Microbial community diversity on the grape epidermis revealing the advantage of planting wine grape in Xinjiang

Feifei Gao[1], Bin Wang[1*], Jing Xiao[2], Xiaoji Zheng[1], Xuewei Shi[1*]

Food college, Shihezi University, Shihezi 832000, Xinjiang Uygur Autonomous Region, P. R. China.

* Corresponding authors

[1]Grape and wine engineering laboratory, Food college, Shihezi University, Shihezi 832000, Xinjiang Uygur Autonomous Region, P. R. China. [2]College of information science and technology, Shihezi 832000, Xinjiang Uygur Autonomous Region, P. R. China. Correspondence and requests for materials should be addressed to B.W. (email: binwang0228@shzu.edu.cn) or X.S. (email: shixuewei@shzu.edu.cn).

E-mail: B.W., binwang0228@shzu.edu.cn; X.S., shixuewei@shzu.edu.cn

Tel.: +86 0993-2058093

Abstract: The native microorganisms on wine grape epidermis contributed to the regional wine characteristic and quality. Xinjiang Uygur Autonoumous Regions in northwest China was one of the eight main wine-producing areas in China. To investigate the relationship between microbial community structure of wine grape epidermis and environment conditions, 16S rDNA and ITS sequences of 48 wine grape samples from 4 wine grape cultivars and 6 wine-growing regions in Xinjiang were sequenced, based on Illumina high-throughput sequencing technology. A total of 691 operational taxonomic units (OTUs) in 16 bacterial phyla and 349 OTUs in 3 fungal phyla were identified. Among them, Proteobacteria and Ascomycota were predominant bacteria and fungi, respectively. The canonical correspondence analysis (CCA) indicated that bacterial community diversity was prominently related with altitude, latitude and longitude, while the fungi was closely related with altitude, dryness, frost-free period, latitude and longitude. Our results suggest that microbial community structure on wine grape epidermis is controlled primarily by environment conditions.

Keywords: Wine grape epidermis; Microbial community structure; Environment conditions; Illumina high-throughput sequencing.

Introduction

Wine is an alcoholic beverage made from grapes. Natural factors, such as the climate, geology, soil and grape cultivars, remarkably contributed to wine quality through affecting grape features and microorganism communities[1,2]. Climate played a key role in wine microbiology, wine chemistry and wine sensory through affecting vine phenology and grape composition[3]. The grape component that affected wine flavors, such as total phenolic, total flavonols and anthocyanins, varied with grape cultivar and vintage [4]. Then, to produce high quality wines, the vineyards were generally limited to choose only a few grape cultivars that were suitable for the local climate and displayed high feature of brewing wine [5,6].

Like other fermented food, microorganisms play a key role in wine quality. In winemaking process, wine need two steps of fermentation, including the alcohol fermentation and the malolactic fermentation[7]. During alcohol fermentation, after fast proliferation in presence of oxygen, the yeast cells degrade glucose and release ethanol to the fermentation broth in the absence of oxygen. The most common yeasts used in wine fermentation belonging to *Brettanomyces*, *Candida*, *Kloeckera*, *Saccharomycodes*, *Schizosaccharomyces*, *Zygosaccharomyces* and *Aureobasidium*[8]. While, during the malolactic fermentation, so-called secondary fermentation, L-malate is transformed into L-lactate and CO_2, which decrease the acidity of the wine because two acidic groups of L-malate have been replaced by only one acidic group of L-lactate [9]. The malolactic fermentation results from metabolisms of some lactic acid bacteria, such as *Lactobacillus*, *Leuconostoc* and *Pediococcus* [10,11]. The flavors of different wine were established during winemaking through a complex role of various microorganisms, especially yeast[12]. Therefore, the wine could be easily distinguished from each other according to the vine cultivar, geographical origin and year of production [13].

Indeed, the structure and the composition of the yeast cells greatly contribute to the sensory features of wine, especially the wine fermented by spontaneous fermentation [14]. During spontaneous fermentation, the microorganisms, consisting of

the added commercial yeast and epiphytic microorganism, always play vital roles in the local flavor and wine quality through producing a large amount of microbial flavor[15]. The epiphytic microorganisms colonized in grape berries, including bacteria, yeast, and filamentous fungi, establish an intricate and kinetic microbe ecosystem[16,17]. These plant-associated microorganisms have also been proved to display a positive interaction with their host plants, such as plant growth promotion and pathogen defense[18,19]. Latest researches have showed that the climate conditions play a key role in the microbial communities in the environment[20,21]. Then, the climate conditions can modify the quality of vines and wines by affecting the geographic delineations of epiphytic microorganisms communities and populations of grape, especially the yeast cells [22]. Furthermore, the technology of high-throughput sequencing has been widely applied to investigate the microbial communities in the environment. Bokulich, et al. [23] demonstrated that regional, site-specific, and grape cultivar factors shaped the fungal and bacterial consortia inhabiting wine-grape surfaces.

Due to the unique geographical environment and climate, Xinjiang Uygur Autonoumous Regions, locating in northwest China, become one of eight main wine producing areas in China. Though grape epidermis microorganisms have been proved to play a key role in wine produce and affected by the local environment and climate, there are few researches investigating the relationship between grape epidermis microorganisms and geographical environment in Xinjiang. In this study, to investigate effects of cultivars，climates and environments on microbial community construction of grape epidermis in different regions, the bacterial and fungal community structures of four wine grape cultivars in six climate regions of China were explored by using 16S rDNA/ITS (internal transcribed spacer) Illumina high-throughput sequencing and related bioinformatics and statistical analyses. This study will provide significantly scientific proof for classifying wine grape quality in different areas of Xinjiang and evaluating wine producing techniques and microbe resources.。

Materials and methods

Sample collection and cells treating

Though Xinjiang has a vast territory and a temperate continental climate, grape cultivation is mainly concentrated in the northern Xinjiang. Then, a total of 48 wine grape samples, belonging to 4 different wine grape cultivars (Cabernet Sauvignon, Merlot, Italian Riesling, Cabenet Franc), were collected from 6 wine-growing region in the northern Xinjiang in China, including Shanshan, Yanqi, Heshuo, Huoerguosi, Fukang and Manasi (Fig. 1). The map of sampling point was drawed by using Adobe Photoshop 8.01 software.

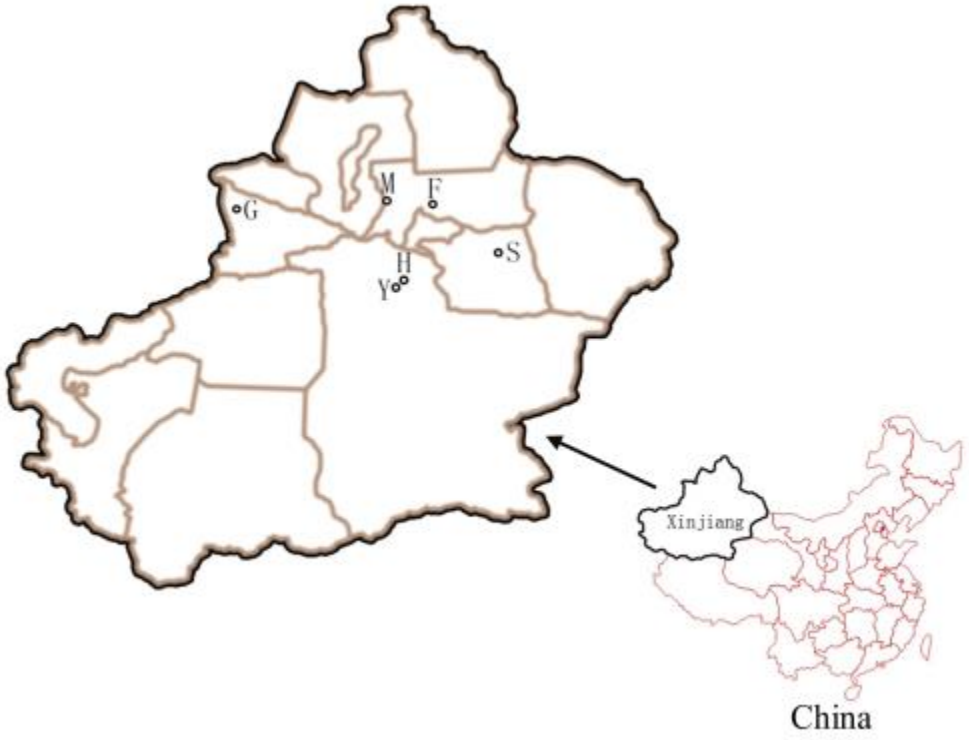

Fig. 1. Geographic location of the sample collection sites. Xinjiang Uygur Autonoumous Regions is located in the northwest of China. F, G, H, M, S and Y respectively represent Fukang (87°59′24″E, 44°9′36″N, altitude 552 m), Huoerguosi (80°24′36″E, 44°12′36″N, altitude 800 m), Heshuo (86°51′36″E, 42°16′12″N, altitude 1094 m), Manasi (86°13′12″E, 44°18′N, altitude 462 m), Shanshan (90°7′12″E, 42°52′12″N, altitude 381 m) and Yanqi (86°34′12″E, 42°3′36″N, altitude 1059 m).

All the samples were collected with sterile 50 mL centrifuge tubes and rapidly stored at -20 ℃. Among the samples, the wine grape cultivars, Cabernet Sauvignon, Merlot, Italian Riesling and Cabenet Franc, were numbered 1, 2, 3 and 4, respectively. While the cultural areas, Shanshan, Yanqi, Heshuo, Huoerguosi, Fukang and Manasi,

were shorted for S, Y, H, G, F and M, respectively (Table 1). For example, the sample M1 meant that the grapes of Cabernet Sauvignon were sampled from Manasi.

Table 1. Grape samples of four cultivars were collected from six different culture areas. Three grape samples of the same variety for each region were collected for duplicates. "-" indicates that the cultivar is not planted in this culture area.

Grape cultivars	Manasi (M)	Fukang (F)	Heshuo (H)	Yanqi (Y)	Huoerguosi (G)	Shanshan (S)
Cabernet Sauvignon	M1	F1	H1	Y1	G1	S1
Merlot	M2	F2	-	-	G2	S2
Italian Riesling	M3	F3	H3	-	-	-
Cabenet Franc	M4	-	-	Y4	G4	-

After unfreezing, 20 g of grape samples were washed with 40 mL sterile water for 5 times. The suspension was collected with 250 mL of sterile Erlenmeyer flask and then filtered with a 0.22 μ m filter. The filtered microorganism was used to extract microbial genome DNA.

