogen-like) rhizobial strains. Previously, *PsSym2* was localized in the linkage group I (LG I) of the pea genome (Kozik et al., 1996). Later, two paralogous genes encoding LysM-RLKs (*PsSym37* and *PsK1*) were found to co-localize there (Zhukov et al., 2008). However, neither of them was shown to play the same role as *PsSym2*, although both were shown to participate in bacterial signal recognition (Zhukov et al., 2008). In *M. truncatula*, the genomic region corresponding to the pea *PsSym2* location (LYK region) contains at least seven highly similar RLK genes; *MtLYK1–MtLYK7* (Limpens et al., 2003). One of these genes, *MtLYK3*, was identified as a symbiotic gene that controls bacterial penetration into the infection thread (Limpens et al., 2003; Smit et al., 2007). Considering these facts, one could expect to find more genes encoding RLKs in pea LG I, including genes related to symbiosis.

The early domestication and subsequent global dissemination of pea has led to significant and multifaceted selection pressure, resulting from both natural and breeding selection (Smýkal et al., 2012). Several loci associated with agronomically important traits have been found to play crucial roles in the domestication process (Weeden, 2007). However, the loci responsible for symbiosis, such as, bacterial recognition genes, are unlikely to be directly affected by breeding, although their influence on plant productivity should result in some selective pressure. Thus, the aim of this work was to estimate the genetic diversity of the genes encoding LysM-RLKs in the *Sym2* region of the pea genome and corresponding genes from the LYK region in

the *M. truncatula* genome, and to evaluate and compare the selection pressures affecting different parts of the sequences. To this end, we (i) examined the region containing *PsSym37* and *PsK1* in pea LG I to identify new LysM-RLK genes; (ii) studied the polymorphism of the first exon sequences (corresponding to the receptor domain of the encoded protein) of three LysM-RLK genes (two known genes and one newly identified gene) in a large set of pea genotypes; and (iii) compared the selection signatures in pea genes with those in their *M. truncatula* homologs. As a result, we found a new gene, which we named *PsLykX* (for *P. sativum* *LysM kinase eXclusive*), and detected signatures of positive, balancing, and negative selection in the gene sequences. The polymorphism levels and patterns of selective signatures were unique to each of the three pea genes, but remarkably similar for the *Medicago* genes *MtLYK3* and *MtLYK4*. The results of the polymorphism analyses indicate that *PsLykX* is a promising candidate for *PsSym2*, the determinant of the unique symbiotic phenotype of cv. Afghanistan.

MATERIALS AND METHODS

Screening of the Pea BAC Library

To find novel LysM-RLK genes in the *Sym2* region, the pea Psa-B-Cam BAC library screening was performed at the INRA-CNRCV Plant Genomic Center (Toulouse, France). The screening was carried out via quantitative PCR, using BAC DNA pools amplified with Phi29 DNA polymerase and primer pairs for *PsSym37* and *PsK1* genes (Table S1). Information about the BAC library can be found online: (<http://cnrcv.toulouse.inra.fr/>).

The whole insertion in the identified BAC Psa-B-Cam 446P4 was sequenced using the 454 GS-Junior system with the Titanium kit in CNRCV, yielding 11697 reads with a mean length of 500 bp. Reads were then quality-trimmed using AdapterRemoval software (v. 2.1.7) (Lindgreen, 2012) and mapped onto the *Escherichia coli* DH5 α strain genome using hisat2 mapper (Kim et al., 2015) to eliminate bacterial contamination. Then, 10,927 non-aligned reads were assembled using SPAdes (v 3.9.1) assembler (Bankevich et al., 2012), resulting in three contigs, which were uploaded to the NCBI database as a single entry under the accession number MF185734. Three genes (*PsSym37*, *PsK1*, and the novel gene *PsLykX*) were found via BLASTn searches (e-value cutoff 10^{-6}) using the sequence of *PsK1* as the query (Figure 1). Additional BLAST searches against the NCBI non-redundant nucleotide database revealed fragmented transposable elements (Figure 1), but no additional genes.

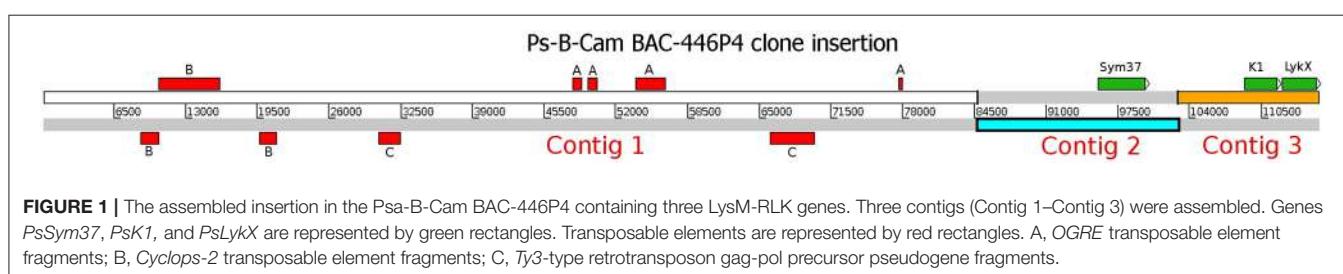
The 3'-and 5'-ends of *PsLykX* cDNA were obtained by rapid amplification of cDNA ends (RACE) using total cDNA from inoculated roots of cv. Cameor. The Mint kit was used for cDNA synthesis and Mint kit primers were used for amplification of cDNA ends (Evrogen, Moscow, Russia). To obtain RNA for cDNA synthesis, five cv. Cameor plants were inoculated with *Rhizobium leguminosarum* bv. *viciae* strain RCAM 1026 and then grown in sterile sand for 10 days (Afonin et al., 2017). The specific primers used for RACE PCR are listed in Table S1.

Plant Material

The pea genotypes were previously selected by Borisov et al. (2002) from the collection of cultivated peas of the VIR (Saint Petersburg, Russia). The criterion for selection was diversity of places of origin. Preference was given to wild genotypes or primitive varieties, the so-called landraces. This set of 99 genotypes has already been tested for the efficiency of symbiosis with arbuscular-mycorrhizal (AM) fungi under the conditions of rhizobia inoculation (Jacobi et al., 2000; Borisov et al., 2002), and also for their responses to the toxic heavy metal cadmium under symbiotic conditions (Belimov et al., 2015). The originally obtained samples were propagated by selfing for several generations at the All-Russia Research Institute for Agricultural Microbiology (ARRIAM, Saint Petersburg, Russia). The genotypes and their places of origin (when known) are shown in Table S2.

Plant DNA Extraction

DNA was extracted from 3- to 4-day old seedlings of each pea genotype. The seedlings were cultivated in Petri dishes filled with sterile vermiculite at 28°C. Three to five seedlings per genotype were bulked and used for DNA extraction using the slightly modified CTAB protocol described elsewhere (Rogers and Bendich, 1985). Briefly, the steps were as follows: seedlings, 3–4 glass beads, and 700 μ l 2 \times CTAB buffer were placed into flat-bottomed microcentrifuge tubes and ground with the Fastprep-24 homogenizer (MP Biomedicals, Irvine, CA USA; 2 \times 60 s at maximum frequency), then incubated in water bath for 2 h at 65°C. Then, 700 μ l chloroform/isoamyl alcohol (24:1 v:v) was added to each tube, the mixture was vortexed, and then centrifuged for 10 min at 16,873 g. The upper fraction was transferred to a fresh microcentrifuge tube, and the previous step was repeated with 300–500 μ l chloroform instead of chloroform/isoamyl alcohol. The upper fraction was again transferred to a fresh tube, and 1,000 μ l 96% ethanol and 20 μ l 5 M NaCl were added. The tubes were gently inverted two to three



times to mix the contents, and then the mixture was centrifuged (10 min, 16,873 g). The supernatant was discarded, and 500 μ L 70% ethanol was added to the pellet. The samples were vortexed, incubated at room temperature for 30 min, and then centrifuged (10 min, 16,873 g). The ethanol was discarded, and the tubes were incubated at 37°C for 20–30 min. Finally, the samples were dissolved in 20–30 μ L TE buffer (pH 8.0). In total, DNA samples from 93 separate genotypes were used in further analyses.

Sequencing of First Exon of Pea Genes Encoding LysM-RLKs

To amplify the first exons of *PsSym37*, *PsK1*, and *PsLykX*, a set of PCR primers was designed (Table S1). For different pea lines, different pairs of primers were used to ensure successful amplification. The quality of primers was verified by the OligoCalc on-line service: <http://biotools.nubic.northwestern.edu/OligoCalc.html> (Kibbe, 2007). The PCRs were performed in 96-well plates on an iCycler (Bio-Rad, Hercules, CA, USA) or Dyad (Bio-Rad, USA) instrument using the ScreenMix-HS kit (Evrogen, Moscow, Russia). The PCR cycling conditions were as follows: 95°C (5 min), 35 \times [95°C (30 s), Tm (varying depending on primers) (30 s), 72°C (1 min)], 72°C (5 min). The PCR fragments were sequenced using the ABI Prism3500xL system (Applied Biosystems, Palo Alto, CA, USA) at the “Genomic Technologies, Proteomics, and Cell Biology” Core Center of the ARRIAM. The sequences of the first exons have been deposited in the NCBI under the accession numbers MF155289–MF155381 (*PsK1*), MF155382–MF155469 (*PsLykX*), and MF155470–MF155549 (*PsSym37*).

Data Collection from *M. truncatula* Hapmap Project

The *Medicago* Hapmap project, a long-term, community-accessible GWA mapping resource, is based on re-sequencing of 384 inbred lines spanning the range of *Medicago* diversity using Illumina next-generation sequencing technology (Stanton-Geddes et al., 2013). To study the polymorphisms of genes encoding LysM-RLKs, sequences of the first exons of *MtLYK2* (Medtr5g086310), *MtLYK3* (Medtr5g086130), and *MtLYK4* (Medtr5g086120) were retrieved from the *Medicago* Hapmap project website, assembly Mt4.0 (<http://www.medicagohapmap.org/>). Single nucleotide variants at particular sites were considered different from the reference *M. truncatula* genome if they were detected in more than 50% of corresponding Illumina reads. In total, 116 sequences for *MtLYK2*, 220 for *MtLYK3*, and 196 for *MtLYK4* were included in subsequent analyses.

Pairwise Comparison of LysM-RLKs

Pairwise comparison of all LysM receptor kinases was performed using the Needleman–Wunsch global alignment algorithm from the EMBOSS suite v.6.3.1 (Rice et al., 2000), using standard parameters. Comparisons of the three pea and three *M. truncatula* LysM-RLKs were performed separately for whole sequences, the kinase part, and the receptor part. The results from each comparison were combined using a custom python script. The results for full genes are shown in Table 1, and the results for the separate regions are shown in Tables S3, S4.

Pairwise comparison of the known pea *Sym2* region and the *M. truncatula* LYK region was performed using BLAST and visualized using ACT software (release 13.0.0; Carver et al., 2005). The sequence of the *Sym2* region was obtained from the BAC Psa-B-Cam 446P4 clone (see Results). The *M. truncatula* LYK region sequence (g5:37100000–g5:537450000) was retrieved from the Phytozome website (www.phytozome.net/). The minimal identity cut-off was set to 70%. The exon/intron structure of genes in these regions was modeled by the Exonerate package (<http://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate>) with the -est2genome option enabled to search for gene mRNAs in the respective genomes.

Polymorphism Analysis and Detection of Selection Pressures

Neighbor-joining (NJ) trees were constructed using MEGA v.6.60 software (Tamura et al., 2013) with the assumption that substitutions followed the Jukes–Cantor model and had uniform rates among sites. Bootstrap tests of NJ trees were performed with 500 bootstrap replications. The branches were cut off at 70% bootstrap support. Trees were visualized using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

We used DnaSP v5.10.01 (Librado and Rozas, 2009) to assess expected haplotype heterozygosity (HHe), and nucleotide diversity (π), including that for synonymous (π_s) and non-synonymous (π_a) sites, and to conduct the following neutrality tests: Tajima’s D test (Tajima, 1989), Fu’s Fs test (Fu, 1997), Fu and Li’s D and H tests, Fay and Wu’s H test (Fay and Wu, 2000), and McDonald–Kreitman MK test (McDonald and Kreitman, 1991). When needed, an outgroup sequence was chosen as follows: *PsSym37* (cv. Cameor) for barrel medic genes and *MtLYK3* (cv. Jemalong) for pea genes. DnaSP v5.10.01 was also used for sliding window analyses (100-bp window size, 25-bp step size), and to determine the significance of departure from the neutrality model by coalescent simulations with 1,000 replicates. Other neutrality tests, namely normalized Fay and Wu’s H (*nH*) test (Zeng et al., 2006) and Ewens–Watterson (EW) test (Zeng et al., 2007), were carried out using DH software (<http://zeng-lab.group.shef.ac.uk>).

The codon-based Z-test considering the average rate of synonymous (*dS*) and nonsynonymous (*dN*) substitutions per site was performed in MEGA v.6.60 using the modified Nei–Gojobori method (Nei and Gojobori, 1986), with Jukes–Cantor correction (Jukes and Cantor, 1969) for multiple substitutions. Standard errors were estimated from 1,000 bootstrap replicates.

In all cases, separate estimates were made for the whole sequences of the coding part of the first exons and for sequences corresponding to LysM modules and intermediate parts. The borders of LysM modules were drawn as described previously (Zhukov et al., 2008).

Inoculation Experiment

An inoculation experiment was performed to determine the symbiotic phenotype of the pea line K-6883 (84). Seeds of K-6883 and cv. Cameor used as the inoculum control were planted in 2-L pots with sterile sand and inoculated with either strain RCAM1026 or A1 (Chetkova and Tikhonovich, 1986) of *R. leguminosarum* bv. *viciae*. Each pot contained five seeds, and the

TABLE 1 | Identity and similarity of *M. truncatula* MtLYK1-7, *P. sativum* PsSym37, PsK1 and PsLykX, *Lotus japonicus* LjNFR1 and *Arabidopsis thaliana* AtCERK1 whole putative proteins calculated by pairwise comparison.

	AtCERK1	LjNFR1	MtLYK1	MtLYK2	MtLYK3	MtLYK4	MtLYK5	MtLYK6	MtLYK7	PsK1	PsLykX	PsSym37
IDENTITY												
AtCERK1	100	55.1	47.6	54.2	54.8	46.7	48.8	53.7	55.7	54.6	48.6	54.8
LjNFR1	55.1	100	52	78.2	77.8	58.1	54.9	54.3	57.4	78.8	60.5	78.2
MtLYK1	47.6	52	100	51.9	50.6	64.5	72.1	57.7	56.5	51.7	61.6	51.6
MtLYK2	54.2	78.2	51.9	100	81	59.9	51.9	54.2	56.2	80.9	64.4	81.2
MtLYK3	54.8	77.8	50.6	81	100	66	50.5	54	57.2	84.4	63.5	83.7
MtLYK4	46.7	58.1	64.5	59.9	66	100	71.5	50.5	52.6	61.2	70	61.1
MtLYK5	48.8	54.5	72.1	51.9	50.5	71.5	100	55.3	58	53.3	64.5	53.2
MtLYK6	53.7	54.3	57.7	54.2	54	50.5	55.3	100	63.8	54.2	51.5	53.5
MtLYK7	55.7	57.4	56.5	56.2	57.2	52.6	58	63.8	100	57.6	52.7	57.7
PsK1	54.6	78.8	51.7	80.9	84.4	61.2	53.3	54.2	57.6	100	65	86.3
PsLykX	48.6	60.5	61.6	64.4	63.5	70	64.5	51.5	52.7	65	100	64.9
PsSym37	54.8	78.2	51.6	81.2	83.7	61.1	53.2	53.5	57.7	86.3	64.9	100
SIMILARITY												
AtCERK1	100	71.2	63	69	70.3	61.8	64.8	68.9	71.5	69.1	63.1	68.1
LjNFR1	71.2	100	65	86.8	86.3	71.1	66.8	68.5	71.9	87.1	72.2	86.6
MtLYK1	63	65	100	65.3	63.4	73.7	81.6	71	69.8	64.6	71.6	63.8
MtLYK2	69	86.8	65.3	100	89.5	73.6	67.1	69.7	71.7	90.2	76.2	88.7
MtLYK3	70.3	86.3	63.4	89.5	100	75.6	64.5	68.1	71.2	91.3	75.3	89.8
MtLYK4	61.8	71.1	73.7	73.6	75.6	100	80.9	65.6	65.9	73.1	79.9	74.1
MtLYK5	64.8	66.7	81.6	67.1	64.5	80.9	100	70.3	71.2	66.8	74.4	67.1
MtLYK6	68.9	68.5	71	69.7	68.1	65.6	70.3	100	75.6	69.1	65.9	66.4
MtLYK7	71.5	71.9	69.8	71.7	71.2	65.9	71.2	75.6	100	72.2	67.3	71.3
PsK1	69.1	87.1	64.6	90.2	91.3	73.1	66.8	69.1	72.2	100	76.3	92.4
PsLykX	63.1	72.2	71.6	76.2	75.3	79.9	74.4	65.9	67.3	76.3	100	76.2
PsSym37	68.1	86.6	63.8	88.7	89.8	74.1	67.1	66.4	71.3	92.4	76.2	100

The “heat map” coloring indicates the degree of identity and similarity.

experiment was carried out with two technical replicates. After 28 days, the plants were removed from the pots and the phenotype of the root system was examined. The average number of nodules was calculated in SigmaPlot 12.0 (Systat Software, Inc., San Jose, CA, USA).

RESULTS

Discovery of a New LysM-RLK Gene in the Sym2 Region Using Pea BAC Library

To detect novel genes encoding LysM-RLKs in the *Sym2* region of pea LG I, the pea Psa-B-Cam BAC library was screened at the INRA-CNRGV Plant Genomic Center using quantitative PCR with primer pairs for *PsSym37* and *PsK1* (see section Materials and Methods). In the BAC clone Psa-B-Cam 446P4, we identified the two previously known LysM-RLK genes *PsSym37* and *PsK1* (used as probes), and a new LysM-RLK gene that we named *PsLykX* (for *LysM kinase eXclusive*; see Figure 1). No other BAC clones were detected during further screening with a part of *PsLykX* as the probe. This may indicate a lack of saturation of the Psa-B-Cam library in this genomic region.

PsLykX was located close to *PsK1*; the stop codon of *PsK1* and the start codon of *PsLykX* were separated by 531 bp (Figure 1,

also see MF185734 at GenBank). Despite this, *PsLykX* appeared to be a fully functional gene, because sequences perfectly matching its cDNA were found in pea transcriptome assemblies generated from nitrogen-fixing nodules (Alves-Carvalho et al., 2015; Sudheesh et al., 2015; Zhukov et al., 2015). The *PsLykX* complete open reading frame was successfully amplified from cDNA generated from RNA extracted from pea roots inoculated with nodule bacteria, and the sequences of *PsLykX* cDNA ends were verified using 3'- and 5'-RACE. Alignment of the genomic sequence and the cDNA of *PsLykX* revealed a gene structure similar to that of *PsSym37*, with 12 exons and 11 introns (Zhukov et al., 2008), which is typical of symbiotic LysM-RLK genes (Figure 2A). The sequence of *PsLykX* has been deposited in GenBank under the accession number MF135533.

The putative *PsLykX* protein contained three LysM modules encoded by the first exon and forming the receptor domain, along with transmembrane and kinase domains (Figure 2B). Multiple alignment between *PsLykX* and known LYKs from the syntenic LYK region of the *M. truncatula* genome and LysM-RLKs from other organisms revealed that *PsLykX* was closest to MtLYK4, MtLYK1 and MtLYK5 (Figure 3). In the neighbor-joining phylogenetic tree, *PsLykX* and MtLYK4 were grouped apart from *PsSym37* and *PsK1* which formed a distinctive clade with

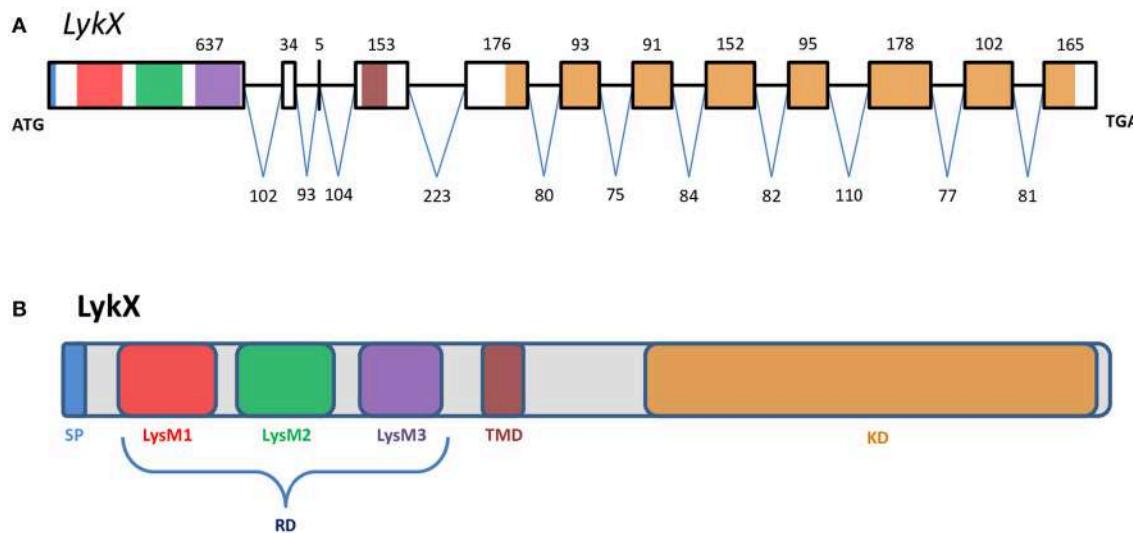


FIGURE 2 | The exon-intron structure of the *PsLykX* gene (**A**) with domain structure of the corresponding LysM-RLK protein (**B**). (**A**) Exons are designated by rectangles, introns—by horizontal lines. Numbers indicate the size of exons and introns. Colors match the colors of protein domains in (**B**) encoded by marked exons. (**B**) Colored blocks represent protein domains. SP, signal peptide; LysM1,2,3, LysM modules of receptor domain; RD, receptor domain; TMD, transmembrane domain; KD, kinase domain.

MtLYK2 and MtLYK3 of *M. truncatula* and symbiotic LysM-RLK NFR1 of *L. japonicus* (Figure 3).

Polymorphism of First Exon Sequences of LysM-RLK Genes in Sym2 Region

To assess the level of polymorphism in parts of pea LysM-RLK genes encoding ligand binding structures, we sequenced the first exons (corresponding to the receptor domain) of the three identified LysM-RLK genes from the *Sym2* region in a subset of cultivated pea accessions (Table S2). To fully characterize the extent of polymorphism, previously obtained sequences of *PsSym37* and *PsK1* (Zhukov et al., 2008) available from online databases were included in the analysis. Overall, 101 sequences were analyzed for *PsK1*, 90 sequences for *PsLykX*, and 89 sequences for *PsSym37* (Table S2).

The nucleotide sequences of the first exons were aligned using the codon-based ClustalW algorithm, and phylogenetic trees based on LysM-RLKs variability were constructed (Figures 4–6 for *PsSym37*, *PsK1*, and *PsLykX*, respectively). The clades of the resulting trees were considered to represent distinct haplotypes. Comparison of amino acid sequences confirmed the haplotype distribution pattern (Table 2). *PsSym37* and *PsK1* sequences formed two main groups, each with several subgroups (Figures 4, 5), while *PsLykX* sequences appeared to be more diverse and formed six distinct groups (Figure 6).

The amino acid alignment analyses showed that *PsLykX* was the most variable among the three genes, with 16 amino acid polymorphisms in the putative protein sequence. According to these, the whole sample was divided into 13 groups, seven of which were represented by single unique genotypes with additional polymorphic sites. *PsK1* had 17 amino acid

polymorphisms in the putative protein sequence, dividing sample into 12 groups (with six single-genotype groups). Interestingly, although group B of *PsK1* split into two subgroups, B1 and B2, according to nucleotide-based phylogeny, the haplotypes in these subgroups were identical at the amino acid level. *PsSym37* appeared to be more uniform, with only six groups based on 10 amino acid polymorphisms (with one group being represented by a sole genotype).

In all three genes, the most polymorphic region was the sequence corresponding to the first LysM domain (LysM1); it contained the majority of non-synonymous substitutions (Table 2). In *PsK1* and *PsLykX*, several non-synonymous variations were detected in the part corresponding to the signal peptide (SP) as well, although these changes were unlikely to affect protein localization. Further research is required to explore the possible functional significance of these changes.

We identified *PsLykX* as a new gene from the *Sym2* region encoding an LysM-RLK. This gene is a promising candidate for the *PsSym2* gene, which is responsible for increased selectivity toward rhizobia in several genotypes of pea originating from Afghanistan and neighboring areas (Lie, 1984; Kozik et al., 1995). Plants carrying the so-called “Afghan” allele of *PsSym2* (*Sym2^A*) form nodules only with strains producing a double-acetylated Nod factor, while plants with the most common “European” allele (*Sym2^E*) are able to perceive the mono-acetylated Nod factor as well (Davis et al., 1988). Bearing this in mind, we sequenced the first exon of *PsLykX* of cv. Afghanistan (NGB2150), and found that it contained a unique haplotype of *PsLykX*, which was shared with only one other line, K-6883 (84). We investigated the symbiotic phenotype of K-6883 in an experiment with two strains of *R. leguminosarum* bv. *viciae*: RCAM1026 (Afonin et al.,

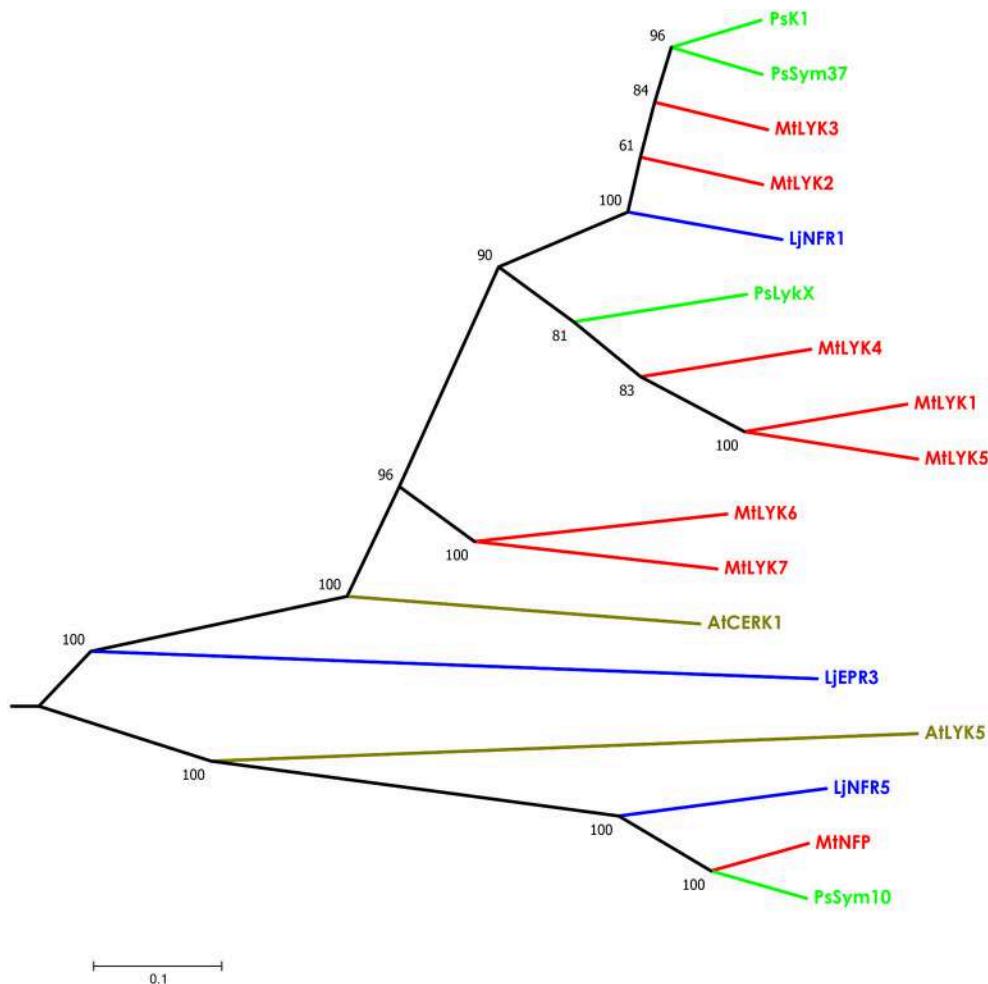


FIGURE 3 | Phylogenetic tree showing the comparison of the whole amino acid sequences of LysM-RLKs from *Sym2* region of pea genome, LysM-RLKs from syntenic LYK region of *Medicago truncatula* genome, and homologous LysM-RLKs from *Lotus japonicus* and *Arabidopsis thaliana*. Branches are colored as follows: pea (Ps)—green, *M. truncatula* (Mt)—red, *Lotus japonicus* (Lj)—blue, *Arabidopsis thaliana* (At)—olive. Tree was built using the Neighbor-Joining method with the assumption that substitutions followed the Jukes–Cantor model and had uniform rates among sites. Bootstrap test of NJ tree was performed with 500 bootstrap replications.

2017), which produces a mono-acetylated Nod factor, and A1, which produces a double-acetylated Nod factor (Chetkova and Tikhonovich, 1986; Ovtynya et al., 1999). K-6883 failed to form nodules with RCAM1026 but formed 105.0 ± 22.3 nodules with A1, while cv. Cameor (control belonging to another *PsLykX* haplotype) formed 151.6 ± 14.1 nodules with RCAM1026 and 105.2 ± 11.7 nodules with A1. Thus, K-6883 showed “Afghan” selectivity toward rhizobia, providing further evidence that *PsLykX* may actually be *PsSym2*.

Identification of Selection Signatures

Investigation into the Correspondence between Pea and *M. truncatula* LysM-RLK Genes

Polymorphism analyses can provide insights into the evolutionary pathways leading to the contemporary state of genes in related species. The fact that pea and barrel medic

LysM-RLK genes are paralogs located close to each other made it difficult to estimate selection signatures in their sequences. The *Sym2* region of pea LG I is syntenic to the *M. truncatula* LYK region on chromosome 5 (Gualtieri et al., 2002), where seven LysM-RLK genes are located (Limpens et al., 2003), while only three genes are known to be present in the *Sym2* region of the pea genome. Thus, to compare the diversity of the pea LysM-RLK genes with that of the *M. truncatula* ones, we first determined the relationship between *PsSym37*, *PsK1*, and *PsLykX* and the *M. truncatula* LYK genes. To this end, we compared the whole sequences of LysM-RLK putative proteins (three from pea, seven from *Medicago*), as well as their receptor and kinase domains separately (Table 1, Tables S3, S4). We included *LjNFR1* as an ortholog of *PsSym37* and *MtLYK3* (Zhukov et al., 2008), and *AtCERK1* as non-symbiotic LysM-RLK of a distant species in these comparisons. The

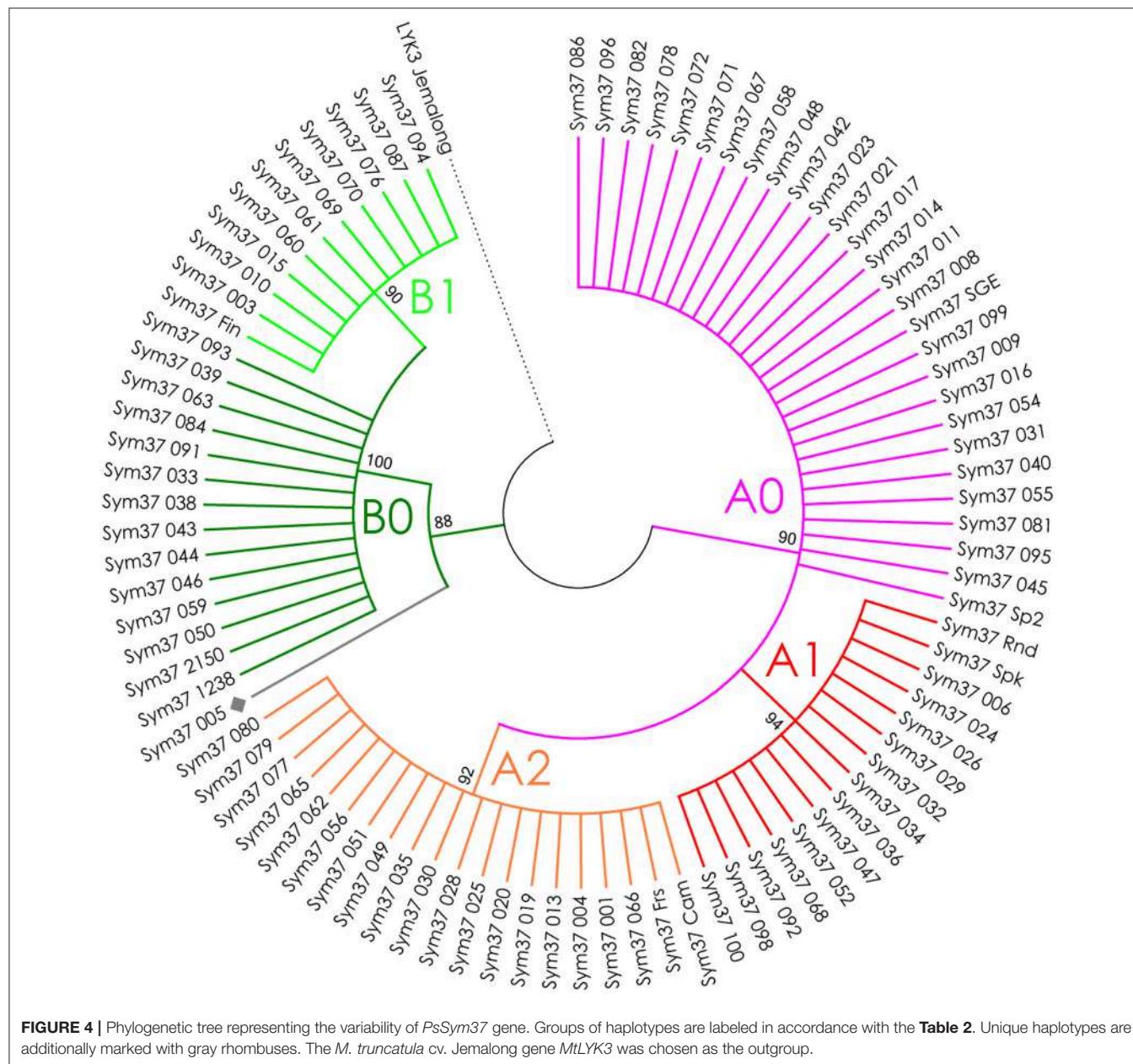
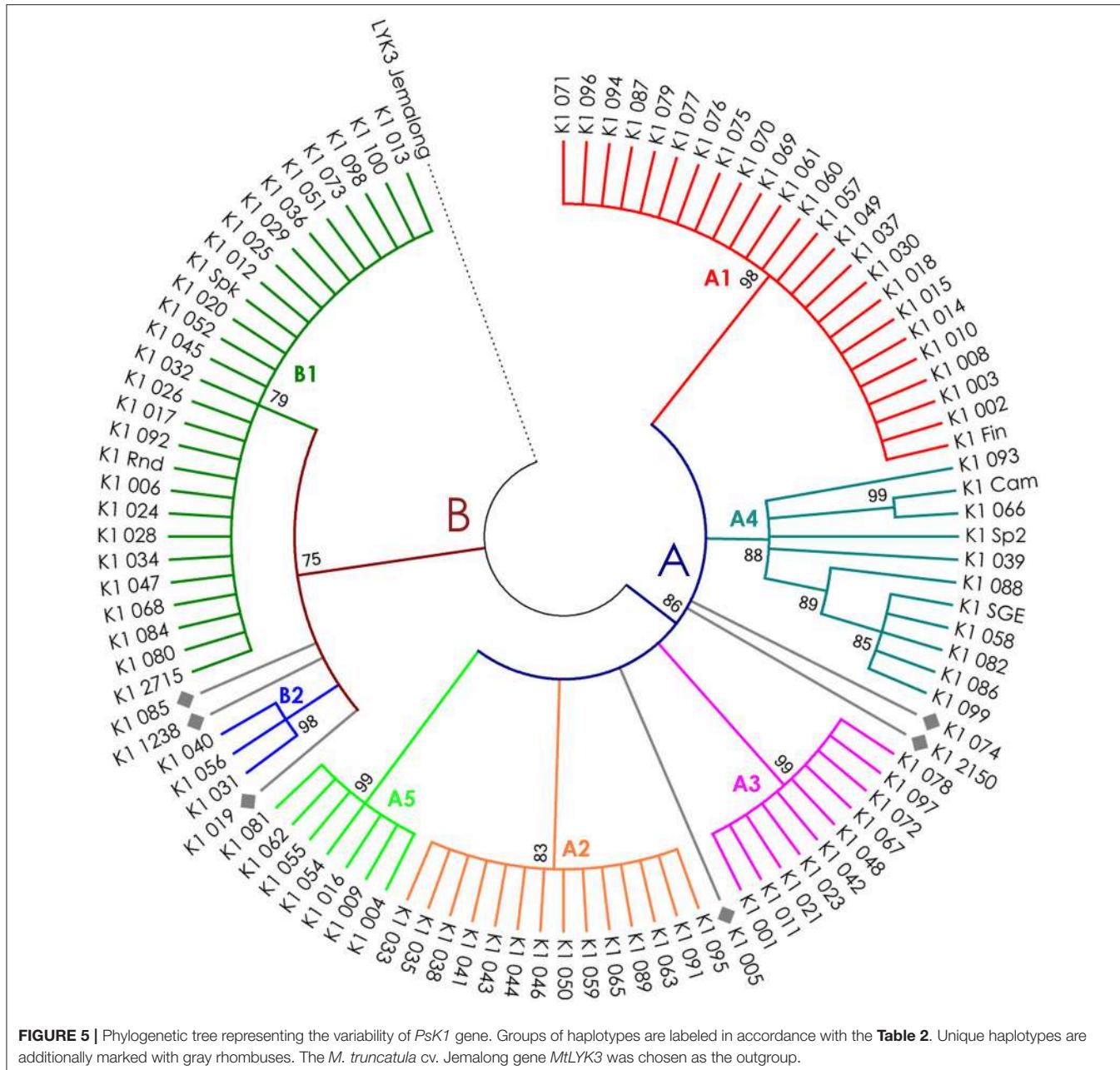


FIGURE 4 | Phylogenetic tree representing the variability of *PsSym37* gene. Groups of haplotypes are labeled in accordance with the **Table 2**. Unique haplotypes are additionally marked with gray rhombuses. The *M. truncatula* cv. Jemalong gene *MtLYK3* was chosen as the outgroup.

identity (strict matching of amino acids in corresponding positions) and similarity (functional analogy of different amino acids in corresponding positions) percentages were taken into account.

In addition, the whole BAC clone Psa-B-Cam 446P4 insertion was compared to the LYK region of the *M. truncatula* genome ver. 4.0 (chr5:37,100,000–37,450,000) using BLAST. The beginning of *PsSym37* showed very strong similarity to *MtLYK3*, even in the upstream region containing the putative promoter (Figure S1A). Exon 2, exon 3, and exons 6–8 of *PsSym37* were similar to corresponding exons in both *MtLYK2* and *MtLYK3*, while exon 5 and exon 11 were similar only to exon 5 and exon 11 of *MtLYK3*, and exon 4 was similar only to exon 4 of *MtLYK2*.

This, together with phenotypic data obtained for mutants (Smit et al., 2007), corroborates the idea that *PsSym37* is the most likely ortholog of *MtLYK3*. For the remaining two pea genes, the similarity patterns were more complicated. In *PsK1*, exons 6–8 and exon 10 were similar to the corresponding exons in *MtLYK3* and *MtLYK2*, while exon 9 was similar only to exon 9 of *MtLYK2* (Figure S1B). Exons 6–9, 10, and 11 of *PsLykX* were similar to the uncharacterized transcribed sequence (Mtr.51442.1.S1_at at MtGEA v. 3.0; <https://mtgea.noble.org/v3/> He et al., 2009) in the LYK region of *M. truncatula* (see Figure S1C), annotated in Mt4.0 as *Medtr5g086080*. However, exons 6–8, exon 10, and exon 11 were also similar to respective exons in *MtLYK4*. It is important to note that the first exon showed significant uniformity (at least

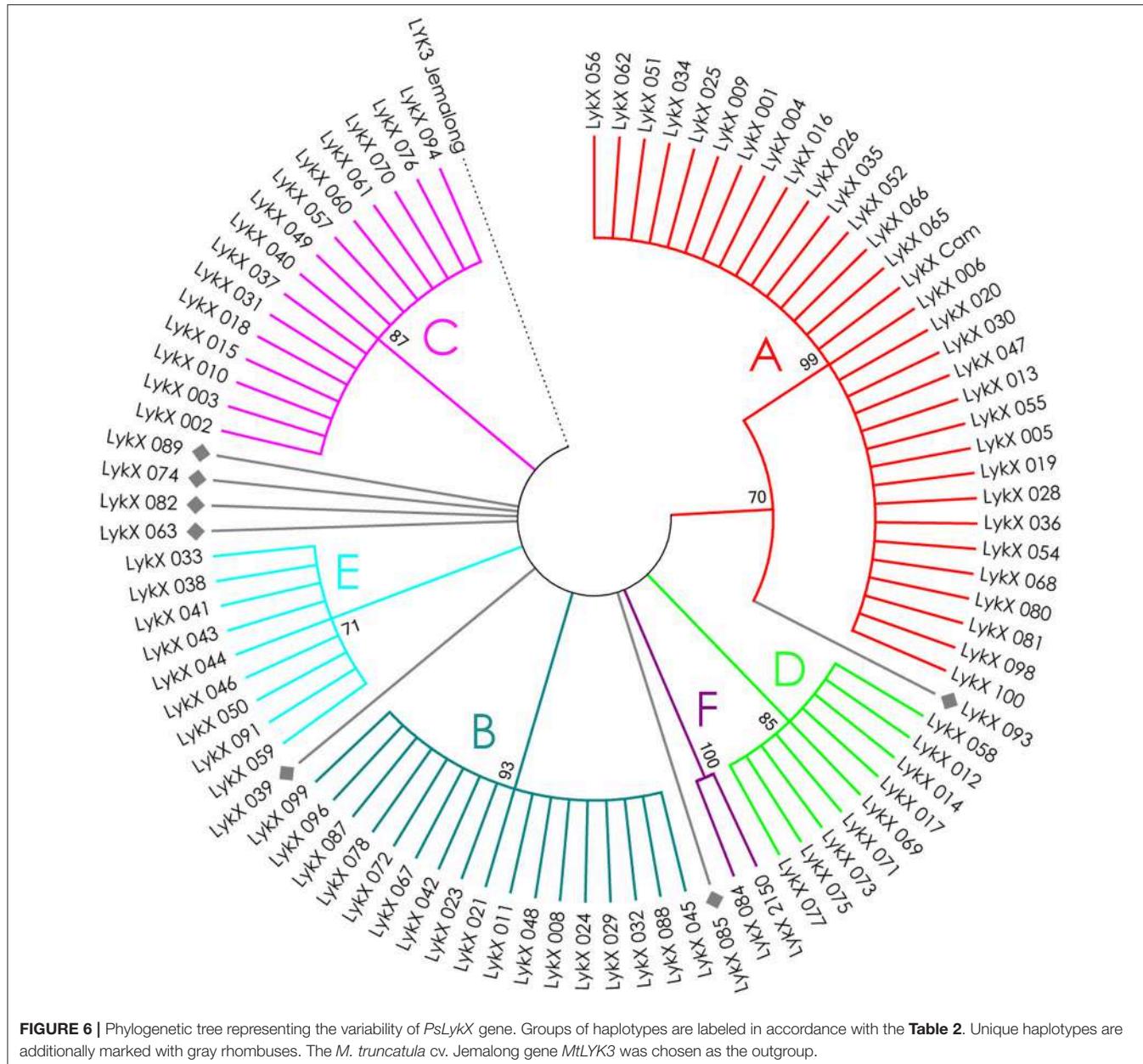


80%) both among the genes in one species (pea or *M. truncatula*) and between genes in different species (Table S3).

Our observations confirmed the previous finding that *PsSym37* is the most likely ortholog of *MtLYK3* (Zhukov et al., 2008). *PsLykX* was more similar to *MtLYK4* than to other LYK genes, whilst *PsK1* showed high similarity to both *MtLYK3* and *MtLYK2* (and also to *PsSym37*) at the putative protein level (Table 1). Taking into consideration the mosaic nature of genes in both the LYK and Sym2 regions, the complex evolutionary history of these genes, and the possibility of concerted evolution of paralogous genes, we are cautious to postulate orthology between pea and *M. truncatula*

LysM-RLK genes, aside from *PsSym37* and *MtLYK3*, which have conserved their basic function of Nod factor recognition and downstream signal transduction. Based on this, and since the first exons of all analyzed genes were >80% identical, we chose outgroups for polymorphism analyses as follows: *PsSym37* for all barrel medic genes, and *MtLYK3* for all pea genes.

The receptor parts of LysM-RLK contain three functionally important modules (LysM domains). Therefore, we assessed overall polymorphism in the first exon and polymorphism of separate parts of the first exon encoding LysM domains, as well as intermediate parts.



Codon-Based Tests for Selection Signatures

The ratio between synonymous and non-synonymous substitutions is often used to detect negative, or purifying, selection (i.e., selection against non-synonymous changes). The codon-based Z-test relies on a statistically significant difference between the rate of synonymous substitutions per synonymous site (dS) and the rate of non-synonymous substitutions per non-synonymous site (dN). We applied this test to the full sequences and to separate LysM module-encoding parts of the first exon of three pea and three *M. truncatula* genes. The results are shown in **Table 3**.

The overall sequences of the first exons of *PsK1* and *PsSym37* showed a clear departure from neutrality in favor

of purifying selection, while the first exon of *PsLykX* showed a tendency toward purifying selection (**Table 3**). Analysis of separate modules indicated that in both *PsK1* and *PsSym37*, the part encoding the LysM2 module (and, for *PsSym37*, also the part preceding LysM1 module) has undergone strong purifying selection. In *PsK1*, the intermediate part between LysM2 and LysM2-encoding regions showed signs of positive selection. There was also a tendency toward purifying selection in LysM3 and between LysM2 and LysM3 of *PsSym37*, which may indicate the importance of this gene. Indeed, mutations in *PsSym37* were shown to markedly decrease the number of nodules in symbiotic conditions (Zhukov et al., 2008). In addition, the site between LysM2 and LysM3 of *PsK1* appeared to be under

TABLE 2 | Haplotypes of pea LysM-RLKs.

	K1																	
	13	31	38	45	46	48	49	50	55	61	65	81	90	104	105	124	181	
A1	F	L	V	P	T	K	N	Y	I	D	S	F	E	T	T	S	N	
A2	F	L	I	A	T	K	N	Y	I	D	S	F	D	T	T	T	N	
A3	F	L	V	P	T	K	N	Y	I	D	S	F	E	T	A	T	I	
A4	F	L	V/I	P	T	K	N/S	Y	I	N	S	F/S	E	I	T	S/T	N	
A5	F	L	V	P	T	K	N	Y	I	D	R	F	D	T	T	S	N	
B1	L	I	V	S	N	T	T	F	L	D	R	F	E	T	T	S	N	
B2	L	I	V	S	N	T	T	F	L	D	R	F	E	T	T	S	N	
K1 005	F	I	I	A	T	K	S	Y	I	D	S	F	E	T	T	T	I	
K1 019	F	I	V	S	N	T	T	F	I	D	R	F	E	T	T	T	K	
K1 074	F	L	V	P	T	K	N	Y	I	D	S	F	E	T	T	T	N	
K1 085	L	I	V	S	N	T	T	F	L	D	R	F	E	T	A	S	N	
K1 1238	L	I	V	S	N	T	T	F	L	D	R	F	E	T	T	T	N	
K1 2150	F	L	V	P	T	K	N	Y	I	D	S	S	E	T	T	T	N	
																158Q:R	172S:F	
	Sym37																	
	43	45	56	60	83	131	141	181	184	205								
A0	E	S	L	S	I	F	A	S	K	I								
A1	E	L	V	S	I	F	A	S	K	I								
A2	E	S	L	S	I	F	A	S	K	I								
B0	Q	S	V	F	V	Y	V/A	R	K	V								
B1	Q	S	V	F	V	Y	V	R	N	V								
Sym37 005	Q	S	V	S	I	F	A	R	K	V	26V:L							
	LykX																	
	9	13	16	18	23	42	44	45	75	76	82	86	111	134	142	184		
A	L	L	V	F	K	L	Q	N	R	A	F	V	T	S	A	F		
B	F	V	V	F	K	K	Q	N	R	A	F	I	T	S	I	S		
C	L	L	F	S	K	K	Q	N	R	A	S	I	S	S	I	S		
D	L	V	F	F	K	K	Q	N	R	A	F	I	T	G	A	S		
E	L	L	F	S	K	K	Q	N	R	A	S	I	S	S	A	S		
F	L	L	V	F	K	K	R	Y	R	D	F	V	T	S	I	S		
LykX 039	L	L	V	F	K	K	Q	N	P	A	F	I	T	S	A	S	128V:F	
LykX 063	L	L	F	S	Q	K	Q	N	R	A	S	I	T	S	A	S		
LykX 074	L	V	F	S	K	K	Q	N	R	A	S	I	T	S	I	S	135H:D	
LykX 082	L	L	F	S	Q	K	Q	N	R	A	S	I	T	G	A	S	191I:M	
LykX 085	L	L	V	F	K	K	Q	N	P	A	F	I	T	S	I	S		
LykX 089	L	L	F	S	K	K	Q	N	R	A	S	I	S	S	I	S		
LykX 093	L	L	V	F	K	L	Q	S	R	A	F	I	T	S	A	S	5F:L 183A:T	

Numbers indicate the amino acid position in sequence of corresponding protein. Letters indicate groups of haplotypes, as shown in **Figures 4–6**. Colors indicate functional parts of corresponding protein (blue—signal peptide, gray—interdomain region, red—LysM1, green—LysM2, violet—LysM3; also see **Figure 2B**).

positive selection, suggesting that the structure of LysM modules is more important than that of intermediate regions in case of this particular LysM-RLK.

In *Medicago*, *MtLYK2* has undergone purifying selection, which has strongly affected LysM2; this was also the case for *PsK1* and *PsSym37*. It is important to note that the function of *MtLYK2*

TABLE 3 | Results of codon-based Z-test for examined regions of pea and *M. truncatula* LysM-RLK genes.

Gene	Part	Z
<i>PsK1</i>	Overall	-3.323** (Purifying)
	Upstream LysM1	-1.315
	LysM1	-1.392
	Int. LysM1-2	1.131
	LysM2	-2.028* (Purifying)
	Int. LysM2-3	2.372* (Positive)
	LysM3	-1.715
<i>PsSym37</i>	Overall	-4.117*** (Purifying)
	Upstream LysM1	-2.163* (Purifying)
	LysM1	-0.691
	Int. LysM1-2	-1.109
	LysM2	-2.263* (Purifying)
	Int. LysM2-3	-1.903
	LysM3	-1.759
<i>PsLykX</i>	Overall	-1.827
	Upstream LysM1	-0.141
	LysM1	0.178
	Int. LysM1-2	-0.973
	LysM2	-1.083
	Int. LysM2-3	-
	LysM3	-1.619
<i>MtLYK2</i>	Overall	-3.306** (Purifying)
	Upstream LysM1	-1.183
	LysM1	-1.487
	Int. LysM1-2	-1.098
	LysM2	-2.025* (Purifying)
	Int. LysM2-3	-0.39
	LysM3	-1.934
<i>MtLYK3</i>	Overall	-0.389
	Upstream LysM1	-0.912
	LysM1	1.632
	Int. LysM1-2	-
	LysM2	-0.895
	Int. LysM2-3	-0.049
	LysM3	-0.239
<i>MtLYK4</i>	Overall	0.617
	Upstream LysM1	1.312
	LysM1	1.904
	Int. LysM1-2	-
	LysM2	-0.33
	Int. LysM2-3	1.286
	LysM3	-0.587

Values of Z-statistic showing the statistically significant difference from the neutral model (meaning that $dN \neq dS$) are given in bold, with asterisks marking the significance level according to P-value: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. The direction of the selection (purifying or positive) is given in brackets.

is unknown; however, its similarity to *MtLYK3* and the results of our analyses suggest that *MtLYK2* may participate in the same symbiotic signal cascade as *MtLYK3*. However, *MtLYK3* clearly demonstrates neutral evolution, even though it is essential for Nod factor perception (Smit et al., 2007). A possible explanation is that *MtLYK3* has a low basic level of polymorphism (Table 4), which results in a statistically insignificant difference between dN and dS . *MtLYK4* has undergone neutral evolution, like *MtLYK3* and pea *PsLykX*. However, the neutral evolution signal may be a result of interference of two oppositely directed signals implying positive or purifying selection in different lineages in the dataset.

We also used the McDonald–Kreitman test (MK-test) as a statistical test of synonymous and non-synonymous changes. This test compares genes in two related species. The MK test did not detect departure from a neutral model, with the exception of the LysM1 region for the gene pair *PsLykX* and *MtLYK4* ($P = 0.0492$).

SNP-Based Tests

We used several molecular evolution tests to analyze the distribution of SNP sites across the dataset, regardless of whether they represented synonymous or non-synonymous changes. These tests can identify either balancing selection (in favor of two or more alleles) or positive selection (in favor of one allele) and distinguish these types of selection from the neutral evolution model. First, nucleotide diversity (π_i or π) was assessed in all sequences of the first exons and their separate parts, considering changes in synonymous and non-synonymous sites. In general, the rate of π was higher in pea genes than in *M. truncatula* ones (Table 4), even though the *M. truncatula* sequence dataset was larger than the pea dataset. In addition, haplotype diversity was higher for *M. truncatula* genes than for pea ones, possibly reflecting the prevalence of rare nucleotide variants (singletons) in the *Medicago* sequence dataset. The π_a/π_s ratios for LysM1 in *MtLYK3* and *MtLYK4* were very high (4.1 and 8.7, respectively), indicating an excess of non-synonymous substitutions. This result indicated that natural selection has tended to support new variants of LysM1 in these genes.

All the sequence sets were subjected to several neutrality tests (see section Materials and Methods, Table 5). In addition to analyses of the whole sequences and parts corresponding to LysM modules, a sliding window approach was used to visualize the obtained criteria values (Figures 7, 8).

In *PsSym37*, the value of Tajima's D suggested a departure from the neutral model for the LysM1 domain, its positive value indicating balancing selection (in favor of several different variants, namely haplotypes A and B; Table 2). No deviations from neutral evolution were identified in other regions, except for a weak signal of positive selection upstream of LysM1 (FW-H criterion = -2.34, $P = 0.046$).

In *PsK1*, the Tajima's D value indicated balancing selection in LysM1, which appeared to be the most polymorphic part of this gene, and the region responsible for the segregation of two main haplotype groups A and B (Table 2). Interestingly, for the region between LysM2 and LysM3, Tajima's D value fell

TABLE 4 | Nucleotide diversity in examined regions of pea and *M. truncatula* LysM-RLK genes.

Gene	Part	Length, bp	Segr. sites	HHe	$\pi^* \times 10^3$	$\pi_s^* \times 10^3$	$\pi_a^* \times 10^3$	π_a/π_s
<i>PsK1</i>	Overall	633	56	0.874	21.57	49.11	13.64	0.278
	Upstream LysM1	105	11	0.748	19.20	46.52	11.39	0.245
	LysM1	159	19	0.815	42.04	71.36	33.30	0.467
	Int. LysM1-2	42	2	0.346	8.33	0.00	10.40	–
	LysM2	141	11	0.830	22.71	68.46	9.16	0.134
	Int. LysM2-3	54	4	0.058	1.45	0.00	1.87	–
	LysM3	126	8	0.874	10.33	37.37	2.76	0.074
<i>PsSym37</i>	Overall	633	39	0.862	18.39	52.62	8.46	0.161
	Upstream LysM1	105	6	0.788	11.14	53.13	0.27	0.005
	LysM1	159	9	0.777	20.30	29.28	17.66	0.603
	Int. LysM1-2	42	2	0.320	7.81	37.72	0.00	0.000
	LysM2	141	10	0.675	19.39	59.87	6.72	0.112
	Int. LysM2-3	54	4	0.455	26.30	104.13	3.96	0.038
	LysM3	126	8	0.483	21.88	59.32	10.98	0.185
<i>PsLykX</i>	Overall	636	42	0.841	15.60	29.52	11.65	0.395
	Upstream LysM1	105	10	0.753	22.48	24.36	21.97	0.902
	LysM1	165	12	0.709	14.94	13.02	15.50	1.190
	Int. LysM1-2	42	1	0.313	7.44	37.51	0.00	0.000
	LysM2	141	10	0.802	20.70	42.42	14.06	0.331
	Int. LysM2-3	54	0	0.000	0.00	0.00	0.00	–
	LysM3	126	9	0.780	15.32	52.02	5.11	0.098
<i>MtLYK2</i>	Overall	639	65	0.915	12.74	30.38	7.69	0.253
	Upstream LysM1	105	9	0.439	14.25	28.52	10.46	0.367
	LysM1	165	16	0.729	14.16	27.32	10.49	0.384
	Int. LysM1-2	42	1	0.051	1.22	5.70	0.00	0.000
	LysM2	141	21	0.329	12.80	27.97	8.24	0.295
	Int. LysM2-3	54	3	0.163	4.24	5.99	3.77	0.629
	LysM3	126	15	0.816	17.50	58.75	5.69	0.097
<i>MtLYK3</i>	Overall	636	63	0.976	11.46	13.41	10.89	0.812
	Upstream LysM1	105	9	0.447	6.20	13.04	4.25	0.326
	LysM1	162	28	0.899	17.73	5.17	21.45	4.149
	Int. LysM1-2	42	0	0.000	0.00	0.00	0.00	–
	LysM2	141	9	0.744	22.13	40.14	16.84	0.420
	Int. LysM2-3	54	5	0.089	2.18	2.28	2.15	0.943
	LysM3	126	11	0.281	3.55	4.15	3.38	0.814
<i>MtLYK4</i>	Overall	639	47	0.971	12.21	12.26	12.19	0.994
	Upstream LysM1	105	5	0.391	5.71	2.57	6.61	2.572
	LysM1	162	13	0.562	10.91	1.58	13.78	8.722
	Int. LysM1-2	42	0	0.000	0.00	0.00	0.00	–
	LysM2	141	12	0.712	22.40	35.32	18.60	0.527
	Int. LysM2-3	54	3	0.564	12.23	0.00	15.57	–
	LysM3	126	13	0.697	10.30	18.61	8.01	0.430

π_a/π_s ratio higher than 1 is indicated in bold.

below the critical value, suggesting rapid spreading of a particular allelic state due to positive selection (selective sweep), or due to a recent population expansion. The FW-H value did not exceed the critical value, suggesting that population expansion

has shaped the DNA sequence in this region. Thus, the low Tajima's D value should be attributed to genetic drift rather than to a selective advantage of a particular allelic state of *PsK1*. In other regions, no departures from the neutral evolution model

TABLE 5 | Results of the molecular evolution tests applied to the examined regions of pea and *M. truncatula* LysM-RLK genes.

Gene	Part	Tajima's D	Fu's Fs	FL-D	FL-F	FW-H	nH	EW
PsK1	Overall	0.649	1.664	-1.133	-0.577	-7.350	-1.904*	0.137
	Upstream LysM1	-0.123	-2.087	-0.208	-0.309	-3.743*	-2.619*	0.283
	LysM1	1.902*	3.682	1.132	1.600*	-4.186	-2.001*	0.209
	Int. LysM1-2	-0.139	0.006	-1.091	-0.934	0.264	0.500	0.658
	LysM2	1.334	-1.659	0.673	0.958	0.370	0.239	0.179
	Int. LysM2-3	-1.775*	-4.971	-3.959***	-3.836***	0.078	0.096	0.942**
PsSym37	LysM3	-0.619	-1.815	-2.380*	-2.046*	0.019	-0.235	0.422
	Overall	1.384	5.805	0.439	1.002	-6.786	-1.676	0.153
	Upstream LysM1	-0.030	-0.786	0.179	0.130	-2.342*	-2.161*	0.221
	LysM1	2.070*	3.403	0.474	1.196	-0.225	-0.169	0.237
	Int. LysM1-2	-0.268	-0.185	0.496	0.583	0.226	0.626	0.720
	LysM2	0.684	1.671	-0.557	-0.138	-0.769	-0.717	0.337
	Int. LysM2-3	1.613	4.028	0.956	1.374	-0.796	-0.966	0.550
PsLykX	LysM3	1.398	3.516	0.474	0.860	-2.880	-2.905*	0.541
	Overall	0.533	2.817	-1.478	-0.828	-9.689	-1.865*	0.169
	Upstream LysM1	0.146	0.234	-1.159	-0.752	0.909	0.650	0.270
	LysM1	0.107	1.213	-2.161*	-1.682	-2.954	-0.687	0.299
	Int. LysM1-2	0.683	1.263	0.496	0.643	0.239	0.662	0.691
	LysM2	1.241	0.595	0.474	0.741	-3.972*	-2.981*	0.297
	Int. LysM2-3	-	-	-	-	-	-	-
MtLYK2	LysM3	0.217	-0.437	-0.930	-0.632	-3.912*	-2.690*	0.229
	Overall	-1.064	-45.063***	0.119	-0.489	-20.660**	-4.385**	0.097***
	Upstream LysM1	-0.263	-10.208***	0.543	0.304	-2.984*	-2.452*	0.565***
	LysM1	-0.613	-5.803*	-1.354	-1.311	-0.491	-0.496	0.646**
	Int. LysM1-2	-0.811	-1.169	0.481	0.107	-1.862	-5.283***	0.949
	LysM2	-1.556	-2.436	-0.216	-0.929	-14.759***	-5.614***	0.674***
	Int. LysM2-3	-1.040	-0.762	0.815	0.264	0.211	0.317	0.839
MtLYK3	LysM3	-0.593	-8.523**	0.911	0.526	-0.774	-0.415	0.197
	Overall	-0.930	-155.381***	0.641	-0.032	-26.204**	4.255**	0.046***
	Upstream LysM1	-1.278	-4.046	-0.594	-1.092	0.405	0.563	0.685
	LysM1	-1.080	-22.651***	0.011	-0.436	-11.986**	-4.468**	0.113
	Int. LysM1-2	-	-	-	-	-	-	-
	LysM2	2.401*	-1.400	1.276	2.017*	-3.713	-3.157*	0.284
	Int. LysM2-3	-1.629	-7.577***	0.976	0.108	-5.802***	-6.069***	0.913**
MtLYK4	LysM3	-1.788*	-7.629**	0.430	0.430	-5.183***	-1.319	0.720*
	Overall	-0.100	-145.577***	0.656	0.522	-15.296*	-3.300**	0.040***
	Upstream LysM1	-0.576	-3.467	0.888	0.552	-1.050	-1.439	0.665*
	LysM1	-0.508	-11.353**	0.827	0.385	-7.239**	-3.520**	0.450***
	Int. LysM1-2	-	-	-	-	-	-	-
	LysM2	1.315	0.502	-0.090	0.630	-2.136	-1.391	0.302
	Int. LysM2-3	0.462	-0.253	0.637	0.901	0.098	0.325	0.449
	LysM3	-1.033	-6.292*	-0.230	-0.562	-5.186**	-3.271*	0.343

In case of statistically significant deviation from the neutral model, the values of criteria are given in bold and marked with asterisks accordingly to the P-value: *P < 0.05; **P < 0.01; ***P < 0.001.

were identified, except for a signal of positive selection upstream of LysM1 (FW-H criterion = -3.74, P = 0.018), similar to PsSym37.

PsLykX, like *M. truncatula* MtLYK3 and MtLYK4, also showed a π_a/π_s ratio >1 in the LysM1-encoding region (Table 4). Remarkably, a clear signal of positive selection was found in

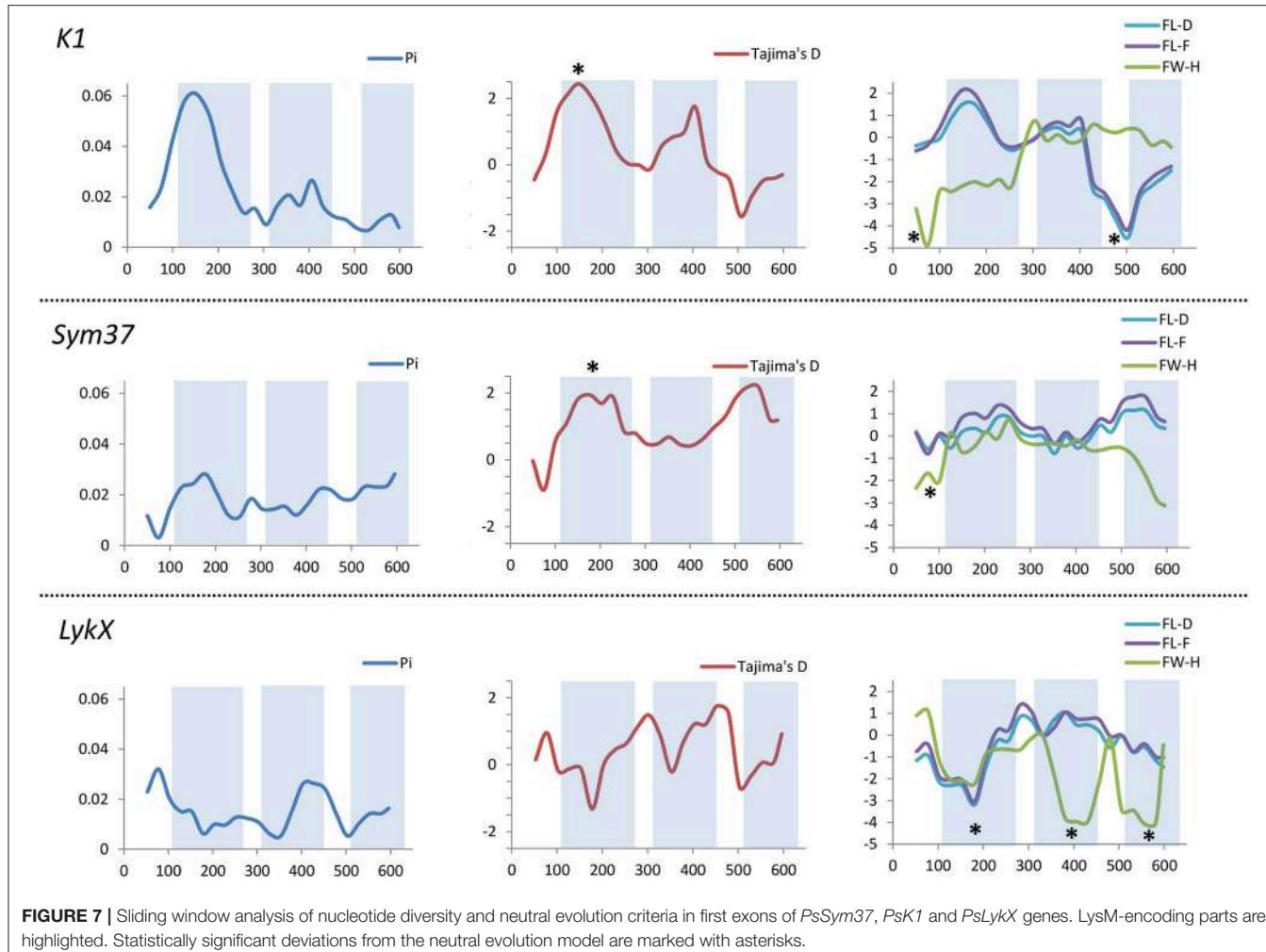


FIGURE 7 | Sliding window analysis of nucleotide diversity and neutral evolution criteria in first exons of *PsSym37*, *PsK1* and *PsLykX* genes. LysM-encoding parts are highlighted. Statistically significant deviations from the neutral evolution model are marked with asterisks.

