

Single seed microbiota: assembly and transmission from parent plant to seedling

Guillaume Chesneau¹, Béatrice Laroche², Anne Prévieux¹, Coralie Marais¹, Martial Briand¹, Brice Marolleau¹, Marie Simonin¹, Matthieu Barret^{1*}

¹Univ Angers, Institut Agro, INRAE, IRHS, SFR QUASAV, F-49000 Angers, France

²Université Paris-Saclay, INRAE, MaIAGE, 78350 Jouy-en-Josas, France

*email : matthieu.barret@inrae.fr

Abstract: Gaining basic understanding of processes involved in seed microbiota assembly is a prerequisite for improving crop establishment. Investigation of microbiota structure during seed development revealed that individual seeds of bean and radish were associated with a dominant bacterial taxon representing more than 75% of all reads. The identity of this taxon was highly variable between plants and within seeds of the same plant. Succession of dominant taxa occurred during seed filling and maturation through *Selection*. In a second step, we evaluated seed to seedling transmission of these dominant seed-borne taxa. We showed that initial bacterial abundance on seeds was not a good predictor of seedling transmission and that the identity of seed-borne taxa can impact seedling phenotype. Altogether this work unveiled that seeds are colonized by few bacterial taxa of highly variable identity, which appears to be important for the early stages of plant development.

Introduction:

Seed microbiota allows the transmission of microbial communities between plant generations¹ and is a key factor influencing both seed vigor and seedling development, two essential steps for crop establishment. Several studies focused on the role of the mature seed microbiota in releasing dormancy², improving germination³ or protecting against damping-off⁴. In addition, seed microbiota may have longer-term consequences on crop establishment by promoting establishment of soil-borne plant-beneficial microorganisms⁵ or limiting the incidence of soil-borne plant pathogens⁶. However, all the above-mentioned studies characterized microbiota of dried mature seeds, while limited knowledge is available on the dynamics of microbiota assembly during seed development⁷. A consequence of this knowledge gap is that the origin, timing of arrival and succession of seed-associated microbial taxa is mostly unknown, as well as their future consequences on seed vigor and seedling development.

To date, the characterization of seed microbiota structure has been performed on about 50 plant species including major crops such as barley⁸, bean⁹, rapeseed¹⁰, rice¹¹, tomato¹² and wheat⁸. *Selection* by the local environment or by the host significantly modulates seed microbiota composition^{9,10,13}. Other ecological processes such as dispersal by pollinators¹⁴ and ecological drift¹⁵ are also important drivers of seed community assembly. However, in most studies seed microbiota was characterized on large seed samples from 10 to 1,000 seeds. One reason for using seed lots rather than individual seeds is related to the technical difficulties of collecting enough microbial DNA from a single seed as these are usually colonized by a low bacterial population size¹⁶. Another rationale for

using seed samples is related to historical studies carried out in seed phytopathology. Indeed, seed-transmitted plant pathogens are generally distributed within a Poisson distribution¹⁷ and increasing sampling size increases the probability of detection. While working on seed lots is valid for improving plant pathogen detection, this sampling procedure artificially increases seed microbiota richness and more importantly does not allow the estimation of seed-to-seed variability.

The few studies carried out at the individual seed level or small seed samples (i.e. less than 10 seeds) through culture-dependent^{16,18} or –independent^{19–21} approaches have revealed a low microbial richness. It has even been proposed that each seed was colonized by none or only one endophytic microorganism¹⁸. The authors of the latter study propose a hypothesis, the “primary symbiont hypothesis”, that endophyte microbial transmission in seed is bottlenecked as a result of plant defenses and interactions among seed-transmitted microorganisms¹⁸. However, this hypothesis has still not been confirmed at the whole individual seed microbiota level, which includes endophyte and epiphytes. In addition, the order and routes of arrival of these “primary symbionts” as well as the ecological processes involved in their assembly remained unexplored.

The presence of a microorganism, on, or in seeds is not a guarantee of transmission to the seedling²². To date studies focusing on the dynamics of microbial communities during germination and emergence have been performed on several subsamples of the same seed lot (e.g : ^{8,12,23–25}). However, in order to truly assess transmission of seed-borne bacterial taxa to the seedling, it is necessary to work at the individual seed-seedling level in a non-destructive way.

In this context, this study aimed to address several fundamental questions on the dynamics of seed microbiota from mother plant to seedling. Firstly, does the order and timing of microbial taxa immigration during seed development impact seed microbiota composition? In other words, are the first microorganisms that colonize the seed during its development, the ones that will be established on mature seeds? What are the main seed transmission pathways employed by these seed-borne taxa? What are the ecological processes involved in seed microbiota assembly? What is the dynamics of seed-borne taxa during seedling emergence? And what is their impact on seed germination and seedling phenotype?

To answer these questions we worked on seeds of common bean (*Phaseolus vulgaris*) and radish (*Raphanus sativus*), two plant species whose seed microbiota has been extensively characterized ^{7,9,15,23}. We estimated microbiota composition of single seeds of different individual plants during seed filling and maturation through amplicon sequencing. We also inferred the origins of seed-associated taxa through sampling and sequencing of flowers, vascular tissues and atmosphere microbiota. In a second step, we evaluated seedling transmission of seed-borne bacteria by sampling, in a non-destructive way, individual seed and seedling microbiota. This second step was performed aseptically in order to specifically focus on the inheritance of the seed microbiota, avoiding influence of the environment. Finally, we assessed the impact of seed microbiota on seedling emergence and confirmed the impact of some selected taxa on seedling fitness. We showed that individual seeds are colonized by one dominant taxon with variable taxonomic identity. These taxa are not necessarily abundant in the future seedling but can have a significant beneficial or detrimental impact on the seedling phenotype, such as *Pantoea agglomerans* on bean seedling.

Results :

Experiment 1 – Culture-based enrichment of seed-borne bacterial strains efficiently captured the diversity of the seed microbiota.

Sequencing of samples with low microbial biomass is subject to a low signal to noise ratio as a result of weak amounts of DNA starting material²⁶. We assessed whether culture-based enrichment of seed-borne bacteria could increase population size without distorting community structure. Firstly, the bacterial diversity obtained after dilution and plating on nine culture media was estimated through amplicon sequencing of two molecular markers that differed in their taxonomic resolution, namely the V4 region of 16S rRNA gene (hereafter referred to as 16S) and a portion of *gyrB*²³. According to Faith's phylogenetic diversity, 1/10 strength Tryptic Soy Agar (hereafter referred to as TSA10) allowed the greatest diversity of bacterial isolates to be obtained (**Fig. S1**), therefore confirming previous results obtained with root-associated bacterial strains²⁷.

In a second step we compared the bacterial community profiles before and after culture-enrichment on TSA10 on 24 bulk seed samples of bean and radish. In total 1,276 ASVs were detected with 45 common to both approaches (**Fig. 1A**) including the most abundant ASVs (**Fig. 1B**). ASVs specifically detected with one of the two approaches were not affiliated to a particular phylogenetic lineage (**Fig. 1C**). Hence, differences in the detection of taxa with or without TSA10 culture-enrichment could be due to differences in amplification of rare taxa. Altogether, these results highlighted that investigating bacterial community assembly is possible with a step of culture-based enrichment and may even allow better detection of rare taxa since 927 ASVs were specifically detected following culture-based enrichment (**Fig. 1A**).

Experiment 2 – Assembly and structure of the microbiota during individual seed development

Percentage of seeds colonized by cultivable bacteria varied between plants.

Seeds from bean and radish were collected aseptically during seed filling and maturation, two stages defined by measuring seed water content (**Fig. 2 A**). As expected the percentage of water content significantly decreased during seed development for both plant species, especially during late maturation. Bacteria were detected on average in 43% to 84% of bean seeds per stage and 30% to 82% of radish seeds per stage (**Fig. 3A**). Variation in the percentage of seed-containing bacteria was associated with the sampling stage for both plant species ($P<0.01$). While no correlation between bean seed development and percentage of seed-containing bacteria was observed, the percentage of seed-containing bacteria gradually increased during radish seed development (**Fig. 3A**). Within each sampling stage, the percentage of detection of seed-borne bacteria was significantly ($P<0.01$) different between individual plants and varied from 25% to 100% per bean plant and 20% to 100% for radish (**Fig. 3A**).

Individual seeds are associated with low bacterial richness

Few bacterial taxa were detected per individual seeds. Indeed, a median richness of eight 16S ASVs in bean and four 16S ASVs in radish were observed (**Fig. 3B**), while a median of nine and three

gyrB ASVs were observed in bean and radish seeds (**Fig. 3C**). Bacterial richness significantly ($P<0.01$) increased during seed development of both plant species regardless of the molecular markers employed. Moreover, within each sampling stage, high variability in richness between seeds was observed (**Fig. 3B and 3C**). For instance, based on 16S, one to 45 ASVs were detected per bean seed at the mature stage (D50, **Fig. 3B**), while one to 51 ASVs were associated with radish seeds at the same developmental stage (D65, **Fig. 3B**). Variation in richness was accompanied by a high variation in taxonomic composition (**Fig. 3D and 3E**). Changes in phylogenetic composition between seed bacterial communities was significantly ($P<0.001$) explained by the seed development stages for bean (16S: 18.3% ; *gyrB* : 16.1%,) and radish (16S: 15.3% ; *gyrB*: 7.4%). In addition, variation in phylogenetic composition was also significantly ($P<0.001$) associated with plant inter-individual variations in bean (16S: 26.5% ; *gyrB*: 26.1%) and radish (16S: 28.5% ; *gyrB*: 20.0%).

A seed is associated with one dominant bacterial taxa

One dominant bacterial taxa was associated with bean and radish seeds. According to 16S (**Fig. 4A**), these dominant taxa had mean relative abundance of 77.1% ($\pm 22.0\%$) and 75.4% ($\pm 22.8\%$) with *gyrB* (**Fig. S2A**). The identity of the dominant taxa was highly variable between seeds of both plant species with 94 16S ASVs and 285 *gyrB* ASVs distributed in more than 15 bacterial orders (**Fig. S3 and S4**).

The phylogenetic diversity of dominant taxa per plant significantly ($P<0.05$) increased during bean seed development (**Fig. 4B**). Moreover, a clear shift in bacterial taxonomic composition was observed during bean seed development with a decrease in the occurrence of dominant taxa affiliated to Enterobacterales (**Fig. 4C and S2C**). Although our destructive sampling does not allow a direct monitoring of microbiota dynamics during seed development, phylogenetic diversity and taxonomic composition indicated succession of dominant taxa over seed filling and maturation of bean. For radish, we did not observed significant change in phylogenetic diversity. However, changes in taxonomic composition of dominant taxa per plant species was observed with an increase in occurrence of Flavobacteriales during the last seed maturation stage (**Fig. 4C**).

Within a sampling stage, a large variation in taxonomic composition was observed between seed bacterial communities from different plants (**Fig. 3D**). The distribution pattern of seed dominant taxa indicated a significant (Chi-squared test, $P<0.01$) association between some dominant taxa and specific plant individuals (**Fig. S3, S4 and Table S1**). For instance at D27 in bean, one dominant taxa affiliated to Erwiniaceae with 16S and *gyrB* was significantly more frequent in plant 1 (**Fig. S3, S4 and Table S1**). Hence the identity of dominant seed-associated taxa appeared to be dependent on the local pool of bacterial taxa found within one specific plant individual.

The origin of seed-associated taxa

To gain more insight on the potential inoculum sources of the seed microbiota, we performed additional sampling of the bacterial fraction of the microbiota after culture enrichment on three distinct habitats, namely flowers, stems and atmosphere (**Fig. 2C**). These three habitats could represent the main sources of the floral, internal and external pathways described for seed

transmission²⁸. Detection of *gyrB* sequence variants was employed in this section as *gyrB* has a better taxonomic resolution (i.e. infraspecific) than 16S.

Observed richness associated with atmosphere, flowers and stems was lower than on seeds for most of the seed developmental stages (**Fig. S5A**). All communities were dominated by bacteria belonging to the Proteobacteria, Actinobacteria and Firmicutes (**Fig. S5B**). The temporal dynamics of the airborne samples over time was explaining 33.7% ($P < 0.001$) of variability in phylogenetic composition (**Fig. S5C**). We also observed significant ($P < 0.05$) variation in phylogenetic composition of bacterial communities associated with bean stem (23.0%) and radish stem (21.4%) over time (**Fig. S5C**). The percentage of seed taxa that were not detected in atmosphere, flowers and stems was at least higher than 80% for both plant species at each sampling time (**Fig. 5**). Overall, we detected more seed-associated ASVs in the atmosphere compared to the other habitats sampled (**Fig. 5**).

We next looked at the origin of the dominant taxa. The most frequent dominant ASV, *Pantoea agglomerans* (30% of all seeds with 289 occurrences, **Fig. S4**), was detected in all sampled habitats for both plant species (**Fig. S6**). Although some taxa specifically associated with one specific plant (**Table S1**) were highly abundant in stem (e.g. ASV5), this observation was not generalized to the whole, therefore suggesting distinct transmission routes (**Fig. S6**).

Dispersion is less important than local seed processes

Since dominant seed-associated taxa were specifically associated with one plant individual, proportions of observed ASVs within each seed were aggregated at the plant level to form a species-proportion distribution. We compared the resulting distribution with the observed species proportions predicted by an adapted neutral model (see Methods). Species-proportion distributions be considered as a first order approximation²⁹ for assessing the relative importance of dispersal over seed local processes (i.e. ecological drift or *Selection*) through the ratio I as well as the size S of the species pool available for these processes (see Methods).

Dispersion was less important than local seed processes in more than 85% of bean and radish plants. For bean, no clear pattern appeared over time but I was less important at D33 than at the following stages (**Fig. 6A and 6C**, $P=0.003$ for 16S, $P=0.063$ for *gyrB*). For radish, I increased significantly ($P=0.012$ for 16S, $P=0.05$ for *gyrB*) between D37 and D65 (**Fig. 6E and 6G**). The species pool size available for these processes was markedly low with median values ranging from 3-6 and 2-5 for bean and 2-6 and 3-4 for radish, for *gyrB* and 16S markers respectively (**Fig. 6B, 6D, 6F and 6H**). These results suggest that the balance between dispersion and local processes is shifted in time.

Selection drive local seed microbiota assembly at the plant level

To estimate the relative importance of *Selection* and ecological drift in the local (e.g. plant level) assembly of the seed microbiota we employed null model or more specifically abundance-based β -null deviation measures³⁰. With respect to the data obtained on bean seeds with 16S and *gyrB*, β -nearest taxon index (β NTI) values negatively deviated from the null expectation (**Fig. 7A and B**). β NTI values less < -2 are considered as significantly different from the null expectation³⁰. Since all β NTI values were below this threshold, this indicated that the phylogenetic composition of bean seed

communities aggregated at the plant level are more similar than expected by chance and that this similarity could be driven by *Selection*³¹.

Regarding the bacterial communities associated with radish seeds, β NTI values were distributed around -2 with 16S and *gyrB* for the two first sampling stages and then less than -2 at the final mature seed stage (**Fig. 7A and B**). Therefore, it seems that *Selection* had less influence on assembly of radish seed communities, at least on the early seed developmental stages, in comparison to bean seed communities. Thanks to quantitative process estimates (QPE) framework³⁰, we were able to estimate the relative importance of other ecological processes during assembly of the seed microbiota. Based on this framework, it seems that ecological drift and dispersal limitation occurred during the early assembly stages (D26 and D37) of radish seed microbiota (**Fig. 7C and D**). At the late seed maturation stage of radish (D65), *Selection* became the principal process of community assembly (**Fig. 7C and D**).

Experiment 3 – Seedling transmission of seed-borne taxa and their influence on seed vigour

Seed microbiota composition is influenced by leaf height

In the previous section, we showed that individual seeds were colonized by a few bacterial taxa of variable identity. Part of the variation was explained by the individual plants. To gain more insight into the factors involved in this variability at the seed level, mature seeds from five individual plants were collected from eight fruits located at four different leaf heights for each plant species. Based on nested PERMANOVA performed on weighted UniFrac distance, leaf height inside each plant explained 12% and 11.8% of variation ($P < 0.05$, **Fig. S7A and S8A**) in bacterial phylogenetic composition of bean and radish seeds, respectively. In contrast, we did not observe a significant pattern of richness and taxonomy according to the position of the seed according to the leaf height (**Fig. S7B and S8B**). We confirm here the hypothesis of a spatial heterogeneity of seed microbiota between plants but also within each plant individual.

Dynamics of seed microbiota during seedling emergence

We next characterized which seed-borne bacterial taxa were transmitted to the seedling. Dynamics of the seed microbiota during emergence were analyzed in soil-less conditions using a non-destructive sampling procedure (see Methods), which makes it possible to monitor these dynamics at the level of individual seedling. We analyzed only the samples where CFU were recovered for seed and seedlings, therefore excluding non-germinating seeds or seeds without detectable CFU (**Fig. 8A**).

During transition from seed to seedling, median bacterial population size increased by 1.4 to 6.0 log₁₀ CFU in bean and by 2.3 to 6.1 log₁₀ CFU in radish (**Fig. 8B**). Increase in bacterial population size during emergence was associated with a significant ($P < 0.001$) decrease in richness in bean (10 to 4 ASVs with 16S and *gyrB*). In contrast, no significant change in bacterial richness was detected in radish with both molecular markers (**Fig. 8C and S9**).

About one fifth (43 out of 223) and one quarter (56 out of 243) of the bacterial ASVs associated with bean and radish seeds were detected in seedlings using 16S (**Fig. 9A**). The use of *gyrB* reduced these proportions to 11% and 5.6% for bean and radish, respectively (**Fig. S10A**). Although seed-borne

ASVs not detected in the seedlings were affiliated with a multitude of bacterial orders, Burkholderiales were almost never detected in seedlings of both plant species (**Fig. S11** and **S12**). More surprisingly some seedling-associated ASVs were not detected in seeds. The absence of CFU detection in our sterile experimental system (**Fig. S9A**) strongly suggested that these seedling-associated ASVs originated from seeds. Among these ASVs specifically associated with seedlings were numerous Bacillales and Rhizobiales for bean and multiple Pseudomonadales for radish (**Fig. S11** and **S12**).

The initial abundance of seed-associated taxa does not explain its transmission to seedlings

Since transmission of some seed-borne plant pathogens to seedling is explained by their initial populations sizes on seeds^{32,33}. We investigated whether the relative abundance of seed-borne ASVs was a predictor of seedling transmission. As what had already been observed on experiment 2, each individual seed of both plant species was associated with one dominant ASV. The same pattern was observed for seedling, where one dominant ASV was associated with bean and radish seedlings (**Fig. 9B** and **S11B**).

However, the dominant seed-borne ASV did not necessarily become dominant in the corresponding seedlings. Indeed, only 11 (out of 57) and six (out of 52) dominant seed-borne ASVs were also dominant on radish seedling according to 16S and *gyrB*, respectively. This value is even lower for bean with four 16S ASVs (out of 43) and one *gyrB* ASV, which were dominant in both habitats. Altogether these results suggest that the initial abundance of seed-borne taxa is not predictive of its transmission to seedlings as many rare seed-borne taxa become dominant on seedlings, especially in bean.

The identity of the seed-borne taxa affected seedling fitness

Since seeds are colonized by few taxa of variable identity, we investigated whether the identity of these ASVs could impact the first phases of crop establishment, namely germination and seedling emergence.

