



Elucidating the interplay between metabolites and microorganisms in the spermosphere of common bean (*Phaseolus vulgaris* L.) seeds

Chandrodhay Saccaram, Marie Simonin, Stéphanie Boutet, Céline Brosse, Shuang Peng, Tracy François, Boris Collet, François Perreau, Delphine Sourdeval, Coralie Marais, et al.

► To cite this version:

Chandrodhay Saccaram, Marie Simonin, Stéphanie Boutet, Céline Brosse, Shuang Peng, et al.. Elucidating the interplay between metabolites and microorganisms in the spermosphere of common bean (*Phaseolus vulgaris* L.) seeds. *mSystems*, 2025, 10 (8), pp.e00707-25. <10.1128/msystems.00707-25>. <hal-05533608>

HAL Id: hal-05533608

<https://hal.science/hal-05533608v1>

Submitted on 2 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-SA 4.0 - Attribution - Non-commercial use - ShareAlike - International License

Elucidating the interplay between metabolites and microorganisms in the spermosphere of common bean (*Phaseolus vulgaris* L.) seeds

Chandrodhay Saccaram^a, Marie Simonin^b, Stéphanie Boutet^a, Céline Brosse^a, Shuang Peng^a, Tracy François^a, Boris Collet^a, François Perreau^a, Delphine Sourdeval^a, Coralie Marais^b, Matthieu Barret^b, Loïc Rajjou^{a#}, Massimiliano Corso^{a#}

^a Université Paris-Saclay, INRAE, AgroParisTech, Institute Jean-Pierre Bourgin for Plant Sciences (IJPB), 78000, Versailles, France.

^b IRHS-UMR1345, Université d'Angers, INRAE, Institut Agro, SFR 4207 QuaSaV, 49071, Beaucouzé, France.

Running Title: Metabolome-Microbiome Interaction in the Spermosphere

Correspondence to: Loïc Rajjou & Massimiliano Corso

loic.rajjou@agroparistech.fr & massimiliano.corso@inrae.fr

Keywords: Seeds, Spermosphere, Germination, Metabolome, Microbiome, *Phaseolus vulgaris*

4338 Words (excluding abstract, figure legends and references)

2 Tables

5 Figures

1 Supplementary Methods

13 Supplementary Tables

8 Supplementary Figures

ABSTRACT (215 words)

The spermosphere, the dynamic interface surrounding germinating seeds, is shaped by the intricate interplay between seed-exuded natural compounds and seed-associated microbial communities. In this work we provide the first comprehensive metabolomic and microbiome characterization of common bean (*Phaseolus vulgaris*) spermosphere of eight genotypes produced in two contrasted production regions. Through an integrated approach, we explored the metabolomic and microbiota composition in the spermosphere of germinating common bean seeds and elucidated their environmental and genotype regulation. We detected and analysed 2,467 metabolite features (Mf) through untargeted metabolomics categorized into fourteen metabolic categories, highlighting the prevalence of amino acids, flavonoids, and terpenoids. Genotype was the key factor influencing the chemical composition of the spermosphere. Furthermore, we identified 19 bacterial families and 23 fungal families inhabiting the spermosphere, with both genotype and seed production location exerting varying degrees of influence on microbial community composition. Through a multiscale integrated approach, we revealed specific associations between metabolites and microorganisms such as negative correlation between flavonoids and *Bacillus* spp., emphasizing the genotype-dependent nature of these interactions. This comprehensive investigation sheds light on the mechanisms underlying seed germination and the complex interactions between plant genotypes, seed exudates, environmental conditions, and microbial communities in the spermosphere. These findings provide a framework for developing innovative strategies to promote seed health and sustainable crop production.

IMPORTANCE (149 words)

The spermosphere, the dynamic interface around germinating seeds, is shaped by the intricate interplay between seed-exuded compounds and microbial communities. Despite the importance of these interactions for eventual seedling emergence and health, little knowledge is available on the subject. We are the first to comprehensively analyse the chemical and microbial diversity of the spermosphere of *Phaseolus vulgaris* (common bean). We identified thousands of primary and specialized metabolites, highlighting their diversity but largely unknown roles in germinating seed-environment interactions. We revealed significant genotype-driven differences in the chemical composition as well as the influence of both genotype and seed production location on microbial community structure in the spermosphere. Our metabolome-microbiome integrative analysis suggests that common bean shapes the spermosphere microbiome through specific seed exudates. This research advances our understanding in the metabolic capabilities and ecological roles of seed microbiota within the spermosphere, contributing to our understanding of seed health and vigour.

1. INTRODUCTION (520 words)

Seeds play a crucial role in plant propagation and agricultural sustainability, with their quality and vigour significantly impacting crop yield and resilience in diverse environmental conditions (1, 2). During germination, seeds release complex mixtures of organic and inorganic compounds into their surrounding environment, creating a localized microenvironment known as the spermosphere, *i.e.* a zone of increased microbial activity around germinating seeds (3–6). While extensive research has focused on plant microbiomes in the phyllosphere and rhizosphere, the spermosphere and its microbial and chemical composition have received limited attention (3, 5, 6). Seeds serve as vital conduits for transmitting microorganisms between plant generations, influencing the establishment of the plant microbiome (7). Previous studies have shown that seed microbial communities are influenced by seed genotype and production environment (8, 9), possibly due to the diversity of primary and specialized metabolites accumulated in seeds (10). Seed exudates during germination further contribute to shaping the spermosphere microbiome, similar to plant-microorganism interactions in the rhizosphere mediated by root exudates (11) but this remains to be investigated. For instance, the isoflavone daidzein influenced microbial community structure in the rhizosphere (12, 13). Furthermore, the flavanone naringenin has induced chemotaxis in rhizobacteria *Aeromonas* sp. (14).

Despite the significance of metabolic transition during seed germination (15), limited information exists on the chemical (16) and microbial composition of the seed spermosphere (17), which are likely to play essential roles in seed interactions with the environment. Seed germination and seedling emergence have been shown to favour microbial taxa with fast-growing abilities (18), suggesting the importance of elucidating interactions in the spermosphere.

Common bean (*Phaseolus vulgaris* L.) is a globally consumed grain legume rich in nutraceutical metabolites and proteins (19, 20). The domestication of common bean led to phenotypic changes affecting both the

seed microbiome community structure and chemical composition (21, 22). Studying common bean interaction with the environment is crucial to improve seed vigor. In this context, seeds of this species provide a suitable model to explore the interactions between metabolites and microorganisms in the spermosphere (23).

In this work we designed a simplified experimental setup to collect the spermosphere upon seed imbibition for studying metabolome and microbiome interplay. We hypothesized that seeds play a central role in shaping the spermosphere by influencing both exudate composition and associated microbial communities. To investigate this, we (i) characterized the chemical and microbial composition of the common bean spermosphere, (ii) identified the factors (namely genotype and seed production environment) that influence spermosphere composition and microbial assembly, and (iii) explored associations between seed-borne microorganisms and seed-exuded metabolites. Using integrated metabolomic and microbiome multi-omics data, we characterized, for the first time, the metabolic and microbial composition of the common bean spermosphere, as well as the complex interactions between metabolites and microorganisms. We demonstrated that the chemical composition of the common bean seed spermosphere is primarily influenced by genotype, while seed-borne microbial communities are shaped by both genotype and the seed production environment. Integration of multi-omics data enabled the identification of potential associations between spermosphere metabolites and both beneficial and pathogenic microorganisms. These findings provide new insights into the ecological roles of spermosphere metabolites and microbiota in common bean, enhancing our understanding of seed health and vigour.

2. RESULTS (1274 words)

2.1. Chemical and microbial diversity in the common bean spermosphere

2.1.1. Chemical diversity in common bean spermosphere

Metabolite diversity of the germinating seed spermosphere of eight common bean genotypes produced in two locations (24, 25) was analysed using untargeted metabolomics (LC–MS/MS) in ElectroSpray Ionization in positive (ESI+) and negative (ESI–) modes. The metabolomic analyses allowed the detection of 2,467 ions, or metabolite features (Mf), with a putative or unknown annotation. LC-MS/MS data were used to build a molecular network (Fig. 1A; Fig. S4) (26, 27) to improve the assignment of each Mf to a metabolic category (Fig. 1B, Table S1). This analysis allowed the assignment of a total of 506 Mf (21% of the spermosphere metabolome) to a known metabolic category. Mf identified in the germinating common bean seed spermosphere have been assigned to fourteen putative major metabolic categories (Fig. 1B) and specific classes (Table S1). The categories in which the highest number of Mf were detected correspond to amino acid (aa) and derivatives [140], flavonoids [126] and terpenoids and derivatives [55]. Amino acids and derivatives were mainly represented by very small peptides [102], flavonoids were mainly represented by flavonols [44] and flavan-3-ols [30], and jasmonic acid and derivatives [7] were predominant in the plant hormone category (Table S1).

2.1.2. Spermosphere microbial assemblages were represented by 19 bacterial and 23 fungal families

Spermosphere bacterial and fungal communities were studied, using the exact same samples employed for metabolomic analysis, by a metabarcoding community profiling approach. (Table S2 and S3).

The bacterial community in the spermosphere of common bean samples primarily comprised two prominent phyla: *Pseudomonadota* and *Bacillota*, with average relative abundances of 75.8% and 23.2% respectively across all samples (Fig. 2A, Table S4). Nineteen bacterial families were identified, with

Erwiniaceae being notably substantial with average relative abundance of 55.0% across all samples (Fig. 2B, Table S4). Within this family, *Pantoea* with average relative abundance of 33.7% across all samples emerged as the most abundant genus (Fig. 2C, Table S4) with *Pantoea agglomerans* (B_ASV2) showing high conservation across samples, being present in over 70% of all samples. This highlights its significance as a core component of the spermosphere taxa.

Ascomycota and *Basidiomycota* with average relative abundance across all samples of 97.1% and 2.9% respectively were two fungal phyla in the common bean spermosphere (Fig. 2E, Table S5). Fungi were further categorised into 23 distinct families, with the *Pleosporaceae* family being notably overrepresented with average relative abundance of 64.3% across all samples of germinating bean spermosphere (Fig. 2F, Table S5). *Alternaria* and *Stemphylium* (*Pleosporaceae* family) emerged as the most abundant fungal genera across the common bean spermosphere, boasting average relative abundances of 34.0% and 29.9% across all samples respectively (Fig. 2G, Table S5). Several fungal species were present in over 70% of the common bean genotype samples, thereby forming the core fungal taxa within the spermosphere (Table S5). These included non-pathogenic and potential pathogenic filamentous fungi, such as *Gibberella intricans*, *Stemphylium herbarum*, *Botrytis fabae*, *Alternaria infectoria*, an *Alternaria* sp., *Cladosporium delicatulum*, as well as two *Cladosporium* spp. (F_ASV8 and F_ASV105), and Basidiomycota yeasts, such as *Vishniacozyma victoriae*.

2.1.3. Contrasted patterns in alpha diversity for bacteria and fungi

Bacterial and fungal richness levels were then assessed. Bacterial richness, determined through *gyrB* sequence analysis, unveiled a median of 25 ASVs, with a large variability spanning from 2 to 46 ASVs among samples (Fig. 2D, Table S6). Likewise, fungal richness exhibited a median richness of 29 ASVs, encompassing a range of 3 to 43 ASVs (Fig. 2H, Table S7).

Bacterial richness was mainly influenced by the seed production location (Fig. 2D; ANOVA, $p < 0.01$), whereas fungal richness was significantly affected by location, genotype, and their interaction (Fig. 2H; ANOVA, $p < 0.05$). Overall, both bacterial and fungal richness were higher in seeds produced in Maine-et-Loire compared to those from the Gers region (Fig. 2D, 2H).