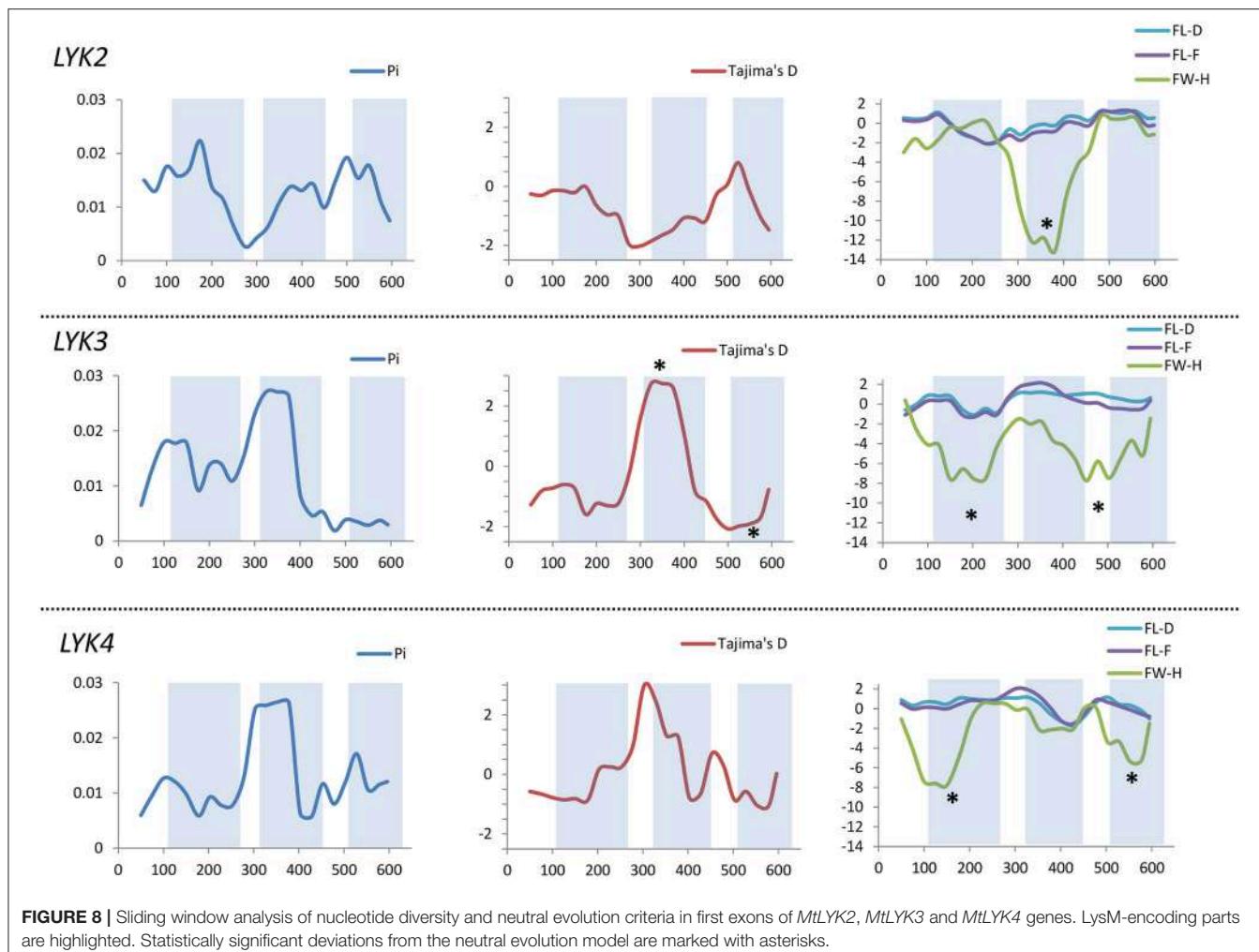
both the LysM2 and LysM3 domains (FW-H criterion below critical values; **Figure 7**), possibly indicating the importance of these parts for binding the Nod factor. Using a sliding window approach, declining values of modified Fu and Li's D and F criteria were detected at the LysM1 domain. However, in the absence of signals from powerful criteria (Fay and Wu's H), this was not considered as a footprint of positive selection. Instead, it could be attributed to a population expansion or a non-random selection of the samples in the sequence dataset.

In *Medicago*, all three examined genes had significantly low Fay and Wu's H values, indicating strong positive selection at the first exons (corresponding to the receptor parts of the LysM-RLKs). The strongest pressure was at LysM2 of MtLYK2 and at LysM1 of MtLYK3. In contrast to the pea sequences, the *M. truncatula* gene sequences showed significantly low values of Fu's Fs (**Table 5**), which is sensitive to population expansion. This implied that recent demographic processes have contributed to the molecular evolution of *M. truncatula* LysM-RLK genes. Interestingly, the similar pattern of nucleotide variant distribution in MtLYK3 and MtLYK4 genes led to comparable Tajima's D profiles in the sliding window analysis.

The remarkable increase in Tajima's D at LysM2-encoding region (**Figure 8**) implied that there has been balancing selection at these sites. Meanwhile, the low value of Tajima's D at LysM3 of MtLYK3 was indicative of recent positive selection pressure at this domain. The FW-H and EW values also indicated that all three examined *M. truncatula* genes have undergone positive selection at the whole first exon, especially the LysM module-encoding parts. Recently, De Mita et al. (2014) pinpointed positive selection signatures at positions 43Q, 45R, and 77G in the MtLYK3 protein sequence. These results are consistent with the positive selection signatures found using other statistical models of nucleotide sequence diversity in the present study (sliding window midpoints 150 and 225; **Figure 8**).

DISCUSSION

Both natural and artificial selection tend to affect sequence variability at selected genomic loci and at neutral loci linked to them. These selection signatures are widely used to identify loci subjected to selection, thereby giving researchers an insight into the evolutionary process, as well as providing functional



information about genes or genomic regions (Kreitman, 2000; Nielsen, 2005).

The most crucial step in the symbiosis between legume plants and nitrogen-fixing bacteria (rhizobia) is the initial mutual recognition of the macro- and micro-symbionts. The importance of this step should not be underestimated, as partnerships with ineffective bacteria can be lethal for plants growing in harsh environments. Thus, plants have developed multicomponent receptor systems for the bacterial signal to identify the most suitable and beneficial partner. As bacteria tend to evolve rapidly, the genes involved in their recognition should be “hot spots” under constant selective pressure. Therefore, one can expect to find footprints of selection in the nucleotide sequences of genes encoding legume symbiotic receptors. In this work, we studied the polymorphism of the first exon of three LysM-RLK genes in large set of pea genotypes, and compared them with their homologs in *M. truncatula*. Although all these genes appeared to be quite similar to each other, our data suggest that they differ significantly in their evolutionary patterns, consistent with the symbiotic partner recognition unique to each legume species.

A single amino acid replacement in the receptor part of the LysM-RLK can change its recognition of the Nod factor (as shown for *L. japonicus* LysM-RLK NFR5, Radutoiu et al., 2007). Therefore, we expected to find a relatively low level of polymorphism in the first exon (corresponding to the receptor domain) of the three LysM-RLK genes from the *Sym2* region. However, several “legal” haplotypes were detected for all three pea genes, implying that the structure of receptors may be variable. Such variability is possibly advantageous. Previously, Li et al. studied the polymorphism of the LysM-RLK gene *PsSym37* in 10 pea genotypes, and detected two groups of haplotypes (at the protein level) correlated with the ability of the plant to perceive Nod factors containing a C18:1 rather than a C18:4 (number of carbon atoms:number of unsaturated bonds) acyl group at the non-reducing end of molecule (Li et al., 2011). Despite our sample set being more than 10 times larger, we were unable to identify any new *PsSym37* haplotypes in addition to those already discovered, which may indicate a certain fundamental base of this division among pea genotypes. However, it is hard to tell exactly how, where, and why these two symbiotic groups may have arisen, as the history of pea

domestication, spreading, and crossing is long and unclear. Although our results indicate generally neutral evolution of *PsSym37*, the overall positive Tajima's D value indicates that the sequence dataset may be not uniform. Independent analyses of two lineages (haplotype groups A and B) may be required to identify footprints of positive selection.

Interestingly, *PsK1* shows a pattern of polymorphism similar to that of *PsSym37*, with two basic groups of haplotypes. However, the haplotypes do not correlate with any symbiotic feature of pea known so far. The function of *PsK1* is also unclear, although preliminary data (MSc. Anna Kirienko, Dr. Elena Dolgikh, personal communication) indicate that it may function prior to *PsSym37*, as suggested by phenotypic analyses of mutants. Since *PsK1* is similar to the *M. truncatula* gene *MtLYK2*, whose function is also unknown, it is possible that these proteins participate in the initial capture and recognition of the Nod factor, together with the "early" symbiotic LysM-RLKs *PsSym10* (in pea) and *MtNFP* (in barrel medic). These kinases are characterized by an inactive kinase domain (lacking the activation loop; Arrighi et al., 2006), while this domain appears to be fully functional in *PsK1* and *MtLYK2*. Therefore, the observed variability of the receptor part of *PsK1* may be essential for recognition of diverse symbiotic bacteria.

The newly discovered pea LysM-RLK gene *PsLykX* is unique in that it has six equivalent major haplotype groups (each containing several genotypes) and about the same number of unique haplotypes (corresponding to a single genotype). This suggests that our sample of 90 *PsLykX* sequences does not cover all the possible "legal" allelic variants of this gene. Another important fact is the correlation of one of the identified *PsLykX* haplotypes with the "Afghan" phenotype. This makes *PsLykX* a promising candidate for *PsSym2*, and, among other things, may indicate that other *PsLykX* haplotypes are also related to nuances of perception of bacterial signals. The function of *MtLYK4*, the closest homologous LYK gene of *PsLykX*, is unclear, despite its significant similarity to *MtLYK3* both in its sequence and in its discernable selection signatures. In roots of *MtLYK4*-knockdown plants, infection thread morphology was affected but nodulation occurred normally (Limpens et al., 2003). Furthermore, evidence of positive selection in the LysM2 and LysM3 modules suggests that *MtLYK4* has recently acquired a new function, probably associated with the expansion of the *M. truncatula* population. Remarkably, the kinase domains of both putative *PsLykX* and *MtLYK4* proteins lack the YAQ motif, which is inherent to symbiotic kinases in *L. japonicus* and is absent from non-symbiotic ones (Nakagawa et al., 2011). On the other hand, the receptor system in legumes with indeterminate nodules (e.g., pea and barrel medic) appears to be more complicated than that of legumes with determinate nodules, such as, *L. japonicus* (Ardourel et al., 1994). The presence of *PsLykX* transcripts in pea nodules, together with the *lyk4* RNA-interference phenotype in *M. truncatula* (disruption of infection process; Limpens et al., 2003), provide evidence for the participation of YAQ-lacking LysM-RLKs in indeterminate nodule development.

For all three genes, the haplotype distribution was not uniform among our samples, resulting in strictly correlating haplotype groups: for instance, group E of *PsLykX* (see Figure 6, Table 2) was found exclusively with group A2 of *PsK1* and B0 of

PsSym37. However, considering that: (i) these genes are clustered within approximately 20 kb; (ii) the genotype sampling was not random; and (iii) the VIR collection does not represent a single population of pea, this non-uniformity can be explained by factors other than the evolutionary advantage of a particular haplotype combination, e.g., genetic linkage or founder effect. On the other hand, the fact that we observed so many haplotype combinations despite the close proximity of these genes suggests that there may be a selective advantage of symbiotic LysM-RLK haplotype shuffling in pea.

Despite the fact that the LYK region of *M. truncatula* and the *Sym2* region of pea are clearly syntenic (Gualtieri et al., 2002), it is nearly impossible to determine the orthologous relationships between genes contained in those regions. As shown in this work, LysM-RLK genes from these genomic regions are characterized by mutual mosaic similarity, with different parts of different genes being similar to each other both within a species and between species. Since the LYK and *Sym2* regions are represented primarily by gene clusters that apparently originated via multiplication of ancestral LysM-RLK gene(s), the genes in those regions have undergone concerted evolution (Liao, 1999) implying genetic conversion. This has led to the diversification of LysM-RLK gene clusters in different legume species. In other words, each cluster represents an "evolution crucible" unique to each species, where genes and gene parts have been shuffled, combined, or broken apart. Thus, divergent evolution may lead to the loss of earlier versions of genes in one species, and the retention of genes and acquisition of unique functions in another species through neofunctionalization. This is a promising explanation for the unique *PsSym2^A* phenotype that is not observed in any other legume species.

CONCLUSION

As a complex multi-stage process, the symbiosis of legumes with rhizobia is not yet fully understood. There is still much to learn about the functions of many genes and the population genetics of the symbiosis. In *M. truncatula* (De Mita et al., 2011; Ho-Huu et al., 2012) and other legumes with well-studied genomes like *L. japonicus*, chickpea (*Cicer arietinum* L.), and soybean [*Glycine max* (L.) Merr.], surveys of large ecotype collections and association genetics analyses have been widely used to identify loci of interest (Kale et al., 2015; Li et al., 2017; Plekhanova et al., 2017). However, knowledge about the diversity of certain genes, especially those related to symbiosis, is still insufficient. The aim of this study was to help to fill this knowledge gap. Accordingly, we evaluated the genetic diversity of the pea LysM-RLK-encoding genes from the *Sym2* region and corresponding barrel medic genes from the LYK region, detected the signatures of positive, balancing and negative selection in the gene sequences, and compared the patterns of selection pressure affecting particular modules in the sequences.

The cultivated pea accessions analyzed in this work have been previously characterized to determine their productivity under symbiotic conditions (Borisov et al., 2002). Therefore, information on symbiotic LysM-RLK gene polymorphisms may be useful for studies on associations between gene variants and the formation of highly effective symbioses with beneficial

microorganisms. Considering the importance and economic value of pea in modern agriculture, the value of such research activities should not be underestimated.

Analyses of the polymorphism of crucial symbiotic genes and the identification of selection signatures have allowed us to formulate new hypotheses about their roles in symbiosis, which will be tested experimentally in the near future. On the basis of our results, we propose that the newly discovered pea gene *PsLykX* could in fact be *PsSym2*, which is responsible for increased selectivity toward a symbiotic partner in plants with the characteristic phenotype of cv. Afghanistan. However, more detailed analyses are required to confirm that *PsLykX* is the elusive *PsSym2* gene. To confirm our hypothesis, the polymorphism of *PsLykX* should be studied in a larger sample of genotypes with high selectivity toward microsymbionts. In addition, it is important to analyze the phenotypes of lines with mutations in the *PsLykX* gene. The pea TILLING mutants collection (Dalmais et al., 2008) provides opportunities for such analyses. Complementation of the *PsLykX* mutations by either “European” or “Afghan” alleles resulting in nodulation or non-nodulation phenotypes would confirm the characteristics of particular groups of pea genotypes.

AUTHOR CONTRIBUTIONS

VZ, IT, and LL designed and conceived the study; AS, AA, and AZ conducted experiments; AS, AZ, AA, VZ, and IT analyzed the

REFERENCES

- Afonin, A., Sulima, A., Zhernakov, A., and Zhukov, V. (2017). Draft genome of the strain RCAM1026 *Rhizobium leguminosarum* bv. *viciae*. *Genomics Data* 11, 85–86. doi: 10.1016/j.gdata.2016.12.003
- Alves-Carvalho, S., Aubert, G., Carrère, S., Cruaud, C., Brochot, A., Jacquin, F., et al. (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. *Plant J.* 84, 1–19. doi: 10.1111/tpj.12967
- Ardourel, M., Demont, N., Debellb, F., Maillet, F., Billy, F., De Promb, J., 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. doi: 10.1105/tpc.6.10.1357
- Arrighi, J.-F., Barre, A., Amor, B., Ben, B. A., Soriano, L. C., Mirabella, R., et al. (2006). The *Medicago truncatula* lysis motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol.* 142, 265–279. doi: 10.1104/pp.106.084657
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Belimov, A. A., Puhalsky, I. V., Safranova, V. I., Shaposhnikov, A. I., Vishnyakova, M. A., Semenova, E. V., et al. (2015). Role of plant genotype and soil conditions in symbiotic plant-microbe interactions for adaptation of plants to cadmium-polluted soils. *Water Air Soil Pollut.* 226:264. doi: 10.1007/s11270-015-2537-9
- Borisov, A. Y., Danilova, T. N., Koroleva, T. A., Naumkina, T. S., Pavlova, Z. B., Pinaev, A. G., et al. (2004). Pea (*Pisum sativum* L.) regulatory genes controlling development of nitrogen-fixing nodule and arbuscular mycorrhiza: fundamentals and application. *Biologia* 59, 137–144.
- Borisov, A. Y., Tsyanov, V. E., Shtark, O. Y., Jacobi, L. M., Naumkina, T. S., Serdyuk, V. P., et al. (2002). “Pea: symbiotic effectiveness,” in *Catalogue of the World Collection of VIR*, Vol. 728 (St.Petersburg: VIR Press), 29.
- Dalmais, M., Schmidt, J., Le Signor, C., Moussy, F., Burstin, J., Savois, V., et al. (2008). UTILLdb, a *Pisum sativum* *in silico* forward and reverse genetics tool. *Genome Biol.* 9:R43. doi: 10.1186/gb-2008-9-2-r43
- Davis, E. O., Evans, I. J., and Johnston, A. W. B. (1988). Identification of *nodX*, a gene that allows *Rhizobium leguminosarum* biovar *viciae* strain TOM to nodulate Afghanistan peas. *Mol. Gen. Genet.* 212, 531–535. doi: 10.1007/BF00330860
- Borisov, A., Vasil'chikov, A., Voroshilova, V., Danilova, T., Zhernakov, A., Zhukov, V., et al. (2007). Regulatory genes of garden pea (*Pisum sativum* L.) controlling the development of nitrogen-fixing nodules and arbuscular mycorrhiza: a review of basic and applied aspects. *Prikl. Biokhim. Mikrobiol.* 43, 265–271. doi: 10.1134/S0003683807030027
- Broghammer, A., Krusell, L., Blaise, M., Sauer, J., Sullivan, J. T., Maolanon, N., et al. (2012). Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13859–13864. doi: 10.1073/pnas.1205171109
- Buendia, L., Wang, T., Girardin, A., and Lefebvre, B. (2016). The LysM receptor-like kinase SLYK10 regulates the arbuscular mycorrhizal symbiosis in tomato. *New Phytol.* 210, 184–195. doi: 10.1111/nph.13753
- Buendía-Clavería, A. M., Moussaid, A., Ollero, F. J., Vinardell, J. M., Torres, A., Moreno, J., et al. (2003). A purL mutant of *Sinorhizobium fredii* HH103 is symbiotically defective and altered in its lipopolysaccharide. *Microbiology* 149, 1807–1818. doi: 10.1099/mic.0.26099-0
- Buist, G., Steen, A., Kok, J., and Kuipers, O. P. (2008). LysM, a widely distributed protein motif for binding to (peptido) glycans. *Mol. Microbiol.* 68, 838–847. doi: 10.1111/j.1365-2958.2008.06211.x
- Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G., and Parkhill, J. (2005). ACT: the Artemis comparison tool. *Bioinformatics* 21, 3422–3423. doi: 10.1093/bioinformatics/bti553
- Chetkova, S. A., and Tikhonovich, I. A. (1986). Isolation and study of *Rhizobium leguminosarum* strains effective on peas from Afghanistan. *Mikrobiologiya* 55, 143–147.
- Frontiers in Plant Science | www.frontiersin.org

data. AS, VZ, and AA drafted the manuscript, IT and LL critically revised the manuscript. All authors read and approved the final manuscript.

FUNDING

The work on pea genes sequencing and analysis was funded by the Russian Science Foundation (grant # 16-16-00118); the work of *M. truncatula* genes analysis was funded by the Russian Foundation for Basic Research (grant # 15-29-02737 ofi_m).

ACKNOWLEDGMENTS

The authors thank Dr. Hélène Bergès and colleagues (INRA-CNRS, Toulouse, France) for analysis of pea BAC library, Dr. Eugene Andronov and Dr. Alexey Borisov (ARRIAM, Saint Petersburg, Russia) for fruitful discussions and invaluable advice, and MSc. Tamara Rychagova and MSc. Tatiana Aksanova (ARRIAM, Saint Petersburg, Russia) for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2017.01957/full#supplementary-material>

- De Mita, S., Chantret, N., Lordin, K., Ronfort, J., and Bataillon, T. (2011). Molecular adaptation in flowering and symbiotic recognition pathways: insights from patterns of polymorphism in the legume *Medicago truncatula*. *BMC Evol. Biol.* 11:229. doi: 10.1186/1471-2148-11-229
- De Mita, S., Streng, A., Bisseling, T., and Geurts, R. (2014). Evolution of a symbiotic receptor through gene duplications in the legume–rhizobium mutualism. *New Phytol.* 201, 961–972. doi: 10.1111/nph.12549
- Denarie, J., Debelle, F., and Prome, J.-C. (1996). Rhizobium lipochitoooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. *Annu. Rev. Biochem.* 65, 503–535. doi: 10.1146/annurev.bi.65.070196.002443
- Fay, J. C., and Wu, C.-I. (2000). Hitchhiking under positive Darwinian selection. *Genetics* 155, 1405–1413. Available online at: <http://www.genetics.org/content/155/3/1405.long>
- Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.
- Gentzbittel, L., Andersen, S. U., Ben, C., Rickauer, M., Stougaard, J., and Young, N. D. (2015). Naturally occurring diversity helps to reveal genes of adaptive importance in legumes. *Front. Plant Sci.* 6:269. doi: 10.3389/fpls.2015.00269
- Gobbato, E. (2015). Recent developments in arbuscular mycorrhizal signaling. *Curr. Opin. Plant Biol.* 26, 1–7. doi: 10.1016/j.pbi.2015.05.006
- Govorov, L. I. (1928). The peas of Afghanistan. *Bull. Appl. Bot.* 19, 497–522.
- Gualtieri, G., Kulikova, O., Limpens, E., Kim, D.-J., Cook, D. R., Bisseling, T., et al. (2002). Microsynteny between pea and *Medicago truncatula* in the SYM2 region. *Plant Mol. Biol.* 50, 225–235. doi: 10.1023/A:1016085523752
- He, J., Benedito, V. A., Wang, M., Murray, J. D., Zhao, P. X., Tang, Y., et al. (2009). The *Medicago truncatula* gene expression atlas web server. *BMC Bioinformatics* 10:441. doi: 10.1186/1471-2105-10-441
- Ho-Huu, J., Ronfort, J., De Mita, S., Bataillon, T., Hochu, I., Weber, A., et al. (2012). Contrasted patterns of selective pressure in three recent paralogous gene pairs in the *Medicago* genus (L.). *BMC Evol. Biol.* 12:195. doi: 10.1186/1471-2148-12-195
- Jacobi, L. M., Kukalev, A. S., Ushakov, K. V., Tsyanov, V. E., Naumkina, T. S., Provorov, N. A., et al. (2000). Polymorphism of garden pea (*Pisum sativum* L.) for the efficiency of symbiosis with endomycorrhizal fungus *Glomus* sp. under the conditions of rhizobia inoculation. *Agric. Biol.* 3, 94–102. doi: 10.1080/09064710510029169
- Jing, R., Vershinin, A., Grzebyta, J., Shaw, P., Smýkal, P., Marshall, D., et al. (2010). The genetic diversity and evolution of field pea (*Pisum*) studied by high throughput retrotransposon based insertion polymorphism (RBIP) marker analysis. *BMC Evol. Biol.* 10:44. doi: 10.1186/1471-2148-10-44
- Jukes, T. H., and Cantor, C. R. (1969). Evolution of protein molecules. *Mamm. Protein Metab.* 3:132. doi: 10.1016/B978-1-4832-3211-9.50009-7
- Kale, S. M., Jaganathan, D., Ruperao, P., Chen, C., Punna, R., Kudapa, H., et al. (2015). Prioritization of candidate genes in “QTL-hotspot” region for drought tolerance in chickpea (*Cicer arietinum* L.). *Sci. Rep.* 5:15296. doi: 10.1038/srep15296
- Kawaharada, Y., Kelly, S., Nielsen, M. W., Hjuler, C. T., Gysel, K., Muszynski, A., et al. (2015). Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* 523, 308–312. doi: 10.1038/nature14611
- Kibbe, W. A. (2007). OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.* 35, W43–W46. doi: 10.1093/nar/gkm234
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Kozik, A., Heidstra, R., Horvath, B., Kulikova, O., Tikhonovich, I., Ellis, T. H. N., et al. (1995). Pea lines carrying *sym1* or *sym2* can be nodulated by *Rhizobium* strains containing *nodX*; *sym1* and *sym2* are allelic. *Plant Sci.* 108, 41–49. doi: 10.1016/0168-9452(95)04123-C
- Kozik, A., Matvienko, M., Scheres, B., Paruvangada, V. G., Bisseling, T., van Kammen, A., et al. (1996). The pea early nodulin gene PsENOD7 maps in the region of linkage group I containing *sym2* and leghaemoglobin. *Plant Mol. Biol.* 31, 149–156. doi: 10.1007/BF00020614
- Kreitman, M. (2000). Methods to detect selection in populations with applications to the human. *Annu. Rev. Genomics Hum. Genet.* 1, 539–559. doi: 10.1146/annurev.genom.1.1.539
- Li, R., Knox, M. R., Edwards, A., Hogg, B., Ellis, T. H. N., Wei, G., et al. (2011). Natural variation in host-specific nodulation of pea is associated with a haplotype of the SYM37 LysM-type receptor-like kinase. *Mol. Plant. Microbe Interact.* 24, 1396–1403. doi: 10.1094/MPMI-01-11-0004
- Li, Y., Ruperao, P., Batley, J., Edwards, D., Davidson, J., Hobson, K., et al. (2017). Genome analysis identified novel candidate genes for ascochyta blight resistance in chickpea using whole genome re-sequencing data. *Front. Plant Sci.* 8:359. doi: 10.3389/fpls.2017.00359
- Liao, D. (1999). Concerted evolution: molecular mechanism and biological implications. *Am. J. Hum. Genet.* 64, 24–30. doi: 10.1086/302221
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Lie, T. A. (1984). Host genes in *Pisum sativum* L. conferring resistance to European *Rhizobium leguminosarum* strains. *Plant Soil* 82, 415–425. doi: 10.1007/BF02184279
- Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* 302, 630–633. doi: 10.1126/science.1090074
- Lindgreen, S. (2012). AdapterRemoval: easy cleaning of next generation sequencing reads. *BMC Res. Notes* 5:337. doi: 10.1186/1756-0500-5-337
- Lohmann, G. V., Shimoda, Y., Nielsen, M. W., Jørgensen, F. G., Grossmann, C., Sandal, N., et al. (2010). Evolution and regulation of the *Lotus japonicus* LysM receptor gene family. *Mol. Plant. Microbe Interact.* 23, 510–521. doi: 10.1094/MPMI-23-4-0510
- Macas, J., Neumann, P., and Navrátilová, A. (2007). Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. *BMC Genomics* 8:427. doi: 10.1186/1471-2164-8-427
- Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., et al. (2003). A receptor kinase gene of the LysM type is involved in legumeperception of rhizobial signals. *Nature* 425, 637–640. doi: 10.1038/nature02045
- McDonald, J. H., and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351:652. doi: 10.1038/351652a0
- Mesnage, S., Dellaroche, M., Baxter, N. J., Rouget, J.-B., Dimitrov, J. D., Wang, N., et al. (2014). Molecular basis for bacterial peptidoglycan recognition by LysM domains. *Nat. Commun.* 5:4269. doi: 10.1038/ncomms5269
- Nakagawa, T., Kaku, H., Shimoda, Y., Sugiyama, A., Shimamura, M., Takanashi, K., et al. (2011). From defense to symbiosis: limited alterations in the kinase domain of LysM receptor-like kinases are crucial for evolution of legume–Rhizobium symbiosis. *Plant J.* 65, 169–180. doi: 10.1111/j.1365-313X.2010.04411.x
- Nei, M., and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3, 418–426.
- Nielsen, R. (2005). Molecular signatures of natural selection. *Annu. Rev. Genet.* 39, 197–218. doi: 10.1146/annurev.genet.39.073003.112420
- Oldroyd, G. E. D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* 11, 252–263. doi: 10.1038/nrmicro2990
- Oldroyd, G. E. D., Murray, J. D., Poole, P. S., and Downie, J. A. (2011). The rules of engagement in the legume–rhizobial symbiosis. *Annu. Rev. Genet.* 45, 119–144. doi: 10.1146/annurev-genet-110410-132549
- Ovtsova, O., Rademaker, G. J., Esser, E., Weinman, J., and Rolfe, B. G., Tikhonovich, I. A., et al. (1999). Comparison of characteristics of the nodX genes from various *Rhizobium leguminosarum* strains. *Mol. Plant. Microbe Interact.* 12, 252–258. doi: 10.1094/MPMI.1999.12.3.252
- Plekhanova, E., Vishnyakova, M. A., Bulyntsev, S., Chang, P. L., Carrasquilla-Garcia, N., Negash, K., et al. (2017). Genomic and phenotypic analysis of Vavilov's historic landraces reveals the impact of environment and genomic islands of agronomic traits. *Sci. Rep.* 7:4816. doi: 10.1038/s41598-017-05087-5
- Ponting, C. P., Aravind, L., Schultz, J., Bork, P., and Koonin, E. V. (1999). Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer. *J. Mol. Biol.* 289, 729–745. doi: 10.1006/jmbi.1999.2827
- Provorov, N. A., and Vorobyov, N. I. (2013). Macroevolution of symbiosis as self-organization of superspecies system controlled by natural selection. *Biol. Bull. Rev.* 3, 274–285. doi: 10.1134/S2079086413040063
- Radutoiu, S., Madsen, L. H., Madsen, E. B., Jurkiewicz, A., Fukai, E., Quistgaard, E. M. H., et al. (2007). LysM domains mediate lipochitin-oligosaccharide

- recognition and Nfr genes extend the symbiotic host range. *EMBO J.* 26, 3923–3935. doi: 10.1038/sj.emboj.7601826
- 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. doi: 10.1038/srep29733
- Razumovskaya, Z. G. (1937). Nodule formation in various pea cultivars. *Russ. J. Microbiol.* 6, 321–328.
- Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16, 276–277. doi: 10.1016/S0168-9525(00)02024-2
- Rogers, S. O., and Bendich, A. J. (1985). Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5, 69–76. doi: 10.1007/BF00020088
- 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. doi: 10.1104/pp.107.100495
- Smýkal, P., Aubert, G., Burstin, J., Coyne, C. J., Ellis, N. T. H., Flavell, A. J., et al. (2012). Pea (*Pisum sativum* L.) in the genomic era. *Agronomy* 2, 74–115. doi: 10.3390/agronomy2020074
- Stanton-Geddes, J., Paape, T., Epstein, B., Briskine, R., Yoder, J., Mudge, J., et al. (2013). Candidate genes and genetic architecture of symbiotic and agronomic traits revealed by whole-genome, sequence-based association genetics in *Medicago truncatula*. *PLoS ONE* 8:e65688. doi: 10.1371/journal.pone.0065688
- Sudheesh, S., Sawbridge, T. I., Cogan, N. O., Kennedy, P., Forster, J. W., and Kaur, S. (2015). De novo assembly and characterisation of the field pea transcriptome using RNA-Seq. *BMC Genomics* 16:611. doi: 10.1186/s12864-015-1815-7
- Suzaki, T., Yoro, E., and Kawaguchi, M. (2015). Leguminous plants: inventors of root Nodules to accommodate symbiotic bacteria. *Int. Rev. Cell Mol. Biol.* 316, 111–158. doi: 10.1016/bs.ircmb.2015.01.004
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Veiga, R. S. L., Faccio, A., Genre, A., Pieterse, C. M. J., Bonfante, P., and van der Heijden, M. G. A. (2013). Arbuscular mycorrhizal fungi reduce growth and infect roots of the non-host plant *Arabidopsis thaliana*. *Plant. Cell Environ.* 36, 1926–1937. doi: 10.1111/pce.12102
- Weeden, N. F. (2007). Genetic changes accompanying the domestication of *Pisum sativum*: is there a common genetic basis to the “domestication syndrome” for legumes? *Ann. Bot.* 100, 1017–1025. doi: 10.1093/aob/mcm122
- Werner, G. D. A., Cornwell, W. K., Cornelissen, J. H. C., and Kiers, E. T. (2015). Evolutionary signals of symbiotic persistence in the legume–rhizobia mutualism. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10262–10269. doi: 10.1073/pnas.1424030112
- Wheatley, R. E., and Sprent, J. I. (2010). Legume nodulation: a global perspective. *Exp. Agric.* 46:568. doi: 10.1017/S0014479710000554
- Young, N. D., and Udvardi, M. (2009). Translating *Medicago truncatula* genomics to crop legumes. *Curr. Opin. Plant Biol.* 12, 193–201. doi: 10.1016/j.pbi.2008.11.005
- Zeng, K., Fu, Y.-X., Shi, S., and Wu, C.-I. (2006). Statistical tests for detecting positive selection by utilizing high-frequency variants. *Genetics* 174, 1431–1439. doi: 10.1534/genetics.106.061432
- Zeng, K., Mano, S., Shi, S., and Wu, C.-I. (2007). Comparisons of site-and haplotype-frequency methods for detecting positive selection. *Mol. Biol. Evol.* 24, 1562–1574. doi: 10.1093/molbev/msm078
- Zhang, X., Dong, W., Sun, J., Feng, F., Deng, Y., He, Z., et al. (2015). The receptor kinase CERK1 has dual functions in symbiosis and immunity signalling. *Plant J.* 81, 258–267. doi: 10.1111/tpj.12723
- Zhukov, V. A., Shtark, O. Y., Nemankin, T. A., Kryukov, A. A., Borisov, A. Y., and Tikhonovich, I. A. (2016). Genetic mapping of pea (*Pisum sativum* L.) genes involved in symbiosis. *Agrobiology* 51, 593–601. doi: 10.1538/agrobiology.2016.5.593eng
- Zhukov, V. A., Zhernakov, A. I., Kulaea, O. A., Ershov, N. I., Borisov, A. Y., and Tikhonovich, I. A. (2015). *De novo* assembly of the pea (*Pisum sativum* L.) nodule transcriptome. *Int. J. Genomics* 2015:695947. doi: 10.1155/2015/695947
- Zhukov, V., Radutoiu, S., Madsen, L. H., Rychagova, T., Ovchinnikova, E., Borisov, A., et al. (2008). The Pea Sym37 receptor kinase gene controls infection-thread initiation and nodule development. *Mol. Plant Microbe Interact.* 21, 1600–1608. doi: 10.1094/MPMI-21-12-1600
- Zipfel, C., and Oldroyd, G. E. D. (2017). Plant signalling in symbiosis and immunity. *Nature* 543, 328–336. doi: 10.1038/nature22009

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Sulima, Zhukov, Afonin, Zhernakov, Tikhonovich and Lutova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Identification and Expression Analysis of *Medicago truncatula* Isopentenyl Transferase Genes (IPTs) Involved in Local and Systemic Control of Nodulation

Mahboobeh Azarakhsh, Maria A. Lebedeva* and Lyudmila A. Lutova

Department of Genetics and Biotechnology, Saint Petersburg State University, Saint Petersburg, Russia

OPEN ACCESS

Edited by:

Nikolai Provorov,
All-Russian Research Institute
of Agricultural Microbiology of the
Russian Academy of Agricultural
Sciences, Russia

Reviewed by:

Oswaldo Valdes-Lopez,
Universidad Nacional Autónoma
de México, Mexico
Frederique Catherine Guinel,
Wilfrid Laurier University, Canada

*Correspondence:

Maria A. Lebedeva
mary_osipova@mail.ru

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 14 April 2017

Accepted: 22 February 2018

Published: 09 March 2018

Citation:

Azarakhsh M, Lebedeva MA and
Lutova LA (2018) Identification
and Expression Analysis of *Medicago*
truncatula Isopentenyl Transferase
Genes (IPTs) Involved in Local
and Systemic Control of Nodulation.
Front. Plant Sci. 9:304.

doi: 10.3389/fpls.2018.00304

Cytokinins are essential for legume plants to establish a nitrogen-fixing symbiosis with rhizobia. Recently, the expression level of cytokinin biosynthesis *IPTs* (ISOPENTENYLTRANSFERASES) genes was shown to be increased in response to rhizobial inoculation in *Lotus japonicus*, *Medicago truncatula* and *Pisum sativum*. In addition to its well-established positive role in nodule primordium initiation in root cortex, cytokinin negatively regulates infection processes in the epidermis. Moreover, it was reported that shoot-derived cytokinin inhibits the subsequent nodule formation through AON (autoregulation of nodulation) pathway. In *L. japonicus*, *LjIPT3* gene was shown to be activated in the shoot phloem via the components of AON system, negatively affecting nodulation. However, in *M. truncatula*, the detailed analysis of *MtIPTs* expression, both in roots and shoots, in response to nodulation has not been performed yet, and the link between *IPTs* and AON has not been studied so far. In this study, we performed an extensive analysis of *MtIPTs* expression levels in different organs, focusing on the possible role of *MtIPTs* in nodule development. *MtIPTs* expression dynamics in inoculated roots suggest that besides its early established role in the nodule primordia development, cytokinin may be also important for later stages of nodulation. According to expression analysis, *MtIPT3*, *MtIPT4*, and *MtIPT5* are activated in the shoots in response to inoculation. Among these genes, *MtIPT3* is the only one the induction of which was not observed in leaves of the *sunn-3* mutant defective in CLV1-like kinase, the key component of AON, suggesting that *MtIPT3* is activated in the shoots in an AON-dependent manner. Taken together, our findings suggest that *MtIPTs* are involved in the nodule development at different stages, both locally in inoculated roots and systemically in shoots, where their expression can be activated in an AON-dependent manner.

Keywords: cytokinin, IPT, legume-rhizobium symbiosis, nodule development, autoregulation of nodulation

INTRODUCTION

Cytokinins (CKs) are involved in different aspects of nodule development, playing a dual role in nodulation, depending on the time and place of their action. The exogenous application of cytokinin to legume roots induces responses similar to those of rhizobial Nod-factor (NF), including cortical cell division and expression of early nodulin genes (Cooper and Long, 1994; Heckmann et al., 2011). Moreover, the gain-of-function mutation in *Lotus histidine kinase1* (*Lhk1*) cytokinin receptor gene causes the formation of spontaneous nodules in *Lotus japonicus* (Tirichine et al., 2007). The loss-of-function mutations in *Lhk1* gene or RNAi-mediated downregulation of its ortholog in *M. truncatula* leads to a dramatic reduction in the nodule formation (Gonzalez-Rizzo et al., 2006; Murray et al., 2007, respectively). All these data suggest that cytokinins have a positive influence positively on the nodule development. However, elevated cytokinin levels can also locally contribute to the negative regulation of nodulation. *ckx3* mutants that are defective in cytokinin oxidase/dehydrogenase gene, involved in cytokinin degradation, exhibit reduced nodulation and infection thread formation together with elevated root levels of tZ (trans-zeatin) and DHZ (dihydrozeatin) cytokinins (Reid et al., 2016). In addition, the *lhk1* mutant that is defective in a cytokinin receptor gene in *L. japonicus* displays epidermal hyperinfection, suggesting that CKs negatively regulate epidermal infection (Held et al., 2014). However, in the *lhk1-1 lhk1a-1 lhk3* triple mutant, the infection thread is unable to enter the root cortex, which indicates that CKs positively regulate the infection thread entry and growth in the cortex (Miri et al., 2016). Moreover, the ectopic expression of *Arabidopsis AtCKX* (pEPI:AtCKX3) gene in the root epidermis of *Medicago truncatula* results in increased nodule number, while the ectopic expression of *AtCKX3* (pCO:AtCKX3) in the root cortex is associated with decreased nodule number (Jardinaud et al., 2016). Thus CKs act both positively and negatively on the nodulation depending on the tissue: while CKs are essential for the nodule primordium initiation and the infection process in the root cortex, they negatively regulate the infection in the epidermis (for review see Gamas et al., 2017).

CKs biosynthesis is a multistep process. The first and rate-limiting step of CKs biosynthesis is catalyzed by the adenosine phosphate isopentenyl transferase (IPT). Nine *IPT* genes (*AtIPT1*–*AtIPT9*) have been identified in *Arabidopsis thaliana*, exhibiting different expression patterns both in the shoots and the roots (Takei et al., 2001; Miyawaki et al., 2004). Among them, seven *IPT* genes (*AtIPT1* and *AtIPT3*–*AtIPT8*) encode ATP/ADP-isopentenyltransferases that have been shown to isopentenylate ATP and ADP (Kakimoto, 2001). *AtIPT2* and *AtIPT9* are tRNA-isopentenyltransferases which are supposed to supply cis-zeatin-type cytokinins in *A. thaliana* (Miyawaki et al., 2004). The production of trans-zeatin (tZ) nucleotide is catalyzed by the cytochrome P450 monooxygenase (CYP735A family) via an hydroxylation of the isopentenyl adenine nucleotide (Takei et al., 2004). Eventually, the formation of active cytokinin nucleobases from nucleotides is mediated by the product of *LONELY GUYs* (*LOGs*) genes (Kurakawa et al., 2007). In legumes, several cytokinin biosynthesis genes are activated in

response to rhizobium inoculation or NF treatment. For example, in *L. japonicus*, the expression of *IPT1/2/3/4* together with *CYP735A* and *LOG1/4* is increased in response to rhizobial inoculation (Chen et al., 2014; Reid et al., 2017). Furthermore, the overexpression of cytokinin biosynthesis pathway genes (*IPT3*, *LOG4*, and *CYP735A*) is sufficient to induce spontaneous nodule formation (Reid et al., 2017). In *M. truncatula*, the expression of some cytokinin biosynthesis genes occurs as early as 3 h after NF treatment (van Zeijl et al., 2015). Two *LOG* genes are activated in response to bacterial inoculation (Mortier et al., 2014). Moreover, RNA-seq analysis shows the expression of some *LOG* genes along with that of *CYP735A1* gene (Medtr6g017325) in the epidermis of inoculated roots (Jardinaud et al., 2016). Activation of *IPT* and *LOG* genes has also been demonstrated in nodules of *Pisum sativum* (Azarakhsh et al., 2015; Dolgikh et al., 2017).

Furthermore, *LjIPT3* is activated in the shoot phloem via the components of the AON (autoregulation of nodulation) system which negatively affects nodulation (Sasaki et al., 2014). AON represents a systemic regulation controlling the number of root nodules; it involves long distance signaling and communication between the roots and the shoot (for review see Reid et al., 2011; Mortier et al., 2012; Soyano et al., 2014). AON was shown to be activated when the first nodule primordia are formed (Caetano-Anolles and Gresshoff, 1991; Li et al., 2009). One of its key components is the product of a CLV1-like (CLAVATA1-like) receptor kinase gene [*SUPER NUMERIC NODULES* (*SUNN*) in *M. truncatula* and *HYPER NODULATION ABERRANT ROOT FORMATION1* (*HARI*) in *L. japonicus*], which acts in the shoot. Mutations in the CLV1-like kinase gene lead to a shoot-controlled supernodulating phenotype in different legume species (Krusell et al., 2002; Nishimura et al., 2002; Searle et al., 2003; Schnabel et al., 2005). According to the present model of AON (for review see Soyano et al., 2014), CLE-peptides produced in nodules are transported to the shoot, where they bind to the CLV receptor complex including CLV1-like kinase; consequently, the response to suppress excessive nodule formation on roots is triggered. A set of studies suggests that the CLV receptor complex may include several proteins. In *M. truncatula*, the proteins *CORYNE* (*CRN*) and *CLAVATA 2* (*CLV2*) interact with *SUNN* (Crook et al., 2016), and a mutation in the *CLV2* gene also leads to a supernodulating phenotype in legume plants (Krusell et al., 2011). In *L. japonicus*, the *KLAVIER* gene, the defect of which is also characterized by a shoot-controlled supernodulating phenotype, encodes a receptor-like kinase, structurally unrelated to CLV1-like kinase, that was shown to interact with *HARI*, suggesting that these proteins may form a receptor complex which would perceive CLE-peptides (Miyazawa et al., 2010). Among the factors acting downstream of CLV1-like kinase in the root, the product of the *TOO MUCH LOVE* (*TML*) gene, an F-box protein, was revealed; its mutation results in a root-controlled supernodulating phenotype (Takahara et al., 2013).

The exact molecular nature of the shoot-derived signal which inhibits nodulation is still not well understood. It was shown that the *Mtsunn* mutant had increased amount of auxin transported from the shoot to the root. This indicates that the activation of CLV1-like kinase in the shoot leads to a reduction of auxin transport to developing nodules, thereby to a reduction in

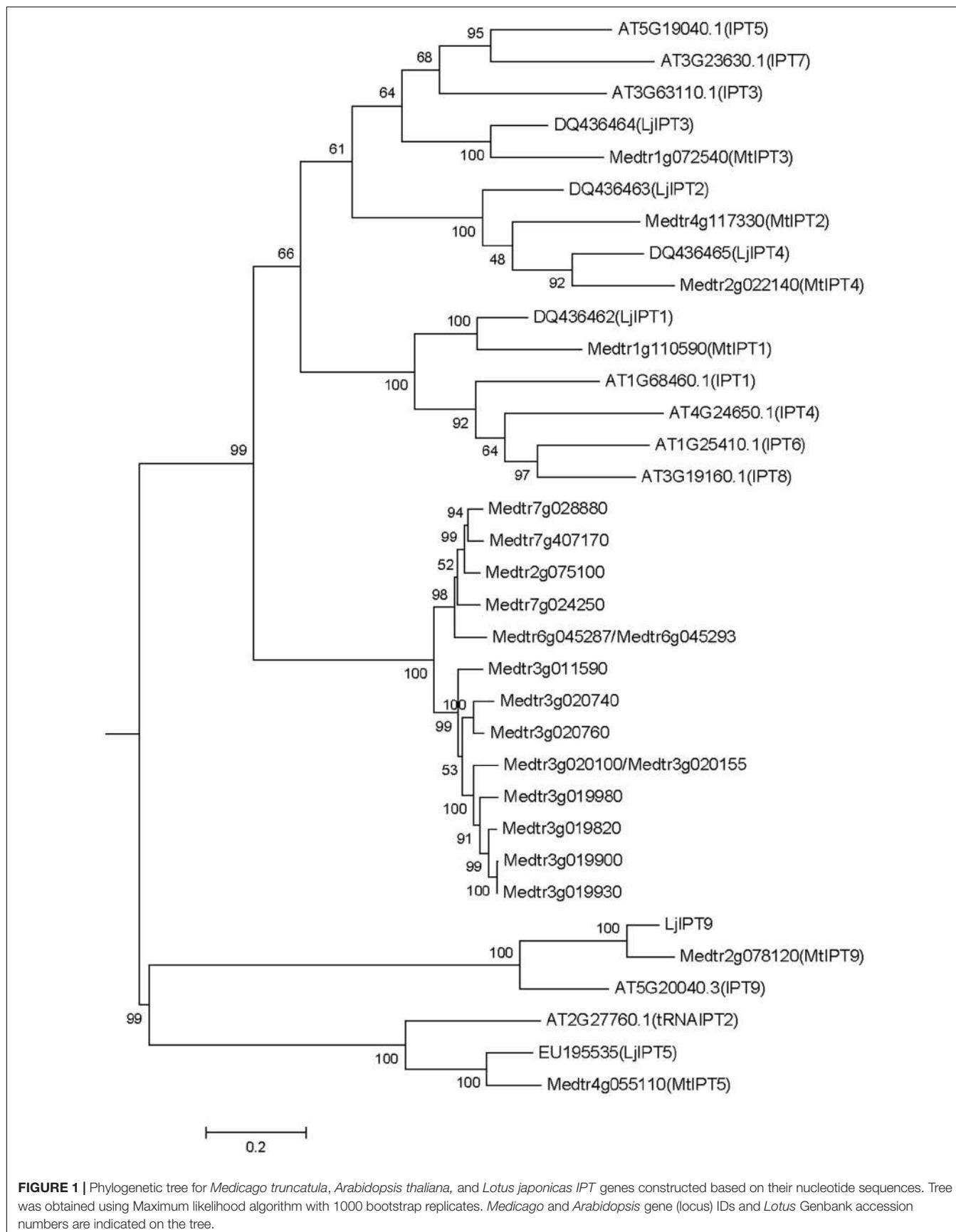


FIGURE 1 | Phylogenetic tree for *Medicago truncatula*, *Arabidopsis thaliana*, and *Lotus japonicas* IPT genes constructed based on their nucleotide sequences. Tree was obtained using Maximum likelihood algorithm with 1000 bootstrap replicates. *Medicago* and *Arabidopsis* gene (locus) IDs and *Lotus* Genbank accession numbers are indicated on the tree.

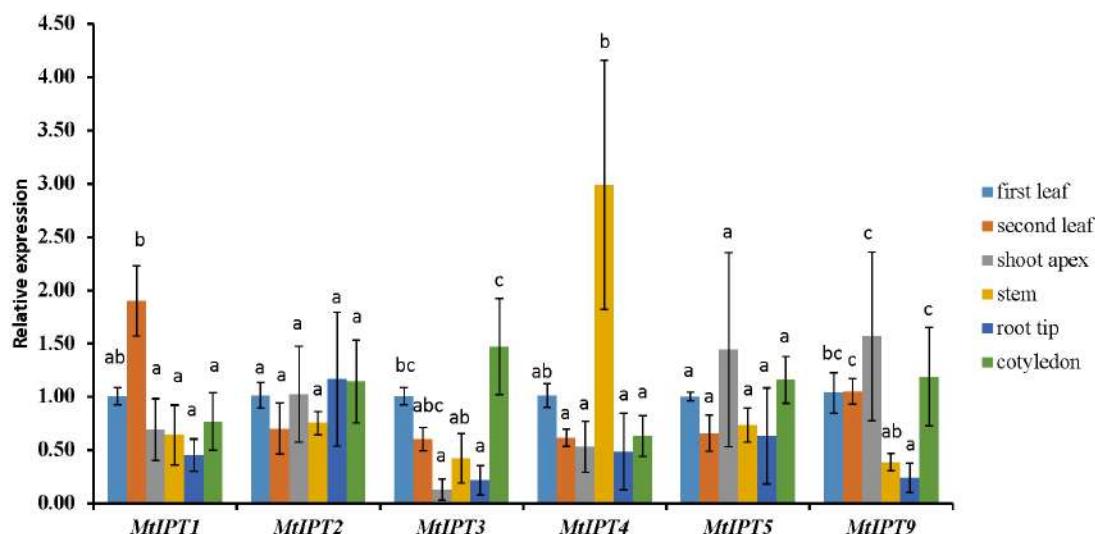


FIGURE 2 | qRT-PCR expression analysis of *MtIPTs* in different organs. Samples are first leaf, second leaf, shoot apex, stem, root tip, and cotyledon. The first leaf for each gene was used as reference sample. The results are means \pm 95% confidence intervals of three biological repeats. Different letters indicate significant differences in relative expression levels according to Kruskal and Wallis test ($P < 0.05$).

nodulation (van Noorden et al., 2006). The activation of *LjIPT3* in the shoot phloem, downstream of CLV1-like kinase, suggests that the shoot-to-root transport of both auxin and cytokinin is targeted by CLV1-like kinase, so that in roots the auxin amount is reduced and the cytokinin level is increased to restrict subsequent nodule formation. Other hormones such as jasmonic acid (JA) have also been implicated in the shoot-to-root communication during AON (Kinkema and Gresshoff, 2008; Reid et al., 2012).

In our previous study we had showed that the orthologs of *LjIPT1* and *LjIPT3* genes in *M. truncatula*, *Medtr1g110590* (*MtIPT1*) and *Medtr1g072540* (*MtIPT3*) respectively, were also upregulated in developing nodules at 7–9 days after rhizobial inoculation (Azarakhsh et al., 2015). However, the detailed analysis of *MtIPTs* expression, both in the roots and the shoots, in response to nodulation has not been performed and the link between *IPTs* and AON has not been studied in *M. truncatula* so far. In this study, we estimated, both in the shoots and the roots, the expression levels of all *MtIPTs* identified in databases at different time points after rhizobial inoculation. Moreover, we analyzed the expression of *MtIPTs* in the *sunn-3* mutant to identify *MtIPTs* potentially involved in the systemic control of nodulation in a SUNN-dependent manner.

MATERIALS AND METHODS

Plant Material, Bacterial Strains, and Growth Conditions

Medicago truncatula Gaertn. Jemalong wild-type A17 and *sunn-3* mutant plants were grown in the growth chambers (16 h/8 h day/night regime, 21°C, and 75% relative humidity). The seeds were surface-sterilized with concentrated sulphuric acid for 10 min and washed five to six times with sterile water. The seeds

were transferred on 1% agar and were kept at 4°C for 24 h; they were germinated at room temperature in darkness for 48 h. For temporal expression analysis *M. truncatula* plants were grown in vermiculite-containing pots moistened with nitrogen-free Fahraeus medium (Fahraeus, 1957). Ten days after germination, individual plants were inoculated with 1 ml of a *Sinorhizobium meliloti* (strain Sm2011) culture (OD600-0.7). Infected sites of the root with developing nodules as well as the shoot (the first leaves, the second leaves) were harvested at different stages after rhizobial inoculation. Non-inoculated plants grown in the same conditions were used as control. In the temporal expression analysis of *MtIPTs* performed at different stages of nodule development [from 1 to 21 days after inoculation (dpi), see Supplementary Figure S2 for microscopy images of the nodule developmental stages], we used 3 control time points [non-inoculated plants (NI) at 3, 5, and 7 days]. To avoid harvesting lateral root primordia, only the segments between emerged lateral roots were collected. Nodules were obtained at different stages after inoculation from the infected sites of the roots. For the expression analysis in different organs, plants were grown under the same conditions as for the temporal expression analysis, and the first leaf, the second leaf, shoot apex, stem, root tip and cotyledons were harvested at 16 days after germination, i.e., 6 days after inoculation.

Quantitative Reverse Transcription PCR (qRT-PCR) Analysis

Total RNA was isolated from the plant tissues with an RNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's instructions. DNase treatment was done using Rapid Out DNA Removal Kit (Thermo Fisher Scientific, United States). The quality of the samples was controlled and quantified with a Nano Drop 2000c UV-Vis Spectrophotometer

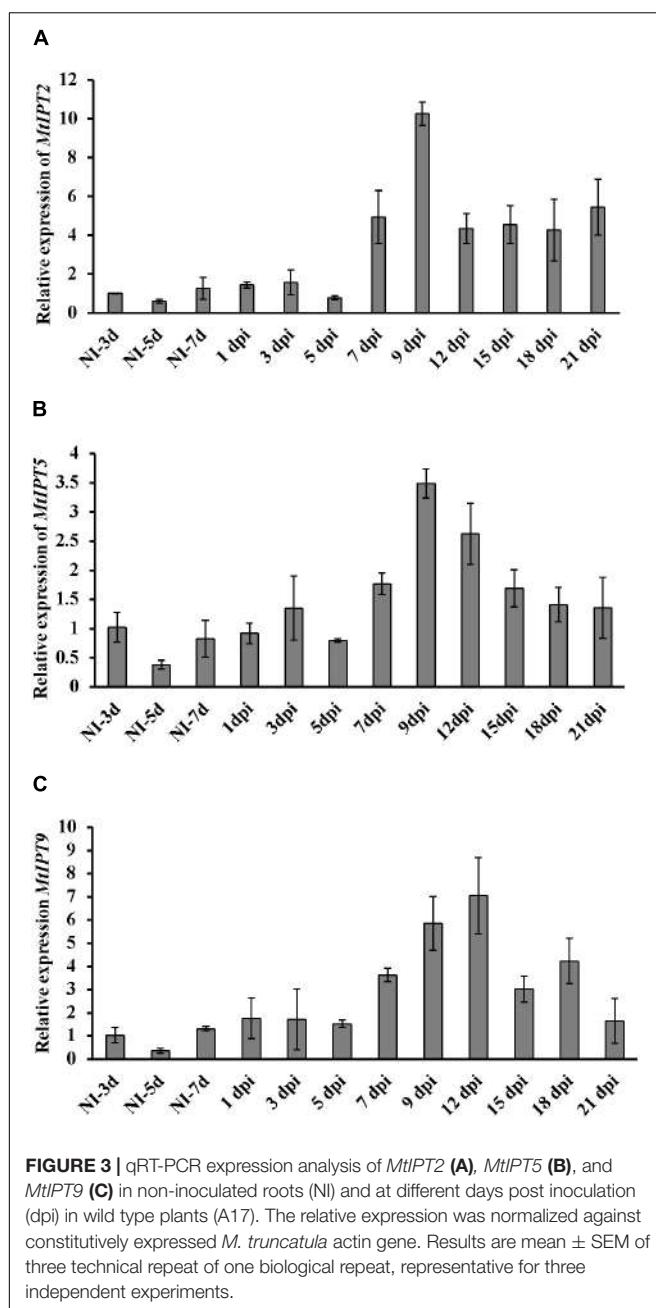


FIGURE 3 | qRT-PCR expression analysis of *MtIPT2* (A), *MtIPT5* (B), and *MtIPT9* (C) in non-inoculated roots (Nl) and at different days post inoculation (dpi) in wild type plants (A17). The relative expression was normalized against constitutively expressed *M. truncatula* actin gene. Results are mean \pm SEM of three technical repeat of one biological repeat, representative for three independent experiments.

(Thermo Fisher Scientific, United States). cDNA synthesis was performed with equal amount of RNA for all the timepoints in each experiment (varying between 400 ng up to 1 μ g of RNA in different experiments), using Revert Aid Reverse Transcriptase kit (Thermo Fisher Scientific, United States). To check DNase treatment efficacy, qRT-PCR analysis of control samples without reverse transcriptase was performed. The qRT-PCR experiments were done on a CFX-96 real-time PCR detection system with a C1000 thermal cycler (Bio-Rad Laboratories, United States). The detection was achieved using SYBR Green and Eva Green intercalating dyes (Bio-Rad Laboratories, United States). All qRT-PCR reactions were done in triplicate. Cycle threshold

(C_t) values were obtained using CFX96 manager software, and the data were analyzed by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The relative expression was normalized against the constitutively expressed actin 11 gene in *Medicago*. cDNA sequences were taken from the *M. truncatula* genome database Mt4.0v1. All primers (Supplementary Table S1) were designed using Vector NTI Advance 10 software (Thermo Fisher Scientific, United States), and were synthesized by Evrogen (Evrogen, Russia). The specificity of PCR amplification was confirmed based on dissociation curve (55–95°C). For each experiment, at least three independent biological repeats were performed. The materials for each biological repeat of the shoot or the root/nodule were taken from 4 plants.

Computer Software and Statistical Methods

Multiple alignment of nucleotide sequences was performed using Clustal W algorithm (Thompson et al., 1994) in Vector NTI Advance 10 software (Thermo Fisher Scientific, United States). For phylogenetic analysis, nucleotide sequences were retrieved from phytozome¹ for *M. truncatula* and *A. thaliana* and from Genbank NCBI database² for *L. japonicus*. Sequences were aligned with the MEGA6 program using Clustal W and the phylogenetic tree was generated using Maximum Likelihood method based on Tamura-Nei model (Tamura and Nei, 1993; Hall, 2013) with 1000 bootstrap replicates.

One-way ANOVA, Kruskal and Wallis test, and Student's *t*-test were used to compare the gene expression levels of different samples. The graphs indicate mean with 95% confidence interval. At least three independent biological repeats were done for each experiment. For each time point (inoculated or non-inoculated) and in each biological repeat, four plants were used.

RESULTS

Identification of *Medicago truncatula* IPT Genes

Medicago truncatula genomic data (Mt4.0v1) contain 23 sequences annotated as isopentenyl transferases (IPT), two pairs of which are exactly the same sequences (*Medtr6g045287/Medtr6g045293* and *Medtr3g020100/Medtr3g020155*). Among the 21 non-repeated sequences, there are two sequences encoding truncated peptides (*Medtr7g007180* and *Medtr7g007190* with 57 and 59 amino acids, respectively). A phylogenetic tree was obtained using the nucleotide sequences of *IPT* genes of *M. truncatula*, *A. thaliana* and *L. japonicus* (Figure 1). According to the phylogenetic tree, *IPT* genes can be divided into five groups, one of which is unique for *M. truncatula*. This group contains 15 members that are highly similar in sequences with 52.6% identity and 95.45% consensus positions (Supplementary Figure S1). Other *M. truncatula* *IPT* genes are clustered with *A. thaliana* and *L. japonicus* genes into four groups. *MtIPT1,3,4* (*Medtr1g110590*, *Medtr1g072540*, and

¹<https://phytozome.jgi.doe.gov/>

²<https://www.ncbi.nlm.nih.gov/>

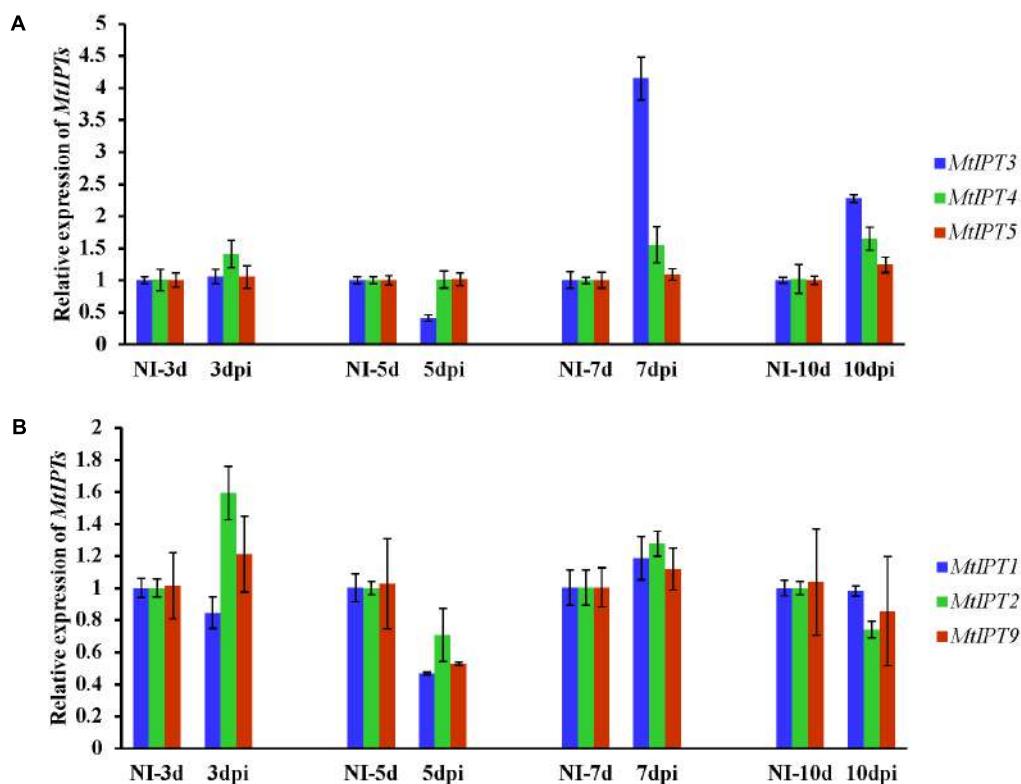


FIGURE 4 | Relative expression of *MtIPT3*, *MtIPT4*, *MtIPT5* (A), and *MtIPT1*, *MtIPT2*, *MtIPT9* (B) genes in the first leaf at different days post inoculation (3, 5, 7, and 10 dpi) in comparison with non-inoculated plants (NI). Results are mean \pm SEM of three technical repeat of one representative biological repeat out of three biological repeats.

Medtr2g022140, respectively) have been identified previously (Azarakhsh et al., 2015), and they were named according to their nearest orthologs in *L. japonicus*. Similarly, here we refer to the other *M. truncatula* genes according to their near orthologs in *L. japonicus* and *A. thaliana*. *Medtr4g117330* is referred to as *MtIPT2*, *Medtr4g055110* as *MtIPT5* and *Medtr2g078120* as *MtIPT9*. Because the two latter genes are clustered with *AtIPT2*, *AtIPT9* and *LjIPT5*, *LjIPT9* which are all known to be tRNA IPTs (Kakimoto, 2001; Chen et al., 2014), *MtIPT5* and *MtIPT9* were annotated as tRNA isopentenyl transferases.

Expression Pattern of *MtIPT* Genes

The expression levels of *MtIPT* genes were examined using qRT-PCR in different organs (the first and the second leaves, shoot apex, stem, root tips, and cotyledons) and the first leaf was used as the reference tissue. *MtIPT1* is expressed in all organs we tested with a slightly higher expression in the second leaf. As for *MtIPT2*, it is expressed in all analyzed organs at the comparable levels. *MtIPT3* exhibits a significantly lower expression in shoot apex and root tip than in the first leaf and cotyledons, and *MtIPT4* shows a higher expression in the stem. *MtIPT5* is expressed in all analyzed organs at the comparable levels, whereas *MtIPT9* has lower expression in the stem and in the root tip in comparison with other organs. (Figure 2). Most genes from the *M. truncatula* unique group show no expression in the different organs analyzed

(no signal at all or Ct values close to water controls) except for *Medtr6g045287*, *Medtr7g028880*, and *Medtr7g407170*, which exhibited low expression levels in all analyzed organs (data not shown).

Expression of *MtIPT* Genes in Response to Rhizobial Inoculation

To study the temporal expression of *MtIPT* genes during the nodule development, relative transcript levels were estimated at different timepoints after the rhizobial inoculation (days post inoculation, dpi) and compared to transcript levels of non-inoculated control roots (see Supplementary Figure S2 for microscopy images of the nodule developmental stages).

Previously, we had analyzed the expression dynamic of three *MtIPTs* (*MtIPT1*, *MtIPT3*, and *MtIPT4*) in response to the rhizobium inoculation (See Figure 5 from Azarakhsh et al., 2015). We had found that *MtIPT1* gene expression dramatically increased at 7 and 9 dpi, compared with non-inoculated control. Moreover, the expression level of *MtIPT3* was also increased during the nodulation. Although van Zeijl et al. (2015) had found activation of *MtIPT4* 3 h after NF treatment, we had not detected a significant *MtIPT4* expression upon nodule development (see Figures 5A–C from Azarakhsh et al., 2015).

Here, we found that the expression level of *MtIPT2* increased at 7 dpi, reached a 10-fold increase at 9 dpi in comparison to

non-inoculated control (NI-3d), and remained high until the late stages of nodule development (up to 4-fold increase at 21 dpi). *MtIPT5* was slightly activated throughout nodule formation with a maximum of 3-fold increase at 9 dpi in comparison to non-inoculated control (NI-3d). Expression of *MtIPT9* was also increased up to 6–7 fold at 9–12 dpi (Figure 3). None of the genes from *M. truncatula* unique group were activated during nodule development (data not shown). According to the *M. truncatula* LCM-RNA-seq data³, *MtIPT1* and *MtIPT4* were expressed mostly in the nodule meristem, while *MtIPT2* shows higher expression in the meristem region, the distal and proximal infection zone (Supplementary Figure S3).

To find *MtIPT* genes potentially involved in AON, we estimated the expression of *MtIPT* genes in the shoot (the first leaf and the second leaf) at different days post inoculation (3, 5, 7, and 10 dpi). The expression levels of *MtIPT3* increased in the leaves in response to the rhizobial inoculation at 7 dpi (Figures 4, 5 and Supplementary Figure S4). However, among different biological repeats, the increased *MtIPT3* expression in the leaves relative to non-inoculated controls was also observed at other time points, in particular at 5 dpi (see Supplementary Figure S5). Nevertheless, it was at the 7 dpi timepoint that a statistically significant activation of *MtIPT3* was observed taking into account three biological repeats (Figure 5). Moreover, there was also a slight but statistically significant activation of *MtIPT4* and *MtIPT5* in the first leaves in response to inoculation, while such activation was not observed for *MtIPT1*, *MtIPT2*, and *MtIPT9* (Figure 5). Activation of *MtIPT4* and *MtIPT5* was confirmed in second leaves as well (Supplementary Figure S4). Activation of *MtIPT3*, *MtIPT4*, and *MtIPT5* expression in the leaves in response to inoculation may indicate the involvement of these genes in AON.

