SCIENTIFIC REPORTS

natureresearch

OPEN

Understanding the phyllosphere microbiome assemblage in grape species (*Vitaceae*) with amplicon sequence data structures

Prashant Singh [ID], Sylvain Santoni, Audrey Weber, Patrice This & Jean-Pierre Péros

Impacts of plant genotype on microbial assemblage in the phyllosphere (above-ground parts of plants, which predominantly consists of the set of photosynthetic leaves) of *Vitis vinifera* cultivars have been studied previously but the impact of grape species (under the grape family *Vitaceae*) was never investigated. Considering the fact, that the phyllosphere microbiome may have profound effects on host plant health and its performance traits, studying the impact of grape species in microbial taxa structuring in the phyllosphere could be of crucial importance. We performed 16S and ITS profiling (for bacteria and fungi respectively) to access genus level characterization of the microflora present in the leaf phyllosphere of five species within this plant family, sampled in two successive years from the repository situated in the Mediterranean. We also performed α and β-diversity analyses with robust statistical estimates to test the impacts of grape species and growing year, over a two-year period. Our results indicated the presence of complex microbial diversity and assemblages in the phyllosphere with a significant effect of both factors (grape species and growing year), the latter effect is being more pronounced. We also compared separate normalization methods for high-throughput microbiome data-sets followed by differential taxa abundance analyses. The results suggested the predominance of a particular normalization method over others. This also indicated the need for more robust normalization methods to study the differential taxa abundance among groups in microbiome research.

Plant phyllosphere is the niche of a disparate group of microbial communities of prokaryotes, eukaryotes, and viruses that interact with each other and with their host plant[1]. Due to the limited nutrient availability and fluctuating climatic conditions, phyllosphere is a dynamic and stressful habitat for its microbial colonizers[2,3]. The knowledge of these colonizers and drivers that might affect their assemblage might explain the mechanisms that control operations at the intersection between plants, microbiome, and the atmosphere. However, phyllosphere has been largely overlooked in most of the grape related microbiome studies and researches were mainly focused on rhizosphere (root and soil colonizers) and endosphere (inside tissue colonizers)[4–9]. Most of these studies including few in phyllosphere[10–12] analyzed the main drivers for microbial assemblage on *Vitis vinifera*, the most widely cultivated species for wine and raisins. They suggested that the environmental conditions at different geographic locations (or terroir) and growing seasons are important drivers in the rhizosphere, endosphere or phyllosphere[10,11]. Cultivars or genotype interaction with the environment also seemed to play a role in this assemblage[10,12]. Since most of the current breeding schemes in grapevines involve other *Vitis* species as pest resistance sources, we wanted to inquire about the effect of the plant species on microbial assemblage. Furthermore, as the selection of microbiome could lead to new plant breeding strategies of next-generation, identifying microbiome differences among grape species could be an interesting opportunity for new grape breeding schemes to develop resistant, healthy and more productive varieties[13]. Moreover, grape associated epiphytes have been recently established as promising biocontrol agents (BCA) against fungal pathogens[14] of *Vitis vinifera*, the species level variation could offer us hints of new potential players for BCA.

In this work, we sampled epiphytes from five different species in the *Vitaceae*[15,16] to find major microbial colonizers in their phyllosphere and compared the impact of grape species and growing year on the phyllosphere microbiome assemblage. To mitigate the confounding effect of environmental conditions at different geographic

AGAP, University of Montpellier, INRA-SupAgro, CIRAD, Montpellier, France. Correspondence and requests for materials should be addressed to P.S. (email: patrice.this@inra.fr)

Number of samples	Reads in	Reads out (Filtered)	Denoised	Chimera removal	OTUs
16S Data					
30	5,568,565	4,538,503	4,139,738	3,9763,42	10,825
ITS Data					
30	2,7429,06	2,674,944	2,411,474	2,407,048	5,252

Table 1. The total number of reads at each step of bacterial microbiome data processing.

locations (or terroir), we sampled all five grape species (*Vitis vinifera cv. Cabernet-Sauvignon, Vitis pentagona, Vitis riparia, Muscadinia rotundifolia*, and *Parthenocissus quinquefolia*) from Montpellier-SupAgro repository of INRA in the Mediterranean. These species were chosen to study the genotype effect at species level within subgenus *Vitis*, at the subgenus level (subg. *Vitis* versus subg. *Muscadinia*) and at the genus level (*Vitis* versus *Parthenocissus*). The three species within subg. *Vitis* are genetically distant[15,17] and cover the distribution range of this subgenus in the Northern Hemisphere with the European *V. vinifera*, the Asian *V. pentagona* and the American *V. riparia*[16,18,19]. If the genotype is a major driver of the phyllosphere microbiome assemblage, we, therefore, expected increasing differences in microbial assemblages according to the genetic distance between plant species.

From a statistical point of view, we compared the differential abundance of microbial genera using three separate data normalization methods to predict the better normalization approach for grape related or other microbiome data. Normalization is the data transformation process enabling a true comparison of statistics from separate measurements by discarding artefactual biases from the original measurements. Most of the microbiome studies including grapevines[9,20] still use the most standard statistical methods for differential abundance analysis without testing the data distribution and transformation. In microbial ecology, rarefying samples to even sequencing depth is standard normalization method but is also not an ideal one, as it potentially reduces statistical power relying upon how much data is filtered and does not address the challenge of compositional data[21]. The log ratio transformation method is also widely accepted by researchers and statisticians and in various high dimensional studies[22,23]. Here, we used few recently published data transformation methods of cumulative sum scaling (CSS)[24], DESeq2[25] and log ratio to normalize our zero-inflated taxa abundance data and compared estimates of differentially abundant genera between two growing years and grape species.

Results

Phyllosphere exhibits diverse bacterial and fungal communities. Millions of amplicon reads were processed from both data-sets (16S and ITS) and 10825 bacterial and 5252 fungal amplicon sequence variants (or operational taxonomic units-OTUs) were obtained (Table 1). After assigning the OTUs to phylum level more than 73% bacterial OTUs and ~95% of fungal OTUs were assigned to phylum level. Unknown sequences corresponded to ~27% and ~5% for bacterial and fungal data respectively, meaning that the taxonomic assignment procedure was not able to assign any microorganism to these sequences. *Proteobacteria* (relative abundance ~15%) and *Cyanobacteria* (~14.8%) were the most dominated phyla across the samples followed by *Firmicutes* (~3%) and *Actinobacteria* (~1.3%). On the other hand, samples were heavily dominated by fungal phyla of *Ascomycota* (~91%) followed by *Basidiomycota* (~9%). After gloming of these OTUs at the genus level, 677 bacterial and 434 fungal genera were recovered. Out of these, *Sphingomonas, Methylobacterium, Rubelimicrobium, Blastococcus* and *Alternaria, Aureobasidium, Cladosporium, Lachnum* were most abundant bacterial and fungal genera, respectively (Fig. 1).