2.2. The metabolic landscape and microbial community structure of the germinating common bean spermosphere is influenced by both the genotype and the site of production

Statistical analyses, including ANOVA ($p < 0.05$; Fig. 1I and 1J), PCA, and PERMANOVA (Fig. 3A and 3G), revealed that genotype is the primary factor influencing spermosphere metabolic diversity across all metabolic categories, whereas location had a comparatively limited effect on the metabolite landscape. Heatmap of relative abundance of metabolites (Fig. 3A) and metabolic classes (Figure S5) in the spermosphere further confirmed separation of samples mainly according to genotypes.

Microbiome analyses of the common bean spermosphere revealed a comparable impact of genotype on bacterial (23%; Fig. 3C, 3D, 3G) and fungal (19%; Fig. 3E, 3F, 3G) community diversity. In contrast, location had a stronger effect on fungal diversity, accounting for 43% of the variation (Fig. 3E, 3G). While hierarchical clustering based on bacterial abundance showed no clear pattern in distribution of samples (Fig. 3D), fungal abundance data showed a clear distribution of samples between the two locations (Fig. 3F).

2.3. Metabolome-microbiome integration analysis

2.3.1. Multiscale network analysis

The relationship between metabolome and microbiome in common bean spermosphere was explored through a correlation network analysis approach (28). Pairwise spearman metabolite-microorganisms and microorganisms-microorganisms correlation ($r \geq |0.5|$; $p < 0.05$) were selected to create a network (Fig. 4A, Table S10). The resulting network encompassed a total of 1,736 nodes, comprising metabolites [1,559],

bacteria [107], and fungi [70], connected by 6,225 edges representing the correlations (Fig. 4A and B). Negative correlation coefficients ranged from -0.825 to -0.5, while the positive coefficients spanned from 0.5 to 0.909. The top 15 nodes having the greatest degree (i.e. the number of connected nodes) are of microbial origin and are mainly represented by *Bacillus* spp. (Table 1).

2.3.2. Complex interactions between microbial communities and metabolites within the germinating seed spermosphere

Specific correlation subnetworks for each major metabolic category (e.g., flavonoids) were constructed. More specifically, nodes of annotated metabolic categories correlated with microorganism nodes (first neighbours) were selected (sub-networks). The three largest correlation subnetworks belonged to amino acids and derivatives, flavonoids, and terpenoids and derivatives (Fig. 4B).

Enrichment analysis (hypergeometric test, $p < 0.05$, fold enrichment > 1.5 , Table 2) was carried out to identify over-represented bacterial and fungal genus for each annotated metabolic category sub-network (positive and negative correlations). *Bacillus* genus was over-represented within the subnetworks of amino acids and derivatives, flavonoids, terpenoids and derivatives, carbohydrates, and alkaloids. On the other hand, *Cladosporium* was over-represented across subnetworks of flavonoids, nucleosides and derivatives, cinnamic acids and derivatives and terpenoids and derivatives sub-networks.

2.3.3. *Bacillus* spp. are negatively correlated to different flavonoids

Bacillus was overrepresented across multiple annotated metabolic categories and ranked among the most abundant genera in the common bean spermosphere. To further investigate its associations, a specific correlation subnetwork was constructed for *Bacillus* spp. (including all ASVs identified as *Bacillus*), capturing both positive and negative correlations. In total, 496 molecular features (Mfs), of which 103 were annotated, showed positive correlations, while 91 Mfs (33 annotated) were negatively correlated with *Bacillus* spp. (Table S10). Flavonoids exhibited mainly negative correlations with *Bacillus* spp.,

whereas aa and peptides showed mainly positive correlations (Fig. 4C). *Bacillus* spp. were less abundant in the spermosphere of the CON and FAC genotypes (Table S2), which showed specific flavonoid accumulation profiles (Fig. 5A). Compared to the other genotypes, CON showed higher accumulation of flavan-3-ols, while FAC showed higher accumulation of anthocyanidins, flavonols, flavones, isoflavones, and aurones. Correlation analysis indicated that *Bacillus* spp. were negatively correlated with flavan-3-ols (including a putative catechin, p89) and anthocyanidins (including a putative malvidin-3-O-galactoside, p1858) (Fig. 4C and 5B). Catechin and malvidin-3-O-galactoside showed differential accumulation in different genotypes of common bean (Fig. 5B) and were negatively correlated with four and six *Bacillus* species, respectively (Fig. 5C). Experimental assay revealed a dose-dependent effect of catechin on *Bacillus megaterium* (B_ASV14) cultured in TSB with 10% DMSO. Indeed at 0.1% catechin, growth rate increased; at 0.3% catechin, growth slowed down; and at 0.5% catechin, growth was completely inhibited (Fig. 5D).

3. DISCUSSION (1170 words)

The seed spermosphere acts as a crucial micro-environment surrounding seeds, displaying a specific microbial and chemical composition favourable for germination *sensu stricto* (5, 6). This environment may promote resilience against external biotic and abiotic stressors, facilitating the successful seedling establishment (5, 6, 29). To our knowledge, our study represents the first comprehensive characterization of both the chemical and microbial diversity of common bean spermosphere using an untargeted metabolomic and metabarcoding approaches.

Our work enabled the detection and characterization of an extensive range of metabolites, substantially enriching the limited existing data on the chemical diversity of the common bean spermosphere (16). Their exudation during germination raises questions about their functions. These metabolites were categorized in fourteen major metabolic categories (Table S1), which were also characterised in dry common bean

seeds (20). Primary and specialised metabolites are well known compounds regulated by abiotic and biotic environments (30) and have been shown to influence microbiota in other plant compartments such as the rhizosphere when they are exuded (31, 32). These compounds included flavonoids (33), terpenoids (34), cinnamic acids (35), lipids (36), plant hormones (37) and alkaloids (38). Moreover, other compounds such as proteins (39, 40) and small peptides (5-50 aa) (24, 25) are also exuded by seeds during germination but not explored here. We hypothesize that these compounds play an important role in shaping the microbiome in the spermosphere during the early phases of germination.

As for microbial community, the most abundant identified bacterial genera (i.e. *Pantoea*, *Leclercia*, *Bacillus*, *Paenibacillus*, *Escherichia*, *Pseudomonas*; see Table S8) and fungal genera (i.e. *Alternaria*, *Stemphylium*, *Gibberella*, *Botrytis*, *Cladosporium*, *Cryptococcus*, *Monographella*, *Rhizoctonia*; see Table S9) in the common bean spermosphere (Fig. 1C and G) have been previously identified in mature seeds (41), spermosphere and young seedlings (17) of various plant species. These microorganisms may be potentially transmitted from dry seeds to the seedling, likely facilitated by their ability to establish and persist in the spermosphere environment (3, 29). Several abundant microbial genera identified in the spermosphere have been previously reported to promote seed germination and seedling growth (e.g., *Bacillus*, *Pantoea*, *Pseudomonas*) (42–44), while others are known to be associated with damping-off diseases in various plant species (e.g., *Alternaria*, *Botrytis*, *Stemphylium*) (45–48). However, the functional roles of microorganisms should be interpreted with caution, as both beneficial and pathogenic traits are context-dependent rather than inherent properties (49). Nonetheless, the high abundance of these genera may reflect their adaptation to the common bean spermosphere, potentially driven by a copiotrophic lifestyle and rapid growth in response to seed exudates (50, 51).

Our results highlighted a significant impact of genotype on both the chemical and microbial composition of the spermosphere (Fig. 3A, C, E). The high genetic variability in common bean, shaped by multiple domestication events and long-term genome evolution across diverse geographical regions (52, 53), may

explain its strong genotype influence on chemical composition (20). In contrast, species with a shorter domestication and evolutionary history, such as rapeseed, show a metabolome more strongly influenced by the environment (54–56). This genetic diversity likely drives genotype-dependent variation in spermosphere chemistry, which, along with other factors, may contribute to the significant genotype effect on microbial composition, as previously observed in the rhizosphere (57–59). However, fungal communities were also and mainly affected by location of seed production in comparison to bacterial communities (Fig.3C, E). This result could be explained by the relatively stable dynamics of fungal communities given their longer turnover times leading to fewer modifications of the community (60) and/or may be the possibility of an "endemism" effect of fungal communities based on the production area. In contrast, the genotype-dependent signature of the bacterial communities during germination could be a result of their faster multiplication and growth rate adapting to the seed exuded compounds in the spermosphere (61). Taken together, these results suggested a shift from predominantly environmental influences on dry seed microbial diversity (9) to a combined effect of genotype and environment in the spermosphere, possibly driven by seed exudate-mediated selection (3, 5, 6). This process may help conserve and transfer seed-associated microbial assemblages across generations (17, 62, 63).

Our multi-scale correlation network analysis revealed strong associations between *Bacillus* spp. and *Cladosporium* spp. with many central and specialised metabolites. Specifically, the negative correlation between flavonoids and *Bacillus* spp., known for their PGP (Plant Growth Promoting) bacteria properties (64), could be due to the composition of common bean seed coats. Specifically, CON and FAC genotypes, showing high accumulation of putative catechin (flavan-3-ol) and malvidin-galactoside (anthocyanidin), respectively, were the genotypes showing the lowest abundance of *Bacillus* ssp. (Fig. 5A) (Fig. S1). In this light, prior reports highlighted the antimicrobial properties of phenylpropanoids against *Bacillus* spp. (65). In addition, catechin has been shown to inhibit the growth of *B. subtilis* and *B. cereus* while anthocyanin rich extract demonstrated antimicrobial properties against *B. cereus* in dose dependant manner (66–68).

Similarly, our experimental works confirmed a dose-dependent hormetic effect of catechin on *Bacillus megaterium* (B_ASV14), where low concentrations (0.1%) stimulated growth rate, while higher doses (0.5%) caused complete growth inhibition (Fig. 5D) as demonstrated with other flavonoids on different biological models (69). These findings reinforce the potential role of genotype-specific flavonoids in shaping *Bacillus* spp. abundance in the spermosphere. Although catechin was used as a representative flavonoid in our *in vitro* assay, it is likely that other structurally related compounds may exert similar effects on microbial growth. Future studies should evaluate a broader set of flavonoids to establish the generalizability of these effects on microbial community structure.

Due to their Plant Growth-Promoting (PGP) properties, many studies have concentrated on using *Bacillus* spp. as seed coating or biopriming (70, 71). However, limited information is available about the factors contributing to their assembly in the spermosphere and promoting their transmission to the seedlings. Further research is needed to clarify the impact of these metabolites on *Bacillus* abundance, notably if flavonoids have a direct effect in shaping microbial communities during germination (5, 6), or indirect if these metabolites affected broader microbial communities, which could, in turn, impact *Bacillus* abundance through competitive interactions, niche exclusion, or microbial signalling (33). Additionally, the multi-scale correlation network analysis suggested a positive association between *Bacillus* and amino acids and small peptides, potentially indicating their role as a carbon source (72). While these findings highlight key metabolic interactions, other correlations observed in the network warrant further investigation to uncover their ecological significance.

Our results support the original hypothesis that germinating seeds, through genotype-specific exudate profiles, play a central role in shaping the spermosphere by modulating both its metabolic composition and the assembly of associated microbial communities. While the approach used in this work does not fully replicate natural soil-microbe interactions in the spermosphere (3, 5, 6), the use of seed-dwelling microbes and water (24) ensured repeatability and enabled more precise metabolite characterization.