Extraction of genome DNA and PCR amplification

Power Water® DNA Isolation Kit (MOBIO firm in America) was used to extract total DNA of the microorganisms on grape surface[24]. The DNA concentrations were evaluated by electrophoresis in a 1.5% agarose gel, and the extracted DNA was stored at -20 ℃ until further tests[25]. Then 1 μL of genomic DNA was added into a centrifuge tube and diluted to 1 ng/μL using sterile water. A total of 48 diluted DNA samples were submitted to Novegene company for 16S rDNA Amplicon Sequencing and Internal Transcribed Spacer (ITS) sequencing based on Illumina Hiseq sequencing platform.

The variable region V4 of the 16S rDNA gene was selected for the construction of the bacterial community library for Illumina sequencing. The specific primers, 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACHVGGGTWTCT AAT-3'), were used to amplify the sequence of the 16S rDNA gene, while ITS5-1737F (5'-GGAAGTAAAAGTCGTAACAAGG-3') and

ITS2-2043R (5'-GCT GCGTTCTTCATCGATGC-3') were used to amplify fungal ITS sequence. PCR was performed in 20 μL reactions in triplicate, with each reaction tube containing 0.2 mM of each primer, 10 ng of template DNA (the diluted genomic DNA), 0.25 mM dNTPs, 1×PCR reaction buffer, 2 U of FastPfu DNA Polymerase. The following PCR condition was used for 16S rDNA: 95 ℃ for 2 min, 95 ℃ 30 seconds, 55 ℃ 30 seconds and 72 ℃ 45 seconds for 30 cycles, and a final extension of 72 ℃ for 10 min. The same PCR conditions were used for ITS, except that the second stage had 35 cycles[26]. Additionally, Phusion® High-Fidelity PCR Master Mix with GC Buffer from the New England Biolabs company, the enzyme of high-performance and high-fidelity was used for PCR to ensure amplification efficiency and accuracy[27]. Then, PCR products mixed same volume of 1X loading buffer (contained SYB green) were detected by electrophoresis in a 2.0% agarose gel, and samples with bright main strip between 400-450 bp were chosen for further experiments.

PCR products purification and library preparation

PCR products were mixed in equal density ratios according to the concentration of PCR products, and then purified with Qiagen Gel Extraction Kit (Qiagen company in Germany)[28]. Sequencing libraries were generated by using TruSeq® DNA PCR-Free Sample Preparation Kit (Illumina, USA) following manufacturer's recommendations, then indices codes were added[29]. The library quality was evaluated on the Qubit@ 2.0 Fluorometer (Thermo Scientific) and Agilent Bioanalyzer 2100 system [30,31]. At last, the library was sequenced on an Illumina HiSeq2500 platform and 250 bp paired-end reads were generated.

Bioinformatics analyses

The paired-end (PE) reads were gained after accomplishing High-throughput sequencing, and were assigned to samples based on their unique barcode and truncated by cutting off the barcode and primer sequence. PE reads were spliced by FLASH V1.2.7 [32] and filtrated according to the QIIME V1.7.0 quality controlled

process[33], which were called the high-quality Clean Tags. The tags were compared with the reference database using UCHIME algorithm to detect chimera sequences, and then the chimera sequences were removed[34]. Then the Effective Tags were finally obtained. Sequences with more than 97% similarity were assigned to the same operational taxonomic units (OTUs). The QIIME suite of programs was used to evaluate alpha diversity including ACE, Chao1 richness, Shannon diversity, Goods-coverage, Simpson indices and Observed species[35]. Rarefaction analysis was used to estimate sequencing depth. The differences in community structure of different samples and groups were analyzed by the principal coordinate analysis (PCoA) based on weighted and unweighted calculation [36,37]. Flower chart and Venn diagrams were showed to explore the common and specific OTUs information between different samples or groups.

The microbial communities of grape epidermis between different regions/cultivars were further compared using analysis of molecular variance (Amova), analysis of similarities (Anosim), Metastats and LDA EffectSize (LEfSe). Amova analysis using mothur software was performed to assess significance of microbial community structure between different regions[38]. Anosim is a distribution-free method of multivariate data analysis to test whether inter-group difference is significantly greater than intra-group one[39]. Metastats analysis is a statistical method for identifying the significantly different bacterial and fungal genera in different wine-growing regions [40].

The LEfSe is a software used to identify high-dimensional biometric identifiers and reveal genomic characteristics, and it was performed for revealing species with significant differences between groups[41]. The correlation between environmental variables and microbial community composition was performed by canonical correspondence analysis (CCA)[42]. Spearman correlation analysis and variance partitioning analysis were used to uncover the correlation between microbial community composition and environmental and climatic factors[43,44], then the significant correlation was further proved by Mental test[45]. After basic analysis, the figures were drawn by using the packages in RStudio (version 2.15.3)[46,47].

Results

Diversity assessment

A total of 48 wine grape samples, belonged to Cabernet Sauvignon, Merlot, Italian Riesling and Cabenet Franc, were collected in the vineyards from Shanshan, Yanqi, Heshuo, Huoerguosi, Fukang and Manasi of Xinjiang in China. The DNA of microorganisms on grape surface was subjected to Novegene company (China) for 16S rDNA and ITS sequencing based on Illumina Hiseq sequencing platform. To evaluate the quality of sequencing results, the raw sequences, raw tags, average length, sequencing error rate and effective tags percentage from samples were concerned. In this study, raw PE, effective tags, avglen (nt), Q20, Q30 and effective % in samples were higher than 53329, 52267, 219, 99.22%, 98.46% and 89.86%, respectively (Supplementary Table S1), which indicated that all of the parameters met the demand of further analysis.

After eliminating chimeral sequences and mismatches, the total number of 16S rDNA reads obtained from the 48 samples was 3,245,424 (average 67,613), which were clustered into 691 operational taxonomic units (OTUs) with at least 97% similarity in nucleotide identity (Supplementary Table S2). However, 2242,800 (average 46,725) ITS reads were clustered into 349 OTUs (Supplementary Table S3). The species accumulation boxplot also agreed with the accuracy of these data (Supplementary Fig. S1). When the sequences reached up to 20000, the rarefaction curves of most samples tended to be complanate, which indicated that the reasonable sequencing depth had been acquired and further sequencing data might only produce a small amount of new OTUs (Supplementary Fig. S2). Shannon-Wiener curve was also in accordance with this claim (Supplementary Fig. S3).

Table 2. Abundance and diversity estimation of the 16S rDNA sequencing libraries from 48 wine grape samples.

Sample	OTUs	observed_ species	shannon	simpson	chao1	ACE	PD_whole_tree
S1	36	20.00±3.61	0.83±0.38	0.35±0.20	23.17±5.32	28.69±7.74	3.22±0.61
S2	38	19.00±3.46	0.34±0.20	0.09±0.05	27.61±3.91	29.35±6.38	2.99±1.90

M1	398	227.33±25.58	0.90±0.62	0.28±0.25	265.19±32.43	275.18±19.20	16.77±1.99
M2	47	22.67±3.51	1.02±0.35	0.40±0.21	29.71±4.75	36.87±6.89	3.43±0.60
M3	56	29.67±8.96	0.28±0.19	0.07±0.05	46.83±18.87	56.76±23.16	4.01±1.22
M4	38	17.00±6.93	0.15±0.05	0.04±0.01	19.55±9.21	22.07±10.54	2.54±1.10
F1	41	23.00±4.36	0.14±0.03	0.03±0.01	23.93±4.76	25.31±4.93	3.08±0.71
F2	51	28.00±3.61	0.43±0.11	0.11±0.03	32.90±4.43	37.42±4.37	4.01±0.88
F3	35	22.33±1.15	0.32±0.02	0.07±0.01	25.42±1.38	27.25±2.57	2.97±0.16
H1	441	254.33±65.96	0.71±0.49	0.16±0.14	291.98±59.47	310.92±56.93	19.38±3.75
H3	311	170.67±41.48	0.64±0.27	0.16±0.08	210.32±49.27	224.77±54.68	14.32±2.09
Y1	251	138.67±14.74	0.36±0.03	0.07±0.01	175.95±31.85	192.37±30.31	12.05±1.21
Y4	336	180.00±64.65	0.96±0.54	0.29±0.20	214.30±80.38	227.81±82.88	15.11±5.32
G1	297	157.33±34.00	0.66±0.56	0.22±0.25	201.90±44.27	207.78±34.47	14.07±3.89
G2	310	168.00±60.02	0.97±0.61	0.31±0.24	220.04±69.80	229.80±72.26	14.27±4.71
G4	392	225.67±37.81	1.13±0.44	0.36±0.20	382.18±194.29	356.00±126.38	18.46±1.44

The Alpha Diversity for different samples at a 97% consistency threshold was calculated via Chao1, ACE, Shannon, Simpson and Coverage indices (Table 2). Chao1 and ACE indices of the samples from the same region samples displayed slight difference. Whereas, the value of Heshuo (H), Yanqi (Y) and Huoerguosi (G) regions were distinctly higher than those of Shanshan (S), Manasi (M) and Fukang (F) regions, which signified that the bacterial community richness on grapes from H, Y and G regions are obviously higher than those from S, M and F. Furthermore, Shannon and Simpson diversity indices, combining evenness and species richness, also indicated that the bacterial community diversity on grapes from H, Y and G regions is remarkably higher than ones of S, M and F regions. Coverage estimates (Sequencing depth indices) were very high for all samples. As for fungi (Supplementary Table S4), the Chao1 and ACE value of the samples from G, H, M and Y region was distinctly higher than those from S and F region, which suggested that the fungal community richness of G, H, M and Y regions were obviously higher than those from S and F region. At the meantime, Shannon and Simpson diversity indices indicated that the fungal community diversity on the grapes from G, H, M and Y regions was remarkably higher than those from S and F regions. However, Coverage estimates (Sequencing depth indices) were very high for all samples.

