scientific reports

Check for updates

OPEN

Comparative analysis towards the identification of genome wide characteristics of a beneficial fungal endophyte

Olga Tsiouri, Antonia Argyri, Vasiliki Skiada, Marianna Avramidou, Elena Dadami, Sotirios Vasileiadis✉ & Kalliope K. Papadopoulou✉

Fusarium solani **strain K (FsK) is an endophytic fungus with a wide host range that protects its host plants against pathogens and environmental stresses. In this study, we performed de novo genome sequencing and annotation, while a phylogenomic analysis confirmed the placement of FsK within the *Fusarium solani* species complex (FSSC). A comprehensive comparative genomics analysis was conducted with the closely-related pathogenic *Fusarium vanettenii* 77–13-4 species as well as two beneficial fungal species, the basidiomycete *Serendipita indica* and the arbuscular mycorrhizal fungus *Rhizophagus irregularis*, both model symbiotic organisms in their respective fungal division. All fungal strains can colonize tomato as a common host. To identify mechanisms of early-stage FsK-plant interaction and fungal adaptation, comparative analysis of secreted effectors, carbohydrate-active enzymes and secondary metabolite clusters, was performed. FsK specific genes implicated in DNA repair and iron acquisition via ferrirhodin synthesis were identified, highlighting adaptation to stress conditions and possible mutualistic functions with plant hosts. These findings provide insights into the genomic determinants of lifestyles of root colonizing fungi and establish a foundation for future functional studies on fungal-plant interactions in agricultural contexts.**

Keywords Fusarium, Beneficial, Effectome, Cazymes, Comparative genomics, Ferrirhodin

Plant-associated fungi exhibit phenotypic plasticity as their interaction with the plant host can span the mutualistic-pathogenic continuum, from mutualistic (both organisms benefit) to parasitic/pathogenic (only one gets harmed)[1,2]. Mutualistic fungi can be true symbionts, such as the arbuscular mycorrhizal fungi (AMF – Glomeromycota)[3] and typical endophytes (Asco- and Basidiomycota) that reside entirely and complete their life cycle within plant tissues, intra- and/or intercellularly[4,5]. Endophytic fungi that asymptomatically invade the plant can enhance plant fitness, as colonized plants show improved growth, mitigation of environmental stresses, induced plant resistance against pathogens and pests[6]. Because of these phenotypes, beneficial endophytes are promising for supporting agricultural production[7] for a wide range of crop plants. Yet, their effective application is still limited by an incomplete understanding of the mechanisms that govern their interactions with plant hosts.

In particular, the genetic and molecular determinants involved in the establishment and maintenance of the symbiosis or in the benefit conferred to the plant are not well-studied. While recent sequencing projects have provided insights into traits associated with various lifestyles in fungi, these are mainly focused either on pathogens or mycorrhizal species[8–10]. These lifestyles can often be inferred by analysing specific proteins groups, such as secreted proteins that interact with plant proteins (effectome)[11] and/or carbohydrate-active enzymes (CAZymes), usually present in the apoplast (apoplastome) and involved in the degradation of plant biopolymers[12], that may play a key role in the establishment of the plant-fungus interaction. In addition, secondary metabolites produced by the plant-associated fungi can have a wide range of functions, acting as phytohormones and antibiotics, to compounds toxic to the plant[13].

The filamentous, endophytic fungus *Fusarium solani strain K*, studied here, was isolated from root tissues of tomato plants and acts protectively against root and foliar pathogens[14] and pests, such as spider mites and zoophytophagous predators[15,16], while its presence in the plant alleviates the response of the plant to drought[17]. The colonization pattern of the fungus displays unique traits, since it can reach and proliferate in the vascular bundle of the root, normally exhibited by severe pathogens of the plants. FsK can also colonize the model

Laboratory of Plant and Environmental Biotechnology, Department of Biochemistry and Biotechnology, University of Thessaly, 41500 Larissa, Biopolis, Greece. ✉email: sovasileiadis@uth.gr; kalpapad@uth.gr

legume *Lotus japonicus*[18,19], a process mediated by Common Symbiotic Signaling Pathway (CSSP), also needed for the establishment of other symbiotic microorganisms, such as rhizobia and AMF[20]. FsK promotes root and shoot length and improves nutrient status of *L. japonicus* under iron deficiency conditions[21]. FsK has been classified[14] as a member of the genus *Fusarium* and specifically of the phylogenetically distinct species, known as 'Fusarium solani species complex (FSSC)[22,23]. *Fusarium* has been traditionally considered of monophyletic nature[24,25], although recent studies based on morphological and phylogenetic data suggested otherwise[26–28]. This has sparked a debate on whether or not FSSC belongs to *Fusarium*[24]. While there have been some general comparative genomics attempts within the FSSC[23,29], a more detailed analysis among members of the group is missing.

In this study, the whole genome de novo sequencing and annotation of FsK is presented and the taxonomic placement of *Fusarium solani* strain K (FsK) was addressed, using state-of-the-art phylogenomics pipelines. Comparative genomics analysis was performed using the predicted proteomes of FsK and of the closely related pathogenic *Fusarium vanettenii* 77–13-4 (previously known as *Nectria haematococca*)[30,31]. In addition, we included in our analysis, the comparison with the symbiotic Glomeromycete *Rhizophagus irregularis* and the beneficial Basidiomycete *Serendipita indica.* The latter two are both considered model organisms in the fungal division they belong to[32,33]. Their capacity to confer protection against abiotic and biotic stresses in *Solanum lycopersicum*[34–37] and growth promotion in *Lotus japonicus*[38,39] is well established. All four fungal species in this study were further selected for their well-characterised ability to colonize tomato and *Lotus japonicus* roots and, thus, share at least two common host-plants. We focused our analysis on identifying (a) FsK species-specific proteins with unique functions, (b) FsK-specific proteins non-homologous with those of the pathogen, (c) similarities and differences on the types of predicted effectors and apoplastic CAZymes and in the secondary biosynthetic clusters present in all four fungal species under study. Our aim was to provide a greater insight on how different fungi establish interactions their host plants and shed light to possible adaptation mechanisms that explain the phenotypes observed during plant colonization by the endophytes.

Results and discussion

FsK genome and functional annotation characteristics

Hybrid Nanopore/Illumina sequencing, yielded a genome assembly of 58 Mbp for the FsK strain, organized into 44 scaffolds, with N50 and N90 values 2 and 0.7 Mbp respectively (Table 1). The 49.2% GC content of the FsK genome identified here was marginally lower than that of fungi from the FSSC complex (49.4–51.4% GC content)[29,40]. Functional genomics analysis resulted in 98% of the proteins being functionally annotated by eggNOG-mapper v2.0, with 65% containing known protein families database (Pfam) domains[41]. Most of the single copy orthologues (BUSCOs- Benchmarking Universal Single-Copy Orthologs) retrieved from the dikarya_odb9 protein database were found to be complete in amino acid length. This is indicative of a good quality genome and transcriptome assembly with no major duplication or mis-assembly events adversely affecting gene prediction during FsK annotation (Table 1). The gene annotation of the genome of this beneficial

Category	Feature number or percentage
Genome assembly statistics	
Number of scaffolds	44
Length (bp)	57,679,002
N50	2,416,816
L50	9
N90	709,252
L90	26
%GC	49.2
Genome annotation statistics	
Genes	16,141
mRNAs	15,830
tRNAs	311
CDS	15,830
Proteins	15,830
Functional annotation statistics	
Eggnog	15,154
PFAM	10,805
KEGG	4,788
%Complete BUSCOs	96
%Single copy BUSCOs	95
MEROPS	549
CAZy	649

Table 1. Genome assembly, annotation and functional annotation statistics of the beneficial fungus *Fusarium solani strain K* (FsK).

Fusarium revealed a broad metabolic capacity, with numerous genes related to carbohydrate metabolism (1021, Fig. 1a) and associated enzymes (1295, Fig. 1b), which warrant further investigation to assess their potential contribution to the endophytic lifestyle of the fungus. The abundance of amino acid metabolism genes (867, Fig. 1a) may indicate for a possible role in nutrient exchange that could support mutualistic interactions. Gene repertoires associated with transcription (588, Fig. 1a), ribosome biogenesis (200, Fig. 1b), and RNA processing (169 mRNA biogenesis; 134 spliceosome) highlight the capacity to fine-tune gene expression that may enable rapid shifts between saprotrophy[42], host colonization[43], and stress adaptation[44]. Among these general traits of the fungal genome, we observed a large number of genes involved in intracellular trafficking and secretion (586, Fig. 1a; 370, Fig. 1b), indicating a robust secretory system. In beneficial strains, this machinery has been associated with the release of plant growth–promoting enzymes, antimicrobial compounds, and secondary metabolites that influence plant signaling[45,46]. Finally, genes involved in replication, repair and recombination (344, Fig. 1a; 150, Fig. 1b) reportedly support genome stability, while contributing to adaptability across hosts and environments[47].

Fusarium solani K phylogenomics

We constructed a phylogenomic tree for FSSC based on all predicted proteins per organism included in this analysis. An all vs all protein comparison strategy was employed to assign proteins that share a common ancestor into orthogroups which were used to construct the phylogenetic trees. We did not restrict to single copy coding genes, given their low representation (1.8% of the total orthogroup numbers) of the total genomes and reduced capacity to achieve high hylogenomic resolution. Finally, all trees were reconciled into a consensus species tree as shown in Fig. 2. *Fusarium solani* strain K clusters firmly within the *Fusarium* genus, specifically within the clade corresponding to the *Fusarium solani* species complex (FSSC), confirming the previous assignment of the strain to this species complex[14].

The closest isolate to FsK within the FSSC species in our analysis is annotated as a *Fusarium falciforme* Fu3.1 strain, an isolate from sea turtle eggs (NCBI biosample id: SAMN26804627) and a *Fusarium falciforme* strain A01-2 isolated from a diseased citrus trees (NCBI biosample id: SAMN30354179) (Fig. 2). The only other *Fusarium* strain in our analysis, able to colonize tomato roots is the pathogenic *Fusarium vanettenii* 77–13-4. The latter was chosen for further comparison with FsK, as it represents the best studied strain, formerly being classified as *Nectria haematococca* mating population VI (MPVI). The genome of strain *F. vanettenii* 77–13-4 was the first sequenced out of the FSSC isolates[22]. FsK has a larger genome size and there is a significant difference in %GC content across scaffold, coding sequence (CDS), exon, and intron regions (Supplementary Table S1, Supplementary Fig. S1). This is in contrast with previous findings of a meta-analysis of 1090 publicly known fungal genomes, which concluded that large genomes are usually characteristics of pathogenic fungi[48]. Other non-pathogenic fungi also tend to have higher %GC content than pathogens. For example, the wood-degrading basidiomycete *Phanerochaete chrysosporium* has a %GC content of 57%[49], while the ascomycete biocontrol agent *Trichoderma reesei* ranges from 51–54% depending on the strain (NCBI data, accessed February 2026).. We have also compared the genomes of FsK and *F. vanettenii* by comparison of the scaffold sequences between to detect putative FsK-specific accessory structures (Supplementary Fig. S2). This revealed extensive genome-wide correspondence and long syntenic blocks retained since divergence (Supplementary Fig. S3a). Despite the depicted structural rearrangements such as inversions, translocations, and lineage-specific duplications, the two species share high nucleotide-level conservation, suggesting genome evolution has primarily involved

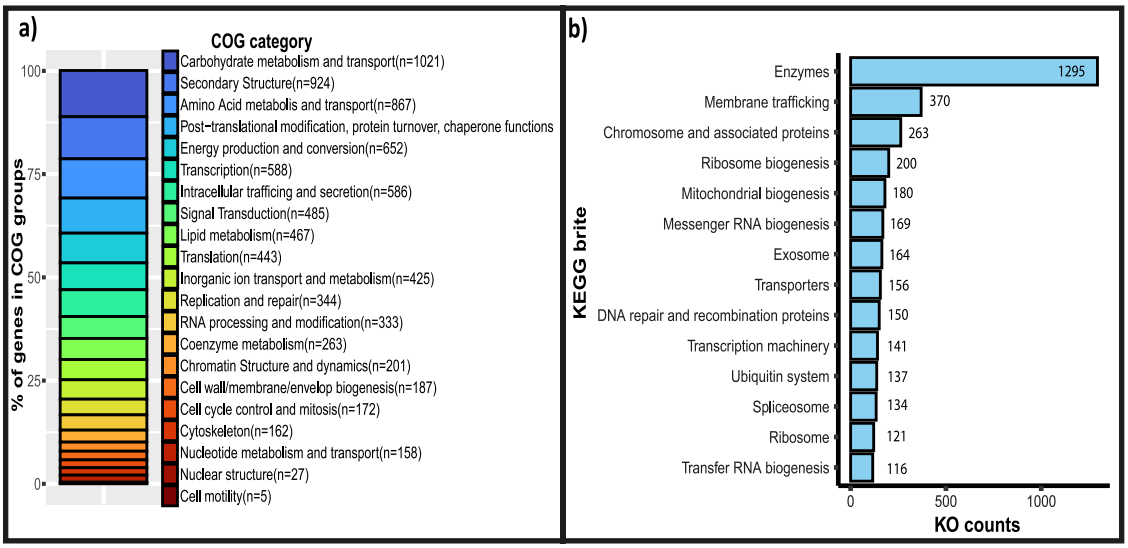

Fig. 1. FsK functional annotation statistics. (**a**) Percentage of genes assigned to specific cluster of orthologous groups (COG) per COG category on FsK (*Fusarium solani strain K*). These data were retrieved using eggNOG-mapper v2.0. (**b**) Number of FsK KEGG orthologs (KO) per KEGG brite category. These data were retrieved using BlastKOALA v3.0 and were aggregated using KEGG Mapper Reconstruct (July 2021 update).

