https://doi.org/10.1038/s44383-026-00026-8

An endosymbiosis in plant shoots based on an ankyrin repeat protein

Namrata Baruah[1], Janne J. Koskimäki[1,2], Habibollah Mohammad Parast Tabas[1], Xavier Rey-Velasco[1], A. Carolin Frank[2] & Anna Maria Pirttilä[1] ✉

Chemical-based, modern agriculture is harmful to the environment, consuming fossil fuels, producing eutrophication, killing pollinators, and threatening human health. Because sustainable alternatives based on soil- or root-associated microorganisms are unreliable, new tools are necessary. *Methylorubrum extorquens* DSM13060 colonizes plant meristematic cells and aggregates around nuclei. The endosymbiont promotes plant growth without producing hormones but encodes putative nucleomodulins. We examined the role of an ankyrin-repeat protein (Ank) in the endosymbiosis. Ank interacted with 46 host proteins, expressed mainly in the cytoplasm and nucleus. Ank had a viral motif and interacted with defense-associated proteins. Deletion strain, Δ*ank*, lacked the capacity for colonization, indicating that interaction of Ank with plant proteins is needed. The plant-growth promotion capacity was lost in Δ*ank*, confirming the role of endosymbiosis behind this effect. Our results characterize Ank as a key that unlocks the beneficial traits of endosymbiosis, portraying a new mechanism of plant-associated bacteria for sustainable agriculture.

Currently, agricultural practices are almost equally harmful to the environment as any anthropogenic process. Monocultures of plants are treated with synthetic fertilizers, which are produced by fossil fuel consumption[1], mined from un-renewable sources[2], and create eutrophication after leaching to the environment[3]. Plant diseases are averted by chemical pesticides, which kill pollinators and pose a threat to human health[4]. Environmentally friendly alternatives, used in organic agriculture, are mainly based on microorganisms. Bacteria and fungi that provide beneficial outcomes in crop plant growth and health, biofertilizers and biocontrol agents, respectively, are available as products for agriculture, based on plant-growth-promoting rhizobacteria (PGPR), as well as soil, epiphytic, and mycorrhizal fungi[5]. However, their use is not straightforward, faced by challenges, such as the inconsistency of field performance[6]. Therefore, there is a dire need for sustainable, reliable, and effective biofertilizers and biocontrol agents. Plant internal microorganisms, endophytes, represent a more intimate symbiotic interaction with potentially more consistent effects on the host plant[5]. Specific members of endophytes are even capable of entering plant cells[7], but their mechanism of growth promotion has been elusive.

Besides the well-known root endosymbionts, rhizobia[8] and arbuscular mycorrhiza[9], intracellular infection of plants is rarely reported[10–12], especially in the plant shoots. The best studied shoot endosymbiont, *Methylorubrum extorquens* DSM13060 (Rhizobiales), colonizes the interior of plant cells[7,13] and positively affects plant growth[14]. The endosymbiont lacks the typical mechanisms of plant growth promotion, such as biosynthesis of plant hormones or providing nutrients for the host[13–15]. *M. extorquens* DSM13060 therefore greatly differs from the epiphytic *Methylobacterium* or *Methylorubrum* strains, which typically produce the plant hormone cytokinin, but the importance of cytokinin biosynthesis for plant growth has been questioned[16]. In contrast to other endophytes[10–12], *M. extorquens* DSM13060 colonizes the plant without damaging host tissue[13,17], and the genome is scarce in genes for plant cell-wall degradation[13]. Such an unusual lifestyle raises questions on how the bacterium benevolently colonizes the host cytoplasm and induces plant growth.

The colonization mechanism of *M. extorquens* DSM13060 shares similarities with stem-colonizing rhizobia[18]. In the symbiosis between rhizobia and legumes, the bacteria typically produce signaling molecules known as Nod-factors to access root cells, although exceptions of plant-rhizobia interactions without Nod signaling exist[19,20]. When present, the recognition of Nod-factors by the plant host leads to polarized root-hair tip growth and invagination of the plasma membrane for bacterial entry, leading to cell divisions in the cortex and nodule formation[21]. Whereas the colonization mechanism of *M. extorquens* DSM13060 does not involve root hair deformation[13,17], the endosymbiont forms infection pockets upon entry and uses infection thread-like mechanisms to advance in the plant tissue[18]. The genome of *M. extorquens* DSM13060 carries *Nod-like* genes, but Nod-factors have not been identified[13,15]. Once the *M. extorquens* DSM13060 cells have colonized the pine host intracellularly, they establish themselves in close contact with plant nuclei[13].

[1]Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland. [2]Environmental Systems, Department of Life and Environmental Science, University of California Merced, Merced, CA, USA. ✉e-mail: am.pirttila@oulu.fi

In general, bacteria that colonize hosts intracellularly often encode effector proteins with eukaryotic domains that interact with host proteins and processes[22]. Effector proteins that target nuclear processes, such as transcription, chromatin dynamics, and histone modification, are referred to as nucleomodulins[23,24], and they are typically delivered to the host cell through a secretion system, such as the type IV secretion system (T4SS)[13,25] Because our earlier work identified both putative nucleomodulins and a T4SS secretion system in the *M. extorquens* DSM13060 genome[13], we hypothesized that these proteins are involved in the intracellular colonization of host cells. Therefore, we bioinformatically predicted T4SS effectors in the *M. extorquens* DSM13060 genome, and based on the analysis, selected an ankyrin repeat protein (Ank) for further studies. We analyzed the structure and interaction targets of the Ank protein, generated a single-gene deletion, tracked colonization of the host, and tested plant growth promotion capacity of the deletion strain.

Results

A search algorithm for type IV effector proteins, S4TE 2.0, was used to predict nucleomodulins delivered to the host cell through the T4SS. The S4TE predicts and ranks the candidate T4SS effectors by using a combination of 14 sequence features, including homology to previously identified effectors, homology to eukaryotic domains, presence of secretion signals or subcellular localization signals. S4TE 2.0 is designed to evolve rapidly with the publication of new experimentally validated T4Es, which will reinforce the predictive power of the algorithm. Computational methodology can be used to identify a wide spectrum of candidate bacterial effectors that lack sequence conservation but have similar amino acid characteristics. The predicted T4SS-secreted proteins are listed in Table S1. Specifically, the S4TE 2.0 score predicted that an ankyrin repeat protein (Ank) of *M. extorquens* DSM13060 is translocated to the host cell through T4SS with a high probability (Table S1).

We then generated a single-gene deletion of the *Ank* gene, and the growth curve of the deletion strain Δ*ank* was evaluated in normal growth conditions and compared with *M. extorquens* 13061 that carries the Green Fluorescent Protein marker gene in the genome[14]. There were no differences, analyzed visually and statistically (not shown). The Δ*ank* deletion strain was identical to the 13061.

Δ*ank* lacks the host colonization potential

In the early process (45 days-post-inoculation, dpi) of colonization, *M. extorquens* 13061 (Fig. 1A-B) had colonized through the root endodermal layer, residing inside the host cells and accumulating around the plant nuclei (Fig. 1B). At this point, only a low number of Δ*ank* cells (Fig. 1C-D) colonized on the root epidermis (Fig. 1C-D). At 90 dpi, the density of 13061 bacteria was high on the epidermis, and numerous cells colonized the non-vascular parenchyma (Fig. 1E-F). In contrast, Δ*ank* had few bacterial cells colonizing the surface of root epidermis (Fig. 1G), and sporadic cells in the cortex (Fig. 1H). At 110 dpi, when bacteria of the 13061-strain were regularly seen accumulating around the nuclei of root cells (Fig. 1I-J), Δ*ank* mainly infected the root epidermal cells (Fig. 1K-L). At the last stage of colonization (150 dpi), the 13061-strain was colonizing in high numbers in the cells of the cortex (Fig. 1M-N), whereas Δ*ank* colonization had remained stagnant, with bacteria forming microcolonies on the epidermis and sporadic cells having entered the root tissues (Fig. 1O-P).

Differences between the 13061-strain and Δ*ank* were more pronounced in the colonization of the shoots. At 90 dpi (Fig. 1Q-R), the 13061-strain had established a dense biofilm on the epidermis, but Δ*ank* was unable to colonize the surface of the epidermis (Fig. 1S-T). At 150 dpi, the 13061-strain colonized the shoot tissues from the cortical layer to the non-vascular parenchyma and xylem vessels (Fig. 1U-V). However, Δ*ank* was still largely absent in the shoots at 150 dpi (Fig. 1X-Y). In contrast to Δ*ank*, when a unique endopolygalacturonase-like gene (*CellW*) of *M. extorquens* DSM13060[13] was deleted and tested for host colonization, the deletion strain Δ*cellW* was able to colonize the host equally to the 13061-strain (Supplementary Fig. S1).

Δ*ank* has lost the plant growth promotion capacity

Due to the low plant colonization potential of Δ*ank*, we tested whether the deletion strain could promote plant growth at the same level as the 13061-strain. The pine seedlings were inoculated as before[14] with Δ*ank*, *M. extorquens* 13061, or water, and seedling dry weight was measured. The growth of the seedlings inoculated with Δ*ank* was lower (Fig. 2A), or at the same level (Fig. 2B) as in the water-inoculated seedlings (Fig. 2C-D, F-G), although the 13061-strain significantly induced pine root (Fig. 2A, E, H) and shoot growth (Fig. 2B, E, H) at all time points.