Microbial communities clustered distinctly with year and grape species. Multidimensional scaling (or principal coordinate analysis; PCoA) on microbial abundance data showed that the samples from each year clustered together and were distinct from each other (more predominant in ITS data, Fig. 2), confirmed the significant impact of the growing year on microbial community structuring in the phyllosphere. Clustering among grape species was not prominent (Fig. 2) but performing PCoA on the subset of the data (i.e. on separate data for each year) displayed a lower but significant impact of grape species in shaping phyllosphere microbiome, especially the fungal microbiome (Fig. 3). Permutational analysis of variance (PERMANOVA) statistics according to Year, Species and the interaction term (Year × Species), further confirmed the hypothesis (Table 2) that the synergy between the environment and the plant genotype could be held responsible for microbial community structuring in the phyllosphere.

Furthermore, *observed α-diversity* estimates of bacterial and fungal OTUs of each grape species (within each growing year, Fig. 4) revealed that the unique OTU-richness in the phyllosphere of each grape species significantly differed between the two years (Table 2). This reconfirmed the major impact of the growing year in shaping phyllosphere microbial assemblage. Although OTU-richness estimates did not vary according to grape species (P~0.05), the interaction term (Year × Species) displayed strong differences in the richness of unique OTUs (Table 2).

Comparing normalization methods. Most of the microbiome data doesn't follow a normal distribution for taxa abundance across the samples[9,20,25] and our data-sets were not the exceptions. Using three separate data transformation methods (square root, log-ratio, and CSS) we were unable to achieve proper normalization of our data-sets, and even after transformation, most of the distribution followed negative binomial distribution except CSS which performed better than others (Fig. 5). We used these transformed data for differential abundance analysis of each taxon (at genus level) according to the grape species and growing year (DESeq2 was applied to square root transformed data, to handle negative binomial distribution).
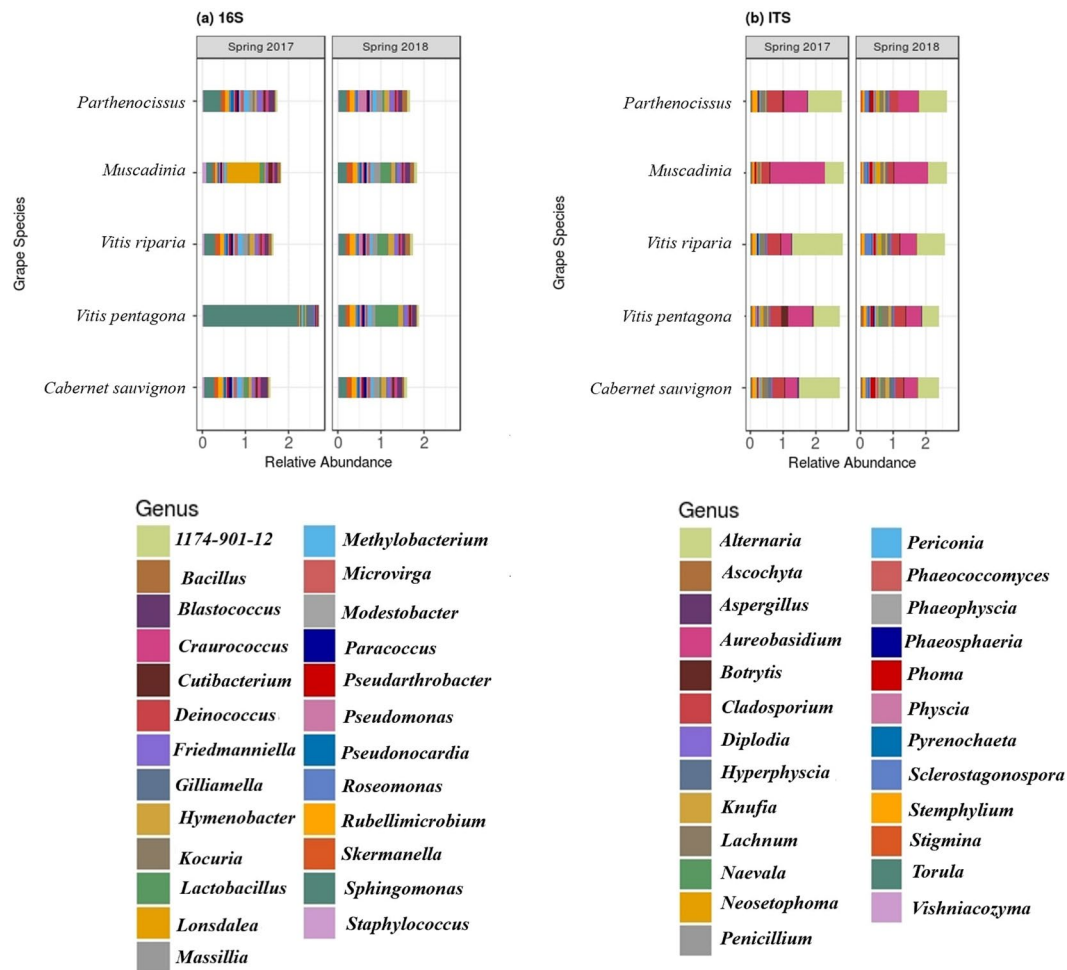
Figure 1. Mean (**a**) bacterial and (**b**) fungal relative abundance of top 25 genus present across samples i.e leaf phyllosphere of five grape species.
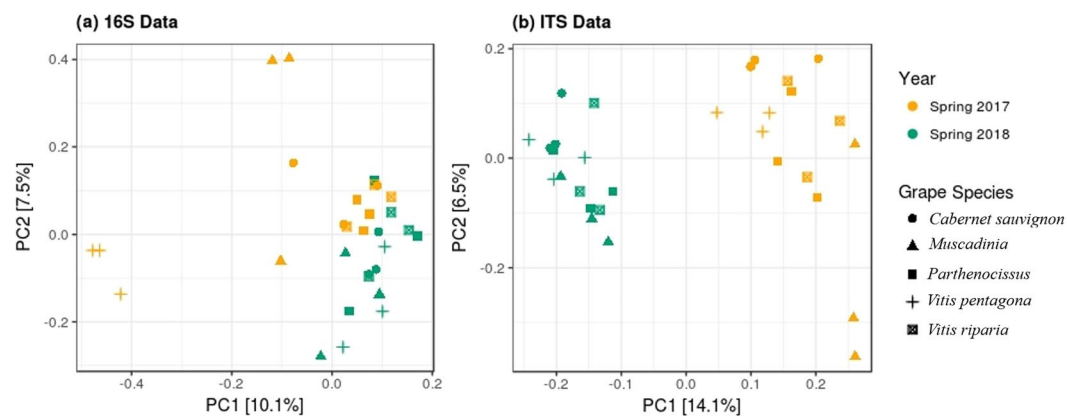
Figure 2. PCoA ordinations of (**a**) bacterial and (**b**) fungal communities derived from leaf phyllosphere at two growing years, using Bray-Curtis distance matrix. Both the axis explains ~20% of variations. The shape represents grape species (N = 30).

