What is the contribution of the phyllosphere microbiota to the health of apricot trees?

Master BS - qBio

Bich Ngoc DO

Supervisors: Hervé Sanguin (CIRAD) & Bénédicte Quilot (INRAE)

Co-supervisors: Karine Adeline (ONERA) & Thomas Defferier (Vegepolys Valley)

Host Lab: PHYTOBIOM – PHIM, CIRAD Baillarguet

Total days worked: 80 days

July, 2026

Remerciements

Bien que le français ne soit pas ma langue maternelle, j'ai souhaité rédiger ces remerciements en français afin d'exprimer plus sincèrement ma gratitude envers chacun. Si certaines formulations manquent d'élégance ou de précision, j'espère que vous voudrez bien m'en excuser.

Je tiens à exprimer ma sincère gratitude à toutes les personnes qui ont contribué au bon déroulement de ce stage au sein de l'unité PHIM, CIRAD.

Mes premiers remerciements s'adressent à mon encadrant principal**, Hervé Sanguin**, chercheur au sein de l'équipe MicroQuar, pour son encadrement précieux et son accompagnement tout au long de ce stage. Au-delà de ses compétences scientifiques, il a toujours su créer un environnement de travail bienveillant et détendu, en me prodiguant des conseils avisés qui dépassaient largement le cadre du laboratoire. Je lui suis particulièrement reconnaissant(e) de m'avoir emmené(e) visiter le verger d'abricotiers et d'avoir favorisé la cohésion d'équipe à travers des activités conviviales qui m'ont aidé(e) à m'intégrer. Son soutien, tant sur le plan des compétences en recherche que des compétences interpersonnelles, a été déterminant dans mon parcours au cours de cette période.

Je remercie chaleureusement ma co-encadrante, **Bénédicte Quilot** (INRAE, Avignon), dont l'implication et la patience ont été remarquables malgré la distance géographique et mes connaissances initiales limitées en biologie de l'abricotier. Sa douceur et sa bienveillance naturelles ont toujours créé un cadre de travail agréable et rassurant. Elle s'est constamment rendue disponible pour me guider et m'a fourni les connaissances et les ressources nécessaires pour mieux appréhender ce système végétal.

J'adresse également mes sincères remerciements à Estelle Tournier, assistante technicienne au sein de l'équipe, pour son encadrement minutieux et patient tout au long des travaux en laboratoire. Sa disponibilité permanente, sa volonté de répondre à toutes mes questions et son engagement à faciliter mes expériences ont été d'une aide précieuse.

Je remercie également **Frédéric Mahé**, chercheur au sein de l'équipe MicroQuar, pour sa contribution essentielle au traitement des données de séquençage à l'aide de son pipeline bioinformatique. Ses explications m'ont permis non seulement de maîtriser les données brutes issues du séquenceur, mais aussi de mieux comprendre les approches modernes utilisées en analyse du microbiome.

Je souhaite exprimer ma gratitude à **Morgane Roth** (INRAE) pour son aide précieuse dans la clarification des métadonnées liées aux abricotiers et pour les perspectives d'analyse stimulantes qu'elle a proposées. La visite du verger où les échantillons ont été collectés a été une expérience particulièrement enrichissante et mémorable.

Je remercie également mes co-encadrants **Karine Adeline** (ONERA) et **Thomas Defferier** (Vegepolys Valley) pour leurs contributions constructives ainsi que pour le financement qui a rendu ce stage possible.

Merci au GIS Fruits pour le financement de la bourse de stage.

Je souhaite également remercier **Pany**, stagiaire au sein de l'équipe, pour avoir été une véritable compagnonne de stage. Bien que notre temps ensemble ait été court, cette expérience partagée a été profondément significative j'ai eu la chance de trouver en toi une amie sincère, malgré la légère barrière de la langue. Merci pour ta présence et ta bonne humeur au quotidien.

Je remercie également **Sophie** et **Manon**, membres de l'équipe, pour leur enthousiasme et leur générosité. Leur bonne humeur constante et leur aide spontanée chaque fois que j'en avais besoin ont rendu le quotidien du laboratoire bien plus agréable.

Enfin, je souhaite exprimer toute ma reconnaissance aux membres de l'équipe MicroQuar ainsi qu'à l'ensemble des chercheurs du BatK, PHIM, CIRAD. Malgré la barrière de la langue, mon français étant encore limité, chacun a fait preuve d'une grande gentillesse, de patience et de bonne volonté à mon égard. Cet environnement chaleureux a fait de ce stage une expérience heureuse et profondément enrichissante, dont je garderai un souvenir durable.

At the end, I would also like to express my heartfelt gratitude to the *Master qBIO programme*. These two years have been truly unforgettable, filled with knowledge, discoveries, and countless memorable experiences both inside and outside the classroom. I am particularly thankful to the programme coordinators, **Prof. Luca Ciandrini** and **Prof. Cherine Bechara**, who were not only exceptional academic mentors but also genuine life guides. Their invaluable advice, their dedication, and their constant care for their students made a world of difference, especially for an international student far from home. That warmth and attentiveness will stay with me long after graduation.

I also wish to thank all the professors and lecturers who contributed to my academic journey throughout this Master's programme. Finally, a special thank you to my fellow qBIO students, with whom I have "survived" and thoroughly enjoyed these two remarkable years. I could not have asked for better companions along the way.

Table of Contents

Abstract

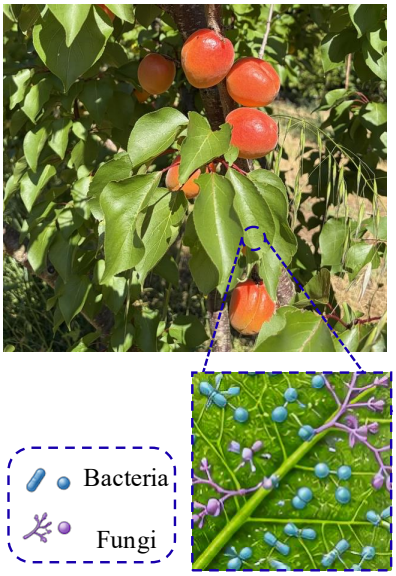
Apricot *(Prunus armeniaca L.)* is an economically important fruit tree whose productivity is strongly affected by foliar diseases and associated microbiome dynamics. The phyllosphere microbiome is increasingly recognized as an integral component of the plant holobiont, shaped by host genotype, tissue niche, and environmental conditions. Here, we characterized the fungal and bacterial phyllosphere of a genetically diverse apricot core collection grown in a common-garden orchard and investigated how host genetic background and foliar disease severity structure microbial communities. Leaves from 145 genotypes (ITS2) and 119 genotypes (16S rRNA V4 region), spanning seven germplasm origins, were analyzed using amplicon sequencing. Diversity and composition were assessed using Hill-number metrics, multivariate variance analyses, ordination, and variance partitioning. The fungal community was dominated by a conserved *Didymella–Aureobasidium* core shared across all origins, with richness and composition showing limited genotypic variation; geographic origin explained only a minor fraction of fungal turnover. Bacterial communities were more variable: richness differed significantly among host origins, and an origin-based compositional signal accounted for a small but highly significant proportion of variation ($R^2 = 4.5\%$, $P = 1\times10^{-4}$), indicating stronger host-genotype filtering of bacteria than fungi. Disease effects were kingdom-specific. *Coryneum* severity was linked to bacterial richness ($R^2 = 13\%$), rust to fungal richness ($R^2 = 6\%$). No significant relationship was detected for *Monilia*. Overall. these findings provide the first integrated characterization of the apricot leaf holobiont, suggesting a stronger relation between bacterial community with host genetic background than fungal communities, within a reciprocal host–microbiome–disease relationship.

Counts: 249 words

Graphical Abstract

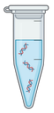
Apricot phyllosphere microbiome

(Prunus armeniaca)

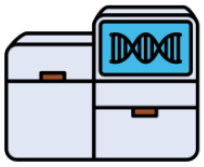


1. Leaf sampling from Apricot
158 genotypes

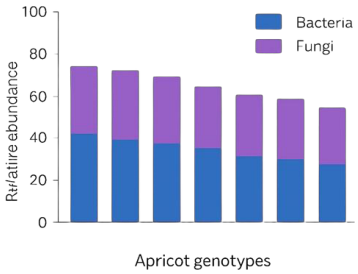
2. DNA extraction



3. **16S rRNA** and **ITS2** amplicon sequencing



4. Bioinformatics analysis



Key findings

1. Who is there?

Fungi

Didymella + unclassified *Didymella*

Aureobasidium

Conserved core, shared across all origins

Bacteria

Mostly unclassified lineages, plus:

Methylpbacterium
Sphingomonas
Hymenobacter

Massilia

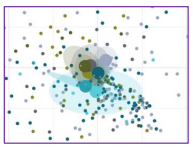
More variable between origins

Fungi form a stable core while **bacteria are more genotype-variable.**

2. Who shapes it?

Does host origin structure the community? (NMDS)

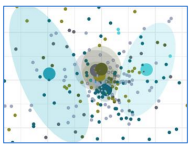
Fungi



PERMANOVA
$R^2 = 0.02$,
P=0.27

Origin NOT significant

Bacteria



PERMANOVA
$R^2 = 0.05$,
$P = 1 \times 10^{-4}$

Origin significant

Host genetic background structures bacteria more strongly than fungi.
(Horn-Morisita dissimilarity)

3. Does disease matter?

Variance in microbial richness explained by each disease

Fungal richness

Rust

6% (P = 0.001)

Bacterial richness

Coryneum

13% (P = 0.001)

Monilia

no detectable effect

Disease leaves kingdom-specific marks

Coryneum tracks bacteria, rust tracks fungi.
Effects are modest and correlative – a reciprocal host-microbiome relationship.