Ank structure has a unique viral motif

Ank was comprised of 194 amino acids and shared similarity with ankyrin proteins of other *Methylorubrum* and *Methylobacterium* strains (Fig. 3A, Supplementary Texts S1-S2). The greatest differences were found at the beginning and the end of the protein. Additionally, a unique sequence-stretch from amino acids 111 to 122 distinguished the Ank of *M. extorquens* from the other ankyrins (Fig. 3B). When this specific region was analyzed by BLAST, it had the best hit (Query Coverage 100%, Identity 100%) with an F-box domain and ankyrin repeat protein (QBK91994.1) of Pithovirus LCPAC304, which had three repeats of the region.

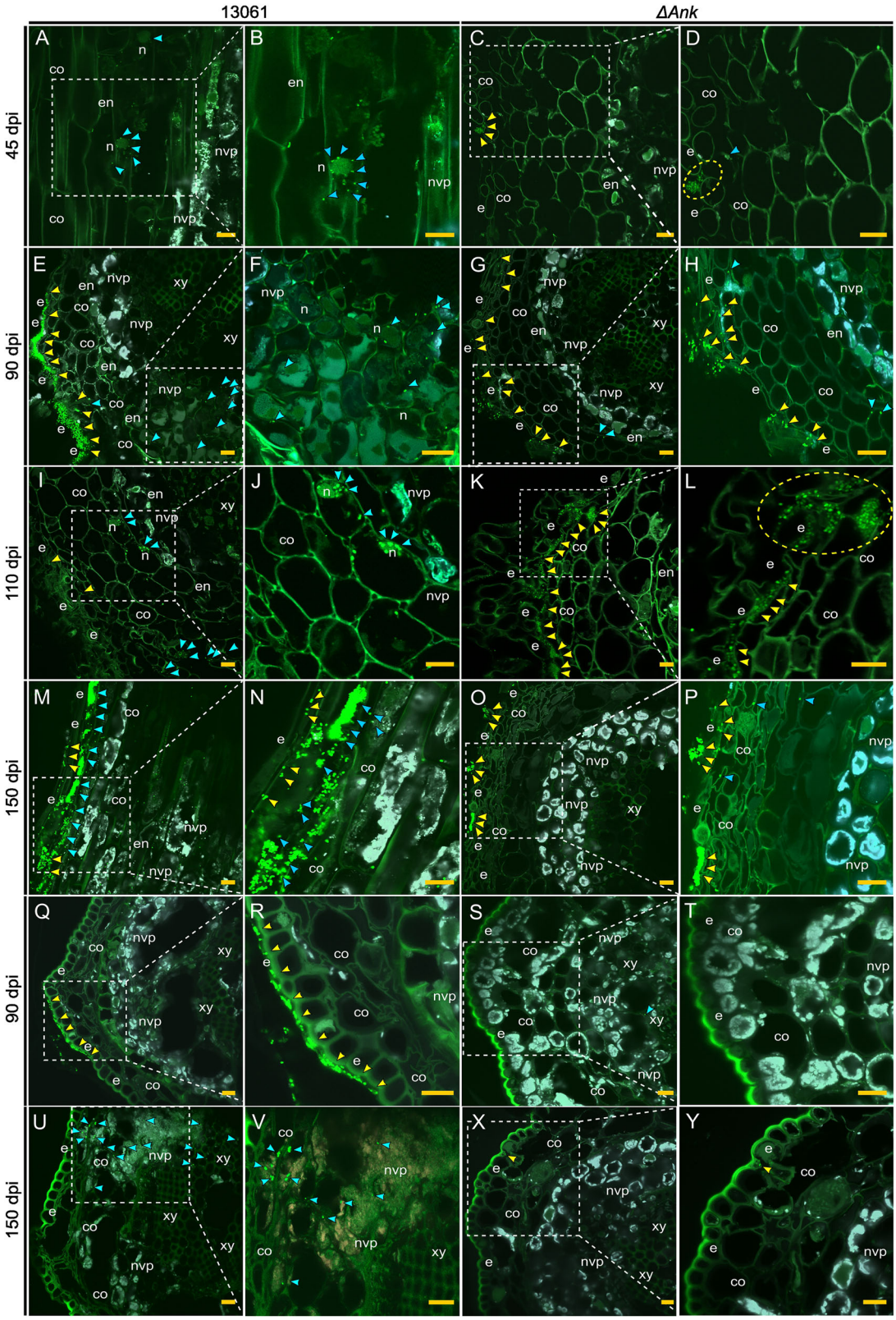
When the tertiary structure of Ank was studied by AlphaFold, the prediction was based on up to 12236 unique sequences (non-gap counts). The Ank structure had five typical ankyrin repeat motifs of alpha helix-turn-helix with a high confidence (Predicted Local Distance Difference Test, pLDDT >90) (Fig. 3C). The turn between ankyrin repeats 4 and 5 at residues 130-132 had a low confidence (70>pLDDT>50, Predicted Aligned Error, PAE 20-30), in addition to the residues 12-22 of the sequence (Fig. 3D, E). The Ank was strikingly different from the AnkB of *Legionella pneumophila*, which is the closest intracellular relative of *M. extorquens* DSM13060, having only three ankyrin repeats in its predicted structure (Fig. 3F).

Ank interaction targets are mainly inside the plant cell

The interactions of Ank with host proteins were analyzed using yeast two-hybrid assay. The positive hits represented 46 different host genes, of which the great majority (>84%) are expressed inside the plant cell (Fig. 4A). Of these, 21 genes are expressed in the nucleus, of which seven are exclusively expressed within the nuclear envelope. Twenty-seven genes are found in the cytoplasm, of which six genes are associated only with chloroplasts, and one gene is expressed only in mitochondria. A very small portion (6%) of the genes is expressed outside of the plant cell, either in the plasma membrane or in the extracellular matrix (Fig. 4A). By function, the genes are mainly associated with defense or stress (36%), cellular processes (40%), and development (24%). Specifically, 10% of the genes are associated with photosynthesis (Fig. 4B).

The interaction was confirmed by co-transformation for 25 out of 30 manually selected genes (Table S2). These included nuclear proteins, such as the CDK inhibitor P21 binding protein involved in cell cycle regulation (BCCIP) (similar with AT2G44510), an adenosine kinase (ADK) (AT2G37250), the basic region/leucine zipper motif protein 49 (BZIP49) (AT3G56660), the nuclear encoded CLP proteinase 5 (NCLPP5) (AT1G49970), and a nuclear pore complex-like protein (NUP54) (AT1G24310). Cytoplasmic proteins that are associated with plant development included the Pumilion 24 (PUM24) (AT3G16810), the GRF1-Interacting factor 1 (GIF1) (AT5G28640), the phosphoribosylanthranilate isomerase 1 (PAI1) (AT1G07780), and the phosphate transporter 2 (ATPT2) (AT2G38940). Furthermore, the PSPB-like protein 1 (PPL1) (AT3G55330) selected as a representative of photosynthesis-associated genes, was confirmed to interact with Ank (Table S2).

One extracellular protein identified among the confirmed targets, endo-1,3-beta glucosidase 7 (GH17b) (AT5G24318), functions in cell wall degradation by gene ontology classification. Another confirmed extracellular target, the seed storage 2S albumin superfamily protein (AT1G43890), has a role in plant defense. There were also several cytoplasmic proteins identified as Ank targets that were associated with plant defense. These include a flavonol synthase (AT2G15490), a chalcone



synthase (CHS) (AT5G13930), the E3 ubiquitin ligase (AT5G20620), the ACT domain repeats protein 11 (ACR11) (AT1G16880), the GmDREB5-interaction protein (TPR1) (AT1G80490), and the beta carbonic anhydrase 4 (BCA4) (AT1G70410).

Another group of confirmed cytoplasmic Ank targets were associated with plant stress responses and stress tolerance. These include alternative oxidase 1 A (AOX1A) (AT3G22370), catalase 2 (CAT2) (AT4G35090), the general regulatory factor 3 (GRF3) (AT5G38480), which is a 14-3-3 protein, the DNAJ heat shock N-terminal domain containing protein (AT3G04980), which is a chaperone, and the early responsive to dehydration 15-like protein (ERD15) (AT2G41430). An Ank target associated with both primary metabolism and redox

Fig. 1 | Laser scanning confocal microscopy (LSCM) of Scots pine roots (A-P) and shoots (Q-Y) colonized by the deletion strain *Δank* and the *M. extorquens* 13061 at 45 (A- D), 90 (E- H, Q- T), 110 (I- L) and 150 (M- P, U- Y) days-post-inoculation (dpi). Bacterial cells labeled with a fluorescent GFP reporter under the control of a constitutive promoter are visualized in bright green. **A** A lateral section of root inoculated with the 13061-strain, magnified in (**B**). Arrowheads indicate bacterial aggregation (blue) near the nuclei of parenchymal cells. **C** A cross-section inoculated with *Δank*, magnified in (**D**). The arrowhead (yellow) and dashed circle show bacteria on the epidermis. **E** A cross-section of root inoculated with the 13061-strain, arrowheads (yellow) show a biofilm on the epidermis, bacteria in non-vascular parenchyma magnified in **F**, where arrowheads (blue) point single bacterial cells. **G** A cross-section of root inoculated with *Δank*, magnified in (**H**). Bacterial cells are present in the epidermis (yellow arrowheads). **I** A cross-section of root inoculated with the 13061-strain, magnified in (**J**). The bacterial cells (blue arrowheads) are aggregating around the host nucleus in the endodermis. **K** A cross-section of the root inoculated with *Δank*, magnified in (**L**). Cells of bacteria colonize the epidermis (yellow arrowheads). **M** A lateral section of the root inoculated with the 13061-strain, magnified in (**N**). Arrowheads (yellow) indicate individual bacterial cells and infection pockets in the epidermis and cortex. **O** A cross-section of root inoculated with *Δank*, magnified in (**P**). Bacterial biofilms are only present on the epidermis (yellow arrowhead). **Q** Cross-section colonized by the 13061-strain showing a bacterial biofilm on the epidermis (yellow arrowheads), magnified in (**R**). **S** Cross-section inoculated with *Δank* showing no bacteria present, magnified in (**T**). **U** Cross-section colonized by the 13061-strain where bacteria have progressed through the epidermis, colonizing in the cortex, non-vascular parenchyma and xylem (yellow arrowheads), magnified in (**V**). **X** Cross-section inoculated with *Δank*, where an arrowhead (yellow) shows a single bacterial cell on the epidermis, magnified in (**Y**). Co cortex; e epidermis; en endodermis; nvp non-vascular parenchyma; xy xylem. Dashed square=magnified area. Scale bars, 5 µm.

