

RESEARCH

Open Access



Grapevine phyllosphere pan-metagenomics reveals pan-microbiome structure, diversity, and functional roles in downy mildew resistance

Jingyun Jin¹, Xiangfeng Wang¹, Xuenan Zhang¹, Junjie Mei¹, Wei Zheng¹, Linling Guo¹, Haisheng Sun¹, Lili Zhang¹, Chonghuai Liu², Wenxiu Ye^{1*} and Li Guo^{1*}

Abstract

Background Grapevines are among the most economically important fruit crops, and the microbiome profoundly influences their health, yield, and quality. However, mechanistic insights into microbiome-orchestrated grapevine biology remain limited.

Results Here, we conduct large-scale pan-metagenomic and pan-metatrascriptomic analyses of the phyllosphere microbiome from 107 grapevine accessions spanning 34 *Vitis* species. We show that the grapevine core microbiome is dominated by phyla Bacillota and Pseudomonadota. Leveraging PacBio sequencing, we assembled 19 high-quality metagenome-assembled genomes (MAGs) from the grapevine pan-microbiome, representing the first MAG reconstruction in plant-associated microbial communities using PacBio reads. These MAGs encode genes associated with antibiotic resistance, secondary metabolism, and carbohydrate-active enzymes (CAZymes), which could potentially influence grapevine biology. During downy mildew (DM) infection, DM-resistant grapevines exhibit significantly higher microbial network complexity than susceptible counterparts. Among the key taxa contributing to this complexity, Bacillota emerged as the dominant phylum, displaying strong abundance correlations with phylum Euglenozoa and Cyanobacteriota, and an isolated Bacillota species from the grapevine leaves, *Bacillus cereus*, demonstrated potent biocontrol activity against DM infection. Pan-metatrascriptomic analysis further revealed significant upregulation of eukaryotic microbial genes involved in primary and secondary metabolism.

Conclusions This pan-metagenomic study offers unprecedented insights into the complex structure, diversity, and functional roles of the grapevine phyllosphere microbiome and presents valuable genomic and microbial resources for microbiome research and engineering to enhance viticulture productivity and quality.

Keywords Grapevine, Microbiome, Pan-metagenome, Pan-metatrascriptome, Downy mildew

*Correspondence:

Wenxiu Ye
wenxiu.ye@pku-iaas.edu.cn
Li Guo
li.guo@pku-iaas.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Grapevines are perennial plants belonging to the Vitaceae family, renowned for their exceptional flavor as well as their significant economic and nutritional value. In 2023, global grape production reached 72 million tons, with China, Italy, and France leading as the top producers, and the global production value of grapes was estimated at 92 billion USD [1]. In addition, grapes are rich in bioactive compounds, including polyphenols, phenolic acids, vitamins, and dietary fiber, which provide a variety of health benefits, such as anticancer, antidiabetic, cardioprotective, and gut-microbiome-regulating properties [2]. Despite remarkable advancements in grape breeding and cultivation in recent decades, grape production worldwide remains significantly threatened by diseases such as downy mildew (DM), powdery mildew, and gray mold, causing substantial economic losses across grape production and related industries each year [3–7]. Especially, DM is the most destructive among the diseases caused by the oomycete *Plasmopara viticola*, severely affecting grape production and quality by causing damage to leaves and fruit clusters [8]. With increasing evidence highlighting the involvement of the grapevine microbiome in DM infection [9, 10], it is crucial to investigate the grapevine-associated microbiome, which could serve as a foundation for developing environmentally friendly and sustainable microbiome-based biocontrol methods in viticulture.

Microbiome, often regarded as the second genome of plants, plays crucial roles in nutrient absorption, disease resistance, and overall plant health [4]. The microbiome profoundly affects grape productivity and the quality of grape-derived products such as wine and raisins [11–13]. The grapevine microbiomes primarily consist of bacteria and fungi, with Acidobacteria, Actinomycetota, Bacillota, Bacteroidota, Chloroflexi, and Pseudomonadota being the most prevalent phyla, based on previous 16S rRNA amplicon sequencing [4, 9, 14–16]. The microbiome composition varies across grapevine tissues, with the highest microbial diversity observed in roots and gradually decreasing for trunks and leaves [4]. Studies show that grapevine genetic backgrounds and geographic origins are primary factors influencing their microbial compositions [11, 14, 17–24]. Despite the advancement, current plant microbiome studies primarily rely on 16S rRNA sequencing, which yields limited insights into the mechanistic understanding of microbiome functions compared to Illumina sequencing. However, Illumina sequencing falls short in assembling the complete microbial genomes, given the limitation of short reads, impeding accurate annotations and functional analysis of the metagenome. So far, long-read technologies mainly have been used in microbiome studies for human-related

samples [25, 26], but remain underutilized in plant and agricultural metagenomics. In addition, metatranscriptomics, which is crucial to understanding the functional implications of the microbiome by revealing microbiome gene expression changes, has not been widely adopted in plant microbiome research. Most importantly, integrating pan-metagenomics and pan-metatranscriptomics of plants to study the diversity of microbial profiles across plant species and cultivars has not been reported. These limitations hinder the exploration of the grapevine microbiome and the application of microbial biocontrol in viticulture.

Here, we conducted a comprehensive pan-metagenomic analysis of grapevine leaves from 107 *Vitis* accessions (each accession corresponds to a unique *Vitis* cultivar, except for a few samples) spanning 34 species, utilizing Illumina, PacBio, RNA-seq, and DM resistance phenome data. Through systematic profiling of the grapevine phyllosphere microbiome, we identified dominant microbial communities under a pan-metagenomic framework. Leveraging PacBio metagenomic reads from the grapevine phyllosphere, we successfully generated 19 high-quality grapevine phyllospheric metagenome-assembled genomes (GPMAGs) and comprehensively annotated genes associated with antibiotic resistance, secondary metabolism, and CAZymes. Based on the phenome data, we revealed that the microbiome from DM-resistant grapevines exhibits significantly more complex microbial networks than susceptible ones. Additionally, we identified key genera potentially conferring resistance to DM infection, with Bacillota emerging as the dominant phylum exhibiting high abundance correlation with Euglenozoa and Cyanobacteriota, and experimentally validated the potent biocontrol activity of two *Bacillus cereus* strains against DM. Pan-metatranscriptomic analysis further highlighted significant changes in the expression of eukaryotic primary and secondary metabolism genes in response to DM infection. These findings provide novel insights into the grapevine microbiome and valuable genomic and microbial resources for viticulture, laying a solid foundation for developing microbiome-based strategies to enhance grape health and promote sustainable production.

Results

Grapevine core microbiome is dominated by Bacillota and Pseudomonadota

To investigate the microbiome profile of the grapevine phyllosphere across grapevines, we collected a total of 107 leaf samples from 34 *Vitis* species covering 100 grapevine cultivars originated from various geographical locations with distinct genotypes and DM resistances [27] (Fig. 1a, b and Table S1) and conducted

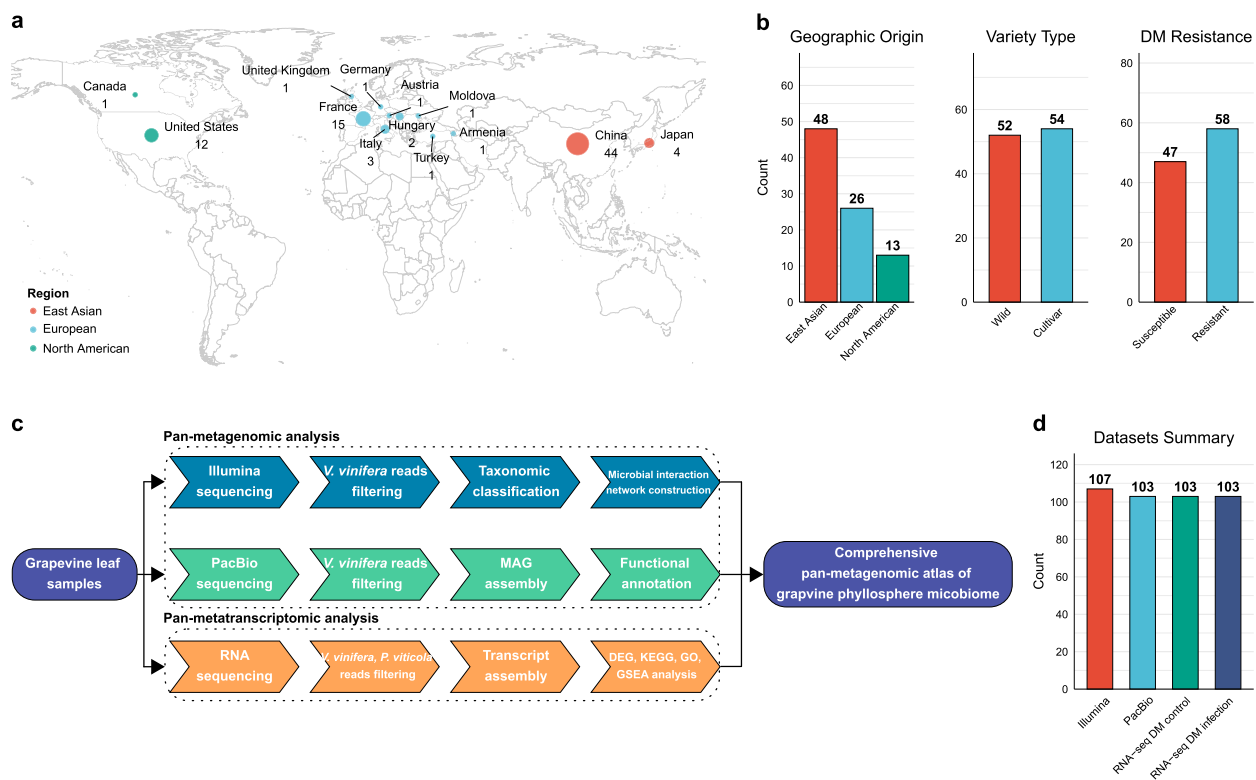


Fig. 1 Research design for grapevine phyllosphere pan-metagenomics and pan-metatrascriptomics. **a** Geographical origins of the sampled grapevine accessions. Numbers of samples were numerically labeled. **b** Classification statistics of grapevine accessions by geographical origin, cultivar type, and DM resistance status. **c** Pan-metagenomic and pan-metatrascriptomic analysis workflow. Colored blocks indicate sequencing methods: blue (Illumina), green (PacBio), orange (RNA-seq), and purple (shared components). **d** Summary of generated datasets

pan-metagenomic analysis as outlined in the workflow (Fig. 1c). The metagenomic DNA from the samples was sequenced using Illumina paired-end sequencing [27], yielding a total of 107 Illumina shotgun metagenomic datasets (Fig. 1d). The reads belonging to the grapevines were completely filtered out by mapping against a complete grapevine genome [27] (VHP-T2T.hap1, NCBI accession: GCA_046255155.1) (Table S2), with on average 4.4 million microbial reads remaining per sample. Taxonomic classification of these microbial reads using Kraken2 [28] identified a total of 1326 genera across the grapevine phyllosphere (Fig. 2a and Table S3). Among these, Bacillota (43.6%) and Pseudomonadota (37.9%) dominated the microbiome profile at the phylum level, while *Lactiplantibacillus* (42.1%) and *Xanthomonas* (22.2%) were the two most abundant genera (Fig. 2b). Furthermore, kingdom-level profiling demonstrated that bacteria dominated the grapevine phyllosphere microbiome (91.3%), and other kingdoms accounted for less than 10% (Fig. 2c). Notably, the non-bacterial kingdoms exhibited dominating phyla, such as Euryarchaeota dominating Archaea, Ascomycota and Euglenozoa dominating Eukaryota, and Cressdnariviricota dominating viruses,

reflecting specific phylum-level dominance for the grapevine microbiome. To examine the worldwide representativeness of our microbial composition, we analyzed worldwide representative accessions from published Illumina data containing a global grapevine phyllosphere collection [29] using the same bioinformatic pipeline. The phylum-level microbial composition of the global microbiome was generally consistent with our Illumina-based composition (Fig. S1), demonstrating the worldwide representativeness of our microbial composition. These results highlight the prevalence of Bacillota and Pseudomonadota and their potential roles in the grapevine microbiome.

