Genome mining of *Streptomyces scabrisporus* NF3 reveals symbiotic features including genes related to plant interactions

Corina Diana Ceapă[1☯]*, **Melissa Vázquez-Hernández**[1☯], **Stefany Daniela Rodríguez-Luna**[1], **Angélica Patricia Cruz Vázquez**[1,2,3], **Verónica Jiménez Suárez**[2], **Romina Rodríguez-Sanoja**[1], **Elena R. Alvarez-Buylla**[2], **Sergio Sánchez**[1]*

1 Departmento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México (UNAM), Ciudad de México, México, **2** Laboratorio de Genética Molecular, Epigenética, Desarrollo y Evolución de Plantas, Instituto de Ecología, Universidad Nacional Autónoma de México (UNAM), Ciudad de México, México, **3** Instituto Tecnológico de Tuxtla Gutiérrez,Tuxtla, Gutiérrez, Chiapas, México

☯ These authors contributed equally to this work.

* corinaceapa@iibiomedicas.unam.mx (CDC); sersan@biomedicas.unam.mx (SS)

Abstract

Endophytic bacteria are wide-spread and associated with plant physiological benefits, yet their genomes and secondary metabolites remain largely unidentified. In this study, we explored the genome of the endophyte *Streptomyces scabrisporus* NF3 for discovery of potential novel molecules as well as genes and metabolites involved in host interactions. The complete genomes of seven *Streptomyces* and three other more distantly related bacteria were used to define the functional landscape of this unique microbe. The *S. scabrisporus* NF3 genome is larger than the average *Streptomyces* genome and not structured for an obligate endosymbiotic lifestyle; this and the fact that can grow in R2YE media implies that it could include a soil-living stage. The genome displays an enrichment of genes associated with amino acid production, protein secretion, secondary metabolite and antioxidants production and xenobiotic degradation, indicating that *S. scabrisporus* NF3 could contribute to the metabolic enrichment of soil microbial communities and of its hosts. Importantly, besides its metabolic advantages, the genome showed evidence for differential functional specificity and diversification of plant interaction molecules, including genes for the production of plant hormones, stress resistance molecules, chitinases, antibiotics and siderophores. Given the diversity of *S. scabrisporus* mechanisms for host upkeep, we propose that these strategies were necessary for its adaptation to plant hosts and to face changes in environmental conditions.

Introduction

The symbiotic relationships between microorganisms and their hosts have come under scrutiny, as beneficial active molecules are likely to be expressed during this type of interactions. In

endophytic relationships, microorganisms gain protection from the environment and easy
access to nutrients, in return providing benefits to their host. These include the transformation
of proteins into forms digestible by the host[1,2], synthesis of essential vitamins or amino acids
[3,4], priming of the host immune system[2,5], xenobiotic degradation[6–8], protection
against pathogens[9–11] and overall fitness to the plant[12–14]. It is envisioned that soon,
plant growth-promoting bacteria will begin to replace the use of chemicals in agriculture, horticulture, silviculture, and environmental cleanup strategies[15–18]. The genus *Streptomyces*
exhibits a remarkable potential to produce secondary metabolites and exposes diverse and
unique functional capacities as well as biological activities. The advent of ~omics technologies
and the ability to analyze the resulting data extended our ability to uncover new molecules for
medical and industrial use and to understand their mechanisms[19–20].

The genomes of terrestrial *Streptomyces*, which are large in comparison to those of other
bacteria, indicate that these organisms possess diverse carbon, nitrogen, and iron uptake systems, multiple genes for developmental regulation, and a remarkable number of gene clusters
to produce biologically active small molecules[21–23]. Ample evidence suggests that *Streptomyces* are quantitatively and qualitatively important in the rhizospheres of plants, where they
may influence plant growth and protect plant roots against invasion by root pathogens[24–
26]. Also, a number of *Streptomyces* have an endophytic lifestyle, living within a plant for at
least part of its life cycle without causing apparent disease[27–29]. Endophytic *Streptomyces*
spp. are documented to promote plant growth[25,30], enable nutrient procurement[31], phytohormone synthesis[32–34], antibiosis and local competition against pathogens[31,35], and
induction of systemic defense responses[24,29,36].

Streptomyces scabrisporus NF3 is an endophyte belonging to the Actinobacteria isolated
from the tree *Amphipterygium adstringens*, endemic of Mexico (Rodriguez-Peña et al., in preparation). Samples of the tree were collected in the town of Barranca Honda in Yautepec, Morelos. The extract of this tree is used in traditional native medicine for the treatment of fever,
fresh wounds, hypercholesterolemia, cholelithiasis, gastritis, gastric ulcers, gastrointestinal
cancer and other inflammatory conditions[37–40].

While few strains of the species have been isolated and characterized in the past (mostly
from soils) [35,41–43], only one strain of *S. scabrisporus* was sequenced and published so far:
DSM 41855, isolated from soil from Japan. The genome of this strain is composed of 199 contigs, making it very difficult to use it in comparative analyses[44]. The few genomes sequenced
of *S. scabrisporus* may be linked to the difficulty in obtaining cultured isolates, due to its slow
growth rate in laboratory conditions and to its occurrence as an endophyte. *S. scabrisporus* has
a great potential to produce diverse antimicrobial compounds against a wide range of microorganisms, among which, characterized so far are okilactomycin[43], hitachimycin[41,44] and
their derivates. Antitumoral activity for alborixin has also been reported[45].

Here were report the exploration of the draft genome sequence of *S. scabrisporus* NF3 (SS
NF3), and its comparison to ten complete genomes of biologically relevant strains, among
which seven are taxonomical relatives from the family *Streptomycetaceae*: *S. venezuelae* ATCC
15439 (SV), *S. coelicolor* A3 (SC), *Kitasatospora setae* KM-6054 *(K. setae, KS)*, *S. avermitilis*
MA-4680 (SA), *S. hygroscopicus* subsp. *limoneus* KCTC 1717 (SH), *Streptomyces* sp. TLI_053
(ST), *S. griseus* subsp. *griseus* NBRC 13350 (SG), and three more distantly related: *Bacillus
anthracis* Sterne (BA) (mammalian intestinal pathogen), *Corynebacterium pseudotuberculosis*
C231 (mammalian blood pathogen) (CP) and *Bifidobacterium breve* DSM 20213 (mammalian
intestinal commensal) (BB). Several comparisons were made including the sequences of the
incomplete *S. scabrisporus* DSM 41855 (SS DSM 41855) (S1G Table). The genomes were purposely selected not to be known endophytes, with the intention to detect an increase in gene
abundance for plant interaction related genes in the genome of interest.

The comparison revealed an enrichment for genes associated with amino acid production and utilization, protein secretion, secondary metabolite and antioxidant production, and xenobiotic degradation, potentially indicating that SS contributes to the metabolic enrichment of its hosts. Exploration in the genome of SS NF3 revealed the presence of a large assortment of genes coding for the production of potential host interaction molecules like plant hormones, stress resistance molecules, chitinases, antibiotics, antibiotic resistance and siderophores. Among these genes, various numerous encode for prospective production of novel molecules.

Results and discussion

Genome assembly and gene prediction

The genome sequence of SS NF3 was previously reported [28]. The genome is linear and has a 71.61% of GC content (Fig 1), falling into the normal range for a streptomycete and a size of 10.95 Mbp, falling in the top 5% of known *Streptomyces* genomes (S1 Fig). The gene prediction and annotation were made with two platforms, Prokaryotic Dynamic Programming Genefinding Algorithm (PRODIGAL) and the Rapid Annotations using Subsystems Technology (RAST) (S1A and S1B Table, respectively) and complemented with manual annotations based on domain and motif searches (S1A and S1C Table). Further analysis of the annotated genome was performed with Pathosystems Resource Integration Center (PATRIC) (Fig 1, S1A Table). The genome is predicted to contain 10,492 coding sequences (CDSs), 156 repeat regions, 60 tRNAs and 12 rRNAs, 5,691 predicted genes, 4,729 genes coding for hypothetical proteins and 382 pseudogenes (Fig 1). The origin of replication *oriC* is located after the *dnaA* gene (PEG.3184) at position 333,394 of scaffold 1, with the *oriC* repeats spanning over a 561 bp
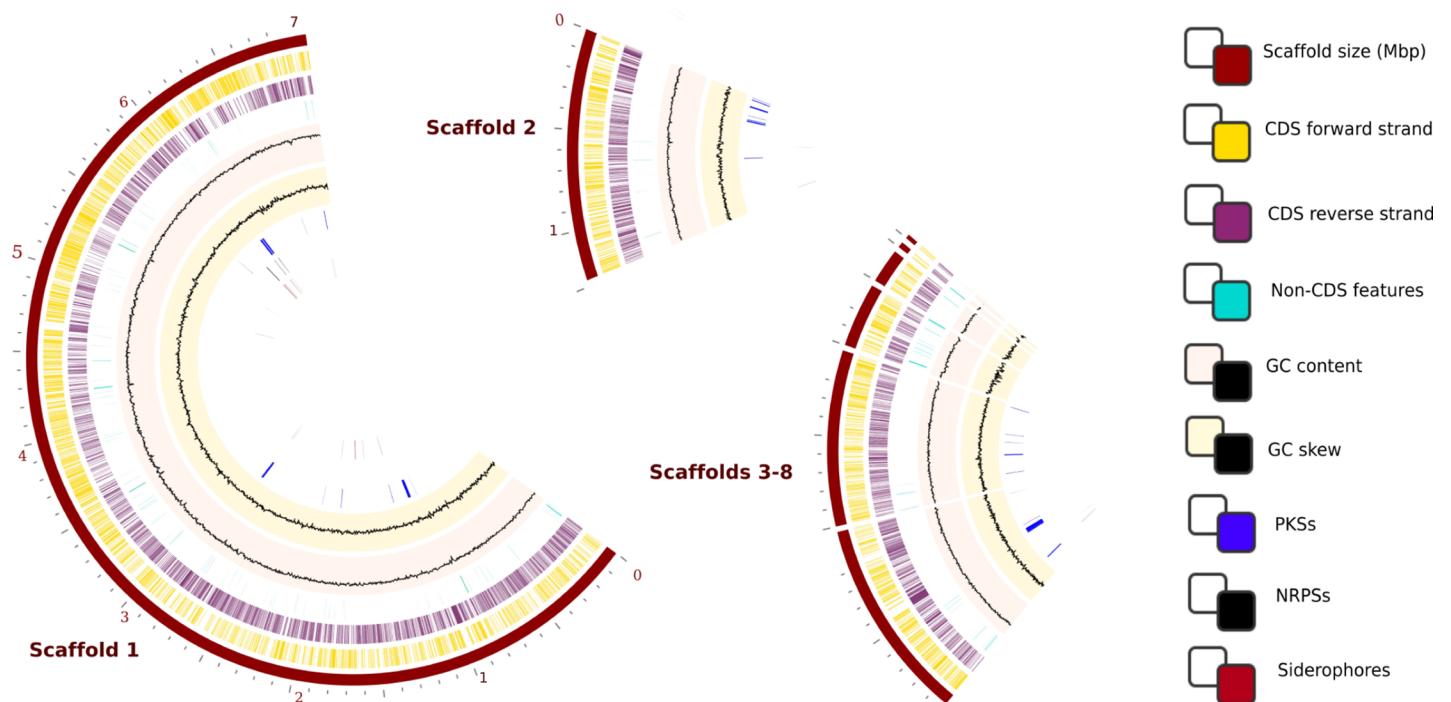
Fig 1. Genome characteristics for *S. scabrisporus* NF3. The circles refer to (from outside to inside): grey—Position label (Mbp): scaffold size in kbp; dark blue–contigs: organization of the scaffolds from the largest to the shortest; green–coding sequences on the forward strand (CDS FWD); purple–coding sequences on the reverse strand (CDS REV); blue–non-CDS features: location of tRNA, rRNA and pseudogenes; light purple–GC content; light orange–GC skew; dark green–location of genes encoding for polyketides; orange–location of genes encoding for nonribosomal peptide synthetases.

https://doi.org/10.1371/journal.pone.0192618.g001

region. Table 1 shows the genome characteristics, as well as information about the genomes used for comparison.

Four highly conserved sequences (16S rRNA, gyrB, rpoB, recG) were used to construct a Neighbor Joining (NJ) phylogenetic tree of all compared genomes (Fig 2), which represents the evolutionary relatedness of the strains. As expected, an outgroup was formed with sequences of BA, CP and BB, while SS strains appear highly related. The position of KS between the other *Streptomyces* species was expected since this strain also belongs to the Streptomycetaceae family.

Orthologous and orphan genes

The proteomes of SS NF3 and of strains from 10 other different species (with complete genomes) with various taxonomic relatedness (Fig 2) were analyzed and compared using the PATRIC FIGFam software to identify orthologous gene families (OGF) (Fig 3, S1G Table). From the total of 5,784 OGF, only 421 families (7.3%) were shared by all 11 genomes (excluding single gene families). The unique OGFs from each genome ranged from 2 (*S. venezuelae* ATCC15439) to 266 (*B. anthracis* Sterne) (S2 Table). The genome with the most similar gene family profile to SS NF3 was *Streptomyces* sp. TLI_053, accounting for 1,842 common OGFs (Fig 3, S2 Table).

