Wine terroir and the soil microbiome: an amplicon sequencing–based assessment of the Barossa Valley and its sub-regions

Jia Zhou[1], Timothy R. Cavagnaro[1], Roberta De Bei[1], Tiffanie M. Nelson[2], John R. Stephen[1], Andrew Metcalfe[3], Matthew Gilliham[1,4], James Breen[5,6,7], Cassandra Collins[1,4], Carlos M. Rodríguez López[1,7*]

[1]School of Agriculture Food and Wine and Waite Research Institute, The University of Adelaide, PMB 1, Glen Osmond, 5064, South Australia, Australia

[2]Queensland Facility for Advanced Bioinformatics, School of Medicine, Menzies Health Institute Queensland, Griffith University, Southport, QLD, Australia

[3]School of Mathematical Sciences, University of Adelaide, Adelaide, SA, Australia

[4]ARC Industrial Transformation Training Centre for Innovative Wine Production, Waite Research Institute, PMB 1, Glen Osmond, 5064, South Australia, Australia.

[5]Bioinformatics Hub, School of Biological Sciences, University of Adelaide, Adelaide, SA, Australia

[6]Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia

[7]South Australian Health & Medical Research Institute, Adelaide, SA, Australia

[8]Environmental Epigenetics and Genetics Group; Department of Horticulture, College of Agriculture, Food and Environment, University of Kentucky, Lexington, KY, USA

*** Correspondence:**
Dr Carlos Rodríguez López
E-mail:carlos.rodriguezlopez@uky.edu

Keywords: Terroir, vineyard soil microbiome, grapevine, Barossa Valley, Illumina, 16SrRNA.

Word count: 5,391 words

Abstract

Soil is an important factor that contributes to the uniqueness of a wine produced by vines grown in specific conditions. Recent data shows that the composition, diversity and function of soil microbial communities may play important roles in determining wine quality and indirectly affect its economic value. Here, we evaluated the impact of environmental variables on the soil microbiomes of 22 Barossa Valley vineyard sites based on the 16S rRNA gene hypervariable region 4. In this study, we report that environmental heterogeneity (soil plant-available P content, elevation, rainfall, temperature, spacing between row and spacing between vine) caused more microbial dissimilarity than geographic distance. Vineyards located in cooler and wetter regions showed lower beta diversity and a higher ratio of dominant taxa. Differences in microbial community composition were significantly associated with differences in fruit traits and in wine chemical and metabolomic profiles, highlighting the potential influence of microbial communities on the phenotype of grapevines. Our results suggest that environmental factors affect wine terroir, and this may be mediated by changes in microbial structure, thus providing a basic understanding of how growing conditions affect interactions between plants and their soil microbiomes.

1 Introduction

Wine price differs considerably depending on its quality (e.g., flavor, color and typicity), which is largely determined by the interactions between the grape and the growing conditions including climate, soil, topography, agricultural management, and the wine making process (Bokulich et al., 2016). These interactions influence the expression of wine's terroir (Bokulich et al., 2016; Jullian Fabres et al., 2017). Research on the drivers of terroir have predominantly focused on abiotic environmental factors, such as climate, soil, viticultural management and wine making process, studied individually (Mira de Orduña, 2010; Romero et al., 2016; Vega-Avila et al., 2015) and simultaneously (Van Leeuwen et al., 2004). However, little research has been done, in the context of terroir, on whether soil microbiomes exhibit distinct patterns of distribution at small geographic scales (e.g. neighboring vineyards), and whether vineyard microbiomes are associated with a wine's terroir (Burns et al., 2015; Bokulich et al., 2016).

Soil microbiomes, especially bacterial species, have been found to be qualitatively and quantitatively different between vineyard systems (Vega-Avila et al., 2015). Environmental factors, such as topography, climate, soil properties, cultivars and agricultural management, combine to affect soil microbial communities (Reeve et al., 2010; Castro et al., 2010; Lamb et al., 2011). It has been shown that climate and topography, including rainfall pattern and temperature, affect these communities through their impacts on soil (Burns et al., 2015). Soil properties such as soil texture, nitrogen (N) content, phosphorus (P) content, carbon to nitrogen (C:N) ratio, water content, and pH show significant effects on the diversity and composition of microbial communities (Girvan et al., 2003; Frey et al., 2004; Rousk et al., 2010; Fierer and Jackson, 2006). Plant genotypes exert an influence on the structural and functional diversity of soil microbiomes by varying root exudates and rhizodeposition (Broeckling et al., 2008; Dias et al., 2013; Philippot et al., 2013). Management practices, land use and varying degrees of stress and disturbance influence the soil microbiome markedly due to specific management objectives (Crowder et al., 2010; Lumini et al., 2011; Reeve et al., 2010; Sugiyama et al., 2010).

Soil microbiomes interact with the vines, and thus affect wine quality (Bokulich et al., 2016; Burns et al., 2015). The interaction between soil microorganisms and plants includes the facilitation of nutrient uptake/utilization, stabilization of soil structure, reduction of disease prevalence by out-competing soil-borne pathogens or increase of disease prevalence by microbial pathogen invasion (Edwards et al., 2014; Zarraonaindia et al., 2015). Soil microbiomes also contribute to the wine fermentation flora, ultimately affecting wine quality (Barata et al., 2012; Compant et al., 2011; Martins et al., 2013). However, microbial assemblage function is intrinsically difficult to measure and define because of its highly changeable nature (Nannipieri et al., 2003). Additionally, due to the complex interactions between soil microbes, the influence of certain microbial communities can be substituted by other microorganisms with the same ecological function (Nannipieri et al., 2003; Crowder et al., 2010; Lamb et al., 2011; Wittebolle et al., 2009).

The primary aim of this project was to assess if there is a relationship between soil microbiomes and terroir. To achieve this, we asked the following questions:

(i) Do wine sub-regions have distinct soil bacterial communities?;
(ii) What environmental conditions and agricultural practices shape soil bacterial community of vineyards?; and
(iii) Do differences in the soil bacterial community correlate with berry and wine characteristics?

In order to answer these questions, we undertook a soil microbiome survey in an iconic wine region, the Barossa in South Australia. The Barossa has a winemaking history of over 160 years and because of its importance as a growing region, has been chosen as a model to investigate terroir previously (Wolf et al., 2003; Edwards et al., 2014; Xie et al., 2017). Besides, the environmental characteristics of the Barossa, including climate, soil and topography have been previously characterized in detail (Robinson and Sandercock, 2014). However, to date, no study has analyzed the soil microbiomes of the Barossa wine region or the possible influence on wine properties. Thus, determining how soil microbiome diversity and composition are influenced by environmental factors, and how microbiome differences correlate with differences in fruit/wine composition, will provide a starting point from which to better understand the (potential) functional role of soil microbial communities in terroir.

2 Materials and methods

2.1 Experimental design and plant material

Twenty-two Barossa vineyards (Figure S1), planted with own-rooted Shiraz (*Vitis vinifera* L.) and representative of the climate, soil and management practices of six Barossa sub-regions (i.e. Eden Valley, Northern Grounds, Central Grounds, Eastern Edge, Western Ridge, Southern Grounds) were selected for this study. Three to four vineyards per sub-region were included and nine vines from three rows from each vineyard were selected for measurement and sampling. Vines within the same row were adjacent to each other. Vines adjacent to missing vines, end of row vines and border rows were excluded from the selection.

2.2 Soil sampling protocol

Three soil cores (0-10cm soil layer) were collected using a (20 mm diameter) soil auger from around each individual plant (approximately 10cm from the trunk) and combined, giving a total nine soil samples per row. A total of 594 soil samples were collected (27 soil samples from each vineyard) on the 2nd of November (Austral Spring) 2015. Soil samples were immediately stored at 4˚C and returned to the laboratory on the same day of collection. Soil samples from the same row were thoroughly mixed to obtain three samples per vineyard, and a total of 66 samples across the study. Coarse debris was removed from each soil sample using a 2mm sieve, and each sample was then divided into three sub-samples (approximately 850 cm^3 each). The first subsample (approx. 20 g) was used for determination of soil gravimetric moisture content. The second was air-dried until a constant mass was achieved and used for analysis of soil texture, pH, electrical conductivity, and plant-available (Colwell) P (phosphorus), as described previously (Cavagnaro, 2015)). The third soil subsample was stored at -80˚C for DNA extraction and downstream genomic analysis (see below).

2.3 Vineyard physical characterization

In this study, the climate was characterized on the basis of rainfall and temperature. The influence of topography was studied through elevation above sea level and vineyard orientation. Soil texture was determined following (Giddings, 2015). Soil pH and electrical conductivity were determined on a 1:5 soil/water mixture and then measured using pH/salinity meter (WP-81 Conductivity-Salinity-pH-mV Meter, v6.0, TPS Pty Ltd). Plant-available phosphorus was extracted and measured using Colwell P method (Rayment and Higginson, 1992) (Table S1). The remaining soil, topographic and climatic data was obtained from the Barossa Grounds project (Robinson and Sandercock, 2014), while vineyard management information was collected from participating growers (Table S2).

2.4 Fruit and wine chemical analysis

Fruit juice pH and total acidity (TA) was measured using an autotitrator (Crison instruments Barcelona, Spain) (Iland et al. 2004). Total soluble solids (TSS) of juice samples were tested with a digital refractometer (BRX-242 Erma inc. Tokyo, Japan). A sample of 50 berries from random bunches on selected vines were collected and frozen at -20°C for anthocyanin, phenolic and tannin analyses. Total grape tannins were measured by the methyl cellulose precipitable (MCP) tannin assay (Sarneckis et al. 2006) using the protocol of Mercurio et al. (2007). Total anthocyanin and phenolics were determined according the method of Iland et al. (2004) (Table S3).

One bottle of commercial wine per vineyard was used for the chemical analysis. Wine pH and TA was determined as described by Iland et al. (2004). Final alcohol levels were determined using an Alcolyzer Wine ME (Anton Paar, Graz, Austria). Wine colour was determined using the modified Somers assay using a high throughput method in 96 well plates [98]. Wine tannin concentration was determined using the methyl cellulose precipitable (MCP) tannin assay of Mercurio et al. (2007) and is expressed as epicatechin equivalents (mg/L) using an 8-point epicatechin standard curve Sarneckis et al. (2006). The modified Somers assay was used to determine; wine colour density (WCD), SO2-corrected WCD, degree of anthocyanin ionisation, phenolic substances and anthocyanins (in mg/L) (Table S4).

