SCIENTIFIC REPORTS

OPEN

Populus trichocarpa encodes small, effector-like secreted proteins that are highly induced during mutualistic symbiosis

Jonathan M. Plett[1,2], Hengfu Yin[3], Ritesh Mewalal[3], Rongbin Hu[3], Ting Li[3], Priya Ranjan[3], Sara Jawdy[3], Henrique C. De Paoli[3], George Butler[4], Tessa Maureen Burch-Smith[4], Hao-Bo Guo[4], Chun Ju Chen[3], Annegret Kohler[1], Ian C. Anderson[2], Jessy L. Labbé[3], Francis Martin[1], Gerald A. Tuskan[3] & Xiaohan Yang[3]

During symbiosis, organisms use a range of metabolic and protein-based signals to communicate. Of these protein signals, one class is defined as 'effectors', i.e., small secreted proteins (SSPs) that cause phenotypical and physiological changes in another organism. To date, protein-based effectors have been described in aphids, nematodes, fungi and bacteria. Using RNA sequencing of *Populus trichocarpa* roots in mutualistic symbiosis with the ectomycorrhizal fungus *Laccaria bicolor*, we sought to determine if host plants also contain genes encoding effector-like proteins. We identified 417 plant-encoded putative SSPs that were significantly regulated during this interaction, including 161 SSPs specific to *P. trichocarpa* and 15 SSPs exhibiting expansion in *Populus* and closely related lineages. We demonstrate that a subset of these SSPs can enter *L. bicolor* hyphae, localize to the nucleus and affect hyphal growth and morphology. We conclude that plants encode proteins that appear to function as effector proteins that may regulate symbiotic associations.

Symbiosis, defined as a durable interaction between two or more organisms, is a complex association that requires both physiological and morphological signaling, reception and alterations on the part of both organisms. Many fungal lineages within the pathogenic/mutualistic continuum have evolved elaborate protein-based signals to influence their hosts in order to support their metabolic requirements during symbiosis[1]. These proteins, called effectors, are typically fungal strain- or species-specific, usually ≤250 amino acids in size, and carry a secretion signal motif that may be cysteine-rich. Some examples of pathogenic effector proteins include: Avr3a of *Phytophthora infestans*[2,3], the ChCEs of *Colletotrichum higginsianum*[4], RTP1 of *Uromyces fabae*[5] and the TIN2 of *Ustilago maydis*[6]. More recently, effector proteins have also been characterized in mutualistic fungi. For example, MiSSP7 of the ectomycorrhizal (ECM) fungus *Laccaria bicolor* stabilizes the *Populus* jasmonic acid signaling repressor PtJAZ6[7] while the SP7 protein of the arbuscular mycorrhizal fungus *Rhizophagus irregularis* (formerly *Glomus intraradices*) blocks the activity of ERF19[8].

When viewed from the perspective of the fungal-encoded effector biology described to date, the outcome of symbiosis would appear to be heavily influenced by the colonizing fungus. Is this true, or could plants encode similar "effector-like" proteins that, in turn, exert a level of influence on the activity of the symbiotic associates? Previous research supporting the influence of plants over microbial biology suggests that: (1) plants utilize a number of secondary metabolites to affect rhizospheric and endopsheric microbes[9], (2) general root exudates act as microbial chemoattractants[10], (3) plants produce strigolactones which act as signals in initiating mycorrhizal formation[11] and (4) flavonoids from plants function as antimicrobial compounds[12,13].

[1]INRA, UMR 1136 INRA-University of Lorraine, Interactions Arbres/Microorganismes, Laboratory of Excellence ARBRE, INRA-Nancy, 54280, Champenoux, France. [2]Hawkesbury Institute for the Environment, University of Western Sydney, Richmond, 2753, NSW, Australia. [3]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA. [4]Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, TN, 37996, USA. Jonathan M. Plett and Hengfu Yin contributed equally to this work. Correspondence and requests for materials should be addressed to X.Y. (email: yangx@ornl.gov)

Recently, attention has turned to investigating proteins present in root exudates and their role in influencing plant-microbe interactions. Proteins, such as peptidases, hydrolases and defensins, have been implicated in affecting symbiosis[14–16]. There is also a precedent for plant small secreted proteins (SSPs; ≤250 amino acids) that are produced by roots and enter the cytosol of nitrogen-fixing bacteria during nodule formation to govern the outcome of these mutualistic interactions[17–20]. In total, 146 plant cysteine-rich SSPs have been identified[17, 19, 21]. It is generally understood that the discovery of SSPs in newly sequenced plant genomes has been hampered by the lack of gene annotation for proteins smaller than 100 amino acids (aa) due to a high false positive discovery rate from computational prediction of small open reading frames (ORFs) and a lack of homology across plant genomes which constrains *ab initio* gene calling[22, 23].

The lack of SSP annotation is changing, however, due to the increasing availability of deep RNA-sequencing data. Recent re-annotation of the *Arabidopsis* and *Populus* genomes found evidence for 2,099 and 1,282 new small proteins, respectively[22, 24]. Moreover, mutation analysis of a number of these new *Arabidopsis* small proteins has demonstrated that a large portion play critical roles in plant developmental processes such as flowering, leaf development and overall plant structure[24]. However, the possibility that a number of these proteins could be secreted and function as effector proteins during symbiotic interactions is still uncertain. The aim of this study was to determine if: (1) *Populus* encoded effector-like proteins are regulated during mutualistic symbiosis with the ectomycorrhizal fungus *L. bicolor* and (2) whether some of these proteins might be able to enter the hyphae of *L. bicolor* and affect the growth of the fungus.

Results

Size bias in proteins predicted to be secreted during a mutualistic interaction.
Using a computational pipeline for discovery of small protein-encoding genes based on transcriptomes analyzed by RNA-seq (Fig. 1a), we identified 2,819 *Populus* protein-encoding genes (designated as the 'All-protein set') that exhibited differential (p-value < 0.05) transcript abundance across all stages of mycorrhizal root tip development during symbiosis between *P. trichocarpa* and *L. bicolor*. The quantitative expression based on RNA-seq data was validated by quantitative RT-PCR analysis of 13 randomly-selected transcripts, which significantly correlated with the data obtained from the RNA-seq transcript analysis (p-value < 0.001; Supplementary Fig. S1; Supplementary Table S1). In the "All-protein set", 2,631 transcripts (93.3%) contained a complete open reading frame (ORF) and were designated as the full-length (FL) set. In the FL set, 1,242 proteins (47.2%) satisfied the criteria for small proteins, i.e., ≤250 amino acids in length, and were designated as the "SmP set". Given that SSPs are typically secreted, the predicted subcellular localizations of the FL set proteins were determined by three differential computational tools: LocTree2[25], CELLO[26] and YLoc[27]. From this analysis, 646 proteins in the FL set were predicted to code for secreted proteins (the SeP set). Within the SeP set, there was an over-representation of small proteins of ≤250 aa in length (64.6% of the SeP set), with the largest number of proteins falling within the size range of 51–100 aa (Fig. 1b). This subgroup of proteins was designated the SSP set (Supplementary Table S2).