Compared to the other genomes, SS NF3 includes 166 unique genes whose orthologues were absent from all the comparative genomes used in this study; they were therefore selected as "orphan genes" (OG). Among these 166 OG, 129 belong to single-member gene families and 37 genes belong to multi-membered gene families (S1D Table). The second largest gene family for the OG (after hypotheticals) includes three glutathione S-transferase (GST) omega paralog genes. The GST omega subfamily is involved in cellular detoxification by catalyzing the GSH dependent reduction of protein disulfides, dehydroascorbate, and monomethylarsonate, activities which are more characteristic of glutaredoxins. Along with this GSTs, one iron-containing redox enzyme (DUF3050) are predicted to be involved in the reduction of oxidative stress. As the reduction of environmental oxidative stress is considered a factor positively contributing to plant growth[17,47,48], the presence of these unique genes could support strain SS NF3's resistance to changing environmental conditions and indicate an endophytic lifestyle. These antioxidant genes are shared with SS DSM 41855. In SS NF3, these genes appear to be complementing the large array of xenobiotic degradation pathways present in the genome, such that NF3 could contribute to bioremediation of soils and water. Interestingly, among the NF3-specific orphan proteins, one unique sigma factor 24 (predicted to control adaptation to high temperatures in *E. coli*[49]), 2 peptidases and three luciferase genes (forming an operon) are present. Many of these unique genes belong to secondary metabolite operons as well as transporters (for instance one branched-chain amino acid ABC transporter). Alternative sigma factors activation was linked to reduction in the secretion of secondary metabolites in other *Streptomyces* sp.[50–53] and it is possible that its removal can allow SS NF3 to be used as a cell factory for its own or for heterologous metabolites production[49]. As no regulatory controller can be identified with the current data, remains to be investigated further if the luciferase operon is active and under which conditions.

The PFAM domains superfamilies that are overrepresented in the orphan genes are the P-loop containing nucleoside triphosphate hydrolase superfamily and amidohydrolase superfamily (S1C Table). These unique proteins may determine the host specificity of this organism, as the group encompasses detoxification enzymes (urease, creatinine amidohydrolase), antibiotic resistance proteins (beta-lactamase) (S1N Table) as well as cell wall modification enzymes

Table 1. General genome characteristics.

NCBI Taxon ID	Genome Name	Genome Status	Culture Collection	BioProject Accession	Assembly Accession	GenBank Accessions	Contigs	Genome Length	GC Content	PATRIC CDS	RefSeq CDS	Cell Shape	Sporulation	Oxygen Requirement	Habitat
85011	*Streptomyces scabrisporus* NF3	WGS	IIBiomedicas, UNAM	PRJNA377273	GCA_002024165	GCF_002024165	8	10958352	71,61	10491	9020	-	Yes	Aerobic	Plant
100226	*Streptomyces coelicolor* A3(2)	Complete		PRJNA242	GCA_000203835	AL645882, AL589148, AL645771	1	9054847	72	8325	8154	Branched filament	Yes	Aerobic	Multiple
1855352	*Streptomyces* sp. TLI_053	Complete		PRJEB16370	GCA_900105395	LT629775	1	9900053	73,63	8757	8065	-	-	Aerobic	-
227882	*Streptomyces avermitilis* MA-4680	Complete	ATCC 31267	PRJNA189	GCA_000009765	BA000030, AP005645	1	9119895	70,7	8106	7676	Branched filament	Yes	Aerobic	Multiple
260799	*Bacillus anthracis* Sterne	Complete		PRJNA236483	GCF_000832635	CP009541, CP009540	1	5409120	35,28	5854	5622	Bacilli	-	Aerobic	-
264445	*Streptomyces hygroscopicus* subsp. *limoneus* KCTC 1717	Complete	KCTC 1717	PRJNA302679	GCF_001447075	CP013219, CP013220	1	10537932	71,96	9983	8983	-	-	Aerobic	Terrestrial
452652	*Kitasatospora setae* KM-6054	Complete	ATCC 33774	PRJNA19951	GCA_000269985	AP010968	1	8783278	73,1	7709	7569		Yes	Aerobic	Terrestrial
455632	*Streptomyces griseus* subsp. *griseus* NBRC 13350	Complete		PRJNA20085	GCA_000010605	AP009493	1	8545929	72,2	7294	7136	Branched filament	Yes	Aerobic	Multiple
518634	*Bifidobacterium breve* DSM 20213	Complete		PRJDB57	GCF_001025175	AP012324	1	2269415	58,89	2016	1929	-	-	Anaerobic	-
54571	*Streptomyces venezuelae* ATCC 15439	Complete	ATCC 15439	PRJEB10583	GCF_001406115	LN881739	1	9034396	71,75	8487	8682	-	-	Aerobic	-
681645	*Corynebacterium pseudotuberculosis* C231	Complete		PRJNA40875	GCA_000144675	CP001829	1	2328208	52,19	2209	2053	Rod	No	Facultative	-

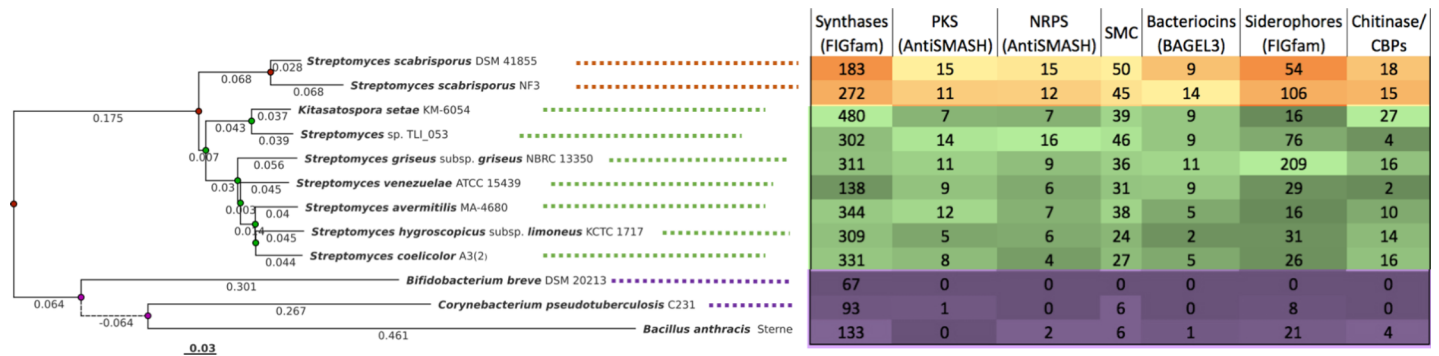|	Synthases (FIGfam)	PKS (AntiSMASH)	NRPS (AntiSMASH)	SMC	Bacteriocins (BAGEL3)	Siderophores (FIGfam)	Chitinase/ CBPs
Streptomyces scabrisporus DSM 41855	183	15	15	50	9	54	18
Streptomyces scabrisporus NF3	272	11	12	45	14	106	15
Kitasatospora setae KM-6054	480	7	7	39	9	16	27
Streptomyces sp. TLI_053	302	14	16	46	9	76	4
Streptomyces griseus subsp. *griseus* NBRC 13350	311	11	9	36	11	209	16
Streptomyces venezuelae ATCC 15439	138	9	6	31	9	29	2
Streptomyces avermitilis MA-4680	344	12	7	38	5	16	10
Streptomyces hygroscopicus subsp. *limoneus* KCTC 1717	309	5	6	24	2	31	14
Streptomyces coelicolor A3(2)	331	8	4	27	5	26	16
Bifidobacterium breve DSM 20213	67	0	0	0	0	0	0
Corynebacterium pseudotuberculosis C231	93	1	0	6	0	8	0
Bacillus anthracis Sterne	133	0	2	6	1	21	4

Fig 2. Phylogenetic tree of analyzed genomes. Phylogenetic trees were based on 4 concatenated proteins (16S rRNA, gyrB, rpoB, recG) conserved in all compared genomes using Kalign [46]. The tree was generated from a multiple sequence alignment using a Neighbor Joining algorithm. Bar shows substitution per nucleotide. The right side of the figure contains information about the abundance of genes of various functionalities discussed in the following paragraphs (synthases, siderophores and chitinases/chitin binding proteins (CBP)–computed using FIGfams, polyketide synthases (PKS), non-ribosomal peptide synthetases (NRPS) and total number of secondary metabolites clusters (SMC)–computed using antiSMASH, and bacteriocins–analyzed using BAGEL3).

https://doi.org/10.1371/journal.pone.0192618.g002

(peptidoglycan N-acetylglucosamine deacetylases), all of which can contribute to plant interactions, as will be discussed further.

Biosynthetic capabilities (KEGG metabolic mapping)

The endophyte's functionality, as revealed by their gene content, is the result of their continuous and complex interactions with the host plant as well as other members of the host microbiome[16]. Understanding these functions at a molecular level and their participation in the global picture of plant-microbe interactions can be used to improve techniques of plant growth, biocontrol and bioremediation[54]. Microbiota research involving endophyte communities suggests high rates of metabolic activity related to carbohydrate and amino acid metabolism for these microorganisms[14]. Therefore, the biosynthetic capabilities of SS NF3 were inferred using metabolic mapping of its genome information onto the KEGG database.

Amino acid metabolism. Genomic analysis of the NF3 strain using the ModelSEED database (based on KEGG mapping) suggested prototrophy for all proteinogenic amino acids except pyrrolysine (S1F Table). From the non-proteinogenic amino acids, there are complete biosynthetic pathways for homocysteine, homoserine, hypotaurine, D-arginine, D-alanine, beta-alanine, D-glutamate, D-glutamine and ornithine. There is a high diversity of amino acid posttranslational modification genes, which means that they can produce extremely highly modified peptides, including D-amino acids, methylated amino acids and dehydrated amino acids and dozens of modified residues. The capacity to grow without needing external amino acids can be an advantage in poor nutrient environments and implies that the SS NF3 strain is likely to, at least in some segments of its life cycle, grow in soil communities; growth of another SS strain on various nitrogen sources was reported for peptone, yeast extract, (NH4)2SO4, NH4NO3 and L-asparagine[35].

Amino acids production is considered essential for establishing plant-bacteria interactions. Comparisons between endophytes and pathogenic strains of the same species indicate that endophytes would use a broader array of available amino acids and other acids as compared to the pathogens[55], suggesting a trend of endophyte prototrophy for amino acids. The biosynthetic capacity of a wide range of amino acids is also indicative of its advantage as an endophyte, since there are zones in plants which are poor in nutrients, becoming able to grow anywhere in the plant and providing nutrients to the microbiota existing there too. Bacterial protein secretion systems were previously successfully used for determining plant-bacterial
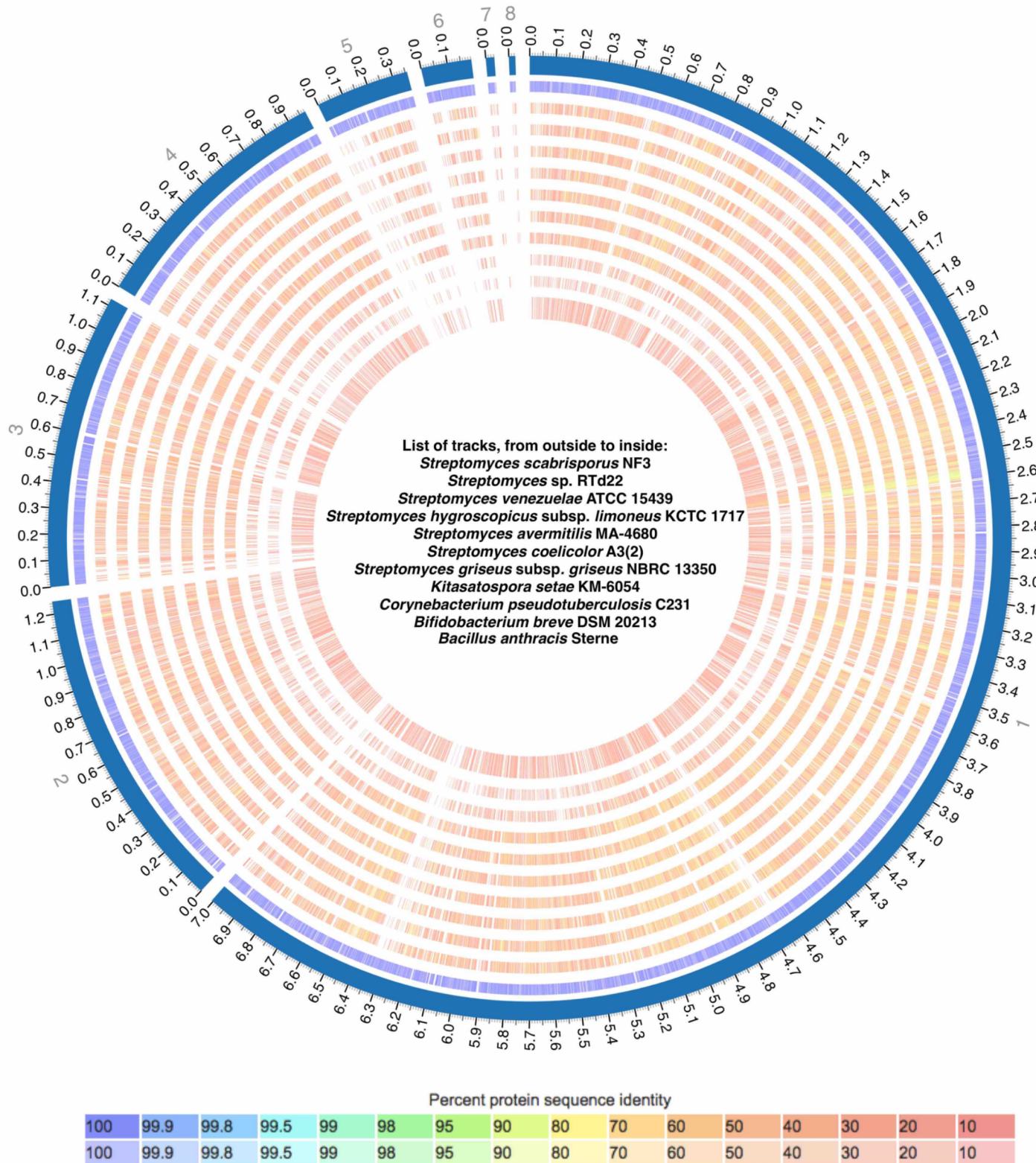
Fig 3. Proteome comparison between NF3 and the 10 complete genomes analyzed. The genomes are (from outside to inside): SS, SR, SV, SH, SA, SC, SG, KS, CP, BB, BA. The exterior solid line circle indicates the scaffolds of the reference genome–SS NF3. Each circle is comprised of lines, each of them indicating one protein homologous to a protein of the reference genome. An identity matrix indicating the relationship between the colors used and protein similarity is placed below the figure. The image was obtained from PATRIC utilizing the tool Proteome Comparison.
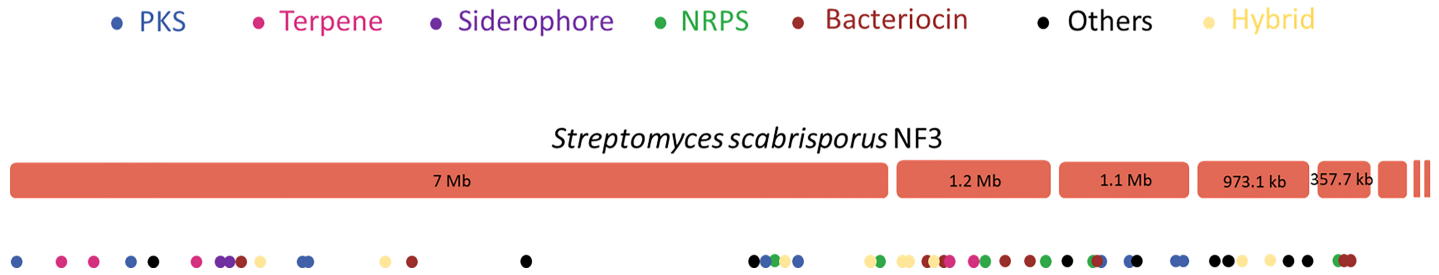
Fig 4. Localization of secondary metabolite clusters in *S. scabrisporus* NF3.

interactions[1]. Transport systems in SS NF3 are dominated by genes involved in amino acid, peptide and protein transport (55% of transport genes are encoded by 92 genes), further supporting the idea that amino acids and peptides are essential for this bacteria's interaction with its environment. An in-depth analysis of the specificities of these transporters by mutagenesis could provide further insights into the lifestyle of SS NF3.