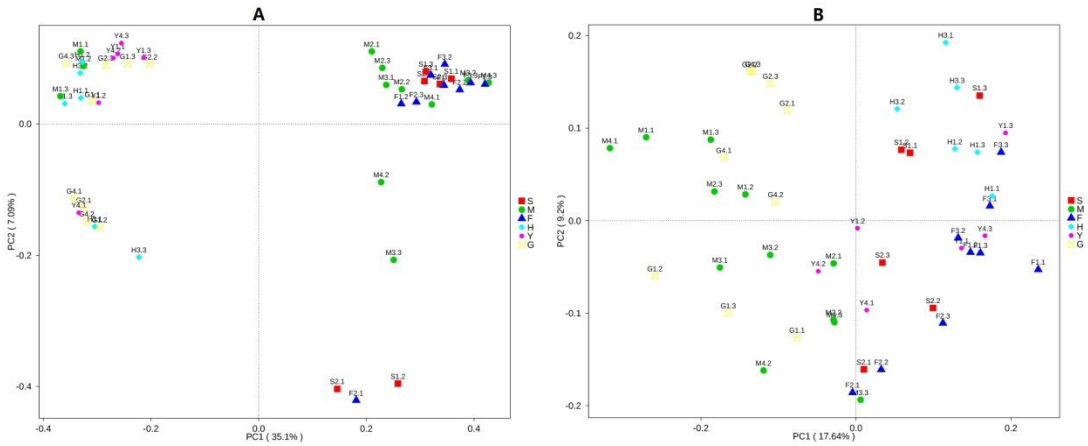

Fig.2. Comparison of microbial community in samples from different cultivation regions. Principal Coordinate Analysis (PCoA) based on Unweighted Unifrac Distance was generated with OTUs (at 97% similarity) present in the different cultivation areas samples. (A) bacteria, (B) fungi. principal coordinate (PC). Different color shapes represent samples of different cultivation areas.

The Principal Coordinates Analysis (PCoA) based on unweighted Unifrac distance showed that the bacterial community composition flocked to gather well than fungal community composition on grapes from different regions, which indicated that the bacterial community was more sensitive to the natural environment and local climate from different regions than fungal community. The bacterial community composition on grapes surface from S, M and F mainly flocked to gather, while those from H and G were divided into two clusters. One of the clusters was similar to Y region (Fig. 2A). There also existed some outlier. For example, S1.2, S2.1 and F2.1 deviated their main clusters. Deviate the contribution value of PC1 is 35.1%, followed by PC2 (9.2%). At the same time, the fungal community composition on grapes surface from H flocked together, while those from other regions were scattered (Fig. 2B). PC1 alone explains 17.64% of variance, followed by PC2 9.2%.

Amova showed that there was significant difference among F, G, H, M, S and Y fungal communities (Amova, $p<0.001$), but the difference of bacterial communities was not significant. Anosim，a non-parametric statistical test，indicated that the difference between F and G (R 0.183, $p<0.05$) bacterial communities was substantially greater than the intra-group difference, additionally F and H (R 0.291, $p<0.05$), F and S (R 0.246, $p<0.05$) , F and Y (R>0), S and Y (R>0), M and G (R>0).

Analogously, the difference between F, G, H, M, S and Y fungal communities was substantially greater than the intra-group difference (Anosim, $R>0$, $p<0.05$).

Taxonomic distribution of bacteria

OTUs of the samples from different regions were further assigned to different taxa and their relative taxonomic richness was evaluated. Besides some unknown groups, 36 bacteria classes were grouped into at least sixteen phyla (Supplementary Table S5). Among these phyla, proteobacteria strains were represented in all samples and the relative abundance was higher than 98.55% of the total bacteria population in each group samples (Fig. 4A).

Furthermore, the common and exclusive bacterial OTUs of the grape epidermis were showed by Venn diagrams. Bacterial OTUs distribution of grape epidermis in F, G, H, M, S and Y culture areas were presented 85 (12.3%), 540 (78.1%), 499 (72.2%), 430 (62.2%), 58 (8.4%) and 405 (58.6%), respectively (Supplementary Fig. S4). Bacterial OTUs distribution of different cultivar grapes in an area displayed a significant difference. For example, the OTUs of Cabernet Sauvignon, Merlot and Cabenet Franc cultivar in Huoerguosi were 297, 310 and 392, respectively, and 163 bacterial OTUs were common (Supplementary Fig. S4C). Similarly, bacterial OTUs distribution of same grape cultivar in different culture areas were also compared. For instance, Cabernet Sauvignon cultivar, exclusive bacteria OTUs of grape epidermis in F, G, H, M, S and Y culture areas were presented 1, 40, 80, 45, 2 and 17, respectively, and 13 bacterial OTUs were common (Fig. 3). The results indicated that bacterial OTUs distribution of grape epidermis affected by the cultivation area and cultivar.

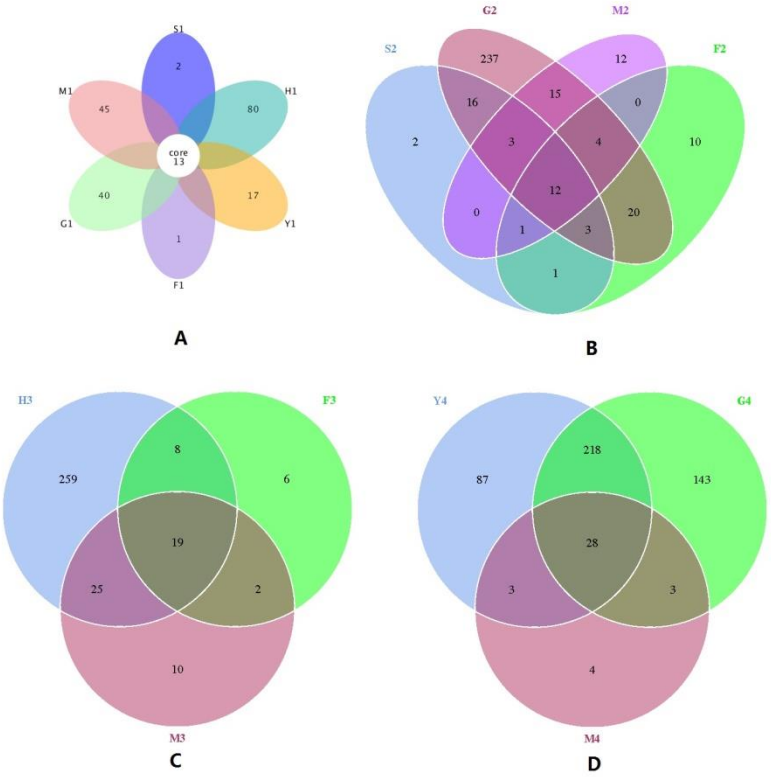

Fig. 3. Venn diagrams showing the common and exclusive bacterial OTUs of the grape epidermis. A: Cabernet Sauvignon; B: Merlot; C: Italian Riesling; D: Cabenet Franc.

However, those common and exclusive bacteria OTUs truly shaped bacteria community constructure. Particularly, mainly bacteria OTUs were presented by the top 10 of relative abundance bacteria genera, including *Pantoea*, *Pseudomor*, *Buchnera*, *Rhodococcus*, *Nitrosospira*, *Massilia*, *Aeromonas*, *Steroidobacter*, *Thermomonas* and *Pedobacter*, which were predominated by *Pantoea* and *Pseudomor* (Fig. 4). Furthermore, the relative abundance of bacteria in different regions showed more significant differences. For example, the relative abundance of *Pantoea* in F, G, H, M, S and Y culture areas were 96.3%, 79.6%, 91.3%, 76.6%, 85.1% and 88.9%, respectively. Therefore, bacterial abundance of grape epidermis was affected by different cultivation area.

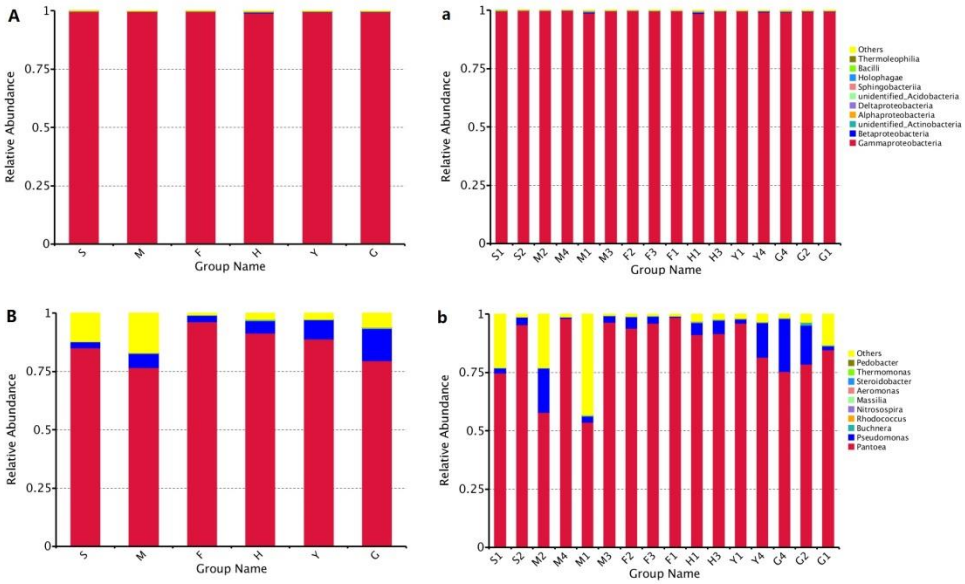

Fig.4. Relative abundance of top 10 bacterial class (A/a) and genus (B/b).

Proteobacteria

Sequence reads of the proteobacteria phylum could be classified into four classes, named Alpha, Beta-, Delta- and Gammaproteobacteria, as well as some unidentified and unclassified classes. Gammaproteobacteria had relatively high reads, followed by Beta-, Alpha- and Deltaproteobacteria (Supplementary Table S5). These four classes were completely present in G and M cultivation areas, whereas Alphan-, Beta- and Gammaproteobacteria were present in F, H, S and Y areas. Gammaproteobacteria dominated over Betaproteobacteria, followed by Alphaproteobacteria and Deltaproteobacteria. Gammaproteobacteria was represented by sixteen orders and twenty-three families, while Betaproteobacteria was represented by eight orders and nine families. The majority of the Gammaproteobacteria reads belong to the orders Aeromonadales, Enterobacteriales, Pseudomonadales and Xanthomonadales, furthermore Enterobacteriales dominated over Pseudomonadales, followed by Xanthomonadales and Aeromonadales (Supplementary Fig. S5). The four orders were present in G, H, M, S and Y. Notably, Oceanospirillales was substituted for Aeromonadales in F area. Most Betaproteobacteria reads were affiliated with the orders Burkholderiales and Nitrosomonadales, in which the former possessed a higher relative abundance than the later. The orders Nitrosomonadaceae and

Oxalobacteraceae were presented in G, H, M and Y areas, but Betaproteobacteria was only represented by the orders Burkholderiales in F and S areas the order. Enterobacteriales was chiefly represented by the families Enterobacteriaceae (Supplementary Table S3). Burkholderiales was principally represented by the families Alcaligenaceae, Burkholderiaceae, Comamonadaceae and Oxalobacteraceae. The relative abundance of many families varied large among F, G, H, M, S and Y areas (Fig. 5A).