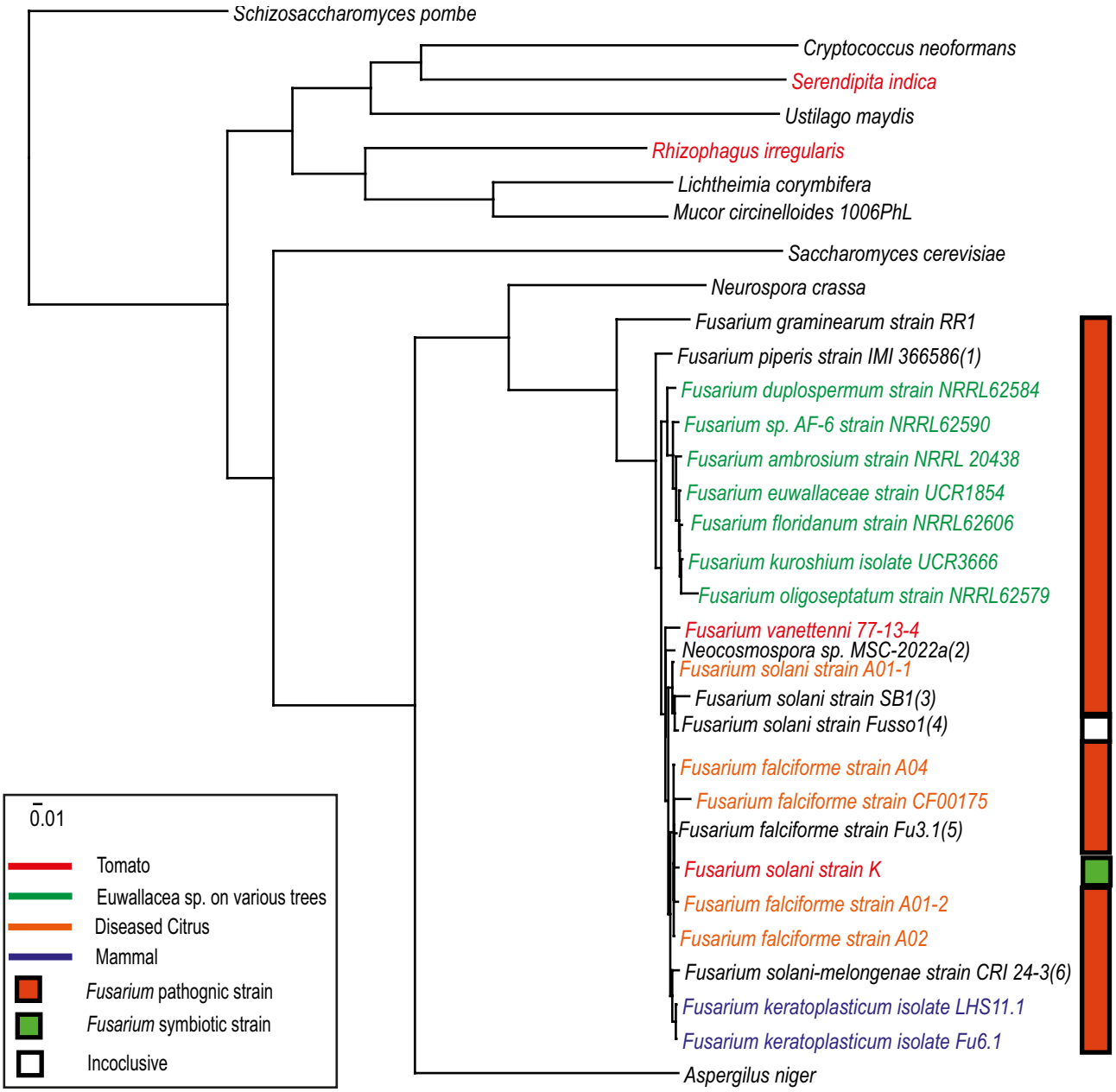

Fig. 2. Fusarium solani species complex phylogenetic tree. The tree was performed with OrthoFinder based on the predicted organism proteomes. Fungal strain hosts are indicated with different colors according to the key. Strains of other hosts are as follows: (1) endophyte from banana and Manilkara bidentata plants; (2) pathogen of sugarbeet; (3) *Arabidopsis thaliana* root isolate; (4) sweet potato pathogen; (5) pathogen of the sea turtle *Eretmochelys imbricate*; (6) pathogen of *Lumnitzera racemosa.*

rearrangement of conserved blocks rather than widespread sequence divergence. From this analysis, a highly probable, 80 Kb long accessory scaffold candidatescaffold_43, was identified with an alignment coverage of only 0.5% against the pathogen's chromosome/scaffold areas (Supplementary Fig. S2b).Based on the genome annotation, the predicted genes of this scaffold encode only tRNAs (Supplementary Table S2). It may be envisaged that these FsK specific tRNAs play a role in the host-fungus interaction that warrants further investigation.

Predicted cytoplasmic and apoplastic effectomes

To identify potential similarities and differences in genomic traits that may be related to the establishment of their interaction with a host plant, we comparatively investigated the predicted secreted proteome of FsK with the one of the pathogenic *F. vanettenii* and of the two beneficial symbionts, *Rhizophagus irregularis* and *Serendipita indica*. This was done via the identification of effector proteins, i.e. small, secreted proteins that can be transferred extracellularly from the fungus to the apoplast/cytoplasm of a plant cell with a putative role in communication with the host plant and the environment. The predicted proteomes of the three fungal strains were retrieved

from Ensembl Fungi (15,705 proteins were predicted for the *F. vanettenii*, 26,143 for *Rhizophagus irregularis* and 11,766 for *Serendipita indica*) and compared with the predicted proteins derived from the genome annotation of FsK (15,830 proteins). Figure 3a depicts the pipeline used for this comparative analysis. The percentage of secreted proteins was 6% for the two *Fusarium* species (947/15,830 for FsK, 866/15,705 for *F. vanettenii*), 1% for *Rhizophagus irregularis* (266/26,143) and 5% for *Serendipita indica* (555/11,766) (Fig. 3b). From the predicted secretome, proteins with length less than 300aa, classified as effectors by the software EffectorP v3.0, were assigned as part of the predicted effectome for all strains, in accordance with methods that have been used to classify effector proteins before[50,51]. FsK contains more predicted proteins belonging in the predicted secretome and effectome, compared to its pathogenic relative (947 vs 866 and 245 vs 230 respectively; Fig. 3b). These findings contrast with recent large scale comparative genomics analysis of 150 fungal species, where the secretome prediction derived from non-pathogenic fungi had on average lower number of secreted proteins compared with the ones from pathogens[52]. The highest number of secreted proteins on average were found on hemibiotrophic and then necrotrophic fungi, such as *F. vanettenii*. Taking into account the documented biocontrol properties and plant growth promoting capacity of FsK as well as its wide range of host plants in at least two plant families (Solanaceae and Leguminosae)[14–16,21,53], the increased number of putative effectors predicted in FsK might be indicative of mechanisms for host adaptation, via providing multiple mechanisms to evade host defenses and modulate host physiology[52,54].

CAZymes were also assigned in the secretome analysis (Fig. 3c) and classified as potential apoplastic or cytoplasmic using EffectorP v3.0 The pathogenic *Fusarium* contains more apoplastic CAZymes than FsK (24 vs 19), probably due to the higher need of pathogens to scavenge nutrients[23,29] compared with symbiotic

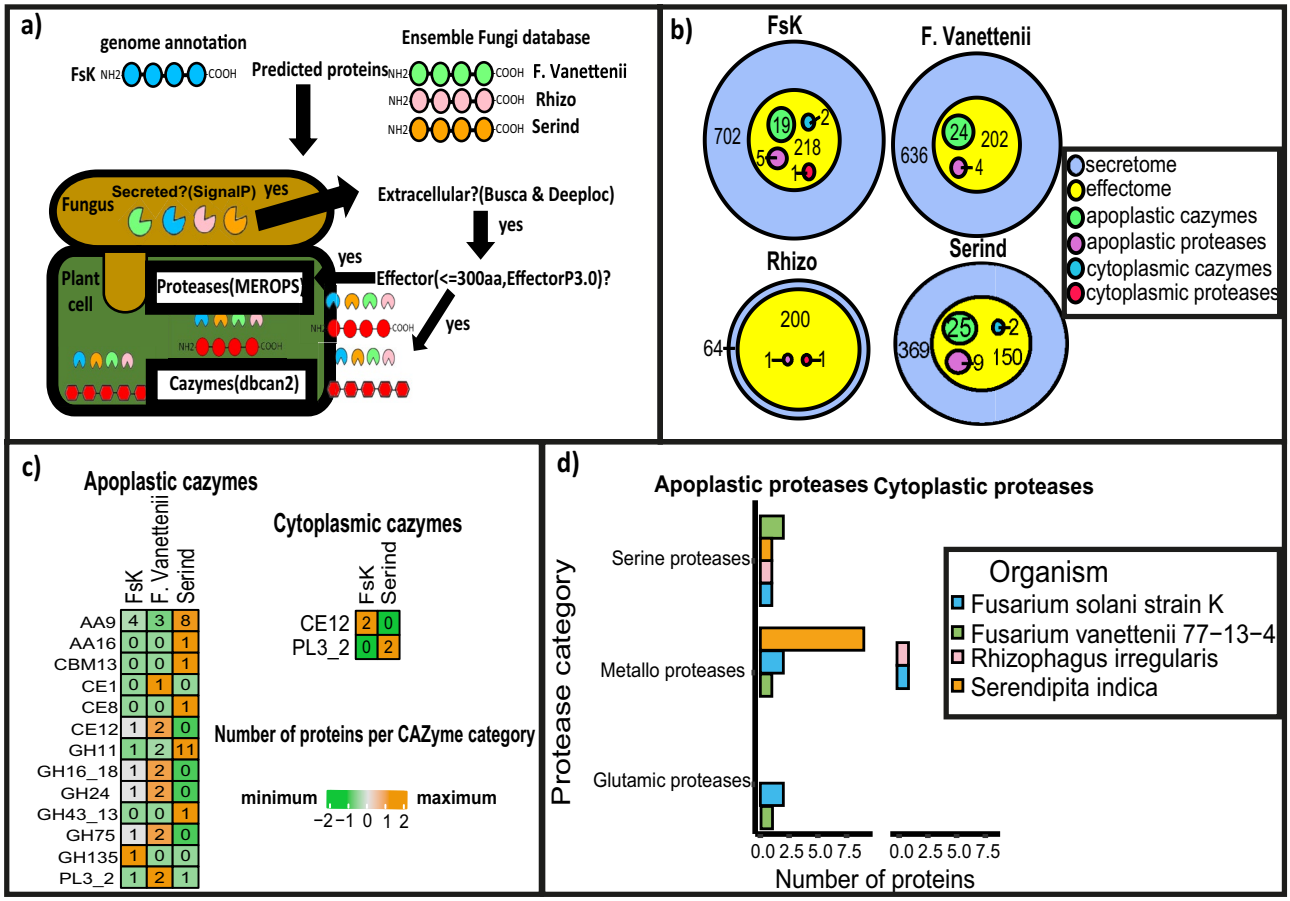

Fig. 3. Fungal comparative genomics. (**a**) Pipeline used for the identification of fungal proteases and carbohydrate active enzymes (CAZymes), that can be secreted in the apoplast or the cytoplast of plant cells. The software and databases used for this analysis are shown in paranthesis. (**b**) Number of proteins per fungus belonging to the effectome, secretome and counterclockwise (as in FsK) apoplastic CAZymes, apoplastic proteases, cytoplasmic proteases and cytoplasmic CAZymes. (**c**) Heatmap depicting the maximum/minimum number of proteins per CAZyme category per organism. *Rhizophagus irregularis* was not found to contain any predicted effector apoplastic CAZymes, thus it was omitted from the heatmap. For the same reason, *Fusarium vanettenii* and *Serendipita indica*, are not present in the heatmap of predicted effector cytoplasmic CAZymes. (**d**) Barplot depicting the number of apoplastic and cytoplastic effectors per organism per protease category. FsK: Fusarium solani strain K, F. vanettenii: *Fusarium vanettenii 77–13-4*, Rhizo: *Rhizophagus irregularis*, Serind: *Serendipita indica*.

fungi. Moreover, the pathogenic *F. vanettinii* may produce additional CAZymes to overcome plant immune responses and cell wall reinforcement, which are important host defense mechanisms[23]. On the contrary, the non-pathogenic strain FsK likely utilizes cytoplasmic CAZymes for saprophytic growth[23,55]. For *R. irregularis,* which, in our analysis presented the lowest presence of secreted proteins and CAZymes, we confirmed previously reported studies that report that arbuscular mycorrhiza fungi lack the production of plant cell wall degrading enzymes[56]. While the number of effector CAZymes in *S. indica*, has not been reported before, it is well-established that, unlike many arbuscular mycorrhizal fungi (AMF) and other root endophytes that require living plant tissue, *S. indica* can proliferate and colonize both living and dead host cells. This flexibility may be supported by its distinct CAZyme repertoire that enables it to modify plant cell walls differently[57,58].