Performing multiple testing with FDR corrected P-values (adj-P value < 0.05) gave nine bacterial genera for DESeq2 as compared to two and three genera for log and CSS transformed data between the two years. Similarly, 45 fungal genera were obtained for DESeq2 as compared to 11 and 13 genera for other methods between the two years (Figs 6 and 7). The same testing was performed on data-sets of Spring 2017 and Spring 2018 separately
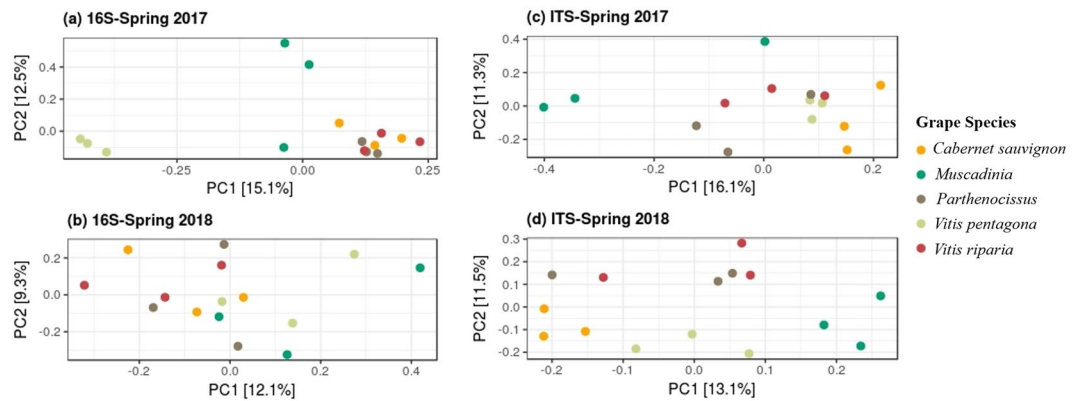
Figure 3. PCoA ordinations of bacterial (**a,b**) and fungal (**c,d**) communities derived from leaf phyllosphere at spring 2017 and spring 2018 separately, using Bray-Curtis distance matrix.

Data	ANOVA (on α-diversity measures)	PERMANOVA (on PCoA clusters)
16S		
Year	at F = 5.725, P = 0.0076**	R² = 0.269, F = 1.811, P = 0.0001***
Year × Grape Species	**at F = 9.022, P = 0.00138****	**R² = 0.154, F = 1.737, P = 0.0001****
Grape Species (Spring 2017)	at F = 3.752, P = 0.041	R² = 0.379, F = 1.525, P = 0.0001***
Grape Species (Spring 2018)	at F = 1.743, P = 0.217	R² = 0.304, F = 1.134, P = 0.031*
ITS		
Year	at F = 49.261, P = 1.24e-07	R² = 0.101, F = 3.532, P = 0.0001***
Year × Grape Species	**at F = 57.340, P = 2.71e-07**	**R² = 0.112, F = 3.767, P = 0.0001****
Grape Species (Spring 2017)	at F = 2.843, P = 0.08	R² = 0.325, F = 1.206, P = 0.0038**
Grape Species (Spring 2018)	at F = 1.274, P = 0.34	R² = 0.334, F = 1.257, P = 0.0001***

Table 2. Components predicting the impacts of Year and Grape Species on the leaf microbiome.

to identify differential taxa abundance among five grape species and the results were similar (Tables 3 and 4). DESeq2 gave a higher number of genera as compared to other methods (probably an overestimation). Combining normalization performance on taxa abundance across samples and identification of differentially abundant genera, indicated that CSS normalization method worked better and was statistically more robust, i.e. fewer false positives and lower false discovery rates, as compared to other methods.

Discussion

Microbial colonizers of endosphere and rhizosphere associated with *Vitis vinifera* cultivars (or grapevines) have been studied before quite extensively[4,8,9]. Recently, grapevine's phyllosphere has also been explored by researchers[10–12], but these habitats were not explored in the context of different grape species of *Vitaceae* prior to this study. At first, we explored the phyllosphere of five different grape species using the culture-independent methods of 16S and ITS profiling. Out of the complex microbial diversity, *Sphingomonas* and *Methylobacterium* were the most dominant bacterial genera. The fungal community was dominated by *Aureobasidium, Cladosporium* and *Alternaria*. These results are in line with the previous phyllosphere related works[26–28], which suggests that the phyllosphere is usually dominated by these genera because of their important functions (e.g- plant growth, plant defense, and biocontrol agent, etc.)[11,29]. For example, phyllosphere is usually quite rich in *Methylobacterium*, which can benefit the plant by absorbing the methanol (gaseous) released by the stomata and produces a growth promoting agent for the seedlings of some crop plant[29]. Biocontrol potential of some of these genera (e.g- *Sphingomonas* and *Aureobasidium*) have been reported previously[30–34] and should be explored in future studies in the context of grapes and in collaboration with other taxa with major or minor abundances. As the microbial communities in the phyllosphere can modulate leaf susceptibility to infections, it can protect the plant from foliar diseases[30,35]. Hence, the microflora of the phyllosphere and the local environmental conditions can thus regulate the fitness of their host plant. The current major challenge is to learn the properties of these microbial communities in the phyllosphere (e.g- taxonomic and functional diversity and microbial network structure) that will be beneficial under changing climatic conditions to cultivate properties of biocontrol and robustness in grape plants to sustain the productivity, yield, and resilience of agricultural systems. Among the diverse and complex communities found in the phyllosphere the genus, *Massilia* is also noteworthy as its presence was quite consistent during the two years. It is a major pollutant of an aerosol with agricultural applications[36,37] and could be an indicator of agricultural practices in the Mediterranean. The genus *Rubelimicrobium* was also one of the major and dominant genera found as a leaf epiphyte. Few species of this genus were isolated from soil[38], and this could be a shred of evidence to support the claim that the soil microbiota may influence the epiphyte compositions[9].
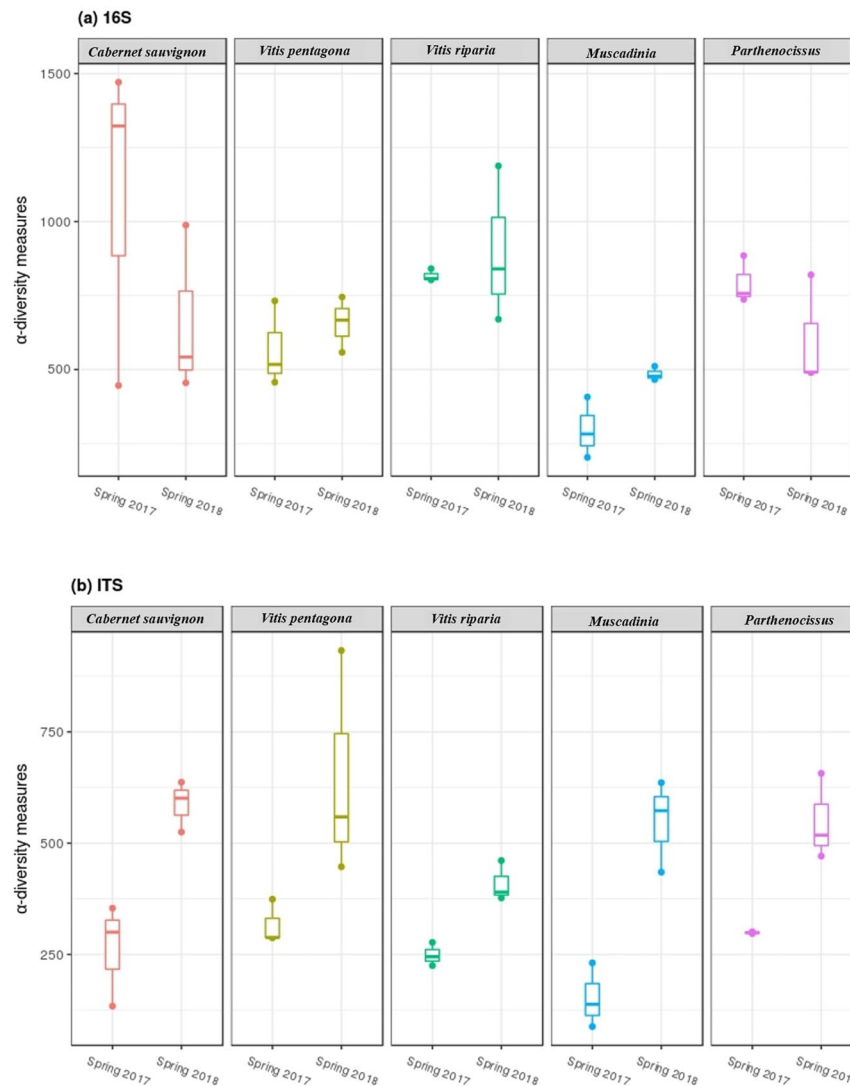
Figure 4. Box plots of mean *observed α-diversity* measures or unique OTU-richness between Spring of 2017 and 2018 for (**a**) bacterial and (**b**) fungal data. For each box plot, the top point is maximum observation, the lower point is minimum observation, top of the box is the third quartile, the bottom of the box is the first quartile, the middle bar is median values and color represent grape species (N = 30).

