Responses to recommenders

We thank the recommenders for their thorough review of our article and their constructive remarks. The paper was significantly revised and the tracked change version can be found at this link (file too large for upload on PCI website – limit size at 5Mb): https://1drv.ms/b/s!AhEwQWtXbYPTi5k2pWXIPegIrog76g?e=C4eKjR
The modified text is highlighted in yellow and our responses point by point are detailed below in blue.

Review 1

Overall, the study describes an elegant approach to cope with variability of natural microbiota using Syncom. The manuscript describes a well designed study with relevant interpretation.Proper controls have been used at different stages of the experiment but their interpretation could be stated/improved (see comment hereunder)

Thank you for your review and constructive comments, we have improved the description and interpretation of the controls (see below).

- Bacterial taxa were added and followed. There is no information on fungal or protists taxa. Do they have also a potential influence ? It might be interesting to add a sentence explaining their relative importance (or not) somewhere in the article (maybe in the introduction).
Thanks for this comment, the diversity of fungal taxa on seeds is well characterized and is often similar to bacterial diversity (Simonin et al. 2022 New phytol), but their role and transmission at the seedling stage is not well known. To our knowledge, there is currently no studies on seed protist communities.
We designed this study with only bacterial isolates because we do not have a fungal strain collection available and our expertise is mainly in bacteriology. However, initially in this study we planned to assess the effect of the bacterial SynComs on the fungal community assembly of seedlings. “Unfortunately”, the surface-disinfection seemed to have been very efficient at removing fungi from our seeds, and most of the ITS amplicon sequencing on seedlings failed (95 out of 112 seedling samples), so we decided to not include these results in the article. In a future study, it would be interesting to replicate this when seeds are grown in soil to see how SynCom inoculation influence fungal recruitment from the soil.

Abstract

Line 31-32: “These results confirm that the plant core microbiome includes pathogenic and not only commensal or mutualistic taxa” -> It is a very interesting and strong statement. The pathogens can be detrimental in some conditions and at least in the conditions tested during the experiment. At evolutionary point of view, how can it be explained/interpreted as the presence of pathogens could negatively impact fitness and survival ?
Yes, good question, we think that most of the time their abundance is not as high as in synthetic communities and their negative impacts are under the control of the other members of the microbiota. Addressing this question will require setting up specific experiments to validate it and this current experimental design do not permit to

address it. Hence for now, we prefer to just state this observation and do not speculate on the evolutionary implications of plant-microbiota interactions in the paper.

Line 33-35: the impact of Syncom on seed and seedling microbiota is underlined. I interpret that the fitness is related to the plant but it can be stated more clearly -> ".. and cause strong plant fitness differences between native…"
The sentence was not clear and was re-written, we meant the "fitness of the strains". New sentences:
Altogether, these results show that SynCom inoculation can effectively manipulate seed and seedling microbiota diversity and modulate plant phenotypes. This study also highlights strong differences between native seed-borne taxa in the colonization and survival on these plant habitats.

Introduction

Lines 77-78: the statement "Despite the crucial role of seeds for food production and maintenance of plant biodiversity, microbiome studies on the seed compartment are still a minority (Shade, Jacques and Barret 2017)". The citation is 6 years-old, is it still the case in 2023 ? Were there many experiments in this timeframe ? Some examples are given further but it could be relevant to use past term if the 2017 reference is maintained or to state it is still the case today. Explaining new research in between would be interesting
Line 81 "Recent studies are starting to be published" -> recent studies have been published… + are there many other studies ? Only two are cited and it is not clear how they were selected for being highlighted

Agreed, we significantly rewrote this section of the introduction. It is true that many studies characterizing the diversity of seed microbiota has been published in the time frame. However, to our knowledge there is only one SynCom study available on seeds (Figueiredo dos Santos et al. 2021) compared to dozens of studies available on the root or phyllosphere compartments. This section now reads as follows:
Recent studies using synthetic microbial communities (SynComs) to assess the role of the microbiota on plant nutrition or resistance to pathogens have been performed on the rhizosphere and phyllosphere compartments (e.g. Kwak et al. 2018; Carlström et al. 2019; Finkel et al. 2020). But so far, the seed compartment and the assembly of microbiota in the early life of the plant has been neglected. One correlative study between seed microbiota structure and seed germination of different rapeseed genotypes offers promising results about the key role of microbiota in seed vigor (Rochefort et al. 2019). One SynCom study by Figueiredo dos Santos et al. (2021) showed that seed disinfection reduced maize germination rates that could be recovered after inoculation of a SynCom on seeds. Additionally, Matsumoto et al. (2021) demonstrated the role of seed endophytes in disease resistance to the seed-borne pathogen *Burkholderia plantarii*. These reports encourage further investigations on the influence of seed microbiota on seedling microbiota assembly to uncover the role of this primary inoculum of plants.