A. Introduction

1. Plants and their microbial partners

Since their colonization of land roughly 450 million years ago, plants have never existed in isolation. Any plant surface from root to leaf is colonized by a dense and functionally diverse community of bacteria, fungi, and other microorganisms. Far from being passive bystanders, these microorganisms interact continuously with the host, shaping its growth, nutrition, and resistance to disease[1]. The growing recognition of this interdependence has prompted a fundamental shift in how we conceptualize the plant: no longer as a discrete biological unit, but as an ecological entity, the "plant holobiont", in which the phenotype and fitness of the host are inseparable from the composition and activity of its associated microbiome.[2]

This holobiont perspective is grounded in an evolutionary argument. Over hundreds of millions of years, plants and their microbial partners have co-evolved a suite of reciprocal mechanisms that consolidate their association. Plants have developed metabolite-mediated strategies to selectively recruit beneficial microorganisms and, upon perception of microbial signals, activate immune responses that modulate community composition. Conversely, community members have evolved host-adapted traits and antagonistic interactions among themselves, collectively preventing the unchecked proliferation of pathogens within the holobiont[2]. The outcome of this co-evolutionary process is a structured and partially heritable microbiome whose composition correlates with plant phylogenetic relatedness across diverse ecosystems[3].

The ecological and agronomic consequences of plant-microbe partnerships are far-reaching. Phyllosphere and rhizosphere microbial communities collectively regulate key ecosystem functions, including atmospheric nitrogen fixation, decomposition, and water cycling[4]. At the level of individual plants, microbiome members have been documented to promote growth, enhance tolerance to drought and nutrient limitation, and suppress foliar or soil-borne pathogens. Despite this multifaceted importance, the majority of plant microbiome research has historically concentrated on the belowground compartment, the rhizosphere and root endosphere[5]. Yet, the aboveground surface of the plant, collectively termed the phyllosphere, represents one of the largest microbial habitats on Earth, yet its ecology, assembly rules, and role in plant health remain comparatively underexplored[6]. Bridging this gap is both a scientific priority and a practical necessity, particularly for perennial crops where foliar disease is a primary threat to productivity[4].

2. The Phyllosphere as a microbial habitat

The phyllosphere is an important niche of the plant which designates far more than a simple leaf surface: it constitutes a dynamic, heterogeneous microenvironment at the plant-climate interface, extending from the outer edge of the aerodynamic leaves, stems, flowers, pollen, nectar, and fruits[7]. This habitat is shaped by a remarkable complexity of abiotic gradients (temperature, UV radiation, water availability, and volatile organic compound fluxes…) and biotic forces (insect herbivory, pollinator visitation, and pathogen pressure), making it one of the most physically and chemically heterogeneous microbial environments on Earth[8]. Microorganisms colonize this habitat as epiphytes on the surface or as endophytes within plant tissues, and these two niches impose fundamentally different selection pressures[9]. Epiphytes must withstand physical stress directly, while endophytes are more tightly filtered by host physiology and immunity. Despite these constraints, microbial densities on leaf surfaces routinely reach 10^6–10^7 cells per cm^2, comprising bacteria, fungi, yeasts, and a growing list of other groups including protists and viruses[10].

Across studies and biomes, the bacterial phyllosphere is consistently dominated by Proteobacteria, Actinobacteria, Bacteroidetes, and Firmicutes, while fungi are overwhelmingly represented by Ascomycota and Basidiomycota with the genera such as *Alternaria, Cladosporium,* and *Candida*

among the most recurrently reported [11]. Yet this apparent taxonomic consistency masks considerable functional and compositional variation, driven by four interacting eco-evolutionary processes: dispersal, diversification, selection, and drift[6]. Host identity and plant functional group consistently emerge as the strongest predictors of phyllosphere community structure, outweighing soil diversity and short-term environmental disturbance[12]. Importantly, bacteria and fungi respond differently to these drivers, fungal communities tend to be more host-specific, while bacterial communities are more sensitive to abiotic gradients, which underscores the importance of characterizing both kingdoms jointly [11].

3. Drivers of phyllosphere microbiome assembly

The composition of the phyllosphere microbiome is not simply a reflection of the surrounding microbial pool. While dispersal from soil, air, and neighbouring plants provides the initial colonist pool, community assembly is ultimately governed by a series of selective filters operating at multiple scales[4]. At the broadest level, abiotic factors such as climate, season, rainfall, and agricultural management modulate leaf surface chemistry and microhabitat conditions, creating temporal and spatial variation in community structure[4]. Yet, among all drivers assessed in recent meta-analyses, host-related traits account for the majority of explained variance: plant species identity, genotype, tissue type, and developmental stage consistently emerge as the strongest predictors of phyllosphere community composition[13]. This host filtering operates through multiple mechanisms such as cuticle composition, stomatal regulation, leaf surface metabolites, and immune signalling pathways, including salicylic acid and ethylene and also results in microbial diversity that decreases progressively from soil to root to leaf, with the phyllosphere harbouring the most host-specific and lowest-diversity communities [14,15].

Host genetic control of microbiome assembly is not uniform across plant species, however, and the relative weight of deterministic versus stochastic processes varies[13]. In some species, host genotype exerts a strong and consistent selective pressure, producing species-specific and even cultivar-level microbiome signatures[3]. In others, stochastic processes such as dispersal and priority effects play a larger role, masking genetic effects[13]. In addition to these baseline dynamics, biotic disturbances like pathogen infection, herbivory, and interspecific microbial competition can reshape community composition abruptly, recruiting new taxa or disrupting existing ones[1]. The degree to which host genotype interacts with disease state to modulate microbiome structure is particularly relevant for perennial fruit cropsmwhere genetic diversity among cultivars is large, and disease pressure is recurrent [4].

4. Phyllosphere microbiome and plant health

The relationship between the phyllosphere microbiome and plant health is bidirectional[16]. A balanced microbial community confers protection against pathogens through competitive exclusion, production of antimicrobial metabolites, and induction of systemic resistance in the host. When this balance is disrupted, a state termed dysbiosis, the plant becomes more vulnerable: putative latent pathogen ‘pathobionts’ proliferate, beneficial taxa decline, and disease-associated phenotypes can emerge even in the absence of a primary pathogen [4]. This concept, initially established in the gut microbiome[17], has now been demonstrated in the phyllosphere of *Arabidopsis* mutants with impaired immune networks, where dysbiosis alone was sufficient to trigger leaf chlorosis and necrosis [4]. Beyond passive protection, plants can also actively reshape their microbiome in response to infection, the so-called “cry for help” strategy, recruiting beneficial taxa from the surrounding environment to reinforce foliar defences[18]. Evidence from citrus melanose [1] and Huanglongbing[19] illustrates this dynamic clearly: infected leaves show marked shifts in community composition, with enrichment of taxa such as *Sphingomonas*, *Pantoea*, and *Methylobacterium* that suppress the pathogen through iron competition and direct antifungal activity. Taken together, these findings position the phyllosphere microbiome not merely as

a community shaped by disease, but as an active participant in its outcome, a perspective with direct implications for how we approach disease management in crops[16].

5. Apricot: Biology, diversity and disease context

Apricot *(Prunus armeniaca* L.*)* is one of the oldest cultivated stone fruit species with a centre of origin in the Ili Valley of Northwest China from which it dispersed westward through Central Asia into the Mediterranean basin along the Silk Road[20]. This long history of dispersal and independent selection under different climatic conditions has produced a genetically diverse crop, structured into at least three major germplasm groups – Central Asian, European, and North Chinese – each carrying distinct adaptive traits[20]. Notably, apricot retains relatively high levels of genetic diversity compared to other domesticated *Prunus* species, due to its limited vegetative propagation, ongoing gene flow with wild relatives, and relatively few domesticated bottlenecks[20]. This genetic breadth makes apricot an ideal system for exploring how host genotype shapes associated microbial communities: a panel of cultivars drawn from divergent genetic backgrounds offers meaningful variation in host traits likely to filter the phyllosphere microbiome, from cuticle chemistry to immune gene expression[3].

Despite its agronomic value, the apricot is susceptible to a wide range of foliar and fruit diseases that cause substantial yield losses each season[21]. The most economically damaging include brown rot caused by *Monilinia laxa* and *M. fructicola*, which infect blossoms and shoots, with up to 80% crop loss recorded in susceptible cultivars; shot hole disease caused by *Wilsonomyces carpophilus*, a widespread polyphagous pathogen with disease prevalence exceeding 65% in affected orchards; powdery mildew (*Podosphaera pannosa*) and leaf rust (*Tranzschelia* spp.), which affect canopy health and reduce photosynthetic capacity, and *Alternaria* leaf spot (*Alternaria alternata*), commonly associated with weakened trees[22]. Chemical fungicides remain the primary management strategy for most of these diseases, but growing concerns about resistance and regulatory restrictions call for alternative, ecologically grounded approaches[5].

6. Objectives

The present study addresses a major knowledge gap in apricot by investigating the phyllosphere microbiome, whose assembly is shaped by host genetics, environmental conditions, and disease state [3,4] Existing work on *Prunus armeniaca* has focused primarily on the rhizosphere or on fruit surface communities assessed through culture-dependent methods[23]; no comprehensive study has examined the leaf microbiome in relation to both host genetic diversity and foliar disease, leaving a clear and timely research gap that this thesis aims to address. Using 16S rRNA- and ITS-based amplicon sequencing on a panel of genetically diverse apricot cultivars grown under natural orchard conditions, this work aims to: (i) characterize the global phyllosphere microbiome of *Prunus armeniaca* for the first time; (ii) assess how host genotype shapes fungal community diversity and composition across cultivars; and (iii) explore associations between microbiome structure and the incidence of foliar diseases. Together, these objectives contribute to a broader understanding of how host genetic diversity mediates plant-microbe interactions in perennial crops and lay the groundwork for microbiome-informed approaches to disease management in apricot.

B. Materials and Methods

1. Plant Material and Leaf Sampling

Leaf samples were collected in 2024 from the apricot *(Prunus armeniaca L.)* core-collection at Domaine de l'Amarine (INRAE), which comprises 158 genotypes representing the broad genetic diversity of the species across its main germplasm groups. For each genotype, three trees distributed in three blocks in the orchard were sampled as independent biological replicates. Three leaves per tree were sampled and grouped to capture within-tree leaf variation. At sampling, phenotypic metadata, including disease status and health observations, were recorded for each tree. Leaf discs (~100 mg fresh tissue) were taken using a sterile punch and immediately stored at −20°C to preserve microbial community integrity.

2. DNA Extraction

Leaf discs were pooled per tree as a first approach, freeze-dried, and ground to a homogeneous powder. Cell lysis was performed using MATAB (Mixed Alkyltrimethylammonium Bromide) buffer, a plant-adapted cationic detergent[24]. DNA purification was then carried out using the NucleoMag Plant Kit (Macherey-Nagel, Hoerdt, France) on a MagnetaPure 32 automated liquid handling robot (Macherey-Nagel) across all 474 samples. DNA concentration was measured by fluorometry using the Quant-iT™ PicoGreen™ dsDNA Assay Kit and dsDNA Reagents (ThermoFisher Scientific, Villebon Courtaboeux, France) and a Tecan Multimode Microplate Reader Spark (Tecan Group, Männedorf, Switzerland). Extracts were normalized to ~5 ng/µL as template input for PCR amplification.