Non-targeted metabolomic analysis of the wine samples was performed using LC-MS/MS. The metabolites were isolated from bottled wine samples using solid-phase extraction (SPE) with Phenomenex Strata-X 33 um 85Å polymeric reverse-phase 60mg/3mL cartridges. A 2 mL aliquot of each sample was evaporated to dryness under nitrogen at 30°C. SPE conditions are presented in Table S5. A pooled mix of all samples was prepared and used to monitor instrument performance. The analysis was performed on an Agilent 1200SL HPLC coupled to a Bruker microTOF-Q II in ESI negative mode. The operating conditions are described in Table S5 (Table S6-7).

Following data acquisition, mass calibration was performed on each file using Bruker Daltonic's DataAnalysisViewer4.1 "Enhanced Quadratic" calibration method (Bruker Singapore, The Helios, Singapore). Each file was exported from DataAnalysis in the mzXML generic file format for further processing. The files were processed using R (statistical programming environment) v3.1.0 and Bioconductor v2.14 under a Debian Linux 64-bot environment. Molecular features were extracted for each file using xcmx package and features that possessed a common mass and retention time across samples were grouped together.

2.5 16S rRNA Gene Next Generation Sequencing library preparation

DNA extractions from soil 66 samples were carried out at the Australian Genome Research Facility (AGRF) (Adelaide node) using MoBio Powersoil kit (MoBio Laboratories, Inc) following the manufacture's protocol. DNA concentrations were estimated using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and normalized to 5 ng/µl using nanopure water.

Primers 515F and 808R (Bates et al., 2011; Caporaso et al., 2011) specific for the Bacterial 16S rRNA gene hypervariable "V4" region were used for PCR amplification of extracted DNA and to prepare amplicon libraries. 515F worked as a universal forward primer for all the samples and 808R included 12-base sample specific barcodes to allow downstream de-multiplexing (Table S8).

Three replicated PCR reactions were performed for each of the 66 samples. Each of these runs included one negative control as 'sample67' with no template DNA added. PCR reactions included 10ng of extracted DNA, 12.5 µl Q5 high-fidelity 2*master mix (New England Biolabs), 8.5 µl dH_2O, 1µl forward and reverse primers (10 µM) in 25µl reaction system. The PCR thermocycler (Bio-Rad T100)

program was 95°C for 6min, followed by 38 cycles of 95°C for 30sec, 50°C for 30sec and 72°C for 1m30sec.

Success of PCR reactions was verified by agarose gel (1.5% w/v) electrophoresis. Samples exhibiting weak bands were reamplified by adjusting the amount of DNA template. The triplicate reactions were then pooled into 67 pools. Individual pools were quantified by Qubit fluorometric double stranded DNA assay (Invitrogen, Carlsbad, CA, USA) and then mixed on an equimolar base to generate six pools each with 11 samples (each containing 5 ul of the water control pool). Pools were size-selected to remove unused primers using Agencourt® AMPure® XP (Beckman Coulter, Brea, CA, USA) following the manufacturer's protocol and mixed to equimolar concentrations to make one final pool. Library concentration and fragment size were estimated using TapeStation (Agilent, Santa Clara, CA, USA) and sequenced on the Illumina MiSeq platform (300 bp PE) (Illumina, San Diego, CA, USA) at the Australian Genome Research Facility-Adelaide node (AGRF).

2.6 Bioinformatics analysis

Raw Illumina sequencing data was quality filtered and demultiplexed at AGRF-Adelaide node. Forward and reverse sequences passing QC filter were merged using *bbmerge* (Bushnell, 2016).

Merged reads were analyzed using Quantitative Insight Into Microbial Ecology (QIIME) (QIIME version 1.8.0) (Navas-Molina et al., 2013). Operational taxonomic units (OTUs) were clustered using open-reference picking with the default *uclust* method (Edgar, 2010) based on 97% sequence similarity to the 16S rRNA Greengenes database (McDonald et al., 2012; DeSantis et al., 2006). OTUs were aligned to the Greengenes core reference database using *PyNAST* (Caporaso et al., 2010). Ribosomal Database Project (RDP) classifier was used to assign taxonomy (Wang et al., 2007). Both closed-reference OTU picking and open-reference OTU picking were performed for later analyses.

A BIOM file was generated after OTU picking, then OTUs identified in the negative control samples were removed from soil sample OTUs, leaving between 37,176 and 114,777 OTUs per sample (mean = 60,147 OTUs). Data with and without rarefaction were used for alpha diversity and beta diversity analyses. 37,176 OTUs (the lowest amount of OTUs in one sample) were randomly selected from each sample for rarefaction.

Alpha diversity (within-sample species richness and evenness) was measured using non-phylogenetic (including the observed number of OTUs and the Chao 1 estimator of the total number that would be observed with infinite sampling) and phylogenetic (Faith's Phylogenetic Diversity) indices (Faith, 1992). Phylogenetic beta diversity (between-sample diversity) was calculated using both weighted and unweighted UniFrac (Lozupone and Knight, 2005) and three-dimensional principal coordinates analysis (PCoA) plots were built through Emperor (Vázquez-Baeza et al., 2013). We then constructed a neighbor joining ultrametric tree in QIIME from the beta diversity UniFrac distance matrix. The generated tree file, as well as the Barossa Valley geographical map, vineyard locations and taxa summary files, were input into GenGIS (Parks et al., 2009, 2013) to visualize the relationship between soil bacterial beta diversity and vineyard location. The statistical significance of this relationship was determined using the Mantel test based on 9,999 random permutations and implemented on GenAlex v6.5 (Peakall and Smouse, 2012).

To identify the association of environmental variables and grape and wine properties (Table S1-S7) with soil bacterial microbiome, bacterial community dissimilarities were visualized with non-metric multidimensional scaling (nMDS) plots. Variables were fitted to the ordination plots using the function *envfit* in the package *Vegan version 2.5-2* (Oksanen et al., 2013) implemented in R version 3.5.0 (Team,

2013). Spearman's rank correlation coefficients were measured between individual taxon abundance and fruit and wine traits using the function *rcorr* in the package *Hmisc*. Grape traits included those from sensory, basic chemistry analyses, while wine traits included basic chemistry, wine fermentation products and amino acids concentration. Those traits and taxa with a significant (p-value <0.05) correlation coefficient larger than 0.80 or lower than -0.80 were deemed as significantly associated.

To identify which variables are important in explaining the composition of the soil microbial community, we performed distance-based redundancy analysis (dbRDA), a form of multivariate multiple regression that we performed directly on a Bray-Curtis dissimilarity matrix of OTUs using the ADONIS function in *Vegan*. We used automatic model building using the function *step* in R. The step function uses Akaike's Information Criterion (AIC) in model choice, which is based on the goodness of fit. The model building proceeds by steps until the 'best' fit is identified. If two predictor variables were highly correlated (>0.85) one, typically that which was more difficult to measure, was removed as well as variables with missing replicates (Variables included in the automatic model building are marked with * in Tables S2-S8). Differential statistic functions within the *edgeR* package (Chen et al., 2008) was used, as in Weis et al. (2017) to determine the significantly different taxa between vineyards separated by the main environmental drivers of beta diversity (i.e. soil type and soil phosphorous content). In order to avoid the influence of taxa showing low counts, a minimum threshold was set up at 100 counts per million.

3 Results

3.1 Analysis of soil properties

Of the three soil physicochemical properties tested, plant-available phosphorous (P) and electrical conductivity (a measure of soil salinity), differed significantly (Kruskal-Wallis: p-value < 0.05) between sub-regions of the Barossa (Table S1). Plant-available P was lowest in the Northern Grounds (11.5±2.7µg P/g soil) and highest in the Eastern Edge (39.0 ±14.2µg P/g soil). Electrical conductivity ranged from 111.0 uS/cm (Northern Grounds, SE=34.2) to 302.5 uS/cm (Central Grounds, SE=123.5). Soil pH did not differ between sub-regions, ranging from 6.2 (Eden Valley, SE=0.4) to 6.8 (Southern Grounds, SE=0.5).

3.2 Barossa Valley soil bacteria community composition

After quality filtering of the raw sequencing results, an average of 130,949 paired sequences remained per sample. Of these an average of 86,835 paired-end sequences per sample (66.3%) could be joined using *bbmerge* (Table S9).

Both bacterial and archaeal DNA was detected in all soil samples. A total of 98.9% of sequences were classifiable at the phylum level (Figure 1A) and 95.2% at the genus level. Of those classifiable at the phylum level, 96.5% were assigned to one of nine dominant groups (relative abundance ≥1.0%) in the samples namely: Actinobacteria (26.9%), Proteobacteria (26.7%), Acidobacteria (12.0%), Planctomycetes (6.2%), Chloroflexi (5.6%), Firmicutes (5.3%), Gemmatimonadetes (3.9%), Bacteroidetes (3.5%), Verrucomicrobia (2.5%) (Figure 1A). The only dominant Archaea group was Crenarchaeota (4.0%). The overall dominant Bacteria and Archaea groups were consistently present in the six regions, but at different ratios (Figure 1A). The phylogenetic inference of microbiome composition differences between sub-regions showed three clusters with Central and Northern Grounds, and Eden Valley and Western Ridge sharing the more similar microbial profiles (Figure 1A).

The number of observed OTUs (Figure 1B) showed significant differences (t-test: p-value < 0.05) between the OTU rich sub-regions (Northern and Central Grounds) and the relatively OTU poor sub-regions (Eden Valley and Western Ridge) (Table S10). Similarly, the Chao1 metric showed that Northern and Central Grounds presented higher levels of OTU richness while Eden Valley and Western Ridge had the lowest (Figure 1C). Pairwise comparison of alpha diversity between sub-regions showed significant differences (t-test, p-value < 0.05) between Northern Grounds and Eden Valley and Western Ridge and between Central Grounds and Eden Valley and Western Ridge (Table S11).

Dissimilarities in microbial communities between samples (i.e. beta diversity) were calculated as weighted and un-weighted UniFrac distances and both methods showed similar patterns, and so only analyses based on weighted results are shown here. For the most part, the three replicates from within a given vineyard were closely grouped on the ordination plot (Figure 2A), indicating that bacterial communities were consistent within sites. Pairwise analysis of the differences between groups (vineyards and sub-regions) showed that all vineyards and sub-regions are significantly different to each other (Adonis, p-value < 0.001). Mantel test analysis of the association between microbiome compositional differences and geographic distance, showed a small but significant correlation (rxy = 0.315; p-value = 0.0001) (Figure 2B).

To further explore dissimilarities among and within regions, neighbor joining analysis was used to cluster samples and to generate a similarity tree in QIIME. This information, along with a geographical map of the regions and their locations, were combined using the GenGIS software package (Parks et al., 2009). This approach showed a low level of clustering of vineyards according to their geographic location (Figure 2C).