The overwhelming majority of CAZymes found in the predicted effectome of this study are predicted as localized in the apoplast (Fig. 3c). FsK lacks 5 of the 13 CAZyme classes as well as enzymes involved in the degradation of xylan (i.e. CE1 with EC number 3.1.1.72, CE12 with EC number 3.1.1.72, GH11 with EC numbers 3.2.1.8 and 3.2.1.27), chitin (GH75 with EC number 3.2.132) and pectin (PL3_2 with EC number 4.2.2.2) (Fig. 3c). This potentially indicates that FsK can colonize the host plant, using less intrusive mechanisms of entry into the plant tissues than degradation of these significant polymers within the plant cell wall[59]. In our comparison, the GH135 fungal CAZyme category is present only in FsK. GH135 enzymes catalyse the breakdown of galactosaminogalactan, an exopolysaccharide composed of galactose and N-acetylgalactosamine that was found to be secreted from fungal pathogens[60]. Galactosaminogalactan functions as the dominant adhesin of the human pathogen *A. fumigatus* and mediates adherence to plastic, fibronectin, and epithelial cells as well as supports biofilm formation[61]. While the role of this CAZyme category plant-fungal interactions has not been investigated as yet, it is possible that FsK uses the extracellular enzymes to attack other fungi during the competition for niche. This is compatible with the biocontrol role of FsK against root pathogens[14]. In comparison with FsK, many of the CAZyme categories are overrepresented or solely present in *S. indica*. For example, the AA9 category of lytic polysaccharide monooxygenases that facilitate the breakdown of parts of cellulose of the CE8 family of carbohydrate esterases, involved in the breaking down of pectin ester bonds into pectate methanol, is represented with a higher number of effectors in *S. indica*. Furthermore, the family AA16 (EC number 11.14.99.54) CAZymes that function as cellulose monooxygenases that create esters of cellulose molecules, are only present in *S. indica*[62]. Although these enzymes are present on many pathogenic species and are upregulated during plant infection[63], their presence in *S. indica* is probably related to covering the nutritional needs of the fungus. In the cytoplasm, only two CAZyme families were predicted as present in FsK and *S. indica*: CE12, and PL3_2. The CE12 family includes pectin acetylesterases (EC 3.1.1.11), which catalyze the deacetylation of pectin to pectate; however, the predicted CE12 protein lacks a clearly defined functional annotation (Supplementary Table S2). The PL3_2 family comprises pectate lyases (EC 4.2.2.2), enzymes involved in the degradation of pectin, a major structural component of the middle lamella and outer layer of the plant cell wall[64,65]. In non-pathogenic fungi, these enzymes may contribute to nutrient acquisition. In addition, cytoplasm-localized effectors may perform alternative, potentially non-enzymatic functions during plant–microbe interactions[66–68]. Finally, we cannot exclude the possibility that the predicted cytoplasmic localization of these proteins may represent false-positive result of the algorithm, as their annotated enzymatic activities are more consistent with apoplastic localization and plant cell wall degradation[59].

In summary, compared with the pathogenic *Fusarium* strain, the beneficial FsK has a reduced number of CAZymes in the apoplast. In addition, the CAZyme repertoire available in the three endophytic beneficial fungal strains varies, which probably depicts their different mode of successful entry and establishment in the host plants.

By comparing the apoplastic proteases of FsK with the identified ones in the other pathogenic and non-pathogenic strains in this study (Fig. 3d), we found that only the two *Fusarium* species encode peptidases of the glutamic family (G1). G1 proteases are characterized as carboxyl-peptidases[69]. Their role in host–pathogen co-evolution, host adaptation and pathogenicity has been shown in the ascomycete wheat pathogen *Zymoseptoria tritici and* its relative fungi[70]. In the study of Krishnan et al.[70], the multiple duplications of the gene encoding G1 proteins were proposed as a mechanism of rapid diversification. It may, thus, be proposed that the extra G1 protein in FsK, not present in the other three fungal species, may be linked with a different host specificity of the strain. Furthermore, a protease family using serine in the catalytic center had more proteins in *F. vanettenii*'s predicted effectome that the other three beneficial fungal strains. These results align with large analysis on 634 fungal proteomes that revealed that serine proteases are overrepresented in pathogenic fungi[71]. An S9 protease, predicted to cleave peptide bonds where proline is involved[72], is only present in the pathogen *F. vanettenii* and not in FsK. Since proline is a rigid molecule, it is utilised for the reconstruction, integrity and elongation of the plant cell wall[73]. Therefore, the absence of this enzyme from FsK might indicate a less violent and invasive approach during its host colonization. Finally, metalloproteases in the apoplastome are overrepresented in *S. indica* and contain proteases of the family M43, that degrades collagen and gelatin. Their presence in *S. indica* is previously documented, induced after the infection of dead barley roots[74]. Furthermore, as saprophytic fungi exist in the genus *Fusarium*[75,76], and as this family is also present in FsK but not in *F. vanettenii*, this may suggest an ecological distinction between the two species. Specifically, the presence of this family in FsK may indicate that FsK can also act as a saprophyte under certain conditions. However, the role of these proteases in saprophytic members of this genus remains poorly characterized.

Identification of FsK-specific proteins

We proceeded with our analysis by creating, by orthologous inference analysis, FsK-specific protein groups: (a) proteins that are present only in FsK when all four fungal species were compared and they do not belong in the same functional categories as proteins of the pathogenic *Fusarium* ("FsK specific with unique function vs *F. vanettenii*", Supplementary Fig. S3c) and (b) FsK proteins that have no homology with the pathogenic

Fusarium but are othologous with either one or both beneficial endophytic fungi (Fig. 4). We first identified FsK proteins that do not share common ancestor with the orthologs of the pathogen *F. vanattennii*. These 1244 proteins were further filtered to contain only those with no homology in the functional annotation assignment in the pathogen, which lead to 904 proteins that were grouped as "FsK specific vs *F. vanettenii*" (Supplementary Fig. S3a). Secondly, the orthology inference analysis was repeated to identify FsK specific proteins that do not share a common ancestor with two beneficial species *S. indica* and *R. irregularis* or the pathogen. This analysis resulted in 1191 FsK specific proteins (named "FsK specific vs All") (Supplementary Fig. S3b). Next, the common proteins derived from the two subsets of "FsK-specific proteins" were further filtered to exclude those with the same functional categories with the 993 pathogen-specific proteins collected in the previous step. This resulted in a subset of only 14 proteins ("FsK specific with unique function vs *F. vanettenii*") that show the greatest functional divergence from those of the other three species (Fig. 4c). The assignment into functional categories is shown in Table 2. Most of those FsK-specific proteins are involved in transcription and DNA damage repair response networks, such as DNA helicases[77]. In plant-pathogenic fungi, DNA damage repair mechanisms were implicated as a response to the host production of reactive oxygen species[78], which, in plants, also act as essential systemic signaling agents during plant-(a)biotic interactions. ROS exert differential actions on microbes[79] and plants can selectively filter out microbial taxa that are not ROS-tolerant[80]. Whether or not such DNA damage repair mechanisms play a role in FsK's colonization in the various plant hosts requires further investigations. Another FsK-specific protein is represented by an N-acetylglucosaminyl transferase, reported as involved in

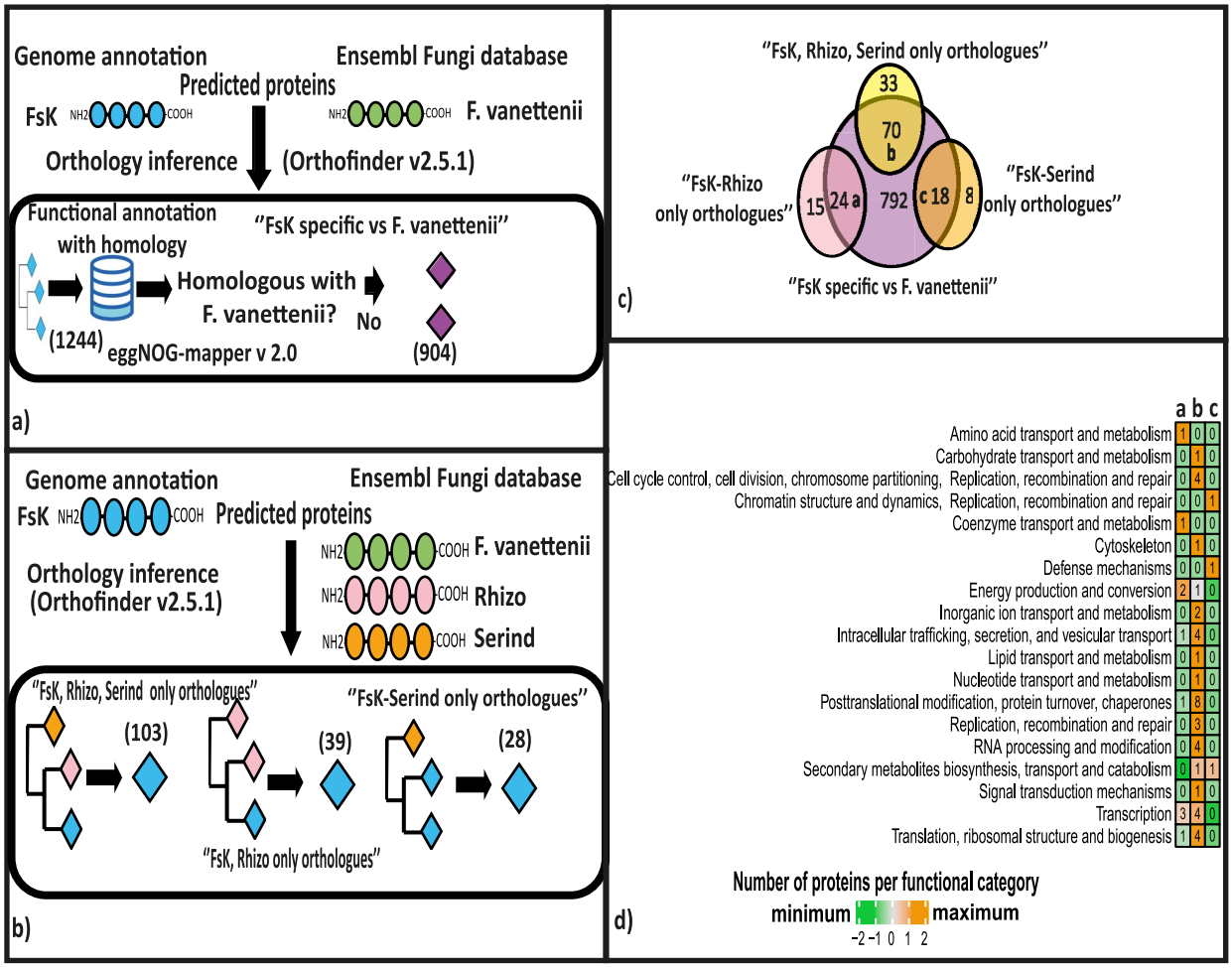

Fig. 4. Identification of FsK proteins non homologous to F. vanettenii but orthologous to either only Rhizo, only Serind or to both Rhizo and Serind proteins. Programs and databases used for this analysis are shown in parentheses. (**a**) Inference of orthologous proteins between the beneficial FsK and the pathogen F. vanettenii according to Orthofinder, and non-homologous proteins between FsK and F. vanettenii, according to Orthofinder or eggNOG-mapper and F. vanettenii BLAST functional annotation. (**b**) Orthofinder analysis of proteins from FsK, F. vanettenii and the beneficials Rhizo and Serind and retrieval of FsK proteins non-homologous to F. vanettenii but orthologous to either only Rhizo, only Serind or with both Rhizo and Serind. (**c**) Merging of the two FsK datasets of (**a**) and (**b**) and retrieval of final FsK protein sets. (**d**) Heatmap depicting the maximum/minimum number of proteins per merged subset a,b,c for each functional category. The numbers in parenthesis show the number of proteins per step in the pipeline. FsK: *Fusarium solani strain K*, F. vanettenii: *Fusarium vanettenii 77–13-4*, Rhizo: *Rhizophagus irregularis*, Serind: *Serendipita indica*.

Id	Functional category	Functional description
FUN_002565	Amino acid transport and metabolism, Coenzyme transport and metabolism	Belongs to the TPP enzyme family
FUN_005499	Carbohydrate transport and metabolism, Cell wall/membrane/envelope biogenesis	3-beta-hydroxy-delta5-steroid dehydrogenase activity
FUN_000300	Cell cycle control, cell division, chromosome partitioning	Structural constituent of cuticle
FUN_007028	Cell wall/membrane/envelope biogenesis, Posttranslational modification, protein turnover, chaperones	N-acetylglucosaminyl transferase component (Gpi1)
FUN_012231	Chromatin structure and dynamics	GAL4-like Zn(II)2Cys6 (or C6 zinc) binuclear cluster DNA-binding domain
FUN_001600	Defense mechanisms	Beta-lactamase
FUN_003870	Defense mechanisms	peptidase activity
FUN_001959	RNA processing and modification	Domain with 2 conserved Trp (W) residues
FUN_003521	RNA processing and modification	DNA helicase
FUN_004479	RNA processing and modification	39S ribosomal protein L53/MRP-L53
FUN_007587	RNA processing and modification	Sec63 Brl domain
FUN_011856	RNA processing and modification	DNA helicase
FUN_012199	RNA processing and modification	DNA helicase
FUN_015535	RNA processing and modification	Spliceosome-associated protein

Table 2. Protein id, Functional category and Functional description of *Fusarium solani strain K* specific proteins. These are proteins from functional categories non present in the species specific proteins of the pathogen *Fusarium vanettenii mpVI 77–13-4*.

posttranslational modifications, as it catalyzes the transfer of a N-acetylglucosamine (GlcNAc) moiety from UDP-GlcNAc to Serine or Threonine residues on cytoplasmic and nuclear proteins in filamentous fungi[81]. An N-acetylglucosamine transferase has been discovered in *Fusarium oxysporum*, playing a role during infection of tomato plants[82]. Finally, the protein Sec63, a component of a pathway involved in the secretion of fungal proteins from the endoplasmatic reticulum (ER) to the membrane or the extracellular space, as well as interacting with chaperones to assist in protein folding, appears to be present in the FsK-specific set of proteins[83,84]. An early-stage role in colonization and/or a role in host selection might be indicative in this case, since based on the type of effector proteins secreted and membrane proteins presented, the fungal interaction with various hosts might be affected.