Additionally, by eliminating soil-related confounders, this method improved the detection of microbes present in seed exudates (73). Moreover, pooling exudates at different germination time-points allowed capturing fluctuating metabolites and microbes and a more exhaustive characterization of seed exudates and microbes. Additionally, LC-MS/MS untargeted metabolomics and molecular networking improved metabolite annotation (74, 75) and, coupled to metabarcoding microbiome characterization, generate an original dataset that allowed the identification of many correlations between microorganisms and metabolites that need to be further explored (Fig.S7,S8; Table S10). Further analysis of this novel dataset and alternative multiscale strategies, or incorporating other network metrics like centrality measures, could provide deeper insights into these relationships (Table S11) (76–78). Overall, this study opens new avenues for further research into the metabolic capabilities and ecological roles of seed microbiota within the spermosphere to seed health and seed vigour.

4. MATERIAL and METHODS (1359 words)

4.1. Biological materials

Seeds were harvested in the year 2020 from eight different common bean genotypes (French commercial cultivars of *Phaseolus vulgaris* L.) cultivated in two different locations from the FNAMS field experimental stations: Gers in South-West of France (43°57'25.2"N 0°23'31.7"E) and Maine et Loire (47°28'13.9"N 0°23'40.3"W) in the Loire Valley. The common bean genotypes namely, VEZ, CON, CAP, FAC, FLA, LIN, DEE and VAN were chosen to represent the highest diversity among breeding programmes and qualitative seed traits in France by the seed company Vilmorin-Mikado SAS (Limagrain group). All seeds used in our experiment were from the same post-harvest age. The seeds underwent micro-cleaning, and a purity test was performed by GEVES (the French official seed testing organization) and eventually stored under identical conditions in accordance with the International Seed Testing Association (ISTA)

recommendations (24). Seeds exhibit contrasting phenotypes in terms of size, weight, coat colour and germination kinetics as shown in Figure S1 and S2.

4.2. Spermosphere production and collection

The experimental design for collection of the spermosphere was as illustrated in Figure S2 according to previously developed protocols (24) with modifications as detailed in Supplementary Methods 1.1 .

4.3. Untargeted LC-MS/MS metabolomic analysis

Metabolite extraction was performed according to previously established protocols (79, 80). Similarly, data acquisition, processing and annotations were also performed according to previously developed protocol with modifications as detailed in Supplementary Methods 1.2 (27).

4.4. Metabarcoding and Microbiome analysis

Microbial DNA extraction, library construction, data acquisition by sequencing and data processing were adapted from a previous report with modifications as detailed in Supplementary Methods 1.3 (8).

4.5. Statistical analysis

4.5.1. Statistical analysis and computational analyses of polar and semi-polar specialised

All statistical analyses were conducted using R (v.4.1.1). The metabolomic dataset, consisting of 48 samples, was normalized and then log₁₀-transformed. A linear model was fitted to the log-transformed data to assess the effects of Genotype, Location, and their Interaction (G x L) on metabolite accumulation. Two-way ANOVA (Analysis of variance) was performed to evaluate for significant effects of Genotype, Location and their Interaction (G x L). To address multiple testing, the p-values from the ANOVA were adjusted using the false discovery rate (FDR) correction according to the Benjamini-Hochberg method. This controlled for Type I errors across multiple tests. The adjusted p-values for the interaction, genotype, and location effects were reported in the final analysis.

Principal Component Analysis (PCA) was conducted on the normalised metabolomic dataset comprising of 48 samples. The results were visualized in PCA plots, where different genotypes were color-coded, and locations were represented by different shapes. Heatmaps and hierarchical clustering analyses of metabolomic data were carried out using the pheatmap R package (v.1.0.12) (<https://cran.r-project.org/package=pheatmap>) (81). The heatmaps were built using the following parameters: Euclidean distance measure and Ward clustering method.

4.5.2. Microbial community analysis

Prior to statistical analysis, rarefaction was performed on the metabarcoding data, resulting in 42 bacterial samples (with 6 samples lost) and 48 fungal samples (Fig. S3). Analyses of microbial diversity were then carried out with the R package “phyloseq” (v.1.38.0) (82). Microbial richness was assessed using the number of amplicon sequence variants (ASVs) per sample. Analysis of variance (ANOVA) was performed to assess the significant influences of genotype, location, and their interaction on microbial richness. To further examine the significant differences in microbial richness attributed to genotype within each location, post-hoc tests were conducted using Tukey's Honestly Significant Difference (HSD) method. Changes in community composition were assessed using Bray-Curtis dissimilarity. The relative contributions of plant genotypes and production region in community composition was estimated with permutational multivariate analysis of variance (PERMANOVA) (83) through the function `adonis2` of the R package `vegan` v.2.6.4 (84). All figures were prepared using `ggplot2` v.3.4.3 (85) and the data management was done using `dplyr` v.1.1.3 and `tidyverse` v.2.0.0. Heatmaps were generated using the `plot_taxa_heatmap` function available in the package `microbiomeutilities` v.1.00.17 (86) from the microbial abundance counts data.

4.6. Integration of metabolome and microbiome analysis

Since both metabolomic and metabarcoding analyses were performed on the same biological samples, we applied a correlation-based multiscale integration approach to explore potential associations between metabolites and microbial communities in the spermosphere (28, 87). While correlation analysis can reveal significant associations, it does not establish causation (88). However, multi-omics integration is a powerful tool for generating hypotheses about potential biological interactions and underlying mechanisms (89).

4.6.1. Correlation-based approaches

To integrate the metabolomic and microbiome results, the two normalised datasets were combined and standardized by scaling, which involved centring each variable to have a mean of 0 and scaling them to have a standard deviation of 1. This step removed the influence of scale differences among variables, enabling meaningful comparisons and analyses. Finally, Spearman correlation coefficients were computed on the standardized dataset to assess relationships between variables while accounting for potential non-linear associations. All p-values in the Spearman's correlation were adjusted by Benjamini and Hochberg using the false discovery rate (FDR) control procedure. We focused on microbial-microbial and microorganism-metabolite correlations ($p < 0.05$) to investigate community dynamics and the influence of seed exuded metabolites on microbial assembly, while metabolite-metabolite correlations were omitted as they were not the primary focus of our analysis (Table S10). Consequently, only microbial-microbial and microorganism-metabolite pairs were retained if their correlation coefficient was $\geq |0.5|$ which were considered moderate to strong associations (Table S10) (90, 91).

4.6.2. Network analysis and enrichment analysis

A pairwise correlation network was visualized from microbial-microbial and microorganism-metabolite pairs of correlation coefficient $\geq |0.5|$ using Cytoscape (v.3.10.0) (92). Metabolic subnetworks were

extracted from the previously created pairwise correlation network comprising of both positive and negative correlations. From the initial network, annotated metabolic nodes of each major metabolic categories were selected along with their direct correlated microorganism node (first neighbours). Enrichment analysis using hypergeometric test was carried out by comparing microbial genera in the different metabolic category correlation subnetworks to those present in the spermosphere. For enrichment analysis, reference sets were bacterial and fungal ASVs identified in spermosphere respectively, and the target sets were bacterial and fungal ASVs in the respective metabolic categories' subnetworks. Network analysis was also carried out using the Cytoscape built-in plugin Network Analyser to calculate node degree, a key metric representing the number of connections (edges) each node has within the network (92). Node degree was chosen as it directly quantifies interaction frequency, allowing us to identify highly connected microbial taxa and metabolites that may play key roles in spermosphere dynamics. Additional network metrics were also calculated with Cytoscape and are provided in the supplementary material, although they were not specifically explored in this study (Table S11).

4.7. *In-vitro* experimental investigation of the effects of catechin on *Bacillus megaterium*

Bacillus megaterium strain B_ASV14 were isolated on Tryptic Soy Agar (TSA) 10% strength from same common bean seed lot used in the study (CON produced in Maine et Loire) and registered in the CIRM-CFBP French Collection for Plant-Associated Bacteria under the accession numbers CFBP9575. The strain was first grown on Tryptic Soy Broth (TSB) agar plates (Fisher BioReagents BP2471-100) and incubated at 28°C for 24 hours. A single colony was then inoculated into 10 mL of liquid TSB medium and incubated under agitation. After 12 hours, the optical density (OD₆₀₀) of the culture was measured using an Eppendorf Biophotometer Plus with square cuvettes. Cultures with an OD₆₀₀ between 0.4 and 0.8 were diluted to an OD₆₀₀ of 0.01, then transferred to a 96-well microplate for growth assays. Experimental conditions included a control (TSB containing 10% DMSO, CAS 67-68-5, without catechin), and catechin treatments (CAS 225937-10-0) at final concentrations of 0.1%, 0.3%, and 0.5% (w/v), each prepared in TSB

containing 10% DMSO to ensure identical solvent conditions across treatments. Each condition was tested in six technical replicates. The microplate was incubated at 28°C for 72 hours with continuous shaking, and absorbance at 600 nm was recorded every 20 minutes using a TECAN Spark spectrophotometer (93, 94).

4.8. Data availability statement

The data that support the findings of this study are openly available in online repositories. Metabolomic LC-MS/MS data are available on MassIVE at <https://massive.ucsd.edu/>, reference number MSV000094600 and the metabarcoding data are available in the European Nucleotide Archive (ENA) at <https://www.ebi.ac.uk/ena/>, reference number PRJEB59579.

5. ACKNOWLEDGEMENTS

This work was also part of the 3rd Programme for Future Investments (France2030) and operated by the SUCSEED project (ANR-20-PCPA-0009) funded by the ‘Growing and Protecting crops Differently’ French Priority Research Program (PPR-CPA), part of the national investment plan operated by the French National Research Agency (ANR). This work has benefited from the support of IJPB's Plant Observatory platform PO-Chem (Metabolomic platform) and PO-Bioch (biochemistry platform). The IJPB benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007).

6. REFERENCES

1. Basu S, Groot SPC. 2023. Seed Vigour and Invigoration, p. 67–89. *In* Seed Science and Technology: Biology, Production, Quality. Springer.
2. Rajjou L, Duval M, Gallardo K, Catusse J, Bally J, Job C, Job D. 2012. Seed Germination and Vigor. *Annu Rev Plant Biol* 63:507–533.

3. Nelson EB. 2018. The seed microbiome: Origins, interactions, and impacts. *Plant Soil* 422:7–34.
4. Verona O. 1958. [The spermosphere]. *Ann Inst Pasteur* 95:795–798.
5. Nelson EB. 2004. Microbial dynamics and interactions in the spermosphere. *Annu Rev Phytopathol* 42:271–309.
6. Schiltz S, Gaillard I, Pawlicki-Jullian N, Thiombiano B, Mesnard F, Gontier E. 2015. A review: what is the spermosphere and how can it be studied? *J Appl Microbiol* 119:1467–1481.
7. Abdelfattah A, Tack AJM, Lobato C, Wassermann B, Berg G. 2023. From seed to seed: the role of microbial inheritance in the assembly of the plant microbiome. *Trends Microbiol* 31:346–355.
8. Rochefort A, Briand M, Marais C, Wagner M-H, Laperche A, Vallée P, Barret M, Sarniguet A. 2019. Influence of Environment and Host Plant Genotype on the Structure and Diversity of the Brassica napus Seed Microbiota. *Phytobiomes J* 3:326–336.
9. Klaedtke S, Jacques M-A, Raggi L, Prévieux A, Bonneau S, Negri V, Chable V, Barret M. 2016. Terroir is a key driver of seed-associated microbial assemblages. *Environ Microbiol* 18:1792–1804.
10. Liang J, Gao G, Zhong R, Liu B, Christensen MJ, Ju Y, Zhang W, Li Y, Li C, Zhang X, Nan Z. 2023. Effect of *Epichloë gansuensis* Endophyte on Seed-Borne Microbes and Seed Metabolites in *Achnatherum inebrians*. *Microbiol Spectr* 11:e01350-22.
11. Yu P, He X, Baer M, Beirinckx S, Tian T, Moya YAT, Zhang X, Deichmann M, Frey FP, Bresgen V, Li C, Razavi BS, Schaaf G, von Wirén N, Su Z, Bucher M, Tsuda K, Goormachtig S, Chen X, Hochholdinger F. 2021. Plant flavones enrich rhizosphere Oxalobacteraceae to improve maize performance under nitrogen deprivation. *Nat Plants* 7:481–499.