3.3 Drivers of soil microbiome differentiation

Model selection was used to identify the combination of variables that explained the greatest variation in the soil microbiome. This approach consistently selected soil plant-available phosphorus (P) and soil texture as the main drivers (Model: p-value = 0.001) of soil microbiomes in the Barossa vineyards tested (Figure 3). Together, both variables explained 19.7% of the observed variability. Independent pairwise analysis of UniFrac distances of vineyards grouped by these soil characteristics, showed that microbial communities in clay soil types were significantly dissimilar from those in sandy soils (PERMANOVA: p-value < 0.001, Figure 4A). Microbial communities in soils with high plant-available phosphorus (P > 30mg/kg) were also dissimilar from those with low plant available phosphorous (PERMANOVA: p-value < 0.001, Figure 4C). Three and eight taxa were significantly more abundant in clay and sandy soils respectively (Figure 4B), while eight taxa were found significantly associated with low plant available phosphorous content, and three associated high levels of plant available phosphorous in soil (Figure 4D).

Envfit analysis identified a number of other environmental factors as individually associated with microbial community composition (Figure 5). Aside from plant available phosphorous (r^2 = 0.3706, p-value < 0.001), these variables were: elevation (r^2 = 0.3609, p-value < 0.001), growing season rainfall (r^2 = 0.2499, p-value < 0.001), mean annual rainfall (r^2 = 0.1621, p-value = 0.004), spacing between rows (r^2 = 0.1512, p-value = 0.006) and between vines (r^2 = 0.1561, p-value = 0.011), and growing season mean temperature (r^2 = 0.1113, p-value = 0.022).

Analysis of the correlation between individual environmental and vineyard management variables and taxa abundance, identified 4 positive (Spearman's >0.80, p-value <0.001) and 3 negative (Spearman's <-0.80, p-value <0.001) significant correlations (Figure S2). Positive correlations with individual taxa included, pH (order iii1-15 and family Pirellulaceae), elevation (family Isosphaeraceae), and plant age

(family Hyphomicrobiaceae); while negative correlations included P (family OPB35), elevation (family Conexibacteraceae), and the spacing between vines on the same row (family Haliangiaceae).

3.4 Terroir and vineyard soil microbiomes

Twenty four of the 75 grape and wine characteristics included in the study displayed a significant correlation with the soil microbial community composition (Table 2). The strongest associations identified for each of the four groups of traits tested were: 50 berry weight and average color per berry (basic berry properties); total anthocyanins and total phenolics (basic wine chemistry); Glycine and Alanine (wine amino acids); and 2-phenyl ethyl ethanol and acetic acid (wine fermentation products).

Significant positive correlation (Spearman's >0.80, p-value <0.05) were identified between the abundance of one taxon (order IS_44) and the average level of total phenolics mg/g berry weight (Figure S3A). Similarly, six wine traits showed positive correlations with the abundance of six microbial taxa (Figure S3B-E). Briefly, the genus *Rhodoplanes* was positively associated with the level of wine total phenolics and the family Chitinophagaceae was associated with color density of SO_2 corrected wine and with the level SO_2 resistant pigments in wine, while the family Kouleothrixaceae was positively correlated with wine color density.

4 Discussion

Previous studies have shown that environmental factors (e.g. climate and soil properties) and crop management may affect microbial populations in vineyards (Fierer, 2008; Burns et al., 2015; Weckert, 2016). To better understand how these variables contribute to vineyard microbial communities and how microbial diversity and composition correlate with fruit and wine quality traits, we studied the soil microbiome composition of 22 commercial vineyards representative of the Barossa Valley wine region in Australia.

4.1 Vineyard soil microbiome composition and diversity

With over 37,176 sequences per sample we reached a sequencing depth higher than those achieved in previous studies deemed sufficient to resolve differences between similar samples (e.g. Liu et al., 2014; Burns et al. 2015). From a species composition point of view, our results indicate that vineyard soil microbiomes present similarities across the six sub-regions studied. All soils analyzed presented both bacteria and archaea. A total of 96.5% of the all identified sequences were allocated in one of nine main dominant phyla (relative abundance >= 1.0%). Of these, eight (Actinobacteria, Proteobacteria, Acidobacteria, Planctomycetes, Chloroflexi, Firmicutes, Gemmatimonadetes, Bacteroidetes, and Verrucomicrobia) were Eubacteria, while only one dominant taxon was from Archeabacteria (Crenarchaeota). Although dominant phyla were consistently found in the six regions tested, they were present in different ratios. This finding is similar to earlier work; for example, investigating the Napa Valley American Viticultural Area (AVA), Burns et al. (2015) found the same nine top dominant bacteria groups, also with different ratios and order for each group. Similarly, Liu's et al. (2014) analysis of agricultural black soils in northeast China found almost the same dominant bacterial groups. Equally, analysis of non-agricultural soils by Faoro et al, (2010) and Lauber et al. (2009) identified the same dominant groups, with the exception of Verrucomicrobia which was replaced by Nitrospira (Faoro et al., 2010) and TM7 and Cyanobacteria replacing Planctomycetes and Chloroflexi (Lauber et al., 2009).

4.2 Location, soil properties, climate and vineyard management contribute are associated with soil microbial community dissimilarity in the Barossa

Although dominant taxa were constant at a regional level, soil microbiome diversity and composition seemed to be a better factor separating soil microbiomes from different sub-regions. The phylogenetic inference of microbiome composition differences between sub-regions showed that OTU richer sub-regions (Northern and Central Grounds) clustered independently from the OTU poorer ones (Eden Valley and Western Ridge).

Previous studies have shown that the major factors determining compositional dissimilarities of soil microbiome between sites are dispersal constraints (which predicts that more distant soils should have greater phylogenetic dissimilarity) and environmental heterogeneity (Fierer, 2008; Liu et al., 2014; Burns et al., 2015). Analysis of the influence of geographical distance on soil microbiome composition differences between Barossa Valley Region vineyards showed a small significant correlation between both parameters. It could be argued that such small contributions to vineyard soil microbiome composition differences could be associated with the relatively small distances between the vineyards in this study (Average distance 11.7 km, minimum distance 0.7 km and maximum distance 26.5 km). However, this correlation was similar to that observed by Burns et al. (2015) when studying 19 vineyards of the Napa Valley AVA that were separated by up to 53 km. This suggests that dispersal constraints contribute to soil microbiome differences at a much smaller scale than previously perceived.

Environmental heterogeneity has been found to be more important than geographic distance in shaping bacterial community at different geographical scales (Fierer and Jackson, 2006; Miura et al., 2017; da C Jesus et al., 2009; Ranjard et al., 2013; Hermans et al., 2017). The main contributors to environmental associated variability in soil communities are differences in climatic conditions, topography, soil properties, and cultivation practices (Mezzasalma et al., 2018; Burns et al., 2015). Microbiome composition similarity analysis results did not show a clear clustering of vineyards according to their geographic location, indicating that even at a close geographic distance, environmental heterogeneity is the dominant factor shaping soil microbiome composition.

To determine which environmental factors contribute to the observed differences in soil microbial communities we used an automatic model building approach. This analysis revealed that when taken in combination, plant-available phosphorous and soil texture were the major contributors to soil microbiome differences between vineyards (approximately 20% of the total observed variability). Differences in plant-available phosphorous have been previously shown to impact microbial communities (Awasthi et al., 2011; Fierer and Jackson, 2006). In the studied vineyards, clay soils tended to show higher plant-available phosphorous content (Figure 4A & C), which is consistent with previous findings (Krogstad et al., 2005). Interestingly, soil particle size has been negatively correlated with microbiome community alpha diversity (Sessitsch et al., 2001) indicating that both variables could be affecting microbiome composition in an, at least partially, independent manner. Moreover, while genera *Streptomyces*, *Rubrobacter* (both Actinobacteria) and *unclassified MND1*, were especially prevalent in clay soils, genera *Streptomyces*, *Pseudomonas* and *unclassified Sinobacteraceae* were found in soils with plant-available phosphorous content higher than 30µg / g soil. *Pseudomonas*, are inorganic P solubilizing bacteria (Awasthi et al., 2011; Goswami et al., 2013; Schmalenberger and Fox, 2016). Conversely, P levels negatively correlated with the abundance of the organic P mineralizing taxon OPB35. Pairwise analysis of individual taxa and environmental variables also identified previously reported strong and positive correlations between soil pH and order iii1-15 (acidobacteria-6) and family Pirellulaceae (Rousk et al., 2010; Hermans et al., 2017; Wu et al., 2017). Interestingly, both P and pH, have been shown to be soil variables that are indicative of anthropogenic activity (Hermans et al., 2017), which highlights the potential use of such taxa as reliable indicators of soil condition.

Previous studies have shown that climatic variables such as rainfall (Wildman, 2015) and temperature (Cong et al., 2015) are major shapers of soil microbial population composition and activity. Our results indicate that cooler and wetter regions (Western Ridge and Eden Valley) had relatively lower soil microbial diversity, and a higher ratio of dominant species, than the warmer and drier sites. Additionally, elevation, which negatively affects air temperature, showed a positive correlation with the families Isosphaeraceae and an unsurprising negative correlation with the thermophilic taxon Conexibacteraceae (Wagner and Wiegel, 2008).

Agricultural lands tend to show similar patterns of dominant bacteria (Burns et al. 2015; Liu et al. 2014; Faoro et al. 2010; Lauber et al. 2009), indicating that microbial community composition can be profoundly affected by cropping practices (Hartman et al., 2018). Our results show that, both spacing between row and vine, which determine the vineyard's planting density, are significantly associated with global differences in soil microbial community. Work in oil palm plantations has shown that planting density affects soil bacteria by altering the level of solar light incidence on soils, which can have dramatic effects on soil temperature and moisture (Tripathi et al., 2016). Pairwise comparisons between agronomical practices and individual taxa showed a negative correlation between spacing among vines on the same row and the abundance of representatives of the Haliangiaceae family. These are mesophilic organisms previously identified to be sensitive to agricultural practices (e.g. (Ding et al., 2014; Kim and Liesack, 2015; Wang et al., 2016)), which abundance could be favored by lower soil temperatures in densely planted vineyards. This highlights the importance of temperature, shown above, in the formation of soil bacterial communities. However, vine density and the use of under-vine cover crops could also cause different levels of interactions between plant roots and soil microbes. This is particularly prominent when comparing sites with similar topography and soil texture, in which spatial patterns of soil biota are assumed to be structured primarily by plant growth, age, growth form and density (Ettema, 2002). Our results indicate that the abundance of taxa from the bacterial family Hyphomicrobiaceae is positively correlated with the vineyard age. Plant age has previously been linked to differences in soil bacterial communities in annual crops (Marques et al., 2014; Walters et al., 2018) and in wild plant species (Wagner et al., 2016; Na et al., 2017). However, how composition and diversity of rhizosphere communities shift with plant age in perennial, long-living crops has received less attention. No-till soil management has been shown to affect community composition (Lewis et al., 2018). It is therefore tempting to speculate that in perennial crops, the effect of plant age on soil bacterial communities, is the result of the prolonged presence of the crop.