This SSP set (417 proteins) was further divided into sub-sets based on three computational 'confidence level' approaches (see Methods section): the high confidence set ("SSP_hc set"; 171 proteins), the medium confidence set ("SSP_mc set"; 126 proteins) and the low confidence set ("SSP_lc set; 120 proteins; Fig. 1a; Supplementary Table S2). Proteins not found in the current *P. trichocarpa* genome annotation (37% of the SSP set) were named with the prefix 'CUFF' while proteins found in the current *Populus* genome annotation v3.0 (www.Phytozome. net) were designated with the prefix 'Potri'. *In silico* analysis of the SSP set found that 79 of these SSPs contained predicted nuclear localization signals as predicted by using YLoc[27] and CELLO[26]. Functional categorization of the differentially expressed SSPs with gene ontology (GO) biological process showed that these proteins are putatively involved in different biological pathways with varying biochemical activity (Supplementary Table S2). Overall, there was enrichment for stress-related GO terms, in particular the "response to fungus" category (Supplementary Table S3).
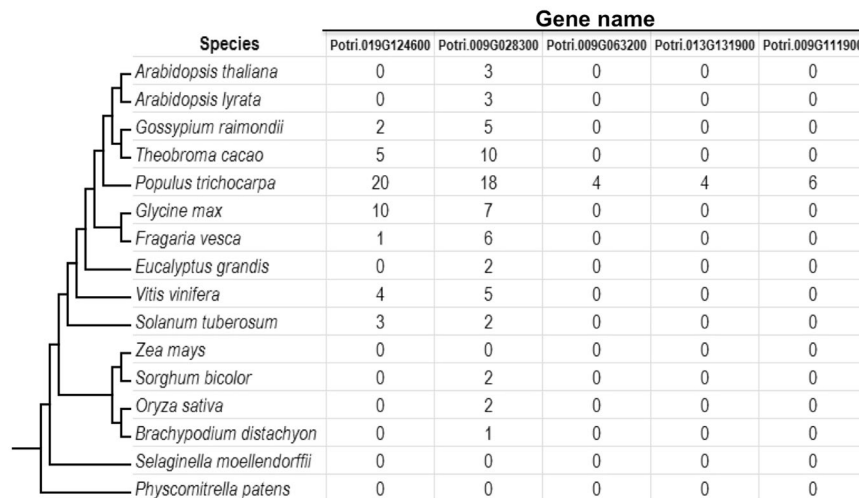
Lineage- or genus-specific set of *P. trichocarpa* SSP genes.
Of the entire 417 *P. trichocarpa* SSP set, 39% (or 161 SSPs) appear to be specific to *Populus*, i.e., without homologs in the other 15 species examined (Fig. 2; Supplementary Table S4), while 15 SSPs have homologs found only within closely related plant genera (Supplementary Table S5). A comparison of SSPs between *P. trichocarpa* accessions (Nisqually-1 and 93–960; strong ECM host plants) and *P. deltoides* accessions (ILL-101 and D124; poor ECM host plants)[28] showed that six SSPs predicted in the *P. trichocarpa* genomes were missing or truncated in the *P. deltoides* accessions (Table 1).

Populus SSPs have functional secretion signals.
To test the secretion of SSPs, a complementation assay in which the survival of the host depends on the secretion of the protein of interest was chosen. Forty full-length proteins (Supplementary Table S6), including their predicted secretion signals, were cloned upstream of a SUC2 invertase that lacked its native secretion signal. The recombinant genes were then expressed in the *Saccharomyces cerevisiae suc2* mutant that cannot grow on media containing sucrose as the sole energy source without the presence of a secreted SUC2 protein. We found that on average 15 of the 40 (38%) SSPs tested complemented the *suc2* mutation (Table 2; Fig. 3). Even though we used full-length genes, including the secretion signal in these tests, it is possible that the secondary structure of the small proteins masked the signal sequence and gave us a number of false negatives. Of the 40 SSPs tested, 16 were from the high confidence set (i.e., the SSP_hc set in Fig. 1a), and of those, 8 SSPs were able to able to rescue the mutant, yielding a 50% confirmation rate. In contrast, the medium confidence set of 14 SSPs (i.e., the SSP_mc set in Fig. 1a) and low confidence set of 10 SSPs (i.e., the SSP_lc set in Fig. 1a) each had a confirmation rate of around 30% (Table 2; Fig. 3). In summation, the signal peptides in some of the *Populus* SSPs were capable of directing secretion in a heterologous system.
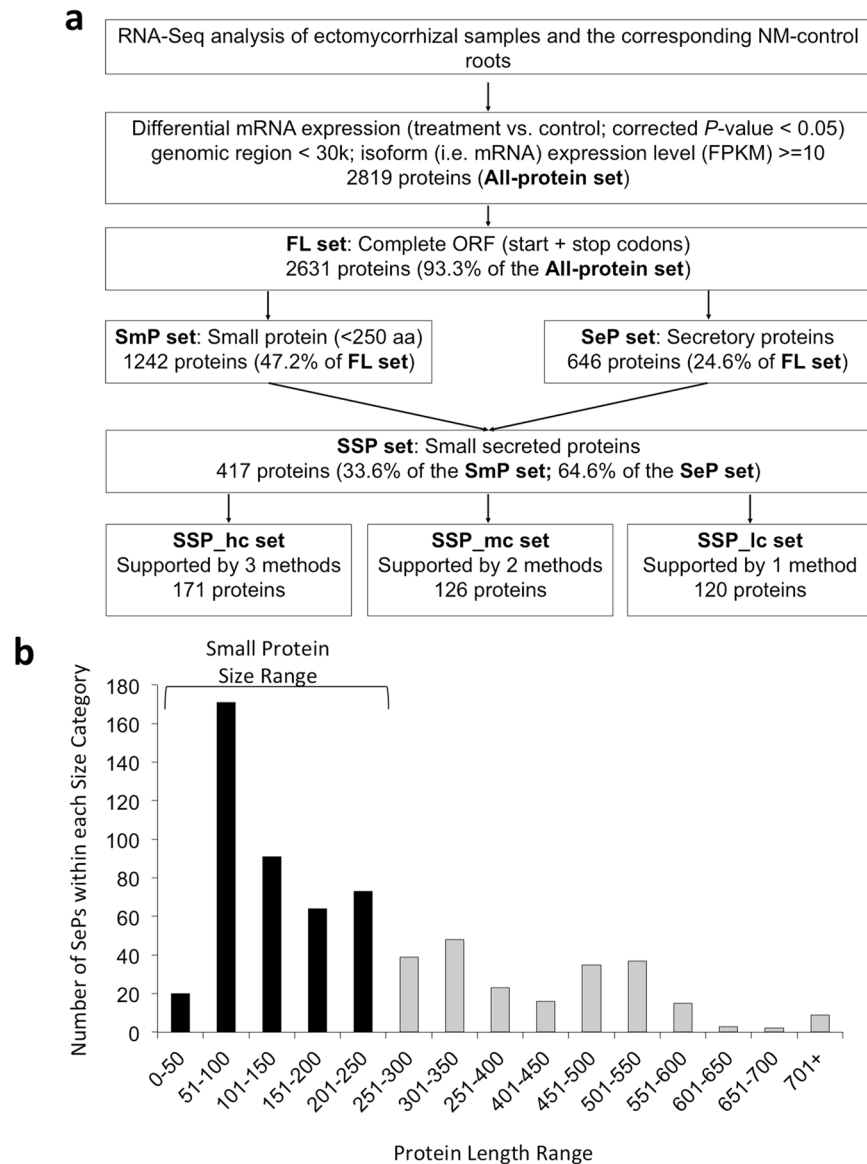
Figure 1. Characteristics of *Populus trichocarpa* genes differentially expressed during mutualistic symbiosis with *Laccaria bicolor*. (**a**) A computational pipeline for discovery of small protein-encoding genes in *P. trichocarpa* in response to *L. bicolor*. (**b**) Sequence length distribution of *P. trichocarpa* computationally-predicted secreted proteins expressed during mutualistic symbiosis with *L. bicolor*. 'Small' proteins are defined as ≤250 amino acids.

Populus SSPs can enter *Laccaria bicolor* hyphae and accumulate in the nucleus. In order to determine if *P. trichocarpa* SSPs could cross the *L. bicolor* hyphal membrane and accumulate within the fungal cells, we used five proteins (CUFF.29946, Potri.009G063200, Potri.010G251000, Potri.019G121200 and Potri.007G006800) produced by cell-free recombinant protein synthesis without the secretion signal and with an amino-terminal covalent addition of the FITC fluorescent molecule. Five biological replicates (i.e., independent *L. bicolor* colonies) were then treated with an individual protein or FITC alone on the leading edge of an *in vitro* fungal colony. Of the five FITC-tagged proteins tested, four (CUFF.29946, Potri.009G063200, Potri.010G251000 and Potri.007G006800) were able to enter the hyphae and concentrate in the nucleus (Fig. 4a–d). The fifth, Potri.019G121200, a negative control, did not localize to the nucleus of *L. bicolor* hyphae (Fig. 4e), indicating that the nuclear localization is not simply due to protein size. There was absence of fluorescent signal in the nuclei of hyphae treated with FITC alone (Fig. 4f). Localization of CUFF.29946, Potri.009G063200 and Potri.010G251000 was further verified using Western blotting of *L. bicolor* treated with *P. trichocarpa* SSPs modified with an HA tag rather than FITC (Fig. 4g). Further, we tested whether scrambling the amino acid sequence of these three proteins would affect their localization. We found that for Potri.009G063200 and CUFF.29946, scrambling the amino acid sequence led to reduced or absent import of the protein into *L. bicolor* hyphae. Scrambled Potri.010G251000 was

