RESEARCH ARTICLE

Temporal dynamics of bacterial communities during seed development and maturation

Guillaume Chesneau¹, Gloria Torres-Cortes¹, Martial Briand¹, Armelle Darrasse¹, Anne Preveaux¹, Coralie Marais¹, Marie-Agnès Jacques¹, Ashley Shade² and Matthieu Barret^{1,*}

¹IRHS-UMR1345, Université d'Angers, INRAE, Institut Agro, SFR 4207 QuaSaV, 49071, Beaucouze, France and

²Department of Microbiology and Molecular Genetics, Program in Ecology, Evolutionary Biology, and Behavior, The DOE Great Lakes Bioenergy Research Center, and The Plant Resilience Institute, Michigan State University, East Lansing, MI 48824, USA

*Corresponding authors: IRHS, 42 rue Georges Morel, 49071 BEAUCOUZE, France. Tel: +33 2 41 22 57 16; E-mail: matthieu.barret@inrae.fr

One sentence summary: Bacterial dynamics during seed development.

Editor: Angela Sessitsch

[†]Matthieu Barret, <http://orcid.org/0000-0002-7633-8476>

ABSTRACT

Seed microbiota acts as a starting point for the assembly of the plant microbiota and contributes to successful plant establishment. To date, the order and timing of microbial taxa immigration during seed development and maturation remained unknown. We investigated the temporal dynamics of seed bacterial communities in bean and radish. A high phylogenetic turnover was observed for both plant species with few taxa associated with all seed developmental stages. Greater heterogeneity in communities structure within each stage was observed for radish. While, about one-third of radish seed bacterial taxa were detected in buds, flowers and fruits, very few taxa seem to be transmitted by the floral route in bean. In the latter species, bacterial populations belonging to the *P. fluorescens* species complex were found either in buds, flowers and fruits or in seeds. The relative phylogenetic proximity of these bacterial populations combined with their habitat specificity led us to explore the genetic determinants involved in successful seed transmission in bean. Comparative genomic analyses of representative bacterial strains revealed dozens of coding sequences specifically associated with seed-transmitted strains. This study provided a first glimpse on processes involved in seed microbiota assembly, which could be used for designing plant-beneficial microbial consortia.

Keywords: seed microbiota; community assembly; bacterial transmission; *Phaseolus vulgaris*; *Raphanus sativus*

INTRODUCTION

Plants host diverse communities of microorganisms, which can impact plant fitness. Part of the plant microbiota is transmitted from one plant generation to the next via seeds (Shade, Jacques and Barret 2017). Although the exact number of plant-associated taxa acquired through seeds is difficult to estimate, numerous cases of seed transmission have been reported. Historically, most of the documented cases of seed transmission

have been related to plant pathogens because this represents an important route for pathogen dispersion and is therefore important for plant diseases emergence (Baker and Smith 1966). For instance, *Acidovorax citrulli*, an important pathogen of watermelon can reach a 100% seed to seedling transmission rate (Dutta et al. 2012). Another well-described example of seed transmission concerned symbiotic fungi of the genus *Epichloë*, that produce bioactive alkaloids, which protect plant against

Received: 8 August 2020; Accepted: 17 September 2020

© The Author(s) 2020. Published by Oxford University Press on behalf of FEMS. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

herbivores but ultimately impact livestock health (Gagic et al. 2018). Finally, diversity surveys of seed microbial communities have detected many bacterial and fungal phyla across seeds of various plant families including *Amaranthaceae* (Lopez-Velasco et al. 2013), *Asteraceae* (Leff et al. 2017), *Cucurbitaceae* (Adam et al. 2018), *Brassicaceae* (Links et al. 2014; Rezki et al. 2016; Rybakova et al. 2017; Rochefort et al. 2019), *Fabaceae* (Klaedtke et al. 2016; Sánchez-López 2018; Gao et al. 2019), *Poaceae* (Links et al. 2014; Yang et al. 2017; Escobar Rodriguez et al. 2020) and *Solanaceae* (Bergna et al. 2018; Chen et al. 2020).

Changes in composition of the seed microbiota can impact plant fitness, notably by modification of seed vigor. For example, decrease of bacterial richness within *Setaria viridis* seeds leads to a reduction in their germination rate (Escobar Rodriguez et al. 2020). Moreover, correlation between seed microbiota composition and germination speed of *Brassica napus* was also recently highlighted (Rochefort et al. 2019). Variation in seed microbiota composition may also influence the structure of plant-associated microbial communities through priority effects, which correspond to the order and timing of species immigration during community assembly (Fukami 2015). For instance, composition of seed microbial community can modify colonization of wheat roots by dark septate endophytes, three weeks following germination (Ridout et al. 2019). Owing to such effects, managing seed microbiota composition is a promising avenue to improve plant growth and health. For example, floral inoculation of the plant-growth beneficial bacterial strain *Paraburkholderia phytofirmans* PsJN can favor its vertical transmission to the next plant generation and increase ear emergence of wheat (Mitter et al. 2017).

Incorporation of plant-beneficial microorganisms within seeds by inoculation of microbial endophytes, requires a good knowledge of the seed transmission pathways employed by these microorganisms. Based on previous works performed with plant pathogenic agents, three main transmission routes have been described: (i) a systemic pathway through hilum infection, (ii) a floral pathway by the stigma, or (iii) an external pathway by contact between the seed and fruits or threshing residues (Maude 1996). For example, *A. citrulli* uses the floral pathway to colonize watermelon seeds (Lessl, Fessehaie and Walcott 2007), as well as several *Xanthomonas* species to colonize *Brassica* or bean seeds (Darrasse et al. 2018; van der Wolf et al. 2019). Other bacteria, such as *Clavibacter michiganensis* can either use systemic or external pathways (Tancos et al. 2013). The genetic determinants required for successful seed transmission remain mostly unknown. The type 3 secretion system (T3SS) and the adhesins PilA and FhaB involved in transmission of *Xanthomonas citri* pv. *fuscans* to bean seeds (Darsonval et al. 2008, 2009), are the only determinants identified to date as required for seed transmission.

Deployment of a microbiota-based solution within seed to improve plant growth and health, requires a deeper understanding of the biological and ecological processes involved in the assembly of the seed microbiota. While some studies have focused on the dynamics of the seed microbiota during germination and emergence (Barret et al. 2015; Yang et al. 2017; Torres-Cortés et al. 2018), the dynamic of microbial communities during seed development is currently unknown. The main objective of this study was therefore to investigate the order and timing of microbial taxa immigration during two key phases of seed development: seed filling (i.e. accumulation of storage compounds) and seed maturation (i.e. loss of water content; Verdier, Leprince and Buitink 2019). A total of two species were chosen as study models, common bean (*Phaseolus vulgaris*) and radish (*Raphanus sativus*) since extensive characterization of their seed microbiota

was previously performed (Barret et al. 2015; Klaedtke et al. 2016; Rezki et al. 2016, Rezki et al. 2018; Torres-Cortés et al. 2018).