A high proportion of seeds (48% for bean, 45% for radish) did not germinate under our experimental conditions (**Fig. 8A**). Bacterial population size (**Fig. S13A**) and richness (**Fig. S13B** and **S14A**) were not significantly different between germinating and non-germinating seeds. Comparison of the relative abundance (RA) of seed-borne 16S ASVs in germinating and non-germinating seeds revealed significant increase ($P < 0.01$, $\log_2FC > 2$) in RA of six and three ASVs in non-germinating seeds of bean and radish, respectively (**Fig. S13 C**). However, according to the taxonomic affiliation at the genus level, none were found in common with the three and eight *gyrB* ASVs enriched in non-germinating seeds (**Fig. S14B**). Unfortunately, we were not able to recover representative bacterial strains for these enriched ASVs and therefore did not confirm the observed phenotype experimentally.

Once the seeds have germinated, the corresponding seedlings can be classified as normal or abnormal according to their morphologies (see Methods). By using this classification scheme, 22 normal and 22 abnormal bean seedlings were obtained as well as 40 normal and 10 abnormal radish seedlings (**Fig. 8A**). Initial bacterial population size and richness observed on seeds were not good predictors of seedling phenotypes (**Fig. 10A** and **10B**). Changes in RA of seed-borne ASVs were next

investigated in bean seeds but not in radish as few seeds produced abnormal seedlings, which does not allow for a balanced distribution. Overall, the RA of six 16S ASVs and eight *gyrB* ASVs were significantly enriched ($P < 0.01$, $\log_2FC > 2$) in seeds that produced abnormal seedlings (**Fig. 10C**). The taxonomic affiliation of four of these 16S and *gyrB* ASVs was in agreement at the genus-level (**Fig. 10C**). On the other hand, the RA of three 16S and six *gyrB* ASVs were significantly enriched in seeds that produced normal seedlings (**Fig. 10C**).

To validate experimentally this phenotype, we isolated representative bacterial strains on the basis of their *gyrB* sequences and inoculated these strains on bean seeds (**Fig. 10D**). We selected strains representative of ASVs (i) significantly enriched in seeds that produced abnormal seedling (ASV1: *Pantoea agglomerans*), (ii) significantly enriched in seeds that produced normal seedling (ASV85: *Paenibacillus amylolyticus* and ASV90: *Bacillus aryabhattai*) and (iii) not significantly enriched (ASV474: *Sphingomonas* sp.). First, we validated that the seed surface-disinfection procedure employed did not impact seedling phenotype (**Fig. 10D**). As a control, seed inoculation of ASV474 did not induce any difference in seedling phenotype. In contrast seed-inoculation with two representative bacterial strains of ASV1 resulted in a significant ($P < 0.05$) increase of abnormal seedling in comparison to surface sterilized seeds, mainly associated with rot symptoms. Inoculation of ASV85 and ASV90 increased the number of normal seedling (**Fig. 10D**). Hence, we confirmed the seedling phenotype expected based on the RA of ASVs.

Discussion :

The aim of this study was to assess the assembly of the seed microbiota during the seed development, its dynamics during seedling emergence and impact on seed vigor. Using a culture-enrichment procedure for performing community profiling at the individual seed level, we confirmed that seed transmission of bacteria is bottlenecked in bean and radish with the colonization of one dominant ASV of variable identity. Although abundance of seed-borne taxa were not predictive of efficient seedling transmission, their identities can significantly impact the bean seedling phenotype.

Community profile of a significant portion of collected seeds (bean 32%, radish 42%) could not be estimated as no CFU was detected on the culture medium employed, an observation already reported for seeds of several plant species¹⁶. The proportion of seed without detectable colony on TSA10 was negatively correlated with seed development stage in radish, while not impacted by seed development in bean. Independent of the development stage, spatial localization of seed was not a good estimator of CFU detection since seeds with or without detected colonies were often co-localized in the same fruit. Could the seeds without detectable colonies be considered as steriles? This is unlikely as CFUs were systematically obtained on each corresponding seedling, which justifies the use of seedlings for the detection of specific plant pathogens in the seed industry (a procedure called grow-out test¹⁷). The absence of CFU detection in some seeds could be explained by a localization in the internal tissues of the seed (i.e. on the surface of the embryo) which would not allow them to be obtained with the soaking procedures employed in this work. However, if this were the case, it is rather difficult to explain why these taxa would be located only in the embryo and not in the other cell layers of the seed since seed-borne bacteria are frequently visualized near the seed coat^{24,34}. Alternatively, it is likely that some seed-associated bacterial populations were in a state of viable but nonculturable (VBNC), which corresponds to a low metabolic activity state without formation of colonies on standard

culture media. VBNC has been described in a number of seed-transmitted bacteria including *Clavibacter michiganensis* pv. *michiganensis*³² and *Acidovorax citrulli*^{35,36}. Why some populations enter the VBNC state in seeds and what are the signals that lift this state during germination-emergence are not yet known.

Overall, individual seeds were associated with a small number of taxa with a median of seven and four bacterial ASVs detected per individual seed of bean and radish, respectively. However, a dominant taxon representing 75 to 100% of all reads was associated with each individual seed, which could be considered as the primary symbiont hypothesis¹⁸. This notion of primary symbiont has been proposed for taxa associated with mature seeds at the end of the seed developmental process¹⁸. But did the early seed colonizers are systematically found on the mature seed? It seems that the answer to this question varies depending on the plant species considered. Although we acknowledge that the destructive sampling procedure employed in this work does not allow direct monitoring of the dynamics of the assembly of the microbiota during seed development, indirect evidence suggested a contrasting response for radish and bean. For radish, we did not observe significant changes in phylogenetic diversity of dominant taxa over time. Moreover, seed microbiota composition was also stable across the radish developmental stage, therefore suggesting that abundant taxa found during seed filling were mostly conserved during seed maturation. In contrast, increase in diversity during bean seed development together with a shift in taxonomic composition indicated that the early dominant taxa associated with bean seeds were progressively replaced during seed filling and maturation. This changes in community composition following colonization of a nearly sterile habitat (like seed) by micro-organisms is similar to primary succession³⁷. Primary succession is influenced (*i*) the quantity of microorganism that can colonize a habitat and (*ii*) and changes in limiting resource through time³⁴. Hence, the distinct microbial succession dynamics observed during bean and radish seed development is probably related to difference in dispersal and/or turnover of nutrient sources.

To estimate the relative importance of the different seed transmission pathways described to date (i.e. internal, floral and external²⁸), atmosphere, flowers and stems were sampled throughout seed development. Overall, less than 20% of seed-borne taxa were detected in these three habitats, which probably reflected an undersampling of the source habitats in comparison to seeds. Although it is difficult to conclude on the relative importance of each habitat as a source of inoculum for the seed microbiota, we could see differences between the two plant species studied. Notably, the co-occurrence of ASVs in bean flower and seed was practically non-existent, which confirmed that seed transmission of bacteria through the floral pathway is weak in bean⁷. More interestingly, the variation in identity of seed-associated dominant ASV within the same sampling stage, especially between plants, appears to indicate that the local species pool varies. Variation in taxa composition is important for airborne communities and stem-associated communities, thus we hypothesize that both pathways might be important for bacterial transmission to the seed. However, airborne taxa have been demonstrated being a low source of microbial inoculum of leaves, fruits and flowers³⁸. We therefore hypothesized that xylem sap could be the most important seed-transmission pathway. If this assumption is correct, then the observed infra-individual variation in seed microbiota composition (i.e. between leaf stages) implies a spatio-temporal dynamics of taxa in the xylem sap³⁹.

Regardless of the origin of seed-borne taxa, what are the ecological processes involved in the assembly of the microbiota at the individual seed level? According to our neutral model, the relative importance of immigration rate was lower than local replacement rate for both plant species during

seed development, even though for radish the relative contribution of immigration increased significantly during seed maturation. Local replacement could be either due to selection or ecological drift. For bean, selection is probably the main driver of community assembly during seed filling and maturation as the phylogenetic composition of seed communities were more similar than expected by chance. It is likely that this selection is due to host-filtering, notably through the production of chemical compounds and physical protective structures that developed during seed filling. From the primary succession dynamics observed during seed filling, the nature of the habitat is likely to change rapidly between seed filling and seed maturation, which appear to counterselect the Enterobacteriales. For radish, the relative importance of the different ecological processes analysed (selection, ecological drift and dispersal) is less clear. Indeed, β NTI values were scattered on both sides of the threshold -2 that is considered as significantly different from the null expectation. It is therefore likely that selection is less important in radish than in bean during seed development and that ecological drift is involved in the replacement rate observed. If this assumption is correct, this would be consistent with the results obtained by Rezki and coworkers, where drift was reported as the key ecological process during bacterial community assembly of mature seeds of radish.

The presence of taxa within mature seeds does not necessarily guarantee their transmission to seedlings²². In the absence of competition with other sources of inoculum such as the soil, we wondered whether taxa abundance on mature seeds is a strong predictor of seedling transmission? That is clearly not the case, since few seed dominant bacterial ASVs were dominant on seedling. This is in agreement with recent results obtained on rapeseed, where rare seed-borne or soil-borne taxa were efficiently transmitted on seedling²⁵. As the initial inoculum size does not guarantee successful transmission to seedlings, then the difference in fitness between taxa (*Selection*) appears as a major driver of assembly of the seedling microbiota. In addition to host-filtering, competition for resources and spaces between taxa could be responsible for the selection of taxa in seedlings. Indeed, copiotrophic taxa were previously reported as dominant in seedling of bean and radish²⁴. However, according to the taxonomic affiliation of ASVs, it seems that seedling-associated bacteria were not exclusively species with high growth rates (e.g. *Rhizobium* sp.), hence other traits are responsible for their selection.

Is seed vigour affected by the identity of the dominant taxon in seeds? In our study, overall seed microbiota structure and bacterial biomass did not influence seed vigour. However, we identified changes in RA of some specific ASVs in seeds that are correlated to non-germinating seeds or abnormal seedling. Moreover, we were able to validate the impact of seed-borne bacterial strains representative of these ASVs on seedling phenotype. Indeed seed-inoculation of two bacterial strains affiliated to *Paenibacillus amylolyticus* and *Bacillus* resulted in an increase of normal seedling Firmicutes are well known for their beneficial impact on plant growth. For instance seed-borne *Bacillus* species^{40,41}, and *Paenibacillus* species⁴² are known to significantly impact the seedling phenotype. In contrast, representative strains of *Pantoea agglomerans* (ASV1) caused an abnormal seedling phenotype, notably through root rot. *P. agglomerans* is a core seed microbiome taxon, highly abundant and observed in almost all plant species studied to date⁴³. In bean, the occurrence of this taxa gradually decreased during seed filling and maturation. Therefore, symptoms caused on seedlings by *P. agglomerans* could occur when succession during seed development is not able to induce a low population level of this bacterial species.

Conclusion Individual seeds are associated with one dominant taxa of variable identity. These changes in identity are partly due to spatial heterogeneity not only between plants but also within the same plant. In addition, primary succession occurring during seed development are also responsible for these variations. Predicting these spatio-temporal changes at the seed level required an in-depth understanding of the processes involved in assembly of the seed microbiota. We showed that Selection is the main driver of community assembly but the nature(s) of this selective process remained to be explored. Research on microbial succession during the assembly of the seed microbiota and its seedling transmission will assist the design of seed microbiota engineering strategies through selection (i) of the most permissive seed transmission pathways and (ii) plant beneficial bacteria that enhance crop establishment.

Methods :

Experiment 1 - Culture-based enrichment of seed-borne bacteria

Common bean (*Phaseolus vulgaris* L. var. Flavert) and radish (*Raphanus sativus* var. Flamboyant5) were at a density of eight bean plants and five radish plants per square meters at the IRHS experimental station (47°28'45.2''N - 0°36'35.6''W). Seeds were aseptically removed from fruits with sterile scalpel and tweezers under a laminar flow hood. Seed samples ($n=24$) were soaked in 2 ml of phosphate-buffered saline (PBS, Sigma-Aldrich) supplemented with Tween® 20 (0.05 % v/v, Sigma-Aldrich) per gram of fresh material. Seed soaking was performed at 4°C under constant agitation (150 rpm) for 2h30 and 16h for radish and bean, respectively. Half of the suspension was stored at -20°C, while the other half was serial-diluted and plated on 1/10 strength Tryptic Soy Agar (TSA10 : 17 g.l⁻¹ tryptone, 3 g.l⁻¹ soybean peptone, 2.5 g.l⁻¹ glucose, 5 g.l⁻¹ NaCl, 5 g.l⁻¹ K₂HPO₄, and 15 g.l⁻¹ agar) supplemented with cycloheximide (50 µg.ml⁻¹, Sigma-Aldrich). Plates were incubated at 18°C during 5 days. Bacterial mats were collected by adding 2mL of PBS Tween and stored at -20°C. DNA extraction was performed on (i) the frozen suspension and (ii) the collected CFU suspension with the NucleoSpin® 96 Food kit (Macherey-Nagel) following the supplier's recommendations.

Experiment 2 – Community assembly at the individual seed level

A second field trial was performed at the experimental station of the National Federation of Seed Multipliers (FNAMS, 47°28'012.42''N - 0°23'44.30''W, Brain-sur-l'Authion, France). Flowers from 5 plants at the blooming stage were collected, all flowers from one plant corresponding to one independent sample. Flowers were soaked in 2 ml of PBS Tween per gram of fresh material, then crushed in a lab blender (Stomacher, Mixwel, Alliance Bio Expertise) for 2 min. Individual seeds from five to seven plants of bean and radish were collected at different stages of the seed development (**Fig. 1**). A total of 96 and 144 seeds were sampled per radish and bean plant, respectively. Seed water content was estimated at each sampling stage on a bulk of 15 seeds by weighing this sample before and after drying (96°C for 3 days). Vascular flow of stem of each individual plant was collected as follows: a 2 cm section of stem from the bottom of the plant was cut and outer-surface-disinfected with 70° ethanol, then first layers of epidermis were removed using a sterile scalpel. The resulting piece of stem was incubated horizontally for 2.5 h in 2 mL of PBS Tween under constant agitation (200 rpm) at 4°C. Finally, airborne bacteria were collected at each sampling stage. To achieve this, a passive sampling scheme was performed with TSA10 plates placed at 10-20cm from the flowers. Six TSA10

plates were positioned homogeneously within the plot during 3 hours and then incubated at 18°C during 5 days.

Similarly, to experiment 1, all suspensions were serial-diluted and plated on TSA10. After 5 days of incubation at 18°C, the number of bacterial cells was estimated by counting the number of colony-forming units (CFU). We then determined a limit of quantification of 20 CFUs and an upper limit of quantification of 250 CFUs per plate⁴⁴. DNA extractions were performed as described in experiment 1.

Experiment 3 - Community dynamics during emergence

At the last sampling point of experiment 2 (D50 bean; D65 radish), additional mature seeds were collected in 2019 at the FNAMS experimental station. Thirty-two seeds per individual plants ($n=5$) were collected from eight fruits located at four different leaf heights. Four seeds were aseptically collected per fruit (**Fig. S2**). Seeds were soaked in PBS Tween. After collecting the resulting suspension, seeds were placed in sterile tubes filled with cotton moistened with 4mL of sterile water. Tubes were incubated in a growth chamber (photoperiod: 16h/8h, temperature 25-22°C) to obtain 4-day-old radish and 6-day-old bean seedlings. Seedlings were placed individually in a plastic bag, crushed and resuspended with 2 ml of sterile water. Suspensions of seeds and seedlings were serial-diluted and plated on TSA10 and DNA was extracted as previously reported. The sterility of our experimental system was assessed by collecting cotton from seedless tubes ($n=30$). No CFU was detected on TSA10 after 5 days of incubation at 18°C.

Construction and sequencing of amplicon libraries

A first PCR amplification was performed with the primer sets 515f/806r⁴⁵ and gyrB_aF64/gyrB_aR553²³, which target the V4 region of 16S rRNA gene and a portion of *gyrB*, respectively. PCR reactions were performed with a high-fidelity Taq DNA polymerase (AccuPrime™ Taq DNA polymerase Polymerase system System, Invitrogen) using 5μL of 10X Buffer, 1μL of forward and reverse primers (gyrB_aF64/gyrB_aR553 [100μM]; 515f/806r [10μM]), 0.2μL of Taq and 5 μL of DNA. Cycling conditions for 515f/806r was composed of an initial denaturation step at 94°C for 3 min, followed by 35 cycles of amplification at 94°C (30 s), 50°C (45 s) and 68°C (90 s), and a final elongation at 68°C for 10 min. The cycling conditions for gyrB_aF64/gyrB_aR553 were identical with the exception of the annealing temperature (55°C instead of 50°C). Amplicons were purified with magnetic beads (Sera-Mag™, Merck). A second PCR amplification was performed to incorporate Illumina adapters and barcodes. PCR cycling conditions were as follows: denaturation at 94°C (2 min), 12 cycles at 94°C (1 min), 55°C (1 min) and 68°C (1 min), and a final elongation at 68°C for 10 min. Amplicons were purified with magnetic beads and pooled. Concentration of the pool was monitored with quantitative PCR (KAPA Library Quantification Kit, Roche). Amplicon libraries were mixed with 5% PhiX and sequenced with four MiSeq reagent kits v2 500 cycles (Illumina). A blank extraction kit control, a PCR-negative control and PCR-positive control (*Lactococcus piscium*, a fish pathogen that is not plant-associated) were included in each PCR plate.

Microbial community analysis.

All scripts and data sets employed in this work are available on GitHub (https://github.com/martialbriand/IRHS_EmerSys). Briefly, primer sequences were removed with cutadapt version 1.8⁴⁶. Fastq files were processed with DADA2 version 1.6.0⁴⁷. Chimeric sequences were removed with the removeBimeraDenovo function. Taxonomic affiliation of amplicon sequence

variants (ASVs) was performed with a naive Bayesian classifier⁴⁸ against the Silva 132 taxonomic training data or an in-house *gyrB* database available upon request.

Diversity analyses were conducted with Phyloseq version 1.32.0⁴⁹. Sequences derived from 16S rRNA gene that were unclassified at the phylum-level, affiliated to Archaea, Chloroplasts and Mitochondria were removed. Since the primer set *gyrB*_aF64/*gyrB*_aR553 primers can sometimes co-amplified *parE*, a paralog of *gyrB*, the *gyrB* taxonomic training data also contained *parE* sequences. ASVs affiliated to *parE* or unclassified at the phylum-level were removed. Sequences were aligned with DECIPHER version 2.16.1 (Wright, 2016) and neighbor joining phylogenetic trees were constructed with Phangorn version 2.5.5⁵⁰. Identification of sequence contaminants was assessed with decontam version 1.8.0⁵¹.

Alpha diversity metrics (richness and Faith's phylogenetic diversity) were measured after rarefaction at 4,000 reads per sample for all datasets and both markers (**Fig. S15**). Changes in bacterial phylogenetic composition were estimated with weighted UniFrac distance⁵² on log10+1 transformed values. Permutational multivariate analysis of variance was performed with the *adonis2* function of *vegan* 2.5.6⁵³. Rank abundance curves were performed BiodiversityR 2.12.3⁵⁴. Number of ASVs shared between specific contrasts were visualized with UpSetR 1.4.0⁵⁵. Annotations of phylogenetic trees were performed with iTol 6.1⁵⁶.

The relative importance of niche- or neutral-based processes in community assembly was estimated with abundance-based β -null deviation measures. Pairwise comparison of seed communities aggregated at the individual plant level was performed with β -mean-nearest taxon distance⁵⁷. Deviation from the null distribution was assessed with the β -nearest taxon index (β NTI³⁰). β NTI values < -2 or > 2 indicates that the observed phylogenetic turnover between a pair of communities is primarily driven by Selection³⁰.