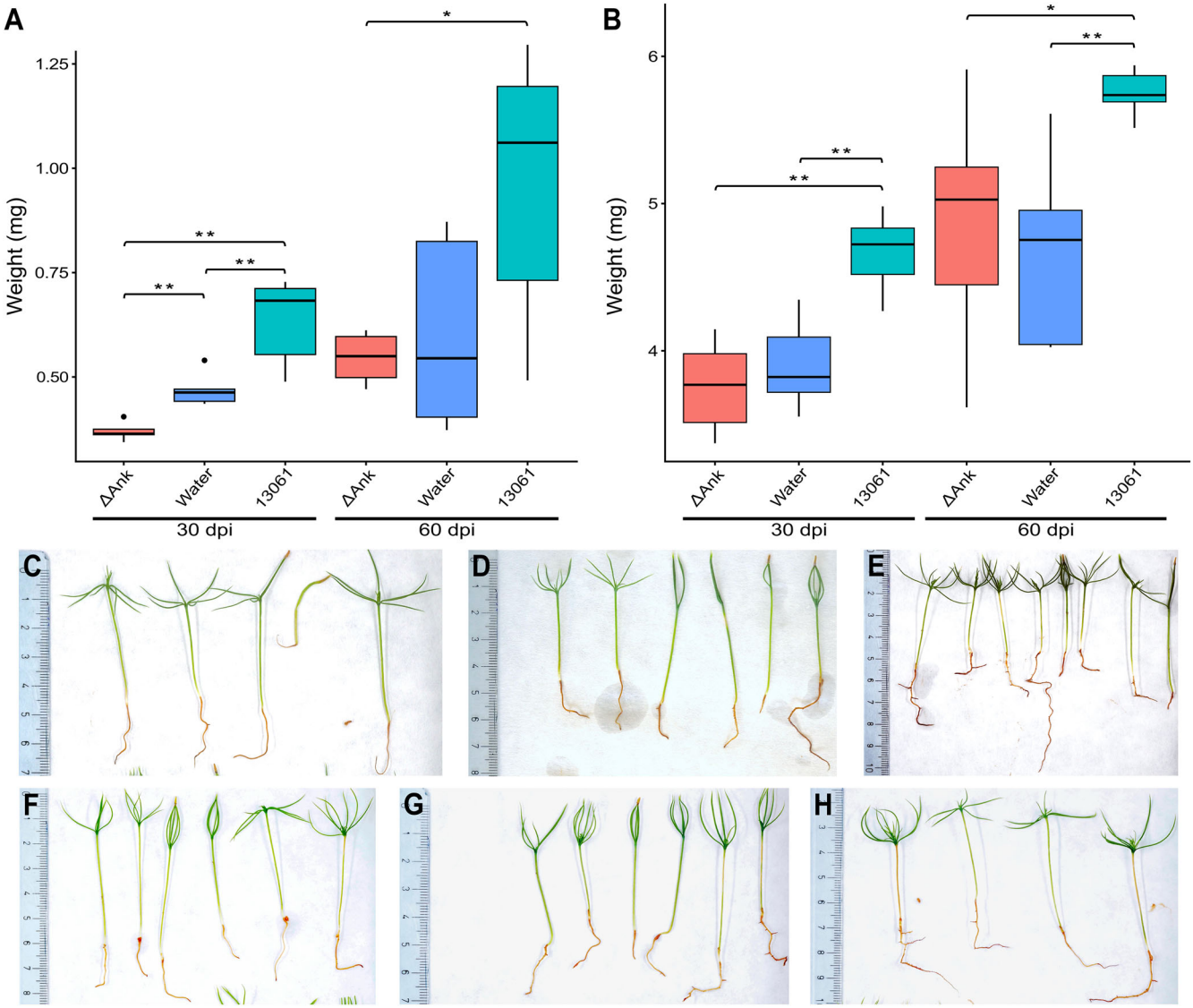


Fig. 2 | Growth of pine seedlings inoculated with the deletion strain *Δank*, water, and *M. extorquens* 13061 at 30- and 60-days-post-inoculation (dpi). **A** Dry weight of pine roots; **B** Dry weight of pine shoots. Five biological replicates were measured, each containing 10 seedlings. The maximum and minimum values are denoted by whiskers and the median by the black line inside each box. *=$p < 0.05$; **=$p < 0.01$ (one-way ANOVA, Tukey's test, t-test). Representatives of seedlings inoculated with (**C**) *Δank*, **D** water and **E** *M. extorquens* 13061 30 dpi, and **F** *Δank*, **G** water and **H** M*M. extorquens* 13061 60 dpi.

signaling, is the malate dehydrogenase (C-NAD-MDH1 or MDH) (AT1G04410) (Table S2).

When the networks of the Ank interacting proteins were studied, the photosynthesis-associated proteins formed a strong cluster (Fig. 4C). The Light-harvesting Chlorophyll B-binding protein 2 (LHCB2.4, AT3G27690) had a further interaction with AOX1A, and the Light-Harvesting Complex of photosystem II (LHCB5, AT4G10340) and Oxygen Evolving Subunit 33 (PSBO2, AT3G50820) had an interaction through non-Ank-interacting