Secondly, we assessed the relative impacts of grape species and the growing year to shape the assemblage of the microbial communities in the phyllosphere. Morphological structures of leaves, its chemistry and physiology may differ according to the grape species as all these traits can be impacted by plant genetics, and this variation may also lead to different blends of microbial community structures[39,40] in the phyllosphere. In previous studies, minor impacts of grapevine genotypes (at cultivar level) have been found[10,12] and we expected greater impacts of genotypes once we sample from different species in order to increase genetic distances between them but our analysis indicated that the growing year had much stronger impacts than grape species in microbial community structuring. At each individual growing year, grape species also showed some influence in shaping this assemblage (especially for fungal assemblage) but that is noticeably weaker in comparison with the growing year. Statistical estimates of α and β-diversity suggest that the plausible hypothesis could be the species-climate interaction that is responsible for recruiting microbes on the leaf surface of different grape species. This interaction could be multilayered as slight changes in meteoclimatic parameters (daylight exposure, increased atmospheric CO_2 concentrations, and temperatures) between two growing years may alter certain important soil characteristics (moisture volume, amount of root exudates, soil temperature etc.), which in turn may induce major changes in soil microbiome and ultimately in epiphytic communities[9,41,42].

Lastly, we also analyzed the impacts of the data normalization methods to detect the differential taxa abundance among different groups. Microbiome datasets like taxonomy reads or OTU counts from amplicon sequencing experiments or differential expression (RNA-Seq) data are often scattered and have frequent zeros. In order to fit these overdispersed microbiome datasets, negative binomial (NB) is often applied[24,43,44]. For example, an NB model was fitted in several microbiome studies related to Parkinson's disease and to analyze the effect of edible
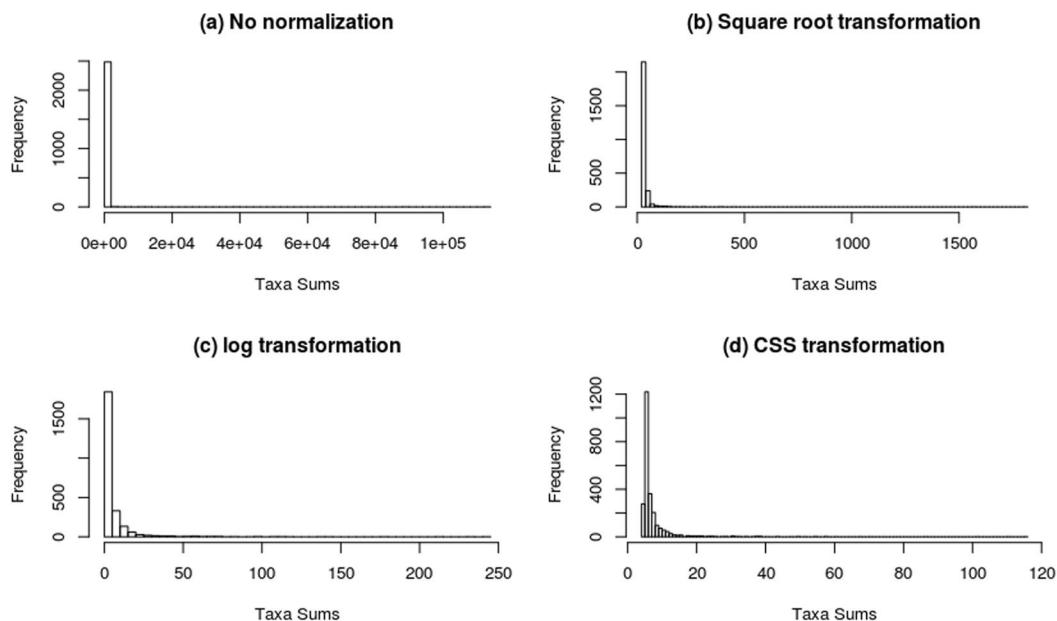
Figure 5. Histograms of the distribution of taxa sums across samples using (**a**) no normalization (**b**) Square root transformation (**c**) log transformation and (**d**) CSS transformations.
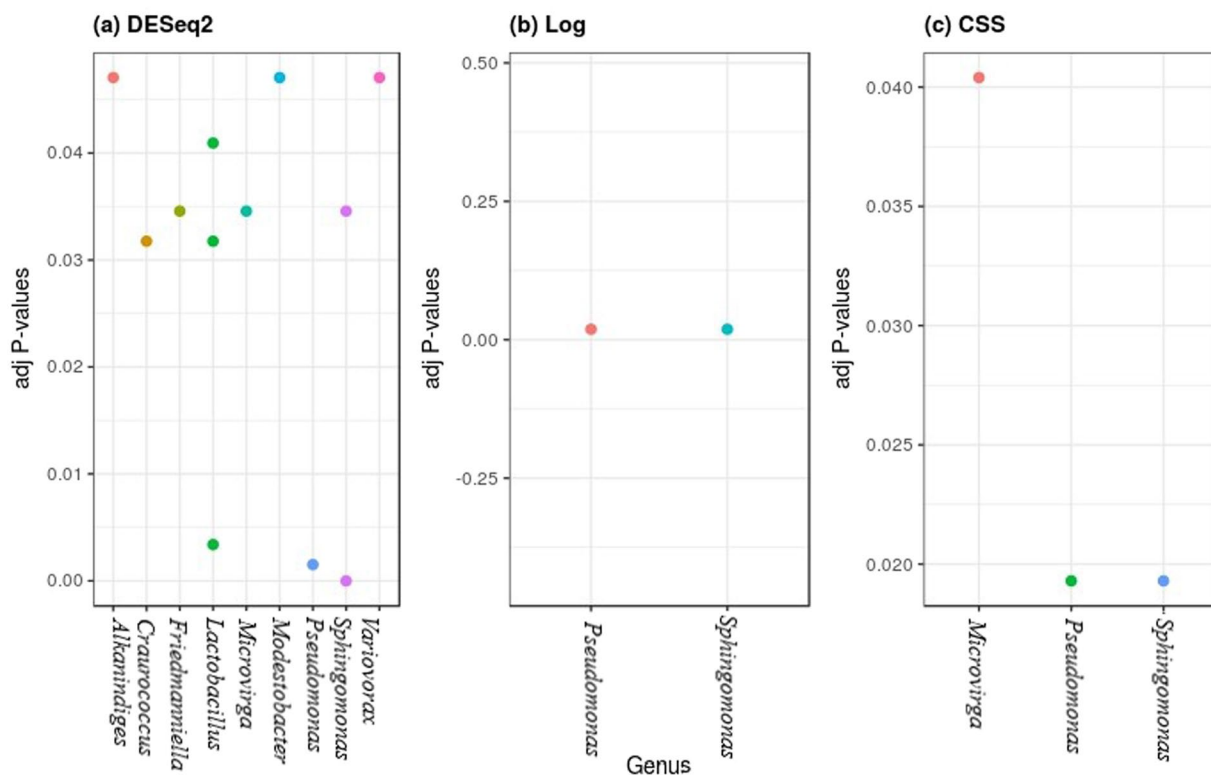
Figure 6. Separate normalization methods identified different bacterial taxa that significantly contributed (adj P < 0.05) to differences between two growing years.