Although several previous studies have profiled grapevine microbiomes [11, 14, 17–24], mostly using 16S rDNA amplicon sequencing, a comprehensive pan-metagenomic analysis leveraging large-scale Illumina datasets had not been conducted prior to this study, which could provide critical insights into the structure and functional roles of the grapevine microbiome. Through our pan-metagenomic analysis of 107 grapevine phyllosphere, we identified seven core genera (present in all accessions), 945 dispensable genera (present in some

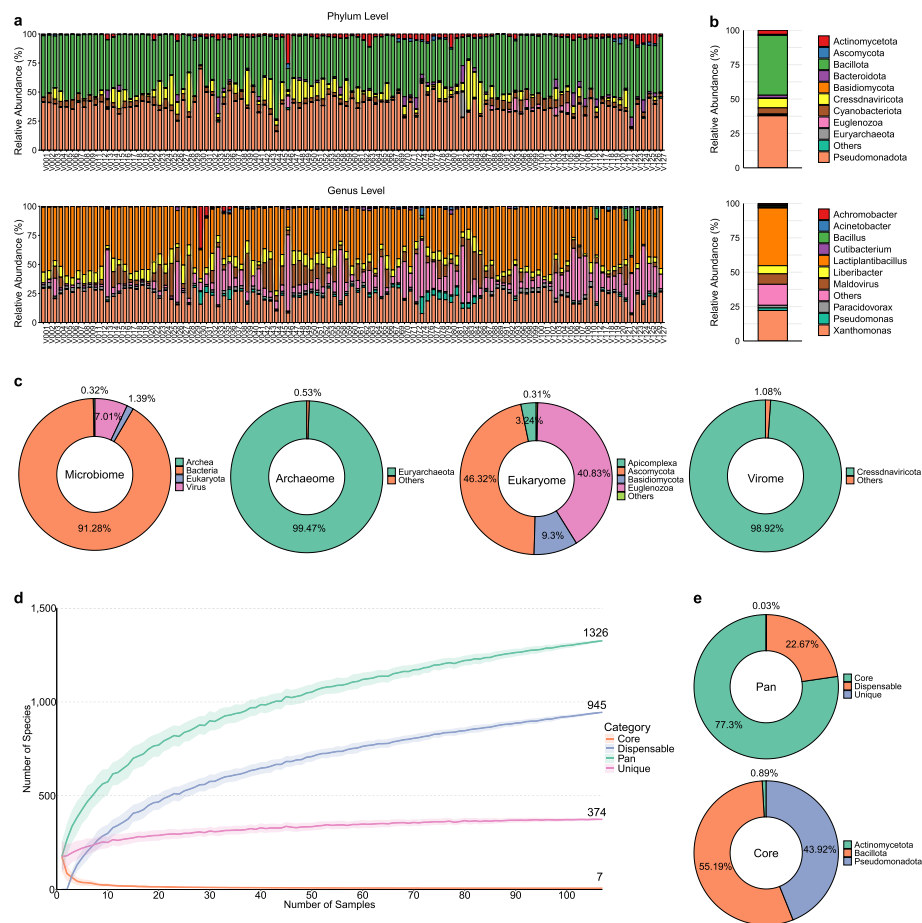


Fig. 2 Pan-metagenomic landscape of the grapevine phyllospheric microbiome. **a** Relative abundances of the 10 most abundant phyla (top) and genera (bottom) across the grapevine accessions. **b** Mean relative abundance of the phyla and genera shown in **a**. **c** Compositional profiles of kingdoms from the grapevine accessions. **d** Pan-metagenomic accumulation curve demonstrating pan, core, dispensable and unique genera across the grapevine accessions. **e** Pan-metagenomic relative abundance at the genus level within the entire pan-metagenome (top) and the top 10 most abundant phyla within the core microbiome (bottom)

but not all accessions), and 374 unique genera (exclusive to a single accession) based on presence and absence patterns (Fig. 2d and Table S3). Strikingly, although the core microbiome accounted for only 0.6% of the detected genera, it dominated 77.3% of the relative abundance, with Bacillota (55.2%) and Pseudomonadota (44.0%) as the major phyla in the core microbiome (Fig. 2e). This result further underscores the central role of the core microbiome (mainly Bacillota and Pseudomonadota) in shaping the grapevine phyllosphere microbiome.

Assembly of grapevine phyllospheric metagenome-assembled genomes (GPMAGs)

To assemble high-quality GPMAGs that provide further insights into the potential functions of the grapevine phyllosphere microbiome that potentially influence the host's biology, we generated a total of 103 PacBio datasets with the PacBio Sequel IIe system [27] (Fig. 1d) for

the grapevine accessions used in the Illumina-based metagenomic analysis above. The PacBio reads were mapped against the grapevine reference genome to completely exclude host-derived sequences, retaining unmapped reads representing microbial data (Table S2). These unmapped reads from all accessions were pooled together and then assembled into contigs by using various assemblers, such as Flye [30], hifiasm [31], hifiasm-meta [32], metaFlye [33], metaMDBG [34], and Verkko [35]. The contigs were subsequently binned using CONCOCT [36], Maxbin 2.0 [37], and MetaBAT 2 [38], followed by dereplication with Galah (<https://github.com/wwood/galah>). This workflow generated a total of 19 non-redundant GPMAGs (Fig. 3a and Table S4) that met the medium-quality threshold based on standard MAG quality assessment criteria [34]. Among these, four were classified as complete, three as near-complete, five as high-quality, and seven as medium-quality MAGs

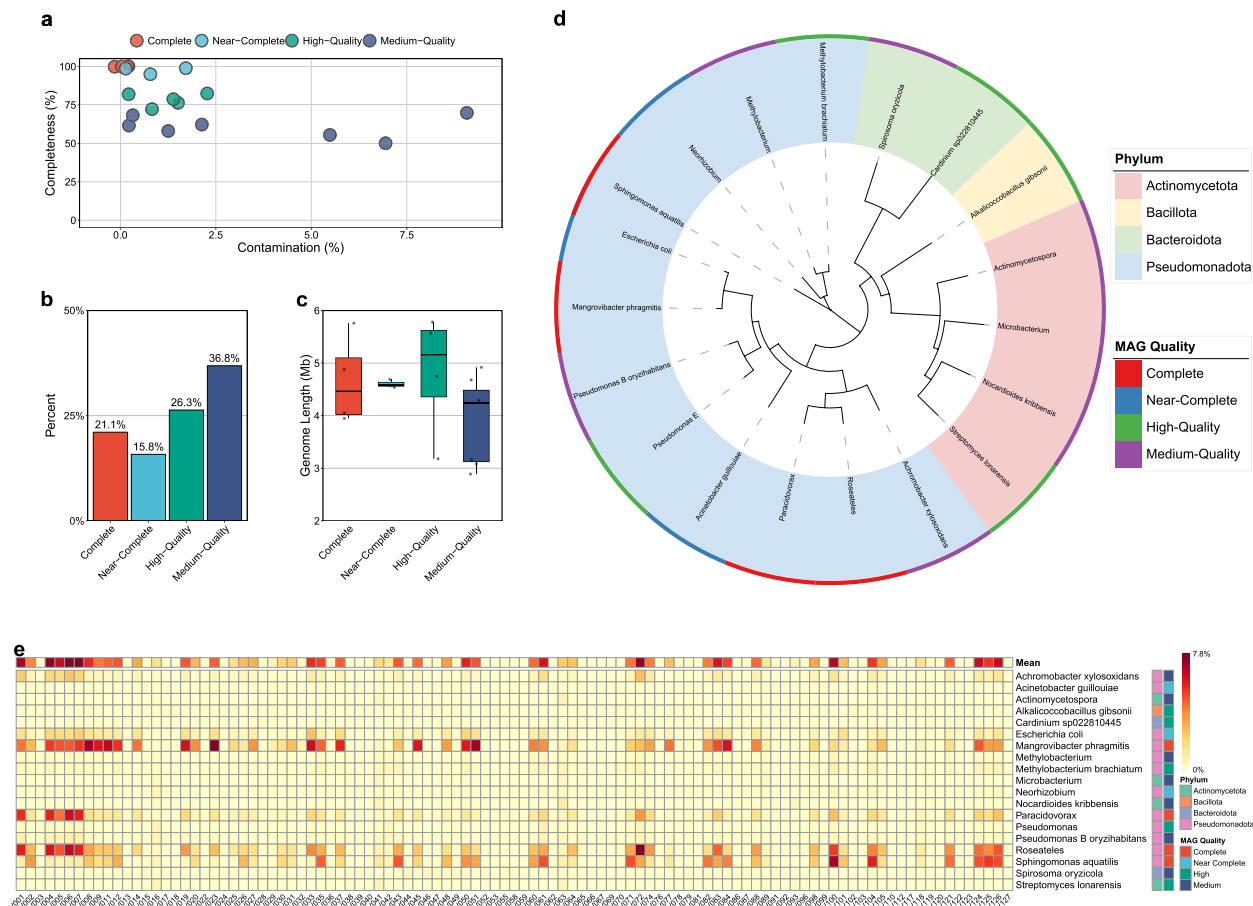


Fig. 3 Assembly and genomic profile of the grapevine phyllospheric metagenome-assembled genomes (GPMAGs). **a** Quality distribution of the GPMAGs in terms of MAG completeness and contamination. **b** Composition of the GPMAG quality categories. **c** Genomic size distribution of the GPMAGs. **d** Phylogenetic tree of the classified GPMAGs, illustrating corresponding phyla and MAG qualities. **e** Pan-genomic prevalence of the GPMAGs across the grapevine accessions. Heatmap depicts the relative abundance of the GPMAGs relative to each grapevine PacBio data accession

(Fig. 3b and Table S4), highlighting the high quality of the GPMAGs. The median genome size of the GPMAGs was 4.5 Mb for the complete, 4.6 Mb for the near-complete, 4.7 Mb for the high-quality, and 4.2 Mb for the medium-quality (Fig. 3c and Table S4), approximating typical bacterial genome sizes. Subsequently, GTDB-tk [39] was employed to identify the microbial profile of the GPMAGs. Among the classified GPMAGs, 63.2% belonged to Pseudomonadota, 5.3% to Bacillota, and 21.1% to Actinomycetota (Fig. 3d and Table S4), which generally aligned with the Illumina-based microbial composition (Fig. 2b), except for Bacillota. To examine the pan-genomic profile of the GPMAGs among the different PacBio accessions, we further investigated the prevalence of the GPMAGs with the different grapevine accessions measured by the PacBio read mapping rates against each GPMAG. Across the grapevine accessions, *Mangrovibacter phragmitis* exhibited the highest prevalence, followed

by *Roseateles* and *Sphingomonas aquatilis* (Fig. 3e). Interestingly, these GPMAGs all belonged to Pseudomonadota, further supporting its prevalence across grapevine accessions as indicated by the Illumina-based microbial composition (Fig. 2b). Moreover, all GPMAGs representing the highest MAG qualities (complete and near-complete) also belong to Pseudomonadota, underscoring a strong association between MAG quality and prevalence. Together, these results highlight the high-quality and comprehensive assembly of PacBio-based GPMAGs representing valuable genomic resources for the grapevine phyllospheric microbiome investigation.