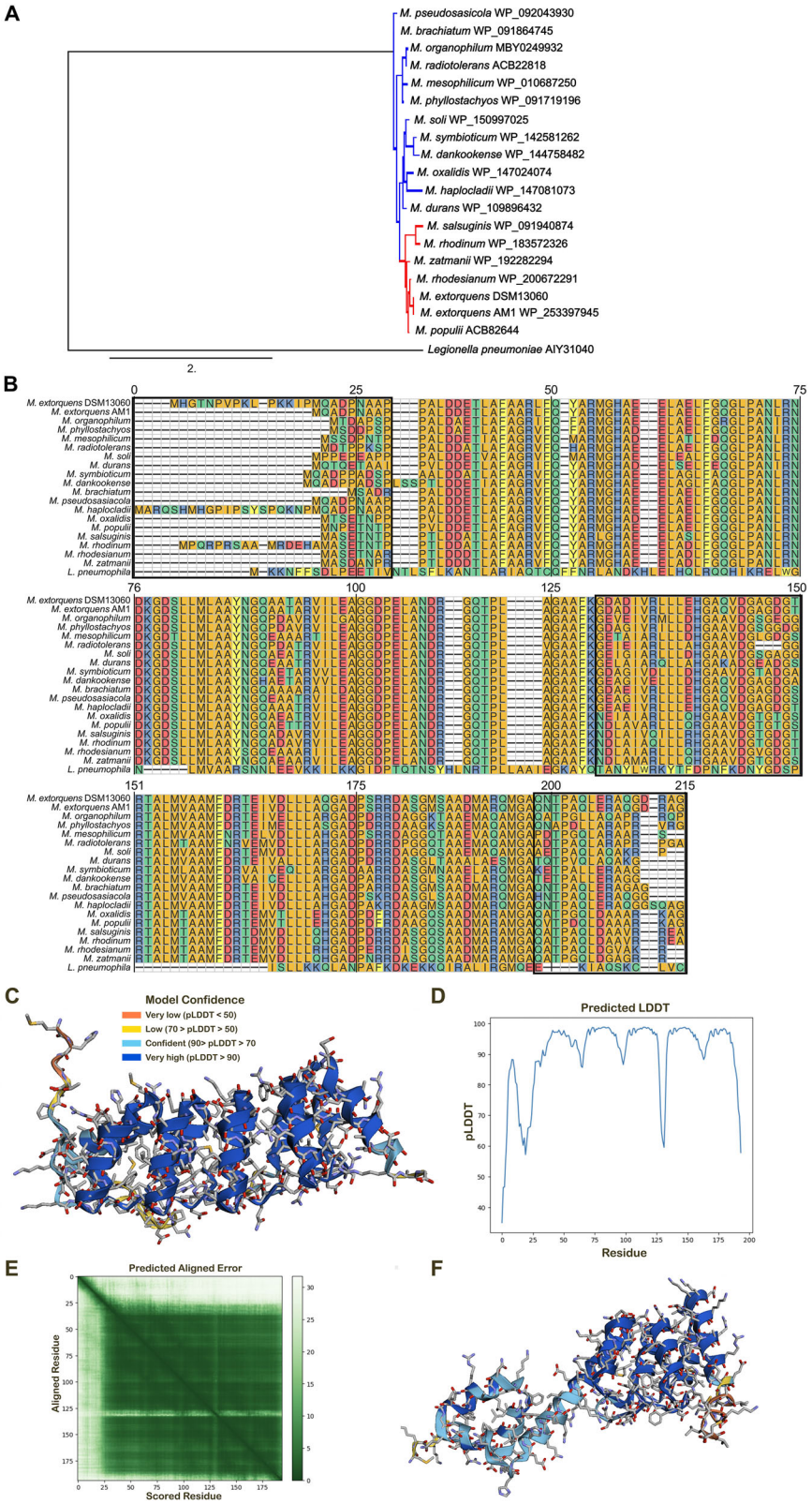


Fig. 3 | Structure of *M. extorquens* DSM13060 ankyrin-repeat protein (Ank). A A phylogenetic tree constructed in the PhyML3.0 program of Methylorubrum extorquens DSM13060 ankyrin-repeat protein Ank and the ankyrins of other Methylorubrum (red) and Methylobacterium (blue) strains, with *Legionella pneumophila* as an outgroup. The GenBank Accession number is followed by the species name. The scale bar indicates the phylogenetic distance of 0.05 nucleotide substitutions per site. A thick line indicates support by >50%. **B** Aligned protein sequences of *M. extorquens* DSM13060 ankyrin-repeat protein Ank and ankyrins of other *Methylorubrum* and *Methylobacterium* strains, as well as *L. pneumophila*. The divergent regions within the amino acid sequences are marked with a box, and the identified unique viral sequence is at the amino acids 131-141. The numbering is based on the ankyrin-repeat protein of *M. haplocladii*. **C** Predicted structure of *M. extorquens* DSM13060 Ank by the AlphaFold program. The prediction was based on up to 12236 unique sequences (non-gap counts). **D** The Predicted Local Distance Difference Test, pLDDT of Ank. **E** Predicted Aligned Error, PAE of Ank. **F** The structure of ankyrin repeat protein (AnkB) of *L. pneumophila*, predicted by AlphaFold program.

proteins Ubiquinol-Cytochrome C Reductase Binding Protein (QCR7-2, AT4G32470.1) and Apocytochrome B (cob, ATMG00220) with Cytochrome C1-2 (CYC1-2, AT5G40810). The CYC1-2 had a further interaction with MDH1, GRF3, and Serine Carboxypeptidase-like 27 (SCPL27, AT3G07990) (Fig. 4C).

Discussion

The modern agriculture should prefer natural, ecologically safe alternatives to chemical fertilizers and pesticides to avoid further degradation of the environment. Alternatives, such as biofertilizers, can improve crop yields through various environmentally friendly mechanisms. However, their

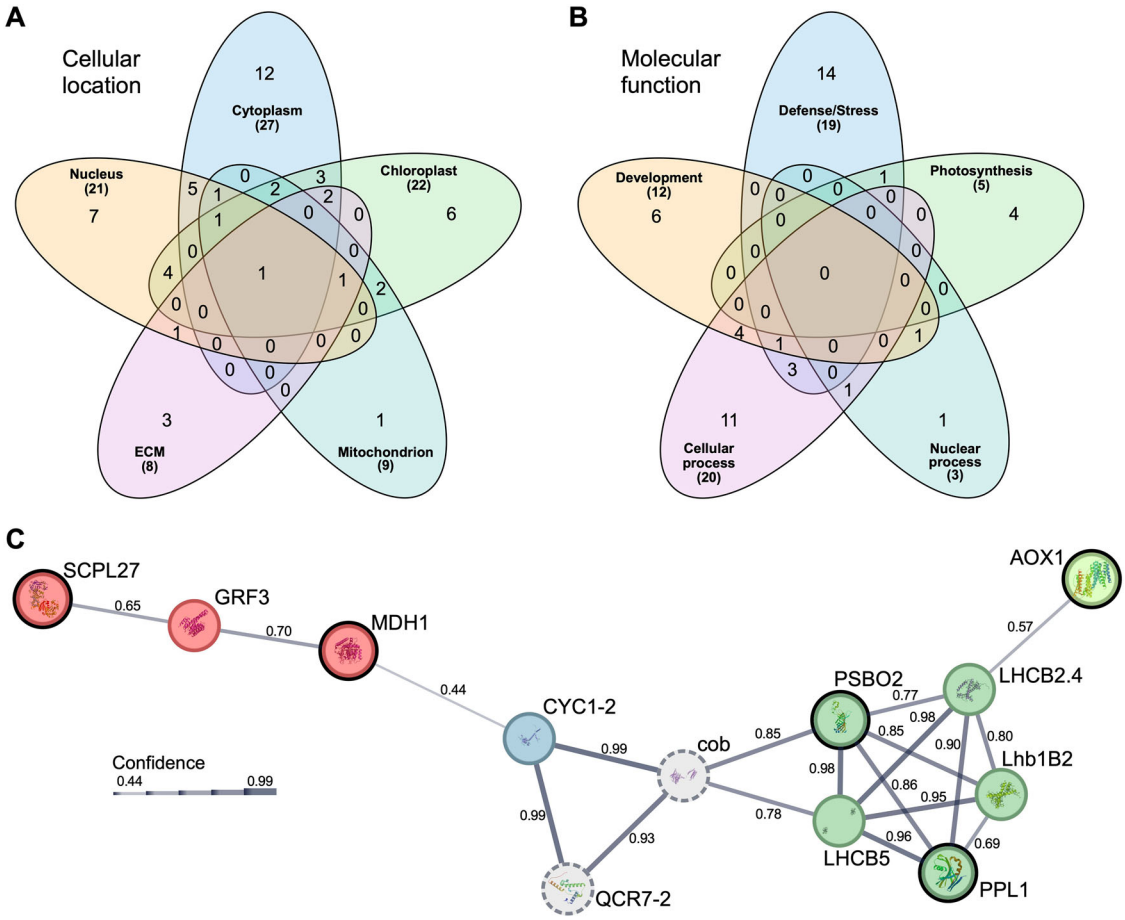


Fig. 4 | Interactions of *M. extorquens* DSM13060 Ank with plant proteins determined by yeast-two hybrid analysis. A A Venn diagram showing the cellular locations of plant proteins interacting with *M. extorquens* DSM13060 ankyrin repeat protein (Ank). ECM = extracellular matrix. **B** A Venn diagram showing biological functions (gene ontology, GO terms) of Ank interacting plant proteins. The numbers in the parenthesis refer to all proteins present in the group; the numbers in the center refer to proteins present exclusively in the group. **C** The STRING protein–protein interaction (PPI) network. The Ank target genes of the plant host, analyzed by yeast-two hybrid method, are highlighted with black halos and genes with light gray fill and dotted lines (cob, QCR7-2) are predicted functional partners added to the STRING network. Thickness of lines represents PPI confidence level, and fill colors of red, blue, green, or light green correspond to the biological processes identified.

application often results in inconsistent field performances due to the unsuitable genotype of the plant, environmental conditions, poor competence of the strains, or interactions with the innate microbial communities. Biofertilizers are typically applied to the soil, where localized microenvironments exist with changing temperatures and humidity[5]. However, shoot endophytic strains, or endosymbionts, such as *M. extorquens* DSM13060, can be directly inoculated to host plants during micropropagation, ensuring a stable interaction[5,26,27]. Strains that improve not only plant growth but also stress tolerance are a priority in biofertilization to confirm survival of the inoculant in the host under stressful growth conditions.

We identified the ankyrin-repeat protein, Ank, as a eukaryotic transcription-like protein in the genome of *M. extorquens* DSM13060 due to the unique trait by this bacterium of accumulating around host nuclei[13]. We predicted that Ank is delivered through T4SS to the host cytoplasm during *M. extorquens* DSM13060 interaction and, therefore, generated the deletion strain Δ*ank*. The Δ*ank* had an extremely poor colonization capacity of the host with only singular bacterial cells occasionally colonizing in the pine root tissue, a low number of bacteria residing on the plant surfaces, and no bacteria present in the shoots.

Earlier, deletion of an ankyrin repeat protein of *Legionella pneumophila*, AnkB, had no effect on bacterial infection, but led to the loss of replication by *L. pneumophila* in eukaryotic cells[28]. *L. pneumophila* is an intracellular human pathogen, which is closely related to the *Methylorubrum* genus. However, the two ankyrin repeat proteins, AnkB and Ank, had no resemblance at the structural level. Ankyrin repeat proteins are also abundant in the genomes of the insect endosymbionts *Spiroplasma, Wolbachia* and *Rickettsia*, where their involvement in interaction with the host has been suggested[29–31].

In general, the protein-protein interaction domain of ankyrin repeats is involved in many cellular tasks, such as transcription, transportation through membranes or cytoskeleton, regulation of cell cycle, and transduction of signals[32]. Due to the specific role in protein-protein interactions by ankyrin proteins, we performed a yeast-two-hybrid screening to identify the host targets of Ank. The results explained the observed effects of *Ank* gene loss in the bacterial deletion strain. Among the Ank targets, the majority were inside the plant cell and only 6% were expressed in the extracellular space or in the plasma membrane. One of the few extracellular targets was involved in plant cell wall degradation (GH17b) (Fig. 5). The *M. extorquens* DSM13060[13] genome lacks in enzymes with the potential of degrading the plant cell wall, and the deletion of a unique endopolygalacturonase gene (*CellW*) had no effect on host colonization. In general, bacterial endopolygalacturonases function in plant cell wall degradation for gaining access to the plant[33]. The endosymbiont therefore employs other mechanisms of invading the host tissue, where Ank plays a role, as deletion of *Ank* resulted in the loss of plant tissue colonization capacity. Ank could be responsible for activating the plant's own cell wall degrading enzyme, GH17b, for gaining access to the plant interior.