Analysis of species-proportion distributions

When considering the set of all ASV observed on a plant, only a few of them are detected in each seed. This is probably caused by censoring induced by the measurement process, but it probably also comes from an important spatial heterogeneity. We therefore proposed a dedicated approach to provide a species-proportion analysis in seeds at the plant level for the most abundant species (> 0.1% of the total ASV abundance). We used the neutral model of Solé and co-workers²⁹ at the seed level, where all ASVs are assumed to be totally equivalent and driven by immigration (probability $0 < \mu < 1$) and internal random replacement (probability $0 < C < 1$). Abundances in each seed follow a beta-binomial (BB) distribution with parameters N_c (the local ASV population size in the seed), α_c, β_c where S is the typical species pool size available for immigration in the vicinity of each seed and I is the ratio of the *per capita* immigration rate over the *per capita* replacement rate $I_c = \frac{\mu}{(1-\mu)C^*} (N_c - 1)$ (Laroche et al. 2020). To obtain a distribution of seed-borne proportions at the plant level, we assumed that N followed a uniform distribution whose range [N_{min}, N_{max}] was determined as the 0.1-0.9 quantiles of the distribution of total ASV number in seeds at a given date. The other parameters ($S, R = \frac{\mu}{(1-\mu)C^*}$) were assumed to be plant dependent. We computed from the data the proportions of ASVs in each seed at a given date on a given plant, evaluated the first three empirical moments of the mean at the plant level and performed a generalized moment estimation to obtain an estimated value for S, R and the expected I , as well as the value of the error criterion (see supplementary method

for more details). All computations were coded in R, Wilcoxon test (resp. pairwise Wilcoxon rank test with Benjamini-Hochberg correction) was used for simple (resp. multiple) comparisons.

Seedling phenotype

Seedling phenotype was estimated *via* rules established by the International seed testing association ⁵⁸ (Briefly, a seedling was considered abnormal if at least 50% of the cotyledons or leaves were necrotic or rotten, if the hypocotyl or epicotyl were deformed, or if the root system was absent, stunted or rotten).

Changes in ASVs relative abundance were assessed with DESeq2 v 1.28 ⁵⁹. Representative bacterial strains of ASVs significantly enriched in abnormal or normal seedling were selected from an in-house bacterial culture collection derived from bean seeds. Bean seeds (*Phaseolus vulgaris* var. Flaver), were surface-sterilized after 1 min sonication (40 Hertz), soaking for 1 min in 96° ethanol, 5 min in 2.6% sodium hypochlorite, 30 sec in 96° ethanol and rinsed 3 times with sterile water. The seeds were dried on paper and placed in a bacterial inoculum from a fresh 24-hour culture calibrated at DO = 0.01 (approximately 10⁷ CFU/mL). The seeds were then dried on paper. In order to check the inoculation of the seeds, half of the dried seeds were individually macerated in a pre-filled plate of PBS Tween. The other half of the seeds were placed in sterile tubes with cotton and moistened with 4mL of water, at the level of one seed per tube. Seedlings phenotypes were evaluated after 6 days in a growth chamber (photoperiod: 16h/8h, temperature 25-22°C). We also analyzed seedling phenotypes of non-disinfected seeds to see the impact of disinfection on seedling phenotype. To confirm the transmission of the inoculated strains, for each phenotype, we gathered seedlings in three plastic bags, one for non-germinated seeds, one for normal seedling and the last one for abnormal phenotypes. Batches were crushed and re suspended with 2 ml of sterile water per individual. Suspensions of seedlings were serial-diluted and plated on TSA 10%. After 5 days, we performed *gyrB* PCR amplification of representative members of each plate with *gyrB_aF64*/*gyrB_aR553* primers. PCR products were sent to GenoScreen for Sanger sequencing.

References :

1. Shade, A., Jacques, M.-A. & Barret, M. Ecological patterns of seed microbiome diversity, transmission, and assembly. *Current Opinion in Microbiology* **37**, 15–22 (2017).
2. Goggin, D. E., Emery, R. J. N., Kurepin, L. V. & Powles, S. B. A potential role for endogenous microflora in dormancy release, cytokinin metabolism and the response to fluridone in *Lolium rigidum* seeds. *Ann Bot* **115**, 293–301 (2015).
3. Escobar Rodriguez, C., Antonielli, L., Mitter, B., Trognitz, F. & Sessitsch, A. Heritability and functional importance of the *Setaria viridis* L. bacterial seed microbiome. *Phytobiomes Journal* (2019) doi:10.1094/PBIOMES-04-19-0023-R.

4. Chen, M.-H., Jack, A. L. H., McGuire, I. C. & Nelson, E. B. Seed-colonizing bacterial communities associated with the suppression of Pythium seedling disease in a municipal biosolids compost. *Phytopathology* **102**, 478–489 (2012).
5. Ridout, M. E., Schroeder, K. L., Hunter, S. S., Styer, J. & Newcombe, G. Priority effects of wheat seed endophytes on a rhizosphere symbiosis. *Symbiosis* **78**, 19–31 (2019).
6. Matsumoto, H. *et al.* Bacterial seed endophyte shapes disease resistance in rice. *Nat. Plants* (2021) doi:10.1038/s41477-020-00826-5.
7. Chesneau, G. *et al.* Temporal dynamics of bacterial communities during seed development and maturation. *FEMS Microbiol Ecol* (2020) doi:10.1093/femsec/fiaa190.
8. Yang, L. *et al.* Dominant Groups of Potentially Active Bacteria Shared by Barley Seeds become Less Abundant in Root Associated Microbiome. *Front. Plant Sci.* **8**, (2017).
9. Klaedtke, S. *et al.* Terroir is a key driver of seed-associated microbial assemblages: Terroir shapes the seed microbiota. *Environmental Microbiology* **18**, 1792–1804 (2016).
10. Rochefort, A. *et al.* Influence of environment and host plant genotype on the structure and diversity of the Brassica napus seed microbiota. *Phytobiomes Journal* (2019) doi:10.1094/PBIOMES-06-19-0031-R.
11. Kim, W.-I. *et al.* Inhibition of Salmonella enterica growth by competitive exclusion during early alfalfa sprout development using a seed-dwelling Erwinia persicina strain EUS78. *International Journal of Food Microbiology* **312**, 108374 (2020).
12. Bergna, A. *et al.* Tomato Seeds Preferably Transmit Plant Beneficial Endophytes. *Phytobiomes Journal* **2**, 183–193 (2018).
13. Morales Moreira, Z. P., Helgason, B. L. & Germida, J. J. Crop, genotype, and field environmental conditions shape bacterial and fungal seed epiphytic microbiomes. *Can J Microbiol* (2020) doi:10.1139/cjm-2020-0306.

14. Prado, A., Marolleau, B., Vaissière, B. E., Barret, M. & Torres-Cortes, G. Insect pollination: an ecological process involved in the assembly of the seed microbiota. *Scientific Reports* **10**, 1–11 (2020).
15. Rezki, S. *et al.* Assembly of seed-associated microbial communities within and across successive plant generations. *Plant and Soil* **422**, 67–79 (2018).
16. Mundt, J. O. & Hinkle, N. F. Bacteria within ovules and seeds. *Appl. Environ. Microbiol.* **32**, 694–698 (1976).
17. Gitaitis, R. & Walcott, R. The Epidemiology and Management of Seedborne Bacterial Diseases. *Annual Review of Phytopathology* **45**, 371–397 (2007).
18. Newcombe, G. A Hypothetical Bottleneck in the Plant Microbiome. *Frontiers in Microbiology* **9**, (2018).
19. Bintarti, A. F., Sulesky-Grieb, A., Stopnisek, N. & Shade, A. Endophytic microbiome variation at the level of a single plant seed. *bioRxiv* 2021.04.27.441449 (2021) doi:10.1101/2021.04.27.441449.
20. Fort, T. *et al.* Maternal effects shape the seed mycobiome in *Quercus petraea*. *New Phytologist* **230**, 1594–1608 (2021).
21. Abdelfattah, A., Wisniewski, M., Schena, L. & Tack, A. J. M. Experimental evidence of microbial inheritance in plants and transmission routes from seed to phyllosphere and root. *Environmental Microbiology* **23**, 2199–2214 (2021).
22. Baker, K. F. & Smith, S. H. Dynamics of Seed Transmission of Plant Pathogens. *Annual Review of Phytopathology* **4**, 311–332 (1966).
23. Barret, M. *et al.* Emergence Shapes the Structure of the Seed Microbiota. *Appl. Environ. Microbiol.* **81**, 1257–1266 (2015).
24. Torres-Cortés, G. Functional Microbial Features Driving Community Assembly During Seed Germination and Emergence. *Frontiers in Plant Science* **9**, (2018).

25. Rochefort, A. *et al.* Asymmetric outcome of community coalescence of seed and soil microbiota during early seedling growth. *bioRxiv* 2020.11.19.390344 (2020) doi:10.1101/2020.11.19.390344.
26. Caruso, V., Song, X., Asquith, M. & Karstens, L. Performance of Microbiome Sequence Inference Methods in Environments with Varying Biomass. *mSystems* **4**, (2019).
27. Zhang, J. *et al.* High-throughput cultivation and identification of bacteria from the plant root microbiota. *Nat Protoc* **16**, 988–1012 (2021).
28. Maude, R. B. Seedborne diseases and their control: principles and practice. *Seedborne diseases and their control: principles and practice*. (1996).
29. Solé, R. V., Alonso, D. & McKane, A. Self-organized instability in complex ecosystems. *Phil. Trans. R. Soc. Lond. B* **357**, 667–681 (2002).
30. Stegen, J. C. *et al.* Quantifying community assembly processes and identifying features that impose them. *The ISME Journal* **7**, 2069–2079 (2013).
31. Vass, M., Székely, A. J., Lindström, E. S. & Langenheder, S. Using null models to compare bacterial and microeukaryotic metacommunity assembly under shifting environmental conditions. *Sci Rep* **10**, 2455 (2020).
32. Darsonval, A. *et al.* Adhesion and Fitness in the Bean Phyllosphere and Transmission to Seed of *Xanthomonas fuscans* subsp. *fuscans*. *MPMI* **22**, 747–757 (2009).
33. Dutta, B., Avci, U., Hahn, M. G. & Walcott, R. R. Location of *Acidovorax citrulli* in Infested Watermelon Seeds Is Influenced by the Pathway of Bacterial Invasion. *Phytopathology* **102**, 461–468 (2012).
34. Glassner, H. *et al.* Bacterial niches inside seeds of *Cucumis melo* L. *Plant Soil* **422**, 101–113 (2018).
35. Jiang, N. *et al.* Induction of the viable but nonculturable state in *Clavibacter michiganensis* subsp. *michiganensis* and in planta resuscitation of the cells on tomato seedlings. *Plant Pathology* **65**, 826–836 (2016).

36. Kan, Y. *et al.* Induction and Resuscitation of the Viable but Non-culturable (VBNC) State in *Acidovorax citrulli*, the Causal Agent of Bacterial Fruit Blotch of Cucurbitaceous Crops. *Front. Microbiol.* **10**, (2019).
37. Fierer, N., Nemergut, D., Knight, R. & Craine, J. M. Changes through time: integrating microorganisms into the study of succession. *Research in Microbiology* **161**, 635–642 (2010).
38. Abdelfattah, A. *et al.* Revealing Cues for Fungal Interplay in the Plant–Air Interface in Vineyards. *Front. Plant Sci.* **10**, (2019).
39. Deyett, E. & Rolshausen, P. E. Temporal Dynamics of the Sap Microbiome of Grapevine Under High Pierce’s Disease Pressure. *Front. Plant Sci.* **10**, (2019).
40. Pitzschke, A. Developmental Peculiarities and Seed-Borne Endophytes in Quinoa: Omnipresent, Robust Bacilli Contribute to Plant Fitness. *Front. Microbiol.* **7**, (2016).
41. Khalaf, E. M. & Raizada, M. N. Taxonomic and functional diversity of cultured seed associated microbes of the cucurbit family. *BMC Microbiol* **16**, 131 (2016).
42. Díaz Herrera, S., Grossi, C., Zawoznik, M. & Groppa, M. D. Wheat seeds harbour bacterial endophytes with potential as plant growth promoters and biocontrol agents of *Fusarium graminearum*. *Microbiological Research* **186–187**, 37–43 (2016).
43. Simonin, M. *et al. unpublished.*
44. Sutton, S. Accuracy of Plate Counts. *Microbiology Topics* **17**, 42–46 (2011).
45. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS* **108**, 4516–4522 (2011).
46. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
47. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**, 581–583 (2016).

48. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* **73**, 5261–5267 (2007).
49. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217 (2013).
50. Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592–593 (2011).
51. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).
52. Lozupone, C. & Knight, R. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology* **71**, 8228–8235 (2005).
53. Oksanen, J. *et al.* vegan: Community Ecology Package version 2.5-6 from CRAN. <https://rdrr.io/cran/vegan/> (2017).
54. Kindt, R. BiodiversityR. Package for Community Ecology and Suitability Analysis. <https://cran.r-project.org/web/packages/BiodiversityR/index.html>. (2015).
55. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).
56. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **44**, W242–W245 (2016).
57. Fine, P. V. A. & Kembel, S. W. Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography* **34**, 552–565 (2011).
58. ISTA, International Rules for Seed Testing. in vol. 2015 Chapter 5, i-5-56 (60) (2015).
59. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

Acknowledgments

This work was supported by the French National Research Agency [ANR-17-CE20-0009-01]. The authors wish to thank Emmanuelle Laurent and Vincent Odeau (FNAMS) for field crop management as well as Muriel Bahut (ANAN platform, SFR Quasav) for amplicon sequencing.

Competing interests

The authors declare no competing interests

Author contributions

G.C. and M.Ba. designed the study program. G.C., B.L., M.S. and M.Ba wrote the manuscript. G.C., M.S., A.P., C.M. and B.M. designed and did the experimental part of the study. G.C., M.B and M.Ba. wrote all computational pipelines with the exception of the “*analysis of species-proportion distributions*”. The neutral model of this latter section was designed and written by B.L. All authors contributed to the editing and final revision of the manuscript.

Data availability

The datasets supporting the conclusions of this article are available in the European Nucleotide Archive under the accession number PRJEB45079

Code availability

All R scripts employed in this work are available on GitHub (https://github.com/martialbriand/IRHS_EmerSys)

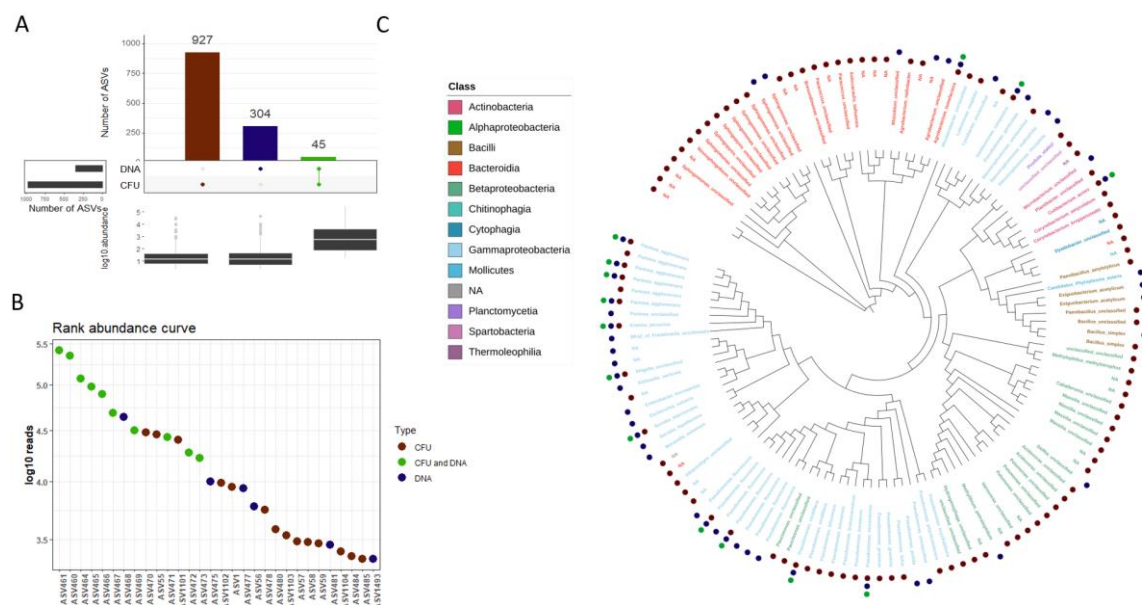


Figure 1 : Comparison of seed bacterial community profiles before and after culture-enrichment on TSA10. A. Visualization of the number of *gyrB* ASVs specifically detected before (DNA) and after (CFU) culture-enrichment, or jointly (CFU and DNA). Log10 abundance of ASVs are displayed in the boxplots. **B.** Rank abundance curve of the 30 most abundant ASVs. **C.** Neighbor-Joining phylogenetic tree of the most abundant ASVs (minimum relative abundance of 1%). Branch tips colors correspond to the Class taxonomic level. Blue and red circles represented ASVs specifically associated with culture -independent and -dependent approaches, respectively. Green circles represented ASVs detected with both approaches.

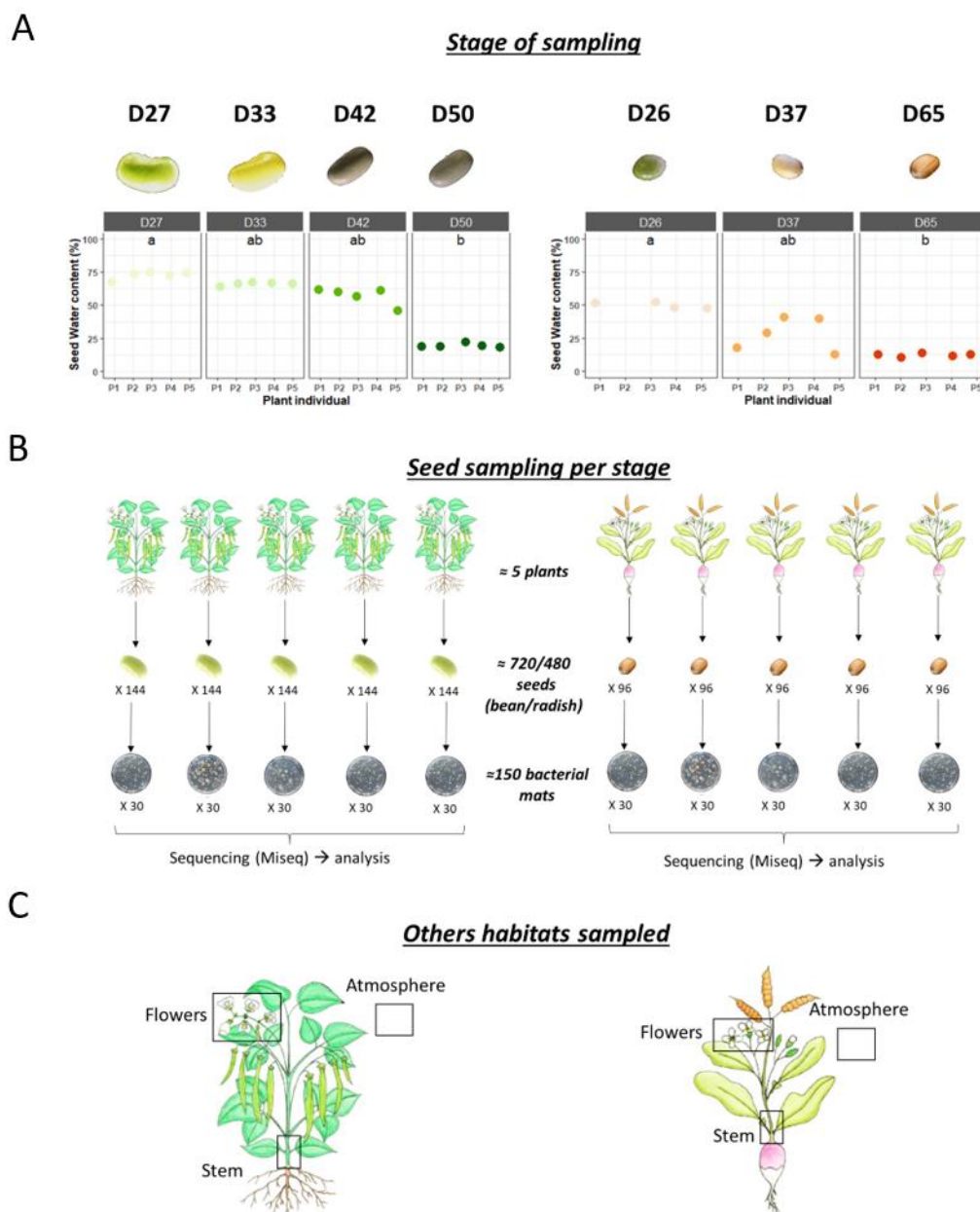


Figure 2 : Experimental design used to monitor the assembly of bacterial communities during seed development. A. Seed water content (%) associated with samples collected during bean (left) and radish (right) seed development. Letters denote significance evaluated from Kruskal-Wallis following Dunn test *post hoc* ($P < 0.01$) **B.** Number of samples processed at each stage. **C.** Sampling of flowers, stems and atmosphere, which could represent the main inoculum sources for seed taxa.