Species	Gene name				
	Potri.019G124600	Potri.009G028300	Potri.009G063200	Potri.013G131900	Potri.009G111900
Arabidopsis thaliana	0	3	0	0	0
Arabidopsis lyrata	0	3	0	0	0
Gossypium raimondii	2	5	0	0	0
Theobroma cacao	5	10	0	0	0
Populus trichocarpa	20	18	4	4	6
Glycine max	10	7	0	0	0
Fragaria vesca	1	6	0	0	0
Eucalyptus grandis	0	2	0	0	0
Vitis vinifera	4	5	0	0	0
Solanum tuberosum	3	2	0	0	0
Zea mays	0	0	0	0	0
Sorghum bicolor	0	2	0	0	0
Oryza sativa	0	2	0	0	0
Brachypodium distachyon	0	1	0	0	0
Selaginella moellendorffii	0	0	0	0	0
Physcomitrella patens	0	0	0	0	0

Figure 2. Lineage-specific expansion of small secreted proteins (SSPs) associated with *Populus trichocarpa-Laccaria bicolor* interaction. Due to space limitation, only five genes (Potri.019G124600, Potri.009G028300, Potri.009G063200, Potri.013G131900 and Potri.009G111900) showing lineage-specific expansion are shown in this figure. See Supplementary Tables S4 and S5 for details about the SSP genes showing lineage-specific expansion in *Populus* and closely related plant genera. The phylogenetic tree was adapted from http://phytozome.jgi.doe.gov.

| Gene name	Confidence Group	Protein length, aa	Genomic location	*Populus trichocarpa*		*Populus deltoides*	
				Nisqually-1	93–968	D124	ILL-101
CUFF.22613.1	HC	216	Chr10:21706385–21717909	100	100	24	33
Potri.009G063200.1	HC	69	Chr09:6460994–6461507	100	99	14	33
CUFF.3798.1	LC	38	Chr01:47088970–47089866	99	100	0	0
Potri.004G136500.1	LC	249	Chr04:15714835–15715724	99	93	28	42
Potri.006G212200.1	LC	70	Chr06:22618830–22619398	100	98	13	45
CUFF.21535.1	MC	139	Chr01:48110837–48112147	99	98	15	23

Table 1. SSP genes that have re-sequencing read coverage of >90% in *Populus trichocarpa* ("93–968" and "Nisqually-1") and <50% in *P. deltoides* ("D124" and "ILL-101"). Confidence levels HC, MC and LC represent three computationally-predicted SSP sets: the high confidence set ("SSP_hc set"), the medium confidence set ("SSP_mc set") and the low confidence set ("SSP_lc set"), respectively.

Confidence Level	Secretion Signals Tested	Number Positive for Secretion*	Percentage
HC	16	8	50.0
MC	14	4	28.6
LC	10	3	30.0
Total/Average	40	15	37.5

Table 2. Percentage of *Populus* SSPs that encode secretion signals recognized in *S. cerevisiae*. A subset from each confidence level of predicted *P. trichocarpa* SSPs was tested. Confidence levels HC, MC and LC represent three computationally-predicted SSP sets: the high confidence set ("SSP_hc set"), the medium confidence set ("SSP_mc set") and the low confidence set ("SSP_lc set"), respectively. *See Supplementary Table S6 for detailed results.

still found to be imported into hyphae and localize to the nucleus, but at a reduced level (Fig. 4g). Therefore, some of the *P. trichocarpa* SSPs can enter *L. bicolor* hyphae.

Populus SSPs can alter the development of multiple fungi.
Finally, in pharmacological studies, where *L. bicolor* hyphae cultures were exposed to the four *Populus* SSPs that were able to move into the fungal nuclei, Potri.009G063200 and Potri.010G251000 significantly increased the distance between hyphal branches (p < 0.05; Fig. 5a). This effect was dose dependent (Fig. 5b). Scrambling of the amino acid sequence of both
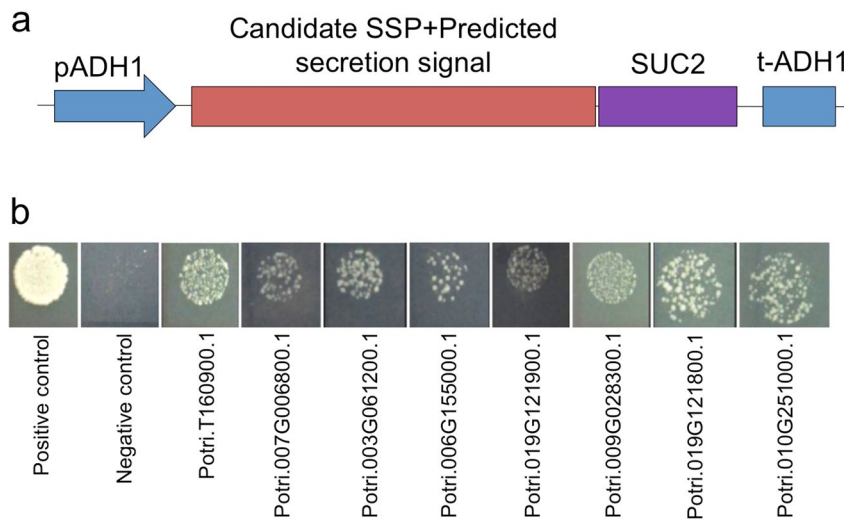
Figure 3. Testing of *Populus trichocarpa* small secreted proteins (SSPs) secretion using a yeast expression system. (**a**) Design of the yeast SUC2-SP vector used for the protein secretion assay[63]. The expression of the candidate *P. trichocarpa* SSP with a positive secretion signal will result in the secretion of the SUC2 protein allowing the growth of yeast on sucrose-containing media. (**b**) Representative examples of positive *P. trichocarpa* SSPs using the yeast secretion assay. The yeast strain growing on medium with glucose supplied was used as a positive control while the strain transformed with empty pGAD-SUC vector was used as a negative control.

Potri.009G063200 and Potri.010G251000 blocked their ability to affect hyphal branching (Fig. 5c). We tested the ability of these two proteins to affect the growth of other ectomycorrhizal fungi; Potri.009G063200 significantly decreased branching distance in *Pisolithus microcarpus* (Boletales) while Potri.010G251000 increased branch distance of *P. microcarpus* and it significantly decreased branching distance in *Suillus granulatus* (Boletales) ($p < 0.05$; Fig. 5d).

L. bicolor hyphae cultures were exposed to the four *Populus* SSPs that were able to move into the fungal nuclei also demonstrated that Potri.009G063200 significantly reduced the growth rate of *L. bicolor* hyphae ($p < 0.05$; Fig. 6a). This effect was dose dependent (Fig. 6b). Scrambling of the amino acid sequence of Potri.009G063200 blocked its effect on fungal growth rate (Fig. 6c). When both Potri.009G063200 and Potri.010G251000 were tested on other fungal species, Potri.009G063200 significantly reduced the growth of the pathogenic Agaricales *Armillaria luteobobulina* and *S. granulatus* (Fig. 6d). Potri.010G251000 increased the growth rate of the ectomycorrhizal *P. microcarpus*.

Discussion

Mutualistic plant-microbe interactions are common among land plants. Interactions with mutualistic fungi and bacteria afford plant hosts enhanced stress tolerance[29–32], increased growth rate and improved reproductive success[33, 34]. Our understanding of the mechanisms used by (1) plants to monitor, encourage and establish these relationships and/or (2) mutualistic organisms to colonize plants is limited. It has recently been established that mutualistic mycorrhizal fungi use SSPs to influence host functioning[7, 8, 35]. In the current study, we used computational predictions and experimental evidence to determine if the model perennial plant *P. trichocarpa* has an analogous genomic complement of effector-like SSPs. Genomic annotation of plant SSPs has lagged behind that of fungal SSPs. To address this limitation, we used a large RNA sequencing data-set sampled from different stages of *P. trichocarpa* root colonization by the mutualistic fungus *L. bicolor* to annotate novel *P. trichocarpa* SSPs. From a pool of differentially expressed transcripts, we identified 417 SSPs in *P. trichocarpa* expressed during mutualistic interaction with the ectomycorrhizal fungus *L. bicolor*. Of these 417 mycorrhiza-induced SSPs, 37% were undocumented in the current *P. trichocarpa* genome annotation v3.0 (www.Phytozome.net).