4.3 Correlations between soil bacterial communities and berry and wine parameters

Berry parameters were found to be significantly associated with both the composition and diversity of soil microbiomes and with the abundance of single taxa. A total of six fruit traits correlated with differences in bacterial community composition and diversity, while one fruit trait was found significantly associated with the abundance of specific taxa. Plant–microbe interactions are known to modify the metabolome of *Arabidopsis thaliana* plants grown under controlled conditions (Badri et al., 2013), however, the modulating effect of soil microbiomes on the metabolome of commercial crops is unexplored. Unfortunately, the non-intervention nature of this research impedes us determining if the relationships observed between vineyard soil microbiomes and fruit traits are causal or simply mere correlations.

Soil microbiomes have previously been described as a contributor to the final sensory properties of wines by affecting wine fermentation. Grape must microbiota was found to be correlated to regional metabolite profiles and was suggested to be potential predictor for the abundance of wine metabolites (Bokulich et al., 2016). Here we identified 19 wine traits correlated with differences in bacterial

community composition and diversity, and seven correlated with the abundance of specific taxa. Vineyard soils may serve as a bacterial reservoir since bacterial communities associated with leaves, flowers, and grapes share a greater proportion of taxa with soil communities than with each other (Zarraonaindia et al., 2015). Unfortunately, the non-intervention nature of this research, the lack of replicability and the use of commercially produced wines, preclude us from determining if the relationships observed between vineyard soil microbiomes and fruit/wine traits are causal or simply mere correlations. Each of these wines was made commercially by different producers so there is potential for a certain level of winemaking effect.

5 Conclusion

Taken collectively our results show that geographic separation between vineyards contributes to bacterial community dissimilarities at a much smaller scale than previously reported. Environmental variables (e.g. climatic, topography, soil properties, and management practices) were the greatest contributor to such differences. Particularly, we found that soil variables are the major shapers of bacterial communities. Also, we show that variables highly affected by soil anthropogenisation (pH, plant available Phosphorous) and agricultural management variables (plant age, planting density) have strong correlations both with the community composition and diversity and the relative abundance of individual taxa. Finally, our results provide an important starting point for future studies investigating the potential influence of microbial communities on the metabolome of grapevines in general, and on the definition of local Terroirs. It will also be important to study a wider range of soil physicochemical properties, and vineyard floor vegetation, on the soil microbiome.

6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

TRC, JRS, AM, MG, JB, CC and CMRL conceived and planned the experiments. CC and RDB contributed to the design of the research project, vineyard selection and fruit and wine chemical characterization. JZ conducted soil physicochemical analysis and the 16S rRNA gene laboratory work. JZ and TMN conducted the bioinformatics analysis. JZ and CMRL took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

8 Funding

This study was funded through a Pilot Program in Genomic Applications in Agriculture and Environment Sectors jointly supported by the University of Adelaide and the Australian Genome Research Facility Ltd. JZ was supported by an Adelaide Graduate Research Scholarship (University of Adelaide). CMRL was supported by a University of Adelaide Beacon Research Fellowship and is currently partially supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, Hatch Program number 2352987000. MG was supported by the Australian Research Council through Centre of Excellence (CE1400008) and Future Fellowship (FT130100709) funding.

9 Acknowledgments

The authors would like to gratefully acknowledge the Barossa Grounds Project and in particular the growers that allowed us to sample material from their properties and supplied information about their vineyards and management strategies. Dr. Kendall R. Corbin contributed to soil sample collection. We are thankful to Dr Hien To, Dr Steve Pederson and Dr Rick Tearle for assistance with data analysis.

10 Data Availability Statement

All the data and supporting information will be made available online.

1. Supplementary Figures S1-S3.
2. Supplementary Tables S1-S11.

The data that support the findings of this study are openly available in NCBI Sequence Read Archive (accession number: PRJNA601984).

11 References

Awasthi, R., Tewari, R., and Nayyar, H. (2011). Synergy between plants and P-solubilizing microbes in soils: effects on growth and physiology of crops. *International Research Journal of Microbiology* 2, 484–503.

Badri, D. V., Zolla, G., Bakker, M. G., Manter, D. K., and Vivanco, J. M. (2013). Potential impact of soil microbiomes on the leaf metabolome and on herbivore feeding behavior. *New Phytologist* 198, 264–273. doi:10.1111/nph.12124.

Barata, A., Malfeito-Ferreira, M., and Loureiro, V. (2012). The microbial ecology of wine grape berries. *International Journal of Food Microbiology* 153, 243–259. doi:10.1016/j.ijfoodmicro.2011.11.025.

Bates, S. T., Berg-Lyons, D., Caporaso, J. G., Walters, W. A., Knight, R., and Fierer, N. (2011). Examining the global distribution of dominant archaeal populations in soil. *ISME Journal* 5, 908–917. doi:10.1038/ismej.2010.171.

Bokulich, N. A., Collins, T. S., Masarweh, C., Allen, G., Heymann, H., Ebeler, S. E., and Mills, D. A. (2016). Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics. *MBio* 7. doi:10.1128/mBio.00631-16.

Broeckling, C. D., Broz, A. K., Bergelson, J., Manter, D. K., and Vivanco, J. M. (2008). Root exudates regulate soil fungal community composition and diversity. *Applied and Environmental Microbiology* 74, 738–744. doi:10.1128/AEM.02188-07.

Burns, K. N., Kluepfel, D. A., Strauss, S. L., Bokulich, N. A., Cantu, D., and Steenwerth, K. L. (2015). Vineyard soil bacterial diversity and composition revealed by 16S rRNA genes: Differentiation

by geographic features. *Soil Biology and Biochemistry* 91, 232–247. doi:10.1016/j.soilbio.2015.09.002.

Bushnell, B. (2016). BBMerge guide. Available at: http://gensoft.pasteur.fr/docs/bbmap/36.11/BBMergeGuide.txt [Accessed October 3, 2016].

Da C Jesus, E., Marsh, T. L., Tiedje, J. M., and de S Moreira, F. M. (2009). Changes in land use alter the structure of bacterial communities in Western Amazon soils. *ISME Journal* 3, 1004–1011. doi:10.1038/ismej.2009.47.

Caporaso, J. G., Bittinger, K., Bushman, F. D., DeSantis, T. Z., Andersen, G. L., and Knight, R. (2010). PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26, 266–267. doi:10.1093/bioinformatics/btp636.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., and Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of National Academy of Sciences of the United States of America* 108 Suppl 1, 4516–4522. doi:10.1073/pnas.1000080107.

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

Cavagnaro T. R. (2016). Life at the interface: above- and below-ground responses of a grazed pasture soil to reforestation. *Applied Soil Ecology* 100, 27-37. doi:10.1016/j.apsoil.2015.12.002

Chen, Y., McCarthy, D., Ritchie, M., Robinson, M., and K.Smyth, G. (2008). *EdgeR:differential expression analysis of digital gene expression data user's guide*. 1st ed.

Compant, S., Mitter, B., Colli-Mull, J. G., Gangl, H., and Sessitsch, A. (2011). Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches of colonization. *Microbial Ecology* 62, 188–197. doi:10.1007/s00248-011-9883-y.

Cong, J., Yang, Y., Liu, X., Lu, H., Liu, X., Zhou, J., Li, D., Yin, H., Ding, J., and Zhang, Y. (2015). Analyses of soil microbial community compositions and functional genes reveal potential consequences of natural forest succession. *Scientific. Reports.* 5, 10007. doi:10.1038/srep10007.

Crowder, D. W., Northfield, T. D., Strand, M. R., and Snyder, W. E. (2010). Organic agriculture promotes evenness and natural pest control. *Nature* 466, 109–112. doi:10.1038/nature09183.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied Environmental Microbiology* 72, 5069–5072. doi:10.1128/AEM.03006-05.

Dias, A. C. F., Dini-Andreote, F., Hannula, S. E., Andreote, F. D., Pereira E Silva, M. de C., Salles, J. F., de Boer, W., van Veen, J., and van Elsas, J. D. (2013). Different selective effects on rhizosphere bacteria exerted by genetically modified versus conventional potato lines. *PLoS*

ONE 8, e67948. doi:10.1371/journal.pone.0067948.

Ding, G.-C., Radl, V., Schloter-Hai, B., Jechalke, S., Heuer, H., Smalla, K., and Schloter, M. (2014). Dynamics of soil bacterial communities in response to repeated application of manure containing sulfadiazine. *PLoS ONE* 9, e92958. doi:10.1371/journal.pone.0092958.

Giddings, J. A. (2015). Determining soil texture using the ribboning technique. NSW Department of Primary Industries. Primefact 1363. 10.1314/RG.2.1.3876.6809.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi:10.1093/bioinformatics/btq461.

Edwards, J., Porter, I., Riches, D., Oliver, D., Bramley, R., Rawnsley, B., and White, B. (2014). Soil Biology: Choosing biological indicators for monitoring vineyard soil quality. *Wine & Viticulture Journal* 29, 48.

Ettema, C. (2002). Spatial soil ecology. *Trends in Ecology & Evolution* 17, 177–183. doi:10.1016/S0169-5347(02)02496-5.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation* 61, 1–10. doi:10.1016/0006-3207(92)91201-3.

Faoro, H., Alves, A. C., Souza, E. M., Rigo, L. U., Cruz, L. M., Al-Janabi, S. M., Monteiro, R. A., Baura, V. A., and Pedrosa, F. O. (2010). Influence of soil characteristics on the diversity of bacteria in the Southern Brazilian Atlantic Forest. *Applied and Environmental Microbiology* 76, 4744–4749. doi:10.1128/AEM.03025-09.

Fierer, N. (2008). "Microbial biogeography: patterns in microbial diversity across space and time," in *Accessing Uncultivated Microorganisms*, ed. K. Zengler. *American Society of Microbiology* 95–115. doi:10.1128/9781555815509.ch6.

Fierer, N., and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America* 103, 626–631. doi:10.1073/pnas.0507535103.