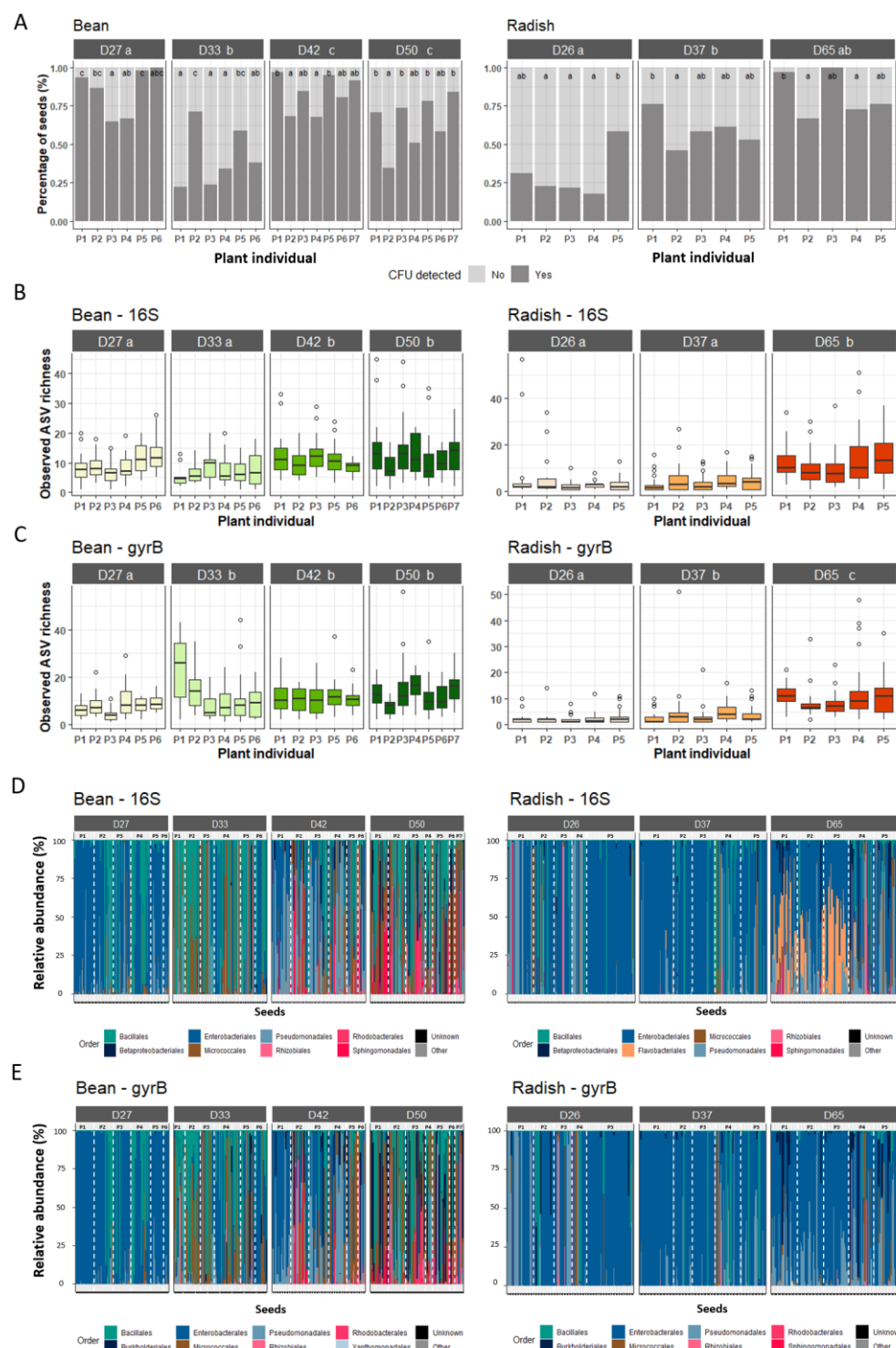


Figure 3 : Structure of the bacterial fraction of the seed microbiota. A. Percentage of seeds with (dark grey) or without (light grey) CFU detected. **B-C.** Number of ASV per seed (richness) estimated with 16S (B) or *gyrB*. Seeds were grouped according to their plant of origin (x-axis). **D-E.** Relative abundance of bacterial Orders for each individual seed according to 16S (D) or *gyrB* (E). White dotted lines represented individual plants. The names in each frame correspond to the stages sampled during seed development. The letters (a, b, c) indicated the significance level ($P < 0.01$).

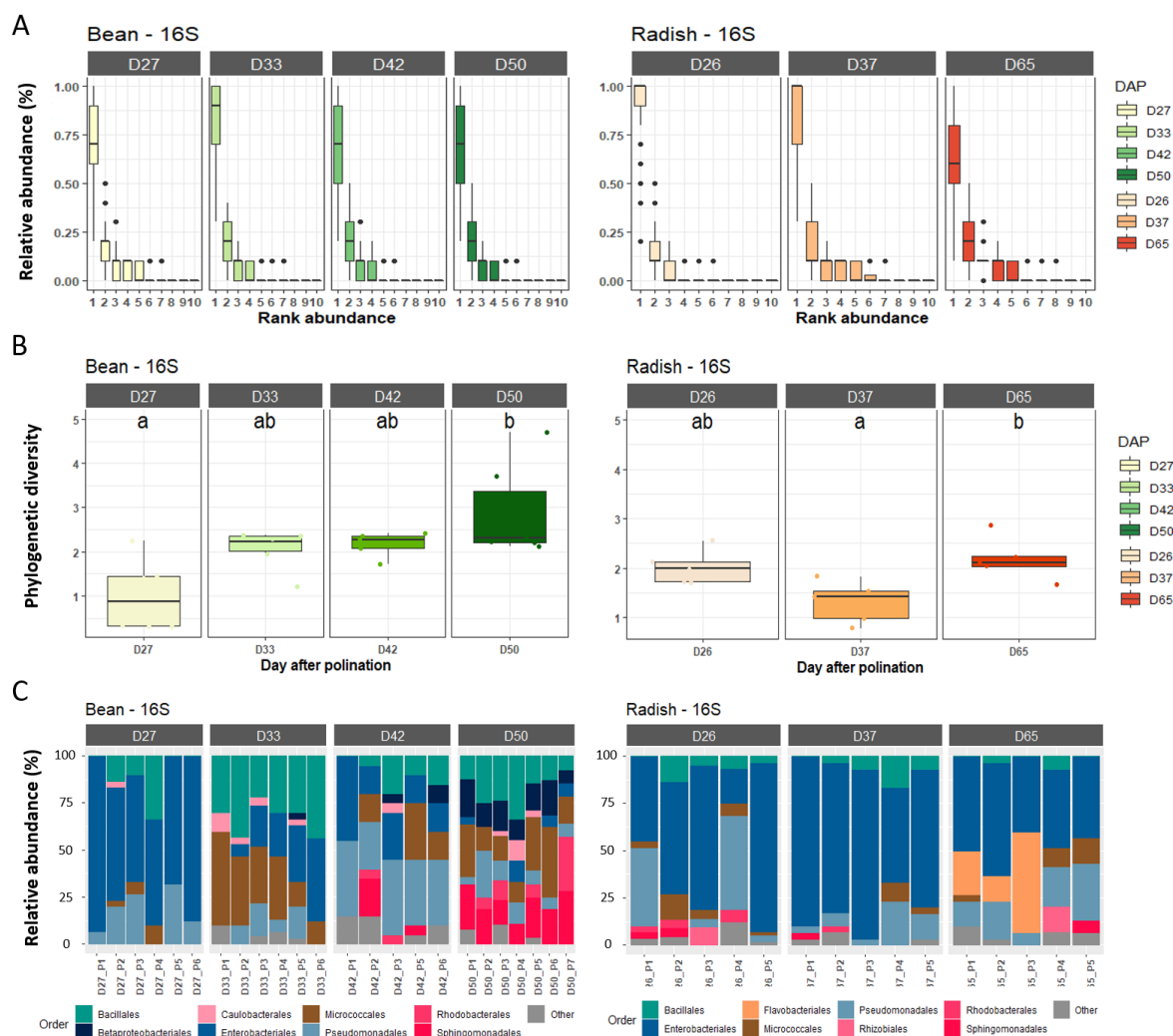


Figure 4 : Dominant bacterial taxa within the seed microbiota based on 16S. A. Rank abundance curves of bacterial ASVs associated with individual seeds. **B.** Faith's phylogenetic diversity of rank1 ASV agglomerated at the plant level. **C.** Relative abundance of bacterial Orders calculated from rank 1 ASV agglomerated at the plant level. The names in each frame correspond to the stages sampled during seed development. The letters (a, b, c) indicated the significance level ($P < 0.01$).

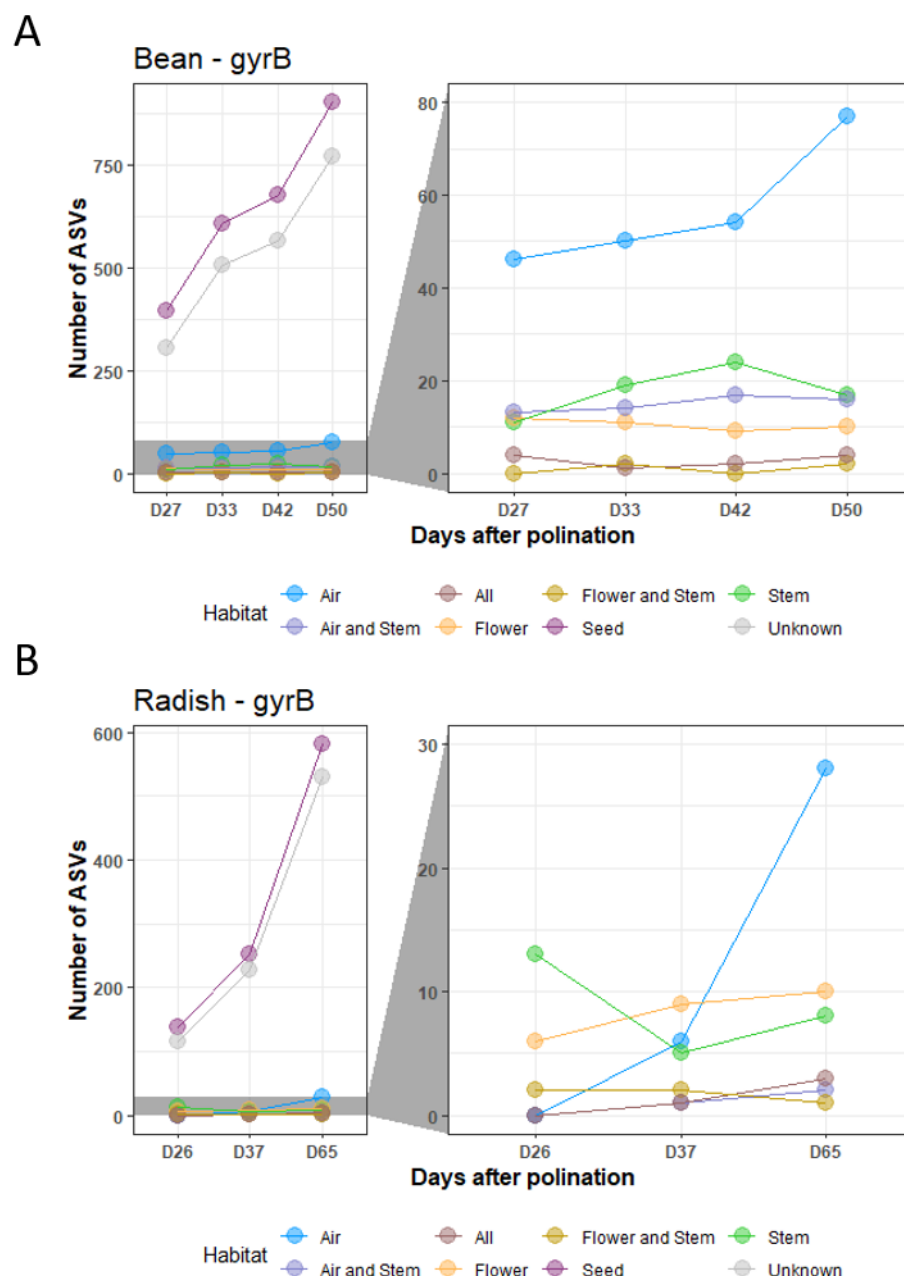


Figure 5 : Origin of seed-borne taxa. A-B. Number of *gyrB* ASVs associated with bean (**A**) and radish (**B**) seeds that were detected in other sampled habitats : flowers, stems and atmosphere. The right panel is a zoom of the left panel.

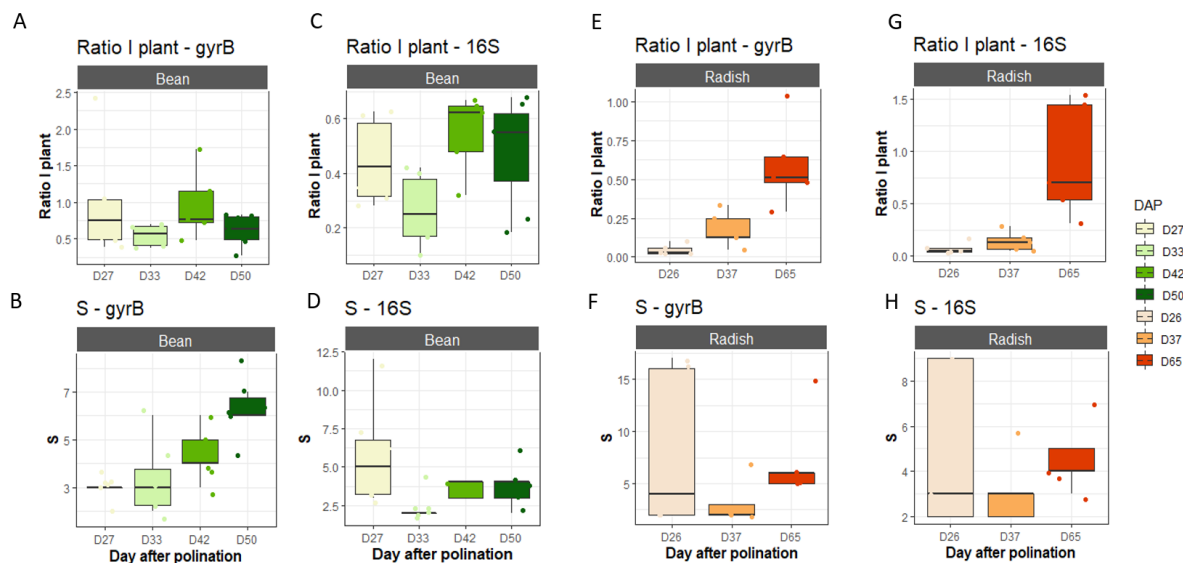


Figure 6 : Relative importance of dispersal over seed internal processes. A-C-E-F. The ratio I is used to estimate the relative importance of dispersal over seed internal processes (see Methods). **B-D-F-H.** Distribution of S (richness) for bean plants (green) and radish plants (red) as a function of the sampling day. Data were estimated with *gyrB* (A-B-E-F) and 16S (C-D-G-H).

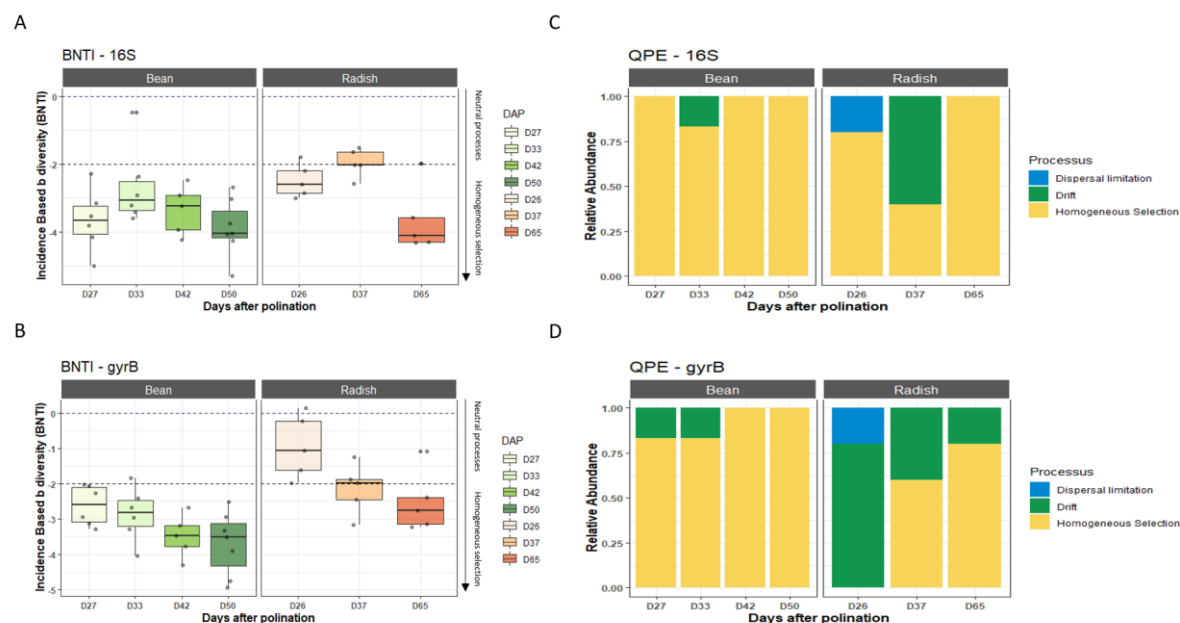


Figure 7 : Ecological processes involved in seed microbiota assembly at the plant level. A-B. β -nearest taxon index (β NTI) values were calculated with 16S (A) and *gyrB* (B) to assess deviation from the null expectation. **C-D.** Relative importance of species-sorting (homogeneous selection), dispersal limitation or ecological drift in assembly of seed bacterial communities at the plant level. Analyses were performed with 16S (C) and *gyrB* (D).