The introduction is clear but I missed some comments on the different compartments where bacteria can be present on seeds (surface or within tissue…). It is indirectly

addressed when mentioning disinfection but it worth refining the information for the readers. Reading material and methods, the reader further understand the work is carried out on microbes present on the surface of the seed but what about the endophytes ? This could be added

Agreed, we now include a sentence on this at the end of the introduction and a paragraph in the method section. Actually, the seeds were surface-disinfected prior to bacterial inoculation but for the bacterial collection isolation and microbiota characterization we studied the entire seed microbiota (endophytes + surface bacteria).

Sentence added in the introduction:

In this study, we considered the entire seed and seedling microbiota, including both endophytes and bacteria living on the plant surfaces.

Paragraph in methods:

A bacterial culture collection was obtained from a seed sample composed of approximately 1000 mature seeds (Torres-Cortés et al. 2019). Seed samples were soaked in 25 ml of phosphate-buffered saline (PBS, Sigma-Aldrich) with Tween® 20 (0.05 % v/v, Sigma-Aldrich) at 6°C under agitation (150 rpm) for 2h30. This seed soaking technique has been developed by the International Seed Testing Association for the detection of bacterial pathogens, including the ones located in seed internal tissues thanks to seed imbibition. Hence, this technique permits to characterize and isolate both the endophytes and the bacteria living on the seed surface. This technique has been validated and compared to grinding in a previous study from our team (Chesneau et al. 2022).

Some information related to the radish seed microbiota can be added in the introduction

Thank you for this suggestion, we added the following sentences:

In particular, our work on radish seeds showed a high variability in richness (1 to 46 bacterial taxa, Figure S1) and composition between individual seeds (Chesneau et al. 2022). An important finding was that seeds were generally dominated by one bacterial taxon (> 75% relative abundance) but its identity was highly variable between parental plants and even between seeds originating from the same plant.

Material and methods

Line 112-114: transferring the number of strains and their phyla, families and genera in part 1 of results

Done, we transferred this sentence to the beginning of the result section

Line 138: how was this concentration selected ? Is it because it allowed to achieve concentration of strains or Syncoms comparable to concentration of native microbiota ?

This concentration was selected based on previous tests at different inoculum concentration in our lab and also based on the SynCom and biostimulant/biocontrol literature on plant inoculation were this is the most common concentration used to achieve a good colonization (e.g Kwak et al. 2018, Pfeilmeier et al. 2021, Vogel et al. 2021). Our previous study on single seed microbiota characterization of radish seeds demonstrated that the bacterial load is extremely variable from one seed to another (10 to 10^6 CFU/seed, Chesneau et al. 2022)

but generally lower than the inoculated concentration here. We are now including this sentence in the Method section to justify our choice:

This concentration was selected based on previous seed inoculation experiments conducted in our team (Chesneau et al. 2022) and multiple reports of this concentration used in synthetic community inoculation on plants (e.g Kwak *et al.* 2018; Pfeilmeier *et al.* 2021).

Line 141: what was (were) the control ?
Thanks for this comment, we included the following sentence:
Control seeds were inoculated with sterile water only.

Line 143: “The bacterial cell density of the inocula and inoculated seeds (pool of 30 seeds) were assessed by plating on TSA 10% (CFU/mL or /seed)” -> at which timing were analysed the inoculated seeds ?
We added the information:
The bacterial cell density of the inocula and inoculated seeds (pool of 30 seeds) were assessed by plating on TSA 10% (CFU/mL or /seed) immediately after inoculating the seeds

Line 153: the link goes to homepage of ISTA and not the protocol
I changed the link to go to the Seedling evaluation handbook used:
https://www.seedtest.org/en/handbooks-calibration-samples/seedling-evaluation-4th-edition-2018-product-1016.html

Line 173: 35 cycles of PCR were carried out. It is more frequent to carry out 25 to 30 cycles to avoid representation bias at the end of exponential phase of the PCR. Can you comment on this specific point ?
Yes, this is a high number of cycles, we optimized this PCR by comparing 15, 25, 35 cycles on a mock community compared to a characterization with shotgun sequencing. We observed that increasing the number of cycles distort more the community structure but it also permits to amplify better communities with low biomass like seed matrices. It is a compromise between accurate community structure and amplification efficiency from low microbial biomass matrices.

