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

Jingyun Jin[1], Xiangfeng Wang[1], Xuenan Zhang[1], Junjie Mei[1], Wei Zheng[1], Linling Guo[1], Haisheng Sun[1], Lili Zhang[1], Chonghuai Liu[2], Wenxiu Ye[1]* and Li Guo[1]*

1. Peking University Institute of Advanced Agricultural Sciences, Shandong Laboratory of Advanced Agricultural Sciences in Weifang, Shandong 261325, China.

2. Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou 450004, China.

*Corresponding author: wenxiu.ye@pku-iaas.edu.cn, li.guo@pku-iaas.edu.cn

Abstract

Grapevines are among the most economically important fruit crops, and microbiome profoundly influence their health, yield, and quality. However, mechanistic insights into microbiome-orchestrated grapevine biology remain limited. Here, we conduct large-scale pan-metagenomic and pan-metatranscriptomic analyses of the phyllosphere microbiome from 107 grapevine accessions spanning 34 *Vitis* species. We show that the grapevine core microbiome is dominated by phylum 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 with 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-metatranscriptomic analysis further revealed significant upregulation of eukaryotic microbial genes involved in primary and secondary metabolism. This pan-metagenomic study offers unprecedented insights into the complex structure, diversity and functional roles of grapevine phyllosphere microbiome, and presents valuable genomic and microbial resources for microbiome research and engineering to enhance viticulture productivity and quality.

Introduction

Grapevines are perennial plants belonging to 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, 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 productions and related industries each year [3–7]. Especially, DM is the most destructive among the diseases that caused by the oomycete *Plasmopara viticola*, severely affecting grape production and quality by causing damages on leaves and fruit clusters [8]. With increasing evidence highlight the involvement of grapevine microbiome in DM infection [9,10], it is crucial to investigate 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]. 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

64 microbiome studies primarily rely on 16S rRNA sequencing which yields limited insights
65 into mechanistic understanding of microbiome functions compared to Illumina sequencing.
66 However, Illumina sequencing falls short in assembling the complete microbial genomes
67 given the limitation of short reads, impeding accurate annotations and functional analysis
68 of metagenome. So far, long-read technologies mainly have been used in microbiome
69 studies for human related samples [25,26], but remain underutilized in plant and agricultural
70 metagenomics. In addition, metatranscriptomics, which crucial to understanding functional
71 implications of microbiome by revealing microbiome gene expression changes, has not been
72 widely adopted in plant microbiome research. Most importantly, integrating pan-
73 metagenomics and pan-metatranscriptomics of plants to study the diversity of microbial
74 profiles across plant species and cultivars have not been reported. These limitations hinder
75 the exploration of grapevine microbiome and the application of microbial biocontrol in
76 viticulture.

77 Here, we conducted a comprehensive pan-metagenomic analysis of grapevine leaves
78 from 107 *Vitis* accessions (each accession corresponds to a unique *Vitis* cultivar, except for
79 a few samples) spanning 34 species, utilizing Illumina, PacBio, RNA-seq and DM
80 resistance phenome data. Through systematic profiling of the grapevine phyllosphere
81 microbiome, we identified dominant microbial communities under pan-metagenomic
82 framework. Leveraging PacBio metagenomic reads from the grapevine phyllosphere, we
83 successfully generated 19 high-quality grapevine phyllospheric metagenome-assembled
84 genomes (GPMAGs), and comprehensively annotated genes associated with antibiotic
85 resistance, secondary metabolism, and CAZymes. Based on the phenome data, we revealed

that the microbiome from DM-resistant grapevines exhibit 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 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 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 metagenomics 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:

107 GCA_046255155.1) (**Table S2**), with on average 4.4 million microbial reads remaining per
108 sample. Taxonomic classification of these microbial reads using Kraken2 [28] identified a
109 total of 1,326 genera across the grapevine phyllosphere (**Fig. 2a and Table S3**). Among
110 these, Bacillota (43.6%) and Pseudomonadota (37.9%) dominated the microbiome profile
111 at the phylum level, while *Lactiplantibacillus* (42.1%) and *Xanthomonas* (22.2%) were the
112 two most abundant genera (**Fig. 2b**). Furthermore, kingdom-level profiling demonstrated
113 that bacteria dominated the grapevine phyllosphere microbiome (91.3%), and other
114 kingdoms accounted for less than 10% (**Fig. 2c**). Notably, the non-bacterial kingdoms
115 exhibited dominating phyla, such as Euryarchaeota dominating Archaea, Ascomycota and
116 Euglenozoa dominating Eukaryota, and Cressdnaviricota dominating viruses, reflecting
117 specific phylum-level dominance for the grapevine microbiome. To examine worldwide
118 representativeness of our microbial composition, we analyzed worldwide representative
119 accessions from published Illumina data containing a global grapevine phyllosphere
120 collection [29] using the same bioinformatic pipeline. The phylum-level microbial
121 composition of the global microbiome generally consistent with our Illumina-based
122 composition (**Fig. S1**), demonstrating the worldwide representativeness of our microbial
123 composition. These results highlight the prevalence of Bacillota and Pseudomonadota and
124 their potential roles in the grapevine microbiome.

125 Although several previous studies have profiled grapevine microbiomes [11,14,17–24],
126 mostly using 16S rDNA amplicon sequencing, a comprehensive pan-metagenomic analysis
127 leveraging large-scale Illumina datasets had not been conducted prior to this study, which
128 could provide critical insights into the structure and functional roles of grapevine

129 microbiome. Through our pan-metagenomic analysis of 107 grapevine phyllosphere, we
130 identified seven core genera (present in all accessions), 945 dispensable genera (present in
131 some but not all accessions), and 374 unique genera (exclusive to a single accession) based
132 on presence and absence patterns (**Fig. 2d and Table S3**). Strikingly, although the core
133 microbiome accounted for only 0.6% of the detected genera, it dominated 77.3% of the
134 relative abundance, with Bacillota (55.2%) and Pseudomonadota (44.0%) as the major phyla
135 in the core microbiome (**Fig. 2e**). This result further underscores the central role of the core
136 microbiome (mainly Bacillota and Pseudomonadota) in shaping the grapevine phyllosphere
137 microbiome.

138

139 **Assembly of grapevine phyllospheric metagenome-assembled genomes (GPMAGs)**

140 To assemble high-quality GPMAGs that provide further insights into the potential functions
141 of the grapevine phyllosphere microbiome that potentially influence the host's biology, we
142 generated a total of 103 PacBio datasets with PacBio Sequel IIe system [27] (**Fig. 1d**) for
143 the grapevine accessions used in Illumina-based metagenomic analysis above. The PacBio
144 reads were mapped against the grapevine reference genome to completely exclude host-
145 derived sequences, retaining unmapped reads representing microbial data (**Table S2**). These
146 unmapped reads from all accessions were pooled together and then assembled into contigs
147 by using various assemblers, such as Flye [30], hifiasm [31], hifiasm-meta [32], metaFlye
148 [33], metaMDBG [34], and Verkko [35]. The contigs were subsequently binned using
149 CONCOCT [36], Maxbin 2.0 [37], MetaBAT 2 [38], followed by dereplication with Galah

150 (https://github.com/wwood/galah). This workflow generated a total of 19 non-redundant
151 GPMAGs (**Fig. 3a and Table S4**) that met the medium-quality threshold based on standard
152 MAG quality assessment criteria [34]. Among these, four were classified as complete, three
153 as near-complete, five as high-quality, and seven as medium-quality MAGs (**Fig. 3b and**
154 **Table S4**), highlighting the high quality of the GPMAGs. The median genome size of the
155 GPMAGs were 4.5 Mb for the complete, 4.6 Mb for the near-complete, 4.7 Mb for the high-
156 quality, 4.2 Mb for the medium-quality (**Fig. 3c and Table S4**), approximating typical
157 bacterial genome sizes. Subsequently, GTDB-tk [39] was employed to identify the
158 microbial profile of the GPMAGs. Among the classified GPMAGs, 63.2 % belonged to
159 Pseudomonadota, 5.3% to Bacillota, and 21.1% to Actinomycetota (**Fig. 3d and Table S4**),
160 which generally aligned with the Illumina-based microbial composition (**Fig. 2b**), except
161 for Bacillota. To examine the pan-genomic profile of the GPMAGs among the different
162 PacBio accessions, we further investigated the prevalence of the GPMAG with the different
163 grapevine accessions measured by the PacBio read mapping rates against each GPMAGs.
164 Across the grapevine accessions, *Mangrovibacter phragmitis* exhibited the highest
165 prevalence, followed by *Roseateles* and *Sphingomonas aquatilis* (**Fig. 3e**). Interestingly,
166 these GPMAGs all belonged to Pseudomonadota, further supporting its prevalence across
167 grapevine accessions as indicated by the Illumina-based microbial composition (**Fig. 2b**).
168 Moreover, all GPMAGs represent the highest MAG qualities (complete and near complete)
169 also belong to Pseudomonadota, underscoring a strong association between MAG quality
170 and prevalence.

171 Together, these results highlight high-quality and comprehensive assembly of PacBio-

172 based GPMAGs representing valuable genomic resources for the grapevine phyllospheric
173 microbiome investigation.

174

175 **Annotation of grapevine MAGs reveals genes encoding key biological processes**

176 To delve deeply into functional roles of the GPMAGs in the grapevine microbiome,
177 particularly those derived from unculturable microorganisms, we performed genomic
178 annotation for the GPMAGs using Prokka [40] pipeline. This analysis identified a total of
179 82,884 genes, comprising 81,801 CDSs, 956 tRNA genes, 96 rRNA genes, 15 tmRNA genes,
180 and 16 repeat regions (**Fig. 4a and Table S5**). Functional annotation of the CDSs was further
181 conducted using eggNOG-mapper v2 [41], and annotated the detailed functions of the CDSs,
182 which identified 67,012 CDSs with various biological process terms. In terms of biological
183 function, proteins with unknown function (26.7%) represented the largest category among
184 the CDSs, and the remaining was primarily associated with core biological functions, such
185 as transcription, metabolism, and transport (**Fig. 4b**). Notably, the top 20 most abundant
186 functional terms accounted for 93.9% of total terms, demonstrating a highly focused
187 functional diversity in the GPMAGs that suggests specialized adaptation to the grape
188 phyllosphere niche. Despite biosynthesis of secondary metabolites and microbial
189 metabolism in diverse environments being the predominant biological pathway terms, the
190 top 20 terms represented only 47.9% of total annotations (**Fig. 4c**), demonstrating a
191 functional versatility in the GPMAGs. These results indicate both functional consistency
192 and expansive metabolic potentials in the GPMAGs, likely facilitating phyllosphere

193 colonization and adaptation in grapevines.

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

206 Plant-associated bacterial endophytes protect their plant hosts from external pathogens
207 by producing secondary metabolites (SMs), including antibiotics and toxins [44–46],
208 establishing them as valuable resources for environmentally friendly and pathogen-effective
209 biocontrol agents. Biosynthetic gene clusters (BGCs) responsible for bacterial SM
210 production typically comprise dozens of related genes distributed across extensive genomic
211 regions, necessitating high-quality MAGs with exceptional completeness and continuity for
212 comprehensive exploration. To leverage our high-quality GPMAGs, we employed
213 antiSMASH [47] to identify a total of 96 BGCs from 26 distinct BGC types in the grapevine
214 phyllospheric microbiome (**Fig. 4a, d, S3 and Table S7**). Among these, terpene, RIPP-like,

215 NRPS clusters were the most abundant, commonly associated with antimicrobial activity
216 critical for host defense [48–50]. Other predominant BGCs, including Ni-siderophore,
217 betalactone, and T3PKS highlight the microbiome's ample potentials in SM productions.
218 These candidate BGCs provide valuable resources for discovering novel natural products
219 and elucidating their roles in grapevine biology and medicinal applications.