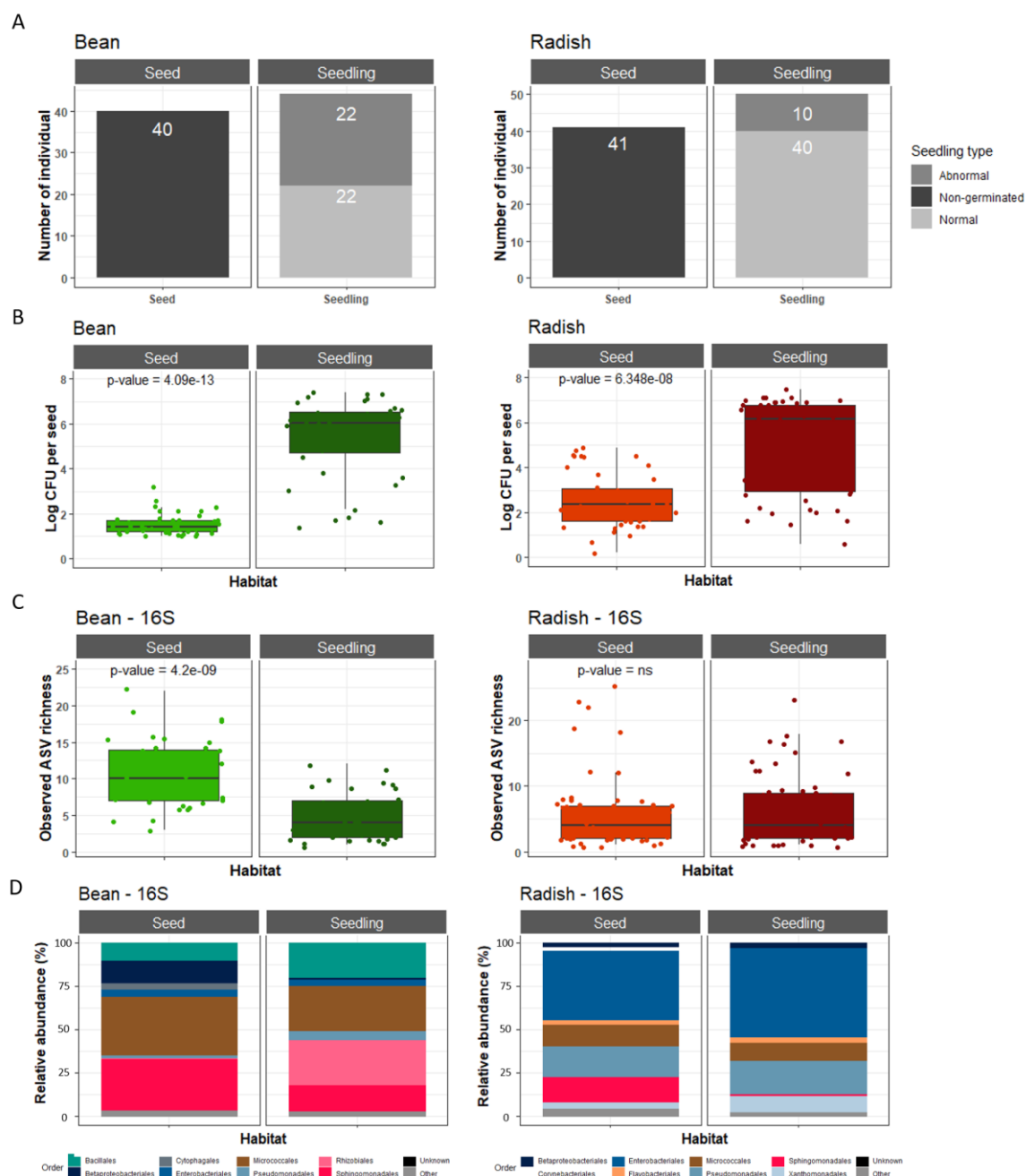


Figure 8 : Comparison of seed and seedling microbiota profiles. A. Number of non-germinated seeds (dark grey), abnormal (grey) and normal seedling (light grey). **B.** Average bacterial population size (log₁₀ CFU) associated with seed and seedling. **C.** Average richness (estimated with 16S ASVs) in seed and seedling. **D.** Taxonomic profile (Order level - 16S) of seeds and seedlings microbiota. *P*-values were derived from a Wilcoxon rank sum test.

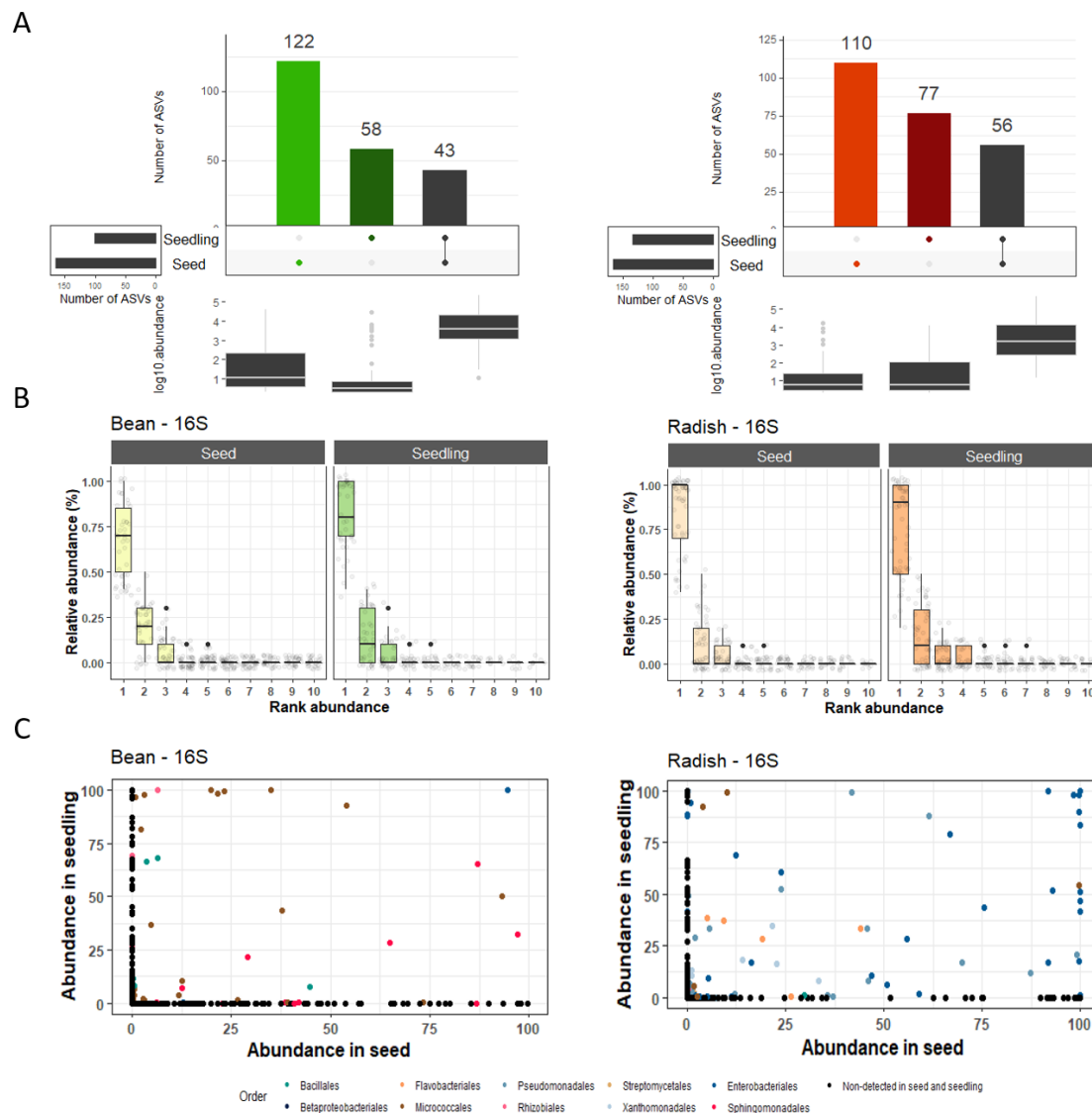


Figure 9: Seed to seedling transmission. **A.** Number of 16S ASVs specifically detected on seeds (light color), seedlings (dark color) or both habitats (black). **B.** Rank abundance curves of 16S ASVs in individual seeds and seedlings. **C.** Comparison of 16S ASV relative abundance in seed and seedling. Each dot corresponded to one ASV, which is colored according to its taxonomic affiliation at the Order level.

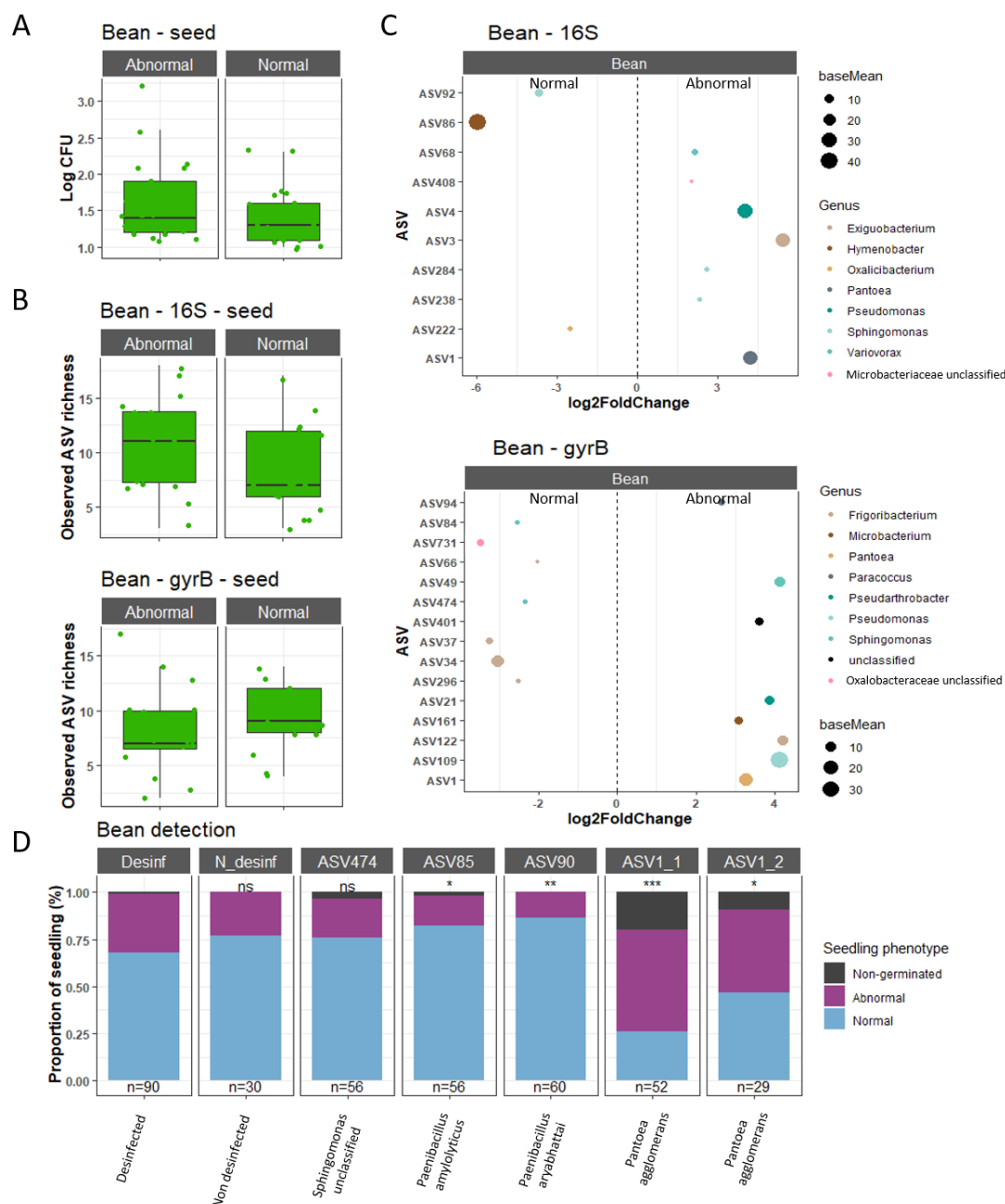


Figure 10: Impact of seed-borne ASVs on bean seedling phenotype. **A.** Correspondence between initial bacterial population size (log₁₀ CFU) and seedling phenotype. **B.** Correspondence between observed richness on seeds and seedling phenotype. **C.** ASVs with significant changes in relative abundance ($P < 0.01$) in normal and abnormal seedlings. Colors corresponded to taxonomic affiliation at the genus level. **D.** Proportion of non-germinated seeds, normal or abnormal seedling after seed inoculation of bacterial strains. All seeds were primarily disinfected before inoculation. Strains were representative of ASVs significantly associated with normal (ASV85 and ASV90) and abnormal seedlings (ASV1) or associated with both phenotypes (ASV474). Statistical tests were performed with Chi-squared test (ns = non-significant ; * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$).

Supplementary method

Species-proportion distribution analysis

We performed a qualitative analysis of the species-proportion distributions for both plant species. We focused on species thought to be actually representative of within seed microbiota, so microbial species whose abundance was less than 1% of the total abundance detected in one seed were discarded.

For each day and each plant, we computed the *empirical moments* of the observed ASV proportions. When considering the set of all ASV observed on a plant, only a few of them are detected in each seed. This is probably caused by censoring induced by the measurement process, but it probably also comes from spatial heterogeneity. This is why we computed the ASV proportions (f) for each seed related to the ASVs detected on that seed, and discarded the null frequencies for unobserved ASVs. Defining I_{plant} and J_{seed} as the set of seeds observed at a given date on a given plant and the set of ASV *observed at least one time* in this seed on that plant at the same date respectively, we obtain for each plant the following expressions for the first three empirical moments

$$\overline{m}_{1,2,3}_{plant} = \sum_{i \in I_{plant}} \frac{1}{|I_{plant}|} \sum_{j \in J_i} \frac{f_{ij}^{1,2,3}}{|J_i|}$$

We compared these moments with those obtained with a neutral distribution derived from [Solé 2002]. These authors proposed the use of neutral distribution as a crude first order approximation of the species-abundance distributions in complex ecosystems. This distribution is the stationary distribution of a neutral model, involving local species communities (for us, a microbial community associated to a seed in a plant) and a “local” regional pool (for us, a plant associated community in the vicinity of a seed). Here, neutrality is taken in the sense of [Laroche et al 2020], that is (i) local per capita rates of birth and death are identical for all individuals, whatever their species, (ii) immigration in the local community occurs through random sampling of individuals from the regional pool, (iii) all these rates depend on the local community size only. Moreover, it is assumed in this model that the local community is saturated (constant size $N > 0$), and the typical local species pool size (the number of ASV available for immigration from the plant and its close environment in the vicinity of a seed) is $S > 0$.

Two phenomena are accounted for, immigration (probability μ , such that $0 < \mu < 1$) and local competition, modelled through a parameter C , representing the percentage of ASV with which an ASV competes, with equal probability to win. This parameter is also called the connectivity of the local community. The corresponding probability to win is $C^* = 1 - (1 - C)^2$.

Standard results in ecological theory (see [Laroche et al. 2020] for a review) lead to the definition of the ratio of the per capita immigration rate (μ) over the per capita replacement rate ($(1 - \mu)C^*/(N_c - 1)$) in a community (seed) c of size N_c .

$$I_c = \frac{\mu}{(1 - \mu)C^*} (N_c - 1)$$

This is the relevant ecological parameter in these models, measuring the relative strength of immigration and internal competition or selection dynamics that shape the species abundance in the community, together with the local species pool size S . The species abundance distribution P is then given by a Beta-binomial distribution with parameters N_c, α_c, β_c expressed as

$$\alpha_c = \frac{I_c}{S}, \quad \beta_c = I_c \frac{(S-1)}{S}, \quad \mathbb{P}\mathbf{Bb}(n) = \binom{N_c}{n} \frac{\Gamma(n + \alpha_c) \Gamma(N_c - n + \beta_c)}{\Gamma(N_c + \alpha_c + \beta_c)} \frac{\Gamma(\alpha_c + \beta_c)}{\Gamma(\alpha_c) \Gamma(\beta_c)}$$

We modified this model to compute the species-abundance distribution at the plant level for a given day. To this end, we hypothesized that for seeds in which at least one microbe was detected, the distribution of total number of detected ASVs in all seeds was an acceptable proxy of the distribution of N_c accounting for natural variability and variability induced by the sampling process, but ignoring cultivation bias. Moreover, after inspection of the data, we assumed that for each observation date d

$$N_c \sim \text{uniform}(N_{\min d}, N_{\max d})$$

where $N_{\min d}, N_{\max d}$ depends on the sampling date but are the same for all plants.

Therefore, the marginal species abundance distribution at date d for seeds of plant p is given by

$$\bar{P}_{N_{\min d}, N_{\max d}, S, \frac{\mu}{(1-\mu)C^*}}(n) = \sum_{N>0, N \geq n} \mathbb{P}\mathbf{Bb}(n|N, S, \frac{\mu}{(1-\mu)C^*}) \mathbb{P}\mathbf{unif}(N|N_{\min d}, N_{\max d})$$

where $\mathbb{P}\mathbf{Bb}$ and $\mathbb{P}\mathbf{unif}$ are the Beta-binomial and uniform (on integer) probability functions.

We may notice that the model actually depends on the ratio $R = \frac{\mu}{(1-\mu)C^*}$ and the parameter S and $N_{\min d}, N_{\max d}$. As we actually only observe proportions of ASVs present within each seed, we compute the three first moments of positive proportion distribution at the plant level for this model:

$$\begin{aligned} m_1 &= E(f|1 \geq f > 0) = \sum_{n>0} f \sum_{N>0, \frac{n}{N}=f} \frac{\mathbb{P}\mathbf{Bb}(n|N, S, R)}{(1 - \mathbb{P}\mathbf{Bb}(0|N, S, R))} \mathbb{P}\mathbf{unif}(N|N_{\min d}, N_{\max d}) \\ &= \sum_{N>0} \left(\sum_{N \geq n>0} \frac{n}{N} \frac{\mathbb{P}\mathbf{Bb}(n|N, S, R)}{(1 - \mathbb{P}\mathbf{Bb}(0|N, S, R))} \right) \mathbb{P}\mathbf{unif}(N|N_{\min d}, N_{\max d}) \\ &= \sum_{N>0} \frac{1/S}{(1 - \mathbb{P}\mathbf{Bb}(0|N, S, R))} \mathbb{P}\mathbf{unif}(N|N_{\min d}, N_{\max d}) \\ m_{2,3} &= E(f^{2,3}|f > 0) = \sum_{N>0} \left(\sum_{N \geq n>0} \left(\frac{n}{N} \right)^{2,3} \frac{\mathbb{P}\mathbf{Bb}(n|N, S, R)}{(1 - \mathbb{P}\mathbf{Bb}(0|N, S, R))} \right) \mathbb{P}\mathbf{unif}(N|N_{\min d}, N_{\max d}) \end{aligned}$$

Based on this model, for each date we used the total number of ASVs to choose $N_{\min d}, N_{\max d}$ separately. After plotting the quantiles for each date, we observed that we could reasonably assume a uniform distribution on the interval corresponding to the 0.1 and 0.9 quantiles. Then, for a given set of parameters we simulated 1000 values of N and computed the moments according to the above formula.

We performed a **generalized moment estimation** of the parameters (S, R) to fit the observed moments for each plant. For this we screened the following parameter grid (whose range was selected by preliminary trial):

$$S = \{2, 3, \dots, 99, 100\},$$

$$R = \{0.000001, 0.000002, 0.000003, \dots, 0.000248, 0.000249, 0.00025\}$$

and selected for each plant the set of parameters that minimize the quantity

$$\sqrt{(m_{1obs} - m_1)^2 + (m_{2obs} - m_2)^2 + (m_{3obs} - m_3)^2}$$

We also tested a weighted version of this criterion, under the form

$$\sqrt{\left(\frac{m_{1obs} - m_1}{m_{1obs}}\right)^2 + \left(\frac{m_{2obs} - m_2}{m_{2obs}}\right)^2 + \left(\frac{m_{3obs} - m_3}{m_{3obs}}\right)^2}$$

The empirical and estimated probability densities of proportions for each day and each plant were tested and plotted for bean and radish with 16S marker (**FigSs1**).