MATERIALS AND METHODS

Site description and sampling of radish and bean

Common bean (*Phaseolus vulgaris* L. var. Flaver) and radish (*Raphanus sativus* L. var. Flamboyant5) were grown in 2016 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). The two plant species were grown in separate blocks (5 m × 10 m), at a density of eight bean plants and five radish plants per square meters. Bean and radish seeds (kindly supplied by Vilmorin-Mikado) were sown at the end of March and May 2016, respectively. Flower buds, open flowers, fruits and seeds were collected weekly until full seed maturity (Fig. 1). More specifically, seeds were aseptically removed from fruits with sterile scalpel and tweezers under a laminar flow hood. Seeds were collected during seed filling (bean: B22, B28; radish: R25, R32 and R39) and seed maturation (bean: B35, B42; radish: R46, R53 and R67), each number corresponds to a number of days after pollination of bean and radish flowers. Since the phases of seed filling and seed maturation take longer for radish, more samples were collected. A total of three independent samples, from three individual plants, were collected at each sampling point. Each sample corresponded to a set of flower buds, open flowers, fruits and seeds (Fig. 1). For each sample, grams of tissues and colonies forming unit (CFU) per gram of sample were measured (Fig. 2A and B).

DNA extraction and construction of amplicon libraries

The collected biological material was soaked in 2 mL of phosphate-buffered saline (PBS, Sigma-Aldrich, St. Louis (USA)) supplemented with Tween® 20 (0.05% v/v, Sigma-Aldrich) per gram of fresh material. Flowers and fruits were crushed in a lab blender (Stomacher, Mixwel, Alliance Bio Expertise, Guipry (France)) for 1 min. Bean and radish seeds were soaked at 4°C under constant agitation (140 rpm) for 16 h and 2 h 30 min, respectively. The suspensions were centrifuged (6000 × g, 10 min, 4°C) and the resulting pellets were suspended in 200 µL of PBS. A total of 100 µL of suspensions were used for DNA extraction while the other 100 µL were dedicated to microbiological analyses (see 'collection of seed-associated bacteria' subsection).

DNA samples were extracted with the Powersoil DNA kit (Mo Bio Laboratories Inc, Carlsbad (USA)) following the supplier's recommendations. PCR reactions were performed with a high-fidelity Taq DNA polymerase (AccuPrime Taq DNA Polymerase System, Invitrogen, Carlsbad (USA)) using 5 µL of 10X Buffer, 1 µL of forward and reverse primers (*gyrB* [100 µM]; 16S rRNA gene [10 µM]), 0.2 µL of Taq and 5 µL of DNA). A first PCR amplification was performed with the primer sets 515f/806r (Caporaso et al. 2011) and *gyrB*.af64/*gyrB*.ar553 (Barret et al. 2015) which target the v4 region of 16S rRNA gene and a portion of *gyrB*, respectively. Cycling conditions for 515f/806r were 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 as followed: initial denaturation at 94°C for 3 min, 35 cycles of amplification at 94°C (30 s), 55°C (45 s) and 68°C (90 s) and final step at 68°C for 10 min. Amplicons were purified with magnetic beads (Sera-Mag, Merck, Darmstadt (Germany)). A second PCR amplification was performed to incor-

Figure 1. Experimental sampling. Flower buds (B01 and R01), open flowers (B02 and R04), fruits (B07 and R11) and seeds at two different stages of development, seed filling (B22, B28, R25, R32 and R39) and seed maturation (B35, B42, R46, R53 and R67), were sampled. A total of three replicates corresponding to a bulk of organs from different individual plants were performed at each sampling stage.

porate Illumina adapters and barcodes. PCR cycling conditions were identical for the two molecular markers: a first denaturation at 94°C (1 min), followed by 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 quantified with the Quant-iT PicoGreen® dsDNA Assay Kit (Invitrogen). All the amplicons were pooled in equimolar concentrations and the concentration of the equimolar pool was monitored with quantitative PCR (KAPA SYBR® FAST, Merck). Amplicon libraries were mixed with 10% PhiX and sequenced with a MiSeq reagent kit v2 500 cycles (Illumina, San Diego (USA)).

Sequence processing

Primer sequences were removed with cutadapt version 1.8 (Martin 2011). Fastq files were processed with DADA2 version 1.6.0 (Callahan et al. 2016), using the following parameters: truncLen = c(200, 150), maxN = 0, maxEE = c(1,1), truncQ = 5. Chimeric sequences were identified and removed with the removeBimeraDenovo function of DADA2. Taxonomic affiliations of amplicon sequence variants (ASVs) were performed with a naive Bayesian classifier (Wang et al. 2007) implemented in DADA2. ASVs derived from 16S rRNA gene were classified with the Silva 132 taxonomic training data (silva_nr_v132_train_set.fa.gz). *gyrB* ASVs were classified with an in-house *gyrB* database available upon request (train_set_gyrB_v4.fa.gz). The datasets supporting the conclusions of this article are available in the ENA database under the accession number [PRJEB38127].

Microbial community analyses

Microbial community analyses were conducted with Phyloseq version 1.22.3 (McMurdie and Holmes 2013). Sequences derived from 16S rRNA gene that were unclassified at the phylum-level, affiliated to Archaeae, 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 2.14.0 (Wright 2016) and neighbor joining phylogenetic trees were constructed with Phangorn version 2.5.5 (Schliep 2011).

Alpha-diversity metrics were calculated after rarefaction at 44 000 *gyrB* and 1000 16S rRNA gene sequences per sample. Observed richness (number of ASVs) and estimated richness (Chao1 index) were calculated with Phyloseq. Faith's (1992) phylogenetic diversity was calculated with picante version 1.7 (Kembel et al. 2010). Differences in alpha-diversity estimators between stages were assessed with Kruskal-Wallis non-parametric analysis of variance followed by Dunn's post-hoc test. Differences were considered as significant at a P-value ≤ 0.05 .

