Tracing the evolutionary routes of plant-microbiota interactions

Carmen Escudero-Martinez and Davide Bulgarelli

Address: University of Dundee, Plant Sciences, School of Life Sciences, Dundee, United Kingdom
Corresponding Author: Bulgarelli, Davide (d.bulgarelli@dundee.ac.uk)

Highlights

• Soil characteristics and stress events are main drivers for plant-microbiota interactions in natural ecosystems;

• Host phylogeny fine-tunes the composition of the microbiota inhabiting wild plants;

• In agricultural ecosystems, selection for crop yield and external inputs have likely reduced the genetic repertoire of the domesticated microbiota;

• Large scale comparative microbial genomics is needed to dissect the full genetic potential of the crop microbiota;

• A research framework encompassing molecular microbiology and crop genomics will be key to identify plant genes shaping the microbiota.
Abstract

The microbiota thriving at the root-soil interface plays a critical role in supporting plant growth, development and health. The interactions between plant and soil microbes can be traced back to the initial plant’s colonisation of dry lands. Understanding the evolutionary drivers of these interactions will be key to re-wire them for the benefit of mankind. Here we critically assess recent insights into the evolutionary history of plant-microbiota interactions in natural and agricultural ecosystems. We identify distinctive features, as well as commonalities, of these two distinct scenarios and areas requiring further research efforts. Finally, we propose strategies that combining advances in molecular microbiology and crop genomics will be key towards a predictable manipulation of plant-microbiota interactions for sustainable crop production.

Keywords: rhizosphere, microbiota, co-evolution, plant phylogeny, crop domestication, sustainable agriculture
**Introduction**

The capacity of establishing interactions with soil microbes was one of the key factors underpinning plant’s transition from water to land: fossil evidence indicates that plants engaged in symbiotic associations with arbuscular mycorrhizal fungi as early as 400 million years ago [1]. The adaptive value of this capacity has been retained throughout plant’s evolutionary history: similar to their animal counterparts, land plants are effectively holobionts hosting a wide variety of microorganisms in the vicinity of and within plant organs, collectively referred to as the plant microbiota [2]. In particular, microbes thriving at the root-soil interface appear critical for enhancing mineral mobilisation from soil for plant uptake and pathogen protection [3-5]. These plant-microbial assemblages are not randomly assorted: their taxonomical and functional composition determines mutualistic, commensal, and parasitic interactions within the plant-defined microbial habitats [6]. Likewise, the plant genome emerged as determinant for, at least a part, of the plant microbiota [7]. Consequently, understanding the evolutionary trajectories of plant-microbiota interactions and our ability to capitalise on them for plant’s adaptation to future climatic scenarios will be critical for sustainable agriculture [8].

In this manuscript, we will evaluate recent studies focusing on the evolutionary relationships between plants and their associated microbiotas. We will compare ‘long-term’ relationships, occurred at an evolutionary scale of millions of years, with ‘short-term’ relationships, i.e., those that have arisen since the inception of agriculture and marked by crop domestication and plant breeding. In addition, we will discuss evidence of microbial evolution within the plant microbiota. An exhaustive appraisal of the current literature is beyond the scope of this manuscript: we will therefore focus on the bacterial
communities thriving at the root soil interface. Finally, we will illustrate how this knowledge can be mined to efficiently integrate plant-microbiota interactions in crop development-breeding programmes.
'Long-term' evolutionary relationships between plants and their microbiotas

Owing to its global distribution and wide range of adaptation, the model plant *Arabidopsis thaliana* represents an ideal system to study how host-microbiota interactions impacted on plant’s adaptation to the environment [9]. A study comparing the root-inhabiting communities of *A. thaliana* and three related *Brassicaceae* species indicated that 17% of the variation in community composition could be attributed to the host species, with the microbiota inhabiting the roots of *A. thaliana* and *Cardamine hirsuta*, a species which diverged from the former ~35 million years ago, being the more distinct. Yet, these differences could be attributed to the enrichment of a limited number of abundant bacterial members of the orders Actinomycetales, Burkholderiales, and Flavobacteriales, and these enrichments are relatively conserved between host plants and more dependent on the soil type [10]. These results indicated that the nature of the soil, which is one of the main sources of inoculum for plant microbiota, can impose a larger selective pressure on plant-associated communities than host phylogeny.