The deletion strain Δ*ank* also lost the capacity of inducing plant growth, as the Δ*ank*-inoculated pine seedlings were equal to, or even

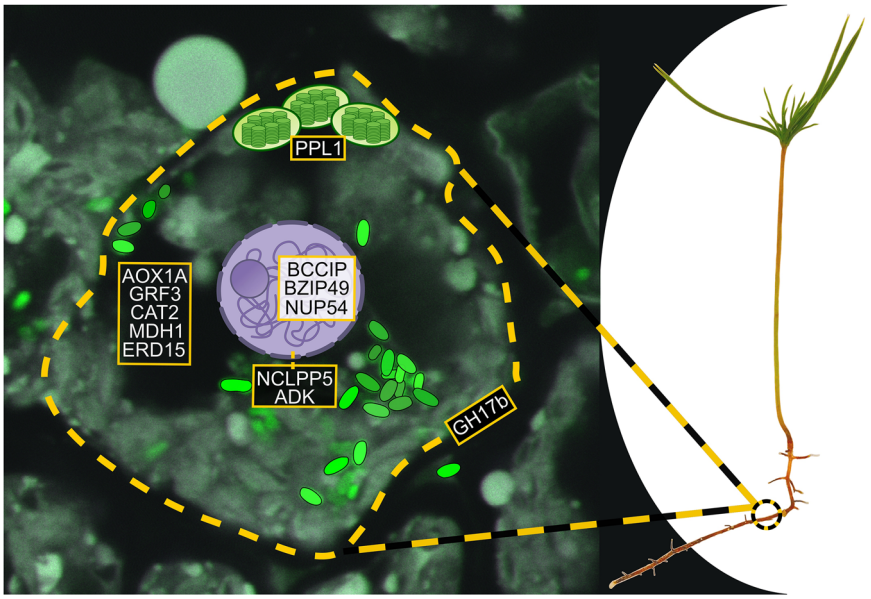


Fig. 5 | A schematic diagram showing the main targets (yellow text box) of *M. extorquens* DSM13060 Ank within the plant cell. A laser scanning confocal microscopy image of pine root cortical tissue in the background. GFP-tagged bacterial cells are illustrated in bright green, chloroplasts in light green with thylakoids, and nucleus in purple color. A dashed yellow line indicates the plant-cell wall, and the yellow-black line, the site of magnification. ADK=adenosine kinase, AOX1-A=alternative oxidase 1A, BCCIP = CDK inhibitor P21 binding protein involved in cell cycle regulation, BZIP49=basic region/leucine zipper motif protein 49, CAT2 = catalase 2, ERD15=early responsive to dehydration 15-like protein, GH17b=endo-1,3-beta glucosidase 7, GRF3 = general regulatory factor 3, MDH = malate dehydrogenase, NCLPP5=nuclear encoded CLP proteinase 5, NUP54=nuclear pore complex-like protein, PPL1 = PSPB-like protein 1.

smaller, in size than the water-inoculated seedlings in vitro, whereas the *M. extorquens* 13061 induced plant growth, as before[14]. Approximately 10% of the Ank targets were chloroplastic, involved in photosynthesis (Fig. 5). Although *M. extorquens* DSM13060 has no effect on the expression of photosynthesis genes[34], it increases the carbon content of the seedlings[14]. Therefore, Ank could be involved in non-transcriptional processes of photosynthesis, such as protein stabilization or influencing protein-protein interactions within the photosynthetic complexes. The network analysis of Ank-interacting proteins revealed a connection between the photosynthesis-associated proteins and those functioning in stress responses, AOX1A, MDH1, GRF3, and SCPL27. The chloroplast is a potent source of reactive oxygen species[35], which explains the interaction between Ank, photosynthetic enzymes, and oxidative stress-associated proteins.

Another Ank target carrying a role both in oxidative stress and primary metabolism was malate dehydrogenase (Fig. 5). In general, this enzyme is responsible for transforming oxaloacetate into malate, and back, where oxaloacetate is responsible for cell protection against hydroxyl radicals[36]. In our previous metabolomic analysis, the levels of malate were elevated in seedlings inoculated with *M. extorquens* DSM13060[34], and malate is a preferred carbon and energy source by *M. extorquens* DSM13060[7]. Therefore, *M. extorquens* DSM13060 likely uses malate as the source of carbon during the endosymbiosis[34], similar to rhizobia[37]. Ank could therefore activate the host malate dehydrogenase to provide a preferred carbon and energy source for the endosymbiont.

A large proportion of the Ank targets were involved in plant defense and stress responses (Fig. 5). This interaction could function to attenuate host defense responses and manage cellular stress during bacterial colonization. There were Ank targets such as AOX1A and CAT2, which are associated with oxidative stress and balancing the mitochondrial function, as well as GRF3, DNAJ heat shock-like protein, and ERD15, which are associated with abiotic stress[38,39]. Our earlier gene expression studies have shown that *M. extorquens* DSM13060 infection reduces expression of several genes involved in plant stress and senescence, such as *xyloglucan endotransglucosylase/hydrolase 24* (*MERI5*), *Chlorophyllase* (*CLH2*), *hydroxyproline-rich glycoprotein family* (*EULS3*), and the *disease resistance protein NHL3*[34], supporting the idea that *M. extorquens* DSM13060 actively suppresses plant stress responses during colonization. Equally, the modification of plant secondary metabolism by Ank through interactions with chalcone and flavonol synthases likely favors *M. extorquens* DSM13060 colonization in the host. At the gene expression level, *M. extorquens* DSM13060 induces biosynthesis of flavones and flavonols, which mainly function in UV protection, while down-regulating biosynthesis of anthocyanins that are typically employed against biotic stress[34,40]. Therefore, Ank likely functions in the processes of down-tuning the plant immune and stress-response mechanisms for a facilitated host entry.

There is accumulating evidence on the role of ankyrin-rich proteins in immune evasion of the host. For example, a eukaryotic-like ankyrin-repeat protein (ARP) of a bacterial sponge symbiont was shown to interfere with the phagocytosis of amoebal cells, resulting in accumulation of bacterial cells in phagosome[41]. We identified a motif in the endosymbiont Ank that differs from the ankyrins of most *Methylorubrum* and *Methylobacterium* strains but is enriched in an F-box domain and ankyrin repeat protein of Pithovirus LCPAC304[42]. Many viral or phage ankyrins have been associated with attenuating host immune responses to aid in bacterial infection. For example, a phage-encoded ankyrin-domain-containing protein (ANKp) modifies the immune responses in sea sponges to facilitate infection by symbiotic bacteria[43]. Therefore, the Ank present in the single megaplasmid of *M. extorquens* DSM13060 could be of viral origin, obtained to facilitate the intracellular colonization of host.

There was a surprisingly large portion (42%) of plant nuclear proteins among the Ank targets (Fig. 5). The trait of *M. extorquens* DSM13060 of aggregating around the host nuclei[13] facilitates the interaction of Ank with its nuclear targets. Among the targets, BCCIP, BZIP49, and NUP54 were strictly nuclear proteins with no cytoplasmic or extracellular occurrence. The NUP54, which forms the central channel of the nuclear pore complex with two other nucleoporins (NUP58 and NUP62), coordinates the activation of various stress-related signaling pathways and inhibits the effector-triggered immunity and programmed cell death pathways in plants[44,45]. Both BZIP49 and NCLPP5 are associated with chaperones[46,47], and the BCCIP is involved in DNA repair[48]. The BZIP49 regulates the DNA-templated transcription[49], and along with ADK, affects plant development[46,50,51].

Ank of *M. extorquens* DSM13060 can therefore influence the host development through interactions with NUP54, ADK and BZIP49 (Fig. 5). Our earlier gene expression studies have shown a deep involvement of *M. extorquens* DSM13060 in plant development[34]. The endosymbiont enhances the growth of pine seedlings without producing any known plant growth hormones[14,15] and induces the expression of several genes associated with auxin responses and plant development. For example, the genes *auxin-response factor 2* (*ARF2*) and *Cullin 1* (*CUL1*) are up-regulated by *M. extorquens* DSM13060 infection. Downstream, genes such as the *CO2 response secreted protease* (*CRSP*) and the *LRR-type receptor protein kinase* (*RFGR1*), associated with leaf and root development, respectively, are

induced[34]. Such changes in gene expression can be mediated through Ank/NUP54, Ank/ADK and Ank/BZIP49 interactions.

M. extorquens DSM13060 is a promising new biofertilizer; an intracellular bacterium that alleviates plant stress responses and induces plant root and shoot development, likely utilizing malate as the main carbon source in the plant endosphere[14,34]. The endosymbiont is capable of infecting several plant hosts and, being intracellular, represents a completely new type of persistent biofertilizers[5]. Here, we showed that the ankyrin repeat protein Ank is crucial for a successful infection by *M. extorquens* DSM13060. Ank interacts with several intracellular proteins associated with modifying the plant defense system for a facilitated host entry, similar to viral ankyrins of sea sponges[43]. Aggregation of *M. extorquens* DSM13060 around host nuclei potentiates the interaction of Ank with nuclear proteins BCCIP, BZIP49, and NUP54, involved in plant development and stress responses. As such, Ank could be characterized as a protein-based key that enables a successful *in planta* phenotype by *M. extorquens* DSM13060, unlocking the positive traits expressed by the endosymbiont in the host (Fig. 5). Our discovery of Ank as a bacterial key to the plant system opens a completely new aspect of mutualistic plant-microbe interactions in the non-rhizobial endosymbiosis.