Annotation of grapevine MAGs reveals genes encoding key biological processes

To delve deeply into the functional roles of the GPMAGs in the grapevine microbiome, particularly those derived from unculturable microorganisms, we performed

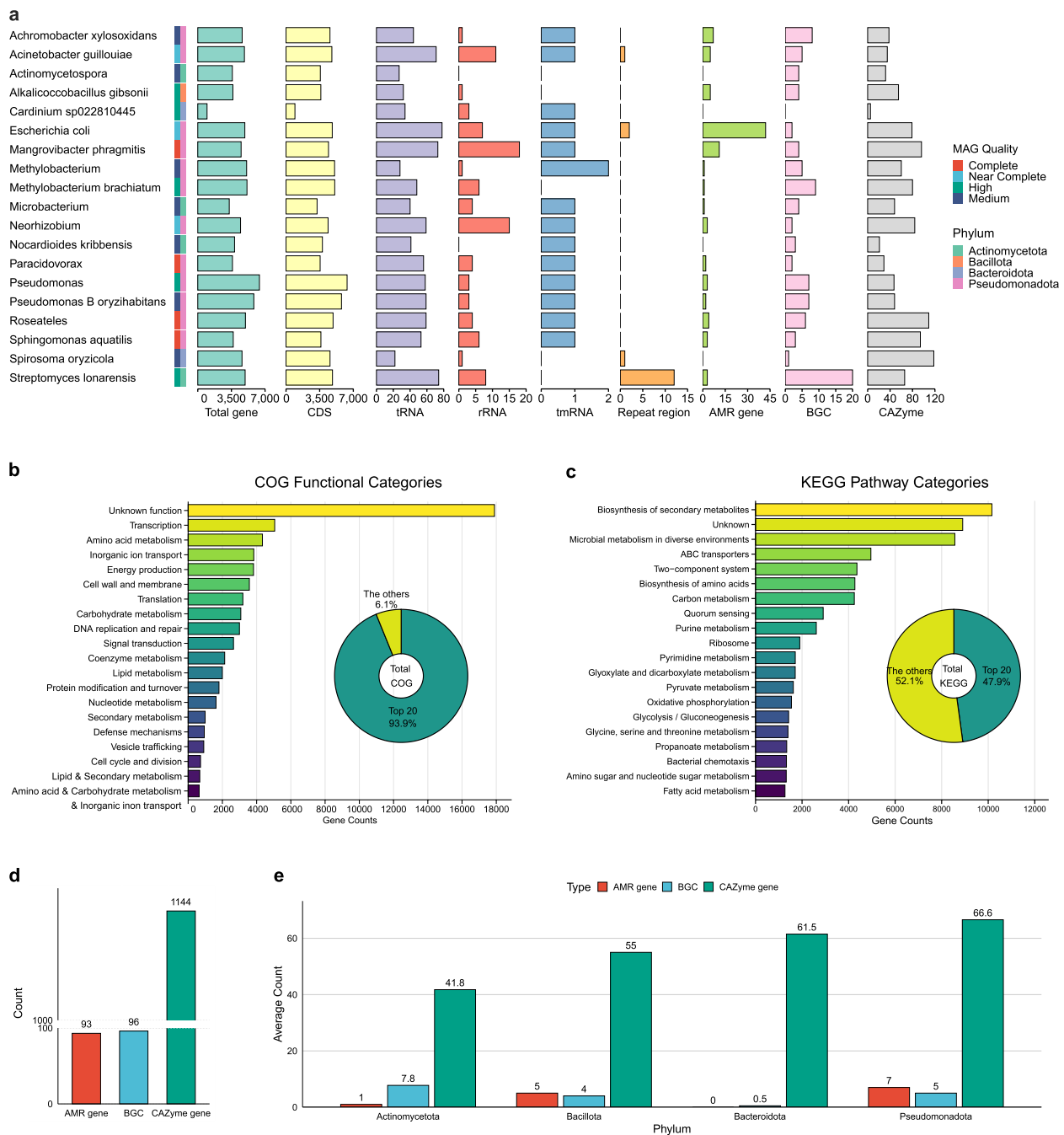


Fig. 4 Functional annotation of the GPMAGs. **a** Gene type distributions in the GPMAGs. **b**, **c** Top 20 COG (**b**) and KEGG pathway (**c**) profiles of the GPMAGs. Donut plots displaying proportions of top 20 terms within total terms. **d** Statistical summary of AMR, BGC, and CAZyme-associated genes. **e** Phylum-level functional capacity displayed as mean AMR/BGC/CAZyme-associated gene counts per phylum. COG Clusters of Orthologous Genes, KEGG Kyoto Encyclopedia of Genes and Genomes, AMR antimicrobial resistance, BGC Biosynthetic Gene Clusters

genomic annotation for the GPMAGs using the Prokka [40] pipeline. This analysis identified a total of 82,884 genes, comprising 81,801 CDSs, 956 tRNA genes, 96 rRNA genes, 15 tmRNA genes, and 16 repeat regions (Fig. 4a and Table S5). Functional annotation of the CDSs was further conducted using eggNOG-mapper v2

[41] and annotated the detailed functions of the CDSs, which identified 67,012 CDSs with various biological process terms. In terms of biological function, proteins with unknown function (26.7%) represented the largest category among the CDSs, and the remaining were primarily associated with core biological functions, such

as transcription, metabolism, and transport (Fig. 4b). Notably, the top 20 most abundant functional terms accounted for 93.9% of the total terms, demonstrating a highly focused functional diversity in the GPMAGs that suggests specialized adaptation to the grape phyllosphere niche. Despite the biosynthesis of secondary metabolites and microbial metabolism in diverse environments being the predominant biological pathway terms, the top 20 terms represented only 47.9% of the total annotations (Fig. 4c), demonstrating a functional versatility in the GPMAGs. These results indicate both functional consistency and expansive metabolic potentials in the GPMAGs, likely facilitating phyllosphere colonization and adaptation in grapevines.

Agriculture such as viticulture heavily relies on pesticide application in disease management regimes, but their misuse may have inadvertently contributed to the emergence of antimicrobial resistance (AMR) in the crop microbiome [42], posing significant threats to both food safety and public health. Notably, the annotated CDSs from GPMAGs frequently encode proteins associated with antibiotic resistance, such as ABC transporter (Fig. 4c), indicating the widespread presence of AMR genes within the GPMAGs. Specifically, we employed RGI [43] to identify 93 AMR genes spanning 16 distinct AMR families (Fig. 4a, d, S2, and Table S6). Notably, a significant proportion of the AMR genes exhibited multi-drug resistance through antibiotic efflux, a predominant resistance mechanism. The abundant presence of AMR genes in the grapevine microbiome implies that antibiotic resistance may have occurred due to prolonged pesticide application and misuse in viticulture.

Plant-associated bacterial endophytes protect their plant hosts from external pathogens by producing secondary metabolites (SMs), including antibiotics and toxins [44–46], establishing them as valuable resources for environmentally friendly and pathogen-effective biocontrol agents. Biosynthetic gene clusters (BGCs) responsible for bacterial SM production typically comprise dozens of related genes distributed across extensive genomic regions, necessitating high-quality MAGs with exceptional completeness and continuity for comprehensive exploration. To leverage our high-quality GPMAGs, we employed antiSMASH [47] to identify a total of 96 BGCs from 26 distinct BGC types in the grapevine phyllospheric microbiome (Fig. 4a, d, S3 and Table S7). Among these, terpene, RIPP-like, and NRPS clusters were the most abundant, commonly associated with antimicrobial activity critical for host defense [48–50]. Other predominant BGCs, including Ni-siderophore, beta-lactone, and T3PKS, highlight the microbiome's ample

potential in SM production. These candidate BGCs provide valuable resources for discovering novel natural products and elucidating their roles in grapevine biology and medicinal applications.

Several studies have demonstrated that microorganism-derived polysaccharides play crucial roles in plant bioprocesses, including immune responses, growth facilitation, stress adaptation, photosynthesis, and plant-bacteria interactions [51–53]. To explore the potential of polysaccharides production in the grapevine microbiome, we employed dbCAN [54] to annotate CAZyme genes from the GPMAGs. The analysis identified a total of 1,144 CAZyme genes spanning 143 unique enzyme families, reflecting the grape phyllospheric microbiome's substantial capacity for polysaccharides metabolism (Fig. 4a, d, S4 and Table S8). Among these genes, glycosyltransferases (GTs) and glycoside hydrolases (GHs) were particularly abundant, with GT2, GT4, GT51, and GH13 emerging as the most prevalent, suggesting a strong focus on polysaccharide synthesis and modification. Additionally, the presence of diverse carbohydrate-binding modules (CBMs) and auxiliary activities (AAs) highlights the microbiome's versatility and its potential involvement in plant–microbe interactions. This extensive CAZyme repertoire indicates that the grape phyllospheric microbiome may influence various host physiological processes, including disease resistance and growth, through polysaccharide synthesis and other host-interactive mechanisms.

To evaluate the functional potentials of each phylum in terms of antibiotic resistance, secondary metabolite biosynthesis, and carbohydrate metabolism that differentially influence the host's biology, we quantified the mean numbers of AMR genes, BGCs, and CAZyme genes of each phylum in the GPMAGs. As a result, Pseudomonadota contained the highest abundance of AMR genes (Fig. 4e), indicating a particularly strong antibiotic resistance potential. Actinomycetota showed the greatest capacity for secondary metabolite biosynthesis, harboring the most BGCs. While all phyla maintained comparable CAZyme gene numbers, Pseudomonadota demonstrated particularly extensive carbohydrate metabolic potential. These results reveal phylum-specific functional potentials that differentially involve the host's biological processes.