Consistently, a study conducted using a natural soil chronosequence revealed that edaphic factors are a primary determinant for the bacterial microbiota of 31 host plant species, including lycopods, ferns, gymnosperms, and angiosperms [11]. Yet, multivariate statistical analysis conducted on the abundances of the plant-associated bacteria revealed a significant signature of host phylogeny in the microbiota, with a bias for members of the genera *Bradyrhizobium*, *Burkholderia*, *Rhizobium* and major uncharacterized lineages such as WPS-2, Ellin329, and FW68. Owing to the fact that lycopods diverged from vascular plants ~400 million years ago, these data provide an
evidence that the assembly of a diverse microbiota is an ancient evolutionary trait in plants [11].

The signature of host phylogeny on the composition of the bacterial microbiota may vary depending on the microhabitat investigated. For instance, a ‘common garden experiment’ conducted using 30 angiosperms spanning 140 million years of evolution revealed 40% of microbial variation in the endosphere, i.e., the communities thriving within the root corpus, as opposed to only 17% of microbial variation in the rhizosphere i.e., the thin layer of soil surrounding plant roots, explained by host species [12].

Consistently, host phylogenetic relatedness correlated with microbial diversity in the endosphere but not in the rhizosphere. Interestingly, the application of a drought stress in the tested plants resulted in a three-fold enrichment of members of the family Streptomycetaceae in the endosphere of stressed-plants regardless of their phylogeny. Of note, this selective enrichment was not triggered in either the cognate rhizosphere samples or in inert wooden samples used as an a control [12].

Strikingly similar results were obtained by a comparative analysis of the microbiota associated with 18 species of the Poaceae family, which showed that host genetic diversity (determined using the sequences of three chloroplast genes) significantly correlate with bacterial diversity in the endosphere but not always with the one retrieved from the rhizosphere compartment [13]. Furthermore, once this panel of plants was exposed to drought stress a 3.1-fold increase in the endosphere populations of Actinobacteria, as compared with 2.3- and 1.5-fold increase in rhizospheres and soils, respectively, was recorded [13].

Taken together, these results indicate that the phylogenetic signatures of the bacterial microbiota are compartment dependent (i.e.,
the different magnitude in either the rhizosphere or endosphere) and suggest that these can be swiftly modulated by abiotic factors towards a stress-adapted microbiota (e.g., the selective enrichment of Actinobacteria under drought conditions).

Whether the host phylogenetic selection on the microbiota thriving at the root-soil interface represents an environmental adaptation or, rather, an evolutionary footprint remains to be elucidated.
‘Short-term’ evolutionary relationships: the domestication of the plant microbiota.

A key feature of cultivated plants is represented by the processes of domestication and breeding, an on-going anthropic selection which interjected the evolutionary history of crops [14]. The net result of these processes is an erosion of the genetic diversity of plants whose growth and development in the field is often promoted with external inputs such as fertilisers and other agrochemicals [15]. Of note, these external inputs may interfere with the establishment of plant-microbe symbiotic assemblages [16] [17].

How did these modifications impact on the recruitment and maintenance of the microbiota thriving at the root-soil interface, considering that modern cultivated varieties and wild ancestors diverged ~10,000 years ago and crops have predominantly been selected for yield traits?

Studies conducted with domesticated food crops indicated that the positioning on the breeding history i.e., wild accessions, ancestral or different modern varieties, significantly impacts the composition of the microbiota in barley [18] bean [19], maize [20] and rice [21], albeit with a proportion of variance explained ranging from ~5% to ~13%. Congruently, a meta-analysis conducted with sequencing information from a broader range of crop species suggested a ‘dichotomy’ in the taxonomic affiliation of the microbiota with the enrichment of members of Actinobacteria and Proteobacteria in modern varieties “opposed” to the enrichment of members of Bacteroidetes in the more ancestral types [22].