Methods

Identification of *M. extorquens* DSM13060 effectors

The Searching Algorithm for Type IV Effector proteins (S4TE2.0) software package[52] was used to identify Type IV secreted effector protein candidates from the *M. extorquens* DSM16030 genome. *M. extorquens* DSM16030 genome was uploaded to S4TE2.0 as a Genbank file and analyzed. The T4SS-dependent translocation of ankyrin-like effectors has previously been validated with intracellular human pathogens, such as *Legionella pneumophila* and *Coxiella burnetii*[53–56].

Bacterial strains, plasmids, and cultivation

Methylorubrum extorquens DSM13060 [Genbank: AGJK00000000] was originally isolated from bud tissues of mature Scots pine in Oulu, Northern Finland (65°0′ N; 25°30′ E)[7,13]. The fluorescent derivative *M. extorquens* 13061 containing two successive chromosomal GFP coding genes under constitutive promoter[14] was used as the parental strain for the construction of gene in-frame deletion mutant[57]. Strains and plasmids used are listed in Supplementary Table S3. For isolation of genomic DNA, mutant construction, and inoculum preparation, *M. extorquens* 13061 and constructed mutants were grown by shaking at 28°C in ammonium mineral salts (AMS) supplemented with 1% (v/v) methanol (MeOH), 1% (w/v) sodium succinate, and appropriate antibiotics when required (kanamycin 50 µg/mL). The AMS medium contained 0.66 g/l $(NH_4)_2SO_4$, 1.0 g/l $MgSO_4$-$7H_2O$, and 0.015 g/l $CaCl_2$-$2H_2O$. In addition, after sterilization, 1.0 ml/l of AMS trace elements solution, 1 ml/l of stock A, and 20 ml/l of 1.0 M phosphate buffer were added. The AMS stock A contained 5.0 g/l Fe-NaEDTA and 2.0 g $NaMoO_4$*$2H_2O$. The AMS trace elements solution contained 0.5 g/l $FeSO_4$*$7H_2O$, 0.4 g/l $ZnSO_4$*$7H_2O$, 0.02 g/l $MnSO_4$*H_2O, 0.015 g/l H_3BO_3, 0.01 g/l $NiCl_2$*$6H_2O$, 0.25 g/l EDTA, 0.05 g/l $CoCl_2$*$6H_2O$, and 0.005 g/l of $CuCl_2$*$2H_2O$. The 1 M phosphate buffer consisted of 113.0 g/l K_2HPO_4 and 47.0 g/l KH_2PO_4.

Escherichia coli strains DH5α and SM10λpir, used for the mutant construction, were grown in Luria-Bertani (LB) medium (1% tryptone, 0.5% yeast-extract, 1% NaCl) at 37°C, supplemented either with 100 µg/ml ampicillin or with 10 µg/ml tetracycline.

Isolation of genomic DNA

Bacterial cells were cultured and grown to the late logarithmic phase (optical density at 600 nm (OD600) 0.8–1.0) and harvested by centrifugation at 6000 × g for 5 min at 4 °C. The bacterial pellet was ground in liquid nitrogen with mortar and pestle. The genomic DNA was then isolated according to the manufacturer's instructions (Qiagen). The concentration and quality of DNA yield were analyzed by Qubit™ dsDNA HS assay kit (Invitrogen™).

Construction of the deletion strains

The genomic data were retrieved from the JGI genome portal (https://img.jgi.doe.gov/) and were used to design primers specific for the gene *Ank* (IMG gene accession 2507328582), and *CellW* (IMG gene accession 2507328925). The targeted deletion strains Δ*ank* and Δ*cellW* of *M. extorquens* 13061 were generated using the method described earlier[58]. In brief, 0.5 kb regions upstream (AB) and downstream (CD) of the target genes were amplified using primers with restriction sites at the overlap sequences (Supplementary Table S3). The up- and downstream regions were combined by overlap extension PCR (OE-PCR) using Phusion High-Fidelity DNA polymerase (Thermo Scientific). The amplified construct was verified on an agarose gel, purified and cloned into the pJET 1.2 vector (CloneJET PCR Cloning Kit Thermo Scientific, MA, USA). The plasmid was then electroporated (Gene Pulser Electroporation system, BioRad) into *E. coli* DH5α. Transformants were selected on LB agar plates supplemented with ampicillin (100 µg/ml) and incubated overnight at 37 °C. The colonies were screened by colony PCR (DreamTaq polymerase, Thermo Scientific) specific for the deletion (Supplementary Table S4), and the constructs were confirmed by Sanger sequencing (Eurofins Genomics, Germany).

The plasmid DNA was isolated with GeneJet Plasmid Miniprep kit (Thermo Scientific). The construct was restricted from pJET 1.2 and cloned into the suicide vector pT18mobsacB. Transformation of the suicide vector was carried out in *E. coli* SM10λ*pir* by electroporation. The transformants were grown on LB agar containing tetracyclin (10 µg/ml) in the presence of 2% X-gal and IPTG. The positive colonies were selected by blue-white screening and confirmed by colony-PCR (primers listed in Supplementary Table S5).

Transposon mutagenesis and mutant screening

Biparental matings were performed to introduce the selected positive colonies of *E. coli* SM10λ*pir::pT18mobsacB-Δank* or *E. coli* SM10λ*pir::pT18-mobsacB-ΔcellW* into *M. extorquens* 13061 by growing for 48-72 hours at 28°C on AMS agar supplemented with MeOH[59,60]. Dilutions of the biparental mating mixtures were then plated onto AMS agar supplemented with MeOH, sodium succinate, and tetracycline (30 µg/ml), and the cells were grown for another 48 h at 28 °C. Counter selections were performed using 10% sucrose for lacZ selection. Correct transformants were identified by colony-PCR (Primers shown in Supplementary Table S5).

Bacterial growth curves

The bacterial growth curve analysis of the *M. extorquens* 13061 and Δ*ank* in normal growth conditions was performed. The bacterial strains were cultured in AMS medium supplemented with 1% MeOH and 1% sodium succinate for up to 5 days. The logarithmic values of CFU/ml of the deletion strain were compared to the 13061-strain at 0.1, 0.3, 0.5, 0.7, and 1.0 OD_{600} at 24 h, 48 h, 72 h, 96 h, and 120 h, respectively.

Seed sterilization, germination, and plant growth conditions

Seeds of Scots pine were prepared by heat treatment for 72 h at 55°C in the dark, followed by incubation in sterile water overnight at room temperature, followed by surface sterilization with 3% calcium hypochlorite (w/v) for 20 min. Once the seeds were surface sterilized, they were rinsed three times with sterile water and grown on moist sterile vermiculite in glass jars. Germination took place after 7–10 days in the plant growth chamber at 24 ± 3°C at 16/8 h photoperiod. For the inoculation, the bacteria were grown in AMS supplemented with 1% (v/v) methanol (MeOH), and 1% (w/v) sodium succinate for 3–4 days. Bacterial cultures were diluted with sterile water to the density of 2.5×10^7 CFU/ml and 100 µl of the inoculum was pipetted onto each germinated pine seedling for the experiment.

Measurement of plant growth parameters

A total of 10–12 seedlings from 5 glass containers per sample type were randomly sampled to assess the plant growth promotion in Scots pine inoculated with the deletion strain Δ*ank*, the *M. extorquens* 13061, or sterile water. The dry weights of roots and shoots were analyzed after drying the

samples for 48 h at +60 °C. The sampling was done at 30- and 60-days post-inoculation (DPI).

Confocal laser scanning microscopy (CLSM)

Seedlings were analyzed at 45-, 90-, 110-, and 150-dpi (days-post-inoculation). A total of 10-12 seedlings/sampling/interval/strain were checked. The experiment comprised three biological replicates conducted over the period of a year and a half.

Roots and shoots were cut into 2–3 mm pieces in diameter and fixed in 4% paraformaldehyde (wt/vol), 0.1% glutaraldehyde (vol/vol), 20% glycerol (vol/vol), and 0.1 M sodium phosphate buffer (pH 7.4) at 4 °C under vacuum. The root samples were fixed for 4 h and the shoots were fixed for 8 h, followed by overnight incubation at 4 °C. The fixed tissues were cut into 20–30 µm sections with a cryomicrotome (Reichert-Jung 2800 Frigocut with 2040 microtome) and mounted on microscopy slides with Vectashield antifade mounting medium with DAPI (Vector Laboratories). The pine tissue sections were studied with CLSM (LSM 5 Pascal; Carl Zeiss) using Plan-Neofluor 40×/1.3 and Plan-Apochromat 63×/1.4 oil objectives. The GFP fluorophore was excited at a wavelength of 488 nm by an argon ion laser, and emissions were detected using a 505-530 nm band-pass filter. The background autofluorescence of the plant tissues was detected using a 670 nm long-pass filter. The projections of the channels were analyzed and merged using the Zeiss LSM Image Browser (ver. 4.2.0.121) and Zeiss ZEN ver. 2.5 software (Blue edition; Carl Zeiss Microscopy GmbH).

RNA isolation

Total RNA of *M. extorquens*-inoculated pine seedlings was isolated 90 days-post-inoculation (DPI) as described previously by Koskimäki et al. 2022[34]. Briefly, 15 pine seedlings were pooled and ground in liquid nitrogen as one sample. Five pooled samples were used for the yeast two-hybrid studies to isolate a minimum of 300 µg of total pine RNA. The RNA was isolated with a two-phase protocol using first a CTAB and chloroform-based extraction[61], and after the $LiCl_2$ precipitation and DNase treatment for 15 min at 25 °C with TURBO DNase (Thermo Fisher Scientific, Waltham, USA), purification of pine total RNA was performed using RNeasy Mini Kit (Qiagen, Hilden, Germany). The RNA concentrations were measured with NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA), and the quality was assessed by agarose gel electrophoresis and Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) with RNA Nano 6000 kit. Then, the RNA was submitted to Creative Biolabs Ltd. for the Yeast 2-Hybrid analysis.