To functionally validate these proteins, we first showed that several of these small proteins contained verified secretion signals (Fig. 3), demonstrating the ability of these proteins to move across the membrane of the plant host cell. Secondly, we were able to demonstrate that four of six *P. trichocarpa* SSPs tested are capable of entering *L. bicolor* hyphae and localizing to the nuclei (Fig. 4). Interestingly, two of these *P. trichocarpa* SSPs significantly modified hyphal growth (Figs 5 and 6). A biological implication of this is that SSPs may be used to slow growth of fungal hyphae to prevent excessive growth within the root and thereby prevent parasitism. Alternatively, plant-based SSPs could be used to induce a switch in hyphal growth patterns from rhizospheric runner hyphae to aggregated hyphae on the root surface during the initial steps of symbiosis.

Plants may be exposed to hundreds of microbes at any given moment[36–38]. While only a small fraction of these organisms will attempt to colonize plant tissues, the plant must be able to distinguish between beneficial vs. detrimental microbes. The degree to which the plant is able to respond differently to organisms of different lifestyles (e.g., pathogenic vs. mutualistic) or between organisms of the same lifestyle (e.g., mutualistic vs. neutral) is poorly understood. Host species specificity in the transcriptomic response of the host to different microbes has been observed in the interaction between *Vitis vinifera* and *Burkholderia phytofirmans* or *Pseudomonas*
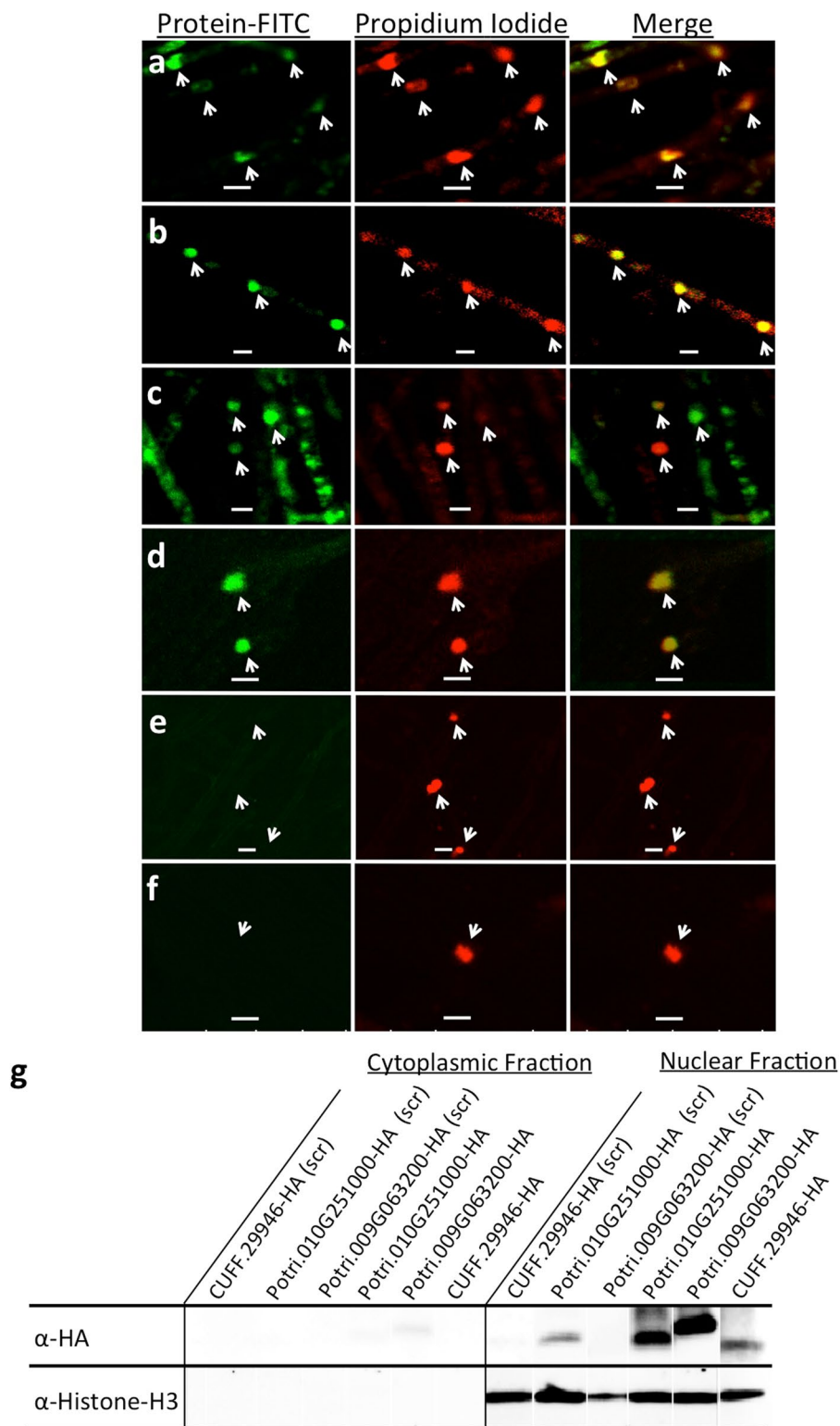
Figure 4. Uptake and sub-cellular localization of *Populus trichocarpa* small secreted proteins (SSPs) into *Laccaria bicolor* hyphae. *L. bicolor* colonies were incubated with 1 μM of either FITC-labelled SSP or FITC alone (control) in PBS (pH 7) for 4 hours after which they were rinsed briefly in PBS, fixed in 4% paraformaldehyde and then stained with 0.1% propidium iodide. Green signal indicates SSP-FITC (mature protein lacking secretion signal) florescent localization, red indicates propidium iodide fluorescence as a nucleus marker. Nuclei are marked by an arrow. (**a**) CUFF.29946; (**b**) Potri.009G063200; (**c**) Potri.010G251000; (**d**) Potri.007G006800; (**e**) Potri.019G121200; (**f**) FITC control; (**g**) Representative Western blot of cytoplasmic and nuclear protein extracted from *L. bicolor* exposed to either scrambled (scr) or native forms of CUFF.29946, Potri.009G063200 and Potri.010G251000 labeled with an HA tag. Probing for histone H3 shows purity of cytoplasmic fraction from nuclear protein.
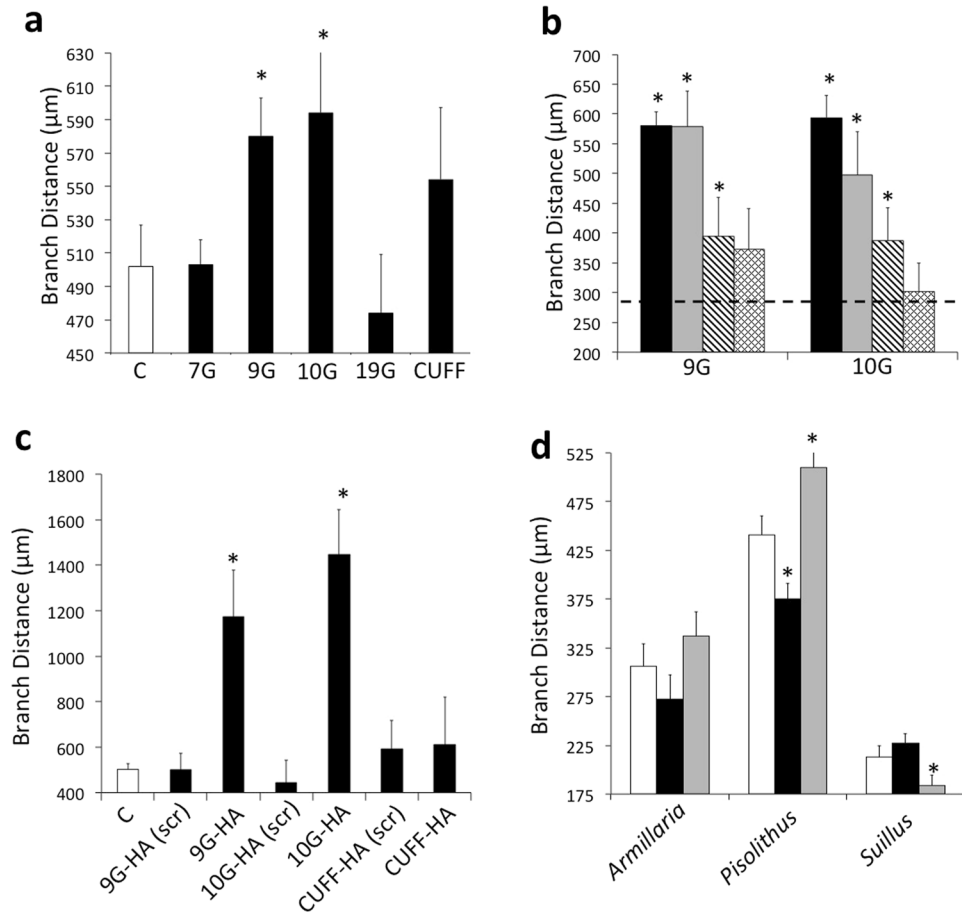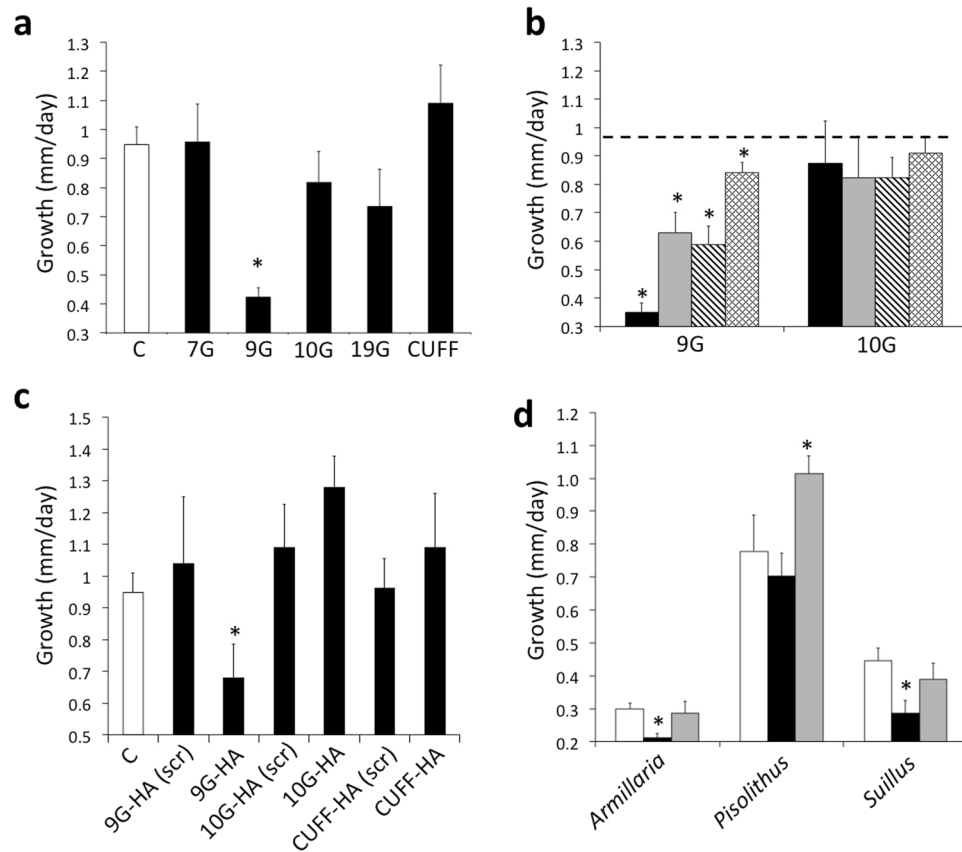
Figure 5. Effect of *Populus trichocarpa* small secreted proteins (SSPs) on hyphal branching. (**a**) Distance between hyphal branches in colonies of *L. bicolor* treated with 1 μM of FITC (C = control), Potri.007G006800 (7G), Potri.009G063200 (9G), Potri.010 G251000 (10G), Potri.019G121200 (19G) or with CUFF.29946.1 (CUFF). (**b**) Distance between hyphal branches in colonies of *L. bicolor* treated with 1 μM (black bars), 0.1 μM (grey bars), 0.01 μM (diagonal striped bars) or 0.001 μM (diamond pattern bars) Potri.009G063200 (9G) or Potri.010G251000 (10G) as compared to control treatment (dashed line, 1% DMSO). (**c**) Distance between hyphal branches in colonies of *L. bicolor* treated with 1 μM of either scrambled (scr) or native forms of Potri.009G063200 and Potri.010G251000 labeled with an HA tag as compared to control treatment ('C', 1% DMSO). (**d**) Impact of Potri.009G063200 (black bars) and Potri.010G25100 (grey bars) as opposed to control treatment (white bars) on the distance between hyphal branches of three different fungal species (*Armillaria luteobobulina*, *Pisolithus microcarpus*, *Suillus granulatus*) ± SE, *Significant difference from control (p < 0.05; Student's t-test).