Identification of FsK proteins with homology only with symbiotic fungi

We have identified three protein subsets that include the orthologous proteins of FsK with each of the beneficial fungi or with both of them but not with the pathogen (Fig. 4b): 39 FsK proteins were orthologous only with *R. irregularis,* 28 only with *S. indica* and 103 FsK proteins were orthologous with both *S. indica* and *R. irregularis*. Finally, we excluded the homologous to the pathogen FsK proteins (Fig. 4a) to get three subsets that are possibly most relevant to the beneficial character of the three phylogenetically distant endophytes: 24 FsK proteins are orthologous to proteins of *R. irregularis* (subset a in Fig. 4c), 18 FsK proteins share the same ancenstral genes with those of *S. indica* (subset c in Fig. 4c) and 70 FsK proteins share the same ancestry with those from both *R. irregularis* and *S. indica* (subset b in Fig. 4c). These were assigned to functional categories as shown in Fig. 5d (annotations shown in Supplementary Tables S4–S6). Several proteins were found with a known physiological role in plant fungal interactions: FUN_008856 in the "Defense mechanisms" functional category, is annotated as an isochorismatase. Interestingly, the absence of predicted secretion signals in both FsK and its S. indica ortholog support a functional divergence of FUN_008856 from canonical isochorismatase effectorsdescribed in plant pathogens, such as *Phytophthora soj*ae and *Verticillium dahliae*, with a conserved role in host salicylic acid suppression[85]. Moreover, divergent functional annotation—including the predicted nicotinamidase activity in S. indica—suggest that these orthologs may encode distinct enzymatic activities and the absence of homologs in the pathogenic Fusarium strain, further support the hypothesis that FUN_008856 has undergone evolutionary divergence and likely fulfills a role unrelated to classical immune evasion in FsK.FUN_006145 in the "RNA processing and modification" functional category, FsK orthologous with both non-pathogenic fungi, interacts with spliceosome-associated proteins, preventing the accumulation of proteins in the spliceosome[86]. The deletion of the gene encoding this protein in *Magnaporthe oryzae* resulted in aberrant intron splicing in nearly 400 genes, 20 of which were important to the fungal development and virulence[86]. A role in beneficial interactions has not as yet been described and this protein represents a possible target for future studies. Finally FUN_003856, a Bola domain-containing protein, homologous in both non-pathogenic fungi, is involved in signal transduction mechanisms by blocking or allowing transcription activators or repressors to bind to promoters of Fe responsive genes and control Fe regulation and uptake in yeast and humans[87] and to have a role in virulence in pathogenic fungi[88]. In FsK, we have previously shown that this gene was induced under low Fe availability conditions, and during the interaction of FsK with *L. japonicus*, where the presence of the beneficial endophyte had a clear benefit in nutrient acquisition for the plant as well[21]. Its presence in the other two phylogenetically distant beneficial endophytes may indicate a common mechanism adopted by symbionts for the acquisition of nutrients, such as iron and other metals.

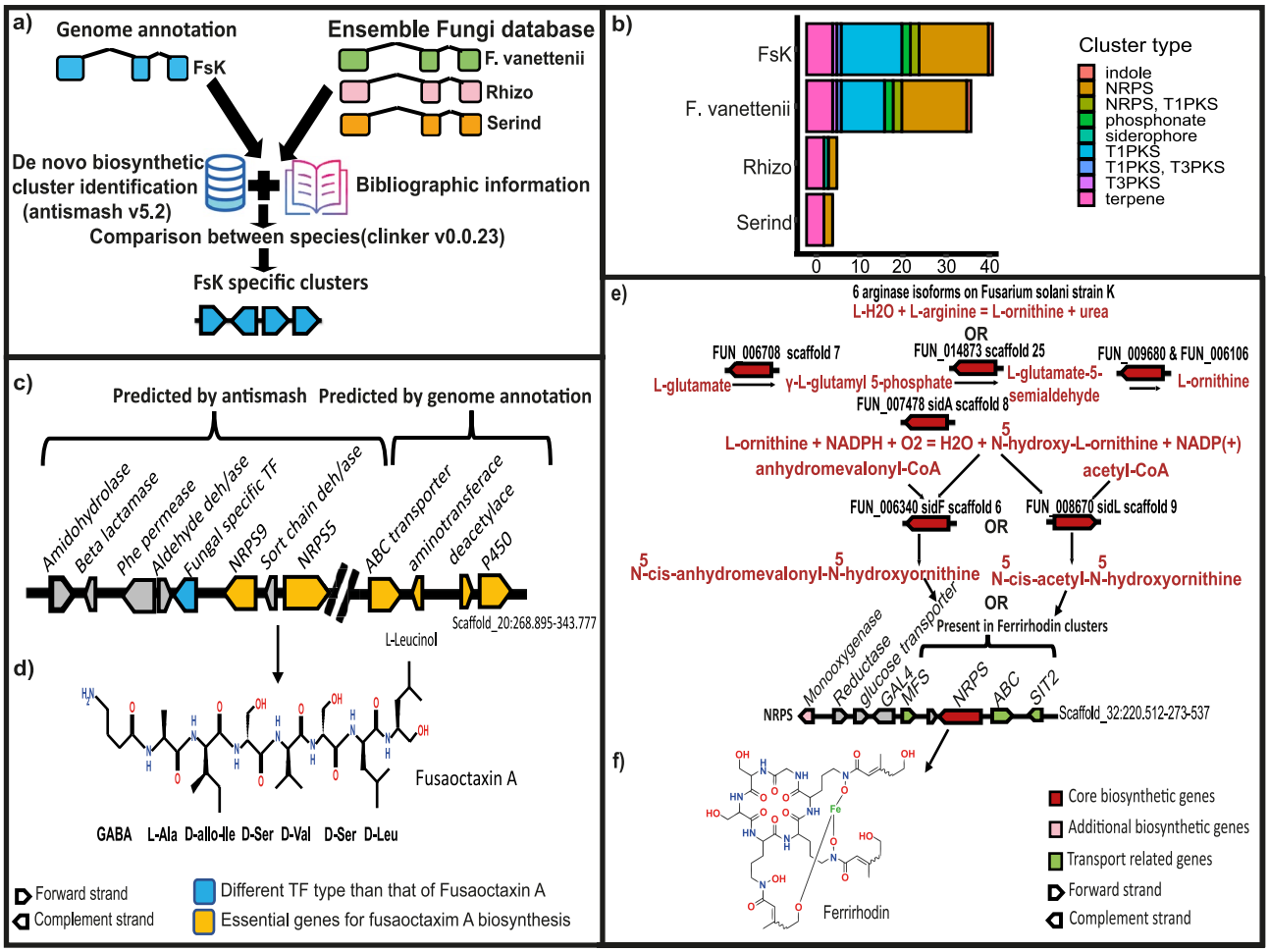

Fig. 5. (**a**) Pipeline for the identification of FsK specific biosynthetic gene clusters or FsK with similarities or differences to F. vanettenii, Rhizo and Serind. Software and databases used for this analysis are shown in parentheses. (**b**) Barplot depicting the number secondary metabolite clusters per organism per cluster category. (**c**) The Fusarium solani strain K specific Fusaoctaxin A-like biosynthetic gene cluster. Genes highlighted with orange are essential for the production of Fusaoctaxin A. A different transcription factor from that of Fusaoctaxin A is present in the currect cluster as shown in blue. Part of the gene cluster was predicted by in silico prediction (antismash v5.2) and a second part 61 kb downstream was identified from the genome annotation of Fusarium solani strain K (**d**) The structure of Fusaoctaxin A. (**e**) Predicted pathways for the biosynthesis of ornithine and siderophore precursor molecules on Fusarium solani strain K (**f**) Fusarium solani strain K specific ferrirhodin-like biosynthetic cluster. Genes are color highlighted based of the secondary metabolite biosynthetic cluster prediction software antismash v5.2, as those responsible for the synthesis of the basic backbone of the metabolite (Core), those responsible for further structural modifications (Additional), and transport of those compounds. NRPS: non-ribosomal peptide synthase NRPS: non-ribosomal peptide synthase, T1PKS, T3PKS: type 1 and type 3 polyketide synthases, MFS: major facilitator superfamily, ABC: ABC transporter family, SIT2: siderochrome iron transporter 2; sidA, L-ornithine N (5)-monooxygenase, sidF: Hydroxyornithine transacylase, sidL: Acyltransferase FsK: *Fusarium solani strain K*, F. vanettenii: *Fusarium vanettenii 77–13-4*, Rhizo: *Rhizophagus irregularis*, Serind: *Serendipita indica*.

Comparison of biosynthetic gene clusters in all four fungi

The comparative genomic analysis of FsK was expanded to investigate putative biosynthetic gene clusters that are common or differentiate the four fungal species under study. Following the bioinformatics analysis depicted in Fig. 5a, a higher number of biosynthetic gene clusters were identified in the genomes of FsK and *F. vanetten*ii compared with the ones present in the genomes of the two distant beneficial fungi (Fig. 5b). *R. irregularis* and *S. indica* share with each other two terpene and one non-ribosomal peptide (NRP) biosynthetic cluster. However, none of their clusters show similarity to those found in the *Fusarium* species included in this study. On the other hand, only the two *Fusarium* species contain putative polyketide (PK) biosynthetic clusters; two clusters are present only in the beneficial FsK, although none of them are characterized before. In addition, two (out of a total of 16) NPR and one (out of 5) terpene biosynthetic clusters were detected only in FsK.

Two of the FsK-specific gene clusters have been previously described: The NPR cluster in scaffold 20 (Fig. 5c) contains the non-ribosomal peptide synthetases (*NRPS*) 9 and 5 that are responsible for the production

of fusaoctaxin A[89]. Fusaoctaxin A is a virulence factor that consists of amino acids and GABA that enables pathogenic fungus *Fusarium graminearum* to colonize wheat[89]. During the fusaoctaxin A biosynthesis, NRPS9 initiates the synthesis by delivering the first amino acid (GABA), while NRPS5 extends the peptide with additional amino acids. *NRPS9* gene knock out mutants lead to greatly reduced lesion sizes on wheat (about 30% of wild-type), impaired cell-to-cell hyphal invasion and loss of full virulence compared to wild-type strains[89]. Importantly, the gene encoding the transcription factor FGM4 that acts as a global activator of the whole cluster expression[90] is not present in FsK cluster. In addition, a part of the biosynthetic gene cluster, encoding for genes not directly responsible for the peptide's production but with a clear role in fungal invasiveness[90], was detected 61 Kb downstream of NRPS5. We have performed limited expression analysis of the two NRPS genes in FsK grown under in vitro conditions and in root tissues colonized by FsK. We detected significant expression of *NRPS5* but not of *NRPS9*, only under in vitro conditions as well as absence of any expression *in planta* colonized roots (Supplementary Fig. S4). The potential difference in structure and/or transcriptional regulation may be key in absence of production of virulence factors or may indicate the production of novel compounds, leading to FsK's non-pathogenic nature and symbiotic phenotype. The absence of expression of *NRPS9* might permit to FsK a sub-optimal level of virulence and plant colonization without excessively destroying plant tissues. Alternatively, it might also be an indicator of FsK's possible opportunistic nature and the product deriving from the gene cluster may be a virulence factor activated only in specific hosts. These results indicate the importance of whole genome analysis towards predicting and identifying putative risk factors in the use of beneficial fungi in agriculture.

The second NRP biosynthetic cluster identified as FsK-specific in our analysis (Fig. 5e) contains a synthase that has 71% identity with a ferrirhodin synthetase from the sugarcane pathogen *Fusarium sacchari* (NCBI id: AFD36451.1 -Supplementary Table S7)[91,92], which is involved in the synthesis of ferrirhodin, a peptide acting as a siderophore. Siderophores are compounds that exist in bacteria, fungi and plants and possess high affinity for iron; therefore, they can function as Fe acquisition and storage units[93]. The FsK-specific cluster contains, also, the transport genes that are usually present in a ferrirhodin cluster, including EmrB transporters, of a major facilitator superfamily protein that can import various compounds, like glucose[94], needed for the synthesis of ferrirhodin and other related siderophores[95]. The other two transport-related genes encode an ABC transporter, which is also present in other siderophore related biosynthetic clusters[95], along with a siderochrome iron transporter 2 (SIT2) that is responsible for the uptake of ferrirhodin into fungal cells[92]. Siderophore compounds are derived from ornithine[91,96]. All the genes necessary to produce precursor molecules[97,98], as deduced from MetaCyc, were identified in the genome of FsK (Supplementary Table S8). Iron acquisition, mediated by fungal siderophores such as ferrirhodin, is essential to homeostasis within iron limiting environments[99]. As mentioned, we have recently showed that FsK is able colonize more efficiently and promote plant growth of the model legume plants *L. japonicus* under iron depleting conditions[21]. Therefore, elucidating the role of ferrirhodin in the capacity of the endophyte to establish and sustain a beneficial phenotype may lead to unexplored mechanisms that underlie plant-fungal interactions.

As referred above, the other FsK strain-specific clusters are poorly characterized (Supplementary Fig. S4 and Supplementary Table S9). The terpene cyclase core gene in the FsK strain-specific cluster has 70% identity with an isoprenoid synthase from the ascomycete *Lyonectria destructans* (KAH7010505.1). Other genes, encoding for SAM-methyltransferases and P450s enzymes are present (Supplementary Table S9). Terpene derivatives have been implicated in many plant-biotic interactions, including beneficial microorganisms[100,101]. Interestingly and indicating that this biosynthetic cluster warrants further investigation, unknown terpene molecules were found to be systemically accumulating in tomato plants, only when the plants were colonized by FsK[16]. A time-course and context-depended gene expression analysis on *L. japonicus*- and *S. lycopersicum*-FsK interaction as well as FsK mutant analyses during host plant colonization, is warranted to validate the roles of these candidate proteins and metabolites in modulating symbiotic versus parasitic interactions and fungal adaptation.