3. Amplicon Library Preparation and Sequencing

Community profiling was performed using a two-step PCR approach targeting the fungal ITS2 subregion and the V4 hypervariable region of the bacterial 16S rRNA gene, with peptide nucleic acid (PNA) clamps incorporated in the 16S amplification to suppress host plastid and mitochondrial amplification.

PCR1 — Target Amplification:

For fungal profiling, the ITS2 subregion was amplified using the tagged primers ITS86F and ITS4R. For bacterial profiling, the 16S V4 region was amplified using primers 515F and 806R combined with two PNA clamps to block co-amplification of host-derived sequences: a mitochondrial PNA and a chloroplast PNA. All primers and PNA sequence were designed according to the scheme provided in Table 1.

Table 1: Primers used in the current study.

Primers	DNA Sequence (5'→ 3')	Cite
ITS-86F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNN**GTGAATCATCGAATCTTTGAA**	[25]
ITS-4R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNN**TCCTCCGCTTATTGATATGC**	
16S-515F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNN**GTGYCAGCMGCCGCGGTAA**	[26]
16S-806R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNN**GGACTACNVGGGTWTCTAAT**	
mPNA	GGCAAGTGTTCTTCGGA	[27]
pPNA	GGCTCAACCCTGGACAG	

Target sequences in **bold and* Illumina sequencing adapters in underline*. mPNA: mitochondrial PNA, pPNA chloroplast PNA*

The first-round PCR (PCR1) was performed in duplicate in a final reaction volume of 25 µl, containing 12.5 µl AmpliTaq Gold 360 Master Mix (Thermo Fisher Scientific, Waltham, MA, USA), 0.2 µM of each forward and reverse primer, and 2 µl of template DNA. For ITS2, the thermocycling program consisted of an initial denaturation at 95°C for 10 min; 35 cycles of 95°C for 30 s, 75°C for 10 s, 62°C for 30 s, and 72°C for 30 s; followed by a final extension at 72°C for 7 min and a hold at 20°C. The 75°C step was included to allow PNA clamp binding before primer annealing, thereby reducing non-target amplification. For the 16S V4 region, the thermocycling program consisted of 95°C for 10 min; 30 cycles of 94°C for 45 s, 55°C for 45 s, and 72°C for 45 s; followed by a final extension at 72°C for 10 min and a hold at 20°C. Each PCR batch included two negative controls. Amplicons showing visible bands on agarose gels were diluted 1:10 (v/v) and used as a template for the second PCR (PCR2).

Because PCR success varied among samples and between the ITS2 and 16S batches, the number of technical replicates selected for sequencing differed between datasets (Table 2).

Table 2: Sequencing input across ITS2 and 16SrRNA-V4 datasets

	ITS2 region batch	**16SrRNA-V4 region batch**
Number of replicates used for sequencing	281	137
Number of unique genotypes	145	119
Number of genotypes with 3 replicates	34	0
Number of genotypes with 2 replicates	68	18
Number of genotypes with 1 replicate	43	101

PCR2 — Indexing and Adapter Ligation

A volume of 2 µl of the diluted PCR1 pool was used as template in a second PCR (PCR2) using Taq 2× Master Mix (New England Biolabs, Ipswich, MA, USA), which incorporated dual barcode indices to Illumina sequencing adapters. PCR quality for both rounds was verified by electrophoresis on 1% agarose gels.

Library Preparation and Sequencing

All final PCR2 products, including negative controls, were quantified by PicoGreen fluorometry. Samples were pooled at equimolar ratios, and the pool was purified using NucleoMag NGS Clean-up and Size Select magnetic beads (Macherey-Nagel, Düren, Germany). Library quality and fragment size distribution were assessed on a TapeStation (Agilent Technologies, Santa Clara, CA, USA), and final quantification was performed by qPCR using SYBR Green (Thermo Fisher Scientific) on a LightCycler 480 instrument (Roche, Basel, Switzerland). The sequencing-ready library was sequenced on an Illumina MiSeq i100 platform in paired-end 300 bp mode (PE300) at the MGX genomics platform (CIRAD-AGAP, Montpellier, France).

4. Bioinformatic Analysis

Raw paired-end sequencing reads were processed with the Nextflow pipeline `'nf-metabarcoding'`[28]. Briefly, reads were first merged using VSEARCH[29](v2.31.0). Adapter sequences were then trimmed with `'cutadapt'`[30] (v5.2), and all reads shorter than 32 nt were discarded to retain only informative sequences for downstream analysis. The remaining reads were clustered into operational taxonomic units (OTUs) using SWARM (v3.1.6), which performs unsupervised, local clustering of amplicon sequences without relying on a global similarity threshold[31]. Chimeric sequences were removed to

reduce artefacts introduced during PCR amplification, using the UCHIME algorithm[32] as implemented in VSEARCH.

Taxonomic assignments were generated: ITS sequences were assigned using the UNITE database[33] (v10.0 2025-02-19), while bacterial 16S rRNA gene sequences were assigned using SILVA[34] (v138.2). Downstream community analyses were then based on the resulting swarm OTU table and taxonomic profiles.

5. Statistical Analysis

Data tables were transformed using the R `tidyverse` package (version 2.0.0)[35] and the plots were generated using the R `ggplot2` package (v4.0.3)[36]. Statistics were All statistical analyses were performed in R using `R phyloseq` package (v1.50.0)[37] and R `vegan` package (v2.7.2)[38]. Prior to downstream analyses, swarm OTU tables were merged by kingdom/phylum and genus for visualisation of relative proportion across apricot origins. Alpha diversity was assessed using Hill numbers using R `Hilldiv` package (v1.5.3)[39] to capture richness and diversity at multiple levels of taxon weighting. Specifically, q = 0 was used to represent observed species richness, q = 1 to represent the Shannon-based effective number of species (common species), and q = 2 to represent the Simpson effective number of species (dominant species). Hill evenness ratios were also calculated based on Shannon and Simpson estimations to evaluate the balance among dominant and rare taxa across origins. Group differences in alpha diversity were visualized using boxplots.

Beta diversity based on community structure analysis was performed on wisconsin-transformed abundance data, and the differences in community structure across apriwere displayed with nonmetric multi-dimensional scaling (NMDS) implemented in the `metaMDS()` function from the R package vegan, based on the Horn–Morisita dissimilarity index. Significance in the variation of community structure was assessed using PERMANOVA in the `adonis3()` function from the R `GUnifrac` package (v1.9). Multivariate dispersion was estimated by HOMOVA using the `betadisper()` and `permutest()` functions as it can affect PERMANOVA results.

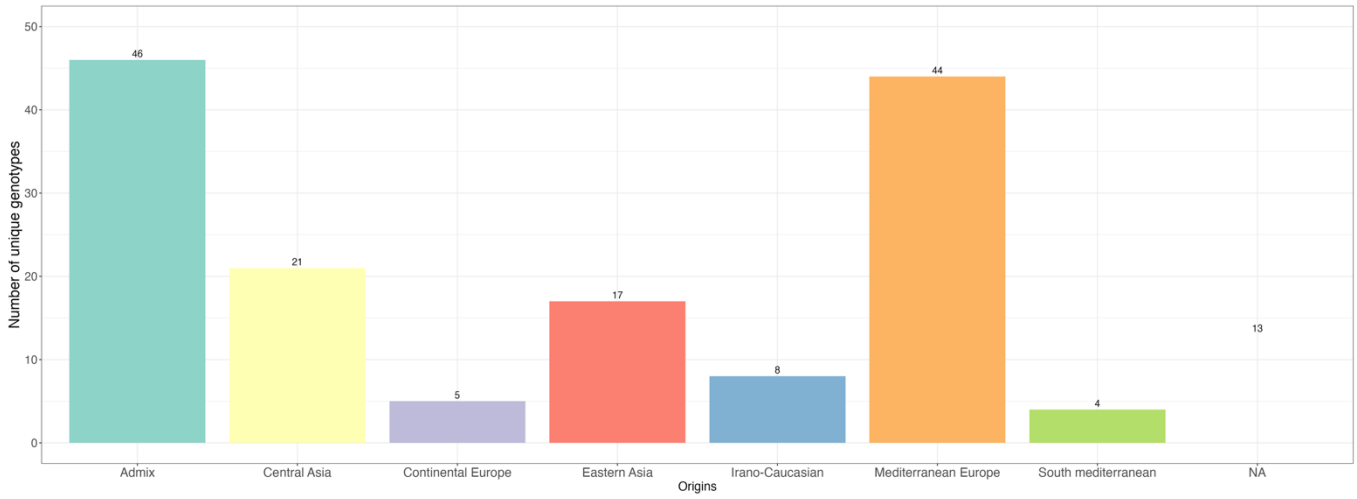
Variance partitioning analyses of community richness regarding disease from *Coryneum* spp., Rust, and *Monilia* spp. were performed using the `varpart()` function from the R package vegan.

C. Results and Discussions

1. Characteristics of the selected subset of the apricot collection

The apricot core collection was unevenly represented across geographic origins, with Mediterranean Europe and Admix contributing the largest number of genotypes (~44 genotypes), whereas Continental Europe and South Mediterranean were represented by only a few samples (Fig.1A). This distribution shows that the available panel was not perfectly balanced, but it still covered all major apricot origin groups. The correlation heatmap of the whole collection (Fig. 1B) showed that tree age was strongly associated with vigor, growth, and fruiting traits, while disease-related traits tended to be weakly or negatively associated with growth-related parameters. Together, these patterns indicate that age, vigor, and disease load are major phenotypic covariates in the collection and should be considered when interpreting downstream microbiome results[4].

Because of time constraints, only a representative subset of the collection was sequenced for the fungal and bacterial surveys. The fungal dataset comprised 145 genotypes and the bacterial dataset 119 genotypes (Fig A1), with samples distributed across all major origins, although the two batches were not identical and should therefore be considered partially overlapping subsets rather than a fully matched design. This design choice was further influenced by marker-specific PCR strategies: 16S rRNA -V4 region amplification included a PNA clamp step to reduce non-target amplification from plant hosts[27], whereas ITS amplification was performed without PNA. Despite the incomplete sequencing of the full collection, the selected genotypes were spread across origins and remained broadly representative of the apricot panel.