Secondary metabolites. *In silico* genome analyses with the Secondary Metabolite Analysis Shell (antiSMASH) algorithm[56], BAGEL[57] and PRISM[58], and metabolic mapping using ModelSeed[59] on the SS NF3 genome revealed the presence of 413 genes in 45 clusters involved in the production of 59 different secondary metabolites, spanning polyketides (PKS–polyketide synthases), non-ribosomal peptides (NRPS–non-ribosomal peptide synthases and bacteriocins (Fig 4, S1K, S1L and S1P Table). The location of most clusters resides in the extremes of the larger scaffold and in the second largest scaffold, which is expected for *Streptomyces*, where most secondary metabolites clusters regularly reside at the extremities of the linear chromosome to prevent loss of central metabolism genes (Fig 4).

A comparison between the abundance of secondary metabolite clusters of distinct types between the two SS strains reveals very similar amounts for the most abundant compounds, with higher numbers of NRPS and PKS for strain DSM 41855. Some clusters are shared with strain DSM 41855 and their products have previously been characterized, for instance alkylresorcinol, griseobactin, steffimycin and streptomycin. A high similarity ($> 40\%$) with clusters previously characterized in other bacteria also exists for clusters potentially encoding sporepigment, polyoxypeptin, piericidin A1 and hopene. From the remaining 38 clusters, 17 show relative similarity to known clusters (5% to 40%) and 20 are likely encoding completely new compounds. Structure prediction was possible using antiSMASH and PRISM for 15 compounds, among which 12 appear to be for molecules newly encoded in this genome (Fig 5).

Xenobiotic degradation. KEGG mapping of the main metabolic pathways for the genome of SS NF3 shows an enrichment of 3% in genes belonging to xenobiotic degradation pathways compared to the other genomes used in this study (data not shown, data from PATRIC pathways).

The xenobiotics predicted to be degraded by SS NF3 can be classified based on the strain's capacity to use them for central metabolism. Directly feeding into the central metabolism are toluene, geranic acid, vanillate, benzene, nitrobenzene, phenol, bisphenol A, aniline, hydroquinone, salicylate, phenyl acetonitrile and phenylacetamide, while compounds that can be degraded to less toxic molecules, but do not supply energy for metabolism are urea, 6-mercaptopurine, fluorouracil, cyclohexanone, 1,1,1-Trichloro-2,2-bis (4-chlorophenyl)ethane (DDT), trichloroethylene, trinitrotoluene (TNT), trichloroethane, phenanthrene, pyrene, anthracene, 4-hydroxyphtalate, styrene, acrylamide, p-cymene and benzoic acid. Certainly, a closer view of the toluene and xylene degradation pathways in SS NF3 indicates the presence of more than one degradation route. The large diversity of genes involved in xenobiotic degradation and
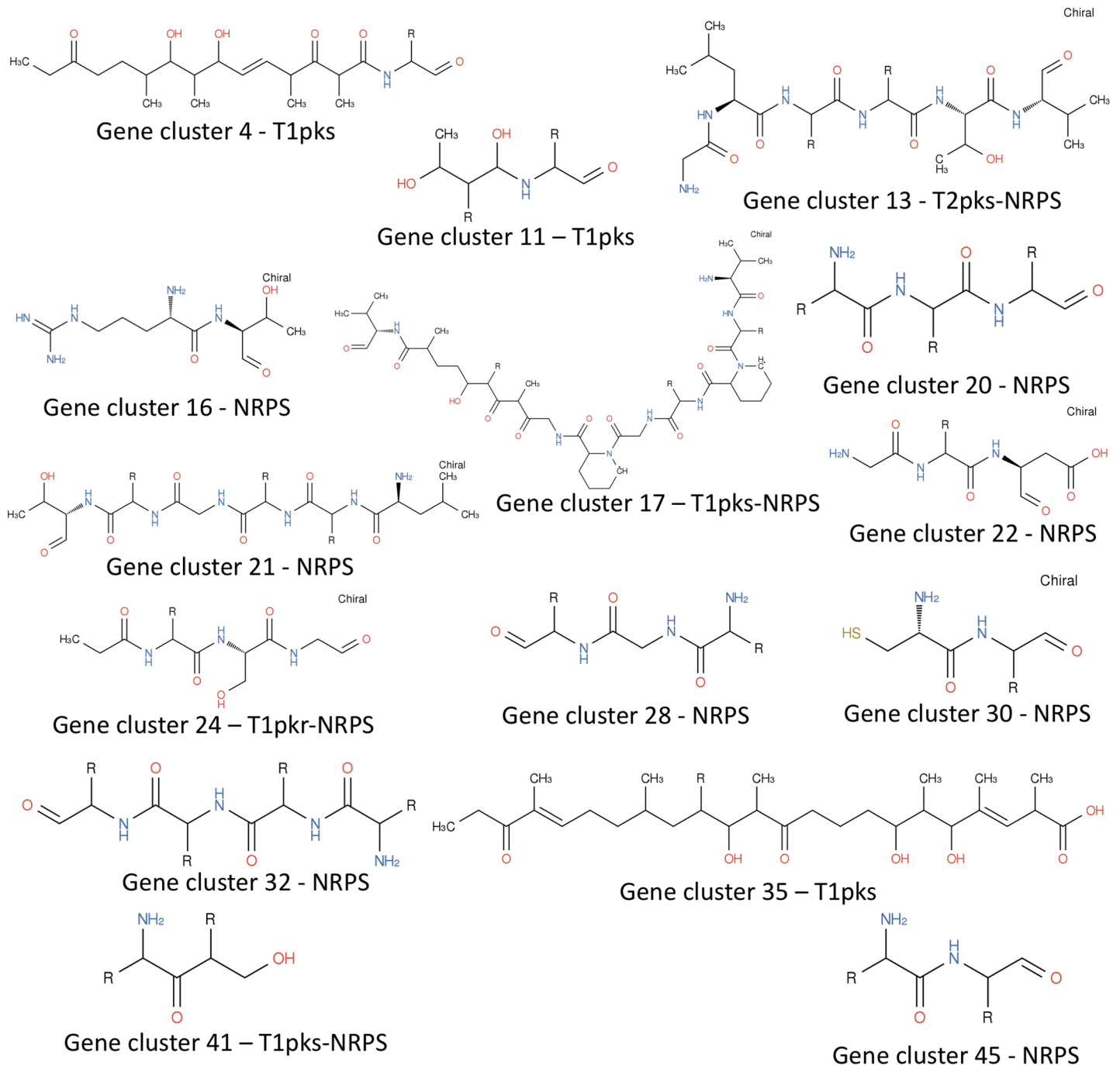
Fig 5. Predicted structures of secondary metabolites of *S. scabrisporus* NF3. The analyses were performed using AntiSMASH and PRISM.

survival in toxic environments serves as an indicator of the extraordinary adaptability of SS NF3 and its elevated potential to be used in bioremediation or bioaugmentation of contaminated sites. In addition, these genes confer a general advantage as an endophyte, as it could help the host plant to degrade toxic compounds absorbed by the roots or feed off carbon sources which the plant does not utilize readily or produces in excess.

Carbohydrate-active enzymes (CAZymes). CAZymes are enzyme families of structurally-related catalytic and carbohydrate-binding modules (or functional domains) of enzymes that degrade, modify, or create glycosidic bonds, like glycoside hydrolases, glycosyl transferases, polysaccharide lyases, carbohydrate esterases and some auxiliary redox enzymes such as polysaccharide mono-oxygenases[60]. In the context of plant interactions, various novel CAZymes with a role in degradation of plant polysaccharides could be identified by mining newly sequenced genomes or metagenomes[14,61–63]. Some studies go as far as using the presence of plant cell-wall degrading enzymes as a basis for bacterial classification as endophytes or as pathogens[14], under the hypothesis that endophytes are positively selected by the plant and only need to produce cellulases and pectinases, while pathogens requiring the degradation of the entire cell wall implement lignocellulose-degrading enzymes. In the SS NF3 genome, 1,071 genes classified in 130 CAZymes families contain carbohydrate-active domains (10.2% of the total gene count), among which the most abundant superfamilies are: histidine phosphatases (197 ORFs), Hsp70 chaperones (56 ORFs), dehydrogenases (50 ORFs), cellulose binding (29 ORFs), beta-lactamases (29 ORFs), response regulators (23 ORFs) and extracellular peptidases and solute-binding proteins (50 ORFs). Among the strain specific genes (76 CAZymes), 21 are predicted to be extracellular, cell wall or membrane attached, and are dominated by Hsp70 chaperones, solute-binding proteins and response regulators.

Root colonizing bacteria are sometimes equipped with enzymes enabling penetration of the root tissues such as lignocellulose or pectin degrading enzymes. The CAZymes involved in lignin degradation include lytic polysaccharide mono-oxygenases (LPMO) as well as auxiliary activities containing both ligninolytic enzymes and lytic polysaccharide mono-oxygenases. At least 29 unique genes from the SS NF3 genome belong to the AA families, while no enzymes of the LPMO type could be found (S1E Table). Using the FOLy database[64] to complement this data, we could establish that the SS NF3 genome contains two endo-1,4-beta-xylanases and 17 potential cellulases (S1I and S1J Table). This complex system of lignin and pectin degrading enzymes indicates a probable point of entry of SS, as a bacterium most commonly isolated from soil, through the plant root system, to reach other distant plant tissues.

Interference with plant metabolism. The close association between an endophyte and a living host is marked by the secretion of effector proteins that augment the host defenses and modify the host metabolism. The best-studied mechanisms of bacterial plant growth promotion include plants with resources/nutrients that they lack such as fixed nitrogen, iron, and phosphorus[65].

Nitrogen fixation is a metabolic property present in many plant-associated bacteria, especially the ones adapted to grow in the rhizosphere[1]. While no indication from the literature exists of *Streptomyces* strains able to fix nitrogen, these bacteria can participate in the formation and enhancement of root nodules[25, 66–68]. In SS NF3, the only nodulation gene we identified so far is *nolO* (PEG.9541), possibly mediating early stage recognition between bacteria and its host[69]. Despite not directly participating in nitrogen fixation, SS NF3 has all the genomic requirements to metabolize fixed nitrogen, through the urea pathway. The hydrolysis of urea is catalyzed by urease, yielding ammonium and carbon dioxide. The ammonium is then converted into glutamine by the glutamine synthetase encoded by the *glnA* gene. The bacterial urease is a multi-subunit protein complex, which consists of three structural subunits and several accessory proteins. We identified a cluster of genes coding for all subunits of the urease complex in the SS NF3 genome (PEG.1320 – 1322—urease, PEG.1318 –ureG, PEG.1319 - ureF). The genome also contains various copies of the *glnA* gene for glutamine synthetase (PEG.2332, PEG.4644, PEG.4669, PEG.6800, PEG.8087, PEG.8680), and a gene for a regulator of the GlnA proteins in response to the levels of nitrogen (PEG.4645).

In addition, second in abundance after amino acids and proteins transport genes, the genome contains a wealth of genes involved in the transport of nitrates/nitrites (PEG.426), iron (PEG.1363-1365; PEG.3259; PEG.5593; PEG.6834; PEG.8844; PEG.9278) and phosphorus (2 operons: PEG.1422 – 1429 and PEG.2798 – 2802). Among the iron transporters, there are some that transport siderophores bound to iron molecules. Diversity in siderophores of SS NF3 is discussed in the following section.

Adaptive genomic changes specific to endophytes recognized in the SS NF3 genome

Symbiosis-related genes. Studies have revealed that beyond metabolism, other factors affect colonization and modulation of the bacteria—host interactions. Variation was observed in the distribution of essential genes related to hormone production and signaling, surface attachment, secretion, and transport between endophytes compared to pathogenic invasive bacteria [14, 65].

