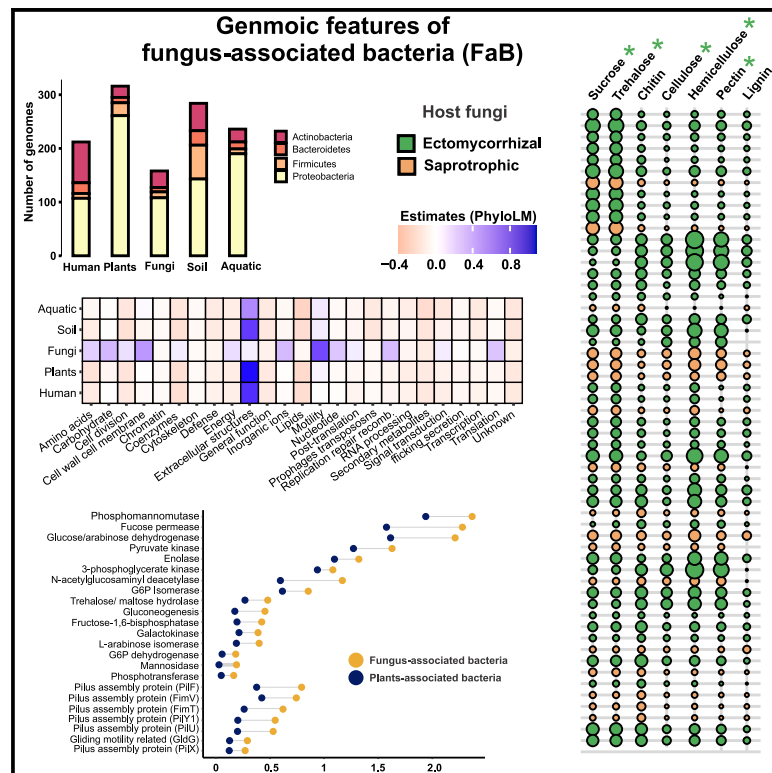


Genomic evidence of symbiotic adaptations in fungus-associated bacteria

Graphical abstract



Authors

Daniyal Gohar, Kadri Põldmaa, Mari Pent, ..., Duncan Y.K. Ng, Falk Hildebrand, Mo Bahram

Correspondence

gohard@oregonstate.edu (D.G.), mohammad.bahram@slu.se (M.B.)

In brief

Genomics; Microbial genomics; Microbial metabolism; Microbiology; Mycology

Highlights

- Fungus-associated bacteria have enriched carbohydrate metabolism and motility genes
- They may aid fungal metabolism through chitin- and trehalose-related genes
- EcM fungal bacteria have more plant- and fungal-substrate-acting genes than SAP
- Bacteria in fungal fruiting bodies lack nitrogen-fixing genes



Article

Genomic evidence of symbiotic adaptations in fungus-associated bacteria

Daniyal Gohar,^{1,2,9,*} Kadri Põldmaa,^{1,3} Mari Pent,¹ Saleh Rahimlou,⁴ Klara Cerk,^{5,6} Duncan Y.K. Ng,⁵ Falk Hildebrand,^{5,6} and Mo Bahram^{7,8,*}

¹Institute of Ecology and Earth Sciences, University of Tartu, J. Liivi St. 2, 50409 Tartu, Estonia

²Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA

³Natural History Museum and Botanical Garden, University of Tartu, Vanemuise 46, 51003 Tartu, Estonia

⁴Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

⁵Gut Microbes & Health, Quadram Institute Bioscience, Norwich Research Park, Norwich, NR4 7UQ Norfolk, UK

⁶Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ Norfolk, UK

⁷Department of Agroecology, Aarhus University, Forsøgsvej 1, 4200 Slagelse, Denmark

⁸Department of Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 756 51 Uppsala, Sweden

⁹Lead contact

*Correspondence: gohard@oregonstate.edu (D.G.), mohammad.bahram@slu.se (M.B.)

<https://doi.org/10.1016/j.isci.2025.112253>

SUMMARY

Fungi harbor diverse bacteria that engage in various relationships. While these relationships potentially influence fungal functioning, their underlying genetic mechanisms remain unexplored. Here, we aimed to elucidate the key genomic features of fungus-associated bacteria (FaB) by comparing 163 FaB genomes to 1,048 bacterial genomes from other hosts and habitats. Our analyses revealed several distinctive genomic features of FaB. We found that FaB are enriched in carbohydrate transport/metabolism- and motility-related genes, suggesting an adaptation for utilizing complex fungal carbon sources. They are also enriched in genes targeting fungal biomass, likely reflecting their role in recycling and rebuilding fungal structures. Additionally, FaB associated with plant-mutualistic fungi possess a wider array of carbon-acquisition enzymes specific to fungal and plant substrates compared to those residing with saprotrophic fungi. These unique genomic features highlight FaB' potential as key players in fungal nutrient acquisition and decomposition, ultimately influencing plant-fungal symbiosis and ecosystem functioning.

INTRODUCTION

Bacteria, a major component of biodiversity on Earth,¹ inhabit diverse environments, including organisms belonging to Eukaryota.^{2,3} Growing evidence suggests that bacteria influence the health of their hosts by promoting immunity, metabolism, and stress tolerance while also acting as pathogens that can negatively affect the host.^{4–6} Genomic studies of host-associated microbes point to their certain evolutionary adaptations. For example, plant symbiotic bacteria have expanded genomes compared to their free-living counterparts.^{7,8} At the same time, obligate symbionts of sap-feeding insects have smaller genomes and consequently encode only essential symbiotic proteins important for their adaptation.⁹ Several studies suggest that symbiotic bacteria may have lost some metabolic genes likely complemented by the host.^{10,11} Furthermore, free-living bacteria may have lost biosynthetic genes and gained genes related to metabolic cross-feeding interactions.¹² Overall, bacterial genomes reflect their adaptation to host and habitat conditions.^{13,14} However, a comprehensive understanding of the genomic features of bacteria across various environmental settings and host-associated habitats remains elusive. Addressing

this gap is crucial to illuminate the complex interplay between bacteria and their hosts, with implications for ecosystem functioning.

Among host-associated bacteria, the genomic features of those inhabiting fungal structures remain largely unexplored. Research has highlighted that fungus-associated bacteria (FaB) play a significant role in enhancing the fitness and resilience of fungi,^{15–17} promoting efficient substrate utilization,^{18,19} enhancing host growth,^{20,21} facilitating mycorrhizae formation,^{22,23} and inducing host sporulation.^{6,24} Moreover, FaB may confer protection against oxidative stress^{25,26} and may also be capable of nitrogen fixation.^{27,28} On the other hand, bacteria can also engage in antagonistic interactions, such as regulating the reproduction of their fungal hosts, with fungi-lacking endosymbiotic bacteria producing more spores than those inhabited by such.²⁹ Furthermore, bacteria can inhibit fungal pathogenesis by producing antifungal compounds such as herbicolin.³⁰ Several studies report that FaB have evolved certain genomic features to establish interactions with their host fungi. Examples include bacteria enhancing the pathogenicity^{31,32} and antibiotic and nutrient uptake ability of host fungi.^{33,34} Comparative phylogenomic studies of the fungi and their endosymbiont bacteria



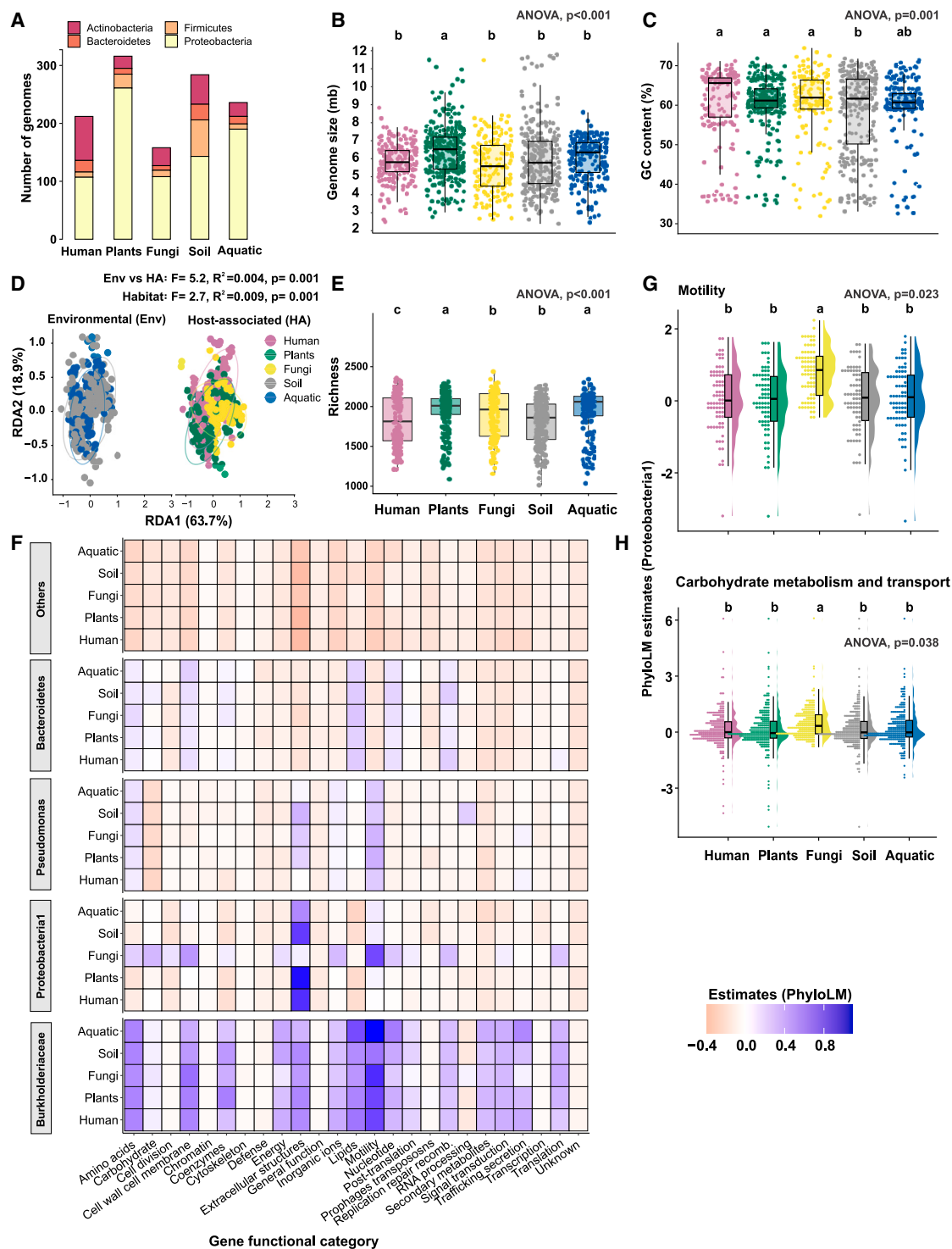


Figure 1. Diversity and enrichment patterns of genes in bacteria from various habitats

(A) The number of genomes analyzed from each habitat.

(B and C) The variations in bacterial genome size and GC content across different habitats.

(D) The composition of functional gene orthologs (COGs) based on the Bray-Curtis distance matrix.