12. Guo Z-Y, Kong C-H, Wang J-G, Wang Y-F. 2011. Rhizosphere isoflavones (daidzein and genistein) levels and their relation to the microbial community structure of mono-cropped soybean soil in field and controlled conditions. *Soil Biol Biochem* 43:2257–2264.
13. Okutani F, Hamamoto S, Aoki Y, Nakayasu M, Nihei N, Nishimura T, Yazaki K, Sugiyama A. 2020. Rhizosphere modelling reveals spatiotemporal distribution of daidzein shaping soybean rhizosphere bacterial community. *Plant Cell Environ* 43:1036–1046.
14. He D, Singh SK, Peng L, Kaushal R, Vílchez JI, Shao C, Wu X, Zheng S, Morcillo RJL, Paré PW, Zhang H. 2022. Flavonoid-attracted *Aeromonas* sp. from the *Arabidopsis* root microbiome enhances plant dehydration resistance. *ISME J* 16:2622–2632.
15. Avezum L, Rondet E, Mestres C, Achir N, Madode Y, Gibert O, Lefevre C, Hemery Y, Verdeil J-L, Rajjou L. 2023. Improving the nutritional quality of pulses via germination. *Food Rev Int* 39:6011–6044.
16. Martins SJ, Medeiros FHV, Lakshmanan V, Bais HP. 2018. Impact of Seed Exudates on Growth and Biofilm Formation of *Bacillus amyloliquefaciens* ALB629 in Common Bean. *Front Microbiol* 8:2631.
17. Johnston-Monje D, Gutiérrez JP, Lopez-Lavalle LAB. 2021. Seed-Transmitted Bacteria and Fungi Dominate Juvenile Plant Microbiomes. *Front Microbiol* 12.
18. Barret M, Briand M, Bonneau S, Prévieux A, Valière S, Bouchez O, Hunault G, Simoneau P, Jacques M-A. 2015. Emergence Shapes the Structure of the Seed Microbiota. *Appl Environ Microbiol* 81:1257–1266.
19. Zhang H, Yasmin F, Song B-H. 2019. Neglected treasures in the wild — legume wild relatives in food security and human health. *Curr Opin Plant Biol* 49:17–26.
20. Mecha E, Erny GL, Guerreiro ACL, Feliciano RP, Barbosa I, Bento da Silva A, Leitão ST, Veloso MM, Rubiales D, Rodriguez-Mateos A, Figueira ME, Vaz Patto MC, Bronze MR. 2022. Metabolomics profile

responses to changing environments in a common bean (*Phaseolus vulgaris* L.) germplasm collection. Food Chem 370:131003.

21. Soldan R, Fusi M, Cardinale M, Homma F, Santos LG, Wenzl P, Bach-Pages M, Bitocchi E, Chacon Sanchez MI, Daffonchio D, Preston GM. 2024. Consistent effects of independent domestication events on the plant microbiota. Curr Biol 34:557-567.e4.

22. Sotelo A, Sousa H, Sánchez M. 1995. Comparative study of the chemical composition of wild and cultivated beans (*Phaseolus vulgaris*). Plant Foods Hum Nutr 47:93–100.

23. Pang Z, Chen J, Wang T, Gao C, Li Z, Guo L, Xu J, Cheng Y. 2021. Linking Plant Secondary Metabolites and Plant Microbiomes: A Review. Front Plant Sci 12:621276.

24. Saccaram C, Brosse C, Collet B, Sourdeval D, François T, Bernay B, Corso M, Rajjou L. 2024. A mass spectrometry-based peptidomic dataset of the spermosphere in common bean (*Phaseolus vulgaris* L.) seeds. Sci Data 11:1202.

25. Saccaram C, Bernay B, Brosse C, Collet B, Peng S, Sourdeval D, François T, Corso M, Rajjou L. 2025. Development of an extraction method for the identification of peptides in the spermosphere of common bean (*Phaseolus vulgaris* L.). Seed Sci Res 1–11.

26. Olivon F, Elie N, Grelier G, Roussi F, Litaudon M, Touboul D. 2018. MetGem Software for the Generation of Molecular Networks Based on the t-SNE Algorithm. Anal Chem 90:13900–13908.

27. Boutet S, Barreda L, Perreau F, Totozafy J-C, Mauve C, Gakière B, Delannoy E, Martin-Magniette M-L, Monti A, Lepiniec L, Zanetti F, Corso M. 2022. Untargeted metabolomic analyses reveal the diversity and plasticity of the specialized metabolome in seeds of different *Camelina sativa* genotypes. Plant J Cell Mol Biol 110:147–165.

28. Li X, Lin Y, Li X, Xu X, Zhao Y, Xu L, Gao Y, Li Y, Tan Y, Qian P, Huang H. 2020. Tyrosine supplement ameliorates murine aGVHD by modulation of gut microbiome and metabolome. *EBioMedicine* 61.
29. Joubert O, Arnault G, Barret M, Simonin M. 2025. Sowing success: ecological insights into seedling microbial colonisation for robust plant microbiota engineering. *Trends Plant Sci* 30:21–34.
30. Barreda L, Brosse C, Boutet S, Perreau F, Rajjou L, Lepiniec L, Corso M. 2024. Specialized metabolite modifications in Brassicaceae seeds and plants: diversity, functions and related enzymes. *Nat Prod Rep* 41:834–859.
31. Bai B, Liu W, Qiu X, Zhang J, Zhang J, Bai Y. 2022. The root microbiome: Community assembly and its contributions to plant fitness. *J Integr Plant Biol* 64:230–243.
32. Sasse J, Martinoia E, Northen T. 2018. Feed Your Friends: Do Plant Exudates Shape the Root Microbiome? *Trends Plant Sci* 23:25–41.
33. Wang L, Chen M, Lam P-Y, Dini-Andreote F, Dai L, Wei Z. 2022. Multifaceted roles of flavonoids mediating plant-microbe interactions. *Microbiome* 10:233.
34. Murphy KM, Edwards J, Louie KB, Bowen BP, Sundaresan V, Northen TR, Zerbe P. 2021. Bioactive diterpenoids impact the composition of the root-associated microbiome in maize (*Zea mays*). *Sci Rep* 11:333.
35. Zhelnina K, Louie KB, Hao Z, Mansoori N, Da Rocha UN, Shi S, Cho H, Karaoz U, Loqué D, Bowen BP, Firestone MK, Northen TR, Brodie EL. 2018. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat Microbiol* 3:470–480.
36. Jiang Y, Wang W, Xie Q, Liu N, Liu L, Wang D, Zhang X, Yang C, Chen X, Tang D, Wang E. 2017. Plants transfer lipids to sustain colonization by mutualistic mycorrhizal and parasitic fungi. *Science* 356:1172–1175.

37. Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald M, Malfatti S, Glavina del Rio T, Jones CD, Tringe SG, Dangl JL. 2015. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* 349:860–864.
38. Lei F, Liu X, Huang H, Fu S, Zou K, Zeng J, Liu H. 2021. The *Macleaya cordata* Symbiont: Revealing the Effects of Plant Niches and Alkaloids on the Bacterial Community. *Front Microbiol* 12.
39. Rocha RO, Morais JKS, Oliveira JTA, Oliveira HD, Sousa DOB, Souza CEA, Moreno FB, Monteiro-Moreira ACO, Antonino de Souza Júnior JD, Grossi de Sá MF, Vasconcelos IM. 2015. Proteome of Soybean Seed Exudates Contains Plant Defense-Related Proteins Active against the Root-Knot Nematode *Meloidogyne incognita*. *J Agric Food Chem* 63:5335–5343.
40. Rose TL, Conceição A da S, Xavier-Filho J, Okorokov LA, Fernandes KVS, Marty F, Marty-Mazars D, Carvalho AO, Gomes VM. 2006. Defense Proteins from *Vigna Unguiculata* Seed Exudates: Characterization and Inhibitory Activity Against *Fusarium Oxysporum*. *Plant Soil* 286:181–191.
41. Simonin M, Briand M, Chesneau G, Rochefort A, Marais C, Sarniguet A, Barret M. 2022. Seed microbiota revealed by a large-scale meta-analysis including 50 plant species. *New Phytol* 234:1448–1463.
42. Dai Y, Li X, Wang Y, Li C, He Y, Lin H, Wang T, Ma X. 2020. The differences and overlaps in the seed-resident microbiome of four Leguminous and three Gramineous forages. *Microb Biotechnol* 13:1461–1476.
43. Garcia-Lemos AM, Großkinsky DK, Saleem Akhtar S, Nicolaisen MH, Roitsch T, Nybroe O, Veierskov B. 2020. Identification of Root-Associated Bacteria That Influence Plant Physiology, Increase Seed Germination, or Promote Growth of the Christmas Tree Species *Abies nordmanniana*. *Front Microbiol* 11:566613.

44. Widnyana IK, Javandira C. 2016. Activities *Pseudomonas* spp. and *Bacillus* sp. to Stimulate Germination and Seedling Growth of Tomato Plants. *Agric Agric Sci Procedia* 9:419–423.
45. Sorrentino R, Lahoz E, Battaglia V, Sorrentino MC, Cerrato D. 2022. First Report of Seedling Damping-Off of industrial Hemp (*Cannabis sativa* L.) caused by seed-transmitted *Alternaria rosae* in Italy. *Plant Dis* <https://doi.org/10.1094/PDIS-06-22-1375-PDN>.
46. Kurose D, Hoang LH, Furuya N, Takeshita M, Sato T, Tsushima S, Tsuchiya K. 2014. Pathogenicity of *Stemphylium lycopersici* isolated from rotted tobacco seeds on seedlings and leaves. *J Gen Plant Pathol* 80:147–152.
47. Ikeda S, Yasuoka S. 2023. Damping-off of onion: a new symptom caused by soilborne *Botrytis byssoides*. *J Gen Plant Pathol* 89:260–265.
48. Halo BA, Al-Yahyai ,Rashid A., and Al-Sadi AM. 2021. Biological control of *Pythium aphanidermatum*-induced cucumber and radish damping-off by an endophytic fungus, *Cladosporium omanense* isolate 31R. *Biocontrol Sci Technol* 31:235–251.
49. Newton AC, Fitt BDL, Atkins SD, Walters DR, Daniell TJ. 2010. Pathogenesis, parasitism and mutualism in the trophic space of microbe–plant interactions. *Trends Microbiol* 18:365–373.
50. Torres-Cortés G, Bonneau S, Bouchez O, Genthon C, Briand M, Jacques M-A, Barret M. 2018. Functional Microbial Features Driving Community Assembly During Seed Germination and Emergence. *Front Plant Sci* 9:902.
51. Barret M, Briand M, Bonneau S, Prévieux A, Valière S, Bouchez O, Hunault G, Simoneau P, Jacques M-A. 2015. Emergence Shapes the Structure of the Seed Microbiota. *Appl Environ Microbiol* 81:1257–1266.

52. Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu S, Song Q, Chavarro C, Torres-Torres M, Geffroy V, Moghaddam SM, Gao D, Abernathy B, Barry K, Blair M, Brick MA, Chovatia M, Gepts P, Goodstein DM, Gonzales M, Hellsten U, Hyten DL, Jia G, Kelly JD, Kudrna D, Lee R, Richard MMS, Miklas PN, Osorno JM, Rodrigues J, Thareau V, Urrea CA, Wang M, Yu Y, Zhang M, Wing RA, Cregan PB, Rokhsar DS, Jackson SA. 2014. A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet* 46:707–713.
53. Cortinovis G, Vincenzi L, Anderson R, Marturano G, Marsh JI, Bayer PE, Rocchetti L, Frascarelli G, Lanzavecchia G, Pieri A, Benazzo A, Bellucci E, Di Vittori V, Nanni L, Ferreira Fernández JJ, Rossato M, Aguilar OM, Morrell PL, Rodriguez M, Gioia T, Neumann K, Alvarez Diaz JC, Gratias A, Klopp C, Bitocchi E, Geffroy V, Delledonne M, Edwards D, Papa R. 2024. Adaptive gene loss in the common bean pan-genome during range expansion and domestication. *Nat Commun* 15:6698.
54. Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Corrêa M, Da Silva C, Just J, Falentin C, Koh CS, Le Clainche I, Bernard M, Bento P, Noel B, Labadie K, Alberti A, Charles M, Arnaud D, Guo H, Daviaud C, Alamery S, Jabbari K, Zhao M, Edger PP, Chelaifa H, Tack D, Lassalle G, Mestiri I, Schnel N, Le Paslier M-C, Fan G, Renault V, Bayer PE, Golicz AA, Manoli S, Lee T-H, Thi VHD, Chalabi S, Hu Q, Fan C, Tollenaere R, Lu Y, Battail C, Shen J, Sidebottom CHD, Wang X, Canaguier A, Chauveau A, Bérard A, Deniot G, Guan M, Liu Z, Sun F, Lim YP, Lyons E, Town CD, Bancroft I, Wang X, Meng J, Ma J, Pires JC, King GJ, Brunel D, Delourme R, Renard M, Aury J-M, Adams KL, Batley J, Snowdon RJ, Tost J, Edwards D, Zhou Y, Hua W, Sharpe AG, Paterson AH, Guan C, Wincker P. 2014. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345:950–953.
55. Bennouna D, Avise J-C, Rosique C, Svilar L, Pontet C, Trouverie J, Fine F, Pinochet X, Fraser K, Martin J-C. 2019. The impact of genetics and environment on the polar fraction metabolome of commercial *Brassica napus* seeds: a multi-site study. *Seed Sci Res* 29:167–178.

56. Hu J, Chen B, Zhao J, Zhang F, Xie T, Xu K, Gao G, Yan G, Li H, Li L, Ji G, An H, Li H, Huang Q, Zhang M, Wu J, Song W, Zhang X, Luo Y, Chris Pires J, Batley J, Tian S, Wu X. 2022. Genomic selection and genetic architecture of agronomic traits during modern rapeseed breeding. *Nat Genet* 54:694–704.
57. Pacheco-Moreno A, Bollmann-Giolai A, Chandra G, Brett P, Davies J, Thornton O, Poole P, Ramachandran V, Brown JKM, Nicholson P, Ridout C, DeVos S, Malone JG. 2024. The genotype of barley cultivars influences multiple aspects of their associated microbiota via differential root exudate secretion. *PLOS Biol* 22:e3002232.
58. Anderson HM, Cagle GA, Majumder EL-W, Silva E, Dawson J, Simon P, Freedman ZB. 2024. Root exudation and rhizosphere microbial assembly are influenced by novel plant trait diversity in carrot genotypes. *Soil Biol Biochem* 197:109516.
59. Pérez-Jaramillo JE, Mendes R, Raaijmakers JM. 2016. Impact of plant domestication on rhizosphere microbiome assembly and functions. *Plant Mol Biol* 90:635–644.
60. Rousk J, Bååth E. 2011. Growth of saprotrophic fungi and bacteria in soil. *FEMS Microbiol Ecol* 78:17–30.
61. Rousk J, Bååth E. 2007. Fungal and bacterial growth in soil with plant materials of different C/N ratios. *FEMS Microbiol Ecol* 62:258–267.
62. Hardoim PR, Hardoim CCP, Overbeek LS van, Elsas JD van. 2012. Dynamics of Seed-Borne Rice Endophytes on Early Plant Growth Stages. *PLOS ONE* 7:e30438.
63. Johnston-Monje D, Raizada MN. 2011. Conservation and Diversity of Seed Associated Endophytes in Zea across Boundaries of Evolution, Ethnography and Ecology. *PLOS ONE* 6:e20396.
64. Radhakrishnan R, Hashem A, Abd_Allah EF. 2017. *Bacillus*: A Biological Tool for Crop Improvement through Bio-Molecular Changes in Adverse Environments. *Front Physiol* 8.

65. Petropoulos SA, Taofiq O, Fernandes Â, Tzortzakis N, Ciric A, Sokovic M, Barros L, Ferreira IC. 2019. Bioactive properties of greenhouse-cultivated green beans (*Phaseolus vulgaris* L.) under biostimulants and water-stress effect. *J Sci Food Agric* 99:6049–6059.
66. Fathima A, Rao JR. 2016. Selective toxicity of Catechin—a natural flavonoid towards bacteria. *Appl Microbiol Biotechnol* 100:6395–6402.
67. Friedman M, Henika PR, Levin CE, Mandrell RE, Kozukue N. 2006. Antimicrobial Activities of Tea Catechins and Theaflavins and Tea Extracts against *Bacillus cereus*†. *J Food Prot* 69:354–361.
68. Trikas ED, Melidou M, Papi RM, Zachariadis GA, Kyriakidis DA. 2016. Extraction, separation and identification of anthocyanins from red wine by-product and their biological activities. *J Funct Foods* 25:548–558.
69. Calabrese EJ, Pressman P, Hayes AW, Dhawan G, Kapoor R, Agathokleous E, Manes P, Calabrese V. 2023. Naringin commonly acts via hormesis. *Sci Total Environ* 896:164728.
70. Junges E, Brião Muniz MF, Bastos B de O, Oruoski P. 2016. Biopriming in bean seeds. *Acta Agric Scand Sect B — Soil Plant Sci* 66:207–214.
71. Miljaković D, Marinković J, Tamindžić G, Đorđević V, Tintor B, Milošević D, Ignjatov M, Nikolić Z. 2022. Bio-Priming of Soybean with *Bradyrhizobium japonicum* and *Bacillus megaterium*: Strategy to Improve Seed Germination and the Initial Seedling Growth. *Plants Basel Switz* 11:1927.
72. Moe LA. 2013. Amino acids in the rhizosphere: From plants to microbes. *Am J Bot* 100:1692–1705.
73. Oburger E, Jones DL. 2018. Sampling root exudates – Mission impossible? *Rhizosphere* 6:116–133.
74. Barreda L, Boutet S, De Vos D, Boulard C, Grain D, Lepiniec L, Corso M. 2025. Specialized metabolome and transcriptome atlas of developing *Arabidopsis thaliana* seed under warm temperatures. *Sci Data* 12:306.

75. Barreda L, Brosse C, Boutet S, Klewko N, De Vos D, Francois T, Collet B, Grain D, Boulard C, Totozafy JC, Bernay B, Perreau F, Lepiniec L, Rajjou L, Corso M. 2025. Multi-omic analyses unveil contrasting composition and spatial distribution of specialized metabolites in seeds of *Camelina sativa* and other Brassicaceae. *Plant J* 121:e17231.
76. Agamah FE, Bayjanov JR, Niehues A, Njoku KF, Skelton M, Mazandu GK, Ederveen THA, Mulder N, Chimusa ER, 't Hoen PAC. 2022. Computational approaches for network-based integrative multi-omics analysis. *Front Mol Biosci* 9.
77. Misra BB, Langefeld C, Olivier M, Cox LA. 2019. Integrated omics: tools, advances and future approaches <https://doi.org/10.1530/JME-18-0055>.
78. Picard M, Scott-Boyer M-P, Bodein A, Périn O, Droit A. 2021. Integration strategies of multi-omics data for machine learning analysis. *Comput Struct Biotechnol J* 19:3735–3746.
79. Routaboul J-M, Kerhoas L, Debeaujon I, Pourcel L, Caboche M, Einhorn J, Lepiniec L. 2006. Flavonoid diversity and biosynthesis in seed of *Arabidopsis thaliana*. *Planta* 224:96–107.
80. Liang W, Zamarreño ÁM, Torres-Montilla S, Torre A de la, Totozafy JC, Kaji T, Ueda M, Corso M, García-Mina JM, Solano R, Chini A. 2024. Conjugation of dn-OPDA with amino acids inhibits its hormonal bioactivity in *Marchantia polymorpha*. *bioRxiv* <https://doi.org/10.1101/2024.08.27.609933>.
81. Kolde R. 2019. pheatmap: Pretty Heatmaps. R package version 1.0. 12.
82. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217.
83. Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26:32–46.

84. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P. 2019. *vegan*. Community Ecol Package 2:5–6.
85. Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis* 2nd ed. 2016. Springer International Publishing : Imprint: Springer, Cham.
86. Shetty SA, Lahti L. 2018. *microbiomeutilities*: An R package for utilities to guide in-depth marker gene amplicon data analysis. Zenodo DOI 10.
87. Theriot CM, Koenigsknecht MJ, Carlson PE, Hatton GE, Nelson AM, Li B, Huffnagle GB, Z. Li J, Young VB. 2014. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to *Clostridium difficile* infection. *Nat Commun* 5:3114.
88. Gershman SJ, Ullman TD. 2023. Causal implicatures from correlational statements. *PLOS ONE* 18:e0286067.
89. Odriozola I, Rasmussen JA, Gilbert MTP, Limborg MT, Alberdi A. 2024. A practical introduction to *holo-omics*. *Cell Rep Methods* 4:100820.
90. Mukaka MM. 2012. A guide to appropriate use of Correlation coefficient in medical research. 3. *Malawi Med J* 24:69–71.
91. Schober P, Boer C, Schwarte LA. 2018. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analg* 126:1763.
92. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang P-L, Lotia S, Pico AR, Bader GD, Ideker T. 2012. A travel guide to Cytoscape plugins. *Nat Methods* 9:1069–1076.
93. Huang AC, Jiang T, Liu Y-X, Bai Y-C, Reed J, Qu B, Goossens A, Nützmann H-W, Bai Y, Osbourn A. 2019. A specialized metabolic network selectively modulates *Arabidopsis* root microbiota. *Science* 364:eaau6389.

94. Ghitti E, Rolli E, Vergani L, Borin S. 2024. Flavonoids influence key rhizocompetence traits for early root colonization and PCB degradation potential of *Paraburkholderia xenovorans* LB400. *Front Plant Sci* 15.

7. Tables

Node	Group	Species/Metabolites	Genus/Category	Degree
B_ASV99	Bacteria	<i>Pantoea ananatis</i>	<i>Pantoea</i>	428
B_ASV19	Bacteria	<i>Bacillus megaterium</i>	<i>Bacillus</i>	397
B_ASV308	Bacteria	Unclassified Actinobacteria	Unclassified <i>Actinobacteria</i>	308
B_ASV162	Bacteria	Unclassified Firmicute	Unclassified Firmicute	285
B_ASV27	Bacteria	<i>Bacillus megaterium</i>	<i>Bacillus</i>	263
B_ASV132	Bacteria	<i>Paenibacillus sp.</i>	<i>Paenibacillus</i>	251
B_ASV10	Bacteria	<i>Bacillus sp.</i>	<i>Bacillus</i>	240
B_ASV91	Bacteria	<i>Pantoea sp.</i>	<i>Pantoea</i>	222
B_ASV14	Bacteria	<i>Bacillus megaterium</i>	<i>Bacillus</i>	217
B_ASV40	Bacteria	<i>Bacillus toyonensis</i>	<i>Bacillus</i>	215
F_ASV3	Fungi	<i>Botrytis fabae</i>	<i>Botrytis</i>	156
F_ASV2	Fungi	<i>Stemphylium herbarum</i>	<i>Stemphylium</i>	155
F_ASV27	Fungi	<i>Gibellulopsis nigrescens</i>	<i>Gibellulopsis</i>	129
F_ASV13	Fungi	Unclassified Dothideomycetes	Unclassified Dothideomycetes	124
F_ASV9	Fungi	<i>Alternaria infectoria</i>	<i>Alternaria</i>	121

Table 1: Network analysis of pairwise metabolite-microorganism correlation network showing top 15 nodes by degree. Node: connection point of the network. It can be either the unique ID of a metabolite or microbial species. Group: specifying the type of variable represented by the node (bacteria, fungi or metabolites). Species/Metabolites: Identity at the level of species for microbial nodes and identity obtained through comparison with metabolic libraries for metabolite nodes. Genus/Category: Identity at the level of genus for microbial nodes and metabolic categories for metabolite nodes. Degree: Number of connections or edges the node has to other nodes.