syringae[39], between *Arabidopsis thaliana* and *Trichoderma asperelloides* or *P. syringae*[40], between *Hordeum vulgare* and *Piriformospora indica* or *Blumeria graminis*[41] and between *A. thaliana* and various viruses[42], suggesting that the plant is able to tailor its SSP transcriptomic response to different lifestyles of microbes.

Effector SSPs are a hallmark in the genomes of microbes that associate with plants[1]. These proteins are typically classified as ≤250 aa in length, are significantly regulated during symbiosis and coded by genes that show genomic evidence of rapid evolution[43, 44]. Upon secretion, these fungal effectors may remain external to plant cells or they may enter the plant cell. Regardless of their final localization, these proteins alter the functioning of the plant cell and their expression is generally considered crucial for the successful colonization of plant tissues[6, 35, 45, 46]. The discovery, described herein, of plant-encoded mycorrhiza-induced SSPs that are: (1) ≤250 aa in size, (2) lineage- or genera-specific, (3) highly expressed during symbiosis, (4) able to enter fungal hyphae and localize to the nucleus and (5) affecting fungal morphology strongly support the conclusion that plants encode proteins that are effector-like. This inference is supported by previous work in the legume:rhizobial interaction where small plant proteins have been reported to enter the symbiotic bacterium and affect its growth[17-21]. *Populus* can respond to mutualistic fungi by secreting their own effector-like proteins that might parallel the role of fungal effectors (Fig. 7). Our observation of effector SSPs in *Populus* implies a novel avenue, apart from the traditional immune response[44], by which plants communicate with (or control) their mutualistic microbial partners.

Our data further supports previous hypotheses which has suggested that plants are not silent, or 'naïve', participants in mutualistic relationships by demonstrating that they too have the genetic tools in the form of effector-like proteins to communicate with and influence their microbial partners. These data are the first steps forward in

Figure 6. Effect of *Populus trichocarpa* small secreted proteins (SSPs) on hyphal growth rate. (**a**) Growth rate of *L. bicolor* hyphae treated with 1 μM of FITC (C = control), Potri.007G006800 (7G), Potri.009G063200 (9G), Potri.010G251000 (10G), Potri.019G121200 (19G) or with CUFF.29946.1 (CUFF). (**b**) Growth rate of *L. bicolor* hyphae treated with 1 μM (black bars), 0.1 μM (grey bars), 0.01 μM (diagonal striped bars) or 0.001 μM (diamond pattern bars) Potri.009G063200 (9G) or Potri.010G251000 (10G) as compared to control treatment (dashed line, 1% DMSO). (**c**) Growth rate of *L. bicolor* hyphae treated with 1 μM of either scrambled (scr) or native forms of Potri.009G063200 and Potri.010G251000 labeled with an HA tag as compared to control treatment ('C', 1% DMSO). (**d**) Impact of Potri.009G063200 (black bars) and Potri.010G25100 (grey bars) as opposed to control treatment (white bars) on the growth rate of hyphae of three different fungal species (*Armillaria luteobobulina*, *Pisolithus microcarpus*, *Suillus granulatus*) ± SE, *Significant difference from control (p < 0.05; Student's t-test).

understanding the role of plant SSPs as effectors in influencing symbiotic relationships. As more transcriptomic and genomic data related to SSPs becomes available, a more complete description of plasticity of plant responses to microbes will be possible. This research highlights the importance of the cross-talk between the plant and fungal effectors. Future studies on the fine-tuned balance between the number and level of the various released effectors could provide deeper insights into the development and the fitness of symbiosis.