**NA: missing data*

Figure 1: Distribution of unique genotypes across their origins of whole collection

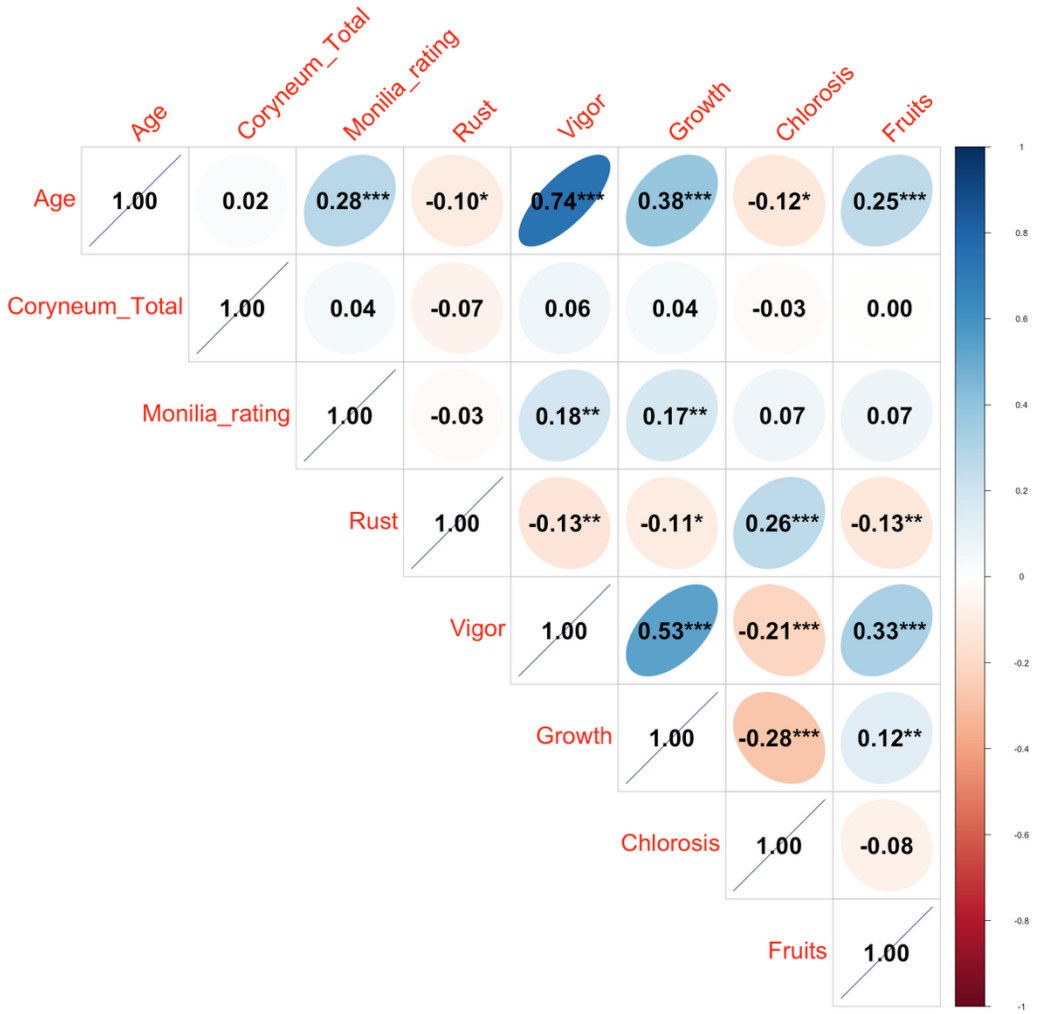


Figure 2: Pairwise Pearson correlation matrix among agronomic and disease traits assessed across 158 apricot genotypes. Each cell displays the correlation coefficient (*r*), with ellipse shape and fill color reflecting the strength and direction of the association (blue = positive; red = negative; color intensity and ellipse narrowness increase with |r|). Significance levels are indicated as: *p < 0.05, **p < 0.01, ***p < 0.001. Non-significant correlations are shown without asterisks.

2. Taxonomic composition and relative abundance

Fungal community. The fungal community of apricot leaves was dominated by a small number of cosmopolitan genera across all geographic origins examined (Fig. 3). Regardless of host origin, *Didymella* was consistently the most abundant taxon (~25-40% of mean relative abundance), followed by group of genera coming from the family *Didymellaceae* (~5-25%) *and Aureobasidium* (~5-25%) with a pooled "Others" fraction and a set of low-abundance taxa (unclassified genus from order *Pleosporales*, *Sporobolomyces*, unclassified genus from *Ascomycotaa* and *Pseudorotiacea*, *Trematеia*, *Alternaria*, unclassified genus from *Rosaceae* family)

Minor compositional differences were visible across origins. Both the Irano-Caucasian and Central Asian origins showed a comparatively higher relative abundance of *Rosaceae gen. Incertae sedis* and unclassified *Ascomycota*, potentially reflecting a richer diversity of Rosaceae-associated fungal lineages in genotypes originating near the primary center of diversity of *Prunus armeniaca*[41]. The South Mediterranean origin showed a relatively elevated abundance of *Aureobasidium*, which may reflect host lineage-specific traits shaped by adaptation to the ancestral climatic conditions of this apricot group, thereby influencing its capacity to support yeast colonization in the phyllosphere[42].

Among the dominant taxa, the consistent dominance of *Didymella* across all origins is particularly noteworthy from both an ecological and plant health perspective. This genus, belonging to the class *Dothideomycetes* (family *Didymellaceae*), is well documented as a versatile leaf-associated fungus capable of occupying both endophytic and epiphytic niches in living plant tissues, with some species

transitioning to necrotrophic pathogenicity under favorable host or environmental conditions[43]. In *Rosaceae* hosts specifically, *Didymella*-related species have been associated with leaf spots and dieback symptoms, suggesting that its persistent and dominant presence across all apricot origins may reflect both its broad ecological competitiveness in the phyllosphere and its potential as a major resident of pathobiont, a taxon that colonizes asymptomatically but retains pathogenic capacity[44]. Similarly, *Aureobasidium*, a highly competitive epiphytic yeast-like fungus, is widely recognized as one of the most ubiquitous phyllosphere colonizers across diverse plant species and climatic zones[45], consistent with its regular detection here. The co-occurrence of *Didymella* and *Aureobasidium* as dominant phyllosphere fungi is not unique to apricot: in plum (*Prunus domestica*), a phylogenetically close stone fruit host, *Aureobasidium* and *Didymella* phacae were similarly among the most consistently detected phyllosphere taxa across cultivars and phenological stages.[46] However, the particularly strong dominance of *Didymella* as the primary genus observed in our apricot dataset appears to represent a host-specific intensification of this broader stone fruit phyllosphere pattern, warranting further investigation of its ecological role and potential implications for apricot leaf health.

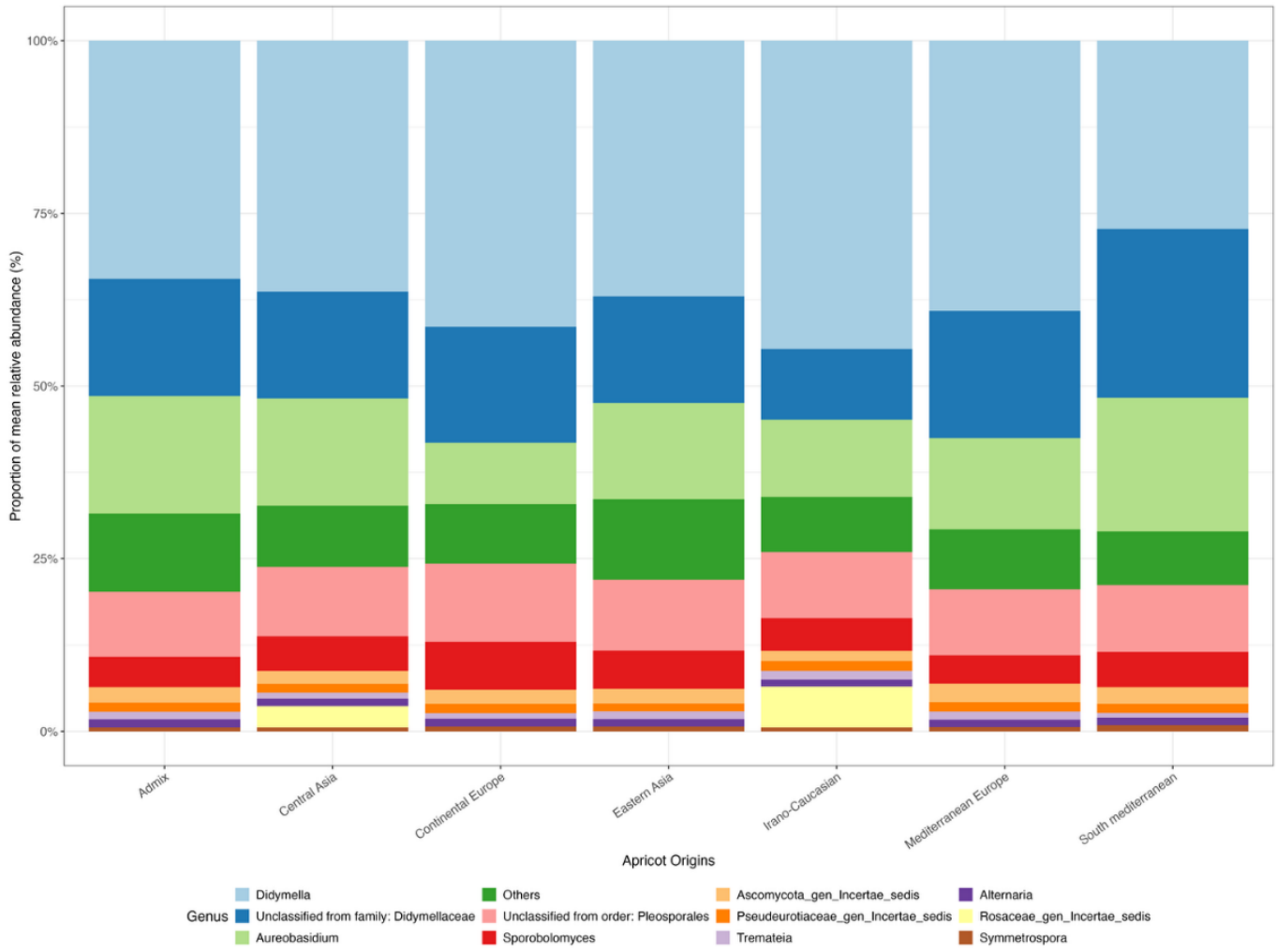


Figure 3: Taxonomic distribution of the 12 most abundant fungal genera across apricot origins. Stacked bar plots represent the mean relative abundance of each genus within each apricot origin, with genera ordered in the legend according to their overall abundance from highest to lowest. The category *"Others"* includes all remaining low-abundance taxa not shown individually.