Conclusion

The elucidation of differences in virulence, host adaptation, and specificity among fungi is considered crucial for the identification of key proteins that may serve as targets for functional characterization or optimization strategies aimed at modulating fungal phenotypes. Such insights are expected to develop efficient and knowledge-driven microbial inocula with capacity for enhancing crop resilience and productivity, thereby promoting sustainable agricultural practices.

In the present study, the taxonomic placement of *F. solani* strain K (FsK) was verified, employing advanced phylogenomic methodologies. FsK, uniquely characterized as the sole symbiotic member of the FSSC capable of colonizing both legumes and tomato plants based on current information, was subjected to comparative genomic analyses. Comparative assessments involving FsK, the closely related pathogenic ascomycete *F. vanettenii* 77–13-4, and the symbiotic fungi *S.indica* (Basidiomycota) and *R. irregularis* (Glomeromycota) —all capable of tomato colonization— recognized putative molecular determinants underpinning fungal colonization and adaptation strategies, including putative effects on phytohormone-mediated plant responses as well as a role in nutrient homeostasis. Distinct genomic features distinguishing FsK from pathogenic and mutualistic fungi have been highlighted. FsK-specific proteins implicated in DNA repair, membrane dynamics, and effector secretion, which may contribute to environmental adaptation and the initiation of fungal-host interactions were identified. Apoplastic and cytoplasmic CAZymes and proteases analysis demonstrated that the pathogenic *F. vanettenii* harbors a more extensive repertoire of enzymes involved in plant cell wall degradation, whereas FsK predominantly relies on cytoplasmic enzymatic activity for nutrient acquisition, indicative of a saprophytic or non-pathogenic lifestyle. Moreover, the discovery of an FsK-specific gene cluster encoding enzymes for the biosynthesis of ferrirhodin, an iron-chelating compound, suggests an adaptive mechanism facilitating iron acquisition under limiting conditions, potentially benefiting both fungus and host. On the other hand, the

detection of the Fusaoctaxin A gene cluster, albeit not functionally characterized, as well as a number of other NRPS and PKS biosynthetic gene clusters in the genome of the beneficial endophyte cautions towards a more thorough investigation of the putative opportunistic nature of the fungus.

The bioinformatic pipelines developed and performed in our case study serves as an example of the highly important information that may be deduced from whole genome sequence efforts for microorganisms with potential as novel biostimulant and biopesticide use in agriculture and calls for the development of appropriate tools for this assessment.

Materials and methods

Fungal material-genome sequencing and assembly

Fusarium solani strain K (FsK)[14] was maintained on Potato Dextrose Agar (PDA; Difco) plates at 25 °C for 5 days. Hyphal propagules were prepared for inoculation by gently scraping the surface of fresh FsK cultures on PDA with a sterile rod. Mycelium was collected with a sterile scalpel and ground using the TissueLyser II system (www.qiagen.com/us) with stainless steel beads. DNA was extracted using CTAB[102]. DNA concentration and integrity were assessed after isolation using Qubit Fluorometer 2.0 and gel electrophoresis (0.8% agarose gel) and by Bioanalyzer 2100 (Agilent Technologies), respectively. Five (5) µg were used for genome sequencing.

Genome assembly was performed using a hybrid strategy, by combining long and short sequence reads. The long sequence read dataset was generated with the MinION sequencing device, the R9.4.1 FLO-MN106 flow cell, and the SQK-LSK109 1D kit (Oxford Nanopore Technologies, Oxford, UK). Resulting reads were assembled with the SmartDeNovo[103] assembly wrapper, and were error-corrected with Pilon v1.23[104], using high quality Illumina reads generated with the MiSeq Illumina instrument for 600 cycles (2 × 300 bp), performed at the MR DNA sequencing facilities (Shallowater, TX, USA), and quality trimmed with Trimmomatic v0.36[105]. The hybrid assembly was performed with Spades v3.13.1[106].

Experimental setup and RNA-seq analysis

For the prediction of gene models in FsK, used for functional analysis and comparative genomics, RNA sequences of the fungus were required. For this purpose, an RNA sequencing data was obtained from plant and fungal tissues during the interaction of FsK with *L. japonicus.* The plant species used is a model legume plant, used in research laboratories globally (see Lotus base, https://lotus.au.dk/). The handling and use of this material is fully compliant with relevant institutional, national, and international guidelines and legislation. The plant material used in this study a Lotus japonicus (ecotype Gifu) plant that was previously obtained through a collaboration with Prof Anne Elisabeth Osbourn (Professor of biology and group leader at the John Innes Centre, U.K.) more than 10 years ago (see Delis et al.[107] and Krokida et al.[108]), and was propagated in our laboratory ever since. *L. japonicus* wild-type (WT) seeds (ecotype Gifu) were chemically scarified with a 12-min sulfuric acid treatment, washed thoroughly with deionized water, surface sterilized for 20 min in a 20% NaOCl (v/v) solution and washed 6 × with sterile deionized water. Seeds were kept in water at 4 °C overnight, to promote seed germination, placed afterwards on half strength MS medium[109], and grown in a growth chamber with a 16 h light/8 h dark photoperiod at 22 °C. Eleven (11) day old plants were transferred to M medium plates[110] and inoculated on the same day with FsK hyphal propagules, by slightly scrapping with a sterile needle the surface of a fresh FsK culture on PDA medium (26 °C, 5 d in the dark), at a distance ~ 1 cm beneath the root tip (Supplementary Fig. 6). Control plants received no inoculum. As a second control, M medium plates were also inoculated with fungal hyphae, in the absence of the plant. Plant and fungal growth was allowed to take place in a growth chamber (16 h light/8 h dark photoperiod at 22 °C) until contact between the two partners occurred (4 dpi; fungal hyphae had sufficiently contacted the lower part of the root). Contact between the two partners was verified via microscopy. At this point, control and inoculated root tissues were collected, washed thoroughly with distilled water sterile to remove the extra-radical fungal mycelium, flash frozen on liquid nitrogen and kept at -80 °C until subsequent RNA extraction. Fungal mycelium growing alone was also harvested, flash frozen, and kept for RNA extraction. *L. japonicus* root tissues were ground with a pestle, under the presence of liquid nitrogen and total RNA was extracted using the Isolate II RNA Plant Kit (BIO-52077, BIOLINE) according to manufacturer's instructions. Total RNA was extracted from fungal mycelium using a modified TRIzol™ (www.thermofisher.com)/TRItidy G™ (www.itwreagents.com) and LiCl precipitation protocol. Samples (100 µg) were lysed in 900 µl freshly prepared buffer containing 0.1 M Tris–HCl (pH 7.5), 10 mM EDTA, 0.5 M LiCl, 1% lithium dodecyl sulfate (LDS), and 5 mM DTT. After centrifugation (13,000 rpm, 10 min, 4 °C), 850 µl of the supernatant was mixed with 1000 µl TRItidy G™ reagent, vortexed, incubated at room temperature (RT) for 5 min, and extracted with 250 µl chloroform. After vortexing and a 2-min RT incubation, samples were centrifuged (13,000 rpm, 10 min, 4 °C). The aqueous phase was mixed with 0.1 volumes each of 3 M sodium acetate and 1 M acetic acid, followed by 2.5 volumes of cold 100% ethanol. After incubation for 30 min at RT and overnight at –20 °C, RNA was pelleted by centrifugation (13,000 rpm, 60 min, 4 °C), washed with 70% ethanol, air-dried, and resuspended in 20–40 µl RNase-free water. Elimination of genomic DNA contamination of RNA samples was performed by DNaseI (www.thermofisher.com) treatment according to the manufacturer's guidelines. DNA elimination was verified by PCR using primers specific to *TEF1a* gene (translation elongation factor 1-alpha), followed by agarose gel electrophoresis. RNA was dissolved by heating at 65 °C for 2 min and immediately cooled on ice. RNA concentration and integrity were assessed after isolation using Qubit Fluorometer 2.0 and gel electrophoresis (2% agarose gel) and by Bioanalyzer 2100 (Agilent Technologies), respectively.

Three biological replicates for each treatment (1 µg of each) were prepared for subsequent RNA sequencing analysis. Briefly, libraries were prepared with the KAPA stranded RNA-Seq Library Preparation kit (KAPA Biosystems, Wilmington, MA, USA), and ~ 10 million paired-end reads of 250 nucleotides each were obtained using an Illumina HiSeq 2500 in Rapid mode. Library prep and sequencing protocols were performed at the DNA

Sequencing Center of the Brigham Young University (BYU, Provo, UT, USA) using a HiSeq 2500 instrument for 500 cycles (2 × 250 bp).

Raw counts per gene were retrieved with htseq-count from package HTSeq v0.12.4[111]. Differential expression analysis was performed using the package package DESeq2 v1.41.1[112]

Genome functional annotation

Gene models of closely related fungal strains along with transcriptome evidence and homology searches were employed for the genome structural and functional annotation as described here. Quality trimming of the Illumina sequencing of mRNA enriched RNA extracts dataset was performed with Trimmomatic v0.36[105] using default parameters, additional filtering steps were applied, including removal of TruSeq3 paired-end adapters, trimming of the first three and last three bases, removal of terminal bases with a sliding-window mean Phred quality score below 20, and discarding of reads shorter than 150 bases. The trimmed files were mapped to FsK genome using Star v020201[113]. From the resulted files, de novo transcriptome and genome-guided assembly was performed using trinity v2.4.0[114]. For the structural gene annontation, the de novo and genome guided transcrips were combined to the pipeline funnanotate v1.8.3[115] as transcript evidence in the command funannotate predict, with the *Fusarium vanettenii* 77–13-4 as AUGUSTUS gene models species[116] and protein evidence; de novo predicted gene models with GeneMark-ES[117]; a soft-masked genome made using the command funannotate mask with default options and –max_intronlen 3000. Functional annotation was made with funannotate annotate and eggNOG mapper v2.0[118]. Metabolic pathways were retrieved using blastKOALA v3.0[119] and results were aggregated using KEGG Mapper Reconstruct[120]. The predicted annotations are provided in the supplementary File S1.

FSSC phylogenetic tree

For the determination of taxonomic position of FsK within the Fusarium solani species complex (FSSC), a phylogenetic tree was constructed using Orthofinder v2.5.2[121], with default parameters. For tree construction, the most recently sequenced strain available for each species was selected based on NCBI GenBank records. According to the NCBI database (May 2025), 110 FSSC strains were genome sequenced and 21 were functionally annotated. Of those, only FsK and the pathogen F*usarium vanettenii 77–13-4* are able to colonize tomato roots. The tree was rooted using a diverse set of fungal outgroups, including *Rhizophagus irregularis* DAOM 181602 = DAOM 197198 (GCA_000439145), *Serendipita indica* DSM 11827 (GCA_000313545), and representatives from key fungal lineages: Ascomycota (closely related to *Fusarium*): *Aspergillus niger, Neurospora crassa*; Basidiomycota (closely related to Serendipita): *Ustilago maydis, Cryptococcus neoformans*; Mucoromycota (closely related to Rhizophagus): *Mucor circinelloides, Lichtheimia corymbifera*; and the model yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. Finally, comparative analysis of the available predicted proteomes of FSSC members was performed using OrthoFinder v2.5.2[121]. Orthology inference results showed that 95.2 percent of proteins were assigned into 23190 orthogroups of which 411 were single-copy. Because only 1.8 percent of all orthogroups were single-copy and only 3% of total proteins assigned to them ((411*32)/441738 where 441738 is the total number of proteins of all genomes), single-copy orthogroups were not used for the construction of the phylogenomic tree for FSSC. Instead, the consensus species tree was constructed as follows: An all-vs-all pairwise comparisons was performed using, DIAMOND v2.0.9 to retrieve sequence similarity scores[122]., Based on those scores, proteins were assigned to ortthogroups using Orthofinder assigns and one tree per orthogroup was inferred using DendroBLAST[123]. Moreover, STAG[124] was used to reconcile all orthogroup trees into one consensus species tree that was rooted via STRIDE[125].

Potential accessory FsK scaffold/s identification

To identify how similar or divergent the *F. solani strain K* and *F. vanettenii* genomes are, the command dnadiff from package MUMmer v3.23 was used[126]. *F. vanettenii*'s genome was downloaded from Ensembl Fungi v4.06[127]. Because the average identity of aligned chromosomes/scaffolds between the 2 species was 90% Minimap2 v2.30-r1287 with the parameter -x asm20 was implemented for microsynteny analysis, accounting for divergence as low as 5%. From the resulting data the alignment coverage per scaffold and the total number of alignments (with or without gaps) were computed and displayed, using the package ggplot2 in R v4.3.0. A Circos plot[128], was created using Galactic Circos v 0.69.8 + galaxy9[129]. Only primary alignments (best alignment per query region) with mapping quality ≥ 40 and alignment length ≥ 1,000,000 bp (including gaps) were shown in the plot (Supplementary Fig. S2b).

Comparison of FsK and Fusarium vanettenii genome statistics

To calculate various general statistics between the two Fusarium species, the coding sequences (CDS) and exon coordinates were obtained from the Genomic Feature Files (GFF3) from the genome annotation of each fungus, while the intronic sequence coordinates were obtained using the extract_intron_gff3_from_gff3.py script (https://github.com/irusri/Extract-intron-from-gff3 , visited on May 2020). Exonic and intronic sequences were retrieved from the GFF3 and genomic fasta file of each fungus using getfasta of bedtools v2.29.2[130]. Finally, CDS, intron, exon and amino acid statistics were calculated with the binary script feature_stats_calculator (https://doi.org/10.5281/zenodo.17552527), which was built with Python using the Python libraries Biopython v1.77[131]. In order to detect statistical significant data, the %GC %AT contents as well as the length of scaffolds, coding sequences, exons and introns were computed using fx2tab from SeqKit v2.9.0[132]. Then the statistical significance between the group means were measured using t.test in R v4.3.0. The scaffolds of FsK where scanned for the presence of dsDNA phage, NCLDV, RNA, ssDNA and lavidaviridae viruses with VirSorter2[133].