Microbial Genus	Target size	Target Occurrence	Target Frequency	Reference size	Reference Occurrence	Reference frequency	Fold Enrichment	P-Value
<u>Enriched in Amino acids and derivatives Subnetwork</u>								
<i>Bacillus</i>	39	13	33.3%	118	24	20.3%	1.6	0.0146
<i>Variovorax</i>	39	3	7.7%	118	3	2.5%	3.0	0.0342
<i>Alternaria</i>	23	4	17.4%	82	6	7.3%	2.4	0.0492
<u>Enriched in Carbohydrates Subnetwork</u>								
<i>Bacillus</i>	24	10	41.7%	118	24	20.3%	2.0	0.0062
<u>Enriched in Cinnamic acids and derivatives Subnetwork</u>								
<i>Cladosporium</i>	15	3	20.0%	82	5	6.1%	3.3	0.0403
<u>Enriched in Coumarins Subnetwork</u>								
<i>Cladosporium</i>	11	4	36.4%	82	5	6.1%	6.0	0.0009
<i>Gibberella</i>	11	2	18.2%	82	3	3.7%	5.0	0.0460
<u>Enriched in Flavonoids Subnetwork</u>								
<i>Bacillus</i>	34	13	38.2%	118	24	20.3%	1.9	0.0031
<i>Cladosporium</i>	24	4	16.7%	82	5	6.1%	2.7	0.0241
<u>Enriched in Lignans Subnetwork</u>								
<i>Pantoea</i>	6	3	50.0%	118	9	7.6%	6.6	0.0056
<u>Enriched in Lipids and derivatives Subnetwork</u>								
<i>Variovorax</i>	9	3	33.3%	118	3	2.5%	13.1	0.0150
<u>Enriched in Organic acid Subnetwork</u>								
<i>Pantoea</i>	8	3	37.5%	118	9	7.6%	4.9	0.0144
<u>Enriched in Terpenoids and derivatives Subnetwork</u>								
<i>Bacillus</i>	29	10	34.5%	118	24	20.3%	1.7	0.0312
<i>Cladosporium</i>	15	3	20.0%	82	5	6.1%	3.3	0.0403

Table 2: Enrichment analysis of microbial genus overrepresented in correlation subnetwork of different metabolic categories as compared to spermosphere. Microbial genus: Enriched microbial genus in specific metabolic category correlation subnetwork. Target size: number of species belonging to specific microbial kingdom present in correlation metabolic subnetwork (Target size > 5 were retained). Target occurrence: Number of species belonging to overrepresented microbial genus in correlation metabolic subnetwork. Target frequency: Frequency of overrepresented genus in correlation metabolic subnetwork. Reference size: total number of species belonging to specific microbial kingdom present in spermosphere. Reference occurrence: Number of species belonging to overrepresented microbial genus in spermosphere. Reference frequency: Frequency of overrepresented genus in spermosphere. Fold enrichment: Enrichment score (Score > 1.5 were retained). P-value: statistical significance of microbial enrichment test (hypergeometric test).

9. Figures

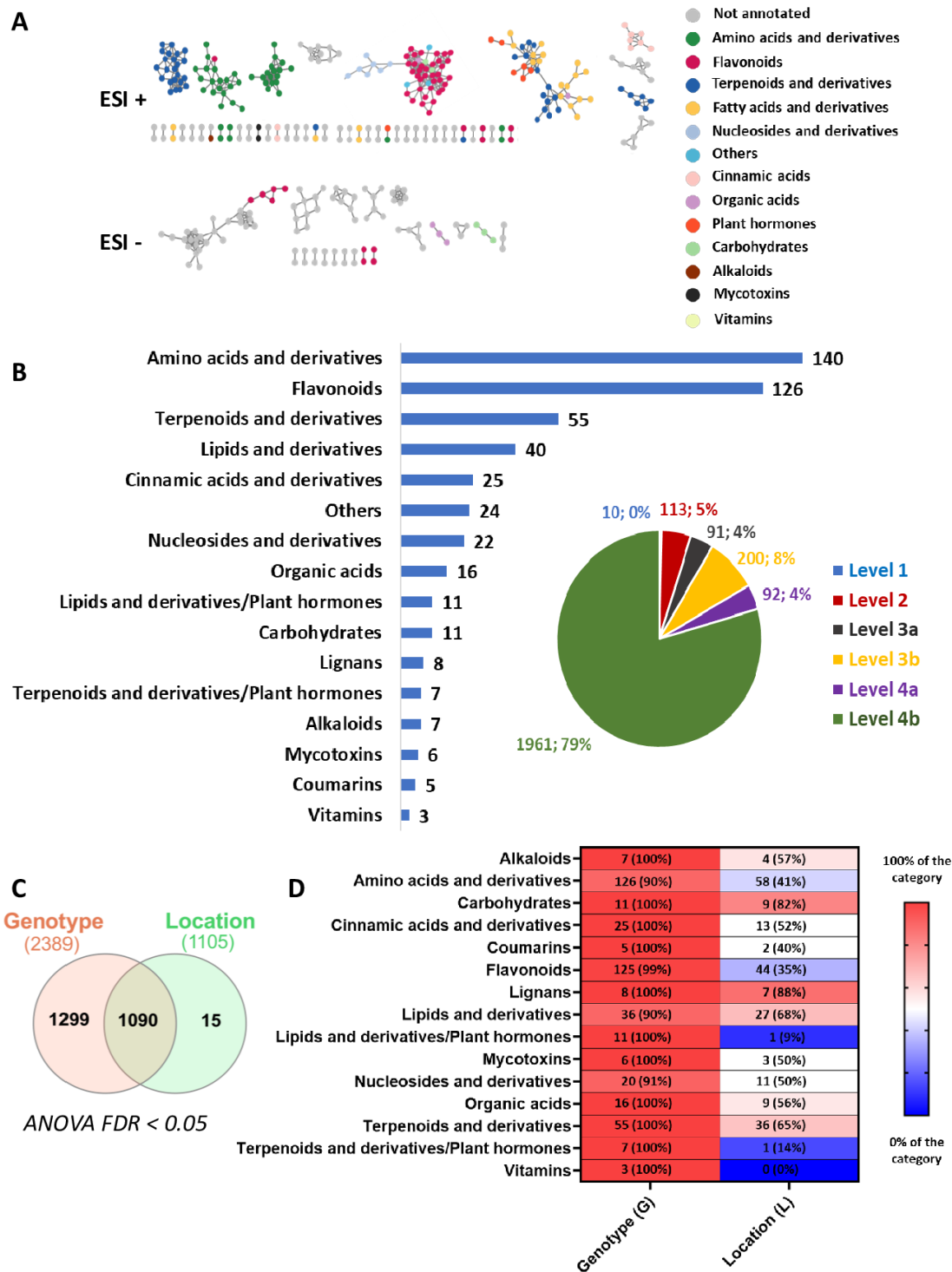


Fig. 1: Chemical diversity in the spermosphere of germinating seeds from 8 genotypes of common beans produced in two different locations. Analysis was performed on a dataset of 48 samples from 8 common bean genotypes grown in two locations (Gers and Maine et Loire). **(A)** A subset of the molecular networks is shown for metabolite analysis carried out in positive (ESI+) and negative (ESI-) ionization modes. Different colours correspond to distinct metabolic categories. Metabolites are grouped based on their chemical structures. Cosine similarity scores of 0.8 and 0.7 were used for ESI+ and ESI-, respectively. *Plant hormones contains two categories: lipid-derived hormones and terpenoid-derived hormones. For representation purposes they were given the same colors in the molecular network. **(B)** Histogram showing the annotated/known metabolic categories identified in common bean seeds. The number of metabolites that belong to each category is indicated. Percentage of annotated metabolites by confidence levels (Level 1-4) are also represented in the sector diagram. Metabolites were annotated by comparing with GNPS public spectral libraries as well as internal homemade libraries. Further putative annotations were carried out by molecular networks. **(C)** Venn diagram of differential metabolite features (FDR adjusted $p < 0,05$) identified by two-way ANOVA on log10 transformed dataset containing 2,467 metabolic features across 48 samples. **(D)** Annotated differential metabolites identified by two-way ANOVA (FDR adjusted $p < 0,05$) on log10 transformed dataset.