(E) The richness of COGs from each habitat and (F) the differences in gene category enrichments across different habitats. The heatmaps indicate the level of enrichment or depletion, with blue and red cells representing overrepresentation and underrepresentation of gene categories, respectively, based on the

(legend continued on next page)

show that the presence or absence of endosymbionts highly modulates host metabolism.¹¹ However, our understanding of the genomic features of FaB is still limited, particularly regarding their adaptation to their host fungi.

Here, we conducted a comparative genomics analysis on 56 newly sequenced and 107 FaB genomes in the public database, as well as available bacterial genomes from other habitats, including plants ($n = 316$), humans ($n = 212$), soil ($n = 284$), and water ($n = 236$). We hypothesized that FaB have evolved distinct genomic features compared to their counterparts in other habitats, enabling them to effectively colonize and thrive with their fungal host. Additionally, we hypothesized that FaB complement the nutrient acquisition strategies of their hosts, which should be reflected in contrasting genomic features of FaB in ectomycorrhizal ($n = 39$) and saprotrophic fungi ($n = 17$).

RESULTS AND DISCUSSION

Genome size and GC content

To determine whether FaB exhibit expanded or reduced genomes compared to bacteria from other habitats, we compared the genome sizes of FaB against bacteria from plants, humans, soil, and water. Our results demonstrate a significant variation in genome size among the habitats studied, with the largest genomes observed for plant-associated bacteria (Figure 1B; Table S1) corroborating previous findings.^{7,8} The average genome size of FaB is comparable to that of soil bacteria, likely due to the predominant facultative symbiotic nature of FaB originating from soil.³⁵ Although a common assumption is that symbiotic microbes have lost certain gene functions that are complemented by their hosts,^{10,11,33} plant-associated bacteria in our data exhibit larger genomes (Figure 1B), which may contribute to their diverse interactions with other microbes and their host plant.^{36–38} However, our data suggest that genome size reduction prevails in obligate endobacteria compared to environmental bacteria (soil and aquatic) that can also inhabit the surface of plants and fungal hyphae and most likely represent facultative symbionts. This is consistent with previous studies showing that the genomes of endohyphal facultative bacteria are more similar to their terrestrial counterparts, while not exhibiting reduction.³⁹ However, our study revealed a similar distribution of GC content across bacteria inhabiting different hosts (humans, plants, and fungi) that is significantly higher than that observed in environmental bacteria (soil and aquatic; Figures 1C and S2B). Although GC content has been correlated with amino acid composition,⁴⁰ nutrient availability,⁴¹ and growth temperature,⁴² it remains to be determined whether higher GC content in host-associated bacteria is functionally relevant.

Bacterial functional diversity and composition reflect their habitat and host specificity

To compare the functional gene diversity and composition of bacteria across habitats, we examined their Pfam,⁴³ COG,⁴⁴ KEGG,⁴⁵ and Tigrfam⁴⁶ profiles using a redundancy analysis (RDA) of a Bray-Curtis distance matrix generated based on orthologous gene counts. The RDA plots reveal some degree of overlap in the clustering, which may reflect the complex and variable distribution of functional genes across the different habitats. Nevertheless, the statistical models indicate notable distinctions between the groups (Figures 1D and S3). These results suggest that the functional gene composition of bacteria can be explained by their habitat type ($R^2 = 0.07$, $p = 0.001$) as well as host type ($R^2 = 0.06$, $p = 0.001$) (Figure S3). The diversity of functional genes also significantly varied across habitats (Figures 1E and S4). The predominant role of bacterial habitat type in defining their functional gene diversity and composition may reflect the adaptive strategies of bacteria in response to the unique environmental conditions of their habitats.⁴⁷ For example, fungi produce diverse secondary metabolites, which may impact the genomic and functional competency of bacteria striving to establish a symbiotic interaction with these fungi.^{48–50} Similarly, bacterial communities may respond to habitat-specific conditions, such as salinity, by activating or deactivating certain metabolic pathways,^{51,52} or they may be capable of transitioning between free-living and host-associated lifestyles, thereby evolving adaptive traits over time.^{53–55} Our results suggest that bacterial functional diversity and composition reflect their habitat and host specificity.

FaB are enriched in carbohydrate metabolism and cell motility genes

To test the hypothesis that FaB may have evolved specific genomic adaptations to colonize and thrive in their fungal habitat effectively, we performed the PhyloLM test.⁵⁶ We evaluated gene enrichment patterns in FaB genomes compared to genomes of bacteria from other habitats. PhyloLM accounts for phylogenetic relationships, thereby eliminating enrichments that may result from shared ancestry, resulting in an accurate assessment of gene enrichments in the given groups.^{8,56} Although we accounted for phylogeny in our analysis, a key limitation was the unequal representation of certain taxa from different habitats, due to the exclusion of genomes for which required data were missing. For instance, Pseudomonadales were more prevalent in aquatic habitats due to a higher number of sequenced genomes from these environments. While our primary goal was to compare FaB with bacteria from other habitats, we included a maximum number of high-quality genomes of similar taxa sequenced from different habitats. Nonetheless, based on our data, the following gene categories were consistently enriched regardless of bacterial phylogeny and habitat

PhyloLM estimates. The full COG category names for the x axis labels are in Table S2. Among the eight taxa, Proteobacteria 2, Bacillales, Alphaproteobacteria, and Actinobacteria were grouped under "Others" to improve the clarity of the figure.

(G and H) The distribution of motility and carbohydrate metabolism and transport-related genes and their enrichment estimates are represented in alphaproteobacteria1, as shown in (F). Different letters above each boxplot indicate significant differences between groups. Similar letters show similarity between the compared groups, whereas different letters denote a statistically significant difference. ANOVA followed by *post hoc* tests (Tukey's HSD) was used to test for differences between group pairs. All *p* values were corrected using the FDR method to control the false discovery rate.

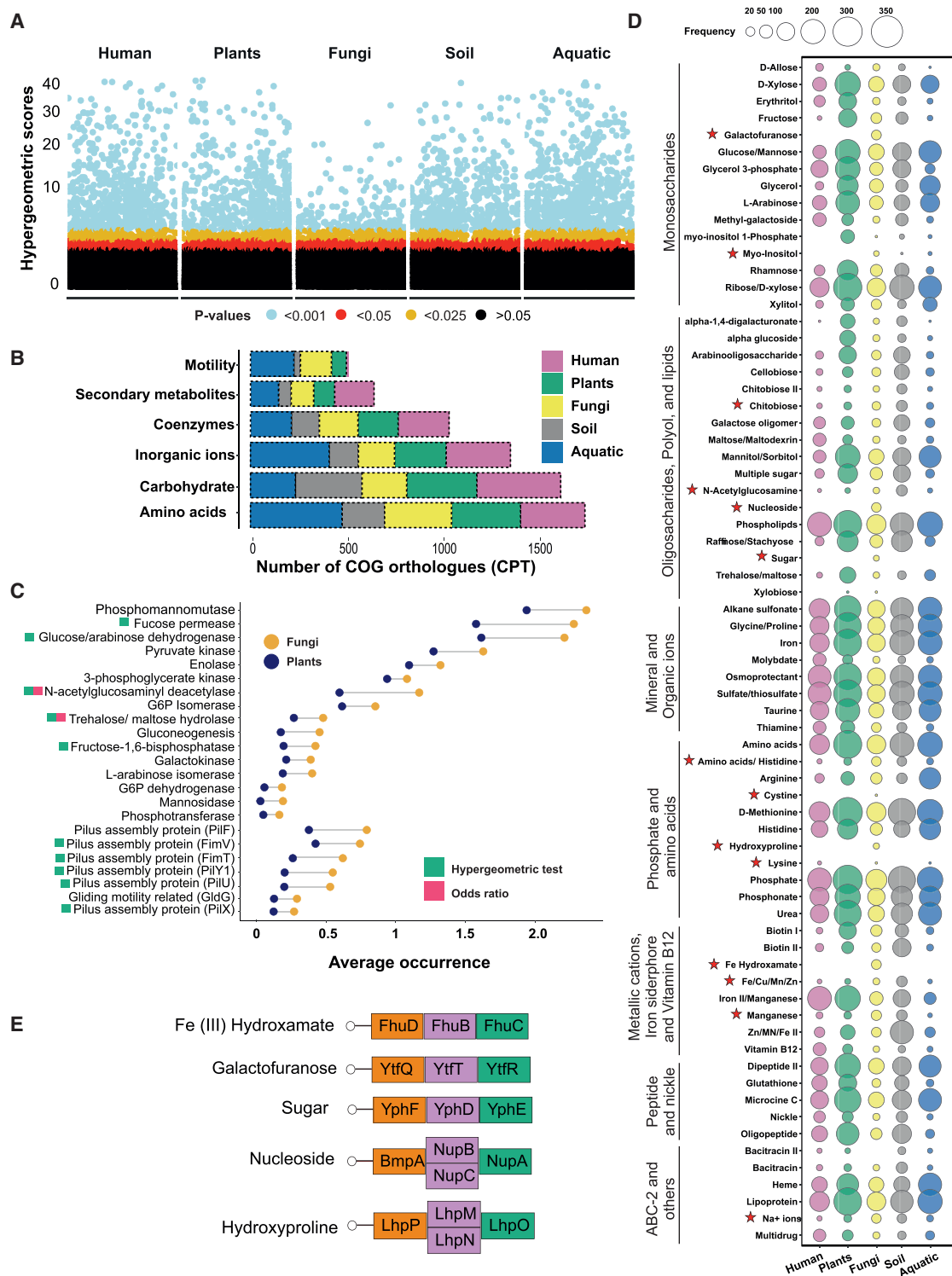


Figure 2. Comprehensive analysis of FaB genes

(A) The analysis of association of gene orthologs (COGs) with their respective habitat was performed using hypergeometric tests. The y axis represents the hypergeometric scores, while the color-coded points signify *p* values indicating a significant gene association with a particular habitat.

(B) The functional categories of habitat-associated genes and the number of genes (COGs) significantly associated with a specific habitat from each functional category. We transformed absolute gene numbers into the gene counts per thousand genomes (CPTs) to account for the varying number of genomes in each

(legend continued on next page)

type: amino acids, lipids, replication and recombination, secondary metabolites, and signal transduction and translation (Figure 1F). Compared to other habitats, the genes related to carbohydrate transport and metabolism, cell wall and cell membrane, and cell motility were significantly more enriched in the genomes of a Proteobacteria cluster (hereafter Proteobacteria1) represented among FaB (Figures 1G and 1H). These genes may enable FaB to efficiently colonize surfaces of fungal cells that receive carbon compounds through symbiotic association with plants or by saprotrophic means and to travel along the hyphae. Moreover, bacteria may employ these genes to locate fungal surfaces and respond to fungal exudates in their surrounding environment. These results provide evidence for the importance of carbohydrate transport and metabolism genes in facilitating bacterial colonization of FaB.