Predicted cytoplasmic and apoplastic effectome comparison

Predicted proteomes of *Rhizophagus irregularis* DAOM 181602 = DAOM 197198 (GCA_000439145), *Serendipita indica* DSM 11827 (GCA_000313545) and *Fusarium vanettenii* 77–13-4 (GCA_000151355.1) were retrieved from Ensembl Fungiv4.06[127] (Fig. 3a). For the secretome analysis, signalP 5.0b[134] was used to identify putative secreted proteins. The subcellular localization prediction of the putative secreted proteins was made using the command line version of DeepLoc[135]. The extracellular proteins predicted by deeploc were submitted to the BUSCA server[136], using the fungi option. Only the proteins predicted as extracellular by both two programs were selected for further analysis. From the predicted secretome, proteins with length less than 300aa were selected and effectome protein prediction and classification for localization as apoplastic or cytoplasmic effector proteins were made using the EffectorP3.0 online version[137]. Peptidases were determined with the command line version of phmmer from the HMMER3.3 package[138] against merops database[139] pepunit.lib file, with all e-value thresholds of 0.001. Moreover, CAZymes were determined from each proteome, using the HMMER tool in dbCAN2 server[140]. From the final lists of proteins, those that were identified as both CAZymes and proteases were removed.

Identification of strain-specific proteins

To identify the most distant FsK proteins with no common ancestor with *F. vanettenii, R.* irregularis and *S. indica*, a two-step approach was created (Fig. 4). This was made to remove biases created by the fact that the two Fusarium species are phylogenetically very close, while *R. irregularis* and *S. indica* are very distant, as well as biases of specific comparative genomics analysis software.

To identify strain-specific proteins from the comparison of FsK with F. vanettenii, predicted protein FASTA files for each genome were analyzed using OrthoFinder v2.5.2[121] (Supplementary Fig. S3a). Sequence similarity searches were carried out with BLAST + in compressed format, providing both speed and sensitivity[141]. For each orthogroup, OrthoFinder constructed multiple sequence alignments, which were then used to build gene trees with IQ-TREE, a maximum-likelihood phylogenetic inference tool[142]. The OrthoFinder config.json file was further modified to enable filtering of low-complexity regions, reducing spurious matches. To improve alignment quality, a custom pipeline called muscle_and_trim was developed. In this pipeline, sequences were first aligned with MUSCLE using default settings[143], and the resulting multiple sequence alignments were trimmed with trimAl using the -automated1 option, which automatically selects the most suitable trimming strategy[144]. FsK species-specific proteins were retrieved from the orthofinder output and functional annotation was made by transferring the functions of orthologous proteins in the eggNOG v5.0 using the software eggNOG -mapper v2.0, with default options[118]. The final set of FsK-specific proteins contained proteins with no common assigned functions with *F. vanettenii* orthologs and the final set of *F. vanettenii* specific proteins contained proteins that do not have assigned functions by FsK orthologs, respectively (dataset name: "FsK specific vs *F. vanettenii*", Supplementary Fig. S3a). This pipeline was repeated with FASTA files with predicted proteins for all four fungi (Supplementary Fig. S3b). The common proteins of the two FsK-specific genes sets were retrieved and only FsK-proteins belonging in functional categories that don not exist *in F. vanettenii*-specific proteins were selected (Supplementary Fig. S3c). Scripts used for this analysis can be found (https://doi.org/10.5281/zenodo.17552527).

Retrieval of FsK-specific proteins orthologous to the non-pathogenic fungi

To identify FsK proteins with no relation of pathogenicity but with possible orthologous relationship with the other two symbiotic fungi, which could possibly indicate adaptation features common between symbiotic fungi, FASTA files with predicted proteins for all four fungi were imported to OrthoFinder v2.5.2[121]. Protein sequence similarity searches were performed using BLAST + in a compressed format, which provides efficient storage and faster processing while retaining sensitivity[141]. For each orthogroup, OrthoFinder generated multiple sequence alignments, allowing more accurate comparisons of homologous sequences across the two genomes. FsK proteins orthologous with each or both of Rhizophagus or Serindipita and non-homologous with *F. vanettenii* were retrieved from the orthofinder output format (Fig. 4b). The common proteins of FsK-specific proteins from the "Fusarium-only" comparison (Fig. 4a) and each of the 3 above sets (Fig. 4b) were retrieved resulting in the final sets of non-*F. vanettennii* homologous FsK proteins (Fig. 4c). Scripts used for this analysis can be found (https://doi.org/10.5281/zenodo.17552527).

Comparison of biosynthetic cluster genes in all four fungi

Genomic FASTA-formatted sequence and genome feature annotation information files version 3 (gff3) were obtained from EnsemblFungi v4.06 repository[127] and used as input files in antiSMASH v5.1.2[145] with default options, to predict secondary metabolite biosynthetic clusters (Fig. 5a). From this output, GenBank files containing all biosynthetic clusters from each species, were collected and bidirectionally searched in an all vs all comparison, using clinker v0.0.21 with default options[146] to retrieve FsK-specific clusters (Fig. 5a).

Sample preparation, RT-qPCR and statistical analysis of FsK specific biosynthetic cluster genes

L. japonicus (ecotype 'Gifu') seeds were prepared as described above[18]. Six 7-day-old seedlings were transferred to each magenta pot containing sterile sand:vermiculite (3:1) substrate, with five biological replicates per condition. Plants were watered with M medium[147] once at planting and thereafter with water only. Seven-day-old plants were inoculated with FsK by applying 10^3 conidia directly to each root. Control plants received no inoculum. Root samples were collected 4 days post inoculation (dpi). FsK was cultured on Potato Dextrose Agar (PDA) at 25 °C for 5 days. Mycelium was collected with a sterile scalpel and ground using the TissueLyser II system (www.qiagen.com/us) with stainless steel beads. RNA was extracted as described above. FsK colonization in root samples was assessed by qPCR using FsK *ITS* primers (101 bp amplicon), with *L. japonicus ubiquitin* (195 bp

amplicon) as internal control. The DNA-free RNA was diluted to 10 ng/µl and used for RT-qPCR with the Luna® Universal One-Step RT-qPCR Kit (www.neb.com) following the manufacturer's instructions. Total reaction volume was reduced to 10 µl and cycling conditions were as follows: 55 °C for 10 min (reverse transcription), 95 °C for 1 min, followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s, with a melting curve step to confirm product specificity. Expression levels of *NRPS5, NRPS9*, *terpene* and *ferrirhodin synthase* genes were measured. The *β-tubulin (TUB2)* gene (107 bp amplicon) served as the internal control. Primer sequences are listed in Supplementary Table S8. Relative expression was calculated using the ΔCT method[148] and data were visualized in RStudio using the ggplot2 package.

Data availability

The raw sequencing reads used for the genome assembly, along with the draft assembled genome sequence can be accessed through the National Center for Biotechnology Information (NCBI) Bioproject number PRJNA796177. The code used for performing all analysis tasks described in this manuscript can be found at https://doi.org/10.5281/zenodo.17552527.

Received: 27 October 2025; Accepted: 2 April 2026
Published online: 11 April 2026

References

1. Schulz, B., Boyle, C., Draeger, S., Römmert, A.-K. & Krohn, K. Endophytic fungi: A source of novel biologically active secondary metabolites. *Mycol. Res.* **106**, 996–1004. https://doi.org/10.1017/S0953756202006342 (2002).
2. Zobel, M., Koorem, K., Moora, M., Semchenko, M. & Davison, J. Symbiont plasticity as a driver of plant success. *New Phyt.* **241**, 2340–2352. https://doi.org/10.1111/nph.19566 (2024).
3. Hawksworth, D. L., Kirk, P. M., Sutton, B. C. & Pegler, D. N. Ainsworth & Bisby's dictionary of the fungi. *Rev. Inst. Med. Trop. São Paulo* **38**, 272 (1996).
4. Bonfante, P. & Genre, A. Mechanisms underlying beneficial plant–fungus interactions in mycorrhizal symbiosis. *Nat. Commun.* **1**, 48. https://doi.org/10.1038/ncomms1046 (2010).
5. Rodriguez, R. J., White, J. F. Jr., Arnold, A. E. & Redman, R. S. Fungal endophytes: Diversity and functional roles. *New Phyt.* **182**, 314–330. https://doi.org/10.1111/j.1469-8137.2009.02773.x (2009).
6. Domka, A. M., Rozpaądek, P. & Turnau, K. Are fungal endophytes merely mycorrhizal copycats? The role of fungal endophytes in the adaptation of plants to metal toxicity. *Front. Microbiol.* https://doi.org/10.3389/fmicb.2019.00371 (2019).
7. Chitnis, V. R. et al. Fungal endophyte-mediated crop improvement: The way ahead. *Front. Plant Sci.* https://doi.org/10.3389/fpls.2020.561007 (2020).
8. Boedi, S. et al. Comparison of Fusarium graminearum Transcriptomes on Living or Dead Wheat Differentiates Substrate-Responsive and Defense-Responsive Genes. *Front. Microbiol.* https://doi.org/10.3389/fmicb.2016.01113 (2016).
9. Handa, A. K., Bressan, R. A., Handa, S. & Hasegawa, P. M. Characteristics of cultured tomato cells after prolonged exposure to medium containing polyethylene glycol. *Plant Physiol.* **69**, 514–521. https://doi.org/10.1104/pp.69.2.514 (1982).
10. Varma, A., Bakshi, M., Lou, B., Hartmann, A. & Oelmueller, R. *Piriformospora indica*: A novel plant growth-promoting mycorrhizal fungus. *Agric Res.* **1**, 117–131. https://doi.org/10.1007/s40003-012-0019-5 (2012).
11. Sugimura, Y. & Saito, K. Transcriptional profiling of arbuscular mycorrhizal roots exposed to high levels of phosphate reveals the repression of cell cycle-related genes and secreted protein genes in Rhizophagus irregularis. *Mycorrhiza* **27**, 139–146. https://doi.org/10.1007/s00572-016-0735-y (2017).
12. Vincent, D., Rafiqi, M. & Job, D. The Multiple Facets of Plant–Fungal Interactions Revealed Through Plant and Fungal Secretomics. *Front. Plant Sci.* **10**, 1626. https://doi.org/10.3389/fpls.2019.01626 (2020).
13. Schmidt-Dannert, C. in *Biotechnology of Isoprenoids* (eds. Schrader, J. & Bohlmann, J.) 19–61 (Springer International Publishing, 2015). https://doi.org/10.1007/10_2014_283
14. Kavroulakis, N. et al. Role of ethylene in the protection of tomato plants against soil-borne fungal pathogens conferred by an endophytic Fusarium solani strain*. *J. Exp. Bot.* **58**, 3853–3864. https://doi.org/10.1093/jxb/erm230 (2007).
15. Garantonakis, N. et al. Tomato Inoculation With the Endophytic Strain Fusarium solani K Results in Reduced Feeding Damage by the Zoophytophagous Predator Nesidiocoris tenuis. *Front. Ecol. Evol.* https://doi.org/10.3389/fevo.2018.00126 (2018).
16. Pappas, M. L. et al. The Beneficial Endophytic Fungus Fusariumsolani Strain K Alters Tomato Responses Against Spider Mites to the Benefit of the Plant. *Front. Plant Sci.* https://doi.org/10.3389/fpls.2018.01603 (2018).
17. Kavroulakis, N., Doupis, G., Papadakis, I. E., Ehaliotis, C. & Papadopoulou, K. K. Tolerance of tomato plants to water stress is improved by the root endophyte Fusarium solani FsK. *Rhizosphere* **6**, 77–85. https://doi.org/10.1016/j.rhisph.2018.04.003 (2018).
18. Skiada, V. et al. Colonization of legumes by an endophytic Fusarium solani strain FsK reveals common features to symbionts or pathogens. *Fungal Genet. Biol.* **127**, 60–74. https://doi.org/10.1016/j.fgb.2019.03.003 (2019).
19. Handberg, K. & Stougaard, J. Lotus japonicus, an autogamous, diploid legume species for classical and molecular genetics. *Plant J.* **2**, 487–496. https://doi.org/10.1111/j.1365-313X.1992.00487.x (1992).
20. Skiada, V., Avramidou, M., Bonfante, P., Genre, A. & Papadopoulou, K. K. An endophytic Fusarium–legume association is partially dependent on the common symbiotic signalling pathway. *New Phytol.* **226**, 1429–1444. https://doi.org/10.1111/nph.16457 (2020).
21. Avramidou, M. et al. A fungal endophyte increases plant resilience to low nutrient availabilities: A case of Fe acquisition in legumes. *Physiol. Plant.* **176**, e14577. https://doi.org/10.1111/ppl.14577 (2024).
22. Coleman, J. J. et al. The genome of *Nectria haematococca*: Contribution of supernumerary chromosomes to gene expansion. *PLoS Genet.* **5**, e1000618. https://doi.org/10.1371/journal.pgen.1000618 (2009).
23. Coleman, J. J. The *Fusarium solani* species complex: Ubiquitous pathogens of agricultural importance. *Mol. Plant Pathol.* **17**, 146–158. https://doi.org/10.1111/mpp.12289 (2016).
24. O'Donnell, K. et al. No to *Neocosmospora*: Phylogenomic and practical reasons for continued inclusion of the *Fusarium solani* species complex in the genus *Fusarium*. *mSphere* **5**, e00810-20. https://doi.org/10.1128/msphere.00810-20 (2020).
25. Geiser, D. M. et al. One fungus, one name: Defining the genus *Fusarium* in a scientifically robust way that preserves longstanding use. *Phytopathology* **103**, 400–408. https://doi.org/10.1094/PHYTO-07-12-0150-LE (2013).
26. Lombard, L., Van Der Merwe, N. A., Groenewald, J. Z. & Crous, P. W. Generic concepts in Nectriaceae. *Stud. Mycol.* **80**, 189–245. https://doi.org/10.1016/j.simyco.2014.12.002 (2015).
27. Sandoval-Denis, M. & Crous, P. W. Removing chaos from confusion: Assigning names to common human and animal pathogens in *Neocosmospora*. *Persoonia* **41**, 109–129. https://doi.org/10.3767/persoonia.2018.41.06 (2018).
28. Sandoval-Denis, M., Lombard, L. & Crous, P. W. Back to the roots: A reappraisal of *Neocosmospora*. *Persoonia* **43**, 90–185. https://doi.org/10.1016/j.simyco.2014.12.002 (2019).