Expression of *MtIPT* Genes in *sunn-3* Mutant

To address the question whether the activation of *MtIPTs* expression can be regulated by the key component of AON, i.e., the CLV1-like receptor, we estimated the expression levels of the *MtIPT* genes in both the shoots and the roots of the *sunn-3* supernodulating mutant. The expression levels of *MtIPT3*, *MtIPT4*, and *MtIPT5* genes that demonstrated activation in wild-type shoots at 7 dpi were estimated in the first and the second leaves of *sunn-3* mutants. In contrast to wild-type, *MtIPT3* did not show any activation in the *sunn-3* first and second leaves, while *MtIPT4* and *MtIPT5* exhibited increased expression at 7 dpi both in the wild-type and *sunn-3* leaves (Figure 6 and Supplementary Figures S6, S7).

To find out if the activation of the *MtIPT* genes in the inoculated roots can be also regulated by CLV1-like kinase, we analyzed the expression level of *MtIPT* genes in the inoculated roots of *sunn-3* mutant at different stages after inoculation. According to qRT-PCR analysis, expression of *MtIPT1-3*, *MtIPT5*, and *MtIPT9* is activated in *sunn-3* mutant as well (Supplementary Figure S8). We did not find any changes in the general pattern of their expression in comparison with

³<https://iant.toulouse.inra.fr/symbimics/>

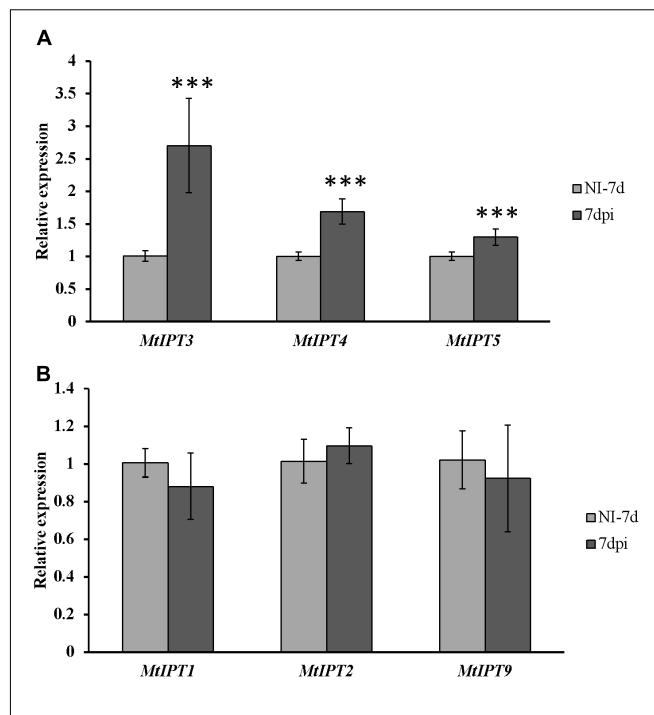


FIGURE 5 | Relative expression of *MtIPT1-5* and *MtIPT9* in the first leaf of A17 plants at 7 dpi. The expression levels of *MtIPT3*, *MtIPT4*, and *MtIPT5* are increased (**A**), while the expression levels of *MtIPT1*, *MtIPT2*, and *MtIPT9* do not change significantly (**B**) in comparison with non-inoculated plants (NI). Asterisks indicate statistically significant differences compared with control (NI): ***P < 0.001, **P < 0.01, *P < 0.05. Error bars indicate the 95% confidence interval of three biological repeats.

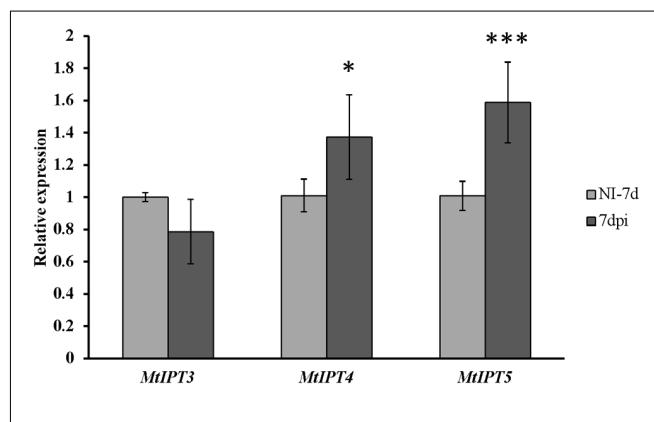


FIGURE 6 | Relative expression of *MtIPT3*, *MtIPT4*, and *MtIPT5* in the first leaf of *sunn-3* mutant at 7 dpi in comparison with non-inoculated plants (NI). Asterisks indicate statistically significant differences compared with control (NI): ***P < 0.001, **P < 0.01, *P < 0.05. Error bars indicate the 95% confidence interval of three biological repeats.

wild-type plants (for comparison see Figure 3 of this article and Figures 5A–C from Azarakhsh et al., 2015), which is consistent with an analogous experiment in *L. japonicus*. This suggests that *MtIPT3* expression in the shoots in response to inoculation may be regulated by CLV1-like kinase MtSUNN, while the activation of *MtIPT3* together with other *MtIPT* genes in the roots occurs independently of CLV1-like kinase MtSUNN.

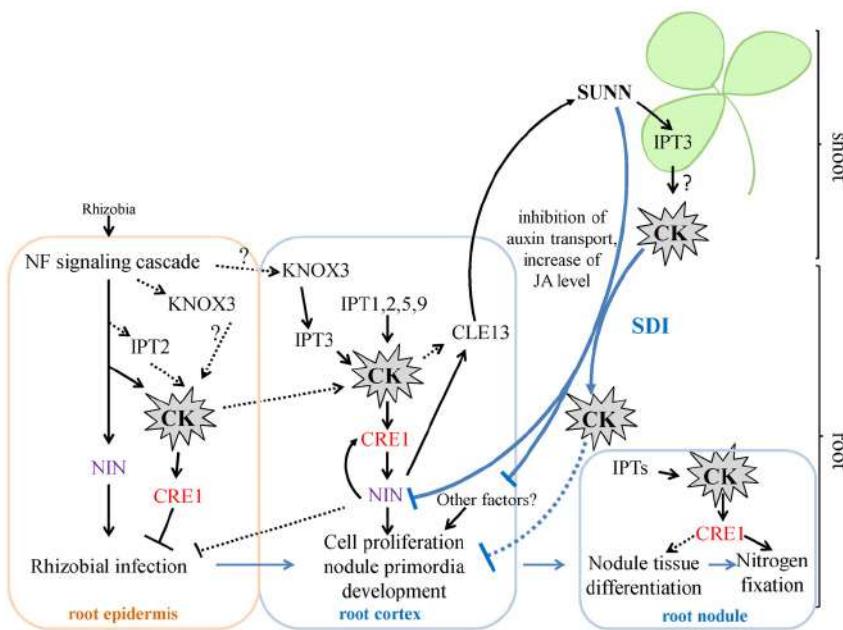


FIGURE 7 | A scheme illustrating the role of *IPTs* and cytokinin in nodulation in *M. truncatula*. Cytokinin (CK) biosynthesis and response are activated both in root epidermis and in root cortex in response to rhizobial inoculation. The induction of *IPT* genes in root epidermis contributes to the negative effect of CK on rhizobial infection, while activation of *IPT* genes in root cortex is important for cortical cell divisions during nodule primordia development, where CK is known to be a positive regulator. The expression of *KNOX3*, a possible activator of cytokinin biosynthesis genes, is induced both in root epidermis and in nodule primordia. CK activates the expression of *NIN* in the root cortex, where *NIN* promotes *CRE1* expression. The *NIN* expressed in root cortical cells negatively regulates rhizobial infection. CK are important for nodule tissue differentiation and nodule meristem development, since *CRE1* was shown to be important for the transition between meristematic and cell differentiation/elongation zones in indeterminate nodules. Moreover, CK is also important for nitrogen fixation, since it was shown that *CRE1* is necessary for this process. The expression of *IPTs* that was observed at later stages of nodule development in the present study is in agreement with this role of CK in subsequent nodule development and functioning. Being involved in AON, CK induces the *CLE13* gene. In its turn, Mt*CLE13* signal peptide activates *SUNN* receptor kinase in the shoot, triggering SDI that is transported to the roots where it inhibits subsequent nodulation. CK was shown to be the part of SDI, since in *L. japonicus* *LjIPT3* is activated in the shoot in HAR1-dependent manner, and shoot-derived cytokinin inhibits nodulation. *IPT3* activation in *SUNN*-dependent manner in the shoot is in agreement with these findings. In addition to CK, SDI involves the inhibition of auxin transport and the increase of JA level. SDI has multiple targets in the root, and *NIN* was shown to be one of them. CK – cytokinin. SDI – shoot-derived inhibitor.

DISCUSSION

In the entire *Medicago* genome Mt4.0v1, 23 sequences were annotated as isopentenyl transferases (IPT). Among them, according to the phylogenetic analysis, 15 sequences were clustered together as a separate group of sequences with a high degree of similarity, and two pairs of sequences from this group appeared to be exactly the same. This suggests that this group of *IPTs* sequences may represent new genes which appeared during evolution as a result of recent duplication events in the *Medicago* genome. However, according to the expression analysis, *IPT* genes from this group exhibited either no or very low level of expression in the *Medicago* tissues tested. Thus, a functional and more comprehensive phylogenetic analysis including *IPT* sequences from other legumes is required to elucidate the possible role of these evolutionary new *IPT* genes in legume development.

Cytokinins have been previously shown to regulate different aspects of nodulation, including rhizobial infection during which CKs act as negative regulators and nodule primordia development and nitrogen fixation for which CKs act as positive regulators (Tirichine et al., 2007; Held et al., 2014; Boivin et al.,

2016; Jardinaud et al., 2016). In *M. truncatula*, accumulation of cytokinins (iP: isopentenyladenine, iPR: isopentenyladenosine, tZ: trans-zeatin) was first observed at 3 h after NF exposure and this occurs in MtCCaMK-dependent manner. (van Zeijl et al., 2015). In our work, we did not study *IPTs* expression at very early stages nodule development, i.e., few hours after rhizobial infection. However, van Zeijl et al. (2015) reported that *Medtr2g022140* (*MtIPT4* according to our nomenclature) induced at 3 h after Nod factor treatment, therefore this *IPT* gene may contribute to cytokinin accumulation at early stages after rhizobial inoculation. Moreover, according to LCM-RNA-seq data, *MtIPT2* is activated in root epidermis 24 h after NF treatment (Jardinaud et al., 2016). Moreover, the expression of the *KNOX3* gene encoding homeodomain-containing TF that activates *IPT3* expression in developing nodules (Azarakhsh et al., 2015) was also induced in root epidermis according to LCM-RNA-seq data (Jardinaud et al., 2016). The early induction of cytokinin biosynthesis in root epidermis may contribute to its negative effect on rhizobial infection reported previously (Held et al., 2014, see Figure 7).

In pea, a two-stage increase in the content of CKs during nodule development was observed (Dolgikh et al., 2017). The

accumulation of tZ and IP at 3 dpi can be associated with cortical cell division where CKs are known to be positive regulators, while the increase of tZ content at later stages suggests that CK also plays a role in subsequent stages of the nodule development and functioning (Dolgikh et al., 2017). The temporal expression dynamics of *MtIPTs* we observed are in line with this observation, suggesting that cytokinin biosynthesis also occurs at later stages of nodule development. According to our previously reported data, the expression of *MtIPT1* and *MtIPT3* increased in the inoculated roots reaching a maximum at 9 dpi (Azarakhsh et al., 2015). Here, we found that the expression of *MtIPT2*, *MtIPT9*, and *MtIPT5* genes were also induced during the nodule development, and the expression maximum of these genes is observed at 9–12 dpi, when the nodule primordia are completely developed.

The nodulation phenotype of the *Mtcre1* mutant defective in a cytokinin receptor suggested that besides having a crucial role in the early steps of nodule initiation, CKs regulate the subsequent stages of nodule development (Plet et al., 2011). The rare nodules which formed on *Mtcre1*-inoculated roots exhibit incomplete differentiation and have frequently multiple lobes, suggesting that a functional MtCRE1 may regulate the transition between meristematic and cell differentiation/elongation zones in the indeterminate nodules (Plet et al., 2011). Taking into account this finding, we suggest that *IPTs* expression that we observed after 7 dpi (at the stages when nodule differentiation occurs and the nodule meristem is being formed) may contribute to cytokinin involvement in the nodule differentiation and the meristem formation (see Figure 7).

Moreover, recently, it was shown that functional cytokinin receptor in *M. truncatula* is required for nitrogen fixation (Boivin et al., 2016). In *cre1* mutant showing reduced and delayed nodule formation, nitrogen fixation was decreased. Notably, the AHK4/CRE1 gene from *Arabidopsis* was able to complement nodule initiation, but not nitrogen fixation in the *cre1* mutant, indicating that legume-specific determinants encoded by MtCRE1 are likely required for nitrogen fixation activity (Boivin et al., 2016). This finding represents one more evidence that CKs act at later stages of nodule development and functioning.

In *L. japonicus*, it was shown that *LjIPT3* gene is activated in the shoot in response to rhizobial inoculation, and its activation is dependent on the CLV1-like kinase LjHAR1 (Sasaki et al., 2014). This increase in *LjIPT3* expression was observed at 3–5 dpi, which is consistent with the timing of AON induction determined previously for this legume in a split-root system (Suzuki et al., 2008). In our experiment we found that the expression of *MtIPT3*, the ortholog of *LjIPT3*, was activated in shoots at 7 dpi. The activation of *MtIPT3* in the leaves of inoculated plants is consistent with the data from phytozome database⁴. *MtIPT4* and *MtIPT5* were also slightly induced in shoots after rhizobial inoculation. Interestingly, the shoot *MtIPT3* activation in response to inoculation was absent in the *sunn-3* mutant

defective in the CLV1-like kinase MtSUNN. These findings indicate that the ortholog of *LjIPT3* gene in *M. truncatula* is also activated in shoots in response to inoculation in a CLV1-like kinase-dependent manner, whereas the expression of *MtIPT3* together with other *MtIPT* genes in roots occurs independently of CLV1-like kinase. The activation of *MtIPT4* and *MtIPT5*, which was still observed in the *sunn-3* mutant in response to inoculation is unlikely to be regulated by MtSUNN, but instead it must be controlled by other regulatory pathways. The timing of AON induction in *Medicago* was determined to be 3 days after inoculation (Kassaw et al., 2015). In our experiments, the statistically significant activation of *MtIPTs* in leaves was observed at 7 dpi, although in one of three biological repeats the increased expression of *MtIPTs* in leaves was found at 5 dpi after rhizobial inoculation (Supplementary Figure S5). We suppose that temporal dynamics of nodule development may vary depending on experimental plant growth conditions, and the timing of AON activation may vary accordingly. Moreover, the expression of *MtCLE13*, a CLE-peptide that triggers AON, was shown to be activated as early as 3 h after NF treatment at a time-point preceding nodule primordium initiation (van Zeijl et al., 2015). According to Mortier et al. (2010), *MtCLE13* was significantly increased in inoculated roots at 4 dpi, and its expression level remained relatively high until 10 dpi, with a slightly pronounced maximum at 8 dpi. These findings allow us to propose that AON may comprise different levels of regulation, involving the perception of different signals. The shoot-derived signal (SDI) in AON also appeared to have a complex molecular nature. AON involves different hormones such as auxins, JA and cytokinins (van Noorden et al., 2006; Kinkema and Gresshoff, 2008; Sasaki et al., 2014), that may also inhibit nodulation at different levels via multiple targets. Among the later ones, the NIN transcription factor was shown to be targeted by AON in inoculated roots in a negative feed-back manner, where NIN also directly induces CLEs expression to trigger AON (Soyano et al., 2014). Together with CKs, NIN represents a crucial regulator of different aspects of nodulation that plays both positive and negative roles in nodulation. Along with the fact that NIN transcription factor is important both for rhizobial infection and nodule primodia development (Schauser et al., 1999; Marsh et al., 2007), NIN expression in the cortex was shown to inhibit rhizobial infection (Yoro et al., 2014). Moreover, NIN expression in the root cortex is activated by CK (Gonzalez-Rizzo et al., 2006; Plet et al., 2011), and NIN can promote CRE1 expression (Vernié et al., 2015) (see Figure 7).

A large body of evidence indicates that CKs play a complex role in nodulation, mediating negative feedback regulatory mechanisms (Mortier et al., 2014; Sasaki et al., 2014). The exact molecular mechanisms underlying the dual effects of CKs on nodulation are still far from being understood. It is likely that CKs acting in a tight crosstalk with other hormones, such as auxins, may exert different effects on nodulation depending on the auxin concentration that is known to fluctuate through different stages of nodulation. Future studies

⁴<https://phytozome.jgi.doe.gov/>

should reveal the molecular pathways regulated by CKs and other hormones to elucidate their complex action during nodulation.

AUTHOR CONTRIBUTIONS

MA performed all the experiments on gene expression analysis. MA and ML analyzed the data and drafted the manuscript. ML and LL planned the experiments and supervised the research. All authors read and approved the manuscript.

FUNDING

This work was supported by Russian Scientific Foundation project No. 16-16-10011 and grants from Russian Foundation for Basic Research Nos. 15-34-20071 and 15-29-02737 and grant of the president of the Russian Federation HIII-9513.2016.4.

REFERENCES

- Azarakhsh, M., Kirienko, A., Zhukov, V., Lebedeva, M., Dolgikh, E., and Lutova, L. (2015). Knotted1-like homeobox 3: a new regulator of symbiotic nodule development. *J. Exp. Bot.* 66, 7181–7195. doi: 10.1093/jxb/erv414
- Boivin, S., Kazmierczak, T., Brault, M., Wen, J., Gamas, P., Mysore, K. S., et al. (2016). Different cytokinin histidine kinase receptors regulate nodule initiation as well as later nodule developmental stages in *Medicago truncatula*. *Plant Cell Environ.* 39, 2198–2209. doi: 10.1111/pce.12779
- Caetano-Anolles, G., and Gresshoff, P. M. (1991). Plant genetic control of nodulation. *Annu. Rev. Microbiol.* 45, 345–382. doi: 10.1146/annurev.mi.45.100191.002021
- Chen, Y., Chen, W., Li, X., Jiang, H., Wu, P., Xia, K., et al. (2014). Knockdown of LjIPT3 influences nodule development in *Lotus japonicus*. *Plant Cell Physiol.* 55, 183–193. doi: 10.1093/pcp/pct171
- Cooper, J. B., and Long, S. R. (1994). Morphogenetic rescue of *Rhizobium meliloti* nodulation mutants by trans-zeatin secretion. *Plant Cell* 6, 215–225. doi: 10.1105/tpc.6.2.215
- Crook, A. D., Schnabel, E. L., and Frugoli, J. A. (2016). The systemic nodule number regulation kinase SUNN in *Medicago truncatula* interacts with MtCLV2 and MtCRN. *Plant J.* 88, 108–119. doi: 10.1111/tpj.13234
- Dolgikh, E., Shaposhnikov, A., Dolgikh, A., Gribchenko, E., Bodyagina, K., Yuzhikhin, O., et al. (2017). Identification of *Pisum sativum* L. cytokinin and auxin metabolic and signaling genes, and an analysis of their role in symbiotic nodule development. *Int. J. Plant Physiol. Biochem.* 9, 22–35. doi: 10.5897/IJPPB2017.0266
- Fahraeus, G. (1957). The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. *Microbiology* 16, 374–381. doi: 10.1099/02221287-16-2-374
- Gamas, P., Brault, M., Jardinaud, M.-F., and Frugier, F. (2017). Cytokinins in symbiotic nodulation: When, where, what for? *Trends Plant Sci.* 22, 792–802. doi: 10.1016/j.tplants.2017.06.012
- Gonzalez-Rizzo, S., Crespi, M., and Frugier, F. (2006). The *Medicago truncatula* CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with *Sinorhizobium meliloti*. *Plant Cell* 18, 2680–2693. doi: 10.1105/tpc.106.043778
- Hall, B. G. (2013). Building phylogenetic trees from molecular data with MEGA. *Mol. Biol. Evol.* 30, 1229–1235. doi: 10.1093/molbev/mst012
- Heckmann, A. B., Sandal, N., Bek, A. S., Madsen, L. H., Jurkiewicz, A., Nielsen, M. W., et al. (2011). Cytokinin induction of root nodule primordia in *Lotus japonicus* is regulated by a mechanism operating in the root cortex. *Mol. Plant Microbe Interact.* 24, 1385–1395. doi: 10.1094/MPMI-05-11-0142
- Held, M., Hou, H., Miri, M., Huynh, C., Ross, L., Hossain, M. S., et al. (2014). *Lotus japonicus* cytokinin receptors work partially redundantly to mediate nodule formation. *Plant Cell* 26, 678–694. doi: 10.1105/tpc.113.119362
- Jardinaud, M.-F., Boivin, S., Rodde, N., Catrice, O., Kisiala, A., Lepage, A., et al. (2016). A laser dissection-RNAseq analysis highlights the activation of cytokinin pathways by nod factors in the *Medicago truncatula* root epidermis. *Plant Physiol.* 171, 2256–2276. doi: 10.1104/pp.16.00711
- Kakimoto, T. (2001). Identification of plant cytokinin biosynthetic enzymes as dimethylallyl diphosphate: ATP/ADP isopentenyltransferases. *Plant Cell Physiol.* 42, 677–685. doi: 10.1093/pcp/pce112
- Kassaw, T., Bridges, J. W., and Frugoli, J. (2015). Multiple autoregulation of nodulation (AON) signals identified through split root analysis of *Medicago truncatula sunn* and *rdn1* mutants. *Plants* 4, 209–224. doi: 10.3390/plants4020209
- Kinkema, M., and Gresshoff, P. M. (2008). Investigation of downstream signals of the soybean autoregulation of nodulation receptor kinase GmNARK. *Mol. Plant Microbe Interact.* 21, 1337–1348. doi: 10.1094/MPMI-21-10-1337
- Krusell, L., Madsen, L. H., Sato, S., Aubert, G., Genua, A., Szczęgowski, K., et al. (2002). Shoot control of root development and nodulation is mediated by a receptor-like kinase. *Nature* 420, 422–426. doi: 10.1038/nature01207
- Krusell, L., Sato, N., Fukuhara, I., Koch, B. E., Grossmann, C., Okamoto, S., et al. (2011). The Clavata2 genes of pea and *Lotus japonicus* affect autoregulation of nodulation. *Plant J.* 65, 861–871. doi: 10.1111/j.1365-313X.2010.04474.x
- Kurakawa, T., Ueda, N., Maekawa, M., Kobayashi, K., Kojima, M., Nagato, Y., et al. (2007). Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature* 445, 652–655. doi: 10.1038/nature05504
- Li, D., Kinkema, M., and Gresshoff, P. M. (2009). Autoregulation of nodulation (AON) in *Pisum sativum* (pea) involves signalling events associated with both nodule primordia development and nitrogen fixation. *J. Plant Physiol.* 166, 955–967. doi: 10.1016/j.jplph.2009.03.004
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Marsh, J. F., Rakocevic, A., Mitra, R. M., Brocard, L., Sun, J., Eschstruth, A., et al. (2007). *Medicago truncatula NIN* is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. *Plant Physiol.* 144, 324–335. doi: 10.1104/pp.106.093021
- Miri, M., Janakirama, P., Held, M., Ross, L., and Szczęgowski, K. (2016). Into the root: how cytokinin controls rhizobial infection. *Trends Plant Sci.* 21, 178–186. doi: 10.1016/j.tplants.2015.09.003
- Miyawaki, K., Matsumoto-Kitano, M., and Kakimoto, T. (2004). Expression of cytokinin biosynthetic isopentenyltransferase genes in *Arabidopsis*: tissue

ACKNOWLEDGMENTS

We thank the Research Resource Center for Molecular and Cell Technologies of Saint Petersburg State University for the equipment used in this study. We thank Sofie Goormachtig (Ghent University, VIB, Ghent, Belgium) for the seeds of *M. truncatula sunn-3* mutant, and Prof. Igor A. Tikhonovich (All-Russia Research Institute for Agricultural Microbiology, Pushkin, Saint Petersburg State University, Saint Petersburg, Russia) for the helpful comments given in the preparation of the manuscript. We also thank reviewers for their constructive comments and criticism, which helped us to improve the manuscript significantly.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00304/full#supplementary-material>

- specificity and regulation by auxin, cytokinin, and nitrate. *Plant J.* 37, 128–138. doi: 10.1046/j.1365-313X.2003.01945.x
- Miyazawa, H., Oka-Kira, E., Sato, N., Takahashi, H., Wu, G.-J., Sato, S., et al. (2010). The receptor-like kinase KLAVIER mediates systemic regulation of nodulation and non-symbiotic shoot development in *Lotus japonicus*. *Development* 137, 4317–4325. doi: 10.1242/dev.058891
- Mortier, V., Den Herder, G., Whitford, R., Van de Velde, W., Rombauts, S., D'Haeseleer, K., et al. (2010). CLE peptides control *Medicago truncatula* nodulation locally and systemically. *Plant Physiol.* 153, 222–237. doi: 10.1104/pp.110.153718
- Mortier, V., Holsters, M., and Goormachtig, S. (2012). Never too many? How legumes control nodule numbers. *Plant Cell Environ.* 35, 245–258. doi: 10.1111/j.1365-3040.2011.02406.x
- Mortier, V., Wasson, A., Jaworek, P., De Keyser, A., Decroos, M., Holsters, M., et al. (2014). Role of LONELY GUY genes in indeterminate nodulation on *Medicago truncatula*. *New Phytol.* 202, 582–593. doi: 10.1111/nph.12681
- Murray, J. D., Karas, B. J., Sato, S., Tabata, S., Amyot, L., and Szczegłowski, K. (2007). A cytokinin perception mutant colonized by *Rhizobium* in the absence of nodule organogenesis. *Science* 315, 101–104. doi: 10.1126/science.1132514
- Nishimura, R., Hayashi, M., Wu, G.-J., Kouchi, H., Imaizumi-Anraku, H., Murakami, Y., et al. (2002). HAR1 mediates systemic regulation of symbiotic organ development. *Nature* 420, 426–429. doi: 10.1038/nature01231
- Plet, J., Wasson, A., Ariel, F., Le Signor, C., Baker, D., Mathesius, U., et al. (2011). MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in *Medicago truncatula*. *Plant J.* 65, 622–633. doi: 10.1111/j.1365-313X.2010.04447.x
- Reid, D. E., Ferguson, B. J., Hayashi, S., Lin, Y.-H., and Gresshoff, P. M. (2011). Molecular mechanisms controlling legume autoregulation of nodulation. *Ann. Bot.* 108, 789–795. doi: 10.1093/aob/mcr205
- Reid, D. E., Hayashi, S., Lorenc, M., Stiller, J., Edwards, D., Gresshoff, P. M., et al. (2012). Identification of systemic responses in soybean nodulation by xylem sap feeding and complete transcriptome sequencing reveal a novel component of the autoregulation pathway. *Plant Biotechnol. J.* 10, 680–689. doi: 10.1111/j.1467-7652.2012.00706.x
- Reid, D. E., Heckmann, A. B., Novak, O., Kelly, S., and Stougaard, J. (2016). Cytokinin oxidase/dehydrogenase3 maintains cytokinin homeostasis during root and nodule development in *Lotus japonicus*. *Plant Physiol.* 170, 1060–1074. doi: 10.1104/pp.15.00650
- Reid, D. E., Nadzieja, M., Novak, O., Heckmann, A. B., Sandal, N., and Stougaard, J. (2017). Cytokinin biosynthesis promotes cortical cell responses during nodule development. *Plant Physiol.* 175, 361–375. doi: 10.1104/pp.17.00832
- Sasaki, T., Suzuki, T., Soyano, T., Kojima, M., Sakakibara, H., and Kawaguchi, M. (2014). Shoot-derived cytokinins systemically regulate root nodulation. *Nat. Commun.* 5:4983. doi: 10.1038/ncomms5983
- Schauser, L., Roussis, A., Stiller, J., and Stougaard, J. (1999). A plant regulator controlling development of symbiotic root nodules. *Nature* 402, 191–195. doi: 10.1038/46058
- Schnabel, E. L., Journet, E.-P., Carvalho-Niebel, D. F., Duc, G., and Frugoli, J. A. (2005). The *Medicago truncatula* SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length. *Plant Mol. Biol.* 58, 809–822. doi: 10.1007/s11103-005-8102-y
- Searle, I. R., Men, A. E., Laniya, T. S., Buzas, D. M., Iturbe-Ormaetxe, I., Carroll, B. J., et al. (2003). Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. *Science* 299, 109–112. doi: 10.1126/science.1077937
- Soyano, T., Hirakawa, H., Sato, S., Hayashi, M., and Kawaguchi, M. (2014). Nodule inception creates a long-distance negative feedback loop involved in homeostatic regulation of nodule organ production. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14607–14612. doi: 10.1073/pnas.1412716111
- Suzuki, A., Hara, H., Kinoue, T., Abe, M., Uchiumi, T., Kucho, K., et al. (2008). Split-root study of autoregulation of nodulation in the model legume *Lotus japonicus*. *J. Plant Res.* 121, 245–249. doi: 10.1007/s10265-007-0145-5
- Takahara, M., Magori, S., Soyano, T., Okamoto, S., Yoshida, C., Yano, K., et al. (2013). Too much love, a novel kelch repeat-containing F-box protein, functions in the long-distance regulation of the legume–*Rhizobium* symbiosis. *Plant Cell Physiol.* 54, 433–447. doi: 10.1093/pcp/pct022
- Takei, K., Sakakibara, H., and Sugiyama, T. (2001). Identification of genes encoding adenylate isopentenyltransferase, a cytokinin biosynthesis enzyme, in *Arabidopsis thaliana*. *J. Biol. Chem.* 276, 26405–26410. doi: 10.1074/jbc.M102130200
- Takei, K., Yamaya, T., and Sakakibara, H. (2004). *Arabidopsis CYP735A1* and *CYP735A2* encode cytokinin hydroxylases that catalyze the biosynthesis of trans-zeatin. *J. Biol. Chem.* 279, 41866–41872. doi: 10.1074/jbc.M406337200
- Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673
- Tirichine, L., Sandal, N., Madsen, L. H., Radutoiu, S., Albrektsen, A. S., Sato, S., et al. (2007). A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science* 315, 104–107. doi: 10.1126/science.1132397
- van Noorden, G. E., Ross, J. J., Reid, J. B., Rolfe, B. G., and Mathesius, U. (2006). Defective long-distance auxin transport regulation in the *Medicago truncatula* super numeric nodules mutant. *Plant Physiol.* 140, 1494–1506. doi: 10.1104/pp.105.075879
- van Zeijl, A., Den Camp, R. H. O., Deinum, E. E., Charnikhova, T., Franssen, H., Den Camp, H. J. O., et al. (2015). Rhizobium lipo-chitoooligosaccharide signaling triggers accumulation of cytokinins in *Medicago truncatula* roots. *Mol. Plant* 8, 1213–1226. doi: 10.1016/j.molp.2015.03.010
- Vernié, T., Kim, J., Frances, L., Ding, Y., Sun, J., Guan, D., et al. (2015). The NIN transcription factor coordinates diverse nodulation programs in different tissues of the *Medicago truncatula* root. *Plant Cell* 27, 3410–3424. doi: 10.1105/tpc.15.00461
- Yoro, E., Suzuki, T., Toyokura, K., Miyazawa, H., Fukaki, H., and Kawaguchi, M. (2014). A positive regulator of nodule organogenesis, NODULE INCEPTION, acts as a negative regulator of rhizobial infection in *Lotus japonicus*. *Plant Physiol.* 165, 747–758. doi: 10.1104/pp.113.233379

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Azarakhsh, Lebedeva and Lutova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Arabidopsis ALA1 and ALA2 Mediate RNAi-Based Antiviral Immunity

Biyun Zhu^{1†}, Hua Gao^{1†}, Gang Xu¹, Dewei Wu¹, Susheng Song², Hongshan Jiang³, Shuifang Zhu³, Tiancong Qi^{1*} and Daoxin Xie^{1*}

¹ Tsinghua-Peking Joint Center for Life Sciences, and MOE Key Laboratory of Bioinformatics, School of Life Sciences, Tsinghua University, Beijing, China, ² Beijing Key Laboratory of Plant Gene Resources and Biotechnology for Carbon Reduction and Environmental Improvement, College of Life Sciences, Capital Normal University, Beijing, China, ³ The Institute of Plant Quarantine, Chinese Academy of Inspection and Quarantine, Beijing, China

OPEN ACCESS

Edited by:

Jari Valkonen,
University of Helsinki, Finland

Reviewed by:

Zonghua Wang,
Fujian Agriculture and Forestry
University, China
Hong-Gu Kang,
Texas State University, USA

*Correspondence:

Daoxin Xie
daoxinlab@tsinghua.edu.cn
Tiancong Qi
qitiancong@163.com

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 28 December 2016

Accepted: 13 March 2017

Published: 07 April 2017

Citation:

Zhu B, Gao H, Xu G, Wu D, Song S, Jiang H, Zhu S, Qi T and Xie D (2017) Arabidopsis ALA1 and ALA2 Mediate RNAi-Based Antiviral Immunity. *Front. Plant Sci.* 8:422. doi: 10.3389/fpls.2017.00422

RNA interreferencing (RNAi) pathway regulates antiviral immunity and mediates plant growth and development. Despite considerable research efforts, a few components in RNAi pathway have been revealed, including ARGONAUTEs (AGO), DICER-LIKEs (DCLs), RNA-dependent RNA polymerase 1 and 6 (RDR1/6), and ALTERED MERISTEM PROGRAM 1 (AMP1). In this study, we performed a forward genetic screening for enhancers of *rdr6* via inoculation of CMV2aTΔ2b, a 2b-deficient Cucumber Mosaic Virus that is unable to suppress RNAi-mediated antiviral immunity. We uncover that the membrane-localized flippase Aminophospholipid ATPase 1 (ALA1) cooperates with RDR6 and RDR1 to promote antiviral immunity and regulate fertility in *Arabidopsis*. Moreover, we find that ALA2, a homolog of ALA1, also participates in antiviral immunity. Our findings suggest that ALA1 and ALA2 act as novel components in the RNAi pathway and function additively with RDR1 and RDR6 to mediate RNAi-based antiviral immunity and plant development.

Keywords: ALA, *Arabidopsis*, 2b, CMV, RNA interference (RNAi), virus

INTRODUCTION

RNA interference (RNAi) mediates plant defense against virus infections (Ding et al., 2004; Incarbone and Dunoyer, 2013; Martinez de Alba et al., 2013). DICER-LIKE ribonucleases (DCLs), such as DCL4, generate viral short interfering RNAs (siRNAs) (Blevins et al., 2006; Parent et al., 2015), which direct the loading of viral RNAs into ARGONAUTE (AGO) proteins (e.g., AGO1) of the RNA-induced silencing complex (RISC) for the cleavage of viral RNAs (Morel, 2002; Adenot et al., 2006; Arribas-Hernandez et al., 2016), resulting in RNAi-mediated antiviral immunity. RNA-dependent RNA polymerases (RDRs) (Xie et al., 2001; Talmor-Neiman et al., 2006; Cao et al., 2014), including RDR1 and RDR6, promote synthesis of siRNAs by synthesizing long double-strand RNAs (dsRNAs), contributing to the antiviral immunity (Qu et al., 2008; Garcia-Ruiz et al., 2010).

Viruses in turn evolve viral suppressor of RNAi (VSR) to suppress host antiviral immunity. For example, Cucumber Mosaic Virus (CMV) utilizes the VSR protein 2b (Zhang et al., 2006; Diaz-Pendon et al., 2007) to suppress host RNAi-based antiviral immunity and causes severe pathogenic responses in wild-type *Arabidopsis*, while CMV2aTΔ2b, a CMV mutant without expression of 2b protein, is unable to cause any obvious viral symptoms in wild-type and the single mutants of *RDR1* or *RDR6*, but is able to cause severe pathogenic responses in the RNAi-deficient double mutant *rdr1 rdr6* (Wang et al., 2010).

Aminophospholipid transporting ATPases (ALAs) are membrane-localized flippases that are responsible for transporting different lipids, which is essential for asymmetry of membrane lipid bilayers (Lopez-Marques et al., 2010, 2012; Botella et al., 2016). There are 12 *Arabidopsis thaliana* ALAs in the IV subfamily of ATPases that control plant development or tolerance to temperature

stresses (Lopez-Marques et al., 2014). ALA1 is required for plant tolerance to chilling (Gomes et al., 2000). ALA3 regulates pollen germination and pollen tube growth, and adaptability to chilling (Poulsen et al., 2008; McDowell et al., 2013). ALA6 and ALA7 control temperate-regulated pollen tube elongation (McDowell et al., 2015). ALA10 affects lipid uptake to regulate root growth and stomatal control (Botella et al., 2016).

In this study, we performed a forward genetic screening for enhancers of *rdr6* with CMV2aTΔ2b infection on M2 population of ethyl methanesulfonate (EMS)-mutagenized *rdr6*. We show that ALA1 and ALA2 act additively with RDR1 and RDR6 to mediate RNAi-mediated antiviral immunity and development. Our findings discover novel roles of ALA1 and ALA2.

MATERIALS AND METHODS

Materials and Growth Conditions

The *Arabidopsis thaliana* mutants *ala1-2* (Salk_056947), *ala3* (GK-317H04), *ala7* (Salk_125598) and *ala10* (Salk_024877) were obtained from Arabidopsis Biological Resource Center. The *Arabidopsis* mutants *rdr1-1* (SAIL_672_F11), *rdr6-15* (SAIL_617H07), *rdr1 rdr6*, the L1 line transgenic for GUS, and the 2b-deficient CMV mutant CMV2aTΔ2b were described as previously (Boutet et al., 2003; Wang et al., 2010). The *ala1-2 rdr1*, *ala1-2 rdr6* and *ala1-2 rdr1 rdr6* were generated via

genetic crossing. *Nicotiana benthamiana* was grown under a 16-h (28°C)/8-h (22°C) light/dark condition.

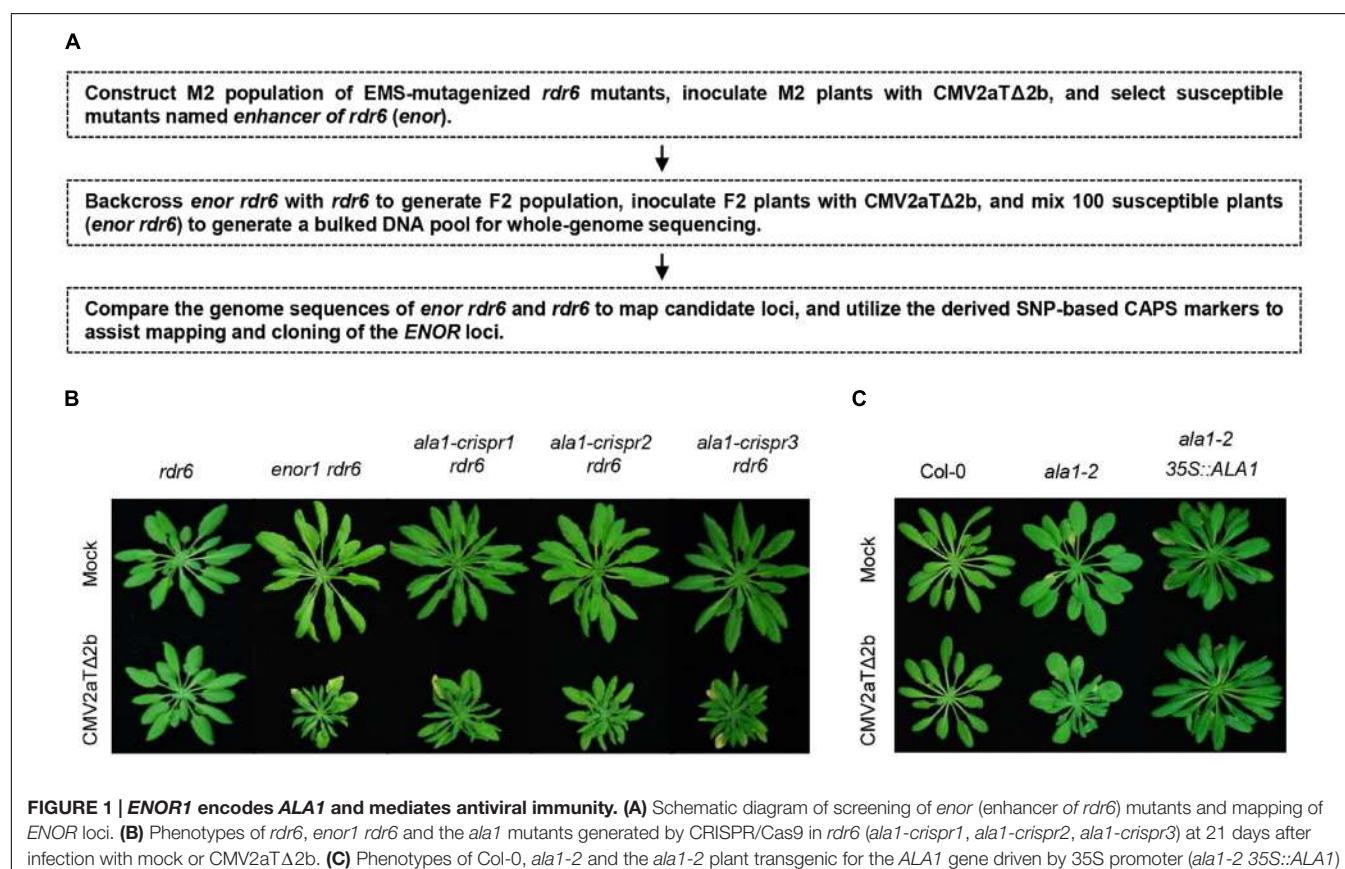
For observation of growth defects in **Figure 5C** and fertility and siliques development in **Figure 3**, *Arabidopsis* seeds were sterilized with 20% bleach, plated on Murashige and Skoog (MS) medium, chilled at 4°C for 3 days, and transferred to a growth room under a 16-h (23–25°C)/8-h (18–20°C) light/dark photoperiod for 9 days. The 9-day-old seedlings were transplanted into soil and grew in the same growth room for another ~3 or ~6 weeks.

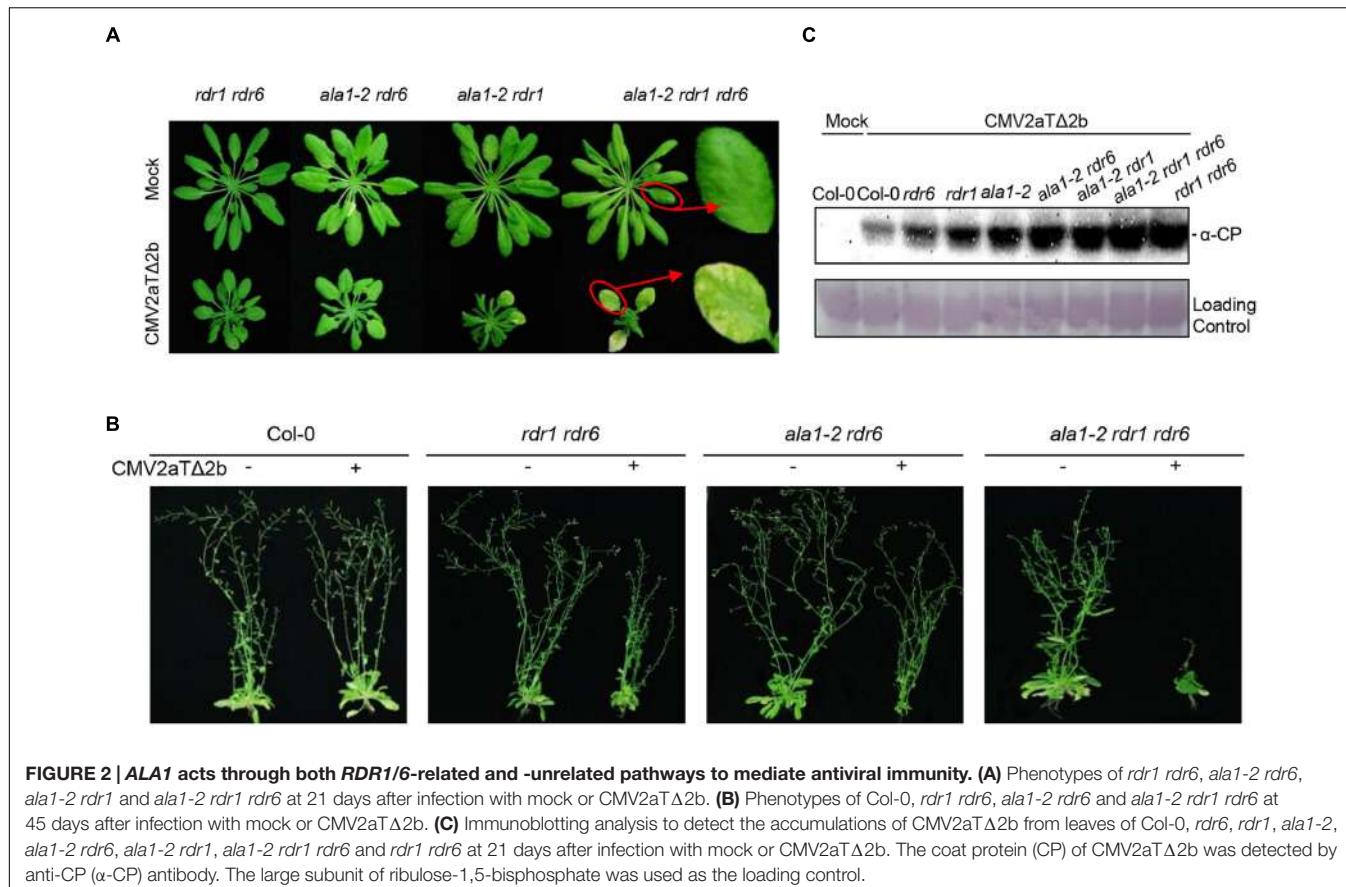
Viral Infection

For viral infection assays, *Arabidopsis* seedlings were sterilized, plated on MS medium, chilled at 4°C for 3 days, and transferred to a growth room under a 16-h (23–25°C)/8-h (18–20°C) light/dark photoperiod for 9 days. The 9-day-old seedlings were transplanted into soil for growth of another 14 days in a growth room under an 8-h (22–24°C)/16-h (16–19°C) light/dark photoperiod. The 23-day-old plants were inoculated with CMV2aTΔ2b as described previously (Wang et al., 2010), and the disease symptoms were recorded at 21 or 45 days after infection.

EMS Mutagenesis

About 20,000 seeds (M1) of the *Arabidopsis* mutant *rdr6-15* (SAIL_617H07) were soaked with 100 mM phosphate buffer





(pH 7.5) overnight at 4°C, washed with sterilized water for five times, and mutagenized with 0.6% ethyl methanesulfonate (EMS) dissolved in phosphate buffer for 8 h at room temperature. The mutagenized seeds were washed with sterilized water for 20 times, and were grown in soil for collection of M2 seeds.

Generation of Mutants and Transgenic Plants

Mutations at 698th (−), 1120th (+) and 2216th (+) bp of coding sequence (CDS) of *ALA1* (Supplementary Figure 1B), and at the 951th (+) bp of CDS of *ALA2* (Supplementary Figure 5B) were introduced into the *rdr6* mutant through CRISPR/Cas9 method (Mao et al., 2013). The guide RNA of the CRISPR target was driven by U6 promoter, and Cas9 was under control of a CaMV35S promoter in a modified pCAMBIA1300 vector (Mao et al., 2013). Primers used for construction of vectors are listed in Supplementary Table 1. The constructs were introduced into *rdr6* mutants through agrobacterium-mediated flower dip method. The transgenic seeds were selected on MS containing 20 mg/L hygromycin, T2 plants were inoculated with CMV2aT Δ 2b. Mutations of *ALA1* or *ALA2* were confirmed by sequencing.

The CDS of *ALA1* was cloned into the pCAMBIA1300 vector through SmaI and XbaI sites for fusion with three FLAG tags under the control of CaMV35S promoter, and introduced into the *ala1-2* using agrobacterium-mediated flower dip method.

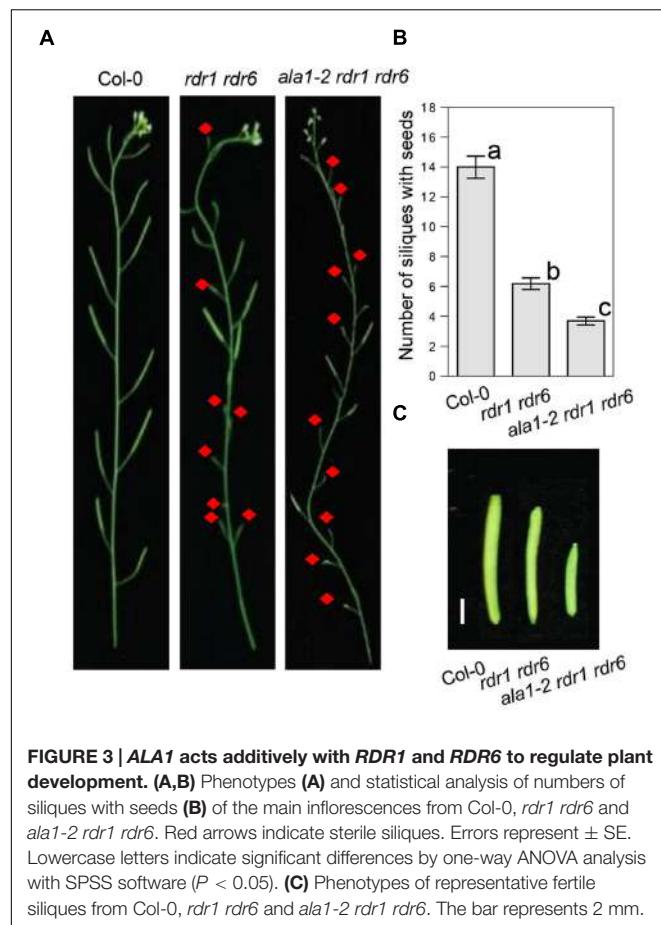
Whole-Genome Sequencing and Gene Cloning of ENOR Loci

The F2 population generated by crossing *enor1 rdr6* or *enor2 rdr6* with *rdr6* were inoculated with CMV2aT Δ 2b. One hundred susceptible plants from F2 population were harvested at 21 days after inoculation to generate a bulked pool for DNA extraction with DNeasy Plant Maxi Kit (QIAGEN, Cat. 68163) and construction of DNA library. Whole genome sequencing was performed with the illumina HiSeq2000 platform. The softwares Skewer, Bowtie2 and SHOREmap were used to analyze the data and isolate mutations (Schneeberger et al., 2009; Sun and Schneeberger, 2015). The SNP-based Cleaved Amplified Polymorphic Sequences (CAPSs) markers generated from comparison of genome sequences of *enor1 rdr6* or *enor2 rdr6* with *rdr6* were used to assist mapping and cloning of *ENOR1* and *ENOR2*.

Immunoblotting Analysis

The total proteins were extracted from plants at 21 days after inoculation with mock or CMV2aT Δ 2b. Fifty microgram of total protein for each sample was quantified and loaded for detection of coat protein (CP) of CMV2aT Δ 2b. The antibody against coat protein (anti-CP) of CMV2aT Δ 2b was produced by Abmart company (Abmart¹) with the recombinant protein

¹<http://www.ab-mart.com.cn/>



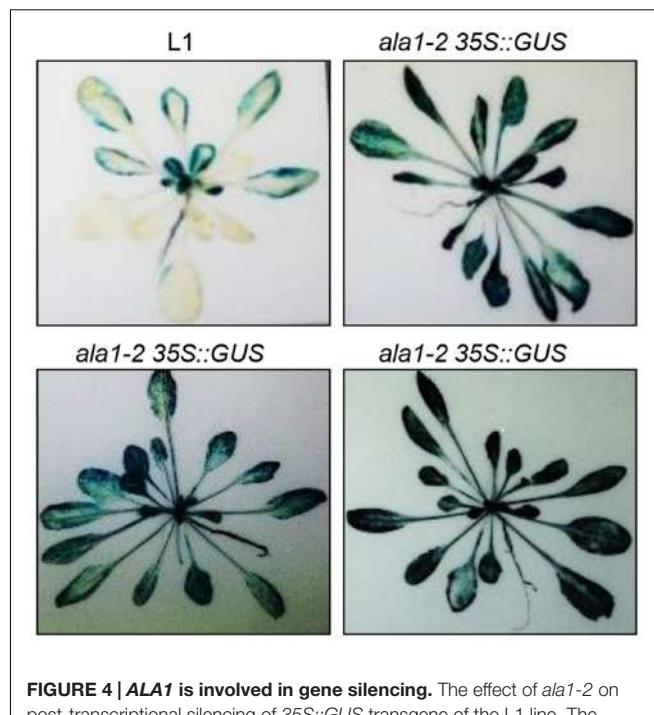
of the 1st to 154th AA of CP. The anti-CP was used as first antibody (1:6000), and anti-rabbit immunoglobulin antibody was used as the secondary antibody respectively (1:3000). All of the experiments were repeated at least three biological times.

GUS Staining

The L1 line transgenic for the β -glucuronidase (*GUS*) gene driven by 35S promoter (35S::GUS), in which the *GUS* activity is very low in all the expanded rosette leaves due to the post transcriptional gene silencing (Boutet et al., 2003), was crossed with *ala1-2* to generate *ala1-2* with L1 transgene (35S::GUS), named *ala1-2 35S::GUS*. Eighteen *ala1-2 35S::GUS* plants were used for histochemical staining of GUS using the method described previously (Shan et al., 2011).

Quantitative Real-Time PCR Analysis

For Figure 6A, the expression of ALA family members was analyzed in Col-0 and *ala1-2* at 21 days after inoculation with mock or CMV2aT Δ 2b. For Supplementary Figure 4, the accumulation of genomic RNA of CMV2aT Δ 2b was analyzed in Col-0 and *ala1-2* at 21 days after CMV2aT Δ 2b inoculation. The primers used for RNA detection of CMV2aT Δ 2b were designed based on the conserved sequences from genomic RNA1 to RNA3 in the 3' prime end. The materials were harvested for RNA extraction using trizol (TRANSGENE, Cat. ET101-01),



and reverse transcription was performed according to the kit (TRANSGENE, Cat. AT311-03). Quantitative real-time PCR was performed with EvaGreen 2*qPCR MasterMix-Low ROX reagents (ABM, Cat. Mastermix-LR) using the ABI7500 real-time PCR system. *ACTIN8* was used as the internal control. All of the experiments were repeated at least three biological times. Primers used for quantitative real-time PCR analysis are listed in Supplementary Table 1.

Phylogenetic Analysis

For the phylogenetic analysis shown in Supplementary Figure 6, the evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The optimal tree with the sum of branch length (4.45679805) is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. The analysis involved all 12 amino acid sequences of ALA family. All positions containing gaps and missing data were eliminated. There were a total of 794 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). The transcripts, including ALA1 (AT5G04930.1), ALA2 (AT5G44240.1), ALA3 (AT1G59820.1), ALA4 (AT1G17500.1), ALA5 (AT1G72700.1), ALA6 (AT1G54280.1), ALA7 (AT3G13900.1), ALA8 (AT3G27870.1), ALA9 (AT1G68710.1), ALA10 (AT3G25610.1),

ALA11 (AT1G13210.1), and ALA12 (AT1G26130.2), were used for phylogenetic analysis.

Subcellular Localization

Coding sequence of *ALA1* was cloned into the pJG054 vector for fusion with YFP under control of CaMV35S promoter (*YFP-ALA1*). The agrobacterium containing *YFP-ALA1* or the mCherry-ER-marker were resuspended in the infiltration buffer (10 mM MgCl₂, 10 mM MES, 0.2 mM acetosyringone) for 3–5 h, and co-infiltrated into leaves of *N. benthamiana*. The fluorescence signals were collected with a Zeiss microscope (LSM710) at ~50 h after co-infiltration. All of the experiments were repeated at least three biological times.

Accession Numbers

The *Arabidopsis* Genome Initiative numbers for genes mentioned in this letter are as follows: ALA1 (AT5G04930), ALA2 (AT5G44240), ALA3 (AT1G59820), ALA4 (AT1G17500), ALA5 (AT1G72700), ALA6 (AT1G54280), ALA7 (AT3G13900), ALA8 (AT3G27870), ALA9 (AT1G68710), ALA10 (AT3G25610), ALA11 (AT1G13210), ALA12 (AT1G26130), RDR1 (AT1G14790), RDR6 (AT3G49500), and ACTIN8 (AT1G49240).

RESULTS

Identification and Mapping of the *enor1* Mutant

We generated M2 population of EMS-mutagenized *rdr6*, and inoculated M2 with CMV2aTΔ2b to identify mutants that enhanced the susceptibility to CMV2aTΔ2b in *rdr6* (referred to as *enhancer of rdr6* [*enor*]), and utilized whole genome sequencing to assist mapping and cloning of *ENOR* loci (Figure 1A).

As shown in Figure 1B, the newly identified mutant *enor1* in the *rdr6* background, named *enor1 rdr6*, exhibited severely stunted and clustered leaves after infection with CMV2aTΔ2b. One fourth of F2 population from the cross between *enor1 rdr6* and *rdr6* were susceptible to CMV2aTΔ2b, demonstrating that *enor1* is a recessive mutation. In order to map the *ENOR* locus, we generated a bulked pool of susceptible plants from the F2 population for whole-genome sequencing, screened mutations by comparing the sequences with SHOREmap methods (Schneeberger et al., 2009; Sun and Schneeberger, 2015), and mapped the *ENOR1* locus using CAPS markers (Figure 1A). We finally found that only a C to T mutation at the 2965th bp of CDS of AT5G04930, which causes a premature stop codon and generates a HaeIII-based CAPS marker, co-segregated with *enor1* (Supplementary Figures 1A,B).

ENOR1 Corresponds to *ALA1* and Is Essential for Antiviral Immunity

AT5G04930 encodes *ALA1* (Lopez-Marques et al., 2014) that co-localizes with the mCherry-ER-marker (Supplementary Figure 2) (Lopez-Marques et al., 2012). To further genetically verify whether AT5G04930 (*ALA1*) corresponds to *ENOR1* and

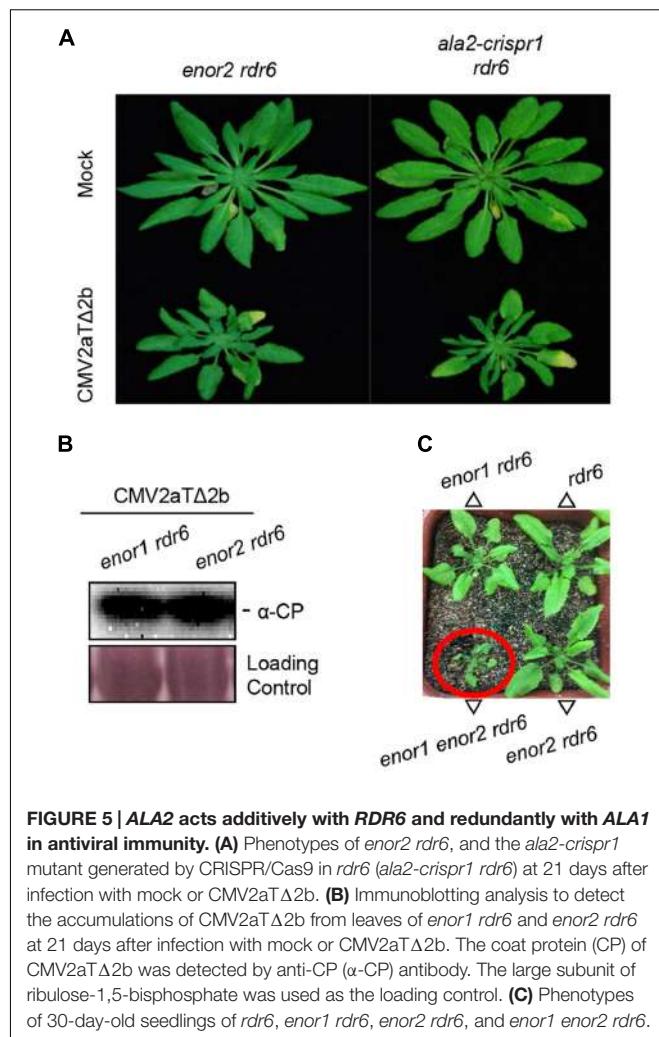


FIGURE 5 | *ALA2* acts additively with *RDR6* and redundantly with *ALA1* in antiviral immunity. **(A)** Phenotypes of *enor2 rdr6*, and the *ala2-crispr1 rdr6* mutant generated by CRISPR/Cas9 in *rdr6* (*ala2-crispr1 rdr6*) at 21 days after infection with mock or CMV2aTΔ2b. **(B)** Immunoblotting analysis to detect the accumulations of CMV2aTΔ2b from leaves of *enor1 rdr6* and *enor2 rdr6* at 21 days after infection with mock or CMV2aTΔ2b. The coat protein (CP) of CMV2aTΔ2b was detected by anti-CP (α -CP) antibody. The large subunit of ribulose-1,5-bisphosphate was used as the loading control. **(C)** Phenotypes of 30-day-old seedlings of *rdr6*, *enor1 rdr6*, *enor2 rdr6*, and *enor1 enor2 rdr6*.

mediates antiviral immunity, we generated *ala1* mutants by the CRISPR/Cas9 genome editing method (Feng et al., 2014; Jia et al., 2016) in the *rdr6* background, and examined whether these *ala1-crispr rdr6* double mutants exhibit the viral symptoms similar to that of *enor1 rdr6* when inoculated with CMV2aTΔ2b. As shown in Figure 1B, all the *ala1-crispr rdr6* double mutants were severely susceptible to CMV2aTΔ2b, demonstrating that *ALA1* corresponds to *ENOR1* and is required for antiviral immunity.

We also obtained a T-DNA insertional mutant (Salk_056947, named *ala1-2*) of *ALA1* (Supplementary Figure 1B), and found that the *ala1-2* single mutant was mildly susceptible to CMV2aTΔ2b, less severe than *ala1-crispr rdr6* (Figures 1B,C), which also supports the *ALA1* function in antiviral immunity. Moreover, we found that transgenic expression of *ALA1* under the control of CaMV 35S promoter fully restored the mutant phenotypes of *ala1-2* (Figure 1C).

ALA1 Acts Additively With *RDR1* and *RDR6* to Regulate Antiviral Immunity

Further analyses of various double mutants and the triple mutant *ala1-2 rdr1 rdr6* showed that all the double mutants including

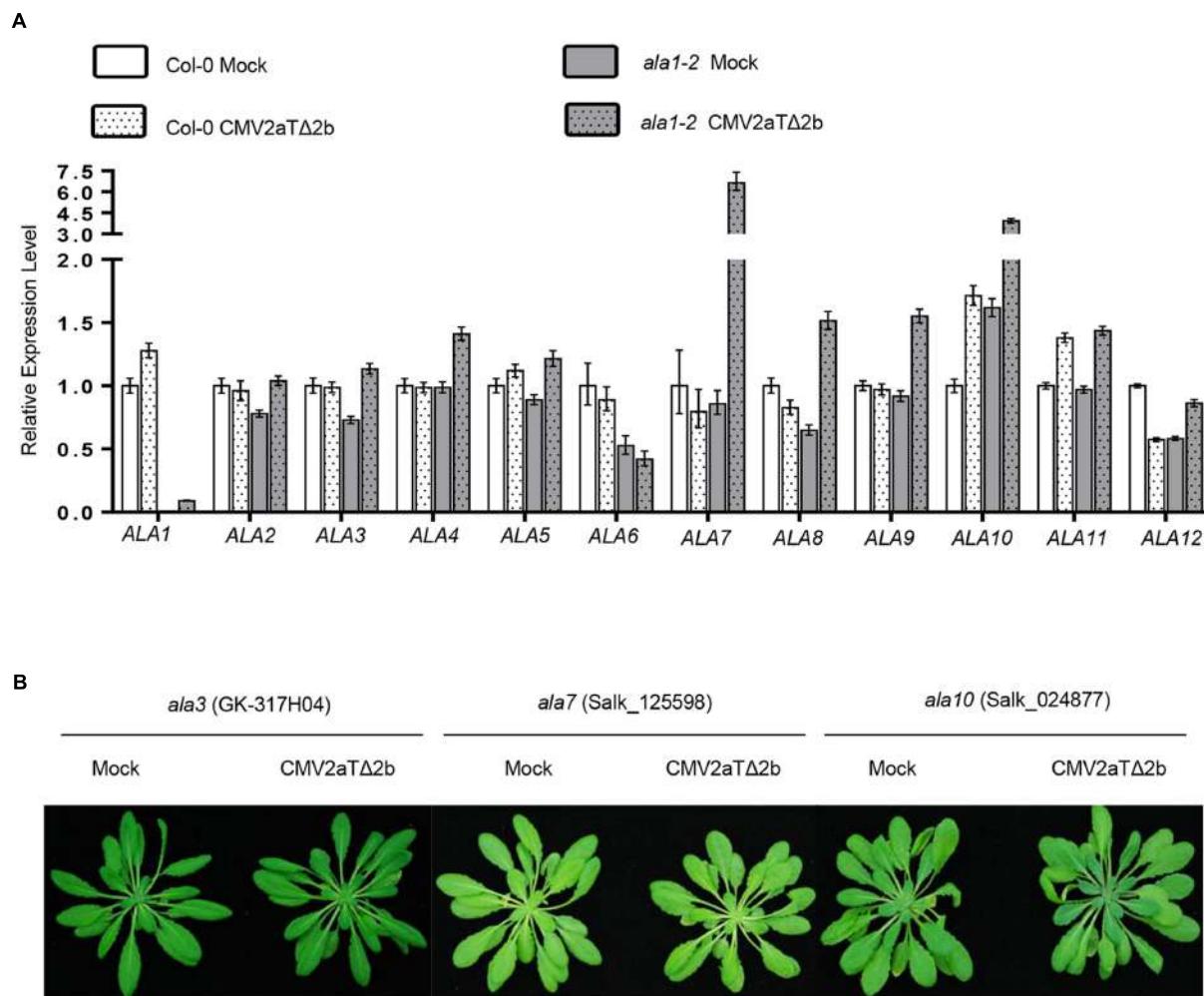


FIGURE 6 | Analysis of ALA members in antiviral silencing. (A) Quantitative real-time PCR showed the relative expression levels of ALA family members in Col-0 and *ala1-2* after inoculation with mock or CMV2aT Δ 2b. The data are means (\pm SE) from three biological repeats. **(B)** Phenotypes of the *ala3*, *ala7* and *ala10* mutants at 21 days after infection with mock or CMV2aT Δ 2b.

ala1-2 rdr6, *ala1-2 rdr1*, *enor1 rdr6*, *ala1-crispr rdr6* and *rdr1 rdr6* exhibited similar symptoms after inoculation with CMV2aT Δ 2b, which were much more severe than the single mutant *ala1-2*, while the triple mutant *ala1-2 rdr1 rdr6* showed the most severe symptoms with over-stunted newly born leaves and yellow old chlorotic leaves (Figures 1B,C, 2A,B and Supplementary Figure 3). These results suggest that *ALA1* functions additively with *RDR1* and *RDR6* to mediate plant immunity.