To further investigate the genes involved in bacterial adaptation to various habitats, we used a hypergeometric test that considers factors such as sample size and gene frequency to identify genes associated with each habitat (Figure 2B). Despite observing some overlap in gene associations among habitat types (Figure S5), the principal coordinate analysis revealed distinct clusters that separated FaB genes from those originating from other habitats, supporting our results from PhyloLM tests (Figure S6). Among broad gene categories, FaB genes mainly belonged to *Unknown function* or *General function*, while *Carbohydrate metabolism and transport* and *Motility-related genes* were also observed (Table S3). To further validate these results, we compared overrepresentation patterns of all genes related to *carbohydrate metabolism* and *cell motility* (Figure 1F). We observed 23 genes significantly overrepresented (Figure 2C) in fungi but depleted in other habitats. Among these, 10 genes were identified as FaB-specific genes, as determined by either hypergeometric or odds ratio tests, including various genes for carbohydrate metabolism and bacterial pilus assembly required for motility. These findings provide evidence for the importance of carbohydrate metabolism and bacterial pilus assembly-related genes for bacterial adaptation to fungal habitats (the complete list of FaB-specific genes is provided in Table S3).

Contribution of FaB to fungal metabolism

Our analysis did not uncover the evidence of nitrogen fixation (nifHDK genes) by FaB, despite previous speculations about their potential nitrogen-fixing abilities.^{28,57,58} However, we found that FaB can potentially influence fungal growth and biomass

production, which is exerted through their impact on the biosynthesis and degradation of trehalose and chitin within the fungal cell wall. This is evidenced by the overrepresentation and significant association of genes coding for enzymes related to trehalose and chitin metabolism with FaB (FDR corrected $p < 0.05$; Figures 2C and 3). Chitin and trehalose are pivotal constituents of fungal biomass, wherein the fungal cell wall is predominantly made up of chitin, an N-acetylglucosamine-based polymer, and trehalose, a disaccharide, serves as a carbohydrate repository that has structural and functional roles.^{59,60} Compared to plant-associated bacteria, the higher abundance of N-acetylglucosaminyl deacetylase (NAGd) genes (Figures 2A and 2C; FDR corrected $p < 0.01$; Figure 3; Table S4), responsible for chitin degradation,⁶¹ in FaB may reflect the fungal need for chitin-active enzymes during developmental processes such as mycorrhizae or fruiting body formation.^{62,63} Hydrolysis of chitin releases monomers or oligomers that can be used as carbon and nitrogen sources by both bacteria and fungi.⁶⁴ Alternatively, chitin-degrading capabilities may enable bacteria to target or degrade their fungal hosts. The presence of NAGd enzymes in bacteria could also enhance ability of fungi to scavenge nitrogen from organic matter, a key nitrogen source for mycorrhizal fungi.^{65–67} Recent research has shown that a significant portion of nitrogen, which would otherwise be unavailable, can be utilized by mycorrhizal fungi in the presence of *Paenibacillus* sp.⁶⁸ These observations suggest that, under certain conditions, FaB may enhance the nutrition of their host fungi. However, such interactions would require further investigation.

Through their role in trehalose metabolism (Figures 2A, 2C, and 3), FaB may boost their hosts' stress tolerance, growth and spore germination. Trehalose is synthesized and stored in fungal hyphae^{69,70} and is utilized as an energy source for spore germination and fruiting body development.^{71–73} Trehalose has been shown to accumulate in the cell wall during certain stages of fungal growth. It may affect the structural properties of the cell wall, potentially contributing to the strength and resilience of the fungal hyphae.^{59,72–75} Furthermore, trehalose secreted by plant-associated microorganisms can act as a sucrose analog, influencing host growth by regulating carbon allocation and potentially influencing assimilate allocation in surrounding cells or tissues.^{76,77} Our analysis of carbohydrate-active enzyme profiles revealed that FaB possess a set of CAZymes, namely GT4, GT20, and GH37, necessary for trehalose synthesis and hydrolysis⁷⁸ (Figure 3; Table S5). These findings suggest that these enzymes may be involved in trehalose metabolism, potentially

habitat. For better visualization, only the most abundant gene functional categories are illustrated (general function and unknown function were omitted). However, a list of the genes associated with habitats can be found in Table S3.

(C) Significantly overrepresented genes (FDR-corrected p value < 0.05 ; Table S4) in FaB compared to plant-associated bacteria. The x axis represents the average occurrence of each gene on the y axis in fungi- and plant-associated bacteria. The color codes on the y-axis labels represent the statistical validation of the association of genes with FaB. To simplify the illustration, we only included genome data of bacteria from plants and fungi, since both of these host-associated habitats occupy a similar niche, in contrast to the habitats in humans.

(D) Comparison of ABC transporter pathways in bacterial genomes across different habitats. Only the complete pathways (with all genes present in a genome) were counted to determine the frequency of complete pathways in genomes from each habitat. The list of genes involved in each pathway is provided in Table S6. The red stars with the pathway name represent where FaB distinguishes itself from other habitats or has more pathway copy numbers than plant-associated bacteria. Frequency of pathway copy number was compared using chi-squared tests followed by p value correction using the FDR method to control for false discovery rate.

(E) Illustrate the list of the pathways and the genes involved in each pathway that were found to be specific to FaB, compared to bacteria from other habitats.

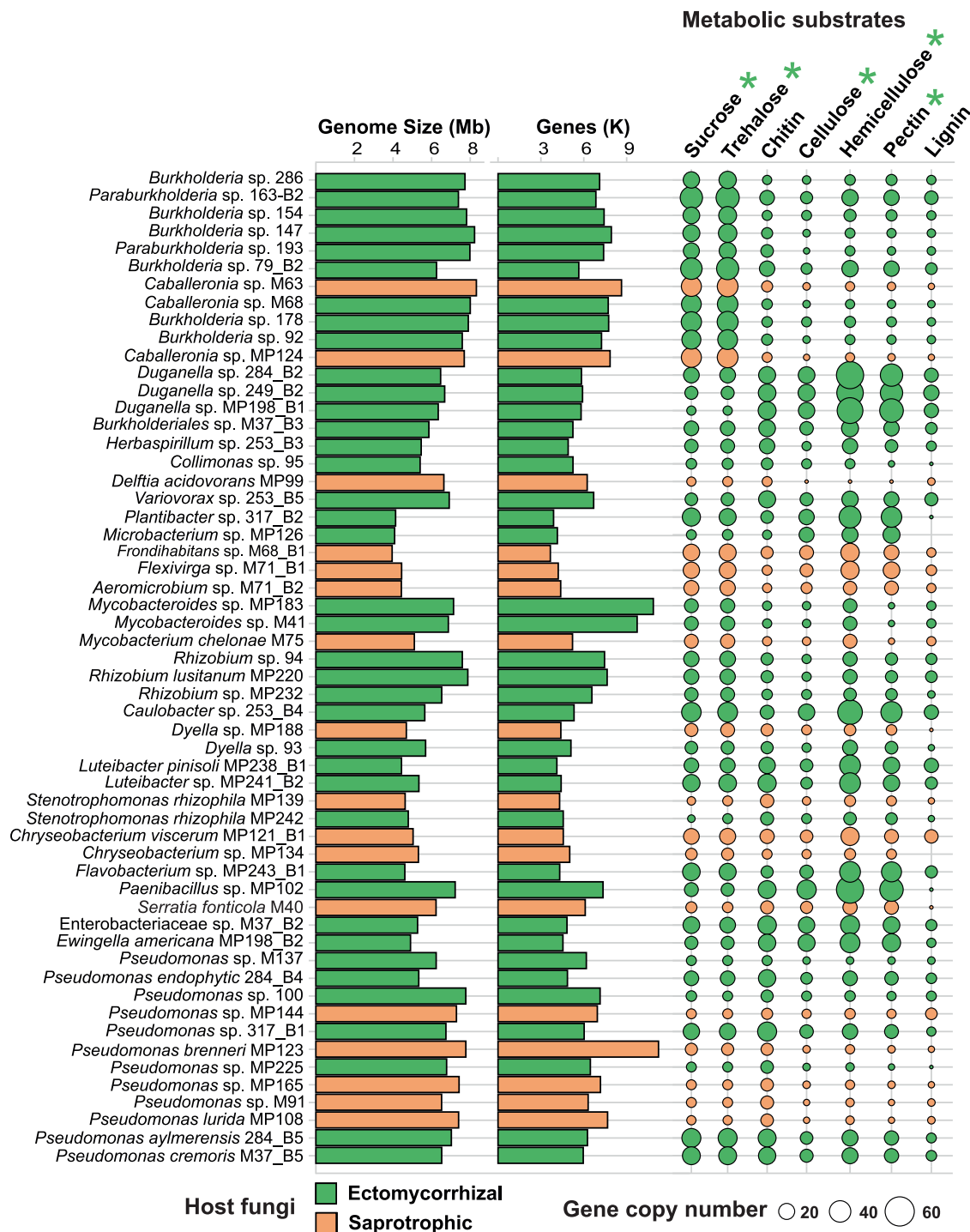


Figure 3. FaB CAZyme repertoires acting on fungal and plant-derived substrates

The number of metabolic enzymes acting on fungal and plant substrates is represented by the point size. The different lifestyles of host fungi (e.g., ectomycorrhizal and saprotrophic) are color-coded. A complete list of CAZyme families involved in the metabolism of each substrate can be found in [Table S5](#). A star symbol indicates that a metabolic substrate is significantly more abundant in ectomycorrhizal than saprotrophic fungi ($p < 0.03$, FDR corrected).

contributing to the growth, stress tolerance, and biomass production of fungi. The presence of GT4, GT20, and GH37 implies that FaB possesses the necessary enzymatic machinery to regu-

late trehalose levels, a process that may be critical for its metabolic flexibility, energy storage, and dynamic response to environmental stress.

Other genes that were overrepresented in FaB encode enzymes involved in the metabolism of oligo- and polysaccharides (Figure 2C; FDR-corrected $p < 0.05$; Table S4), including galactose (Galactokinase; EC 2.7.1.6) and glucose (glucose dehydrogenase), as well as of transporters and hydrolases (Table S4). Furthermore, seven of these genes encode enzymes fundamental to glycolysis in bacteria, including enolase (EC 4.2.1.11), glucose dehydrogenase (EC 1.1.1.47), pyruvate kinase (EC 2.7.1.40), and phosphoglycerate kinase (EC 2.7.2.3). The abundance of simple sugars like glucose in fungal habitats suggests that glycolysis plays a crucial role in FaB metabolism, allowing for direct energy production from readily available nutrients.

The high dispersal ability of FaB

Pilus assembly is critical for bacterial adhesion to surfaces and host-bacterial interactions.^{79,80} This mechanism is a significant determinant of bacterial virulence, movement along surfaces, and colonization in host organisms.^{81–83} Enrichment of genes related to motility, such as pilus assembly in FaB (Figures 1G and 2C), suggests that these features may allow bacteria to reach areas that would otherwise be inaccessible and to travel along the surface of fungal hyphae. Fungal hyphae can create a relatively more aqueous environment due to water films around and within the hyphal network^{84,85} compared to the surrounding environment. Bacteria potentially move more easily on hyphae due to the water films, which reduce friction and provide a pathway for bacterial movement.^{81,82,86} This is in line with a previous study showing that mycorrhizal networks facilitate the colonization of legume roots by symbiotic nitrogen-fixing bacteria.⁸⁷ Our results provide genomic evidence that cell mobility-related genes may play a crucial role in bacterial adaptation to fungi and that the dominant mode of bacterial movement on fungal highways is facilitated through pili.