Bacterial community. The bacterial community of apricot leaves was dominated by poorly resolved lineages across all origins (Fig. 4). Consistent with known limitations of 16S amplicon sequencing in phyllosphere systems, the most abundant category in most origins was sequences classified only as

genus *Incertae sedis* (*Incertae sedis* ~5-25%), or Unclassified genus from Class *Alphaproteobacteria* and *Actinobacteria*, reflecting the persistent difficulty in resolving phyllosphere-associated lineages to genus level using short-read marker gene approaches[47].

Among the well-characterize genera, *Methylobacteria and Sphingomonas* were among the most consistently detected named taxa across origins. Both genera are canonical members of the *Alphaproteobacteria* and are widely recognized as common phyllosphere colonizers, often associated with efficient use of leaf-derived carbon sources, stress tolerance, and potential roles in plant defense modulation[48]. Their recurring presence therefore supports the existence of a shared apricot leaf bacterial core, but the relative abundance of other genus-level groups indicates that each origin departs from this core in a different direction. The almost origins showed a mixed profile in which *Hymenobacter* co-occurred with Alphaproteobacteria and Gammaproteobacteria. This pattern suggests a community that still retains the canonical phyllosphere core, but also includes taxa with greater tolerance to radiation, desiccation, and other surface-associated stresses[49]. In comparison, the South Mediterranean origin was even more strongly enriched in *Hymenobacter* relative to Alphaproteobacteria, indicating a shift toward a more stress-tolerant assemblage rather than a classical Alphaproteobacteria-rich phyllosphere. Because all trees were grown in the same orchard, this contrast is more plausibly explained by host lineage-specific filtering than by current environmental differences[3].

The Central Asia origin stood out because Bacilli-related genera were the most prominent bacterial group among the compared regions. This suggests a phyllosphere assemblage enriched in Firmicutes-like taxa, which may reflect host traits that favor Bacilli recruitment. This interpretation is supported by recent evidence showing that phyllosphere Bacilli are selectively promoted by host genotype and structural or chemical leaf cues, and that they can contribute to plant immunity and disease resistance[50]. By contrast, the Continental Europe origin showed a stronger representation of Alphaproteobacteria, indicating a more canonical phyllosphere assemblage and a closer resemblance to the common leaf-associated bacterial core[49]. The absence of *Massilia* in both Continental Europe and South Mediterranean are also informative. The absence of *Massilia* in Continental Europe and South Mediterranean suggests that this genus is a conditionally recruited phyllosphere taxon rather than a universal apricot-associated core member, consistent with recent studies showing that *Massilia* abundance can depend on host genotype, nutrient conditions, and disease-related leaf traits[51].

Overall, the genus-level profile indicates that apricot origins differ in the ecological strategy of their bacterial communities: some retain a canonical Alphaproteobacteria-rich phyllosphere, whereas others shift toward stress-tolerant *Hymenobacter, Massilia* or Bacilli-dominated assemblages.

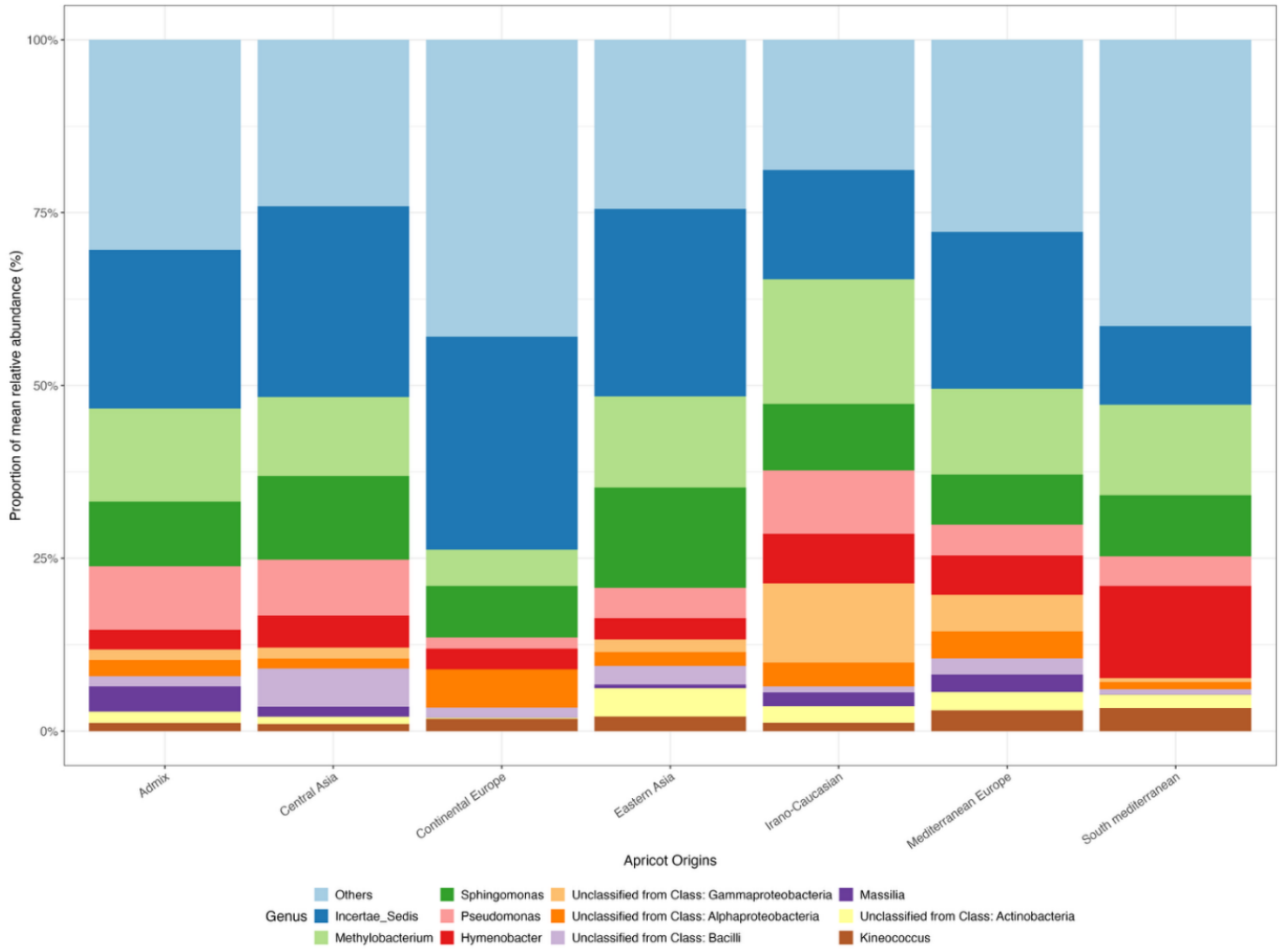


Figure 4: Taxonomic distribution of the 12 most abundant bacterial genera across apricot origins. Stacked bar plots represent the mean relative abundance of each genus within each apricot origin, with genera ordered in the legend according to their overall abundance from highest to lowest. The category *"Others"* includes all remaining low-abundance taxa not shown individually.

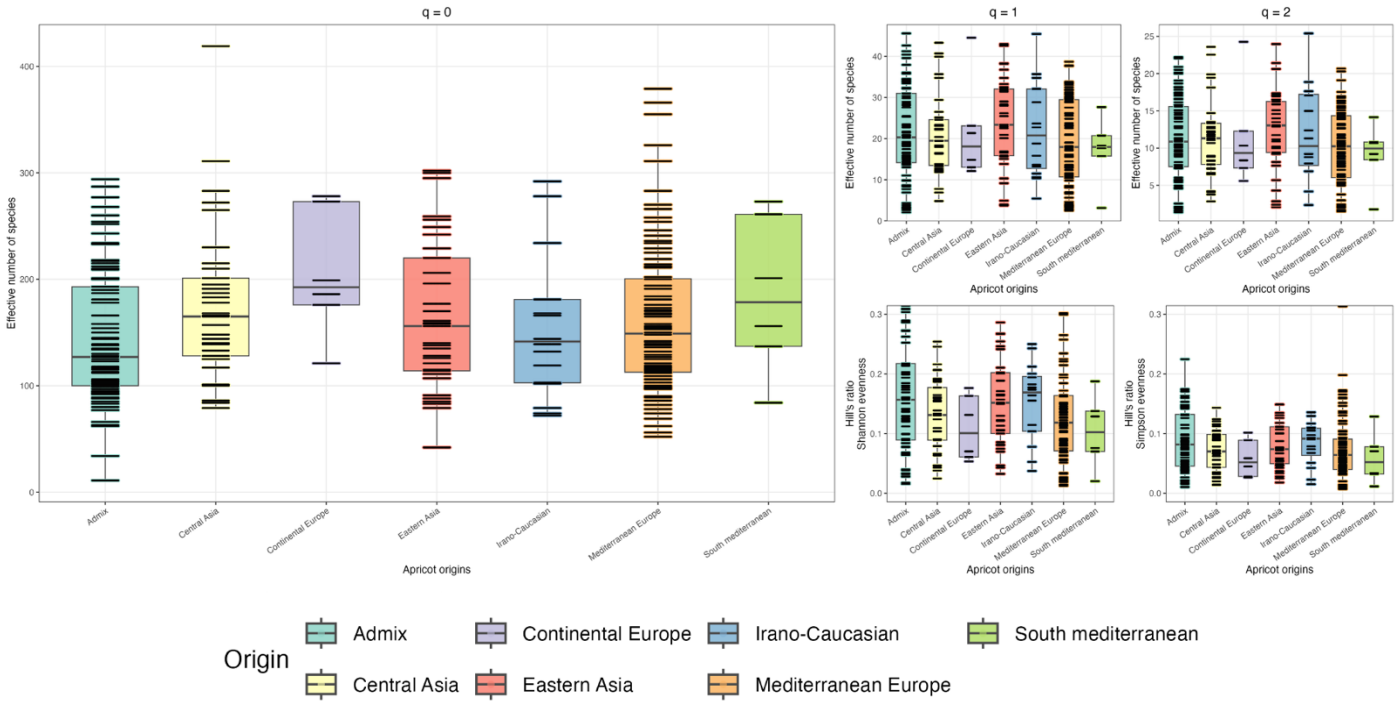
3. Alpha diversity and host genetic background

Alpha diversity of both communities was estimated using Hill numbers of order q=0 (richness), q=1 (Shannon effective number of species), and q=2 (Simpson effective number of species), together Hill's ratio evenness indices[52], in which groups were compared across apricot origins.