Yeast 2-Hybrid (Y2H) assay

The Y2H assay was performed by Creative Biolabs Ltd. (Shirley, NY, USA). Briefly, the pGBKT7 vector was used as the bait construct, and the pGADT7 as the prey construct. The bait and the prey vectors were co-transformed using the GAL4 system. All transformants were spread on SD/-Trp/-Leu/-His/-Ade/AbA/X-a-Gal) selection medium plates and screened for positive clones. A total of 68 positive clones were observed, and after the toxicity and self-activation tests, 51 positive hits were identified, representing 46 different genes. The gene sequences were compared by the BlastX program against the annotated protein sequences of *Pinus taeda* (v. 1.01) and *Arabidopsis thaliana* (TAIR10) genomes, and the functions (gene ontology, GO terms) of each known protein were retrieved from Enzyme Commission (EC), InterPro (IPR) and UniProt (https://www.uniprot.org). Of the 46 genes, 30 were manually selected for interaction validation by co-transformation. All Y2H protocols were performed by Creative Biolabs Y2H standard methods.

Protein-protein interaction analysis and visualization

The Search Tool for Recurring Instances of Neighboring Genes (STRING v. 11.0)[62] was used to identify protein–protein interaction (PPI) networks of yeast two-hybrid (Y2H) candidate genes. The yeast two-hybrid PPI network was constructed with a confidence score between 0.44 (44%; medium confidence) and 0.99 (99%; highest confidence). The confidence level indicates the approximate likelihood of a predicted link between two genes. All 41 Y2H genes, except the five candidates (5/30; 16,67% fail rate) that failed the validation, were selected for the protein-protein interaction analysis with the STRING web platform (https://string-db.org/).

Phylogeny and structural analysis

The Ank primary structure was analyzed with Basic Local Alignment Tool (BLAST) at National Institute of Health Database (www.nih.gov) and with ggmsa tool[63] in *R* v.4.4.0[64]. Maximum-likelihood (ML) trees were constructed with PhyML v3.0[65]. The alignment was performed by Muscle, where contiguous non-conserved positions were minimized, and curation was done by Gblocks 0.91b (Supplementary Texts S1-S2). For the phylogenetic tree, both bootstrapping method, and various approximate Likelihood-Ratio Tests (SH-like, Chi2-based parametric, minimum of SH-like and Chi2) with WAG substitution model were tested.

The three-dimensional structure of Ank and AnkB was predicted using the machine-learning protocol AlphaFold[66]. The AlphaFold creates a per-residue confidence metric, which is named "predicted local distance difference test" (pLDDT) that ranges between 0-100. The pLDDT is used to estimate the agreement of the prediction with an experimentally determined structure, and it is based on a local distance difference test Cα (lDDT-Cα). The program also produces Predicted Aligned Error (PAE). The pLDDT predicts best intra-domain confidences, and the PAE is best fitted for confidence analyses between domains or chains[66].

Statistical analyses

Data for measurement of plant growth promotion was collected and analyzed statistically. The normal distribution was determined with the Shapiro-Wilk's test for normality, and equality of variances was studied by the Levene's test in R. For the normally distributed datasets with equal variances, a one-way analysis of variance (ANOVA) was performed. For the datasets that showed statistical significance, Tukey's test or t-test was performed to reveal the statistical difference between individual pairs. The box plot for pine growth promotion study was constructed in RStudio[64], version 2022.07.2 using packages ggpubr and ggplot2.

Data availability

All data are available in the main text or the supplementary materials. The genome of *M. extorquens* DSM13060 is available in Genbank (https://www.ncbi.nlm.nih.gov/nuccore/AGJK00000000).

Received: 13 August 2025; Accepted: 28 January 2026;
Published online: 01 April 2026

References

1. Van Rikxoort, H., Schroth, G., Läderach, P. & Rodríguez-Sánchez, B. Carbon footprints and carbon stocks reveal climate-friendly coffee production. *Agron. Sustain Dev.* **34**, 887–897 (2014).
2. Reijnders, L. Phosphorus resources, their depletion and conservation, a review. *Resour. Conserv Recycl.* **93**, 32–49 (2014).
3. Tilman, D. Global environmental impacts of agricultural expansion: the need for sustainable and efficient practices. *Proc. Natl. Acad. Sci. USA* **96**, 5995–6000 (1999).
4. Rani, L. et al. An extensive review on the consequences of chemical pesticides on human health and environment. *J. Clean. Prod.* **283**, 124657 (2021).
5. Pirttilä, A. M., Mohammad Parast Tabas, H., Baruah, N. & Koskimäki, J. J. Biofertilizers and biocontrol agents for agriculture: How to identify and develop new potent microbial strains and traits. *Microorganisms* **9**, 817 (2021).
6. Timmusk, S., Behers, L., Muthoni, J., Muraya, A. & Aronsson, A.-C. Perspectives and challenges of microbial application for crop improvement. *Front. Plant Sci.* **8**, 49 (2017).
7. Pirttilä, A. M., Laukkanen, H., Pospiech, H., Myllylä, R. & Hohtola, A. Detection of intracellular bacteria in the buds of Scotch pine (Pinus