The immunoblot analysis with antibody against the CP of CMV2aT Δ 2b showed that CMV2aT Δ 2b accumulated much more in *ala1-2* than in wild-type, and that the double mutants (*ala1-2 rdr6*, *ala1-2 rdr1* and *rdr1 rdr6*) accumulated much more CP than the corresponding single mutants (Figure 2C). These results further demonstrate that *ALA1* acts additively with *RDR1* and *RDR6* to mediate RNAi-based antiviral immunity. Interestingly, the triple mutant *ala1-2 rdr1 rdr6* showed enhanced susceptibility compared with the double mutant *rdr1 rdr6* when inoculated with CMV2aT Δ 2b, however, the accumulation of

CMV2aT Δ 2b was indistinguishable between the triple mutant *ala1-2 rdr1 rdr6* and the double mutant *rdr1 rdr6*, implying that *ALA1* mediates plant immunity through both a *RDR1/6*-related RNAi pathway and *RDR1/6*-unrelated pathways.

Further phenotypic analysis showed that the *ala1-2 rdr1 rdr6* triple mutant also displays developmental defects, including shorter siliques and less fertile siliques (Figures 3A–C). These results imply that *ALA1* may function additively with *RDR1* and *RDR6* to mediate RNAi-regulated plant development, consistent with the previous observations that RNAi, in addition to the RNAi-mediated plant immunity, also mediates plant developmental processes (Yoshikawa et al., 2005).

ALA1 Is Required for Gene Silencing

Having shown that *ALA1* acts additively with *RDR1* and *RDR6* in RNAi-based antiviral immunity and development, we further verified whether *ALA1* affects gene silencing via genetic cross of the *ala1-2* mutant with the L1, a transgenic silencing marker

line where the *GUS* transgene driven by the CaMV35S promoter (35S::GUS) was silenced and expressed at low level (Boutet et al., 2003). As shown in **Figure 4**, the *GUS* activity was obviously increased in *ala1-2* (named *ala1-2* 35S::GUS). These data demonstrate that mutation in *ALA1* abolished the gene silencing on the *GUS* transgene driven by the 35S promoter, suggesting that *ALA1* is indeed required for gene silencing. Consistently, our quantitative real-time PCR analysis showed that the accumulation of CMV2aTΔ2b RNA in *ala1-2* was much higher than that in Col-0, further supporting the essential roles of *ALA1* in gene silencing and antiviral defense.

ALA2 Also Participates in Antiviral Immunity

During the screening, we isolated a second enhancer mutant *enor2 rdr6* (**Figure 5A**), in which CP accumulation was similar with that in *enor1 rdr6* (**Figure 5B**). We further found that *ENOR2* encodes *ALA2* by performing the same mapping and identification procedures as *ENOR1* (Supplementary Figure 5A). The *ALA2* gene in *enor2 rdr6* contained a G to A mutation at the 1995th bp, leading to a premature stop codon (Supplementary Figure 5B), and mutation of *ALA2* by CRISPR/Cas9 in *rdr6* also resulted in severe susceptibility to CMV2aTΔ2b (**Figure 5A** and Supplementary Figure 5B), suggesting that *ALA2* mediates antiviral immunity. Moreover, we generated the *enor1 enor2 rdr6* triple mutant, and found that *enor1 enor2 rdr6* displayed severe developmental defects, including stunted leaves, which is similar with CMV2aTΔ2b-infected *enor1 rdr6* and *enor2 rdr6* (**Figure 5C**). This results (**Figure 5**) indicate that both *ALA1* and *ALA2* act additively with *RDR6* to mediate antiviral immunity and plant growth.

Analysis of Other ALAs in Antiviral Immunity

Phylogenetic analysis of the ALA family proteins showed that *ALA1* and *ALA2* are the closest members, and other members are less related (Supplementary Figure 6). We observed that CMV2aTΔ2b infection dramatically induced the expression of *ALA7* and *ALA10* in *ala1-2*, but could not obviously affect the expression of other ALAs in wild-type and *ala1-2* (**Figure 6A**). We next investigated whether other ALA members play a role in antiviral immunity. The T-DNA insertion mutants of *ALA3* to *ALA12* were inoculated with CMV2aTΔ2b, and the results showed that none of these mutants were susceptible (**Figure 6B** and data not shown). It remains to be elucidated whether these ALAs function redundantly to mediate RNAi-based antiviral immunity and plant development.

DISCUSSION

It is well known that the RNAi pathway regulates plant growth, development and immunity. Previous studies have revealed that AGOs, DCLs, RDR1 and RDR6 are essential components of RNAi pathway (Ding and Voinnet, 2007; Qu et al., 2008; Wang et al., 2010; Cao et al., 2014). In this

study, we developed an effective forward genetic screening using 2b-deficient CMV2aTΔ2b, and defined *ALA1* and *ALA2*, membrane-localized proteins (**Figures 1, 5** and Supplementary Figure 2) (Lopez-Marques et al., 2010, 2012), as the new components in the RNAi pathway. *ALA1* plays an essential role in gene silencing, and acts additively with *RDR1/6* to mediate RNAi-based antiviral immunity and plant development (**Figures 2–4**). *ALA2* also participates in antiviral defense and development, and acts redundantly with *ALA1* in regulation of plant development in *rdr6* background (**Figure 5C**).

A recent study showed that AMP1, a novel key component in RNAi pathway, associates with AGO1 and mediates miRNA-targeted translational inhibition of mRNA on ER membrane (Li et al., 2013). miRNA-guided cleavage can also occur on ER membrane-bound polysomes (Li et al., 2016). These studies take ER into a central stage of small RNAs-mediated silencing (Ma et al., 2013; Li et al., 2016). On the other hand, viruses recruit ER membrane and manipulate lipid synthesis, transport and metabolism to form a circumstance essential for viral replication and morphogenesis (Fernández de Castro et al., 2016). Our finding that the ER membrane-localized *ALA1* and *ALA2* are essential players in silencing pathway and antiviral immunity would help to study and understand both the small RNAs machinery on ER membrane and the roles of lipid transport in silencing and antiviral defense. It would be interesting to investigate whether *ALA1* and *ALA2* associate with AMP1 and/or AGO1 to mediate gene silencing and antiviral immunity.

AUTHOR CONTRIBUTIONS

DX designed the study; BZ, HG, DW, and TQ performed experiments; DX, TQ, BZ, HG, GX, SS, HJ, and SZ analyzed the data. BZ, HG, TQ, and DX wrote the manuscript.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (Grant NO. 31230008 and 31630085).

ACKNOWLEDGMENTS

We thank Dr. Shouwei Ding (University of California, Berkeley) for sharing the *rdr1*, *rdr6* and *rdr1 rdr6* mutants, and Dr. Jiankang Zhu (Center for Plant Stress Biology, Chinese Academy of Science) for the CRISPR/Cas9 vector.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.00422/full#supplementary-material>

REFERENCES

- Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouche, N., Gasciolli, V., et al. (2006). DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. *Curr. Biol.* 16, 927–932. doi: 10.1016/j.cub.2006.03.035
- Arribas-Hernandez, L., Marchais, A., Poulsen, C., Haase, B., Hauptmann, J., Benes, V., et al. (2016). The slicer activity of ARGONAUTE1 is required specifically for the phasing, not production, of trans-acting short interfering RNAs in *Arabidopsis*. *Plant Cell* 28, 1563–1580. doi: 10.1105/tpc.16.00121
- Blevins, T., Rajeswaran, R., Shivaprasad, P. V., Beknazarians, D., Si-Amour, A., Park, H. S., et al. (2006). Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res.* 34, 6233–6246. doi: 10.1093/nar/gkl886
- Botella, C., Sautron, E., Boudiere, L., Michaud, M., Dubots, E., Yamaryo-Botte, Y., et al. (2016). ALA10, a phospholipid flippase, controls FAD2/FAD3 desaturation of phosphatidylcholine in the ER and affects chloroplast lipid composition in *Arabidopsis thaliana*. *Plant Physiol.* 170, 1300–1314. doi: 10.1104/pp.15.01557
- Boutet, S., Vazquez, F., Liu, J., Béclin, C., Fagard, M., Gratias, A., et al. (2003). *Arabidopsis* HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. *Curr. Biol.* 13, 843–848. doi: 10.1016/S0960-9822(03)00293-8
- Cao, M., Du, P., Wang, X., Yu, Y. Q., Qiu, Y. H., Li, W., et al. (2014). Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14613–14618. doi: 10.1073/pnas.1407131111
- Diaz-Pendon, J. A., Li, F., Li, W. X., and Ding, S. W. (2007). Suppression of antiviral silencing by cucumber mosaic virus 2b protein in *Arabidopsis* is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell* 19, 2053–2063. doi: 10.1105/tpc.106.047449
- Ding, S. W., Li, H., Lu, R., Li, F., and Li, W. X. (2004). RNA silencing: a conserved antiviral immunity of plants and animals. *Virus Res.* 102, 109–115. doi: 10.1016/j.virusres.2004.01.021
- Ding, S. W., and Voinnet, O. (2007). Antiviral immunity directed by small RNAs. *Cell* 130, 413–426. doi: 10.1016/j.cell.2007.07.039
- Felsenstein, J. (1985). Confidence-limits on phylogenies - an approach using the Bootstrap. *Evolution* 39, 783–791. doi: 10.2307/2408678
- Feng, Z. Y., Mao, Y. F., Xu, N. F., Zhang, B. T., Wei, P. L., Yang, D. L., et al. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4632–4637. doi: 10.1073/pnas.1400822111
- Fernández de Castro, I., Tenorio, R., and Risco, C. (2016). Virus assembly factories in a lipid world. *Curr. Opin. Virol.* 18, 20–26. doi: 10.1016/j.coviro.2016.02.009
- Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Brembelis, K. J., et al. (2010). *Arabidopsis* RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *Plant Cell* 22, 481–496. doi: 10.1105/tpc.109.073056
- Gomes, E., Jakobsen, M. K., Axelsen, K. B., Geisler, M., and Palmgren, M. G. (2000). Chilling tolerance in *Arabidopsis* involves ALA1, a member of a new family of putative aminophospholipid translocases. *Plant Cell* 12, 2441–2453. doi: 10.1105/tpc.12.12.2441
- Incarbone, M., and Dunoyer, P. (2013). RNA silencing and its suppression: novel insights from *in planta* analyses. *Trends Plant Sci.* 18, 382–392. doi: 10.1016/j.tplants.2013.04.001
- Jia, Y. X., Ding, Y. L., Shi, Y. T., Zhang, X. Y., Gong, Z. Z., and Yang, S. H. (2016). The cbf3 triple mutants reveal the essential functions of CBFs in cold acclimation and allow the definition of CBF regulons in *Arabidopsis*. *New Phytol.* 212, 345–353. doi: 10.1111/nph.14088
- Li, S., Le, B., Ma, X., Li, S., You, C., Yu, Y., et al. (2016). Biogenesis of phased siRNAs on membrane-bound polysomes in *Arabidopsis*. *Elife* 5:e22750. doi: 10.7554/elife.22750
- Li, S., Liu, L., Zhuang, X., Yu, Y., Liu, X., Cui, X., et al. (2013). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in *Arabidopsis*. *Cell* 153, 562–574. doi: 10.1016/j.cell.2013.04.005
- Lopez-Marques, R. L., Poulsen, L. R., Hanisch, S., Meffert, K., Buch-Pedersen, M. J., Jakobsen, M. K., et al. (2010). Intracellular targeting signals and lipid specificity determinants of the ALA/ALIS P4-ATPase complex reside in the catalytic ALA alpha-subunit. *Mol. Biol. Cell* 21, 791–801. doi: 10.1091/mbc.E09-08-0656
- Lopez-Marques, R. L., Poulsen, L. R., and Palmgren, M. G. (2012). A putative plant aminophospholipid flippase, the *Arabidopsis* P4 ATPase ALA1, localizes to the plasma membrane following association with a beta-subunit. *PLoS ONE* 7:e33042. doi: 10.1371/journal.pone.0033042
- Lopez-Marques, R. L., Theorin, L., Palmgren, M. G., and Pomorski, T. G. (2014). P4-ATPases: lipid flippases in cell membranes. *Pflugers Arch.* 466, 1227–1240. doi: 10.1007/s00424-013-1363-4
- Ma, X., Cao, X. F., Mo, B. X., and Chen, X. M. (2013). Trip to ER MicroRNA-mediated translational repression in plants. *RNA Biol.* 10, 1586–1592. doi: 10.4161/rna.26313
- Mao, Y., Zhang, H., Xu, N., Zhang, B., Gou, F., and Zhu, J. K. (2013). Application of the CRISPR-Cas system for efficient genome engineering in plants. *Mol. Plant* 6, 2008–2011. doi: 10.1093/mp/sst121
- Martinez de Alba, A. E., Elvira-Matelet, E., and Vaucheret, H. (2013). Gene silencing in plants: a diversity of pathways. *Biochim. Biophys. Acta* 1829, 1300–1308. doi: 10.1016/j.bbagen.2013.10.005
- McDowell, S. C., Lopez-Marques, R. L., Cohen, T., Brown, E., Rosenberg, A., Palmgren, M. G., et al. (2015). Loss of the *Arabidopsis thaliana* P4-ATPases ALA6 and ALA7 impairs pollen fitness and alters the pollen tube plasma membrane. *Front. Plant Sci.* 6:197. doi: 10.3389/fpls.2015.00197
- McDowell, S. C., Lopez-Marques, R. L., Poulsen, L. R., Palmgren, M. G., and Harper, J. F. (2013). Loss of the *Arabidopsis thaliana* P(4)-ATPase ALA3 reduces adaptability to temperature stresses and impairs vegetative, pollen, and ovule development. *PLoS ONE* 8:e62577. doi: 10.1371/journal.pone.0062577
- Morel, J. B. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell Online* 14, 629–639. doi: 10.1105/tpc.010358
- Parent, J. S., Bouteiller, N., Elmayan, T., and Vaucheret, H. (2015). Respective contributions of *Arabidopsis* DCL2 and DCL4 to RNA silencing. *Plant J.* 81, 223–232. doi: 10.1111/tpj.12720
- Poulsen, L. R., Lopez-Marques, R. L., McDowell, S. C., Okker, J., Licht, D., Schulz, A., et al. (2008). The *Arabidopsis* P-4-ATPase ALA3 localizes to the Golgi and requires a beta-subunit to function in lipid translocation and secretory vesicle formation. *Plant Cell* 20, 658–676. doi: 10.1105/tpc.107.054767
- Qu, F., Ye, X., and Morris, T. J. (2008). *Arabidopsis* DRB4,AGO1,AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14732–14737. doi: 10.1073/pnas.0805760105
- Saitou, N., and Nei, M. (1987). The Neighbor-Joining Method - a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A. H., Nielsen, K. L., et al. (2009). SHOREmap: simultaneous mapping and mutation identification by deep sequencing. *Nat. Methods* 6, 550–551. doi: 10.1038/nmeth.0809-550
- Shan, X., Wang, J., Chua, L., Jiang, D., Peng, W., and Xie, D. (2011). The role of *Arabidopsis* Rubisco activase in jasmonate-induced leaf senescence. *Plant Physiol.* 155, 751–764. doi: 10.1104/pp.110.166595
- Sun, H., and Schneeberger, K. (2015). SHOREmap v3.0: fast and accurate identification of causal mutations from forward genetic screens. *Methods Mol. Biol.* 1284, 381–395. doi: 10.1007/978-1-4939-2444-8_19
- Talmor-Neiman, M., Stav, R., Klipcan, L., Buxdorf, K., Baulcombe, D. C., and Arazi, T. (2006). Identification of trans-acting siRNAs in moss and an RNA-dependent RNA polymerase required for their biogenesis. *Plant J.* 48, 511–521. doi: 10.1111/j.1365-313X.2006.02895.x
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Wang, X. B., Wu, Q., Ito, T., Cillo, F., Li, W. X., Chen, X., et al. (2010). RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 484–489. doi: 10.1073/pnas.0904086107
- Xie, Z., Fan, B., Chen, C., and Chen, Z. (2001). An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6516–6521. doi: 10.1073/pnas.111440998

- Yoshikawa, M., Peragine, A., Park, M. Y., and Poethig, R. S. (2005). A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. *Genes Dev.* 19, 2164–2175. doi: 10.1101/gad.1352605
- Zhang, X., Yuan, Y. R., Pei, Y., Lin, S. S., Tuschl, T., Patel, D. J., et al. (2006). Cucumber mosaic virus-encoded 2b suppressor inhibits *Arabidopsis* Argonaute1 cleavage activity to counter plant defense. *Genes Dev.* 20, 3255–3268. doi: 10.1101/gad.1495506
- Zuckerkandl, E., and Pauling, L. (1965). Molecules as documents of evolutionary history. *J. Theor. Biol.* 8, 357–366. doi: 10.1016/0022-5193(65)90083-4

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Zhu, Gao, Xu, Wu, Song, Jiang, Zhu, Qi and Xie. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Convergent Evolution of Pathogen Effectors toward Reactive Oxygen Species Signaling Networks in Plants

Nam-Soo Jwa^{1*} and Byung Kook Hwang²

¹ Division of Integrative Bioscience and Biotechnology, College of Life Sciences, Sejong University, Seoul, South Korea,

² Laboratory of Molecular Plant Pathology, College of Life Sciences and Biotechnology, Korea University, Seoul, South Korea

OPEN ACCESS

Edited by:

Tatiana Matveeva,
Saint Petersburg State University,
Russia

Reviewed by:

Jin-Long Qiu,
Institute of Microbiology (CAS), China
Pierre Pétriacq,
University of Sheffield,
United Kingdom

***Correspondence:**

Nam-Soo Jwa
nsjwa@sejong.ac.kr

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 19 June 2017

Accepted: 13 September 2017

Published: 29 September 2017

Citation:

Jwa N-S and Hwang BK (2017)
Convergent Evolution of Pathogen Effectors toward Reactive Oxygen Species Signaling Networks in Plants.
Front. Plant Sci. 8:1687.

doi: 10.3389/fpls.2017.01687

Microbial pathogens have evolved protein effectors to promote virulence and cause disease in host plants. Pathogen effectors delivered into plant cells suppress plant immune responses and modulate host metabolism to support the infection processes of pathogens. Reactive oxygen species (ROS) act as cellular signaling molecules to trigger plant immune responses, such as pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) and effector-triggered immunity. In this review, we discuss recent insights into the molecular functions of pathogen effectors that target multiple steps in the ROS signaling pathway in plants. The perception of PAMPs by pattern recognition receptors leads to the rapid and strong production of ROS through activation of NADPH oxidase Respiratory Burst Oxidase Homologs (RBOHs) as well as peroxidases. Specific pathogen effectors directly or indirectly interact with plant nucleotide-binding leucine-rich repeat receptors to induce ROS production and the hypersensitive response in plant cells. By contrast, virulent pathogens possess effectors capable of suppressing plant ROS bursts in different ways during infection. PAMP-triggered ROS bursts are suppressed by pathogen effectors that target mitogen-activated protein kinase cascades. Moreover, pathogen effectors target vesicle trafficking or metabolic priming, leading to the suppression of ROS production. Secreted pathogen effectors block the metabolic coenzyme NADP-malic enzyme, inhibiting the transfer of electrons to the NADPH oxidases (RBOHs) responsible for ROS generation. Collectively, pathogen effectors may have evolved to converge on a common host protein network to suppress the common plant immune system, including the ROS burst and cell death response in plants.

Keywords: pathogen effector, reactive oxygen species, PAMP-triggered immunity, effector-triggered immunity, respiratory burst oxidase homolog, mitogen-activated protein kinase

INTRODUCTION

Plants have evolved sophisticated defense mechanisms to resist potential attacks by microbial pathogens (Grant and Loake, 2000). The first line of defense is triggered in plants by the perception of microbe- or pathogen-associated molecular patterns (MAMPs or PAMPs) via membrane-bound pattern recognition receptors (PRRs), leading to basal immunity, known as PAMP-triggered immunity (PTI) (Gómez-Gómez and Boller, 2000; Jones and Dangl, 2006). PAMPs include conserved cell surface structures including bacterial flagellin, lipopolysaccharides,

and peptidoglycan, or fungal cell wall components such as glucan or chitin (Zipfel et al., 2004; Torres, 2010). Plants may disrupt numerous non-host or host pathogen attacks via PTI; however, adapted pathogens can overcome the PTI-dependent defense response to cause disease on their host plants (Collins et al., 2003; He et al., 2006). PTI requires signal transduction from receptors to downstream components via mitogen-activated protein kinase (MAPK) cascade pathways (Pitzschke et al., 2009). Known PAMPs activate MAP kinases in plant cells. In the second line of defense, plants have acquired a cell-based surveillance system using intracellular nucleotide-binding leucine-rich repeat (NLR) receptors to recognize specific pathogen effectors, leading to resistance (*R*) gene-mediated effector-triggered immunity (ETI) (Jones and Dangl, 2006; Dangl et al., 2013). The two phases of plant immunity may be spatiotemporally distinct but are intimately related to the reactive oxygen species (ROS) burst (Grant and Loake, 2000; Torres et al., 2006; Kadota et al., 2015). Production of ROS in plant cells is a hallmark of successful recognition of plant pathogens and activation of plant defenses (Lamb and Dixon, 1997; Torres, 2010). Pathogen-induced apoplastic ROS production was first demonstrated in potato tuber tissues by Doke (1983) and indeed ROS play important roles in plant immune responses as signaling molecules (Torres, 2010; Mittler et al., 2011; O'Brien et al., 2012a; Frederickson Matika and Loake, 2014; Lehmann et al., 2015). In plant cells, ROS also occur in response to many physiological stimuli (Mori and Schroeder, 2004; Mittler et al., 2011; Baxter et al., 2014).

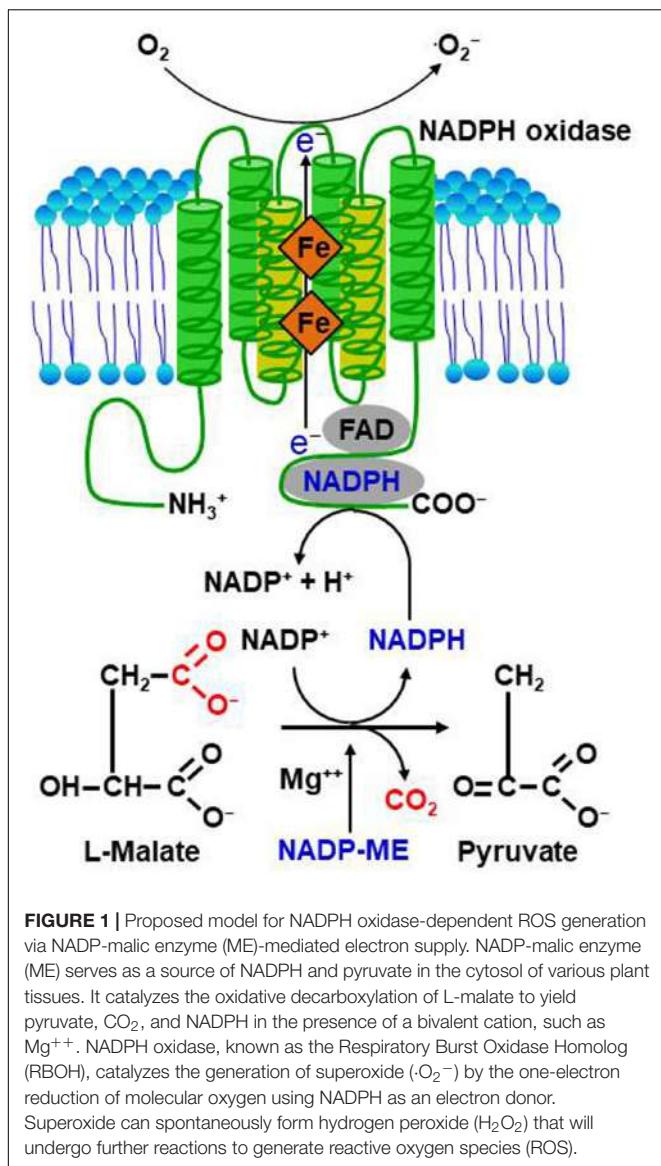
Reactive oxygen species are highly reactive reduced oxygen molecules, such as superoxide ($\cdot\text{O}_2^-$), hydrogen peroxide (H_2O_2), and hydroxyl radical ($\cdot\text{OH}$) (Grant and Loake, 2000; Mori and Schroeder, 2004; Sagi and Fluhr, 2006). There is comprehensive evidence that ROS can act as cellular signaling molecules, mediating various important responses of plant cells to different physiological stimuli, including pathogen attack, abiotic stress, hormone signaling, and polar growth (Mori and Schroeder, 2004; Torres and Dangl, 2005). ROS are formed intracellularly during certain redox reactions in the cell membranes, cytoplasm, nuclei, mitochondria, chloroplasts, peroxisomes, and endoplasmic reticulum (Ashtamker et al., 2007; Liu et al., 2007; Torres, 2010; La and López-Huertas, 2016). The endomembrane and nuclear compartments are likely targets or sources for ROS signaling (Ashtamker et al., 2007). Pathogen-induced ROS generation in chloroplasts is known to play a crucial role in the signaling for and/or execution of hypersensitive response (HR) cell death in plants (Liu et al., 2007). In addition, mitochondrial ROS associated with alteration in respiration are likely to activate defense responses, but could not be directly involved in plant cell death (Vidal et al., 2007). The generation of extracellular ROS, including H_2O_2 , requires the extracellular activities of cell wall peroxidases and plasma membrane NADPH oxidases in plant cells (Bienert and Chaumont, 2014). NADPH oxidase-dependent ROS generation with electron supply provided by NADP-malic enzyme (ME) is presented in **Figure 1**. H_2O_2 is one of the most abundant and stable ROS in plants, and ROS generated inside cells can move apoplastically as H_2O_2 into neighboring cells (Allan and Fluhr, 1997). ROS signal transduction activates Ca^{2+} channels

in the plasma membrane and may be a central step in many ROS-mediated processes regulating the physiology of plant cells (Mori and Schroeder, 2004). It has also been proposed that biomembrane channels (aquaporins) mediate H_2O_2 transport across biological membranes to control ROS signaling in plant cells (Bienert and Chaumont, 2014; Tian et al., 2016). The primary ROS bursts after perception of pathogens occur in the apoplasts; however, ROS produced in different compartments inside the plant cell may function in plant defenses to pathogen invasion (Torres, 2010).

Reactive oxygen species are an effective weapon that can be produced rapidly and utilized against pathogen infection (Levine et al., 1994). H_2O_2 and O_2^- are mainly produced at the site of attempted pathogen invasion in plant cells (Apostol et al., 1989). They are secreted from the cell within 3 min of recognition of MAMPs or PAMPs (Chinchilla et al., 2007; Nühse et al., 2007; Hedrich, 2012). The function of H_2O_2 in inhibiting pathogen growth is fully understood (Chi et al., 2009; Huang et al., 2011; Park et al., 2013). Plant-derived ROS act as a powerful weapon against pathogen invasion. By contrast, pathogens have evolved strategies to reduce plant ROS bursts in different ways during infection. For example, pathogens may avoid the risk of the ROS burst by using effectors. However, in the presence of effector-aware and defensible plants, pathogens are often faced with mutation pressure for virulence (Jones and Dangl, 2006). The competition between pathogens and host plants, called 'arms races' (Boller and He, 2009), may lead to the creation of numerous types of effectors in microbial pathogens and also R- or defense-related proteins in plants. Effectors from these evolutionarily diverse pathogens are predicted to converge on common host plant proteins, which are characterized by a high degree of interaction in host plant protein networks (Weßling et al., 2014; Rovenich et al., 2016). Here, we review and discuss recent advances in understanding how microbial pathogen effectors have evolved toward the suppression of plant apoplastic ROS bursts during infection.

PAMP-MEDIATED ROS BURST THAT TRIGGERS BASAL IMMUNE RESPONSES

Perception of PAMPs by plants via PRRs triggers ROS production through activation of NADPH oxidases as well as peroxidases, leading to PTI-dependent basal defenses that inhibit invading pathogens (**Figure 2**). The apoplastic ROS bursts generated in elicited plant cells are sufficiently cytotoxic to kill invading pathogens (Legendre et al., 1993; Chi et al., 2009; Park et al., 2013). ROS also act as signaling molecules, triggering plant immune and cell death responses (Tenhaken et al., 1995; Jabs, 1999; Torres, 2010). Thus, pathogens need to take steps to avoid exposure to toxic ROS. NADPH oxidases, also known as Respiratory Burst Oxidase Homologs (RBOHs), are responsible for production of ROS in plants during pathogen infection (Torres and Dangl, 2005; Torres, 2010). NADPH oxidase (RBOHD) phosphorylation by the PRR-associated kinase BIK1 has been proposed to be essential for PAMP-triggered ROS



production (Kadota et al., 2014). In addition, the apoplastic peroxidase-dependent ROS burst plays an important role in *Arabidopsis* PTI mediated by the recognition of PAMPs (Daudi et al., 2012). Antisense expression of a heterologous French bean (*Phaseolus vulgaris*) peroxidase (*FBP1*) cDNA in *Arabidopsis* diminishes the expression of *Arabidopsis* peroxidases PRX33 and PRX34, blocking the ROS burst in response to a *Fusarium oxysporum* cell wall elicitor, and leading to enhanced susceptibility to fungal and bacterial pathogens (Bindschedler et al., 2006; Daudi et al., 2012; O'Brien et al., 2012b). Similarly, pepper extracellular peroxidase CaPO2 generates ROS bursts, activating local and systemic cell death and defense response to bacterial pathogens (Choi et al., 2007). Recently, the plant aquaporin AtPIP4 has been demonstrated to trigger cytoplasmic import of apoplastic H_2O_2 into plant cells, activating systemic acquired resistance (SAR) and PTI pathways in response to *Pseudomonas syringae* pv. *tomato* DC3000 and two typical

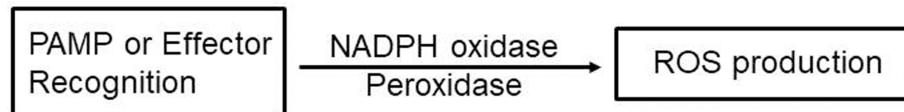
PAMPs (flagellin and chitin), respectively (Tian et al., 2016). This suggests a pivotal role for aquaporins in apoplastic ROS signal transduction in disease immunity pathways.

Plants defend themselves against invading pathogens through cell wall reinforcement. Cell wall fortifications are facilitated by an apoplastic H_2O_2 burst, cell wall cross-linking, and callose deposition at the site of infection (Bradley et al., 1992; Deepak et al., 2007; Luna et al., 2011; Ellinger et al., 2013). Effector (Elicitor)-induced oxidative cross-linking of plant cell wall structural proteins is essential for cell maturation and toughening of cell walls in the initial stages of plant defense (Bradley et al., 1992). Callose, a (1,3)- β -glucan, is a major component of cell wall thickening at sites of fungal penetration in plants (Ellinger et al., 2013). The *Arabidopsis* GTPase RabA4c physically interacts with its effector PMR4 to enhance PMR4-dependent callose biosynthesis, which ultimately results in complete penetration resistance to powdery mildew (*Golovinomyces cichoracearum*) (Ellinger et al., 2014). Hydroxyproline-rich glycoproteins (HRGPs) are involved in cell wall strengthening by formation of intra- and intermolecular cross-links (Deepak et al., 2007). Exogenous application of H_2O_2 rescues the callose deposition-deficient phenotype of peroxidase knockdown *Arabidopsis* lines treated with the bacterial flagellin, Flg22 (Daudi et al., 2012). This suggests that cell wall peroxidase-dependent H_2O_2 production is required for PAMP-triggered immune responses, such as callose deposition.

EFFECTOR-MEDIATED ROS BURST THAT INDUCES HR AND CELL DEATH RESPONSES

Reactive oxygen species bursts are monitored during infection by avirulent pathogens (Grant and Loake, 2000). Pathogen avirulence (Avr) effectors interact directly or indirectly with NLR proteins, leading to a strong ROS burst and the HR cell death response, both key components of ETI (Figure 2; Gabriel and Rolfe, 1990; McHale et al., 2006; van der Hoorn and Kamoun, 2008; Spoel and Dong, 2012; Cesari et al., 2014; Han and Hwang, 2017). However, whether NLR-mediated ROS themselves activate HR cell death and immune responses is not fully understood. Rice resistance to incompatible rice blast fungus (*Magnaporthe oryzae*) isolates is suppressed by inhibiting the accumulation of apoplastic ROS, even in the presence of the *R* gene (Singh et al., 2016). The resistant rice cultivar Hwajeonbyeo carrying *Pib* exhibits a compatible response to *M. oryzae* INA168 carrying *AvrPii* and *AvrPib* via the deletion of *OsNADP-ME2*. In the absence of normal production of ROS, cell death and immune responses are severely suppressed, although the Avr effector and cognate NLR proteins are not impaired (Singh et al., 2016). Increased ROS production during infection is essential for NLR-mediated cell death and immunity as well as disease-associated cell death (Greenberg and Yao, 2004; Choi et al., 2013). Perception of plant NLR receptors by specific pathogen effectors triggers a strong ROS burst through activation of RBOHs receiving an adequate supply of NADPH via the activity of NADP-ME (Figure 1; Singh et al., 2016).

A ROS production in plant-pathogen interactions



B ROS signaling in response to pathogens

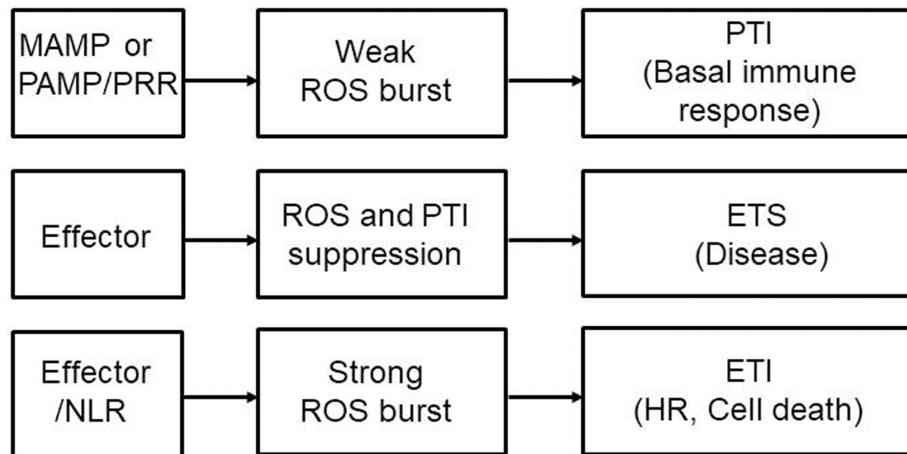


FIGURE 2 | Reactive oxygen species (ROS) production and signaling in plant-pathogen interactions. **(A)** Perception of microbe- or pathogen-associated molecular patterns (MAMPs or PAMPs) or effectors by plants via membrane-bound pattern recognition receptors (PRRs) or intracellular nucleotide-binding leucine-rich repeat (NLR) receptors, respectively, activates NADPH oxidases as well as peroxidases, triggering ROS production and the ROS burst. **(B)** Perception of MAMPs or PAMPs by PRRs induces weak ROS bursts, leading to PAMP-triggered immunity (PTI)-dependent basal defense responses. Adapted pathogens secrete virulent effector proteins into plant cells to suppress the ROS burst and PTI, resulting in effector-triggered susceptibility (ETS) to cause disease in their respective host plants. Pathogen avirulence (Avr) effectors interact directly or indirectly with intracellular NLR proteins, leading to a strong ROS burst and HR cell death response, key components of effector-triggered immunity (ETI). ETI is an accelerated and amplified PTI response, resulting in disease resistance and, usually, a resistance (*R*) gene-mediated HR cell death response at the infection site in plants.

Host cysteine proteases targeted by the *Ustilago maydis* effector Pit2 are likely to be crucial determinants in apoplastic maize immune responses such as the ROS burst (Mueller et al., 2013). The *Cladosporium fulvum* effector Avr2 binds and inhibits Rcr3, a extracellular tomato cysteine protease, which is required for *Cf-2*-dependent disease resistance (Rooney et al., 2005). It has been proposed that the Rcr3-Avr2 complex enables the *Cf-2* protein to activate a HR, including the ROS burst and cell death (Table 1). The defense-related protease Rcr3 may act as a decoy for Avr2 perception in tomato plants carrying the *Cf-2* resistance gene (Shabab et al., 2008).

The *Xanthomonas campestris* pv. *vesicatoria* effector AvrBsT induces a H_2O_2 burst and HR cell death in pepper (*Capsicum annuum*) (Kim et al., 2013). The AvrBsT-triggered HR cell death response is similar to the resistance (*R*) gene-mediated defense response in plants (Eitas and Dangl, 2010; Kim et al., 2010). AvrBsT physically binds to pepper arginine decarboxylase (CaADC1) (Kim et al., 2013), pepper aldehyde dehydrogenase (CaALDH1) (Kim and Hwang, 2015a), pepper heat shock protein 70a (CaHSP70a) (Kim and Hwang, 2015b), and pepper suppressor of the G2 allele of skp1 (CaSGT1) (Kim et al., 2014).

in planta to promote the ROS burst, defense gene expression, cell death, and defense responses (Han and Hwang, 2017). AvrBsT and CaPIK1 directly bind to CaSGT1 in yeast and *in planta*. AvrBsT is subsequently phosphorylated by CaPIK1 and forms the active AvrBsT-SGT1-SGT1-PIK1 complex, which promotes the ROS burst, HR cell death, and defense responses in plants (Kim et al., 2014).

EFFECTOR-MEDIATED ROS SUPPRESSION THAT CAUSES DISEASE ON HOST PLANTS

Adapted microbial pathogens have evolved their effector proteins as virulence factors to suppress the ROS burst and PTI, causing disease on their respective host plants (Figure 2). Secreted effector proteins are delivered into host cells to protect pathogen cell walls against plant-derived hydrolytic enzymes and suppress PAMP-triggered host immunity, leading ultimately to the successful colonization of host plants (de Jonge et al., 2010; Mentlak et al., 2012; Lee et al., 2014; Sánchez-Vallet et al., 2014; Wawra et al., 2016). Various

components of fungal cell walls such as glucans, chitin, and proteins, acting as PAMPs to trigger basal immune responses, are degraded by plant-derived hydrolytic enzymes, such as β -1,3-glucanases, chitinases, and serine and cysteine proteases (Jones and Dangl, 2006; van den Burg et al., 2006).

The *C. fulvum* effector Avr4, a chitin-binding lectin, binds to its own cell walls to protect chitin against hydrolysis by plant chitinases during infection of tomato, suggesting that Avr4 is a virulence factor (Table 1; van den Burg et al., 2006). The LysM domain-containing effector protein Ecp6 of *C. fulvum* mediates virulence through suppression of chitin-triggered immunity in plants (de Jonge et al., 2010). *C. fulvum* Ecp6 (CfEcp6) is secreted at high levels during plant infection and binds chitin, thereby blocking chitin-triggered immunity responses through sequestering chitin fragments and preventing their recognition by plant chitin receptors. By contrast to *C. fulvum* Ecp6, both Mg1LysM and Mg3LysM from the fungus *Mycosphaerella graminicola* protect fungal hyphae against plant-derived hydrolytic enzymes, such as chitinases (Table 1; Marshall et al., 2011). As a virulence determinant in the rice blast fungus *M. oryzae*, the secreted LysM effector (Slp1) binds to chitin inside the fungal cell wall, suppressing chitin-triggered plant immune responses, including the ROS burst and plant defense gene expression (Mentlak et al., 2012). The effector Slp1 inhibits the chitin-induced ROS burst in rice suspension cells. The lectin FGB1 (Fungal Glucan Binding 1), secreted from the root endophyte *Piriformospora indica*, specifically interacts with β -1, 6-linked glucan, altering

cell wall composition and suppressing glucan-triggered ROS production in plants (Wawra et al., 2016). The presence of *P. indica* in the roots of barley inhibits laminarin-induced ROS production. Laminarin-triggered ROS production is also delayed when compared with that observed following chitin elicitation.

Some pathogen effectors specifically bind to plant proteases and may activate downstream signaling components (van der Hoorn and Jones, 2004). Apoplast-localized plant proteases can play an important role in defense responses to microbial pathogens. The *U. maydis* effector Pit2 physically interacts with and inhibits apoplastic maize cysteine proteases, suppressing host immunity (Table 1; Mueller et al., 2013). The secreted effector protein Pit2 is essential for maintenance of biotrophy and induction of tumors in maize. The biotrophic interaction of maize with *U. maydis* depends on inhibition of apoplastic cysteine proteases by the effector Pit2 with a conserved inhibitor domain. The *U. maydis* effector Pep1 (Protein essential during penetration-1) interferes with maize apoplastic peroxidases at the plant-pathogen interface to scavenge ROS (Table 1; Hemetsberger et al., 2012). As an inhibitor of plant peroxidases, Pep1 effectively inhibits the peroxidase-triggered ROS burst, which thereby suppresses the early immune response of maize leading to the establishment of a biotrophic interaction. Pep1 localizes to the plant apoplast, where it accumulates at sites of cell-to-cell passage of biotrophic *U. maydis* hyphae (Doehlemann et al., 2009). The obligate biotrophic fungal pathogen of barley, *Blumeria graminis* f. sp. *hordei*, secretes an extracellular catalase B to

TABLE 1 | Effector targets associated with ROS signaling in plants.

Effector	Pathogen	Host	Target	Reference
Plant-pathogen interface				
Pep1	<i>Ustilago maydis</i>	Maize	Apoplastic peroxidase	Hemetsberger et al., 2012
Catalase B	<i>Blumeria graminis</i>	Barley	Apoplastic H ₂ O ₂	Zhang et al., 2004
Avr4	<i>Cladosporium fulvum</i>	Tomato	Fungal cell wall (Chitin)	van den Burg et al., 2006
Mg3LysM	<i>Mycosphaerella graminicola</i>	Wheat	Fungal cell wall (Chitin)	Marshall et al., 2011
Pit2	<i>U. maydis</i>	Maize	Cysteine protease	Mueller et al., 2013
Avr2	<i>Cladosporium fulvum</i>	Tomato	Cysteine protease Rcr3	Shabab et al., 2008
MAPK signaling pathway				
AvrPtoB	<i>Pseudomonas syringae</i> pv. <i>tomato</i>	Arabidopsis	LysM receptor kinase	Gimenez-Ibanez et al., 2009
HopAO1	<i>P. syringae</i> pv. <i>tomato</i>	Arabidopsis	EF-TU RECEPTOR	Macho et al., 2014
Avr Pto	<i>P. syringae</i> pv. <i>tomato</i>	Arabidopsis	BAK1	Shan et al., 2008
HopF2	<i>P. syringae</i> pv. <i>tomato</i>	Arabidopsis	BAK1, MAPKs	Wu et al., 2011; Zhou et al., 2014
HopF2	<i>P. syringae</i> pv. <i>tomato</i>	Arabidopsis	MKK5	Wang et al., 2010
HopAI1	<i>P. syringae</i> pv. <i>tomato</i>	Arabidopsis	MAPKs	Zhang et al., 2007, 2012
PopP2	<i>Ralstonia solanacearum</i>	Arabidopsis	WRKY TFs	Le Roux et al., 2015
AvrAC	<i>Xanthomonas campestris</i> pv. <i>campestris</i>	Arabidopsis	BIK1	Feng et al., 2012
Vesicle trafficking				
HopM1	<i>P. syringae</i> pv. <i>tomato</i>	Arabidopsis	AtMIN7	Nomura et al., 2006, 2011
Cryptogein	<i>Phytophthora cryptogea</i>	Tobacco	RBOHD	Noirot et al., 2014
Metabolic priming				
AvrPii	<i>Magnaporthe oryzae</i>	Rice	NADP-malic enzyme2	Singh et al., 2016
AvrBsT	<i>X. c.</i> pv. <i>vesicatoria</i>	Pepper	Arginine decarboxylase	Kim et al., 2013
AvrRxo1	<i>X. oryzae</i> pv. <i>oryzicola</i>	Rice	NAD	Shidore et al., 2017

scavenge H_2O_2 at sites of fungal germ tube invasion during infection (Table 1; Zhang et al., 2004). A large number of *catB* transcripts accumulate during the mature primary germ tube and appressorium germ tube stages of fungal development on the susceptible barley plant, suggesting the upregulation of an extracellular catalase gene early during fungal invasion.

EFFECTOR-MEDIATED ROS SUPPRESSION THAT TARGETS MAPK SIGNALING PATHWAYS

During pathogen infection, perception of PRRs by PAMPs leads to PTI by causing a rapid ROS burst via receptor-like cytoplasmic kinases (RLCKs) such as PBL1 and BIK1 (Feng et al., 2012; Shi et al., 2013; Ranf et al., 2014) or through MAPK cascades (Zhou et al., 2014). Pathogens overcome basal

immune responses through inactivation of PAMP-induced signaling pathways that target MAPK cascade components (Pitzschke et al., 2009; Bi and Zhou, 2017). MAPKs are major targets for inactivation by pathogen effector proteins (Figure 3). MAPK cascades are highly conserved modules and are implicated in pathogen signaling during multiple defense responses against pathogen invasion in plants (Yang et al., 2001; Colcombet and Hirt, 2008; Tena et al., 2011; Singh et al., 2012, 2014, 2016). In particular, the MAPK cascade regulates transcriptional reprogramming via the WRKY transcription factor in the early signaling events following PAMP recognition in plants (Adachi et al., 2015). MAPK signaling is involved in the expression of genes required for apoplastic ROS production in plant defense responses. For example, WRKY transcription factors phosphorylated by MAPKs upregulate the RBOH, an NADPH oxidase, inducing pathogen-responsive ROS bursts in *Nicotiana benthamiana* (Adachi et al., 2015). However, because of

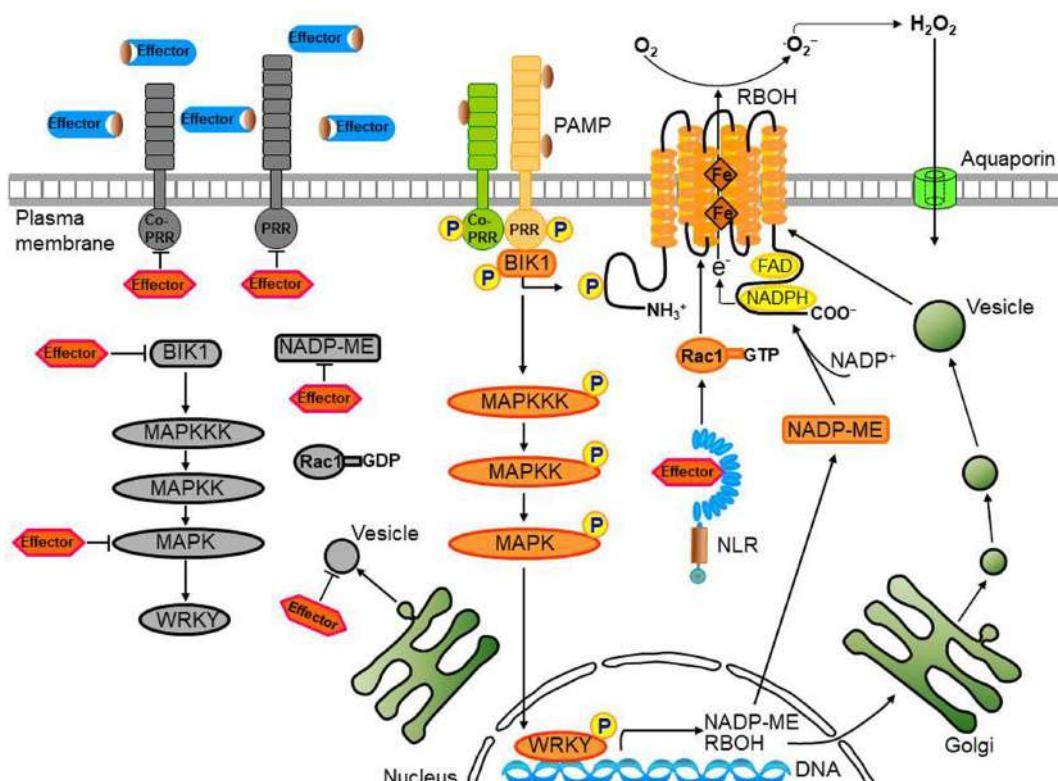


FIGURE 3 | Convergent effector targeting of ROS signaling networks in plant cells. Plant pathogens secrete effectors into the apoplastic area and cytoplasm of plant cells. These apoplastic effectors interfere with PRR-mediated PAMP recognition, ultimately leading to the inactivation of plasma membrane-bound RBOHs. Cytoplasmic effectors target the kinase domains of PRRs, the receptor-like cytoplasmic kinase (RLCK) BIK1, MAPK cascades, and WRKY transcription factors, inhibiting the transcription of *RBOHs* and *NADP-ME*, both essential for robust ROS generation. A bacterial type III effector inhibits BIK1 downstream of the PRR (BAK1), enhancing virulence. Some cytoplasmic effectors target vesicle trafficking, suppressing the transport of ROS-producing RBOH enzymes to the plasma membrane during infection. Secreted pathogen effectors block the metabolic coenzyme NADP-ME, inhibiting the transfer of electrons to the NADPH oxidases (RBOHs) responsible for ROS generation. The biomembrane channels, aquaporins, mediate H_2O_2 transport across biological membranes. However, when specific pathogen effectors recognize intracellular NLR proteins, the effector no longer functions as a virulence factor; here, the effector-NLR complex leads to a strong apoplastic ROS burst and HR cell death response. The PRR-associated kinase BIK1 directly phosphorylates the NADPH oxidase RBOHD to enhance RBOHD-mediated ROS production. Rac1 directly interacts with both NLR and RBOH to activate RBOH. Apoplastic ROS are toxic to pathogens and also activate MAPK cascades and RBOH enzymes in their role as immune signal molecules.

the complicated nature of the downstream MAPK cascade, how MAPK signaling promotes ROS generation is not fully understood.

Pathogen effectors act on plant host target proteins, interfering with PTI-mediated defense signaling cascades, such as ROS and MAPK cascades, ultimately causing disease in host cells (Jones and Dangl, 2006). For example, the *P. syringae* effector HopAO1 targets the *Arabidopsis* receptor kinase EF-TU RECEPTOR (EFR), reducing EFR phosphorylation, thereby preventing subsequent early immune responses, such as the ROS burst and MAPK activation (Table 1; Macho et al., 2014). The *P. syringae* effector HopF2 interacts directly with the plasma membrane-localized receptor-like kinase (RLK) BAK1 and suppresses early signaling events triggered by multiple PAMPs, including BIK1 phosphorylation, MAPK activation, and defense gene expression (Table 1; Wu et al., 2011; Zhou et al., 2014). In fact, BAK1 can directly phosphorylate the plasma membrane-localized RLCK BIK1 (Lu et al., 2010). FLAGELLIN SENSING2 (FLS2)/BAK1-induced MAPK signaling (Zhou et al., 2014) enhances gene expression for plant immune responses via recognition of specific target genes by WRKY transcription factors (Chi et al., 2013). Multiple WRKYS bind to and activate the NADPH oxidase RBOHB promoter, followed by enhanced *RBOHB* expression, which subsequently leads to the RBOHB-dependent ROS burst (Adachi et al., 2015). Type III effectors also target specific host plant proteins, such as major MAPK and WRKY modules (Feng and Zhou, 2012; Le Roux et al., 2015). The *P. syringae* effector HopAI1, for example, inactivates MAPKs by removing the phosphate group from phosphothreonine through a unique phosphothreonine lyase activity (Table 1; Zhang et al., 2007, 2012). Another type III effector, HopF2, interacts with *Arabidopsis* MAP Kinase Kinase 5 (MKK5), and likely other MAPKKs (MKK1, MKK3, MKK4, MKK6, and MKK10), suppressing MAPKs and PAMP-triggered defenses (Table 1; Wang et al., 2010). The *Ralstonia solanacearum* acetyltransferase effector PopP2 localizes to the plant cell nucleus and acetylates lysine residues of WRKY transcription factors, blocking DNA binding (Table 1; Deslandes et al., 2003; Le Roux et al., 2015). The *X. campestris* uridylyl transferase effector AvrAC targets and inhibits BIK1 downstream of BAK1, enhancing virulence (Table 1; Feng et al., 2012). In particular, AvrAC acts upstream of MAPK cascades and ROS production to suppress PTI signaling. The PRR-associated kinase BIK1 directly phosphorylates the NADPH oxidase RBOHD, enhancing RBOHD-mediated ROS production (Figure 3; Kadota et al., 2014; Li et al., 2014). The *Phytophthora infestans* RXLR effector PexRD2 interacts with a specific host MAPKKK, suppressing MAPKKK signaling-dependent cell death (King et al., 2014). In addition, multiple RXLR effectors from *Hyaloperonospora arabidopsis* and *P. infestans* suppress the PAMP-elicited ROS burst (Fabro et al., 2011; Zheng et al., 2014). Functionally redundant effectors from different pathogen species may evolve virulence strategies to target PTI signal transduction processes, such as MAPK cascades, the ROS burst, and defense gene activation (Figure 3; Kvitko et al., 2009; Win et al., 2012; Zheng et al., 2014).

EFFECTOR-MEDIATED ROS SUPPRESSION THAT TARGETS VESICLE TRAFFICKING AND METABOLIC PRIMING

Vesicle trafficking is an important cellular function in plants and is required for the transport of immune receptors and associated proteins, and for the extracellular secretion of immune-related molecules and antimicrobial compounds upon pathogen attack (Inada and Ueda, 2014; Macho and Zipfel, 2015). Visible vesicle-like bodies aggregate directly beneath sites of fungal attack in the barley-*B. graminis* f. sp. *hordei* pathosystem (An et al., 2006) and vesicle incidence is positively associated with levels of resistance to *B. graminis* f. sp. *hordei* penetration (Collins et al., 2003). Vesicles contain phytoalexins, phenolics, or ROS (Hückelhoven et al., 1999; Collins et al., 2003; An et al., 2006). The presence of ROS in the vesicles at penetration sites of *B. graminis* f. sp. *hordei* (Collins et al., 2003) suggests that the endomembrane-associated immune response is associated with ROS signaling. One constituent of the vesicles is H₂O₂, a plant defense compound involved in antimicrobial, cell wall cross-linking, and signaling functions (Lamb and Dixon, 1997; Collins et al., 2003).

Pathogen effectors target vesicle trafficking, suppressing PTI and potentially mediating immune-related ROS signaling (Figure 3; Macho and Zipfel, 2015). The *Phytophthora cryptogea* effector cryptogein induces an increase in abundance of the ROS-producing enzyme RBOHD, as well as ROS production at the plasma membrane of tobacco cells (Table 1; Noiroit et al., 2014). Plant NADPH oxidases (RBOHDs) localize to the plasma membrane and endomembranes, and vesicle trafficking may contribute to the increase in RBOH abundance at the plasma membrane. The *P. syringae* pv. *tomato* virulence effector HopM1 interacts with and degrades an immunity-associated protein AtMIN7 via the host proteasome, suppressing PTI (Table 1; Nomura et al., 2006). AtMIN7 localizes to the *trans*-Golgi network/early endosome of plant cells and mediates immune-associated vesicle trafficking (Nomura et al., 2011). Some type III effector proteins from *X. campestris* pv. *vesicatoria* (Xcv) target plant protein secretion pathways, suppressing PTI (Macho and Zipfel, 2015). The Xcv type III effector XopJ, a member of the YopJ family of SUMO peptidases and acetyltransferases, is attached to the plasma membrane of plant cells through a myristylation motif and interferes with host-cell protein secretion, inhibiting immune-associated callose deposition at the cell wall (Bartetzko et al., 2009; Üstün et al., 2013). In addition, XopB and XopS contribute to Xcv virulence, suppressing PAMP-triggered gene expression (Schulze et al., 2012). XopB localizes to the Golgi vesicles and cytoplasm and interferes with plant cell protein secretion (Schulze et al., 2012). However, specific PTI-related targets of these Xcv effectors are unknown. Overall, better understanding of the role of vesicle trafficking in promoting plant immune responses related to ROS signaling requires further experimental evidence.

Metabolic priming by secreted pathogen effectors has emerged as a common strategy for host manipulation (Figure 3).

U. maydis secretes chorismate mutase (Cmu1) as a virulence factor into plant cells, suppressing plant defense responses associated with pathogen-induced salicylic acid biosynthesis (Djamei et al., 2011). During *U. maydis* infection in maize, Cmu1 is translocated into plant cells, spreads to neighboring cells, and can change the metabolic status of plant cells through metabolic priming. Pathogen effectors may target plant enzymes required for ROS production in the key metabolic pathways. For instance, the chloroplastic enzyme aspartate oxidase, which is involved in NAD metabolism, is required for the NADPH oxidase RBOHD-mediated ROS burst that is triggered by the perception of several unrelated PAMPs (Macho et al., 2012). Notably, inducible NAD overproduction in *Arabidopsis* transcriptionally up-regulates aspartate oxidase during the incompatible infection with *P. syringae* pv. *tomato* (*avrRpm1*) (Pétriacq et al., 2012). Intracellular NAD acts as an integral regulator of multiple defense layers to trigger the production of ROS and defense hormones (Pétriacq et al., 2016). Manipulation of abscisic acid (ABA) content in *Arabidopsis* modulates ROS production via the control of peroxidase activity in response to *Dickeya dadantii* infection (van Gijsegem et al., 2017). Increased ABA contents seem likely to correlate with reduced ROS production and with enhanced disease susceptibility. The *M. oryzae* effector AvrPii specifically interacts with rice NADP-malic enzyme2 (OsNADP-ME2) to suppress the host ROS burst (Table 1; Singh et al., 2016). Indeed, purified AvrPii proteins inhibit *in vitro* NADP-ME activity. NADP-ME, also known as a coenzyme, catalyzes the oxidative decarboxylation of malate and NADP⁺, and provides NADPH as an electron donor for plasma membrane-bound NADPH oxidase (Figure 1), which is essential for the apoplastic ROS burst at the infection site (Doubnerová and Ryšlavá, 2011). NADP-ME is involved in the production of ROS during early plant basal defense against the hemi-biotrophic fungal pathogen *Colletotrichum higginsianum* (Voll et al., 2012). The *X. oryzae* pv. *oryzicola* effector AvRxo1 targets and phosphorylates the central metabolite and redox carrier NAD *in planta*, and this catalytic activity is required for suppression of the ROS burst (Shidore et al., 2017). The *X. campestris* pv. *vesicatoria* (*Xcv*) effector AvrBsT physically interacts with pepper arginine decarboxylase (CaADC1), mediating polyamine metabolism for ROS signaling, cell death, and defense responses in plants (Table 1; Kim et al., 2013). CaADC1 silencing in pepper plants greatly reduces ROS and nitric oxide (NO) bursts, as well as the cell death response during *Xcv* infection, suggesting that arginine decarboxylase is required for polyamine and ROS signaling in the HR cell death response.

CONCLUDING REMARKS

Microbial pathogens have evolved efficient strategies to overcome plant innate immunity for the establishment of compatible plant-pathogen interactions. Despite their evolution over a long period of time, pathogen effectors target a common host protein network, suppressing the common immune system of all plants (Weßling et al., 2014). In particular,

adapted pathogens have developed effective weaponry to compete with the host plants and defeat evolving plant immunity (Jones and Dangl, 2006). In this arms race, the most frequent target of pathogen attack is a powerful plant weapon system that inflicts immense damage to invading pathogens in a short period of time (Doehlemann et al., 2014).

Effector targeting toward ROS signaling networks in plants is proposed in Figure 3. Adapted pathogens have evolved virulence effectors to inhibit both the generation and accumulation of ROS in the apoplastic space, ultimately leading to the inhibition of the intracellular signaling required for the powerful second ROS burst. Plant pathogens deliver virulence effectors into the apoplastic area and cytoplasm of plant cells. Apoplastic effectors interfere with the perception of PRRs by MAMPs or PAMPs, preventing activation of membrane-bound NADPH oxidase RBOHs. Removal of ROS from the apoplastic area reduces direct toxicity to pathogens and blocks plant cell wall reinforcement, which may be beneficial for the successful intracellular invasion and colonization of microbial pathogens. Cytoplasmic effectors target plant PRRs, the RLCK BIK1, MAPK cascades, and WRKY transcription factors, suppressing the expression of RBOHs and NADP-MEs that are essential for robust ROS generation (Figure 3; Singh et al., 2016). During pathogen infection, certain cytoplasmic effectors interfere with vesicle trafficking to suppress the transport of ROS-producing RBOH enzymes to the plasma membrane (Macho and Zipfel, 2015). Secreted pathogen effectors block the metabolic coenzyme NADP-ME, inhibiting the transfer of electrons to NADPH oxidases (RBOHs). However, when specific pathogen effectors interact directly or indirectly with intracellular NLR proteins, these effectors function as avirulence factors to trigger resistance (R) gene-mediated immunity, the so-called ETI (Dangl et al., 2013). The effector-NLR complex leads to a strong apoplastic ROS burst and HR cell death responses. Apoplastic ROS are toxic to pathogens and activate MAPK cascades and NADPH oxidase (RBOH) enzymes. The first phase of ROS production by PTI following PAMP recognition may eventually have a feedback effect on the second phase of ROS production, resulting in a more powerful ETI (Figure 2; Yoshioka et al., 2003; Ishihama et al., 2011; Adachi et al., 2015).

Pathogens have developed virulence effectors to circumvent the newly changed plant immune system during the coevolution of both pathogens and host plants. In host plants, effector-targeted host proteins inside the plant cell can be modified to restrict pathogen invasion. Virulence effectors interact with different target host proteins in the plant cell, which can ultimately result in the suppression of ROS production in the apoplastic area. Once NADP-ME fails to supply NADPH to the NADPH oxidase (RBOH) as a result of its mutation, apoplastic ROS production is inhibited and pathogens can overcome plant immunity (Singh et al., 2016). These results suggest that ROS contribute directly to plant immunity. PTI and ETI are closely related and a strong ROS burst is essential for ETI. The metabolic cellular processes related to ROS generation are required to sustain PTI and reinforce the immune response through ETI. The key proteins and/or enzymes involved in ROS production

may be supplied by activation of MAPK signaling pathways (Ishihama et al., 2011; Adachi et al., 2015). The biomembrane channels, aquaporins, mediate H₂O₂ transport across biological membranes. However, when effectors interfere with MAPK cascades and WRKY transcription factors (Zhang et al., 2007, 2012; Wang et al., 2010; Feng and Zhou, 2012; Le Roux et al., 2015), the second phase ROS burst and accompanying ETI are severely inhibited (Adachi et al., 2015).

Microbial pathogens have evolved versatile effectors to target the cellular processes associated with plant ROS production. Despite recent advances in our knowledge of pathogen effector targets in ROS signaling networks in plants (**Table 1**), the key host factors directly linking ROS signaling to sites of attempted pathogen invasion are not fully understood. Further elucidation of the molecular and cellular functions of pathogen effectors and host factors underlying the ROS-mediated innate immune system will provide important clues to understand better how versatile effectors have evolved to converge on ROS signaling networks in plants.

REFERENCES

- Adachi, H., Nakano, T., Miyagawa, N., Ishihama, N., Yoshioka, M., Katou, Y., et al. (2015). WRKY transcription factors phosphorylated by MAPK regulate a plant immune NADPH oxidase in *Nicotiana benthamiana*. *Plant Cell* 27, 2645–2663. doi: 10.1105/tpc.15.00213
- Allan, A. C., and Fluhr, R. (1997). Two distinct sources of elicited reactive oxygen species in tobacco epidermal cells. *Plant Cell* 9, 1559–1572. doi: 10.1105/tpc.9.9.1559
- An, Q., Hückelhoven, R., Kogel, K. H., and van Bel, A. J. E. (2006). Multivesicular bodies participate in a cell wall-associated defence response in barley leaves attacked by the pathogenic powdery mildew fungus. *Cell Microbiol.* 8, 1009–1019. doi: 10.1111/j.1462-5822.2006.00683
- Apostol, I., Heinstein, P. F., and Low, P. S. (1989). Rapid stimulation of an oxidative burst during elicitation of cultured plant cells : role in defense and signal transduction. *Plant Physiol.* 90, 109–116.
- Ashtamker, C., Kiss, V., Sagi, M., Davydov, O., and Fluhr, R. (2007). Diverse subcellular locations of cryptogein-induced reactive oxygen species production in tobacco bright yellow-2 cells. *Plant Physiol.* 143, 1817–1826. doi: 10.1104/pp.106.090902
- Bartetzko, V., Sonnewald, S., Vogel, F., Hartner, K., Ruth Stadler, R., Hammes, U. Z., et al. (2009). The *Xanthomonas campestris* pv. *vesicatoria* type III effector protein XopJ inhibits protein secretion: evidence or interference with cell wall-associated defense responses. *Mol. Plant Microbe Interact.* 22, 655–664. doi: 10.1094/MPMI-22-6-0655
- Baxter, A., Mittler, R., and Suzuki, N. (2014). ROS as key players in plant stress signalling. *J. Exp. Bot.* 65, 1229–1240. doi: 10.1093/jxb/ert375
- Bi, G., and Zhou, J. M. (2017). MAP kinase signaling pathways: a hub of plant-microbe interactions. *Cell Host Microbe* 21, 270–273. doi: 10.1016/j.chom.2017.02.004
- Bienert, G. P., and Chaumont, F. (2014). Aquaporin-facilitated transmembrane diffusion of hydrogen peroxide. *Biochim. Biophys. Acta* 1840, 1596–1604. doi: 10.1016/j.bbagen.2013.09.017
- Bindschedler, L. V., Dewdney, J., Blee, K. A., Stone, J. M., Asai, T., Plotnikov, J., et al. (2006). Peroxidase-dependent apoplastic oxidative burst in *Arabidopsis* required for pathogen resistance. *Plant J.* 47, 851–863. doi: 10.1111/j.1365-313X.2006.02837
- Boller, T., and He, S. Y. (2009). Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science* 324, 742–744. doi: 10.1126/science.1171647
- Bradley, D. J., Kjellbom, P., and Lamb, C. J. (1992). Elicitor- and wound-induced oxidative cross-linking of a proline-rich plant cell wall protein: a novel, rapid defense response. *Cell* 70, 21–30.
- Cesari, S., Bernoux, M., Moncquet, P., Kroj, T., and Dodds, P. N. (2014). A novel conserved mechanism for plant NLR protein pairs: the “integrated decoy” hypothesis. *Front. Plant Sci.* 5:606. doi: 10.3389/fpls.2014.00606
- Chi, M. H., Park, S. Y., Kim, S., and Lee, Y. H. (2009). A novel pathogenicity gene is required in the rice blast fungus to suppress the basal defenses of the host. *PLOS Pathog.* 5:e1000401. doi: 10.1371/journal.ppat.1000401
- Chi, Y., Yang, Y., Zhou, Y., Zhou, J., Fan, B., Yu, J. Q., et al. (2013). Protein-protein interactions in the regulation of WRKY transcription factors. *Mol. Plant* 6, 287–300. doi: 10.1093/mp/sst026
- Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J. D., et al. (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* 448, 497–500. doi: 10.1038/nature05999
- Choi, H. W., Kim, D. S., Kim, N. H., Jung, H. W., Ham, J. H., and Hwang, B. K. (2013). *Xanthomonas* filamentous hemagglutinin-like protein Fhal interacts with pepper hypersensitive-induced reaction protein CaHIR1 and functions as a virulence factor in host plants. *Mol. Plant Microbe Interact.* 26, 1441–1454. doi: 10.1094/MPMI-07-13-0204-R
- Choi, H. W., Kim, Y. J., Lee, S. C., Hong, J. K., and Hwang, B. K. (2007). Hydrogen peroxide generation by the pepper extracellular peroxidase CaPO2 activates local and systemic cell death and defense response to bacterial pathogens. *Plant Physiol.* 145, 890–904. doi: 10.1104/pp.107.103325
- Colombet, J., and Hirt, H. (2008). *Arabidopsis* MAPKs: a complex signalling network involved in multiple biological processes. *Biochem. J.* 413, 217–226. doi: 10.1042/BJ20080625
- Collins, N. C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J. L., et al. (2003). SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* 425, 973–977. doi: 10.1038/nature02076
- Dangl, J. L., Horvath, D. M., and Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science* 341, 746–751. doi: 10.1126/science.1236011
- Daudi, A., Cheng, Z., O'Brien, J. A., Mammarella, N., Khan, S., Ausubel, F. M., et al. (2012). The apoplastic oxidative burst peroxidase in *Arabidopsis* is a major component of pattern-triggered immunity. *Plant Cell* 24, 275–287. doi: 10.1105/tpc.111.093039
- de Jonge, R., van Esse, H. P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., et al. (2010). Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science* 329, 953–955. doi: 10.1126/science.1190859
- Deepak, S., Shailasree, S., Kini, R. K., Hause, B., Shetty, S. H., and Mithöfer, A. (2007). Role of hydroxyproline-rich glycoproteins in resistance of pearl millet against downy mildew pathogen *Sclerospora graminicola*. *Planta* 226, 323–333. doi: 10.1007/s00425-007-0484-4
- Deslandes, L., Olivier, J., Peeters, N., Feng, D. X., Khounlotham, M., Boucher, C., et al. (2003). Physical interaction between RRS1-R, a protein conferring

AUTHOR CONTRIBUTIONS

NSJ designed the outline of the manuscript and wrote major part of the story and BKH added his previous bacterial data and expertise to upgrade the whole text and Figures.