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

236 To evaluate the functional potentials of each phylum in terms of antibiotic resistance,

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

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

249

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

252 To investigate the role of grapevine microbiome during *P. viticola* infection, we profiled the
253 microbiome composition of the grapevine accessions based on their DM disease phenotypes
254 (susceptible or resistant). Unexpectedly, the microbiome compositions of the susceptible
255 and resistant microbiomes were nearly identical (**Fig. S5a**). Additionally, β-diversity
256 analysis further revealed no overall compositional difference between the two groups (**Fig.**
257 **S5b**), suggesting that microbiome composition unlikely accounts for the differences in DM

258 resistance. Therefore, we hypothesized that interactions within the microbial community
259 might contribute to the observed resistance. To test this hypothesis, we analyzed the
260 complexity of the microbial interaction networks in the susceptible and resistant grape
261 microbiome using Sparse Inverse Covariance estimation for Ecological Association and
262 Statistical Inference (SpiecEasi) [55] analysis. SpiecEasi uses sparse inverse covariance
263 estimation to identify true associations between microorganisms based on correlations
264 derived from the covariance of their relative abundances. Positive correlations indicate
265 cooperative interactions, while negative correlations represent antagonistic interactions.
266 Remarkably, the microbiome associated with resistant plants exhibited a significantly more
267 complex network than the susceptible ones (**Fig. 5a**), with a higher frequency of negative
268 interactions per genus (**Fig. S6**), indicating the involvement of microbial network
269 complexity in modulating DM-resistance, especially the negative interactions play crucial
270 role for the complexity.

271 Notably, the microbiome associated with resistant plants was highly compartmentalized
272 and centralized around a few genera (e.g., *Lactiplantibacillus*, *Liberibacter, Paracidovorax*,
273 and *Xanthomonas*), with relatively much higher abundance (**Fig. 5a**), suggesting the
274 presence of key genera with pivotal roles in network dynamics and resistance modulation.
275 To identify these key genera, we first profiled the top 100 most abundant genera in the
276 resistant microbiome based on relative abundance and microbial interaction metrics. These
277 profiles were then systematically compared with those from the susceptible microbiome to
278 evaluate their respective contributions to the network complexity. Among them,
279 *Liberibacter* and *Pseudomonas* emerged as the most influential genera, with significantly

280 increased interaction metrics compared to those of the susceptible microbiome (**Fig. 5b**).
281 Intriguingly, all these key genera belong to the core microbiome (**Table S3**), highlighting
282 the importance of the core microbiome in modulating the complexity. Additionally, genera
283 such as *Bacillus* exhibit increased abundance and interaction metrics compared to the
284 susceptible microbiome, indicating contribution to the complexity via abundance and
285 microbial interaction. Despite relatively low abundances in the resistant microbiome, genera
286 such as Bacteroides, Agathobacter, Bifidobacterium, Anaerobutyricum, and Neisseria
287 exhibit the highest interactions (8-10 interactions/genus) in the resistant microbiome,
288 demonstrating the important role of low abundant genera in forming the complexity. To
289 evaluate phylum-level contributions to the increased complexity, we compared the
290 proportions of microbial interaction between the resistant and susceptible microbiomes for
291 each phylum. As a result, Bacillota was identified as the primary contributors exhibiting the
292 most significantly increased proportion in the resistant microbiome (33.5%) compared to
293 that of the susceptible one (23.9%), while other phyla show minimal fluctuation in the
294 change (**Fig. S7**). Furthermore, we analyzed Spearman correlations between Bacillota and
295 other major phyla in the resistant microbiomes based on their abundances, identifying those
296 with the strongest interactions with this key phylum. The correlation analysis revealed that
297 Euglenozoa, one of the major eukaryotes in the grapevine microbiome (**Fig. 2c**) and
298 Cyanobacteriota exhibits the strongest positive interaction with Bacillota, while the other
299 phyla display no significant positive interactive dynamics (**Fig. S8**), indicating that Bacillota
300 positively interacts with these phyla in the resistant microbiome. These findings suggest that
301 not all genera contribute equally to the modulation of the network complexity; rather,

302 specific key genera play a primary role along with specific interactions.

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

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

320

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

323 Given that network complexity is primarily driven by Bacillota, which exhibits the strong
324 positive abundance correlation with one of the dominating Eukaryotes-Euglenozoa, we
325 hypothesize that key bacterial phyla involve eukaryotic transcriptomes in the grapevine
326 microbiome, potentially underlying defensive responses against *P. viticola* infection. To
327 examine this, we conducted pan-metatranscriptomic sequencing and analysis to investigate
328 the transcriptional regulation of the grapevine phyllospheric microbiome in response to the
329 infection. In brief, we collected samples from both infected (one day post-inoculation) and
330 uninfected leaves from 103 of the grapevine accessions and subjected them to RNA-seq [27]
331 (**Fig. 1d**). The RNA-seq reads were mapped to the grapevine and *P. viticola* reference
332 genomes (NCBI accession: GCA_001695595.3), obtaining the unmapped reads that
333 represent eukaryotic organisms other than grapevine and *P. viticola* (**Table S2**). The
334 unmapped reads were then *de novo* assembled using Trinity [57], resulting in 192,675
335 eukaryotic transcripts. Differentially expressed transcript (DET) analysis identified 40,546
336 significantly up-regulated and 5,081 down-regulated transcripts (**Fig. 6a**), indicating a
337 considerable overall upregulation of the microbial transcripts upon the infection. Moreover,
338 KEGG and GO enrichment revealed that the up-regulated genes were primarily involved in
339 primary and secondary metabolisms (**Fig. 6b and c**), with secondary metabolism being
340 particularly prominent in the microbiome's response to the infection. Consistently, Gene Set
341 Enrichment Analysis (GSEA) further confirmed significant enrichment of secondary
342 metabolism-related gene sets among the upregulated transcripts (**Fig. 6d**). Notably, these
343 upregulated transcripts encoded the key proteins of BGCs, including ketoacyl reductase (KR)
344 domains and flavin mononucleotide-dependent dehydrogenases, both of which are integral

345 to the biosynthesis of polyketides and antibiotics [58,59]. Together, these findings
346 demonstrate that primary and secondary metabolisms in eukaryotic microorganisms are
347 significantly induced during *P. viticola* infection, suggesting a potential role for the
348 grapevine phyllospheric microbiome in defending against pathogen invasion through the
349 competitive exclusion and SM productions that associated with key bacterial genera.

350

351 **Discussion**

352 Microbiome have profound impact on physiology and immunity of host organisms ranging
353 from human, animals and plants. Mounting evidence underscores the critical roles of the
354 microbiome in the health, yield, and quality [4,11–13] of grapevine which is one of the most
355 economically important fruit crops cultivated worldwide for production of fresh fruits, wine,
356 juice and raisin etc. Despite a huge body of knowledge gained in microbiome research, pan-
357 microbiome diversity across host genera is poorly understood, lacking a large-scale pan-
358 metagenomic approach, especially from long-read sequencing. In this study, combining
359 short, long-read sequencing and RNA-seq, we conducted a pan-metagenomic and pan-
360 metatranscriptomic study of 107 accessions from 34 *Vitis* species to understand the structure,
361 diversity and functional roles of pan-microbiome in grapevine phyllosphere, a key
362 battleground for host-microbe interactions. We identified the core microbiome dominated
363 by Bacillota and Pseudomonadota. Leveraging PacBio sequencing of grapevine
364 phyllosphere, we successfully assembled 19 high-quality GPMAGs including four near-
365 complete MAGs, with detailed functional annotations, representing the first assembly of

366 MAGs derived from plant PacBio datasets. Furthermore, pan-metagenomic and pan-
367 metatranscriptomic analyses elucidated the underlying defense mechanisms of grape
368 microbiome against *P. viticola* infection. Our findings demonstrate the profound influence
369 of the microbiome on grapevine biology, particularly involvement in the resistance of *P.*
370 *viticola* infection, and provide a foundation and valuable genomic and microbial resources
371 for developing microbiome-based biocontrol strategies in viticulture.

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

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

401 It is widely recognized that through interactions microorganisms, plants modulate their
402 immunity to combat pathogen invasion [68,69]. This mechanism primarily relies on
403 microbiome-derived SM production and competitive exclusion to protect host plants from
404 various pathogens. Consistent with this concept, our findings demonstrate the pivotal roles
405 of bacteriome and eukaryome, of which bacteriome re-orchestrates microbial community,
406 especially facilitates interactions with microorganisms to involve their SM production and
407 competitive exclusion in response to *P. viticola* infection (**Fig. 7**). Intriguingly, we observed
408 a significant increase in network complexity, driven by the contributions of key genera from
409 bacteriome (**Fig. 5 and Fig. S6, 7**), which involves the defense mechanism against *P.*

410 *viticola* infection in grapevines through interacting with eukaryome. In olive, *Pseudomonas*
411 *simiae* PICF7, a biological control agent for Verticillium wilt, was reported to significantly
412 increase microbial network complexity upon the infection, resulting in a more stable
413 network that plays a crucial role in defense [70]. Thus, it is likely that grapevines employ a
414 similar strategy to modulate their microbial network, optimizing SM production and
415 competitive exclusion capacities of eukaryome to effectively combat *P. viticola* infection.
416 Additionally, the profiling of key genera in the DM-resistant microbiome provides a
417 valuable resource for exploring biocontrol microorganisms against *P. viticola* infection
418 exemplified by the experimental validation of two *B. cereus* strains, laying the foundation
419 for environmentally friendly and sustainable biocontrol strategies in viticulture.

420

421 **Materials and methods**

422 **Collection of grapevine leaves for metagenome sequencing**

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

428

429 **Downy mildew infection assay**

430 DM infection was performed as described previously [71]. Each grapevine leaf disc (8 mm

431 in diameter) was inoculated with 20 μL of *P. viticola* suspension (1×10^5 sporangia/mL) on
432 the abaxial surface and incubated at 18 ± 2°C with 80 ± 10% relative humidity in a growth
433 chamber under a 12 h light (60 $\mu mol \cdot m^{-2} \cdot s^{-1}$)/12 h dark cycle. Leaf discs were collected 24
434 hours post-inoculation for transcriptome analysis. At least three biological replicates were
435 conducted, with each replicate comprising five leaf discs from different grapevines.

436

437 **DNA and RNA extraction**

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

453

454 **Metagenome sequencing**

455 For Illumina library preparation, 15 μg of the genomic DNA was sheared using g-Tubes
456 (Covaris, Woburn, USA) and concentrated with AMPure PB magnetic beads (Pacific
457 Biosciences, Menlo Park, USA). The library construction was performed with 1 μg of the
458 DNA template following TruSeq DNA Sample Preparation Guide (Illumina, San Diego,
459 USA). Paired-end sequencing was conducted on a HiSeq X Ten platform (Biomarker
460 Technologies, Qingdao, China), generating 150 bp Illumina reads. For PacBio sequencing,
461 15 μg of the purified HMW genomic DNA was used to construct a standard SMRTbell
462 library with SMRT Express Template Prep Kit 2.0 (Pacific Biosciences, Menlo Park, USA).
463 Circular consensus sequencing (CCS) producing PacBio reads was performed on a PacBio
464 Sequel IIe system at Biomarker Technologies (Qingdao, China). To prepare RNA
465 sequencing libraries, 1-3 μg of the total RNA was used as input for library construction with
466 VAHTS Universal V6 RNA-seq Library Prep Kit for Illumina® (Vazyme, Nanjing, China)
467 according to the manufacturer's instructions. The RNA concentration was initially
468 quantified with the Qubit® RNA Assay Kit on a Qubit® 3.0 (Thermo Fisher Scientific,
469 Waltham, USA) and diluted to 1 ng/μl. Insert size was measured with an Agilent
470 Bioanalyzer 2100 (Agilent Technologies, Santa Clara, USA), and final library concentration
471 (>10 nM) was verified with a Bio-RAD CFX 96 quantitative PCR system (Bio-Rad,
472 Hercules, USA). Paired-end sequencing was then performed on a HiSeq X Ten platform
473 (Biomarker Technologies, Qingdao, China), generating 150 bp RNA-seq reads.

474

475 **Quality control and removal of host reads**

476 To ensure the quality of the short reads, quality control was performed on the Illumina and
477 RNA-seq reads using fastp [73] (v0.23.4) with default parameters, resulting in the quality-
478 controlled reads. To completely filter out reads from the grapevine hosts, the Illumina,
479 PacBio, and RNA-seq reads were mapped to the complete reference genome of *Vitis vinifera*
480 [27] (VHP-T2T.hap1, NCBI accession: GCA_046255155.1) using Minimap2 [74] (v2.28)
481 with parameters (Illumina, RNA-seq: -ax sr, PacBio: -ax map-pb). The mapped results were
482 further filtered using SAMtools [75] (v1.21) with parameters (Illumina, RNA-seq: view -u
483 -f 12 -F 256, PacBio: view -f 4 -F 256) to retain microbial reads. Additionally, the RNA-seq
484 reads were further filtered by mapping to *P. viticola* reference genome INRA_Pvit_2 (NCBI
485 accession: GCA_001695595.3) with Minimap2 and SAMtools to exclude *P. viticola* reads.