Frey, S. D., Knorr, M., Parrent, J. L., and Simpson, R. T. (2004). Chronic nitrogen enrichment affects the structure and function of the soil microbial community in temperate hardwood and pine forests. *Forest Ecology and Management* 196, 159–171.

Girvan, M. S., Bullimore, J., Pretty, J. N., Osborn, A. M., and Ball, A. S. (2003). Soil type is the primary determinant of the composition of the total and active bacterial communities in arable soils. *Applied and Environmental Microbiology* 69, 1800–1809.

Goswami, D., Vaghela, H., Parmar, S., Dhandhukia, P. and Thakker, J.N. (2013). Plant growth promoting potentials of Pseudomonas spp. strain OG isolated from marine water. *Journal of Plant Interactions* 8, 281-290. doi:10.1080/17429145.2013.768360

Hartman, K., van der Heijden, M. G. A., Wittwer, R. A., Banerjee, S., Walser, J.-C., and Schlaeppi, K. (2018). Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. *Microbiome* 6, 14. doi:10.1186/s40168-017-0389-

9.

Hermans, S. M., Buckley, H. L., Case, B. S., Curran-Cournane, F., Taylor, M., and Lear, G. (2017). Bacteria as emerging indicators of soil condition. *Applied and Environmental Microbiology* 83. doi:10.1128/AEM.02826-16.

Fabres, P., Collins, C., Cavagnaro, T. R., and Rodríguez López, C. M. (2017). A Concise review on multi-omics data integration for *terroir* analysis in *Vitis vinifera*. *Frontiers in Plant Science* 8, 1065. doi:10.3389/fpls.2017.01065.

Kim, Y., and Liesack, W. (2015). Differential assemblage of functional units in paddy soil microbiomes. *PLoS ONE* 10, e0122221. doi:10.1371/journal.pone.0122221.

Krogstad, T., Sogn, T. A., Asdal, Å., and Sæbø, A. (2005). Influence of chemically and biologically stabilized sewage sludge on plant-available phosphorous in soil. *Ecological Engineering* 25, 51–60. doi:10.1016/j.ecoleng.2005.02.009.

Lamb, E. G., Kennedy, N., and Siciliano, S. D. (2011). Effects of plant species richness and evenness on soil microbial community diversity and function. *Plant and Soil* 338, 483–495. doi:10.1007/s11104-010-0560-6.

Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and Environmental Microbiology* 75, 5111–5120. doi:10.1128/AEM.00335-09.

Lewis, R., Barth, V., Coffey, T., McFarland, C., Huggins, D., and Sullivan, T. (2018). Altered bacterial communities in long-term no-till soils associated with stratification of soluble aluminum and soil pH. *Soil Systems* 2, 7. doi:10.3390/soils2010007.

Liu, J., Sui, Y., Yu, Z., Shi, Y., Chu, H., Jin, J., Liu, X., and Wang, G. (2014). 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.

Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* 71, 8228–8235. doi:10.1128/AEM.71.12.8228-8235.2005.

Lumini, E., Vallino, M., Alguacil, M. M., Romani, M., and Bianciotto, V. (2011). Different farming and water regimes in Italian rice fields affect arbuscular mycorrhizal fungal soil communities. *Ecological Applications* 21, 1696–1707. doi:10.1890/10-1542.1.

Marques, J. M., da Silva, T. F., Vollu, R. E., Blank, A. F., Ding, G.-C., Seldin, L., and Smalla, K. (2014). Plant age and genotype affect the bacterial community composition in the tuber rhizosphere of field-grown sweet potato plants. *FEMS Microbiology Ecology* 88, 424–435. doi:10.1111/1574-6941.12313.

Martins, G., Lauga, B., Miot-Sertier, C., Mercier, A., Lonvaud, A., Soulas, M.-L., Soulas, G., and Masneuf-Pomarède, I. (2013). Characterization of epiphytic bacterial communities from grapes, leaves, bark and soil of grapevine plants grown, and their relations. *PLoS ONE* 8,

e73013. doi:10.1371/journal.pone.0073013.

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., Andersen, G. L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal* 6, 610–618. doi:10.1038/ismej.2011.139.

Mezzasalma, V., Sandionigi, A., Guzzetti, L., Galimberti, A., Grando, M. S., Tardaguila, J., and Labra, M. (2018). Geographical and cultivar features differentiate grape microbiota in Northern Italy and Spain vineyards. *Frontiers in Microbiology* 9, 946. doi:10.3389/fmicb.2018.00946.

Mira de Orduña, R. (2010). 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.

Miura, T., Sánchez, R., Castañeda, L. E., Godoy, K., and Barbosa, O. (2017). Is microbial terroir related to geographic distance between vineyards? *Environmental Microbiology Reports* 9, 742–749. doi:10.1111/1758-2229.12589.

Na, X., Li, X., Zhang, Z., Li, M., Kardol, P., Xu, T. T., Wang, M., Cao, X., and Ma, F. (2017). Bacterial community dynamics in the rhizosphere of a long-lived, leguminous shrub across a 40-year age sequence. *Journal of Soils and Sediments* 18, 1–9. doi:10.1007/s11368-017-1745-x.

Nannipieri, P., Ascher, J., Ceccherini, M. T., Landi, L., Pietramellara, G., and Renella, G. (2003). Microbial diversity and soil functions. *European Journal of Soil Science* 54, 655–670. doi:10.1046/j.1351-0754.2003.0556.x.

Navas-Molina, J. A., Peralta-Sánchez, J. M., González, A., McMurdie, P. J., Vázquez-Baeza, Y., Xu, Z., Ursell, L. K., Lauber, C., Zhou, H., Song, S. J., et al. (2013). Advancing our understanding of the human microbiome using QIIME. *Methods in Enzymology* 531, 371–444. doi:10.1016/B978-0-12-407863-5.00019-8.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., H. Stevens, M.H., and Wagner, H. (2013). Package "vegan." *Community Ecology Package 2, 1-295*

Parks, D. H., Mankowski, T., Zangooei, S., Porter, M. S., Armanini, D. G., Baird, D. J., Langille, M. G. I., and Beiko, R. G. (2013). GenGIS 2: geospatial analysis of traditional and genetic biodiversity, with new gradient algorithms and an extensible plugin framework. *PLoS ONE* 8, e69885. doi:10.1371/journal.pone.0069885.

Parks, D. H., Porter, M., Churcher, S., Wang, S., Blouin, C., Whalley, J., Brooks, S., and Beiko, R. G. (2009). GenGIS: a geospatial information system for genomic data. *Genome Research* 19, 1896–1904. doi:10.1101/gr.095612.109.

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28, 2537–2539. doi:10.1093/bioinformatics/bts460.

Philippot, L., Raaijmakers, J. M., Lemanceau, P., and van der Putten, W. H. (2013). Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews. Microbiology* 11, 789–799.

649 doi:10.1038/nrmicro3109.

650 Ranjard, L., Dequiedt, S., Chemidlin Prévost-Bouré, N., Thioulouse, J., Saby, N. P. A., Lelievre, M.,
651 Maron, P. A., Morin, F. E. R., Bispo, A., Jolivet, C., et al. (2013). Turnover of soil bacterial
652 diversity driven by wide-scale environmental heterogeneity. *Nature Communications* 4, 1434.
653 doi:10.1038/ncomms2431.

654 Rayment, G. E., and Higginson, F. R. (1992). *Australian laboratory handbook of soil and water*
655 *chemical methods*. Inkata Press Pty Ltd.

656 Reeve, J. R., Schadt, C. W., Carpenter-Boggs, L., Kang, S., Zhou, J., and Reganold, J. P. (2010). Effects
657 of soil type and farm management on soil ecological functional genes and microbial activities.
658 The *ISME Journal* 4, 1099–1107. doi:10.1038/ismej.2010.42.

659 Robinson, S., and Sandercock, N. (2014). An analysis of climate, soil and topographic information to
660 aid the understanding of barossa terroir. Adelaide, Australia: SA: Primary Industries and
661 Regions South Australia.

662 Romero, P., Fernandez-Fernandez, J., and Botia Ordaz, P. (2016). Interannual climatic variability
663 effects on yield, berry and wine quality indices in long-term deficit irrigated grapevines,
664 determined by multivariate analysis. *International Journal of Wine Research* 8, 3–17.
665 doi:10.2147/IJWR.S107312.

666 Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., Knight, R., and
667 Fierer, N. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil.
668 The *ISME Journal* 4, 1340–1351. doi:10.1038/ismej.2010.58.

669 Schmalenberger, A., and Fox, A. (2016). Bacterial mobilization of nutrients from biochar-amended
670 soils. *Advances in Applied Microbiology* 94, 109–159. doi:10.1016/bs.aambs.2015.10.001.

671 Sessitsch, A., Weilharter, A., Gerzabek, M. H., Kirchmann, H., and Kandeler, E. (2001). Microbial
672 population structures in soil particle size fractions of a long-term fertilizer field experiment.
673 *Applied and Environmental Microbiology* 67, 4215–4224. doi:10.1128/AEM.67.9.4215-
674 4224.2001.

675 Sugiyama, A., Vivanco, J. M., Jayanty, S. S., and Manter, D. K. (2010). Pyrosequencing assessment
676 of soil microbial communities in organic and conventional potato farms. *Plant disease* 94,
677 1329–1335. doi:10.1094/PDIS-02-10-0090.

678 Team, R. C. (2013). R: A language and environment for statistical computing.

679 Tripathi, B. M., Edwards, D. P., Mendes, L. W., Kim, M., Dong, K., Kim, H., and Adams, J. M. (2016).
680 The impact of tropical forest logging and oil palm agriculture on the soil microbiome.
681 *Molecular Ecology* 25, 2244–2257. doi:10.1111/mec.13620.

682 Van Leeuwen, C., Friant, P., Chone, X., Tregoat, O., Koundouras, S., and Dubourdieu, D. (2004).
683 Influence of climate, soil, and cultivar on terroir. *American Journal of Enology and Viticulture*
684 *55, 207-217*.doi: ajevonline.org/cgi/reprint/55/3/207

685 Vázquez-Baeza, Y., Pirrung, M., Gonzalez, A., and Knight, R. (2013). EMPeror: a tool for visualizing

high-throughput microbial community data. *Gigascience* 2, 1. doi: 10.1186/2047-217X-2-16.

Vega-Avila, A. D., Gumiere, T., Andrade, P. A. M., Lima-Perim, J. E., Durrer, A., Baigori, M., Vazquez, F., and Andreote, F. D. (2015). Bacterial communities in the rhizosphere of Vitis vinifera L. cultivated under distinct agricultural practices in Argentina. *Antonie Van Leeuwenhoek* 107, 575–588. doi:10.1007/s10482-014-0353-7.