Interestingly, these recruitment patterns display a stress-inducible component: drought stress promoted the enrichment of Actinobacteria in the root-rhizosphere and
root communities of *Oryza sativa* and *Oryza glaberrima*, two domesticated rice species, in three distinct soil types [23].

Furthermore, field trials conducted with several inbred maize lines identified a subset of ‘hereditable bacteria’, i.e., bacteria whose abundance was significantly associated to the plant genotypes, in the rhizosphere microbiota, although soil and seasonal variation significantly impacted on these plant-bacterial assemblages [20] [24].

Despite this host-genotype specificity, and unlike what observed for wild species, no obvious relationships between host phylogeny and microbial diversity have yet been reported within the same lineage of a given crop. Examples from maize using either a high resolution single nucleotide polymorphisms (SNPs) information [20] or microsatellite sequences [25] failed to identify a significant correlation between plant genetic relatedness and bacterial diversity in the rhizosphere.

A possible explanation for these observations is that the microbial community assembly in domesticated plants is governed, at least in part, by a few major alleles, rather than by many alleles of small effect located throughout the genome [20]. Consistently, mono-mendelian mutations in a specific root trait, root hairs, perturbed ~18% of the rhizosphere communities in barley [26]. Similarly, the rice gene NRT1.1B, encoding a nitrate transporter and sensor whose sequence differs in the *indica* and *japonica* type, shapes both the taxonomic and functional composition of the rice microbiota. Of note, this effect displays a bias for microbial genes implicated in the nitrogen biogeochemical cycle [27].

Taken together, the results discussed in these sections point to a scenario where domesticated plants have not lost the capacity to shape the soil biota per se. Rather, these relationships seems to follow the same pattern observed for
natural ecosystems, whereby the soil type and the occurrence of stress events are capable of shifting the composition of plant-associated communities. Yet, using the variance explained by the host genotype in amplicon sequencing surveys as a readout, domesticated plants appear to exert a relatively limited selection on their microbiota (compared to wild counterparts. Of note, this selection) and this could be traced to a few major genes in the plant genome.

A prediction of these observations we predict that the genetic diversity of the crop microbiota is likely reduced compared to the one of the microbial communities associated to wild plants. Coupled with the application of anthropic inputs to crop, this undermines the resilience and sustainability of agroecosystems to multiple stressors, including climate change. That the genetic diversity of the domesticated microbiota likely mirrors the reduced, compared to wild plants, genetic diversity of their host. In this scenario, the resilience of the agroecosystems to stress events is intrinsically linked to external inputs. We therefore propose that an increased genetic diversity of the crop-associated microbiota will contribute to conjugate sustainable yield with a reduced pressure-footprint of agriculture on the environment (Figure 1).

The evolution of the microbes within the plant microbiota

It is worth considering an intrinsic limitation of the presented studies, which predominantly relied on amplicon sequencing surveys. The building blocks of these studies are represented by the so-called Operational Taxonomic Units (OTUs) [28] or Amplicon Sequencing Variants (ASVs) [29] of the 16S rRNA gene. These may fail to recapitulate the full extent of genetic diversity encoded by the plant microbiota.
This has been elegantly demonstrated by a recent study which compared the genomes of 1,524 *Pseudomonas* strains associated to a single bacterial OTU retrieved from *Arabidopsis thaliana* leaves across seasons and multiple natural host populations. Strikingly, this study revealed that within the same OTU, co-existed *Pseudomonas* strains that diverged ~300,000 years ago [30]. A distinctive feature of these strains is that, despite being potential pathogenic on their natural host, these were assembled into genetically diverse populations as opposed to what is often observed in agricultural settings, where pathogens give rise to genetically identical microbial lineages [31]. These observations support the notion that the ‘wild microbiota’ may be genetically less homogenous than the domesticated microbiota regardless of the apparent lack of qualitative differences in amplicon sequencing surveys.

Similarly, a comparative genomics study of 944 novel genomes of bacterial representative of the Rhizobiales, a core lineage of the plant microbiota [7] and isolated from multiple legume and non-legume host plants, revealed that commensal lifestyle exhibited by these strains predated the acquisition of genes required for nodulation [32]. Thus, being part of a plant-associated microbiota can act as a catalyst for microbial diversification.