Virulence genes. Genes possibly involved in pathogen–host interactions (PHIs) were identified by their homologs in the PHI-base database [70]. About 7% (739 OGs) of SS NF3 genes had a hit in the PHI database, but only 2.5% (261 OGs) are extracellular or cell wall associated. Interproscan analysis of motifs present in the extracellular or cell wall associated virulence factors indicates the majority to be transporters and oxidoreductases (S1M Table), showing that SS NF3 acquired proteins could support symbiotic interspecies interactions.

Non-metabolic plant-endophyte interactions. In addition to metabolic collaboration, described in the previous paragraphs, endophytes use other strategies to modulate plant responses which encompass production of: a) compounds interfering with plant hormone signaling; b) antioxidants; c) chitin degradation molecules as a defense mechanism against fungi; d) siderophores for iron-storage; and e) antibacterial molecules such as bacteriocins to combat pathogens. Molecules belonging to these categories are abundant products of genes in the SS NF3 genome, as will be discussed below.

Plant hormones play key roles in plant growth and development and in the response of plants to their environment[13,71]. Microorganisms may also produce or modulate phytohormones under *in vitro* conditions[13,72]. Several microorganisms including endophytes can alter phytohormone levels and thereby affect the plant's hormonal balance and its response to stress[25,73–74]. For instance, plant growth regulators (PGRs) such as auxins, cytokinins, gibberellins, strigolactones, brassinosteroids can be secreted by endophytes to enhance the host defense[75]. Soil or plant-associated microorganisms could produce IAA (indole-3-acetic acid) to promote plant growth[76–78]. In SS NF3, IAA can be produced from intermediates of the precursor tryptophan using the indole-3-pyruvic acid (IPyA) pathway, for which the genes are present in the genome (aldehyde dehydrogenase—PEG.8491 and indole-3-pyruvate decarboxylase—PEG.7861) (Fig 6).

Furthermore, in SS NF3, three individual genes and one operon could potentially modulate plant responses, by the synthesis and degradation of phytoene to carotenoid compounds. Carotenoids are natural pigments as well as lipid-soluble antioxidants[79]. In plants, carotenoids play important regulatory roles in optical system assembly, light harvesting and protection, photo-morphogenesis, non-photochemical quenching, lipid peroxidation, and seed dormancy and aging[79]. Phytoene synthase (PSY) is the first step in the biosynthesis pathway of carotenoids in plants, followed by phytoene dehydrogenase, an enzyme that converts phytoene into zeta-carotene. In SS NF3, an operon comprising the three genes coding for phytoene synthase, phytoene desaturase and squalene synthase is associated with the synthesis of carotenoids and hopanoids (Fig 6). Hopanoids act similarly to sterols in eukaryotes in that
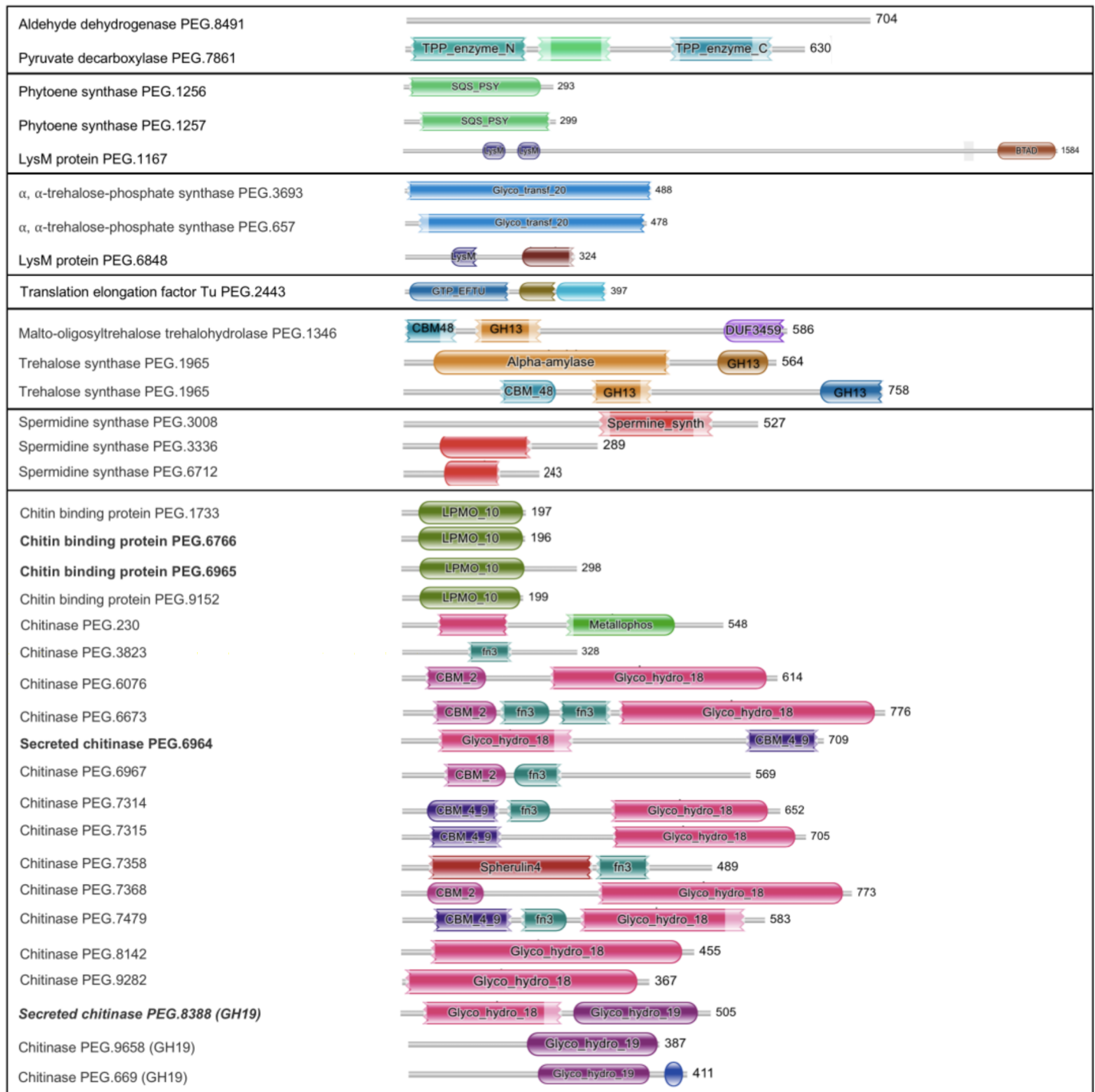
Fig 6. Domain composition for proteins from SS NF3 potentially involved in plant interactions. Legend: green–phytoene synthase; dark purple–LysM domain; dark blue–EF-Tu domain; pink–chitinase domain (GH18), red–spermidine synthase; orange–amylase. The analysis was performed using the HMMER scan tool from EBI. Only domains relevant for potential host interactions were annotated.

they condense lipid membranes and reduce permeability. In the case of prokaryotes, they provide stability under of high temperatures and extreme acidity due to the rigid ring structures [80,81]. Indeed, upregulation of squalene-hopene cyclase occurs in certain bacteria in the

presence of hot or acidic environments[82]. In this case, carotenoids and hopanoids are likely to be shared with the host plant when conditions require it.

Rovenich et al. (2014) have recently discussed the possibility of ecological roles for plant effector proteins, especially those with the LysM/chitin-binding domain[83]. LysM domains are involved in non-covalently binding to the peptidoglycan of other bacteria[84]. Two genes with associated LysM domains in the genome of SS NF3 appear to be of interest to this discussion (Fig 6). One of them, PEG.1167, belongs to a protein of the tetratrico peptide repeat superfamily (TPR), which was implicated in plant hormone signaling in gibberellin, cytokinin and auxin responses as well as ethylene biosynthesis through direct interaction of its TPR domains with a 1-aminocyclopropane-1-carboxylate synthase isoform of plants[85]. The other lysM domain is associated with a cysteine peptidase Clan CA (PEG.6848). This class of enzymes is mainly used in plants to mobilize storage proteins in seeds. Protein bodies of seeds contain protease precursors. They are activated after germination and start degradation of the stored proteins[86]. Involvement of plant endophytes in seed development could explain instances in which association of plants with certain microbes lead to better crop yields[1,12].

Another type of signaling involves recognition by the plant immune system and induction of an elevated state of immunity that acts as prevention in the case of a pathogenic attack. EF-Tu is the most abundant bacterial protein that is recognized as MAMP by the receptor EFR in *Arabidopsis* sp. and other Brassicaceae family members[87]. One gene producing the translation elongation factor Tu (PEG.2443) is annotated in the SS NF3 genome (Fig 6), indicating that plant colonization by this Actinobacterium could further elevate plant resistance to pathogens.

Plant colonization by bacteria requires a high tolerance for environmental stresses, especially osmotic stress, and production of trehalose as one of the main factors contributing to such resistance[73]. Trehalose of bacterial origin was among the most strongly induced metabolites produced following inoculation of soybean (*Glycine max*) root hairs by the nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum*[73], functioning as an inducer of the endophytic relationship. In the genome of SS NF3, two operons for trehalose synthases, two independent synthases (PEG.3693 and PEG.657) and two degradation genes are contributing to trehalose metabolism. The first operon is formed by a malto-oligosyl-trehalose-synthase and a malto-oligosyl-trehalose trehalohydrolase (PEG.1346-1347). The second operon (PEG.1965 – 1967) codes for a protein with an alpha amylase with a malt amylase domain, one phosphotransferase and two trehalose synthases. The presence of seven genes involved in at least three different pathways for trehalose production denotes a high dependence of the bacteria for this carbohydrate and its probable importance in SS NF3's lifestyle. Other synthase complexes involved in bacterial stress responses with many homologs in SS NF3 (PEG.3008, PEG.3336 and PEG.6712) are spermidine synthases. Spermidine derivates contribute to tolerance to drought and salinity in plants.

Another essential mechanism for bacteria to become useful partners in their interactions with plants is through fighting off pathogenic attacks. Siderophores, bacteriocins and antibiotics are three of the most effective and well-known mechanisms that an antagonist can employ to minimize or prevent phytopathogenic proliferation[88]. In addition, chitinases target fungal cell walls to release chitin fragments that activate immune receptors, leading to further chitinase accumulation to induce hyphal lysis. Chitinases are enzymes that catalyze the hydrolysis of the beta-1,4-N-acetyl-D-glucosamine linkages in chitin polymers present in the cell wall of microorganisms [89]. They are produced mostly in filamentous fungi, which are known to produce up to 20 different chitinases. *Serratia marcescens* is one of the bacteria with the highest reported chitinolytic activity, counting five different chitinases[90–92]. The SS NF3 genome contains 14 chitinases belonging to the glycoside hydrolase (GH) 18 family and 4 chitinases

from the GH19 family (S1O Table, Fig 6). A BLAST analysis shows that these chitinases are unique for the species SS but shared with the other sequenced strain of the species: DSM 41855. Eight of them have a carbohydrate binding domain. Among them, four genes (PEG.6964-6968) form a cluster containing one secreted chitinase with a degradation domain (Fig 6), two chitin-binding proteins and an esterase. Exceptionally, the SS NF3 genome also contains four chitin-binding proteins, with a total of 20 chitin degradation or binding genes (Fig 6, S1O Table). In addition, a NJ tree based on a MSA produced using MUSCLE containing all the chitinase genes from the compared genomes (116 sequences) (Fig 7A) indicates that the SS NF3 chitinases are not paralogs, encompassing a large structural variation and belonging to different branches of the tree.

This raises the question of what the necessity could be for this high number of various chitinases. The foreseen explanation is that they could be involved in the biological control against particular fungal or nematode plant diseases[88,93]. In support of this hypothesis, recent research evaluating the chitinase activity of 63 different strains of Actinobacteria isolated from soil samples of China found that the strain identified as SS Bn035 displayed the best chitinase activity, indicating that chitin degradation could be a species characteristic. The antifungal activity of strain Bn035 was tested against four pathogenic fungi (*Bipolaris sorokiniana*, *Fusarium oxysporum*, *Rhizoctonia solani* and *Pythium capsici*), involved in turfgrass root rot disease, with highly positive results[35]. The chitinolytic activity of SS NF3 in laboratory conditions was assayed in solid plates containing colloidal chitin as a sole carbon source. We could observe that the bacteria present chitinolytic activity after 7 days of incubation and supports the formation of spores after 14 days of incubation (Fig 7B).

Many bacteria closely interacting with plants produce secondary metabolites as agents needed for nutrient uptake (Fig 8). For instance, iron bioavailability is reduced in aerobic environments, including soil. To overcome this limitation, microorganisms have developed different strategies, such as iron chelation by siderophores. Bacteria producing siderophores have the ability to suppress phytopathogens and could be of significant agronomic importance[18]. Ten siderophore synthethase clusters containing 19 siderophore synthetases potentially involved in iron acquisition were found in the SS NF3 genome, as well as various iron-siderophore transporters (S1P Table).

After mining the diversity and abundance of the genes present in SS NF3 in comparison to the group of 10 other complete genomes and DSM 41855, we conclude that overall, this organism has a larger potential for host interaction, secondary metabolites production and protein degradation (S1P Table). The sum of all genes included in PFAM families related to the production of secondary metabolites is highest in the strain NF3 (87 families) compared to all the other genomes (<75 families), indicating a significant enrichment for these genes (by 16%). The ability of SS NF3 to express this remarkable genomic potential remains to be investigated by proteomics/metabolomics, in different growth conditions, in future studies.