Wagner, I. D., and Wiegel, J. (2008). Diversity of thermophilic anaerobes. *Annals of the New York Academy of Sciences* 1125, 1–43. doi:10.1196/annals.1419.029.

Wagner, M. R., Lundberg, D. S., Del Rio, T. G., Tringe, S. G., Dangl, J. L., and Mitchell-Olds, T. (2016). Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nature Communications* 7, 12151. doi:10.1038/ncomms12151.

Walters, W. A., Jin, Z., Youngblut, N., Wallace, J. G., Sutter, J., Zhang, W., González-Peña, A., Peiffer, J., Koren, O., Shi, Q., et al. (2018). Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proceedings of the National Academy of Sciences of the United States of America* 115, 7368–7373. doi:10.1073/pnas.1800918115.

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* 73, 5261–5267. doi:10.1128/AEM.00062-07.

Wang, W., Wang, H., Feng, Y., Wang, L., Xiao, X., Xi, Y., Luo, X., Sun, R., Ye, X., Huang, Y., et al. (2016). Consistent responses of the microbial community structure to organic farming along the middle and lower reaches of the Yangtze River. *Scientific Reports* 6, 35046. doi:10.1038/srep35046.

Weckert, M. (2016). Soil microbiology: Vineyard terroir – is there a link with soil microbiology? *Wine & Viticulture Journal* 31, 39-40.

Wildman, H. (2015). Improving mine rehabilitation success through microbial management. *Journal of Environmental Solutions for Oil, Gas, and Mining* 1, 32–46. doi:10.3992/1573-2377-374X-1.1.32.

Wittebolle, L., Marzorati, M., Clement, L., Balloi, A., Daffonchio, D., Heylen, K., De Vos, P., Verstraete, W., and Boon, N. (2009). Initial community evenness favours functionality under selective stress. *Nature* 458, 623–626. doi:10.1038/nature07840.

Wolf, T. K., Dry, P. R., Iland, P. G., Botting, D., Dick, J., Kennedy, U., and Ristic, R. (2003). Response of Shiraz grapevines to five different training systems in the Barossa Valley, Australia. *Australian Journal of Grape and Wine Research* 9, 82–95.

Wu, Y., Zeng, J., Zhu, Q., Zhang, Z., and Lin, X. (2017). pH is the primary determinant of the bacterial community structure in agricultural soils impacted by polycyclic aromatic hydrocarbon pollution. *Scientific Reports* 7, 40093. doi:10.1038/srep40093.

Xie, H., Konate, M., Sai, N., Tesfamicael, K. G., Cavagnaro, T., Gilliham, M., Breen, J., Metcalfe, A., DeBei, R., Collins, C., et al. (2017). Environmental conditions and agronomic practices induce consistent global changes in DNA methylation patterns in grapevine (Vitis vinifera cv Shiraz).

BioRxiv. doi:10.1101/127977.

Zarraonaindia, I., Owens, S. M., Weisenhorn, P., West, K., Hampton-Marcell, J., Lax, S., Bokulich, N. A., Mills, D. A., Martin, G., Taghavi, S., et al. (2015). The soil microbiome influences grapevine-associated microbiota. *mBio* 6. doi:10.1128/mBio.02527-14.

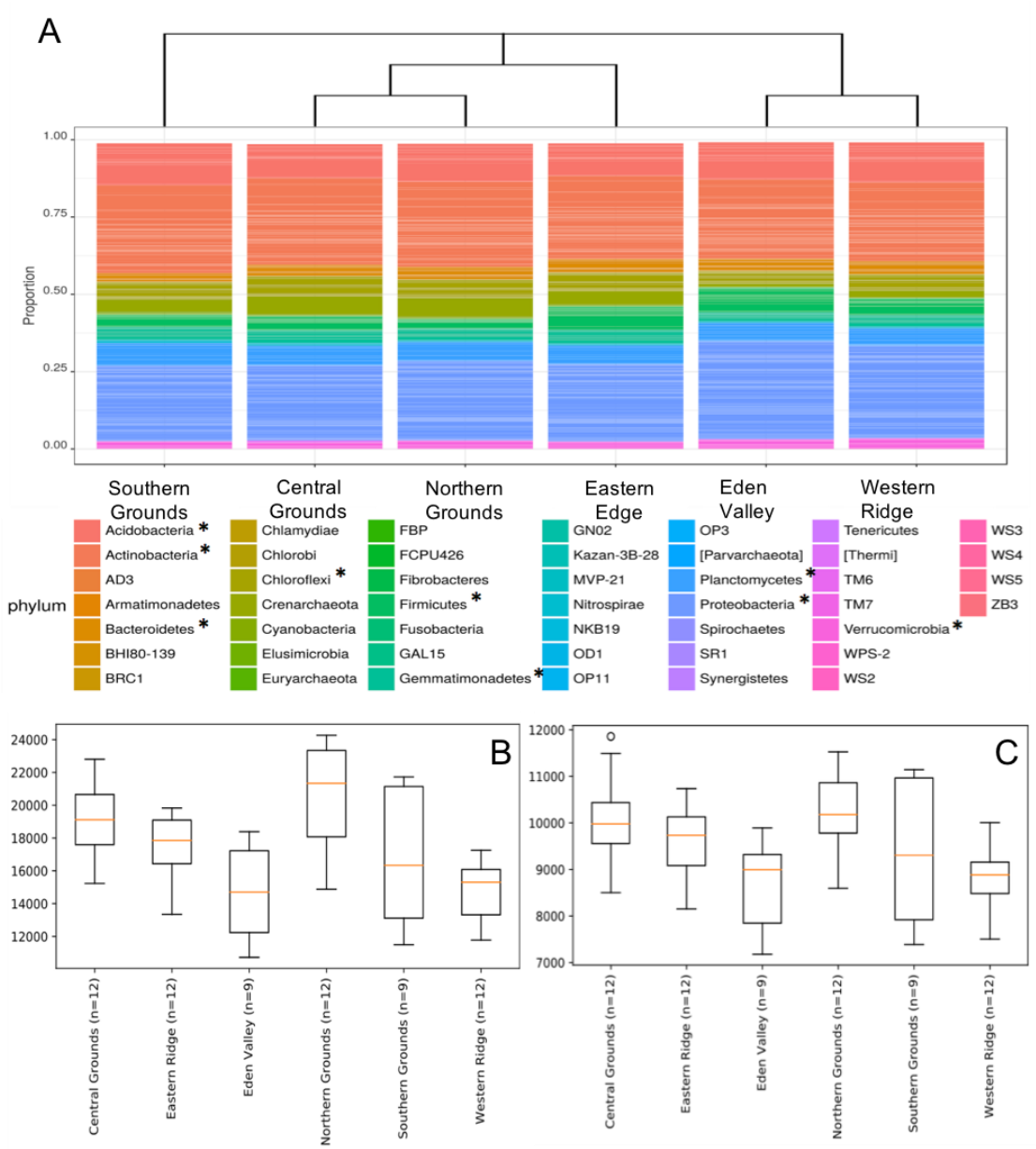
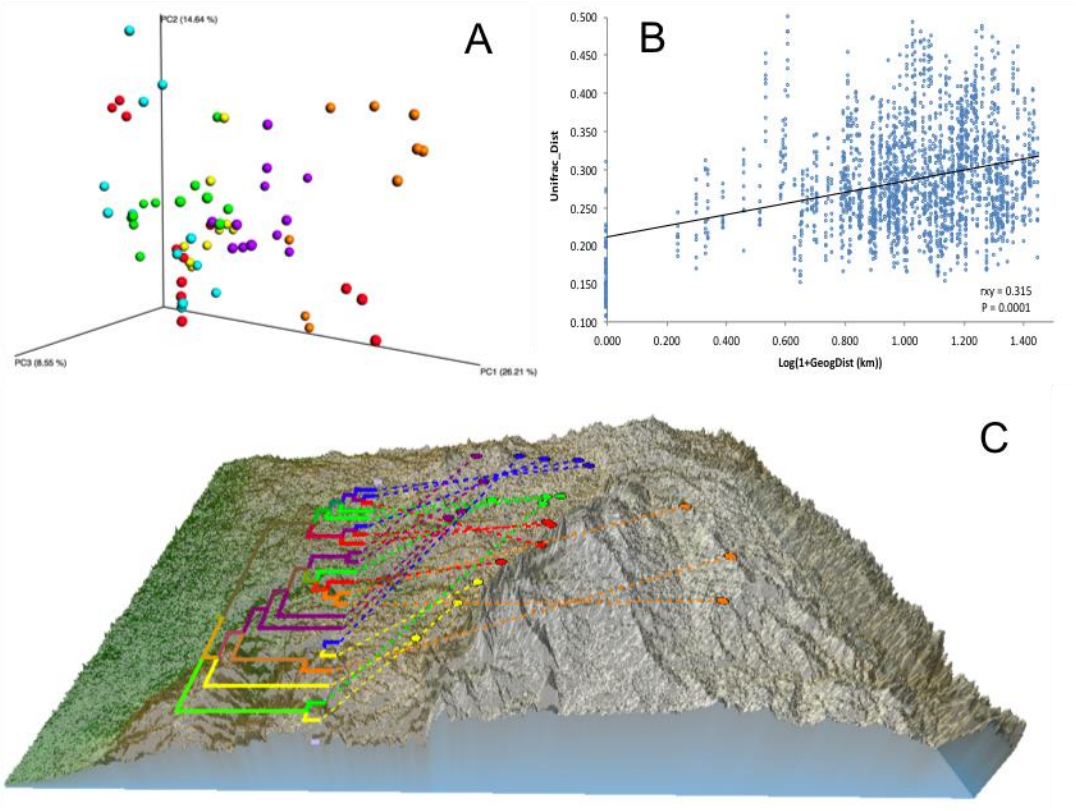


Figure 1. Soil bacteria community composition and diversity in 6 Barossa sub-regions. **A)** Phylogenetic inference of microbiome composition differences between Barossa sub-regions. Neighbour joining tree was generated with weighted UniFrac distances calculated with sequences classifiable at the phylum level (98.9% of total). 96.5% of all sequences were assigned to one of nine

733 main dominant groups (relative abundance >= 1.0%) (indicated here by *). **B)** Alpha diversity: Chao1
734 diversity comparison, **C)** and observed species diversity comparison. Alpha diversity values were
735 calculated based on rarefied data was established using 16S sequencing reads from 3 soil samples per
736 vineyard.