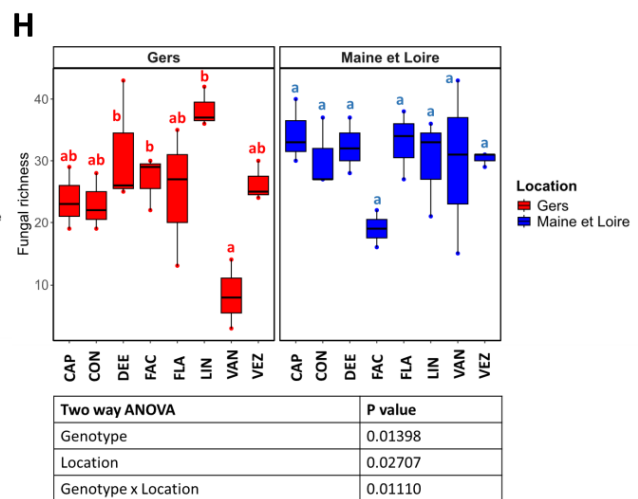
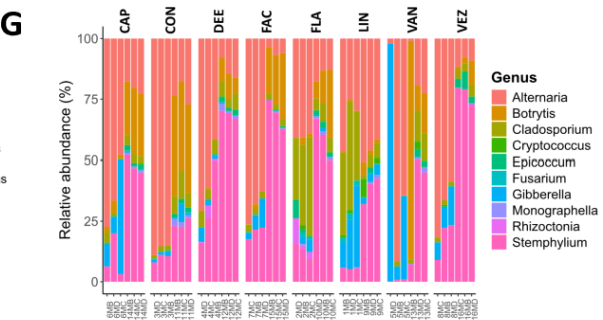
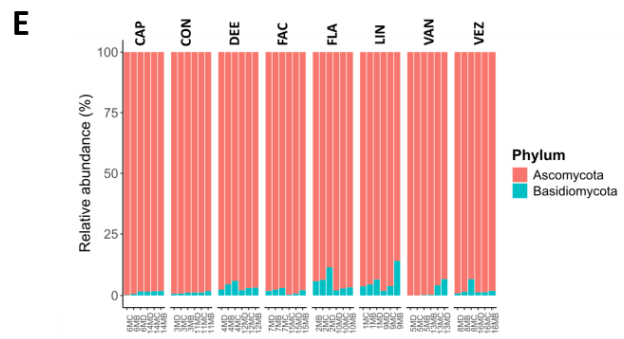
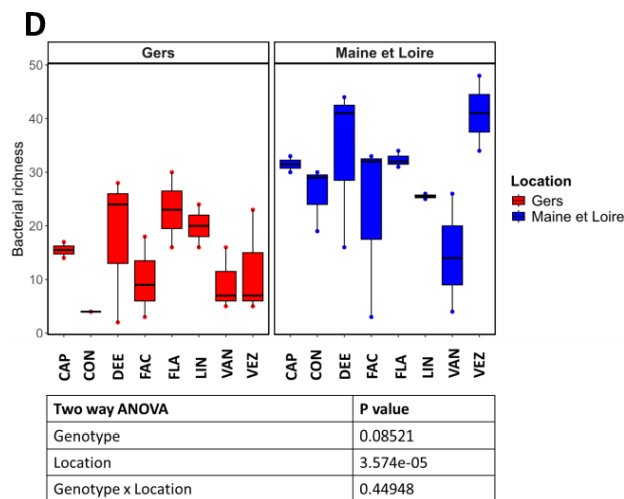
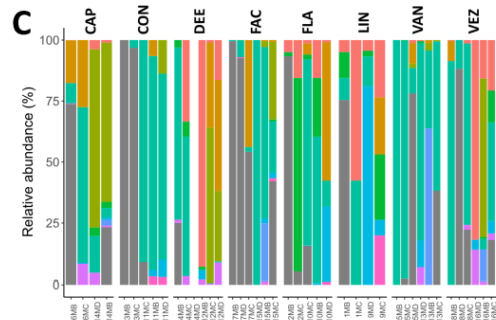
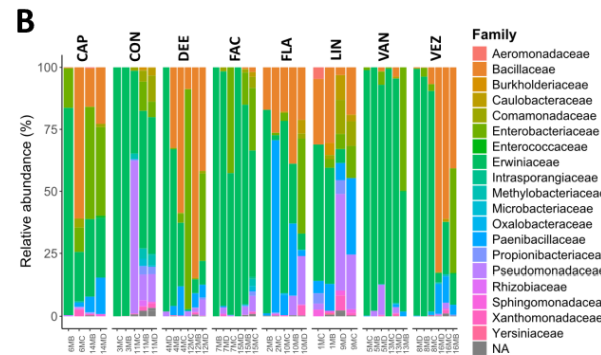
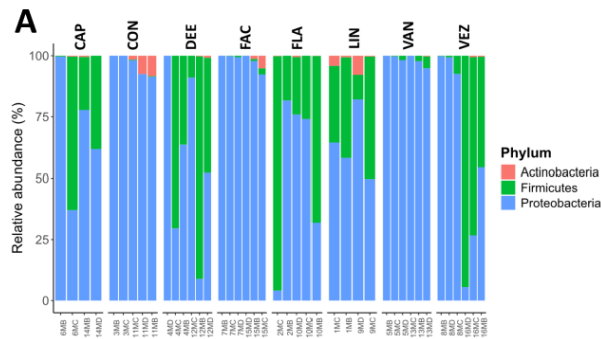
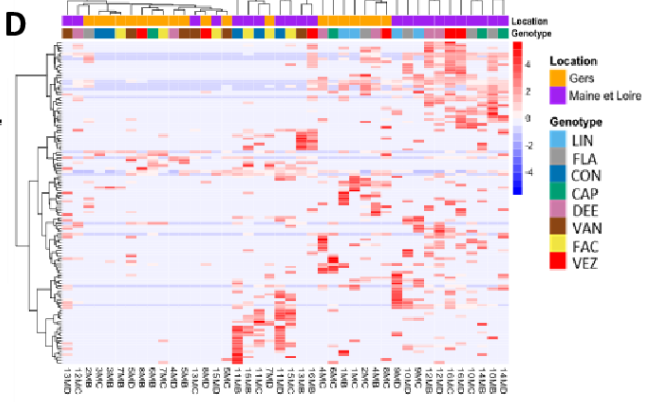
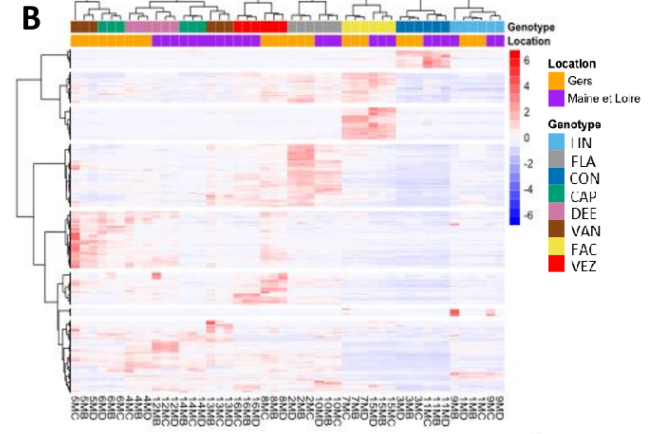
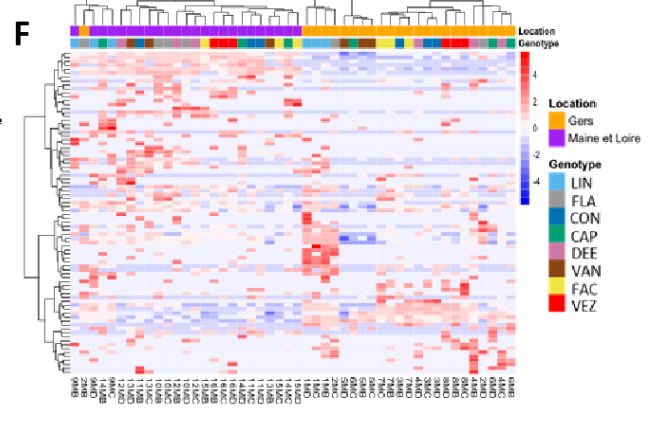


Fig. 2: Microbial diversity in the spermosphere of germinating seeds from 8 genotypes of common beans produced in two different locations. Analyses were conducted on rarefied datasets: *gyrB* (bacteria) at 1250 reads per sample (resulting in 42 samples) and ITS (fungi) at 5000 reads per sample (retaining all 48 samples). **(A)** and **(E)** Relative abundance of bacterial and fungal phyla respectively. **(B)** and **(F)** Relative abundance of bacterial and fungal families respectively. **(C)** and **(G)** Relative abundance of ten most represented bacterial and fungal genus respectively. **(D)** and **(H)** Measure of bacterial and fungal α -diversity through observed richness (number of predicted species) produced in Gers (red box plots) and Maine & Loire (blue box plots). Box plots display the data distribution through their quartiles and whiskers are used to indicate variability outside the upper and lower quartiles. The median marks the midpoint of the data and is indicated by the black line.



Page 10 of 10



PERMANOVA	Factors	Metabolome		Bacterial communities		Fungal communities	
		Variance R ²	Significance	Variance R ²	Significance	Variance R ²	Significance
	Genotype	47.95%	***	22.9%	***	18.7%	***
	Location	11.57%	***	8.9%	***	43.1%	***
	Genotypex Location	22.63%	***	21.3%	**	18.2%	***
Significance code: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$, NS $p \geq 0.1$							

Fig. 3: Factors influencing microbial structure and metabolomic landscape in common bean spermosphere. Analyses were conducted on rarefied datasets: *gyrB* at 1250 reads per sample (resulting in 42 samples) and ITS at 5000 reads per sample (retaining all 48 samples), alongside metabolomic data from 48 samples. **(A)** Principal Component Analysis (PCA) scores plot of original normalised data of spermosphere metabolome distribution of 8 common bean seed genotypes (represented by different colors) produced in two different locations (represented by different shapes). **(B)** Heatmap of relative abundance of Mf identified in the spermosphere clustered by Ward method. Each row represents a distinct metabolite feature and each column represents a distinct sample. The color intensity in each cell corresponds to the relative abundance of the respective Mf within the sample, with darker shades indicating higher abundance. **(C and E)** Bacterial and fungal community structure (β -diversity) respectively. Genotypes are represented by different colors and location of seed production by shapes. The variation in bacterial and fungal community structures explained by each PCoA axis is given in parentheses. **(D and F)** Heatmap of relative abundance of bacterial and fungal species in the spermosphere clustered by Ward method respectively. Each row in the heatmap represents a distinct microbial species, while each column represents an individual sample. The color intensity in each cell corresponds to the relative abundance of the respective microbial species within the sample, with darker shades indicating higher abundance. **(G)** Permutational multivariate analysis of variance (PERMANOVA) showing the effects of genotype, location, and their interaction (G x L) on metabolome, bacterial, and fungal community structures. R^2 values represent the proportion of explained variance; asterisks denote significance ($p < 0.05$).

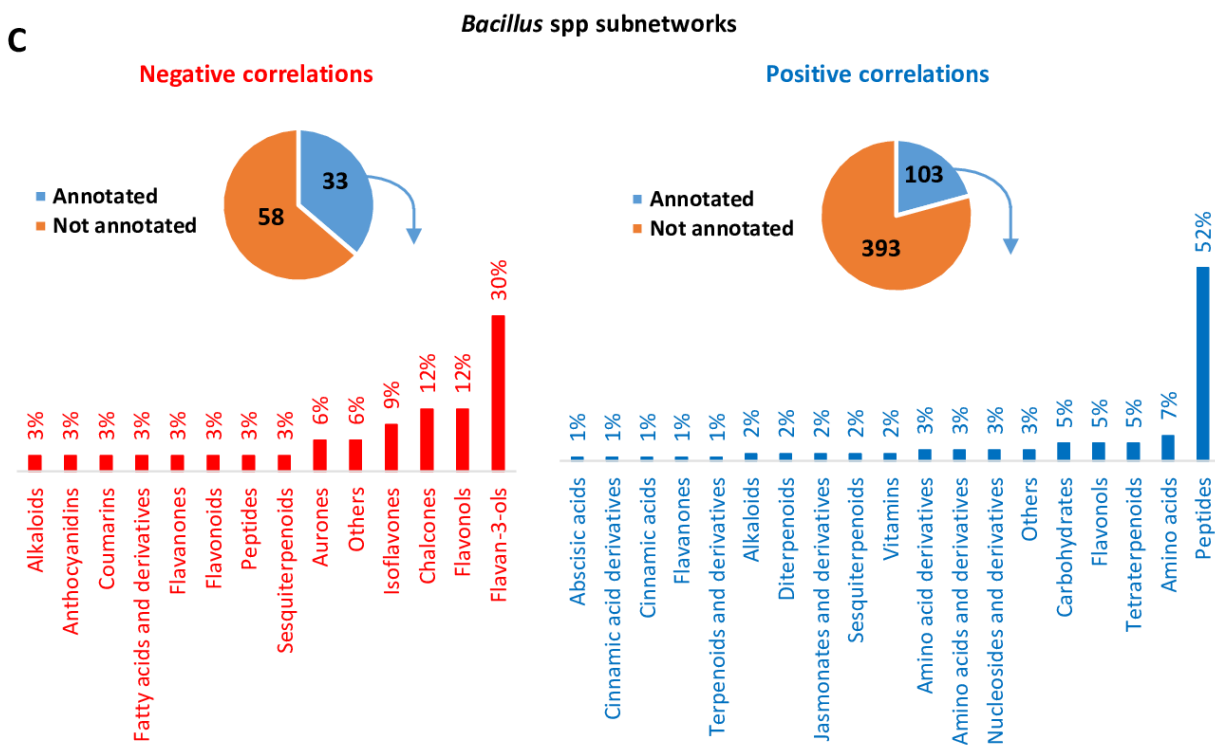
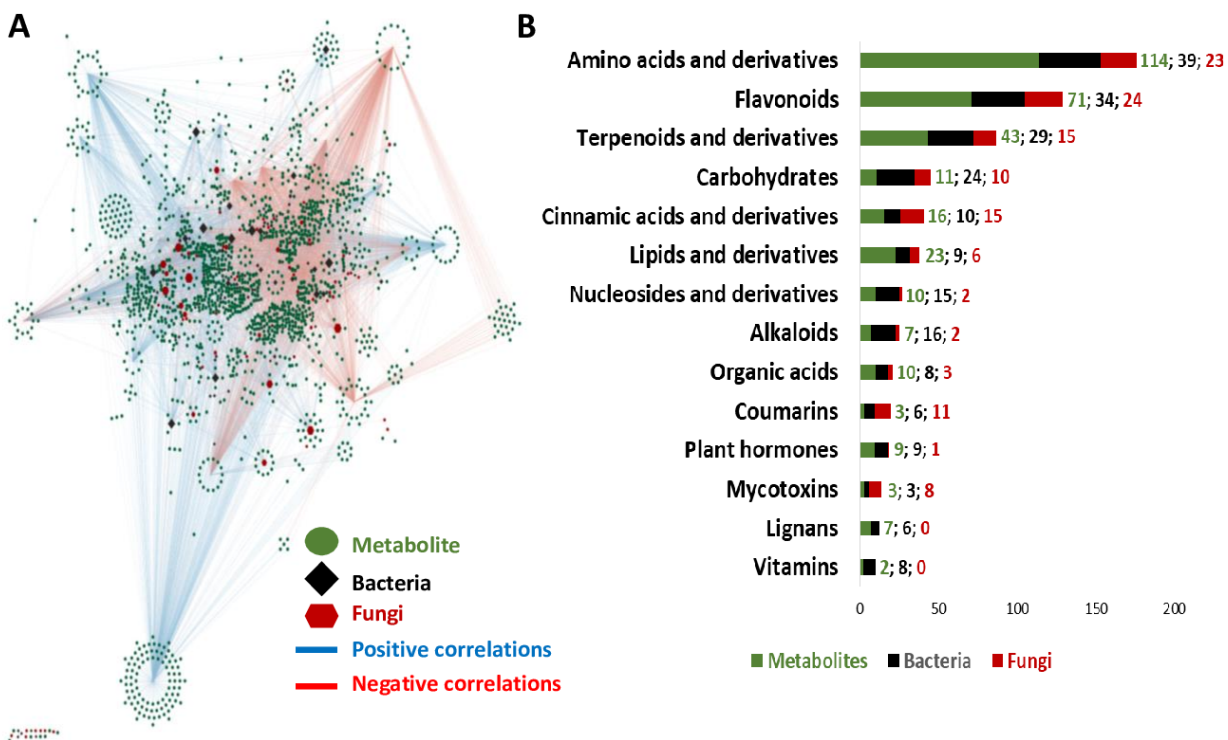