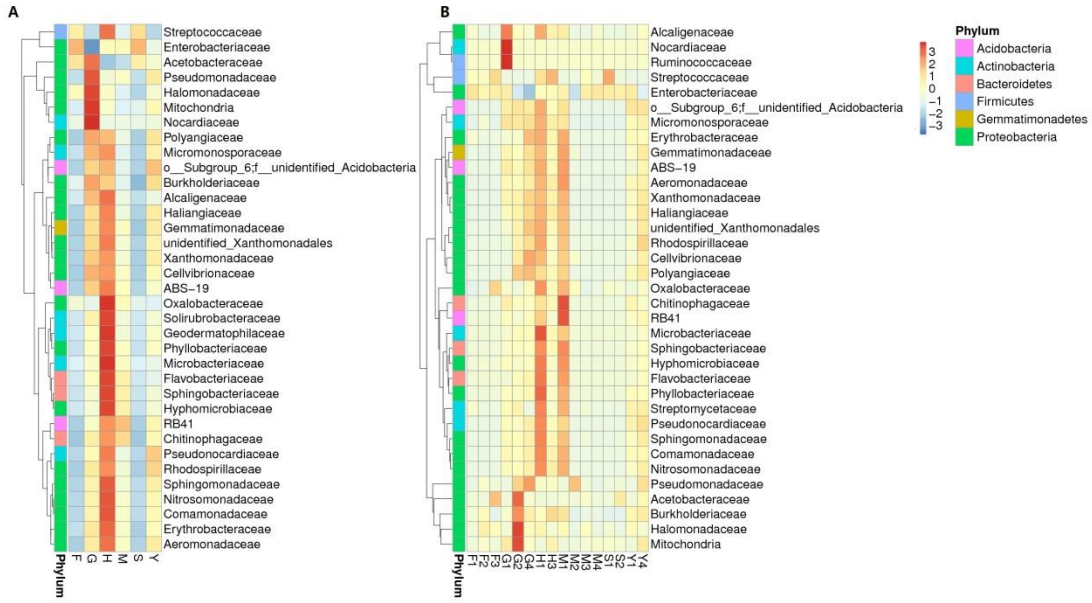

Fig.5. Heat maps showing bacterial family frequency distribution in six areas. The different color intensities represent the relative bacterial abundance in each groups.

Taxonomic distribution of fungi

Seventeen classes were classified into at least 3 phyla (Ascomycota, Basidiomycota and Zygomycota), as well as some unidentified groups (Supplementary Table S6). Among these phyla, Ascomycota dominated over Basidiomycota, followed by Zygomycota. These phyla were represented in all grape epidermis. The Ascomycota had the largest number of reads in each grape epidermis, together constituting more than 70% of the total fungi population in each group of samples (Fig. 7A).

Furthermore, the common and exclusive fungi OTUs of the grape epidermis were showed by Venn diagrams. Fungal OTUs distribution of grape epidermis in F, G, H, M, S and Y culture areas were presented 164 (47.0%), 228 (65.3%), 174 (49.9%), 253 (72.5%), 156 (44.7%) and 163 (46.7%), respectively (Supplementary Fig. 6). As for

different grape cultivar in the same area, the difference of fungal OTUs distribution was obviously showed. For instance, the OTUs of Cabernet Sauvignon, Merlot, Italian Riesling and Cabenet Franc cultivar in Manasi was 176, 139, 127 and 169, respectively, and 79 fungal OTUs were common (Supplementary Fig. 6A). Similarly, fungal OTUs distribution of same grape cultivar in different culture areas were also compared (Fig. 6). For example Cabernet Sauvignon cultivar, exclusive fungi OTUs of grape epidermis in F, G, H, M, S and Y culture areas were presented 4, 27, 10, 33, 12 and 16, respectively, and 64 fungal OTUs were common (Fig. 6A). The results indicated that fungal OTUs distribution of grape epidermis affected by the cultivation area and cultivar.

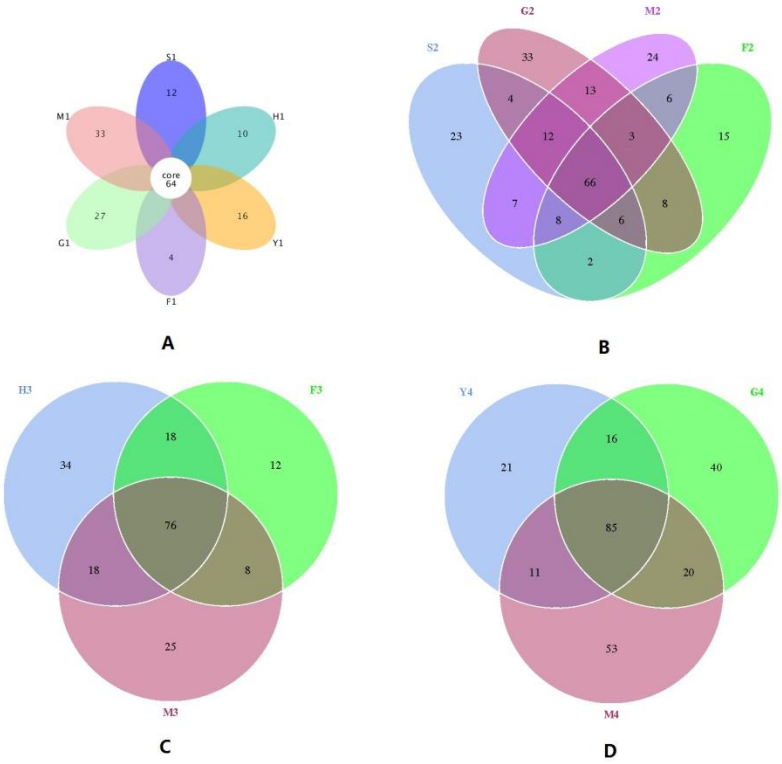

Fig.6. Venn diagrams showing the common and exclusive fungal OTUs of the grape epidermis. A: Cabernet Sauvignon; B: Merlot; C: Italian Riesling; D: Cabenet Franc.

However, those common and exclusive fungi OTUs truly shaped fungal community constructure. Particularly, mainly fungi OTUs were presented by the top 10 of relative abundance fungi genera, including *Aureobasidum*, *Alternaria*, *Botrytis*, *Rhodotorula*, *Cryptococcus*, *Hanseniaspora*, *Mucor*, *Fusarium*, *Chaetopyrena*,

Cladosporium. The relative abundance of fungi in different regions showed more significant differences. For example, the relative abundance of *Aureobasidum* in F, G, H, M, S and Y culture areas were 48.8%, 45.4%, 13.3%, 36.7%, 12.4% and 7.4%, respectively. Thus, fungal abundance of grape epidermis was affected by different cultivation area (Fig. 7).

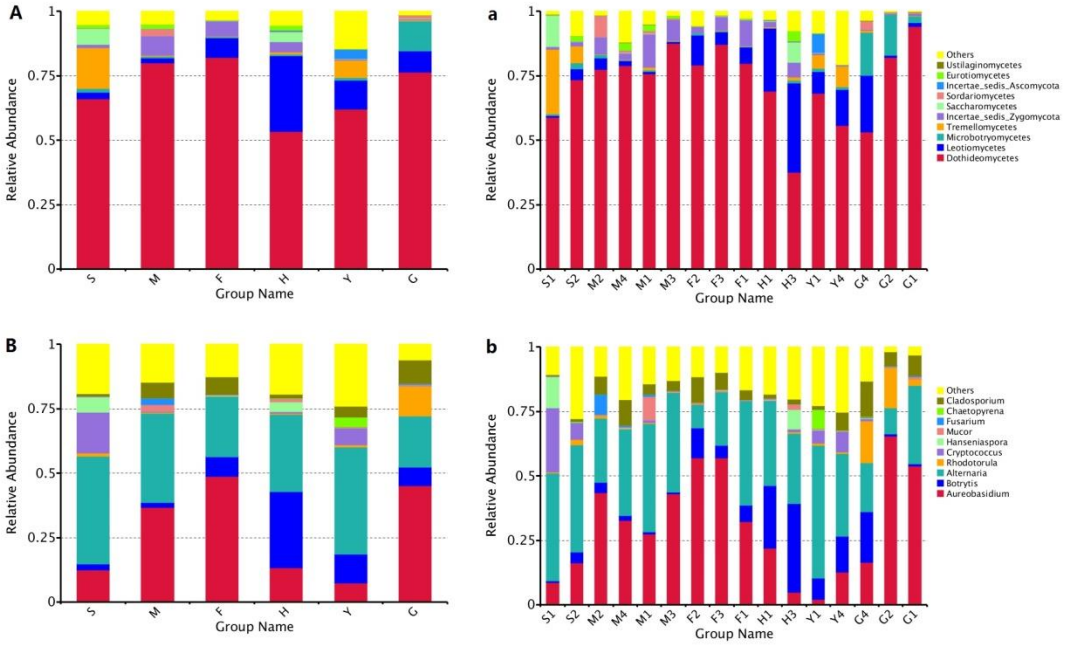

Fig.7. Relative abundance of top 10 fungal OTUs in fungal class (A/a) and genus (B/b) level.