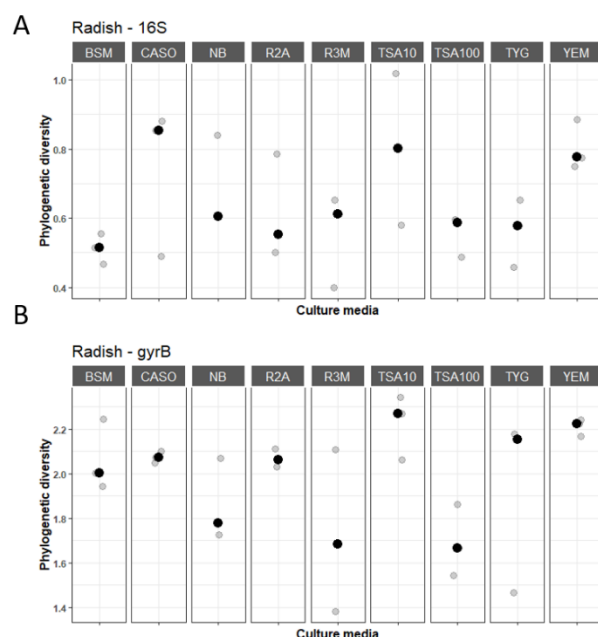


Figure S1: Phylogenetic diversity of seed bacterial communities after culture-enrichment on different synthetic media. Bacterial suspensions derived from radish seeds after soaking in PBS Tween (three subsamples of 1,000 seeds each) were diluted and plated on the following nine culture media: bacterial screening medium 523 (BSM); casein peptone soybean flour peptone agar (CASO); nutrient broth agar (NB); Reasoner's 2A (R2A); r3 medium (R3M) (Dione et al., 2016) ; tryptic soy agar 1/10 strength (TSA10); tryptic soy agar (TSA100); tryptone yeast extract glucose (TYG) and yeast extract mannitol agar (YEM). Phylogenetic diversity (Faith's PD) of bacterial communities was estimated with the v4 region of 16S rRNA gene (**A**) and *gyrB* (**B**). Black circles represented the average phylogenetic diversity of three replicates, which are displayed in grey.

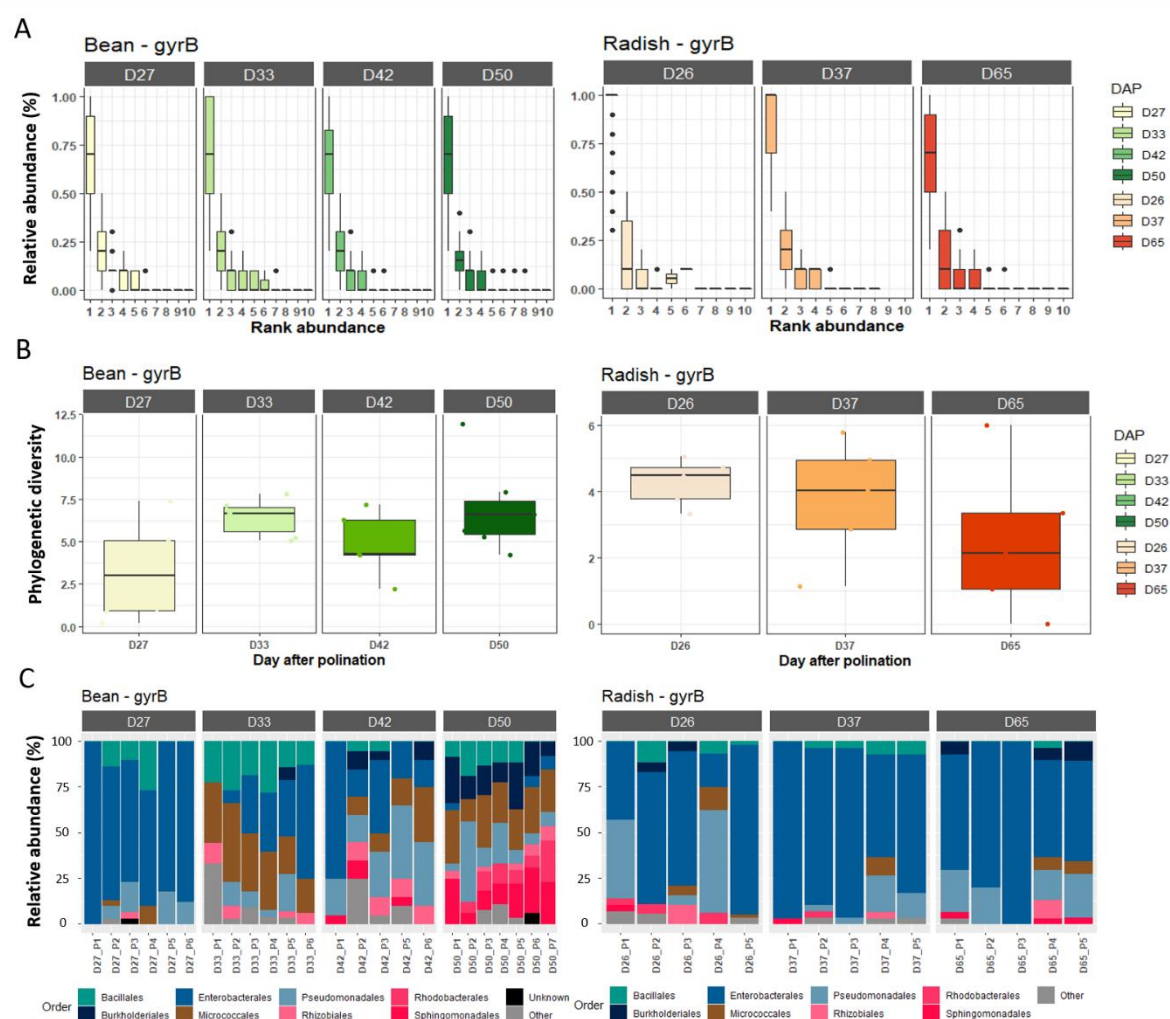


Figure S2 : Dominant bacterial taxa within the seed microbiota based on *gyrB*. **A.** Rank abundance curves of bacterial ASVs associated with individual seeds. **B.** Faith's phylogenetic diversity of rank1 ASV agglomerated at the plant level. **C.** Relative abundance of bacterial Orders calculated from rank 1 ASV agglomerated at the plant level. The names in each frame correspond to the stages sampled during seed development. The letters (a, b, c) indicated the significance level ($P < 0.01$).

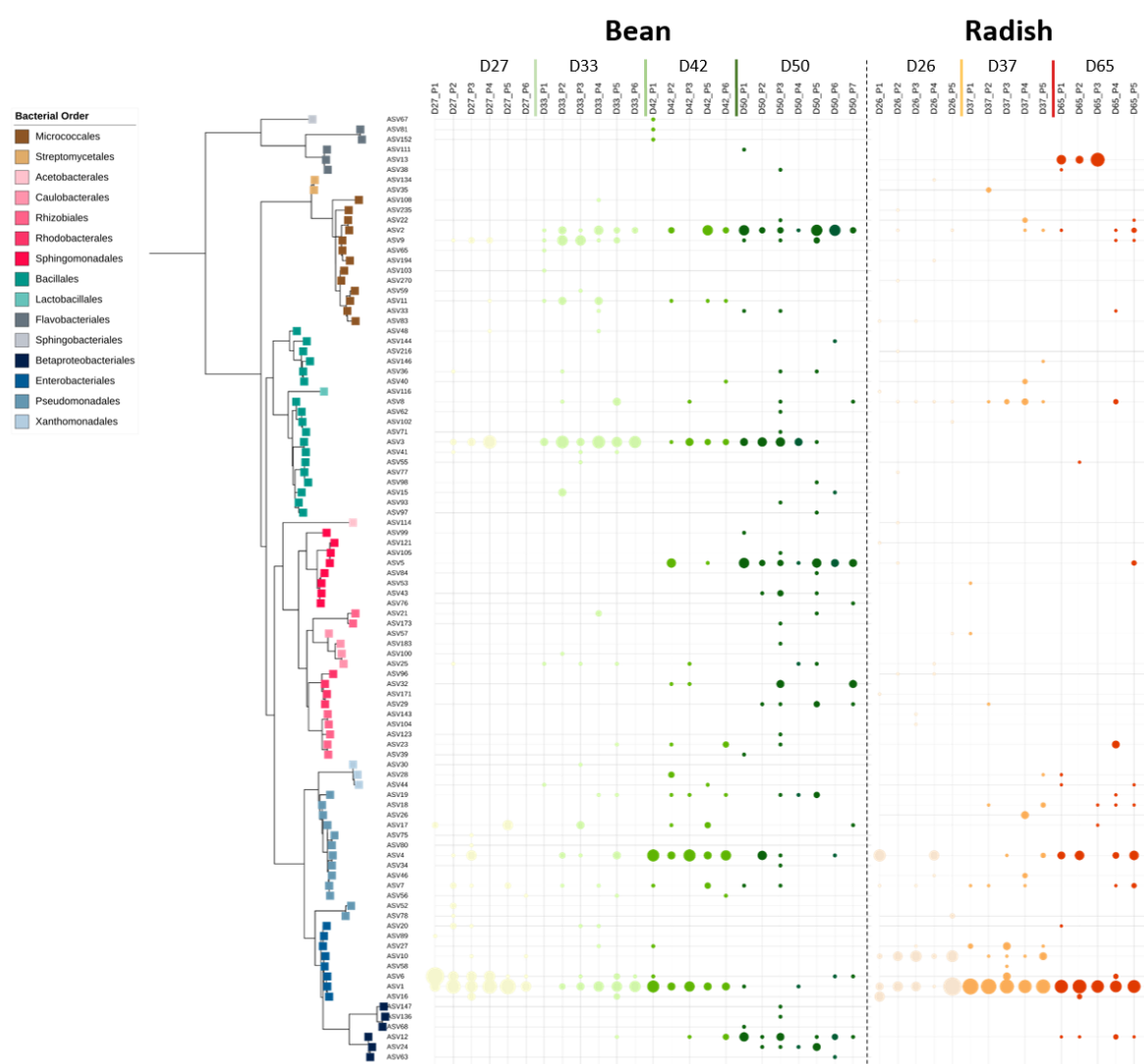


Figure S3 : Phylogenetic diversity of seed-borne dominant taxa based on 16S. Neighbor-joining phylogenetic tree of dominant 16S ASV. Occurrence of seed-borne ASVs within a specific plant is represented by a circle whose size is proportional to its relative frequency. Tips labels represented taxonomy at the Order level.

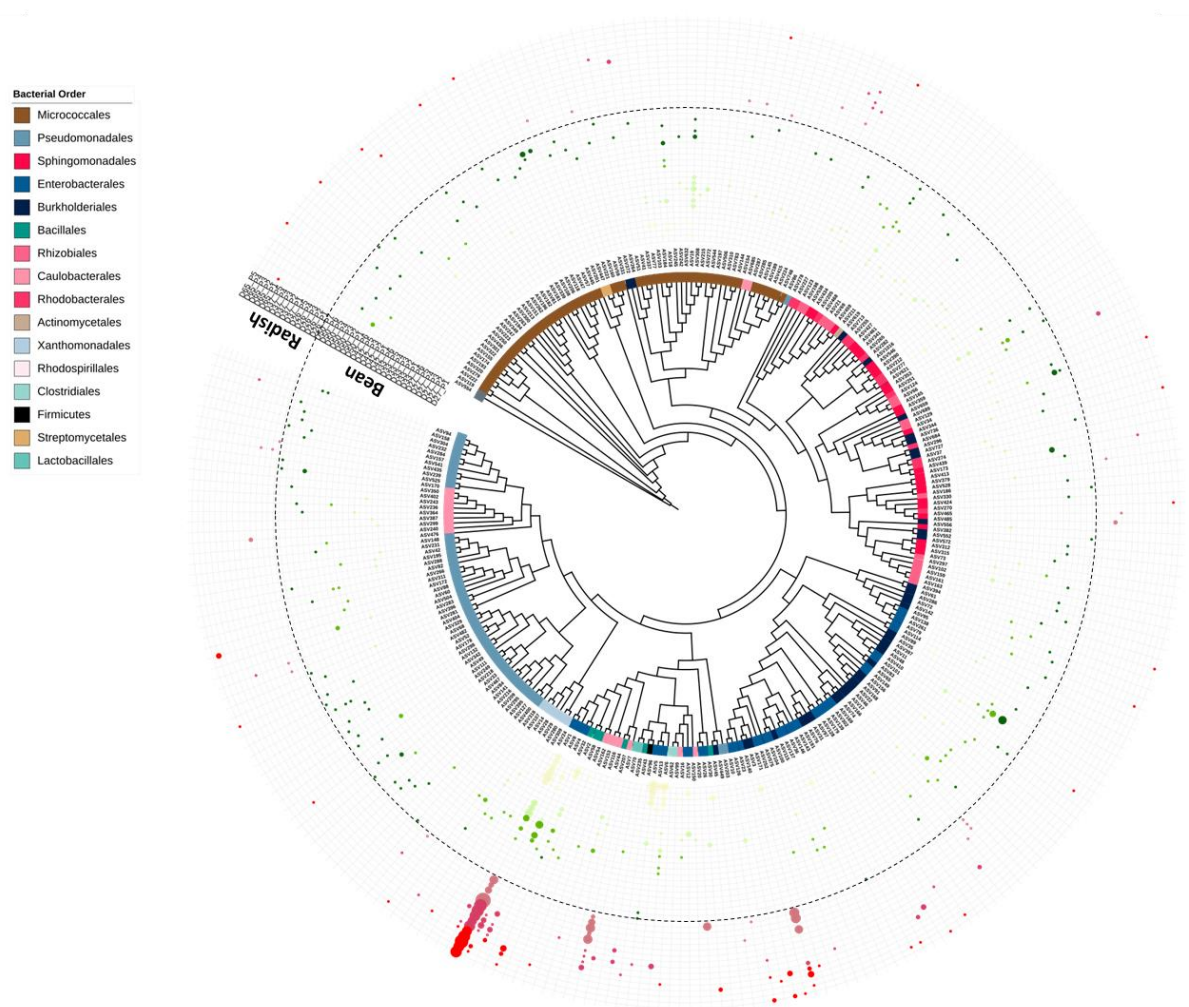
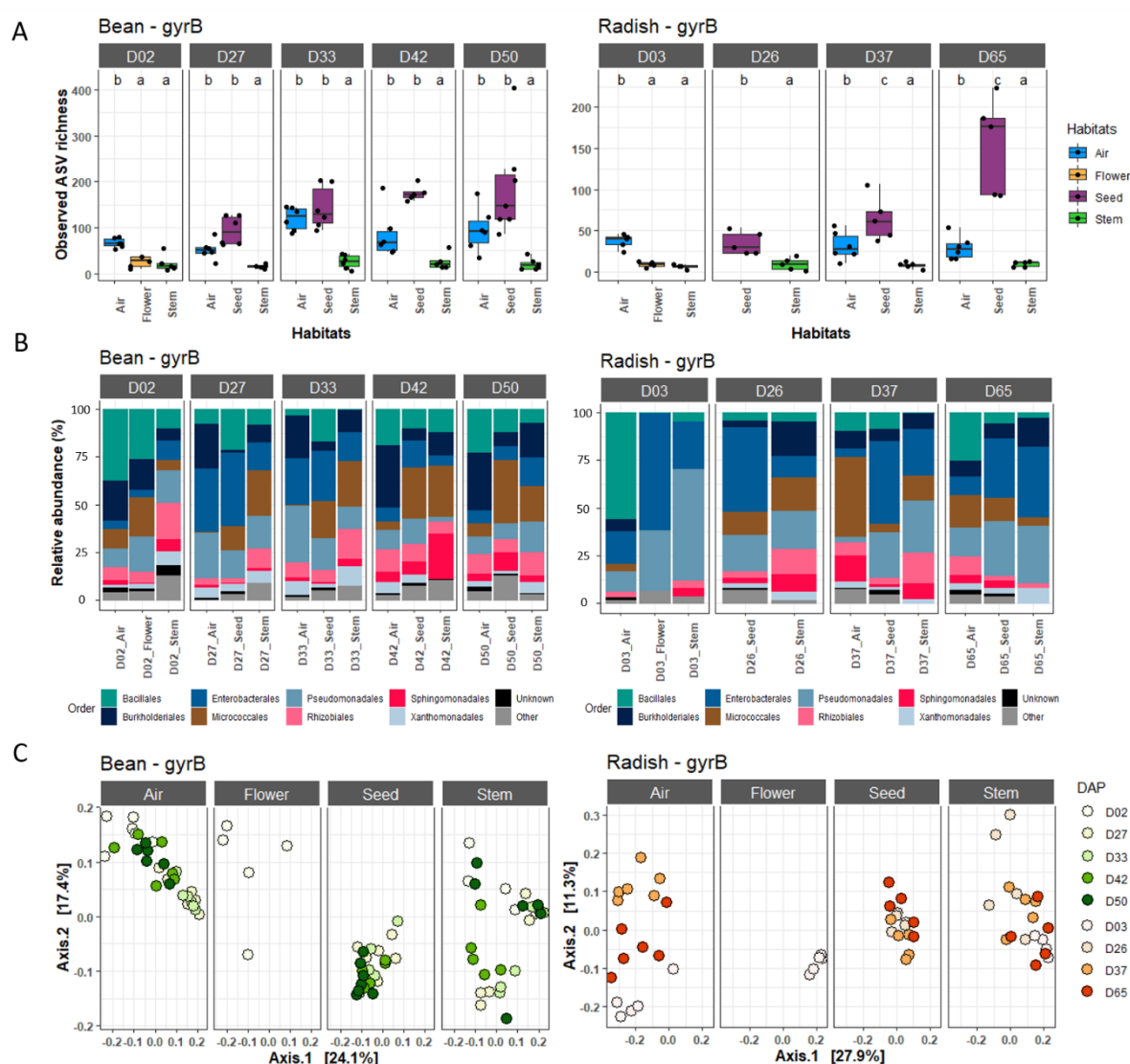


Figure S4 : Phylogenetic diversity of seed-borne dominant taxa based on *gyrB*. Neighbor-joining phylogenetic tree of dominant *gyrB* ASV. Occurrence of seed-borne ASVs within a specific plant is represented by a circle whose size is proportional to its relative frequency. Tips labels represented taxonomy at the Order level.



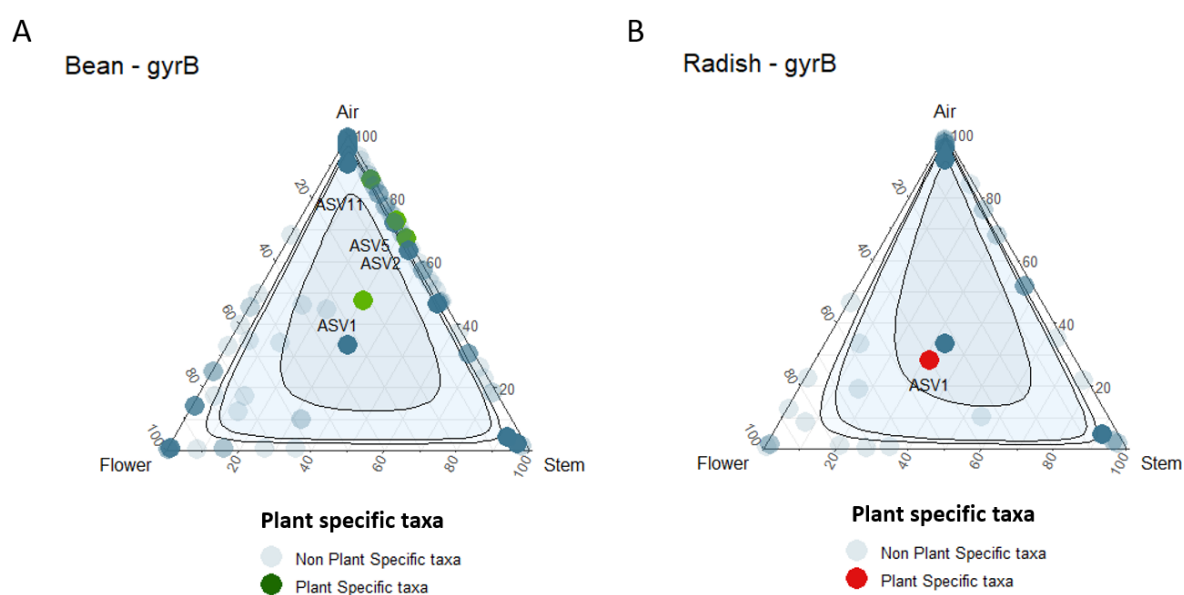


Figure S6 : Origin of dominant seed-borne taxa. A and B. Detection of dominant seed-borne ASVs (*gyrB*) in other sampled habitats: flowers, stems and atmosphere. Intervals of confidence are indicated by a blue colour gradient ranging from 0.50 to 0.95. ASVs significantly associated with a specific plant (see **Table S1**) are highlighted in green (**A**) and red (**B**) for bean and radish. Of note, ASV20 is not represented for radish because it was not detected in any sampled habitats.