Changes in phylogenetic membership and composition were measured with unweighted and weighted UniFrac distances (Lozupone and Knight 2005), respectively. Changes in phylogenetic membership were measured with unweighted UniFrac distances, which is sensitive to low-abundant taxa. In contrast we measured changes in composition with weighted UniFrac distances which take into account the relative abundance of each taxa. The relative contribution of seed developmental stages on phylogenetic structure was estimated with canonical analysis of principal coordinates through capscale function of vegan 2.4.2 (Oksanen et al. 2017) followed with permutational multivariate analysis of variance (PERMANOVA; Anderson 2001). The dispersion of bacterial phylogenetic composition for each sampling stage was monitored with weighted UniFrac distance rather than unweighted UniFrac distance, in order to limit the impact of rare bacterial ASVs. Difference in dispersion between plants species was assessed with Wilcoxon non-parametric test.

Decomposition of phylogenetic turnover (species replacement) and phylogenetic diversity gradient (nestedness) were estimated with the phy.beta.pair function of Betapart version 1.5.1 (Baselga and Orme 2012). On the one hand, phylogenetic turnover (phy.beta.jtu) reflects the proportion of lineages that would be replaced between communities if both communities had the same number of taxa. Phylogenetic turnover was measured through unweighted UniFrac derived pair-wise phylogenetic dissimilarity. On the other hand, phylogenetic diversity gradient (phy.beta.jne) indicated phylogenetic dissimilarity, between communities due to differences in species richness. Phylogenetic diversity gradient was measured as the nestedness-fraction of unweighted UniFrac derived pairwise phylogenetic dissimilarity (Leprieur et al. 2012).

Hierarchical clustering of bacterial ASVs was performed on a subset of ASVs that were detected in at least two samples

Figure 2. Temporal diversity of bacterial communities associated with reproductive organs of common bean and radish. (A) Grams of tissues and (B) log CFU monitored for each sample. (C) Observed richness and (D) phylogenetic diversity (Faith's phylogenetic diversity index, Faith 1992) estimated with *gyrB* amplicon sequence variants (ASVs). (E) Estimated richness (Chao1) and (F) phylogenetic diversity estimated with 16S rRNA ASVs. Small light dots represented replicates while large opaque dots represented medians. Letters a and b denoted significant changes between conditions considered at a P -value ≤ 0.05 (Kruskal-Wallis with post hoc Dunn test). No significant differences were observed for (A), (B), (E) and (F).

with a relative abundance $>0.1\%$. Hierarchical clustering was performed via cosine similarity of log-transformed ASV counts ($\log_{10}(x + 1)$). Cosine similarity measures the cosine of the angle between two vectors (ASVs) projected in a multi-dimensional space (samples). This measure ignores the magnitude (differences in sequences counts between ASVs) and targets the orientation of the vector. Heatmaps were constructed with pheatmap version 1.0.12 (Kolde 2019) to visualize ASVs co-occurrence. Visualization of phylogenetic tree was done with Interactive Tree of Life version 5.5 (Letunic and Bork 2016). Scripts and data sets employed in this work are available in GitHub: https://github.com/martialbriand/IRHS_EmerSys/blob/master/Chesneau.etal.2020/Chesneau.etal.2020.Final.Version.R.

Collection of seed-associated bacteria and reconstruction of genomic sequences

Suspension were serial-diluted and plated on 1/10 strength Tryptic Soy Agar (17 g/L tryptone, 3 g/L soybean peptone, 2.5 g/L

glucose, 5 g/L NaCl, 5 g/L K₂HPO₄ and 15 g/L agar, Oxoid, Waltham (USA)) supplemented with cycloheximide (50 μ g/mL, Sigma-Aldrich). After 5 days of incubation at 18 °C, colony forming units (CFUs) were counted for each sample. A total of 24 CFUs were randomly picked at each sampling stage, thereby resulting in 180 and 196 bacterial strains for bean and radish, respectively. Molecular typing of each bacterial strain was performed through *gyrB* amplification and subsequent Sanger sequencing (Genoscreen). Strains were associated with ASVs if their *gyrB* sequences were strictly identical (100% identity over 100% of their length) to the ASV sequence.

DNA of 39 bacterial strains representing the most prevalent seed-associated bacterial populations on community sequencing data (Table S1, Supporting Information), were extracted with the Wizard® Genomic DNA Purification Kit (Promega, Madison (USA)). DNA was sequenced with DNBSeq using a PCR free library protocol (BGI). Paired-end reads were assembled with SOAPdenovo version 2.04 (Li et al. 2010) and VELVET version 1.2.10 (Zerbino and Birney 2008). The bacterial strains were

deposited in the CIRM-Plant Associated Bacteria collection (<https://www6.inrae.fr/cirm.eng/CFBP-Plant-Associated-Bacteria>).

Taxonomic classification of genome sequences was performed through calculation of overall genome relatedness indices. The phylogenetic neighbors were selected from genomes sequences of the NCBI WGS database ($n = 154\,480$ as of July 2018) with the one2all mode of Kmer-db version 1.6.2 (Deorowicz et al. 2019). Genomes sequences were affiliated at the species level when they shared at least 50% of 15-mers with the genomes of type strains. The percentage of shared 15-mers was also used to investigate the relatedness between the 39 genomes sequences.

Structural and functional annotations were carried out with prokka version 1.2 (Seemann 2014). Orthology assignment was performed with DIAMOND version 0.9.10 (Buchfink, Xie and Huson 2015) on the eggNOG4.5 database (Huerta-Cepas et al. 2016). Orthologous groups (OGs) significantly (Bonferroni adjusted P -value ≤ 0.05) associated with seeds were identified with Scoary version 1.6.16 (Brynildsrud et al. 2016) using a presence/absence table of OGs.

RESULTS

Estimation of bacterial community profiles with 16S rRNA gene and *gyrB*

Although the v4 region of 16S rRNA gene (hereafter 16S) is routinely employed in PCR-based surveys of microbial communities, this region has a poor discriminatory power at the species level. To circumvent this limitation, an alternative bacterial marker based on a portion of *gyrB* that encodes the β subunit of the bacterial gyrase was recently developed and employed for estimating the structure of plant-associated bacterial communities (e.g. Barret et al. 2015; Bartoli et al. 2018; Rochefort et al. 2019). To assess the potential amplification bias of *gyrB* primers, we compared the bacterial community profiles estimated with 16S or *gyrB*. A total of four main phyla, *Proteobacteria*, *Actinobacteria*, *Firmicutes* and *Bacteroidetes*, were consistently detected with both molecular markers (Figures S1A and S1B, Supporting Information). Phyla specifically detected with 16S (e.g. *Chloroflexi*; Figure S1A, Supporting Information) or *gyrB* (e.g. *Nitrospirae*; Figure S1B, Supporting Information) represented ASVs of low prevalence with low sequences counts. According to Mantel test, unweighted UniFrac distances derived from 16S and *gyrB* were moderately correlated ($R = 0.37$; Figure S1C, Supporting Information), while weighted UniFrac distances were highly correlated ($R = 0.83$; Figure S1D, Supporting Information). This reflected differences in detection of ASVs of low abundance. The number of reads per sample was 10- to 100-fold lower with 16S (Figure S1E, Supporting Information) in comparison to *gyrB* (Figure S1F, Supporting Information) as a result of co-amplification of chloroplast and mitochondrial DNA with the primer set that target the v4 region of 16S gene. In the rest of the manuscript, estimates of bacterial diversity are presented with *gyrB* and compared where possible with the 16S dataset.