Linear discriminant analysis Effect Size (LEfSe) has been widely used to discover high-dimensional biomarkers and reveal genomic features. In this study, it was used to estimate effects of the abundance of each component on the microbial population difference between various samples[48]. The differential features were identified on the OTU level. The six cultivation areas were used as the class of subjects. LEfSe revealed 6, 9, 6, 11, 13 and 5 fungal clades on samples from F, G, H, M, S and Y, respectively, as well as the statistically significant differences of fungal communities between six areas (Fig.8). The most differentially abundant fungal taxa in F, G, H, M, S and Y belonged to Dothideomycetes (Ascomycota), Microbotryomycetes (Basidiomycota), Leotiomycetes (Ascomycota), Sordariomycetes (Ascomycota), Tremellomycetes (Basidiomycota) and Incertae-sedis-Ascomycota (Ascomycota), respectively (Fig.7A). The overrepresented

clades of M also included Choanephora (Mucorales, Zygomycota) and Fusarium (Hypocreales, Sordariomycetes), which were different from those of F (Aureobbasidium, Dothioraceae, Dothideomycetes), G (Rhodotorula, Sporidiobolales, Microbotryomycetes; Cladosporium, Capnodiales, Dothideomycetes), H (Botrytis, Helotiales, Leotiomycetes), S (Cryptococcus, Tremellales, Tremellomycetes; Hanseniaspora, Saccharomycetales, Saccharomycetes) and Y (Chaetopyrena, Ascomycota). The Beta diversity of these communities was also approved by Metastats results at the genus level.

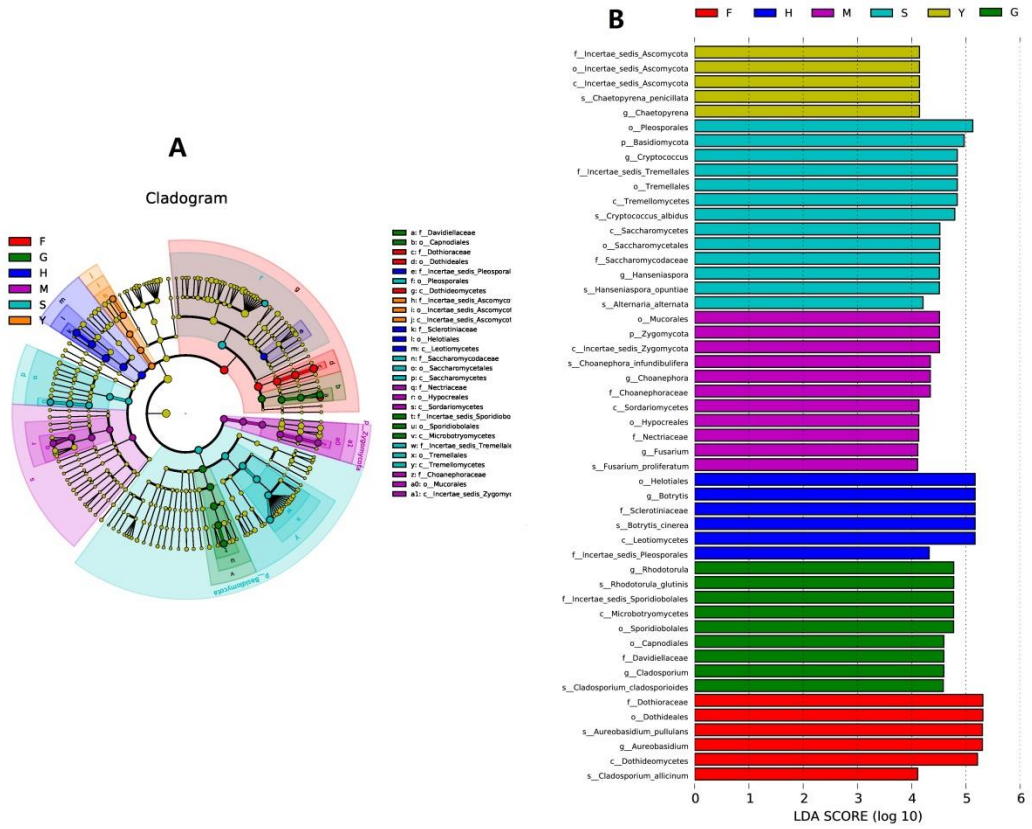

Fig. 8. LEfSe results on grape epidermis microbiomes. The cladogram reports the taxonomic representation of statistically and biologically consistent differences between F, G, H, M, S and Y fungal communities.

Ascomycota

Sequence reads of Ascomycota could be classified into nine classes, including 24 orders, 51 families and 104 genera, in addition to some unknown classes.

Dothideomycetes, Leotiomycetes, Saccharomycetes, Sordariomycetes, Incertae sedis Ascomycota and Eurotiomycetes classes were together present in six areas. Among the classes, Dothideomycetes represented Ascomycota in all samples. Of course, other classes with few richness on grape epidermis were also incompletely present in all cultivation regions. For instance, both Orbiliomycetes and Pezizomycetes classes were present in G and M areas, but only Taphrinomycetes was uniquely shown in S (Fig. 8B). The mean relative abundance of Leotiomycetes in H (29.55%; Fig. 7A) was significantly higher than those in F (7.59%), G(8.04%), M(1.97%), S(2.43%) and Y(11.19%) (AMOVA, p 0.001), so was that of Saccharomycetes (S 6.11%, M 0.09%, F 0.12%, H 0.22%, Y 0.05%, G 0.07%; $p<0.05$). In contrast, the mean relative abundance of Sordariomycetes, Incertae sedis Ascomycotawas and Eurotiomycetes were considerably lower in each aeras.

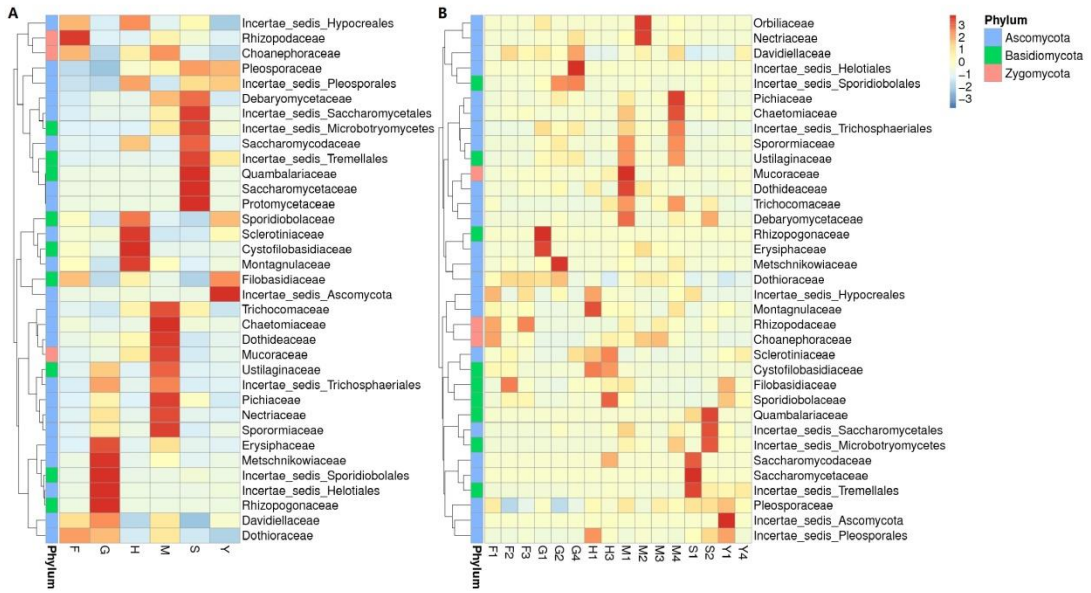

Fig. 9. Heat maps showing fungal family frequency distribution in six areas. The different color intensities represent the relative bacterial abundance in each groups.

The dominating Dothideomycetes orders were Dothideales, Pleosporales and Capnodiales. The abundances of Capnodiales and Dothideales in F, G and M were significantly higher than those in H, S and Y, which was contrary to Pleosporales. Helotiales and Erysiphales were predominant Leotiomycetes orders, while Saccharomycetales and Hypocreales separately dominated Saccharomycetes and

Sordariomycetes. There were 4, 5, 5, 5, 4 and 5 orders in F, G, H, M, S and Y, respectively. The relative abundance of prevalent orders in different grape epidermis was extremely distinct. For example, Helotiales was shared in each area, while 29.54% abundance was shared in H, following Y (11.18%), G (7.85%), F (7.59%), S (2.41%) and M (1.87%).

There were 4, 5, 6, 5, 5 and 6 dominating families in F, G, H, M, S and Y, respectively. The relative abundance of the same family varied greatly in different grape surface (Fig. 9). For instance, the mean relative abundance of Dothioraceae (Dothideales) was significantly higher in F (50.26%), G (45.33%) and M (36.77%) than in H (14.01%), S(13.02%), and Y(8.28%) ($p < 0.05$; Figs. 9A and Supplementary Figs. S7D), which was contrary that of Incertae_sedis_Pleosporales (Pleosporales; H 5.07%, Y 4.31%, S 3.95%, M 0.66%, F 0.46%, G 0.20%; Fig. 9A).

Basidiomycota

Seven classes were identified, and 4, 7, 3, 4, 5 and 5 classes were separately found in F, G, H, M, S and Y. Basidiomycota was mainly represented by Microbotryomycetes and Tremellomycetes classes, which were completely presented in each region (Fig. 7A). Atractiellomycetes, Cystobasidiomycetes, Exobasidiomycetes, Ustilaginomycetes, and Incertae_sedis_Basidiomycota were extremely low abundance. Microbotryomycetes class could be classified into three families belonging to two orders and some unclassfied groups. Analogously, Tremellomycetes was represented by four orders and four families. The predominated Microbotryomycetes order was Sporidiobolales, and its relative abundance was significantly higher in G (11.70%) than in S (1.32%), Y (0.88%), M (0.55%), F (0.46%), and H (0.44%) ($P<0.05$, Supplementary Figs. S7E). Tremellales, dominating Tremellomycetes order, was obviously richer in S (15.80%) than in Y (6.58%), H (0.79%), M (0.38%), G (0.37%) and F (0.13%) ($P<0.05$, Supplementary Figs. S7F). Sporidiobolales and Tremellales were respectively controlled by Incertae sedis Sporidiobolales and Incertae sedis Tremellales, and their abundance distribution bring into correspondence with that of

class (Fig. 7A).

Cultivars and environment conditions shaping microbial community composition

To elucidate relation between growing region, cultivar, climate, and microbial biogeography, the climate and geography data of six growing region in 2016 year were collected from Statistical Yearbook of Xinjiang Uygur Autonomous Region, Meteorological Data Center of China Meteorological Administration (http://data.cma.cn) and Altitude Information Inquiry Network (http://haiba.qhdi.com) (Table 3).

Table 3 The environment conditions of collecting sample site.