486

487 **Taxonomic classification**

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

492

493 **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 vegan [76] (v2.6-8) package in R. All statistical analysis and visualization 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] (v 2.17) 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 phylogenetic tree using IQ-TREE2 [78] (v2.1.4_beta) with

515 default parameters, and visualized by iTOL [79] (https://itol.embl.de/).

516

517 **Functional annotation of the GPMAGs**

518 To analyze different gene types from the GPMAGs, Prokka [40] (v1.14.6) was used for gene
519 type prediction with default parameters. The predicted CDSs were further applied for
520 functional annotation using eggNOG-mapper [80] (2.1.12) with parameters (-m diamond -
521 -evalue 1e-5 --query_cover 50 --subject_cover 50). To annotate AMR genes, BGCs, and
522 CAZyme genes of the GPMAGs, RGI [43] (v6.0.3) (default parameters), antiSMASH [47]
523 (v7.1.0) (parameters: --genefinding-tool prodigal --taxon bacteria –fullhmmer --cb-
524 knownclusters --cb-subclusters --cc-mibig --asf --rre --pfam2go), and dbCAN [54] (v4.1.4)
525 (default parameters) were employed, respectively.

526

527 **Construction of microbial interaction network**

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

531

532 **Isolation of grapevine phyllosphere microorganisms**

533 To isolate microorganisms from grapevine leaves, we collected 10 leaves from each of the
534 following cultivars: *V. vinifera* cv. Chardonnay, *V. vinifera* cv. Muscat Hamburg, *V. vinifera*

535 x *V. labrusca* cv. Kyoho, and *V. vinifera* x *V. labrusca* cv. Vidal Blanc at the vineyards. After
536 rinsing with sterile water, the leaves were homogenized in 1x PBS buffer using an electric
537 mixer to obtain a microbial suspension. 2 ml of the suspension were inoculated onto Nutrient
538 Agar medium (Coolaber Technology Co., Ltd., Beijing, China) and incubated at 30°C for
539 24 hours. Single colonies were isolated and further purified by streaking on Nutrient Agar
540 medium. Species identification was performed by 16S rRNA sequencing (Tsingke
541 Biotechnology Co., Ltd., Beijing, China).

542

543 **Experimental validation of biocontrol activity**

544 To validate the efficacy of isolated microorganisms against *P. viticola* infection, the *B.*
545 *cereus* strains were cultured in Nutrient Broth medium (Coolaber Technology Co., Ltd.,
546 Beijing, China) at 30°C and 220 rpm for 12 hours. The cultures were diluted to OD of 0.02
547 at 600 nm using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham,
548 USA). The diluted suspensions were sprayed onto 1 cm grape leaf discs and incubated at
549 30°C and 220 rpm for 24 hours. Subsequently, *P. viticola* was co-inoculated onto the discs
550 using the method described above and incubated at 19°C for 6 days. Disease severity was
551 assessed 6 days post-inoculation using the following grading scale: Grade 0, no lesions (0%
552 of leaf disc area affected); Grade 1, lesions covering less than 10% of the leaf disc area;
553 Grade 2, lesions covering 10-30% of the leaf disc area; Grade 3, lesions covering 31-60%
554 of the leaf disc area; Grade 4, lesions covering 61-80% of the leaf disc area; and Grade 5,
555 lesions covering more than 80% of the leaf disc area. The disease index (DI) was calculated

556 as follows:

557 $$DI = \frac{\sum(s_i\, n_{i2})}{N_2 5}\ 100$$

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

561

562 **Transcriptomic analysis of *P. viticola* infection**

563 Transcriptomic profiling in response to *P. viticola* infection was conducted using the filtered
564 RNA-seq reads. First, *De novo* assembly of transcripts was performed with Trinity [57]
565 (v2.15.2) using default parameters. The generated transcripts were dereplicated using
566 VSEARCH [81] (2.29.0) with parameters (--id 0.95), selecting the longest transcripts as the
567 representatives of duplicates. Transcript expression was quantified using Salmon [82]
568 (v1.10.3) with default parameters and the filtered RNA-seq reads. Differential expression
569 analysis was performed with the DESeq2 [83] (v1.45) package in R. Subsequently, the
570 significantly up-regulated genes were further annotated with eggNOG-mapper, followed by
571 KEGG, GO, and GSEA enrichment analyses using AnnotationForge (v1.47.1),
572 AnnotationDbi (v1.67.0), ClusterProfiler [84] (v4.13.3) and enrichplot (v1.25.2) packages
573 in R with default parameters.

574

575 **Acknowledgments**

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

Author 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) and the Taishan Scholars Program of Shandong Province. 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).

References

1. Food and Agriculture Organization of the United Nations. FAOSTAT. Rome; 2024.

2. Zhou DD, Li J, Xiong RG, Saimaiti A, Huang SY, Wu SX, et al. Bioactive Compounds, Health Benefits and Food Applications of Grape. Foods. MDPI; 2022.

596 3. Buonassisi D, Colombo M, Migliaro D, Dolzani C, Peressotti E, Mizzotti C, et al.
597 Breeding for grapevine downy mildew resistance: a review of "omics" approaches.
598 Euphytica. Springer Netherlands; 2017.
599 4. Bettenfeld P, Cadena i Canals J, Jacquens L, Fernandez O, Fontaine F, van Schaik E, et
600 al. The microbiota of the grapevine holobiont: A key component of plant health. J Adv Res.
601 Elsevier B.V.; 2022. p. 1–15.
602 5. Campos G, Chialva C, Miras S, Lijavetzky D. New Technologies and Strategies for
603 Grapevine Breeding Through Genetic Transformation. Front Plant Sci. Frontiers Media S.A.;
604 2021.
605 6. Cantu D, Walker MA. Compendium of Plant Genomes The Grape Genome [Internet].
606 Available from: http://www.springer.com/series/11805
607 7. Magon G, De Rosa V, Martina M, Falchi R, Acquadro A, Barcaccia G, et al. Boosting
608 grapevine breeding for climate-smart viticulture: from genetic resources to predictive
609 genomics. Front Plant Sci. Frontiers Media SA; 2023.
610 8. Peng J, Wang X, Wang H, Li X, Zhang Q, Wang M, et al. Advances in understanding
611 grapevine downy mildew: From pathogen infection to disease management. Mol Plant
612 Pathol. John Wiley and Sons Inc; 2024.
613 9. Fournier P, Pellan L, Jaswa A, Cambon MC, Chataigner A, Bonnard O, et al. Revealing
614 microbial consortia that interfere with grapevine downy mildew through microbiome
615 epidemiology. Environ Microbiome. 2025;20.
616 10. Duret M, Wallner A, Besaury L, Aziz A. Diversity and functional features of the root-
617 associated bacteriome are dependent on grapevine susceptibility to *Plasmopara viticola*.

618 Environ Microbiome. 2025;20.

619 11. Bokulich NA, Thorngate JH, Richardson PM, Mills DA. Microbial biogeography of
620 wine grapes is conditioned by cultivar, vintage, and climate. Proceedings of the National
621 Academy of Sciences. 2014;111.

622 12. Ding Y, Wei R, Wang L, Wang W, Wang H, Li H. Exploring the ecological characteristics
623 of natural microbial communities along the continuum from grape berries to winemaking.
624 Food Research International. 2023;167.

625 13. Bokulich NA, Collins TS, Masarweh C, Allen G, Heymann H, Ebeler SE, et al.
626 Associations among wine grape microbiome, metabolome, and fermentation behavior
627 suggest microbial contribution to regional wine characteristics. mBio. 2016;7.

628 14. Coller E, Cestaro A, Zanzotti R, Bertoldi D, Pindo M, Larger S, et al. Microbiome of
629 vineyard soils is shaped by geography and management. Microbiome. 2019;7:140.

630 15. Cui S, Zhou L, Fang Q, Xiao H, Jin D, Liu Y. Growth period and variety together drive
631 the succession of phyllosphere microbial communities of grapevine. Science of the Total
632 Environment. 2024;950.

633 16. Marasco R, Rolli E, Fusi M, Michoud G, Daffonchio D. Grapevine rootstocks shape
634 underground bacterial microbiome and networking but not potential functionality.
635 Microbiome. 2018;6.

636 17. Awad M, Giannopoulos G, Mylona P V., Polidoros AN. Comparative Analysis of
637 Grapevine Epiphytic Microbiomes among Different Varieties, Tissues, and Developmental
638 Stages in the Same Terroir. Applied Sciences. 2022;13:102.

639 18. Wei T-L, Zheng Y-P, Wang Z-H, Shang Y-X, Pei M-S, Liu H-N, et al. Comparative

640 microbiome analysis reveals the variation in microbial communities between 'Kyoho' grape
641 and its bud mutant variety. PLoS One. 2023;18:e0290853.
642 19. Vitulo N, Lemos WJF, Calgaro M, Confalone M, Felis GE, Zapparoli G, et al. Bark and
643 Grape Microbiome of Vitis vinifera: Influence of Geographic Patterns and Agronomic
644 Management on Bacterial Diversity. Front Microbiol. 2019;9.
645 20. Bao L, Sun B, Wei Y, Xu N, Zhang S, Gu L, et al. Grape Cultivar Features Differentiate
646 the Grape Rhizosphere Microbiota. Plants. 2022;11:1111.
647 21. Awad M, Giannopoulos G, Mylona P V., Polidoros AN. Genotype May Influence
648 Bacterial Diversity in Bark and Bud of *Vitis vinifera* Cultivars Grown under the Same
649 Environment. Applied Sciences. 2020;10:8405.
650 22. Swift JF, Migicovsky Z, Trello GE, Miller AJ. Grapevine bacterial communities display
651 compartment-specific dynamics over space and time within the Central Valley of California.
652 Environ Microbiome. 2023;18:84.
653 23. Fernandes P, Afonso IM, Pereira J, Rocha R, Rodrigues AS. Epiphitic Microbiome of
654 Alvarinho Wine Grapes from Different Geographic Regions in Portugal. Biology (Basel).
655 2023;12:146.
656 24. Zhang J, Wang ET, Singh RP, Guo C, Shang Y, Chen J, et al. Grape berry surface
657 bacterial microbiome: impact from the varieties and clones in the same vineyard from
658 central China. J Appl Microbiol. 2019;126:204–14.
659 25. Kim CY, Ma J, Lee I. HiFi metagenomic sequencing enables assembly of accurate and
660 complete genomes from human gut microbiota. Nat Commun. 2022;13.
661 26. Ding Y, Zhao L, Wang G, Shi Y, Guo G, Liu C, et al. PacBio sequencing of human fecal

662 samples uncovers the DNA methylation landscape of 22 673 gut phages. Nucleic Acids Res.
663 2023;51:12140–9.
664 27. Guo L, Wang X, Ayhan DH, Rhaman MS, Yan M, Jiang J, et al. Super pangenome of
665 *Vitis* empowers identification of downy mildew resistance genes for grapevine improvement.
666 Nat Genet. 2025;57:741–53.
667 28. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome
668 Biol. 2019;20.
669 29. Liang Z, Duan S, Sheng J, Zhu S, Ni X, Shao J, et al. Whole-genome resequencing of
670 472 *Vitis* accessions for grapevine diversity and demographic history analyses. Nat
671 Commun. 2019;10.
672 30. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using
673 repeat graphs. Nat Biotechnol. 2019;37:540–6.
674 31. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo
675 assembly using phased assembly graphs with hifiasm. Nat Methods. 2021;18:170–5.
676 32. Feng X, Cheng H, Portik D, Li H. Metagenome assembly of high-fidelity long reads
677 with hifiasm-meta. Nat Methods. 2022;19:671–4.
678 33. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, et al.
679 metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods.
680 2020;17:1103–10.
681 34. Benoit G, Raguideau S, James R, Phillippy AM, Chikhi R, Quince C. High-quality
682 metagenome assembly from long accurate reads with metaMDBG. Nat Biotechnol.
683 2024;42:1378–83.