Colonization of the plant *Arabidopsis thaliana* by SS NF3

Genome mining of the SS NF3 genome as well as its isolation source seems to indicate a potential for plant colonization, hypothesis which we tested under laboratory conditions by inoculating plants of *A. thaliana* (AT) with the bacteria and comparing with an untreated plant group. After a full growth cycle (21 days), plants initially treated with SS NF3 continued to carry the bacteria, despite it needing to compete with the soil microbiota and plant defense systems (Fig 8A). In addition, dilutions of plant material after a 90 seconds ethanol treatment, plated on YMG, showed (for some plants uniquely) colonies of SS NF3 (Fig 8B). Control experiments using only mycelia of NF3 treated with ethanol for 90 seconds did not result in
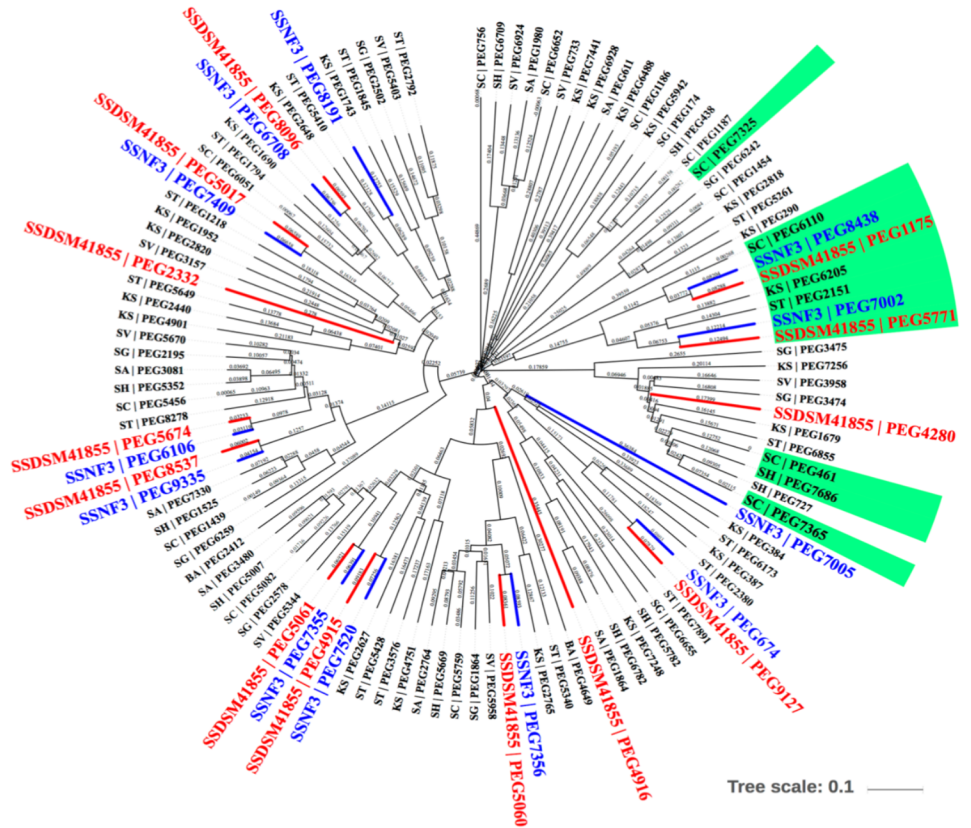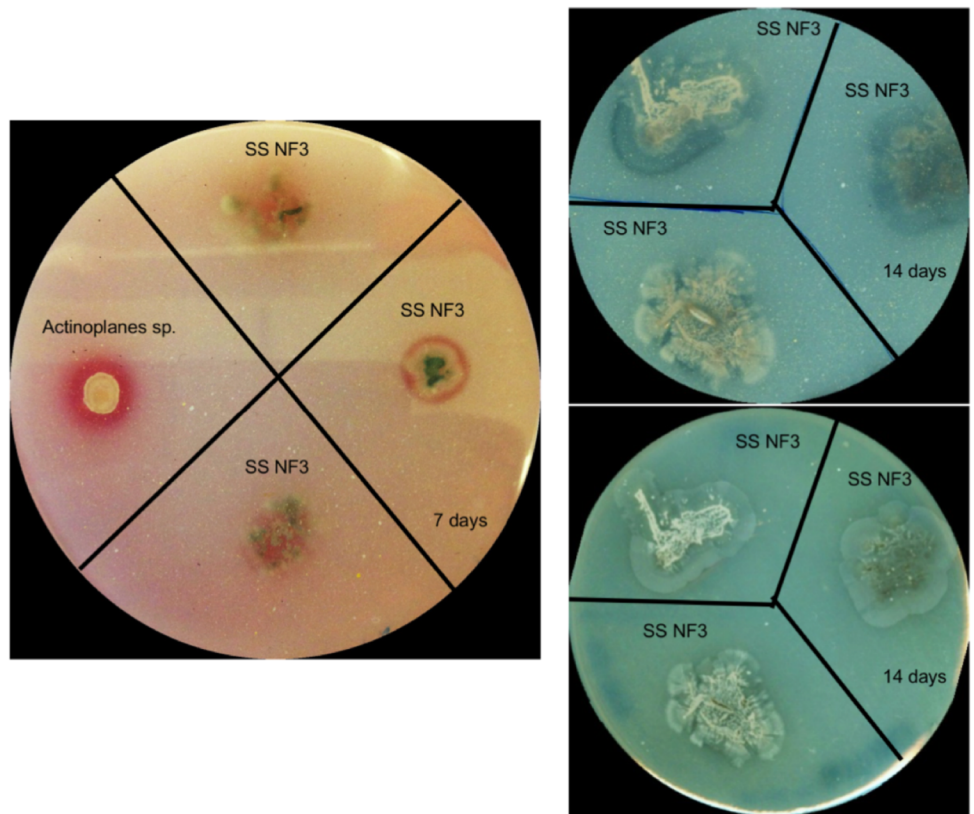
Fig 7. Chitinase genes and chitinolytic activity. A. NJ tree of all chitinase genes in the compared genomes. The MSA was produced using MUSCLE on all 116 chitinase sequences. Branches from proteins belonging to genomes S. scabrisporus NF3 and DSM 41855, as well as all secreted chitinases, are color labeled, respectively: blue, red and green. B. Expression of chitin degrading genes in SS NF3. Left side—chitin degradation at 7 days after inoculation. Right side upper level—chitin degradation at 14 days after inoculation. Ride side lower level—spore formation after growth on chitin for 14 days.

colonies (data not shown), indicating that the bacteria were residing at the interior of the plants and not on their surface. Plants treated with SS NF3 showed a different phenotype in both MS plates (at 10 days) but not after the period of growth in the soil. The observed changes in plant physiology included visibly smaller plants with shorter roots and a reduction in lateral roots formation for growth in MS plates (Fig 8C and 8D). The changes in plant physiology in MS plates could be due to a competition for nutrients between the bacteria and the plants, since colonies of SS NF3 are clearly visible at the inoculation site. This effect is not maintained when plants are transplanted in soil, where, in the presence of full soil microbiota and despite its presence at the end of the experiment, SS NF3 does not appear to have an effect on plant physiology. This a preliminary conclusion and further analysis of soil microbiota changes, production of plant hormones by plants and bacteria, and discrete changes in the plant need to be investigated. Another explanation for the data is that AT is not the most common host/ type of host for SS NF3 and larger physiological effects could be observed when changing the plant model.

In summary, the SS NF3 genome contains a diverse range of genes encoding for metabolites potentially important for this strain's adaptation to an endophytic lifestyle (Fig 9), and investigation of its singular activities could provide a platform for understanding mechanisms of plant—microbe cross-talk.

Conclusions

In this study, we investigated the genome of the endophytic SS NF3 strain, which presents remarkable biotechnological capabilities, with the interest to reveal its functional attributes in microbe-plant interactions. Proteome comparison with other taxonomically related as well as distant bacteria revealed an enrichment in metabolic pathways related to amino acid and protein synthesis and carbohydrate degradation as well as the presence of more than 50 clusters encoding potentially active compounds against plant pathogens (S1F Table). Metabolic mapping revealed prototrophy for 28 amino acids, nitrogen fixation capacity through the urea pathway and the potential for synthesis of hormones able to modulate plant growth (Fig 8). Experimental work confirmed the hypotheses that resulted from genome mining, such as the potential for plant colonization and chitin degradation.

Further research into the unique gene pool of this species with plant growth potential and a diversity of xenobiotic degradation and chitinolytic enzymes could provide valuable ways to improve agricultural practices by use of plant probiotics and decreasing the use and persistence of chemicals currently used for plant growth.

Materials and methods

Strains, media and culture conditions

The strain SS NF3 is deposited and available from the Institute of Biomedical Research Culture Collection, UNAM. The bacteria were grown in YMG (4 g/L yeast extract [Difco], 10 g/L malt extract [Difco], 4 g/L dextrose) in liquid cultures and incubated at 26°C for mycelia production and in YMG with 15 g/L agar [Difco] for growth on solid plates; cryogenic stocks were stored YMG with 15% glycerol at -80°C.
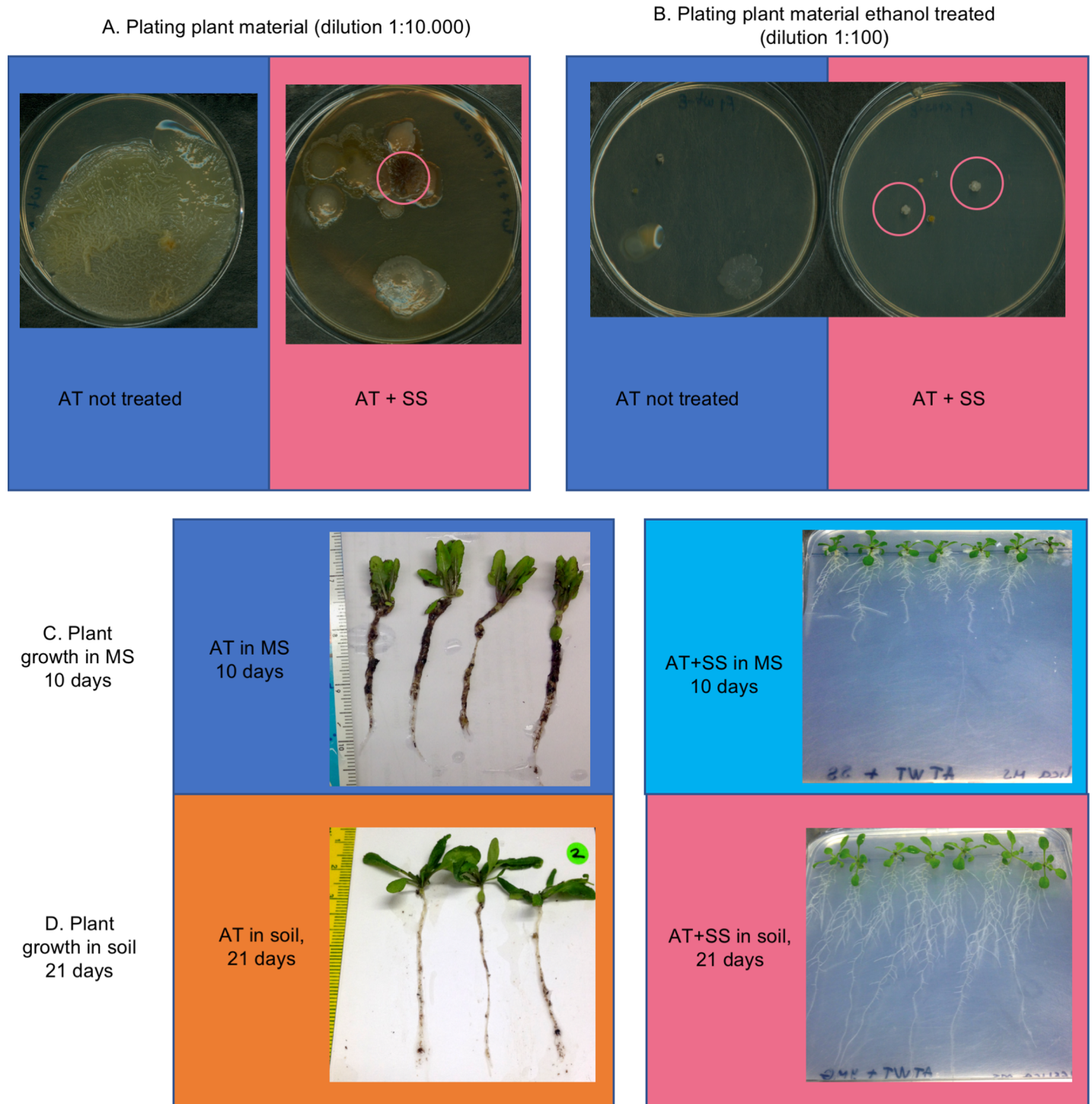
Fig 8. Colonization by SS NF3 of the plant *A. thaliana* (AT). Pink circles are used to indicate presence of SS NF3 colonies, which display a characteristic crumpled surface and production of red/orange secondary metabolites. A. Dilutions of plant material after 10 and 21 days of growth of plants inoculated with the bacteria (dilution 1:10.000). B. Dilutions of the same plant material after ethanol treatment for 90 seconds (dilution 1:100). C. Plant growth after 10 days in MS plates. Left, control plants with uninoculated plants. Right, plants whose seeds were inoculated with SS NF3. D. Plant growth after 21 days in regular soil. Left, control plants with uninoculated plants. Right, plants whose seeds were inoculated with SS NF3. (data shown is representative for 2 independent experiments).
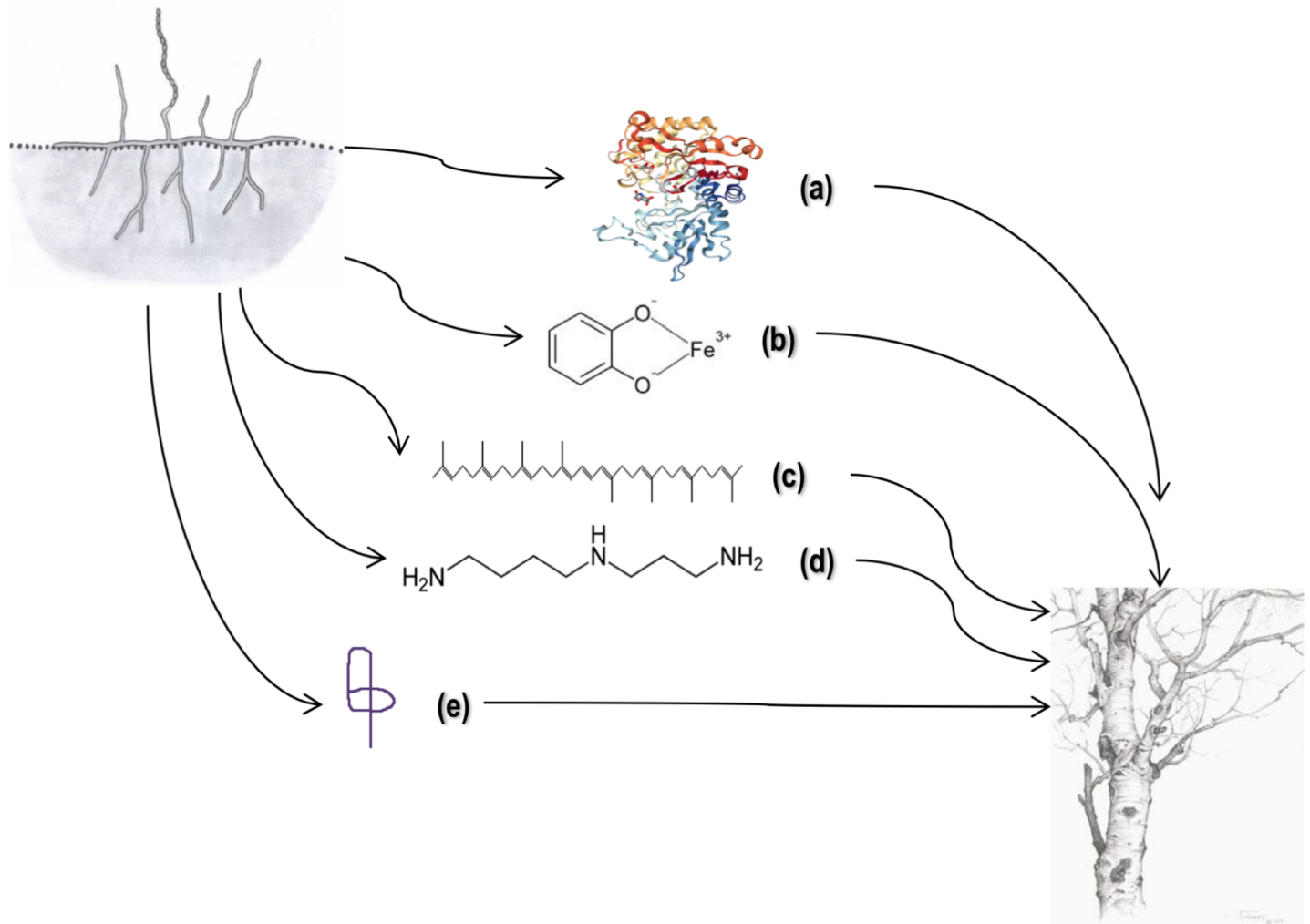
Fig 9. Genes involved in symbiotic interactions between SS NF3 and *A. adstringens*. (a) Chitinase family 18 (PDB 4Q22), (b) Siderophore, (c) Phytoene, (d) Spermidine, (e) Bacteriocins (here a lassopeptide).