Fungi. No clear differentiation in fungal alpha diversity was observed across origins for any Hill order (Fig. 5A). Median richness (q=0) ranged from ~130 to ~200 effective species, q=1 and q=2 converged around ~15-35 and ~5-15 effective species, respectively. Evenness was consistently low (Hill's ratio median ~0.05-0.18), confirming communities strongly dominated by a few taxa, consistent with the dominance of *Didymella* and *Aureobasidium* in the taxonomic data.

Bacteria. In contrast, bacterial richness varied more across origins, particularly at q = 0. Median richness ranged from ~25–30 (Admix, Eastern Asia) to ~40–50 (Central Asia, South Mediterranean; Continental Europe also high), with South Mediterranean showing the widest spread (up to ~80 effective species) (Fig. 5B). At q = 1 and q = 2 the differences attenuated (~8–18 and ~5–12), indicating that the dominant bacterial taxa are relatively consistent across origins even when total richness differs. Evenness was low-to-moderate (Shannon ~0.25–0.35, Simpson ~0.15–0.25), higher than in the fungal community, bacteria are dominated by few taxa, but less sharply than the mycobiome.

A. Fungi



B. Bacteria

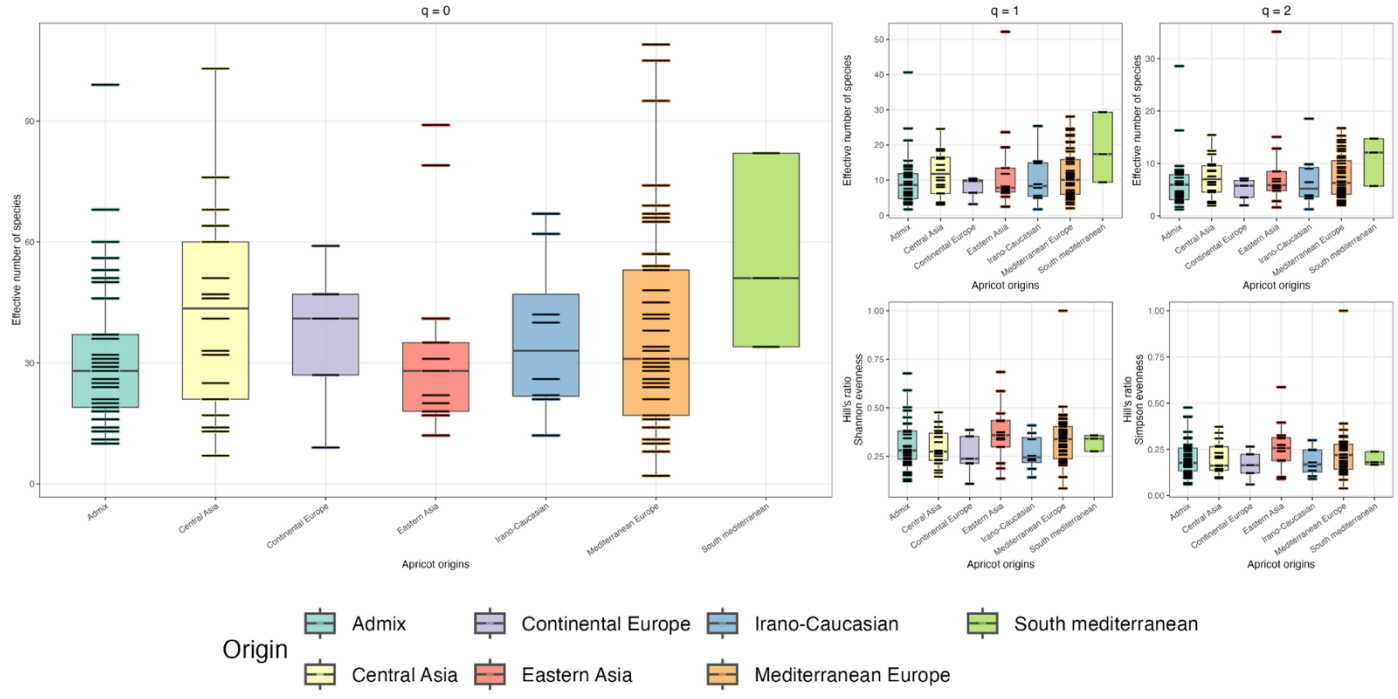


Figure 5: Alpha diversity of apricot leaf fungal (A) and bacterial (B) communities across geographic origins, estimated using Hill numbers. Diversity was expressed as the effective number of species for q=0, q=1, and q=2, corresponding to taxon richness, Shannon diversity, and Simpson diversity, respectively. Boxplots show the distribution of alpha diversity values across apricot origins, with individual samples overlaid as points.

Joint interpretation. The two kingdoms behave differently with respect to the host background. Fungal richness is essentially flat across origins, whereas bacterial richness shows a gradient, with South

Mediterranean and Central Asian genotypes, ancestrally divergent backgrounds, tending to support richer bacterial communities. Because all trees share one orchard, such differences are best read as host-genotype effects (e.g. cuticle chemistry, stomatal density, secondary-metabolite profiles) rather than environmental ones, in line with evidence that host leaf traits are among the strongest filters of phyllosphere bacterial colonization[3]. The relatively low bacterial richness of Eastern-Asian genotypes is consistent with their more distinct taxonomic profile.

4. Beta diversity and host genetic background

Beta diversity was assessed by non-metric multidimensional scaling (NMDS) on Wisconsin-standardized abundance data using the Horn–Morisita dissimilarity index, with significance tested by PERMANOVA and multivariate dispersion by HOMOVA.

Fungi. The fungal ordination had a stress of 0.2393, indicating moderate-to-high distortion in two dimensions, so spatial patterns are interpreted cautiously (Fig. 6A). Communities from all origins overlapped substantially with no visual separation. PERMANOVA confirmed that origin did not significantly structure fungal communities (R^2 = 0.0228, P = 0.271). However, HOMOVA was significant (P = 0.027): although mean compositions are not distinct, within-group dispersion differs among origins, with Admix and Continental Europe occupying broader regions of ordination space and Eastern Asia showing tighter dispersion. Notably, this dispersion is not related to the number of accessions per origin.

Bacteria. The bacterial ordination had a comparable stress (0.2627) (Fig. 6B). Origins overlapped broadly — particularly Mediterranean Europe, Continental Europe and Admix — while Eastern Asia and Central Asia were more dispersed. Here, however, PERMANOVA was highly significant: origin explained 4.53 % of total compositional variation (R^2 = 0.0453, P = 1×10^{-4}), while HOMOVA was not significant (P = 0.275). The groups therefore differ in their average community structure, not in their internal variability. The modest R^2 is typical of field-collected phyllosphere data, where environmental stochasticity, local inoculum and individual leaf variation dominate the variance.

Joint interpretation. The contrast between kingdoms is the central beta-diversity result: host origin leaves a significant compositional signature on bacteria but not on fungi. This suggests bacteria are more responsive to host genotype-level filtering than fungi in the apricot phyllosphere, consistent with stronger bacterial phylosymbiosis with host genetic distance in other plant systems[53]. A likely mechanism is the tighter coupling between leaf-associated bacteria and genotype-specific host immune signalling (pattern-triggered immunity, receptor-mediated recognition)[48,50,51], whereas fungal colonization may be governed more by leaf-surface physical properties and airborne dispersal that override host-specific filtering, the same broad-amplitude, cosmopolitan taxa (*Didymella, Aureobasidium*) that homogenize fungal communities across hosts[44].