FUNDING

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1D1A1A09918756).

ACKNOWLEDGMENT

We thank Y. Kadota for valuable comments on the manuscript.

- resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8024–8029. doi: 10.1073/pnas.1230660100
- Djamei, A., Schipper, K., Rabe, F., Ghosh, A., Vincon, V., Kahnt, J., et al. (2011). Metabolic priming by a secreted fungal effector. *Nature* 478, 395–398. doi: 10.1038/nature10454
- Doehlemann, G., Requena, N., Schaefer, P., Brunner, F., O'Connell, R., and Parker, J. E. (2014). Reprogramming of plant cells by filamentous plant-colonizing microbes. *New Phytol.* 204, 803–814. doi: 10.1111/nph.12938
- Doehlemann, G., van der Linde, K., Assmann, D., Schwammbach, D., Hof, A., Mohanty, A., et al. (2009). Pep1, a secreted effector protein of *Ustilago maydis*, is required for successful invasion of plant cells. *PLOS Pathog.* 5:e1000290. doi: 10.1371/journal.ppat.1000290
- Doke, N. (1983). Involvement of superoxide anion generation in the hypersensitive response of potato-tuber tissues to infection with an incompatible race of Phytophthora-infestans and to the hyphal wall components. *Physiol. Plant Pathol.* 23, 345–357. doi: 10.1016/0048-4059(83)90019-X
- Doubnerová, V., and Ryšlavá, H. (2011). What can enzymes of C₄ photosynthesis do for C₃ plants under stress? *Plant Sci.* 180, 575–583. doi: 10.1016/j.plantsci.2010.12.005
- Eitas, T. K., and Dangl, J. L. (2010). NB-LRR proteins: pairs, pieces, perception, partners, and pathways. *Curr. Opin. Plant Biol.* 13, 472–477. doi: 10.1016/j.pbi.2010.04.007
- Ellinger, D., Glöckner, A., Koch, J., Naumann, M., Stürtz, V., Schütt, K., et al. (2014). Interaction of the Arabidopsis GTPase RabA4c with its effector PMR4 results in complete penetration resistance to powdery mildew. *Plant Cell* 26, 3185–3200. doi: 10.1105/tpc.114.127779
- Ellinger, D., Naumann, M., Falter, C., Zwirkowics, C., Jamrow, T., Manisseri, C., et al. (2013). Elevated early callose deposition results in complete penetration resistance to powdery mildew in Arabidopsis. *Plant Physiol.* 161, 1433–1444. doi: 10.1104/pp.112.211011
- Fabro, G., Steinbrenner, J., Coates, M., Ishaque, N., Baxter, L., Studholme, D. J., et al. (2011). Multiple candidate effectors from the oomycete pathogen *Hyaloperonospora arabidopsis* suppress host plant immunity. *PLOS Pathog.* 7:e1002348. doi: 10.1371/journal.ppat.1002348
- Feng, F., Yang, F., Rong, W., Wu, X., Zhang, J., Chen, S., et al. (2012). A *Xanthomonas* uridine 5'-monophosphate transferase inhibits plant immune kinases. *Nature* 485, 114–118. doi: 10.1038/nature10962
- Feng, F., and Zhou, J. M. (2012). Plant-bacterial pathogen interactions mediated by type III effectors. *Curr. Opin. Plant Biol.* 15, 469–476. doi: 10.1016/j.pbi.2012.03.004
- Frederickson Matika, D. E., and Loake, G. J. (2014). Redox regulation in plant immune function. *Antioxid. Redox Signal.* 21, 1373–1388. doi: 10.1089/ars.2013.5679
- Gabriel, D. W., and Rolfe, B. G. (1990). Working models of specific recognition in plant-microbe interactions. *Annu. Rev. Phytopathol.* 28, 365–391.
- Gimenez-Ibanez, S., Hann, D. R., Ntoukakis, V., Petutschnig, E., Lipka, V., and Rathjen, J. P. (2009). AvrPtoB targets the LysM receptor kinase CERK1 to promote bacterial virulence on plants. *Curr. Biol.* 19, 423–429. doi: 10.1016/j.cub.2009.01.054
- Gómez-Gómez, L., and Boller, T. (2000). FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Mol. Cell* 5, 1003–1011.
- Grant, J. J., and Loake, G. J. (2000). Role of reactive oxygen intermediates and cognate redox signaling in disease resistance. *Plant Physiol.* 124, 21–29.
- Greenberg, J. T., and Yao, N. (2004). The role and regulation of programmed cell death in plant-pathogen interactions. *Cell Microbiol.* 6, 201–211. doi: 10.1111/j.1462-5822.2004.00361
- Han, S. W., and Hwang, B. K. (2017). Molecular functions of *Xanthomonas* type III effector AvrBsT and its plant interactors in cell death and defense signaling. *Planta* 245, 237–253. doi: 10.1007/s00425-016-2628-x
- He, P., Shan, L., Lin, N. C., Martin, G. B., Kemmerling, B., Nürnberger, T., et al. (2006). Specific bacterial suppressors of MAMP signaling upstream of MAPKKK in Arabidopsis innate immunity. *Cell* 125, 563–575. doi: 10.1016/j.cell.2006.02.047
- Hedrich, R. (2012). Ion channels in plants. *Physiol. Rev.* 92, 1777–1811. doi: 10.4161/psb.23009
- Hemetsberger, C., Herrberger, C., Zechmann, B., Hillmer, M., and Doehlemann, G. (2012). The *Ustilago maydis* effector Pep1 suppresses plant immunity by inhibition of host peroxidase activity. *PLOS Pathog.* 8:e1002684. doi: 10.1371/journal.ppat.1002684
- Huang, K., Czymmek, K. J., Caplan, J. L., Sweigard, J. A., and Donofrio, N. M. (2011). HYR1-mediated detoxification of reactive oxygen species is required for full virulence in the rice blast fungus. *PLOS Pathog.* 7:e1001335. doi: 10.1371/journal.ppat.1001335
- Hückelhoven, R., Fodor, J., Preis, C., and Kogel, K. H. (1999). Hypersensitive cell death and papilla formation in barley attacked by the powdery mildew fungus are associated with hydrogen peroxide but not with salicylic acid accumulation. *Plant Physiol.* 119, 1251–1260.
- Inada, N., and Ueda, T. (2014). Membrane trafficking pathways and their roles in plant-microbe interactions. *Plant Cell Physiol.* 55, 672–686. doi: 10.1093/pcp/pcu046
- Ishihama, N., Yamada, R., Yoshioka, M., Katou, S., and Yoshioka, H. (2011). Phosphorylation of the *Nicotiana benthamiana* WRKY8 transcription factor by MAPK functions in the defense response. *Plant Cell* 23, 1153–1170. doi: 10.1105/tpc.110.081794
- Jabs, T. (1999). Reactive oxygen intermediates as mediators of programmed cell death in plants and animals. *Biochem. Pharmacol.* 57, 231–245.
- Jones, J. D., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Kadota, Y., Shirasu, K., and Zipfel, C. (2015). Regulation of the NADPH oxidase RBOHD during plant immunity. *Plant Cell Physiol.* 56, 1472–1480. doi: 10.1093/pcp/pcv063
- Kadota, Y., Sklenar, J., Derbyshire, P., Stransfeld, L., Asai, S., Ntoukakis, V., et al. (2014). Direct regulation of the NADPH oxidase RBOHD by the PRR-associated kinase BIK1 during plant immunity. *Mol. Cell* 54, 43–55. doi: 10.1016/j.molcel.2014.02.021
- Kim, N. H., Choi, H. W., and Hwang, B. K. (2010). *Xanthomonas campestris* pv. *vesicatoria* effector AvrBsT induces cell death in pepper, but suppresses defense responses in tomato. *Mol. Plant Microbe Interact.* 23, 1069–1082. doi: 10.1094/MPMI-23-8-1069
- Kim, N. H., and Hwang, B. K. (2015a). Pepper aldehyde dehydrogenase CaALDH1 interacts with *Xanthomonas* effector AvrBsT and promotes effector-triggered cell death and defence responses. *J. Exp. Bot.* 66, 3367–3380. doi: 10.1093/jxb/erv147
- Kim, N. H., and Hwang, B. K. (2015b). Pepper heat shock protein 70a interacts with the type III effector AvrBsT and triggers plant cell death and immunity. *Plant Physiol.* 167, 307–322. doi: 10.1104/pp.114.253898
- Kim, N. H., Kim, B. S., and Hwang, B. K. (2013). Pepper arginine decarboxylase is required for polyamine and upgamma-aminobutyric acid signaling in cell death and defense response. *Plant Physiol.* 162, 2067–2083. doi: 10.1104/pp.113.217372
- Kim, N. H., Kim, D. S., Chung, E. H., and Hwang, B. K. (2014). Pepper suppressor of the G2 allele of skp1 interacts with the receptor-like cytoplasmic kinase1 and type III effector AvrBsT and promotes the hypersensitive cell death response in a phosphorylation-dependent manner. *Plant Physiol.* 165, 76–91. doi: 10.1104/pp.114.238840
- King, S. R., McLellan, H., Boevink, P. C., Armstrong, M. R., Bukharova, T., Sukarta, O., et al. (2014). *Phytophthora infestans* RXLR effector PexRD2 interacts with host MAPKKK ε to suppress plant immune signaling. *Plant Cell* 26, 1345–1359. doi: 10.1105/tpc.113.120055
- Kvitko, B. H., Park, D. H., Velásquez, A. C., Wei, C. F., Russell, A. B., Martin, G. B., et al. (2009). Deletions in the repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion effector genes reveal functional overlap among effectors. *PLOS Pathog.* 5:e1000388. doi: 10.1371/journal.ppat.1000388
- La, D. R., and López-Huertas, E. (2016). ROS generation in peroxisomes and its role in cell signaling. *Plant Cell Physiol.* 57, 1364–1376. doi: 10.1093/pcp/pcw076
- Lamb, C., and Dixon, R. A. (1997). The oxidative burst in plant disease resistance. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48, 251–275. doi: 10.1146/annurev.aplant.48.1.251
- Le Roux, C., Huet, G., Jauneau, A., Camborde, L., Trémousaygue, D., Kraut, A., et al. (2015). A receptor pair with an integrated decoy converts pathogen disabling of transcription factors to immunity. *Cell* 161, 1074–1088. doi: 10.1016/j.cell.2015.04.025

- Lee, W. S., Rudd, J. J., Hammond-Kosack, K. E., and Kanyuka, K. (2014). *Mycosphaerella graminicola* LysM effector-mediated stealth pathogenesis subverts recognition through both CERK1 and CEBiP homologues in wheat. *Mol. Plant Microbe Interact.* 27, 236–243. doi: 10.1094/MPMI-07-13-0201-R
- Legendre, L., Rueter, S., Heinstein, P. F., and Low, P. S. (1993). Characterization of the oligogalacturonide-induced oxidative burst in cultured soybean (*Glycine max*) cells. *Plant Physiol.* 102, 233–240.
- Lehmann, S., Serrano, M., L'Haridon, F., Tjamos, S. E., and Metraux, J. P. (2015). Reactive oxygen species and plant resistance to fungal pathogens. *Phytochemistry* 112, 54–62. doi: 10.1016/j.phytochem.2014.08.027
- Levine, A., Tenhaken, R., Dixon, R., and Lamb, C. (1994). H₂O₂ from the oxidative burst orchestrates the plant hypersensitive disease resistance response. *Cell* 79, 583–593.
- Li, L., Li, M., Yu, L., Zhou, Z., Liang, X., Liu, Z., et al. (2014). The FLS2-associated kinase BIK1 directly phosphorylates the NADPH oxidase RbohD to control plant immunity. *Cell Host Microbe* 15, 329–338. doi: 10.1016/j.chom.2014.02.009
- Liu, Y., Ren, D., Pike, D., Pallardy, S., Gassmann, W., and Zhang, S. (2007). Chloroplast-generated reactive oxygen species are involved in hypersensitive response-like cell death mediated by a mitogen-activated protein kinase cascade. *Plant J.* 51, 941–954. doi: 10.1111/j.1365-313X.2007.03191
- Lu, D., Wu, S., Gao, X., Zhang, Y., Shan, L., and He, P. (2010). A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 496–501. doi: 10.1073/pnas.0909705107
- Luna, E., Pastor, V., Robert, J., Flors, V., Mauch-Mani, B., and Ton, J. (2011). Callose deposition: a multifaceted plant defense response. *Mol. Plant Microbe Interact.* 24, 183–193. doi: 10.1094/MPMI-07-10-0149
- Macho, A. P., Boutrot, F., Rathjen, J. P., and Zipfel, C. (2012). Aspartate oxidase plays an important role in *Arabidopsis* stomatal immunity. *Plant Physiol.* 159, 1845–1856. doi: 10.1104/pp.112.199810
- Macho, A. P., Schwessinger, B., Ntoukakis, V., Brutus, A., Segonzac, C., Roy, S., et al. (2014). A bacterial tyrosine phosphatase inhibits plant pattern recognition receptor activation. *Science* 343, 1509–1512. doi: 10.1126/science.1248849
- Macho, A. P., and Zipfel, C. (2015). Targeting of plant pattern recognition receptor-triggered immunity by bacterial type-III secretion system effectors. *Curr. Opin. Microbiol.* 23, 14–22. doi: 10.1016/j.mib.2014.10.009
- Marshall, R., Kombrink, A., Motteram, J., Loza-Reyes, E., Lucas, J., Hammond-Kosack, K. E., et al. (2011). Analysis of two *in planta* expressed LysM effector homologs from the fungus *Mycosphaerella graminicola* reveals novel functional properties and varying contributions to virulence on wheat. *Plant Physiol.* 156, 756–769. doi: 10.1104/pp.111.176347
- McHale, L., Tan, X., Koehl, P., and Michelmore, R. W. (2006). Plant NBS-LRR proteins: adaptable guards. *Genome Biol.* 7:212. doi: 10.1186/gb-2006-7-4-212
- Mentlak, T. A., Kombrink, A., Shinya, T., Ryder, L. S., Otomo, I., Saitoh, H., et al. (2012). Effector-mediated suppression of chitin-triggered immunity by *Magnaporthe oryzae* is necessary for rice blast disease. *Plant Cell* 24, 322–335. doi: 10.1105/tpc.111.092957
- Mittler, R., Vanderauwera, S., Suzuki, N., Miller, G., Tognetti, V. B., Vandepoele, K., et al. (2011). ROS signaling: the new wave? *Trends Plant Sci.* 16, 300–309. doi: 10.1016/j.tplants.2011.03.007
- Mori, I. C., and Schroeder, J. I. (2004). Reactive oxygen species activation of plant Ca²⁺ channels. A signaling mechanism in polar growth, hormone transduction, stress signaling, and hypothetically mechanotransduction. *Plant Physiol.* 135, 702–708. doi: 10.1104/pp.104.042069
- Mueller, A. N., Ziemann, S., Treitschke, S., Aßmann, D., and Doeblemann, G. (2013). Compatibility in the *Ustilago maydis*-maize interaction requires inhibition of host cysteine proteases by the fungal effector Pit2. *PLOS Pathog.* 9:e1003177. doi: 10.1371/journal.ppat.1003177
- Noiroit, E., Der, C., Lherminier, J., Robert, F., Moricova, P., Kiêu, K., et al. (2014). Dynamic changes in the subcellular distribution of the tobacco ROS-producing enzyme RBOHD in response to the oomycete elicitor cryptogein. *J. Exp. Bot.* 65, 5011–5022. doi: 10.1093/jxb/eru265
- Nomura, K., Debroy, S., Lee, Y. H., Pumplin, N., Jones, J., and He, S. Y. (2006). A bacterial virulence protein suppresses host innate immunity to cause plant disease. *Science* 313, 220–223. doi: 10.1126/science.1129523
- Nomura, K., Mecey, C., Lee, Y. N., Imboden, L. A., Chang, J. H., and He, S. Y. (2011). Effector-triggered immunity blocks pathogen degradation of an immunity-associated vesicle traffic regulator in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10774–10779. doi: 10.1073/pnas.1103338108
- Nühse, T. S., Bottrell, A. R., Jones, A. M., and Peck, S. C. (2007). Quantitative phosphoproteomic analysis of plasma membrane proteins reveals regulatory mechanisms of plant innate immune responses. *Plant J.* 51, 931–940. doi: 10.1111/j.1365-313X.2007.03192
- O'Brien, J. A., Daudi, A., Butt, V. S., and Bolwell, G. P. (2012a). Reactive oxygen species and their role in plant defence and cell wall metabolism. *Planta* 236, 765–779. doi: 10.1007/s00425-012-1696-9
- O'Brien, J. A., Daudi, A., Finch, P., Butt, V. S., Whitelegge, J. P., Souda, P., et al. (2012b). A peroxidase-dependent apoplastic oxidative burst in cultured *Arabidopsis* cells functions in MAMP-elicited defense. *Plant Physiol.* 158, 2013–2027. doi: 10.1104/pp.111.190140
- Park, S. Y., Choi, J., Lim, S. E., Lee, G. W., Park, J., Kim, Y., et al. (2013). Global expression profiling of transcription factor genes provides new insights into pathogenicity and stress responses in the rice blast fungus. *PLOS Pathog.* 9:e1003350. doi: 10.1371/journal.ppat.1003350
- Pétriacq, P., de Bont, L., Hager, J., Didierlaurent, L., Mauve, C., Guérard, F., et al. (2012). Inducible NAD overproduction in *Arabidopsis* alters metabolic pools and gene expression correlated with increased salicylate content and resistance to *Pst*-*AvrRpm1*. *Plant J.* 70, 650–665. doi: 10.1111/j.1365-313X.2012.04920
- Pétriacq, P., Ton, J., Patriti, O., Tcherkez, G., and Gakiere, B. (2016). NAD acts as an integral regulator of multiple defense layers. *Plant Physiol.* 172, 1465–1479.
- Pitzschke, A., Schikora, A., and Hirt, H. (2009). MAPK cascade signalling networks in plant defence. *Curr. Opin. Plant Biol.* 12, 421–426. doi: 10.1016/j.pbi.2009.06.008
- Ranf, S., Eschen-Lippold, L., Fröhlich, K., Westphal, L., Scheel, D., and Lee, J. (2014). Microbe-associated molecular pattern-induced calcium signaling requires the receptor-like cytoplasmic kinases, PBL1 and BIK1. *BMC Plant Biol.* 14:374. doi: 10.1186/s12870-014-0374-4
- Rooney, H. C. E., van't Klooster, J. W., van der Hoorn, R. A. L., Joosten, M. H. A. J., Jones, J. D. G., and de Wit, P. J. G. M. (2005). *Cladosporium Avr2* inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science* 308, 1783–1786. doi: 10.1126/science.1111404
- Rovenich, H., Zuccaro, A., and Thomma, B. P. (2016). Convergent evolution of filamentous microbes towards evasion of glycan-triggered immunity. *New Phytol.* 212, 896–901. doi: 10.1111/nph.14064
- Sagi, M., and Fluhr, R. (2006). Production of reactive oxygen species by plant NADPH oxidases. *Plant Physiol.* 141, 336–340. doi: 10.1104/pp.106.07809
- Sánchez-Vallet, A., Mesters, J. R., and Thomma, B. P. H. J. (2014). The battle for chitin recognition in plant-microbe interactions. *FEMS Microbiol. Rev.* 39, 171–183. doi: 10.1093/femsre/fuu003
- Schulze, S., Kay, S., Büttner, D., Egler, M., Eschen-Lippold, L., Hause, G., et al. (2012). Analysis of new type III effectors from *Xanthomonas* uncovers XopB and XopS as suppressors of plant immunity. *New Phytol.* 195, 894–911. doi: 10.1111/j.1469-8137.2012.04210
- Shabab, M., Shindo, T., Gu, C., Kaschani, F., Pansuriya, T., Chintha, R., et al. (2008). Fungal effector protein AVR2 targets diversifying defense-related cys proteases of tomato. *Plant Cell* 20, 1169–1183. doi: 10.1105/tpc.107.056325
- Shan, L., He, P., Li, J., Heese, A., Peck, S. C., Nürnberger, T., et al. (2008). Bacterial effectors target the common signaling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. *Cell Host Microbe* 4, 17–27. doi: 10.1016/j.chom.2008.05.017
- Shi, H., Shen, Q., Qi, Y., Yan, H., Nie, H., Chen, Y., et al. (2013). BR-SIGNALING KINASE1 physically associates with FLAGELLIN SENSING2 and regulates plant innate immunity in *Arabidopsis*. *Plant Cell* 25, 1143–1157. doi: 10.1105/tpc.112.107904
- Shidore, T., Broeckling, C. D., Kirkwood, J. S., Long, J. J., Miao, J., Zhao, B., et al. (2017). The effector AvrRxo1 phosphorylates NAD in planta. *PLOS Pathog.* 13:e1006442. doi: 10.1371/journal.ppat.1006442
- Singh, R., Dangol, S., Chen, Y., Choi, J., Cho, Y. S., Lee, J. E., et al. (2016). *Magnaporthe oryzae* effector AVR-Pii helps to establish compatibility by inhibition of the rice NADP-malic enzyme resulting in disruption of oxidative burst and host innate immunity. *Mol. Cells* 39, 426–438. doi: 10.14348/molcells.2016.0094
- Singh, R., Lee, J. E., Dangol, S., Choi, J., Yoo, R. H., Moon, J. S., et al. (2014). Protein interactome analysis of 12 mitogen-activated protein kinase kinase

- kinase in rice using a yeast two-hybrid system. *Proteomics* 14, 105–115. doi: 10.1002/pmic.201300125
- Singh, R., Lee, M. O., Lee, J. E., Choi, J., Park, J. H., Kim, E. H., et al. (2012). Rice mitogen-activated protein kinase interactome analysis using the yeast two-hybrid system. *Plant Physiol.* 160, 477–487. doi: 10.1104/pp.112.200071
- Spoel, S. H., and Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *Nat. Rev. Immunol.* 12, 89–100. doi: 10.1038/nri3141
- Tena, G., Boudsocq, M., and Sheen, J. (2011). Protein kinase signaling networks in plant innate immunity. *Curr. Opin. Plant Biol.* 14, 519–529. doi: 10.1016/j.pbi.2011.05.006
- Tenhaken, R., Levine, A., Brisson, L. F., Dixon, R. A., and Lamb, C. (1995). Function of the oxidative burst in hypersensitive disease resistance. *Proc. Natl. Acad. Sci. U.S.A.* 92, 4158–4163. doi: 10.1073/pnas.92.10.4158
- Tian, S., Wang, X., Li, P., Wang, H., Ji, H., Xie, J., et al. (2016). Plant aquaporin AtPIP1;4 links apoplastic H₂O₂ induction to disease immunity pathways. *Plant Physiol.* 171, 1635–1650. doi: 10.1104/pp.15.01237
- Torres, M. A. (2010). ROS in biotic interactions. *Physiol. Plant.* 138, 414–429. doi: 10.1111/j.1399-3054.2009.01326
- Torres, M. A., and Dangl, J. L. (2005). Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. *Curr. Opin. Plant Biol.* 8, 397–403. doi: 10.1016/j.pbi.2005.05.014
- Torres, M. A., Jones, J. D. G., and Dangl, J. L. (2006). Reactive oxygen species signaling in response to pathogens. *Plant Physiol.* 141, 373–378. doi: 10.1104/ pp.106.079467
- Üstün, S., Bartetzko, V., and Börnke, F. (2013). The *Xanthomonas campestris* type III effector XopJ targets the host cell proteasome to suppress salicylic-acid mediated plant defence. *PLOS Pathog.* 9:e1003427. doi: 10.1371/journal.ppat.1003427
- van den Burg, H. A., Harrison, S. J., Joosten, M. H., Vervoort, J., and de Wit, P. J. (2006). *Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Mol. Plant Microbe Interact.* 19, 1420–1430. doi: 10.1094/MPMI-19-1420
- van der Hoorn, R. A., and Kamoun, S. (2008). From guard to decoy: a new model for perception of plant pathogen effectors. *Plant Cell* 20, 2009–2017. doi: 10.1105/tpc.108.060194
- van der Hoorn, R. A. L., and Jones, J. D. G. (2004). The plant proteolytic machinery and its role in defence. *Curr. Opin. Plant Biol.* 7, 400–407. doi: 10.1016/j.pbi.2004.04.003
- van Gijsegem, F., Pétron, J., Patriti, O., Simond-Côte, E., Maia-Grondard, A., Pétriacq, P., et al. (2017). Manipulation of ABA content in *Arabidopsis thaliana* modifies sensitivity and oxidative stress response to *Dickeya dadantii* and influences peroxidase activity. *Front. Plant Sci.* 8:456. doi: 10.3389/fpls.2017.00456
- Vidal, G., Ribas-Carbo, M., Garmier, M., Dubertret, G., Rasmussen, A. G., Mathieu, C., et al. (2007). Lack of respiratory chain complex I impairs alternative oxidase engagement and modulates redox signaling during elicitor-induced cell death in tobacco. *Plant Cell* 19, 640–655.
- Voll, L. M., Zell, M. B., Engelsdorf, T., Saur, A., Gerrard Wheeler, M. C., Drincovich, M. F., et al. (2012). Loss of cytosolic NADP-malic enzyme 2 in *Arabidopsis* is associated with enhanced susceptibility towards *Colletotrichum higginsianum*. *New Phytol.* 195, 189–202. doi: 10.1111/j.1469-8137.2012.04129
- Wang, Y., Li, J., Hou, S., Wang, X., Li, Y., Ren, D., et al. (2010). A *Pseudomonas syringae* ADP-ribosyltransferase inhibits *Arabidopsis* mitogen-activated protein kinase kinases. *Plant Cell* 22, 2033–2044. doi: 10.1105/tpc.110.075697
- Wawra, S., Fesel, P., Widmer, H., Timm, M., Seibel, J., Leson, L., et al. (2016). The fungal-specific β-glucan-binding lectin FGB1 alters cell-wall composition and suppresses glucan-triggered immunity in plants. *Nat. Commun.* 7:13188. doi: 10.1038/ncomms13188
- Weßling, R., Epple, P., Altmann, S., He, Y., Yang, L., Henz, S. R., et al. (2014). Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. *Cell Host Microbe* 16, 364–375. doi: 10.1016/j.chom.2014.08.004
- Win, J., Chaparro-Garcia, A., Belhaj, K., Saunders, D. G., Yoshida, K., Dong, S., et al. (2012). Effector biology of plant-associated organisms: concepts and perspectives. *Cold Spring Harb. Symp. Quant. Biol.* 77, 235–247. doi: 10.1101/sqb.2012.77.015933
- Wu, S., Lu, D., Kabbage, M., Wei, H. L., Swingle, B., Records, A. R., et al. (2011). Bacterial effector HopF2 suppresses *Arabidopsis* innate immunity at the plasma membrane. *Mol. Plant Microbe Interact.* 24, 585–593. doi: 10.1094/MPMI-07-10-0150
- Yang, K. Y., Liu, Y., and Zhang, S. (2001). Activation of a mitogen-activated protein kinase pathway is involved in disease resistance in tobacco. *Proc. Natl. Acad. Sci. U.S.A.* 98, 741–746. doi: 10.1073/pnas.98.2.741
- Yoshioka, H., Numata, N., Nakajima, K., Katou, S., Kawakita, K., Rowland, O., et al. (2003). *Nicotiana benthamiana* gp91phox homologs NbrbohA and NbrbohB participate in H₂O₂ accumulation and resistance to *Phytophthora infestans*. *Plant Cell* 15, 706–718. doi: 10.1105/tpc.008680
- Zhang, J., Shao, F., Li, Y., Cui, H., Chen, L., Li, H., et al. (2007). A *Pseudomonas syringae* effector inactivates MAPKs to suppress PAMP-induced immunity in plants. *Cell Host Microbe* 1, 175–185. doi: 10.1016/j.chom.2007.03.006
- Zhang, Z., Henderson, C., and Gurr, S. J. (2004). *Blumeria graminis* secretes an extracellular catalase during infection of barley: potential role in suppression of host defence. *Mol. Plant Pathol.* 5, 537–547. doi: 10.1111/j.1364-3703.2004.00251
- Zhang, Z., Wu, Y., Gao, M., Zhang, J., Kong, Q., Liu, Y., et al. (2012). Disruption of PAMP-induced MAP kinase cascade by a *Pseudomonas syringae* effector activates plant immunity mediated by the NB-LRR protein SUMM2. *Cell Host Microbe* 11, 253–263. doi: 10.1016/j.chom.2012.01.015
- Zheng, X., McLellan, H., Fraiture, M., Liu, X., Boevink, P. C., Gilroy, E. M., et al. (2014). Functionally redundant RXLR effectors from *Phytophthora infestans* act at different steps to suppress early flg22-triggered immunity. *PLOS Pathog.* 10:e1004057. doi: 10.1371/journal.ppat.1004057
- Zhou, J., Wu, S., Chen, X., Liu, C., Sheen, J., Shan, L., et al. (2014). The *Pseudomonas syringae* effector HopF2 suppresses *Arabidopsis* immunity by targeting BAK1. *Plant J.* 77, 235–245. doi: 10.1111/tpj.12381
- Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E. J., Jones, J. D. G., Felix, G., et al. (2004). Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* 428, 764–767. doi: 10.1038/nature02485

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Jwa and Hwang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Interplant Aboveground Signaling Prompts Upregulation of Auxin Promoter and Malate Transporter as Part of Defensive Response in the Neighboring Plants

Connor Sweeney^{1,2,3}, Venkatachalam Lakshmanan^{1,2†} and Harsh P. Bais^{1,2*}

¹ Delaware Biotechnology Institute, Newark, DE, USA, ² Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA, ³ Wilmington Charter School, Wilmington, DE, USA

OPEN ACCESS

Edited by:

Tatiana Matveeva,
Saint Petersburg State University,
Russia

Reviewed by:

Paul W. Paré,
Texas Tech University, USA
Oswaldo Valdés-López,
National Autonomous University
of Mexico, Mexico

*Correspondence:

Harsh P. Bais
hbais@udel.edu

†Present address:

Venkatachalam Lakshmanan,
Plant Biology Division, The Samuel
Roberts Noble Foundation, Ardmore,
OK, USA

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 24 January 2017

Accepted: 31 March 2017

Published: 19 April 2017

Citation:

Sweeney C, Lakshmanan V and
Bais HP (2017) Interplant
Aboveground Signaling Prompts
Upregulation of Auxin Promoter
and Malate Transporter as Part
of Defensive Response
in the Neighboring Plants.
Front. Plant Sci. 8:595.
doi: 10.3389/fpls.2017.00595

When disrupted by stimuli such as herbivory, pathogenic infection, or mechanical wounding, plants secrete signals such as root exudates and volatile organic compounds (VOCs). The emission of VOCs induces a response in the neighboring plant communities and can improve plant fitness by alerting nearby plants of an impending threat and prompting them to alter their physiology for defensive purposes. In this study, we investigated the role of plant-derived signals, released as a result of mechanical wounding, that may play a role in intraspecific communication between *Arabidopsis thaliana* communities. Plant-derived signals released by the wounded plant resulted in more elaborate root development in the neighboring, unwounded plants. Such plant-derived signals also upregulated the Aluminum-activated malate transporter (*ALMT1*) responsible for the secretion of malic acid (MA) and the DR5 promoter, an auxin responsive promoter concentrated in root apex of the neighboring plants. We speculate that plant-derived signal-induced upregulation of root-specific *ALMT1* in the undamaged neighboring plants sharing the environment with stressed plants may associate more with the benign microbes belowground. We also observed increased association of beneficial bacterium *Bacillus subtilis* UD1022 on roots of the neighboring plants sharing environment with the damaged plants. Wounding-induced plant-derived signals therefore induce defense mechanisms in the undamaged, local plants, eliciting a two-pronged preemptive response of more rapid root growth and up-regulation of *ALMT1*, resulting in increased association with beneficial microbiome.

Keywords: beneficial microbes, *Bacillus subtilis*, malic acid, microbiome, VOCs, wounding

INTRODUCTION

Studies have also shown that aboveground pathogen and herbivore attack shifts microbiome activity at the belowground level (Yang et al., 2011; Song et al., 2016). We have shown previously that plants under attack by pathogenic bacteria induce a shoot-to-root systemic signal, inducing roots to recruit benign, protective microbes (Rudrappa et al., 2008; Lakshmanan et al., 2012). The shoot-to-root systemic signal triggers a malate transporter (*ALMT1*), which has been shown

to also be activated in response to other abiotic responses (Kochian, 1995; Kobayashi et al., 2007). The *ALMT1* transporter prompts the secretion of tricarboxylic acid cycle intermediate L-malic acid (MA) from *Arabidopsis thaliana* roots, which augments recruitment of the beneficial rhizobacterium *Bacillus subtilis* UD1022-a plant–microbial interaction that decreases susceptibility to many foliar pathogens (Rudrappa et al., 2008; Kumar et al., 2012; Lakshmanan et al., 2012, 2014; Lakshmanan and Bais, 2013). Like most Gram-positive bacteria, *B. subtilis* creates an extracellular matrix composed mainly of proteins and exopolysaccharides (Marvasti et al., 2010). It is documented that the ability of *B. subtilis* to colonize plant roots via biofilm formation is an important feature that adds to the plant growth promotion and biocontrol activity (Lakshmanan et al., 2014; Allard-Massicotte et al., 2016). When colonized on plant roots, *B. subtilis* forms a sort of protective armor around its host by secreting antimicrobial compounds, namely the lipopeptide surfactin, that inhibit the growth of fungi, nematodes, and pathogenic bacteria like *Pseudomonas syringae* (Vlamakis et al., 2013). It is also known that both biotic and abiotic stress may modulate the root microbiome (Erlacher et al., 2015; Lakshmanan, 2015). In addition to root-exuded chemicals, plants are known to signal other plants, microbes, nematodes, and insects via emission of non-polar volatile organic compounds (VOCs) (Delory et al., 2016). The root secretions and VOCs serve as a plant’s arsenal of chemical signals that induce change in inter/intraplant interactions (Baldwin et al., 2006; Owen et al., 2007; Delory et al., 2016). It is known that plant-derived chemical compounds impact plants response against microbes and also mediate changes in plant development via upregulation of growth regulator response (Dudareva et al., 2013).

One important plant growth hormone is the indole-3-acetic acid, a natural auxin, which is responsible for plant cell division and elongation and serves as a signaling molecule in the process of organ and root offshoot initiation (Vanneste and Friml, 2009). The role of auxin in mitigating plant stress has also been noted, specifically to inhibit photo-respiratory-dependent cell death in *Arabidopsis thaliana* (Kerchev et al., 2015). Root growth and differentiation is important for plant survival and its adaption to the extreme environment (Villordon et al., 2014). It is known that root branching and architecture is mediated by both biotic and abiotic factors (Villordon et al., 2014; Khan et al., 2016). Endogenous factors such as growth regulators and auxins play a critical role in root branching and differentiation (Malamy, 2005). The phytohormone auxin is considered to be one of the main growth regulator that triggers the lateral root formation (Bainbridge et al., 2008; Nibau et al., 2008). To monitor auxin activity in response to both biotic and abiotic factors, a *DR5* auxin-inducible promoter (Ulmasov et al., 1995; Chen et al., 2013) fused either to a *GUS* or a *GFP* reporter gene is used. It is also shown that microbes both pathogens and benign bacteria modulate root growth and differentiation (López-Bucio et al., 2006; Zolobowska and Van Gijsegem, 2006). Recently, it was shown that few beneficial microbes such as *Pseudomonas* sp. induce root developmental changes via secretion of diffused compounds (Zamioudis et al., 2013). It is argued that root-derived chemicals mediate belowground microbiome, but it is

tempting to speculate that both biotic and abiotic factors may temporally change root-derived chemical synthesis and secretion (Badri and Vivanco, 2009).

Many biotic and abiotic stress regimes cause defensive responses in the affected plants. These responses are categorized based on the directness of their approach to alleviate the stressor. Direct defenses repel and kill enemies through the secretion of toxins, whereas indirect defenses, including the release of plant-derived chemicals, deter enemies by increasing predation pressure on an attacking herbivore (Kessler and Baldwin, 2001; Baldwin et al., 2002). However, most plants only increase production of the chemicals used in these defensive strategies when they are actually under attack. Documented in interspecies and intra-plant (within a single organism) systems, plant-derived chemicals including VOCs change plant transcriptional patterns of defense-related genes and can increase production of growth regulators related to defending against a certain stressor (Heil and Kost, 2006). Previous studies have investigated the complex chemical conduits active in the interconnected role between aboveground and belowground signaling of plants (Bezemer and van Dam, 2005). Belowground organisms can induce aboveground defense responses and vice versa. Exposure to damaging belowground organisms, such as insects, nematodes, root pathogens, and mycorrhizal fungi, impact the aboveground defense responses and induce indirect defenses that attract carnivores or enhance the effectiveness with which those carnivores consume the attacking herbivores. Similarly, above-ground herbivory can influence the concentration of defense-related compounds in belowground root structures (van Dam, 2009). It is clearly shown that plants can sense microbial neighbors and modify the root-derived chemicals (Badri and Vivanco, 2009). It is shown that *Arabidopsis* and *Medicago* each treated with a pathogen (*Pseudomonas syringae* DC3000) and a benign (*Sinorhizobium meliloti*) microbe trigger secretions of different proteins, indicating plants use different chemicals to signal different neighbors (De-la-Peña et al., 2008). It is appropriate to speculate that plants may have similar kind of machinery to sense the neighboring plants. On the similar lines, *Arabidopsis* plant grown in larger monocultures produce more defense metabolites (glucosinolates) compared to smaller monocultures (Wentzell and Kliebenstein, 2008).

In the present study, we speculated that plants sharing the space with a mechanically injured neighbor may show differences in root plasticity. We also questioned how the recipient community perceives damaged-derived chemical signals and evaluated its impact on root growth and root–microbe interactions. The current study relies on measurements of root growth rate and fluorescence assays using the β -glucuronidase (GUS) reporter gene in two transgenic reporter lines of *Arabidopsis thaliana* (*ALMT1::GUS* and *DR5::GUS*). These transgenic reporter lines offer insight on two belowground, induced-defense mechanisms observed in unwounded plants exposed in close proximity to injured neighbors. The triggered defense responses include the upregulation of the *ALMT1* gene and an auxin-responsive *DR5* gene, and accelerated lateral and primary root growth. We report an unusual shoot-to-root

interplant communication leading to altered belowground root responses and benign biotic associations.

MATERIALS AND METHODS

Plant Growth Conditions

Seeds of wild-type *Arabidopsis thaliana* ecotype Columbia (Col-0) were obtained from the *Arabidopsis* Biological Resource Center (ABRC) and surface sterilized using 50% sodium hypochlorite for 1 min and then thrice washed with sterile water. *ALMT1::GUS* and *DR5::GUS* transgenic lines were obtained from Hiroyuki Koyama (Gifu University, Japan) and Wendy Peer (University of Maryland). The seeds were cultured on Murashige and Skoog (MS) (Murashige and Skoog, 1962) solid agar with 3% sucrose in petri dishes and were incubated at $21 \pm 2^\circ\text{C}$ with 12/12 h of light and dark photoperiod and illuminated with cool fluorescent light with an intensity of $120 \mu\text{Em}^{-2}\text{s}^{-1}$. At 8 days post-germination, seedlings were individually transferred to either undivided (on which two seedlings were positioned either 2 or 4 cm apart) or partition petri plates (with one seedling on each side of the partition).

For root colonization, *ALMT1::GUS*, and *DR5::GUS* assays, 12-days-old seedlings were transferred from solid MS media to 6-well culture plates (Fisher Scientific) containing 2.5 mL of 0.5x MS liquid medium with 0.05 mM MES and 3% sucrose. Each 6-well plate contained two seedlings, placed in corner wells opposite and diagonal from one another to maximize distance apart. Plants were grown for 12 days with constant shaking at 90 rpm.

Mechanical Wounding

Sterilized, room-temperature forceps created 4 distinct punctures to the lamina of 2 of the first true leaves on each “donor” *Arabidopsis* plant. In both the Col-0, and *ALMT1::GUS* assay experiments, one seedling in each petri plate was designated the “Donor” community and was mechanically wounded, while its adjacent seedling was left untouched. Mechanical wounding occurred on the same day as seedling transfer. Non-invasive (non-puncturing) contact of the forceps on the seedlings established control trials in which neither seedling was wounded. Primary root growth rate was measured and calculated as $\mu\text{m h}^{-1}$ over 8 days post-wounding.

ALMT1::GUS Assay and Analysis

Partition plates sealed with Parafilm M® film (Bemis) divide agar but allow for shared airspace. The plates used for the *ALMT1::GUS* assay were half-filled with a solid MS agar with 3% sucrose, while the other halves of the plates were filled with MS agar with 10 μM AlCl₃ (Sigma-Aldrich). Two seedlings, 8-days post-germination, were transferred to the plates and allowed to grow per the earlier growth conditions. Mechanical wounding of the randomly selected “donor” seedling (that which “donates” VOCs) occurred on the same day as transfer. Eight days after transfer, the unwounded seedling in each plate was processed per the published description of the β -Glucuronidase Reporter Gene Staining assay and stored at 4°C in a 4% formaldehyde solution until microscopy on an AxioCam color dissecting microscope.

Gene Staining assay (Sigma-Aldrich) and stored at 4°C in a 4% formaldehyde solution until microscopy on an AxioCam color dissecting microscope.

Bacillus subtilis UD1022 Biofilm Formation

Bacillus subtilis UD1022 was streaked from a -80°C glycerol stock onto a plate of low-salt Luria Bertani (LB) medium (10 g L^{-1} Tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl, pH = 7.0) and grown for 24 h at 28°C with shaking at 180 RPM. A subculture was started in 200 mL of LB liquid culture from the previous streak. After shaking for 24 h at 28°C , the subculture was diluted 1:1000 and incubated further at 28°C . When the subculture OD₆₀₀ reached 0.6–1.0, 10 μL of inoculum (OD₆₀₀ = 0.007 of UD1022) were added to the existing 0.5X MS liquid medium in the wells of the “recipient” *A. thaliana* plants.

Microscopy

Adherent UD1022 cells and biofilm on root surface were imaged using laser scanning confocal microscopy. After 24 h shaking at 90 rpm and 6 h stationary at $21 \pm 2^\circ\text{C}$ under the photoperiod described for growth conditions, UD1022-inoculated plants were removed from media and roots were sliced from aboveground plant body. Root samples were then placed in sterile 1-mL tubes (Eppendorf), rinsed once with Phosphate Buffer (2.5 mM), and then suspended in 1-mL of buffer. Histological staining relied on 1:1000 concentration STYO®13 (Invitrogen, Molecular Probes, Eugene, OR, USA) and 1:500 concentration Calcofluor (Sigma-Aldrich), which were left in contact with the roots for 20 ± 3 min and then rinsed once with sterile water. Images were captured with a 25X C-Apochromat objective on a Zeiss LSM 710. Spectral data was collected on the 710 spectral detector. Collected spectral data was used in online fingerprinting and images were post-processed channel unmixed resulting in blue (calcofluor), green (SYTO®13 in UD1022 biofilm) and red (auto-fluorescence) layers. Limited amounts of SYTO®13 are suspected to have penetrated root vascular tissue and cause increased green fluorescence outside of the UD1022 biofilm.

DR5::GUS Assay and Analysis

Two seedlings, 15-days post-germination, were transferred to partition plates with a solid MS agar with 3% sucrose (Sigma-Aldrich) and allowed to grow per the earlier growth conditions. Mechanical wounding of the randomly selected “donor” seedling occurred on the same day as transfer. Five days after transfer, the unwounded seedling in each plate was processed per the published description of the β -Glucuronidase Reporter Gene Staining assay and stored at 4°C in a 4% formaldehyde solution until microscopy on an AxioCam color dissecting microscope.

Statistical Analysis

The data were analyzed by a one-way analysis of variance (ANOVA) using JMP® Pro, Version 11 (SAS Institute, Inc., Cary,

NC, USA 1989–2007). When necessary to compare two means, Student's two-tailed *t*-test were also generated using JMP® Pro, Version 11.

RESULTS

Mechanical wounding of *A. thaliana* plant facilitated the release of airborne VOCs that induced an elaborate series of defense-mechanisms in the neighboring seedlings. The VOCs upregulated the root-specific malate transporter (*ALMT1*) gene, increasing recruitment of a beneficial bacterium, and the *DR5* auxin promoter, which accelerated the root growth.

Airborne Volatile Organic Compounds (VOCs) from Mechanical Wounding Accelerate Root Growth

We designed a prototype to test if wounding neighboring plants changes belowground phenotype in the neighboring communities (Figures 1A,C). The Recipient (the unwounded seedling) and Donor (the seedling that is wounded and releases VOCs) seedlings were positioned either at 4 or 2 cm apart from each other, in an effort to identify any changes in response that may be due to weakened potency of the VOC signal over greater distance between plants. We used partition- and no-partition-prototypes to check if the Donor community releases both VOCs and root exudates to trigger change in phenotype in the Recipient communities (Figure 1B). We speculated that the no other signals besides VOCs could be exchanged in the partition plates, hence VOCs may play a critical role in signaling between injured donors and recipient communities. Mechanical wounding of the Donor seedling was followed by the observation that, at 4 cm apart, the primary lateral roots of Recipients next to wounded Donors grew significantly faster at $224 \pm 63 \mu\text{m h}^{-1}$ than the $137 \pm 29 \mu\text{m h}^{-1}$ of the seedlings next to unwounded Donors. There was, however, no significant difference in the growth rate of primary lateral root ("PR") between the 2 and 4 cm spaced trials (Figure 2). This result necessitates further investigation into the potency of the VOC signal over larger distance. The number of lateral root extensions ("LR") on the Recipient seedlings was also counted, with Recipient communities next to wounded Donors showing an average of 4.25 more lateral roots than Recipients with unwounded Donors (Figure 2).

Partition plates were used to assess whether the noted VOC signal functioned as a diffusion signal and would still affect the Recipient community without sharing the same medium (Figure 1B). Recipient plants next to wounded Donors exhibited accelerated root growth consistent with that observed in the undivided petri plates: the mean primary root growth rate was $220 \pm 9 \mu\text{m h}^{-1}$ compared to the control growth rate of $164 \pm 11 \mu\text{m h}^{-1}$ (Figure 2). A strong causal pattern is established here between exposure to the VOCs elicited by mechanical wounding and acceleration of growth in PR and LR, suggesting a commensalistic relationship in which wounded plants signal potentially vulnerable neighboring

plants of the same species to mitigate damage by increasing biomass.

Auxin Response Upregulated in Presence of VOCs

Both Donor and Recipient groups in the wounded treatment exhibited longer PR and a greater number of LR than the control Donor and Recipient groups, suggesting that the VOCs released by mechanical wounding may upregulate the auxin response that results in increased accumulation of auxin in the apical meristem of primary roots in both the Recipient and Donor plants. To confirm the involvement of auxin interplay in the VOC triggered Recipient communities, we used an auxin reporter *DR5*:GUS line. Recipient communities described in Figure 1 adjacent to the wounded/unwounded Donor communities were replaced by the *DR5*:GUS lines. 24 h post-wounding, *DR5*:GUS expression of the roots of the Recipients next to wounded Donors exhibited deeper blue coloration, suggesting greater auxin accumulation than in the roots of the control (Figure 3A). Blue staining was concentrated in gradients toward root apex of the primary and lateral root extensions. The significant difference in the magnitude of gradient in the GUS staining suggests that a signaling cascade for the *DR5* auxin reporter is triggered by the VOCs documented in this study; the *DR5* upregulation is consistent with the earlier observation that primary root and lateral root extensions were longer and more numerous in our earlier experiments.

VOCs Upregulate Malate Transporter in the Recipient Communities

Previously, we have shown that root specific malate transporters (such as *ALMT1*) play a vital role in shaping root microbiome in plants infected with an aboveground pathogen (Rudrappa et al., 2008). We also show that the wound-induced VOCs change root phenotype in the Recipient communities, which may involve auxin interplay. Here we argued that the wound-induced VOCs may also modulate *ALMT1* expressions in the neighboring communities. To this end, the neighboring plants exposed to wounded and unwounded Donors were replaced by *ALMT1*:GUS expressing reporter lines. *ALMT1*:GUS expressing reporter lines exposed to wounded/unwounded Donors were harvested post 24 h of exposure. *ALMT1*:GUS expressing reporter lines were stained for GUS activity per the published protocol (Sigma-Aldrich, Co.; Rudrappa et al., 2008). Recipients next to unwounded Donors served as controls. *ALMT1*:GUS expression of the roots of the Recipients next to wounded Donors stained deeper blue than the roots of the control, which exhibited no GUS fluorescence (Figure 3B). GUS expression was concentrated in lateral root extensions and the apical meristem of the primary root, suggesting activation of malate transporter by wounded-induced VOCs.

The intensity of the GUS staining in the Recipient plants next to wounded Donors strongly indicates the existence of a signal transduction pathway that upregulates the *ALMT1* in the presence of the VOCs elicited by mechanical wounding.

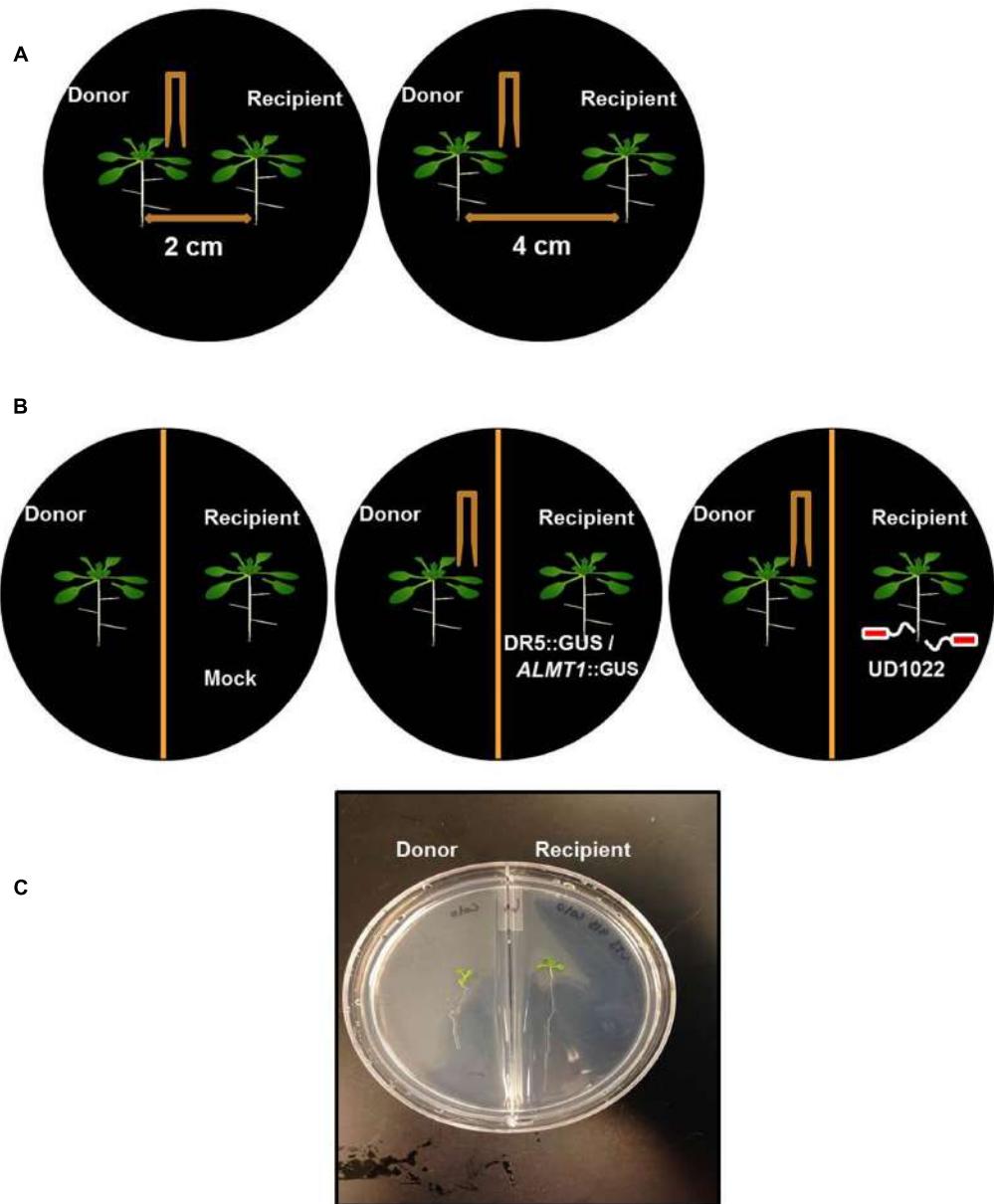


FIGURE 1 | Schematic of *Arabidopsis thaliana* seedlings of donor and recipient communities in different experiments. (A) Twelve day old, two uniform *Arabidopsis* seedlings were transferred to a petri plate with 2 cm and 4 cm apart from each other on petri plates. Donor seedling was wounded to induce VOC release and growth of the seedling was recorded after 8 days of wounding. **(B)** Similarly, *Arabidopsis* seedlings were transferred on the either side of partition petri plates. The donor plants were Col-0 and wounded and recipient's plants were DR5::GUS or ALMT1::GUS or Col-0 rhizo-inoculated with UD1022. **(C)** An image of the recipient and the donor community of *Arabidopsis* in a partition plate.

Recipient Plants Exhibited Increased Association with Benign *Bacillus subtilis* UD1022 in Presence of Wound-induced VOCs

Having shown that wound-induced VOCs trigger malate transporter expression in the Recipient communities, we tested whether the increased malate transporter expression triggered more root colonization by benign bacterium *B. subtilis* UD1022

in the Recipient communities. Recipient communities exposed to wounded/unwounded plants were subjected to UD1022 inoculum ($OD_{600} = 0.007$ of UD1022). Post 24 h of exposure to the wounded/unwounded plants' VOCs, Recipient communities were checked for *B. subtilis* colonization and biofilm formation using confocal microscopy. Confocal imaging of Recipient plant roots exposed to the wounded Donors revealed significantly more UD1022 biofilm development than existent on control roots (exposed to the unwounded Donors) inoculated with identical

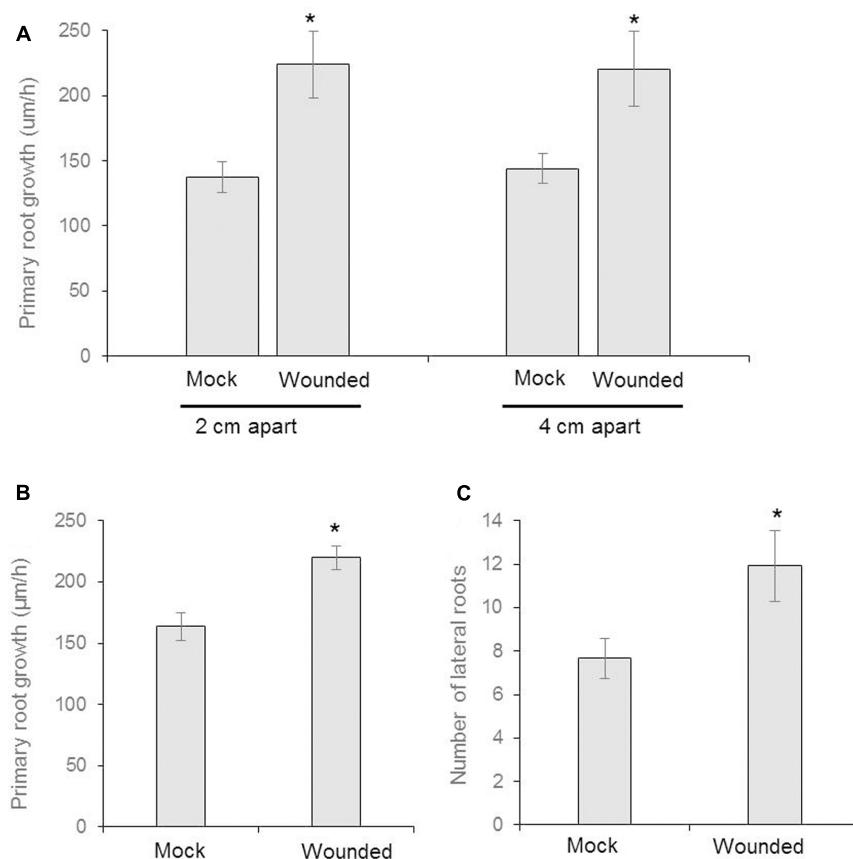


FIGURE 2 | Seeds of *Arabidopsis* (Col-0) were germinated and two uniform seedlings of 12 days old were transferred to the diffused plate or partitioned plate and placed 2 cm or 4 cm apart each other (as shown in schematic Figure 1A). At the same time, the donor plants were wounded mechanically and incubated for 8 days. **(A)** The growth rate of primary root of recipient plant in diffused plate and presented as cm/h ($n = 6$). **(B)** The measurement of growth rate of recipient plant in partitioned petri plate ($n = 12$). **(C)** Number of lateral roots in recipient plants ($n = 12$). Asterisks denote significant differences as analyzed by Student's *t*-test. * $P \leq 0.05$, Student's two-tailed *t*-test. Error bar standard error mean.

levels of UD1022 (Figure 3C). This result suggests that malate upregulation induced by wounded Donors may also associate more with the beneficial bacterium UD1022.

DISCUSSION

The impact of aboveground to belowground signaling in intra-plant communications is a field that has recently gained a lot of attention. Both biotic and abiotic stress trigger mobile signals between the aerial and root parts of a plant (Pangesti et al., 2013). Various studies have shown a two-way communication conduit in plants, wherein plants exposed to aerial pathogens and herbivory alter root phenotype and roots exposed to both biotic and abiotic stress change aboveground physiology in plants (Pangesti et al., 2013). Similarly, few lines of studies have shown that interplant communications changing aboveground physiology in plants exposed to both biotic and abiotic stress agents (Heil and Bueno, 2007). Earlier work explored the concept of “talking trees,” introducing the prey-parasitoid concept triggered by release of VOCs from the stressed plants (De

Moraes et al., 2001). The majority of work related to interplant communication relates to VOC-inducible defense responses in plants (Paré and Tumlinson, 1996). It has been demonstrated that VOCs can attract predatory parasitoids, thus mitigating the threat of attacking herbivores (Thaler, 1999; De Moraes et al., 2001; Kessler and Baldwin, 2001; Karban, 2011). In contrast, it has also been shown that VOCs can help herbivores locate hosts, leading to plant damage (Horiuchi et al., 2003). Most interestingly, VOCs can be used by neighboring, yet-undamaged plants in proximity to damaged plants to adjust their defensive phenotypes (Heil and Bueno, 2007). So far, leaf-derived VOCs in interplant communication have resulted in changes to aboveground physiology. Similarly, involvement of leaf-derived VOCs in intraplant communication has shown microbiome shifts and plant defense response (Song et al., 2016). There exists a gap in our knowledge of how VOCs may modulate interplant communication by changing belowground plasticity and root-microbe interactions. The current study shows that VOCs derived from a damaged plant change root plasticity and root-microbe interaction in the neighboring, yet undamaged plants.

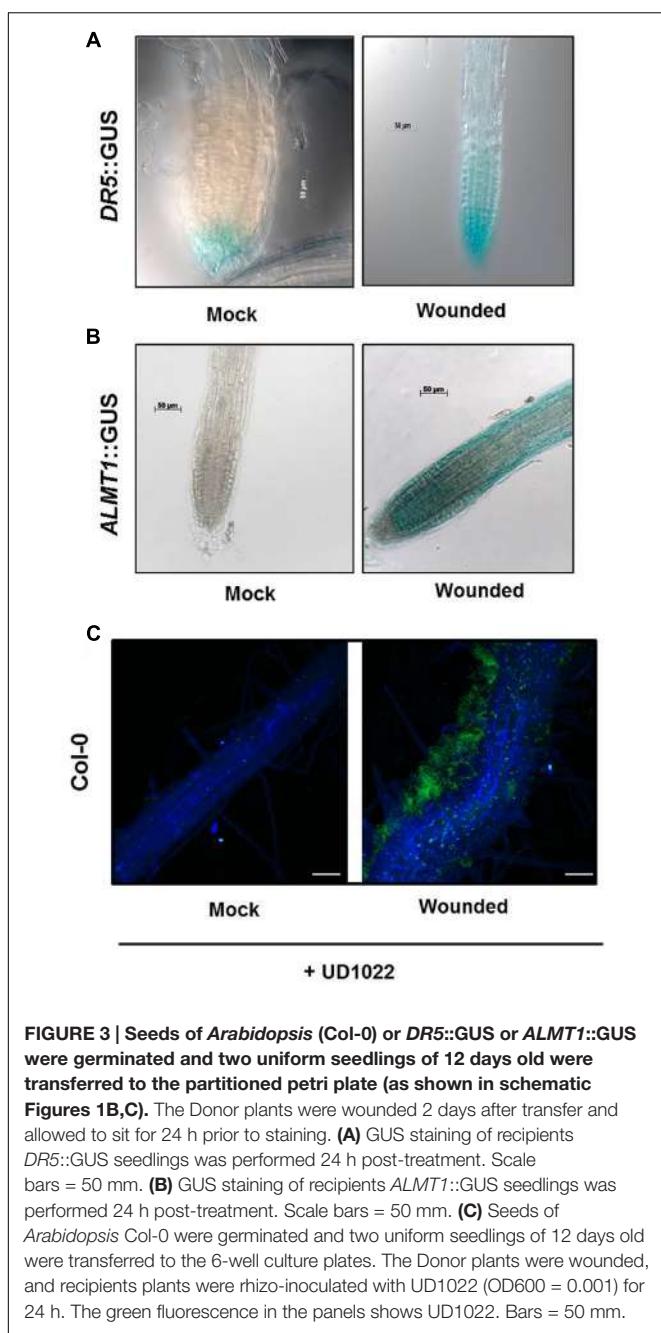


FIGURE 3 | Seeds of *Arabidopsis* (Col-0) or DR5::GUS or ALMT1::GUS were germinated and two uniform seedlings of 12 days old were transferred to the partitioned petri plate (as shown in schematic Figures 1B,C). The Donor plants were wounded 2 days after transfer and allowed to sit for 24 h prior to staining. (A) GUS staining of recipients DR5::GUS seedlings was performed 24 h post-treatment. Scale bars = 50 mm. (B) GUS staining of recipients ALMT1::GUS seedlings was performed 24 h post-treatment. Scale bars = 50 mm. (C) Seeds of *Arabidopsis* Col-0 were germinated and two uniform seedlings of 12 days old were transferred to the 6-well culture plates. The Donor plants were wounded, and recipient plants were rhizo-inoculated with UD1022 (OD600 = 0.001) for 24 h. The green fluorescence in the panels shows UD1022. Bars = 50 mm.

Wounding response at the intraplant level is very well-characterized. It has been shown that wounded plants trigger both local and systemic responses at the intra-plant level (León et al., 2001). It has also been shown that various growth regulators play a part in wound signaling in plants (León et al., 2001). Agents such as oligosaccharides (OGAs), ethylene, jasmonic acid (JA), and abscisic acid (ABA) play a critical role in those local and systemic responses during wound signaling; the signal peptide systemin has also been demonstrated to influence defensive responses in a wounded plant system (León et al., 2001). Systemin is an 18-amino acid peptide generated from a larger protein

precursor called prosystemin (McGurl and Pearce, 1992) and is known to modulate growth regulators in wounded plants. It was reported that wounding in plants causes an increase in ethylene concentration, leading to altered defense responses in plants (Liu et al., 1993; Bouquin et al., 1997). In contrast, wounded tobacco plants show decreased auxin responses (a drop in the endogenous levels of indole acetic acid) (Thornburg and Li, 1991). It has been proposed that the recovery of the initial levels of active auxins serves as a mechanism to limit the duration of the response to wounding (Rojo et al., 1998). Our results showed a contrasting phenotype in the undamaged plants exposed to the wounded neighbors. Plants exposed to the wounded neighbors showed an increase in root growth compared to plants exposed to undamaged neighbors. The shift in plasticity at the interplant level showed that plants respond to aboveground VOCs and alter belowground phenotype. Though auxins are reported to show an inverse relationship with wound response at the intraplant level, our results showed that auxin may play a different role in the interplant signaling response. The DR5 gene marked with the GUS reporter gene in this study is a synthetic auxin-responsive promoter that indicates high auxin accumulation (Chen et al., 2013). In our study, the increase in DR5::GUS staining evident in the root cap and apical meristem of Recipient seedlings next to the wounded Donors indicates that the VOCs initiate the signal transduction responsible for the upregulation of DR5. This histochemical evidence complements the accelerated growth rate (a sign of augmented auxin accumulation) seen in the Recipients in proximity to wounded Donors. The increase in the auxin-responsive promoter DR5 in roots of neighboring plants exposed to the damaged neighbors suggests that auxin activity operates differently under interplant wound signaling compared to intraplant wound signaling.

Our work adds a layer of complexity to the previously documented above- and belowground interactions involving VOCs and microbiome interactions. Recipient plants exposed to damaged neighbors showed increased ALMT1 expression followed by increased colonization by UD1022. The matrix of cells (see green in Figure 3C) surrounding only the Recipient root next to a wounded neighbor illustrates the increased recruitment of *B. subtilis* associated with the upregulation of ALMT1. Previous research has shown that this strain is an effective plant growth-promoting rhizobacteria and reduces foliar entry of deleterious pathogens (Kumar et al., 2012). It is also shown that colonization by beneficial microbes on the root surface increases plant growth promotion and biocontrol activity in plants (Lakshmanan et al., 2014; Allard-Massicotte et al., 2016). The literature suggests that the impact of beneficial microbe-derived volatiles on plants may play a critical role in inducing plant growth promotion and biocontrol activity (Ryu et al., 2003; Rudrappa et al., 2010; Kanchiswamy et al., 2015; Zamioudis et al., 2015). Conversely, there is a gap in our understanding in terms of plant sentinels that may trigger association of benign microbial association with plants (Lakshmanan et al., 2014). We have evidence of how plants manipulate belowground microbiome (Lakshmanan, 2015; Wagner et al., 2016), but we still lack data of plant-derived factors that may modulate the microbiome diversity. There is also evidence which shows association of beneficial microbes in

plants is not a straight-forward process and involves suppression of defense response in plants by benign microbes (Lakshmanan et al., 2012). It would be interesting to see if suppression of defense response by benign microbes also exists in studies involving plant communities. Our work showed that plants may relay a stress-induced sentinel which attracts belowground benign microbes in the neighboring yet undamaged plants. At this juncture we do not fully understand how this association may inflict plant growth promotion phenotype in the undamaged neighboring plants.

In this study, we report that aboveground mechanical wounding elicits substantial belowground changes in plant phenotypic and genotypic characteristics. Previous literature has shown the defense-catalyzing capabilities of VOCs on intraplant and interspecific systems (Holopainen and Gershenson, 2010). Likewise, the effect of mechanical wounding has been noted to induce upregulation of genes in local, undamaged seedlings (Heil and Kost, 2006). The novelty of the current study lies therein the observation, that those priming VOCs reflect altruistic evolutionary developments, as the plant is purposefully designed to warn its neighbors in light of its own damage. The upregulation of malate transporter and auxin responsive genes in neighboring plants exposed to wounded neighbors suggests that *A. thaliana* evolved to anticipate abiotic and biotic stress and survived most when root systems matured at a faster rate, allowing for adequate nutrient and moisture uptake even in potentially contaminated soil.