These examples clearly indicate the power of comparative microbial genomics to dissect the full extent of the genetic potential of the microbiota. Therefore, further studies on the plant microbiota will benefit from a) the development of indexed microbial collection of given hosts, similar to the ones available for model plants [33], the integration of this resource with b) amplicon sequencing survey and whole genome comparison [34] and c) attempts at genome reconstruction from metagenomics datasets [35].
Re-wiring the evolutionary trajectories of plant-microbiota interactions for sustainable agriculture

The knowledge extracted from the different evolutionary trajectories (i.e., long versus short term) and the reconstruction of the relatedness between host phylogeny and microbial diversity can assist the breeding for the plant microbiota, resulting in future crops better equipped for climate-smart scenarios [36]. For instance, this can be achieved by crosses between wild relatives and modern varieties among interfertile species which can serve in genetic mapping analyses to discover gene/loci putatively shaping the microbiota. Examples of these approaches are mainly available for the phyllosphere of Arabidopsis [37] and maize [38]. It would be interesting exploring these approaches also for microbial communities thriving at the root-soil interface (Figure 2a). Owing to the impact of the soil type on the microbiota, the discovery of these genes/loci can be expedited by the availability of genome-annotated, geographically referenced genotypes of wild and domesticated plants which is now available for crop species with complex genomes such as barley [39] or wheat [40] [41].

In parallel, a ‘candidate gene approach’ can be deployed for genes putatively implicated in microbiota recruitment. This has recently been demonstrated for the model plant A. thaliana where a series of root metabolites, thalianyn, thalianin and arabidin, derived from the triterpene biosynthetic pathway, were implicated in microbiota recruitment using both mutant plants and by direct application of triterpene-derived metabolites [41].

This initial gene discovery phase can be the complemented by integration of the gene/gene variants of interest into the genome of modern crop varieties. Novel gene
editing tools such as CRISPR-Cas9 can produce targeted insertions, deletions, amino acid exchanges or regulate gene expression [42] (Figure 2b). As proof of concept, gene editing enabled a de novo domestication of a wild tomato by introgressing the introgression of up to six loci involved in tomato domestication from an elite tomato variety, while maintaining most of the wild ancestor traits [43]. The development of these novel plant genotypes can now be accelerated by ‘speed breeding’ which consists of creating optimal abiotic conditions under a controlled environment for a determinate crop to minimize its life cycle, reducing the time between generations [44] [45] (Figure 2).

With the availability of indexed crop-specific microbial collections, synthetic communities of a limited number of strains (SynComs) can be developed. SynComs can be used to mimic an entire microbiota and validate the impact of given host genes on the plant phenotype. This approach was pioneered in the model plant A. thaliana for the identification of host genetic traits shaping the phyllosphere microbiota [46], to study the impact of the host immune system [47] and phosphorus nutrition [48] on the root-inhabiting communities. Interestingly, SynComs have been applied also to crop plants such as maize and rice to identify key metabolic properties of their microbiotas, [27,49]. In this scenario, the application of SynComs with specific attributes can be used, for instance, to increase the access to the limited soil nutrients and/or to modulate the host immune responses against pathogens [5, 47-49]. Groups of bacteria isolated from plants containing the genes/loci responsible for the microbial phenotype can readily be grown in a gnotobiotic system to confirm that the plant genetics together with selected microbes can induce the phenotype of interest prior further validation under soil conditions (Figure 2c).
Conclusions

The recent history of crop domestication and breeding has diverted crop plants from the evolutionary trajectories of their wild counterparts by selecting genes mainly associated with productivity under high-input conditions. This approach neglected the contribution of the microbiota to plant growth, development and health. Thus, domestication and breeding have likely eroded the genetic diversity of the crop-associated microbial communities although the full impact of these processes on the crop microbiota remains to be fully elucidated. We argue that current crop selection based on artificial inputs is unsustainable on the long term. It is therefore necessary to dissect the breeding history of crops and their environment to accurately determine microbial-associated traits available in the wild and cultivated germplasm and the plant genes shaping these traits [50]. A novel research framework embracing state-of-the-art approaches in molecular microbiology and crop genomics can expedite the achievements of these tasks.