cricket consumption on gut microbiome[45–47]. Similarly, an NB model was also used for identifying differential sequence tag abundance and for detecting differentially abundant features in clinical metagenomic samples[44,48]. Zero-inflated Gaussian (ZIG) mixture model was proposed[24] in 2013, which uses a new cumulative sum scaling (CSS) normalization technique to correct the bias in measuring the differential abundance introduced by total sum normalization. This model precisely evaluates the probability that an observed zero is generated due to the
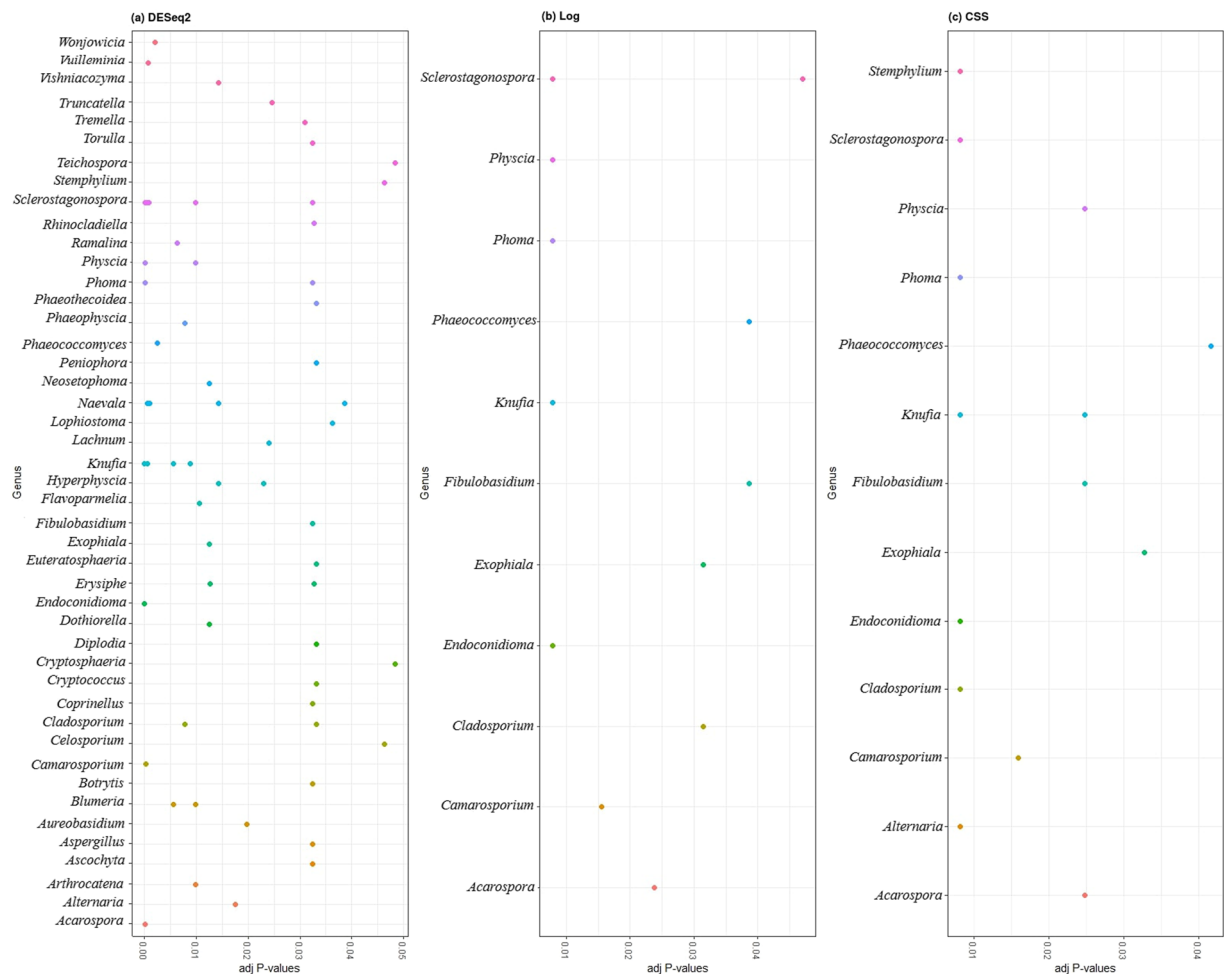
Figure 7. Separate normalization methods identified different fungal taxa that significantly contributed (adj P < 0.05) to differences between two growing years.

undersampling or from the count distribution (absence of the taxonomic feature in the samples). We evaluated one NB method (DESeq2), CSS method and log transformation of the data to generate normalized counts and performed multiple testing on differential taxa abundance. Our results suggested that the CSS method worked finer than other methods in obtaining the normalized counts and gave statistically more robust estimates of differential taxa abundance.