CONCLUSION

Our hypothesis that the same defense response elicited by VOCs in intraplant and interspecific systems would be induced between neighboring, but anatomically separate plants was correct. We identified the benefits of the *ALMT1* malate transporter in increasing biofilm development and the *DR5* auxin reporter in accelerating root growth. These findings contribute to a

REFERENCES

- Allard-Massicotte, R., Tessier, L., Lécyer, F., Lakshmanan, V., Lucier, J.-F., Garneau, D., et al. (2016). *Bacillus subtilis* early colonization of *Arabidopsis thaliana* roots involves multiple chemotaxis receptors. *mBio* 7:e01664-16. doi: 10.1128/mBio.01664-16
- Badri, D. V., and Vivanco, J. M. (2009). Regulation and function of root exudates. *Plant Cell Environ.* 32, 666–681. doi: 10.1111/j.1365-3040.2009.01926.x
- Bainbridge, K., Guyomarc'h, S., Bayer, E., Swarup, R., Bennett, M., Mandel, T., et al. (2008). Auxin influx carriers stabilize phyllotactic patterning. *Genes Dev.* 22, 810–823. doi: 10.1101/gad.462608
- Baldwin, I. T., Halitschke, R., Paschold, A., Von Dahl, C. C., and Preston, C. A. (2006). Volatile signaling in plant-plant interactions: “talking trees” in the genomics era. *Science* 311, 812–815. doi: 10.1126/science.1118446
- Baldwin, I. T., Kessler, A., and Halitschke, R. (2002). Volatile signaling in plant-plant-herbivore interactions: what is real? *Curr. Opin. Plant Biol.* 5, 351–354. doi: 10.1016/S1369-5266(02)00263-7
- Bezemer, T. M., and van Dam, N. M. (2005). Linking aboveground and belowground interactions via induced plant defenses. *Trends Ecol. Evol.* 20, 617–624. doi: 10.1016/j.tree.2005.08.006
- Bouquin, T., Lasserre, E., Pradier, J., Pech, J.-C., and Balagué, C. (1997). Wound and ethylene induction of the ACC oxidase melon gene CM-ACO1 occurs via two direct and independent transduction pathways. *Plant Mol. Biol.* 35, 1029–1035. doi: 10.1023/A:1005902226054
- Chen, Y., Yordanov, Y. S., Ma, C., Strauss, S., and Busov, V. B. (2013). DR5 as a reporter system to study auxin response in *Populus*. *Plant Cell Rep.* 32, 453–463. doi: 10.1007/s00299-012-1378-x
- De-la-Peña, C., Lei, Z., Watson, B. S., Sumner, L. W., and Vivanco J. M. (2008). Root-microbe communication through protein secretion. *J. Biol. Chem.* 283, 25247–25255. doi: 10.1074/jbc.M801967200
- De Moraes, C. M., Mescher, M. C., and Tumlinson, J. H. (2001). Caterpillar-induced nocturnal plant volatiles repel conspecific females. *Nature* 410, 577–580. doi: 10.1038/35069058
- Delory, B. M., Delaplace, P., Fauconnier, M. L., and du Jardin, P. (2016). Root-emitted volatile organic compounds: can they mediate belowground plant-plant interactions? *Plant Soil* 402, 1–26. doi: 10.1007/s11104-016-2823-3
- Dudareva, N., Klempien, A., Muhlemann, J. K., and Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytol.* 198, 16–32. doi: 10.1111/nph.12145
- Erlacher, A., Cardinale, M., Grube, M., and Berg, G. (2015). Biotic stress shifted structure and abundance of *Enterobacteriaceae* in the lettuce microbiome. *PLoS ONE* 10:e0118068. doi: 10.1371/journal.pone.0118068

growing body of research on root-microbe interactions and expand the agricultural applications of VOCs as factors for pathogen protection and plant growth promotion. Our study demonstrates that the volatiles released by damaged plants elicit belowground changes related to root plasticity and root-microbe interactions in neighboring plants under controlled conditions. Many questions still remain regarding the capabilities of this specific aboveground-belowground VOC relationship: how concentrated does the VOC signal have to be to effectively upregulate auxin-responsive and malate transporter genes under natural conditions? What is the specific chemical composition of the VOC signal at play in this interplant interaction? What other genotypic transductions do these VOCs cause, aside from the two genes of interest reported in the present study? Do VOCs derived from wounded plants play a role in belowground interspecies signaling? The answers to these inquiries ultimately march closer to a new breed of organic crop primers that secure more robust, disease-free crop yields without relying on unsustainable industrial fertilizers.

AUTHOR CONTRIBUTIONS

CS conducted all the experiments described in the manuscript. VL and CS analyzed the data. CS performed the statistics on the root growth analysis. CS and HB drafted the manuscript. HB conceived the study and CS participated in its design and coordination. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We acknowledge support from the University of Delaware Research Foundation (UDRF) and NSF-EPSCoR. We are also thankful to Debbie Powell and Mike Moore, technical faculty at the Delaware Bioimaging Center, for assistance with microscopy.

- Heil, M., and Bueno, J. C. S. (2007). Herbivore-induced volatiles as rapid signals in systemic plant responses: how to quickly move the information? *Plant Signal. Behav.* 2, 191–193. doi: 10.4161/psb.2.3.4151
- Heil, M., and Kost, C. (2006). Priming of indirect defences. *Ecol. Lett.* 9, 813–817. doi: 10.1111/j.1461-0248.2006.00932.x
- Holopainen, J. K., and Gershenson, J. (2010). Multiple stress factors and the emission of plant VOCs. *Trends Plant Sci.* 15, 176–184. doi: 10.1016/j.tplants.2010.01.006
- Horiuchi, J.-I., Arimura, G.-I., Ozawa, R., Shimoda, T., Takabayashi, J., and Nishioka, T. (2003). A comparison of the responses of *Tetranychus urticae* (Acarı: Tetranychidae) and *Phytoseiulus persimilis* (Acarı: Phytoseiidae) to volatiles emitted from lima bean leaves with different levels of damage made by *T. urticae* or *Spodoptera exigua* (Lepidoptera: Noctuidae). *Appl. Entomol. Zool.* 38, 109–116. doi: 10.1303/aez.2003.109
- Kanchiswamy, C. N., Malnoy, M., and Maffei, M. E. (2015). Bioprospecting bacterial and fungal volatiles for sustainable agriculture. *Trends Plant Sci.* 20, 206–211. doi: 10.1016/j.tplants.2015.01.004
- Karban, R. (2011). The ecology and evolution of induced resistance against herbivores. *Funct. Ecol.* 25, 339–347. doi: 10.1111/j.1365-2435.2010.01789.x
- Kerchev, P., Mühlensbock, P., Denecker, J., Morreel, K., Hoeberichts, F. A., Van Der Kelen, K., et al. (2015). Activation of auxin signalling counteracts photorespiratory H2O2-dependent cell death. *Plant Cell Environ.* 38, 253–265. doi: 10.1111/pce.12250
- Kessler, A., and Baldwin, I. T. (2001). Defensive function of herbivore-induced plant volatile emissions in nature. *Science* 291, 2141–2144. doi: 10.1126/science.291.5511.2141
- Khan, M. A., Gemenet, D. C., and Villordon, A. (2016). Root system architecture and abiotic stress tolerance: current knowledge in root and tuber crops. *Front. Plant Sci.* 7:1584. doi: 10.3389/fpls.2016.01584
- Kobayashi, Y., Hoekenga, O. A., Itoh, H., Nakashima, M., Saito, S., Shaff, J. E., et al. (2007). Characterization of *ALMT1* expression in aluminum-inducible malate release and its role for rhizotoxic stress tolerance in *Arabidopsis*. *Plant Physiol.* 145, 843–852. doi: 10.1104/pp.107.102335
- Kochian, L. V. (1995). Cellular mechanisms of aluminum toxicity and resistance in plants. *Annu. Rev. Plant Biol.* 46, 237–260. doi: 10.1146/annurev.pp.46.060195.001321
- Kumar, A. S., Lakshmanan, V., Caplan, J. L., Powell, D., Czymmek, K. J., Levi, D. F., et al. (2012). Rhizobacteria *Bacillus subtilis* restricts foliar pathogen entry through stomata. *Plant J.* 72, 694–706. doi: 10.1111/j.1365-313X.2012.05116.x
- Lakshmanan, V. (2015). Chapter Three- Root microbiome assemblage is modulated by plant host factors. *Adv. Bot. Res.* 75, 57–79. doi: 10.1016/bs.abr.2015.09.004
- Lakshmanan, V., and Bais, H. P. (2013). Factors other than root secreted malic acid that contributes toward *Bacillus subtilis* FB17 colonization on *Arabidopsis* roots. *Plant Signal. Behav.* 8, 657–668. doi: 10.4161/psb.27277
- Lakshmanan, V., Kitto, S. L., Caplan, J. L., Hsueh, Y.-H., Kearns, D. B., Wu, Y.-S., et al. (2012). Microbe-associated molecular patterns-triggered root responses mediate beneficial rhizobacterial recruitment in *Arabidopsis*. *Plant Physiol.* 160, 1642–1661. doi: 10.1104/pp.112.200386
- Lakshmanan, V., Selvaraj, G., and Bais, H. P. (2014). Functional soil microbiome: belowground solutions to an aboveground problem. *Plant Physiol.* 166, 689–700. doi: 10.1104/pp.114.245811
- León, J., Rojo, E., and Sánchez-Serrano, J. J. (2001). Wound signalling in plants. *J. Exp. Bot.* 52, 1–9. doi: 10.1093/jxb/52.354.1
- Liu, D., Li, N., Dube, S., Kalinski, A., Herman, E., and Mattoo, A. K. (1993). Molecular characterization of a rapidly and transiently wound-induced soybean (*Glycine max* L.) gene encoding 1-amino cyclopropane-1-carboxylate synthase. *Plant Cell Physiol.* 34, 1151–1157.
- López-Bucio, J., Acevedo-Hernández, G., Ramírez-Chávez, E., Molina-Torres, J., and Herrera-Estrella, L. (2006). Novel signals for plant development. *Curr. Opin. Plant Biol.* 9, 523–529. doi: 10.1016/j.pbi.2006.07.002
- Malamy, J. E. (2005). Intrinsic and environmental response pathways that regulate root system architecture. *Plant Cell Environ.* 28, 67–77. doi: 10.1111/j.1365-3040.2005.01306.x
- Marvasi, M., Visscher, P. T., and Martinez, L. C. (2010). Exopolymeric substances (EPS) from *Bacillus subtilis*: polymers and genes encoding their synthesis. *FEMS Microbiol. Lett.* 313, 1–9. doi: 10.1111/j.1574-6968.2010.02085.x
- McGurl, B., and Pearce, G. (1992). Structure, expression, and antisense inhibition of the systemin precursor gene. *Science* 255, 1570–1573. doi: 10.1126/science.1549783
- Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x
- Niba, C., Gibbs, D. J., and Coates, J. C. (2008). Branching out in new directions: the control of root architecture by lateral root formation. *New Phytol.* 179, 595–614. doi: 10.1111/j.1469-8137.2008.02472.x
- Owen, S. M., Clark, S., Pompe, M., and Semple, K. T. (2007). Biogenic volatile organic compounds as potential carbon sources for microbial communities in soil from the rhizosphere of *Populus tremula*. *FEMS Microbiol. Lett.* 268, 34–39. doi: 10.1111/j.1574-6968.2006.00602.x
- Pangesti, N., Pineda, A., Pieterse, C. M., Dicke, M., and Van Loon, J. J. (2013). Two-way plant mediated interactions between root-associated microbes and insects: from ecology to mechanisms. *Front. Plant Sci.* 4:414. doi: 10.3389/fpls.2013.00414
- Paré, P. W., and Tumlinson, J. H. (1996). Plant volatile signals in response to herbivore feeding. *Fla. Entomol.* 79, 93–103. doi: 10.2307/3495807
- Rojo, E., Titarenko, E., Leon, J., Berger, S., Vancanneyt, G., and Sánchez-Serrano, J. J. (1998). Reversible protein phosphorylation regulates jasmonic acid-dependent and-independent wound signal transduction pathways in *Arabidopsis thaliana*. *Plant J.* 13, 153–165. doi: 10.1046/j.1365-313X.1998.0020.x
- Rudrappa, T., Biedrzycki, M. L., Kunjeti, S. G., Donofrio, N. M., Czymmek, K. J., Paré, P. W., et al. (2010). The rhizobacterial elicitor acetoin induces systemic resistance in *Arabidopsis thaliana*. *Commun. Integr. Biol.* 3, 130–138.
- Rudrappa, T., Czymmek, K. J., Paré, P. W., and Bais, H. P. (2008). Root-secreted malic acid recruits beneficial soil bacteria. *Plant Physiol.* 148, 1547–1556. doi: 10.1104/pp.108.127613
- Ryu, C. M., Farag, M. A., Hu, C. H., Reddy, M. S., Wei, H. X., Paré, P. W., et al. (2003). Bacterial volatiles promote growth in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4927–4932. doi: 10.1073/pnas.0730845100
- Song, G. C., Sim, H.-J., Kim, S.-G., and Ryu, C.-M. (2016). Root-mediated signal transmission of systemic acquired resistance against above-ground and below-ground pathogens. *Ann. Bot.* 118, 821–831. doi: 10.1093/aob/mcw152
- Thaler, J. S. (1999). Jasmonate-inducible plant defences cause increased parasitism of herbivores. *Nature* 399, 686–688. doi: 10.1038/21420
- Thornburg, R. W., and Li, X. (1991). Wounding *Nicotiana tabacum* leaves causes a decline in endogenous indole-3-acetic acid. *Plant Physiol.* 96, 802–805. doi: 10.1104/pp.96.3.802
- Ulmasov, T., Liu, Z. B., Hagen, G., and Guilfoyle, T. J. (1995). Composite structure of auxin response elements. *Plant Cell* 7, 1611–1623. doi: 10.1105/tpc.7.10.1611
- van Dam, N. M. (2009). Belowground herbivory and plant defenses. *Annu. Rev. Ecol. Evol. Syst.* 40, 373–391. doi: 10.1146/annurev.ecolsys.110308.120314
- Vanneste, S., and Friml, J. (2009). Auxin: a trigger for change in plant development. *Cell* 136, 1005–1016. doi: 10.1016/j.cell.2009.03.001
- Villordon, A. Q., Ginzberg, I., and Firon, N. (2014). Root architecture and root and tuber crop productivity. *Trends Plant Sci.* 19, 419–425. doi: 10.1016/j.tplants.2014.02.002
- Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., and Kolter, R. (2013). Sticking together: building a biofilm the *Bacillus subtilis* way. *Nat. Rev. Microbiol.* 11, 157–168. doi: 10.1038/nrmicro2960
- Wagner, M. R., Lundberg, D. S., Tijana, G., Tringe, S. G., Dangl, J. L., and Mitchell-Olds, T. (2016). Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat. Commun.* 7, 1–15. doi: 10.1038/ncomms12151
- Wentzell, A. M., and Kliebenstein, D. J. (2008). Genotype, age, tissue, and environment regulate the structural outcome of glucosinolate activation. *Plant Physiol.* 147, 415–428. doi: 10.1104/pp.107.115279
- Yang, J. W., Yi, H. S., Kim, H., Lee, B., Lee, S., Ghim, S. Y., et al. (2011). Whitefly infestation of pepper plants elicits defence responses against bacterial pathogens in leaves and roots and changes the below-ground microflora. *J. Ecol.* 99, 46–56. doi: 10.1111/j.1365-2745.2010.01756.x
- Zamioudis, C., Korteland, J., Van Pelt, J. A., Hamersveld, M., Dombrowski, N., Bai, Y., et al. (2015). Rhizobacterial volatiles and photosynthesis-related

signals coordinate *MYB72* expression in *Arabidopsis* roots during onset of induced systemic resistance and iron-deficiency responses. *Plant J.* 84, 309–322. doi: 10.1111/tpj.12995

Zamioudis, C., Mastranesti, P., Dhonukshe, P., Blilou, I., and Pieterse, C. M. (2013). Unraveling root developmental programs initiated by beneficial *Pseudomonas* spp. bacteria. *Plant Physiol.* 162, 304–318. doi: 10.1104/pp.112.212597

Zlobowska, L., and Van Gijsegem, F. (2006). Induction of lateral root structure formation on petunia roots: a novel effect of GMI1000 *Ralstonia solanacearum* infection impaired in Hrp mutants. *Mol. Plant Microbe Interact.* 19, 597–606. doi: 10.1094/MPMI-19-0597

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Sweeney, Lakshmanan and Bais. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



***Bacillus amyloliquefaciens* Confers Tolerance to Various Abiotic Stresses and Modulates Plant Response to Phytohormones through Osmoprotection and Gene Expression Regulation in Rice**

Shalini Tiwari^{1,2}, Vivek Prasad², Puneet S. Chauhan¹ and Charu Lata^{1*}

¹ Council of Scientific & Industrial Research–National Botanical Research Institute, Lucknow, India, ² Department of Botany, University of Lucknow, Lucknow, India

OPEN ACCESS

Edited by:

Tatiana Matveeva,
Saint Petersburg State University,
Russia

Reviewed by:

Maria Carolina Quecine,
University of São Paulo, Brazil
Klára Kosová,
Crop Research Institute, Czechia

***Correspondence:**

Charu Lata
charulata@nbri.res.in;
charulata14@gmail.com

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 25 May 2017

Accepted: 16 August 2017

Published: 29 August 2017

Citation:

Tiwari S, Prasad V, Chauhan PS and
Lata C (2017) *Bacillus amyloliquefaciens* Confers Tolerance
to Various Abiotic Stresses and Modulates Plant Response
to Phytohormones through Osmoprotection and Gene
Expression Regulation in Rice.
Front. Plant Sci. 8:1510.
doi: 10.3389/fpls.2017.01510

Being sessile in nature, plants have to withstand various adverse environmental stress conditions including both biotic and abiotic stresses. Comparatively, abiotic stresses such as drought, salinity, high temperature, and cold pose major threat to agriculture by negatively impacting plant growth and yield worldwide. Rice is one of the most widely consumed staple cereals across the globe, the production and productivity of which is also severely affected by different abiotic stresses. Therefore, several crop improvement programs are directed toward developing stress tolerant rice cultivars either through marker assisted breeding or transgenic technology. Alternatively, some known rhizospheric competent bacteria are also known to improve plant growth during abiotic stresses. A plant growth promoting rhizobacteria (PGPR), *Bacillus amyloliquefaciens* NBRI-SN13 (SN13) was previously reported by our lab to confer salt stress tolerance to rice seedlings. However, the present study investigates the role of SN13 in ameliorating various abiotic stresses such as salt, drought, desiccation, heat, cold, and freezing on a popular rice cv. Saryu-52 under hydroponic growth conditions. Apart from this, seedlings were also exogenously supplied with abscisic acid (ABA), salicylic acid (SA), jasmonic acid (JA) and ethephon (ET) to study the role of SN13 in phytohormone-induced stress tolerance as well as its role in abiotic and biotic stress cross-talk. All abiotic stresses and phytohormone treatments significantly affected various physiological and biochemical parameters like membrane integrity and osmolyte accumulation. SN13 also positively modulated stress-responsive gene expressions under various abiotic stresses and phytohormone treatments suggesting its multifaceted role in cross-talk among stresses and phytohormones in response to PGPR. To the best of our knowledge, this is the first report on detailed analysis of plant growth promotion and stress alleviation by a PGPR in rice seedlings subjected to various abiotic stresses and phytohormone treatments for 0, 1, 3, 10, and 24 h.

Keywords: abiotic stress, cross-talk, expression, osmolytes, phytohormones, rhizobacteria

INTRODUCTION

Rice (*Oryza sativa* L.) is the second most important staple crop across the globe having high caloric value. Various abiotic stresses such as salt, drought, and extreme temperatures, and biotic stresses are major threats for agricultural production and productivity worldwide. In fact abiotic stresses are reported to adversely affect the yields of staple food crops by 70% (Kaur et al., 2008; Mantri et al., 2012). Crop responses to these environmental stresses are manifested at physiological, biochemical and molecular levels from the early stage of seed germination to maturity and senescence. Expression of several thousand genes is known to be altered during various individual and multiple abiotic stresses (Miller et al., 2010; Dahro et al., 2016). It has now been well documented that the genes expressed under various abiotic stresses not only help in improving cellular tolerance by maintaining osmotic homeostasis but also by regulating stress responsive gene expression (Lata et al., 2015; Tiwari et al., 2017).

Several components of abiotic stress signaling have also been found to be regulated through plant growth regulators (PGRs) or phytohormones (Großkinsky et al., 2016). These are signal molecules which are produced within plants at very low concentrations with the ability to regulate various biological processes both locally and distally. Phytohormones namely auxins (AUX), cytokinins (CK), gibberellins (GA), abscisic acid (ABA), brassinosteroids (BR), salicylic acid (SA), jasmonates (JA) and ethylene play important roles in the regulation of plant developmental processes and signaling networks as they are involved either directly or indirectly in a wide range of biotic and abiotic stress responses and tolerance (Khan et al., 2012; Asgher et al., 2015; Srivastava et al., 2016; Tiwari et al., 2016). Numerous studies have shown the positive effects of phytohormones on the growth and development of a variety of crop plants (Cabello-Conejo et al., 2014). However, several phytohormones are also reported to be involved in pathogen-induced defense pathways and hence, have also been extensively used as biotic stress elicitors (Kim et al., 2004; Zhao et al., 2010; Koramutla et al., 2014). Each of these phytohormones is involved in different biological processes, thus affecting the growth and development of plants in a unique way. The effects of phytohormones on the plant vary with applied concentration, environmental factors, and on the physiological status of the plant at the time of application (Cabello-Conejo et al., 2014). Key plant hormones such as ABA, SA, JA, and ET are well known for their regulatory response under stress. As for example, various ethylene response factors such as *SlERF5* and *AtERF6* were established as master regulators of salt and drought stress tolerance in tomato and *Arabidopsis*, respectively (Pan et al., 2012; Dubois et al., 2013). It has also been anticipated that phytohormone synthesis and signaling play central role in response and adaptation to adverse environmental conditions (Bari and Jones, 2009; Lata et al., 2011b; Nafisi et al., 2015).

Interestingly, various plant growth promoting rhizobacteria (PGPR) are known for production of phytohormones as AUX, CK, GA and inhibit the synthesis of ethylene (Park et al., 2017). PGPR are also reported to increase the plant

growth by reducing susceptibility to various environmental stresses (Nautiyal et al., 2013; Tiwari et al., 2016). The use of these beneficial microorganisms considered as one of the most promising methods for safe crop-management practices. Many soil microorganisms like *Azospirillum*, *Agrobacterium*, *Pseudomonas* and *Bacillus* produce phytohormones and also known to modulate endogenous level of phytohormones in plants thereby modulating the overall plant's hormonal balance and its response to stress (Glick et al., 2007; Kundan et al., 2015). Members of *Bacillus* genus are among the most naturally abundant PGPR in the soil. A *Bacillus amyloliquefaciens* strain (NBRI-SN13, referred to as SN13) has been isolated from alkaline soil of Banthara, Lucknow and its characterization for various plant growth promotional attributes and stress tolerance such as, auxin and ACC deaminase production, solubilisation of tricalcium phosphate and proline accumulation under salt stress were carried out earlier in our laboratory (Nautiyal et al., 2013). This PGPR was also reported as a biocontrol agent for *Rhizoctonia solani* infection in rice (Srivastava et al., 2016). However, to the best of our knowledge the role of this PGPR strain in various abiotic stresses and PGRs simultaneously has not been studied till date. Further, numerous molecular studies on phytohormones and stress related genes in rice have suggested the role of various stress signaling and regulatory pathways to play key roles in the cross-talk between phytohormone and biotic/abiotic stresses for plant protection (Du et al., 2013; Srivastava et al., 2016). Therefore, the aim of this study was to investigate the temporal effects of SN13 inoculation on various biochemical and molecular parameters in rice under different short-term abiotic stresses and phytohormone treatments, and delineating the mechanism of a possible cross-talk among them.

MATERIALS AND METHODS

Plant Material, Inoculation and Stress Treatments

The experiment was conducted in a plant growth chamber at CSIR-NBRI, Lucknow, India with temperature oscillating between $25 \pm 2^\circ\text{C}$ (day) and $20 \pm 2^\circ\text{C}$ (night). A popular rice cultivar Saryu-52 was used for this study. The experiment was designed with two parameters control and 1% SN13 inoculated seedlings. Seeds of rice were surface sterilized with 0.1% HgCl_2 , transplanted and grown in hydroponics for 1 week in Hewitt medium. After 24 h of SN13 inoculation, both inoculated and uninoculated seedlings were subjected to various stresses. For salt and osmotic/drought stresses, 100 mM NaCl and 20% polyethylene glycol (PEG) was supplied, respectively (Srivastava et al., 2012; Nautiyal et al., 2013). For desiccation, seedlings were transferred to sterile Whatman filter paper and for heat, cold and freeze stresses, rice seedlings were transferred to 45°C , $\sim 4^\circ\text{C}$, and $\leq 0^\circ\text{C}$, respectively (Mishra et al., 2015). For various phytohormone treatments, both inoculated and uninoculated seedlings 100 μM each of ABA, SA, JA, and ethephon was supplied to the Hewitt medium as described elsewhere (Van Bockhaven et al., 2015). After respective treatments the root and shoot samples were harvested at 1, 3, 10, and 24 h

for further studies. Unstressed seedlings (both inoculated and uninoculated) were maintained as control. All biochemical studies were performed with whole seedlings on the day of harvesting. Samples for qRT-PCR analyses were snap frozen in liquid nitrogen and stored at -80°C until further use. All experimental data are means of at least four independent biological replicates, while three biological as well three technical replicates were used for qRT-PCR analyses.

Proline

Proline content was analyzed using ethanolic extract prepared by homogenizing ~ 100 mg fresh tissue in 1 ml of 70% ethanol (Carillo and Gibbon, 2011; Tiwari et al., 2016). The reaction mixture constituted 1% w/v ninhydrin in 60% v/v acetic acid and 20% v/v ethanol, mixed with ethanolic extract in the ratio of 2:1. The 100 μl reaction mixture was then incubated in a water bath at 95°C for 20 min, cooled to room temperature, and absorbance was recorded at 520 nm in a microplate reader (Spectrum max plus; Molecular devices, Sunnyvale, CA, United States).

Total Soluble Sugar

Total soluble sugar (TSS) content of rice seedlings was determined according to DuBois et al. (1956) with some modifications. About 100 mg of sample was homogenized in 3 ml of 80% methanol and was incubated at 70°C for 30 min. After incubation, equal volume (500 μl) of extract and 5% phenol each was mixed with 1.5 ml of 95% H_2SO_4 and further incubated in dark for 15–20 min. Absorbance was then measured in spectrophotometer (Spectrum max plus; Molecular devices, Sunnyvale, CA, United States) at 490 nm wavelength.

Lipid Peroxidation

The level of lipid peroxidation (LP) in control and treated tissues was determined by measuring malondialdehyde (MDA) content via 2-thiobarbituric acid (TBA) reaction using modified protocol described by Heath and Packer (1968). About 100 mg of leaf tissues were homogenized in 500 μl of 0.1% (w/v) TCA and centrifuged for 10 min at 13,000 g at 4°C . Further 500 μl of supernatant was then mixed with 1.5 ml 0.5% TBA and incubated in water bath at 95°C for 25 min. Mixture was then incubated on ice for 5 min for termination of reaction. Absorbance of mixture was measured at 532 and 600 nm in a microplate reader (Spectrum max plus; Molecular devices, Sunnyvale, CA, United States).

Quantitative Real Time (qRT) PCR Analysis of Stress-Responsive Genes from Rice

Total RNA was isolated from 1-week-old rice seedlings subjected to different durations of various abiotic stresses and phytohormones treatments with or without SN13-inoculation, using Tri-Reagent (Sigma, United States). DNase treatment was done using TURBO DNase (Ambion, United States) to remove DNA contamination from total RNA samples. The first strand of cDNA was synthesized using 1 μg of DNase free total RNA primed with oligo dT primers in a 20 μl reaction mix

using Maxima H Minus M-MuLV reverse transcriptase (Thermo Scientific, United States) following manufacturer's instructions. Before using as a template in quantitative real time-polymerase chain reactions (qRT-PCR), the cDNA products were five fold diluted with deionized water. qRT-PCR was performed using 2X Brilliant III SYBR[®] Green QPCR (Agilent Technologies, United States) on Stratagene Mx3000P (Agilent Technologies, United States) in triplicates. A constitutive gene actin from rice was used as an internal control (Verma et al., 2016). The amount of transcript accumulated for each target gene normalized to the internal control was examined using $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen, 2001). The primers used for qRT-PCR analysis were designed from sequences of the respective genes downloaded from the National Center for Biotechnology Information (NCBI) using the IDT Primer Quest software (Table 1). The qRT-PCR cycling conditions were: initial denaturation at 95°C for 10 min, 95°C for 30 s, and 60°C for 1 min for 40 cycles followed by melt curve analysis at 95°C for 1 min, 60°C for 30 s, and 95°C for 30 s. The heat map for gene expression profiles were generated using TIGR MultiExperiment viewer (MeV 4) software package (Saeed et al., 2003).

Statistical Analysis

All experimental results are expressed as mean with standard deviation (mean \pm SD). To test the significance between mean values of control and stressed plants or SN13-inoculated unstressed and stressed plants, one way analysis of variance (ANOVA) was performed, and comparison among means was carried out using Duncan multiple range test (DMRT) at $P < 0.05$ with the help of SPSS software version 16.0 (SPSS Inc./IBM Corp., Chicago, IL, United States). All results were graphically presented using Graph Pad Prism software (version 5.03, San Diego, CA, United States). Principal component analysis (PCA) to delineate biochemical traits and gene expression differentiation among the treatments was performed using R 3.4.1 package.

RESULTS

Modulation in Biochemical Parameters under SN13 Inoculation

To combat the adverse effects of various environmental stresses, plants have evolved complex mechanisms for their better survival, growth and adaptation. PGPR also regulate morphophysiological, biochemical and molecular responses in plants. Therefore, in order to analyze the effects of SN13 on various biochemical parameters, rice seedlings were collected at different time points of various abiotic stresses namely, salt, drought, desiccation, heat, cold, and freeze, and phytohormone treatments viz. ABA, SA, JA, and ethephon after 1, 3, 10, and 24 h with or without SN13-inoculation.

Proline

Proline is an osmoprotectant and has been suggested to contribute in osmoregulation during stress tolerance in plants. In inoculated and uninoculated rice seedlings, accumulation of proline was determined under all stress treatments. Interestingly,

TABLE 1 | List of primers used in qRT-PCR analysis.

Gene(s)	NCBI accession no.	Primer sequence	Amplicon length
DHN	XM_015762124	F-5' CTCCAGCTCATCCTCTGA 3' R-5' GAGCTTCCTCTGATCTTCTC 3'	79
GST	XM_015759043	F-5' CCTACGTCAACCAGAGTGAA 3' R-5' TCTTGTGCGGAGATCCT 3'	81
LEA	XM_015782086	F-5' TGAGCAGGTGAAGAGCA 3' R-5' GCAGAGGTGTCTTGTG 3'	111
NAM	XM_015760382	F-5' GGAGTCCATGGAATTGGAAAG 3' R-5' GGCACTGCCAACAGGAGTT 3'	101
GRAM	XM_015772803	F-5' GCCCGCCTCCAAGAACATA 3' R-5' TCTCCAAACCTCTCCCCATCT 3'	69
NRAMP6	XM_015792141	F-5' CTACTCCCTCAGCACAAAGC 3' R-5' TTCCCTGAACGTCAGGTAGAT 3'	138
Actin	XM_015761709	F-5' GAGTATGATGAGTCGGGTCCAG 3' R-5' ACACCAACAATCCCAAACAGAG 3'	143

in all abiotic stresses and phytohormone treatments proline content was found to be significantly higher in SN13-inoculated seedlings in comparison to uninoculated. Salt, drought, desiccation, heat, and cold stress treatments led to significantly progressive increase in proline content at all stress durations in inoculated seedlings than uninoculated ones with maximum accumulation at 3 h in most of the stresses (**Figures 1A–J**). In comparison to control, uninoculated seedlings at 24 h under salt, drought, desiccation, heat, and cold showed percentage increment of 37, 56, 100, 243, and 31% while SN13-inoculated 24 h stressed seedlings showed increase in proline content by 81, 43, 180, 176, and 62%, respectively (**Figures 1A–E**). On the other hand, inoculated rice seedlings subjected to salt, drought, desiccation, and cold stresses showed enhanced proline content by 72, 20, 84, and 62%, respectively, at 24 h in comparison to the respective stressed uninoculated seedlings (**Figures 1A–C,E**). However, at 24 h of heat stress only 5% increase in proline content was recorded in inoculated seedlings as compared to uninoculated (**Figure 1D**). Furthermore, SN13 inoculated seedlings under freeze and all phytohormone treatments namely, ABA, SA, JA, and ethephon showed comparatively higher proline content viz. 86, 66, 24, 33, and 119%, respectively, at 3 h of stress as compared to inoculated controls (**Figures 1F–J**). While at 24 h of abovementioned stresses, a decline in proline content by 30, 34, 23, 22, and 39%, respectively, was observed in inoculated seedlings. However, at the same time point, a less significant difference in uninoculated rice seedlings was recorded for proline content under freeze and phytohormone treatments.

Total Soluble Sugar

Total soluble sugar is also a well-known compatible osmolyte that help plants to withstand various environmental stresses via maintaining the stability of membranes. Accumulation of soluble sugar was significantly higher in SN13-inoculated plants in comparison to the uninoculated seedlings at all durations of all applied stresses. Both inoculated and uninoculated seedlings during salt, drought, desiccation, heat, and cold stress showed

significantly progressive increment in TSS content till 3 h and at 24 h of stress treatments (**Figures 2A–E**). In comparison to inoculated unstressed seedlings, the inoculated seedlings subjected to abovementioned stresses, showed enhancement in TSS content by 126, 328, 157, 161, and 185%, respectively, at 3 h while an increment between ~200 and 400% was recorded for each stress at 24 h in inoculated stressed seedlings. Similarly, maximum TSS content (~180%) was found at 3 h of ABA and ethephon treatments in inoculated seedlings as compared to the inoculated control while after a decline at 10 h, an increment by 167 and 84%, respectively, was observed at 24 h (**Figures 2G,J**). Uninoculated seedlings under salt, drought, desiccation and heat stresses followed similar pattern as inoculated seedlings at all time points of stress treatments but degree of accumulation of TSS was comparatively lower than inoculated seedlings. Furthermore, both inoculated and uninoculated seedlings under freeze, SA and JA treatments showed progressive increase in level of soluble sugar with increasing duration. Under these treatments maximum level of soluble sugar was recorded at 24 h (~170, ~210, and ~200%, respectively) in both inoculated and uninoculated rice seedlings, as compared to their respective controls (**Figures 2F,H,I**).

Lipid Peroxidation

To check the membrane integrity lipid peroxidation analysis was performed by measuring total MDA content. During stress, MDA works as an indicator of extent of lipid peroxidation in living tissues, i.e., an increase in MDA denotes more membrane damage and vice versa. At all given treatments membrane destruction significantly increased with stress progression (**Figures 3A–J**). Under controlled conditions, MDA content was found to be slightly higher in inoculated seedlings as compared to the uninoculated seedlings, while under salt, drought, desiccation, heat, cold, and freeze stresses, SN13 inoculated seedlings showed significantly equivalent or lower MDA content in comparison to the uninoculated seedlings till 10 h (**Figures 3A–F**). However, at 24 h of salt, drought, desiccation, and heat stress, SN13-inoculated seedlings showed slightly higher MDA content by 17,

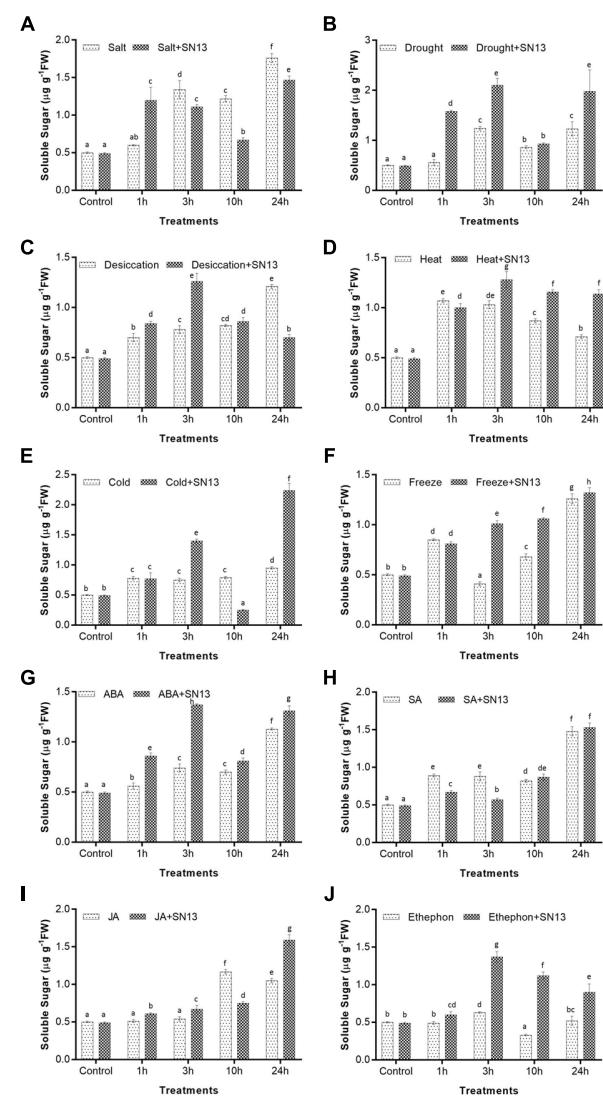
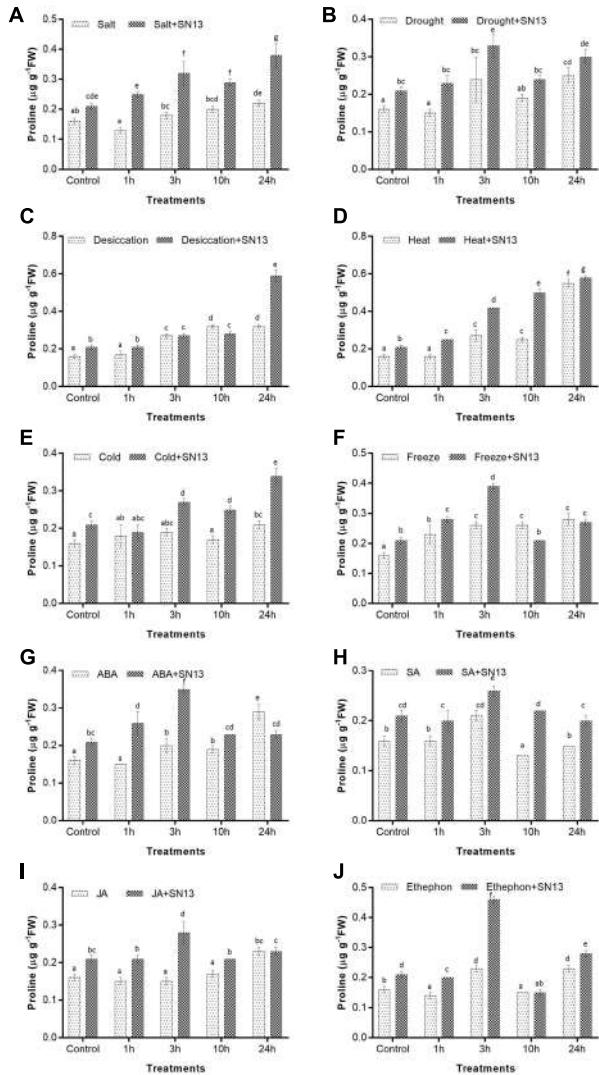


FIGURE 1 | Determination of proline in rice exposed to salt (A), drought (B), desiccation (C), heat (D), cold (E), freeze (F), ABA (G), SA (H), JA (I), and ethephon (J) stress at 1, 3, 10, and 24 h in the presence or absence of SN13. Data represent the means \pm SD of four independent experiments. Different letters on the graph indicate significant differences according to Duncan's test ($P \leq 0.05$).

11, 48, and 38%, respectively, in comparison to the uninoculated seedlings, while the same was not observed under cold and freezing stresses where the MDA content were lower by $\sim 5\%$ than uninoculated seedlings. On the other hand, the inoculated and uninoculated rice seedlings treated with ABA and ethephon at 3 h showed improved membrane integrity on account of lower MDA content, i.e., by 3 and 12% in ABA and 6 and 34% in ethephon, respectively, while an increase in MDA accumulation was observed till 24 h in both (Figures 3G,J). At the same time both inoculated and uninoculated JA-treated seedlings, showed irregular pattern for membrane ion leakage as MDA content was found to increase till 3 h (~ 10 and 117%, respectively) and at

24 h (~ 12 and $\sim 117\%$, respectively) (Figure 3I). Both inoculated and uninoculated seedlings under desiccation and SA treatment showed gradual increase in the level of MDA content until 10 h with an abrupt increase by $\sim 200\%$ at 24 h (Figures 3C,F).

SN13 Inoculation Alters Stress-Responsive Gene Expression

In order to verify our results for biochemical analyses, qRT-PCR analysis of six stress-responsive genes namely, dehydrin (DHN), glutathione S-transferase (GST), late embryogenesis abundant (LEA), no apical meristem (NAM), glucosyltransferases, Rab-like GTPase activators, myotubularin

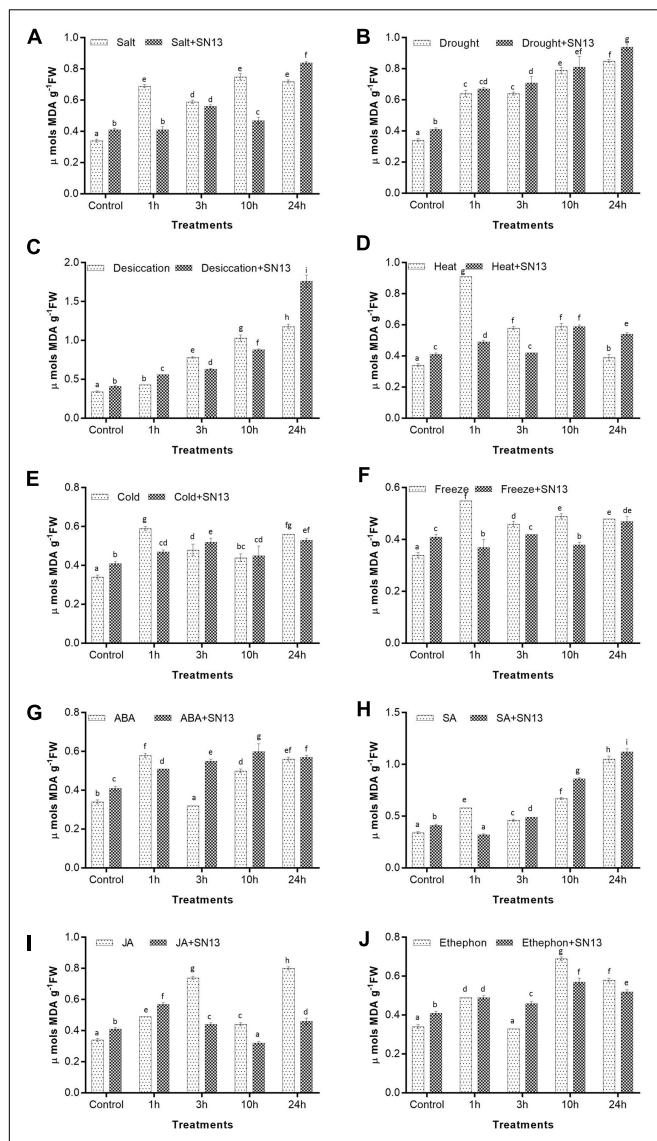


FIGURE 3 | Determination of MDA content in rice exposed to salt (A), drought (B), desiccation (C), heat (D), cold (E), freeze (F), ABA (G), SA (H), JA (I), and ethephon (J) stress at 1, 3, 10, and 24 h in the presence or absence of SN13. Data represent the means \pm SD of four independent experiments. Different letters on the graph indicate significant differences according to Duncan's test ($P \leq 0.05$).

(GRAM) and natural resistance-associated macrophage protein 6 (NRAMP6) genes were carried out at all-time points in rice roots under two representative abiotic stresses, i.e., salt and heat, and two representative phytohormone treatments, i.e., ABA and JA. All genes showed significant differential expression under abovementioned treatments and the gene expression profiling data was also found to be considerably in correlation with our biochemical results under these four stresses. Under unstressed conditions, the inoculated seedlings showed \sim 3-fold upregulation in DHN expression than uninoculated ones (Figures 4A–D). In both inoculated and uninoculated seedlings expression of DHN

in salt stressed plants increased gradually till 3 h with a \sim 11-fold induction as compared to control. In case of ABA treated inoculated seedlings maximum expression of DHN was observed at 3 h (\sim 11-fold) followed by 24 h (\sim 9-fold) as compared to control (Figure 4C). In contrast, uninoculated seedlings showed upto \sim 9-fold increase in DHN expression after 1 and 10 h of ABA treatment. Heat and JA stressed seedlings showed somewhat similar pattern of DHN expression (Figures 4B,D). Under heat and JA treatments, uninoculated seedlings showed maximum DHN expression (\sim 12-fold) at 3 h whereas SN13-inoculated seedlings showed maximum upregulation of \sim 12- and \sim 9-fold, respectively, at 1 h for both stresses.

Under unstressed conditions, GST expression was found to be upregulated by \sim 4.5-fold in inoculated seedlings as compared to uninoculated (Figures 4A–D). As compared to control, GST expression was found to be highest (\sim 7-fold) in inoculated seedlings at 3 h and in uninoculated seedlings at 10 h of salt stress. Further, 1 h heat stressed inoculated seedlings showed highest GST expression (\sim 8-fold) while a \sim 9-fold expression was observed at 24 h in uninoculated rice seedlings (Figure 4B). In case of ABA treatment, \sim 7-fold increase was observed in inoculated seedlings at 3 and 24 h as compared to control while uninoculated seedlings showed \sim 5-fold of upregulation at 10 h (Figure 4C). JA treated seedlings with SN13- inoculation showed \sim 6-fold upregulation in GST expression at 1 h while uninoculated seedlings showed \sim 8-fold expression at 3 h as compared to control (Figure 4D). However, a gradual decline in GST expression was recorded till 24 h in JA-treated seedlings under both inoculated and uninoculated seedlings.

On the other hand, SN13-inoculated control seedlings showed upto two fold upregulation in NAM and GRAM gene expression while no significant difference was found in the expression of LEA gene as compared to uninoculated control (Figures 4A–D). Under salt stress, LEA, NAM and GRAM genes showed highest expression (\sim 4- to \sim 5.5-fold) in SN13-inoculated seedlings at 3 and 10 h of stress as compared to control (Figure 4A). ABA-treated seedlings showed similar expression patterns for LEA, NAM, and GRAM genes as under salt stress (Figure 4C). Inoculated seedlings showed highest expression of LEA, NAM, and GRAM, i.e., \sim 7.8-, \sim 4.7- and \sim 5.3-fold, respectively, after 3 h of ABA application, while a gradual decline was observed in both inoculated and uninoculated seedlings till 24 h of stress. However, at 3 h of JA treatment, uninoculated seedlings showed upto six fold increment in LEA, NAM, and GRAM expression (Figure 4D). Furthermore, no significant change in expression of these genes was recorded till 24 h of stress. Unlike other stresses, all three genes showed distinct expression pattern under heat stress (Figure 4B). Maximum expression of NAM and GRAM genes was recorded to be \sim 4- and \sim 5-fold, respectively, in both inoculated and uninoculated seedlings at 3 h of heat stress, thereafter a slight reduction was observed in their expression with stress progression till 24 h time point. Maximum expression of LEA (\sim 5-fold) was recorded at 1 h of heat stress as compared to control. Interestingly, its expression was found to be more or less maintained till 24 h in both inoculated and uninoculated seedlings (Figure 4B).

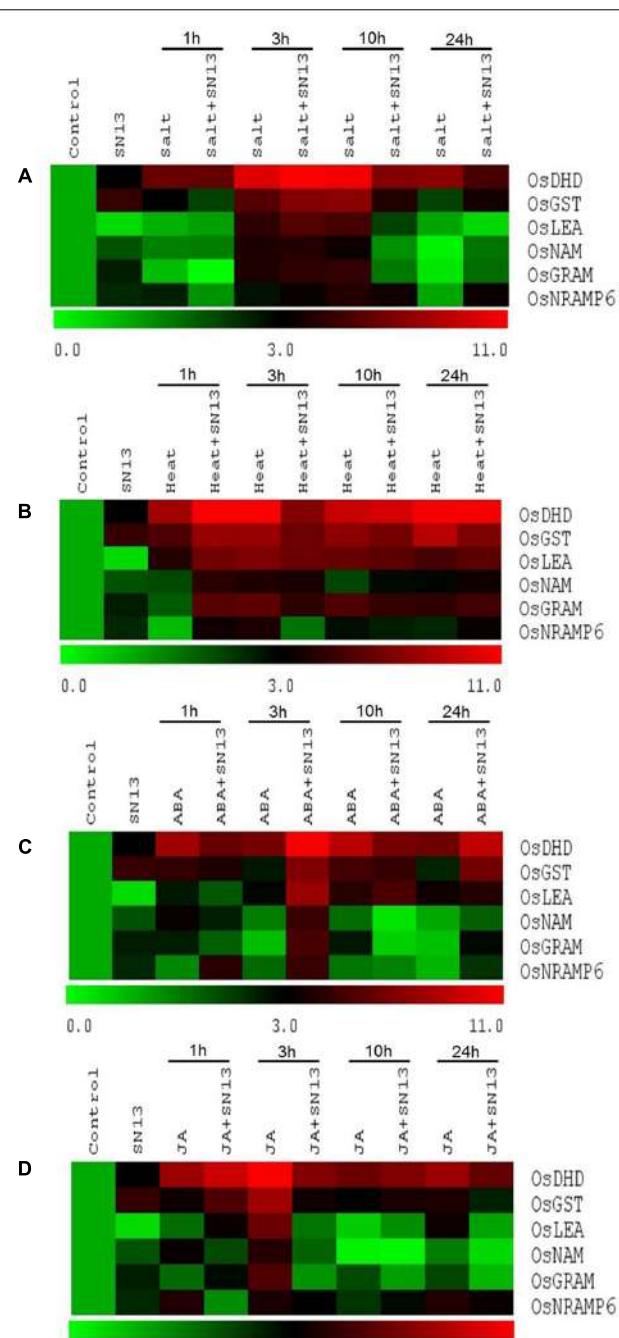


FIGURE 4 | Differential expression of genes in rice exposed to salt (A), heat (B), ABA (C), and JA (D) stress at 1, 3, 10, and 24 h in the presence or absence of SN13. The heat map has been generated based on the fold-change values in the treated sample when compared with its unstressed control sample. The color scale for fold-change values is shown at the bottom.

Expression of *NRAMP6* gene was ~2-fold higher in inoculated seedlings as compared to uninoculated seedlings under unstressed condition (Figures 4A–D). Under salt stress *NRAMP6* expression was highest at 10 h in both uninoculated and inoculated seedlings (~4- and ~3.5-fold, respectively) in

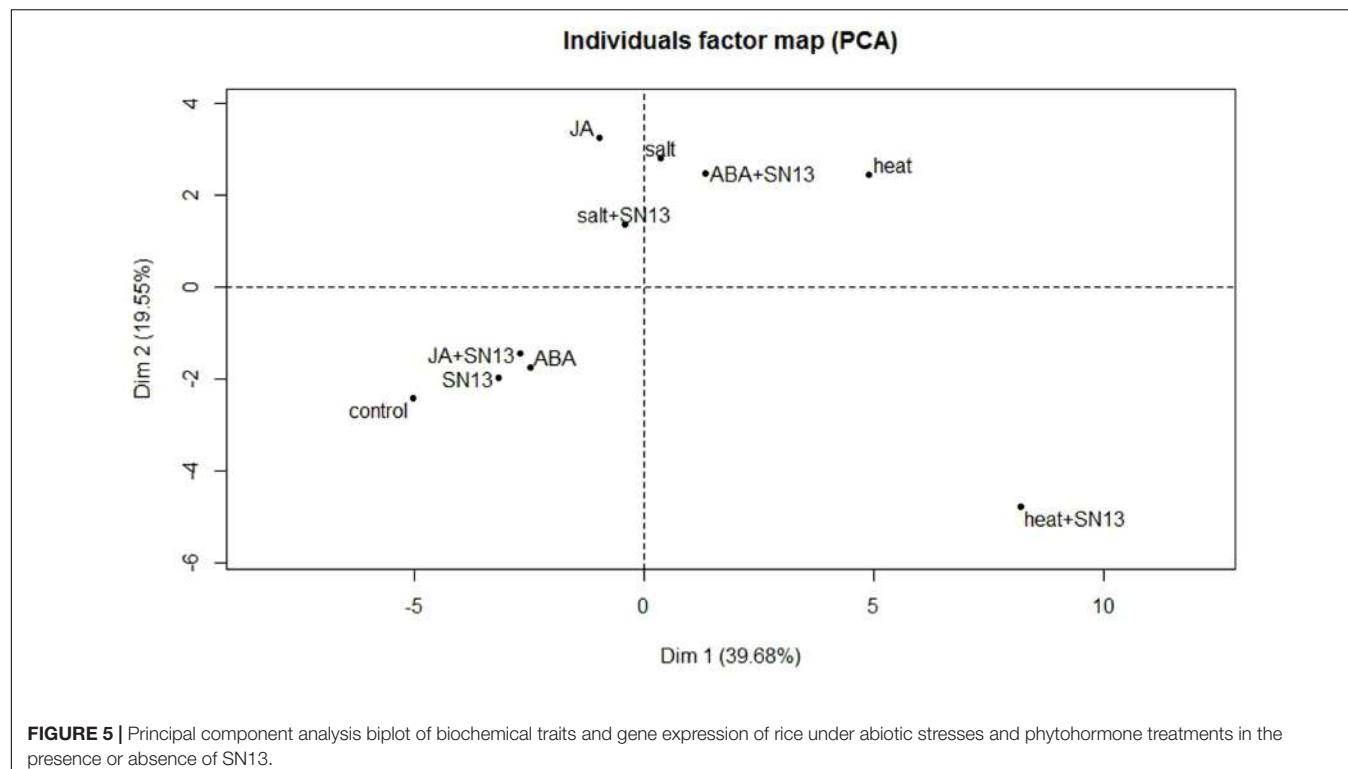
comparison to control (Figure 4A). After 10 h of salt stress, a ~3-fold reduction in *NRAMP6* expression was observed in uninoculated seedlings while no significant difference was found in inoculated seedlings. Expression of *NRAMP6* under heat stress was highest at 1 h in inoculated seedlings and at 3 h in uninoculated seedlings (~4-fold) in comparison to control (Figure 4B). A gradual decline in *NRAMP6* expression was observed in uninoculated seedlings while inoculated seedlings showed reduction only at 3 h of heat stress thereafter its expression increased upto ~3-fold at 24 h. In ABA-treated uninoculated seedlings, a ~1.5-fold expression of *NRAMP6* was observed at all-time points while SN13-inoculated seedlings showed ~4-fold upregulation after 1 and 3 h of ABA treatment with a significant reduction (~2-fold) at 10 and 24 h (Figure 4C). On the other hand, uninoculated JA-treated seedlings showed ~4-fold *NRAMP6* expression at 1, 3, and 24 h and ~2-fold at 10 h of stress as compared to control. While inoculated seedlings showed ~3-fold up-regulation in *NRAMP6* expression at 3, 10, and 24 h in comparison to control (Figure 4D). Expression patterns of all six genes, i.e., DHN, GST, LEA, NAM, GRAM and *NRAMP6* under salt, heat, ABA and JA treatments in both inoculated and uninoculated seedlings at all-time points have also been provided as individual graphical representations in supplementary information (Supplementary Figures S1A–F).

PCA of Abiotic Stresses and Phytohormone Treatments

In order to better understand the relationships, similarities and dissimilarities among the results for biochemical traits and gene expression, a multivariate PCA was carried out. The multivariate PCA lets a large number of variables to be lessened to only a few which largely account for majority of the variance in the observed experimental results. PCA was applied among biochemical parameters (proline, TSS and LP) and gene expression of all six genes to determine the interaction between abiotic stresses (salt and heat) and phytohormones (ABA and JA) under the influence of PGPR at all-time intervals (Supplementary Figures S2A–D). Dimension 1 (Dim1) and Dim2 accounted for ~80% of the total variance at all four time intervals. More specifically after assembling all-time points, Dim1 accounted for 39.68%, and Dim2 was responsible for 19.55% of the total variance (Figure 5). Mainly two clusters were formed in the biplot. Cluster including salt, heat, JA, salt+SN13 and ABA+SN13 have positive values at both the axis while the cluster containing control, SN13, ABA and JA+SN13 lied at negative values of both the axis. Heat+SN13 is located far from the clusters indicating dissimilar response from other clusters. In accordance, distinct clusters formed in biplot clearly exhibited correlation between abiotic stresses and phytohormones under influence of SN13 in rice.

DISCUSSION

In the present scenario of global climate change, adverse environmental conditions lead to significant reduction in growth, development and yield of crop plants (Lata et al., 2015). The development of new crop varieties is one of



the most established methods of crop improvement for stress management. Since, transgenic technology and molecular breeding are time consuming and labor intensive processes; use of plant growth promoting microbes is gaining wide popularity these days as an alternate strategy for improving stress tolerance of crop plants (Glick, 2014; Tiwari et al., 2016). Among various PGPR genus, *Bacillus* and *Pseudomonas* are most extensively studied rhizobacteria that promote plant growth and development (Kumar et al., 2011). Earlier studies have reported several strains of *Bacillus* spp. viz. *B. amyloliquefaciens*, *B. licheniformis*, *B. megaterium*, *B. pumilus*, and *B. subtilis*, etc., as well known rhizosphere residents of many crops with plant growth promoting activities (Klopper et al., 2004; Kumar et al., 2011). Nautiyal et al. (2013) studied PGPR traits of *B. amyloliquefaciens* and its effects on rice during salt stress. Srivastava et al. (2016) also reported *B. amyloliquefaciens* mediated enhanced production of rice under *Rhizoctonia* infection. However, our study demonstrates the regulatory role of PGPR, *B. amyloliquefaciens* NBRISN13 under various abiotic stresses and phytohormone treatments through biochemical studies and gene expression analyses of six stress responsive genes in 1-week-old rice seedlings. Unfavorable environmental conditions during early seedling stage in rice result in drastic reduction in growth resulting in lower yield potential and poor grain quality (Manikavelu et al., 2006; Farooq et al., 2009). To the best of our knowledge, until now none of the studies have been conducted to evaluate the effects of various abiotic stresses and phytohormones on stress tolerance abilities of rice at early seedling stage in the presence of a PGPR.

Accumulation of compatible osmolytes such as proline, soluble sugars, glycine betaine, trehalose, etc., help plants to overcome abiotic stresses by maintaining osmotic turgor (Tiwari et al., 2016; Zandalinas et al., 2017). Elevated levels of proline, TSS, betaine in plants have also been correlated with enhanced stress tolerance in previous studies (Lata et al., 2015; Tiwari et al., 2016). Accordingly this study also reports an increase in proline and TSS content in rice cv. Saryu-52 subjected to six abiotic stresses and four phytohormone treatments. In general, proline and TSS content of inoculated rice plants showed significant time-dependent increase as compared to the non-inoculated seedlings under all stresses. This increase in the level of proline and TSS upon SN13 inoculation can be associated with improved plant health under various stresses resulting in better stress tolerance of rice seedlings. Similar increase in proline and TSS content have also been reported in salt stressed wheat seedlings inoculated with halo-tolerant PGPR *Dietzia natronolimnaea* (Bharti et al., 2016). Accordingly, Khan et al. (2016) observed that the inoculation with *Bacillus pumilus* improved proline content of rice seedlings subjected to salt stress. Nautiyal et al. (2013) also reported enhanced proline content in 1-month-old rice seedlings inoculated with SN13 when subjected to salt stress. Similar observations have also been reported in maize under drought stress (Kandowangko et al., 2009; García et al., 2017). Increased level of soluble sugar in maize on inoculation with *Pseudomonas* spp. under drought stress was correlated with better stress tolerance (Sandhya et al., 2010). Interestingly, high proline and TSS accumulation was observed after 3 h of stress progression in most of the abiotic stresses and all phytohormone treatments. This could be due to an osmotic adjustment as a result

of increased synthesis of osmolytes. Similar observation was made by Jain and Chattopadhyay (2010) in chickpea subjected to drought stress. Further our results indicated a more or less similar pattern of proline accumulation under drought, salt, ABA, SA, JA, and ethephon treatments with a higher accumulation in inoculated seedlings as compared to uninoculated ones. Likewise, TSS accumulation was found to follow a similar pattern as that of proline accumulation in salt, drought, dessication, cold, ABA, and ethephon treatments indicating a complex SN13-mediated cross-talk among various abiotic stresses and phytohormones. Though such observations have not been reported earlier for any plant-PGPR interaction involving so many abiotic stresses and phytohormones, however, there are numerous reports on extensive cross-talk among various abiotic stresses and phytohormones (Fahad et al., 2015; Tiwari et al., 2017). Several phytohormones like ABA, SA, JA, and ET have been reported to be central to drought, salt, cold, and heat stress responses in various plants (Lata et al., 2011b) while other phytohormones such as gibberellins, brassinosteroids, auxin, cytokinins, etc., interact with other phytohormones and stress-related genes to maintain a balanced plant growth and development (Kohli et al., 2013).

Malondialdehyde (MDA) is one of the end products of polyunsaturated fatty acids peroxidation in phospholipids and is responsible for cell membrane damage (Sharma et al., 2012). MDA accumulation is an indication of stress-induced LP of cellular membrane lipids and is often considered a marker for increased oxidative damage (Lata et al., 2011a). In our findings LP was significantly lower at all durations of short-term salt, drought, and desiccation stresses upon SN13 inoculation except at late stress, i.e., 24 h. This might be possible that initially layer of mucilage around root protects plant by direct exposure to stress but later bacterial adherence to root or root cortex changes membrane permeability to some extent and leads to slight increase in membrane damage. Ongena and Jacques (2008) have also reported interaction of *Bacillus* with plant membrane as a biocontrol via altering membrane structure. Pandey et al. (2016) and Meena et al. (2017) reported increased MDA content in rice treated with *Trichoderma*. However, estimation of LP at subsequent later durations (beyond 24 h) of bacterial colonization may be an interesting subject area of study in these stresses. One of the previous studies also reported an increase in MDA content at initial stages of *Burkholderia phytophaga* inoculation in *Vitis vinifera* under cold stress and a decrease in MDA content was recorded at later stages suggesting the stress ameliorating properties of the PGPR (Theocharis et al., 2012). On the other hand, heat, cold, and freeze stresses did not show significant alteration in MDA content upon SN13 inoculation at all durations. This may be due to the fact SN13 may not be directly involved in maintaining membrane integrity under these stresses owing to their poor tolerance levels to extreme temperatures.

Further, PGPR-mediated activation of numerous genes in response to abiotic stresses has recently been reported in many crop plants including rice (Nautiyal et al., 2013; Kim et al., 2014; Tiwari et al., 2016). However, molecular basis of PGPR-plant interactions with respect to abiotic stress tolerance and

phytohormone treatments in rice remain largely unknown. Therefore, in order to understand the changes at molecular level during rice-SN13 interaction under various stresses, expression analyses of a few stress-responsive genes through qRT-PCR was performed.

LEA and *DHN* are mainly involved in stress tolerance and hence, act as marker genes for plant stress response (Tiwari et al., 2016). Overexpression of these genes has been reported to provide tolerance to various abiotic stresses in several crop plants (RoyChoudhury et al., 2007; Kumar et al., 2014). Further, overexpression of dehydration responsive element binding (*DREB*) genes in *Arabidopsis* and rice is also known to increase the expression of *LEA* and dehydrins (Lata and Prasad, 2011). In present study, the expression of these genes increase at all-time points with maximum expression at 3 and 10 h of all four applied treatments in comparison to control, indicating their correlation with an increased osmolyte synthesis at these durations. While SN13-inoculation relatively down regulates the expression of *LEA* and *DHN* under salt and heat stress as well as ABA treatment at all durations indicating the crucial role of SN13 in stress alleviation in 1-week-old rice seedlings. However, *LEA* and *DHN* expression is significantly higher at 3 h of the abovementioned stresses which may be due to their active role in osmolyte biosynthesis and subsequently osmotic adjustment. This result is also in accordance to our biochemical results for proline and TSS. Recently, *Trichoderma harzianum*, a rhizosphere occupants reported in stress mitigation in rice genotypes due to upregulation of dehydrin and other genes (Pandey et al., 2016; Meena et al., 2017). Similarly, the expression of *LEA* increases on application of *B. subtilis* in *Brachypodium* under drought stress (Gagne-Bourque et al., 2015) as well as in chickpea upon *P. putida* inoculation (Tiwari et al., 2016). Further, increased expression of *DHN* was reported by Kumar et al. (2014) in rice and Kosová et al. (2014) in barley and wheat under salt, drought and cold stresses. Interestingly, many *DHNS* have been identified in plants including *Arabidopsis thaliana* and wheat that are up-regulated by exogenous ABA under drought stress (Lv et al., 2017). Richard et al. (2000) reported exogenous application of JA leads to increase in *DHN* expression in white spruce. Differential expression of *DHN* and *LEA* under salt, heat, ABA, and JA treatments in our study suggests an extensive SN13-mediated cross-talk among them. *P. putida* treated tolerant and sensitive chickpea cultivars also showed differential gene expression for abiotic stress-responsive genes *DHN* and *LEA* as well as for *MYC2* and *PR1* genes which involved in JA and SA signaling, respectively, under drought stress (Tiwari et al., 2016).

Glutathione S-transferases (GSTs) are ubiquitous enzymes with antioxidant properties that help in detoxification via converting oxidatively produced compounds to reduced glutathione, thus facilitating their removal, sequestration, or metabolism (Dalton et al., 2009). Increased expression of GST in SN13-inoculated rice seedlings at all stress durations with maximum expression at 3 h indicates an induction of this defense enzyme due to SN13 colonization. Similar observation was also reported by Srivastava et al. (2012) in *Arabidopsis* upon *P. putida* inoculation. Kandasamy et al. (2009) also reported

upregulation in GST expression on treatment of *P. fluorescens* in rice.

NAM TFs have been reported to play significant role in abiotic stress tolerance in various crop plants (Nakashima et al., 2009). Increased expression of NAM gene on exposure to various stresses in rice is in accordance to previous studies (Nguyen et al., 2015; Tiwari et al., 2016). Further, its relatively increased transcript accumulation in SN13-inoculated rice seedlings at all stress durations with a few exceptions shows a positive regulation of NAM by SN13. Wang et al. (2005) and Tiwari et al. (2016) have also demonstrated *Pseudomonas* spp.-induced expression of NAM in *Arabidopsis* and chickpea, respectively, via gene expression profiling studies, and their role were speculated in PGPR-mediated stress tolerance.

GRAM domain containing genes are likely to be involved in membrane associated processes such as intracellular protein or lipid binding signaling pathways (Doerks et al., 2000; Jiang et al., 2008). Baron et al. (2014) also reported that GRAM-domain containing genes show responsiveness to several phytohormone and abiotic stresses. Upregulation of GRAM by ABA was also reported (Liu et al., 1999; Jiang et al., 2008). In our findings, increased expression of GRAM in both inoculated and uninoculated seedlings at early stages of stress treatments suggest an SN13-mediated gene expression modulation during initial stress signal transduction events in rice.