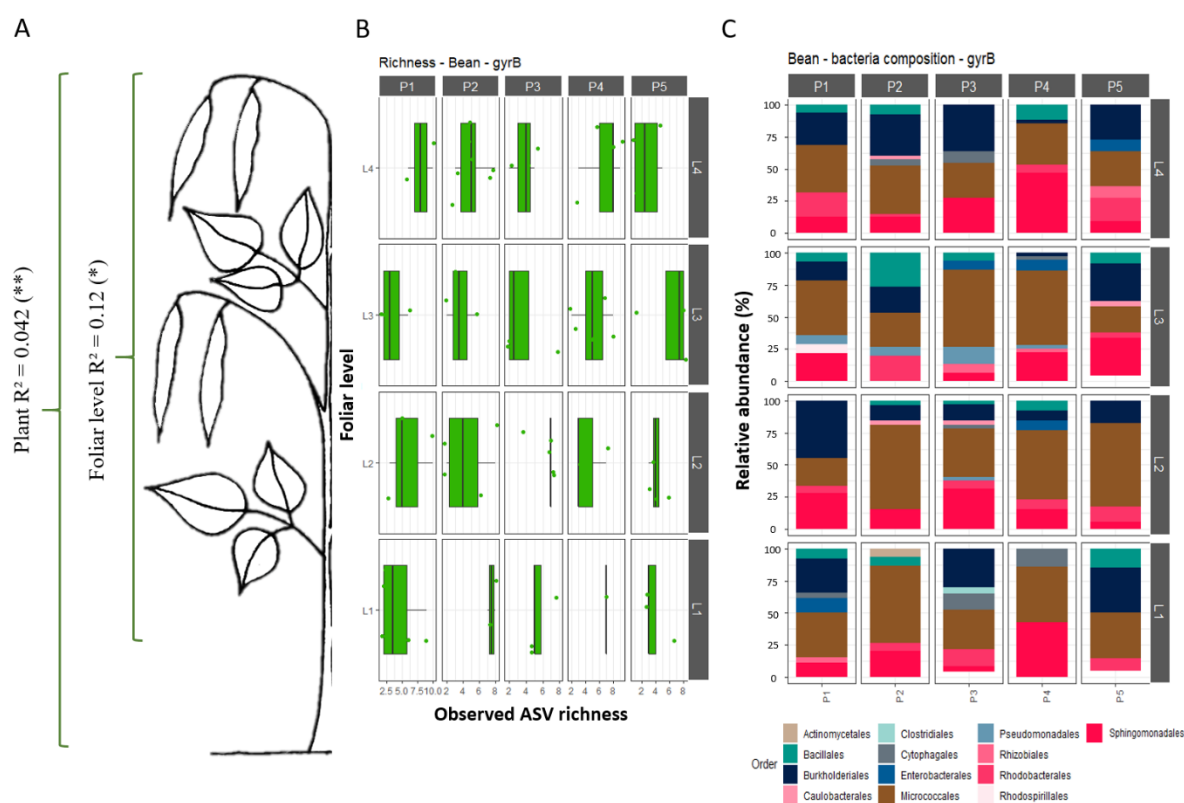


Figure S7 : Influence of foliar level on the structure of bean seed microbiota. A. Variation in seed bacterial phylogenetic composition (weighted Unifrac distance) according to individual plants and foliar level inside each plant (PERMANOVA; * = $P < 0.05$; ** = $P < 0.01$). **B.** Average richness of seed microbiota per foliar level inside each plant. **C.** Relative abundance of bacterial Orders per foliar level. Foliar levels are represented as a gradient from the bottom of the plant (L1) to the top of the plant (L4), y-axis. All analyses were performed with *gyrB*.

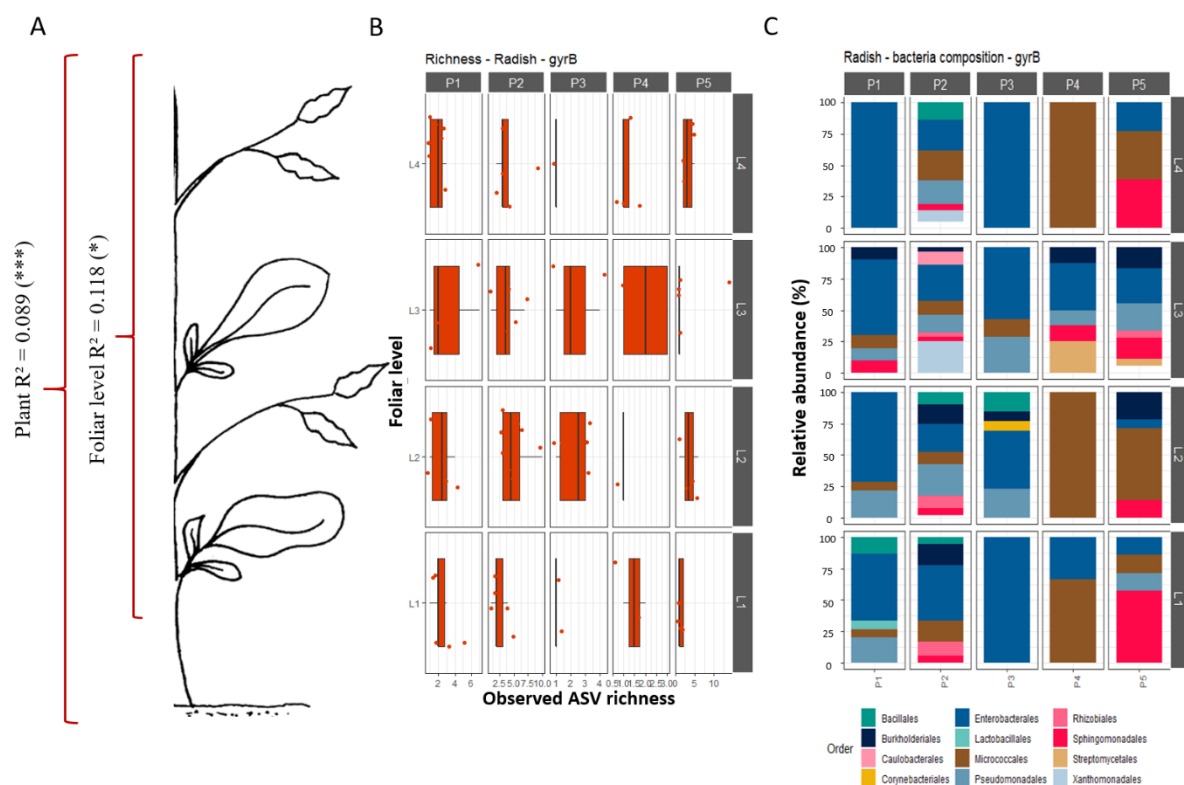


Figure S8 : Influence of foliar level on the structure of radish seed microbiota. A. Variation in seed bacterial phylogenetic composition (weighted Unifrac distance) according to individual plants and foliar level inside each plant (PERMANOVA; * = $P < 0.05$; ** = $P < 0.01$). **B.** Average richness of seed microbiota per foliar level inside each plant. **C.** Relative abundance of bacterial Orders per foliar level. Foliar levels are represented as a gradient from the bottom of the plant (L1) to the top of the plant (L4), y-axis. All analyses were performed with *gyrB*.

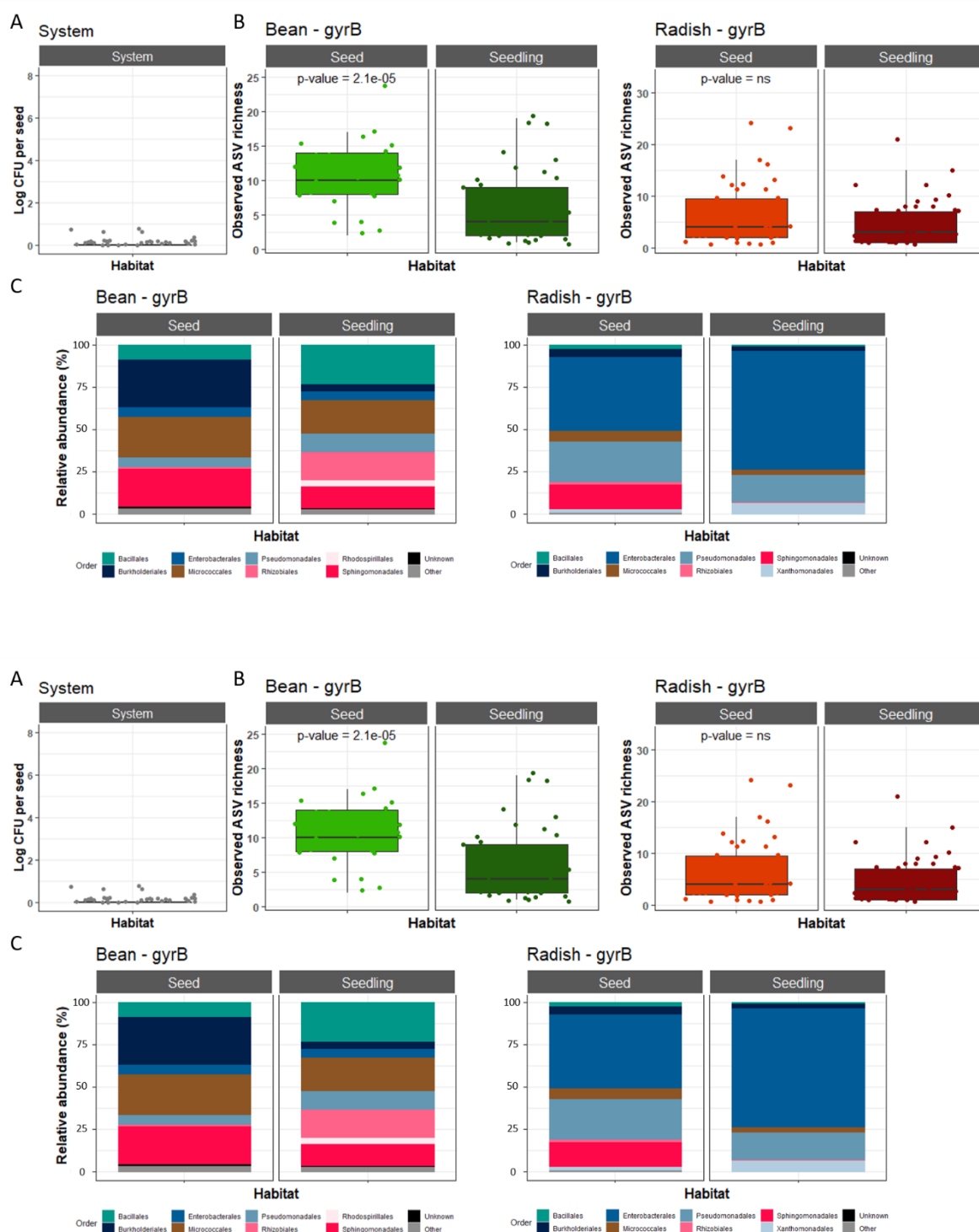


Figure S9 : Comparison of seed and seedling microbiota profiles. A. Log₁₀ CFU detected after seven days in our experimental system composed of tubes filled with cotton and sterile water. **B.** Average richness (estimated with *gyrB* ASVs) in seed and seedling. **C.** Taxonomic profile (Order level - *gyrB*) of seeds and seedlings microbiota. *P*-values were derived from a Wilcoxon rank sum test.

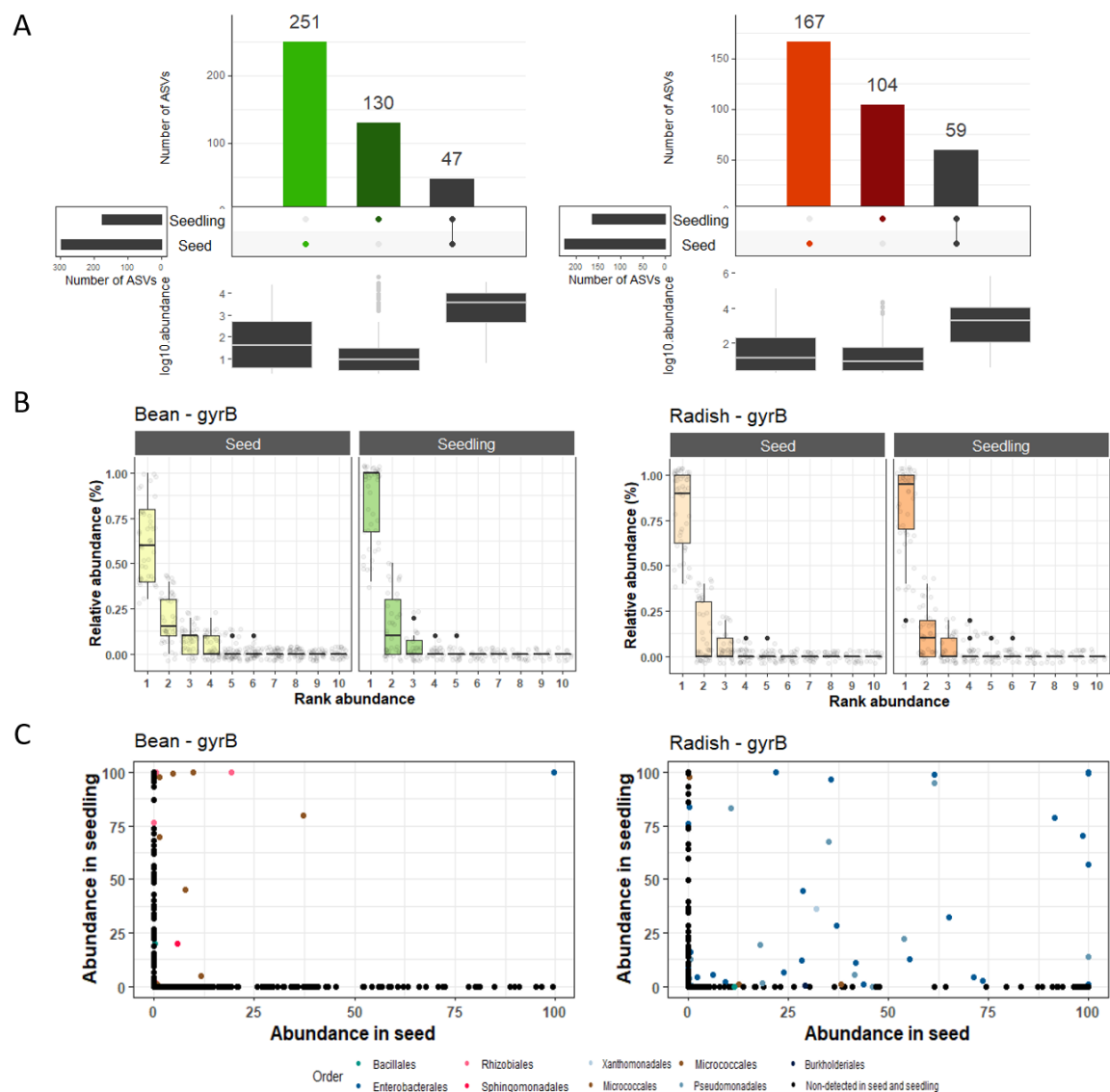


Figure S10: Seed to seedling transmission. **A.** Number of *gyrB* ASVs specifically detected on seeds (light color), seedlings (dark color) or both habitats (black). **B.** Rank abundance curves of *gyrB* ASVs in individual seeds and seedlings. **C.** Comparison of *gyrB* ASV relative abundance in seed and seedling. Each dot corresponded to one ASV, which is colored according to its taxonomic affiliation at the Order level.

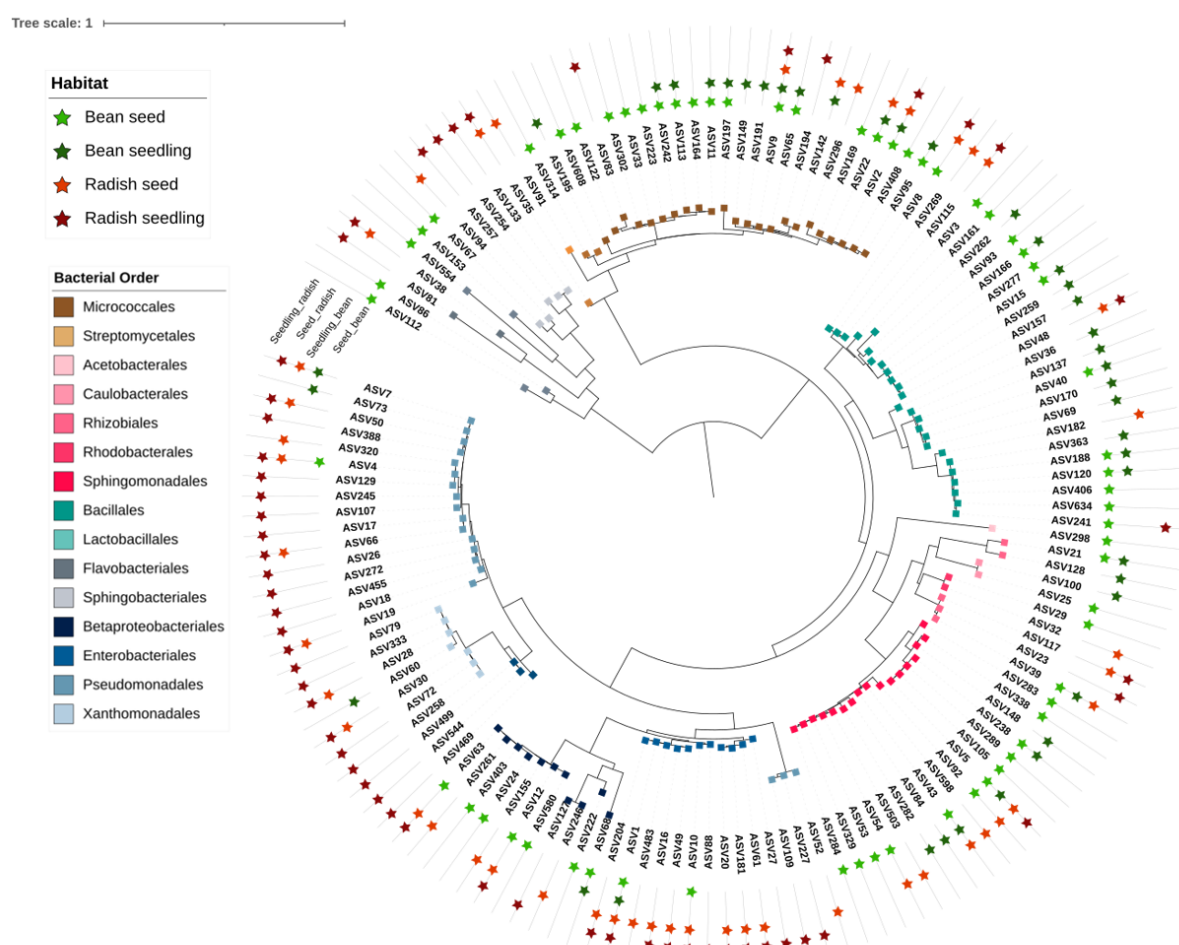


Figure S11 : Phylogenetic diversity of 16S ASVs associated with seeds and seedlings. Neighbor-joining phylogenetic tree of bacterial ASVs. Stars represent ASV present in bean seed (light green), bean seedling (dark green), radish seed (light red) or radish seedling (dark red). Colors represent taxonomy at the Order level.