Line 204: 14,116 reads

There is no information on the way the alien control (Lactococcus piscium) is used to monitor contamination among the samples. Could you explain how you took into account these results in the global analysis (for example to set up contamination threshold as recommended in the following guidelines on HTS use for plant pest detection : https://peercommunityjournal.org/articles/10.24072/pcjournal.181/) ?
We use this alien control as an estimate of index hopping and also as an additional negative control to identify potential cross-contaminations. The index hopping rate of this sequencing run was similar to expected Illumina rates (~0.01%).

Results

-point 1 – comment 1 : The selection process is well described but you did not mention before in M&m that 523 strains were isolated and sanger sequenced for species identification (see comment for line 112-114)
Done see previous response

- Point 1 – comment 2: Did you also took into account already known beneficial or pathogenic effect related to the selected species ?
Thanks for this question, our strain selection was purely performed based on prevalence-abundance profiles and phylogenetic diversity without any a priori on functions or known effects on plant phenotypes. Our objective was to study the transmission and effect of seed microbiota as it can occur in nature, with commensals, mutualists or pathogens. We knew that *Pseudomonas viridiflava* was a leaf pathogen but it is also extremely abundant and prevalent in seeds. We were interested in seeing if it had any effect at the germination and seedling stage in isolation and community settings. The result section has been updated to clarify this point:
The twelve strains were selected to be representative of the diversity and prevalence-abundance profiles of radish seed microbiota (Figure 1A), without any a priori on their potential functions or effects on plant phenotype. The objective of the study was to study realistic seed bacterial communities that could include any type of taxa (pathogens, commensals, mutualistic), as in nature.

- Point 2: a native microbiome (e.g. non disinfected seed) could have been useful as additional conditions. You pointed the high variability observed elsewhere (and discuss it) but it might have added interesting information to compare naive and Syncoms (as for the other analyses). It might be relevant to discuss this point (maybe first part of discussion) ?
Thank you for this suggestion. We also characterized the native microbiota of seeds before surface-disinfection and we are now including these results in Figure S2. We included the following sentence in the results:
The microbiota comparison of native and surface-disinfected seeds suggests that remaining endophytes are mainly dominated by Pseudomonas species and Pantoea agglomerans likely have high epiphytic abundance (Figure S2).
Regarding your point on the high variability, Figure S2 suggests that surface-disinfected seeds have a higher variability than native seeds, so it is not changing the conclusion of Figure 3B on the reduced beta-dispersion following SynCom inoculation. Following this comment and other reviewers' suggestion, we have now developed the beginning of the discussion to address this topic.

- Figure 3A: I interpret the figure with the fact that seed and inoculum are very close to each other (superposition). It seems confirmed by fig 3D. Is it the case ?
Yes, this is the case, the inoculum and seed symbols are superposed, we increased the size of the figure make it more apparent.

- Line 300: "The control seeds were surface-disinfected but they still harbored a low bacterial diversity likely of endophytic bacteria, including ASVs of strains included in the SynComs because the strains selected have been isolated from the same radish genotype". Does it also explain the much higher variability observed on fig 4A ?

Yes, the remaining diversity of endophytic bacteria in the control seeds was variable between samples (5 to 12 ASVs), the full taxonomic description of control seeds is now presented in Figure S2.

Line 316: the genera names can be abbreviated (please check on the complete document)
We have decided to not abbreviate the genera names to facilitate the reader understanding of the paper, especially due to several genera starting with the same letter (Pseudomonas, Pantoea, Paenibacillus Erwinia, Enterobacter…).

Figure 4B: when looking the results of the control for the seedlings, it seems there are more than a median of 3 taxa. I understand therefore that "other taxa" corresponds to a single or two ASV ? If so, is it the same taxa between replicates or it varies ? It might worth to consider these "other taxa" for better overview of the results.
Thanks for this comment, we are now including the taxonomic description of control seedlings in Figure S3. The control seedlings generally had a low diversity (most dominated by 2 to 5 ASVs) and the composition was quite variable, due to the fact that each disinfected seed likely have a distinct seed endophytic microbiota. Still, the seedlings were generally dominated by 3 species (*Pantoea agglomerans*, *Pseudomonas viridiflava* and *Erwinia persicina*) but that did not necessarily have the same ASV than the strains included in the SynComs. We are now including these sentences in the results:
A complete taxonomic description of the control seedlings (including the "other taxa") is available in Figure S3. It shows that control seedling bacterial community is highly variable between individuals and is generally dominated by 2 to 9 ASVs.

Discussion

- Lines 450-456: discussion on P. fluorescens is interesting but It can be noted that there are pathogenic strains as well as beneficial strains from this species that are studied from a long time. Here, a more subtle distinction is raised.
Actually, none of the 4 *P. fluorescens* strains had detrimental effects, so this aspect is not discussed in the paper. We mainly developed our discussion on their differences in colonization and transmission success to seedlings that was highly strain-specific.