The *Arabidopsis thaliana* plants used are Columbia-0 ecotype. Seedlings were grown on vertical plates with 0·2× Murashige Skoog salts (MP Biomedicals) and 1% sucrose, as previously described[94].

Genome and proteome comparisons

The closest taxonomic relative of *S. scabrisporus* NF3 with an annotated genome is *S. scabrisporus* DSM 41855, with which it shares approximately 98% of predicted coding sequences (CDSs). Due to the inadequate quality of the genome sequence for the latter, it could not be included in many of the genome analyses. A bi-directional BLAST analysis using the two *S. scabrisporus* genomes was performed using PATRIC's Proteome comparison tool. The strains used for proteome comparisons with PATRIC's Proteome comparison tool as well as orthologous group families comparison using FIGFams are: *Streptomyces coelicolor* A3 (2), *Streptomyces* sp. TLI_053, *Streptomyces avermitilis* MA-4680, *Streptomyces hygroscopicus* subsp. *limoneus* KCTC 1717, *Kitasatospora setae* KM-6054, *Streptomyces griseus* subsp. *griseus* NBRC 13350, *Streptomyces venezuelae* ATCC 15439, *Corynebacterium pseudotuberculosis* C231, *Bacillus*

anthracis Sterne and *Bifidobacterium breve* DSM 20213 = JCM 1192. Comparison of the gene families was performed in PATRIC using a Pearson pairwise average linkage correlation.

Other genomic comparisons and data analysis were performed using the online tools BLAST at the NCBI and Sanger web sites (www.ncbi.nlm.nih.gov/BLAST/ and http://www.sanger.ac.uk/), the Kyoto Encyclopedia of Genes and Genomes (www.genome.jp/kegg/), and ModelSEED (http://modelseed.org/) for metabolic pathway analysis.

For the enrichment analysis for genes potentially involved in host interaction, the PFAM family search was performed for family annotations including: polyketide, non-ribosomal, bacteriocin, lantipeptide, spermidine, indole, phytoene, chitin, chitinase, siderophore, choosing families with more than 2 members and more than 3 genomes. The results of the bacteriocin presence in the first column were obtained with Bagel3 (S1P Table).

PFAM domain annotation

InterProScan 5 was used to annotate the Pfam domains in the *S. scabrisporus* NF3 and comparative genomes. An enrichment test was performed using Fisher's exact test embedded in the Blast2Go module. Secreted proteins were predicted using PSORTb and SignalP[95] (S1A Table).

Secondary metabolites operons detection

Sequence analysis was performed by combining antiSMASH version 3.0.4[56], BAGEL3[57] and PRISM [58] genomic analysis platforms.

Phylogenetic analyses

In order to evaluate the taxonomical relatedness among the compared organisms, we used a multiple sequence alignment built with Kalign[46] of 4 concatenated protein sequences from each organism: 16S rRNA, gyrB, rpoB and recG. These proteins were chosen based on previous reports of phylogenomic analyses in prokaryotes according to their presence in all selected species, absence of additional fused domains, no subjection to Horizontal Gene Transfer (HGT), and completeness. The tree was built by calculating pairwise distances using PHYLIP Neighbor Joining with a Kimura distance matrix model in UGENE[96].

The relatedness of all chitinase genes from all compared genomes was evaluated using a multiple sequence alignment built with MUSCLE[97], and visualized by building an unrooted tree by calculating pairwise distances using PHYLIP Neighbor Joining[98] with a Kimura distance matrix model in UGENE. Branches from proteins belonging to genomes *S. scabrisporus* NF953 and DSM 41855, as well as secreted chitinases, were color labeled, respectively: blue, red and green.

Chitin degradation assay

Chitinolytic activity was determined as described[99]. Agar plates (2%) supplemented with colloidal chitin at 0.5% (SIGMA), partly hydrolyzed by stirring in 0.5 M HCl for 2h, were inoculated with strain SS NF3 and incubated at 28˚C. Hydrolysis zones were detected as cleared zones after 7 and 14 days of incubation.

Colonization of *Arabidopsis thaliana*

Seeds of Arabidopsis *thaliana* Col-0 used in this study were disinfected in 20% sodium hypochlorite and 0.01% of Tween-20 for 15 min. Then seeds were stratified at 4˚C for 4 days under dark conditions and sown on square petri dishes containing 0.2X Murashige and Skoog salts

(MS) (MP Biomedicals), 0.05% MES (SigmaAldrich), 1% sucrose (SigmaAldrich) and 1% agar (Becton, Dickinson and Company) at pH = 5.6. At the moment of inoculation in MS plates, one group of seeds were inoculated with 10^7 CFU of bacterial mycelium previously grown in YMG for three days and washed (3x) in phosphate saline buffer (PBS). Bacteria was used as mycelia since spores could have survived in MS and soil and confounded the experiment. The plates were vertically incubated in a growth chamber at 22˚C under long-day (LD; 16 h light/8 h dark) conditions with a light intensity of 110 μEm-2 s-1 for 4 days. Plants were transplanted at day 10 in complete unsterilized soil (3 plants/pot) and grown for an additional 11 days. After a complete 21 days of growth, plants were harvested, homogenized and a set of serial dilutions between 100 and 10000x was plated on YMG. One set of inoculated plants were pretreated with ethanol for 90 seconds before homogenization with the purpose of eliminating plant surface microbiota.

Supporting information

S1 Fig. Position of the genome size of *S. scabrisporus* NF3 within the genome length distribution of sequenced *Streptomyces* isolates. All *Streptomyces* genomes from the PATRIC database with more than 30x coverage and a size larger than 6Mb (the smallest reported *Streptomyces* has 6.8Mb [99]) were included in the analysis.
(TIF)

S1 Table. Supplementary analyses performed on the analyzed genomes. a. PATRIC annotation data and supplementary analyses summary: PSORTb, SignalP, PHI hit (virulence), PFAM hmm domain, PFAM superfamily, CAZymes families BESC (PFAM based), CAZymes annotation BESC (PFAM based), CAZymes families dbCAN (HMM based), CAZymes description.
b. RAST annotations for the *Streptomyces scabrisporus* NF3 genome.
c. PFAM annotations for the *Streptomyces scabrisporus* NF3 genome.
d. Unique gene families (FIGfam) for the *Streptomyces scabrisporus* NF3 genome.
e. Secreted unique proteins for *Streptomyces scabrisporus* NF3 as resulting from the comparative proteomes analysis.
f. KEGG mapping for secondary metabolites pathways in *Streptomyces scabrisporus* NF3.
g. Proteome comparisons for *Streptomyces scabrisporus* strains NF3 and DSM 41855.
h. All selected genomes, determination of OG.
i. CAZymes classification of genes of *Streptomyces scabrisporus* NF3.
j. CAZymes unique for the *Streptomyces scabrisporus* NF3 genome.
k. Detection of secondary metabolite clusters using automated engines: AntiSMASH results.
l. Detection of secondary metabolite clusters using automated engines: BAGEL3 results.
m. Positive BLAST results for genes from the *Streptomyces scabrisporus* NF3 genome in the PHI virulence database.
n. Antibiotic resistance genes recognized in the *Streptomyces scabrisporus* NF3 genome.
o. Genes annotated as chitinases in the *Streptomyces scabrisporus* NF3 genome.
p. Enrichment analysis for genes belonging to PFAM families with annotations including: polyketide, non-ribosomal, bacteriocin, lantipeptide, spermidine, indole, phytoene, chitin, chitinase, siderophore, choosing families with more than 2 members and more than 3 genomes. The results of the bacteriocin presence in the first column were obtained with Bagel3.
(XLSX)

S2 Table. Orthologous gene families mapping for all compared genomes.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Corina Diana Ceapă, Melissa Vázquez-Hernández, Verónica Jiménez Suárez, Romina Rodríguez-Sanoja, Elena R. Alvarez-Buylla, Sergio Sánchez.

Data curation: Corina Diana Ceapă, Melissa Vázquez-Hernández, Stefany Daniela Rodríguez-Luna, Angélica Patricia Cruz Vázquez, Verónica Jiménez Suárez.

Formal analysis: Corina Diana Ceapă, Melissa Vázquez-Hernández, Stefany Daniela Rodríguez-Luna, Angélica Patricia Cruz Vázquez, Verónica Jiménez Suárez.

Funding acquisition: Corina Diana Ceapă, Romina Rodríguez-Sanoja, Elena R. Alvarez-Buylla, Sergio Sánchez.

Investigation: Corina Diana Ceapă, Melissa Vázquez-Hernández, Stefany Daniela Rodríguez-Luna, Angélica Patricia Cruz Vázquez, Verónica Jiménez Suárez, Sergio Sánchez.

Methodology: Corina Diana Ceapă, Melissa Vázquez-Hernández, Angélica Patricia Cruz Vázquez, Verónica Jiménez Suárez.

Project administration: Corina Diana Ceapă, Melissa Vázquez-Hernández, Romina Rodríguez-Sanoja, Sergio Sánchez.

Resources: Corina Diana Ceapă, Melissa Vázquez-Hernández, Angélica Patricia Cruz Vázquez, Sergio Sánchez.

Software: Corina Diana Ceapă.

Supervision: Corina Diana Ceapă, Verónica Jiménez Suárez, Romina Rodríguez-Sanoja, Elena R. Alvarez-Buylla, Sergio Sánchez.

Validation: Corina Diana Ceapă, Melissa Vázquez-Hernández, Stefany Daniela Rodríguez-Luna, Romina Rodríguez-Sanoja.

Visualization: Corina Diana Ceapă, Melissa Vázquez-Hernández, Stefany Daniela Rodríguez-Luna.

Writing – original draft: Corina Diana Ceapă, Melissa Vázquez-Hernández, Stefany Daniela Rodríguez-Luna, Sergio Sánchez.

Writing – review & editing: Corina Diana Ceapă, Melissa Vázquez-Hernández, Stefany Daniela Rodríguez-Luna, Angélica Patricia Cruz Vázquez, Verónica Jiménez Suárez, Romina Rodríguez-Sanoja, Elena R. Alvarez-Buylla, Sergio Sánchez.

References

1. Sessitsch A, Hardoim P, Döring J, Weilharter A, Krause A, Woyke T, et al. (2012) Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. Mol Plant-Microbe Interactions MPMI 25: 28–36. Available: http://apsjournals.apsnet.org/doi/abs/10.1094/MPMI-08-11-0204.