Methods

Plant material, RNA extraction and sequencing. For colonization experiments, *L. bicolor* (Maire) P.D. Orton isolate S238N inoculum was grown on a substrate of peat moss:vermiculite (3:1) containing liquid Pachlewski medium for 3 months[47, 48]. Dormant hardwood stems from *Populus trichocarpa* clone 101–74 were collected from stool beds in January. Dormant stems were then cut into 25-cm cuttings with six to seven nodes per cutting. These cuttings were sealed in polyethylene bags and kept in cold storage (0 °C) until they were removed for rooting. The experiments in which these cuttings were used were then performed during the following year. Three node stem cuttings of *P. trichocarpa* clone 101–74 were pre-rooted for 1 week in a solution of 2.5 mM KNO$_3$, 0.8 mM KH$_2$PO$_4$, 1 mM MgSO$_4$·7 H$_2$O, 2.3 mM Ca(NO$_3$)$_2$·4 H$_2$O, 23 μM H$_3$BO$_3$, 4.6 μM MnCl$_2$·4 H$_2$O, 0.4 μM ZnSO$_4$·7 H$_2$O, 0.09 μM (NH$_4$)$_2$ MoO$_4$, 0.18 μM CuSO$_4$·5 H$_2$O, 20 μM FeNa·EDTA, pH 5.8 and then planted in a 1 L-pot containing a 9:1 mixture (v:v) of Terra-Green to *L. bicolor* inoculum as previously described[49]. The colonization experiment was replicated twice (technical replicate) under climate controlled greenhouse conditions maintaining a 16-hr photoperiod at 22 °C.

Tissues used for transcriptomic profiling were sampled from fine roots of *P. trichocarpa* between 10 am and 2 pm each day to avoid confounding effects of circadian rhythm on gene expression. Plant host roots, grown in the same substrate and with the same nutrient regime but without fungal inoculum, were harvested as a non-mycorrhizal control for transcriptomic analyses. Three biological replicates of between 50–100 mg of
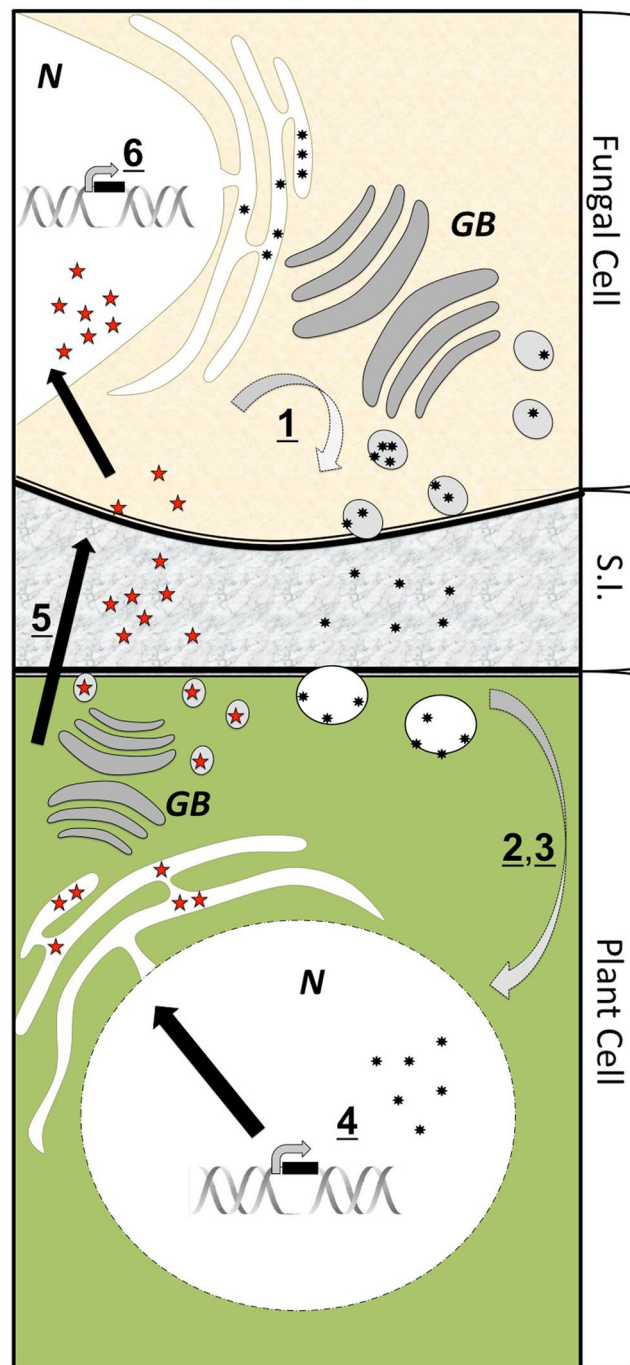
Figure 7. Proposed model of reciprocal small secreted proteins (SSP) feedback between ectomycorrhizal fungi and their host plants. During colonization of receptive plant hosts, fungal hyphae (brown cell) produce effector-like SSPs (black stars) that are secreted from the hyphal cell (#1). A portion of these proteins cross the symbiotic interface (grey shaded area) and enter the plant cell (green cell; #2) where they localize to different sub-cellular compartments (#3). These effectors may induce a change in the transcription of host genes (#4; adapted from Plett *et al.*[35]). Plant cells, in turn, produce small effector-like SSPs (red stars) that are secreted, cross the symbiotic interface (#5) where some may enter the hyphal cell and localize to different sub-cellular compartments. Some effectors may induce a change in the transcription of host genes (#6). GB = Golgi bodies; N = nucleus; S.I. = symbiotic interface. Note: we do not currently know the priority of this signaling network (i.e. if the fungus signals and the plant responds or vice-versa). Rather, numbering here is used to enable description of the different steps in a feedback system that is likely reciprocal in nature.

colonized root tips (equivalent to approximately 100 colonized root tips) or of uncolonized lateral roots (control) were harvested at 2, 4 and 12 weeks after *L. bicolor* inoculation, washed thoroughly, snap frozen in liquid nitrogen

and stored at −80 °C until total RNA extraction was performed using the RNAeasy kit (Qiagen) per the manufacturer's instructions. An on-column DNA digestion step with DNAse I (Qiagen) was also included to avoid DNA contamination. RNA quality was verified using Experion HighSens capillary gels (Bio-Rad). For *P. trichocarpa* RNA sequencing (RNA-seq), three biological replicates of colonized or un-colonized tissues were pooled for sequencing. As mycorrhizal tissues are very recalcitrant to RNA extraction, quantities of RNA recovered were too low for quantities needed for RNA-seq analysis. Therefore, we amplified 250 ng total RNA using the Clonetech SMARTer amplification kit (Ozyme, St Quentin en Yvelines, France), according to manufacturer's instructions and being sure to maintain the amplification in the logarithmic phase such that RNA pools were still proportional to original quantities of transcript as described previously[49]. Each pooled sample was sequenced (36-bp single-end reads) using Illumina GAII technology by IGA Technology services (Udine, Italy; GEO Accession Series GSE54789). Also, 100-bp single-end sequencing was also performed using Illumina HiSeq2000 platform (Beckman Coulter Genomics, Danvers, MA, USA) for samples collected 12 weeks after *Laccaria* inoculation (GEO Accession Series GSE75437).

RNA-seq data analysis and prediction of *P. trichocarpa* small secreted proteins. The computational pipeline for RNA-seq data analysis and prediction of *P. trichocarpa* SSPs is illustrated in Fig. 1a. RNA-seq reads were mapped to the *P. trichocarpa* genome[50] using TopHat[51] and novel transcripts were discovered using Cufflinks[52, 53]. Differentially expressed genes between treatment (i.e., *P. trichocarpa* roots with *L. bicolor* ectomycorrhizas) and control (i.e., *P. trichocarpa* roots without *L. bicolor* inoculation) were identified using Cuffdiff[52]. The open reading frames (ORFs) were annotated using six-frame translation based on standard genetic code with a length range of 10–10,000 aa. The best ORF for each transcript was chosen based on those ORFs with the highest score in a BLASTp search using the default setting against the UniRef90 database (http://www.uniprot.org) or the longest ORF if there were no BLASTp hits.

The subcellular localizations of proteins were predicted using three different methods: (1) LocTree2[25], (2) CELLO[26] and (3) YLoc[27] using default parameters. Proteins with full-length sequences of ≤250 aa were defined as small proteins (SmPs). The predicted small secreted proteins (SSPs) were divided into three sets: high-confidence set (SSP_hc) including proteins with subcellular localization as "secreted" predicted by method 1 and "extracellular space" by both method 2 and method 3; medium-confidence set (SSP_mc) proteins with subcellular localization as "secreted" predicted by method 1 and "extracellular space" by method 2 or method 3; and low-confidence set (SSP_lc) proteins with subcellular localization as "secreted" predicted by method 1, but no "extracellular" prediction by either method 2 or method 3.

Protein domain and Gene Ontology analysis. A protein-domain search for all SSP sequences against the Pfam 28.0 database[54] was performed with an E-value cutoff of 1e-10. Gene Ontology (GO) annotation was performed using Blast2GO with a BLASTp E-value hit filter of 1e-6, an annotation cutoff value of 55, and GO weight of 5[55]. GO enrichment analysis was performed using BiNGO[56].