Climatic conditions	F	G	H	M	S	Y
Frost-free	172	181	187	175	216	184
Dryness	8.32	5.92	41.35	10.64	97.84	42.58
Altitude	552	800	1094	462	381	1059
Longitude	87.99	80.41	86.86	86.22	90.12	86.57
Latitude	44.16	44.21	42.27	44.3	42.87	42.06
Average temperature	21.6	21.69	20.22	19.97	27.17	21.47

The former analysis showed that alpha fungi diversity might be correlated with environment factors in six regions. Then, spearman correlation analysis was used to further investigate correlation between microbial species richness (alpha diversity) and environment factors. ACE, chao1, observed_species and Goods_coverage indices were significantly affected by altitude (A), frost-free (FR) and longitude (LO). LO were negatively correlated with ACE, chao1 and observed_species indices, while similarly Goods_coverage indices were also correlated with A. In contrast, ACE, chao1 and observed_species indices were significantly correlated with A, whereas Goods_coverage indices and A were negatively correlated. As for other indices, the correlation was not significant (supplementary Fig. 14A).

The relationship between bacterial community composition and environment factors in each area was explored by the canonical correspondence analysis (CCA). The CCA indicated that the A, LA and LO showed significant relationships with

bacterial taxonomy composition (supplementary Fig. 15A). The first two axes explained 87.85% of the taxonomic information. Variance partitioning CCA further suggested that env2 (including A, V, LA and LO factors) were the most important factors affecting bacterial taxonomy composition. Furthermore, Variance partitioning of the CCA showed that 15.16% of the total variability in the bacterial community composition were explained by the environment variables (supplementary Fig. 16A).

Similarly, alpha fungi diversity was also affected by those factors. For instance, LO were subjected to negative effect on ACE, chao1 and observed_species indices, opposite to Shannon, Simpson and Goods_coverage indices. However, Degree of dryness (D) was positively correlated with Shannon, Simpson and Goods_coverage indices (supplementary Fig. 14B). The canonical correspondence analysis (CCA) also indicated that the A, D, FR, LA and LO showed significant relationships with fungal taxonomy composition (supplementary Fig. 15B). The first two axes explained 59.88% of the taxonomic information. What's more, Variance partitioning CCA further suggested that env2 (including A, V, LA and LO factors) may be the most important factors affecting bacterial taxonomy composition. Furthermore, Variance partitioning of the CCA showed that 34.73% of the total variability in the fungal community composition were explained by the environment variables (supplementary Fig. 16B). Likewise, Mantel test indicated that the fungal community composition was significantly correlated to LO, FR, D and A rather than V and T，which was consistent with overlapping areas in variance partitioning analysis.

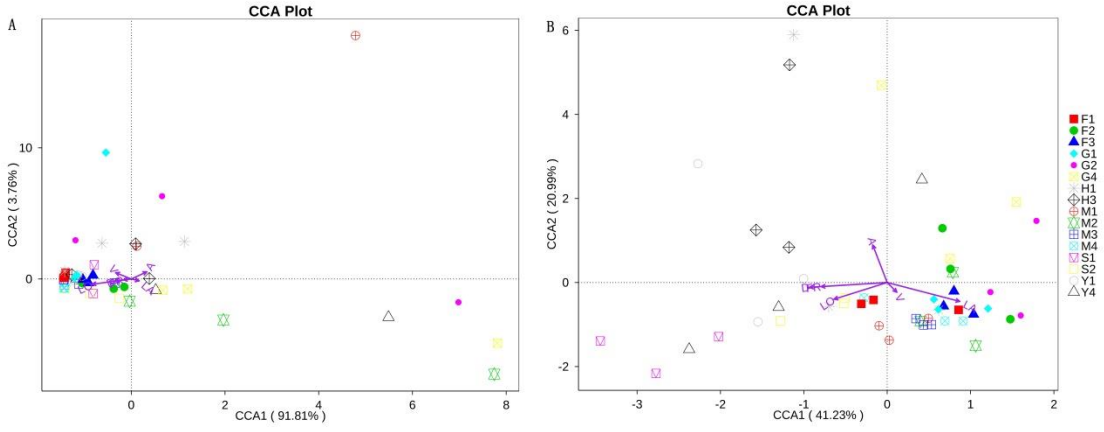

Fig. 9 Based on bacterial (A) and fungal (B) OTUs for CCA analysis. The different shapes represent the sample groups in different environments or conditions; the arrows indicate the environmental and climatic factors; the angle between the species and the environmental and climatic factor represents the positive and negative correlation between the species and the environmental and climatic factors (acute angle: positive correlation; obtuse: negative correlation; right angle: no correlation).

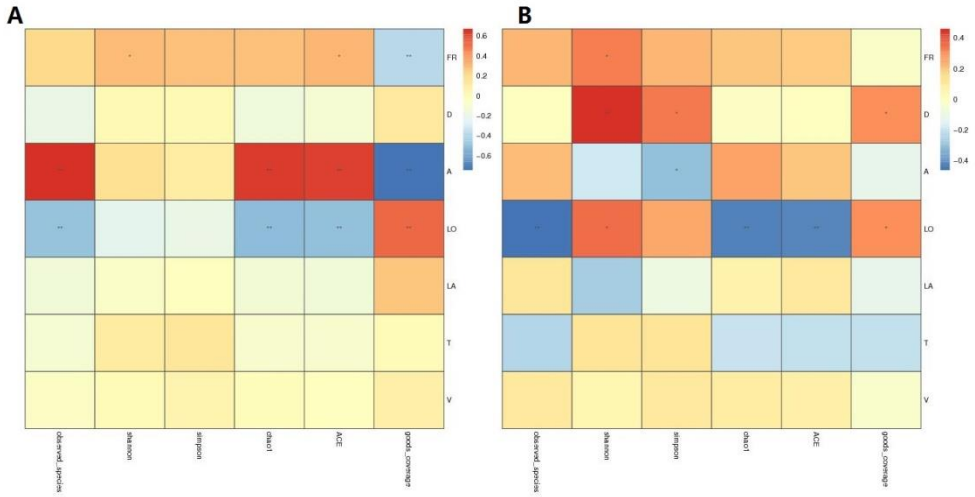

Fig. 10 Spearman correlation analysis studying correlation between microbial species richness (alpha diversity) and environment factors. The different color intensities represent species richness. A, bacteria; B, fungi. *, significant; **, extremely significant. Frost-free period, FR; Degree of dryness, D; Altitude, A; Degree of longitude, LO; Degree of latitude, LA; Average temperature, T; Cultivar, V.

Discussion

Grape berries were colonized by a wide array of epiphytic microorganisms, such as yeast, fungi and filamentous fungi, which played a major role in crop health and contributed to wine quality through participating in winemaking [16,49,50]. For instance, the non-Saccharomyces yeasts isolated from grapes, such as *Hanseniaspora*, *Candida*, *Rhodotorula* and *Cryptococcus*, have been proved to contribute to composition, sensory properties and flavor of wines [51,52]. In addition，some strains from *Botrytis cinerea* and *Neofusicoccum parvum* could devastate fungal pathogens of grapevines [53]. Later research also showed that the microbial diversity was subjected to affect from other factors such as geographic patterns, environment and climate [23]. Therefore，we attempted to study the bacterial/fungal community structure of wine grape surface and reveal the relation between the microbial diversity and these factors by using Illumina high-throughput sequencing and subsequent analysis.

This study showed a culture independent analysis of the microbes on wine grape epidermis from different culture regions and multiple cultivars. We comprehensively investigated the microbial population of wine grapes surface. In 48 wine grape samples from 4 wine grape cultivars and 6 wine-growing regions, 691 bacterial OTUs and 349 fungal OTUs were identified. Comparing to epidermis of apples, peaches and strawberries, the wine grape epidermis had a relatively higher diversity of fungi and bacteria [54]. Comparing to other fruits, the bacterial OTUs richness of wine grape epidermis in some regions (H, Y and G) was obviously higher, but bacterial community diversity was lower according to Shannon indices [55]. Although the fungal OTUs richness of each wine grape epidermis was generally lower than those of apples and blackcurrant, fungal community abundance was unanimously closed to apples and blackcurrant according to Shannon indices [56]. Additionally, the fungal OTUs richness of the wine grape epidermis was comparable to those of the grape berries in previous study, but the very high Shannon indices of the latter suggested the high community diversity [57]. These results suggested the abundant microbial diversity of the wine grape epidermis samples in this study.

The diversity of microbe inhabiting wine grape epidermis varied with different regions and grape cultivars. The bacterial community composition was similar across different samples of the same regions, but exceedingly differed from the different regions (Fig. 1 and Supplementary Fig. 4). The bacterial community composition of Y region was similar to that of H and G regions, while the bacterial community composition of S region was similar to that of M and F regions, which might be influenced by their culture regions. Previous studies of grape surface have presented that fungi diversity profiles varied with grapevine culture regions[55]. Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria phyla as the most main phylum in grapes surface were previously reported [54], while our study only presents Proteobacteria as the most dominated phylum in each grape sample.

There are some studies of the community structure of grape surface associated fungi. Our study only presents Ascomycota, Basidiomycota, and Zygomycota as the most abundant fungi phylum in each grape sample.