In conclusion, these findings provide valuable insights into the functional potential of the grapevine phyllospheric microbiome, emphasizing its involvement in shaping the host's health and resilience through microbial-host interactions.

Downy mildew resistance relates to complexity of microbial interaction network with presence of key genera

To investigate the role of the grapevine microbiome during *P. viticola* infection, we profiled the microbiome composition of the grapevine accessions based on their DM disease phenotypes (susceptible or resistant). Unexpectedly, the microbiome compositions of the susceptible and resistant microbiomes were nearly identical (Fig. S5a). Additionally, β -diversity analysis further revealed no overall compositional difference between the two groups (Fig. S5b), suggesting that microbiome composition unlikely accounts for the differences in DM resistance. Therefore, we hypothesized that interactions within the microbial community might contribute to the observed resistance. To test this hypothesis, we analyzed the complexity of the microbial interaction networks in the susceptible and resistant grape microbiomes using Sparse Inverse Covariance estimation for Ecological Association and Statistical Inference (SpiecEasi) [55] analysis. SpiecEasi uses sparse inverse covariance estimation to identify true associations between microorganisms based on correlations derived from the covariance of their relative abundances. Positive correlations indicate cooperative interactions, while negative correlations represent antagonistic interactions. Remarkably, the microbiome associated with resistant plants exhibited a significantly more complex network than the susceptible ones (Fig. 5a), with a higher frequency of negative interactions per genus (Fig. S6), indicating the involvement of microbial network complexity in modulating DM resistance, especially the negative interactions that play a crucial role in the complexity.

Notably, the microbiome associated with resistant plants was highly compartmentalized and centralized around a few genera (e.g., *Lactiplantibacillus*, *Liberibacter*, *Paracitovorax*, and *Xanthomonas*), with relatively much higher abundance (Fig. 5a), suggesting the presence of key genera with pivotal roles in network dynamics and resistance modulation. To identify these key genera, we first profiled the top 100 most abundant genera in the resistant microbiome based on relative abundance and microbial interaction metrics. These profiles were then systematically compared with those from the susceptible microbiome to evaluate their respective contributions to the network complexity. Among them, *Liberibacter* and *Pseudomonas* emerged as the most influential genera, with significantly increased interaction metrics compared to those of the susceptible microbiome (Fig. 5b). Intriguingly, all these key genera belong to the core microbiome (Table S3), highlighting the importance of the core microbiome in modulating the complexity. Additionally, genera such as *Bacillus* exhibit increased abundance and interaction metrics compared to the susceptible microbiome,

indicating a contribution to the complexity via abundance and microbial interaction. Despite relatively low abundances in the resistant microbiome, genera such as *Bacteroides*, *Agathobacter*, *Bifidobacterium*, *Anaerobutyricum*, and *Neisseria* exhibit the highest interactions (8–10 interactions/genus) in the resistant microbiome, demonstrating the important role of low abundant genera in forming the complexity. To evaluate phylum-level contributions to the increased complexity, we compared the proportions of microbial interaction between the resistant and susceptible microbiomes for each phylum. As a result, Bacillota was identified as the primary contributor, exhibiting the most significantly increased proportion in the resistant microbiome (33.5%) compared to that of the susceptible one (23.9%), while other phyla show minimal fluctuation in the change (Fig. S7). Furthermore, we analyzed Spearman correlations between Bacillota and other major phyla in the resistant microbiomes based on their abundances, identifying those with the strongest interactions with this key phylum. The correlation analysis revealed that Euglenozoa, one of the major eukaryotes in the grapevine microbiome (Fig. 2c), and Cyanobacteriota exhibit the strongest positive interaction with Bacillota, while the other phyla display no significant positive interactive dynamics (Fig. S8), indicating that Bacillota positively interacts with these phyla in the resistant microbiome. These findings suggest that not all genera contribute equally to the modulation of the network complexity; rather, specific key genera play a primary role along with specific interactions.

To validate the effectiveness of key genera profiling and the role of key genera during *P. viticola* infection, we screened microorganisms isolated from the grapevine leaves. Among the key genera, we successfully identified two *Bacillus* strains which belong to Bacillota, *B. cereus* 117 gite and *B. cereus* SW9SE. Given our observation that *Bacillus* is a key genus belonging to Bacillota with an increased proportion and microbial interactions in the microbiome associated with resistant plants (Fig. 5b and S5a), and considering a prior report on *B. cereus* efficacy against grape ripe rot disease [56], we hypothesized and tested the biocontrol potential of *B. cereus* against *P. viticola* infection. To this end, we inoculated grapevine leaf discs with varying concentrations of these strains followed by *P. viticola* infection. Strikingly, both strains exhibited apparent DM inhibition at a low concentration (e.g., OD=0.02) (Fig. 5c and d), demonstrating the potent biocontrol activity of the strains against *P. viticola* infection. These findings underscore the effectiveness of our key genera profiling and pave the way for *B. cereus* as a novel biocontrol resource for managing DM infection in viticulture.

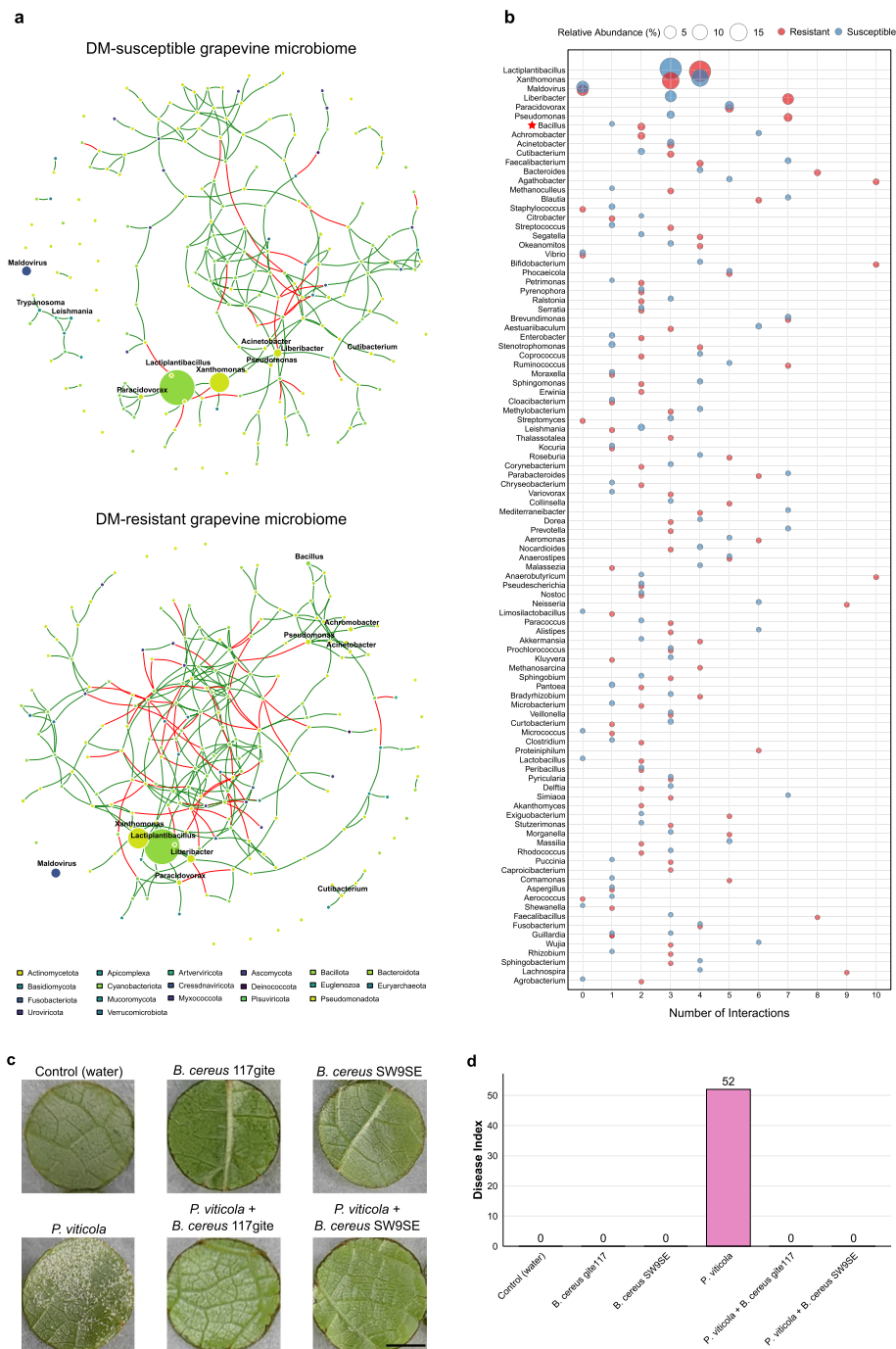


Fig. 5 Downy-mildew resistance shapes complex microbial interaction networks with key genera. **a** Microbial co-occurrence networks based on the DM resistance types. The networks displaying the top 200 most abundant genera within each type. Node sizes indicate genera' relative abundances, with colors denoting different phyla as shown in bottom panel legend. Green edges represent positive correlations, while red edges indicate negative correlations between nodes. The top 10 abundant genera within each microbiome were labeled above corresponding nodes. **b** Profiles of top 100 abundant key microorganisms in the DM-resistant grapevine microbiome, illustrating relative abundances and numbers of microbial interactions. Colored bubbles represent the profiles in the corresponding microbiome. The top 100 genera of the resistant microbiome which not present in the top 200 genera in the susceptible microbiome were not shown in the plot. Red star indicates a genus that experimentally isolated and validated in this study. **c** Biocontrol assay of two *B. cereus* strains against *P. viticola* infection. The top and bottom rows show mono- and co-inoculation, respectively. Black bar indicates 5 mm. **d** Bar plot of the disease index shown in **c** where higher values indicate more severe infection

Collectively, our pan-metagenomic analysis and experiments highlight the association of the microbial network complexity with DM resistance, driven by intricate interactions involving specific key genera with specific interactions.

Downy mildew triggers eukaryotic primary and secondary metabolism of the grapevine microbiome

Given that network complexity is primarily driven by Bacillota, which exhibits a strong positive abundance correlation with one of the dominating Eukaryotes—Euglenozoa, we hypothesize that key bacterial phyla involve eukaryotic transcriptomes in the grapevine microbiome, potentially underlying defensive responses against *P. viticola* infection. To examine this, we conducted pan-metatranscriptomic sequencing and analysis to investigate the transcriptional regulation of the grapevine phyllospheric microbiome in response to the infection.