NRAMP genes in plants are known to encode intracellular metal transporters with capacity to transport both the metal nutrient iron (Fe) and the toxic metal cadmium (Cd) (Thomine et al., 2000; Cailliatte et al., 2009). However, their role in other abiotic stresses and phytohormones has not been elucidated till date. Interestingly, NRAMP6 was found to be highly up-regulated in one of our SN13-induced salt stress transcript profiling study of rice seedlings (Unpublished). Accordingly, in this study this

gene was found to be differentially expressed under all stresses with maximum expression at 3 h. It indicates possible NRAMP-regulated stress alleviation in rice. Our results and previous evidences regarding stress alleviation by PGPR suggest the crucial role of SN13 in positively modulating gene expression under various stresses and also indicate a possible gene-regulated cross-talk among all stresses and phytohormones treatments.

CONCLUSION AND FUTURE PERSPECTIVES

This study highlights a beneficial bipartite plant–microbe interaction between rice seedlings and *B. amyloliquefaciens* SN13 under short-term abiotic stresses and phytohormone treatments. Taken together our results indicate that the abiotic stress amelioration capacity of rice seedlings have been significantly improved with SN13-inoculation under all stresses. Stress-induced symptoms in rice such as membrane integrity, accumulation of osmoprotectants, and expression of marker genes were significantly improved in presence of SN13. However, a more detailed study on the role of SN13 in improving stress tolerance of rice at subsequent developmental stages can be an interesting topic for further investigation. Based on differential responses of rice seedlings to abiotic stresses and phytohormones, PCA analysis confirmed basis for a holistic view on SN13 inoculation effects on rice response to different abiotic stresses and phytohormone treatments. This highlighted a possible PGPR-induced cross-talk among abiotic stresses and phytohormones (Figure 6). It can be deduced that SN13-responsive cross-talk is most extensive among all four phytohormones and salt and drought stresses as compared to heat, desiccation, cold, and freeze. Our results thus paves way for

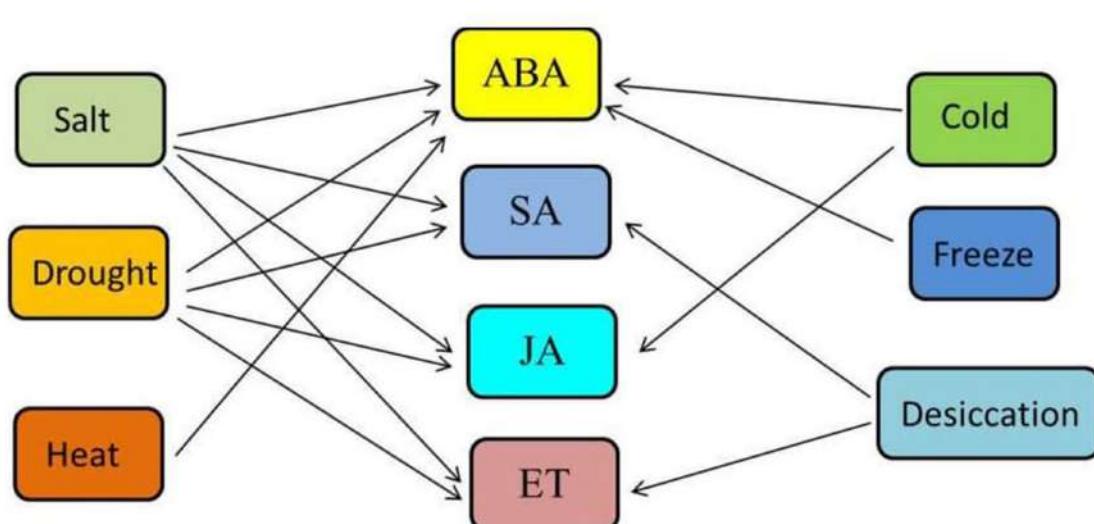


FIGURE 6 | A model of the SN13-responsive cross-talk among various abiotic stresses and phytohormones operating in rice is created based on biochemical analyses and gene expression profiling. The phytohormones and abiotic stress interactions shown here is only a fraction of stress signal cross-talk occurring in rice when exposed to SN13 and different other cross-talk points are yet to be discovered.

a more detailed understanding of various cross-talk points among phytohormones and stress signaling cascades in response to beneficial microbe(s) for their effective utilization in developing crop varieties with improved stress tolerance.

AUTHOR CONTRIBUTIONS

CL conceived and designed research. ST conducted experiments. ST, CL, VP, and PC analyzed data. CL and ST wrote the manuscript.

ACKNOWLEDGMENTS

The study was a part of the In-house project “Plant growth promoting rhizobacteria mediated stress management for

REFERENCES

- Asgher, M., Khan, M. I. R., Anjum, N. A., and Khan, N. A. (2015). Minimizing toxicity of cadmium in plants-role of plant growth regulators. *Protoplasma* 252, 399–413. doi: 10.1007/s00709-014-0710-4
- Bari, R., and Jones, J. D. G. (2009). Role of plant hormones in plant defence responses. *Plant Mol. Biol.* 69, 473–488. doi: 10.1007/s11103-008-9435-0
- Baron, K. N., Schroeder, D. F., and Stasolla, C. (2014). GEM-Related 5 (GER5), an ABA and stress-responsive GRAM domain protein regulating seed development and inflorescence architecture. *Plant Sci.* 223, 153–166. doi: 10.1016/j.plantsci.2014.03.017
- Bharti, N., Pandey, S. S., Barnawal, D., Patel, V. K., and Kalra, A. (2016). Plant growth promoting rhizobacteria *Dietzia natronolimnaea* modulates the expression of stress responsive genes providing protection of wheat from salinity stress. *Sci. Rep.* 6:34768. doi: 10.1038/srep34768
- Cabello-Conejo, M. I., Prieto-Fernández, Á., and Kidd, P. S. (2014). Exogenous treatments with phytohormones can improve growth and nickel yield of hyperaccumulating plants. *Sci. Total Environ.* 494, 1–8. doi: 10.1016/j.scitotenv.2014.06.102
- Cailliatte, R., Lapeyre, B., Briat, J. F., Mari, S., and Curie, C. (2009). The NRAMP metal transporter contributes to cadmium toxicity. *Biochem. J.* 422, 217–228. doi: 10.1042/BJ20090655
- Carillo, P., and Gibon, Y. (2011). *Protocol: Extraction and Determination of Proline*. PrometheusWiki01/2011. Available at: http://www.researchgate.net/publication/211353600_PROTOCOL_Extraction_and_determination_of_proline
- Dahro, B., Wang, F., Peng, T., and Liu, J. H. (2016). PtrA/NINV, an alkaline/neutral invertase gene of *Poncirus trifoliata*, confers enhanced tolerance to multiple abiotic stresses by modulating ROS levels and maintaining photosynthetic efficiency. *BMC Plant Biol.* 16:76. doi: 10.1186/s12870-016-0761-0
- Dalton, D. A., Boniface, C., Turner, Z., Lindahl, A., Kim, H. J., Jelinek, L., et al. (2009). Physiological roles of glutathione S-transferases in soybean root nodules. *Plant Physiol.* 150, 521–530. doi: 10.1104/pp.109.136630
- Doerks, T., Strauss, M., Brendel, M., and Bork, P. (2000). GRAM, a novel domain in glucosyltransferases, myotubularins and other putative membrane-associated proteins. *Trends Biochem. Sci.* 25, 483–485. doi: 10.1016/S0968-0004(00)01664-9
- Du, H., Liu, H., and Xiong, L. (2013). Endogenous auxin and jasmonic acid levels are differentially modulated by abiotic stresses in rice. *Front. Plant Sci.* 4:397. doi: 10.3389/fpls.2013.00397
- DuBois, M., Gilles, K., Hamilton, J., Rebers, P., and Smith, F. (1956). Colorimetric method for determination of sugars and related substances. *Anal. Chem.* 28, 350–356. doi: 10.1021/ac60111a017
- increasing crop productivity” (OLP0091) supported by core grant from the Council of Scientific and Industrial Research (CSIR), New Delhi, India.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01510/full#supplementary-material>

FIGURE S1 | Differential expression of DHN (**A**), GST (**B**), LEA (**C**), NAM (**D**), GRAM (**E**), and NRAMP6 (**F**) in rice exposed to salt, heat, ABA, and JA at 1, 3, 10, and 24 h in the presence or absence of SN13. Data represent the means ± SD of three independent experiments. Different letters on the graph indicate significant differences according to Duncan's test ($P \leq 0.05$).

FIGURE S2 | Principal component analysis biplot of biochemical traits and gene expression of rice at 1 h (**A**), 3 h (**B**), 10 h (**C**) and 24 h (**D**) under abiotic stresses and phytohormone treatments in the presence or absence of SN13.

- Dubois, M., Skirycz, A., Claeys, H., Maleux, K., Dhondt, S., De Bodt, S., et al. (2013). ETHYLENE RESPONSE FACTOR6 acts as a central regulator of leaf growth under water-limiting conditions in *Arabidopsis*. *Plant Physiol.* 162, 319–332. doi: 10.1104/pp.113.216341
- Fahad, S., Hussain, S., Bano, A., Saud, S., Hassan, S., Shan, D., et al. (2015). Potential role of phytohormones and plant growth-promoting rhizobacteria in abiotic stresses: consequences for changing environment. *Environ. Sci. Pol. Res.* 22, 4907–4921. doi: 10.1007/s11356-014-3754-2
- Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., and Basra, S. M. A. (2009). “Plant drought stress: effects, mechanisms and management” in *Sustainable Agriculture*, eds E. Lichtfouse, M. Navarrete, P. Debaeke, V. Souchere, and C. Alberola (Dordrecht: Springer), 153–188.
- Gagne-Bourque, F., Mayer, B. F., Charron, J. B., Vali, H., Bertrand, A., and Jabaji, S. (2015). Accelerated growth rate and increased drought stress resilience of the model grass *Brachypodium distachyon* colonized by *Bacillus subtilis* B26. *PLoS ONE* 10:e0130456. doi: 10.1371/journal.pone.0130456
- García, J. E., Maroniche, G., Creus, C., Suárez-Rodríguez, R., Ramírez-Trujillo, J. A., and Groppa, M. D. (2017). In vitro PGPR properties and osmotic tolerance of different Azospirillum native strains and their effects on growth of maize under drought stress. *Microbiol. Res.* 202, 21–29. doi: 10.1016/j.micres.2017.04.007
- Glick, B. R. (2014). Bacteria with ACC deaminase can promote plant growth and help to feed the world. *Microbiol. Res.* 169, 30–39. doi: 10.1016/j.micres.2013.09.009
- Glick, B. R., Cheng, Z., Czarny, J., and Duan, J. (2007). Promotion of plant growth by ACC deaminase-producing soil bacteria. *Eur. J. Plant Pathol.* 119, 329–339. doi: 10.1007/s10658-007-9162-4
- Großkinsky, D. K., van der Graaff, E., and Roitsch, T. (2016). “Regulation of abiotic and biotic stress responses by plant hormones,” in *Plant Pathogen Resistance Biotechnology*, ed. D. B. Collinge (New York, NY: Wiley and Sons), 131–134. doi: 10.1002/978111867716.ch7
- Heath, R. L., and Packer, L. (1968). Photoperoxidation in isolated chloroplasts. I. Kinetics and stoichiometry of fatty acid peroxidation. *Arch. Biochem. Biophys.* 125, 189–198. doi: 10.1016/0003-9861(68)90654-1
- Jain, D., and Chattopadhyay, D. (2010). Analysis of gene expression in response to water deficit of chickpea (*Cicer arietinum* L.) varieties differing in drought tolerance. *BMC Plant Biol.* 10:24. doi: 10.1186/1471-2229-10-24
- Jiang, S. Y., Ramamoorthy, R., and Ramachandran, S. (2008). Comparative transcriptional profiling and evolutionary analysis of the GRAM domain family in eukaryotes. *Dev. Biol.* 314, 418–432. doi: 10.1016/j.ydbio.2007.11.031
- Kandasamy, S., Loganathan, K., Muthuraj, R., Duraisamy, S., Seetharaman, S., Thiruvengadam, R., et al. (2009). Understanding the molecular basis of plant

- growth promotional effect of *Pseudomonas fluorescens* on rice through protein profiling. *Proteome Sci.* 7:47. doi: 10.1186/1477-5956-7-47
- Kandowangko, N. Y., Suryatmana, G., Nurlaeny, N., and Simanungkalit, R. D. M. (2009). Proline and abscisic acid content in droughted corn plant inoculated with *Azospirillum* sp. and Arbuscular mycorrhizae fungi. *Hayati J. Biosci.* 16, 15–20. doi: 10.4308/hjb.16.1.15
- Kaur, G., Kumar, S., Nayyar, H., and Upadhyaya, H. D. (2008). Cold stress injury during the pod-filling phase in Chickpea (*Cicer arietinum* L.): effects on quantitative and qualitative components of seeds. *J. Agron. Crop Sci.* 194, 457–464.
- Khan, A., Zhao, X. Q., Javed, M. T., Khan, K. S., Bano, A., Shen, R. F., et al. (2016). *Bacillus pumilus* enhances tolerance in rice (*Oryza sativa* L.) to combined stresses of NaCl and high boron due to limited uptake of Na+. *Environ. Exp. Bot.* 124, 120–129. doi: 10.1016/j.envexpbot.2015.12.011
- Khan, N. A., Nazar, R., Iqbal, N., and Anjum, N. A. (2012). *Phytohormones and Abiotic Stress Tolerance in Plants*. Berlin: Springer. doi: 10.1007/978-3-642-25829-9
- Kim, K., Jang, Y. J., Lee, S. M., Oh, B. T., Chae, J. C., and Lee, K. J. (2014). Alleviation of salt stress by *Enterobacter* sp. Ej01 in tomato and *Arabidopsis* is accompanied by up-regulation of conserved salinity responsive factors in plants. *Mol. Cells* 37, 109–117. doi: 10.14348/molcells.2014.2239
- Kim, S. T., Kim, S. G., Hwang, D. H., Kang, S. Y., Koo, S. C., Cho, M. J., et al. (2004). Expression of a salt-induced protein (SALT) in suspension-cultured cells and leaves of rice following exposure to fungal elicitor and phytohormones. *Plant Cell Rep.* 23, 256–262. doi: 10.1007/s00299-004-0836-5
- Kloepper, J. W., Ryu, C. M., and Zhang, S. (2004). Induced systemic resistance and promotion of plant growth by *Bacillus* spp. *Phytopathology* 94, 1259–1266. doi: 10.1094/PHYTO.2004.94.11.1259
- Kohli, A., Sreenivasulu, N., Lakshmanan, P., and Kumar, P. P. (2013). The phytohormone crosstalk paradigm takes center stage in understanding how plants respond to abiotic stress. *Plant Cell Rep.* 32, 945–957. doi: 10.1007/s00299-013-1461-y
- Koramutla, M. K., Kaur, A., Negi, M., Venkatachalam, P., and Bhattacharya, R. (2014). Elicitation of jasmonate-mediated host defense in *Brassica juncea* (L.) attenuates population growth of mustard aphid *Lipaphis erysimi* (Kalt.). *Planta* 240, 177–194. doi: 10.1007/s00425-014-2073-7
- Kosová, K., Vítámvás, P., and Prášil, I. T. (2014). Wheat and barley dehydrins under cold, drought, and salinity—what can LEA-II proteins tell us about plant stress response? *Front. Plant Sci.* 5:343. doi: 10.3389/fpls.2014.00343
- Kumar, A., Prakash, A., and Johri, B. N. (2011). “*Bacillus* as PGPR in crop ecosystem,” in *Bacteria in Agrobiology: Crop Ecosystems*, ed. D. K. Maheshwari (Berlin: Springer), 37–59. doi: 10.1007/978-3-642-18357-7_2
- Kumar, M., Lee, S. C., Kim, J. Y., Kim, S. J., and Kim, S. R. (2014). Over-expression of dehydrin gene, OsDhn1, improves drought and salt stress tolerance through scavenging of reactive oxygen species in rice (*Oryza sativa* L.). *J. Plant Biol.* 57, 383–393. doi: 10.1007/s12374-014-0487-1
- Kundan, R., Pant, G., Jadon, N., and Agrawal, P. K. (2015). Plant growth promoting rhizobacteria: mechanism and current prospective. *J. Fertil. Pestic.* 6, 155. doi: 10.4172/2471-2728.1000155
- Lata, C., Muthamilarasan, M., and Prasad, M. (2015). “Drought stress responses and signal transduction in plants,” in *Elucidation of Abiotic Stress Signaling in Plants*, ed. G. K. Pandey (New York, NY: Springer), 195–225. doi: 10.1007/978-1-4939-2540-7_7
- Lata, C., and Prasad, M. (2011). Role of DREBs in regulation of abiotic stress responses in plants. *J. Exp. Bot.* 62, 4731–4748. doi: 10.1093/jxb/err210
- Lata, C., Jha, S., Dixit, V., Sreenivasulu, N., and Prasad, M. (2011a). Differential antioxidative responses to dehydration-induced oxidative stress in core set of foxtail millet cultivars (*Setaria italica* L.). *Protoplasma* 248, 817–828. doi: 10.1007/s00709-010-0257-y
- Lata, C., Yadav, A., and Prasad, M. (2011b). “Role of plant transcription factors in abiotic stress tolerance,” in *Abiotic stress/ Book2*, eds A. Shanker and B. Venkateswarlu (Rijeka: INTECH Open Access Publishers), 269–296.
- Liu, J. H., Luo, M., Cheng, K. J., Mohapatra, S. S., and Hill, R. D. (1999). Identification and characterization of a novel barley gene that is ABA-inducible and expressed specifically in embryo and aleurone. *J. Exp. Bot.* 50, 727–728. doi: 10.1093/jexbot/50.334.727
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_t}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Lv, A., Fan, N., Xie, J., Yuan, S., An, Y., and Zhou, P. (2017). Expression of CdDHN4, a novel YSK2-type dehydrin gene from bermuda grass, responses to drought stress through ABA-dependent signal pathway. *Front. Plant Sci.* 8:748. doi: 10.3389/fpls.2017.00748
- Manikavelu, A., Nadarajan, N., Ganesh, S. K., Gnanamalar, R. P., and Babu, R. C. (2006). Drought tolerance in rice: morphological and molecular genetic consideration. *Plant Growth Regul.* 50, 121–138. doi: 10.1007/s10725-006-9109-3
- Mantri, N., Patade, V., Penna, S., and Ford, R. (2012). “Abiotic stress responses in plants: present and future,” in *Abiotic Stress Responses in Plants*, eds P. Ahmad and M. N. V. Prasad (New York, NY: Springer), 1–19. doi: 10.1007/978-1-4614-0634-1-1
- Meena, K. K., Sorty, A. M., Bitla, U. M., Choudhary, K., Gupta, P., Pareek, A., et al. (2017). Abiotic stress responses and microbe-mediated mitigation in plants: the omics strategies. *Front. Plant Sci.* 8:172. doi: 10.3389/fpls.2017.00172
- Miller, G., Suzuki, N., Ciftci-Yilmaz, S., and Mittler, R. (2010). Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant Cell Environ.* 33, 453–467. doi: 10.1111/j.1365-3040.2009.02041.x
- Mishra, M. K., Singh, G., Tiwari, S., Singh, R., Kumari, N., and Misra, P. (2015). Characterization of *Arabidopsis* sterol glycosyltransferase TTG15/UGT80B1 role during freeze and heat stress. *Plant Signal. Behav.* 10:e1075682. doi: 10.1080/15592324.2015.1075682
- Nafisi, M., Fimognari, L., and Sakuragi, Y. (2015). Interplays between the cell wall and phytohormones in interaction between plants and necrotrophic pathogens. *Phytochemistry* 112, 63–71. doi: 10.1016/j.phytochem.2014.11.008
- Nakashima, K., Ito, Y., and Yamaguchi-Shinozaki, K. (2009). Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses. *Plant Physiol.* 149, 88–95. doi: 10.1104/pp.108.129791
- Nautiyal, C. S., Srivastava, S., Chauhan, P. S., Seem, K., Mishra, A., and Sopory, S. K. (2013). Plant growth-promoting bacteria *Bacillus amyloliquefaciens* NBRISN13 modulates gene expression profile of leaf and rhizosphere community in rice during salt stress. *Plant Physiol. Biochem.* 66, 1–9. doi: 10.1016/j.plaphy.2013.01.020
- Nguyen, K. H., Ha, C. V., Watanabe, Y., Tran, U. T., Esfahani, M. N., Nguyen, D. V., et al. (2015). Correlation between differential drought tolerability of two contrasting drought-responsive chickpea cultivars and differential expression of a subset of CaNAC genes under normal and dehydration conditions. *Front. Plant Sci.* 6:449. doi: 10.3389/fpls.2015.00449
- Onsgena, M., and Jacques, P. (2008). *Bacillus* lipopeptides: versatile weapons for plant disease biocontrol. *Trends Microbiol.* 16, 115–125. doi: 10.1016/j.tim.2007.12.009
- Pan, Y., Seymour, G. B., Lu, C., Hu, Z., Chen, X., and Chen, G. (2012). An ethylene response factor (ERF5) promoting adaptation to drought and salt tolerance in tomato. *Plant Cell Rep.* 31, 349–360. doi: 10.1007/s00299-011-1170-3
- Pandey, V., Ansari, M. W., Tula, S., Yadav, S., Sahoo, R. K., Shukla, N., et al. (2016). Dose-dependent response of *Trichoderma harzianum* in improving drought tolerance in rice genotypes. *Planta* 243, 1251–1264. doi: 10.1007/s00425-016-2482-x
- Park, Y. G., Mun, B. G., Kang, S. M., Hussain, A., Shahzad, R., Seo, C., et al. (2017). *Bacillus aryabhatai* SRB02 tolerates oxidative and nitrosative stress and promotes the growth of soybean by modulating the production of phytohormones. *PLoS ONE* 12:e0173203. doi: 10.1371/journal.pone.0173203
- Richard, S., Morency, M. J., Drevet, C., Jouanin, L., and Séguin, A. (2000). Isolation and characterization of a dehydrin gene from white spruce induced upon wounding, drought and cold stresses. *Plant Mol. Biol.* 43, 1–10. doi: 10.1023/A:1006453811911
- RoyChoudhury, A., Roy, C., and Sengupta, D. N. (2007). Transgenic tobacco plants overexpressing the heterologous lea gene Rab16A from rice during high salt and water deficit display enhanced tolerance to salinity stress. *Plant Cell Rep.* 26, 1839–1859. doi: 10.1007/s00299-007-0371-2
- Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003). TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 34, 374–378.
- Sandhya, V., Ali, S. Z., Grover, M., Reddy, G., and Venkateswarlu, B. (2010). Effect of plant growth promoting *Pseudomonas* spp. on compatible solutes,

- antioxidant status and plant growth of maize under drought stress. *Plant Growth Regul.* 62, 21–30. doi: 10.1007/s10725-010-9479-4
- Sharma, P., Jha, A. B., Dubey, R. S., and Pessarakli, M. (2012). Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *J. Bot.* 2012:217037. doi: 10.1155/2012/217037
- Srivastava, S., Bist, V., Srivastava, S., Singh, P. C., Trivedi, P. K., Asif, M. H., et al. (2016). Unraveling aspects of *Bacillus amyloliquefaciens* mediated enhanced production of rice under biotic stress of *Rhizoctonia solani*. *Front. Plant Sci.* 7:587. doi: 10.3389/fpls.2016.00587
- Srivastava, S., Chaudhry, V., Mishra, A., Chauhan, P. S., Rehman, A., Yadav, A., et al. (2012). Gene expression profiling through microarray analysis in *Arabidopsis thaliana* colonized by *Pseudomonas putida* MTCC5279, a plant growth promoting rhizobacterium. *Plant Signal. Behav.* 7, 235–245. doi: 10.4161/psb.18957
- Theocharis, A., Bordiec, S., Fernandez, O., Paquis, S., Dhondt-Cordelier, S., and Baillieul, F. (2012). *Burkholderia phytofirmans* PsJN primes *Vitis vinifera* L. and confers a better tolerance to low non-freezing temperatures. *Mol. Plant Microbe Interact.* 25, 241–249. doi: 10.1094/MPMI-05-11-0124
- Thomine, S., Wang, R., Ward, J. M., Crawford, N. M., and Schroeder, J. I. (2000). Cadmium and iron transport by members of a plant metal transporter family in *Arabidopsis* with homology to Nramp genes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4991–4996. doi: 10.1073/pnas.97.9.4991
- Tiwari, S., Lata, C., Chauhan, P. S., and Nautiyal, C. S. (2016). *Pseudomonas putida* attunes morphophysiological, biochemical and molecular responses in *Cicer arietinum* L. during drought stress and recovery. *Plant Physiol. Biochem.* 99, 108–117. doi: 10.1016/j.plaphy.2015.11.001
- Tiwari, S., Lata, C., Prasad, M., Prasad, V., and Chauhan, P. S. (2017). A functional genomic perspective on drought signalling and its cross-talk with various phytohormone-mediated signalling pathways in plants. *Curr. Genomics* doi: 10.2174/1389202918666170605083319 [Epub ahead of print].
- Van Bockhaven, J., Spíchal, L., Novák, O., Strnad, M., Asano, T., Kikuchi, S., et al. (2015). Silicon induces resistance to the brown spot fungus *Cochliobolus miyabeanus* by preventing the pathogen from hijacking the rice ethylene pathway. *New Phytol.* 206, 761–773. doi: 10.1111/nph.13270
- Verma, P. K., Verma, S., Pande, V., Mallick, S., Tripathi, R. D., Dhankher, O. P., et al. (2016). Overexpression of rice glutaredoxin OsGrx_C7 and OsGrx_C2.1 reduces intracellular arsenic accumulation and increases tolerance in *Arabidopsis thaliana*. *Front. Plant Sci.* 7:740. doi: 10.3389/fpls.2016.00740
- Wang, Y., Ohara, Y., Nakayashiki, H., Tosa, Y., and Mayama, S. (2005). Microarray analysis of the gene expression profile induced by the endophytic plant growth-promoting rhizobacteria, *Pseudomonas fluorescens* FPT9601-T5 in *Arabidopsis*. *Mol. Plant Microbe Interact.* 18, 385–396. doi: 10.1094/MPMI-18-0385
- Zandalinas, S. I., Mittler, R., Balfagón, D., Arbona, V., and Gómez-Cadenas, A. (2017). Plant adaptations to the combination of drought and high temperatures. *Physiol. Plant.* doi: 10.1111/ppl.12540 [Epub ahead of print].
- Zhao, J. L., Zhou, L. G., and Wu, J. Y. (2010). Effects of biotic and abiotic elicitors on cell growth and tanshinone accumulation in *Salvia miltiorrhiza* cell cultures. *Appl. Microbiol. Biotechnol.* 87, 137–144. doi: 10.1007/s00253-010-2443-4

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Tiwari, Prasad, Chauhan and Lata. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Endophytic *Paecilomyces formosus* LHL10 Augments *Glycine max* L. Adaptation to Ni-Contamination through Affecting Endogenous Phytohormones and Oxidative Stress

Saqib Bilal¹, Abdul L. Khan², Raheem Shahzad¹, Sajjad Asaf¹, Sang-Mo Kang¹ and In-Jung Lee^{1*}

¹ School of Applied Biosciences, Kyungpook National University, Daegu, South Korea, ² UoN Chair of Oman's Medicinal Plants and Marine Natural Products, University of Nizwa, Nizwa, Oman

OPEN ACCESS

Edited by:

Tatiana Matveeva,
Saint Petersburg State University,
Russia

Reviewed by:

Vladislav V. Yemelyanov,
Saint Petersburg State University,
Russia

Andrzej Bajguz,
University of Białystok, Poland

*Correspondence:

In-Jung Lee
ijlee@knu.ac.kr

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 27 January 2017

Accepted: 10 May 2017

Published: 29 May 2017

Citation:

Bilal S, Khan AL, Shahzad R, Asaf S, Kang S-M and Lee I-J (2017)
Endophytic *Paecilomyces formosus* LHL10 Augments *Glycine max* L. Adaptation to Ni-Contamination through Affecting Endogenous Phytohormones and Oxidative Stress.
Front. Plant Sci. 8:870.
doi: 10.3389/fpls.2017.00870

This study investigated the Ni-removal efficiency of phytohormone-producing endophytic fungi *Penicillium janthinellum*, *Paecilomyces formosus*, *Exophiala* sp., and *Preussia* sp. Among four different endophytes, *P. formosus* LHL10 was able to tolerate up to 1 mM Ni in contaminated media as compared to copper and cadmium. *P. formosus* LHL10 was further assessed for its potential to enhance the phytoremediation of *Glycine max* (soybean) in response to dose-dependent increases in soil Ni (0.5, 1.0, and 5.0 mM). Inoculation with *P. formosus* LHL10 significantly increased plant biomass and growth attributes as compared to non-inoculated control plants with or without Ni contamination. LHL10 enhanced the translocation of Ni from the root to the shoot as compared to the control. In addition, *P. formosus* LHL10 modulated the physio-chemical apparatus of soybean plants during Ni-contamination by reducing lipid peroxidation and the accumulation of linolenic acid, glutathione, peroxidase, polyphenol oxidase, catalase, and superoxide dismutase. Stress-responsive phytohormones such as abscisic acid and jasmonic acid were significantly down-regulated in fungal-inoculated soybean plants under Ni stress. LHL10 Ni-remediation potential can be attributed to its phytohormonal synthesis related genetic makeup. RT-PCR analysis showed the expression of *indole-3-acetamide hydrolase*, *aldehyde dehydrogenase* for indole-acetic acid and *geranylgeranyl-diphosphate synthase*, *ent-kaurene oxidase* (*P450-4*), *C13-oxidase* (*P450-3*) for gibberellins synthesis. In conclusion, the inoculation of *P. formosus* can significantly improve plant growth in Ni-polluted soils, and assist in improving the phytoremediation abilities of economically important crops.

Keywords: soybean, nickel stress, endophytic fungi, phytohormones, antioxidant enzymes, fatty acids

INTRODUCTION

Rapid industrialization has contributed to the increased heavy metal pollution in the environment (Qin et al., 2015; Wu et al., 2015). Effluents from industry mix with rivers and drains, which enter the food chain through crops grown in contaminated water. Consequently, by consuming contaminated foods, heavy metals accumulate in the human body and lead to severe health problems, including cardiac disorders, digestive disorders, and kidney, stomach,

liver, and lung cancer (Qin et al., 2015). Heavy metal pollutants in the soil are toxic and non-biodegradable and are thus extremely persistent and stable in the environment (Wu et al., 2015). Therefore, their removal or conversion to less toxic forms in the water-soil is deemed crucial in order to provide land that is safe and free from heavy metal pollution for agriculture purposes (Amari et al., 2014).

Serious concerns have been raised in developing countries regarding environmental pollution caused by heavy metals. Among heavy metals, trace amounts of nickel (Ni) are required by plants for the regulation of different metabolic pathways (Yusuf et al., 2011; Amari et al., 2014). However, exposure to excess amounts of Ni is toxic to plants, and induces lethal alterations in plant metabolism resulting in the inhibition of plant growth, wilting, necrosis, and chlorosis (Amari et al., 2014). Consequently, plant growth, yield, and quality are drastically affected, and the consumption of Ni-contaminated food can lead to serious health problems (Amari et al., 2014; Clemens and Ma, 2016). Therefore, bio-remediation is an efficient and inexpensive technique used to detoxify or remove Ni contamination from the soil and to provide land that is free from Ni toxicity for crop cultivation. This represents an alternative way of providing safe food for humans, thus preventing the development of lethal disorders induced by Ni toxicity (Kamran et al., 2015; Stella et al., 2017). Hence, the development of economical and reliable strategies is required for the removal or detoxification of heavy metals in order to avoid metal pollution of the environment and to provide safe food for human consumption.

In this regard, interactions between microbes and plants in soil may lead to the development of symbiosis, owing to the potential role of microorganisms in eliminating the toxic effects of metal-contaminated soil and their possible role in augmenting plant growth in metal-contaminated soil by providing nutrients and metabolites to the plant (Khan and Lee, 2013). Endophytes can play a pivotal role in the bioremediation of soils rich in heavy metals, and are considered a cost-effective and favorable replacement of conventional physical and chemical based treatments (Deng et al., 2011; Khan and Lee, 2013). Various studies have recognized that endophytic microbes can significantly enhance the host plant's potential to grow in soil polluted with heavy metal (Deng et al., 2016). The ability of plants to survive in heavy metal-contaminated soil is attributable to the positive role played by endophytic fungi in the detoxification and degradation/removal of heavy metals from soil, as well as the promotion of enhanced host plant growth due to the production of growth regulators (Deng et al., 2016). Different endophytic microbes such as *Rhizoderma veluwensis* and *Phialocephala fortinii* fungi, and *Cryptococcus* sp., *Enterobacter* sp., and *Bacillus thuringiensis* bacteria have been identified for enhancing bioremediation in metal-polluted soil (Deng et al., 2016; Ma et al., 2016). Such symbiotic interactions between plants and phytohormone-producing endophytes have drawn significant attention owing to their biotechnological potential for mitigating and degrading heavy metals from polluted media and promoting growth and yield in metal contaminated soils.

Among the endophytic microbes, fungi are considered superior to bacteria owing to their ubiquitous, multifaceted,

morphological diverse nature and higher capacity of tolerance to environmental stresses (Kaushik and Malik, 2010, 2011; Mishra and Malik, 2014). Moreover, endophytic fungi tend to produce plant growth regulators such as auxins and gibberellins (GAs), and protect plants from both biotic and abiotic stresses (Khan et al., 2015a; Deng and Cao, 2017). Similarly, they have been known to produce exopolysaccharides, proteins, extracellular enzymes, organic acids, and other metabolites that aid the removal of soil pollutants by enhancing the phytoremediation capacity of host plants (Khan et al., 2011a; Redman et al., 2011). Endophytic fungi possess appropriate metal chelation or sequestration abilities, which boost their level of tolerance to heavy metals as well as their sophisticated multicellular biomass making them suitable for use in bioremediation (Aly et al., 2011). In addition, endophytic fungi have coined for the production of phytohormones such as GA, auxin (indole acetic acid – IAA) and abscisic acid etc, in recent decade or so, owing to their potential benefits in plant stress physiology (Khan et al., 2015a). Despite these traits, the role of plant growth-promoting endophytic fungi in the context of bioremediation has not been extensively explored (Bourdel et al., 2016). Therefore, understanding the association of bioaccumulating metal from contaminated soil, phytohormone-producing endophytic fungi with plants, will not only excel metal-tolerance capabilities/phytoremediation but will also promote plant growth and yield.

Glycine max (Soybean) growth from germination through to yield is usually affected by the lethal effects of various metal/metalloids including Ni toxicity in the soil, which consequently inhibits normal plant growth by adversely targeting adsorption, translocation, and synthesis processes (Pérez-Chaca et al., 2014; Miransari, 2015; Abdel Latef et al., 2016). Soybean is one of the best agricultural crops, especially in East-Asian countries, owing to its outstanding medicinal and nutritional values (Miransari, 2015). Therefore, we aimed to screen various phytohormone-producing endophytic fungi in order to assess their potential for bioaccumulating metals in the contaminated medium, as well as their ability to alleviate heavy metal toxicity in host plants. The following fungal endophytes were selected: *Penicillium janthinellum* LK5, *Paecilomyces formosus* LHL10, *Exophiala* sp. LHL08, and *Preussia* sp. BSL10, and screened for their metal accumulation potential against different heavy metals (Table 1). Based on the production of GA and other phytohormones/secondary metabolites associated with plant defense, such as IAA, abscisic acid, and proline, the most bioactive fungal strain in terms of metal accumulation was selected for interaction with host soybean plants in metal-polluted soil in order to elucidate its metal-removing capacity and its effects on soybean plant growth and physiology.

MATERIALS AND METHODS

Endophytes Growth and Their Tolerance against Heavy Metal

Penicillium janthinellum, *P. formosus*, *Exophiala* sp., and *Preussia* sp. were isolated from the roots of tomato, cucumber, and frankincense plant leaf, respectively (Table 1). The endophytes

TABLE 1 | List of endophytic fungi with their traits used for initial screening of assessing their metal tolerance.

Endophytic fungi	Isolated host	Bioactive metabolites	Accession no.	Effect on plant growth under abiotic stress	Reference
<i>Penicillium janthinellum</i> LK5	Tomato	Gibberellins	JX111908	Growth promoting	Khan et al., 2013, 2014
<i>Paecilomyces formosus</i> LHL10	Cucumber	Auxin and gibberellins	HQ444388	Growth promoting	Khan et al., 2012a
<i>Exophiala</i> sp. LHL08	Cucumber	Gibberellins	HM017066	Growth promoting	Khan et al., 2011a
<i>Preussia</i> sp. BSL10	Frankincense	Auxin and extracellular enzymes (gibberellins not published)	KR231682	Growth promoting	Khan et al., 2016

were previously identified by extracting genomic DNA, PCR amplification of the internal transcribed spacer (ITS) region, and constructing a phylogenetic tree (Khan et al., 2011b, 2012a, 2014; 2016). The universal primers used for the identification were ITS-1 (5'-TCC GTA GGT GAA CCT GCG G-3') and ITS-4 (5'-TCC TCC GCT TAT TGA TAT GC-3') and LR0R (F) (5'-ACC CGC TGA ACT TA AGC-3') and TW13(R) (5'-GGT CCG TGT TTC AAG ACG-3').

The aforementioned strains were grown on potato dextrose agar (PDA) plates supplemented with 1 mM copper (CuSO_4), cadmium (CdSO_4), and nickel (NiSO_4), respectively. The PDA plates were then incubated at 28°C in for 7 days in darkness, after which the growth rate of endophytes was measured as described by Khan et al. (2017a). The most active metal resistant fungal strain was grown in broth and subjected for ICP-MS analysis to determine metal uptake potential. The bioactive strain was selected based on primary screening and grown in Czapek broth for 10 days at 30°C in a shaking incubator with 120 rpm for further experiments.

RNA Isolation and Reverse Transcription-PCR

In order to validate IAA and GA producing capability, the expression of IAA and GAs related genes of most bioactive fungal strain was carried out. For this purpose, the extraction of total RNA was performed from endophytic fungal according to the method described by Marcos et al. (2016). Briefly, fungal mycelia (0.1–0.2 g) were disrupted in 1 ml of TRI reagent (Sigma) with 1.5 g of zirconium beads by using a cell homogenizer. Cell debris was removed by centrifugation and supernatants were obtained with chloroform. RNA was precipitated with isopropanol and treated with DNase I (USB) prior to use in RT-PCR experiments. To synthesize cDNA, 1.0 µg of total RNA was then used by using a DiaStar™ RT Kit (SolGent, South Korea) according to manufacturer's standard protocol. The expression of IAA and GAs related genes were compared with relative expression of actin as internal control. The detail list of genes and their primers are given in Supplementary Table S1.

Endophyte and Host-Plant Interactions under Heavy Metal Stress

Seeds of soybean (Taekwang Cultivar) were surface-sterilized by washing with 70% ethanol for 30 s followed by 2.5% sodium hypochlorite for 30 min and several rinsing steps with autoclaved double distilled water. Then, seeds were incubated at 28°C for

24 h to obtain equal germination. To investigate the plant-microbe association under metal toxicity, the germinated seeds were sown in autoclaved pots containing 1 kg of horticulture soil composed of cocopeat (68%), perlite (11%), zeolite (8%), as well as micronutrients available as NH_4^+ ~0.09 mg g⁻¹; P_2O_5 ~0.35 mg g⁻¹; NO_3^- ~0.205 mg g⁻¹; and K_2O ~0.1 mg g⁻¹.

The metal-resistant bioactive strain grown in Czapek broth was applied to the plants (25 ml per pot, with 3–4 g of fungal mycelia) sown in pots as described by Khan and Lee (2013) to initiate the symbiotic association between fungus and plant. The non-inoculated plants were treated with the same amount of fungal-free Czapek broth in order to stabilize the effect of additional nutrients on plants. Thus, the plants were grown inside a growth chamber (day/night cycle: 14 h; 28°C/10 h; 24°C; relative humidity 60–70%; light intensity 1000 Em-2-s Natrium lamps) for 15 days. Then, different concentrations of Ni (0.5, 1, and 5 mM) were applied to the soybean plants. After every 24 h, 80 ml Ni solution of each respective concentration was applied to the plants for 2 weeks. To avoid the leaching of metal, each plant was irrigated before Ni treatment. The experimental design involved the following treatments: (i) control fungi free media treated plants without metal stress; (ii) endophytic fungal-treated plants without metal stress, (iii) metal-treated plants (0.5, 1, or 5 mM); (iv) fungal-inoculated plants with in addition to Ni stress (as above). The experiment was replicated three times, with 15 replicates per treatment.

Growth Attributes of Plants under Nickel Stress

After 15 days of stress, the chlorophyll content of plants subjected to each treatment was measured by employing a chlorophyll meter (SPAD-502 Minolta, Japan). Immediately after harvesting, plant shoot and root lengths, as well as fresh weight, were measured and samples were stored at -80°C for different biochemical analyses. The dry weights were measured after drying in an oven at 65°C for 72 h.

Analysis of Ni in Fungal Mycelium and Plant Roots and Shoots through ICP-MS

To analyze the root and shoot nickel contents well as in fungal mycelium, the freeze-dried samples were processed into a powder and subjected to the quantification of Ni by inductively coupled plasma atomic emission spectroscopy (ICP) (Optima 7900DV, PerkinElmer, United States). The translocation efficiency of Ni from the root-to-shoot was measured by calculating the translocation factor (TF) using the following formula: TF = total

concentration of Ni in shoots mg plant⁻¹)/total concentration of Ni in root mg plant⁻¹. Similarly, the nickel tolerance index (TI) of plants was assessed as determined by Khan et al. (2017b) using the following formula: TI % = (root length in Ni treatment/root length in control) × 100.

Lipid Peroxidation and Fatty Acid Quantification of Plants under Ni Stress

The level of lipid peroxidation was analyzed as reported by Khan et al. (2013). Briefly, plant shoots ground with liquid nitrogen were extracted with 10 mM phosphate buffer at pH 7. The reaction mixture was prepared by adding 0.2 ml 8.1% sodium dodecyl sulfate, 1.5 ml 20% acetic acid (pH 3.5), and 1.5 ml 0.81% thiobarbituric acid aqueous solution to the supernatant. The reaction mixture was heated in boiling water for 60 min, immediately after cooling to room temperature, 5 ml solution butanol: pyridine (15:1 v/v) was added. The upper layer of organic acid was removed, and the optical density of the resulting pink color was measured at 532 nm using a spectrophotometer. The level of lipid peroxidation was expressed as the amount of malondialdehyde (MDA) formed per gram tissue weight. The experiment was performed in triplicate.

The fatty acid profile (relative percentage of total fatty acid) of randomly selected plants from each treatment was determined by following the protocol reported by Khan et al. (2012b). Briefly, each plant sample (1 g) was treated with 10 ml of hexane and kept in a shaking incubator (150 rpm) at 50°C for 2 days. Centrifugation (1200 × g at 25°C) was outperformed to separate the supernatant, which was then transferred into new tubes. Hexane was evaporated from each sample by passing air through an evaporating unit. The extracted material from each sample was placed in a screw-capped vial, and 5 ml of methylation solution (H₂SO₄:methanol:toluene = 1:2:10 ml) was added. Then, the sealed vials were placed in a water bath (100°C) for 60 min for heating, followed by cooling at room temperature. Then, water (5 ml) was added and the samples were shaken. Two layers formed, and were separated by taking the upper layer and subjecting it to dryness using anhydrous sodium sulfate for 5 min. The sample (1 µl) was directly injected to the GC using an automatic sampler (Agilent 7683B). GC-MS analysis was carried out on an Agilent Model 7890A series (Agilent, Dover, DE, United States) equipped with an Agilent 5975C MS detector, an Agilent 7683 autosampler, and a MS ChemStation Agilent v. A.03.00. GC-MS was equipped with a DB-5MS capillary column (30 m × 0.25 mM i.d. × 0.25 µm film thickness; J&W Scientific-Agilent, Folsom, CA, United States) while helium was used as a carrier gas with a flow rate of 0.6 ml/min and a split mode [1:50]. The injector and detector temperature were 120 and 200°C, respectively. The column temperature was programmed from 50 to 200°C at 10°C/min and then finally held at 200°C for 5 min. The mass conditions were: ionization voltage, 70 eV; scan rate, 1.6 scan/s; mass range, 30–450; ion source temperature, 180°C. The components were identified based on the comparison of their relative retention time and mass spectra with those of standards, Wiley7N, NIST library data of the GC-MS system, and published data. The samples were

assessed for palmitic acid, stearic acid, oleic acid, and linolenic acid.

Analysis of Antioxidant and Related Enzymes in Inoculated and Non-inoculated Plants

Reduced glutathione (GSH) content was measured by Ellman (1959) protocol as described by Khan et al. (2015b) with slight modifications. Briefly, 100 mg fresh leaf tissue was ground in 3 ml of 5% (v/v) trichloroacetic acid using a chilled mortar and pestle. The homogenate was subjected to centrifugation at 12,000 rpm for 15 min at 4°C. The obtained supernatants were used to analyze the GSH contents by taking 0.1 ml of the sample supernatant and mixing with 3.0 ml 150 mM monosodium phosphate buffer (pH 7.4) and 0.5 ml of Ellman's reagent. For each reaction, the mixture was incubated at 30°C for 5 min. The absorbance was measured at 412 nm and the GSH content was calculated using a standard curve.

The activity of antioxidant enzymes such as peroxidase (POD) and polyphenol oxidase (PPO) was analyzed by the method reported by Khan et al. (2017a) with slight modifications. Briefly, leaf samples (400 mg) were ground using a chilled mortar and pestle. Then, the samples were homogenized with 0.1 M potassium phosphate buffer (pH 6.8) and centrifuged at 4°C for 15 min at 5000 rpm in a refrigerated centrifuge. The reaction mixture for the POD assay contained 0.1 M potassium phosphate buffer (pH 6.8), 50 µl pyrogallol (50 µM), 50 µl H₂O₂ (50 µM), and 100 µl of the sample crude extract. The reaction mixture was incubated for 5 min at 25°C, followed by the addition of 5% H₂SO₄ (v/v) in order to stop the enzymatic reaction. The level of purpurogallin formed was determined by the absorbance at 420 nm. For the PPO activity assay, the same reaction mixture containing the same components as that used for POD excluding H₂O₂ was used and the resulting assay was measured at 420 nm. One unit of POD or PPO was measured as a 0.1-unit increase in absorbance. Superoxide dismutase (SOD) activity was measured according to method described by Sirhindi et al. (2016), following the photo reduction of nitroblue tetrazolium (NBT). The absorbance was measured by spectrophotometer at 540 nm. SOD unit is the quantity of enzyme that hampers 50% photo reduction of NBT and is expressed as U/mg protein. Catalase (CAT) activity was assessed by the method reported by Sirhindi et al. (2016). The absorbance was measured by spectrophotometer at 240 nm and the activity was expressed as U/mg protein.

Endogenous Abscisic Acid (ABA) and Jasmonic Acid (JA) Quantification

The endogenous ABA contents were extracted following the method described by Qi et al. (1998). ABA was extracted from the plant roots and shoots by an extraction solution containing 95% isopropanol, 5% glacial acetic acid, and 20 ng of [(±)-3,5,5,7,7-d₆]-ABA. The extract was filtered and then concentrated via a rotary evaporator. The residue was dissolved in 4 ml 1 N sodium hydroxide solution, and then rinsed three times with 3 ml of methylene chloride in order to remove lipophilic materials.

After reducing the pH of the aqueous phase to approximately 3.5 by adding 6 N HCl, the sample was partitioned through solvent-solvent extraction with ethyl acetate (EtOAc) three times. EtOAc extracts were then combined and evaporated. The nearly dry residue was dissolved in phosphate buffer solution (pH 8.0), which was run through a polyvinylpolypyrrolidone (PVPP) column. The eluted phosphate buffer solution was adjusted to pH 3.5 with 6 N HCl and again partitioned three times into EtOAc. All three EtOAc extracts were combined again and evaporated through a rotary evaporator. The residue was dissolved in dichloromethane (CH₂Cl₂), and passed through a silica cartridge (Sep-Pak; Water Associates, Milford, MA, United States), which was pre-washed with 10 ml of diethyl ether: methanol (3:2, v/v) and 10 ml of dichloromethane. ABA was recovered from the cartridge by elution with 10 ml of diethyl ether (CH₃-CH₂)₂O: methanol (MeOH) (3:2, v/v). The resulting extract was dried with N₂ gas and subsequently methylated by adding diazomethane for GC-MS analysis using selected ion monitoring (SIM) 6890N network GC system, and the 5973 network mass-selective detector; Agilent Technologies, Palo Alto, CA, United States). The monitor responses to ions of *m/z* of 190 and 162 for Me-ABA, and 194 and 166 for Me-[²H₆]-ABA, were obtained using Lab-Base (ThermoQuest, Manchester, United Kingdom) data system software.

For the quantification of endogenous JA, the protocol reported by McCloud and Baldwin (1997) was followed. The freeze-dried stem and root tissues were separately ground to powder with a chilled mortar and pestle, and 0.1 g of the ground powder was suspended in a mixture of acetone and 50 mM citric acid (70:30, v/v). Internal standard [9,10-²H₂]-9,10-dihydro-JA (20 ng) was also added to the suspension. The extracts were left overnight at a low temperature to permit the highly volatile organic solvent to evaporate, retaining the less volatile fatty acids. The resulting aqueous solution was filtered, and then extracted three times with 10 mL diethyl ether. The combined extracts were then loaded onto a solid-phase extraction cartridge (500 mg of sorbent, aminopropyl), and the cartridges were washed using 7.0 mL of trichloromethane and 2-propanol (2:1, v/v). The exogenous JA and relevant standard were eluted with 10 mL of diethyl ether and acetic acid (98:2, v/v). After evaporation of the solvents, the residue was esterified with excess diazomethane, the volume was adjusted to 50 μ L with dichloromethane, and the extracts were analyzed by GC-MS (6890N network GC system and the 5973 network mass selective detector; Agilent Technologies, Palo Alto, CA, United States) in the selected ion mode. The ion fragment was monitored at *m/z* = 83 amu corresponding to the base peaks of JA and [9,10-²H₂]-9,10-dihydro-JA; the amount of endogenous JA was estimated from the peak areas compared with the respective standards. The whole experiment was performed three times.

Statistical Analysis

Experiments were independently performed in triplicate and the values obtained are presented as the means \pm standard deviation (SD). Data obtained showing the effect of Ni toxicity on the growth attributes of soybean were subjected to *t*-test using online GraphPad Prism software to determine the significant difference

among treatment means at *P* < 0.05. The biochemical analyses were analyzed with two-way analysis of variance (ANOVA) using GraphPad Prism software (version 6.01, San Diego, CA, United States).

RESULTS

In Vitro Screening of Endophytic Fungi against Heavy Metals Uptake

The results revealed that among four endophytic fungal strains, *Exophiala* sp. and *Preussia* sp. are highly sensitive to Ni, Cu, and Cd toxicity, which significantly inhibited their growth, with the exception of *Preussia* sp. showed mild growth in Ni-amended PDA medium with 120 mm² growth area (**Figure 1**). Contrarily, *P. formosus* and *P. janthinellum* had substantial tolerance to Ni toxicity followed by Cu with growth area of 424 \pm 0.53, 135 \pm 0.62, 230 \pm 0.54, and 190 \pm 0.32 mm², respectively. In terms of Cd toxicity, only *P. formosus* exhibited trivial resistance with a growth area of 120 mm². Thus, following initial investigations of tolerance, and owing to its low-to-high growth trend under Cd, Cu, and Ni toxicity, *P. formosus* was further analyzed in plant-microbe associations under Ni stress. Upon significant growth of *P. formosus* LHL10 in Ni-amended PDA medium, we grow it in PDB supplemented with Ni (**Figure 1B**) in order to evaluate its metal accumulating potential through ICP-MS quantification. Results revealed that fungal grown in PDB medium had significantly accumulated heavy metal in mycelia (**Figure 1C**). Along with heavy metal accumulation, fungal mycelium exhibited significant potential of accumulating sulfur (S) and potassium (K) contents. Since, such improved growth of fungi during stress conditions have been attributed to their potential to produce bioactive secondary metabolites and previous studies have shown this (Khan et al., 2015b), therefore, to further validate the phytohormones producing role of *P. formosus* LHL10, we have carried out RT-PCR analysis to know the existence of IAA and GA related biosynthesis pathways.

Expression of IAA and GA Biosynthesis Related Genes of *P. formosus* LHL10

The results related to the expression of IAA and GA biosynthesis genes in *P. formosus* LHL10 is shown in (**Figure 1D**). The IAA producing ability of *P. formosus* was previously quantified by HPLC (Khan et al., 2012a), which was further validated in the current study by the real time RT-PCR analysis. The results confirmed that the LHL10 showed the ability to produce both IAA and GA. In case of IAA, we observed the expression of its biosynthesis related genes, i.e., *aldehyde dehydrogenase (ALD)* and *indole-3-acetamide hydrolase (IAAH)*. Similarly, in case of GA biosynthesis *geranylgeranyl-diphosphate synthase (GGS2)*, *ent-kaurene oxidase (P450-4)*, *GA20-oxidase*, and *C13-oxidase* expressions were found and the results exhibited induced transcript accumulation in RT-PCR, recommending the existence of GA genes cluster in *P. formosus*. However, the expression of (*GGS2*) and (*P450-4*) was comparatively higher than *GA20-oxidase* and *C13-oxidase*.

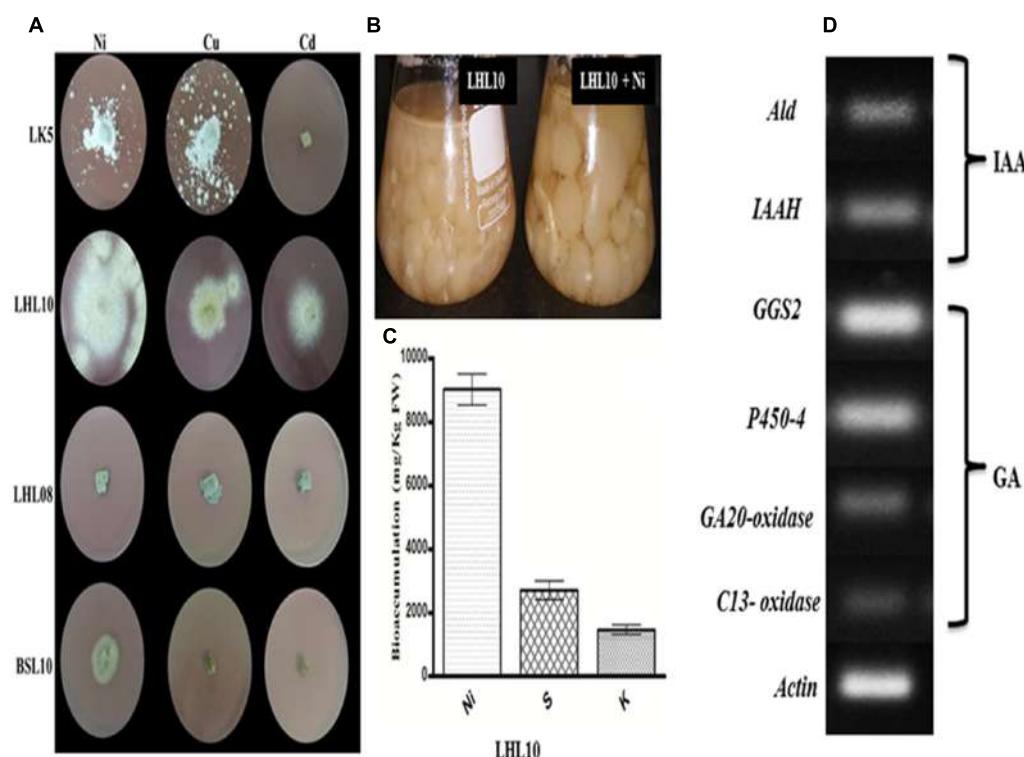


FIGURE 1 | Growth dynamics of selected fungal endophytes in 1.0 mM Ni, Cu, Cd amended media. (A) Each fungus was screened against metals toxicity and their ability to survive. The diameter of growing mycelia was measured after 7 days of incubation at 25°C. **(B)** On the base of significant growth in Ni stress, LHL10 was grown in PDB medium for metal and nutrient uptake analysis. **(C)** Metal uptake analysis via ICP-MS. **(D)** Genes expression related to gibberellins (GAs) and indole acetic acid (IAA) biosynthesis of *P. formosus*. LHL10. *Indole-3-acetamide hydrolase* (*IAAH*), *aldehyde dehydrogenase* (*Ald*) are IAA related genes, while GAs related genes are *geranylgeranyl-diphosphate synthase* (*GGS2*), *ent-kaurene oxidase* (*P450-4*), GA 20-oxidase and *C13-oxidase* (*P450-3*). The **(A)** is representative of five replications whereas **(B)** and **(C)** are repeated three times.

Association of *P. formosus* LHL10 under Ni Stress Enhance Plant Growth and Chlorophyll Content

The results of the current study clearly demonstrated a significant growth effect ($p < 0.05$) for soybean plants inoculated with endophytic fungi under different levels of Ni stress. The association of soybean plants with *P. formosus* significantly enhanced growth by stimulating an increase in the length and fresh/dry weight of shoot and root well as chlorophyll content in comparison to non-inoculated plants (Figure 2 and Table 2). Soybean plants subjected to Ni stress exhibited considerable retardations in shoot, root length, and fresh and dry weight, which adversely affected endophytic fungal-free plants. Furthermore, the interaction between fungus and plant mitigated Ni toxicity and promoted enhanced shoot and root lengths, and fresh and dry weights in plants grown with low, medium, and high concentrations of Ni compared with those of non-endophytic-infected plants. With heavy metal application, shoot and root lengths, and fresh and dry weights were significantly ($p < 0.05$) higher (ranging from 11 to 47%) in endophytic fungal-treated plants compared to control plants under varying concentrations of Ni stress (Table 2). The chlorophyll contents in both endophytic fungal-inoculated and

non-inoculated plants under different concentrations of soil Ni were screened. *P. formosus* plants exhibited an approximate 10% increase in chlorophyll levels compared with the non-inoculated plants in the absence of Ni toxicity. A marked reduction in the chlorophyll content was detected as the level of Ni toxicity in the soil increased. Non-inoculated plants displayed reductions of 13, 26, and 46% under low, medium, and high levels of Ni toxicity, respectively, in the soil compared with inoculated plants (Table 2).

Influence of *P. formosus* Inoculation on Metal Uptake and Distribution in Soybean Plant

Determination of metal accumulation is important for delineating the role of plants in remediation. Ni was not found to accumulate in the roots or shoots of endophytic fungal-treated and non-treated plants grown in the absence of metal. Non-inoculated plants grown in the soil containing Ni application, accumulated a substantial amount ($p < 0.05$) of Ni in their roots followed by their shoots (Figure 3 and Table 3). However, soybean plants inoculated with *P. formosus* presented lower Ni uptake as compared to inoculated plants, which was 483.66 to 694.66 mg/kg in roots, followed by 284 to



FIGURE 2 | Effect of endophytic fungus on plant growth under Ni stress. Each plant represents 15 replicates of eight treatments.

TABLE 2 | Growth promoting effect of *Paecilomyces formosus* LHL10 in soybean under various nickel stress.

Treatment	SL (cm)	RL (cm)	SFW (g)	SDW (g)	RFW (g)	RDW (g)	CC (SPAD)	TF	TI%
Control	24.46 ± 0.38 ^b	13.27 ± 0.58 ^b	14.71 ± 0.61 ^b	5.47 ± 0.62 ^b	6.19 ± 0.52 ^b	2.16 ± 0.22 ^b	31.74 ± 1.52 ^b	—	—
Fungus	26.59 ± 0.86 ^a	16.63 ± 0.46 ^a	16.74 ± 0.73 ^a	6.22 ± 0.49 ^a	7.25 ± 0.61 ^a	2.87 ± 0.35 ^a	34.91 ± 1.12 ^a	—	108.67
0.5 mM Ni	20.44 ± 0.25 ^b	12.77 ± 0.63 ^b	12.99 ± 0.65 ^b	3.97 ± 0.38 ^b	5.25 ± 0.68 ^b	1.71 ± 0.27 ^b	24.37 ± 1.14 ^b	0.58	83.55
Fungus + 0.5 mM Ni	23.13 ± 0.53 ^a	14.81 ± 0.61 ^a	17.18 ± 0.78 ^a	4.82 ± 0.51 ^a	8.92 ± 0.45 ^a	2.69 ± 0.41 ^a	29.77 ± 1.35 ^a	0.62	94.56
1 mM Ni	19.43 ± 0.84 ^b	9.15 ± 0.89 ^b	11.42 ± 0.47 ^b	3.55 ± 0.32 ^b	4.75 ± 0.56 ^b	1.57 ± 0.43 ^b	20.54 ± 1.39 ^b	0.61	79.41
Fungus + 1 mM Ni	22.01 ± 0.81 ^a	13.17 ± 0.63 ^a	15.46 ± 0.01 ^a	4.50 ± 0.62 ^a	8.74 ± 0.46 ^a	2.54 ± 0.49 ^a	27.12 ± 0.76 ^a	0.52	89.98
5 mM Ni	16.4 ± 0.89 ^b	6.21 ± 0.81 ^b	10.51 ± 0.76 ^b	2.98 ± 0.43 ^b	3.88 ± 0.52 ^b	1.15 ± 0.37 ^b	12.52 ± 0.79 ^b	0.56	67.02
Fungus + 5 mM Ni	19.04 ± 0.74 ^a	11.85 ± 1.32 ^a	14.25 ± 0.98 ^a	3.94 ± 0.49 ^a	6.98 ± 0.60 ^a	2.07 ± 0.15 ^a	22.56 ± 1.14 ^a	0.53	77.84

SL, shoot length; RL, root length; SFW, seedlings fresh weight; SDW, shoot dry weight; CC, chlorophyll content; TF, translocation factor; TI, percentage of tolerance index. Each value represents mean ± SD of 15 replicates from three independent experiments. Values in the columns followed by different letters are significantly different at $p \leq 0.05$ based on t-test.

395 mg/kg in shoots under lower and higher stress, respectively. The results revealed that non-inoculated plants accumulated a 30–42 and 35–35% higher amount of Ni in their shoots and roots, respectively, under varied Ni concentrations. The results indicated that Ni accumulated in the order of root > shoot.

In order to further validate the Ni-detoxifying role of endophytic fungi, nMDS ordination was conducted on the Ni-uptake data for inoculated and non-inoculated plants (Figure 3). Control and only fungal-infected plants were found to have the same cluster in the absence of Ni stress. However, as the level of Ni stress increased, the clusters of fungal-treated and non-treated plants became distinct, as shown in Figure 3. At 1 mM Ni toxicity, the fungal-treated and non-treated plants did not form a cluster, showing a significant effect of fungal treatment. Both 1 NiF and 5 NiF exhibited a strong correlation by showing a similar distance from each other as compared to other treatments (Figure 3). Translocation factor was determined to evaluate the Ni translocation efficiency of soybean plants from root to shoot. TF extended from 0.528 to 0.622 and showed

different orders of effect, regardless of metal application and fungal inoculation in soybean plants (Table 2). Similarly, there was a decrease in Ni tolerance capability as metal concentration increased. *P. formosus*-inoculated plants had a significantly high TI%, which was around 11.01, 10.57, and 10.82% higher in their respective metal stress (0.5, 1.0, and 5.0 mM) than those of non-inoculated plants (Table 2). These results clearly indicate that the interaction of endophytic *P. formosus* with soybean alleviates stress induced by Ni toxicity.

Association of *P. formosus* Reduces Lipid Peroxidation and Regulates Fatty Acids Composition

Malondialdehyde content, which is an indicator of lipid peroxidation, was evaluated using a spectrophotometer. Normally, metal-induced stress alters the MDA content, which induces lipid peroxidation. Likewise, in the current experiment, the MDA contents in soybean leaves increased as the level

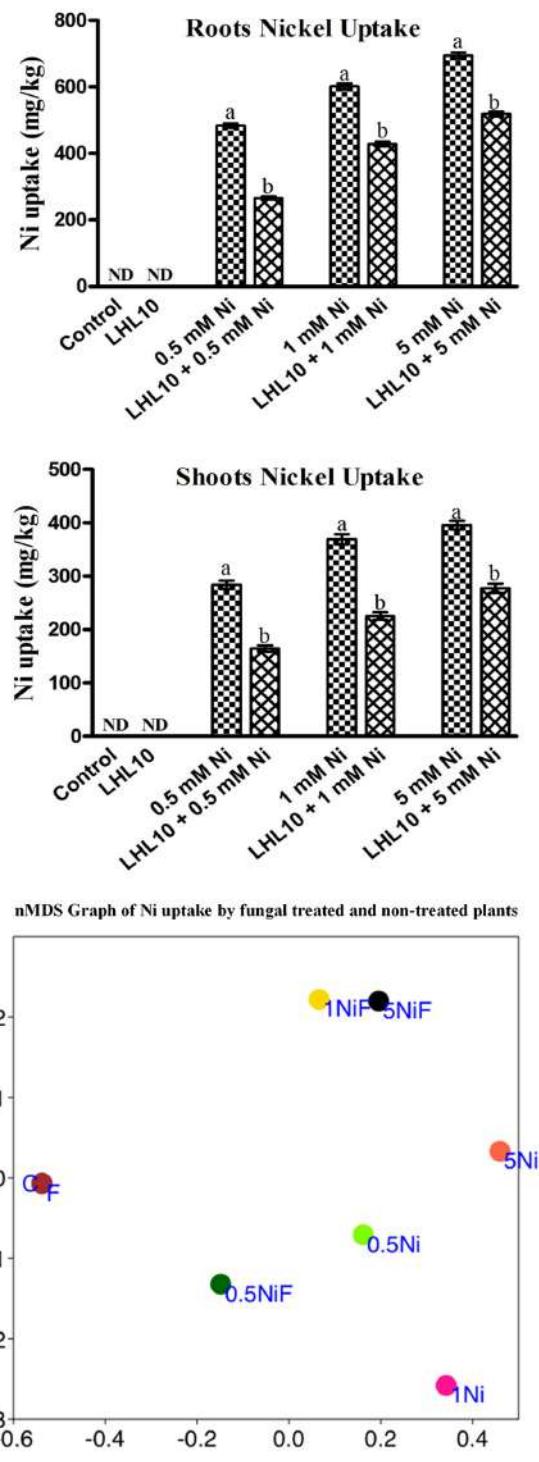


FIGURE 3 | Ni accumulation in roots and shoots of soybean plants inoculated with and without the fungal endophyte LHL10. Each value represents the mean \pm standard error. Different letters on the bars represent significant difference between the treatments (Tukey's HSD, $p < 0.05$). A non-metric multidimensional scaling plot portrays Ni uptake by fungal-inoculated plants and non-inoculated plants (C, control; F, only fungal treated; 0.5, 1.0, 5.0 mM; Ni, Nickel-treated plants; NiF, fungal- and nickel-treated plants).

of Ni toxicity increased both in endophytic fungal-treated and non-treated plants. However, the MDA content was significantly reduced in plants inoculated with *P. formosus* ($p < 0.05$) under all Ni-stress treatments compared with the non-inoculated plants (Figure 4). Under normal conditions, both control and inoculated plants exhibited similar levels of MDA. These results showed that *P. formosus* treatment inhibited the lipid peroxidation process, and consequently, enhanced plant tolerance against Ni toxicity by protecting cell membranes from metal attack. To rectify the reduction in MDA content induced by *P. formosus* association, alteration in essential fatty acids composition under different levels of Ni toxicity was assessed (Table 4). The results revealed that Ni toxicity at varying concentrations affects the analyzed fatty acids, including palmitic, stearic, oleic, and linolenic acids. Under the control condition, fungal-inoculated plants had statistically ($p < 0.05$) higher percentages of palmitic acid and stearic acid as compared to non-inoculated plants. Levels of linolenic acid were not statistically significant between *P. formosus*-treated and non-treated plants in the control condition. In contrast, inoculated plants showed a significantly ($p < 0.05$) lower level of palmitic and oleic acid under Ni toxicity as compared to non-inoculated plants. The level of stearic acid was not statistically different between fungal-inoculated and as non-inoculated plants under Ni stress. In the case of linolenic acid, *P. formosus*-infected plants displayed a significantly marked accumulation, with increases of 43, 44, and 66%, respectively, compared with non-treated plants under the respective concentrations of Ni stress.