- Lines 464-470: the detrimental effect of P. viridiflava might not be surprising as it is a known plant pathogen. Has it been selected on purpose to include a pathogenic strain ? If so, it could be mentioned in when explaining species selection
Agreed, we were not surprised of P. viridiflava detrimental effects but this strain was selected based on its high prevalence-abundance profiles in radish seeds that made it a "core strain". We now present more clearly our rationale for the strain selection earlier.

- Line 473-475: how were the core taxa defined ? It worth reminding the criteria for being considered as core taxa (prevalence threshold ?) to better contextualize this term (maybe adding the number of species considered as belonging to core taxa as a comparison: three bacterial strains among … species of the core taxa ? This comment

is also related with previous comment on giving more details on core microbiota of seeds/seedlings.
Thanks for these comments, we are now included more information on the core taxa analysis. In the initial meta-analysis, the core ASV had to be detected in a minimum of two independent studies (out of the 7) and be present in at least half of the samples (minimum prevalence 50%). Here, we strengthened this criterium for the strain selection and included core strains with > 80% prevalence. We selected 3 core strains out of the 9 identified.

- when mentioning Paenibacillus, “spp.” should be always added in the text (correction needed at least for line 488)
Done

Reviews 2
Reviewed by Sebastian Pfeilmeier, 11 Mar 2023 21:54
Review of manuscript by Simonin et al. for PCIMicrobiology

Simonin et al. used a synthetic community approach to study the transmission of radish-associated bacteria from seeds to seedlings and examine the bacterial community composition under different conditions. In addition, the authors characterize the impact of those bacteria on seed germination and plant health.

The study is well designed and of interest to the plant microbiota research community and beyond. It covers a relevant topic as the seed microbiota can have an impact on all stages of the lifecycle of plants. The data was carefully analysed and interpreted. I agree with the main conclusions of the authors, but still have a few suggestions to improve the manuscript:

· The figures are nicely composed. However, often the figure quality is quite low and data points or labels blurry, which makes it difficult to read. I recommend increasing figure resolution and size of text labels in the figures throughout the manuscript.
Thanks for this comment, we improved the quality and size of the text. We also split Figure 5 in two different figures to improve legibility.

· Figure 6: use statistical test to show significance of treatments.
Thanks for this comment, we are now including symbols in now Figure 7 showing the treatments with significant effects relative to the control condition. We have updated the figure caption:
Figure 7: A) Effect of the inoculation of single bacterial strains or synthetic bacterial community on germination and seedling phenotypes. The asterisks represent the significance of the Fisher exact test that compares the proportions of normal and abnormal seedling phenotypes between control and inoculated seeds: ***P<0.001, **P<0.01. B) Photography of the typical phenotypes observed in the experiment (Credit: Guillaume Chesneau).

· Figure 7: panel labels are missing.
Good catch, we updated the figure (now Figure 8)

· Line 21: it did not get clear to me whether the entire seedling, or only shoot was harvested.
Thanks for this comment, we clarified in the method that we analyzed the entire seedling (shoot + root) to analyze the overall transmission of the SynCom. Our previous seedling microbiota analyses under in vitro conditions demonstrated no differentiation between the root and shoot compartments on the colonization of the SynComs after 4 days of growth (not statistically different). So, in this study we analyzed the microbiota of the entire seedlings. Updated text:
To measure the bacterial cell density and the bacterial communities of seedlings, each entire plant (shoot and root together) was crushed in a sterile plastic bag (n=30 seedlings by condition), resuspended in 2 mL of sterile water and homogenized.

· Line 29: if the authors base the definition of those strains as “core taxa” on a previous study, I would add “previously identified” in the text. Otherwise, mention that this finding is a result of a meta-analysis of 16S datasets (because based on a SynCom of only 12 strains, I would not be comfortable to deduce “core strains”).
Thanks, we made the change in the abstract, and the core strains were identified in the previous meta-analysis based on gyrB metabarcoding (better taxonomic resolution than 16S at the species / sub-species level) based on 7 independent radish seed studies.

· Line 30: maybe highlight that core strains showed increased abundance in diseased seedlings, without being virulent themselves (at least for Pantoea and Erwinia)
Not sure to understand this comment, sorry. We have no way to know if the strains were virulent themselves in SynComs, especially for Erwinia that was found to be enriched in abnormal seedlings but did not cause symptoms on its own. Pantoea agglomerans was not enriched in abnormal seedlings.