In brief, we collected samples from both infected (1 day post-inoculation) and uninfected leaves from 103 of the grapevine accessions and subjected them to RNA-seq [27] (Fig. 1d). The RNA-seq reads were mapped to the grapevine and *P. viticola* reference genomes (NCBI accession: GCA_001695595.3), obtaining the unmapped reads that represent eukaryotic organisms other than grapevine and *P. viticola* (Table S2). The unmapped reads were then de novo assembled using Trinity [57], resulting in 192,675 eukaryotic transcripts. Differentially expressed transcript (DET) analysis identified 40,546 significantly up-regulated and 5081 downregulated transcripts (Fig. 6a), indicating a considerable overall upregulation of the microbial transcripts upon the infection. Moreover, KEGG and GO enrichment revealed that the up-regulated genes were primarily involved in primary and secondary metabolisms (Fig. 6b and c), with secondary metabolism being particularly prominent in the

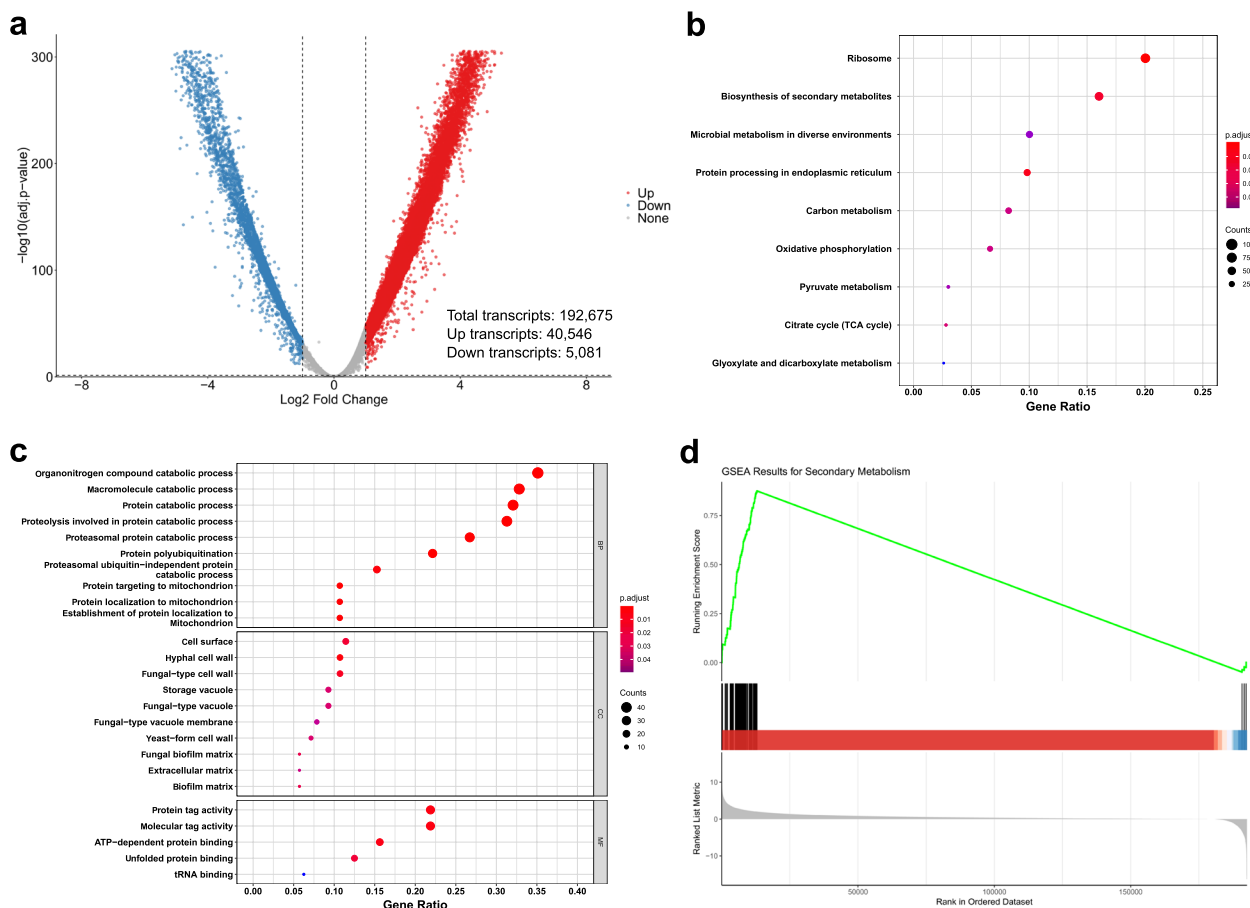


Fig. 6 Eukaryotic transcriptome profile of the grapevine phyllospheric microbiome in response to the oomycete *Plasmopara viticola* infection. **a** Volcano plot of the grape microbiome upon infection by *P. viticola*. The dashed lines indicate thresholds for statistical significance ($\log_2$ fold change > 1, adjusted $p < 0.05$). **b**, **c** KEGG pathway (**b**) and GO enrichment analysis (**c**) of the up-regulated DETs. GO enrichment shows the top 10 terms in each category. **d** GSEA of the secondary metabolism-related transcripts. The graph illustrates the enrichment score (green line) for the transcript set associated with secondary metabolism across the ranked transcript list (bottom) by $\log_2$ fold change values. The black bars in the middle represent the positions of transcripts from the set in the ranked list with significances indicated by red and blue

microbiome's response to the infection. Consistently, Gene Set Enrichment Analysis (GSEA) further confirmed significant enrichment of secondary metabolism-related gene sets among the upregulated transcripts (Fig. 6d). Notably, these up-regulated transcripts encoded the key proteins of BGCs, including ketoacyl reductase (KR) domains and flavin mononucleotide-dependent dehydrogenases, both of which are integral to the biosynthesis of polyketides and antibiotics [58, 59]. Together, these findings demonstrate that primary and secondary metabolisms in eukaryotic microorganisms are significantly induced during *P. viticola* infection, suggesting a potential role for the grapevine phyllospheric microbiome in defending against pathogen invasion through the competitive exclusion and SM productions that are associated with key bacterial genera.

Discussion

Microbiomes have a profound impact on the physiology and immunity of host organisms ranging from humans, animals, and plants. Mounting evidence underscores the critical roles of the microbiome in the health, yield, and quality [4, 11–13] of grapevines, which are one of the most economically important fruit crops cultivated worldwide for the production of fresh fruits, wine, juice, and raisins, etc. Despite a huge body of knowledge gained in microbiome research, pan-microbiome diversity across host genera is poorly understood, lacking a large-scale pan-metagenomic approach, especially from long-read sequencing. In this study, combining short-read, long-read sequencing, and RNA-seq, we conducted a pan-metagenomic and pan-metatrascriptomic study of 107 accessions from 34 *Vitis* species to understand the structure, diversity, and functional roles of the pan-microbiome in the grapevine phyllosphere, a key battleground for host-microbe interactions. We identified the core microbiome dominated by Bacillota and Pseudomonadota. Leveraging PacBio sequencing of the grapevine phyllosphere, we successfully assembled 19 high-quality GPMAGs, including four complete MAGs, with detailed functional annotations, representing the first assembly of MAGs derived from plant PacBio datasets. Furthermore, pan-metagenomic and pan-metatrascriptomic analyses elucidated the underlying defense mechanisms of the grape microbiome against *P. viticola* infection. Our findings demonstrate the profound influence of the microbiome on grapevine biology, particularly its involvement in the resistance of *P. viticola* infection, and provide a foundation and valuable genomic and microbial resources for developing microbiome-based biocontrol strategies in viticulture.

Structural analysis of the grapevine phyllosphere microbiome revealed the core microbiome dominated by specific phyla, primarily Bacillota and Pseudomonadota, which constitute the dominant proportion of the microbial community (Fig. 2). This composition aligns well with previous studies on grapevine phyllosphere microbiota across different cultivars and geographic regions [4, 15, 60, 61], highlighting the prevalence and pivotal roles of the core microbiome. However, the mechanisms underlying core microbiome recruitment in the grapevine phyllosphere remain largely unknown. In plants, it is well established that metabolic secretions, particularly SMs such as flavonoids, salicylic acid, and strigolactones, shape microbiome composition through plant-microorganism interactions [62–65]. In turn, the modulated microbiome significantly influences host biological processes via plant-microbe interaction, including growth, disease resistance, and stress responses [62–65]. As a major plant part interacting with microbial communities, common SMs might accumulate in grapevine leaves of diverse genetic backgrounds, thereby helping establish the core microbiome. This core microbiome plays critical roles in growth, disease resistance, and stress responses, underscoring the importance of grapevine-microbe interactions in maintaining normal plant development.

Compared to amplicon sequencing analysis, MAGs provide novel and pivotal insights into the plant microbiome. Our GPMAGs demonstrate high quality and prevalence, serving as ideal genomic resources with detailed annotations such as AMR, BGC, and CAZyme information for further applications. To our knowledge, this is the first study to generate MAGs from PacBio reads of a large collection of plant accessions. However, the rare presence of dominant phyla such as Bacillota in the GPMAGs (Fig. 3d and Table S4) highlights the limitations of long-read technology in the number of microbial genomes recovered from the microbiome, primarily due to the higher cost of sequencing and thus lower data coverage compared to Illumina data, and the limitation of bioinformatic methods [66, 67]. Decreasing sequencing costs and further optimization and advancement of these methods will improve PacBio-based MAG assemblies in the future. Together, this study establishes a novel and robust method for assembling MAGs from plants, leveraging the contiguity and accuracy of PacBio reads to accelerate research on the plant microbiome.

It is widely recognized that, through interactions, microorganisms and plants modulate their immunity to combat pathogen invasion [68, 69]. This mechanism primarily relies on microbiome-derived SM production and competitive exclusion to protect host plants from