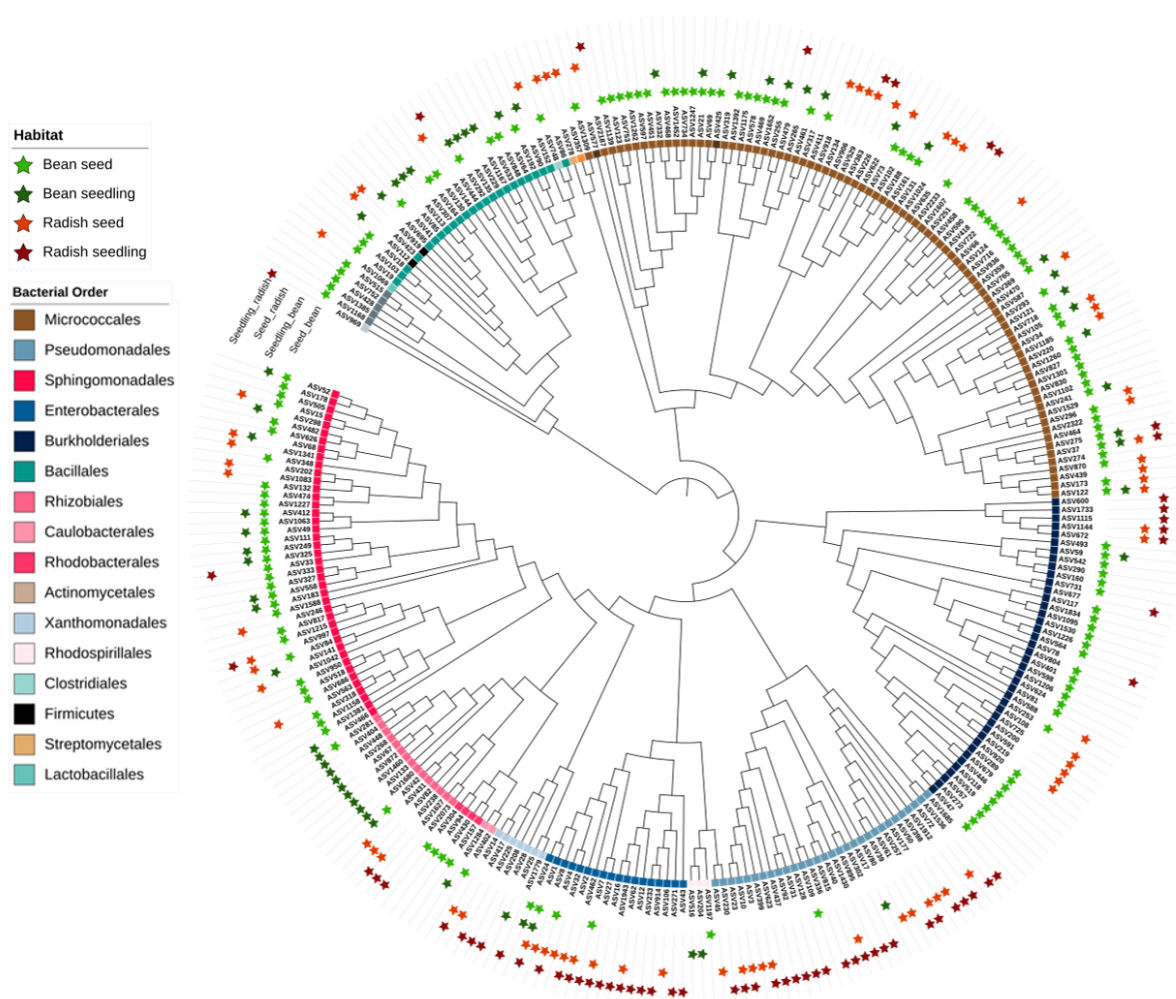


Figure S12 : Phylogenetic diversity of *gyrB* ASVs associated with seeds and seedlings. Neighbor-joining phylogenetic tree of bacterial ASVs. Stars represent ASV present in bean seed (light green), bean seedling (dark green), radish seed (light red) or radish seedling (dark red). Colors represent taxonomy at the Order level.

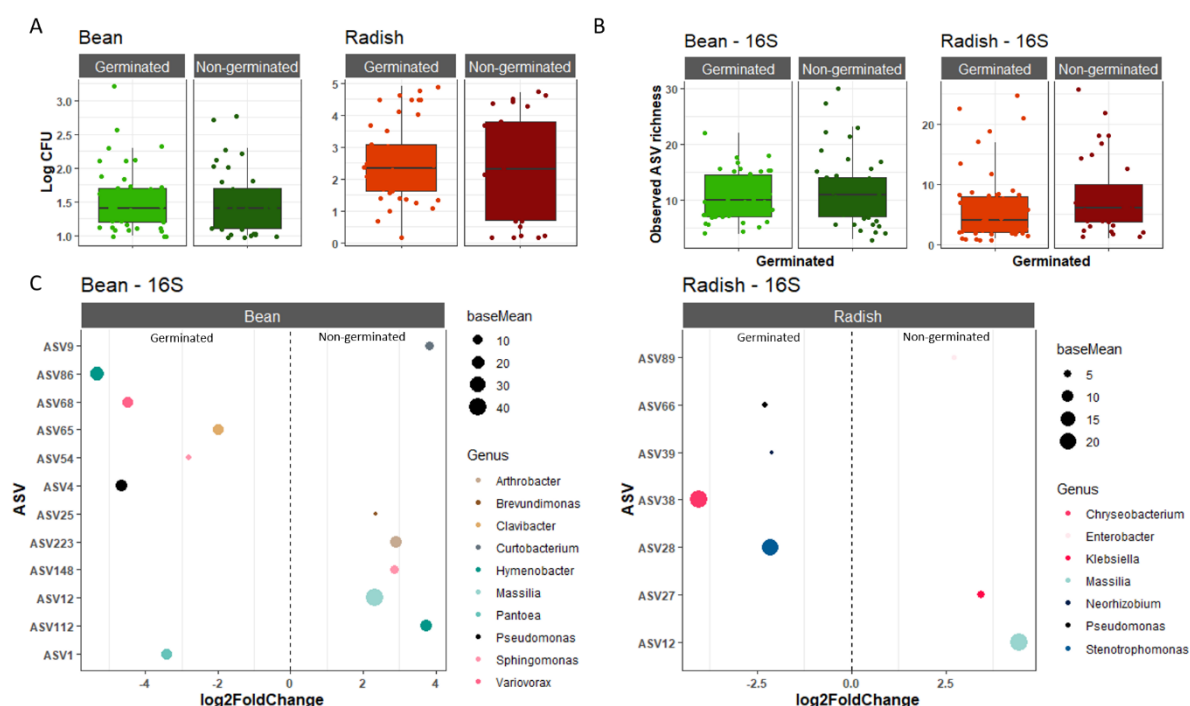


Figure S13 : Differences in seed microbiota profile (16S) between germinated and non-germinated seeds. A. Correspondence between initial bacterial population size (log10 CFU) and seed germination. **B.** Average richness in germinated and non-germinated seeds. **C.** ASVs significantly ($P < 0.01$) enriched in germinated seeds (left) or non-germinated seeds (right). ASVs are colored according to their taxonomic affiliations at the genus level.

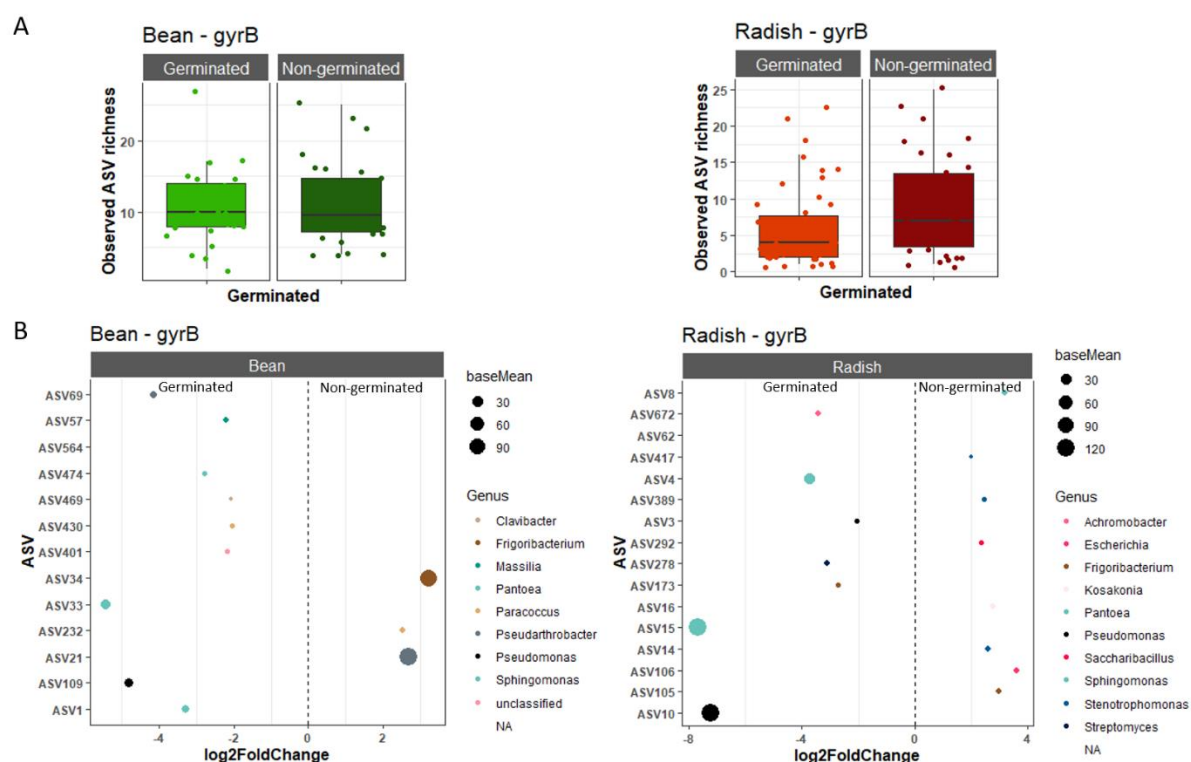


Figure S14 : Differences in seed microbiota profile (*gyrB*) between germinated and non-germinated seeds. A. Average richness in germinated and non-germinated seeds. B. ASVs significantly ($P < 0.01$) enriched in germinated seeds (left) or non-germinated seeds (right). ASVs are colored according to their taxonomic affiliations at the genus level.

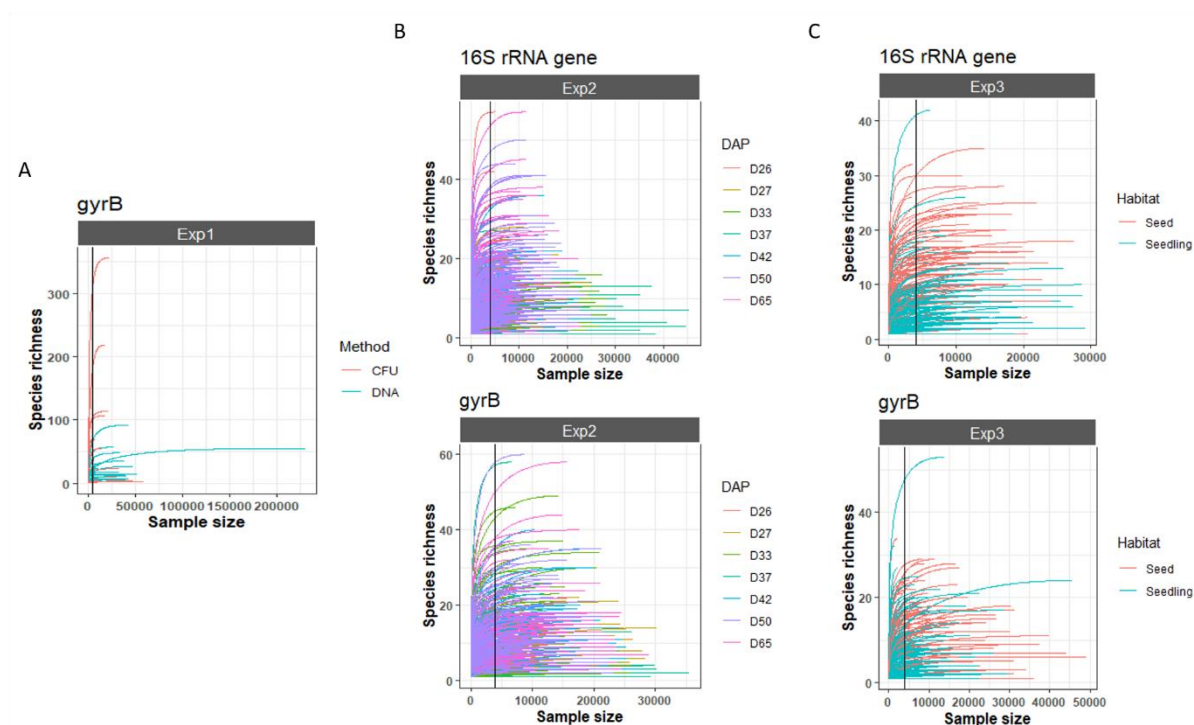


Figure S15 : Rarefaction curves for all data sets. A. data set 1 B. data set 2 C. data set3. Black vertical lines correspond to a rarefaction threshold of 4,000 reads.

Plant species	Plant	16S rRNA gene					gyrB				
		ASV	Taxonomy	Frequency	Residual	Stem (RA%)	ASV	Taxonomy	Frequency	Residual	Stem (RA%)
Bean	D27_P1	ASV6	Erwiniaceae	0,73	4,95	24,52	ASV5	Erwiniaceae	0,47	3,27	58,23
Bean	D27_P2	-	-	-	-	-	ASV2	<i>P. agglomerans</i>	0,20	3,13	nd
Bean	D27_P5	-	-	-	-	-	ASV1	<i>P. agglomerans</i>	0,64	3,57	45,18
Bean	D33_P6	-	-	-	-	-	ASV1	<i>P. agglomerans</i>	0,44	3,82	20,09
Bean	D50_P2	ASV4	<i>Pseudomonas</i>	0,35	4,12	nd	ASV11	<i>P. putida</i>	0,44	7,02	nd
Radish	D26_P1	ASV4	<i>Pseudomonas</i>	0,38	3,87	8,98	-	-	-	-	-
Radish	D26_P1	ASV16	Erwiniaceae	0,24	4,70	nd	ASV20	<i>P. ananatis</i>	0,25	4,66	nd
Radish	D26_P4	ASV4	<i>Pseudomonas</i>	0,44	3,53	99,97	-	-	-	-	-
Radish	D26_P5	ASV1	<i>Pantoea</i>	0,71	3,67	nd	ASV1	<i>P. agglomerans</i>	0,77	3,73	nd
Radish	D65_P3	ASV13	<i>Chryseobacterium</i>	0,53	4,70	NA	-	-	-	-	-

Table S1 : ASVs significantly associated with one individual plant at each stage of development.

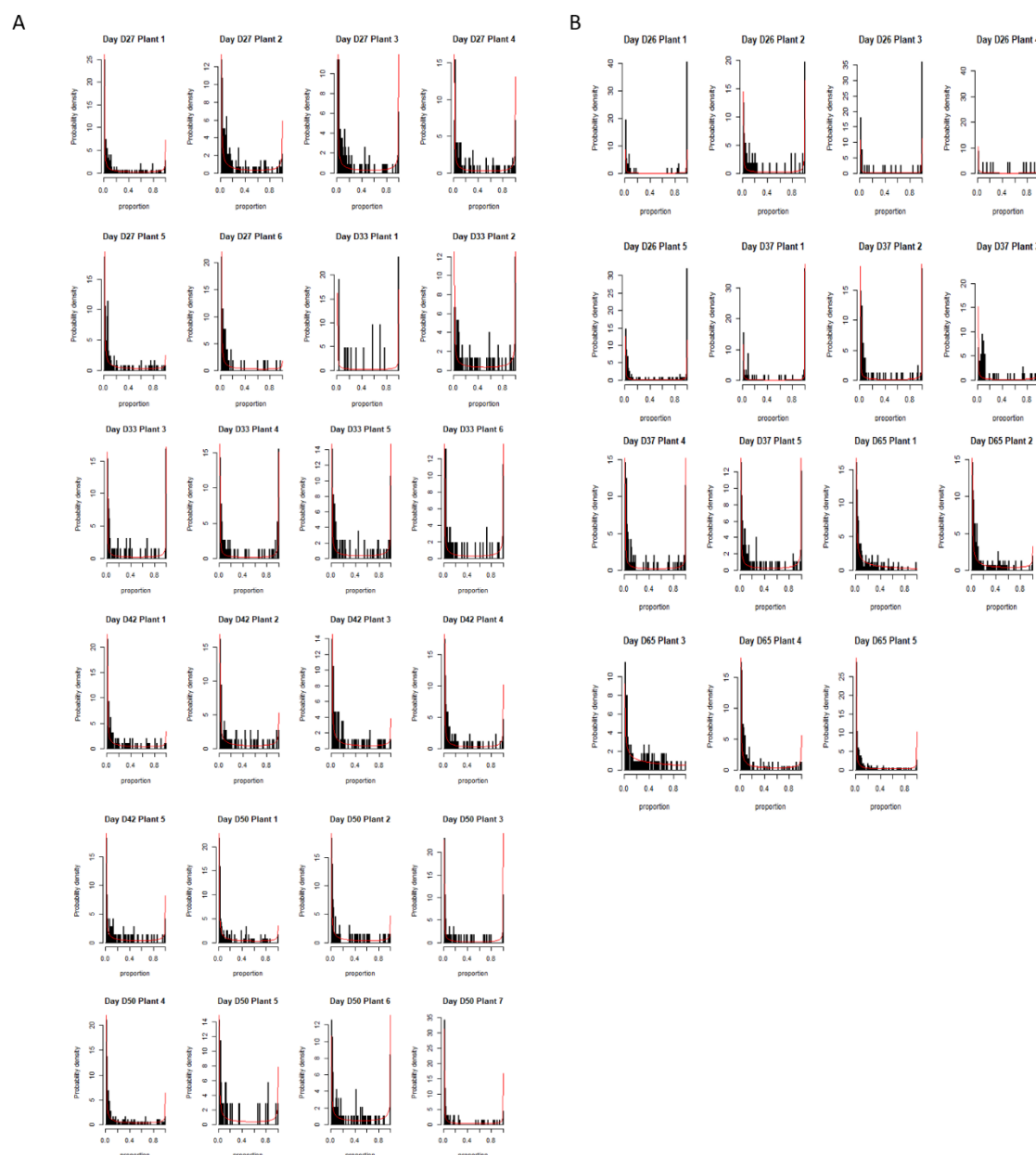


Figure Ss1 : The empirical and estimated probability densities of proportions. Probabilities were estimated for each day and each plant of bean (A) and radish (B).