Flower to seed dynamics of bacterial communities

A total of 361 and 740 bacterial *gyrB* ASVs were detected on bean and radish samples, while 64 and 84 ASVs were obtained with 16S. However, the variation between replicates within the same sampling stage was important since 15 (4% of *gyrB* ASVs detected in bean) and 42 (6% of *gyrB* ASVs detected in radish) ASVs were

Figure 3. Temporal structure of bacterial communities associated with reproductive organs of common bean and radish. Dispersion in bacterial phylogenetic composition (weighted UniFrac distances) at each stage of seed development estimated with *gyrB* ASVs (A) and 16S ASVs (B), respectively. (C) Decomposition of phylogenetic beta-diversity into (i) nestedness (phylo.beta.jne; dark gray) and (ii) phylogenetic turnover (phylo.beta.jtu; gray), between successive stages sampled. Small light dots represented replicates while large opaque dots represented medians.

associated with all replicates of a sampling stage (Table S1, Supporting Information). With regard to 16S ASVs, a higher percentage of ASVs was systematically detected at each sampling stage (17 ASVs-26% in bean, 24 ASVs-29% in radish, Table S2, Supporting Information). In bean samples, richness and phylogenetic diversity significantly (P -value ≤ 0.05) increased at the end of seed maturation with both molecular markers (Fig. 2C–F). These increases in richness and phylogenetic diversity were neither a consequence of the amount of tissue sampled (Fig. 2A) nor an increase in the size of the cultivable bacterial population (Fig. 2B). In radish samples, richness and phylogenetic diversity significantly decreased during the transition from fruit to developing seeds and then remained stable over the course of seed filling and maturation (Fig. 2C–F). This decrease was independent of the amount of tissue sampled or the size of the cultivable population (Fig. 2A and B).

Changes of bacterial communities structure were subsequently monitored by analyzing phylogenetic dissimilarity among all sampled habitats. According to distance from centroid, variation in bacterial phylogenetic composition was significantly lower in bean in comparison to radish for all sampled habitats (Fig. 3A and B). In fact, the bean microbiota was almost exclusively composed of bacterial ASVs affiliated to *Pseudomonadales* whilst the radish microbiota was made up of ASVs related

to *Enterobacterales* and *Pseudomonadales* (Figure S2, Supporting Information). Phylogenetic membership (unweighted UniFrac distance) was significantly (P -value ≤ 0.001) impacted in bean and radish by habitat type sampled with both molecular markers (Figure S3, Supporting Information). Overall, habitat type explained 22.0% and 33.6% of variance in bean samples with *gyrB* and 16S sequences and 22.2% and 31.9% in radish samples (Figure S3, Supporting Information). Phylogenetic composition (weighted UniFrac distance) was also significantly impacted by habitat type in bean (95.0% and 45.1% of variance with *gyrB* and 16S sequences) but not in radish (Figure S3, Supporting Information).

To assess whether changes in phylogenetic membership across habitats were driven by phylogenetic turnover or nestedness, we employed a framework initially proposed for assessing spatial turnover of lineages (Baselga and Orme 2012; Leprieux et al. 2012). Changes in beta-diversity between two communities can be attributed to nestedness, which is the acquisition of new members while retaining some original members, or to turnover, which is the replacement of original members with new members. Phylogenetic turnover was the major component of beta-diversity in bean samples (>90%) up to the seed maturation stage where the total amount of phylogenetic beta-diversity became half between turnover and nestedness (Fig. 3C). With respect to radish samples, phylogenetic turnover represented more than 75% of phylogenetic beta-diversity at all stages sampled with the exception of the transition from fruit to developing seeds (R11–R25, Fig. 3C). During this transition nestedness represented 74% of the total amount of phylogenetic beta-diversity.

Influence of the floral pathway in bacterial seed transmission

To estimate the relative importance of the floral pathway in seed bacterial communities memberships, we measured the number of ASVs detected in buds, flowers and fruits that were associated with seeds. As the *gyrB* read set is composed of a larger number of sequences than the 16S dataset, we used the former for subsequent analyses. Only 10% of the ASVs associated with bean seeds were detected in buds, flowers and fruits. This percentage was more than three times higher in radish, with 36% seeds ASVs detected in buds, flowers and fruits.

To assess whether some groups of ASVs displayed similar changes in their abundances over time a hierarchical clustering based on cosine similarity was performed. Radish-associated taxa were not clustered according to the nature of the habitat sampled (Figure S4, Supporting Information). For instance, we did not detect any taxa specifically associated with radish seeds. However, a set of eight persistent taxa was found at every stage sampled (cluster 3; Figure S4, Supporting Information). These persistent taxa were affiliated to *Pseudomonas*, *Pantoea*, *Erwinia* and *Sphingomonas*. Concerning bean samples, one cluster (cluster 2) was mainly composed of ASVs associated with seeds while another cluster (cluster 4) contained ASVs primarily detected on bud, flower and fruit (Fig. 4).

Investigation of genetic determinants involved in transmission to bean seeds

Isolation of bacterial strains from buds, flowers, fruits and seeds on a single synthetic media resulted in 180 and 196 isolates in bean and radish, respectively. Of these isolates, 88% and 83% possess 100% identity to ASVs obtained through community

profiling approach. Bean isolates were distributed in 13 ASVs (96.9% of all reads) with a mean of 12 isolates (min 1, max 52) per ASV while 28 ASVs (93.5% of all reads) were associated with radish isolates with a mean of 4 isolates (min 1, max 26) per ASV. A total of 39 isolates representative of the most abundant ASVs (Figure S5, Supporting Information) were selected for whole genome sequencing. These 39 genomic sequences (Table S3, Supporting Information) were clustered according to the percentage of shared k-mer, resulting in 12 bacterial groups at a 50% threshold (Fig. 5). A total of two groups, related to ASV1 (*Pseudomonas orientalis*) and ASV28 (*Pseudomonas coleopterorum*), both belonging to the *P. fluorescens* subgroup (Hesse et al. 2018), were composed of bean seed-associated taxa (bean cluster 2). Of note, ASV1 group (*P. orientalis*) was composed of genomes sequences derived from bean and radish seed isolates, which suggested no host specificity (Fig. 5).