737



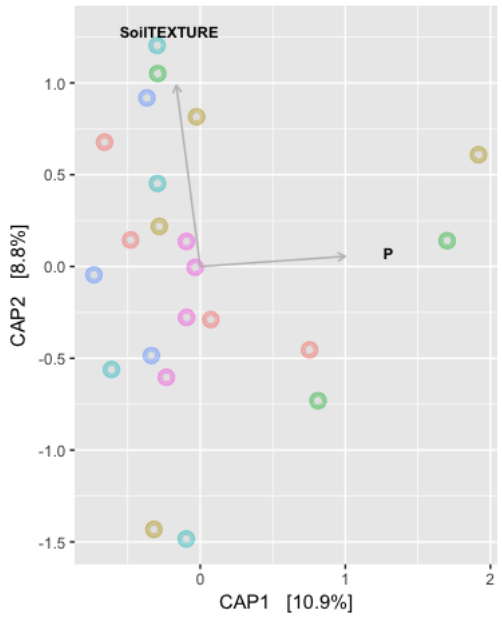
738

739 **Figure 2. Effect of vineyard location on soil microbiome differentiation. A)** PCoA based on Beta
740 diversity of soil bacterial communities calculated using weighted UniFrac distances. Values were
741 calculated based on rarefied data to 37,176 sequences per sample. **B)** Relationship between
742 phylogenetic Beta diversity and geographic distance. Unifrac_dist indicates weighted UniFrac
743 distances. Geographic distances were calculated from latitude/longitude coordinates using GenAlex
744 v6.5 *geographic distance* function implemented as Log(1+distances in Kilometres). The relationship
745 was tested using Mantel's correlation coefficient (rxy) with its probability estimate for significance (P)

746 based on 9,999 random permutations and implemented using GenAlex v6.5. **C)** Neighbour joining
747 ultrametric tree calculated from Beta diversity weighted UniFrac distance matrix between 22 vineyards
748 located in six sub-regions: Northern Grounds (blue); Southern Grounds (yellow); Central Grounds
749 (green); Eastern Edge (red); Western Ridge (purple); Eden Valley (orange). Tree was overlayed with
750 the Barossa Region elevation map using GenGIS. Beta diversity was established using 16S sequencing
751 reads from 3 soil samples per vineyard.

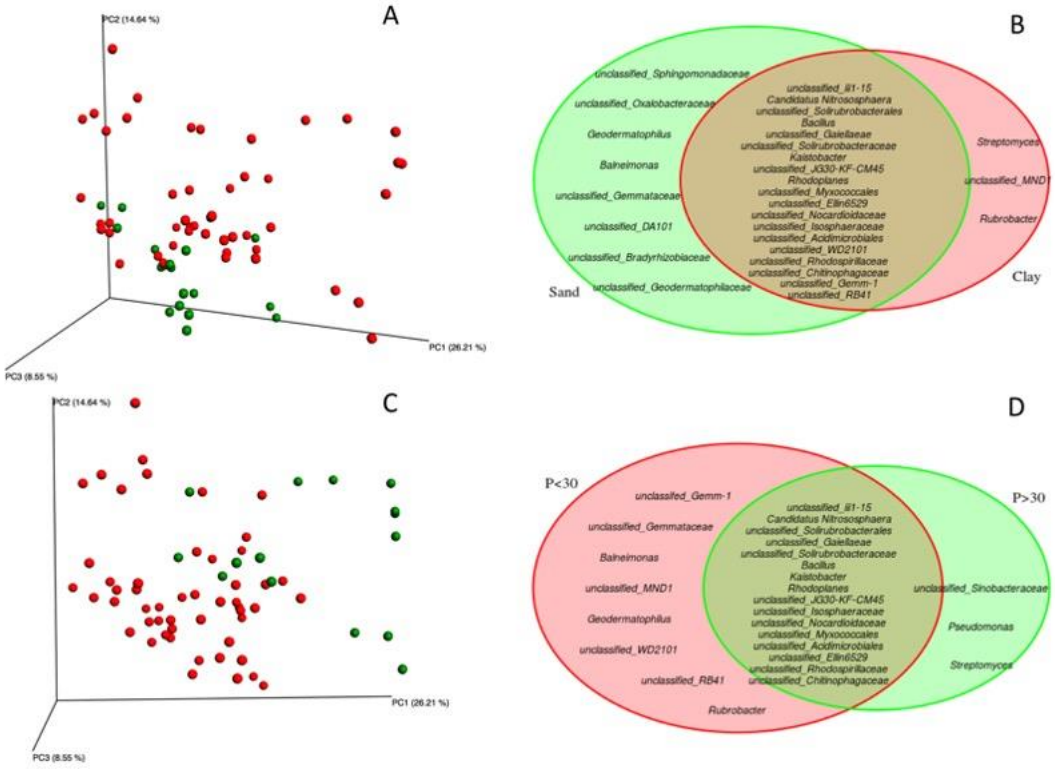
752

753



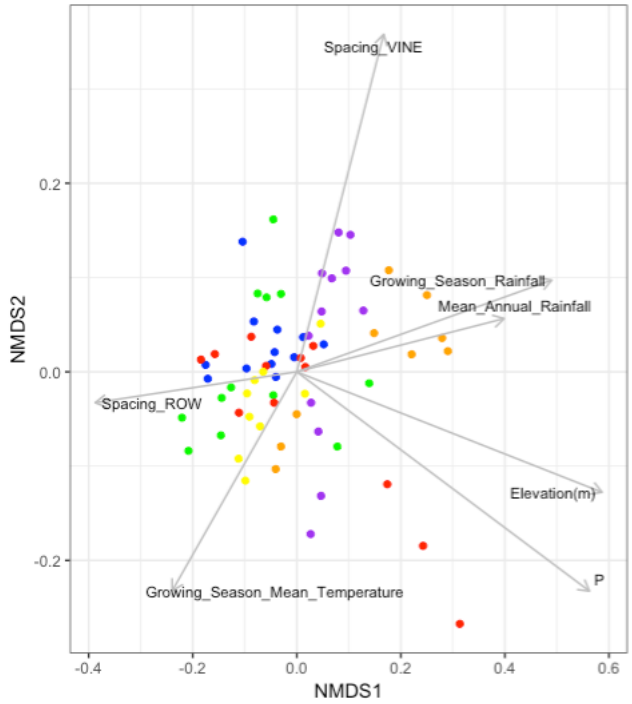
754

755 **Figure 3. Main drivers of soil microbiome differentiation between Barossa Region vineyards.** The
756 observed important soil factors that affect soil microbial community groups in combinations. CAP plot
757 displays the combination of variables that explained the greatest variation in the soil microbiome
758 through model selection (full results Table 1). The correlation test was carried out on environmental
759 variables following the removal of the highly correlated variables (>0.85) using the function *ordisten*,
760 in the package *Vegan*. The variables implemented in the final model were soil phosphorous and soil
761 texture, which explained 19.7% of variation in the soil microbiome. Distance based redundancy
762 analysis (dbRDA) with Bray-Curtis dissimilarity matrix of OTUs was used to examine the influence
763 of these predictor variables using the function capscale in the package Vegan in R.



764

765 **Figure 4. Identification of microbial genera associated to soil texture and plant-available**
766 **phosphorous in Barossa Region vineyards.** Principal coordinate analysis plots display weighted
767 UniFrac distances of soil samples from 22 vineyards in six sub-regions of Barossa Valley. Venn
768 Diagrams show significantly different ($P > 0.01$) genera. Plots and diagrams are grouped by (A/B) soil
769 type (clay (red) versus sandy soils (green)), and (C/D) plant-available Phosphorous (P) ($P < 30$ µg P /
770 g soil (red), $P > 30$ µg P / g soil (green). Beta diversity was established using 16S sequencing reads
771 from 3 soil samples per vineyard.



772

773 **Figure 5. Environmental and vineyard management factors significantly associated with soil**
774 **microbial community composition in Barossa Region vineyards.** Non-metric multidimensional
775 scaling plot displays the microbial community composition of 22 vineyards located in six sub-regions:
776 Northern Grounds (blue); Southern Grounds (yellow); Central Grounds (green); Eastern Edge (red);
777 Western Ridge (purple); Eden Valley (orange). Vector arrows indicate the association with
778 environmental variables with p-value < 0.05. Arrow heads indicate the direction and length indicates
779 the strength of the variable and nMDS correlation. Analysis was conducted using 999 permutations
780 with variables deemed significant where p-value < 0.05.

781

782

783

784

785 **Table 1. Main drivers of soil microbiome differentiation between Barossa Region vineyards.**
786 Variables that explained the greatest variation in the soil microbiome through model selection. The
787 correlation test was carried out on environmental variables following the removal of the highly
788 correlated variables (>0.85) using the function *ordisten*, in the package *Vegan*.

789

Step	Df	Deviance	Resid. Df	Resid. Dev	AIC
1	NA	NA	21	75.29566	29.06836
Soil P	-1	8.170918	20	67.12475	28.54123
Soil texture	-1	6.620899	19	60.50386	28.25662