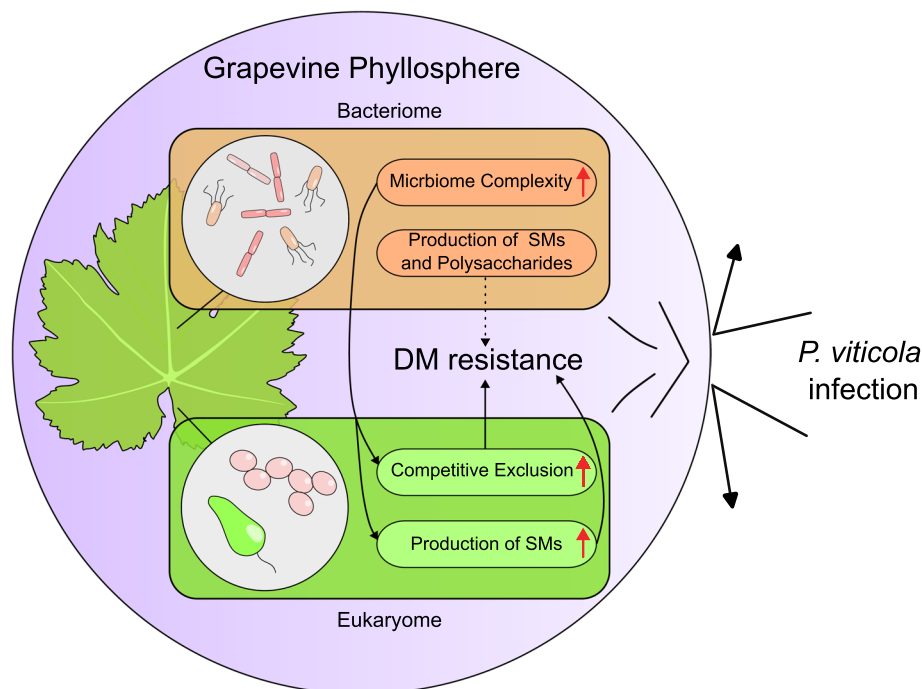


Fig. 7 Schematic diagram depicting grapevine microbiome's potential deployment of bacteriome and eukaryome to combat *P. viticola* infection. This pan-metagenomic and pan-metatranscriptomic study demonstrates that the grapevine bacteriome orchestrates microbiome assembly, with key taxa (e.g., Bacillota) interacting selectively with eukaryotes (e.g., Euglenozoa) to involve eukaryotic competitive exclusion and secondary metabolites (SMs). The bacteriome may further utilize SMs and polysaccharides for enhanced DM resistance (dashed line, speculative)

various pathogens. Consistent with this concept, our findings demonstrate the pivotal roles of the bacteriome and eukaryome, of which the bacteriome re-orchestrates microbial community, especially facilitating interactions with microorganisms to involve their SM production and competitive exclusion in response to *P. viticola* infection (Fig. 7). Intriguingly, we observed a significant increase in network complexity, driven by the contributions of key genera from the bacteriome (Fig. 5 and Fig. S6, S7), which involve the defense mechanism against *P. viticola* infection in grapevines through interactions with the eukaryome. In olive, *Pseudomonas simiae* PICF7, a biological control agent for Verticillium wilt, was reported to significantly increase microbial network complexity upon the infection, resulting in a more stable network that plays a crucial role in defense [70]. Thus, it is likely that grapevines employ a similar strategy to modulate their microbial network, optimizing SM production and competitive exclusion capacities of the eukaryome to effectively combat *P. viticola* infection. Additionally, the profiling of key genera in the DM-resistant microbiome provides a valuable resource for exploring biocontrol microorganisms against *P. viticola* infection, exemplified by the experimental validation of two *B. cereus* strains, laying the foundation for environmentally friendly and sustainable biocontrol strategies in viticulture.

Materials and methods

Collection of grapevine leaves for metagenome sequencing

The leaf samples from 107 grapevine accessions used in this study were collected from vineyards at the Zhengzhou Fruit Research Institute, Chinese Academy of Sciences (Zhengzhou, Henan, China), and the Institute of Advanced Agricultural Sciences, Peking University (Weifang, Shandong, China). All grapevines were maintained under standard management practices, including watering, fertilization, pruning, and disease control.

Downy mildew infection assay

DM infection was performed as described previously [71]. Each grapevine leaf disc (8 mm in diameter) was inoculated with 20 μ L of *P. viticola* suspension (1×10^5 sporangia/mL) on the abaxial surface and incubated at 18 ± 2 °C with $80 \pm 10\%$ relative humidity in a growth chamber under a 12 h light ($60 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$)/12 h dark cycle. Leaf discs were collected 24 h post-inoculation for transcriptome analysis. At least three biological replicates were conducted, with each replicate comprising five leaf discs from different grapevines.

DNA and RNA extraction

Genomic DNA for Illumina was extracted from the grapevine leaves using a DNeasy Plant Mini Kit

(QIAGEN, Hilden, Germany) according to the manufacturer's instructions. The DNA integrity was assessed with an Agilent 4200 Bioanalyzer (Agilent Technologies, Santa Clara, USA). High-molecular-weight (HMW) DNA for PacBio sequencing was prepared from the fresh grapevine leaves using the cetyltrimethylammonium bromide (CTAB) method [72]. The DNA quality was verified with 1% agarose gel electrophoresis, and purity was assessed using a NanoPhotometer[®] spectrophotometer (IMPLEN, Westlake Village, USA). The DNA concentration was measured with a Qubit[®] 2.0 Fluorometer (Life Technologies, Carlsbad, USA). Total RNA was extracted using TRIzol reagent (Thermo Fisher Scientific, Waltham, USA) following the manufacturer's instructions. The RNA concentration was measured with a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, USA), and the purity was evaluated using a NanoPhotometer[®] (IMPLEN, Westlake Village, USA). The RNA concentration and integrity were further analyzed with an Agilent 2100 Bioanalyzer using the RNA Nano 6000 assay kit (Agilent Technologies, Santa Clara, USA).

Metagenome sequencing

For Illumina library preparation, 15 µg of the genomic DNA was sheared using g-Tubes (Covaris, Woburn, USA) and concentrated with AMPure PB magnetic beads (Pacific Biosciences, Menlo Park, USA). The library construction was performed with 1 µg of the DNA template following the TruSeq DNA Sample Preparation Guide (Illumina, San Diego, USA). Paired-end sequencing was conducted on a HiSeq X Ten platform (Biomarker Technologies, Qingdao, China), generating 150 bp Illumina reads. For PacBio sequencing, 15 µg of the purified HMW genomic DNA was used to construct a standard SMRTbell library with the SMRT Express Template Prep Kit 2.0 (Pacific Biosciences, Menlo Park, USA). Circular consensus sequencing (CCS) producing PacBio reads was performed on a PacBio Sequel IIe system at Biomarker Technologies (Qingdao, China). To prepare RNA sequencing libraries, 1–3 µg of the total RNA was used as input for library construction with the VAHTS Universal V6 RNA-seq Library Prep Kit for Illumina[®] (Vazyme, Nanjing, China) according to the manufacturer's instructions. The RNA concentration was initially quantified with the Qubit[®] RNA Assay Kit on a Qubit[®] 3.0 (Thermo Fisher Scientific, Waltham, USA) and diluted to 1 ng/µl. The insert size was measured with an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, USA), and the final library concentration (>10 nM) was verified with a Bio-RAD CFX 96 quantitative PCR system (Bio-Rad, Hercules, USA). Paired-end sequencing was then performed on a HiSeq X Ten platform (Biomarker Technologies, Qingdao, China), generating 150 bp RNA-seq reads.

Quality control and removal of host reads

To ensure the quality of the short reads, quality control was performed on the Illumina and RNA-seq reads using fastp [73] (v0.23.4) with default parameters, resulting in the quality-controlled reads. To completely filter out reads from the grapevine hosts, the Illumina, PacBio, and RNA-seq reads were mapped to the complete reference genome of *Vitis vinifera* [27] (VHP-T2T.hap1, NCBI accession: GCA_046255155.1) using Minimap2 [74] (v2.28) with parameters (Illumina, RNA-seq: -ax sr, PacBio: -ax map-pb). The mapped results were further filtered using SAMtools [75] (v1.21) with parameters (Illumina, RNA-seq: view -u -f 12 -F 256, PacBio: view -f 4 -F 256) to retain microbial reads. Additionally, the RNA-seq reads were further filtered by mapping to *P. viticola* reference genome INRA_Pvit_2 (NCBI accession: GCA_001695595.3) with Minimap2 and SAMtools to exclude *P. viticola* reads.

Taxonomic classification

To profile the taxonomic composition of the grapevine microbiome, the filtered Illumina reads were taxonomically classified using Kraken2 [28] (v2.1.3) and k2_pluspf_20240904 database with parameters (-paired -confidence 0.1 -minimum-hit-groups 2), yielding the compositional statistics of the grapevine microbiome.

Pan-metagenomic and diversity analysis

The identification of the core, dispensable, and unique microbiome was conducted with R using the microbial composition statistics. To assess microbiome diversity, PCoA analysis was performed with the vegan [76] (v2.6–8) package in R. All statistical analyses and visualizations were performed with R.

Assembly of the GPMAGs

To generate the GPMAGs, the filtered PacBio reads were assembled using Flye [30] (v2.9.5), Hifiasm [31] (v0.19.9), Hifiasm-meta [32] (vhamtv0.3.2), metaFlye [33] (v2.9.5), MetaMDBG [34] (v1.0), and Verkko [35] (v2.2.1) with default parameters to produce initial contig assemblies. The resulting contigs were then binned with CONCOCT [36] (v1.1.0), MaxBin2 [37] (v2.2.7), and MetaBAT2 [38] (v2.17) with default parameters. The bins were quality-checked using CheckM2 [77] (v1.0.2), and those meeting the medium-quality criterion [34] were dereplicated to generate non-redundant GPMAGs using Galah (v0.4.2) (<https://github.com/wwood/galah>) with parameters (-ani 95 -precluster-ani 90 -min-aligned-fraction 30 -min-completeness 50 -max-contamination 10).

Taxonomic and phylogenetic analysis of the GPMAGs

To reveal the taxonomic composition of the GPMAGs, GTDB-Tk [39] (v2.4.0) was employed with the R220 database and default parameters, generating the taxonomic composition and a multiple sequence alignment (MSA) file. The MSA file was further subjected to the construction of a phylogenetic tree using IQ-TREE2 [78] (v2.1.4-beta) with default parameters and visualized by iTOL [79] (<https://itol.embl.de/>).

Functional annotation of the GPMAGs

To analyze different gene types from the GPMAGs, Prokka [40] (v1.14.6) was used for gene type prediction with default parameters. The predicted CDSs were further applied for functional annotation using eggNOG-mapper [41] (2.1.12) with parameters (-m diamond -evaluate 1e-5 -query_cover 50 -subject_cover 50). To annotate AMR genes, BGCs, and CAZyme genes of the GPMAGs, RGI [43] (v6.0.3) (default parameters), antiSMASH [47] (v7.1.0) (parameters: -genefinding-tool prodigal -taxon bacteria -fullhmmer -cb-knownclusters -cb-subclusters -cc-mibig -asf -rre -pfam2go), and dbCAN [54] (v4.1.4) (default parameters) were employed, respectively.

Construction of microbial interaction network

Microbial interaction network was constructed using the microbial composition statistics and SpiecEasi [55] (v1.1.3) package in R with parameters (lambda.min.ratio = 0.5, nlambdas = 30, rep.num = 100).

Isolation of grapevine phyllosphere microorganisms

To isolate microorganisms from grapevine leaves, we collected 10 leaves from each of the following cultivars: *V. vinifera* cv. Chardonnay, *V. vinifera* cv. Muscat Hamburg, *V. vinifera* × *V. labrusca* cv. Kyoho, and *V. vinifera* × *V. labrusca* cv. Vidal Blanc at the vineyards. After rinsing with sterile water, the leaves were homogenized in 1 × PBS buffer using an electric mixer to obtain a microbial suspension. 2 ml of the suspension was inoculated onto Nutrient Agar medium (Coolaber Technology Co., Ltd., Beijing, China) and incubated at 30 °C for 24 h. Single colonies were isolated and further purified by streaking on Nutrient Agar medium. Species identification was performed by 16S rRNA sequencing (Tsingke Biotechnology Co., Ltd., Beijing, China).