To assess whether some genetic determinants could be associated with transmission of bacterial taxa to bean seed, a comparative genomic analysis was performed between isolates specifically associated to seeds ASVs (bean cluster 2) and isolates associated with buds, flowers and fruits ASVs (bean cluster 4). Since *Pseudomonadales* represented the main bacterial order detected on bean seeds (Figure S2, Supporting Information), we focused our analysis on *Pseudomonas* isolates associated with bean cluster 2 ($n = 8$) and bean cluster 4 ($n = 12$). A total of four and 12 eggNOG orthologous groups (OGs) were significantly (P -value ≤ 0.05) associated with bean cluster 2 and bean cluster 4, respectively (Table 1). The four OGs shared between *P. orientalis* (ASV1) and *P. coleopterorum* (ASV28) were predicted to encode a LysR family transcriptional regulator, a DUF262 domain-containing protein, an aldose 1-epimerase family protein and a sugar major facilitator superfamily (MFS) transporter (Table 1).

Although *P. orientalis* (ASV1) and *P. coleopterorum* (ASV28) were both detected during bean seed filling and maturation, the former (ASV1) was the dominant taxa of bacterial communities (Fig. 4). The high relative abundance of *P. orientalis* could be potentially related to specific genetic determinants that provided a better fitness. In addition to the four OGs shared with *P. coleopterorum*, 68 OGs were specifically detected in the representative genome sequences ($n = 6$) of *P. orientalis*. According to eggNOG classification approximately 75% of these OGs were associated with proteins of unknown function (Table S4, Supporting Information). Closer inspection of protein sequence similarity through BLASTp searches revealed however that some of these OGs of unknown function were associated with secondary metabolites biosynthetic clusters (B207_01600, B207_01601, B207_01607, B207_01611 and B207_02823) and putative type III effector (B207_02170).

DISCUSSION

Origin of seed-associated taxa

According to our sampling, the arrival of bacteria at the early stages of seed development was different between the two plant species investigated. In radish, 36% of the bacterial taxa detected in seeds were also associated with buds, flowers and fruits, therefore suggesting that these taxa were seed-transmitted by the floral pathway. In contrast, taxa associated with bean seed were rarely (~10%) detected on flower and fruit, which could then imply that the systemic pathway is the preferred route of seed transmission. The difference in seed transmission pathways between bean and radish could be

Figure 4. Bacterial taxa dynamics during seed development of common bean. Abundant ASVs ($>1\%$ in relative abundance) detected in at least two samples were clustered through cosine similarity of log-transformed ASV counts (y-axis). A total of four distinct clusters were highlighted, (i) cluster 1 and 3 for transient ASVs, (ii) cluster 2 for seed-specific ASVs and (iii) cluster 4 for flower/fruit specific ASVs. The different habitats, harvesting days and replicates were represented on the x-axis by different colors: red (buds), green (open flowers), blue (fruits) and purple (seeds). The heatmap color gradient represented the log10 of read count.

related to their distinct fertilization modes. In common bean, pollination is mainly autogamous with a very low occurrence of cross pollination. Fertilization occurs before the period during which flower open (Frankel and Galun 1977). This pre-anthesis self-pollination could limit the number of micro-organisms that are transmitted through the floral pathway. Although seed transmission of *Xanthomonas citri* pv. *fuscans* has been already reported via this pathway in common bean under controlled conditions (Darsonval et al. 2008; Darrasse et al. 2018), this mode of transmission is probably not widely employed by bacteria in natural setting. In contrast to bean, radish is a self-incompatible annual plant species that relies on anthesis synchrony for successful cross-fertilization. Effective radish reproduction also depends on insect pollinators (Kercher and Conner 1996). During nectar collection, pollinating insects such as small bees and honey-bees (Conner, Sahli and Karoly 2009) deliver microorganisms onto flower (Ushio et al. 2015; Manirajan et al. 2016) and some of these microorganisms are latter incorporated in the seed microbiota (Compant et al. 2011; Prado et al. 2020). Hence, some radish seed-associated taxa detected at the early stage of seed development could be introduced by pollinators.

Dynamics of bacterial communities

Buds, flowers and fruits habitats are associated with high bacterial diversity, which is consistent with previous observations of promotion of bacterial diversity in floral microhabitats such as nectar, stamina and styles (Junker et al. 2011). In contrast, a limited phylogenetic diversity was observed within seed bacterial

communities of both plant species during seed filling and seed maturation. Such observations of low bacteria diversity were observed in previous studies on seeds of various plant species such as *Brassicaceae* (Barret et al. 2015; Rezki et al. 2016), *Poaceae* (Eyre et al. 2019) or *Cucurbitaceae* (Adam et al. 2018). These preliminary findings were in agreement with the hypothesis of a high population bottleneck in seeds (Newcombe et al. 2018) as a result of accumulation of defense compounds within seeds (Mel-dau, Erb and Baldwin 2012) and to the low number of microhabitats associated with the seed (Junker et al. 2011). For instance, bean seed bacterial communities were composed of only four abundant taxa consistently detected during seed development and taxa of low abundance transiently associated with few seed samples. Higher variability of bacterial communities structure was observed on all reproductive organs of radish. This higher variability could be related to a number of plant traits including pollination mode but also plant phenology or plant chemistry.

A significant amount of beta-diversity was associated with phylogenetic turnover during seed filling and seed maturation. Although this taxa turnover is mainly driven by rare taxa that replaced members of the same bacterial order, it does, however, raises the question of their origin. If we cannot rule out the possibility of these taxa being incorporated through the systematic or external pathway during seed development, the most plausible explanation is a limit of the experimental design employed in this work. Indeed, the destructive sampling prevented temporal tracking of bacterial communities within one plant. Since the composition of seed bacterial communities is highly variable between radish plants (Rezki et al. 2018), the important species

Figure 5. Overall genome relatedness of bacterial strains. Genome sequences of 39 bacterial strains representing the most abundant ASVs were obtained. The frequency of shared 15-mers between genomes sequences was used as a proxy of genome relatedness. The branch length corresponded to the frequency of shared 15-mers. Branches were colored according to the taxonomic affiliation of each genome sequence. Green and red circles represented bacterial strains isolated from bean and radish, respectively. Bacterial strains associated with bean cluster 2 (seed-associated) and cluster 4 (flower- and fruit-associated) were represented by purple and orange squares, respectively. Dashed line indicates the threshold employed for species delineation (50% of 15-mers, Briand et al. 2019).