684 35. Rautiainen M, Nurk S, Walenz BP, Logsdon GA, Porubsky D, Rhie A, et al. Telomere-
685 to-telomere assembly of diploid chromosomes with Verkko. Nat Biotechnol. 2023;41:1474–
686 82.
687 36. Alneberg J, Bjarnason BS, De Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning
688 metagenomic contigs by coverage and composition. Nat Methods. 2014;11:1144–6.
689 37. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: An automated binning algorithm to
690 recover genomes from multiple metagenomic datasets. Bioinformatics. 2016;32:605–7.
691 38. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: An adaptive
692 binning algorithm for robust and efficient genome reconstruction from metagenome
693 assemblies. PeerJ. 2019;2019.
694 39. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: A toolkit to classify
695 genomes with the genome taxonomy database. Bioinformatics. 2020;36:1925–7.
696 40. Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics.
697 2014;30:2068–9.
698 41. Cantalapiedra CP, Herņandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-
699 mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the
700 Metagenomic Scale. Mol Biol Evol. 2021;38:5825–9.
701 42. Miller SA, Ferreira JP, Lejeune JT. Antimicrobial Use and Resistance in Plant
702 Agriculture: A One Health Perspective. Agriculture (Switzerland). 2022;12.
703 43. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, et al. CARD
704 2023: expanded curation, support for machine learning, and resistome prediction at the
705 Comprehensive Antibiotic Resistance Database. Nucleic Acids Res. 2023;51:D690–9.

44. Singh K, Birinchiikumarrsarmaaeditors S. Emerging Trends in Plant Pathology. 2020.

45. Narayanan Z, Glick BR. Secondary Metabolites Produced by Plant Growth-Promoting Bacterial Endophytes. Microorganisms. MDPI; 2022.

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;15.

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, Prazdnova EV, Gurnani M, Bhardwaj P, Sharma S, et al. Nature's Antimicrobial Arsenal: Non-Ribosomal Peptides from PGPB for Plant Pathogen Biocontrol. Fermentation. Multidisciplinary Digital Publishing Institute (MDPI); 2023.

49. Ninkuu V, Zhang L, Yan J, Fu Z, Yang T, Zeng H. Biochemistry of terpenes and recent advances in plant protection. Int J Mol Sci. MDPI; 2021.

50. Li Y, Rebuffat S. The manifold roles of microbial ribosomal peptide-based natural products in physiology and ecology. Journal of Biological Chemistry. American Society for Biochemistry and Molecular Biology Inc.; 2020. p. 34–54.

51. Salachna P, Mizielińska M, Soból M. Exopolysaccharide gellan gum and derived oligo-gellan enhance growth and antimicrobial activity in *eucomis* plants. Polymers (Basel). 2018;10.

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. Exopolysaccharides Synthesized by Rhizospheric Bacteria: A Review Focused on Their Roles in Protecting Plants against Stress. Appl Microbiol. Multidisciplinary Digital Publishing Institute (MDPI); 2023. p. 1249–61.

54. Zheng J, Ge Q, Yan Y, Zhang X, Huang L, Yin Y. DbCAN3: 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;11.

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, Grininger M. Engineering strategies for rational polyketide synthase design. Nat Prod Rep. Royal Society of Chemistry; 2018. p. 1070–81.

59. Valton J, Filisetti L, Fontecave M, Nivière V. A two-component flavin-dependent monooxygenase involved in actinorhodin biosynthesis in *Streptomyces coelicolor*. ArXiv [Internet]. 2015 [cited 2024 Dec 13]; Available from: https://arxiv.org/abs/1501.01469

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.

750 Sci Rep. 2019;9.

751 61. Singh P, Santoni S, This P, Péros JP. Genotype-environment interaction shapes the
752 microbial assemblage in grapevine's phyllosphere and carposphere: An ngs approach.
753 Microorganisms. 2018;6.

754 62. Compant S, Cassan F, Kostić T, Johnson L, Brader G, Trognitz F, et al. Harnessing the
755 plant microbiome for sustainable crop production. Nat Rev Microbiol. Nature Research;
756 2024.

757 63. Musilova L, Ridl J, Polivkova M, Macek T, Uhlik O. Effects of secondary plant
758 metabolites on microbial populations: Changes in community structure and metabolic
759 activity in contaminated environments. Int J Mol Sci. MDPI AG; 2016.

760 64. Rodriguez PA, Rothballer M, Chowdhury SP, Nussbaumer T, Gutjahr C, Falter-Braun P.
761 Systems Biology of Plant-Microbiome Interactions. Mol Plant. Cell Press; 2019. p. 804–21.

762 65. Pang Z, Chen J, Wang T, Gao C, Li Z, Guo L, et al. Linking Plant Secondary Metabolites
763 and Plant Microbiomes: A Review. Front Plant Sci. Frontiers Media S.A.; 2021.

764 66. Gand M, Bloemen B, Vanneste K, Roosens NHC, De Keersmaecker SCJ. Comparison
765 of 6 DNA extraction methods for isolation of high yield of high molecular weight DNA
766 suitable for shotgun metagenomics Nanopore sequencing to detect bacteria. BMC
767 Genomics. 2023;24.

768 67. Portik DM, Brown CT, Pierce-Ward NT. Evaluation of taxonomic classification and
769 profiling methods for long-read shotgun metagenomic sequencing datasets. BMC
770 Bioinformatics. 2022;23.

771 68. Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. Plant–microbiome interactions: from

community assembly to plant health. Nat Rev Microbiol. Nature Research; 2020. p. 607–21.

69. Vannier N, Agler M, Hacquard S. Microbiota-mediated disease resistance in plants. PLoS Pathog. 2019;15.

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;18.

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 Report.

73. Chen S, Zhou Y, Chen Y, Gu J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. 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. Cantalapiedra CP, Herṇandez-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.

81. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: A versatile open source tool for metagenomics. PeerJ. 2016;2016.

82. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9.

83. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15.

84. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: An R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.

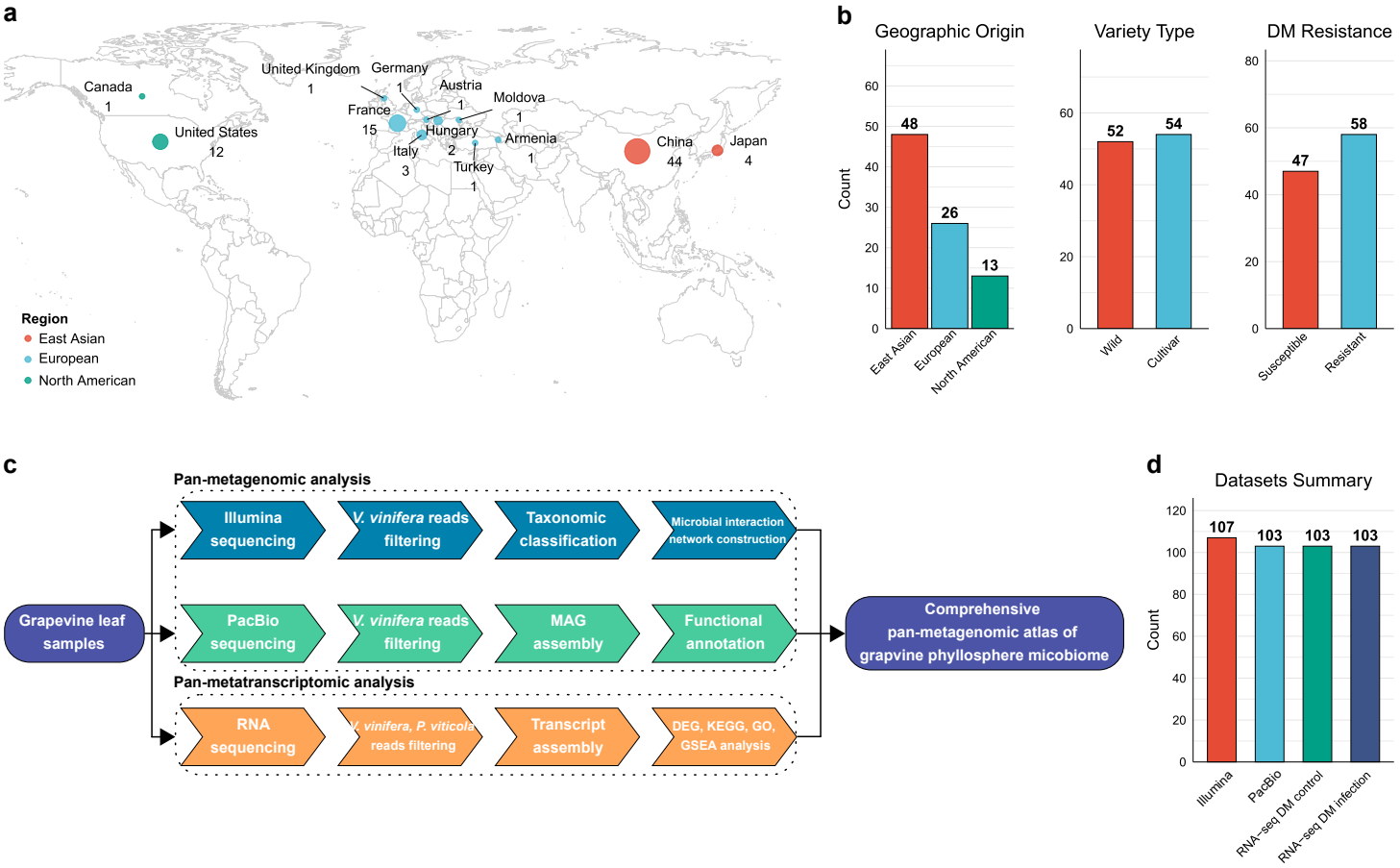

Fig. 1 Research design for grapevine phyllosphere pan-metagenomics and pan-metatranscriptomics. a, Geographical origins of the sampled grapevine accessions. Numbers of samples were numerically labeled. **b**, Classification statistics of grapevine accessions by geographical origin, cultivar type, and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

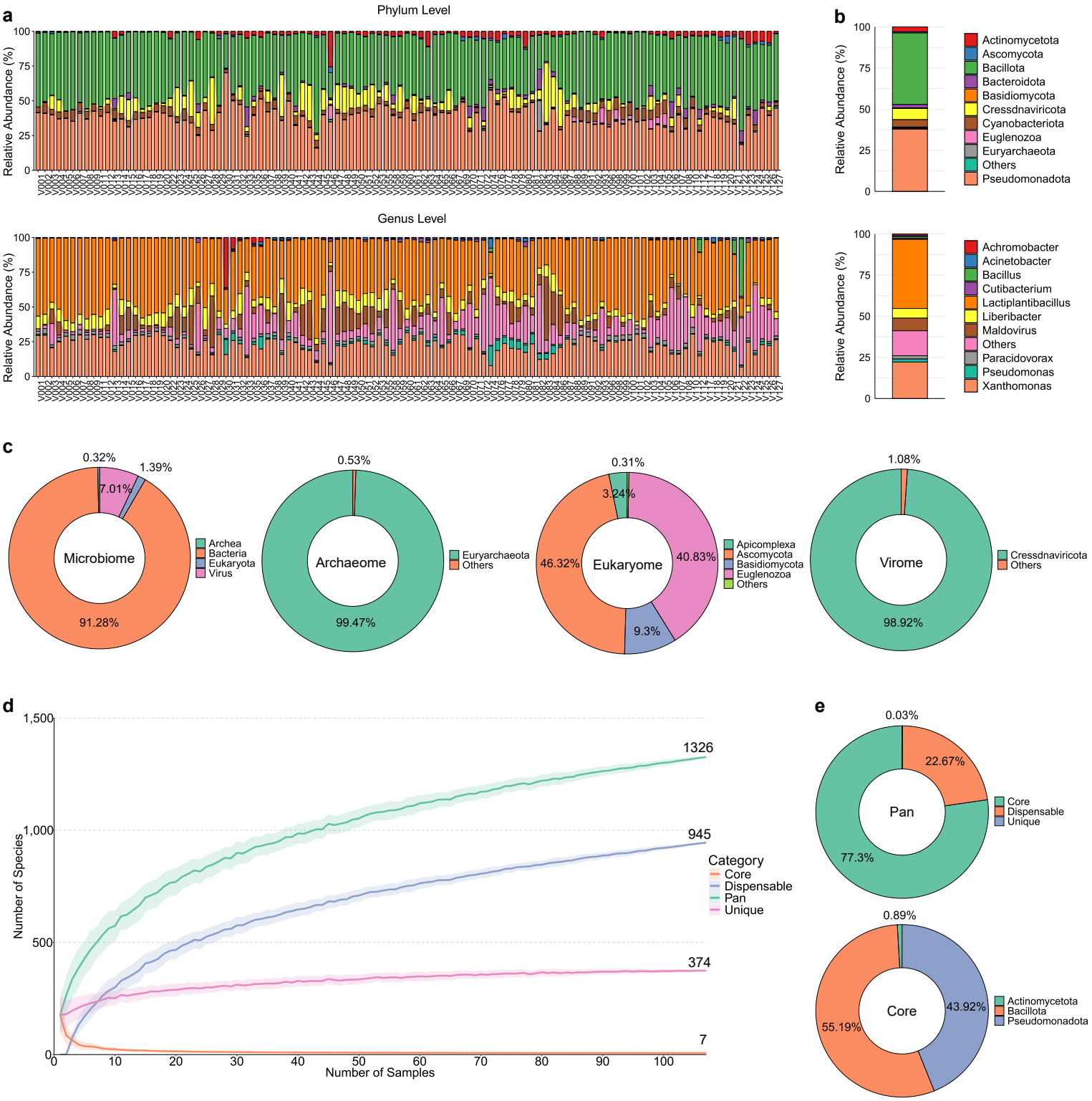

Fig. 2 Pan-metagenomic landscape of the grapevine phyllospheric microbiome. **a**, Relative abundances of the 10 most abundant phyla (top) and genera (bottom) across the grapevine accessions. **b**, Mean relative abundance of the phyla and genera shown in (**a**). **c**, Compositional profiles of kingdoms 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).