Nutrient acquisition capability of FaB

We further found that FaB possess diverse nutrient acquisition capabilities as reflected by the presence of an array of ATP-binding cassette (ABC) transporter pathways (Figure 2D; Table S6), which are responsible for the uptake and transport of carbohydrates, minerals, amino acids, and vitamins. In particular, compared to bacteria associated with other environments, FaB exhibited a significantly higher abundance of transporter pathways for certain essential carbohydrates, amino acids and ions, such as galactofuranose and Fe hydroxamate (FDR corrected, chi-squared $p < 0.05$; Figures 2D and 2E), which suggests that FaB possess a distinctive ability to obtain nutrients from its surroundings. Furthermore, FaB likely use galactofuranose, a key component of fungal hyphae,⁸⁸ as a means of lipopolysaccharide communication,^{89,90} to recognize, infect, and colonize their fungal host, as reflected in the exclusive presence of pathways encoding proteins responsible for galactofuranose uptake in FaB (Figure 2D; Table S6). Similar interactions involving bacterial attachment to fungal hyphae through lipopolysaccharides have been reported for *Agaricus bisporus*⁹¹ and arbuscular mycorrhizal fungi,⁹² and the use of polysaccharides by bacteria has been shown to enhance their ability to evade recognition by the host immune system.⁸⁸ Furthermore, the transporters may

facilitate the exchange of nutrients between partners and optimize resource utilization in symbiosis. For instance, bacterial Fe (III) hydroxamates, iron-chelating compounds,⁹³ may contribute to iron transport by fungi to their plant hosts.^{94–96}

FaB may complement the functions of their hosts

Our analysis further revealed contrasting genomic features of FaB depending on the nutritional modes of their host fungi (Figure 3), suggesting a potential role of FaB in the functioning of their host fungi. Specifically, we found that metabolic enzymes acting on certain plant substrates, including sucrose, cellulose, and hemicellulose, were significantly more abundant in ectomycorrhizal fungi-associated bacteria than in bacteria associated with saprotrophic fungi (Figure 3; Table S5). This contrasts comparative genomics studies of fungi reporting fewer CAZymes for polysaccharides in ectomycorrhizal fungal genomes compared to saprotrophic fungi.⁹⁷ Ectomycorrhizal fungi are usually adapted to environments with low litter quality and thus need extensive enzymes for obtaining nutrients from organic matter.^{98,99} It is, therefore, tempting to speculate that a greater C-degradation potential of bacteria may allow ectomycorrhizal fungi to develop diverse nutrient acquisition strategies when the supply from plant hosts suffers.

To investigate the genomic potential of FaB to degrade sucrose, which may play a role in fungal-plant interactions, we performed an *in vitro* experiment. Specifically, we cultivated three analyzed bacterial strains, irrespective of the guild of their associated fungi, with genes for GH32 enzymes (encoding for Invertase: EC 3.2.1.26) on media containing sucrose as the carbon source. We focused on sucrose as it is the primary source of C for mycorrhizal fungi in the apoplast of the root-fungal interface.^{100–102} Our results confirmed the capacity of tested FaB strains (in two out of three) to utilize sucrose as a carbon source (Figure S8). Unlike saprotrophic fungi, ectomycorrhizal fungi lack sucrose invertases^{103–105} and rely on plant cell wall-bound invertases to obtain digestible monosaccharides.^{101,106} By regulating the activity of sucrose invertases, plants can modulate the allocation of sugars for fungi, influencing fungal growth.^{100,107–109} However, our analysis of CAZyme repertoires (Figure 3) suggests that most FaB have the potential to play a key role in the inversion of sucrose, augmenting host functions and contributing to the gradient of sucrose concentration at the root-fungal interface. Previous studies have reported the observation of similar complementary behavior of bacteria in tripartite interactions of plants, fungi, and bacteria.^{15,110} For example, FaB enhance their host fungi's responsiveness to strigolactones exuded by roots, leading to increased hyphal elongation and branching relevant to mycorrhizal establishment.^{21,111} Our findings, therefore, underscore the potential of FaB in metabolizing sucrose and support that FaB symbiotically complement their fungal hosts in the degradation and utilization of carbohydrates.

Conclusions

This study identified distinctive genomic features that may enhance the colonization of fungi by bacteria and uncovered the potential key role of carbohydrate metabolism and cell mobility in bacterial-fungal interactions. Moreover, our findings suggest that FaB have evolved more specialized adaptations

for interacting with mycorrhizal fungi than saprotrophic fungi and that FaB may exploit fungal hyphae for nutrient acquisition and interacting with plants. These findings underscore the importance of genomic adaptations in facilitating bacterial-fungal interactions and point to the important role of FaB in enhancing ecosystem functioning through their support to host fungi.

Limitations of the study

The shortcomings of this study stem from the number of available FaB genomes and the limited range of environments analyzed. Another limitation was the incomplete metadata for genomes retrieved from public databases, particularly the lack of host taxonomy and geographical information. However, for the analysis of CAZyme associated with ectomycorrhizal-versus saprotrophic fungi, we focused on the newly generated genomes, which helped to reduce biases that could be caused by geographical variation and ensured that genomes with only well-characterized host associations were included. Finally, while this study provides evidence for the distinct genomic potential of FaB, their realized functions remain to be examined in future studies.

RESOURCE AVAILABILITY

Lead contact

Any further information or requests for resources and methods should be directed to the lead contact, Daniyal Gohar (gohard@oregonstate.edu).

Materials availability

The isolates analyzed in this study are preserved in the microbiological collection of the Natural History Museum and the Collection of Environmental and Laboratory Microbial Strains (CELMS) at the University of Tartu. The isolates can be made available upon request to the collections' curators.

Data and code availability

- The data supporting the results of this study are available in the public database JGI/IMG, with accession numbers provided in Table S1. Alternatively, the data can be accessed as separate datasets at <https://doi.org/10.5281/zenodo.10889199>.

Dataset 1: Nucleotide FASTA files (assembled).

Dataset 2: Annotations.

Dataset 3: CAZyme annotations.

Dataset 4: Metadata.

- The script used for this study's bioinformatics and statistical analyses can be accessed at <https://github.com/daniyalgohar/Genomic-features-of-FaB>.

ACKNOWLEDGMENTS

We thank Jessie Uehling and Age Brauer for their constructive comments on an earlier version of the manuscript. M.B. was supported by the Swedish Research Council Vetenskapsrådet (grant 2021–03724) and the Novo Nordisk Foundation (NNF). F.H. was supported by the European Research Council H2020 StG (erc-stg-948219, EPYC), and D.Y.K.N. and F.H. by H2020-EU.3.2.2.3. (grant no. 863059 – www.fns-cloud.eu). K.C., D.Y.K.N., and F.H. were supported by the Biotechnology and Biological Sciences Research Council Institute Strategic Program Food Microbiome and Health BB/X011054/1 and its constituent project BBS/E/F/000PR13631 as well as Decoding Biodiversity BB/X011089/1 and its constituent project(s) BBS/E/T/000PR13638.

This work was supported by the Estonian Research Council grant PRG1170 and the European Union through the European Regional Development Fund (the Centers of Excellence: EcolChange and AgroCropFuture), Swedish

Research Council Vetenskapsrådet (grant 2021–03724), and Formas (grant 2020–00807).

AUTHOR CONTRIBUTIONS

D.G., K.P., and M.B. set up and designed the study. D.G. performed the most of the data analysis and wrote the manuscript. M.P. conducted bacterial culturing, DNA extraction, and library preparation for sequencing. S.R. contributed to quality filtering and genome assembly. K.C., D.Y.K.N., and F.H. helped with bioinformatics and statistical analysis. M.B. and K.P. supervised the study. D.G. wrote the first version of the manuscript with input from M.B. All authors contributed to the manuscript revision and approved the submitted version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **METHOD DETAILS**
 - Study sites and fungal sampling
 - Culturing, DNA extraction and library preparation
 - Sequencing
 - Genome data selection and filtration
 - Construction of the phylogenetic tree
 - Clustering of 1211 genomes into eight clusters
 - Gene category enrichment analysis
 - ABC transporters
 - Genome-wide analysis of CAZyme domains
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2025.112253>.

Received: May 8, 2024

Revised: July 18, 2024

Accepted: March 17, 2025

Published: March 20, 2025

REFERENCES

1. Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., Harend, H., et al. (2018). Structure and function of the global topsoil microbiome. *Nature* 560, 233–237. <https://doi.org/10.1038/s41586-018-0386-6>.
2. Emmett, B.D., Lévesque-Tremblay, V., and Harrison, M.J. (2021). Conserved and reproducible bacterial communities associate with extraradical hyphae of arbuscular mycorrhizal fungi. *ISME J.* 15, 2276–2288. <https://doi.org/10.1038/s41396-021-00920-2>.
3. Pierce, E.C., Morin, M., Little, J.C., Liu, R.B., Tannous, J., Keller, N.P., Poglian, K., Wolfe, B.E., Sanchez, L.M., and Dutton, R.J. (2020). Bacterial-fungal interactions revealed by genome-wide analysis of bacterial mutant fitness. *Nat. Microbiol.* 6, 87–102. <https://doi.org/10.3410/f.738956054.793583962>.
4. Tsukamoto, T., Murata, H., and Shirata, A. (2002). Identification of Non-Pseudomonad Bacteria from Fruit Bodies of Wild Agaricales Fungi That Detoxify Tolaasin Produced by *Pseudomonas tolaasii*. *Biosci. Biotechnol. Biochem.* 66, 2201–2208. <https://doi.org/10.1271/bbb.66.2201>.