Conclusion
Our study revealed the existence of heterogeneous microbiome in the leaf phyllosphere of *Vitaceae* and we speculated that the multilayer species-climate interaction could be the reason for the microbiome assemblage in the phyllosphere. We have also inferred the advantages of CSS normalization over other methods, however, this method needs to be further evaluated with the sufficient amount of other microbiome studies.

Material and Methods
Sampling, isolation of phyllosphere microbes and DNA extraction. Leaf samples (from each of the 5 grape species) were collected from the repository of the agronomic school SupAgro at Montpellier, France (Mediterranean). 18–20 fully developed asymptomatic leaves were randomly sampled from 3–4 plants of each grape species in the Spring season (mid of May 2017 and 2018, before fungicide spraying) to make three replicates. The sample washing procedure was adopted by our previously published protocol[12] and was performed with isotonic sodium chloride solution (0.15 M) with 0.01% Tween 20 in 50 ml propylene tubes using a horizontal shaker at 100 RPM for 1hr[10]. To maximize the microbial recovery from the leaf surface, samples were given an ultrasonic bath for 7–10 minute using Ultrasonic Cleaner (Branson 5510). Afterward, the remaining solution was centrifuged at 4,000 g and microbial pellets were transferred in 2-ml Eppendorf tubes and stored at −20 °C. DNA was extracted from each sample using the ZymoBiomics DNA MicroPrep Kit (Zymo Research, USA).

PCR amplification and amplicon sequence library preparation. The 16S ribosomal gene (V4 region) was amplified using 515 F and 806 R primers to characterize bacterial communities. Modified ITS9 and ITS4 primers targeting the ITS2 region were used to access fungal community diversity and abundances. Sequencing

Method	Genus	Adj P-value	FDR
Spring 2017			
CSS	Cutibacterium	0.033	0.001
	Rubelimicrobium	0.0171	0.001
	Sphingomonas	0.034	0.018
Log	Cutibacterium	0.038	0.017
	Rubelimicrobium	0.168	0.001
DESeq2	Cutibacterium	0.038	0.128
	Rubelimicrobium	0.0192	0.212
	Sphingomonas	0.042	0.201
	Microvirga	0.022	0.113
	Rhodococcus	0.022	0.113
Spring 2018			
CSS	Hymenobacter	0.036	0.061
	Rubelimicrobium	0.036	0.061
Log	Hymenobacter	0.034	0.045
	Rubelimicrobium	0.034	0.045
	Methylobacterium	0.039	0.045
DESeq2	Hymenobacter	0.027	0.118
	Rubelimicrobium	0.027	0.118
	Paracoccus	0.027	0.118
	Methylobacterium	0.038	0.221
	Pedobacter	0.038	0.221
	Kineococcus	0.038	0.221
	Spirosoma	0.043	0.312
	Blastococcus	0.043	0.312

Table 3. Differential taxa abundance (bacterial taxa, genus level) among Grape Species at each year.

Method	Genus	Adj P-value	FDR
Spring 2017			
CSS	Aspergillus	0.0011	0.022
	Alternaria	0.0171	0.033
Log	Aspergillus	0.0012	0.017
	Alternaria	0.0012	0.017
	Botrytis	0.0131	0.112
DESeq2	Hyphodontia	0.006	0.081
	Aspergillus	0.026	0.119
	Alternaria	0.026	0.119
	Botrytis	0.026	0.119
Spring 2018			
CSS	Alternaria	0.024	0.038
	Sclerostagonospora	0.024	0.038
	Truncatella	0.032	0.048
Log	Alternaria	0.027	0.0199
	Sclerostagonospora	0.027	0.0199
	Truncatella	0.034	0.0231
DESeq2	Knulfa	0.021	0.132
	Alternaria	0.021	0.132
	Sclerostagonospora	0.021	0.132
	Truncatella	0.037	0.172
	Botrytis	0.037	0.172

Table 4. Differential fungal taxa abundance (fungal taxa, genus level) among Grape Species at each year.

libraries were prepared using two-step PCR reactions. Amplification of the target regions and the addition of Illumina Nextera transposase sequence to the amplicons was performed in step one PCR (or PCR1). Both PCR1 primers (forward and reverse) were amended with frameshift (FS) sequences in their 5′ overhang to improve

overall read quality and sequence diversity[49]. 25 μL reactions (with 30 ng of sample DNA) using the KAPA HiFi HotStart (KAPA Biosystems, USA) PCR mix (initial denaturing at 95 °C followed by 30 cycles of denaturing at 95 °C for 30 s, primer annealing at 57 °C for 60 s and primer extension at 68 °C for 60 s) were used to perform PCR1. Purification of amplicons was achieved using Agencourt AMPure XP beads (Beckman Coulter, USA) at a bead-to-DNA ratio of 0.7:1. Amplicons were then resuspended in 30 μL MilliQ water and evaluated in agarose gels. In PCR2 (Illumina dual indexing PCR), each cleaned PCR1 product within the same sample received a unique combination of forward and reverse primers (respectively, N7 and S5 Illumina dual index oligos). Samples were again cleaned afterward, using AmPure XP magnetic beads, pooled in equimolar concentrations and sequenced using 2 × 250 bp MiSeq v2 sequencing (Illumina Inc., San Diego, CA, USA). ZymoBIOMICS microbial community standard and DNA extraction buffer were used for this library preparation as a positive and negative control respectively following the manufacturer's instructions.

Data analysis. Paired-end sequence reads from 16S and ITS sequences were filtered, trimmed and processed with the *dada2* v1.8 (R Bioconductor package)[50]. The fastqPairedFilter function of the *dada2* was used to trim primers and discard bases with low-quality scores (q < 11). A core Divisive Amplicon Denoising Algorithm (DADA) was performed on these filtered files and amplicon sequence variants (or OTUs) were inferred[10,50]. Chimeras were removed afterward using the removeBimeraDenovo function of the same *dada2* package (Table 1 represents the total number of reads retained after performing each of these steps). Taxonomy was assigned to bacterial and fungal OTU sequences using the RDP classifier[51] and UNITE data base[52] respectively with k-mer size 8 and 100 bootstrap replicates.

Later, the *phyloseq* package[53] was used to get the α and β-diversity estimates. The estimate_richness function of the *phyloseq* gave *observed α-diversity* measures within sample categories. Relative abundances of microbial genera were plotted using the *ggplot2* package[54] after gloming the data at the genus level (using the tax_glom function of the *phyloseq*) and transforming genus abundance data into relative counts. PCoA ordination was performed on variance stabilized log-transformed data using the Bray-Curtis dissimilarity matrix and visualized by using their base functions in the *phyloseq* package.

Data transformations and statistical analysis. Square root transformation of the data was performed on taxa counts using sqrt(1 + x) function. Log transformation was done using log(1 + x) function on taxa counts. DESeq2 normalization was performed using the phyloseq_to_deseq2 function of the *DESeq2* package[25]. CSS normalization was done using the *metagenomeSeq* package[24]. Multiple testing was performed using the mt function of the phyloseq package after each data transformations to identify differentially abundant taxa between groups.