Identification of homologs of *Populus* SSP genes in other plant species. The representative protein sequences (i.e., the longest protein sequence in case of multiple transcripts annotated for one gene locus) in 16 species (*Arabidopsis thaliana*, *Arabidopsis lyrata*, *Gossypium raimondii*, *Theobroma cacao*, *Populus trichocarpa*, *Glycine max*, *Fragaria vesca*, *Eucalyptus grandis*, *Vitis vinifera*, *Solanum tuberosum*, *Zea mays*, *Sorghum bicolor*, *Oryza sativa*, *Brachypodium distachyon*, *Selaginella moellendorffii*, *Physcomitrella patens*) were downloaded from Phytozome v9.0 (www.Phytozome.net). The homologs were identified using BLASTp[57] with e-value cutoff of 1e-10.

Comparative analysis of SSP genes among different *Populus* genotypes. *Populus* genome resequencing data[58, 59] was used to reveal genomic differences in the region of selected genes. Two *P. trichocarpa* genotypes (Nisqually-1 and 93–968) and two *P. deltoides* genotypes (ILL-101 and D124) were selected for the analysis. The Illumina reads were aligned to the reference *P. trichocarpa* genome sequence[50] to generate bam files using the MAQ software package[60]. Per base sequence coverage for mapped reads in the region corresponding to the selected genes were obtained using bedtools[61].

Testing of protein secretion using a yeast expression system. To construct a Gateway compatible vector for yeast secretion assay, a DNA fragment containing the *SUC* gene without signal peptide was synthesized by GeneScript (Piscataway, NJ, USA) and validated by sequencing. The DNA fragment was ligated into pGAD-424 vector by S*phI*. Candidate SSPs were cloned into the secretion vector, with the SSP gene fused with the N-terminus of *SUC* gene, by Gateway cloning and transformed into the *suc2* yeast mutant (strain ATCC-96100). Positive transformants were selected and validated on a synthetic dropout medium. Transformants were assayed on plates with glucose (10 mM) and sucrose (10 mM) as carbon supplies respectively and growth activity was measured after growing 2–3 days at 28 °C.

L. *bicolo*r uptake of *P. trichocarpa* SSPs and growth analysis tests. In order to determine if *P. trichocarpa* SSPs could cross the *L. bicolor* hyphal membrane, five proteins (CUFF.29946, Potri.009G063200, Potri.010G251000, Potri.019G121200 and Potri.007G006800) were produced by cell-free recombinant peptide synthesis (Eurogentec) without the secretion signal and amino-terminal covalent addition of FITC using standard amino acid coupling techniques or with the addition of an HA tag (C-terminal addition of the following amino acid sequence: YPYDVPDYA). CUFF.29946, Potri.009G063200 and Potri.010G251000 were also produced in this method as a 'scrambled' (scr) version whereby the amino acids were randomly re-ordered and synthesized in a new sequence. Potri.019G121200 was used as a negative control, which contains a FASCICLIN-LIKE ARABINOGALACTAN PROTEIN 12 domain (www.phytozome.net), and it was predicted to be a cell adhesion

protein and was not expected to be imported by fungal hyphae. The peptides were purified using HPLC fractionation to ensure the purity of the synthesis product. A control of the FITC fluorophore or of dilution buffer was also used throughout the experiments to ensure that this tag nor the buffer were altering our results.

To test subcellular localizations of these *P. trichocarpa* SSPs in *L. bicolor* hyphae and the role of these proteins in affecting fungal morphology (for *L. bicolor, P. microcarpus, A. luteobobulina, S. granulatus*), fungal colonies grown on cellophane membranes on ½ MMN agar plates at 25 °C until the colonies were 2–3 cm in diameter (2 weeks). Five biological replicates (i.e., independent colonies) were treated with an individual synthetic protein or FITC or buffer alone (controls) by dripping 10 μL of a solution of the protein (concentrations varying between 1 μM and 0.001 μM as outlined in the text) on the leading edge of the fungal colony. This treatment was repeated daily for four days after which the growth of the leading fungal edge and the hyphal branching were measured using ImageJ. The protein, FITC or buffer solutions were produced fresh daily and never re-used to avoid degradation of the protein.

For the localization experiment, the same aged *L. bicolor* colonies were treated by placing a 10 μL drop of either 1 μM protein or 1 μM FITC (PBS, pH7) on the leading edge of an in-tact, un-perturbed fungal colony grown as described above. After 4 hour of incubation, colonies were washed in excess PBS two times for five minutes each. The treated hyphae were then cut away from the rest of the colony and were fixed overnight in 4% (w/v) paraformaldehyde at 4 °C. Post fixing, samples were washed 3 × 10 min in PBS and then stained in 0.1% propidium iodide for 20 min then imaged using an inverted Leica TCS SP5 laser scanning confocal microscope. Fluorescent excitation was achieved using a 488 nm beam (10% power) for FITC and a 561 nm beam for propidium iodide. Emission was recovered between 510–530 nm for FITC and between 610–630 nm for propidium iodide. To corroborate fluorescent cell entry tests, *L. bicolor* colonies were exposed to HA-tagged proteins as above. Rather than fixing the samples, they were ground in liquid nitrogen and cytoplasmic and nuclear proteins were extracted as per Wang and colleagues[62]. Equal quantities of total protein were separated on a 1D Bio-Rad gradient gel (4–20%) and then transferred to a PVDF membrane. Membranes were blocked and then probed with either α-HA (to detect the presence of the SSP) or α-Histone H3 primary antibody (to ascertain the purity of the cytoplasmic vs. nuclear protein fraction). Bands were developed using the Clarity ECL Blotting substrate (Bio-Rad) as per manufacturer's instructions and chemiluminescent images were taken using the Bio-Rad ChemiDoc Imaging system with an exposure of between 15–20 s.

References

1. Martin, F. & Kamoun, S. Effectors in plant-microbe interactions. (John Wiley & Sons, 2011).
2. Bos, J. I. *et al.* Phytophthora infestans effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 9909–9914 (2010).
3. Bos, J. I. *et al.* The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*. *Plant Journal* **48**, 165–176 (2006).
4. Kleemann, J. *et al.* Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen *Colletotrichum higginsianum*. *PLoS Pathogens* **8**, e1002643 (2012).
5. Kemen, E. *et al.* Identification of a protein from rust fungi transferred from *Haustoria* into infected plant cells. *Molecular Plant-Microbe Interactions* **18**, 1130–1139 (2005).
6. Tanaka, S. *et al.* A secreted *Ustilago maydis* effector promotes virulence by targeting anthocyanin biosynthesis in maize. *eLife* **3**, e01355 (2014).
7. Plett, J. M. *et al.* Effector MiSSP7 of the mutualistic fungus *Laccaria bicolor* stabilizes the *Populus* JAZ6 protein and represses jasmonic acid (JA) responsive genes. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8299–8304 (2014).
8. Kloppholz, S., Kuhn, H. & Requena, N. A secreted fungal effector of *Glomus intraradices* promotes symbiotic biotrophy. *Current Biology* **21**, 1204–1209 (2011).
9. Lebeis, S. L. *et al.* Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349**, 860–864 (2015).
10. de Weert, S. *et al.* Flagella-driven chemotaxis towards exudate components is an important trait for tomato root colonization by *Pseudomonas fluorescens*. *Molecular Plant-Microbe Interactions* **15**, 1173–1180 (2002).
11. Akiyama, K., Matsuzaki, K. & Hayashi, H. Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature* **435**, 824–827 (2005).
12. Kneer, R., Poulev, A. A., Olesinski, A. & Raskin, I. Characterization of the elicitor-induced biosynthesis and secretion of genistein from roots of *Lupinus luteus* L. *Journal of Experimental Botany* **50**, 1553–1559 (1999).
13. Maróti, G., Kereszt, A., Kondorosi, E. & Mergaert, P. Natural roles of antimicrobial peptides in microbes, plants and animals. *Research in Microbiology* **162**, 363–374 (2011).
14. De-la-Peña, C., Lei, Z., Watson, B. S., Sumner, L. W. & Vivanco, J. M. Root-microbe communication through protein secretion. *Journal of Biological Chemistry* **283**, 25247–25255 (2008).
15. De-la-Pena, C. & M. Loyola-Vargas, V. The hidden chemical cross-talk between roots and microbes: A proteomic approach. *Current Proteomics* **9**, 103–117 (2012).
16. Sagaram, U. S. *et al.* Structural and functional studies of a phosphatidic acid-binding antifungal plant defensin MtDef4: identification of an RGFRRR motif governing fungal cell entry. *PLoS One* **8**, e82485 (2013).
17. Farkas, A. *et al. Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 5183–5188 (2014).
18. Penterman, J. *et al.* Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 3561–3566 (2014).
19. Van de Velde, W. *et al.* Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **327**, 1122–1126 (2010).
20. Wang, D. *et al.* A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* **327**, 1126–1129 (2010).
21. Durgo, H. *et al.* Identification of nodule-specific cysteine-rich plant peptides in endosymbiotic bacteria. *Proteomics* **15**, 2291–2295 (2015).
22. Yang, X. *et al.* Discovery and annotation of small proteins using genomics, proteomics, and computational approaches. *Genome Research* **21**, 634–641 (2011).
23. Ghorbani, S. *et al.* Expanding the repertoire of secretory peptides controlling root development with comparative genome analysis and functional assays. *Journal of Experimental Botany* **66**, 5257–5269 (2015).
24. Hanada, K. *et al.* Small open reading frames associated with morphogenesis are hidden in plant genomes. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 2395–2400 (2013).