Experimental validation of biocontrol activity

To validate the efficacy of isolated microorganisms against *P. viticola* infection, the *B. cereus* strains were cultured in Nutrient Broth medium (Coolaber Technology Co., Ltd., Beijing, China) at 30 °C and 220 rpm

for 12 h. The cultures were diluted to an OD of 0.02 at 600 nm using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, USA). The diluted suspensions were sprayed onto 1 cm grape leaf discs and incubated at 30 °C and 220 rpm for 24 h. Subsequently, *P. viticola* was co-inoculated onto the discs using the method described above and incubated at 19 °C for 6 days. Disease severity was assessed 6 days post-inoculation using the following grading scale: Grade 0, no lesions (0% of leaf disc area affected); Grade 1, lesions covering less than 10% of the leaf disc area; Grade 2, lesions covering 10–30% of the leaf disc area; Grade 3, lesions covering 31–60% of the leaf disc area; Grade 4, lesions covering 61–80% of the leaf disc area; and Grade 5, lesions covering more than 80% of the leaf disc area. The disease index (DI) was calculated as follows:

$$DI = \frac{\sum(s_i n_{i2})}{N_2 5} 100$$

where s_i represents the disease grade (ranging from 0 to 5), n_{i2} is the number of leaf discs with disease grade s_i , and N_2 is the total number of leaf discs assessed. The DI value was rounded to one decimal place.

Transcriptomic analysis of *P. viticola* infection

Transcriptomic profiling in response to *P. viticola* infection was conducted using the filtered RNA-seq reads. First, De novo assembly of transcripts was performed with Trinity [57] (v2.15.2) using default parameters. The generated transcripts were dereplicated using VSEARCH [80] (2.29.0) with parameters (-id 0.95), selecting the longest transcripts as the representatives of duplicates. Transcript expression was quantified using Salmon [81] (v1.10.3) with default parameters and the filtered RNA-seq reads. Differential expression analysis was performed with the DESeq2 [82] (v1.45) package in R. Subsequently, the significantly up-regulated genes were further annotated with eggNOG-mapper, followed by KEGG, GO, and GSEA enrichment analyses using AnnotationForge (v1.47.1), AnnotationDbi (v1.67.0), ClusterProfiler [83] (v4.13.3) and enrichplot (v1.25.2) packages in R with default parameters.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-025-02287-4>.

Supplementary Material 1. Figure S1. Worldwide comparison of grapevine phyllosphere microbiome compositions. a, Geographical distribution of *Vitis* accessions used for comparative analysis. One representative *Vitis* accession was selected per country. b, Top 10 phylum level composition of grapevine phyllosphere microbiome across worldwide *Vitis* accessions. PKU represents mean phylum composition of this study.

Supplementary Material 2. Figure S2. Resistome profile of the GPMAGs. Identified AMR gene frequencies in the GPMAGs were categorized by gene family, resistance mechanism, and target antibiotics.

Supplementary Material 3. Figure S3. BGC profile of the GPMAGs. Identified BGCs from the GPMAGs were displayed by bar chart.

Supplementary Material 4. Figure S4. CAZyme profile of the GPMAGs. Bar chart exhibits identified CAZyme types of the GPMAGs.

Supplementary Material 5. Figure S5. Microbiome profiles associated with DM resistance. a, Mean relative abundances of the top 10 most abundant genera based on the types of DM resistance. b, PCoA of the grape microbiomes at genus level. Each point represents a grapevine accession with colored according to its variety. Ellipses denote the 95% confidence intervals for each group. Percentages on the axes indicate the variance explained by PCoA1 and PCoA2. PERMANOVA analysis revealed significant differences among the groups, with R^2 values represent the proportion of variance explained by the grouping factor and p -values indicate the statistical significance of these differences.

Supplementary Material 6. Figure S6. Microbial interaction profiles associated with DM-resistance. Violin plots depict the median microbial interactions for each node in the networks shown in Fig. 5a. White diamonds indicate medians. Significance levels are denoted as ** ($p < 0.01$), **** ($p < 0.0001$), and ns (not significant).

Supplementary Material 7. Figure S7. Phylum-level microbial interaction proportions between microbiomes. Bubble plots display interaction frequencies among the top 10 abundant phyla in the susceptible and resistant microbiomes, with bubble sizes corresponding to proportional values (numerically labeled).

Supplementary Material 8. Figure S8. Abundance correlation between Bacillota and major phyla in the resistant microbiome. Heatmap depicts Spearman coefficients between Bacillota and major phyla based on their abundances in the resistant microbiome. Numerical values and significance levels * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$), ns (not significant) are shown in cells.

Supplementary Material 9. Table S1 Summary of the grapevine accessions. Table S2 Statistical summary of the filtered Illumina, PacBio, and RNA-seq datasets. Table S3 Microbial composition of the grape phyllospheric microbiome. Table S4 Statistical summary of the GPMAGs. Table S5 Statistical summary of the GPMAG gene types. Table S6 Summary of AMR genes in the GPMAGs. Table S7 Summary of BGCs in the GPMAGs. Table S8 Summary of CAZymes in the GPMAGs.

Acknowledgements

We would like to thank the Bioinformatics Platform at Peking University Institute of Advanced Agricultural Sciences for providing the high-performance computing resources.

Authors' contributions

LG and WXY conceived and supervised the study. JYJ, XNZ, JJM, WZ, LLG, HSS, CHL and LLZ prepared the plant samples and performed the experiments. JYJ and XFW conducted the bioinformatics analysis. JYJ, LG and WXY wrote and revised the manuscript. All authors read and approved the final manuscript.

Funding

This project was supported by Key R&D Program of Shandong Province (Grant No. 2024CXPT031 and ZR202211070163), the Taishan Scholars Program of Shandong Province and Weifang Key Laboratory of Grapevine Improvement and Utilization. LG is also supported by the Natural Science Foundation for Distinguished Young Scholars of Shandong Province (Grant No. ZR2023JQ010). JYJ is supported by the Shandong Provincial Natural Science Foundation Youth Project (Grant No. ZR2024QC241).

Data availability

The filtered datasets of Illumina and PacBio reads were deposited in European Nucleotide Archive (ENA) under accession number PRJEB85252 and PRJEB86040, respectively. The FASTA files of the 19 GPMAGs generated in this study were deposited at the figshare under <https://doi.org/10.6084/m9.figshare.28307192.v1>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Shandong Key Laboratory of Precision Molecular Crop Design and Breeding, Peking University Institute of Advanced Agricultural Sciences, Shandong Laboratory of Advanced Agricultural Sciences in Weifang, Weifang, Shandong 261325, China. ²Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou 450004, China.

Received: 16 April 2025 Accepted: 22 October 2025

Published online: 01 December 2025

References

- Food and Agriculture Organization of the United Nations. FAOSTAT. Rome; 2024.
- Zhou DD, Li J, Xiong RG, Saimaiti A, Huang SY, Wu SX, et al. Bioactive compounds, health benefits and food applications of grape. *Foods*. 2022. <https://doi.org/10.3390/foods11182755>.
- Buonassisi D, Colombo M, Migliaro D, Dolzani C, Peressotti E, Mizzotti C, et al. Breeding for grapevine downy mildew resistance: a review of "omics" approaches. *Euphytica*. 2017. <https://doi.org/10.1007/s10681-017-1882-8>.
- Bettenfeld P, Cadena i Canals J, Jacquens L, Fernandez O, Fontaine F, van Schaik E, et al. The microbiota of the grapevine holobiont: a key component of plant health. *J Adv Res*. 2022. <https://doi.org/10.1016/j.jare.2021.12.008>.
- Campos G, Chialva C, Miras S, Lijavetzky D, Frontiers Media S.A. New technologies and strategies for grapevine breeding through genetic transformation. *Front Plant Sci*. 2021. <https://doi.org/10.3389/fpls.2021.767522>.
- Cantu D, Walker MA. Compendium of plant genomes the grape genome. Available from: <http://www.springer.com/series/11805>
- Magon G, De Rosa V, Martina M, Falchi R, Acquadro A, Barcaccia G, et al. Boosting grapevine breeding for climate-smart viticulture: from genetic resources to predictive genomics. *Front Plant Sci*. 2023. <https://doi.org/10.3389/fpls.2023.1293186>.
- Peng J, Wang X, Wang H, Li X, Zhang Q, Wang M, et al. Advances in understanding grapevine downy mildew: from pathogen infection to disease management. *Mol Plant Pathol*. 2024. <https://doi.org/10.1111/mpp.13401>.
- Fournier P, Pellán L, Jaswa A, Cambon MC, Chataigner A, Bonnard O, et al. Revealing microbial consortia that interfere with grapevine downy mildew through microbiome epidemiology. *Environ Microbiome*. 2025. <https://doi.org/10.1186/s40793-025-00691-9>.
- Duret M, Wallner A, Besaury L, Aziz A. Diversity and functional features of the root-associated bacteriome are dependent on grapevine susceptibility to *Plasmopara viticola*. *Environ Microbiome*. 2025. <https://doi.org/10.1186/s40793-025-00690-w>.
- Bokulich NA, Thorngate JH, Richardson PM, Mills DA. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proc Natl Acad Sci U S A*. 2014. <https://doi.org/10.1073/pnas.1317371110>.
- Ding Y, Wei R, Wang L, Wang W, Wang H, Li H. Exploring the ecological characteristics of natural microbial communities along the continuum from grape berries to winemaking. *Food Res Int*. 2023. <https://doi.org/10.1016/j.foodres.2023.112718>.
- Bokulich NA, Collins TS, Masarweh C, Allen G, Heymann H, Ebeler SE, et al. Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics. *MBio*. 2016. <https://doi.org/10.1128/mBio.00631-16>.