5. Bahram, M., and Netherway, T. (2022). Fungi as mediators linking organisms and ecosystems. *FEMS Microbiol. Rev.* 46, fuab058. <https://doi.org/10.1093/femsre/fuab058>.
6. Frey-Klett, P., Garbaye, J., and Tarkka, M. (2007). The mycorrhiza helper bacteria revisited. *New Phytol.* 176, 22–36. <https://doi.org/10.1111/j.1469-8137.2007.02191.x>.
7. Rahimlou, S., Bahram, M., and Tedersoo, L. (2021). Phylogenomics reveals the evolution of root nodulating alpha- and beta-Proteobacteria (rhizobia). *Microbiol. Res.* 250, 126788. <https://doi.org/10.1016/j.micres.2021.126788>.
8. Levy, A., Salas Gonzalez, I., Mittelviehhaus, M., Clingenpeel, S., Herrera Paredes, S., Miao, J., Wang, K., Devescovi, G., Stillman, K., Monteiro, F., et al. (2017). Genomic features of bacterial adaptation to plants. *Nat. Genet.* 50, 138–150. <https://doi.org/10.1038/s41588-017-0012-9>.
9. McCutcheon, J.P., and Moran, N.A. (2011). Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* 10, 13–26. <https://doi.org/10.1038/nrmicro2670>.
10. Ullah, A., Bano, A., and Janjua, H.T. (2020). Microbial Secondary Metabolites and Defense of Plant Stress. *Microbial Services in Restoration Ecology*, 37–46. <https://doi.org/10.1016/b978-0-12-819978-7.00003-8>.
11. Uehling, J., Gryganskyi, A., Hameed, K., Tschaplinski, T., Misztal, P.K., Wu, S., Desirò, A., Vande Pol, N., Du, Z., Zienkiewicz, A., et al. (2017). Comparative genomics of *Mortierella elongata* and its bacterial endosymbiont *Mycosporium cysteinexigens*. *Environ. Microbiol.* 19, 2964–2983. <https://doi.org/10.1111/1462-2920.13669>.
12. D'Souza, G., and Kost, C. (2016). Experimental Evolution of Metabolic Dependency in Bacteria. *PLoS Genet.* 12, e1006364. <https://doi.org/10.1371/journal.pgen.1006364>.
13. Holt, R.D. (2009). Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proc. Natl. Acad. Sci. USA* 106, 19659–19665. <https://doi.org/10.1073/pnas.0905137106>.
14. Sheppard, S.K., Guttman, D.S., and Fitzgerald, J.R. (2018). Population genomics of bacterial host adaptation. *Nat. Rev. Genet.* 19, 549–565. <https://doi.org/10.1038/s41576-018-0032-z>.
15. Partida-Martinez, L.P., de Loos, C.F., Ishida, K., Ishida, M., Roth, M., Buder, K., and Hertweck, C. (2007). Rhizonin, the first mycotoxin isolated from the zygomycota, is not a fungal metabolite but is produced by bacterial endosymbionts. *Appl. Environ. Microbiol.* 73, 793–797. <https://doi.org/10.1128/aem.01784-06>.
16. Pakvaz, S., and Soltani, J. (2016). Endohyphal bacteria from fungal endophytes of the Mediterranean cypress (*Cupressus sempervirens*) exhibit in vitro bioactivity. *For. Pathol.* 46, 569–581. <https://doi.org/10.1111/efp.12274>.
17. Hoffman, M.T., Gunatilaka, M.K., Wijeratne, K., Gunatilaka, L., and Arnold, A.E. (2013). Endohyphal Bacterium Enhances Production of Indole-3-Acetic Acid by a Foliar Fungal Endophyte. *PLoS One* 8, e73132. <https://doi.org/10.1371/journal.pone.0073132>.
18. Pavić, A., Stanković, S., Saljnikov, E., Krüger, D., Buscot, F., Tarkka, M., and Marjanović, Ž. (2013). Actinobacteria may influence white truffle (*Tuber magnatum* Pico) nutrition, ascocarp degradation and interactions with other soil fungi. *Fungal Ecol.* 6, 527–538. <https://doi.org/10.1016/j.funeco.2013.05.006>.
19. Shaffer, J.P., U'Ren, J.M., Gallery, R.E., Baltrus, D.A., and Arnold, A.E. (2017). An endohyphal bacterium (Chitinophaga, Bacteroidetes) alters carbon source use by *Fusarium keratoplasticum* (F. solani species complex, Nectriaceae). *Front. Microbiol.* 8, 350. <https://doi.org/10.3389/fmicb.2017.00350>.
20. Lackner, G., Moebius, N., and Hertweck, C. (2010). Endofungal bacterium controls its host by an hrp type III secretion system. *ISME J.* 5, 252–261. <https://doi.org/10.1038/ismej.2010.126>.
21. Lumini, E., Bianciotto, V., Jargeat, P., Novero, M., Salvioli, A., Faccio, A., Bécard, G., and Bonfante, P. (2007). Presymbiotic growth and sporul morphology are affected in the arbuscular mycorrhizal fungus *Gigaspora margarita* cured of its endobacteria. *Cell. Microbiol.* 9, 1716–1729. <https://doi.org/10.1111/j.1462-5822.2007.00907.x>.
22. Poole, E.J., Bending, G.D., Whipps, J.M., and Read, D.J. (2001). Bacteria associated with *Pinus sylvestris*-*Lactarius rufus* ectomycorrhizas and their effects on mycorrhiza formation in vitro. *New Phytol.* 151, 743–751. <https://doi.org/10.1046/j.0028-646x.2001.00219.x>.
23. Zhang, C., van der Heijden, M.G.A., Dodds, B.K., Nguyen, T.B., Spooren, J., Valzano-Held, A., Cosme, M., Berendsen, R.L., and Berendsen, R.L. (2024). A tripartite bacterial-fungal-plant symbiosis in the mycorrhiza-shaped microbiome drives plant growth and mycorrhization. *Microbiome* 12, 13. <https://doi.org/10.52843/cassyni.0kd91f>.
24. Partida-Martinez, L.P., Monajembashi, S., Greulich, K.O., and Hertweck, C. (2007). Endosymbiont-Dependent Host Reproduction Maintains Bacterial-Fungal Mutualism. *Curr. Biol.* 17, 773–777. <https://doi.org/10.1016/j.cub.2007.03.039>.
25. Vannini, C., Carpentieri, A., Salvioli, A., Novero, M., Marsoni, M., Testa, L., de Pinto, M.C., Amoresano, A., Ortolani, F., Bracale, M., and Bonfante, P. (2016). An interdomain network: the endobacterium of a mycorrhizal fungus promotes antioxidative responses in both fungal and plant hosts. *New Phytol.* 211, 265–275. <https://doi.org/10.1111/nph.13895>.
26. Venice, F., de Pinto, M.C., Novero, M., Ghignone, S., Salvioli, A., and Bonfante, P. (2017). *Gigaspora margarita* with and without its endobacterium shows adaptive responses to oxidative stress. *Mycorrhiza* 27, 747–759. <https://doi.org/10.1007/s00572-017-0790-z>.
27. Ruiz-Herrera, J., León-Ramírez, C., Vera-Núñez, A., Sánchez-Arreguín, A., Ruiz-Medrano, R., Salgado-Lugo, H., Sánchez-Segura, L., and Peña-Cabriaes, J.J. (2015). A novel intracellular nitrogen-fixing symbiosis made by *Ustilago maydis* and *Bacillus* spp. *New Phytol.* 207, 769–777. <https://doi.org/10.1111/nph.13359>.
28. Gohar, D., Pent, M., Pöldmaa, K., and Bahram, M. (2020). Bacterial community dynamics across developmental stages of fungal fruiting bodies. *FEMS Microbiol. Ecol.* 96, 175. <https://doi.org/10.1093/femsec/fiaa175>.
29. Mondo, S.J., Lastovetsky, O.A., Gaspar, M.L., Schwardt, N.H., Barber, C.C., Riley, R., Sun, H., Grigoriev, I.V., and Pawlowska, T.E. (2017). Bacterial endosymbionts influence host sexuality and reveal reproductive genes of early divergent fungi. *Nat. Commun.* 8, 1843. <https://doi.org/10.1038/s41467-017-02052-8>.
30. Xu, S., Liu, Y.X., Cernava, T., Wang, H., Zhou, Y., Xia, T., Cao, S., Berg, G., Shen, X.X., Wen, Z., et al. (2022). *Fusarium* fruiting body microbiome member *Pantoea agglomerans* inhibits fungal pathogenesis by targeting lipid rafts. *Nat. Microbiol.* 7, 831–843. <https://doi.org/10.1038/s41564-022-01131-x>.
31. Partida-Martinez, L.P., and Hertweck, C. (2005). Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* 437, 884–888. <https://doi.org/10.1038/nature03997>.
32. Lackner, G., Partida-Martinez, L.P., and Hertweck, C. (2009). Endofungal bacteria as producers of mycotoxins. *Trends Microbiol.* 17, 570–576. <https://doi.org/10.1016/j.tim.2009.09.003>.
33. Richter, I., Radosa, S., Cseresnyés, Z., Ferling, I., Büttner, H., Niehs, S.P., Gerst, R., Scherlach, K., Figge, M.T., Hillmann, F., and Hertweck, C. (2022). Toxin-Producing Endosymbionts Shield Pathogenic Fungus against Micropredators. *mBio* 13, e0144022. <https://doi.org/10.1128/mbio.01440-22>.
34. Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., Prill, R.J., Tripathi, A., Gibbons, S.M., Ackermann, G., et al. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551, 457–463. <https://doi.org/10.1038/nature24621>.
35. Pent, M., Pöldmaa, K., and Bahram, M. (2017). Bacterial communities in boreal forest mushrooms are shaped both by soil parameters and host identity. *Front. Microbiol.* 8, 836. <https://doi.org/10.3389/fmicb.2017.00836>.