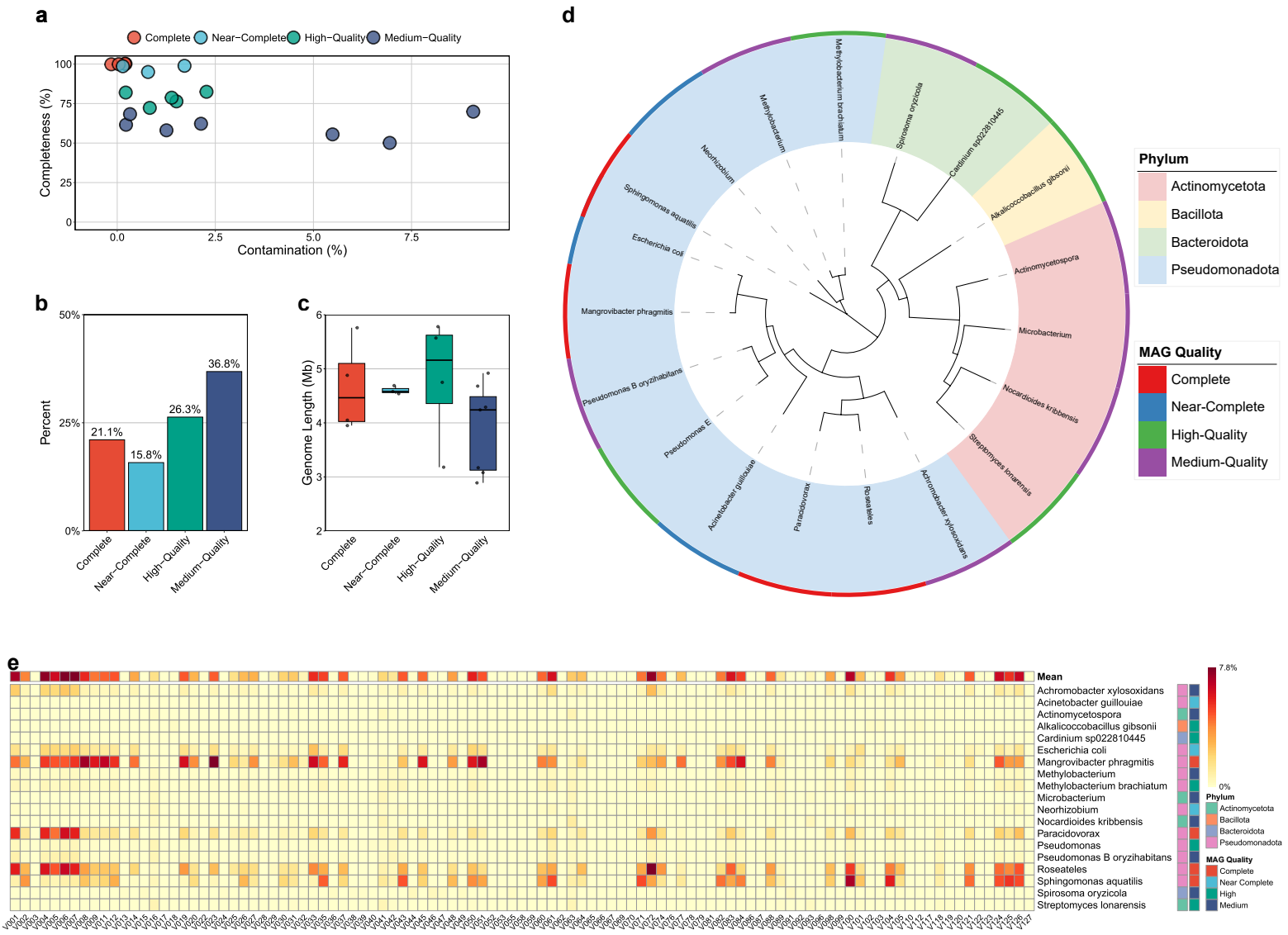

Fig. 3 Assembly and genomic profile of the grapevine phyllospheric metagenome-assembled genomes (GPMAGs). **a**, Quality distribution of the GPMAGs in terms of MAG completeness and contamination. **b**, Composition of the GPMAG quality categories. **c**, Genomic size distribution of the GPMAGs. **d**, 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.

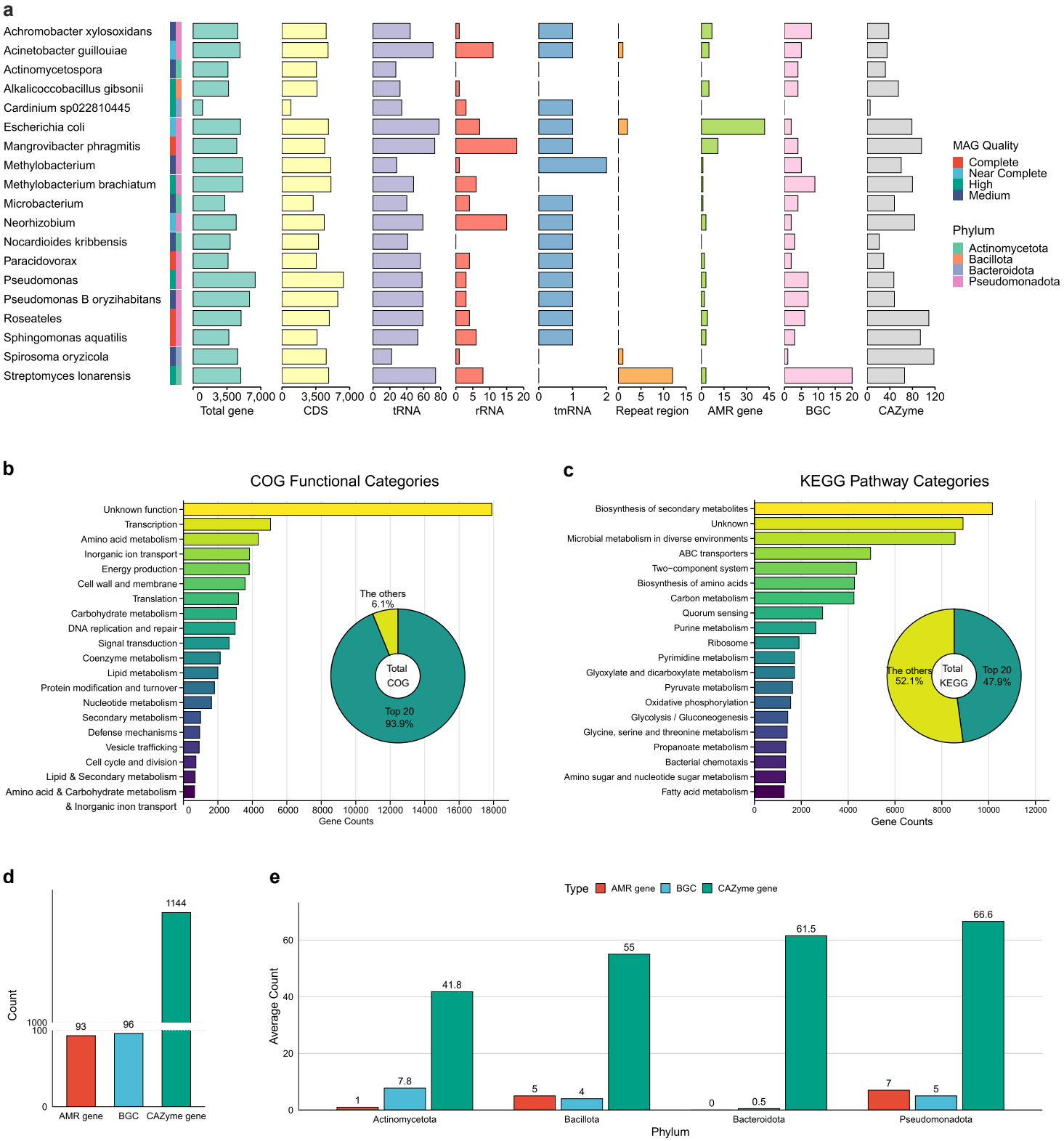

Fig. 4 Functional annotation of the GPMAGs. **a**, Gene type distributions in the GPMAGs. **b-c**, Top 20 COG (**b**) and KEGG pathway (**c**) profiles of the GPMAGs. Donut plots displaying proportions of top 20 terms within total terms. **d**, Statistical summary of AMR, BGC, and CAZyme-associated genes.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