All computational analyses were performed in R-environment v3.3.4 (R Core Team, 2017) and the statistical significance was assessed at P < 0.05 throughout, and we adjusted P-values for multiple comparisons (whenever necessary) according to the Benjamini and Hochberg method to control False Discovery Rate (FDR)[55]. We performed ANOVA (analysis of variance)[56] among sample categories while measuring the observed estimates of α-diversity. To assess the statistical significance of PCoA clusters according to the sample categories, stratified permutational multivariate analysis of variance (PERMANOVA) with 9999 permutations was conducted (at α = 0.05) on all principal coordinates obtained during PCoA ordinations with the adonis command (with appropriate model matrix) of the *vegan* package[57].

Data Availability

Data is provided fully in the result section of this paper and the sequence data is uploaded at the institutional server of INRA and can be obtained upon reasonable request to authors.

References

1. Lindow, S.E., & Brandl, M.T. Microbiology of the phyllosphere. *Appl. Environ. Microbiol.* **69**(4), 1875–83, https://doi.org/10.1128/AEM.69.4.1875-1883 (2003).
2. Couée, I., Sulmon, C., Gouesbet, G. & El Amrani, A. Involvement of soluble sugars in reactive oxygen species and responses to oxidative stress in plants. *J. Exp. Bot.* **57**, 449–459, https://doi.org/10.1093/jxb/erj027 (2006).
3. Trouvelot, S. *et al.* Carbohydrates in plant immunity and plant protection: roles and potential application as foliar sprays. *Front. Plant Sci.* **5**, 592, https://doi.org/10.3389/fpls.2014.00592 (2014).
4. Bokulich, N. A., Thorngate, J. H., Richardson, P. M. & Mills, D. A. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *PNAS.* **111**, E139–148, https://doi.org/10.1073/pnas.1317377110 (2014).
5. Burns, K. N. *et al.* Vineyard soil bacterial diversity and composition revealed by 16S rRNA genes: Differentiation by geographic features. *Soil Bio.* **91**, 232–247, https://doi.org/10.1016/j.soilbio.2015.09.002 (2015).
6. Setati, M. E., Jacobson, D., Andong, U. C. & Bauer, F. F. The vineyard yeast microbiome, a mixed model microbial map. *PLoS One* **7**, e52609, https://doi.org/10.1371/journal.pone.0052609 (2012).
7. Corneo, P. E. *et al.* Microbial community structure in vineyard soils across altitudinal gradients and in different seasons. *FEMS Microbiol Ecol* **84**, 588–602, https://doi.org/10.1111/1574-6941.12087 (2013).
8. Pinto, C. *et al.* Unravelling the Diversity of Grapevine Microbiome. *PLoS ONE* **9**(1), e85622, https://doi.org/10.1371/journal.pone.0085622 (2014).
9. Zarraonaindia, I. *et al.* The soil microbiome influences grapevine-associated microbiota. *mBio.* **6**(2), e02527–14, https://doi.org/10.1128/mBio.02527-14 (2015).
10. Singh, P., Santoni, S., This, P. & Péros, J. P. Genotype-Environment interaction shapes the microbial assemblage in grapevine phyllosphere: An NGS Approach. *Microorganisms* **6**(4), 96, https://doi.org/10.3390/microorganisms6040096 (2018).
11. Perazzolli, M. *et al.* Resilience of the natural phyllosphere microbiota of the grapevine to chemical and biological pesticides. *Appl. Environ. Microbiol.* **80**, 3585–3596, https://doi.org/10.1128/AEM.00415-14 (2014).
12. Singh, P. *et al.* Assessing the impact of plant genetic diversity in shaping the microbial community structure of Vitis vinifera phyllosphere in the Mediterranean. *Front. Life Sci.* **11**, 35–46, https://doi.org/10.1080/21553769.2018.1552628 (2018).
13. Gopal, M. & Gupta, A. Microbiome Selection Could Spur Next-Generation Plant Breeding Strategies. *Front. Microbiol.* **7**, 1971, https://doi.org/10.3389/fmicb.2016.01971 (2016).

14. Cordero-Bueso, G. *et al*. Wild Grape-Associated Yeasts as Promising Biocontrol Agents against Vitis vinifera Fungal Pathogens. *Front. Microbiol.* **8**, 2025, https://doi.org/10.3389/fmicb.2017.02025 (2017).

15. Wen, J. *et al*. Transcriptome Sequences Resolve Deep Relationships of the Grape Family *PloS ONE* **8**(9), e74394, https://doi.org/10:1371/journal.pone.0074394 (2013).

16. Zhang, N., Wen, J. & Zimmer, E. A. Congruent Deep Relationships in the Grape Family (Vitaceae) Based on Sequences of Chloroplast Genomes and Mitochondrial Genes via Genome Skimming. *PLoS ONE* **10**(12), e0144701, https://doi.org/10.1371/journal.pone.0144701 (2015).

17. Wen, J. *et al*. A new phylogenetic tribal classification of the grape family (Vitaceae). *J. Syst. Evol.* **56**, 262–272, https://doi.org/10.1111/jse.12427 (2018).

18. Péros, J.-P. *et al*. Genetic variation and biogeography of the disjunct Vitis subg. Vitis (Vitaceae). *J. Biogeogr.* **38**, 471–486, https://doi.org/10.1111/j.1365-2699.2010.02410.x (2011).

19. Wan, Y. *et al*. A phylogenetic analysis of the grape genus (Vitis L.) reveals broad reticulation and concurrent diversification during neogene and quaternary climate change. *BMC Evol. Biol.* **13**, 141, https://doi.org/10.1186/1471-2148-13-141 (2013).

20. Shi, S. *et al*. Successional trajectories of rhizosphere bacterial communities over consecutive seasons. *mBio* **6**(4), e00746–15, https://doi.org/10.1128/mBio.00746-15 (2015).

21. McMurdie, P. J. & Holmes, S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol.* **10**(4), e1003531, https://doi.org/10.1371/journal.pcbi.1003531 (2014).

22. Aitchison, J. The statistical analysis of compositional data (with discussion). J R Stat Soc Ser B *Stat Methodol.* **44**(2), 139e177, https://www.jstor.org/stable/2345821 (1982).

23. Pawlowsky-Glahn, V., Egozcue, J. J. & Tolosana-Delgado, R. Modeling and Analysis of Compositional Data. Springer. London, UK. *John Wiley & Sons* (2015).

24. Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. Robust methods for differential abundance analysis in marker gene surveys. *Nat Methods.* **10**(12), 1200e1202, https://doi.org/10.1038/nmeth.2658 (2013).

25. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**(12), 550, https://doi.org/10.1186/s13059-014-0550-8 (2014).

26. Kembel, S. W., Connor, T. K. O., Arnold, H. K., Hubbell, S. P. & Wright, S. J. Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *PNAS* **111**, 13715–20, https://doi.org/10.1073/pnas.1216057111 (2014).