2. Lòpez-Fernàndez S, Sonego P, Moretto M, Pancher M, Engelen K, Pertot I, et al. (2015) Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria. Front Microbiol 6: 419. https://doi.org/10.3389/fmicb.2015.00419 PMID: 26074885

3. Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG (2005) Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature 438: 90–93. https://doi.org/10.1038/nature04056 PMID: 16267554

4. Martens JH, Barg H, Warren MJ, Jahn D (2002) Microbial production of vitamin B12. Appl Microbiol Biotechnol 58: 275–285. Available: http://www.ncbi.nlm.nih.gov/pubmed/11935176. https://doi.org/10.1007/s00253-001-0902-7 PMID: 11935176

5. Garay-Arroyo A, De La Paz Sanchez M, Garcia-Ponce B, Azpeitia E, Alvarez-Buylla ER (2012) Hormone symphony during root growth and development. Dev Dyn 241: 1867–1885. https://doi.org/10.1002/dvdy.23878 PMID: 23027524

6. Agrawal N, Shahi SK (2015) An environmental cleanup strategy—microbial transformation of xenobiotic compounds. Int J Curr Microbiol Appl Sci 4: 429–461.

7. Khan Z, Roman D, Kintz T, Delas Alas M, Yap R, Doty S (2014) Degradation, phytoprotection and phytoremediation of phenanthrene by endophyte *Pseudomonas putida* PD1. Environ Sci Technol 48: 12221–12228. https://doi.org/10.1021/es503880t PMID: 25275224

8. Rylott EL (2014) Endophyte consortia for xenobiotic phytoremediation: the root to success? Plant Soil 385: 389–394. https://doi.org/10.1007/s11104-014-2296-1

9. Card S, Johnson L, Teasdale S, Caradus J (2016) Deciphering endophyte behaviour: The link between endophyte biology and efficacious biological control agents. FEMS Microbiol Ecol 92. https://doi.org/10.1093/femsec/fiw114 PMID: 27222223

10. Wang X, Qin J, Chen W, Zhou Y, Ren A, Gao Y (2016) Pathogen resistant advantage of endophyte-infected over endophyte-free *Leymus chinensis* is strengthened by pre-drought treatment. Eur J Plant Pathol 144: 477–486. https://doi.org/10.1007/s10658-015-0788-3

11. Ardanov P, Sessitsch A, Häggman H, Kozyrovska N, Pirttilä AM (2012) Methylobacterium-induced endophyte community changes correspond with protection of plants against pathogen attack. PLoS One 7. https://doi.org/10.1371/journal.pone.0046802 PMID: 23056459

12. Yang B, Ma HY, Wang XM, Jia Y, Hu J, Li X, et al. (2014) Improvement of nitrogen accumulation and metabolism in rice (*Oryza sativa* L.) by the endophyte *Phomopsis liquidambari*. Plant Physiol Biochem 82: 172–182. https://doi.org/10.1016/j.plaphy.2014.06.002 PMID: 24972305

13. Straub D, Yang H, Liu Y, Tsap T, Ludewig U (2013) Root ethylene signalling is involved in *Miscanthus sinensis* growth promotion by the bacterial endophyte *Herbaspirillum frisingense* GSF30 T. J Exp Bot 64: 4603–4615. https://doi.org/10.1093/jxb/ert276 PMID: 24043849

14. Tian B-Y, Cao Y, Zhang K-Q (2015) Metagenomic insights into communities, functions of endophytes, and their associates with infection by root-knot nematode, *Meloidogyne incognita*, in tomato roots. Sci Rep 5: 17087. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4658523/pdf/srep17087.pdf. https://doi.org/10.1038/srep17087 PMID: 26603211

15. Compant S, Clément C, Sessitsch A (2010) Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. Soil Biol Biochem 42: 669–678.

16. Maheshwari DK (2012) Bacteria in agrobiology: Plant probiotics. 1–371 p. https://doi.org/10.1007/978-3-642-27515-9

17. Armada E, Probanza A, Roldán A, Azcón R (2016) Native plant growth promoting bacteria *Bacillus thuringiensis* and mixed or individual mycorrhizal species improved drought tolerance and oxidative metabolism in *Lavandula dentata* plants. J Plant Physiol 192: 1–12. https://doi.org/10.1016/j.jplph.2015.11.007 PMID: 26796423

18. Beneduzi A, Ambrosini A, Passaglia LMP (2012) Plant growth-promoting rhizobacteria (PGPR): Their potential as antagonists and biocontrol agents. Genet Mol Biol 35: 1044–1051. PMID: 23411488

19. Tan GY, Deng K, Liu X, Tao H, Chang Y, Chen J, et al. (2017) Heterologous biosynthesis of spinosad: an omics-guided large polyketide synthase gene cluster reconstitution in *Streptomyces*. ACS Synth Biol 6: 995–1005. https://doi.org/10.1021/acssynbio.6b00330 PMID: 28264562

20. van Baarlen P, Kleerebezem M, Wells JM (2013) Omics approaches to study host-microbiota interactions. Curr Opin Microbiol 16: 270–277. https://doi.org/10.1016/j.mib.2013.07.001 PMID: 23891019

21. Chaudhary AK, Dhakal D, Sohng JK (2013) An insight into the "-omics" based engineering of streptomycetes for secondary metabolite overproduction. Biomed Res Int 2013. https://doi.org/10.1155/2013/968518 PMID: 24078931

22. Nah JH, Kim HJ, Lee HN, Lee MJ, Choi SS, Kim ES (2013) Identification and biotechnological application of novel regulatory genes involved in *Streptomyces* polyketide overproduction through reverse engineering strategy. Biomed Res Int 2013. https://doi.org/10.1155/2013/549737 PMID: 23555090

23. Seipke RF, Kaltenpoth M, Hutchings MI (2012) *Streptomyces* as symbionts: An emerging and widespread theme? FEMS Microbiol Rev 36: 862–876. https://doi.org/10.1111/j.1574-6976.2011.00313.x PMID: 22091965

24. Salla TD, Ramos da Silva T, Astarita LV, Santarém ER (2014) *Streptomyces* rhizobacteria modulate the secondary metabolism of Eucalyptus plants. Plant Physiol Biochem 85: 14–20. https://doi.org/10.1016/j.plaphy.2014.10.008 PMID: 25394796

25. Gopalakrishnan S, Srinivas V, Alekhya G, Prakash B (2015) Effect of plant growth-promoting *Streptomyces* sp. on growth promotion and grain yield in chickpea (Cicer arietinum L). 3 Biotech 5: 799–806. Available: http://www.ncbi.nlm.nih.gov/pubmed/28324533. Accessed 25 April 2017. https://doi.org/10.1007/s13205-015-0283-8 PMID: 28324533

26. Gopalakrishnan S, Srinivas V, Sree Vidya M, Rathore A (2013) Plant growth-promoting activities of Streptomyces spp. in sorghum and rice. Springerplus 2: 574.

27. Rey T, Dumas B (2017) Plenty is no plague: *Streptomyces* symbiosis with crops. Trends Plant Sci 22: 30–37. https://doi.org/10.1016/j.tplants.2016.10.008 PMID: 27916552

28. Vazquez-Hernandez M, Ceapa CD, Rodríguez-Luna SD, Rodríguez-Sanoja R, Sánchez S (2017) Draft genome sequence of *Streptomyces scabrisporus* NF3, an endophyte isolated from *Amphipterygium adstringens*. Genome Announc 5: e00267–17. Available: http://www.ncbi.nlm.nih.gov/pubmed/28450524. Accessed 31 July 2017. https://doi.org/10.1128/genomeA.00267-17 PMID: 28450524

29. Lin L, Hui Ge M, Yan T, Yan RX (2012) Thaxtomin A-deficient endophytic *Streptomyces* sp. enhances plant disease resistance to pathogenic *Streptomyces scabies*. Planta. 2012 Dec; 236(6):1849–61. https://doi.org/10.1007/s00425-012-1741-8 PMID: 22922880

30. Joseph B, Sankarganesh P, Justinraj S, Edwin BT (2011) Need exploration of endophytic *Streptomyces*—Review. J Pure Appl Microbiol 5: 153–160.

31. Jaiswal AK, Elad Y, Paudel I, Graber ER, Cytryn E, Frenkel O (2017) Linking the belowground microbial composition, diversity and activity to soilborne disease suppression and growth promotion of tomato amended with biochar. Sci Rep 7: 44382. https://doi.org/10.1038/srep44382 PMID: 28287177

32. Sadeghi A, Karimi E, Dahaji PA, Javid MG, Dalvand Y, Askari H (2012) Plant growth promoting activity of an auxin and siderophore producing isolate of *Streptomyces* under saline soil conditions. World J Microbiol Biotechnol 28: 1503–1509. https://doi.org/10.1007/s11274-011-0952-7 PMID: 22805932

33. Jog R, Nareshkumar G, Rajkumar S (2012) Plant growth promoting potential and soil enzyme production of the most abundant *Streptomyces* spp. from wheat rhizosphere. J Appl Microbiol 113: 1154–1164. https://doi.org/10.1111/j.1365-2672.2012.05417.x PMID: 22849825

34. Berg G (2009) Plant-microbe interactions promoting plant growth and health: Perspectives for controlled use of microorganisms in agriculture. Appl Microbiol Biotechnol 84: 11–18. https://doi.org/10.1007/s00253-009-2092-7 PMID: 19568745

35. Wang Q, Duan B, Yang R, Zhao Y, Zhang L (2015) Screening and identification of chitinolytic *Actinomycetes* and study on the inhibitory activity against turfgrass root rot disease fungi. J Biosci Med 3: 56–65.

36. Kurth F, Mailänder S, Bönn M, Feldhahn L, Herrmann S, Große I, et al. (2014) *Streptomyces*—induced resistance against oak powdery mildew involves host plant responses in defense, photosynthesis, and secondary metabolism pathways. Mol Plant-Microbe Interact 27: 891–900. https://doi.org/10.1094/MPMI-10-13-0296-R PMID: 24779643

37. Castillo-Juárez I, Rivero-Cruz F, Celis H, Romero I (2007) Anti-Helicobacter pylori activity of anacardic acids from *Amphipterygium adstringens*. J Ethnopharmacol 114: 72–77. https://doi.org/10.1016/j.jep.2007.07.022 PMID: 17768020

38. Olivera Ortega AG, Soto Hernández M, Martínez Vázquez M, Terrazas Salgado T, Solares Arenas F (1999) Phytochemical study of cuachalalate (*Amphiptherygium adstringens*, Schiede ex Schlecht). J Ethnopharmacol 68: 109–113. https://doi.org/10.1016/S0378-8741(99)00047-1 PMID: 10624869

39. Navarrete A, Oliva I, Sánchez-Mendoza ME, Arrieta J, Cruz-Antonio L, Castañeda-Hernández G. (2005) Gastroprotection and effect of the simultaneous administration of cuachalalate (*Amphipterygium adstringens*) on the pharmacokinetics and anti-inflammatory activity of diclofenac in rats. J Pharm Pharmacol 57: 1629–1636. Available: http://www.ncbi.nlm.nih.gov/pubmed/16354407. https://doi.org/10.1211/jpp.57.12.0013 PMID: 16354407

40. Oviedo-Chávez I, Ramírez-Apan T, Soto-Hernández M, Martínez-Vázquez M (2004) Principles of the bark of *Amphipterygium adstringens* (Julianaceae) with anti-inflammatory activity. Phytomedicine 11: 436–445. https://doi.org/10.1016/j.phymed.2003.05.003 PMID: 15330500

41. Ping X (2004) *Streptomyces scabrisporus* sp. nov. Int J Syst Evol Microbiol 54: 577–581. Available: http://ijs.sgmjournals.org/content/54/2/577.full. https://doi.org/10.1099/ijs.0.02692-0 PMID: 15023978

42. Charousová I, Medo J, Halenárová E, Javoreková S (n.d.) Antimicrobial and enzymatic activity of acti-nomycetes isolated from soils of coastal islands. J Adv Pharm Technol Res 8: 46–51. Available: http://www.ncbi.nlm.nih.gov/pubmed/28516055. Accessed 5 June 2017. https://doi.org/10.4103/japtr.JAPTR_161_16 PMID: 28516055

43. Zhang C, Ondeyka JGG, Zink DLL, Basilio A, Vicente F, Salazar O, et al. (2009) Discovery of okilacto-mycin and congeners from *Streptomyces scabrisporus* by antisense differential sensitivity assay target-ing ribosomal protein S4. J Antibiot 62: 55–61. https://doi.org/10.1038/ja.2008.8 PMID: 19132063

44. Kudo F, Kawamura K, Uchino A, Miyanaga A, Numakura M, Takayanagi R, et al. (2015) Genome min-ing of the hitachimycin biosynthetic gene cluster: involvement of a phenylalanine-2,3-aminomutase in biosynthesis. ChemBioChem 16: 909–914. https://doi.org/10.1002/cbic.201500040 PMID: 25786909

45. Shah AM, Wani A, Qazi PH, Rehman S, Mushtaq S, Ali SA, et al. (2016) Isolation and characterization of alborixin from *Streptomyces scabrisporus*: A potent cytotoxic agent against human colon (HCT-116) cancer cells. Chem Biol Interact 256: 198–208. https://doi.org/10.1016/j.cbi.2016.06.032 PMID: 27378626

46. Lassmann T, Sonnhammer EL. (2015) Kalign—an accurate and fast multiple sequence alignment algo-rithm. BMC Bioinformatics 6:298.