· Line 130: if I understand correctly, the seed surface sterilization does not kill all bacteria. This only became clear to me later in the manuscript. I recommend to mention here or early in the manuscript (the latest in line 275).
Thank you for this suggestion, we added this information early in the M&M section: While, the surface-sterilization removed all surface bacteria as expected, this method cannot eliminate the seed endophytes that colonize the internal tissues of the seed without impairing seed physiology and germination potential. Hence, the experiments were performed on seeds that carried endophytes that were subsequently monitored in non-inoculated control seeds and seedlings.

· Line 231: “phylogenetic diverse” is relative to the overall diversity. Would you consider your SynCom strains representative of the radish microbiome composition?
Yes, we think that our SynComs are representative of the radish seed microbiota because the taxa chosen (core, sub-abundant, rare) and the community richness levels are similar to what we observed on single seeds (Chesneau et al. 2022 mBio). Moreover, all the strains were isolated from the same radish genotype, so they are adapted to the colonization of this seed habitat. Of course, we selected only 12 strains out of hundreds, but all this diversity is never observed on a single seed, so here we reconstruct one possible community assembly out of many possibilities. In response to several previous comment, we tried to justify our rationale.

· Line 234: it would be great to make the gyrB gene sequences available in the supplement.
Thanks for this suggestion, we are now including the sequences and the following sentence:
The *gyrB* sequences of the selected strains and of all bacteria detected in the study (seeds and seedlings) are available in Supplementary Data.

· Line 394: the authors identified phenotype as significant driver of microbiota composition. I recommend to mention the variance explained as shown in the figure 7 also in the text.
Thanks, we added the R^2 in the text.

· Line 415: I am a bit confused about the statement, that the 12 ASVs of the SynCom are also part of the "sterilization-surviving/endophytic" strains. Why do these strains not appear in the richness analysis, as all 12 strains would also be present in the 8 member SynCom. Please clarify.
· Line 417: Why do the authors think that the inoculum becomes dominant if all 12 strains are also naturally present in the seeds. Is it due to low abundance levels in seeds that lead to high variability?
Thanks for these 2 comments that are related, we understand your confusion that is due to the fact that we are looking at community profiles without considering the microbial biomass of the seeds. Some of the ASVs of the strains are detected in surface-sterilized as they are likely endophytes (Figure 4B, Pseudo viridiflava, Pantoea, Erwinia). However, the CFU measured on control seedlings show that the bacterial population size is close to 100 CFU/seedling, which is really low and the addition of the SynCom completely take over these native endophytes in terms of abundance. So, the addition of the SynComs completely "submerge" or "swamp" the native microbiota and enables to homogenize and reduce the natural variability from seed to seed.
We modified this part of the discussion to clarify this point:
This homogenization of community structure is due to the high abundance of inoculated strains (10^7 CFUs per seedling), which makes the members of the native endophytic community (10^2 CFUs per seedling) undetectable. This similarity of seed microbiota between seed replicates is a prerequisite to study the transmission of microorganisms from seed to seedling and establish causal relationship between seed microbiota and plant microbiota or phenotype.

· Line 427: What does "ex" mean?
It means "example", I changed it to "e.g"

· Line 435: I would rephrase "show" to "suggest", as it has not been directly tested in an experiment with varying "dominance/inoculation levels".
Done
· Line 448: I would rephrase "difficult to identify in community context", because a major advantage of SynCom experiments is the possibility to add or remove individual strains and test the impact on the community. This has been shown in various studies, e.g. from lab of Julia Vorholt (ETH Zurich)
Thanks, we have rephrased the sentence:

These SynCom-specific patterns of each strain are likely driven by exploitative and interference competition between individuals that necessitate subsequent SynCom experiments to identify the specific mechanisms expressed in planta.

· Line 466: A striking example how prevalent pathogenic P. viridiflava can be in the leaf microbiome of plants: Karasov et al. 2018, Cell Host Microbe; Lundberg et al 2022, PNAS
Thanks, we rephrased this section and included the references.

· Line 477: The findings are in line with another study showing that opportunistic pathogens can be part of the microbiota from healthy looking plants. It would be worth to discuss Pfeilmeier et al. 2021, Nat. Microbiol.
Thanks for this suggestion, we have rewritten the end of the paragraph as follows:
Another important observation was that the three bacterial strains identified as seed core taxa (Pseudomonas viridiflava, Pantoea agglomerans, Erwinia persicina) were associated with detrimental effects on seedling phenotypes either in isolation or in SynComs. These results are in line with other studies showing that the plant core microbiota includes pathogens or conditional pathogens that have a pathogenic potential but are generally kept in check by other members of the microbiota and the host (Pfeilmeier et al. 2021, Simonin et al. 2020).