29. Hoh, D. Z. et al. Fusarium solani species complex genomes reveal bases of compartmentalisation and animal pathogenesis. *Preprint at* https://doi.org/10.1101/2022.03.22.485422 (2022).
30. Navasca, A. et al. Dispensable genome and segmental duplications drive the genome plasticity in *Fusarium solani*. *Front. Fungal Biol.* **6**, 1432339. https://doi.org/10.3389/ffunb.2025.1432339 (2025).
31. Xie, S.-Y. et al. Whole-genome sequencing and comparative genome analysis of *Fusarium solani-melongenae* causing *Fusarium* root and stem rot in sweetpotatoes. *Microbiol. Spectr.* **10**, e00683-22. https://doi.org/10.1128/spectrum.00683-22 (2022).
32. Ceballos, I. et al. The in vitro mass-produced model mycorrhizal fungus, *Rhizophagus irregularis*, significantly increases yields of the globally important food security crop cassava. *PLoS ONE* **8**, e70633. https://doi.org/10.1371/journal.pone.0070633 (2013).
33. Prasad, R., Bagde, U. S., Puspangadan, P. & Varma, A. Bacopa monniera L.: Pharmacological Aspects and Case Study Involving Piriformospora indica. *J. Integr. Biol.* **3**, 2–100 (2018).
34. Saleem, S., Sekara, A. & Pokluda, R. *Serendipita indica*—A review from agricultural point of view. *Plants* **11**, 3417. https://doi.org/10.3390/plants11243417 (2022).
35. Heidarianpour, M. B., Aliasgharzad, N. & Olsson, P. A. Positive effects of co-inoculation with *Rhizophagus irregularis* and *Serendipita indica* on tomato growth under saline conditions, and their individual colonization estimated by signature lipids. *Mycorrhiza* **30**, 455–466. https://doi.org/10.1007/s00572-020-00962-y (2020).
36. Bidellaoui, B., Segarra, G., Hakkou, A. & Isabel Trillas, M. Beneficial effects of *Rhizophagus irregularis* and *Trichoderma asperellum* strain T34 on growth and fusarium wilt in tomato plants. *J. Plant. Pathol.* **101**, 121–127. https://doi.org/10.1007/s42161-018-0159-y (2019).
37. Calvo-Polanco, M., Molina, S., Zamarreño, A. M., García-Mina, J. M. & Aroca, R. The symbiosis with the arbuscular mycorrhizal fungus *Rhizophagus irregularis* drives root water transport in flooded tomato plants. *Plant Cell Physiol.* **55**, 1017–1029. https://doi.org/10.1093/pcp/pcu035 (2014).
38. Xu, Y. et al. Improvement of *Lotus japonicus* hairy root induction and development of a mycorrhizal symbiosis system. *Appl. Plant Sci.* **6**, e1141. https://doi.org/10.1002/aps3.1141 (2018).
39. Banhara, A., Ding, Y., Kühner, R., Zuccaro, A. & Parniske, M. Colonization of root cells and plant growth promotion by *Piriformospora indica* occurs independently of plant common symbiosis genes. *Front. Plant Sci.* **6**, 667. https://doi.org/10.3389/fpls.2015.00667 (2015).
40. Navasca, A. et al. First report and draft genome resource of a unique opportunistic *Fusarium solani* pathogen associated with unknown dark galls in sugarbeet. *PhytoFrontiers™* https://doi.org/10.1094/PHYTOFR-01-23-0006-A (2023).
41. Finn, R. D. et al. Pfam: the protein families database. *Nuc. Acids Res.* **42**, D222–D230. https://doi.org/10.1093/nar/gkt1223 (2014).
42. de Vries, S., de Vries, J., Archibald, J. M. & Slamovits, C. H. Comparative analyses of saprotrophy in *Salisapilia sapeloensis* and diverse plant pathogenic oomycetes reveal lifestyle-specific gene expression. *FEMS Microbiol. Ecol.* **96**, fiaa184. https://doi.org/10.1093/femsec/fiaa184 (2020).
43. Mengistu, A. A. Endophytes: Colonization, behaviour, and their role in defense mechanism. *Int. J. Microbiol.* **2020**, 6927219. https://doi.org/10.1155/2020/6927219 (2020).
44. Hogues, H. et al. Transcription factor substitution during the evolution of fungal ribosome regulation. *Mol. Cell* **29**, 552–562. https://doi.org/10.1016/j.molcel.2008.02.006 (2008).
45. Wang, Q. et al. Exploring plant growth-promoting traits of endophytic fungi isolated from *Ligusticum chuanxiong* Hort and their interaction in plant growth and development. *J. Fungi* **10**, 713. https://doi.org/10.3390/jof10100713 (2024).
46. Adedayo, A. A. & Babalola, O. O. Fungi that promote plant growth in the rhizosphere boost crop growth. *J. Fungi* **9**, 239. https://doi.org/10.3390/jof9020239 (2023).
47. Maher, R. L., Branagan, A. M. & Morrical, S. W. Coordination of DNA replication and recombination activities in the maintenance of genome stability. *J. Cell. Biochem.* **112**, 2672–2682. https://doi.org/10.1002/jcb.23211 (2011).
48. Aylward, J. et al. A plant pathology perspective of fungal genome sequencing. *IMA Fungus* **8**, 1–15. https://doi.org/10.5598/imafungus.2017.08.01.01 (2017).
49. Raeder, U. & Broda, P. Comparison of the lignin-degrading white rot fungi Phanerochaete chrysosporium and Sporotrichum pulverulentum at the DNA level. *Curr. Genet.* **8**, 499–506 (1984).
50. Carreón-Anguiano, K. G., Islas-Flores, I., Vega-Arreguín, J., Sáenz-Carbonell, L. & Canto-Canché, B. EffHunter: A tool for prediction of effector protein candidates in fungal proteomic databases. *Biomolecules* **10**, 712. https://doi.org/10.3390/biom10050712 (2020).
51. Chang, W., Li, H., Chen, H., Qiao, F. & Zeng, H. Identification of mimp-associated effector genes in *Fusarium oxysporum* f. sp. *cubense* race 1 and race 4 and virulence confirmation of a candidate effector gene. *Microbiol. Res.* **232**, 126375. https://doi.org/10.1016/j.micres.2019.126375 (2020).
52. Jia, M. et al. Identification and analysis of the secretome of plant pathogenic fungi reveals lifestyle adaptation. *Front. Microbiol.* **14**, 1171618. https://doi.org/10.3389/fmicb.2023.1171618 (2023).
53. Dalakouras, A. et al. A beneficial fungal root endophyte triggers systemic RNA silencing and DNA methylation of a host reporter gene. *RNA Biol.* **20**, 20–30. https://doi.org/10.1080/15476286.2022.2159158 (2023).
54. Feldman, D., Yarden, O. & Hadar, Y. Seeking the roles for fungal small-secreted proteins in affecting saprophytic lifestyles. *Front. Microbiol.* **11**, 10.3389/fmicb.2020.00455 (2020).
55. Zhao, Z., Liu, H., Wang, C. & Xu, J.-R. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* **14**, 274. https://doi.org/10.1186/1471-2164-14-274 (2013).
56. Tisserant, E. et al. Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20117–20122. https://doi.org/10.1073/pnas.1313452110 (2013).
57. Khalid, M. et al. Mutualistic fungus *Piriformospora indica* modulates cadmium phytoremediation properties of host plant via concerted action of enzymatic and non-enzymatic biochemicals. *Pedosphere* **32**, 256–267. https://doi.org/10.1016/S1002-0160(21)60014-0 (2022).
58. Hallasgo, A. M., Spangl, B., Steinkellner, S. & Hage-Ahmed, K. The fungal endophyte *Serendipita williamsii* does not affect phosphorus status but carbon and nitrogen dynamics in arbuscular mycorrhizal tomato plants. *J. Fungi* **6**, 233. https://doi.org/10.3390/jof6040233 (2020).
59. Alberts, B. et al. in *Molecular Biology of the Cell*. 4th edition (Garland Science, 2002). https://www.ncbi.nlm.nih.gov/books/NBK26928/
60. Fontaine, T. et al. Galactosaminogalactan, a new immunosuppressive polysaccharide of *Aspergillus fumigatus*. *PLoS Pathog.* **7**, e1002372. https://doi.org/10.1371/journal.ppat.1002372 (2011).
61. Gravelat, F. N. et al. Aspergillus galactosaminogalactan mediates adherence to host constituents and conceals hyphal β-glucan from the immune system. *PLoS Pathog.* **9**, e1003575. https://doi.org/10.1371/journal.ppat.1003575 (2013).
62. Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M. & Henrissat, B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6**, 41. https://doi.org/10.1186/1754-6834-6-41 (2013).
63. Jagadeeswaran, G., Veale, L. & Mort, A. J. Do lytic polysaccharide monooxygenases aid in plant pathogenesis and herbivory?. *Trends Plant Sci.* **26**, 142–155. https://doi.org/10.1016/j.tplants.2020.09.013 (2021).
64. Flutto, L. Pectin | properties and determination. In *Encyclopedia of Food Sciences and Nutrition* Second Edition (ed. Caballero, B.) 4440–4449 (Academic Press, 2003). https://doi.org/10.1016/B0-12-227055-X/00901-9.
65. Fry, S. C. The structure and functions of xyloglucan.. *J. Exp. Bot.* **40**, 1–11. https://doi.org/10.1093/jxb/40.1.1 (1989).