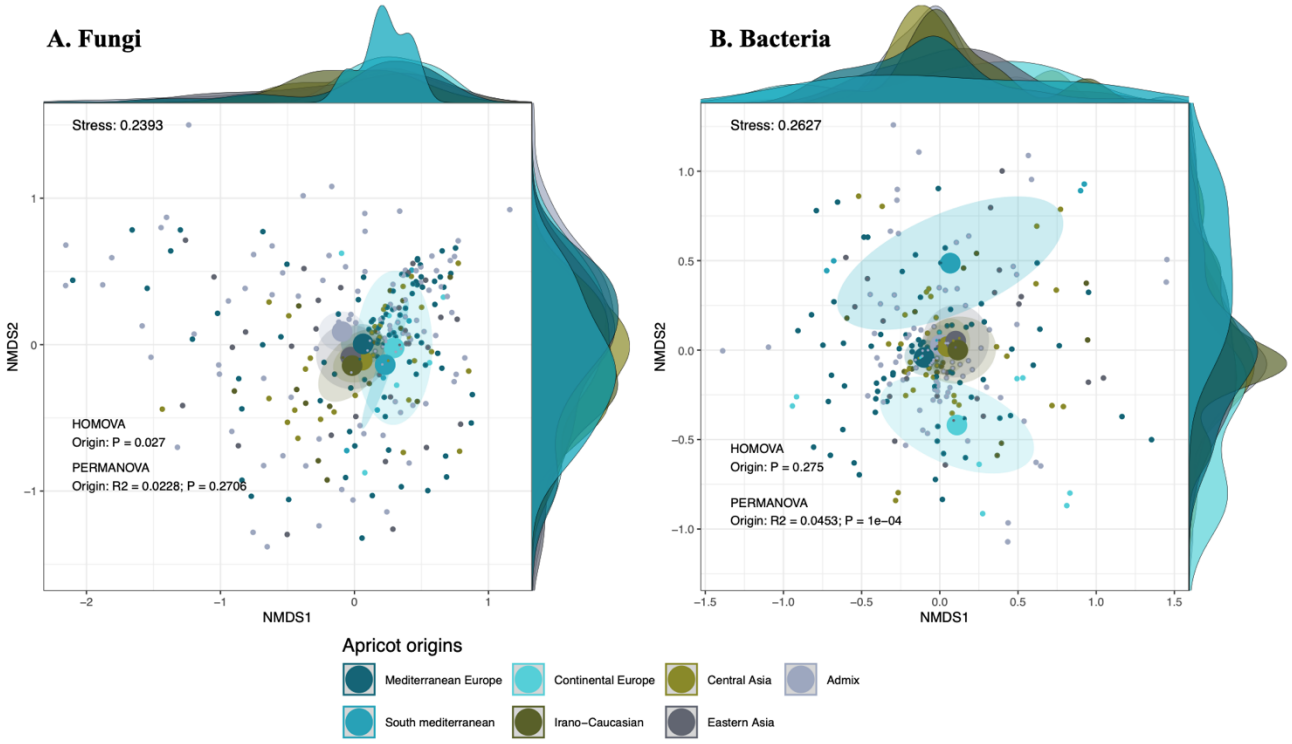


Figure 6: Non-metric multidimensional scaling (NMDS) ordination of apricot leaf fungal (A) and bacterial (B) communities across geographic origins. Community dissimilarities were calculated using Horn–Morisita distances after Wisconsin double standardization of abundance data. Points represent individual samples, and shaded ellipses indicate the clustering of samples by apricot origin. Marginal density plots along the top and right axes show the distribution of samples within each origin. Community differences among origins were evaluated using PERMANOVA, and homogeneity of multivariate dispersions was assessed using HOMOVA.

5. Disease-associated variation in microbial richness

To explore relation between apricot health and microbiota, variance partitioning was performed, revealing that the three major diseases scored were differently associated with bacterial and fungal richness in the apricot phyllosphere. For bacterial richness, coryneum disease explained a significant proportion of the variation, indicating that samples with higher coryneum severity tended to harbor distinct bacterial richness patterns (Fig 7A). For fungal richness, rust showed a weak but statistically significant contribution, suggesting a more modest but detectable link between rust severity and fungal community richness (Fig 7B). In contrast, monilia did not appear to explain a meaningful proportion of richness variation in either kingdom, implying that its effect on phyllosphere diversity was comparatively limited in this dataset.

This pattern is ecologically interesting because the associations do not imply a strictly one-way effect of disease on microbiota. In a holobiont framework, host genotype and the resident phyllosphere community can also influence disease development by shaping microbial assembly, recruiting potentially protective taxa, and contributing to colonization resistance or immune priming[54]. Therefore, the observed patterns most likely reflect reciprocal feedback between apricot health and microbiome structure[55], rather than disease acting independently of the host-microbe system.

Ecologically, this means that coryneum may be more closely linked to bacterial community restructuring, potentially through changes in leaf tissue condition, exudation, or microhabitat availability, whereas rust may be more associated with fungal richness through stronger fungal–fungal or host-defense-mediated interactions. However, because the explained variance was small, these

effects likely represent subtle signatures within a much larger background of environmental and host-driven variation.

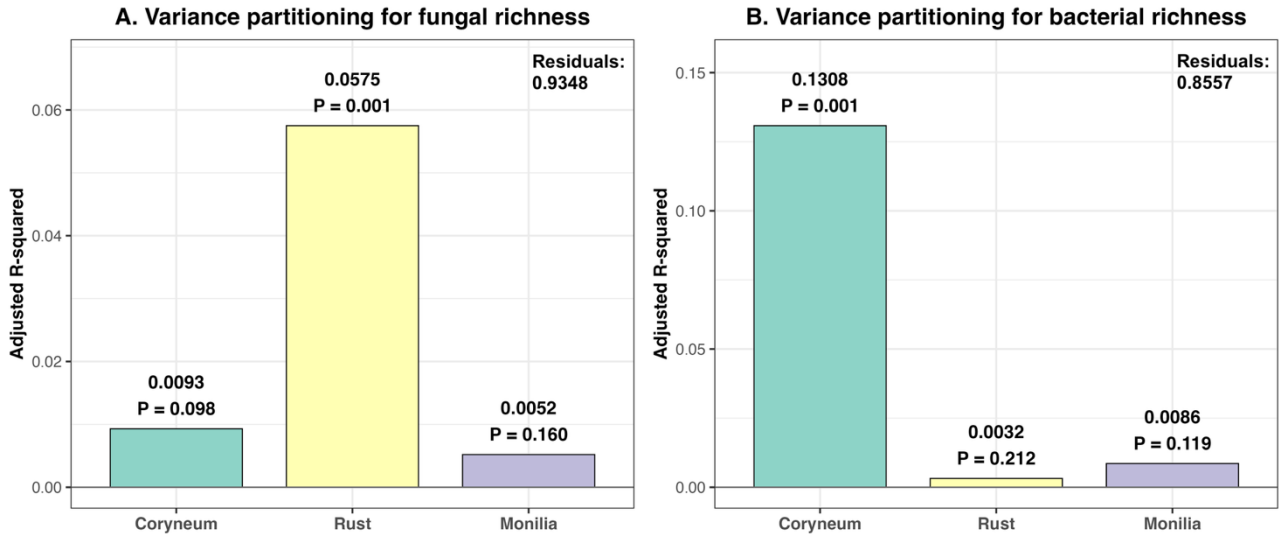


Figure 7. Relative contribution of apricot diseases (*Monilia*, *Rust*, *Coryneum*) on leaf fungal (A) and bacterial community richness (B) estimated by variance partitioning analyses. Residuals indicate the percentage of variance unexplained by the environmental parameters, estimated by variance partitioning analyses.

D. Conclusions and Perspectives

Conclusions. This work gives a first integrated picture of the apricot leaf microbiome and already tells a coherent story. The fungal side is remarkably stable: a small *Didymella–Aureobasidium* core dominates every genetic background, and neither richness nor composition tracks host origin. The bacterial side behaves differently in which origin leaves a small but reproducible signature on community structure and a gradient in richness that points to host genotype filtering bacteria more tightly than fungi, plausibly through immune and leaf-surface traits that act more strongly on bacterial colonizers. On the disease side, the signal is kingdom-specific rather than diffuse: bacterial richness is associated mainly with coryneum and fungal richness with rust, while monilia leaves little trace. Importantly, because all trees were grown in one orchard, these are best read as host effects, and the disease association should be treated as a two-way, holobiont-level relationship rather than a one-way effect of disease on the microbiome.

Perspectives. The most immediate gain will come from scale and resolution. Only a subset of the collection was sequenced here; extending the analysis to the full panel, and to the second sampling year with a third on the way will let us separate the genuinely stable core from year-to-year and sampling noise, and test whether the bacterial origin signal holds when statistical power is no longer limiting.

The second priority is to replace "origin" with genetic distance or GWAS analysis. Geographic origin is only a coarse stand-in for relatedness, and we now have kinship data giving the genetic proximity between genotypes**.** Using that kinship matrix directly, that is, rather than seven discrete groups, turns the phylosymbiosis question into a quantitative one: does microbiome dissimilarity increase with host genetic distance? Is that relationship stronger for bacteria than for fungi, as the present results hint? Are microbiota taxa associated with loci variants?

Third, the expanded set of phenotypic traits now available should move the disease work beyond variance partitioning of richness. Modelling specific taxa or community axes against individual traits (vigour, age, each disease score) with appropriate regression or GLM frameworks would let us ask which microbial features track health and begin to identify candidate protective or pathobiont taxa like *Didymella,* being the obvious one to follow up given its dominance and known pathogenic potential.

Finally, all of the disease associations here are correlative. Confirming the direction of whether a particular community protects the tree or simply reflects its disease state, thus will eventually need longitudinal sampling across the multi-year dataset and, ideally, controlled inoculation or synthetic-community experiments. Together, these steps would take the project from a descriptive survey toward a mechanistic, genetics-aware view of how the apricot holobiont contributes to tree health.

E. Supplementary figures

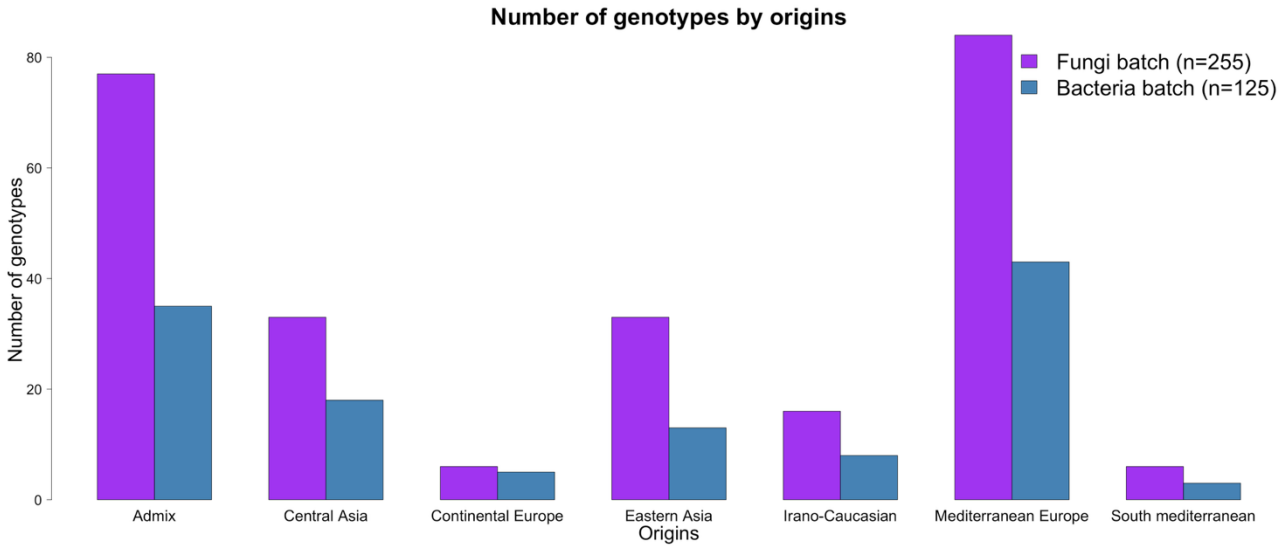


Figure A1. Distribution of genotypes included in the bacterial and fungal community analyses according to the geographic origins of apricot accessions. Purple bars represent genotypes from the fungal dataset (n = 255), and blue bars represent genotypes from the bacterial dataset (n = 125).