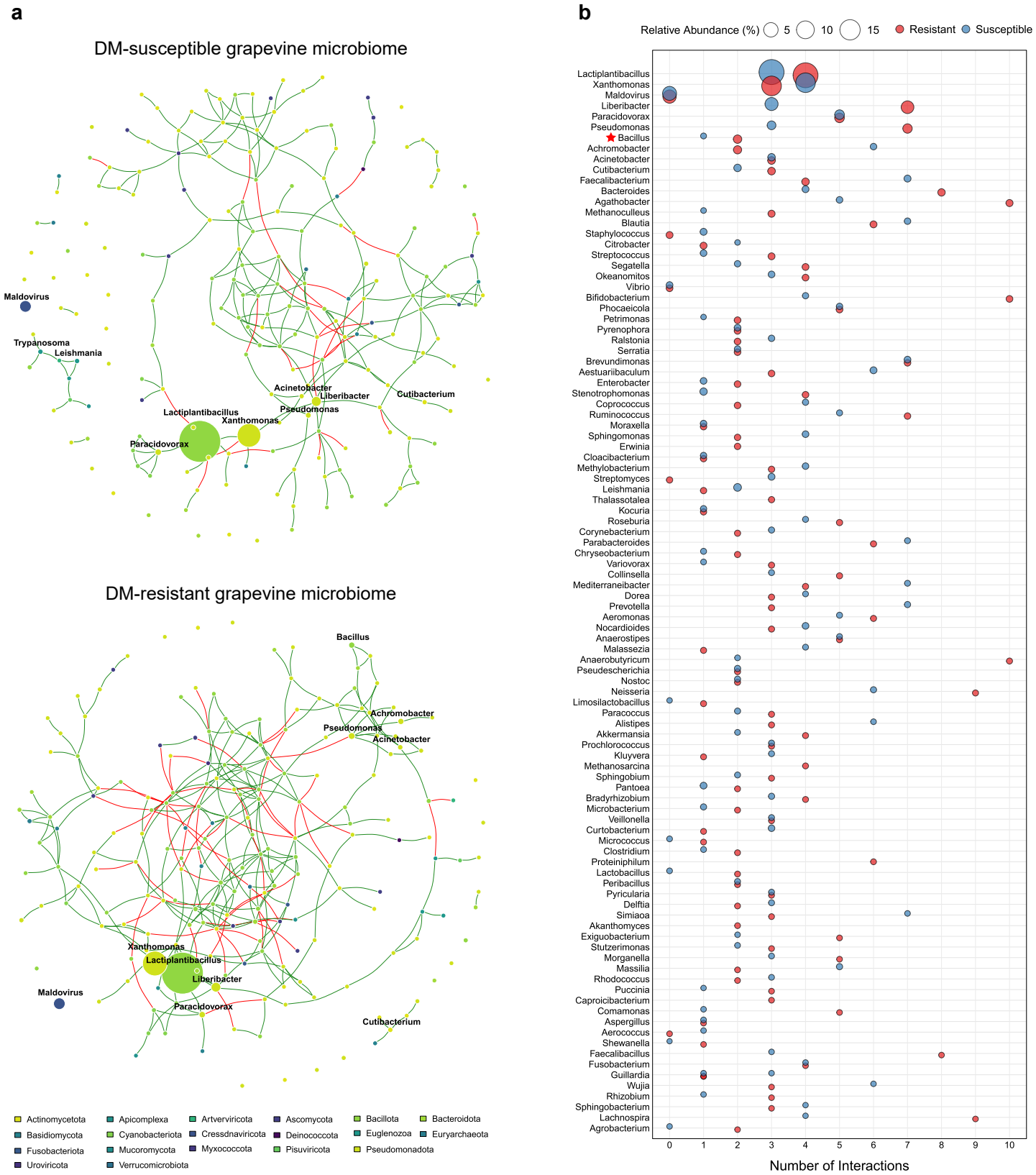
a
DM-susceptible grapevine microbiome
Maldovirus
Trypanosoma
Leishmania
Acinetobacter
Liberibacter
Cutibacterium
Pseudomonas
Lactiplantibacillus
Xanthomonas
Paracidovorax
DM-resistant grapevine microbiome
Bacillus
Achromobacter
Pseudomonas
Acinetobacter
Xanthomonas
Lactiplantibacillus
Liberibacter
Maldovirus
Paracidovorax
Cutibacterium
Actinomycetota
Basidiomycota
Fusobacteriota
Uroviricota
Apicomplexa
Cyanobacteriota
Mucoromycota
Verrucomicrobiota
Artverviricota
Cressdnaviricota
Myxococcota
Ascomycota
Deinococcota
Pisuviricota
Bacillota
Euglenozoa
Pseudomonadota
Bacteroidota
Euryarchaeota
b
Relative Abundance (%) 5 10 15
Resistant
Susceptible
Lactiplantibacillus
Xanthomonas
Maldovirus
Liberibacter
Paracidovorax
Pseudomonas
Bacillus
Achromobacter
Acinetobacter
Cutibacterium
Faecalibacterium
Bacteroides
Agathobacter
Methanoculleus
Blautia
Staphylococcus
Citrobacter
Streptococcus
Segatella
Okeanomitos
Vibrio
Bifidobacterium
Phocaeicola
Petrimonas
Pyrenophora
Ralstonia
Serratia
Brevundimonas
Aestuariibaculum
Enterobacter
Stenotrophomonas
Coprococcus
Ruminococcus
Moraxella
Sphingomonas
Erwinia
Cloacibacterium
Methylobacterium
Streptomyces
Leishmania
Thalassotalea
Kocuria
Roseburia
Corynebacterium
Parabacteroides
Chryseobacterium
Variovorax
Collinsella
Mediterraneibacter
Dorea
Prevotella
Aeromonas
Nocardioides
Anaerostipes
Malassezia
Anaerobutyricum
Pseudescherichia
Nostoc
Neisseria
Limosilactobacillus
Paracoccus
Alistipes
Akkermansia
Prochlorococcus
Kluyvera
Methanosarcina
Sphingobium
Pantoea
Bradyrhizobium
Microbacterium
Veillonella
Curtobacterium
Micrococcus
Clostridium
Proteiniphilum
Lactobacillus
Peribacillus
Pyricularia
Delftia
Simiaoa
Akanthomyces
Exiguobacterium
Stutzerimonas
Morganella
Massilia
Rhodococcus
Puccinia
Caproicibacterium
Comamonas
Aspergillus
Aerococcus
Shewanella
Faecalibacillus
Fusobacterium
Guillardia
Wujia
Rhizobium
Sphingobacterium
Lachnospira
Agrobacterium
0 1 2 3 4 5 6 7 8 9 10
Number of Interactions
c
Control (water)
B. cereus 117gite
B. cereus SW9SE
P. viticola
P. viticola + B. cereus 117gite
P. viticola + B. cereus SW9SE
d
Disease Index
0 10 20 30 40 50
52
0 0 0 0 0
Control (water)
B. cereus gite117
B. cereus SW9SE
P. viticola
P. viticola + B. cereus gite117
P. viticola + B. cereus SW9SE

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.

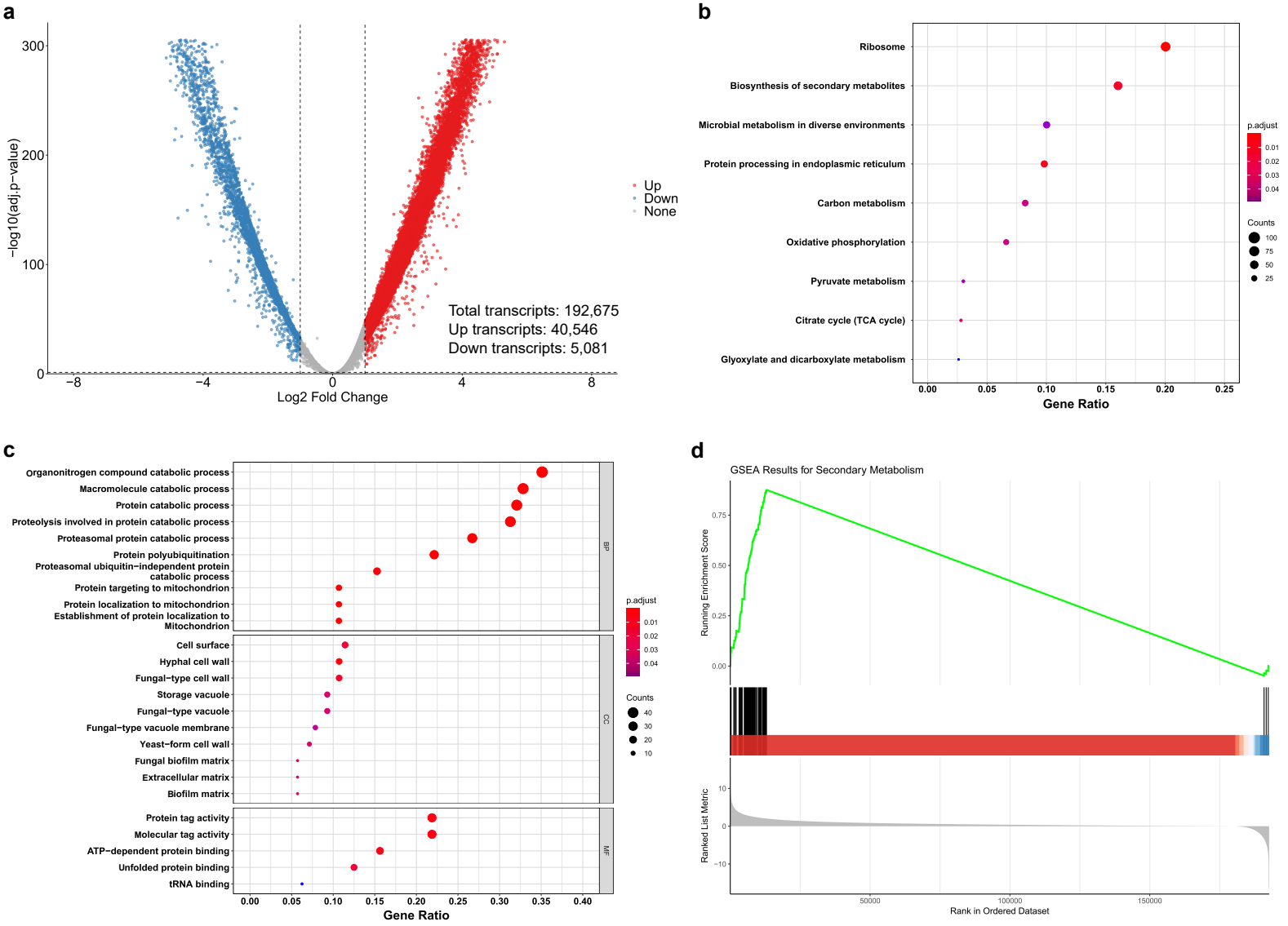

Fig. 6 Eukaryotic transcriptome profile of the grapevine phyllospheric microbiome in response to the oomycete *Plasmopara viticola* infection. a, Volcano plot of the grape microbiome upon infection by *P. viticola*. The dashed lines indicate thresholds for statistical significance (log2 fold change > 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 log2 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.

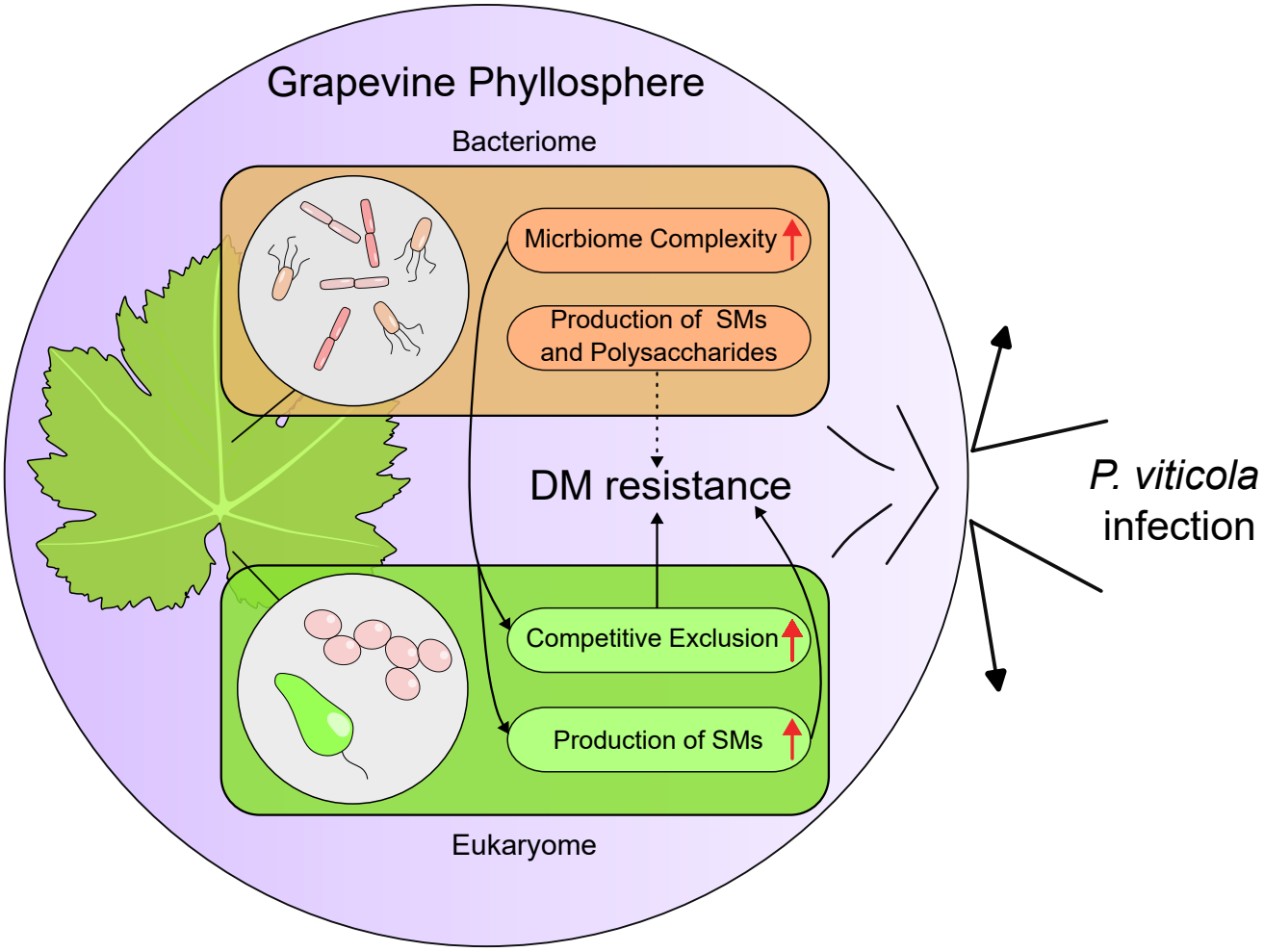

Fig. 7 Schematic diagram depicting grapevine microbiome's potential deployment of bacteriome and eukaryome to combat *P. viticola* infection. This pan-metagenomic and pan-metatranscriptomic study demonstrates that the grapevine bacteriome orchestrates microbiome assembly, with key taxa (e.g., Bacillota) 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).