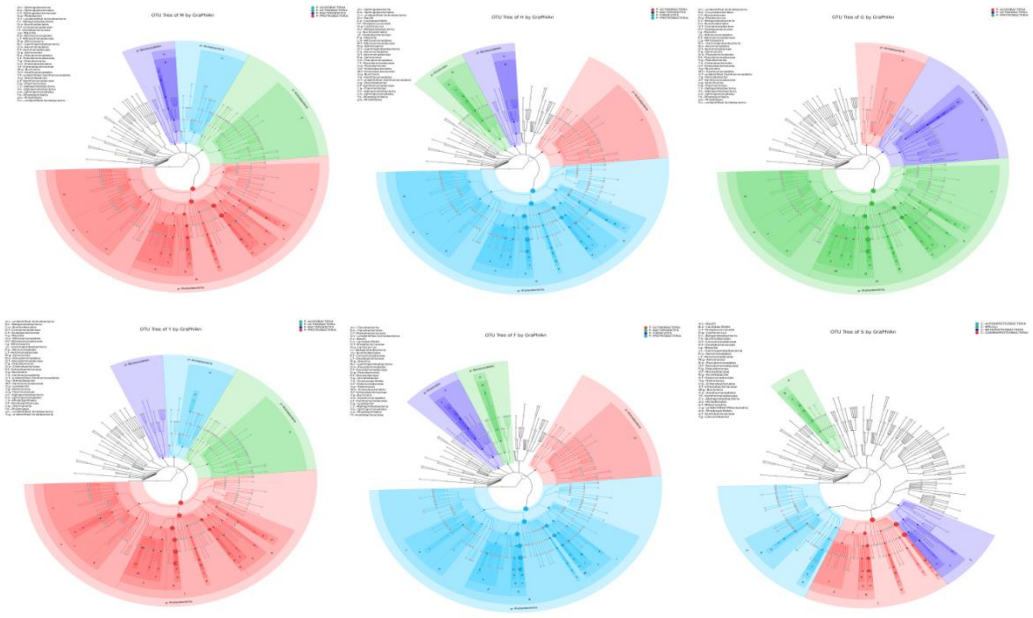

Fig. 11 Bacteria OTU tree graph of six cultivation areas based on GraPhlAn

We found that altitude (A), cultivar (V), degree of latitude (LA) and degree of longitude (LO) factors might be the most important factors affecting bacterial taxonomy composition, and the fungal community composition was significantly correlated to degree of longitude (LO), frost-free period (FR), degree of dryness (D) and altitude (A) rather than cultivar (V) and average temperature (T). Consequently, our study revealed the connection between the microbial diversity and these factors (geographic patterns, environment and climate) using Illumina high-throughput sequencing, which is in accord with previous study.

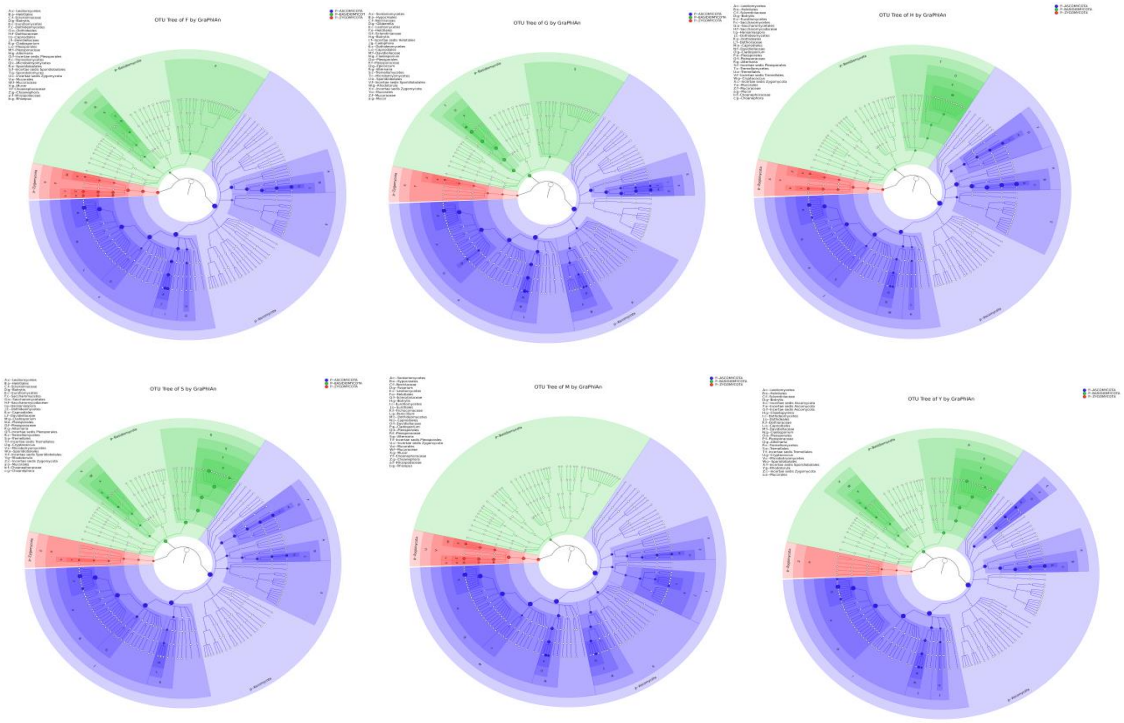

Fig. 12 Fungal OTU tree graph of six cultivation areas based on GraPhlAn

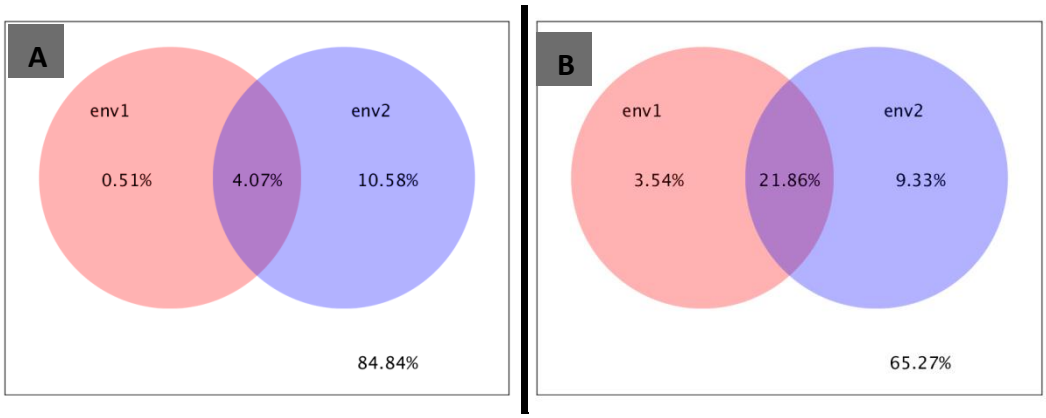

Fig. 13 Venn diagram of the variance partitioning analysis shows the relative effects of multiple variables on the composition of bacterial (A) and fungal (B) taxa. The impact factors of community composition are made up of env1 (including D, FR and T factors) and env2 (including A, V, LA and LO factors). The areas correspond to the amount of variance explained by each factor. Overlapping areas indicate shared variation of the parameter effect on community composition (A).

References

1 Karna L. Sacchi & Linda F. Bisson, a. D. O. A. A Review of the Effect of Winemaking Techniques on Phenolic Extraction in Red Wines. *American Journal of Enology and Viticulture* **56**（**3**）, 10 (2005).

2 Mirás-Avalos, J. M. *et al.* Influence of irrigation on consumer acceptability of Albariño and Godello wines. *LWT - Food Science and Technology* **85**, 345-352, doi:10.1016/j.lwt.2017.01.076 (2017).

3 Mira de Orduña, R. Climate change associated effects on grape and wine quality and production. *Food Research International* **43**, 1844-1855, doi:10.1016/j.foodres.2010.05.001 (2010).

4 Van Leeuw, R., Kevers, C., Pincemail, J., Defraigne, J. O. & Dommes, J. Antioxidant capacity and phenolic composition of red wines from various grape varieties: Specificity of Pinot Noir. *Journal of Food Composition and Analysis* **36**, 40-50, doi:10.1016/j.jfca.2014.07.001 (2014).

5 Hannah, L. *et al.* Climate change, wine, and conservation. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6907-6912, doi:10.1073/pnas.1210127110 (2013).

6 Gilbert, J. A., van der Lelie, D. & Zarraonaindia, I. Microbial terroir for wine grapes. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 5-6, doi:10.1073/pnas.1320471110 (2014).

7 Berbegal, C. *et al.* Starter cultures as biocontrol strategy to prevent Brettanomyces bruxellensis proliferation in wine. *Applied microbiology and biotechnology*, doi:10.1007/s00253-017-8666-x (2017).

8 Hierro, N., Gonzalez, A., Mas, A. & Guillamon, J. M. Diversity and evolution of non-Saccharomyces yeast populations during wine fermentation: effect of grape ripeness and cold maceration. *FEMS Yeast Res* **6**, 102-111, doi:10.1111/j.1567-1364.2005.00014.x (2006).

9 Wibowo, D., Eschenbruch, R., Davis, C. R., Fleet, G. H. & Lee, T. H. Occurrence and Growth of Lactic Acid Bacteria in Wine: A Review. *American Journal of Enology and Viticulture* **36**, 12 (1985).

10 Patarata, L., Pimentel, M. S., Pot, B., Kersters, K. & Faia, A. M. Identification of lactic acid bacteria isolated from Portuguese wines and musts by SDS-PAGE. *Journal of Applied Bacteriology* **76**, 288-293, doi:10.1111/j.1365-2672.1994.tb01630.x (1994).

11 Masqué, M. C. & Bordons, A. Isolation and selection of malolactic bacteria from southern Catalan wines. *Journal of Wine Research* **7**, 91-101, doi:10.1080/09571269608718069 (2007).

12 Christiaens, J. F. *et al.* The fungal aroma gene ATF1 promotes dispersal of yeast cells through insect vectors. *Cell Rep* **9**, 425-432, doi:10.1016/j.celrep.2014.09.009 (2014).

13 Tarantilis, P. A., Troianou, V. E., Pappas, C. S., Kotseridis, Y. S. & Polissiou, M. G. Differentiation of Greek red wines on the basis of grape variety using attenuated total reflectance Fourier transform infrared spectroscopy. *Food Chemistry* **111**, 192-196, doi:10.1016/j.foodchem.2008.03.020 (2008).

14 Romano, P., Fiore, C., Paraggio, M., Caruso, M. & Capece, A. Function of yeast species and strains in wine flavour. *International journal of food microbiology* **86**, 169-180, doi:10.1016/s0168-1605(03)00290-3 (2003).

15 Cordero-Bueso, G. *et al.* Influence of the farming system and vine variety on yeast communities associated with grape berries. *International journal of food microbiology* **145**,

132-139, doi:10.1016/j.ijfoodmicro.2010.11.040 (2011).

16 Martins, G. *et al.* Influence of the farming system on the epiphytic yeasts and yeast-like fungi colonizing grape berries during the ripening process. *International journal of food microbiology* **177**, 21-28, doi:10.1016/j.ijfoodmicro.2014.02.002 (2014).

17 Liu, Y. *et al.* Wine microbiome: A dynamic world of microbial interactions. *Crit Rev Food Sci Nutr* **57**, 856-873, doi:10.1080/10408398.2014.983591 (2017).

18 Garbeva, P., van Veen, J. A. & van Elsas, J. D. Microbial diversity in soil: selection microbial populations by plant and soil type and implications for disease suppressiveness. *Annu Rev Phytopathol* **42**, 243-270, doi:10.1146/annurev.phyto.42.012604.135455 (2004).