· Line 489: It is interesting indeed that other “non pathogenic” taxa benefit from a diseased plant. Similar observation has been made in Pfeilmeier et al. 2021, Nat. Microbiol.
Thanks for this comment, we included the citation

· Line 498: The protection by commensal microbiota members and small SynComs against pathogenic strains have been experimentally tested and can be discussed, e.g. Vogel et al. 2021, Nat. Microbiol.; Shalev et al. Nat Ecol Evol. 2022
Thanks for these suggestions, we rewrote this part of the discussion that now includes these citations:
The reduction of detrimental effects in the 12-strain SynCom experimentally confirms the importance of bacterial interactions in plant disease development in a pathobiome context (Bass et al. 2019). This observation is supported by other SynCom experiments on the phyllosphere demonstrating the role of single or consortia of commensal microorganisms (e.g Pseudomonas) in disease protection (Vogel et al. 2021, Shalev et al. 2022).

Review 3

The topic of this paper is very interesting and is a priority concern for downstream applications such as beneficial microorganisms application. The methodology is well-designed, and the presented results are interesting. Nonetheless, some recommendations to improve the submitted work are described hereunder.
General comments:

• What I retained from the introduction is that we know nothing. It underlines the interest of the study, but the topic could be better introduced as the literature is growing in this area. The problem is underlined, but its context could be developed.

Agreed, we rewrote great parts of the introduction to address this comment and some from the other reviewers. Here is the end of the intro rewritten on the context:

Recent studies using synthetic microbial communities (SynComs) to assess the role of the microbiota on plant nutrition or resistance to pathogens have been performed on the rhizosphere and phyllosphere compartments (e.g. Kwak et al. 2018; Carlström et al. 2019; Finkel et al. 2020). But so far, the seed compartment and the assembly of microbiota in the early life of the plant has been neglected. One correlative study between seed microbiota structure and seed germination of different rapeseed genotypes offers promising results about the key role of microbiota in seed vigor (Rochefort et al. 2019). One SynCom study by Figueiredo dos Santos et al. (2021) showed that seed disinfection reduced maize germination rates that could be recovered after inoculation of a SynCom on seeds. Additionally, Matsumoto et al. (2021) demonstrated the role of seed endophytes in disease resistance to the seed-borne pathogen Burkholderia plantarii. These reports encourage further investigations on the influence of seed microbiota on seedling microbiota assembly to uncover the role of this primary inoculum of plants.

• What part of the seed is considered for the microbial study? Spermosphere, spermoplane, or endosphere? This should be described in the introduction (the different niches existing in seeds)

Thanks for this comment. In the study, we considered the entire seed microbiota, we didn't attempt to separate just the bacteria living in the surface or inside seed tissues. To do so, our bacterial collection and microbiota characterization (including in the meta-analysis presented) were performed using a seed soaking technique that permit to have access to both endophytes and spermoplane bacteria. Similarly, for seedling microbiota we didn't separate the different compartments.

We added the following sentence at the end of the introduction and also a more detailed explanation in the Method section (see response in Comments line by line):

In this study, we considered the entire seed and seedling microbiota, including both endophytes and bacteria living on the plant surfaces.

• It is not clear what part of the seedling is used for the microbial study (i.e., leafy part, seed, or root). Crushing the entire seedling, including the remaining inoculated seed, may introduce a bias. This should be better described, and the potential bias of sampled tissue discussed.

Thank you for this comment, we clarified this point in the M&M:
To measure the bacterial cell density and the bacterial communities of seedlings, each entire plant (shoot and root together) without the remaining seed and tegument was crushed in a sterile plastic bag (n=30 seedlings by condition), resuspended in 2 mL of sterile water and homogenized.

• Texts in figures are sometimes too small to be read.

Thanks for this comment, we improved the quality of the figure and size of the text.

• Titles and subtitles could be shortened and standardized. For example, line 371 explains a result directly.

Thanks for this comment, the figure captions were revised to have similar structure and a shorter text

• Passive form should be preferred instead of "we"

We rewrote most sentences at the passive form

• The discussion part tends to repeat results which might be better compared to the literature instead.

• In the discussion, you often write your results, and at the end of the sentence, you add references. What does it mean? That they obtain the same results as yours? See comments line by line for example.

Sorry about this, yes the references added at the end of the sentence were showing support for our observation. We now have rephrased the sentence to make this clearer.

• Overall, the bibliographical support for the introduction and the discussion could be improved.

Thanks for this comment, we worked on this aspect as also suggested by reviewer 2. The end of the introduction and great parts of the discussion were significantly rewritten and we include now more discussion of our results with 9 new citations added.