Table 1. Orthologous groups specifically associated with genome sequences from bean cluster 2 and bean cluster 4.

Group	Locus tag	Functional class	Annotation	Bean cluster 2	Bean cluster 4	P-value
0XPF7@NOG	B144.00351	P	TonB-dependent receptor family protein	0	12	0.02
1141G@NOG	B144.00599	S	Hypothetical protein	0	12	0.02
0YHTZ@NOG	B144.00885	E	ABC transporter substrate-binding protein	0	12	0.02
COG2423@NOG	B144.00945	E	Ornithine cyclodeaminase family protein	0	12	0.02
COG3803@NOG	B144.01198	S	DUF924 domain-containing protein	0	12	0.02
0XRNF@NOG	B144.01771	K	AraC family transcriptional regulator	0	12	0.02
11HN1@NOG	B144.02140	S	Hypothetical protein	0	12	0.02
COG2378@NOG	B144.04315	K	YafY family transcriptional regulator	0	12	0.02
0XR7A@NOG	B144.04334	M	Fatty acid cis/trans isomerase	0	12	0.02
0XRA7@NOG	B144.04582	K	LysR family transcriptional regulator	0	12	0.02
COG5281@NOG	B144.04708	S	Phage-related minor tail protein	0	12	0.02
COG4733@NOG	B144.04719	S	Phage-related protein, tail component	0	12	0.02
105ZK@NOG	B207.00203	K	LysR family transcriptional regulator	8	0	0.02
COG1479@NOG	B207.02513	S	DUF262 domain-containing protein	8	0	0.02
0XPUG@NOG	B207.02938	S	Aldose 1-epimerase family protein	8	0	0.02
COG0738@NOG	B207.02939	G	Sugar major facilitator superfamily (MFS) transporter	8	0	0.02

Orthologous groups (OGs) significantly (Bonferroni adjusted P-value ≤ 0.05) associated with bean cluster 2 ($n = 8$ genome sequences) and bean cluster 4 ($n = 12$) were obtained with Scoary. OGs were classified into broad functional categories: Amino acid transport and metabolism (E), Carbohydrate transport and metabolism (G), Transcription (K), Cell wall/membrane/envelope biogenesis (M), Inorganic ion transport and metabolism (P) and Unknown function (S).

turnover observed could be solely due to differences in bacterial community structure between individuals.

In the last stages of seed maturation, an increase of phylogenetic diversity was observed. This increase is independent of the 'true taxa turnover' (Leprieur et al. 2012) reported in the above paragraph since it is related to gain of taxa within communities, which contributed to nestedness. This gain of taxa at late seed maturation could be provided by the external pathway as a result of anatomical changes in fruit structure. Indeed, in the course of fruit maturation, fruit undergo programmed cell death by the action of enzymes inducing walls breakdown and dehiscence of fruits (Meakin and Roberts 1990; Ogutcen et al. 2018). Alternatively, the physiological changes occurring during late seed maturation (Leprince et al. 2017) could also increase the relative abundance of seed-associated taxa that were very rare or even below the detection limit. For instance, the low seed moisture, which corresponded to a water potential of around -150 Mp (Leprince et al. 2017) could favor the growth of bacteria that can tolerate high osmotic stress.

Genetic determinants of seed transmission in common bean

To date, seed transmission routes have been described for some plant pathogenic bacteria (Darrasse et al. 2018; van der Wolf et al. 2019) but remain unclear for most seed associated bacteria. Seed-transmitted taxa have not only to overcome rapid host physiological changes occurring during seed filling and maturation but also to compete for resources with other microbial populations (Hibbing et al. 2010). Therefore, successful seed transmission required a number of genetic adaptations to overcome these challenges. Our study revealed significant alteration in phylogenetic membership and composition between bean reproductive habitats (buds, flowers, fruits and seeds). Thus, the presence of two clusters of ASVs (i) either associated with bean flowers and fruits but rarely detected on bean seeds or (ii) mainly seed-associated provided an interesting opportunity to identify these determinants through comparative genome analyses. Only four OGs were specifically associated with *P. orientalis* and *P. coelestis*, two species detected in bean during seed filling and maturation. This low number of OGs detected in both seed-associated species could indicate that the genetic determinants of seed transmission are not conserved across these species. Two OGs specifically associated with seed isolates, B207.02938 and B207.02939 (Table 1), were related to hexose metabolism and transport. If the product of these genes are involved in successful seed colonization, then they may be required during seed embryogenesis when hexose concentration is high (Weber, Borisjuk and Wobus 2005). These initial observations still need to be confirmed with a larger sampling and sequencing of seed-associated bacterial populations.

Given the dominance of *P. orientalis* within seed bacterial communities of bean, the OGs specifically associated with the representative genome sequences of this species were analyzed. Although enrichment of most of these OGs were plausibly explained by shared ancestry, some could be involved in the successful colonization of bean seed. For instance, one protein of unknown function (B207.02170) shared 28% identity at the C-terminus with the type III effector HopB1 from *Pseudomonas syringae* pv. *tomato* DC3000. HopB1 perturbs *Arabidopsis thaliana* immune signaling through the cleavage of BAK1, a protein that together with FLS2 are involved in perception of bacterial flagellins (Li et al. 2016). While this type III effector protein is not

present in the genome of *X. citri* pv. *fuscans* CFBP4834-R, a functional T3SS is mandatory for seed transmission of this plant pathogen to bean (Darsonval et al. 2008), which thus indicate that bacteria must circumvent the plant immune system for being seed transmitted. In addition to determinant involved in plant-bacterial interactions, OGs of *P. orientalis* associated with secondary metabolites production could provide a competitive advantage to this bacterial species over other microbial populations. For example, one OG (B207.02823) was located within a phenazine biosynthetic cluster closely related to *P. orientalis* F9 (Zengerer et al. 2018). Phenazine could increase the competitiveness of *P. orientalis* in nutrient poor environment (such as seed) thanks to its antimicrobial functions (Price-Whelan, Dietrich and Newman 2006; Hibbing et al. 2010).

In conclusion, this study revealed that seed bacterial communities were composed of few dominant taxa that are presented at the early seed filling stages and persisted during maturation. Difference in structure of seed communities between *Phaseolus vulgaris* and *Raphanus sativus* could be related to the seed transmission pathway employed as a result of the plant pollination mode. The mode of pollination is not the only difference between radish and bean seeds, a subsequent analysis must be carried out to answer this question. The dominance of *P. orientalis* within bean seed microbial communities could be due to specific bacterial determinants such as T3 effector and secondary metabolites. Future works should assess the *in planta* expression of these genetic determinants during seed transmission.