19 Xie, X., Zhang, H. & Pare, P. Sustained growth promotion in Arabidopsis with long-term exposure to the beneficial soil bacteriumBacillus subtilis(GB03). *Plant Signaling & Behavior* **4**, 948-953, doi:10.4161/psb.4.10.9709 (2014).

20 Castro, H. F., Classen, A. T., Austin, E. E., Norby, R. J. & Schadt, C. W. Soil microbial community responses to multiple experimental climate change drivers. *Applied and environmental microbiology* **76**, 999-1007, doi:10.1128/AEM.02874-09 (2010).

21 Wallenstein, M. D. & Hall, E. K. A trait-based framework for predicting when and where microbial adaptation to climate change will affect ecosystem functioning. *Biogeochemistry* **109**, 35-47, doi:10.1007/s10533-011-9641-8 (2011).

22 Gayevskiy, V. & Goddard, M. R. Geographic delineations of yeast communities and populations associated with vines and wines in New Zealand. *ISME J* **6**, 1281-1290, doi:10.1038/ismej.2011.195 (2012).

23 Bokulich, N. A., Thorngate, J. H., Richardson, P. M. & Mills, D. A. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E139-148, doi:10.1073/pnas.1317377110 (2014).

24 Goldberg, C. S. *et al.* Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution* **7**, 1299-1307, doi:10.1111/2041-210x.12595 (2016).

25 Ubeda, J., Maldonado Gil, M., Chiva, R., Guillamon, J. M. & Briones, A. Biodiversity of non-Saccharomyces yeasts in distilleries of the La Mancha region (Spain). *FEMS Yeast Res* **14**, 663-673, doi:10.1111/1567-1364.12152 (2014).

26 Hao, D. C., Song, S. M., Mu, J., Hu, W. L. & Xiao, P. G. Unearthing microbial diversity of Taxus rhizosphere via MiSeq high-throughput amplicon sequencing and isolate characterization. *Sci Rep* **6**, 22006, doi:10.1038/srep22006 (2016).

27 Heydenreich, F. M. *et al.* High-throughput mutagenesis using a two-fragment PCR approach. *Sci Rep* **7**, 6787, doi:10.1038/s41598-017-07010-4 (2017).

28 Zetsche, B. *et al.* Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat Biotechnol* **35**, 31-34, doi:10.1038/nbt.3737 (2017).

29 Nones, K. *et al.* Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. *Nat Commun* **5**, 5224, doi:10.1038/ncomms6224 (2014).

30 Kao, C. M. *et al.* The change of microbial community from chlorinated solvent-contaminated groundwater after biostimulation using the metagenome analysis. *J Hazard Mater* **302**, 144-150, doi:10.1016/j.jhazmat.2015.09.047 (2016).

31 Zhai, Y., Wang, W., Tan, H. & Cao, L. A New Approach to Analyzing Endophytic Actinobacterial

Population in the Roots of Banana Plants (Musa sp., AAA). *Journal of Biochemistry and Molecular Biology Research* **2**, 180-184, doi:10.17554/j.issn.2313-7177.2016.02.30 (2016).

32 El-Ashram, S. & Suo, X. Exploring the microbial community (microflora) associated with ovine Haemonchus contortus (macroflora) field strains. *Sci Rep* **7**, 70, doi:10.1038/s41598-017-00171-2 (2017).

33 Chen, J. *et al.* Characterization of the gut microbiota in the golden takin (Budorcas taxicolor bedfordi). *AMB Express* **7**, 81, doi:10.1186/s13568-017-0374-5 (2017).

34 Li, J. *et al.* Distribution of picoplankton in the northeastern South China Sea with special reference to the effects of the Kuroshio intrusion and the associated mesoscale eddies. *Sci Total Environ* **589**, 1-10, doi:10.1016/j.scitotenv.2017.02.208 (2017).

35 Marsh, A. J., O'Sullivan, O., Hill, C., Ross, R. P. & Cotter, P. D. Sequencing-based analysis of the bacterial and fungal composition of kefir grains and milks from multiple sources. *PLoS One* **8**, e69371, doi:10.1371/journal.pone.0069371 (2013).

36 Liu, W. *et al.* High-throughput sequencing for the detection of the bacterial and fungal diversity in Mongolian naturally fermented cow's milk in Russia. *BMC Microbiol* **15**, 45, doi:10.1186/s12866-015-0385-9 (2015).

37 Horn, S., Hempel, S., Verbruggen, E., Rillig, M. C. & Caruso, T. Linking the community structure of arbuscular mycorrhizal fungi and plants: a story of interdependence? *ISME J* **11**, 1400-1411, doi:10.1038/ismej.2017.5 (2017).

38 Klaedtke, S. *et al.* Terroir is a key driver of seed-associated microbial assemblages. *Environmental microbiology* **18**, 1792-1804, doi:10.1111/1462-2920.12977 (2016).

39 Hollister, E. B., Schadt, C. W., Palumbo, A. V., James Ansley, R. & Boutton, T. W. Structural and functional diversity of soil bacterial and fungal communities following woody plant encroachment in the southern Great Plains. *Soil Biology and Biochemistry* **42**, 1816-1824, doi:10.1016/j.soilbio.2010.06.022 (2010).

40 Yan, D. *et al.* Diversity and Composition of Airborne Fungal Community Associated with Particulate Matters in Beijing during Haze and Non-haze Days. *Front Microbiol* **7**, 487, doi:10.3389/fmicb.2016.00487 (2016).

41 Nicola Segata *et al.* Metagenomic biomarker discovery and explanation.pdf>. *Genome Biology* **12:R60**, 18 (2011).

42 Sterkenburg, E., Bahr, A., Brandstrom Durling, M., Clemmensen, K. E. & Lindahl, B. D. Changes in fungal communities along a boreal forest soil fertility gradient. *New Phytol* **207**, 1145-1158, doi:10.1111/nph.13426 (2015).

43 Liu, J. *et al.* High throughput sequencing analysis of biogeographical distribution of bacterial communities in the black soils of northeast China. *Soil Biology and Biochemistry* **70**, 113-122, doi:10.1016/j.soilbio.2013.12.014 (2014).

44 Xiao, E. *et al.* Depth-resolved microbial community analyses in two contrasting soil cores contaminated by antimony and arsenic. *Environmental pollution* **221**, 244-255, doi:10.1016/j.envpol.2016.11.071 (2017).

45 Jian Yang, Li'an Ma, Hongchen Jiang, Geng Wu & Dong, H. Salinity shapes microbial diversity and community structure in surface sediments of the Qinghai-Tibetan Lakes.pdf. *Scientific Reports*, 6, doi:10.1038/srep25078 (2016).

46 Hayden, C. J. & Beman, J. M. Microbial diversity and community structure along a lake elevation gradient in Yosemite National Park, California, USA. *Environmental microbiology* **18**,

1782-1791, doi:10.1111/1462-2920.12938 (2016).

47 Jia, H. R. *et al.* The effects of Bt Cry1Ie toxin on bacterial diversity in the midgut of Apis mellifera ligustica (Hymenoptera: Apidae). *Sci Rep* **6**, 24664, doi:10.1038/srep24664 (2016).

48 Chenhong Zhang *et al.* Structural modulation of gut microbiota in life-long calorie-restricted mice. *Nature Communications*, doi:DOI: 10.1038/ncomms3163 (2013).

49 Verginer, M., Leitner, E. & Berg, G. Production of volatile metabolites by grape-associated microorganisms. *J Agric Food Chem* **58**, 8344-8350, doi:10.1021/jf100393w (2010).

50 Nisiotou, A. A., Rantsiou, K., Iliopoulos, V., Cocolin, L. & Nychas, G. J. Bacterial species associated with sound and Botrytis-infected grapes from a Greek vineyard. *International journal of food microbiology* **145**, 432-436, doi:10.1016/j.ijfoodmicro.2011.01.017 (2011).

51 Hong, Y. A. & Park, H. D. Role of non-Saccharomyces yeasts in Korean wines produced from Campbell Early grapes: potential use of Hanseniaspora uvarum as a starter culture. *Food microbiology* **34**, 207-214, doi:10.1016/j.fm.2012.12.011 (2013).

52 Quiros, M., Rojas, V., Gonzalez, R. & Morales, P. Selection of non-Saccharomyces yeast strains for reducing alcohol levels in wine by sugar respiration. *International journal of food microbiology* **181**, 85-91, doi:10.1016/j.ijfoodmicro.2014.04.024 (2014).

53 Haidar, R. *et al.* Multi-organ screening of efficient bacterial control agents against two major pathogens of grapevine. *Biological Control* **92**, 55-65, doi:10.1016/j.biocontrol.2015.09.003 (2016).

54 Jonathan W. Leff & Fierer, N. Bacterial Communities Associated with the Surfaces of Fresh Fruits and Vegetables. *Plos One* **8**, 9, doi:10.1371/journal.pone.0059310 (2013).

55 Portillo Mdel, C., Franques, J., Araque, I., Reguant, C. & Bordons, A. Bacterial diversity of Grenache and Carignan grape surface from different vineyards at Priorat wine region (Catalonia, Spain). *International journal of food microbiology* **219**, 56-63, doi:10.1016/j.ijfoodmicro.2015.12.002 (2016).

56 Vepštaitė-Monstavičė, I. *et al.* Distribution of apple and blackcurrant microbiota in Lithuania and the Czech Republic. *Microbiological Research* **206**, 1-8, doi:10.1016/j.micres.2017.09.004 (2018).

57 Mezzasalma, V. *et al.* Grape microbiome as a reliable and persistent signature of field origin and environmental conditions in Cannonau wine production. *PLoS One* **12**, e0184615, doi:10.1371/journal.pone.0184615 (2017).

Acknowledgements

This work was supported by the National Natural Science Foundations of China (No. 31560446 and 31500092), and the Project of Science and Technology, Xinjiang Production and Construction Corps (No. 2015AB016, 2016AB009 and 2016AD024).

Author Contributions

F.G. collected wine grape samples and environment data, performed DNA extractions and PCRs, accomplished the bioinformatics analysis of the Illumina sequencing data, compiled the figures and tables, and wrote the complete first draft of the manuscript. B.W. contributed to advices and constructive critiques, reviewed results and corrected language style in the final manuscript. J.X., X.J. and X.S. corporately supervised and guided the research project. All authors reviewed the manuscript.