Figure legends

Figure 1: The evolutionary trajectory of plant-microbiota interactions from natural ecosystems to future agricultural scenarios

The transition from natural ecosystems to agroecosystem has been marked by plant domestication and the breeding for crop yield, with no recognition of plant-microbiota interactions and their impact on host and microbial genetic diversity. An increased understanding of these interactions will contribute to develop novel crops whose yield will be less dependent on external inputs. Boxes depict the key features of each scenario with focus on the contribution to ecosystem’s resilience of plant-microbiota interactions.
Figure 2. A common research framework for dissecting and capitalising on plant-microbiota interactions.

(a). Crops wild relatives are a main source of genetic variability translated in traits that can be introgressed into elite varieties. Crosses between elite crops and wild relatives can be used to discover new microbial plant traits in genetic mapping experiments. The speed breeding technique can accelerate the achievement of this task.

(b). Gene editing techniques can be used to manipulate plant genes shaping the microbiota previously identified by mapping experiments and/or by candidate gene approach. The speed breeding technique can accelerate the achievement of this task.

(c) Synthetic communities (SynComs) can be inoculated in plants generated by plant breeding or gene-edited plants to gauge the impact of host genes on microbial recruitment and host performance prior field trials validation.

Conflicts of interest

The authors declare no conflict of interest

Acknowledgements

This work was supported by a BBSRC grant (BB/S002871/1) awarded to DB. We thank Dr Rodrigo Alegria Terrazas (University of Dundee) and Dr Alessandro Passera (University of Milan) for the critical comments on the manuscript.
References and recommended reading


microbiome across plant phyla along a tropical soil chronosequence.

*Nature Communications* 2017, 8.

12. Fitzpatrick CR, Copeland J, Wang PW, Guttman DS, Kotanen PM, Johnson MTJ: 

This manuscript identified a correlation between endosphere microbial diversity across a group of phylogenetically distant angiosperms. Drought shifted the composition of root microbiomes, where Actinobacteria were predominant. Plant drought tolerance was correlated with increased relative abundances of Streptomyces in the endosphere.


The relatens between host phylogenetic distance and the microbiome diversity within root the endophytic community of 18 different members of the family Poaceae. These plant hosts displayed different degrees of drought resistance and the application of this stress produced shifts in bacterial community towards Actinobacteria.


NRT1.1B, a rice nitrate transporter and sensor, was responsible for the differential enrichment of bacteria between the rice varieties indica compared with japonica types. Metagenomic approaches showed that mutants plants for this transporter seem to be depleted for ammonification. SynComs isolated from indica varieties inoculated into an indica variety improved plant growth performance evidencing an intertwined relationships between plant breeding and the microbiota.


The authors identified in A. thaliana the interaction of different metabolites derived from the triterpene biosynthetic pathway (thalianyn, thalianin and arabidin), with root microbiota recruitment. Functional validations were conducted using mutant plants and direct application of these triterpene-derived metabolites.


The authors performed a de novo domestication of the wild tomato species Solanum pimpinellifolium. This domestication was based upon the introgression of up to six yield-related genes by CRISPR-Cas9 technology (targeted reverse genetic engineering). The edited plants displayed characteristics of increased fruit size and number, along with an increment in lycopene similar to a domesticated tomato.


Comprehensive review of the lastest breeding technologies available (speed breeding, high-throughput phenotyping, genotyping and marker-assisted selection, genome editing, genomic selection and de novo domestication) that will accelerate the generation of novel crops with improved characteristic for feeding an increasing population.


(a) Plant breeding

- Elite crop
- Wild relative
- Mapping population
- Speed breeding
- Fine mapping
- Parental alleles
- Near isogenic lines

Top-down approach

(b) Gene editing

- Cas9
- gRNA
- Target site
- DNA template
- Homologous recombination
- GOI
- Mock microbiota
- Selected microbiota

Bottom-up approach

(c) Microbial functional validation

Mock microbiota
Selected microbiota