Fig. 4: Metabolite–microorganism correlation network analysis. (A) Correlation network based on Spearman correlation ($r > |0.5|$, FDR p-value < 0.05) containing 1736 nodes and 6225 edges (Table S10). Only pairwise correlation microorganism-microorganism and metabolite-microorganism retained. Nodes in the network represents the metabolites (green circles), bacteria (black diamonds) and fungi (brown octagons). The size of each nodes in the network is proportional to their importance in terms of direct first neighbours they are connected to (degree). Positive correlations are represented by red edges and inversely blue edges represents negative correlations. **(B)** Size of metabolic category subnetwork in terms of number of total nodes (metabolites in green, bacteria in black and fungi in brown). Metabolic subnetwork was extracted from whole network and only microbial nodes directly correlated to metabolic nodes (first neighbors) were selected. **(C)** *Bacillus* spp. subnetwork analysis revealed negative and positive correlations with annotated metabolites. Network based on Spearman correlation ($r > |0.5|$, $p < 0.05$). Positive and negative correlation network are represented by orange and blue, respectively. *Bacillus* spp. positive correlation subnetwork consisted of 496 metabolites of which 103 were assigned to a metabolic class. *Bacillus* spp. negative correlation subnetwork consisted of 91 metabolites of which 33 were assigned to a metabolic class. Only the metabolites that have been assigned to a metabolic class are represented as a proportion in the graphic.

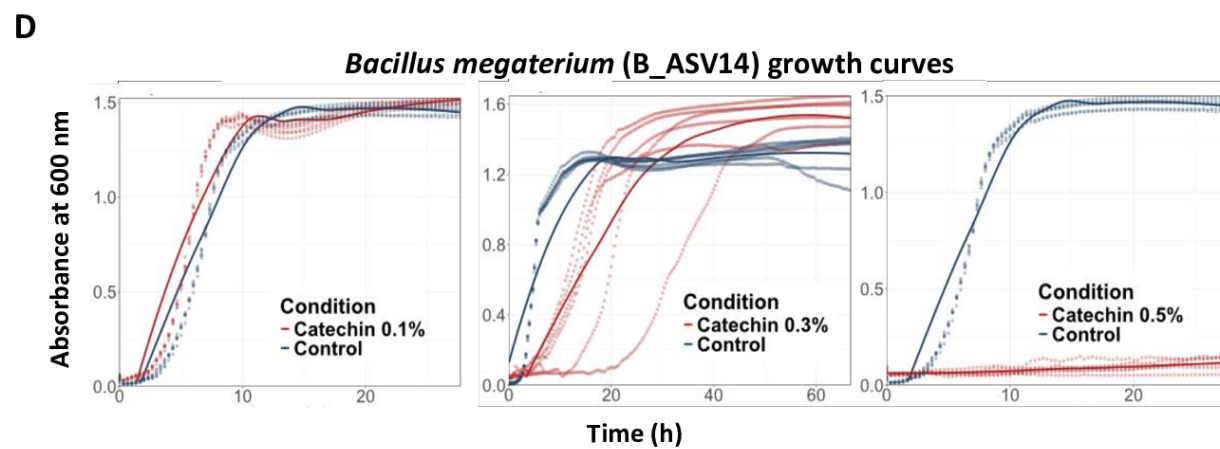
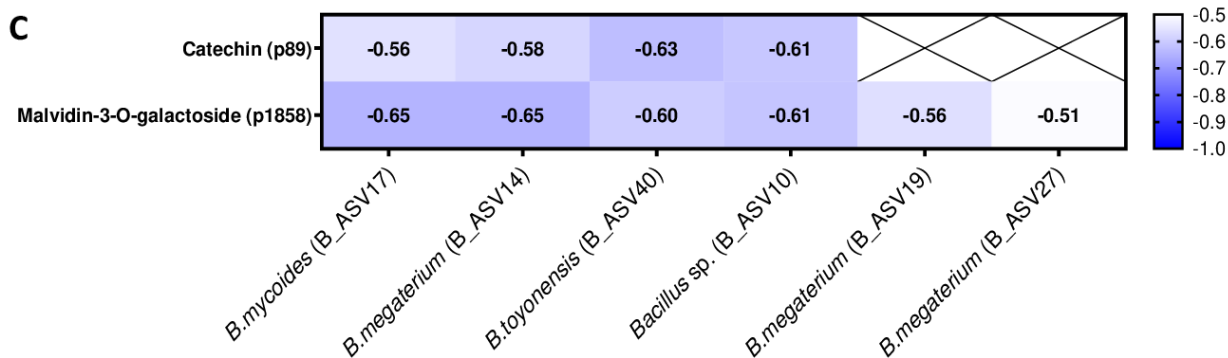
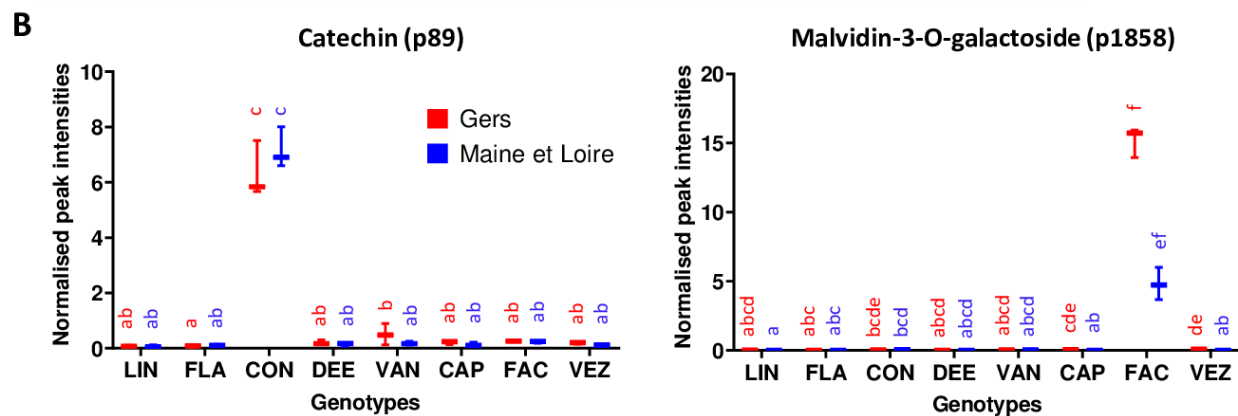
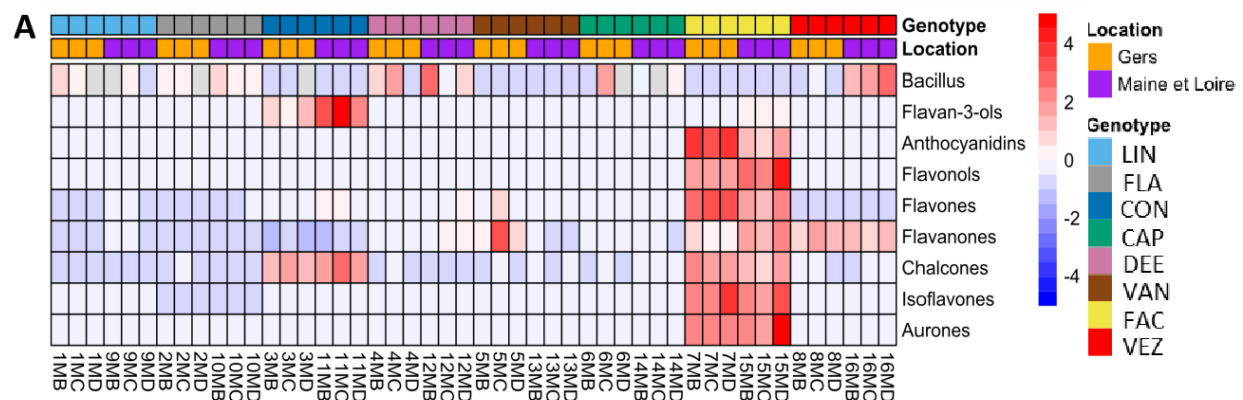


Fig. 5: Relationship between flavonoids and *Bacillus* spp. **(A)** Heatmap of relative abundance of *Bacillus* spp. and different flavonoid classes in the spermosphere of eight common bean genotypes produced in two locations. **(B)** Differential abundance of catechin (p89) and malvidin-3-O-galactoside (p1858) in the common bean spermosphere samples. Post-hoc Sidak test ($p < 0.01$) were carried out to assess differences in catechin and malvidin-3-O-galactoside composition among genotypes and locations. **(C)** Correlation analysis heatmap between flavonoids (catechin and malvidin-3-O-galactoside) and *Bacillus* spp. Spearman correlation coefficient of each metabolite-microorganisms pair shown on heatmap. Catechin and malvidin-3-O-galactoside are negatively correlated to four and six *Bacillus* species respectively. **(D)** *In vitro* growth assay was conducted using the *Bacillus megaterium* strain (B_ASV14), isolated from common bean seed samples, to assess its response to catechin. Growth curves illustrate the response of the strain to commercial catechin at concentrations of 0.1%, 0.3%, and 0.5% (red), compared to a control condition with no catechin (blue). Each condition included six technical replicates, represented by dotted lines. The smooth solid curves correspond to LOESS-fitted average trends across these replicates for each condition. Absorbance was measured at 600 nm over time.