66. Kloppholz, S., Kuhn, H. & Requena, N. A secreted fungal effector of *Glomus intraradices* promotes symbiotic biotrophy.. *Curr. Biol.* **21**, 1204–1209. https://doi.org/10.1016/j.cub.2011.06.044 (2011).
67. Park, C.-H. et al. The *Magnaporthe oryzae* effector AvrPiz-t targets the RING E3 ubiquitin ligase APIP6 to suppress pathogen-associated molecular pattern–triggered immunity in rice. *Plant Cell* **24**, 4748–4762. https://doi.org/10.1105/tpc.112.105429 (2012).
68. Plett, J. M. et al. Effector MiSSP7 of the mutualistic fungus *Laccaria bicolor* stabilizes the *Populus* JAZ6 protein and represses jasmonic acid (JA) responsive genes.. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8299–8304. https://doi.org/10.1073/pnas.1322671111 (2014).
69. Fujinaga, M., Cherney, M. M., Oyama, H., Oda, K. & James, M. N. G. The molecular structure and catalytic mechanism of a novel carboxyl peptidase from *Scytalidium lignicolum*.. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 3364–3369. https://doi.org/10.1073/pnas.0400246101 (2004).
70. Krishnan, P., Ma, X., McDonald, B. A. & Brunner, P. C. W. Widespread signatures of selection for secreted peptidases in a fungal plant pathogen.. *BMC Evol. Biol.* **18**, 7. https://doi.org/10.1186/s12862-018-1123-3 (2018).
71. Muszewska, A. et al. Fungal lifestyle reflected in serine protease repertoire.. *Sci. Rep.* **7**, 9147. https://doi.org/10.1038/s41598-017-09644-w (2017).
72. Alkin, N. et al. Proline-specific fungal peptidases: Genomic analysis and identification of secreted DPP4 in alkaliphilic and alkalitolerant fungi.. *J. Fungi* **7**, 744. https://doi.org/10.3390/jof7090744 (2021).
73. Kavi Kishor, P. B., Hima Kumari, P., Sunita, M. S. L. & Sreenivasulu, N. Role of proline in cell wall synthesis and plant development and its implications in plant ontogeny. *Front. Plant Sci.* **6,** https://doi.org/10.3389/fpls.2015.00544 (2015).
74. Zuccaro, A. et al. Endophytic life strategies decoded by genome and transcriptome analyses of the mutualistic root symbiont *Piriformospora indica*.. *PLoS Pathog.* **7**, e1002290. https://doi.org/10.1371/journal.ppat.1002290 (2011).
75. Karim, N. F. A., Mohd, M., Nor, N. M. I. M. & Zakaria, L. Saprophytic and Potentially Pathogenic Fusarium Species from Peat Soil in Perak and Pahang. *Trop. Life Sci. Res.* **27**, 1–20 (2016).
76. Korniłłowicz-Kowalska, T., Wojdyło-Kotwica, B., Bohacz, J. & Możejko, M. Occurrence and distribution of *Fusarium* communities in the root zone in a post-bog permanent meadow in relation to mineral fertilization and growing seasons. *Pathogens* **11**, 341. https://doi.org/10.3390/pathogens11030341 (2022).
77. Steenwyk, J. L. Evolutionary divergence in DNA damage responses among fungi. *MBio* **12**, 10.1128/mbio.03348-20 (2021).
78. Singh, Y., Nair, A. M. & Verma, P. K. Surviving the odds: From perception to survival of fungal phytopathogens under host-generated oxidative burst. *Plant Commun.* **2**, 100142. https://doi.org/10.1016/j.xplc.2021.100142 (2021).
79. Entila, F., Han, X., Mine, A., Schulze-Lefert, P. & Tsuda, K. Commensal lifestyle regulated by a negative feedback loop between Arabidopsis ROS and the bacterial T2SS. *Nat. Commun.* **15**, 456. https://doi.org/10.1038/s41467-024-44724-2 (2024).
80. Berrios, L. & Rentsch, J. D. Linking reactive oxygen species (ROS) to abiotic and biotic feedbacks in plant microbiomes: The dose makes the poison. *Int. J. Mol. Sci.* **23**, 4402. https://doi.org/10.3390/ijms23084402 (2022).
81. Konopka, J. B. N-Acetylglucosamine functions in cell signaling. *Scientifica* **2012**, e489208. https://doi.org/10.6064/2012/489208 (2012).
82. López-Fernández, L. et al. The *Fusarium oxysporum* gnt2, encoding a putative N-Acetylglucosamine transferase, is involved in cell wall architecture and virulence. *PLoS ONE* **8**, e84690. https://doi.org/10.1371/journal.pone.0084690 (2013).
83. Jermy, A. J., Willer, M., Davis, E., Wilkinson, B. M. & Stirling, C. J. The Brl Domain in Sec63p Is Required for Assembly of Functional Endoplasmic Reticulum Translocons *. *J. Biol. Chem.* **281**, 7899–7906 (2006).
84. Jung, S. & Kim, H. Emerging view on the molecular functions of Sec62 and Sec63 in protein translocation. *Int. J. Mol. Sci.* **22**, 12757. https://doi.org/10.3390/ijms222312757 (2021).
85. Liu, T. et al. Unconventionally secreted effectors of two filamentous pathogens target plant salicylate biosynthesis. *Nat. Commun.* **5**, 4686. https://doi.org/10.1038/ncomms5686 (2014).
86. Liu, X. et al. Prp19-associated splicing factor Cwf15 regulates fungal virulence and development in the rice blast fungus. *Environ. Microbiol.* **23**, 10.1111/1462-2920.15616 (2021).
87. Rey, P., Taupin-Broggini, M., Couturier, J., Vignols, F. & Rouhier, N. Is there a role for glutaredoxins and BOLAs in the perception of the cellular iron status in plants?. *Front. Plant Sci.* **10**, 10.3389/fpls.2019.00712 (2019).
88. Attarian, R. et al. The monothiol glutaredoxin Grx4 regulates iron homeostasis and virulence in *Cryptococcus neoformans*. *MBio* **9**, e02377-18. https://doi.org/10.1128/mBio.02377-18 (2018).
89. Jia, L.-J. et al. A linear nonribosomal octapeptide from *Fusarium graminearum* facilitates cell-to-cell invasion of wheat. *Nat. Commun.* **10**, 922. https://doi.org/10.1038/s41467-019-08726-9 (2019).
90. Tang, Z. et al. Biosynthesis of a new Fusaoctaxin virulence factor in *Fusarium graminearum* relies on a distinct path to form a guanidinoacetyl starter unit priming nonribosomal octapeptidyl assembly. *J. Am. Chem. Soc.* **143**, 19719–19730. https://doi.org/10.1021/jacs.1c07770 (2021).
91. Munawar, A., Marshall, J. W., Cox, R. J., Bailey, A. M. & Lazarus, C. M. Isolation and characterisation of a Ferrirhodin synthetase gene from the sugarcane pathogen *Fusarium sacchari*. *ChemBioChem* **14**, 388–394. https://doi.org/10.1002/cbic.201200587 (2013).
92. Aguiar, M. et al. The siderophore transporters Sit1 and Sit2 are essential for utilization of Ferrichrome-, Ferrioxamine- and Coprogen-type siderophores in *Aspergillus fumigatus*. *J. Fungi (Basel)* **7**, 768. https://doi.org/10.3390/jof7090768 (2021).
93. Hider, R. C. & Kong, X. Chemistry and biology of siderophores. *Nat. Prod. Rep.* **27**, 637 (2010).
94. Pao, S. S., Paulsen, I. T. & Saier, M. H. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **62**, 1–34. https://doi.org/10.1128/mmbr.62.1.1-34.1998 (1998).
95. von Döhren, H. A survey of nonribosomal peptide synthetase (NRPS) genes in Aspergillus nidulans. *Fungal Genet. Biol.* **46**, S45–S52. https://doi.org/10.1016/j.fgb.2008.08.008 (2009).
96. Blatzer, M. et al. SidL, an *Aspergillus fumigatus* transacetylase involved in biosynthesis of the siderophores Ferricrocin and Hydroxyferricrocin ∇. *Appl. Environ. Microbiol.* **77**, 4959–4966. https://doi.org/10.1128/AEM.00182-11 (2011).
97. Dzikowska, A. et al. KAEA (SUDPRO), a member of the ubiquitous KEOPS/EKC protein complex, regulates the arginine catabolic pathway and the expression of several other genes in *Aspergillus nidulans*. *Gene* **573**, 10.1016/j.gene.2015.07.066 (2015).
98. Gründlinger, M. et al. Fungal siderophore biosynthesis is partially localized in peroxisomes. *Mol. Microbiol.* **88**, 10.1111/mmi.12225 (2013).
99. Forester, N. T., Lane, G. A., Steringa, M., Lamont, I. L. & Johnson, L. J. Contrasting roles of fungal siderophores in maintaining iron homeostasis in Epichloë festucae. *Fungal Genet. Biol.* **111**, 60–72 (2018).
100. Zeilinger, S. et al. Friends or foes? Emerging insights from fungal interactions with plants. *FEMS Microbiol. Rev.* **40**, 182–207. https://doi.org/10.1093/femsre/fuv045 (2016).
101. Galindo-Solís, J. M. & Fernández, F. J. Endophytic fungal terpenoids: Natural role and bioactivities. *Microorganisms* **10**, 339. https://doi.org/10.3390/microorganisms10020339 (2022).
102. Doyle, J. J. & Doyle, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* (1987). at https://worldveg.tind.io/record/33886
103. Istace, B. et al. De novo assembly and population genomic survey of natural yeast isolates with the Oxford Nanopore MinION sequencer. *Gigascience* **6**, giw018. https://doi.org/10.1093/gigascience/giw018 (2017).
104. Walker, B. J. et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963. https://doi.org/10.1371/journal.pone.0112963 (2014).

105. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 (2014).
106. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes de novo assembler. *Curr. Protoc. Bioinformatics.* **70**, e102. https://doi.org/10.1002/cpbi.102 (2020).
107. Delis, C. et al. Role of lupeol synthase in *Lotus japonicus* nodule formation. *New Phytol.* **189**, 335–346. https://doi.org/10.1111/j.1469-8137.2010.03463.x (2011).
108. Krokida, A. et al. A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. *New Phytol.* **200**, 675–690. https://doi.org/10.1111/nph.12414 (2013).
109. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473–497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x (1962).
110. Bécard, G. & Fortin, J. A. Early events of vesicular–arbuscular mycorrhiza formation on Ri T-DNA transformed roots. *New Phytol.* **108**, 211–218. https://doi.org/10.1111/j.1469-8137.1988.tb03698.x (1988).
111. Anders, S., Pyl, P. T. & Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169. https://doi.org/10.1093/bioinformatics/btu638 (2015).
112. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. https://doi.org/10.1186/s13059-014-0550-8 (2014).
113. Dobin, A. et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21. https://doi.org/10.1093/bioinformatics/bts635 (2013).
114. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652. https://doi.org/10.1038/nbt.1883 (2011).
115. Palmer, J. M. & Stajich, J. Funannotate v1.8.1: Eukaryotic genome annotation. *Zenodo* https://doi.org/10.5281/zenodo.4054262 (2020).
116. Stanke, M. et al. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucl. Acids Res.* **34**, W435–W439. https://doi.org/10.1093/nar/gkl200 (2006).
117. Borodovsky, M. & Lomsadze, A. Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES. *Curr. Protoc. Bioinformatics* **35**, 4.6.1-4.6.10. https://doi.org/10.1002/0471250953.bi0406s35 (2011).
118. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829. https://doi.org/10.1093/molbev/msab293 (2021).
119. Kanehisa, M., Sato, Y. & Morishima, K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* **428**, 726–731. https://doi.org/10.1016/j.jmb.2015.11.006 (2016).
120. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**, 587–592. https://doi.org/10.1093/nar/gkac963 (2023).
121. Emms, D. M. & Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238. https://doi.org/10.1186/s13059-019-1832-y (2019).
122. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60. https://doi.org/10.1038/nmeth.3176 (2015).
123. Kelly, S. & Maini, P. K. DendroBLAST: Approximate phylogenetic trees in the absence of multiple sequence alignments. *PLoS ONE* **8**, e58537. https://doi.org/10.1371/journal.pone.0058537 (2013).
124. Emms, D. M. & Kelly, S. STAG: Species Tree Inference from All Genes. 267914 Preprint at https://doi.org/10.1101/267914 (2018).
125. Emms, D. M. & Kelly, S. STRIDE: Species tree root inference from gene duplication events. *Mol. Biol. Evol.* **34**, 3267–3278. https://doi.org/10.1093/molbev/msx259 (2017).
126. Kurtz, S. et al. Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12. https://doi.org/10.1186/gb-2004-5-2-r12 (2004).
127. Yates, A. D. et al. Ensembl 2020. *Nuc. Acids Res.* **48**, 682–688. https://doi.org/10.1093/nar/gkz966 (2020).
128. Krzywinski, M. et al. Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645. https://doi.org/10.1101/gr.092759.109 (2009).
129. Rasche, H. & Hiltemann, S. Galactic Circos: User-friendly Circos plots within the Galaxy platform. *GigaScience* **9**, giaa065. https://doi.org/10.1093/gigascience/giaa065 (2020).
130. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842. https://doi.org/10.1093/bioinformatics/btq033 (2010).
131. Cock, P. J. A. et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423. https://doi.org/10.1093/bioinformatics/btp163 (2009).
132. Shen, W., Sipos, B. & Zhao, L. SeqKit2: A Swiss army knife for sequence and alignment processing. *iMeta* **3**, e191. https://doi.org/10.1002/imt2.191 (2024).
133. Guo, J. et al. VirSorter2: A multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* **9**, 37 (2021).
134. Almagro Armenteros, J. J. et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423. https://doi.org/10.1038/s41587-019-0036-z (2019).
135. Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H. & Winther, O. DeepLoc: Prediction of protein subcellular localization using deep learning. *Bioinformatics* **33**, 3387–3395. https://doi.org/10.1093/bioinformatics/btx431 (2017).
136. Savojardo, C., Martelli, P. L., Fariselli, P., Profiti, G. & Casadio, R. BUSCA: An integrative web server to predict subcellular localization of proteins. *Nucleic Acids Res.* **46**, 459–466. https://doi.org/10.1093/nar/gky320 (2018).
137. Sperschneider, J. & Dodds, P. N. EffectorP 3.0: Prediction of apoplastic and cytoplasmic effectors in Fungi and Oomycetes. *Mol. Plant-Microbe Interact.* **35**, 146–156. https://doi.org/10.1094/MPMI-08-21-0201-R (2022).
138. Eddy, S. R. & the HMMER development team. (2019). HMMER (Version 3.3). Howard Hughes Medical Institute. http://hmmer.org, https://gensoft.pasteur.fr/docs/hmmer/3.3/Userguide.pdf.
139. Rawlings, N. D., Waller, M., Barrett, A. J. & Bateman, A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **42**, 503–509. https://doi.org/10.1093/nar/gkt953 (2014).
140. Zhang, H. et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nuc. Acids Res.* **46**, 95–101. https://doi.org/10.1093/nar/gky418 (2018).
141. Camacho, C. et al. BLAST+: Architecture and applications. *BMC Bioinformatics* **10**, 421. https://doi.org/10.1186/1471-2105-10-421 (2009).
142. Minh, B. Q. et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534. https://doi.org/10.1093/molbev/msaa015 (2020).
143. Edgar, R. C. 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 (2004).
144. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348 (2009).
145. Blin, K. et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nuc. Acids Res.* **47**, 81–87. https://doi.org/10.1093/nar/gkz310 (2019).

146. Gilchrist, C. L. M. & Chooi, Y.-H. clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics* **37**, 2473–2475. https://doi.org/10.1093/bioinformatics/btab007 (2021).
147. Boisson-Dernier, A. et al. *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol. Plant-Microbe Interact.* **14**, 695–700. https://doi.org/10.1094/MPMI.2001.14.6.695 (2001).
148. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. *Methods* **25**, 402–408. https://doi.org/10.1006/meth.2001.1262 (2001).

Author contributions

O.T, S.V. and KKP conceived the project, O.T. performed the bioinformatics analysis, with supervision by S.V.; O.T. created all figures; A.A., V.S., M.A. and E.D. performed experimental work; K.K.P. supervised the work; O.T., S.V. and K.K.P. wrote the manuscript, with input from all authors.

Funding

OT received support by the Hellenic Foundation for Research and Innovation (HFRI) under the 3rd Call for HFRI PhD Fellowships (Fellowship Number:6236). KKP and SV acknowledge funding from EU's Horizon Europe programme under grant agreement No. 101084163 (RATION).

Declarations

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/s41598-026-47792-0.

Correspondence and requests for materials should be addressed to S.V. or K.K.P.

Reprints and permissions information is available at www.nature.com/reprints.

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