36. Nishad, R., Ahmed, T., Rahman, V.J., and Kareem, A. (2020). Modulation of Plant Defense System in Response to Microbial Interactions. *Front. Microbiol.* **11**, 1298. <https://doi.org/10.3389/fmicb.2020.01298>.
37. Trouvelot, S., Héloir, M.C., Poinssot, B., Gauthier, A., Paris, F., Guillier, C., Combiér, M., Trdá, L., Daire, X., and Adrian, M. (2014). Carbohydrates in plant immunity and plant protection: roles and potential application as foliar sprays. *Front. Plant Sci.* **5**, 592. <https://doi.org/10.3389/fpls.2014.00592>.
38. Trivedi, P., Leach, J.E., Tringe, S.G., Sa, T., and Singh, B.K. (2020). Plant-microbiome interactions: from community assembly to plant health. *Nat. Rev. Microbiol.* **18**, 607–621. <https://doi.org/10.1038/s41579-020-0412-1>.
39. Baltrus, D.A., Dougherty, K., Arendt, K.R., Huntemann, M., Clum, A., Pillay, M., Palaniappan, K., Varghese, N., Mikhailova, N., Stamatis, D., et al. (2017). Absence of genome reduction in diverse, facultative endophytic bacteria. *Microb. Genom.* **3**, e000101. <https://doi.org/10.1099/mgen.0.000101>.
40. Du, M.Z., Zhang, C., Wang, H., Liu, S., Wei, W., and Guo, F.B. (2018). The GC content as a main factor shaping the amino acid usage during bacterial evolution process. *Front. Microbiol.* **9**, 2948. <https://doi.org/10.3389/fmicb.2018.02948>.
41. Zhang, H.Y., Bissett, A., Aguilar-Trigueros, C.A., Liu, H.W., and Powell, J.R. (2023). Fungal genome size and composition reflect ecological strategies along soil fertility gradients. *Ecol. Lett.* **26**, 1108–1118. <https://doi.org/10.1111/ele.14224>.
42. Hu, E.Z., Lan, X.R., Liu, Z.L., Gao, J., and Niu, D.K. (2022). A positive correlation between GC content and growth temperature in prokaryotes. *BMC Genom.* **23**, 110. <https://doi.org/10.1186/s12864-022-08353-7>.
43. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., et al. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419. <https://doi.org/10.1093/nar/gkaa913>.
44. Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E.V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36. <https://doi.org/10.1093/nar/28.1.33>.
45. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. <https://doi.org/10.1093/nar/28.1.27>.
46. Haft, D.H., Selengut, J.D., and White, O. (2003). The TIGRFAMs database of protein families. *Nucleic Acids Res.* **31**, 371–373. <https://doi.org/10.1093/nar/gkg128>.
47. Wani, A.K., Akhtar, N., Sher, F., Navarrete, A.A., and Américo-Pinheiro, J.H.P. (2022). Microbial adaptation to different environmental conditions: molecular perspective of evolved genetic and cellular systems. *Arch. Microbiol.* **204**, 144. <https://doi.org/10.1007/s00203-022-02757-5>.
48. Rangel, L.I., Hamilton, O., de Jonge, R., and Bolton, M.D. (2021). Fungal social influencers: secondary metabolites as a platform for shaping the plant-associated community. *Plant J.* **108**, 632–645. <https://doi.org/10.1111/tpj.15490>.
49. Brakhage, A.A. (2012). Regulation of fungal secondary metabolism. *Nat. Rev. Microbiol.* **11**, 21–32. <https://doi.org/10.1038/nrmicro2916>.
50. Zhang, C., van der Heijden, M.G.A., Dodds, B.K., Nguyen, T.B., Spooren, J., Valzano-Held, A., Cosme, M., and Berendsen, R.L. (2023). A tripartite bacterial-fungal-plant symbiosis in the mycorrhiza-shaped microbiome drives plant growth and mycorrhization. Preprint at bioRxiv. <https://doi.org/10.52843/cassyni.0kd91f>.
51. Scheuerl, T., Hopkins, M., Nowell, R.W., Rivett, D.W., Barraclough, T.G., and Bell, T. (2020). Bacterial adaptation is constrained in complex communities. *Nat. Commun.* **11**, 754.
52. Li, S.J., Hua, Z.S., Huang, L.N., Li, J., Shi, S.H., Chen, L.X., Kuang, J.L., Liu, J., Hu, M., and Shu, W.S. (2014). Microbial communities evolve faster in extreme environments. *Sci. Rep.* **4**, 6205. <https://doi.org/10.1038/srep06205>.
53. Castillo, D.M., and Pawlowska, T.E. (2010). Molecular evolution in bacterial endosymbionts of fungi. *Mol. Biol. Evol.* **27**, 622–636. <https://doi.org/10.1093/molbev/msp280>.
54. Cases, I., De Lorenzo, V., and Ouzounis, C.A. (2003). Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol.* **11**, 248–253. [https://doi.org/10.1016/s0966-842x\(03\)00103-3](https://doi.org/10.1016/s0966-842x(03)00103-3).
55. Cao, M., and Goodrich-Blair, H. (2017). Ready or not: Microbial adaptive responses in dynamic symbiosis environments. *J. Bacteriol.* **199**, e00883-16. <https://doi.org/10.1128/jb.00883-16>.
56. Ho, L.S.T., and Ané, C. (2014). A Linear-Time Algorithm for Gaussian and Non-Gaussian Trait Evolution Models. *Syst. Biol.* **63**, 397–408. <https://doi.org/10.1093/sysbio/syu005>.
57. Pinto-Tomás, A.A., Anderson, M.A., Suen, G., Stevenson, D.M., Chu, F.S.T., Cleland, W.W., Weimer, P.J., Currie, C.R., and Currie, C.R. (2009). Symbiotic nitrogen fixation in the fungus gardens of leaf-cutter ants. *Science* **326**, 1120–1123. <https://doi.org/10.1126/science.1173036>.
58. Barbieri, E., Ceccaroli, P., Saltarelli, R., Guidi, C., Potenza, L., Basaglia, M., Fontana, F., Baldan, E., Casella, S., Ryahi, O., et al. (2010). New evidence for nitrogen fixation within the Italian white truffle *Tuber magnatum*. *Fungal Biol.* **114**, 936–942. <https://doi.org/10.1016/j.funbio.2010.09.001>.
59. Thammahong, A., Puttikamonkul, S., Perfect, J.R., Brennan, R.G., and Cramer, R.A. (2017). Central Role of the Trehalose Biosynthesis Pathway in the Pathogenesis of Human Fungal Infections: Opportunities and Challenges for Therapeutic Development. *Microbiol. Mol. Biol. Rev.* **81**, e00053-16. <https://doi.org/10.1128/mmb.00053-16>.
60. Vanaporn, M., and Titball, R.W. (2020). Trehalose and bacterial virulence. *Virulence* **11**, 1192–1202. <https://doi.org/10.1080/21505594.2020.1809326>.
61. Moebius, N., Üzü, Z., Dijksterhuis, J., Lackner, G., and Hertweck, C. (2014). Active invasion of bacteria into living fungal cells. *Elife* **3**, e03007. <https://doi.org/10.7554/eLife.03007>.
62. Martin, F., and Selosse, M.A. (2008). The *Laccaria* genome: a symbiont blueprint decoded. *New Phytol.* **180**, 296–310. <https://doi.org/10.1111/j.1469-8137.2008.02613.x>.
63. Krizsán, K., Almási, É., Merényi, Z., Sahu, N., Virágh, M., Kószó, T., Mondo, S., Kiss, B., Bálint, B., Kües, U., et al. (2019). Transcriptomic atlas of mushroom development reveals conserved genes behind complex multicellularity in fungi. *Proc. Natl. Acad. Sci. USA* **116**, 7409–7418. <https://doi.org/10.1073/pnas.1817822116>.
64. Beier, S., and Bertilsson, S. (2013). Bacterial chitin degradation-mechanisms and ecophysiological strategies. *Front. Microbiol.* **4**, 149. <https://doi.org/10.3389/fmicb.2013.00149>.
65. Tunlid, A., Floudas, D., Op De Beeck, M., Wang, T., and Persson, P. (2022). Decomposition of soil organic matter by ectomycorrhizal fungi: Mechanisms and consequences for organic nitrogen uptake and soil carbon stabilization. *Front. For. Glob. Change* **5**, 150. <https://doi.org/10.3389/ffgc.2022.934409>.
66. Op De Beeck, M., Troein, C., Peterson, C., Persson, P., and Tunlid, A. (2018). Fenton reaction facilitates organic nitrogen acquisition by an ectomycorrhizal fungus. *New Phytol.* **218**, 335–343. <https://doi.org/10.1111/nph.14971>.
67. Bois Maillard, F., Kohler, A., Morin, E., Hossann, C., Miyauchi, S., Ziegler-Devlin, I., Gérard, D., Angeli, N., Lipzen, A., Keymanesh, K., et al. (2023). Functional genomics gives new insights into the ectomycorrhizal degradation of chitin. *New Phytol.* **238**, 845–858. <https://doi.org/10.1111/nph.18773>.
68. Rozmoš, M., Bukovská, P., Hršelová, H., Kotianová, M., Dudáš, M., Gančarčíková, K., and Jansa, J. (2022). Organic nitrogen utilisation by an arbuscular mycorrhizal fungus is mediated by specific soil bacteria