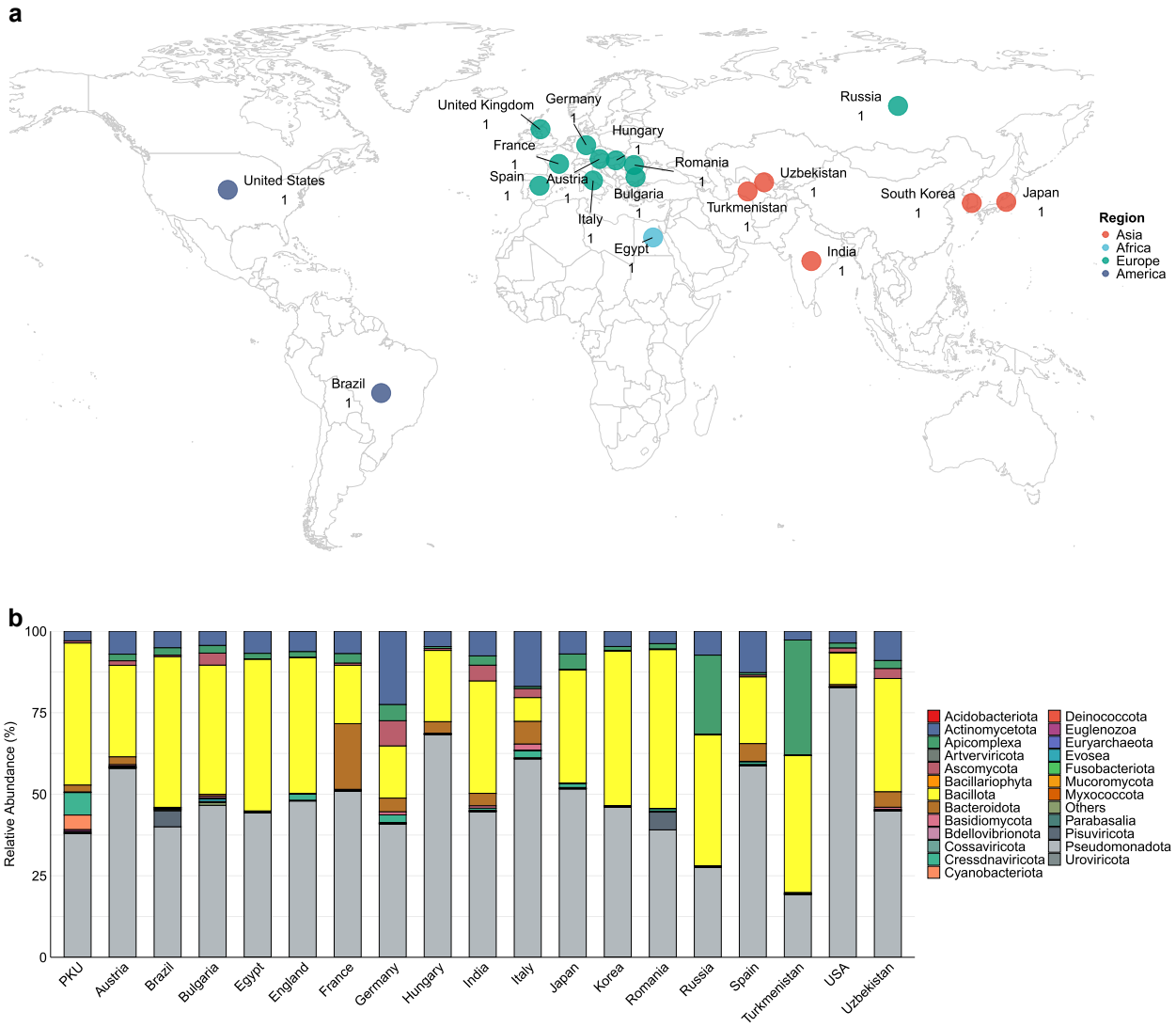

Fig. S1 Worldwide comparison of grapevine phyllosphere microbiome compositions. a, Geographical distribution of Vitis accessions used for comparative analysis. One representative Vitis accession was selected per country. **b**, Top 10 phylum level composition of grapevine phyllosphere microbiome across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

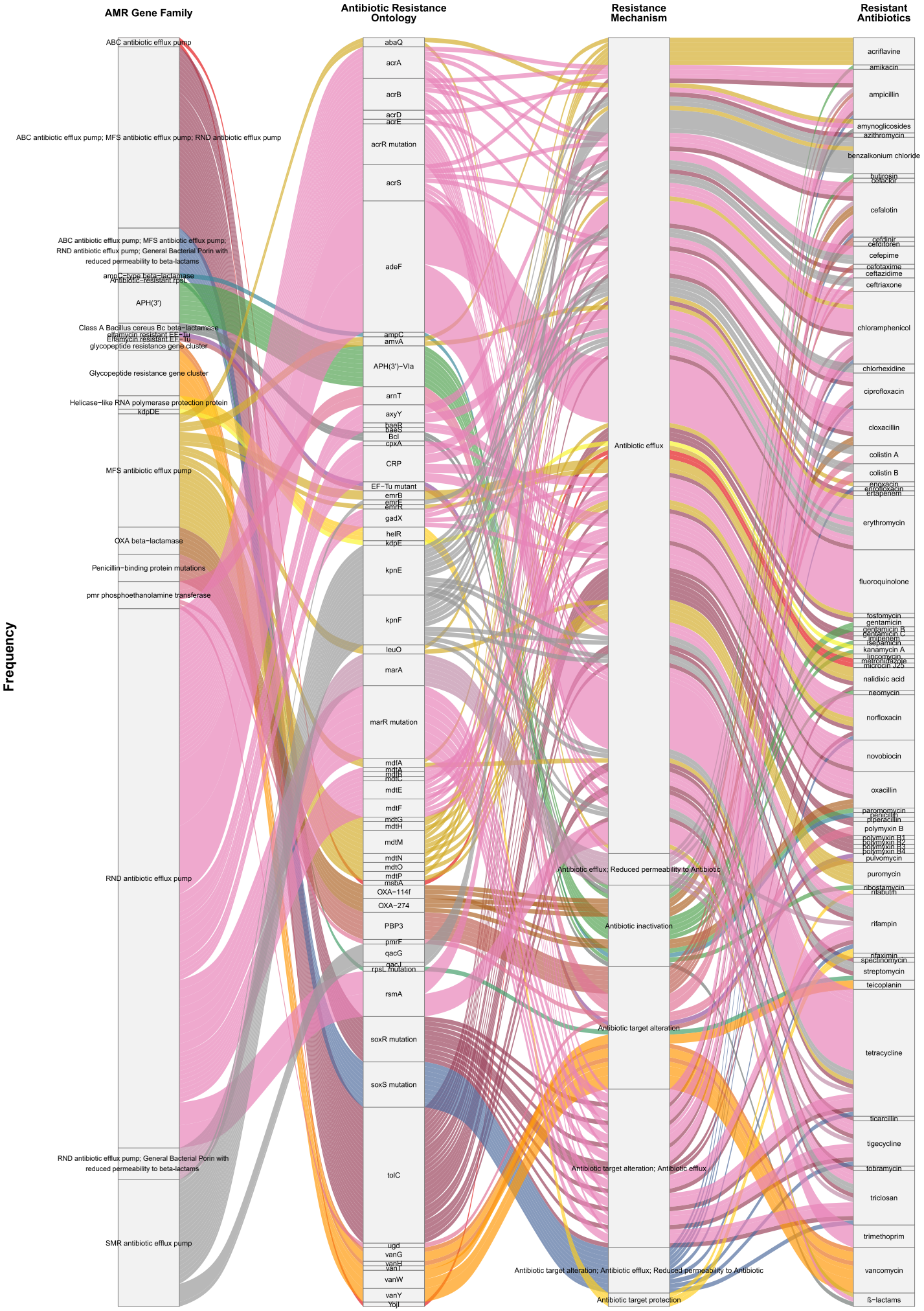

Fig. S2 Resistome profile of the GPMAGs. Identified AMR gene frequencies in the GPMAGs were categorized by gene family, resistance mechanism, and target antibiotics.

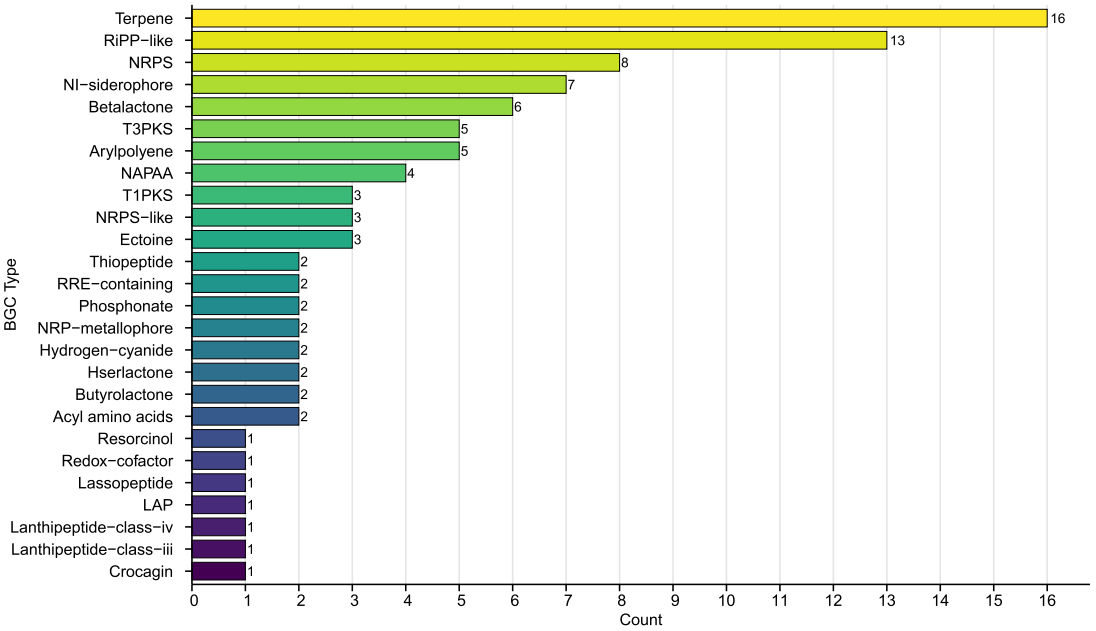

Fig. S3 BGC profile of the GPMAGs. Identified BGCs from the GPMAGs were displayed by bar chart.

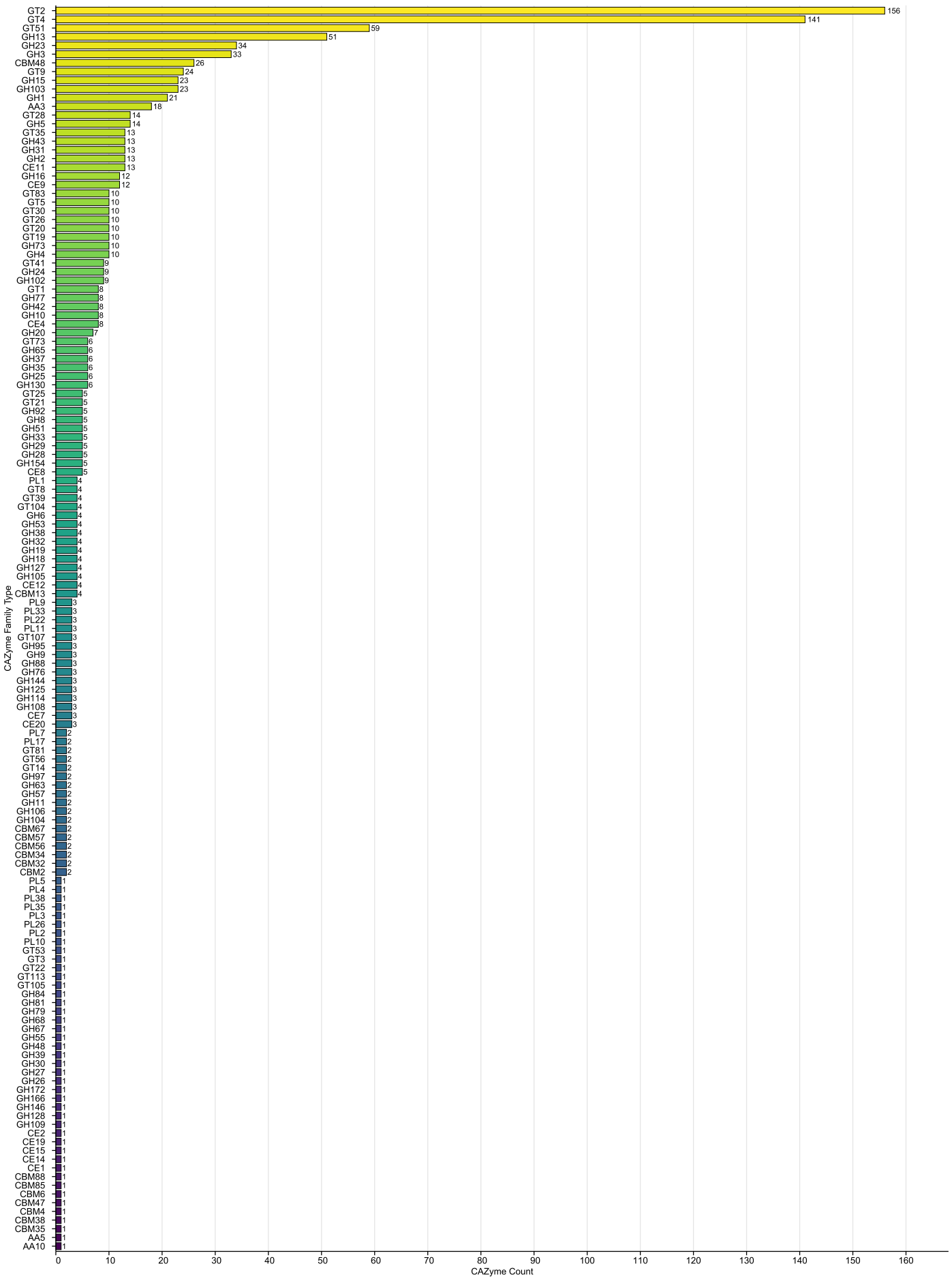

Fig. S4 CAZyme profile of the GPMAGs. Bar chart exhibits identified CAZyme types of the GPMAGs.