27. Delmotte, N. *et al*. Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *PNAS* **106**(38), 16428–33, https://doi.org/10.1073/pnas.0905240106 (2009).

28. Coince, A. *et al*. Leaf and root-associated fungal assemblages do not follow similar elevational diversity patterns. *PLOS ONE* **9**(6), e100668, https://doi.org/10.1371/journal.pone.0100668 (2014).

29. Abanda-Nkpwatt, D., Musch, M., Tschiersch, J., Boettner, M. & Schwab, W. Molecular interaction between Methylobacterium extorquens and seedlings: growth promotion, methanol consumption, and localization of the methanol emission site. *J. Exp. Bot.* **57**(15), 4025–32, https://doi.org/10.1093/jxb/erl173 (2006).

30. Innerebner, G., Knief, C. & Vorholt, J. A. Protection of Arabidopsis thaliana against leaf-pathogenic Pseudomonas syringae by Sphingomonas strains in a controlled model system. *Appl. Environ. Microbiol.* **77**, 3202–3210, https://doi.org/10.1128/AEM.00133-11 (2011).

31. Grube, M., Schmid, F. & Berg, G. Black fungi and associated bacterial communities in the phyllosphere of grapevine. *Fun. Biol.* **115**, 978–986, https://doi.org/10.1016/j.funbio.2011.04.004 (2011).

32. Garijo, P. *et al*. Presence of enological microorganisms in the grapes and the air of a vineyard during the ripening period. *Eur Food Res Technol.* **233**, 359–365, https://doi.org/10.1007/s00217-011-1528-3 (2011).

33. Araújo, W. L. *et al*. Diversity of endophytic bacterial populations and their interaction with Xylella fastidiosa in citrus plants. *Appl. Environ. Microbiol.* **68**(10), 4906–4914 (2002).

34. Lacava, P. T., Araújo, W. L., Marcon, J., Maccheroni, W. Jr. & Azevedo, J. L. D. (2004). Interaction between endophytic bacteria from citrus plants and the phytopathogenic bacteria *Xylella fastidiosa*, causal agent of citrus-variegated chlorosis. *Lett Appl Microbiol.* **39**(1), 55–9 (2004).

35. Ritpitakphong, U. *et al*. The microbiome of the leaf surface of Arabidopsis protects against a fungal pathogen. *New Phytol.* **210**, 1033–43, https://doi.org/10.1111/nph.13808 (2016).

36. Lopez-Velasco, G., Welbaum, G. E., Boyer, R. R., Mane, S. P. & Ponder, M. A. Changes in spinach phylloepiphytic bacteria communities following minimal processing and refrigerated storage described using pyrosequencing of 16S rRNA amplicons. *J. Appl. Microbiol.* **110**, 1203–1214, https://doi.org/10.1111/j.1365-2672.2011.04969 (2011).

37. Rastogi, G. *et al*. Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *ISME J.* **6**, 1812–1822, https://doi.org/10.1038/ismej.2012.32 (2012).

38. Dastager, S. G., Lee, J. C., Ju, Y. J., Park, D. J. & Kim, C. J. *Rubellimicrobium mesophilum sp.* nov., a mesophilic, pigmented bacterium isolated from soil. *Int. J. Syst. Evol. Microbiol.* **58**, 1797–1800, https://doi.org/10.1099/ijs.0.65590-0 (2008).

39. Balint, M. *et al*. Host genotype shapes the foliar fungal microbiome of balsam poplar (Populus balsamifera). *PLOS ONE* **8**(1), e53987, https://doi.org/10.1371/journal.pone.0053987 (2013).

40. Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M. & Vorholt, J. A. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genet.* **10**(4), e1004283, https://doi.org/10.1371/journal.pgen.1004283 (2014).

41. Alkorta, I., Epelde, L. & Garbisu, C. Environmental parameters altered by climate change affect the activity of soil microorganisms involved in bioremediation. *FEMS Microbiol. Lett.* **364**, 19, https://doi.org/10.1093/femsle/fnx200 (2017).

42. Fahimipour, A. K. *et al*. Daylight exposure modulates bacterial communities associated with household dust. *Microbiome.* **6**, 175, https://doi.org/10.1186/s40168-018-0559-4 (2018).

43. Alekseyenko, A. V. *et al*. Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome* **1**, 31, https://doi.org/10.1186/2049-2618-1-31 (2013).

44. White, J. R., Nagarajan, N. & Pop, M. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Comput Biol.* **5**(4), e1000352, https://doi.org/10.1371/journal.pcbi.1000352 (2009).

45. Falkenhorst, G. *et al*. Serological cross sectional studies on salmonella incidence in eight European countries: no correlation with incidence of reported cases. *BMC Public Health.* **12**, 523, https://doi.org/10.1186/1471-2458-12-523 (2012).

46. Yin, X. *et al*. Structural changes of gut microbiota in a rat non-alcoholic fatty liver disease model treated with a Chinese herbal formula. *Syst. Appl. Microbiol.* **36**(3), 188e196, https://doi.org/10.1016/j.syapm.2012.12.009 (2013).

47. Stull, V. J. *et al*. Impact of Edible Cricket Consumption on Gut Microbiota in Healthy Adults, a Double-blind, Randomized Crossover Trial. *Sci. Rep.* **8**, 10762, https://doi.org/10.1038/s41598-018-29032-2 (2018).

48. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* **26**(1), 139e140, https://doi.org/10.1093/bioinformatics/btp616 (2010).

49. Souza, R. S. C. *et al*. Unlocking the bacterial and fungal communities assemblages of sugarcane microbiome. *Sci. Rep.* **6**, 28774, https://doi.org/10.1038/srep28774 (2016).

50. Callahan, B. J. *et al*. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods.* **13**, 581–583, https://doi.org/10.1038/nmeth.3869 (2016).

51. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267, https://doi.org/10.1128/AEM.00062-07 (2007).
52. Abarenkov, K. *et al.* The UNITE database for molecular identification of fungi - recent updates and future perspectives. *New Phytol.* **186**, 281–285, https://doi.org/10.1111/j.14698137.2009.03160.x (2010).
53. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217, https://doi.org/10.1371/journal.pone.0061217 (2013).
54. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. New York, USA: *SpringerVerlag* (2009).
55. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. Royal Stat. Soc.* **57**: 289–300, https://www.jstor.org/stable/2346101 (1995).
56. Chambers, J. M., Freeny, A. E. & Heiberger, R. M. Analysis of variance; designed experiments. Chapter 5 of *Statistical Models in S* (1992).
57. Oksanen, J. F. *et al.* vegan: community Ecology Package. R package version 2.4-3, https://CRAN.R-project.org/package=vegan (2017).

Acknowledgements

Author Contributions

Prashant Singh, Jean-Pierre Péros and Patrice This developed the project. Prashant Singh, Sylvain Santoni & Audrey Weber contributed to the experimental design and execution. Prashant Singh performed data analysis and interpretation. Prashant Singh, Jean-Pierre Péros and Patrice This contributed to the manuscript writing.

Additional Information

Competing Interests: The authors declare no competing interests.

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