Comments line by line:

• 33-35: It should be good to rewrite this sentence.

This last sentence was rewritten as follows:
Altogether, these results show that SynCom inoculation can effectively manipulate seed and seedling microbiota diversity and modulate plant phenotypes. This study also highlights strong differences between native seed-borne taxa in the colonization and survival on these plant habitats.

• 87-91 and 92-95: it is a little redundant.

Agreed we rephrased the paragraph as follows:

In this context, we set up seed inoculation experiments under in vitro conditions to address the following objectives:

- Characterize the transmission of individual seed-borne bacteria and synthetic bacterial communities from seed to seedlings (here on radish plants)
- Determine whether individual seed-borne bacteria or synthetic bacterial communities can impact seedling phenotype.

Twelve bacterial strains representative of radish seed microbiota were selected based on their abundance – prevalence in seeds without any a priori on their potential functions or effects on plants. The strains were studied either individually or in communities (mix of 6, 8, or 12 strains). In this study, we considered the entire seed and seedling microbiota, including both endophytes and bacteria living on the plant surfaces.

• 100: strain of what? Bacteria, fungi, both?

Just bacteria, we updated the text

• 105: so, it was spermoplane that was harvested.

The seed soaking enables to characterize the entire seed microbiota, including both the endosphere and spermoplane. We clarified the text in the introduction and in the material and method section:

This seed soaking technique has been developed by the International Seed Testing Association for the detection of bacterial pathogens, including the ones located in seed internal tissues thanks to seed imbibition. Hence, this technique permits to characterize and isolate both the endophytes and the bacteria living on the seed surface. This technique has been validated and compared to grinding in a previous study from our team (Chesneau et al. 2022).

• 136: how the SynCom was composed? Were the bacterial strains added in equal proportion? The final concentration is 1x107 cfu/ml individually or for all the syncom? Rare taxa stay rare in the Syncom, and the inverse too?

Thanks for these comments, our experience on SynCom inoculation on seeds is that the initial profile of the community in the inoculum is completely modified at the moment of seed colonization (see Figure 4 B) because of differences in strain adhesion. Hence, we prefer to inoculate the strains in equal proportion to give them the same chances to initially colonize seeds and then seedlings.

We improved the description of our methods in the following sentences:

The bacterial inoculations were performed on subsamples of 30 seeds by placing them for 30 minutes under agitation (70 rpm) at 20°C in a suspension at a final concentration of 10^7 CFU/mL for both single strains and SynComs (Optical density 600 nm = 0.01) from fresh 24/48-hour cultures. For the SynComs, each strain was added in equal proportion in the suspension to reach the final concentration of 10^7 CFU/mL to do not provide advantages to some strains through mass effects.

• 133: what were the criteria for selection? Abundant or rare species? (Ok, it is explained after)

• 138: Why this concentration? Is it what we naturally find on the seed surface?

This concentration was selected based on previous tests at different inoculum concentration in our lab and also based on the SynCom and biostimulant/biocontrol literature on plant inoculation were this is the most common concentration used to achieve a good colonization (e.g Kwak et al. 2018, Pfeilmeier et al. 2021, Vogel et al. 2021). Our previous study on single seed microbiota characterization of radish seeds demonstrated that the bacterial load is extremely variable from one seed to another (10 to 10^6 CFU/seed, Chesneau et al. 2022) but generally lower that the inoculated concentration here. We are now including this sentence in the Method section to justify our choice:

This concentration was selected based on previous seed inoculation experiments conducted in our team (Chesneau et al. 2022) and multiple reports of this concentration used in synthetic community inoculation on plants (e.g Kwak *et al.* 2018; Pfeilmeier *et al.* 2021).

• 157: if you crush the seed, you also will have endospheric microorganisms. Could you distinguish the different strains of SynCom in plate? Was it recorded?

• 157: did you remove the seed before crushing? If not, how to know that your inoculum was transmitted to the seedlings (leaf and/or root) or if it was the remaining inoculum on your seed?

We respond to these 2 related comments on Line 157 here: Yes, the seed and remaining teguments were removed as they fell quite quickly from the seedling and as you said, to make sure that the inoculum truly colonized the seedlings. The text was clarified:

To measure the bacterial cell density and the bacterial communities of seedlings, each entire plant (shoot and root together) without the remaining seed and tegument was crushed in a sterile plastic bag (n=30 seedlings by condition), resuspended in 2 mL of sterile water and homogenized.

Unfortunately, we couldn't differentiate the 12 different strains on plate as some have very similar morphs. This is something that we are pursuing and would like to do in future SynCom experiments.