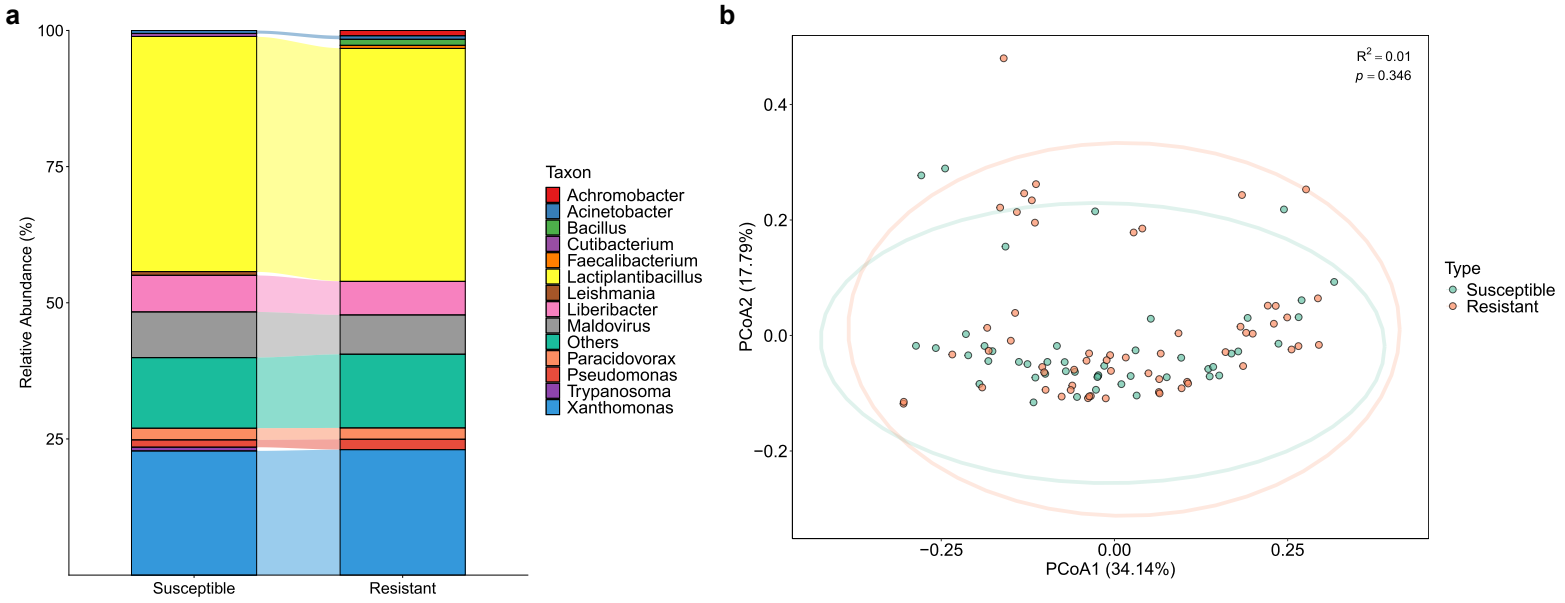

Fig. S5 Microbiome profiles associated with DM resistance. a, Mean relative abundances of the top 10 most abundant genera based on the types of DM resistance. **b**, PCoA of the grape microbiomes at genus level. Each point represents a grapevine accession with colored according to its variety. Ellipses 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.

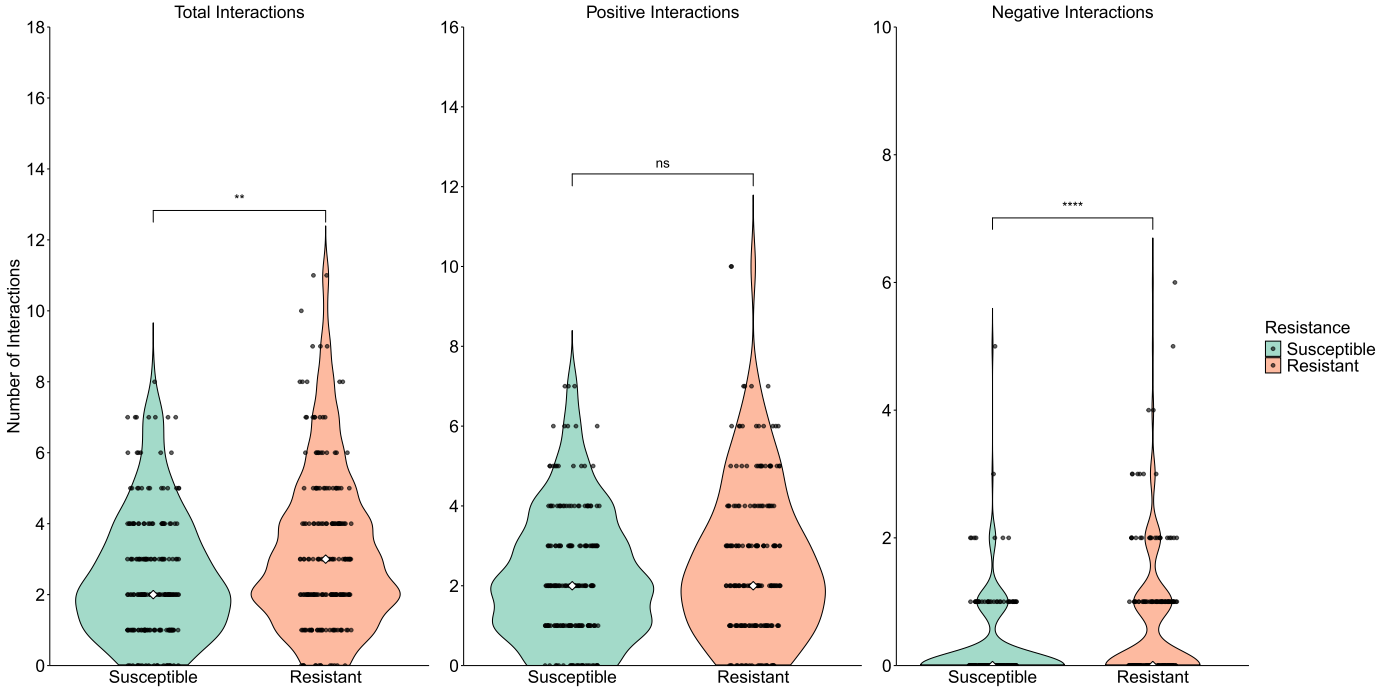

Fig. S6 Microbial interaction profiles associated with DM-resistance. Violin plots depict the median microbial interactions for each node in the networks shown in Fig. 5a. White diamonds indicate medians. Significance levels are denoted as ** ($p < 0.01$), **** ($p < 0.0001$), and ns (not significant).

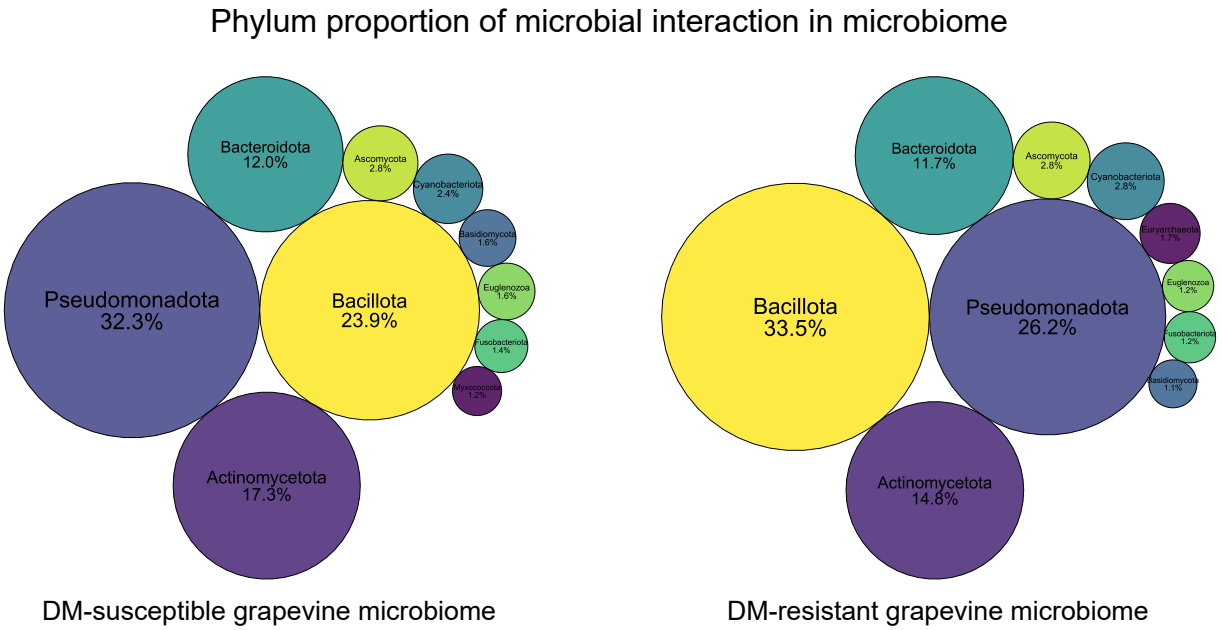

Fig. S7 Phylum-level microbial interaction proportions between microbiomes. Bubble plots display interaction frequencies among the top 10 abundant phyla in the susceptible and resistant microbiomes, with bubble sizes corresponding to proportional values (numerically labeled).

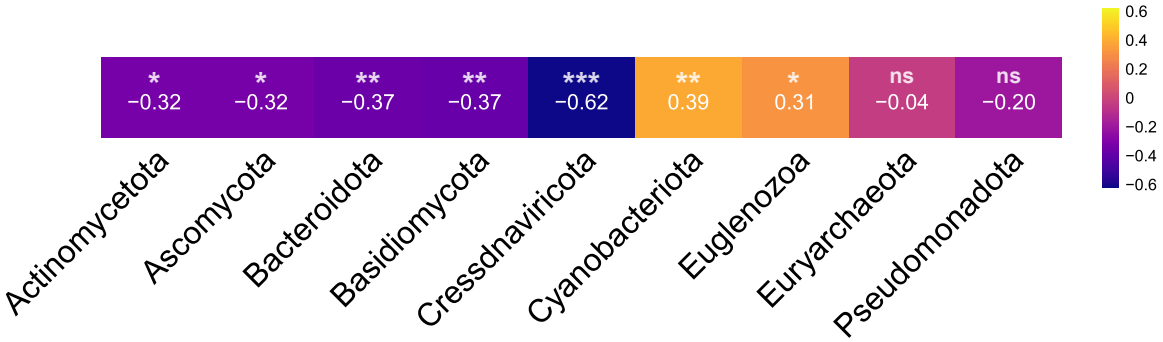

Fig. S8 Abundance correlation between Bacillota and major phyla in the resistant microbiome. Heatmap depicts Spearman coefficients between Bacillota and major phyla based on their abundances in the resistant microbiome. Numerical values and significance levels * ($p < 0.05$), ** ($p < 0.01$), *** ($p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