790

791

792

793

794

795

796

797

798

799

800

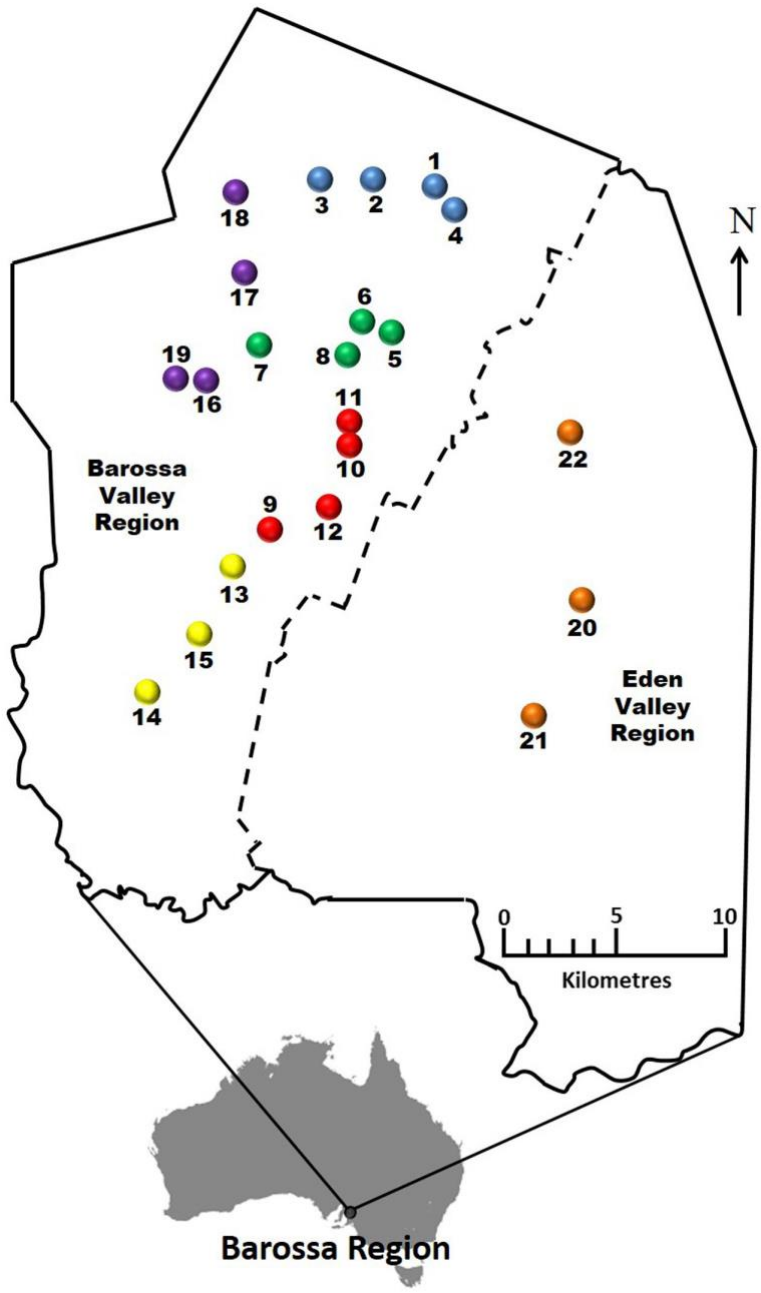
801

802

Table 2. Fruit and wine characteristics significantly associated with microbial community composition in Barossa Region vineyards. Table shows the *envfit* output that was carried out the correlation test between grape and wine characteristics variables that fitted onto an ordination of nonmetric multidimensional scaling (nMDS) plots of microbial community data from soils in 22 vineyard sites. Analysis was conducted using 999 permutations with variables deemed significant where p-value < 0.05.

Variables		NMDS1	NMDS2	r2	Pr(>r)
Basic berry properties	50 berries weight	-0.87544	-0.48332	0.1612	0.008**
	TA berry	0.9369	0.3496	0.1119	0.029*
	Average colour	0.76859	0.63974	0.1337	0.008**
	Average total phenolics berry	0.76558	0.64334	0.135	0.015*
	Malic acid	-0.90493	0.42557	0.104	0.03*
Basic wine chemistry	Total phenolics	**0.83761**	**0.54627**	**0.2132**	**0.002****
	Total anthocyanins	0.99519	0.09801	0.2507	0.001***
	Colour density (so2 corrected)	0.72894	0.68457	0.1449	0.006**
	Hue	-0.78985	0.61331	0.1314	0.011*
Wine amino acids	Alanine	**0.11831**	**0.99298**	**0.124**	**0.017***
	Asparagine	0.55124	0.83435	0.1156	0.023*
	Glutamate	0.39571	0.91837	0.103	0.031*
	Glycine	0.62968	0.77685	0.1847	0.002**
	Serine	0.54731	0.83693	0.0936	0.04*
	Threonine	0.20213	0.97936	0.0934	0.049*
	Tryptophan	0.56228	0.82695	0.119	0.025*

Wine ferment. products	Acetic acid	-0.99914	-0.04153	0.1689	0.003**
	Propanoic acid	-0.96827	0.24991	0.118	0.012*
	3-methylbutanol	0.99079	0.13538	0.1184	0.02*
	2-methylbutanol	0.98968	0.14332	0.108	0.034*
	Butanoic acid	-0.89212	0.4518	0.1298	0.013*
	2-phenyl ethyl ethanol	0.70731	0.7069	0.2064	0.001***
	2-phenyl ethyl acetate	0.82725	0.56183	0.1249	0.013*



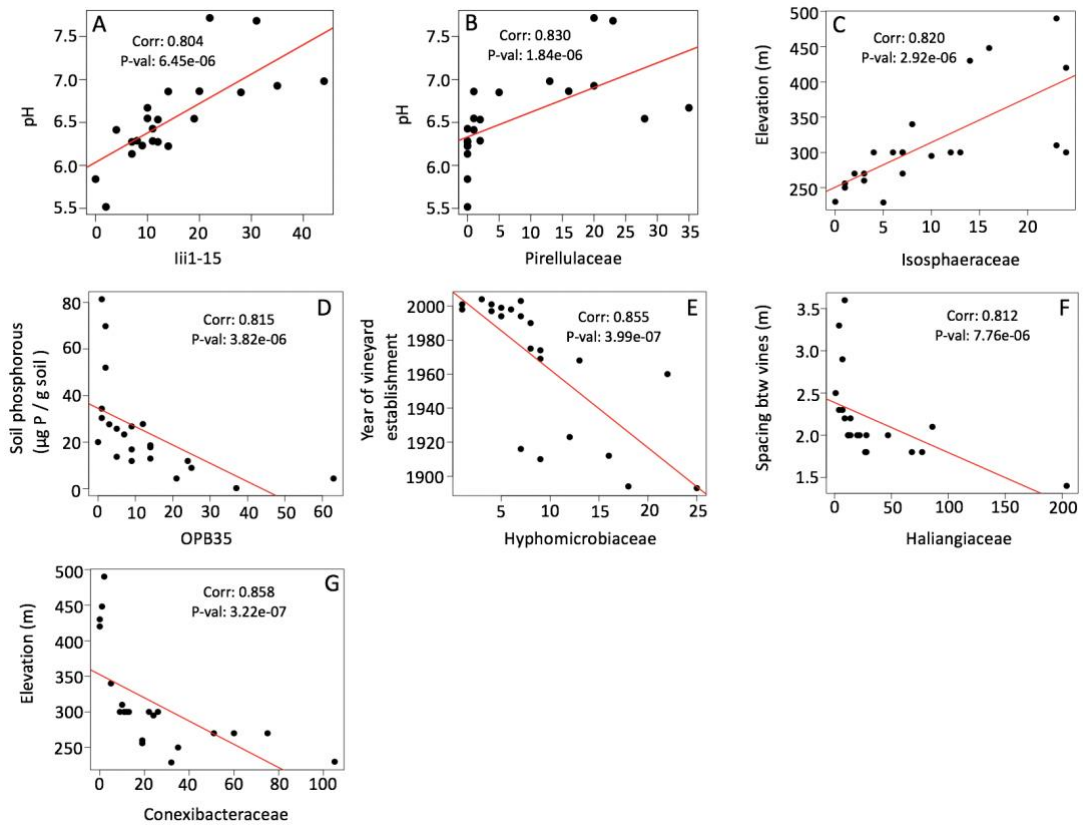
809

810 **Supplementary Figure 1. Location of 22 Barossa vineyard sites.** Vineyards are color coded
811 according to the six wine sub-regions as defined in Xie et al. (2017): Northern Grounds: Blue, Southern
812 Grounds: Yellow, Central Grounds: Green, Eastern Ridge: Red, Western Ridge: Purple, Eden Valley:
813 Orange. Map modified from Xie et al. (2017).

814

815

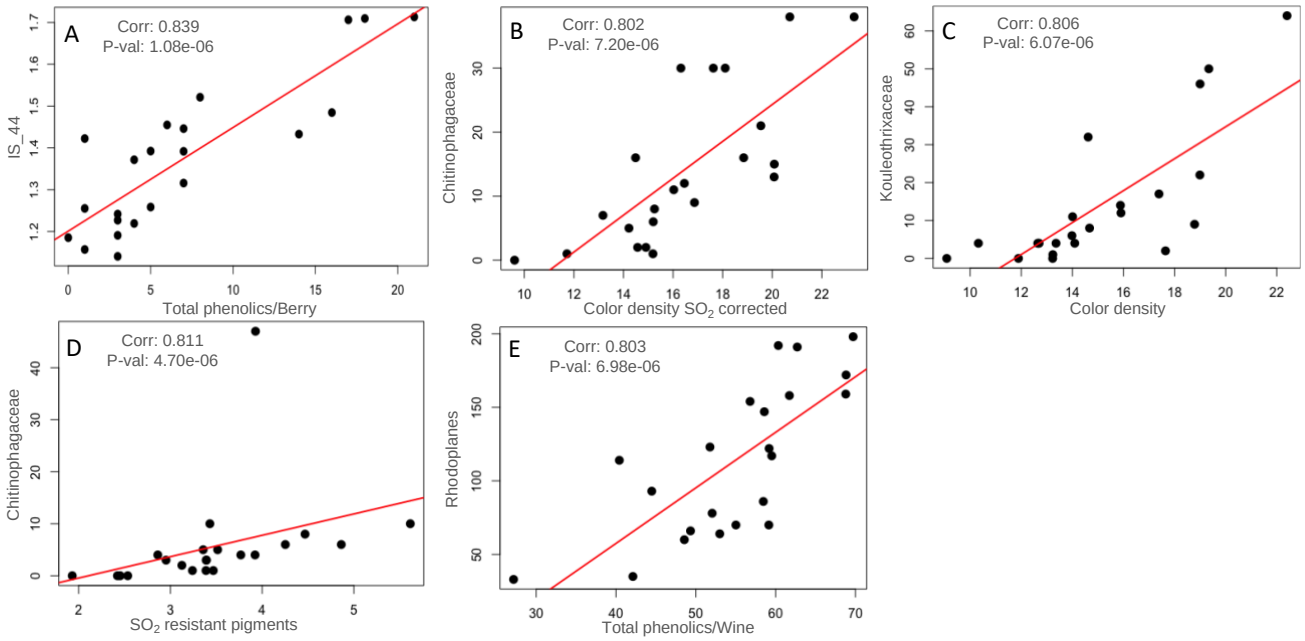
816



817

818 **Supplementary Figure 2. Association between taxon abundance and environmental/agronomical**
819 **variables in Barossa Region vineyard soil bacteria communities.** Correlations were tested using
820 Spearman's rank correlation coefficient with its probability estimate for significance (P) and
821 implemented using the function *rcorr* in the R package *Hmisc*. Correlation coefficient and P values for
822 each of the comparisons are included in each inset.

823



824

825 **Supplementary Figure 3. Association between taxon abundance and fruit/wine traits in Barossa**
826 **Region vineyard soil bacteria communities.** Relationship between taxon abundance and fruit (A) and
827 wine (B-F) traits. Correlations were tested using Spearman's rank correlation coefficient with its
828 probability estimate for significance (P) and implemented using the function *rcorr* in the R package
829 *Hmisc*. Correlation coefficient and P values for each of the comparisons are included in each inset.