Regulation of Soybean Antioxidant System by *P. formosus* in Response to Ni Exposure

Nickel toxicity in plants triggers oxidative stress through the excessive production of reactive oxygen species (ROS), thus leading to lethal irreversible damage to plant growth and functionality. Therefore, the antioxidant system of soybean plants, including the action of enzymatic antioxidants, was evaluated through spectrophotometry, as they can scavenge ROS efficiently and mitigate their adverse effects. Likewise, in the current experiment, the GSH contents in soybean leaves were found to increase as the Ni toxicity level increased both in endophytic fungal-treated and non-treated plants (Figure 4). However, plants inoculated with *P. formosus* had significantly ($p < 0.05$) increased levels of GSH contents under all Ni-stressed treatments compared with non-inoculated plants (Figure 4 and Table 3). Under normal conditions, inoculated plants exhibited slightly increased levels of GSH compared with non-inoculated plants. ROS-induced stress in plant tissues generated by Ni toxicity is mitigated by the synthesis and regulation of antioxidants and related enzymes. Reduced GSH production was significantly boosted in the shoots of endophytic fungal-inoculated plants when exposed to varying concentrations of Ni toxicity, as well as under normal conditions. In terms of the regulation of antioxidant enzymes by endophytic fungal, POD activity was similar in both inoculated and non-inoculated plants in the absence of Ni stress (Figure 4). Fungal-infected plant

TABLE 3 | Two-way ANOVA table of the biochemical analysis performed for inoculated and non-inoculated soybean under different concentration of Ni toxicity.

Analysis	Non-inoculated and LHL 10			Ni stress			Interaction		
	MS	P-value	variance %	MS	P-value	variance %	MS	P-value	variance %
ABA shoots	3520	0.0001	8.62	115500	0.0001	84.83	8750	0.0001	6.43
ABA Roots	58130	0.0001	9.28	171900	0.0001	82.34	17310	0.0001	8.29
JA Roots	30110	0.0001	12.62	45150	0.0001	56.78	24030	0.0001	30.23
JA Shoots	32510	0.0001	18.90	38950	0.0001	67.93	7449	0.0001	12.99
Root Ni uptake	120600	0.0001	8.33	427900	0.0001	88.71	14030	0.0001	2.91
Shoot Ni uptake	58130	0.0001	11.37	135700	0.0001	84.56	6284	0.0001	3.92
LPO	6.848	0.0001	24.78	5.516	0.0001	59.87	0.5414	0.0472	5.88
GSH	4.815	0.0001	45.81	1.149	0.0001	32.80	0.6473	0.0001	18.48
POD	7083	0.0001	17.11	9973	0.0001	72.26	1399	0.0001	10.14
PPO	41.11	0.0001	80.41	0.05457	0.8841	0.32	1.934	0.0022	11.32

Two-way ANOVA was followed by a Bonferroni post hoc test with a $p \leq 0.05$ for Ni stress level, LHL10-inoculated and non-inoculated plants, and their interaction, through GraphPad Prism software (version 6.01, San Diego, CA, United States).

exhibited significantly higher (13–36%) POD activity than non-fungal-treated plants under Ni-induced stress. PPO activity in inoculated plants was increased significantly under Ni-toxicity, with 27 to 44% greater activity than that in non-fungal-treated plants (Figure 4). Under the control condition, endophytic fungal-treated plants exhibited 19.9% higher PPO activity than non-inoculated plants. In case of CAT activity, fungal-inoculated plant displayed significant stimulated activity both under control and stress condition. A significant enhancement in CAT activity of inoculated plants was detected at 5.0 mM Ni stress level, which was approximately 50% higher than non-inoculated plants. In terms of SOD activity, inoculated as well as non-inoculated plants showed statistically non-significant activity both under control and 0.5 mM Ni stress. However, significant enhancement ($p < 0.05$) in SOD activity of fungal inoculated plants than non-inoculated plants was observed at 1.0 and 5.0 mM Ni stress levels. The SOD activity in non-inoculated plants went to its maximum at 1 mM Ni stress, but the decreased was observed at highest Ni stress level.

Regulation of Plant ABA and JA by *P. formosus* under Ni Stress

Abscisic acid and jasmonic acid are plant phytohormones known for their responsiveness to endogenous stress behavior and enhanced expressed under abiotic stress conditions. In the current experiment, *P. formosus* symbiosis with plants significantly increased the ABA level ($p < 0.05$) in shoots as compared to control plants in the absence of Ni stress (Figure 5 and Table 3). However, upon exposure to Ni stress, endophytic fungal-treated plants exhibited significantly reduced content (32–36%) of ABA compared with non-fungal-infected plants. Likewise, ABA content in the roots was promoted in *P. formosus*-treated plants compared with non-inoculated plants under normal growth conditions. When exposed to Ni stress, soybean roots appeared to be more sensitive and produced higher levels of ABA both in fungal-treated and non-treated plants (Figure 5). However, fungal-inoculated plants had significantly

reduced levels of ABA in the root during Ni stress compared with non-inoculated plants. These results indicate that interaction with *P. formosus* helped to mitigate the toxic effect of Ni in plants. The JA content in plants followed the same trend as the ABA content. The JA content was significantly ($p < 0.05$) increased in response to Ni stress both in endophyte-treated and non-treated plants. During Ni stress, non-fungal-treated plants exhibited a higher concentration of JA in the roots and shoots, which was increased by 44, 54, and 51%, and 49, 22, and 23%, respectively, compared to fungal-treated plants at concentrations of 0.5, 1.0, and 5.0 mM, Ni, respectively. The down-regulation of endogenous hormones in soybean clearly shows that the interaction with *P. formosus* was beneficial and aided soybean to ameliorate stress induced by heavy metals by triggering their defense system.

DISCUSSION

In the current study, we assessed the remediation of heavy metals by different endophytic fungal strains (*Penicillium janthinellum* LK5, *P. formosus* LHL10, *Exophiala* sp. LHL08, and *Preussia* sp. BSL10) in Ni-, Cu-, and Cd-contaminated media. The initial investigation suggested that *P. formosus* LHL10 is the most effective strain owing to its higher growth rate in all media contaminated with metal as compared to other strains. *P. formosus* showed higher tolerance to Ni pollution than Cu and Cd. Growth of *P. formosus* in liquid broth indicated that the fungal strain was very effective in Ni removal. The remarkable Ni uptake by *P. formosus* mycelium demonstrates its metal bio-remediating ability which could be attributed either to intercellular absorption or extracellular adsorption (Li et al., 2017). Generally, surface adsorption via ion exchange, extracellular precipitation, hydrolytic adsorption, metal transformation, and sequestration are the favorable biological approaches, adopted by fungi for survival against metal toxicity (Fomina and Gadd, 2014). The remediating process by *P. formosus* is further facilitated with the extracellular

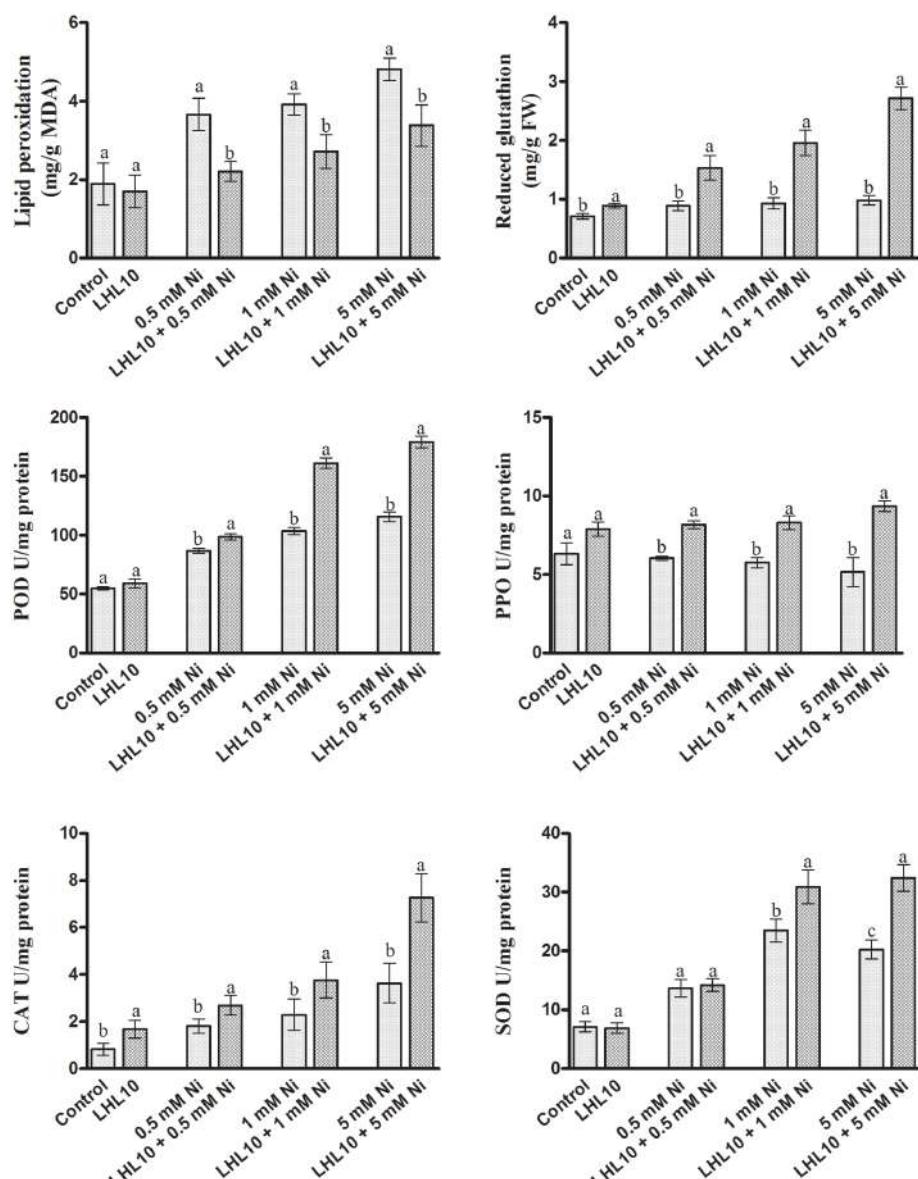


FIGURE 4 | Effect of LHL10 inoculation on MDA, GSH concentration and POD, PPO, SOD, and CAT activity in soybean shoots under different concentrations of Ni. Each value represents the mean \pm standard error. Different letters on the bars represent significant differences between the treatments (Tukey's HSD, $p < 0.05$).

production of chemically active substances, in current case phytohormones. Therefore in current study, both inter and extracellular mechanisms seem to be implied simultaneously by *P. formosus* for metal tolerance. Inter-cellular presence of Ni accumulation by *P. formosus* through ICP-MS proves the deposition in vacuole. Whereas, extracellular production of phytohormones (GA and IAA) and existence of biosynthetic genes further re-affirm the metal stress resistance ability. This was also previously validated by the reports of Rajkumar et al. (2012).

The current results showed that *P. formosus* LHL10 showed the existence of GA and IAA biosynthesis related genetic expression. GAs are synthesized from converting mevalonic

acid into geranylgeranyl diphosphate via hydroxymethylglutaryl coenzyme A, farnesyl diphosphate and geranylgeranyl diphosphate (GGDP). We observed the expression of *GGS2* was detected in RT-PCR, which encodes GGDP synthase principally accountable for providing GGDP for GA production. Similarly, *ent-kaurene oxidase* (*P450-4*) expression indicates that *P. formosus* has cluster of P450 monooxygenase-encoding genes which play crucial role in GAs biosynthesis. *GA20-oxidase* expression was also analyzed in *P. formosus* in the current study and the result showed expression in RT-PCR. *GA20-oxidase* has been reported for in the synthesis of active GAs in *F. fujikuroi* (Tudzynski et al., 2001; Bömké and Tudzynski, 2009). *C13-oxidase* (*P450-3*) was also

TABLE 4 | Regulation of fatty acids by *Paecilomyces formosus* LHL10 in soybean under Ni- stress.

Treatment	Palmitic	Stearic	Oleic	Linolenic
Control	15.74 ± 0.29 ^b	10.76 ± 0.21 ^b	40.5 ± 1.91 ^a	33.0 ± 2.13 ^a
LHL10	17.26 ± 0.35 ^a	11.96 ± 0.36 ^a	36.02 ± 1.85 ^b	34.7 ± 2.53 ^a
0.5 mM Ni	20.64 ± 0.58 ^a	15.02 ± 1.35 ^a	43.41 ± 1.57 ^a	20.93 ± 1.83 ^c
LHL10 + 0.5 mM Ni	18.67 ± 0.34 ^b	13.4 ± 1.72 ^a	31.7 ± 1.79 ^c	36.23 ± 2.46 ^a
1 mM Ni	24.07 ± 0.53 ^a	14.9 ± 0.56 ^a	43.02 ± 1.58 ^a	18.01 ± 2.17 ^c
LHL10 + 1 mM Ni	21.34 ± 0.72 ^b	13.54 ± 1.67 ^a	33.02 ± 1.43 ^c	32.1 ± 2.23 ^a
5 mM Ni	25.68 ± 0.61 ^a	17.95 ± 1.04 ^a	47.36 ± 2.05 ^a	9.01 ± 1.41 ^c
LHL10 + 5 mM Ni	22.87 ± 0.58 ^b	15.38 ± 1.32 ^a	35.15 ± 2.35 ^b	26.6 ± 2.45 ^a

Each value represents the mean ± SD of 11 replicates from three independent experiments. Values in the columns followed by different letters are significantly different at $p \leq 0.05$ based on the t-test.

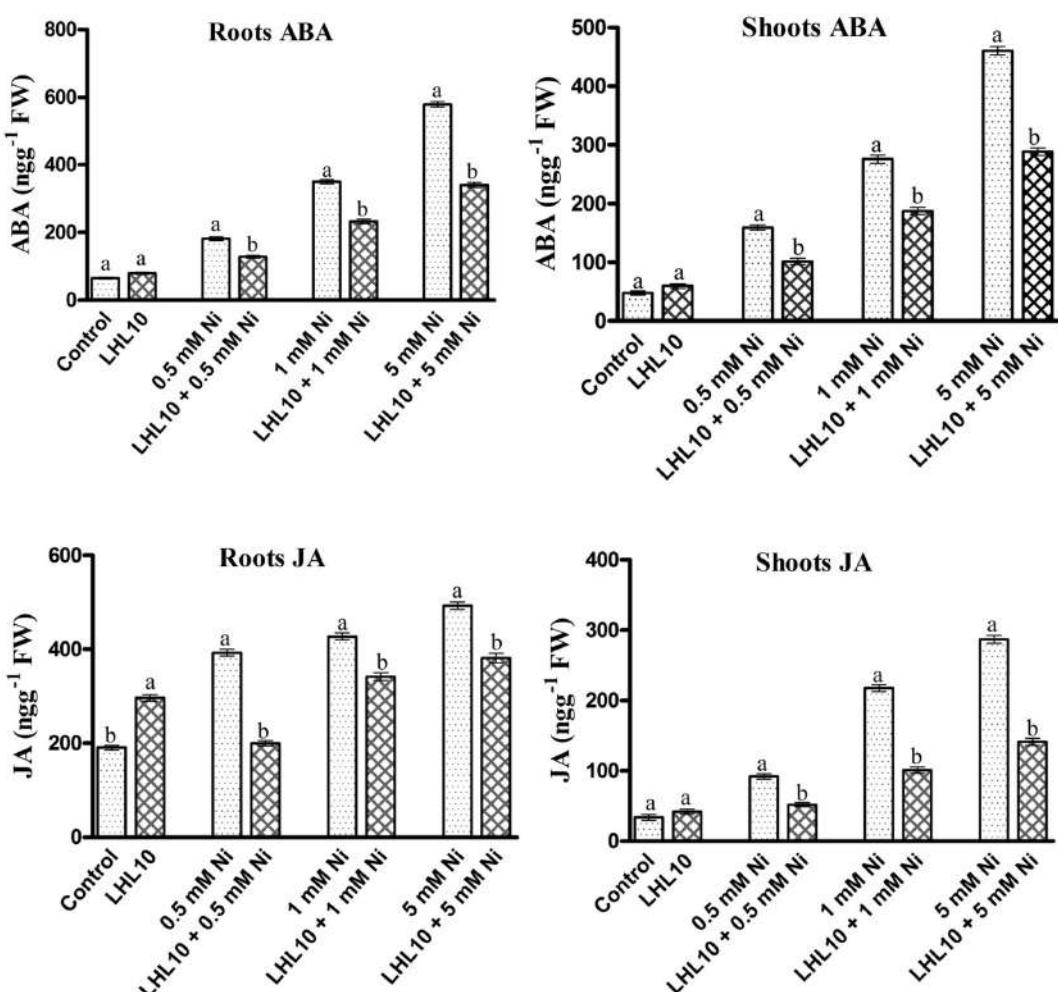


FIGURE 5 | Regulation of endogenous abscisic acid (ABA) and jasmonic acid (JA) in inoculated and non-inoculated soybean roots and shoots under different levels of Ni stress. Each value represents the mean ± standard error. Different letters on the bars represent significant differences between treatments (Tukey's HSD, $p < 0.05$).

analyzed in *P. formosus* and was comparatively less expressed. *C13-oxidase* (*P450-3*) has the role to catalyze C₁₃-hydroxylation of GA₄ to produce the minor GA1 (Böhmke and Tudzynski, 2009). The presence of aforementioned genes involved in the GAs

biosynthesis pathway validated the GA producing capability of endophytic *P. formosus* LHL10. Previously, same GA gene cluster has also been reported in *Fusarium fujikuroi*, *Gibberella fujikuroi*, and *Fusarium proliferatum* (Tudzynski et al., 2001; Albermann

et al., 2013; Rim et al., 2013). However, there is no information on GA gene biosynthetic pathway in *Paecilomyces* sp. Current report has elucidated it for the first time.

The IAA producing ability of *P. formosus* was rectified via RT-PCR analysis. The IAA biosynthesis related genes expression, i.e., *ALD* and *IAAH* exhibited transcript expressions. The *ALD* genes are reported to be required in tryptophan synthesis, which is the precursor of IAA synthesis as well as involved in stress response (Henke et al., 2016). Similarly the expression of *IAAH* further authenticates the IAA producing ability of endophytic *P. formosus*. As *IAAH* encodes indole-3-acetamide hydrolase that constitute indole-3-acetamide pathway, which leads to the synthesis of IAA through converting tryptophan to indole-3-acetamide, encoded by *IaaM* tryptophan-2-monooxygenase. Indole-3-acetamide is further converted into IAA by the help of *IAM* hydrolase which is encoded by *IaaH* (Tsavkelova et al., 2012). Previously, *Tricholoma vaccinum*, *Fusarium verticillioides*, *Fusarium proliferatum*, and *Saccharomyces cerevisiae* have been reported for possessing IAA biosynthetic pathway (Rao et al., 2010; Tsavkelova et al., 2012; Krause et al., 2015).

The results of this study revealed that the endophytic fungal *P. formosus* has the potential to remediate metal toxicity in metal-polluted environments, and can thus be deployed for mitigating heavy-metal stress in plants in order to improve growth and yield and to clean up the metal-contaminated sites. The production of various biochemical such as GA and IAA by *P. formosus* may contribute to the bioremediation ability of endophytic fungi (Khan et al., 2015a; Saleem et al., 2015). Fungal endophytes that produce IAA have been repeatedly recognized for their ability to counteract the toxic effects of heavy metals (Visioli et al., 2015). Fungal endophytes with metal-resistant capabilities employ different metabolic pathways as well as physiological and molecular strategies for effective intracellular and extracellular detoxification of heavy metals (Deng et al., 2016).

The interaction between endophytes and plants in metal-contaminated soil allows them to grow well and mitigates the toxic effect of the metal (Babu et al., 2014; Sura-de Jong et al., 2015). Fungal endophytes such as *Aspergillus*, *Penicillium*, *Fusarium*, *Paecilomyces*, *Cladosporium*, *Lasiodiplodia*, *Glomerella*, and *Phomopsis* have been recognized for their growth in metal-contaminated soil and their ability to counteract stress induced by heavy metals, such as Cd, Al, Zn, Pb, and Cu (Deng et al., 2014, 2016; Khan et al., 2016). To date, very little information is available regarding endophytic fungi, particularly *P. formosus*, and their symbiotic role in crop plants, especially soybean, and whether they provide any protection or extend tolerance to soybean under Ni toxicity. In this study, we observed severe toxicity induced by Ni contamination in soybean, which significantly retarded plant growth attributes. Our findings showing effects of Ni on soybean growth parameters were in concordance with previous studies by Sirhindi et al. (2016) in *Glycine max*, Siddiqui et al. (2011) in *T. aestivum*, and Rehman et al. (2016) in maize. Conversely, the roots, shoot dry and fresh weight of soybean were remarkably enhanced following inoculation with *P. formosus*. Improvements in biomass and other growth parameters of *Solanum nigrum* were observed following inoculation with *P. lilacinus* under Cd and

Pb toxicity (Gao et al., 2010, 2012). The direct vulnerability of non-inoculated soybean roots to Ni stress resulted in poor root growth, probably due to the inhibition of mitosis, which adversely affects the growth parameters of the whole plant (Sirhindi et al., 2016).

Interaction of the plant with *P. formosus* mitigated Ni toxicity and favored plant growth owing to the secretion of plant growth regulatory metabolites such as IAA, and GA, which help to counteract stress induced by metal toxicity. Previous studies have revealed roles for GA and IAA in the modulation and adaptation of plants under severe heavy metal stress (Bashri and Prasad, 2016; Liu et al., 2016). Furthermore, fungal endophytes possess different metal degradation pathways, as well as chelation or sequestration systems, and therefore increase the tolerance of the host plant to heavy metals and assist their growth in metal-polluted soils (Aly et al., 2011; Card et al., 2015). Increases in Ni concentration markedly affected the pigment system of soybean in the current study and resulted in reduced chlorophyll contents. Inhibition of chlorophyll contents by Ni toxicity in tomato and cotton has been reported (Khaliq et al., 2016; Mosa et al., 2016). Decrease in the chlorophyll content can be attributed to the inhibition of proto chlorophyllide reductase and ALA dehydratase enzymes following Ni contamination, which play crucial roles in chlorophyll biosynthesis (Noriega et al., 2007; Gill et al., 2015). In our study, *P. formosus*-inoculated plants exhibited significant increases in chlorophyll content as compared to non-inoculated plants under different concentrations of Ni. Enhanced growth attributes and chlorophyll content in inoculated plants subjected to Ni stress may be due to the reduced levels of MDA and H₂O₂ following inoculation with *P. formosus*. Our results are in accordance with those of previous studies (Babu et al., 2014; Padash et al., 2016), which found increased chlorophyll contents in maize and lettuce under Pb and Zn stress following inoculation with endophytic *Trichoderma virens* and *Piriformospora indica* strains, respectively.

To further assess the beneficial effect of *P. formosus* inoculation, it is important to determine the distribution and accumulation of metal uptake in the roots and shoots when evaluating the survival of plants subjected to metal stress (Gao et al., 2015). Here, we found mitigating effects in inoculated soybean plants in terms of a lower accumulation in shoots followed by roots compared with non-inoculated plants. As roots are the primary site of the plant in contact with the soil, they may accumulate more toxic metals (Czerpak et al., 2006). The increased accumulation of nickel in roots might be due to its compartmentalization in root vacuoles (Sharma et al., 2016). The decreased metal accumulation observed in fungal-treated plants further strengthens the finding that *P. formosus* has a higher capability of Ni sorption or degradation during interaction with soybean plants. Furthermore, inoculation or plants with growth-promoting fungal endophytes that are able to produce IAA, organic acids, and siderophores minimized the phytotoxic effects of heavy metals and enhanced metal uptake (Deng et al., 2016). Thus the lower accumulation of Ni indicates that the interaction between the endophytic fungus *P. formosus* and soybean under varying levels of Ni toxicity alleviates the adverse impact of metal-induced stress.

In addition, the present study showed that excessive Ni toxicity alters the activity of metabolic enzymes and indirectly causes oxidative stress. Lipid membrane integrity and activities related to membrane-associated enzymes are markedly affected by heavy metals (Kamran et al., 2015; Sirhindi et al., 2015). Plants tightly regulate their GSH network, which is considered to have a key role in reducing cellular damages (Jozefczak et al., 2012). Disruption of cellular membranes can be ascribed to the production of superoxide radicals $O_2^{\bullet-}$ in response to Ni stress. Abundant MDA is generated as a product of polyunsaturated fatty acids during peroxidation of lipid membranes. We found that Ni stress boosted MDA contents in the control plants, which was indicative of serious damage to lipid membranes in response to elevated Ni stress. As reported by Llamas and Sanz (2008) and Gajewska et al. (2012), increased Ni toxicity induces structural damage in rice and wheat lipid membranes, which severely alter their properties and enhance K^+ efflux from cells. However, MDA accumulation was successfully minimized in *P. formosus*-treated soybean under Ni stress, providing strong evidence that fungal inoculation has the ability to mitigate oxidative injuries.

Our results are in line with previous findings (Khan and Lee, 2013; Khan et al., 2015b) who found the protective effect of the endophytic fungi *P. funiculosum* and *Penicillium janthinellum* on lipid bilayers breakdown under Cu and Al toxicity in tomato plants, respectively. However, it might be assumed from the results of our study that the secretion of various endogenous phytohormones might directly or indirectly enhance plant tolerance through the low generation of ROS. Ni toxicity significantly altered the fatty acid composition of non-treated plants by enhancing the saturation and breakdown of fatty acids, with the exception of linolenic acid. Polyunsaturated fatty acids are extremely sensitive to peroxidation; therefore, their decreased accumulation indicates a low level of Ni-induced oxidative reactions in cellular membranes. Our results are in accordance with those of Gajewska et al. (2012) and Kim et al. (2014) who reported high accumulations of palmitic acid, stearic acid, and oleic acid, and lower accumulation of linolenic in wheat and rice under Ni and Cd stress. Whilst in *P. formosus* treated plants, high levels of linolenic acid were detected in response to high nickel toxicity. From these findings, we can conclude that inoculation with *P. formosus* may lead to the activation of phospholipases as well as phospholipid-derived molecules, which are thought to be involved in plant defense signaling. Phospholipase activation causes the release of alpha-linolenic acid from membrane lipids, which acts as a precursor of JA, a signaling molecule that enhances plant defense systems (Dar et al., 2015; Hou et al., 2016).

Subsequently, our result showed that Ni stress seriously impairs antioxidant components in soybean, such as POD, PPO, CAT, and SOD as well as the low-molecular weight non-protein antioxidant GSH. However, inoculation with *P. formosus* significantly induces antioxidant activities, including POD, PPO, GSH, CAT, and SOD with increasing levels of Ni toxicity. This implies the beneficial interaction with endophytic fungal enables plants to cope with oxidative stress induced by Ni toxicity. The findings of the present study are consistent with those reported by Degola et al. (2015) and Wang et al. (2016), who found

enhanced antioxidant activities in maize and *Nicotiana tabacum* plants inoculated with endophytic fungi under heavy-metal stress. GSH is an important antioxidant owing to its reducing potential, which can respond directly as a free radical scavenger. GSH is also a precursor for the synthesis of metal chelating phytochelatins. Hence, its accumulation during conditions of oxidative stress can help protect plant macromolecules, such as lipids, proteins, and DNA either by acting as an electron donor to scavenge ROS, or as organic free radicals, or by the formation of direct adducts with reactive electrophiles (Asada et al., 1994; Sharma et al., 2012). As a precursor of phytochelatins, the higher production of GSH in fungal-inoculated plants might have resulted in their increased biosynthesis. Higher production of phytochelatins and homophytochelatins are considered crucial for detoxification and homeostasis of heavy metals and metalloid toxicity in soybean plants (Vázquez et al., 2009). Therefore, in our study, the enhanced accumulation of GSH in *P. formosus*-treated plants may signify a protective role of endophytic fungal in nickel-induced stress. Similarly, increasing POD activity was observed with increased Ni concentration in inoculated plants, while non-treated plants showed decreased POD activity under nickel toxicity. As reported, metal toxicity leads to the suppression of POD (Hossain and Komatsu, 2013; Xie et al., 2015; Wang et al., 2016). The high level of POD level in endophyte-treated plants under Ni stress is consistent with the findings of Yang et al. (2015), Jiang et al. (2016), and Mendarte-Alquisira et al. (2016), who observed a remarkable increase in the POD concentration of fungal-inoculated *Robinia pseudoacacia* L., *S. nigrum*, and *Festuca arundinacea* plants, respectively, in response to different heavy-metal treatments. POD is considered vital for the dismutation of H_2O_2 to water and molecular oxygen. In the current experiment, increased POD activity suggests that the application of endophytic fungi scavenged ROS and helped soybean plants to reduce oxidation-induced damage. Likewise, PPO is thought to play a vital role in abiotic stresses; consequently, the results reported herein show that endophytic fungal-treated plants possess high PPO activity in under Ni toxicity. PPO chelates metals ions and directly encounters ROS during heavy metal stress and can prevent lipid peroxidation by scavenging the lipid alkoxy radical (Arora et al., 2000; Sharma et al., 2012). Our results were in line with those reported by Hashem et al. (2016) who observed increased PPO activity in fungal-inoculated plants under cadmium stress.

In current study, the SOD activity was detected to enhance significantly in fungal inoculated plants than those of non-inoculated plants under Ni toxicity. Our findings coincides with the results of Rozpdeka et al. (2014) and Yang et al. (2015), who detected remarkable SOD activity in endophytic fungal inoculated *Cichorium intybus* L. and *R. pseudoacacia* L. plants, respectively, under Cd, Zn, and Pb stress. The results of the current study demonstrate that fungal inoculation triggered the SOD activity in soybean plant to cope with the super oxide radicals. The higher SOD activity under Ni stress could possibly be related to the high production of superoxide radicals due to the excess of heavy metals ions as well as the *de novo* synthesis of the enzymatic protein (Garg and Kaur, 2013; Yang et al.,

2015). Likewise, CAT activity was also significantly enhanced with the increasing Ni stress level in fungal treated than those of non-treated plants. CAT is considered crucial enzyme for the dismutation of H₂O₂ to H₂O and molecular oxygen in the cells. Thus, increase in CAT production in inoculated plants could possibly be the response to the over production of hydrogen peroxide under Ni stress condition – suggesting fungal inoculation successfully alleviated the Ni induced oxidative stress in soybean plants. A number of studies have demonstrated that symbiotic interaction of fungal with host plant maximizes the antioxidant activities in order to strengthen heavy metal tolerance and alleviate oxidative stresses. For instance, Tan et al. (2015) and Khan et al. (2017a) demonstrated that endophytic *A. alternata* RSF-6L and AM *Glomus versiforme* inoculation enhanced CAT activity of *S. nigrum* and *S. photeinocarpum*, respectively, under severe Cadmium contaminated soil. In the current study, the upsurge in the antioxidant activities of inoculated plants under Ni toxicity suggests that endophytic fungal may regulate and activate the genes, which are responsible for encoding for antioxidant enzymes (Wang et al., 2016).

During heavy metal toxicity, plants accumulate high levels of endogenous phytohormones, such as ABA and JA in order to cope with metal-induced stress (Yan et al., 2015; Carrión-Seguí et al., 2016). However, the enhanced biosynthesis of ABA under severe stress leads to leaf senescence and directly inhibits photosynthesis. Hence, the growth rate of plants is suppressed. Deng et al. (2016) reported that high levels of ABA accumulation under multi-metal stress resulted in the inhibition of maize seed germination. Interestingly, in our experiment, *P. formosus*-treated plants resulted in reduced accumulation of ABA in both roots and shoot as compared to non-treated plants during Ni stress. This suggests that endophytic association with plants ameliorates the effects of metal stress. The level of ABA in the root in inoculated plants was higher than that in the shoot, which might be due the direct exposure of roots to metal stress, as it is transported to shoots via apoplastic route to regulate stomatal behavior (Xu et al., 2010). Previous reports rectify our findings that fungal application lowers the accumulation of ABA in plants under abiotic stresses (Khan et al., 2012b; Aroca et al., 2013; Waqas et al., 2015). Maintaining reduced ABA levels in inoculated plants under Ni stress can be accredited to the potential of *P. formosus* to produce GAs, because a similar trend was exhibited by plants in combination with the GA-secreting endophyte *Penicillium reseedanum* under abiotic stresses (Khan et al., 2015c). Reduced ABA has often coupled with increased endogenous GA contents. This could be true as the *P. formosus* inoculated plants had significantly higher biomass and shoot length, suggesting pivotal role of GA activation during stress. A similar conclusion was also drawn by Khan and Lee (2013) and Kang et al. (2014), where the GA producing microbes have resulted in reduced ABA and increased GA during abiotic stress conditions.

While assessing the effect of fungal inoculation under nickel toxicity on endogenous JA, which acts as a key regulator in plant defense systems, we observed a tendency for JA to be down-regulated, as compared to that in non-inoculated plants. The reduced level of stress-related hormones, such as ABA and JA

compared to control plants suggest that stress is managed better in the fungal-treated plants. The elevated level of JA might inhibit plant growth, as well as suppress GSH synthesis, which is an effective ROS scavenger during stress (Khan A.L. et al., 2017). We studied the increased production of GSH in fungal-treated plants under varying levels of nickel stress. Thus, enhanced GSH synthesis might have helped plants to limit the devastating effects of metal stress by counteracting ROS, and as a result, stress has been resolved and less JA is biosynthesized for plant defense mechanisms. The present results are consistent with those of Kim et al. (2014) who observed decreased levels of endogenous JA in rice under cadmium stress. The current results suggest that endophytic fungus protects plants under Ni stress. Whereas, the cross talk between endogenous hormonal viz. GA, ABA and JA signaling are extremely punitive (Pacifici et al., 2015), that may vary against different environmental stimuli. However, the exact mechanisms of ABA and JA modulation by endophytic fungi under metal stress require further study.

CONCLUSION

The current study reports, for the first time, that the GA-secreting endophyte *P. formosus* is a promising alternative for not only improving plant biomass but also significantly protecting host plant from the adverse effects of metal toxicity. Hence, its ability to produce IAA, as reported previously, in combination with GAs might confer tolerance to Ni phytotoxicity by reducing the levels of Ni toxicity in roots and shoots. In addition, it also has induces remarkable increases in growth attributes. Our results revealed that *P. formosus* inoculation enhanced soybean tolerance to Ni via a mechanism affecting the distribution of Ni in soybean tissue and via the induction of hormonal regulation and antioxidant systems. Thus, the association between phytohormone-inducing fungi and soybean may represent a promising strategy for achieving safer and profitable crop production, and to eliminate toxicity from Ni-contaminated soil. However, plant systems are complex and are responsible for controlling intracellular ion levels, including essential nutrients as well as non-essential metals. Therefore, in future studies, emphasis should be given to the role of *P. formosus* in enhancing macro and micro-nutrients under metal toxicity, because these nutrients have repeatedly been reported to enhance metals tolerance or accumulation potential in plants. Additionally, further molecular studies and investigating of transcriptional level work studies are recommended for an in-depth understanding of *P. formosus* association with host plants in metal-contaminated soil as well as Ni-polluted field trials to assess the role of endophytic fungal on a large scale.

AUTHOR CONTRIBUTIONS

SB, AK, RS, S-MK conceived and designed the experiments; SB and SA performed the experiments; AK and I-JL analyzed the data; I-JL contributed reagents/materials/analysis tools; SB, AK, and RS wrote the paper.

ACKNOWLEDGMENT

This work was financially supported by the Agenda Program (Project No. PJ01228603), Rural Development Administration, South Korea.

REFERENCES

- Abdel Latef, A. A. M., Jan, S., Abd Allah, E. F., Rashid, B., John, R., and Ahmad, P. (2016). "Soybean under abiotic stress: proteomic approach," in *Plant-Environment Interaction: Responses and Approaches to Mitigate Stress*, eds M. M. Azooz and P. Ahmad (Hoboken, NJ: John Wiley and Sons, Inc), 28–42. doi: 10.1002/9781119081005.ch2
- Albermann, S., Elter, T., Teubner, A., Krischke, W., Hirth, T., and Tudzynski, B. (2013). Characterization of novel mutants with an altered gibberellin spectrum in comparison to different wild-type strains of *Fusarium fujikuroi*. *Appl. Microbiol. Biotechnol.* 97, 7779–7790. doi: 10.1007/s00253-013-4917-7
- Aly, A. H., Debbab, A., and Proksch, P. (2011). Fungal endophytes: unique plant inhabitants with great promises. *Appl. Microbiol. Biotechnol.* 90, 1829–1845. doi: 10.1007/s00253-011-3270-y
- Amari, T., Ghnaya, T., Debez, A., Taamali, M., Youssef, N. B., Lucchini, G., et al. (2014). Comparative Ni tolerance and accumulation potentials between *Mesembryanthemum crystallinum* (halophyte) and *Brassica juncea*: metal accumulation, nutrient status and photosynthetic activity. *J. Plant Physiol.* 171, 1634–1644. doi: 10.1016/j.jplph.2014.06.020
- Aroca, R., Ruiz-Lozano, J. M., Zamarreño, ÁM., Paz, J. A., García-Mina, J. M., Pozo, M. J., et al. (2013). Arbuscular mycorrhizal symbiosis influences strigolactone production under salinity and alleviates salt stress in lettuce plants. *J. Plant Physiol.* 170, 47–55. doi: 10.1016/j.jplph.2012.08.020
- Arora, A., Byrem, T. M., Nair, M. G., and Strasburg, G. M. (2000). Modulation of liposomal membrane fluidity by flavonoids and isoflavonoids. *Arch. Biochem. Biophys.* 373, 102–109. doi: 10.1006/abbi.1999.1525
- Asada, K., Foyer, C. H., and Mullineaux, P. M. (1994). "Production and action of active oxygen species in photosynthetic tissues," in *Causes of Photooxidative Stress and Amelioration of Defense Systems in Plants*, eds C. H. Foyer and P. Mullineaux (Boca Raton, FL: CRC Press Inc), 77–104.
- Babu, A. G., Shim, J., Bang, K. S., Shea, P. J., and Oh, B. T. (2014). Trichoderma virens PDR-28: a heavy metal-tolerant and plant growth-promoting fungus for remediation and bioenergy crop production on mine tailing soil. *J. Environ. Manage.* 132, 129–134. doi: 10.1016/j.jenvman.2013.10.009
- Bashri, G., and Prasad, S. M. (2016). Exogenous IAA differentially affects growth, oxidative stress and antioxidants system in Cd stressed *Trigonella foenum-graecum* L. seedlings: toxicity alleviation by up-regulation of ascorbate-glutathione cycle. *Ecotoxicol. Environ. Saf.* 132, 329–338. doi: 10.1016/j.ecotox.2016.06.015
- Böhmke, C., and Tudzynski, B. (2009). Diversity, regulation, and evolution of the gibberellin biosynthetic pathway in fungi compared to plants and bacteria. *Phytochemistry* 70, 1876–1893. doi: 10.1016/j.phytochem.2009.05.020
- Bourdel, G., Roy-Bolduc, A., St-Arnaud, M., and Hijri, M. (2016). Concentration of petroleum-hydrocarbon contamination shapes fungal endophytic community structure in plant roots. *Front. Microbiol.* 7:685. doi: 10.3389/fmicb.2016.00685
- Card, S. D., Hume, D. E., Roodi, D., McGill, C. R., Millner, J. P., and Johnson, R. D. (2015). Beneficial endophytic microorganisms of *Brassica* – A review. *Biol. Control* 90, 102–112. doi: 10.1016/j.biocntrol.2015.06.001
- Carrió-Seguí, Á., Romero, P., Sanz, A., and Peñarrubia, L. (2016). Interaction between ABA signaling and copper homeostasis in *Arabidopsis thaliana*. *Plant Cell Physiol.* 57, 1568–1582. doi: 10.1093/pcp/pcw087
- Clemens, S., and Ma, J. F. (2016). Toxic heavy metal and metalloid accumulation in crop plants and foods. *Annu. Rev. Plant Biol.* 67, 489–512. doi: 10.1146/annurev-arplant-043015-112301
- Czepak, R., Piotrowska, A., and Szulecka, K. (2006). Jasmonic acid affects changes in the growth and some components content in alga *Chlorella vulgaris*. *Acta Physiol. Plant.* 28, 195–203. doi: 10.1007/BF02706531
- Dar, T. A., Uddin, M., Khan, M. M. A., Hakeem, K. R., and Jaleel, H. (2015). Jasmonates counter plant stress: a review. *Environ. Exp. Bot.* 115, 49–57. doi: 10.1093/pcp/pcu158
- Degola, F., Fattorini, L., Bona, E., Sprimuto, C. T., Argese, E., Berta, G., et al. (2015). The symbiosis between *Nicotiana tabacum* and the endomycorrhizal fungus *Funneliformis mosseae* increases the plant glutathione level and decreases leaf cadmium and root arsenic contents. *Plant Physiol. Biochem.* 92, 11–18. doi: 10.1016/j.plaphy.2015.04.001
- Deng, B., Yang, K., Zhang, Y., and Li, Z. (2016). Can heavy metal pollution defend seed germination against heat stress? Effect of heavy metals (Cu²⁺, Cd²⁺ and Hg²⁺) on maize seed germination under high temperature. *Environ. Pollut.* 216, 46–52. doi: 10.1016/j.envpol.2016.05.050
- Deng, Z., and Cao, L. (2017). Fungal endophytes and their interactions with plants in phytoremediation: a review. *Chemosphere* 168, 1100–1106. doi: 10.1016/j.chemosphere.2016.10.097
- Deng, Z., Cao, L., Huang, H., Jiang, X., Wang, W., Shi, Y., et al. (2011). Characterization of Cd- and Pb-resistant fungal endophyte *Mucor* sp. CBRF59 isolated from rapes (*Brassica chinensis*) in a metal-contaminated soil. *J. Hazard. Mat.* 185, 717–724. doi: 10.1016/j.jhazmat.2010.09.078
- Deng, Z., Zhang, R., Shi, Y., Tan, H., and Cao, L. (2014). Characterization of Cd-, Pb-, Zn-resistant endophytic *Lasiodiplodia* sp. MXSF31 from metal accumulating *Portulaca oleracea* and its potential in promoting the growth of rape in metal-contaminated soils. *Environ. Sci. Pollut. Res. Int.* 21, 2346–2357. doi: 10.1007/s11356-013-2163-2
- Ellman, G. L. (1959). Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* 82, 70–77. doi: 10.1016/0003-9861(59)90090-6
- Fomina, M., and Gadd, G. M. (2014). Biosorption: current perspectives on concept, definition and application. *Bioresour. Technol.* 160, 3–14. doi: 10.1016/j.biortech.2013.12.102
- Gajewska, E., Bernat, P., Dugoński, J., and Skodowska, M. (2012). Effect of nickel on membrane integrity, lipid peroxidation and fatty acid composition in wheat seedlings. *J. Agron. Crop Sci.* 198, 286–294. doi: 10.1111/j.1439-037X.2012.00514.x
- Gao, W., Nan, T., Tan, G., Zhao, H., Tan, W., Meng, F., et al. (2015). Cellular and subcellular immunohistochemical localization and quantification of cadmium ions in wheat (*Triticum aestivum*). *PLoS ONE* 10:e0123779. doi: 10.1371/journal.pone.0123779
- Gao, Y., Miao, C., Mao, L., Zhou, P., Jin, Z., and Shi, W. (2010). Improvement of phytoextraction and antioxidative defense in *Solanum nigrum* L. under cadmium stress by application of cadmium-resistant strain and citric acid. *J. Hazard. Mater.* 181, 771–777. doi: 10.1016/j.jhazmat.2010.05.080
- Gao, Y., Miao, C., Wang, Y., Xia, J., and Zhou, P. (2012). Metal-resistant microorganisms and metal chelators synergistically enhance the phytoremediation efficiency of *Solanum nigrum* L. in Cd-and Pb-contaminated soil. *Environ. Technol.* 33, 1383–1389. doi: 10.1080/09593330.2011.629006
- Garg, N., and Kaur, H. (2013). Response of antioxidant enzymes, phytochelatins and glutathione production towards Cd and Zn stresses in *Cajanus cajan* (L.) Millsp. genotypes colonized by arbuscular mycorrhizal fungi. *J. Agron. Crop Sci.* 199, 118–133. doi: 10.1111/j.1439-037X.2012.00533.x
- Gill, R. A., Ali, B., Islam, F., Farooq, M. A., Gill, M. B., Mwamba, T. M., et al. (2015). Physiological and molecular analyses of black and yellow seeded *Brassica napus* regulated by 5-aminolivulinic acid under chromium stress. *Plant Physiol. Biochem.* 94, 130–143. doi: 10.1016/j.plaphy.2015.06.001
- Hashem, A., Abd_Allah, E. F., Alqarawi, A. A., Malik, J. A., Wirth, S., and Egamberdieva, D. (2016). Role of calcium in AMF-mediated alleviation of the adverse impacts of cadmium stress in *Bassia indica* [Wight] AJ Scott. *Saudi J. Biol. Sci.* (in press). doi: 10.1016/j.sjbs.2016.11.003
- Henke, C., Jung, E. M., Voit, A., Kothe, E., and Krause, K. (2016). Dehydrogenase genes in the ectomycorrhizal fungus *Tricholoma vaccinum*: a role for Ald1 in mycorrhizal symbiosis. *J. Basic Microbiol.* 56, 162–174. doi: 10.1002/jobm.201500381

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.00870/full#supplementary-material>

- Hossain, Z., and Komatsu, S. (2013). Contribution of proteomic studies towards understanding plant heavy metal stress response. *Front. Plant Sci.* 3:310. doi: 10.3389/fpls.2012.00310
- Hou, Q., Ufer, G., and Bartels, D. (2016). Lipid signalling in plant responses to abiotic stress. *Plant Cell Environ.* 39, 1029–1048. doi: 10.1111/pce.12666
- Jiang, Q. Y., Tan, S. Y., Zhuo, F., Yang, D. J., Ye, Z. H., and Jing, Y. X. (2016). Effect of *Funnelformis* mosseae on the growth, cadmium accumulation and antioxidant activities of *Solanum nigrum*. *Appl. Soil Ecol.* 98, 112–120. doi: 10.1016/j.apsoil.2015.10.003
- Jozefczak, M., Remans, T., Vangronsveld, J., and Cuypers, A. (2012). Glutathione is a key player in metal-induced oxidative stress defenses. *Int. J. Mol. Sci.* 13, 3145–3175. doi: 10.3390/ijms13033145
- Kamran, M. A., Syed, J. H., Eqani, S. A. M. A. S., Munis, M. F. H., and Chaudhary, H. J. (2015). Effect of plant growth-promoting rhizobacteria inoculation on cadmium (Cd) uptake by *Eruca sativa*. *Environ. Sci. Pollut. Res.* 22, 9275–9283. doi: 10.1007/s11356-015-4074-x
- Kang, S. M., Radhakrishnan, R., Khan, A. L., Kim, M. J., Park, J. M., Kim, B. R., et al. (2014). Gibberellin secreting rhizobacterium, *Pseudomonas putida* H-2-3 modulates the hormonal and stress physiology of soybean to improve the plant growth under saline and drought conditions. *Plant Physiol. Biochem.* 84, 115–124. doi: 10.1016/j.plaphy.2014.09.001
- Kaushik, P., and Malik, A. (2010). Alkali, thermo and halo tolerant fungal isolate for the removal of textile dyes. *Colloids Surf. B Biointerfaces* 81, 321–328. doi: 10.1016/j.colsurfb.2010.07.034
- Kaushik, P., and Malik, A. (2011). Process optimization for efficient dye removal by *Aspergillus lentinus* FJ172995. *J. Hazard. Mat.* 185, 837–843. doi: 10.1016/j.jhazmat.2010.09.096
- Khaliq, A., Ali, S., Hameed, A., Farooq, M. A., Farid, M., Shakoor, M. B., et al. (2016). Silicon alleviates nickel toxicity in cotton seedlings through enhancing growth, photosynthesis, and suppressing Ni uptake and oxidative stress. *Arch. Agron. Soil Sci.* 62, 633–647. doi: 10.1080/03650340.2015.1073263
- Khan, A. L., Al-Harrasi, A., Al-Rawahi, A., Al-Farsi, Z., Al-Mamari, A., Waqas, M., et al. (2016). Endophytic fungi from Frankincense tree improves host growth and produces extracellular enzymes and indole acetic acid. *PLoS ONE* 11:e0158207. doi: 10.1371/journal.pone.0158207
- Khan, A. L., Hamayun, M., Ahmad, N., Waqas, M., Kang, S. M., Kim, Y. H., et al. (2011a). *Exophiala* sp. LHL08 reprograms *Cucumis sativus* to higher growth under abiotic stresses. *Physiol. Plant.* 143, 329–343. doi: 10.1111/j.1399-3054.2011.01508.x
- Khan, A. L., Hamayun, M., Kang, S. M., Kim, Y. H., Jung, H. Y., Lee, J. H., et al. (2012a). Endophytic fungal association via gibberellins and indole acetic acid can improve plant growth under abiotic stress: an example of *Paecilomyces formosus* LHL10. *BMC Microbiol.* 12:3. doi: 10.1186/1471-2180-12-3
- Khan, A. L., Hamayun, M., Kim, Y. H., Kang, S. M., and Lee, I. J. (2011b). Ameliorative symbiosis of endophyte (*Penicillium funiculosum* LHL06) under salt stress elevated plant growth of *Glycine max* L. *Plant Physiol. Biochem.* 49, 852–861. doi: 10.1016/j.plaphy.2011.03.005
- Khan, A. L., Hamayun, M., Radhakrishnan, R., Waqas, M., Kang, S. M., Kim, Y. H., et al. (2012b). Mutualistic association of *Paecilomyces formosus* LHL10 offers thermotolerance to *Cucumis sativus*. *Antonie van Leeuwenhoek* 101, 267–279. doi: 10.1007/s10482-011-9630-x
- Khan, A. L., Hussain, J., Al-Harrasi, A., Al-Rawahi, A., and Lee, I. J. (2015a). Endophytic fungi: resource for gibberellins and crop abiotic stress resistance. *Crit. Rev. Biotechnol.* 35, 62–74. doi: 10.3109/07388551.2013.800018
- Khan, A. L., and Lee, I. J. (2013). Endophytic *Penicillium funiculosum* LHL06 secretes gibberellin that reprograms *Glycine max* L. growth during copper stress. *BMC Plant Biol.* 13:86. doi: 10.1186/1471-2229-13-86
- Khan, A. L., Waqas, M., Asaf, S., Kamran, M., Shahzad, R., Bilal, S., et al. (2017). Plant growth-promoting endophyte *Sphingomonas* sp. LK11 alleviates salinity stress in *Solanum pimpinellifolium*. *Environ. Exp. Bot.* 133, 58–69. doi: 10.1016/j.jenvexpbot.2016.09.009
- Khan, A. L., Waqas, M., Hussain, J., Al-Harrasi, A., Hamayun, M., and Lee, I. J. (2015b). Phytohormones enabled endophytic fungal symbiosis improve aluminum phytoextraction in tolerant *Solanum lycopersicum*: an examples of *Penicillium janthinellum* LK5 and comparison with exogenous GA 3. *J. Hazard. Mater.* 295, 70–78. doi: 10.1016/j.jhazmat.2015.04.008
- Khan, A. L., Waqas, M., Hussain, J., Al-Harrasi, A., and Lee, I. J. (2014). Fungal endophyte *Penicillium janthinellum* LK5 can reduce cadmium toxicity in *Solanum lycopersicum* (Sitiens and Rhe). *Biol. Fertil. Soils* 50, 75–85. doi: 10.1007/s00374-013-0833-3
- Khan, A. L., Waqas, M., Khan, A. R., Hussain, J., Kang, S. M., Gilani, S. A., et al. (2013). Fungal endophyte *Penicillium janthinellum* LK5 improves growth of ABA-deficient tomato under salinity. *World J. Microbiol.* 29, 2133–2144. doi: 10.1007/s11274-013-1378-1
- Khan, A. L., Waqas, M., and Lee, I. J. (2015c). Resilience of *Penicillium resedanum* LK6 and exogenous gibberellin in improving *Capsicum annuum* growth under abiotic stresses. *J. Plant Res.* 128, 259–268. doi: 10.1007/s10265-014-0688-1
- Khan, A. R., Ullah, I., Waqas, M., Park, G. S., Khan, A. L., Hong, S. J., et al. (2017a). Host plant growth promotion and cadmium detoxification in *Solanum nigrum*, mediated by endophytic fungi. *Ecotoxicol. Environ. Saf.* 136, 180–188. doi: 10.1016/j.ecoenv.2016.03.014
- Khan, A. R., Waqas, M., Ullah, I., Khan, A. L., Khan, M. A., Lee, I. J., et al. (2017b). Culturable endophytic fungal diversity in the cadmium hyperaccumulator *Solanum nigrum* L. and their role in enhancing phytoremediation. *Environ. Exp. Bot.* 135, 126–135. doi: 10.1016/j.jenvexpbot.2016.03.005
- Kim, Y. H., Khan, A. L., Kim, D. H., Lee, S. Y., Kim, K. M., Waqas, M., et al. (2014). Silicon mitigates heavy metal stress by regulating P-type heavy metal ATPases, *Oryza sativa* low silicon genes, and endogenous phytohormones. *BMC Plant Biol.* 14:13. doi: 10.1186/1471-2229-14-13
- Krause, K., Henke, C., Asiimwe, T., Ulbricht, A., Klemmer, S., Schachtschabel, D., et al. (2015). Biosynthesis and secretion of indole-3-acetic acid and its morphological effects on *Tricholoma vaccinum-spruce* ectomycorrhiza. *Appl. Environ. Microbiol.* 81, 7003–7011. doi: 10.1128/AEM.01991-15
- Li, X., Wang, Y., Pan, Y., Yu, H., Zhang, X., Shen, Y., et al. (2017). Mechanisms of Cd and Cr removal and tolerance by macrofungus *Pleurotus ostreatus* HAU-2. *J. Hazard. Mater.* 330, 1–8. doi: 10.1016/j.jhazmat.2017.01.047
- Liu, Z., Ding, Y., Wang, F., Ye, Y., and Zhu, C. (2016). Role of salicylic acid in resistance to cadmium stress in plants. *Plant Cell Rep.* 35, 719–731. doi: 10.1007/s00299-015-1925-3
- Llamas, A., and Sanz, A. (2008). Organ-distinctive changes in respiration rates of rice plants under nickel stress. *Plant Growth Regul.* 54, 63–69. doi: 10.1007/s10725-007-9229-4
- Ma, Y., Rajkumar, M., Zhang, C., and Freitas, H. (2016). Beneficial role of bacterial endophytes in heavy metal phytoremediation. *J. Environ. Manage.* 174, 14–25. doi: 10.1016/j.jenvman.2016.02.047
- Marcos, A. T., Ramos, M. S., Marcos, J. F., Carmona, L., Strauss, J., and Cánovas, D. (2016). Nitric oxide synthesis by nitrate reductase is regulated during development in *Aspergillus*. *Mol. Microbiol.* 99, 15–33. doi: 10.1111/mmi.13211
- McCloud, E. S., and Baldwin, I. T. (1997). Herbivory and caterpillar regurgitants amplify the wound-induced increases in jasmonic acid but not nicotine in *Nicotiana sylvestris*. *Planta* 203, 430–435. doi: 10.1007/s004250050210
- Mendarte-Alquisira, C., Gutiérrez-Rojas, M., González-Márquez, H., and Volke-Sepúlveda, T. (2016). Improved growth and control of oxidative stress in plants of *Festuca arundinacea* exposed to hydrocarbons by the endophytic fungus *Lewia* sp. *Plant Soil* 411, 347–358. doi: 10.1007/s11104-016-3035-6
- Miransari, M. (2015). *Soybean Production and Heavy Metal Stress - Abiotic and Biotic Stresses in Soybean Production*. San Diego, CA: Academic Press, 197–216. doi: 10.1016/B978-0-12-801536-0.00009-8
- Mishra, A., and Malik, A. (2014). Novel fungal consortium for bioremediation of metals and dyes from mixed waste stream. *Bioresour. Technol.* 171, 217–226. doi: 10.1016/j.biortech.2014.08.047
- Mosa, A., El-Banna, M. F., and Gao, B. (2016). Biochar filters reduced the toxic effects of nickel on tomato (*Lycopersicon esculentum* L.) grown in nutrient film technique hydroponic system. *Chemosphere* 149, 254–262. doi: 10.1016/j.chemosphere.2016.01.104
- Noriega, G. O., Balestrasse, K. B., Batlle, A., and Tomaro, M. L. (2007). Cadmium induced oxidative stress in soybean plants also by the accumulation of δ-aminolevulinic acid. *Biometals* 20, 841–851. doi: 10.1007/s10534-006-9077-0
- Pacifci, E., Polverari, L., and Sabatini, S. (2015). Plant hormone cross-talk: the pivot of root growth. *J. Exp. Bot.* 66, 1113–1121. doi: 10.1093/jxb/eru534
- Padash, A., Shahabivand, S., Behtash, F., and Aghaei, A. (2016). A practicable method for zinc enrichment in lettuce leaves by the endophyte fungus *Piriformospora indica* under increasing zinc supply. *Sci. Hortic.* 213, 367–372. doi: 10.1016/j.scienta.2016.10.040

- Pérez-Chaca, M. V., Rodríguez-Serrano, M., Molina, A. S., Pedranzani, H. E., Zirulnik, F., Sandalio, L. M., et al. (2014). Cadmium induces two waves of reactive oxygen species in *Glycine max* (L.) roots. *Plant Cell Environ.* 37, 1672–1687. doi: 10.1111/pce.12280
- Qi, Q., Rose, P. A., Abrams, G. D., Taylor, D. C., Abrams, S. R., and Cutler, A. J. (1998). Abscisic acid metabolism, 3-ketoacyl-coenzyme a synthase gene expression, and very-long-chain monounsaturated fatty acid biosynthesis in *Brassica napus* embryos. *Plant Physiol.* 117, 979–987. doi: 10.1104/pp.117.3.979
- Qin, Y. Y., Leung, C. K. M., Lin, C. K., and Wong, M. H. (2015). The associations between metals/metalloids concentrations in blood plasma of Hong Kong residents and their seafood diet, smoking habit, body mass index and age. *Environ. Sci. Pollut. Res.* 22, 13204–13211. doi: 10.1007/s11356-015-4417-7
- Rajkumar, M., Sandhya, S., Prasad, M. N. V., and Freitas, H. (2012). Perspectives of plant-associated microbes in heavy metal phytoremediation. *Biotechnol. Adv.* 30, 1562–1574. doi: 10.1016/j.biotechadv.2012.04.011
- Rao, R. P., Hunter, A., Kashpur, O., and Normanly, J. (2010). Aberrant synthesis of indole-3-acetic acid in *Saccharomyces cerevisiae* triggers morphogenic transition, a virulence trait of pathogenic fungi. *Genetics* 185, 211–220. doi: 10.1534/genetics.109.112854
- Redman, R. S., Kim, Y. O., Woodward, C. J., Greer, C., Espino, L., Doty, S. L., et al. (2011). Increased fitness of rice plants to abiotic stress via habitat adapted symbiosis: a strategy for mitigating impacts of climate change. *PLoS ONE* 6:e14823. doi: 10.1371/journal.pone.0014823
- Rehman, M. Z. U., Rizwan, M., Ali, S., Fatima, N., Yousaf, B., Naeem, A., et al. (2016). Contrasting effects of biochar, compost and farm manure on alleviation of nickel toxicity in maize (*Zea mays* L.) in relation to plant growth, photosynthesis and metal uptake. *Ecotoxicol. Environ. Saf.* 133, 218–225. doi: 10.1016/j.ecoenv.2016.07.023
- Rim, S. O., Yoon, H. J., Lee, J. H., Kim, C. M., and Kim, J. G. (2013). Characterization of gibberellin biosynthetic gene cluster from *Fusarium proliferatum*. *J. Microbiol. Biotechnol.* 23, 623–629. doi: 10.4014/jmb.1212.12029
- Rozpdeć, P., Węowicz, K., Stojakowska, A., Malarz, J., Surówka, E., Anielska, T., et al. (2014). Mycorrhizal fungi modulate phytochemical production and antioxidant activity of *Cichorium intybus* L. (Asteraceae) under metal toxicity. *Chemosphere* 112, 217–224. doi: 10.1016/j.chemosphere.2014.04.023
- Saleem, M., Asghar, H. N., Khan, M. Y., and Zahir, Z. A. (2015). Gibberellic acid in combination with pressmud enhances the growth of sunflower and stabilizes chromium (VI)-contaminated soil. *Environ. Sci. Pollut. Res.* 22, 10610–10617. doi: 10.1007/s11356-015-4275-3
- Sharma, P., Jha, A. B., Dubey, R. S., and Pessarakli, M. (2012). Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *J. Bot.* 2012:217037. doi: 10.1155/2012/217037
- Sharma, S. S., Dietz, K. J., and Mimura, T. (2016). Vacuolar compartmentalization as indispensable component of heavy metal detoxification in plants. *Plant Cell Environ.* 39, 1112–1126. doi: 10.1111/pce.12706
- Siddiqui, M. H., Al-Whaibi, M. H., and Basalah, M. O. (2011). Interactive effect of calcium and gibberellin on nickel tolerance in relation to antioxidant systems in *Triticum aestivum* L. *Protoplasma* 248, 503–511. doi: 10.1007/s00709-010-0197-6
- Sirhindī, G., Mir, M. A., Abd-Allah, E. F., Ahmad, P., and Gucel, S. (2016). Jasmonic acid modulates the physio-biochemical attributes, antioxidant enzyme activity, and gene expression in *Glycine max* under nickel toxicity. *Front. Plant Sci.* 7:591. doi: 10.3389/fpls.2016.00591
- Sirhindī, G., Mir, M. A., Sharma, P., Gill, S. S., Kaur, H., and Mushtaq, R. (2015). Modulatory role of jasmonic acid on photosynthetic pigments, antioxidants and stress markers of *Glycine max* L. under nickel stress. *Physiol. Mol. Biol. Plants* 21, 559–565. doi: 10.1007/s12298-015-0320-4
- Stella, T., Covino, S., Ěvaněarová, M., Filipová, A., Petruccioli, M., D'Annibale, A., et al. (2017). Bioremediation of long-term PCB-contaminated soil by white-rot fungi. *J. Hazard. Mater.* 324, 701–710. doi: 10.1016/j.jhazmat.2016.11.044
- Sura-de Jong, M., Reynolds, R. J., Richterova, K., Musilova, L., Staicu, L. C., Chocholata, I., et al. (2015). Selenium hyperaccumulators harbor a diverse endophytic bacterial community characterized by high selenium resistance and plant growth promoting properties. *Front. Plant Sci.* 6:113. doi: 10.3389/fpls.2015.00113
- Tan, S. Y., Jiang, Q. Y., Zhuo, F., Liu, H., Wang, Y. T., Li, S. S., et al. (2015). Effect of inoculation with *Glomus versiforme* on cadmium accumulation, antioxidant activities and phytochelatins of *Solanum photinocarpum*. *PLoS ONE* 10:e0132347. doi: 10.1371/journal.pone.0132347
- Tsavkelova, E., Oeser, B., Oren-Young, L., Israeli, M., Sasson, Y., Tudzynski, B., et al. (2012). Identification and functional characterization of indole-3-acetamide-mediated IAA biosynthesis in plant-associated *Fusarium* species. *Fungal Genet. Biol.* 49, 48–57. doi: 10.1016/j.fgb.2011.10.005
- Tudzynski, B., Hedden, P., Carrera, E., and Gaskin, P. (2001). The *P450-4* gene of *Gibberella fujikuroi* encodes ent-kaurene oxidase in the gibberellin biosynthesis pathway. *Appl. Environ. Microbiol.* 67, 3514–3522. doi: 10.1128/AEM.67.8.3514-3522.2001
- Vázquez, S., Goldsborough, P., and Carpene, R. O. (2009). Comparative analysis of the contribution of phytochelatins to cadmium and arsenic tolerance in soybean and white lupin. *Plant Physiol. Biochem.* 47, 63–67. doi: 10.1016/j.plaphy.2008.09.010
- Visioli, G., Vamerali, T., Mattarozzi, M., Dramis, L., and Sanangelantoni, A. M. (2015). Combined endophytic inoculants enhance nickel phytoextraction from serpentine soil in the hyperaccumulator *Noccaea caerulescens*. *Front. Plant Sci.* 6:638. doi: 10.3389/fpls.2015.00638
- Wang, J. L., Li, T., Liu, G. Y., Smith, J. M., and Zhao, Z. W. (2016). Unraveling the role of dark septate endophyte (DSE) colonizing maize (*Zea mays*) under cadmium stress: physiological, cytological and genic aspects. *Sci. Rep.* 6:22028. doi: 10.1038/srep22028
- Waqas, M., Khan, A. L., Shahzad, R., Ullah, I., Khan, A. R., and Lee, I. J. (2015). Mutualistic fungal endophytes produce phytohormones and organic acids that promote japonica rice plant growth under prolonged heat stress. *J. Zhejiang Univ. Sci. B* 16, 1011–1018. doi: 10.1631/jzus.B1500081
- Wu, Q., Cui, Y., Li, Q., and Sun, J. (2015). Effective removal of heavy metals from industrial sludge with the aid of a biodegradable chelating ligand GLDA. *J. Hazard. Mater.* 283, 748–754. doi: 10.1016/j.jhazmat.2014.10.027
- Xie, P. P., Deng, J. W., Zhang, H. M., Ma, Y. H., Cao, D. J., Ma, R. X., et al. (2015). Effects of cadmium on bioaccumulation and biochemical stress response in rice (*Oryza sativa* L.). *Ecotoxicol. Environ. Saf.* 122, 392–398. doi: 10.1016/j.ecoenv.2015.09.007
- Xu, Z., Zhou, G., and Shimizu, H. (2010). Plant responses to drought and rewetting. *Plant Signal. Behav.* 5, 649–654. doi: 10.4161/psb.5.6.11398
- Yan, Z., Li, X., Chen, J., and Tam, N. F. Y. (2015). Combined toxicity of cadmium and copper in *Avicennia marina* seedlings and the regulation of exogenous jasmonic acid. *Environ. Saf.* 113, 124–132. doi: 10.1016/j.ecoenv.2014.11.031
- Yang, Y., Han, X., Liang, Y., Ghosh, A., Chen, J., and Tang, M. (2015). The combined effects of arbuscular mycorrhizal fungi (AMF) and lead (Pb) stress on Pb accumulation, plant growth parameters, photosynthesis, and antioxidant enzymes in *Robinia pseudoacacia* L. *PLoS ONE* 10:e0145726. doi: 10.1371/journal.pone.0145726
- Yusuf, M., Fariduddin, Q., Hayat, S., and Ahmad, A. (2011). Nickel: an overview of uptake, essentiality and toxicity in plants. *Bull. Environ. Contam. Toxicol.* 86, 1–17. doi: 10.1007/s00128-010-0171-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer VVY and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2017 Bilal, Khan, Shahzad, Asaf, Kang and Lee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read for greatest visibility and readership



FAST PUBLICATION

Around 90 days from submission to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative, and constructive peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers acknowledged by name on published articles



REPRODUCIBILITY OF RESEARCH

Support open data and methods to enhance research reproducibility



DIGITAL PUBLISHING

Articles designed for optimal readership across devices



FOLLOW US
@frontiersin



IMPACT METRICS
Advanced article metrics track visibility across digital media



EXTENSIVE PROMOTION
Marketing and promotion of impactful research



LOOP RESEARCH NETWORK
Our network increases your article's readership

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00