25. Goldberg, T., Hamp, T. & Rost, B. LocTree2 predicts localization for all domains of life. *Bioinformatics* **28**, i458–i465 (2012).

26. Yu, C. S., Chen, Y. C., Lu, C. H. & Hwang, J. K. Prediction of protein subcellular localization. *Proteins: Structure, Function, and Bioinformatics* **64**, 643–651 (2006).

27. Briesemeister, S., Rahnenfuhrer, J. & Kohlbacher, O. YLoc–an interpretable web server for predicting subcellular localization. *Nucleic Acids Research* **38**, W497–W502 (2010).

28. Labbé, J. *et al.* Identification of quantitative trait loci affecting ectomycorrhizal symbiosis in an interspecific F1 poplar cross and differential expression of genes in ectomycorrhizas of the two parents: *Populus deltoides* and *Populus trichocarpa. Tree Genetics & Genomes* **7**, 617–627 (2011).

29. Colpaert, J., Wevers, J. L., Krznaric, E. & Adriaensen, K. How metal-tolerant ecotypes of ectomycorrhizal fungi protect plants from heavy metal pollution. *Annals of Forest Science* **68**, 17–24 (2011).

30. Li, J. *et al. Paxillus involutus* strains MAJ and NAU mediate K^+/Na^+ homeostasis in ectomycorrhizal *Populus ×* cane*scens* under sodium chloride stress. *Plant Physiology* **159**, 1771–1786 (2012).

31. Pena, R. & Polle, A. Attributing functions to ectomycorrhizal fungal identities in assemblages for nitrogen acquisition under stress. *The ISME journal* **8**, 321–330 (2014).

32. Rooney, D. *et al.* Effect of arbuscular mycorrhizal colonisation on the growth and phosphorus nutrition of *Populus euramericana* cv Ghoy. *Biomass and Bioenergy* **35**, 4605–4612 (2011).

33. Ceballos, I. *et al.* The *in vitro* mass-produced model mycorrhizal fungus, *Rhizophagus irregularis*, significantly increases yields of the globally important food security crop cassava. *PLoS One* **8**, e70633 (2013).

34. Erman, M. *et al.* Effects of *Rhizobium*, arbuscular mycorrhiza and whey applications on some properties in chickpea (*Cicer arietinum* L.) under irrigated and rainfed conditions 1-Yield, yield components, nodulation and AMF colonization. *Field Crops Research* **122**, 14–24 (2011).

35. Plett, J. M. *et al.* A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. *Current Biology* **21**, 1197–1203 (2011).

36. Berendsen, R. L., Pieterse, C. M. & Bakker, P. A. The rhizosphere microbiome and plant health. *Trends in Plant Science* **17**, 478–486 (2012).

37. Hortal, S. *et al.* Beech roots are simultaneously colonized by multiple genets of the ectomycorrhizal fungus *Laccaria amethystina* clustered in two genetic groups. *Molecular Ecology* **21**, 2116–2129 (2012).

38. Lau, J. A. & Lennon, J. T. Evolutionary ecology of plant–microbe interactions: soil microbial structure alters selection on plant traits. *New Phytologist* **192**, 215–224 (2011).

39. Bordiec, S. *et al.* Comparative analysis of defence responses induced by the endophytic plant growth-promoting rhizobacterium *Burkholderia phytofirmans* strain PsJN and the non-host bacterium *Pseudomonas syringae* pv. *pisi* in grapevine cell suspensions. *Journal of Experimental Botany* **62**, 595–603 (2011).

40. Brotman, Y. *et al.* Transcript and metabolite analysis of the *Trichoderma*-induced systemic resistance response to *Pseudomonas syringae* in *Arabidopsis thaliana. Microbiology* **158**, 139–146 (2012).

41. Molitor, A. *et al.* Barley leaf transcriptome and metabolite analysis reveals new aspects of compatibility and *Piriformospora indica*-mediated systemic induced resistance to powdery mildew. *Molecular Plant-Microbe Interactions* **24**, 1427–1439 (2011).

42. Postnikova, O. A. & Nemchinov, L. G. Comparative analysis of microarray data in *Arabidopsis* transcriptome during compatible interactions with plant viruses. *Virology Journal* **9**, 101 (2012).

43. Jiang, R. H. Y., Tripathy, S., Govers, F. & Tyler, B. M. RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4874–4879 (2008).

44. Jones, J. D. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).

45. Doehlemann, G. *et al.* Pep1, a secreted effector protein of *Ustilago maydis*, is required for successful invasion of plant cells. *PLoS Pathogens* **5**, e1000290 (2009).

46. Grant, S. R., Fisher, E. J., Chang, J. H., Mole, B. M. & Dangl, J. L. Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria. *Annual Review of Microbiology* **60**, 425–449 (2006).

47. Pachlewski, R. & Pachlewska, J. Studies on symbiotic properties of mycorrhizal fungi of pine (*Pinus silvestris* L.) with the aid of the method of mycorrhizal synthesis in pure cultures on agar. (Forest Research Institute, Warsaw, Poland; 1974).

48. Mortier, F., Le Tacon, F. & Garbaye, J. Effect of inoculum type and inoculation dose on ectomycorrhizal development, root necrosis and growth of Douglas fir seedlings inoculated with *Laccaria laccata* in a nursery. *Annales des Sciences Forestières* **45**, 301–310 (1988).

49. Plett, J. M. *et al.* The mutualist *Laccaria bicolor* expresses a core gene regulon during the colonization of diverse host plants and a variable regulon to counteract host-specific defenses. *Molecular Plant-Microbe Interactions* **28**, 261–273 (2015).

50. Tuskan, G. A. *et al.* The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596–1604 (2006).

51. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).

52. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562–578 (2012).

53. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28**, 511–515 (2010).

54. Finn, R. D. *et al.* Pfam: the protein families database. *Nucleic Acids Research* **42**, D222–D230 (2014).

55. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).

56. Maere, S., Heymans, K. & Kuiper, M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449 (2005).

57. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

58. Evans, L. M. *et al.* Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nature Genetics* **46**, 1089–1096 (2014).

59. Slavov, G. T. *et al.* Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa. New Phytologist* **196**, 713–725 (2012).

60. Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Research* **18**, 1851–1858 (2008).

61. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

62. Wang, W. *et al.* An importin β protein negatively regulates microRNA activity in *Arabidopsis. Plant Cell* **23**, 3565–3576 (2011).

63. Lee, S.-J., Kim, B.-D. & Rose, J. K. C. Identification of eukaryotic secreted and cell surface proteins using the yeast secretion trap screen. *Nature Protocols* **1**, 2439–2447 (2006).

Acknowledgements

Author Contributions

Jonathan M. Plett, Xiaohan Yang, Gerald A. Tuskan and Francis Martin designed research. Jonathan M. Plett and Annegret Kohler performed RNA sequencing. Ting Li, Priya Ranjan, Hao-Bo Guo and Xiaohan Yang performed computational analysis. Jonathan M. Plett, Hengfu Yin, Ritesh Mewalal, Rongbin Hu, Sara Jawdy, Chun Ju Chen, Henrique C. De Paoli, George Butler and Tessa Maureen Burch-Smith performed experimental analysis. Jonathan M. Plett, Hengfu Yin and Ritesh Mewalal wrote the paper. Ian C. Anderson and Jessy L. Labbé, Xiaohan Yang, Gerald A. Tuskan and Francis Martin discussed the results and edited the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-00400-8

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.