- and a protist. ISME J. 16, 676–685. <https://doi.org/10.1038/s41396-021-01112-8>.
69. López, M.F., Männer, P., Willmann, A., Hampp, R., and Nehls, U. (2007). Increased trehalose biosynthesis in Hartig net hyphae of ectomycorrhizas. *New Phytol.* 174, 389–398. <https://doi.org/10.1111/j.1469-8137.2007.01983.x>.
 70. Nehls, U. (2008). Mastering ectomycorrhizal symbiosis: the impact of carbohydrates. *J. Exp. Bot.* 59, 1097–1108. <https://doi.org/10.1093/jxb/erm334>.
 71. Wannet, W.J., Op den Camp, H.J., Wisselink, H.W., van der Drift, C., Van Griensven, L.J., and Vogels, G.D. (1998). Purification and characterization of trehalose phosphorylase from the commercial mushroom *Agaricus bisporus*. *Biochim. Biophys. Acta* 1425, 177–188. [https://doi.org/10.1016/s0304-4165\(98\)00066-x](https://doi.org/10.1016/s0304-4165(98)00066-x).
 72. Liu, J.H., Shang, X.D., Liu, J.Y., and Tan, Q. (2016). Changes in trehalose content, enzyme activity and gene expression related to trehalose metabolism in *flammulina velutipes* under heat shock. *Microbiology* 162, 1274–1285. <https://doi.org/10.1099/mic.0.000324>.
 73. Lei, M., Wu, X., Huang, C., Qiu, Z., Wang, L., Zhang, R., and Zhang, J. (2019). Trehalose induced by reactive oxygen species relieved the radial growth defects of *Pleurotus ostreatus* under heat stress. *Appl. Microbiol. Biotechnol.* 103, 5379–5390. <https://doi.org/10.1007/s00253-019-09834-8>.
 74. Liu, T., Zhu, L., Zhang, Z., Huang, H., Zhang, Z., and Jiang, L. (2017). Protective role of trehalose during radiation and heavy metal stress in *Aureobasidium subglaciale* F134. *Sci. Rep.* 7, 17586. <https://doi.org/10.1038/s41598-017-15489-0>.
 75. Ocón, A., Hampp, R., and Requena, N. (2007). Trehalose turnover during abiotic stress in arbuscular mycorrhizal fungi. *New Phytol.* 174, 879–891. <https://doi.org/10.1111/j.1469-8137.2007.02048.x>.
 76. Müller, J., Wiemken, A., and Aeschbacher, R. (1999). Trehalose metabolism in sugar sensing and plant development. *Plant Sci.* 147, 37–47. [https://doi.org/10.1016/s0168-9452\(99\)00094-1](https://doi.org/10.1016/s0168-9452(99)00094-1).
 77. Müller, J., Aeschbacher, R.A., Wingler, A., Boller, T., and Wiemken, A. (2001). Trehalose and Trehalase in *Arabidopsis*. *Plant Physiol.* 125, 1086–1093. <https://doi.org/10.1104/pp.125.2.1086>.
 78. Duponnois, R., and Kisa, M. (2006). The possible role of trehalose in the mycorrhiza helper bacterium effect. *Can. J. Bot.* 84, 1005–1008. <https://doi.org/10.1139/b06-053>.
 79. Kline, K.A., Dodson, K.W., Caparon, M.G., and Hultgren, S.J. (2010). A tale of two pili: assembly and function of pili in bacteria. *Trends Microbiol.* 18, 224–232. <https://doi.org/10.1016/j.tim.2010.03.002>.
 80. Telford, J.L., Barocchi, M.A., Margarit, I., Rappuoli, R., and Grandi, G. (2006). Pili in Gram-positive pathogens. *Nat. Rev. Microbiol.* 4, 509–519. <https://doi.org/10.1038/nrmicro1443>.
 81. Marathe, R., Meel, C., Schmidt, N.C., Dewenter, L., Kurre, R., Greune, L., Schmidt, M.A., Müller, M.J.I., Lipowsky, R., Maier, B., and Klumpp, S. (2014). Bacterial twitching motility is coordinated by a two-dimensional tug-of-war with directional memory. *Nat. Commun.* 5, 3759. <https://doi.org/10.1038/ncomms4759>.
 82. Furuno, S., Pätzelt, K., Rabe, C., Neu, T.R., Harms, H., and Wick, L.Y. (2010). Fungal mycelia allow chemotactic dispersal of polycyclic aromatic hydrocarbon-degrading bacteria in water-unsaturated systems. *Environ. Microbiol.* 12, 1391–1398. <https://doi.org/10.1111/j.1462-2920.2009.02022.x>.
 83. Wadhwa, N., and Berg, H.C. (2021). Bacterial motility: machinery and mechanisms. *Nat. Rev. Microbiol.* 20, 161–173. <https://doi.org/10.1038/s41579-021-00626-4>.
 84. Appels, F.V.W., van den Brandhof, J.G., Dijksterhuis, J., de Kort, G.W., and Wösten, H.A.B. (2020). Fungal mycelium classified in different material families based on glycerol treatment. *Commun. Biol.* 3, 334. <https://doi.org/10.1038/s42003-020-1064-4>.
 85. Worrich, A., Stryhanyuk, H., Musat, N., König, S., Banitz, T., Centier, F., Frank, K., Thullner, M., Harms, H., Richnow, H.H., et al. (2017). Mycelium-mediated transfer of water and nutrients stimulates bacterial activity in dry and oligotrophic environments. *Nat. Commun.* 8, 15472. <https://doi.org/10.1038/ncomms15472>.
 86. Warmink, J.A., Nazir, R., Corten, B., and van Elsas, J.D. (2011). Hitchhikers on the fungal highway: The helper effect for bacterial migration via fungal hyphae. *Soil Biol. Biochem.* 43, 760–765. <https://doi.org/10.1016/j.soilbio.2010.12.009>.
 87. de Novais, C.B., Sbrana, C., da Conceição Jesus, E., Rouws, L.F.M., Giovannetti, M., Avio, L., Siqueira, J.O., Saggin Júnior, O.J., da Silva, E.M.R., and de Faria, S.M. (2020). Mycorrhizal networks facilitate the colonization of legume roots by a symbiotic nitrogen-fixing bacterium. *Mycorrhiza* 30, 389–396. <https://doi.org/10.1007/s00572-020-00948-w>.
 88. Leone, M.R., Lackner, G., Silipo, A., Lanzetta, R., Molinaro, A., and Hertweck, C. (2010). An Unusual Galactofuranose Lipopolysaccharide That Ensures the Intracellular Survival of Toxin-Producing Bacteria in Their Fungal Host. *Angew. Chem.* 122, 7638–7642. <https://doi.org/10.1002/ange.201003301>.
 89. Mathis, R., Van Gijsegem, F., De Rycke, R., D'Haese, W., Van Maelsaeke, E., Anthonio, E., Van Montagu, M., Holsters, M., and Vereecke, D. (2005). Lipopolysaccharides as a communication signal for progression of legume endosymbiosis. *Proc. Natl. Acad. Sci. USA* 102, 2655–2660. <https://doi.org/10.1073/pnas.0409816102>.
 90. Jones, N., Blagih, J., Zani, F., Rees, A., Hill, D.G., Jenkins, B.J., Bull, C.J., Moreira, D., Bantan, A.I.M., Cronin, J.G., et al. (2021). Fructose reprogrammes glutamine-dependent oxidative metabolism to support LPS-induced inflammation. *Nat. Commun.* 12, 1209. <https://doi.org/10.1038/s41467-021-21461-4>.
 91. Rainey, P.B. (1991). Phenotypic variation of *Pseudomonas putida* and *P. tolaasii* affects attachment to *Agaricus bisporus* mycelium. *J. Gen. Microbiol.* 137, 2769–2779. <https://doi.org/10.1099/00221287-137-12-2769>.
 92. Bianciotto, V., Andreotti, S., Balestrini, R., Bonfante, P., and Perotto, S. (2001). Extracellular polysaccharides are involved in the attachment of *Azospirillum brasilense* and *Rhizobium leguminosarum* to arbuscular mycorrhizal structures. *Eur. J. Histochem.* 45, 39–49. <https://doi.org/10.4081/1612>.
 93. Al Shaer, D., Al Musaimi, O., de la Torre, B.G., and Albericio, F. (2020). Hydroxamate siderophores: Natural occurrence, chemical synthesis, iron binding affinity and use as Trojan horses against pathogens. *Eur. J. Med. Chem.* 208, 112791. <https://doi.org/10.1016/j.ejmech.2020.112791>.
 94. Rahman, M.A., Parvin, M., Das, U., Ela, E.J., Lee, S.H., Lee, K.W., and Kabir, A.H. (2020). Arbuscular Mycorrhizal Symbiosis Mitigates Iron (Fe)-Deficiency Retardation in Alfalfa (*Medicago sativa* L.) Through the Enhancement of Fe Accumulation and Sulfur-Assisted Antioxidant Defense. *Int. J. Mol. Sci.* 21, 2219. <https://doi.org/10.3390/ijms21062219>.
 95. Caris, C., Hördt, W., Hawkins, H.J., Römhild, V., and George, E. (1998). Studies of iron transport by arbuscular mycorrhizal hyphae from soil to peanut and sorghum plants. *Mycorrhiza* 8, 35–39. <https://doi.org/10.1007/s005720050208>.
 96. Kabir, A.H., Debnath, T., Das, U., Prity, S.A., Haque, A., Rahman, M.M., and Parvez, M.S. (2020). Arbuscular mycorrhizal fungi alleviate Fe-deficiency symptoms in sunflower by increasing iron uptake and its availability along with antioxidant defense. *Plant Physiol. Biochem.* 150, 254–262. <https://doi.org/10.1016/j.plaphy.2020.03.010>.
 97. Pellegrin, C., Morin, E., Martin, F.M., and Veneault-Fourrey, C. (2015). Comparative Analysis of Secretomes from Ectomycorrhizal Fungi with an Emphasis on Small-Secreted Proteins. *Front. Microbiol.* 6, 1278. <https://doi.org/10.3389/fmicb.2016.01734>.
 98. Netherway, T., Bengtsson, J., Krab, E.J., and Bahram, M. (2021). Biotic interactions with mycorrhizal systems as extended nutrient acquisition strategies shaping forest soil communities and functions. *Basic Appl. Ecol.* 50, 25–42. <https://doi.org/10.1016/j.baae.2020.10.002>.
 99. Bahram, M., Netherway, T., Hildebrand, F., Pritsch, K., Drenkhan, R., Loit, K., Anslan, S., Bork, P., and Tedersoo, L. (2020). Plant

- nutrient-acquisition strategies drive topsoil microbiome structure and function. *New Phytol.* 227, 1189–1199. <https://doi.org/10.1111/nph.16598>.
100. Nehls, U., Mikolajewski, S., Magel, E., and Hampp, R. (2001). Carbohydrate metabolism in ectomycorrhizas: gene expression, monosaccharide transport and metabolic control. *New Phytol.* 150, 533–541. <https://doi.org/10.1046/j.1469-8137.2001.00141.x>.
 101. Salzer, P., and Hager, A. (1991). Sucrose Utilization of the Ectomycorrhizal Fungi *Amanita muscaria* and *Hebeloma crustuliniforme* Depends on the Cell Wall-bound Invertase Activity of their Host *Picea abies*. *Bot. Acta* 104, 439–445. <https://doi.org/10.1111/j.1438-8677.1991.tb00256.x>.
 102. Salzer, P., and Hager, A. (1993). Characterization of wall-bound invertase isoforms of *Picea abies* cells and regulation by ectomycorrhizal fungi. *Physiol. Plant.* 88, 52–59. <https://doi.org/10.1111/j.1399-3054.1993.tb01759.x>.
 103. Read, D.J., and Perez-Moreno, J. (2003). Mycorrhizas and nutrient cycling in ecosystems – a journey towards relevance? *New Phytol.* 157, 475–492. <https://doi.org/10.1046/j.1469-8137.2003.00704.x>.
 104. Wiese, J., Kleber, R., Hampp, R., and Nehls, U. (2000). Functional characterization of the *Amanita muscaria* monosaccharide transporter, AmMst1. *Plant Biol.* 2, 278–282. <https://doi.org/10.1055/s-2000-12984>.
 105. Deveau, A., Kohler, A., Frey-Klett, P., and Martin, F. (2008). The major pathways of carbohydrate metabolism in the ectomycorrhizal basidiomycete *Laccaria bicolor* S238N. *New Phytol.* 180, 379–390. <https://doi.org/10.1111/j.1469-8137.2008.02581.x>.
 106. Nehls, U., Grunze, N., Willmann, M., Reich, M., and Küster, H. (2007). Sugar for my honey: Carbohydrate partitioning in ectomycorrhizal symbiosis. *Phytochemistry* 68, 82–91. <https://doi.org/10.1016/j.phytochem.2006.09.024>.
 107. Wright, D.P., Scholes, J.D., Read, D.J., and Rolfe, S.A. (2000). Changes in carbon allocation and expression of carbon transporter genes in *Betula pendula* Roth. colonized by the ectomycorrhizal fungus *Paxillus involutus* (Batsch) Fr. *Plant Cell Environ.* 23, 39–49. <https://doi.org/10.1046/j.1365-3040.2000.00518.x>.
 108. Schaarschmidt, S., González, M.C., Roitsch, T., Strack, D., Sonnewald, U., and Hause, B. (2007). Regulation of arbuscular mycorrhization by carbon. The symbiotic interaction cannot be improved by increased carbon availability accomplished by root-specifically enhanced invertase activity. *Plant Physiol.* 143, 1827–1840. <https://doi.org/10.1104/pp.106.096446>.
 109. Schaeffer, C., Wallenda, T., Guttenberger, M., and Hampp, R. (1995). Acid invertase in mycorrhizal and non-mycorrhizal roots of Norway spruce (*Picea abies* [L.] Karst.) seedlings. *New Phytol.* 129, 417–424. <https://doi.org/10.1111/j.1469-8137.1995.tb04312.x>.
 110. Itabangi, H., Sephton-Clark, P.C.S., Tamayo, D.P., Zhou, X., Starling, G.P., Mahamoud, Z., Insua, I., Probert, M., Correia, J., Moynihan, P.J., et al. (2022). A bacterial endosymbiont of the fungus *Rhizopus microsporus* drives phagocyte evasion and opportunistic virulence. *Curr. Biol.* 32, 1115–1130.e6. <https://doi.org/10.1016/j.cub.2022.01.028>.
 111. Anca, I.A., Lumini, E., Ghignone, S., Salvioli, A., Bianciotto, V., and Bonfante, P. (2009). The *ftsZ* gene of the endocellular bacterium '*Candidatus Glomeribacter gigasporarum*' is preferentially expressed during the symbiotic phases of its host mycorrhizal fungus. *Mol. Plant Microbe Interact.* 22, 302–310. <https://doi.org/10.1094/mpmi-22-3-0302>.
 112. Chen, I.M.A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., Huntemann, M., Varghese, N., White, J.R., Seshadri, R., et al. (2019). IMG/M v. 5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* 47, D666–D677. <https://doi.org/10.1093/nar/gky901>.
 113. Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. <https://doi.org/10.7287/peerj.preprints.554v1>.
 114. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu033>.
 115. Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
 116. Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G., and Toth, I.K. (2016). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal. Methods* 8, 12–24. <https://doi.org/10.1039/c5ay02550h>.
 117. Katoh, K., Misawa, K., Kuma, K.I., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. <https://doi.org/10.1093/nar/gk436>.
 118. Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One* 5, e9490. <https://doi.org/10.1371/journal.pone.0009490>.
 119. Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. <https://doi.org/10.1093/bioinformatics/bty633>.
 120. Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Series B Stat. Methodol.* 63, 411–423. <https://doi.org/10.1111/1467-9868.00293>.
 121. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Marchler, G.H., Song, J.S., et al. (2020). CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 48, D265–D268. <https://doi.org/10.1093/nar/gkz991>.
 122. Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, W445–W451. <https://doi.org/10.1093/nar/gks479>.
 123. Na, S.I., Kim, Y.O., Yoon, S.H., Ha, S.M., Baek, I., and Chun, J. (2018). UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J. Microbiol.* 56, 280–285. <https://doi.org/10.1007/s12275-018-8014-6>.
 124. Falcon, S., Gentleman, R., Gentleman, R., Falcon, S., Falcon, S., and Gentleman, R. (2008). Hypergeometric testing used for gene set enrichment analysis. In *Bioconductor case studies* (Springer), pp. 207–220. https://doi.org/10.1007/978-0-387-77240-0_14.
 125. Szumilas, M. (2010). Explaining Odds Ratios. *J. Can. Acad. Child Adolesc. Psychiatry* 19, 227–229.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Genome data of 56 fruiting body associated bacterial Isolates	This study	https://doi.org/10.5281/zenodo.10889199
Software and algorithms		
DOE-JGI Microbial Genome Annotation Pipeline (MGAP)	Chen et al. ¹¹²	https://img.jgi.doe.gov/docs/pipelineV5/
CheckM	Parks et al. ¹¹³	https://doi.org/10.7287/peerj.preprints.554v1
Trimmomatic	Bolger et al. ¹¹⁴	http://www.usadellab.org/cms/?page=trimmomatic
Megahit	Li et al. ¹¹⁵	https://github.com/voutcn/megahit
pyANI	Pritchard et al. ¹¹⁶	https://pyani.readthedocs.io/en/latest/about.html
UBCG pipeline	Na et al. ¹¹³	https://help.ezbiocloud.net/ubcg-users-manual/
MAFFT	Katoh et al. ¹¹⁷	https://mafft.cbrc.jp/alignment/software/
FastTree	Price et al. ¹¹⁸	https://morgannprice.github.io/fasttree/
ape	Paradis et al. ¹¹⁹	https://cran.r-project.org/web/packages/ape/index.html
hclust	Tibshirani et al. ¹²⁰	https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/hclust
RPS-blast	Lu et al. ¹²¹	https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd_help.shtml#RPSBWhat
pfam database	Mistry et al. ⁴³	http://pfam.xfam.org/
COG database	Tatusov et al. ⁴⁴	https://www.ncbi.nlm.nih.gov/research/cog
KEGG Pathway database	Kanehisa and Goto ⁴⁵	https://www.genome.jp/kegg/
Tigrfam	Haft et al. ⁴⁶	https://tigrfams.jcvi.org/cgi-bin/index.cgi
dbCAN pipeline	Yin et al. ¹²²	https://github.com/linnabrown/run_dbcan