sylvestris L.) by in situ hybridization. *Appl. Environ. Microbiol.* **66**, 3073–3077 (2000).
8. Parniske, M. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nat. Rev. Microbiol.* **6**, 763–775 (2008).
9. Poole, P., Ramachandran, V. & Terpolilli, J. Rhizobia: from saprophytes to endosymbionts. *Nat. Rev. Microbiol* **16**, 291–303 (2018).
10. Chaintreuil, C. et al. Photosynthetic bradyrhizobia are natural endophytes of the African wild rice Oryza breviligulata. *Appl. Environ. Microbiol.* **66**, 5437–5447 (2000).
11. Van der Meij, A. et al. Inter-and intracellular colonization of Arabidopsis roots by endophytic actinobacteria and the impact of plant hormones on their antimicrobial activity. *Antonie Van. Leeuwenhoek* **111**, 679–690 (2018).
12. Cocking, E. C., Stone, P. J. & Davey, M. R. Intracellular colonization of roots of Arabidopsis and crop plants by Gluconacetobacter diazotrophicus. *Vitr. Cell. Dev. Biol. Plant* **42**, 74–82 (2006).
13. Koskimäki, J. J., Pirttilä, A. M., Ihantola, E.-L., Halonen, O. & Frank, A. C. The intracellular scots pine shoot symbiont Methylobacterium extorquens DSM13060 aggregates around the host nucleus and encodes eukaryote-like proteins. *mBio* **6**, e00039–15 (2015).
14. Pohjanen, J. et al. Interaction with ectomycorrhizal fungi and endophytic Methylobacterium affects nutrient uptake and growth of pine seedlings in vitro. *Tree Physiol.* **34**, 993–1005 (2014).
15. Pirttilä, A. M., Joensuu, P., Pospiech, H., Jalonen, J. & Hohtola, A. Bud endophytes of Scots pine produce adenine derivatives and other compounds that affect morphology and mitigate browning of callus cultures. *Physiol. Plant* **121**, 305–312 (2004).
16. Koenig, R. L., Morris, R. O. & Polacco, J. C. tRNA is the source of low-level trans-zeatin production in Methylobacterium spp. *J. Bacteriol.* **184**, 1832–1842 (2002).
17. Koskimäki, J. J. et al. Methyl-esterified 3-hydroxybutyrate oligomers protect bacteria from hydroxyl radicals. *Nat. Chem. Biol.* **12**, 332–338 (2016).
18. Koskimäki, J. J., Pohjanen, J., Ihantola, E.-L., Sutela, S. & Pirttilä, A. M. A shoot endosymbiont colonizes pine host by unique and rhizobia-like mechanisms boosted by surface-fixed methanol. *Plant Cell Physiol*. https://doi.org/10.1093/pcp/pcaf135 (2025).
19. Guha, S. et al. Nod factor-independent 'crack-entry' symbiosis in dalbergoid legume *Arachis hypogaea*. *Environ. Microbiol.* **24**, 2732–2746 (2022).
20. Quilbé, J. et al. Genetics of nodulation in Aeschynomene evenia uncovers mechanisms of the rhizobium–legume symbiosis. *Nat. Commun.* **12**, 829 (2021).
21. Oldroyd, G. E. D., Murray, J. D., Poole, P. S. & Downie, J. A. The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev. Genet* **45**, 119–144 (2011).
22. Frank, A. C. Molecular host mimicry and manipulation in bacterial symbionts. *FEMS Microbiol. Lett.* **366** (2019).
23. Hanford, H. E., Von Dwingelo, J. & Abu Kwaik, Y. Bacterial nucleomodulins: A coevolutionary adaptation to the eukaryotic command center. *PLoS Pathog.* **17**, e1009184 (2021).
24. Bierne, H. & Pourpre, R. Bacterial factors targeting the nucleus: the growing family of nucleomodulins. *Toxins* **12**, 220 (2020).
25. Bierne, H. & Cossart, P. When bacteria target the nucleus: the emerging family of nucleomodulins. *Cell Microbiol* **14**, 622–633 (2012).
26. Ardanov, P., Sessitsch, A., Häggman, H., Kozyrovska, N. & Pirttilä, A. M. Methylobacterium-induced endophyte community changes correspond with protection of plants against pathogen attack. (2012).
27. Podolich, O., Laschevskyy, V., Ovcharenko, L., Kozyrovska, N. & Pirttilä, A. M. *Methylobacterium* sp. resides in unculturable state in potato tissues in vitro and becomes culturable after induction by *Pseudomonas fluorescens* IMGB163. *J. Appl Microbiol.* **106**, 728–737 (2009).
28. Al-Khodor, S., Price, C. T., Habyarimana, F., Kalia, A. & Abu Kwaik, Y. A Dot/Icm-translocated ankyrin protein of *Legionella pneumophila* is required for intracellular proliferation within human macrophages and protozoa. *Mol. Microbiol.* **70**, 908–923 (2008).
29. Lehman, S. S. et al. The Rickettsial Ankyrin Repeat Protein 2 Is a Type IV Secreted Effector That Associates with the Endoplasmic Reticulum. *mBio* **9** (2018).
30. Papafotiou, G., Oehler, S., Savakis, C. & Bourtzis, K. Regulation of Wolbachia ankyrin domain encoding genes in Drosophila gonads. *Res. Microbiol.* **162**, 764–772 (2011).
31. Harumoto, T. & Lemaitre, B. Male-killing toxin in a bacterial symbiont of Drosophila. *Nature* **557**, 252–255 (2018).
32. Chakrabarty, B. & Parekh, N. Sequence and structure-based analyses of human Ankyrin repeats. *Molecules* **27**, 423 (2022).
33. De Lorenzo, G., Cervone, F., Hahn, M. G., Darvill, A. & Albersheim, P. Bacterial endopectate lyase: evidence that plant cell wall pH prevents tissue maceration and increases the half-life of elicitor-active oligogalacturonides. *Physiol. Mol. Plant Pathol.* **39**, 335–344 (1991).
34. Koskimäki, J. J. et al. The meristem-associated endosymbiont *Methylorubrum extorquens* DSM13060 reprograms development and stress responses of pine seedlings. *Tree Physiol.* **42**, 391–410 (2022).
35. Asada, K. Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol.* **141**, 391–396 (2006).
36. Oh, T. J., Kim, I. G., Park, S. Y., Kim, K. C. & Shim, H. W. NAD-dependent malate dehydrogenase protects against oxidative damage in Escherichia coli K-12 through the action of oxaloacetate. *Environ. Toxicol. Pharm.* **11**, 9–14 (2002).
37. Mitsch, M. J., diCenzo, G. C., Cowie, A. & Finan, T. M. Succinate transport is not essential for symbiotic nitrogen fixation by Sinorhizobium meliloti or Rhizobium leguminosarum. *Appl Environ. Microbiol* **84**, e01561–17 (2018).
38. Bao, H. et al. Catalase protects against nonenzymatic decarboxylations during photorespiration in Arabidopsis thaliana. *Plant Direct* **5** (2021).
39. Garmash, E. V., Velegzhaninov, I. O., Ermolina, K. V., Rybak, A. V. & Malyshev, R. V. Altered levels of AOX1a expression result in changes in metabolic pathways in Arabidopsis thaliana plants acclimated to low dose rates of ultraviolet B radiation. *Plant Sci.* **291**, 110332 (2020).
40. Martens, S., Preuß, A. & Matern, U. Multifunctional flavonoid dioxygenases: Flavonol and anthocyanin biosynthesis in Arabidopsis thaliana L. *Phytochemistry* **71**, 1040–1049 (2010).
41. Nguyen, M. T. H. D., Liu, M. & Thomas, T. Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis. *Mol. Ecol.* **23**, 1635–1645 (2014).
42. Bäckström, D. et al. Virus genomes from deep sea sediments expand the ocean megavirome and support independent origins of viral gigantism. *mBio* **10** (2019).
43. Jahn, M. T. et al. A phage protein aids bacterial symbionts in eukaryote immune evasion. *Cell Host Microbe* **26**, 542–550 (2019).
44. Gu, Y. et al. Nuclear pore permeabilization is a convergent signaling event in effector-triggered immunity. *Cell* **166**, 1526–1538 (2016).
45. Boeglin, M. et al. Reduced expression of AtNUP62 nucleoporin gene affects auxin response in Arabidopsis. *BMC Plant Biol.* **16**, 1–11 (2016).
46. Che, P. et al. Signaling from the endoplasmic reticulum activates brassinosteroid signaling and promotes acclimation to stress in Arabidopsis. *Sci. Signal* **3**, ra69–ra69 (2010).
47. Peltier, J.-B., Ytterberg, J., Liberles, D. A., Roepstorff, P. & van Wijk, K. J. Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of Arabidopsis thaliana. *J. Biol. Chem.* **276**, 16318–16327 (2001).
48. Lu, H., Yue, J., Meng, X., Nickoloff, J. A. & Shen, Z. BCCIP regulates homologous recombination by distinct domains and suppresses spontaneous DNA damage. *Nucleic Acids Res.* **35**, 7160–7170 (2007).

49. De Backer, J., Van Breusegem, F. & De Clercq, I. Proteolytic activation of plant membrane-bound transcription factors. *Front. Plant Sci.* **13**, 927746 (2022).
50. Feng, X., Yang, R., Zheng, X. & Zhang, F. Identification of a novel nuclear-localized adenylate kinase 6 from Arabidopsis thaliana as an essential stem growth factor. *Plant Physiol. Biochem.* **61**, 180–186 (2012).
51. Slovak, R. et al. Ribosome assembly factor Adenylate Kinase 6 maintains cell proliferation and cell size homeostasis during root growth. *N. Phytol.* **225**, 2064–2076 (2020).
52. Noroy, C., Lefrançois, T. & Meyer, D. F. Searching algorithm for Type IV effector proteins (S4TE) 2.0: Improved tools for Type IV effector prediction, analysis and comparison in proteobacteria. *PLoS Comput. Biol.* **15**, e1006847 (2019).
53. Burette, M. et al. Modulation of innate immune signaling by a Coxiella burnetii eukaryotic-like effector protein. *Proc. Natl. Acad. Sci. USA* **117**, 13708–13718 (2020).
54. Chen, C. et al. Large-scale identification and translocation of type IV secretion substrates by Coxiella burnetii. *Proc. Natl. Acad. Sci. USA* **107**, 21755–21760 (2010).
55. Voth, D. E. et al. The Coxiella burnetii ankyrin repeat domain-containing protein family is heterogeneous, with C-terminal truncations that influence Dot/Icm-mediated secretion. *J. Bacteriol.* **191**, 4232–4242 (2009).
56. Garcia-Garcia, J. C., Rennoll-Bankert, K. E., Pelly, S., Milstone, A. M. & Dumler, J. S. Silencing of host cell CYBB gene expression by the nuclear effector AnkA of the intracellular pathogen Anaplasma phagocytophilum. *Infect. Immun.* **77**, 2385–2391 (2009).
57. Xi, C., Lambrecht, M., Vanderleyden, J. & Michiels, J. Bi-functional gfp-and gusA-containing mini-Tn5 transposon derivatives for combined gene expression and bacterial localization studies. *J. Microbiol Methods* **35**, 85–92 (1999).
58. Quelas, J. I., Mongiardini, E. J., Pérez-Giménez, J., Parisi, G. & Lodeiro, A. R. Analysis of two polyhydroxyalkanoate synthases in Bradyrhizobium japonicum USDA 110. *J. Bacteriol.* **195**, 3145–3155 (2013).
59. Tian, J., Sinskey, A. J. & Stubbe, J. Kinetic studies of polyhydroxybutyrate granule formation in Wautersia eutropha H16 by transmission electron microscopy. *J. Bacteriol.* **187**, 3814–3824 (2005).
60. Pfeiffer, D., Wahl, A. & Jendrossek, D. Identification of a multifunctional protein, PhaM, that determines number, surface to volume ratio, subcellular localization and distribution to daughter cells of poly (3-hydroxybutyrate), PHB, granules in Ralstonia eutropha. *H16. Mol. Microbiol.* **82**, 936–H951 (2011).
61. Jaakola, L., Pirttilä, A. M., Halonen, M. & Hohtola, A. Isolation of High Quality RNA from Bilberry (Vaccinium myrtillus L.) Fruit. *Mol. Biotechnol.* **19**, 201–204 (2001).
62. Szklarczyk, D. et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
63. Zhou, L. et al. ggmsa: a visual exploration tool for multiple sequence alignment and associated data. *Brief Bioinform.* **23** (2022).
64. Team, R. C. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. *(No Title)* (2021).
65. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
66. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

Acknowledgements

We would like to thank Dr. S. Sutela for designing the primers for *CellW* deletion. Research Council of Finland is thanked for financial support, MoPoW project no. 343565, InCell project no. 308766, and METCOM project no. 355127.

Author contributions

Conceptualization: J.J.K., A.M.P., A.C.F. Methodology: J.J.K., A.M.P. Investigation: J.J.K., N.B., H.M.P.T., X.R.V. Visualization: J.J.K., A.M.P., N.B. Funding acquisition: A.M.P., J.J.K. Project administration: A.M.P., J.J.K. Supervision: A.M.P., J.J.K. Writing–original draft: N.B., A.M.P., J.J.K. Writing–review & editing: A.M.P., J.J.K., A.C.F.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44383-026-00026-8.

Correspondence and requests for materials should be addressed to Anna Maria Pirttilä.