ACKNOWLEDGMENTS

This work was supported by the French National Research Agency [ANR-17-CE20-0009-01] and AgreenSkills+ fellowship programme, which has received funding from the EU's Seventh Framework Programme [FP7-609398]. The authors wish to thank Emmanuel Laurent and Vincent Odeau (FNAMS) for field crop management, Jérôme Gouzy (CATI BBRIC) for the genome assembly pipeline and Muriel Bahut (ANAN platform, SFR Quasav) for amplicon sequencing. AS acknowledges support from Michigan State University AgBioResearch (Hatch) and USDA [NIFA-2019-67019-29305].

SUPPLEMENTARY DATA

Supplementary data are available at [FEMSEC](#) online.

Conflicts of interest. None declared.

REFERENCES

- Adam E, Bernhart M, Müller H et al. The *Cucurbita pepo* seed microbiome: genotype-specific composition and implications for breeding. *Plant Soil* 2018;**422**:35–49.
- Anderson MJ. Permutation tests for univariate or multivariate analysis of variance and regression. *Can J Fish Aquat Sci* 2001;**58**:626–39.
- Baker KF, Smith SH. Dynamics of seed transmission of plant pathogens. *Annu Rev Phytopathol* 1966;**4**:311–32.
- Bartoli C, Frachon L, Barret M et al. In situ relationships between microbiota and potential pathobiota in *Arabidopsis thaliana*. *ISME J* 2018;**8**:2024–38.
- Baselga A, Orme CDL. betapart: an R package for the study of beta diversity. *Methods Ecol Evol* 2012;**3**:808–12.

- Bergna A, Cernava T, Rändler M et al. Tomato seeds preferably transmit plant beneficial endophytes. *Phytobiomes J* 2018;**2**:183–93.
- Briand M, Bouzid M, Hunault G et al. A rapid and simple method for assessing and representing genome sequence relatedness. *bioRxiv* 2019:569640.
- Brynildsrud O, Bohlin J, Scheffer L et al. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol* 2016;**17**:238.
- Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;**12**:59–60.
- Callahan BJ, McMurdie PJ, Rosen MJ et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;**13**:581–3.
- Caporaso JG, Lauber CL, Walters WA et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* 2011;**108**:4516–22.
- Chen X, Krug L, Yang H et al. *Nicotiana tabacum* seed endophytic communities share a common core structure and genotype-specific signatures in diverging cultivars. *Comput Struct Biotechnol J* 2020;**18**:287–95.
- Compant S, Mitter B, Colli-Mull JG et al. Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches of colonization. *Microb Ecol* 2011;**62**:188–97.
- Conner JK, Sahli HF, Karoly K. Tests of adaptation: functional studies of pollen removal and estimates of natural selection on anther position in wild radish. *Ann Bot* 2009;**103**:1547–56.
- Darrasse A, Barret M, Cesbron S et al. Niches and routes of transmission of *Xanthomonas citri* pv. *fuscans* to bean seeds. *Plant Soil* 2018;**422**:115–28.
- Darsonval A, Darrasse A, Durand K et al. Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Mol Plant Microb Interact* 2009;**22**:747–57.
- Darsonval A, Darrasse A, Meyer D et al. The type III secretion system of *Xanthomonas fuscans* subsp. *fuscans* Is Involved in the phyllosphere colonization process and in transmission to seeds of susceptible beans. *Appl Environ Microbiol* 2008;**74**:2669–78.
- Deorowicz S, Gudyś A, Długosz M et al. Kmer-db: instant evolutionary distance estimation. *Bioinformatics* 2019;**35**:133–6.
- Dutta B, Avci U, Hahn MG et al. Location of *Acidovorax citrulli* in infested watermelon seeds is influenced by the pathway of bacterial invasion. *Phytopathology* 2012;**102**:461–8.
- Escobar Rodriguez C, , Antonielli L, Mitter B, Sessitsch A et al. Heritability and functional importance of the *Setaria viridis* L. bacterial seed microbiome. *Phytobiomes J* 2020;**4**:40–52.
- Eyre AW, Wang M, Oh Y et al. Identification and characterization of the core rice seed microbiome. *Phytobiomes J* 2019;**3**:148–57.
- Faith DP. Conservation evaluation and phylogenetic diversity. *Biol Conserv* 1992;**61**:1–10.
- Frankel R, , Galun E. Pollination mechanisms, reproduction and plant breeding. *Monographs on Theoretical and Applied Genetics* 1977;**2**.
- Fukami T. Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annu Rev Ecol Syst* 2015;**46**:1–23.
- Gagic M, Faville MJ, Zhang W et al. Seed transmission of epichloë endophytes in *Lolium perenne* is heavily influenced by host genetics. *Front Plant Sci* 2018; **9**:1580.
- Gao W, Zheng C, Lei Y et al. Analysis of bacterial communities in white clover seeds via high-throughput sequencing of 16S rRNA gene. *Curr Microbiol* 2019;**76**:187–93.
- Hesse C, Schulz F, Bull CT et al. Genome-based evolutionary history of *Pseudomonas* spp. *Environ Microbiol* 2018;**20**:2142–59.
- Hibbing ME, Fuqua C, Parsek MR et al. Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* 2010;**8**:15–25.
- Huerta-Cepas J, Szklarczyk D, Forslund K et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 2016;**44**:D286–93.
- Junker RR, Loewel C, Gross R et al. Composition of epiphytic bacterial communities differs on petals and leaves. *Plant Biol* 2011;**6**:918–24.
- Kembel SW, Cowan PD, Helmus MR et al. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 2010;**26**:1463–4.
- Kercher S, Conner J. Patterns of genetic variability within and among populations of wild radish, *Raphanus raphanistrum* (Brassicaceae). *Am J Bot* 1996;**83**:1416–21.
- Klaedtke S, Jacques M-A, Raggi L et al. Terroir is a key driver of seed-associated microbial assemblages: terroir shapes the seed microbiota. *Environ Microbiol* 2016;**18**:1792–804.
- Kolde R. Pheatmap: Pretty Heatmaps, 2019.
- Leff JW, Lynch RC, Kane NC et al. Plant domestication and the assembly of bacterial and fungal communities associated with strains of the common sunflower, *Helianthus annuus* *New Phytologist* 2017;**214**:412–23.
- Leprieux F, Albouy C, De Bortoli J et al. Quantifying phylogenetic beta diversity: distinguishing between ‘True’ turnover of lineages and phylogenetic diversity ggradients. Shawkey M (ed.). *PLoS One* 2012;**7**:e42760.
- Leprince O, Pellizzaro A, Berriri S et al. Late seed maturation: drying without dying. *J Exp Bot* 2017;**68**:827–41.
- Lessl JT, Fessehaie A, Walcott RR. Colonization of female watermelon blossoms by *Acidovorax avenae* ssp. *citrulli* and the relationship between blossom inoculum dosage and seed infestation. *J Phytopathol* 2007;**155**:114–21.
- 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* 2016;**44**:W242–5.
- Li L, Kim P, Yu L et al. Activation-dependent destruction of a coreceptor by a *Pseudomonas syringae* effector dampens plant immunity. *Cell Host Microb* 2016;**20**:504–14.
- Links MG, Demeke T, Gräfenhan T et al. Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on Triticum and Brassica seeds. *New Phytol* 2014;**202**:542–53.
- Li R, Zhu H, Ruan J et al. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* 2010;**20**:265–72.
- Lopez-Velasco G, Carder PA, Welbaum GE et al. Diversity of the spinach (*Spinacia oleracea*) spermosphere and phyllosphere bacterial communities. *FEMS microbiology letters* 2013;**346**:146–54.
- Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 2005;**71**:8228–35.
- Manirajan BA, Ratering S, Rusch V et al. Bacterial microbiota associated with flower pollen is influenced by pollination type, and shows a high degree of diversity and species-specificity. *Environ Microbiol* 2016;**18**:5161–74.
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 2011;**17**:10–2.