• Fig 1a: It is fine, but the meaning of the triangle is a little bit difficult to understand. Perhaps directly put triangles or circles in the legend? Only watching the legend will help to see which taxa belong to the core microbiota. But you will have a grey triangle and a grey circle for “other taxa” in the legend.

Thanks for this comment, we updated the legend of the figure following your suggestion.

• Fig 1: Why do you not select taxa belonging to the core microbiota for low relative abundant taxa (<1%)? Are Pseudomonas preferably selected instead of other bacteria, and why?

Thanks for these questions, we considered these 2 core taxa with a lower relative abundance but they belonged to

- a *Paenibacillus sp*. closely related to the one already selected

- a *Pseudomonas syringae,* that is a potential pathogen and several Pseudomonas strains were already included in the selection

Pseudomonas strains were not necessarily preferably selected instead of other taxa, they are just very abundant and diverse in seed microbiota. We added the following sentences to justify our choice:

In our meta-analysis on radish seeds, the Pseudomonas genus was the most diverse by far with 380 ASVs (25% of the diversity) that collectively represent 33.3% of the relative abundance of the dataset. Among Pseudomonads, the most diverse species was *Pseudomonas fluorescens* with 62 ASVs that justified its selection to study intra-specific variations.

• 272-274: Avoid “we” as much as possible. Prefer the use of passive form.

The sentences were rephrased.

• 262: if you have 2 log CFU/seedling, does the disinfection fail, or do you have bacteria inside the seed endosphere?

• 280: How do you explain these results for the control while your seeds were disinfected?

Thank you for these 2 comments. As now better stated in the method and result section, this disinfection is just a “surface-disinfection” and it is almost impossible to remove endophytes without completely impairing the germination capacity of the seeds. We are also now including the microbiota comparison of native vs surface-disinfected seeds, showing a shift in composition from Pantoea-dominated (Epiphytes) to Pseudomonas-dominated (Endophytes) seed communities (Figure S2).

• 335: Too much “we”. Perhaps it should be in material and methods.

The sentences were rephrased.

• Fig5: You will see with the editing process, but figures are numerous inside the same caption, making the text size too small.

Thanks for this comment, we have now split the Figure 5 in two different figures to improve the legibility.

• 371: Is it a title? It must be changed.

The title was changed into:

a. Identification of strains and SynComs increasing the proportion of abnormal seedlings and non-germinated seeds

• 443: “major environmental changes” → they are not so major compared to real germination conditions (soil,…). For example, if your experiments were conducted in a real environment, would your microorganisms be found in the seedlings?

Agreed, we removed “major”. We intended to say that the transition from a dry inert seed to a growing seedling is a large change of habitat for the seed microbes (burst of nutrients/sugars, oxidative stress…). We conducted the same experiments in a soil (the results will be soon submitted as Arnault et al) and actually we found that the inoculated strains represent 80% of the seedling microbiota after 7 days of growth, so they have an excellent transmission success.

• 420: Is it your results or those of the references cited?

• 417-420: Don’t forget that you disinfect your seed. Give more nuance in your affirmation.

• 422-427: Too many results and no discussion.

• 427-430: Quite poor as explication/discussion.

• 431-436: Too many results and no (or not enough) discussion.

• 440-459: ibidem

• 476-477: Is it your results or those of the references cited?

• 496: Is it your results or those of the references cited?

We addressed all these different comments by rewriting great parts of the discussion to remove some results and include more references to previous studies in order to interpret/discuss our results. We gave more nuance to our affirmation on the impact of SynCom inoculation on disinfected seeds.

• 476: microbiome or microbiota? Make the distinction.

We changed it to microbiota to be consistent in the text

• 486-488: what does it mean “different”?

We clarified the sentence but the identity of the strains associated with the abnormal phenotypes was not the same between the individual strains and the SynComs:
One is that the identity of the strains responsible for the abnormal phenotypes are different when inoculated in a community (*Enterobacter cancerogenus* and *Erwinia persicina)* or in isolation (*Pseudomonas viridiflava* and *Paenibacillus sp.)*.

• 486-490: do you have references to straighten your hypothesis?

Thanks for this comment, we added references that demonstrate the cry for help or emergence of opportunistic taxa in diseased plant:

Another hypothesis is that Enterobacter cancerogenus and Erwinia persicina are opportunistic taxa that are enriched in abnormal seedlings (e.g necrotrophy or “cry for help” of the plant) but are not responsible for the altered phenotype. Previous reports on diseased plants caused by a known pathogen support these hypotheses as the increase in abundance of beneficial protective strains or opportunistic copiotrophic taxa was frequently observed (Masson et al. 2020, Pfeilmeier et al. 2021, Arnault et al. 2022).