References

1. Li, P.-D. *et al.* The phyllosphere microbiome shifts toward combating melanose pathogen. *Microbiome* **10**, 56 (2022).
2. F, M., S, H. & Bp, T. Co-evolution within the plant holobiont drives host performance. *PubMed* https://pubmed.ncbi.nlm.nih.gov/37471099/ (2023).
3. Singh, P. K. *et al.* Leaf microbiome assembly is linked to plant phylogeny. *Plant Soil* **520**, 1177–1192 (2026).
4. Liu, H., Brettell, L. E. & Singh, B. Linking the Phyllosphere Microbiome to Plant Health. *Trends Plant Sci.* **25**, 841–844 (2020).
5. Beilsmith, K. *et al.* Genome-wide association studies on the phyllosphere microbiome: Embracing complexity in host–microbe interactions. *Plant J.* **97**, 164–181 (2019).
6. Vacher, C. *et al.* The Phyllosphere: Microbial Jungle at the Plant–Climate Interface. *Annu. Rev. Ecol. Evol. Syst.* **47**, 1–24 (2016).
7. Ja, V. Microbial life in the phyllosphere. *PubMed* https://pubmed.ncbi.nlm.nih.gov/23154261/.
8. Khalil, H., Tedersoo, L. & Hagh-Doust, N. Ecological roles of leaf-inhabiting fungi and their host-associated predictors: a review. *Fungal Biol. Rev.* **55**, 100475 (2026).
9. Vorholt, J. A. Microbial life in the phyllosphere. *Nat. Rev. Microbiol.* **10**, 828–840 (2012).
10. Schimann, H. *et al.* Determinants of the Vertical Distribution of the Phyllosphere Differ Between Microbial Groups and the Epi- and Endosphere in a Tropical Forest. *Phytobiomes J.* **7**, 312–323 (2023).
11. Li, M., Hong, L., Ye, W., Wang, Z. & Shen, H. Phyllosphere bacterial and fungal communities vary with host species identity, plant traits and seasonality in a subtropical forest. *Environ. Microbiome* **17**, 29 (2022).
12. Cadavid, I. C., Capelari, É. F., Salvati, C. & Margis, R. Phyllosphere microbiomes uncovered: Research trends, geographic disparities, and key microbial players. *Genet. Mol. Biol.* **49**, e20250083.
13. Noble, A. S., Abbaszadeh, J. & Lee, C. K. Host selection is not a universal driver of phyllosphere community assembly among ecologically similar native New Zealand plant species. *Microbiome* **13**, 35 (2025).
14. Yang, C.-X. *et al.* Plant exudates-driven microbiome recruitment and assembly facilitates plant health management. *FEMS Microbiol. Rev.* **49**, fuaf008 (2025).
15. X, M. *et al.* Chemical Communication Between Plant and Microbe in the Phyllosphere. *PubMed* https://pubmed.ncbi.nlm.nih.gov/41327627/.
16. De Mandal, S. & Jeon, J. Phyllosphere Microbiome in Plant Health and Disease. *Plants* **12**, 3481 (2023).
17. Shen, Y. *et al.* Gut Microbiota Dysbiosis: Pathogenesis, Diseases, Prevention, and Therapy. *MedComm* **6**, e70168 (2025).
18. Liu, Y. *et al.* Nonpathogenic Pseudomonas syringae derivatives and its metabolites trigger the plant “cry for help” response to assemble disease suppressing and growth promoting rhizomicrobiome. *Nat. Commun.* **15**, 1907 (2024).
19. Ginnan, N. A. *et al.* Disease-Induced Microbial Shifts in Citrus Indicate Microbiome-Derived Responses to Huanglongbing Across the Disease Severity Spectrum. *Phytobiomes J.* **4**, 375–387 (2020).
20. Li, W. *et al.* Genetic diversity, population structure, and relationships of apricot (Prunus) based on restriction site-associated DNA sequencing. *Hortic. Res.* **7**, 69 (2020).

21. Muhammad, M. *et al.* Geostatistical Analysis of Apricot Shot Hole Disease and Influence Factors in District Nagar, Gilgit-Baltistan, Pakistan. *Int. J. Phytopathol.* **11**, 227–238 (2022).
22. Soylu, S., Oguz, M., Kurt, Ş., Soylu, E. & Uysal, A. Determination of fungal and bacterial disease agents of apricot trees growing in Hatay province. *Acta Hortic.* 111–114 (2020) doi:10.17660/ActaHortic.2020.1290.20.
23. Bourguiba, H. *et al.* Exploration of bacterial diversity in leaves and rhizosphere soil of flood affected and unaffected apricot trees. *Biologia (Bratisl.)* **78**, 217–227 (2022).
24. Santos, R. *et al.* A protocol for large scale genomic DNA isolation for cacao genetics analysis. *Afr. J. Biotechnol.* **13**, 814–820 (2014).
25. Op De Beeck, M. *et al.* Comparison and Validation of Some ITS Primer Pairs Useful for Fungal Metabarcoding Studies. *PLoS ONE* **9**, e97629 (2014).
26. Walters, W. *et al.* Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* **1**, e00009-15 (2015).
27. Hussain, U. *et al. Peptide Nucleic Acid (PNA) Clamps Reduce Amplification of Host Chloroplast and Mitochondria rRNA Gene Sequences and Increase Detected Diversity in 16S rRNA Gene Profiling Analysis of Oak-Associated Microbiota*. (2024). doi:10.21203/rs.3.rs-4991502/v1.
28. Mahé, F. frederic-mahe/nf-metabarcoding. (2026).
29. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
30. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
31. Mahé, F. *et al.* Swarm v3: towards tera-scale amplicon clustering. *Bioinformatics* **38**, 267–269 (2021).
32. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
33. Abarenkov, K. *et al.* The UNITE database for molecular identification and taxonomic communication of fungi and other eukaryotes: sequences, taxa and classifications reconsidered. *Nucleic Acids Res.* **52**, D791–D797 (2024).
34. Yilmaz, P. *et al.* The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* **42**, D643–D648 (2014).
35. Wickham, H. *et al.* Welcome to the Tidyverse. https://joss.theoj.org/papers/10.21105/joss.01686.
36. Wickham, H. *et al.* ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. (2026).
37. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217 (2013).
38. Oksanen, J. *et al.* Vegan: Community Ecology Package. *R Package Version 22-1* **2**, 1–2 (2015).
39. Alberdi, A. & Gilbert, M. *Hilldiv: An R Package for the Integral Analysis of Diversity Based on Hill Numbers*. (2019). doi:10.1101/545665.
40. Pipe-Friendly Framework for Basic Statistical Tests • rstatix. https://rpkgs.datanovia.com/rstatix/.
41. Moustafa, K. & Cross, J. Production, pomological and nutraceutical properties of apricot. *J. Food Sci. Technol.* **56**, 12–23 (2019).

42. Di Francesco, A., Zajc, J. & Stenberg, J. A. Aureobasidium spp.: Diversity, Versatility, and Agricultural Utility. *Horticulturae* **9**, 59 (2023).

43. Hou, L. W. *et al.* The phoma-like dilemma. *Stud. Mycol.* **96**, 309–396 (2020).

44. Luo, X. *et al.* Morphological and Phylogenetic Analyses Reveal Three New Species of Didymella (Didymellaceae, Pleosporales) from Jiangxi, China. *J. Fungi* **10**, 75 (2024).

45. Bozoudi, D. & Tsaltas, D. The Multiple and Versatile Roles of Aureobasidium pullulans in the Vitivinicultural Sector. *Fermentation* **4**, 85 (2018).

46. Janakiev, T. *et al.* Phyllosphere Fungal Communities of Plum and Antifungal Activity of Indigenous Phenazine-Producing Pseudomonas synxantha Against Monilinia laxa. *Front. Microbiol.* **10**, 2287 (2019).

47. Henderson, G. *et al.* Improved taxonomic assignment of rumen bacterial 16S rRNA sequences using a revised SILVA taxonomic framework. https://pmc.ncbi.nlm.nih.gov/articles/PMC6407505/.

48. Lidstrom, M. E. & Chistoserdova, L. Plants in the Pink: Cytokinin Production by Methylobacterium. *J. Bacteriol.* **184**, 1818–1818 (2002).

49. S, L., B, D., L, P., F, D. & A, L. Humic substances increase tomato tolerance to osmotic stress while modulating vertically transmitted endophytic bacterial communities. *PubMed* https://pubmed.ncbi.nlm.nih.gov/39628527/ (2024).

50. Fessia, A., Barra, P., Barros, G. & Nesci, A. Could Bacillus biofilms enhance the effectivity of biocontrol strategies in the phyllosphere? *J. Appl. Microbiol.* **133**, 2148–2166 (2022).

51. Huang, Q. *et al.* Low-nitrogen input enriches Massilia bacteria in the phyllosphere to improve blast resistance in rice. *New Phytol.* **248**, 3151–3167 (2025).

52. Alberdi, A. & Gilbert, M. T. P. hilldiv: an R package for the integral analysis of diversity based on Hill numbers. https://doi.org/10.1101/545665 (2019) doi:10.1101/545665.

53. C, R., Mfa, L., A, P. & Ee, K. Rhizosphere microbiome response to host genetic variability: a trade-off between bacterial and fungal community assembly. *PubMed* https://pubmed.ncbi.nlm.nih.gov/35595468/ (2022).

54. Wang, Z., Yin, J. & Tsuda, K. Harnessing Aspergillus and host M genes for sustainable phyllosphere microbiome engineering. *Crop Health* **3**, 6 (2025).

55. Cesaro, P., Gamalero, E., Zhang, J. & Pivato, B. Frontiers | Editorial: The Plant Holobiont Volume I: Microbiota as Part of the Holobiont; Challenges for Agriculture. https://doi.org/10.3389/fpls.2021.799168 doi:10.3389/fpls.2021.799168.

56. Haddad, R. Rapid and Efficient Isolation of High Quality Nucleic Acids from Plant Tissues Rich in Polyphenols and Polysaccharides. https://www.academia.edu/61671867/Rapid_and_Efficient_Isolation_of_High_Quality_Nucleic_Acids_from_Plant_Tissues_Rich_in_Polyphenols_and_Polysaccharides (2025).