14. Coller E, Cestaro A, Zanzotti R, Bertoldi D, Pindo M, Larger S, et al. Microbiome of vineyard soils is shaped by geography and management. *Microbiome*. 2019;7:140.
15. Cui S, Zhou L, Fang Q, Xiao H, Jin D, Liu Y. Growth period and variety together drive the succession of phyllosphere microbial communities of grapevine. *Sci Total Environ*. 2024. <https://doi.org/10.1016/j.scitotenv.2024.175334>.
16. Marasco R, Rolli E, Fusi M, Michoud G, Daffonchio D. Grapevine rootstocks shape underground bacterial microbiome and networking but not potential functionality. *Microbiome*. 2018. <https://doi.org/10.1186/s40168-017-0391-2>.
17. Awad M, Giannopoulos G, Mylona PV, Polidoros AN. Comparative analysis of grapevine epiphytic microbiomes among different varieties, tissues, and developmental stages in the same terroir. *Appl Sci*. 2022;13:102.
18. Wei T-L, Zheng Y-P, Wang Z-H, Shang Y-X, Pei M-S, Liu H-N, et al. Comparative microbiome analysis reveals the variation in microbial communities between 'Kyoho' grape and its bud mutant variety. *PLoS One*. 2023;18:e0290853.
19. Vitulo N, Lemos WJF, Calgaro M, Confalone M, Felis GE, Zapparoli G, et al. Bark and grape microbiome of *Vitis vinifera*: influence of geographic patterns and agronomic management on bacterial diversity. *Front Microbiol*. 2019. <https://doi.org/10.3389/fmicb.2018.03203>.
20. Bao L, Sun B, Wei Y, Xu N, Zhang S, Gu L, et al. Grape cultivar features differentiate the grape rhizosphere microbiota. *Plants*. 2022;11:1111.
21. Awad M, Giannopoulos G, Mylona PV, Polidoros AN. Genotype may influence bacterial diversity in bark and bud of *Vitis vinifera* cultivars grown under the same environment. *Appl Sci*. 2020;10:8405.
22. Swift JF, Migicovsky Z, Trello GE, Miller AJ. Grapevine bacterial communities display compartment-specific dynamics over space and time within the Central Valley of California. *Environ Microbiome*. 2023;18:84.
23. Fernandes P, Afonso IM, Pereira J, Rocha R, Rodrigues AS. Epiphytic microbiome of Alvarinho wine grapes from different geographic regions in Portugal. *Biology*. 2023;12:146.
24. Zhang J, Wang ET, Singh RP, Guo C, Shang Y, Chen J, et al. Grape berry surface bacterial microbiome: impact from the varieties and clones in the same vineyard from central China. *J Appl Microbiol*. 2019;126:204–14.
25. Kim CY, Ma J, Lee I. HiFi metagenomic sequencing enables assembly of accurate and complete genomes from human gut microbiota. *Nat Commun*. 2022. <https://doi.org/10.1038/s41467-022-34149-0>.
26. Ding Y, Zhao L, Wang G, Shi Y, Guo G, Liu C, et al. Pacbio sequencing of human fecal samples uncovers the DNA methylation landscape of 22 673 gut phages. *Nucleic Acids Res*. 2023;51:12140–9.
27. Guo L, Wang X, Ayhan DH, Rhaman MS, Yan M, Jiang J, et al. Super pangenome of *Vitis* empowers identification of downy mildew resistance genes for grapevine improvement. *Nat Genet*. 2025;57:741–53.
28. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019. <https://doi.org/10.1186/s13059-019-1891-0>.
29. Liang Z, Duan S, Sheng J, Zhu S, Ni X, Shao J, et al. Whole-genome resequencing of 472 *Vitis* accessions for grapevine diversity and demographic history analyses. *Nat Commun*. 2019. <https://doi.org/10.1038/s41467-019-09135-8>.
30. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol*. 2019;37:540–6.
31. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods*. 2021;18:170–5.
32. Feng X, Cheng H, Portik D, Li H. Metagenome assembly of high-fidelity long reads with hifiasm-meta. *Nat Methods*. 2022;19:671–4.
33. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, et al. Metaflye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods*. 2020;17:1103–10.
34. Benoit G, Raguideau S, James R, Phillippy AM, Chikhi R, Quince C. High-quality metagenome assembly from long accurate reads with metaMDBG. *Nat Biotechnol*. 2024;42:1378–83.
35. Rautiainen M, Nurk S, Walenz BP, Logsdon GA, Porubsky D, Rhie A, et al. Telomere-to-telomere assembly of diploid chromosomes with Verkko. *Nat Biotechnol*. 2023;41:1474–82.
36. Alneberg J, Bjarnason BS, De Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. *Nat Methods*. 2014;11(11):1144–6.
37. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 2016;32:605–7.
38. Kang DD, Li F, Kirtan E, Thomas A, Egan R, An H, et al. MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;2019.
39. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. 2020;36:1925–7.
40. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9.
41. Cantalapiedra CP, Hernandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38:5825–9.
42. Miller SA, Ferreira JP, Lejeune JT. Antimicrobial use and resistance in plant agriculture: a one health perspective. *Agriculture*. 2022. <https://doi.org/10.3390/agriculture12020289>.
43. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, et al. CARD 2023: expanded curation, support for machine learning, and resistance prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res*. 2023;51:D690–9.
44. Singh K, Birnchiikumarasarmaeditors S. Emerging Trends in Plant Pathology. 2020.
45. Narayanan Z, Glick BR, MDPI. Secondary metabolites produced by plant growth-promoting bacterial endophytes. *Microorganisms*. 2022. <https://doi.org/10.3390/microorganisms10102008>.
46. Su P, Kang H, Peng Q, Wicaksono WA, Berg G, Liu Z, et al. Microbiome homeostasis on rice leaves is regulated by a precursor molecule of lignin biosynthesis. *Nat Commun*. 2024. <https://doi.org/10.1038/s41467-023-44335-3>.
47. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. AntiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res*. 2023;51:W46–50.
48. Ranjan A, Rajput VD, Prazdova EV, Gurnani M, Bhardwaj P, Sharma S, et al. Nature's antimicrobial arsenal: non-ribosomal peptides from PGPB for plant pathogen biocontrol. *Fermentation*. 2023. <https://doi.org/10.3390/fermentation9070597>.
49. Ninkuu V, Zhang L, Yan J, Fu Z, Yang T, Zeng H, et al. Biochemistry of terpenes and recent advances in plant protection. *Int J Mol Sci*. 2021. <https://doi.org/10.3390/ijms22115710>.
50. Li Y, Rebuffat S, American Society for Biochemistry and Molecular Biology Inc. The manifold roles of microbial ribosomal peptide-based natural products in physiology and ecology. *J Biol Chem*. 2020. <https://doi.org/10.1074/jbc.REV119.006545>.
51. Salachna P, Mizielińska M, Soból M. Exopolysaccharide gellan gum and derived oligo-gellan enhance growth and antimicrobial activity in *Eucomis* plants. *Polymers*. 2018. <https://doi.org/10.3390/polym10030242>.
52. Castiblanco LF, Sundin GW. New insights on molecular regulation of biofilm formation in plant-associated bacteria. *J Integr Plant Biol*. 2016;58:362–72.
53. Carezzano ME, Alvarez Strazzi FB, Pérez V, Bogino P, Giordano W. Multidisciplinary Digital Publishing Institute (MDPI). Exopolysaccharides synthesized by rhizospheric bacteria: a review focused on their roles in protecting plants against stress. *Appl Microbiol*. 2023. <https://doi.org/10.3390/applmicrobiol3040086>.
54. Zheng J, Ge Q, Yan Y, Zhang X, Huang L, Yin Y. Dbcn3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res*. 2023;51:W115–21.
55. Kurtz ZD, Müller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA. Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol*. 2015. <https://doi.org/10.1371/journal.pcbi.1004226>.
56. Aoki T, Aoki Y, Ishiai S, Otoguro M, Suzuki S. Impact of *Bacillus cereus* NRKT on grape ripe rot disease through resveratrol synthesis in berry skin. *Pest Manag Sci*. 2017;73:174–80.
57. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
58. Klaus M, Grninger M, Royal Society of Chemistry. Engineering strategies for rational polyketide synthase design. *Nat Prod Rep*. 2018. <https://doi.org/10.1039/C8NP00030A>.

59. Valton J, Filisetti L, Fontecave M, Nivière V. A two-component flavin-dependent monooxygenase involved in actinorhodin biosynthesis in *Streptomyces coelicolor*. ArXiv. 2015; Available from: <https://arxiv.org/abs/1501.01469>. Cited 2024 Dec 13.
60. Singh P, Santoni S, Weber A, This P, Péros JP. Understanding the phyllosphere microbiome assemblage in grape species (Vitaceae) with amplicon sequence data structures. Sci Rep. 2019. <https://doi.org/10.1038/s41598-019-50839-0>.
61. Singh P, Santoni S, This P, Péros JP. Genotype-environment interaction shapes the microbial assemblage in grapevine's phyllosphere and carposphere: an NGS approach. Microorganisms. 2018. <https://doi.org/10.3390/microorganisms6040096>.
62. Compant S, Cassan F, Kostić T, Johnson L, Brader G, Trognitz F, et al. Harnessing the plant microbiome for sustainable crop production. Nat Rev Microbiol. 2024. <https://doi.org/10.1038/s41579-024-01079-1>.
63. Musilova L, Ridl J, Polivkova M, Macek T, Uhlik O, MDPI AG. Effects of secondary plant metabolites on microbial populations: changes in community structure and metabolic activity in contaminated environments. Int J Mol Sci. 2016. <https://doi.org/10.3390/ijms17081205>.
64. Rodriguez PA, Rothballer M, Chowdhury SP, Nussbaumer T, Gutjahr C, Falter-Braun P. Systems biology of plant-microbiome interactions. Mol Plant. 2019. <https://doi.org/10.1016/j.molp.2019.05.006>.
65. Pang Z, Chen J, Wang T, Gao C, Li Z, Guo L, et al. Linking plant secondary metabolites and plant microbiomes: a review. Front Plant Sci. 2021. <https://doi.org/10.3389/fpls.2021.621276>.
66. Gand M, Bloemen B, Vanneste K, Roosens NHC, De Keersmaecker SCJ. Comparison of 6 DNA extraction methods for isolation of high yield of high molecular weight DNA suitable for shotgun metagenomics nanopore sequencing to detect bacteria. BMC Genomics. 2023. <https://doi.org/10.1186/s12864-023-09537-5>.
67. Portik DM, Brown CT, Pierce-Ward NT. Evaluation of taxonomic classification and profiling methods for long-read shotgun metagenomic sequencing datasets. BMC Bioinformatics. 2022. <https://doi.org/10.1186/s12859-022-05103-0>.
68. Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK, Nature Research. Plant-microbiome interactions: from community assembly to plant health. Nat Rev Microbiol. 2020. <https://doi.org/10.1038/s41579-020-0412-1>.
69. Vannier N, Agler M, Hacquard S. Microbiota-mediated disease resistance in plants. PLoS Pathog. 2019. <https://doi.org/10.1371/journal.ppat.1007740>.
70. Cardoni M, Fernández-González AJ, Valverde-Corredor A, Fernández-López M, Mercado-Blanco J. Co-occurrence network analysis unveils the actual differential impact on the olive root microbiota by two *Verticillium* wilt biocontrol rhizobacteria. Environ Microbiome. 2023. <https://doi.org/10.1186/s40793-023-00480-2>.
71. Ma T, Chen S, Liu J, Fu P, Wu W, Song S, et al. *Plasmopara viticola* effector PvRXLR111 stabilizes VvWRKY40 to promote virulence. Mol Plant Pathol. 2021;22:231–42.
72. Porebski S, Bailey LG, Baum BR. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Rep. 1997. <https://doi.org/10.1007/BF02772108>.
73. Chen S, Zhou Y, Chen Y, Gu J. Fastp: An ultra-fast all-in-one FASTQ pre-processor. Bioinformatics. Oxford University Press; 2018. p. i884–90.
74. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34:3094–100.
75. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.
76. Dixon P. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science. Opulus Press AB; 2003. p. 927–30.
77. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat Methods. 2023;20:1203–12.
78. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. Iq-tree 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37:1530–4.
79. Letunic I, Bork P. Interactive tree of life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 2024;52:W78–82.
80. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: A versatile open source tool for metagenomics. PeerJ. 2016;2016.
81. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417–9.
82. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014. <https://doi.org/10.1186/s13059-014-0550-8>.
83. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.