47. López MA, Vicente J, Kulasekaran S, Vellosillo T, Martínez M, Irigoyen ML, et al. (2011) Antagonistic role of 9-lipoxygenase-derived oxylipins and ethylene in the control of oxidative stress, lipid peroxidation and plant defence. Plant J 67: 447–458. https://doi.org/10.1111/j.1365-313X.2011.04608.x PMID: 21481031

48. Schouten A, van den Berg G, Edel-Hermann V, Steinberg C, Gautheron N, Alabouvette C, et al. (2004) Defense responses of *Fusarium oxysporum* to 2,4-diacetylphloroglucinol, a broad-spectrum antibiotic produced by *Pseudomonas fluorescens*. Mol Plant-Microbe Interact 17: 1201–1211. https://doi.org/10.1094/MPMI.2004.17.11.1201 PMID: 15559985

49. Raina S, Missiakas D, Georgopoulos C (1995) The rpoE gene encoding the sigma E (sigma 24) heat shock sigma factor of *Escherichia coli*. EMBO J 14: 1043–1055. PMID: 7889935

50. Jones GH, Paget MS, Chamberlin L, Buttner MJ (1997) Sigma-E is required for the production of the antibiotic actinomycin in *Streptomyces antibioticus*. Mol Microbiol 23: 169–178. PMID: 9004230

51. Lonetto MA, Brown KL, Rudd KE, Buttner MJ (1994) Analysis of the *Streptomyces coelicolor* sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regu-lation of extracytoplasmic functions. Proc Natl Acad Sci 91: 7573–7577. PMID: 8052622

52. Sevcikova B, Benada O, Kofronova O, Kormanec J (2001) Stress-response sigma factor sigma(H) is essential for morphological differentiation of *Streptomyces coelicolor* A3(2). Arch Microbiol 177: 98–106. https://doi.org/10.1007/s00203-001-0367-1 PMID: 11797050

53. Otani H, Higo A, Nanamiya H, Horinouchi S, Ohnishi Y (2013) An alternative sigma factor governs the principal sigma factor in *Streptomyces griseus*. Mol Microbiol 87: 1223–1236. https://doi.org/10.1111/mmi.12160 PMID: 23347076

54. Kaul S, Sharma T, K Dhar M (2016) "Omics":Tools for better understanding the plant-endophyte interac-tions. Front Plant Sci 7: 955. https://doi.org/10.3389/fpls.2016.00955 PMID: 27446181

55. Blumenstein K, Macaya-Sanz D, Martín JA, Albrectsen BR, Witzell J (2015) Phenotype microarrays as a complementary tool to next generation sequencing for characterization of tree endophytes. Front Microbiol 6: 1033. https://doi.org/10.3389/fmicb.2015.01033 PMID: 26441951

56. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, et al. (2015) AntiSMASH 3.0-A comprehen-sive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43: W237–W243. https://doi.org/10.1093/nar/gkv437 PMID: 25948579

57. van Heel AJ, de Jong A, Montalbán-López M, Kok J, Kuipers OP (2013) BAGEL3: Automated identifica-tion of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. Nucleic Acids Res 41. https://doi.org/10.1093/nar/gkt391 PMID: 23677608

58. Skinnider M.A., Dejong C.A., Rees P.N., Johnston C.W., Li H., Webster A.L.H., Wyatt M.A., and Magar-vey N.A. (2015). Genomes to natural products prediction informatics for secondary metabolomes (PRISM). Nucleic Acids Research, 43, 9645–9662. https://doi.org/10.1093/nar/gkv1012 PMID: 26442528

59. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. (2010) High-throughput genera-tion, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 28: 977–982. Avail-able: http://www.nature.com/nbt/journal/v28/n9/abs/nbt.1672.html. https://doi.org/10.1038/nbt.1672 PMID: 20802497

60. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels 6: 41. Available: http://biotechnologyforbiofuels.biomedcentral.com/articles/10.1186/1754-6834-6-41. https://doi.org/10.1186/1754-6834-6-41 PMID: 23514094

61. Min B, Park J-H, Park H, Shin H-D, Choi I-G (2017) Genome analysis of a zygomycete fungus *Choanephora cucurbitarum* elucidates necrotrophic features including bacterial genes related to plant colonization. Sci Rep 7: 40432. https://doi.org/10.1038/srep40432 PMID: 28091548

62. Nishizawa T, Miura T, Harada C, Guo Y, Narisawa K, Ohta H, et al. (2016) Complete genome sequence of *Streptomyces parvulus* 2297, integrating site-specifically with Actinophage R4. Genome Announc 4: e00875–16. https://doi.org/10.1128/genomeA.00875-16 PMID: 27563047

63. Artzi L, Bayer EA, Moraïs S (2016) Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. Nat Rev Microbiol 15: 83–95. https://doi.org/10.1038/nrmicro.2016.164 PMID: 27941816

64. Levasseur A, Piumi F ßois F, Coutinho PM, Rancurel C, Asther MM, Delattre M, et al. (2008) FOLy: An integrated database for the classification and functional annotation of fungal oxidoreductases potentially involved in the degradation of lignin and related aromatic compounds. Fungal Genet Biol 45: 638–645. https://doi.org/10.1016/j.fgb.2008.01.004 PMID: 18308593

65. Martín JF, Rodríguez-García A, Liras P (2017) The master regulator PhoP coordinates phosphate and nitrogen metabolism, respiration, cell differentiation and antibiotic biosynthesis: comparison in *Streptomyces coelicolor* and *Streptomyces avermitilis*. J Antibiot (Tokyo) 15: 1–8.

66. Gregor AK, Klubek B, Varsa EC (2003) Identification and use of actinomycetes for enhanced nodulation of soybean co-inoculated with *Bradyrhizobium japonicum*. Can J Microbiol 49: 483–491. https://doi.org/10.1139/w03-061 PMID: 14608383

67. Tokala RK, Strap JL, Jung CM, Crawford DL, Salove MH, Deobald LA, et al. (2002) Novel plant-microbe rhizosphere interaction involving *Streptomyces lydicus* WYEC108 and the pea plant (*Pisum sativum*). Appl Environ Microbiol 68: 2161–2171. https://doi.org/10.1128/AEM.68.5.2161-2171.2002 PMID: 11976085

68. Tarkka MT, Lehr NA, Hampp R, Schrey SD (2008) Plant behavior upon contact with *Streptomycetes*. Plant Signal Behav 3: 917–919. PMID: 19513192

69. Economou A, Hamilton WD, Johnston AW, Downie JA (1990) The Rhizobium nodulation gene nodO encodes a Ca2(+)-binding protein that is exported without N-terminal cleavage and is homologous to haemolysin and related proteins. EMBO J 9: 349–354. PMID: 2303029

70. Urban M, Pant R, Raghunath A, Irvine AG, Pedro H, Hammond-Kosack K (2015). The Pathogen-Host Interactions database: additons and future developments. Nucleic Acids Res 43 (Database Issue): D645–655. https://doi.org/10.1093/nar/gku1165 PMID: 25414340

71. Glick BR, Glick BR (2012) Plant growth-promoting bacteria: mechanisms and applications. Scientifica (Cairo) 2012: 1–15.

72. Cabanás CGL, Schilirò E, Valverde-Corredor A, Mercado-Blanco J (2014) The biocontrol endophyt0ic bacterium *Pseudomonas fluorescens* PICF7 induces systemic defense responses in aerial tissues upon colonization of olive roots. Front Microbiol 5: 427. https://doi.org/10.3389/fmicb.2014.00427 PMID: 25250017

73. Brechenmacher L, Lei Z, Libault M, Findley S, Sugawara M, Sadowsky MJ, et al. (2010) Soybean metabolites regulated in root hairs in response to the symbiotic bacterium *Bradyrhizobium japonicum*. Plant Physiol 153: 1808–1822. https://doi.org/10.1104/pp.110.157800 PMID: 20534735

74. Newton AC, Fitt BDLL, Atkins SD, Walters DR, Daniell TJ (2010) Pathogenesis, parasitism and mutualism in the trophic space of microbe–plant interactions. Trends Microbiol 18: 365–373. https://doi.org/10.1016/j.tim.2010.06.002 PMID: 20598545

75. Yakhin OI, Lubyanov AA, Yakhin IA, Brown P (2016) Biostimulants in plant science: a global perspective. Front Plant Sci 7: 2049. https://doi.org/10.3389/fpls.2016.02049 PMID: 28184225

76. Yu X, Li Y, Cui Y, Liu R, Li Y, Chen Q, et al. (2017) An indoleacetic acid-producing Ochrobactrum sp. MGJ11 counteracts cadmium effect on soybean by promoting plant growth. J Appl Microbiol 122: 987–996. https://doi.org/10.1111/jam.13379 PMID: 27995689

77. Palacios OA, Gomez-Anduro G, Bashan Y, De-Bashan LE (2016) Tryptophan, thiamine and indole-3-acetic acid exchange between *Chlorella sorokiniana* and the plant growth-promoting bacterium *Azospirillum brasilense*. FEMS Microbiol Ecol 92: 1–11.

78. Lee KE, Radhakrishnan R, Kang SM, You YH, Joo GJ, Lee IJ, et al. (2015) *Enterococcus faecium* LKE12 cell-free extract accelerates host plant growth via gibberellin and indole-3-acetic acid secretion. J Microbiol Biotechnol 25: 1467–1475. https://doi.org/10.4014/jmb.1502.02011 PMID: 25907061

79. Han Y, Zheng QS, Wei YP, Chen J, Liu R, Wan HJ. (2015) In silico identification and analysis of phytoene synthase genes in plants. Genet Mol Res 14: 9412–9422. https://doi.org/10.4238/2015.August.14.5 PMID: 26345875

80. Schmerk CL, Bernards MA, Valvano MA (2011) Hopanoid production is required for low-pH tolerance, antimicrobial resistance, and motility in *Burkholderia cenocepacia*. J Bacteriol 193: 6712–6723. Available: http://www.ncbi.nlm.nih.gov/pubmed/21965564. Accessed 26 April 2017. https://doi.org/10.1128/JB.05979-11 PMID: 21965564

81. Silipo A, Vitiello G, Gully D, Sturiale L, Chaintreuil C, Fardoux J, et al. (2014) Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. Nat Commun 5: 5106. https://doi.org/10.1038/ncomms6106 PMID: 25355435

82. Kulkarni G, Busset N, Molinaro A, Gargani D, Chaintreuil C, Silipo A, et al. (2015) Specific hopanoid classes differentially affect free-living and symbiotic states of *Bradyrhizobium diazoefficiens*. MBio 6. https://doi.org/10.1128/mBio.01251-15 PMID: 26489859

83. Rovenich H, Boshoven JC, Thomma BPHJ (2014) Filamentous pathogen effector functions: Of pathogens, hosts and microbiomes. Curr Opin Plant Biol 20: 96–103. https://doi.org/10.1016/j.pbi.2014.05.001 PMID: 24879450

84. Buist G, Steen A, Kok J, Kuipers OP (2008) LysM, a widely distributed protein motif for binding to (peptido)glycans. Mol Microbiol 68: 838–847. https://doi.org/10.1111/j.1365-2958.2008.06211.x PMID: 18430080

85. Rosado A, Schapire AL, Bressan RA, Harfouche AL, Hasegawa PM, Valpuesta V, et al. (2006) The Arabidopsis tetratricopeptide repeat-containing protein Ttl1 is required for osmotic stress responses and abscisic acid sensitivity. Plant Physiol 142: 1113–1126. https://doi.org/10.1104/pp.106.085191 PMID: 16998088

86. Wiederanders B (2003) Structure-function relationships in class CA1 cysteine peptidase propeptides. Acta Biochim Pol 50: 691–713. https://doi.org/035003691 PMID: 14515150

87. Li Y, Huang F, Lu Y, Shi Y, Zhang M, Fan J, et al. (2013) Mechanism of plant-microbe interaction and its utilization in disease-resistance breeding for modern agriculture. Physiol Mol Plant Pathol 83: 51–58.

88. Hjort K, Bergström M, Adesina MF, Jansson JK, Smalla K, Sjöling S. (2010) Chitinase genes revealed and compared in bacterial isolates, DNA extracts and a metagenomic library from a phytopathogen-suppressive soil. FEMS Microbiol Ecol 71: 197–207. https://doi.org/10.1111/j.1574-6941.2009.00801.x PMID: 19922433

89. Igarashi K, Uchihashi T, Uchiyama T, Sugimoto H, Wada M, Suzuki K, et al. (2014) Two-way traffic of glycoside hydrolase family 18 processive chitinases on crystalline chitin. Nat Commun 5: 3975. https://doi.org/10.1038/ncomms4975 PMID: 24894873

90. Purushotham P, Arun PVPS, Prakash JSS, Podile AR (2012) Chitin binding proteins act synergistically with chitinases in *Serratia proteamaculans* 568. PLoS One 7: 1–10. https://doi.org/10.1371/journal.pone.0036714 PMID: 22590591

91. Song YS, Oh S, Han YS, Seo DJ, Park RD, Jung WJ. (2013) Detection of chitinase ChiA produced by *Serratia marcescens* PRC-5, using anti-PrGV-chitinase. Carbohydr Polym 92: 2276–2281. https://doi.org/10.1016/j.carbpol.2012.12.018 PMID: 23399288

92. Rathore AS, Gupta RD, Rathore AS, Gupta RD (2015) Chitinases from bacteria to human: properties, applications, and future perspectives. Enzyme Res 2015: 1–8. https://doi.org/10.1155/2015/791907 PMID: 26664744

93. Kielak AM, Cretoiu MS, Semenov A V., Sørensen SJ, Van Elsas JD (2013) Bacterial chitinolytic communities respond to chitin and pH alteration in soil. Appl Environ Microbiol 79: 263–272. https://doi.org/10.1128/AEM.02546-12 PMID: 23104407

94. Tapia-Lopez R, Garcia-Ponce B, Dubrovsky JG (2008). An AGAMOUS-related MADS-box gene, XAL1 (AGL12), regulates root meristem cell proliferation and flowering transition in *Arabidopsis*. Plant Physiology 146: 1182–92. https://doi.org/10.1104/pp.107.108647 PMID: 18203871

95. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8: 785–786. https://doi.org/10.1038/nmeth.1701 PMID: 21959131

96. Okonechnikov K, Golosova O, Fursov M, UGENE team (2012) Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28(8):1166–7. https://doi.org/10.1093/bioinformatics/bts091 PMID: 22368248

97. Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5: 113. https://doi.org/10.1186/1471-2105-5-113 PMID: 15318951

98. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. Jul; 4(4):406–25. https://doi.org/10.1093/oxfordjournals.molbev.a040454 PMID: 3447015

99. Zaburannyi N., Rabyk M., Ostash B., Fedorenko V., & Luzhetskyy A. (2014). Insights into naturally minimised *Streptomyces albus* J1074 genome. BMC Genomics, 15: 97. https://doi.org/10.1186/1471-2164-15-97 PMID: 24495463