METHOD DETAILS

Study sites and fungal sampling

Fungal fruiting bodies from different genera were collected from boreal forests at three nature reserves in Eastern Estonia. Only mature fruiting bodies were sampled, excluding the decayed or damaged fruiting bodies. All fruiting bodies were individually packed in aluminum foil, transported to the lab using a cold container and stored at 4°C until processed in a laminar hood.

Culturing, DNA extraction and library preparation

Fruiting bodies underwent careful cutting and sterilization procedures. Subsequently, 5 mm³ inner tissue sections of the fruiting bodies were used to liberate bacteria from these. Detailed methodologies pertaining to sample handling and bacterial culturing are described in Pent et al., 2017.³⁵

After pure cultures of bacteria were established, their colonies were subjected to DNA extraction, followed by PCR amplification of 16S gene for identification, as outlined in Pent et al., 2017.³⁵ The isolated strains were preserved at -80°C in 50% glycerol solution in the microbiological collection of the Natural History Museum and the Collection of Environmental and Laboratory Microbial Strains (CELMS) at the University of Tartu.

Based on the 16S rDNA-based identification of bacterial strains, we selected the 56 potential nitrogen-fixing bacteria for whole-genome sequencing. The sequencing library for each DNA extract was prepared using the NEXTRA XT kit, which fragments DNA and tags it with sequencing adapters in a single-tube enzymatic reaction.

Sequencing

DNA from libraries was sequenced on a next-generation sequencing platform using Illumina Novoseq 2xPE150 technology at Novogen corporation Inc. (U.K). Sequenced DNA was quality-filtered and assembled *de novo* via Trimmomatic¹¹⁴ and Megahit,¹¹⁵ respectively. The assembly quality was assessed using BUSCO v3 and CheckM with default parameters. Scaffolds smaller than 500 bp were removed. Assembled genomes were further annotated using the DOE-JGI Microbial Genome Annotation Pipeline (MGAP).

Genome data selection and filtration

We retrieved 4,388 bacterial genomes (belonging to similar genera or to the lowest taxonomic rank in common to strains included from our collection) inhabiting different habitats such as the human body, soil, fungi (mycelia, fruiting-bodies, mycorrhizal root tips), plants (roots, leaves, stems, branches), and water bodies (oceans, lakes) from the IMG database.¹¹² Isolation sites were identified through manual curation, including scanning IMG metadata and NCBI biosample data and literature where required. Bacteria were classified as either host-associated or environmental bacteria, based on the available information on the host, habitat type or isolation source. Bacteria isolated from structures of humans, plants, and fungi were classified as host-associated bacteria regardless of the nature of their interactions with their host. Similarly, those isolated from terrestrial and aquatic habitats were classified as environmental bacteria. We found 138 genomes of fungus-associated bacteria from the public database. Following the quality filtering, only 107 genomes were included in the subsequent analysis, 83 taxonomically corresponding to newly generated genomes while others were included to increase the sample size for FaB. While metadata for the source of host fungi was missing for many FaB genomes, the majority of FaB included in this study originate from either fungal fruiting bodies or the mycelia in the soil, as indicated by the metadata in the public database. However, for a comprehensive comparison of fungus-associated bacteria, we added 56 newly sequenced genomes from fruiting bodies representing different fungal lineages, bringing the total number of FaB genomes to 163 (Table S1).

We applied the following rigorous quality control measures to ensure the selection of a quality and unbiased set of genomes in the subsequent analysis: (1) Based on the metadata and taxonomy, we selected up to 5 random genomes of the same species from the same habitat; (2) Genomes with missing or ambiguous isolation sites were discarded; (3) we used CheckM¹¹³ to assess the genome completeness and contamination, genomes that were predicted to be less than 95% complete and have more than 5% contamination were removed. In addition, we removed the genomes that did not have at least 90% of the 92 single-copy orthologous genes⁹; (4) To evaluate high-quality annotation, genomes with less than 85% protein-coding sequences were discarded; and (5) We used pyANI¹¹⁶ to compute the average nucleotide identity and alignment fraction values for each pair of genomes. We marked the genome pair as redundant when the alignment fraction exceeded 90%, and the average nucleotide identity was greater than 99.995%. In such a case, one genome was randomly selected, and the other was filtered out. This led to the exclusion of many obligate and endosymbiotic bacteria due to their smaller genome sizes, which did not meet our quality filtering criteria. However, we chose to apply the same criteria uniformly across all habitats, even though this resulted in the exclusion of some data. Finally, we selected 1211 genomes from four different phyla, as shown in Figure 1A, for the final analysis.

Construction of the phylogenetic tree

To generate a phylogenetic tree of 1211 high-quality non-redundant genomes, we retrieved 92 single-copy homologous genes from each genome with a UBCG pipeline.¹²³ Each marker gene codon-based alignment (amino acid sequences) was aligned using MAFFT¹¹⁷ with default parameters. All 92 gene alignments were then concatenated into an overall merged alignment, which was used to build a maximum-likelihood phylogenetic tree with FastTree.¹¹⁸

Clustering of 1211 genomes into eight clusters

We divided 1211 genomes into eight taxa to allow downstream gene enrichment analysis. We first converted the phylogenetic tree into a distance matrix using a cophenetic function from the “ape” package in R.¹¹⁹ We then clustered 1211 genomes into eight clusters using the hierarchical clustering method using the function “hclust”. The algorithm clusters n objects based on pre-defined k -mer numbers. To determine the optimal value of k we used silhouette and gap statistics.¹²⁰ We selected $k=8$ because it yielded a maximum silhouette coefficient (0.77) (Figure S1). The resulting clusters have few overlapping genomes that might have appeared because of false taxonomic identification provided with their metadata. Our clustering divided the Proteobacteria genomes into two taxa that we renamed, for simplicity, Proteobacteria1 and Proteobacteria2.

Gene category enrichment analysis

We retrieved protein-coding genes, for each genome using the program RPS-blast¹²¹ at an e-value cutoff of $1e-4$ and alignment length of at least 70% of the consensus sequence length, and mapped them to COG (Cluster of Orthologous) IDs. Each COG ID was translated to at least one COG category (Table S2). We then counted the number of genes from a given category for individual genomes. We then assessed the enrichment patterns of gene categories across different habitats using Phylogenetic linear Regression (PhyloLM) and Hypergeometric tests. PhyloLM incorporates phylogenetic information to account for enrichments that may arise due to shared ancestry,^{8,49} while the hypergeometric test focuses solely on statistical enrichment based on categorical data without considering evolutionary relationships.¹²⁴ Both methods were used to validate and complement the identified enrichment patterns, ensuring robust results. Furthermore, to reduce the bias, we used standardized gene copy numbers as input and predicted the effect size for individual genes. Additionally, we utilized the median of genes for each category to present our results, providing a comprehensive analysis of gene enrichments across the habitats.

ABC transporters

To identify the ABC transporter pathways within bacterial genomes across habitats, we sourced information on transporter genes associated with each pathway from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.⁴⁵ We specifically tallied

complete pathways, which contain all necessary genes within a genome, to assess the prevalence of complete pathways in bacterial genomes from different habitats. The list of genes involved in each pathway is presented in [Table S6](#).

Genome-wide analysis of CAZyme domains

To analyze the genome-wide carbohydrate-active enzyme (CAZyme) related genes, we used the dbCAN pipeline.¹²² The protein sequences were annotated using DIAMOND, HMMER, and eCAMI algorithms integrated within dbCAN against CAZy and dbCAN databases. To avoid ambiguous annotations, we filtered out the genes not annotated by at least two algorithms. We then used custom R scripts that included basic data manipulation of results obtained from the dbCAN. First, we extracted all the CAZyme domains present in each genome collated them by domain family, and then summarized them into a table including the count of all the CAZyme domains in each family per genome. These collated and counted values of domain frequency per CAZyme domain family per genome were visualized using R.

QUANTIFICATION AND STATISTICAL ANALYSIS

Unless otherwise mentioned, we used R, using the latest available version at the time of analysis, to statistically analyze the data. We distinguished functional repertoires between habitat types using distance-based redundancy analyses (dbRDA) and permutational multivariate analysis of variance (PERMANOVA) based on Bray-Curtis dissimilarity indices between genomes calculated from gene copy numbers represented in The Protein families (Pfam), COG, KEGG, and The Institute for Genomic Research's database of Protein Families (Tigrfam) annotations (host-associated and environmental). Effects of estimates of PhyloLM models were predicted using the "Predict" function, and median values of predicted residuals for each category and habitat were compared using pairwise comparison tests. All of the pairwise comparisons in this study were performed using Wilcoxon tests and t-tests, Kruskal-Wallis, and Mann-Whitney U test. P-values resulting from all tests were corrected using the False Discovery Rate (FDR) method. Furthermore, the overrepresentation of genes was also compared using pairwise tests. Hypergeometric and Odds Ratio tests were used to test the association of genes with their respective habitat. The Odds Ratio test is a statistical method used to measure the strength of association between two variables by comparing their probability of occurrence in different groups.¹²⁵