- Maude RB. Seedborne diseases and their control: principles and practice. CAB International. Wallingford, UK, 1996.
- McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;8:e61217.
- Meakin PJ, Roberts JA. Dehiscence of fruit in oilseed rape (*Brassica napus* L.)I. ANATOMY OF POD DEHISCENCE. *J Exp Bot* 1990;41:995–1002.
- Meldau S, Erb M, Baldwin IT. Defence on demand: mechanisms behind optimal defence patterns. *Ann Bot (Lond)* 2012;110:1503–14.
- Mitter B, Pfaffenbichler N, Flavell R et al. A new approach to modify plant microbiomes and traits by introducing beneficial bacteria at flowering into progeny seeds. *Front Microbiol* 2017;8:11.
- Newcombe G, Harding A, Ridout M et al. A hypothetical bottleneck in the plant microbiome. *Front Microbiol* 2018;9:1645.
- Ogutcen E, Pandey A, Khan MK et al. Pod shattering: a homologous series of variation underlying domestication and an avenue for crop improvement. *Agronomy* 2018;8:137.
- Oksanen J, Blanchet F, Friendly M et al. vegan: Community Ecology Package version 2.5-6 from CRAN ; R. 2019.
- Prado A, Marolleau B, Vaissière BE et al. Insect pollination: an ecological process involved in the assembly of the seed microbiota. *Sci Rep* 2020;10:1–11.
- Price-Whelan A, Dietrich LEP, Newman DK. Rethinking “secondary” metabolism: physiological roles for phenazine antibiotics. *Nat Chem Biol* 2006;2:71–8.
- Rezki S, Champion C, Iacomì-Vasilescu B et al. Differences in stability of seed-associated microbial assemblages in response to invasion by phytopathogenic microorganisms. *PeerJ* 2016;4:e1923.
- Rezki S, Champion C, Simoneau P et al. Assembly of seed-associated microbial communities within and across successive plant generations. *Plant Soil* 2018;422:67–79.
- Ridout ME, Schroeder KL, Hunter SS et al. Priority effects of wheat seed endophytes on a rhizosphere symbiosis. *Symbiosis* 2019;78:19–31.
- Rochefort A, Briand M, Marais C et al. Influence of environment and host plant genotype on the structure and diversity of the *Brassica napus* seed microbiota. *Phytobiomes J* 2019;3:326–36.
- Rybakova D, Mancinelli R, Wikström M et al. The structure of the *Brassica napus* seed microbiome is cultivar-dependent and affects the interactions of symbionts and pathogens. *Microbiome* 2017;5:104.
- Schliep KP. phangorn: phylogenetic analysis in R. *Bioinformatics* 2011;27:592–3.
- Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–9.
- Shade A, Jacques M-A, Barret M. Ecological patterns of seed microbiome diversity, transmission, and assembly. *Curr Opin Microbiol* 2017;37:15–22.
- Sánchez-López AS. Community structure and diversity of endophytic bacteria in seeds of three consecutive generations of *Crotalaria pumila* growing on metal mine residues. *Plant Soil* 2018;422:51–66.
- Tancos MA, Chalupowicz L, Barash I et al. Tomato fruit and seed colonization by *Clavibacter michiganensis* subsp. *michiganensis* through external and internal routes. *Appl Environ Microbiol* 2013;79:6948–57.
- Torres-Cortés G, Bonneau S, Bouchez O et al. Functional microbial features driving community assembly during seed germination and emergence. *Front Plant Sc* 2018;9:902.
- Ushio M, Yamasaki E, Takasu H et al. Microbial communities on flower surfaces act as signatures of pollinator visitation. *Sci Rep* 2015;5:1–7.
- van der Wolf J, Kastelein P, da Silva Júnior TAF et al. Colonization of siliques and seeds of rapid cycling *Brassica oleracea* plants by *Xanthomonas campestris* pv. *campestris* after spray-inoculation of flower clusters. *Eur J Plant Pathol* 2019;154:445–61.
- Verdier J, Leprince O, Buitink J. A physiological perspective of late maturation processes and establishment of seed quality in *Medicago truncatula* seeds. The model legume *Medicago Truncatula*. John Wiley & Sons Ltd 2019;44–54.
- Wang Q, Garrity GM, Tiedje JM et al. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007;73:5261–7.
- Weber H, Borisjuk L, Wobus U. Molecular physiology of legume seed development. *Annu Rev Plant Biol* 2005;56:253–79.
- Wright ES. Using DECIPHER v2.0 to analyze big biological sequence data in R. *R Journal* 2016;8:352.
- Yang L, Danzberger J, Schöler A et al. Dominant groups of potentially active bacteria shared by barley seeds become less abundant in root associated microbiome. *Front Plant Sci* 2017;8:1005.
- Zengerer V, Schmid M, Bieri M et al. *Pseudomonas orientalis* F9: a potent antagonist against phytopathogens with phytotoxic effect in the apple flower. *Front Microbiol* 2018;9:145.
- Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008;18:821–9.